



**Victor Eduardo Leite de Almeida Duca**

**Uma Nova Abordagem para Geração de  
Cenários de Séries Temporais Periódicas  
Via Modelos Não Lineares**

**Dissertação de Mestrado**

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

Orientador: Prof. Reinaldo Castro Souza  
Co-orientador: Prof. Pedro Guilherme Costa Ferreira



**Victor Eduardo Leite de Almeida Duca**

**Uma Nova Abordagem para Geração de Cenários de Séries Temporais Periódicas Via Modelos Não Lineares**

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

**Prof. Reinaldo Castro Souza**  
Orientador  
DEE – PUC-Rio

**Prof. Pedro Guilherme Costa Ferreira**  
Co-orientador  
FGV

**Prof. Fernando Luiz Cyrino Oliveira**  
PUC-Rio

**Prof. Joari Paulo da Costa**  
ONS

**Prof. Márcio da Silveira Carvalho**  
Coordenador Setorial do Centro  
Técnico Científico

Rio de Janeiro, 29 de fevereiro de 2016

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

### **Victor Eduardo Leite de Almeida Duca**

Graduou-se em Estatística pela Universidade Federal Fluminense (UFF). Mestre em Engenharia Elétrica (Métodos de Apoio à Decisão) pela PUC-Rio. Atualmente é aluno de doutorado em Estatística na Universidade Federal do Rio de Janeiro (UFRJ). Trabalhou como professor substituto no Departamento de Estatística da Universidade Federal Fluminense (UFF).

#### Ficha Catalográfica

Uma Nova Abordagem para Geração de Cenários de Séries Temporais Periódicas Via Modelos Não Lineares / Victor Eduardo Leite de Almeida Duca; orientador: Reinaldo Castro Souza; co-orientador: Pedro Guilherme Costa Ferreira. — Rio de Janeiro : PUC-Rio, Departamento de Engenharia Elétrica, 2016.

v., 94 f: il. ; 29,7 cm

1. Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Inclui referências bibliográficas.

1. Engenharia Elétrica – Tese. 2. Modelo Periódico Autor-regressivo. 3. PGAR(1). 4. PAGAR(1). 5. PMGAR(1). 6. Simulação de Séries Sintéticas. I. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. II. Título.

## Agradecimentos

Gostaria de agradecer primeiramente a Deus por todo o meu caminho trilhado até o momento.

Agradeço em especial à Bruna Melo Guimarães por sempre estar ao meu lado me apoiando e por me fazer cada dia mais feliz.

Agradeço à minha família pelo apoio dado em todas as minhas decisões.

Agradeço à minha grande amiga Silnelly Caldeira Mattos por todo apoio ao longo da minha vida.

Agradeço aos meus amigos e professores da Universidade Federal Fluminense pelo apoio e ensinamento dado ao longo dos anos.

Agradeço ao CNPq e à PUC-Rio pelo apoio financeiro dado ao longo do meu mestrado.

Agradeço aos professores do Departamento de Engenharia Elétrica da PUC-Rio por todo ensinamento adquirido ao longo do meu mestrado.

Agradeço ao meu orientador, Prof. Reinaldo Castro Souza, por aceitar me orientar e pelo ensinamento adquirido durante meu mestrado

Agradeço ao meu co-orientador, Prof. Pedro Guilherme Costa Ferreira por me co-orientar e pelo ensinamento durante minha dissertação.

Agradeço aos membros da banca, Prof. Fernando Luiz Cyrino Oliveira e Dr. Joari Paulo da Costa, por aceitarem o convite da minha defesa de dissertação.

Agradeço aos meus amigos que conheci durante a minha trajetória no Mestrado e que agora fazem parte da minha vida.

## Resumo

Duca, Victor Eduardo Leite de Almeida; Souza, Reinaldo Castro (Orientador); Ferreira, Pedro Guilherme Costa (Co-orientador). **Uma Nova Abordagem para Geração de Cenários de Séries Temporais Periódicas Via Modelos Não Lineares**. Rio de Janeiro, 2016. 94p. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Os modelos autorregressivos são comumente encontrados dentro do contexto de séries hidrológicas, especificamente em séries de vazões e/ou ENA (Energia Natural Afluente). Muitos destes modelos são de ordem 1, possuem parâmetros constantes ou periódicos e necessitam do requisito de normalidade. Segundo a literatura, séries de vazões anuais podem ser aproximadas para distribuições normais, porém em períodos de tempo curtos como diário, semanal e mensal esta característica não é observada, especialmente pelo problema de assimetria. Devido a isto, uma nova classe de modelo de ordem 1 foi estudada na tentativa de suprir tal problema. O novo modelo mantém estrutura autorregressiva, pode ser aditivo, multiplicativo ou híbrido, onde incorpora propriedades aditivas e multiplicativas conjuntamente, porém suas marginais assumirão distribuição gama. Além disso, a modelagem parte do pressuposto que os Métodos de Momentos são eficientes para estimação de seus parâmetros. Recentemente esta abordagem, sob a forma híbrida, não demonstrou sucesso para o contexto do despacho hidrotérmico brasileiro. O presente trabalho foca na análise completa do modelo híbrido para as séries do Setor Elétrico Brasileiro, trazendo como novidade a estimação via máxima verossimilhança além dos estudos isolados de modelos aditivos e multiplicativos. Os resultados revelaram uma linha de pesquisa promissora, abrindo um campo de possibilidades para que novas ordens superiores a primeira ou distribuições assimétricas possam ser estudadas partindo deste princípio.

## Palavras-chave

Modelo Periódico Autorregressivo; PGAR(1); PAGAR(1); PMGAR(1); Simulação de Séries Sintéticas.

## Abstract

Duca, Victor Eduardo Leite de Almeida; Souza, Reinaldo Castro (Advisor); Ferreira, Pedro Guilherme Costa (Co-advisor). **A New Approach to generate time series periodical scenarios via Non-Linear Models**. Rio de Janeiro, 2016. 94p. MSc. Dissertation - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Autoregressive models are commonly found in the context of hydrological series, specifically in streamflow and/or ANE series (Affluent Natural Energy). Most of them are models of order 1, which have constant or periodic parameters and need the requirement of normality. According to the literature, annual streamflow series can be approximated for normal distributions, however, in short periods of time, like daily, weekly and monthly, this feature is not observed, especially because of the asymmetry issue. Due to this reason, a new class of model of order 1 was studied for attempting to solve such problem. The new model keeps autoregressive structure and can be additive, multiplicative or hybrid, in which embodies additive and multiplicative properties together, but its marginals will assume gamma distribution. Moreover, this modeling departs from the presupposition that Methods of Moments are efficient to the estimation of its parameters. Recently, this approach, under the hybrid way, has not proved to be successful to the context of Brazilian hydrothermal dispatch. This work focuses on the complete analysis of hybrid model to the series of the Brazilian Electric Sector, bringing, as novelty, Maximum Likelihood Estimation, besides isolated studies of additive and multiplicative models. The results revealed a prosperous line of research, opening a field of possibilities for new orders or asymmetric distributions to be studied starting from this point.

## Keywords

Periodic Autoregressive; PGAR(1); PAGAR(1); PMGAR(1); Synthetic Time Series Simulation.

## Sumário

1	Introdução	12
2	A Metodologia Box e Jenkins	16
3	O Modelo Periódico Autorregressivo Gama	21
3.1	O Modelo Periódico Autorregressivo Aditivo Gama de Ordem 1	22
3.2	Modelo Periódico Autorregressivo Multiplicativo Gama de Ordem 1	25
3.3	Casos particulares para a modelagem PMGAR(1)	28
3.4	Modelo Periódico Autorregressivo Gama de Ordem 1	29
4	Simulação de Modelos Não Lineares Gama	34
4.1	Estimação dos Parâmetros	37
4.2	Geração dos resíduos do Modelo PGAR(1)	38
4.3	Geração de Séries Sintéticas	38
5	Resultados	41
5.1	Análise Preliminar das Séries ENAs dos Subsistemas	41
5.2	Simulação de Séries Sintéticas do Modelo PGAR(1) para os Subsistemas	51
5.3	Teste de Diagnóstico	58
5.4	A Modelagem Aditiva e Multiplicativa	71
6	Considerações Finais e Trabalhos futuros	80
7	Apêndice	85

## Lista de figuras

1.1	Problema de Decisão da Operação - Fonte: (ONS, 2015)	13
1.2	Problema de Decisão da Operação - Fonte:(Pereira, et al., 1998)	14
2.1	Filtro no contexto de Séries Temporais - Fonte: (Cavalcanti Netto, M. A., 2015)	16
2.2	Filtro para o modelo e resíduos - - Fonte: (Adaptado de Cavalcanti Netto, M. A., 2015)	16
2.3	Filtros do modelo ARIMA(p,d,q) - Fonte: (Cavalcanti Netto, M. A., 2015)	17
2.4	Etapas da Metodologia Box & Jenkins - Fonte: Oliveira(2010)	18
4.1	Etapas da Modelagem	34
4.2	Série do Subsistema Norte	35
4.3	Boxplot Periódico para o Subsistema Norte	35
4.4	Histograma para o Subsistema Norte	36
4.5	Ruído gerado para o Subsistema Norte	39
4.6	Simulação de 500 séries sintéticas para o Subsistema Norte	39
5.1	Série do Subsistema Sul	42
5.2	Boxplot Periódico para o Subsistema Sul	42
5.3	Série do Subsistema Sudeste/Centro-Oeste	43
5.4	Boxplot Periódico para o Subsistema Sudeste/Centro-Oeste	43
5.5	Série do Subsistema Nordeste	44
5.6	Boxplot Periódico para o Subsistema Nordeste	44
5.7	Histograma para o Subsistema Sul	45
5.8	Histograma para o Subsistema Sudeste/Centro-Oeste	45
5.9	Histograma para o Subsistema Nordeste	46
5.10	Geração dos "Betas" para os Subsistemas	51
5.11	Ruído gerado para o Subsistema Sul	52
5.12	Ruído gerado para o Subsistema Sudeste/Centro-Oeste	52
5.13	Ruído gerado para o Subsistema Nordeste	53
5.14	Ruído gerado para o Subsistema Norte	53
5.15	PGAR no Subsistema Sul	54
5.16	PGAR no Subsistema Sudeste/Centro-Oeste	55
5.17	PGAR no Subsistema Nordeste	55
5.18	PGAR no Subsistema Norte	56
5.19	Desvio-Padrão PGAR no Subsistema Sul	56
5.20	Desvio-Padrão PGAR no Subsistema Sudeste/Centro-Oeste	57
5.21	Desvio-Padrão PGAR no Subsistema Nordeste	57
5.22	Desvio-Padrão PGAR no Subsistema Norte	58
5.23	P-valor do Teste de Levene para o modelo no Subsistema Sul	59
5.24	P-valor do Teste de Levene para o modelo PGAR no Subsistema Sudeste/Centro-Oeste	59
5.25	P-valor do Teste de Levene para o modelo PGAR no Subsistema Nordeste	60

5.26	P-valor do Teste de Levene para o modelo PGAR no Subsistema Norte	60
5.27	Proporção de aceitação do Teste de Mann-Whitney - Subsistema Sul	64
5.28	Proporção de aceitação do Teste de Mann-Whitney - Subsistema Sudeste/Centro-Oeste	64
5.29	Proporção de aceitação do Teste de Mann-Whitney - Subsistema Nordeste	65
5.30	Proporção de aceitação do Teste de Mann-Whitney - Subsistema Norte	65
5.31	Avaliação da restrição do modelo PGAR no Subsistema Sul	66
5.32	Avaliação da restrição do modelo PGAR no Subsistema Sudeste/Centro-Oeste	66
5.33	Avaliação da restrição do modelo PGAR no Subsistema Nordeste	67
5.34	Avaliação da restrição do modelo PGAR no Subsistema Norte	67
5.35	Box-plot da Simulação Sul dos meses 1 à 60	68
5.36	Box-plot da Simulação Sul dos meses 61 à 120	68
5.37	Box-plot da Simulação Sudeste/Centro-Oeste dos meses 1 à 60	69
5.38	Box-plot da Simulação Sudeste/Centro-Oeste dos meses 61 à 120	69
5.39	Box-plot da Simulação Nordeste dos meses 1 à 60	70
5.40	Box-plot da Simulação Nordeste dos meses 61 à 120	70
5.41	Box-plot da Simulação Norte dos meses 1 à 60	71
5.42	Box-plot da Simulação Norte dos meses 61 à 120	71
5.43	Teste de Levene para o PAGAR(1) e PMGAR(1) no subsistema Sul	72
5.44	Teste de Levene para o PAGAR(1) e PMGAR(1) no subsistema Sudeste/Centro-Oeste	72
5.45	Teste de Levene para o PAGAR(1) e PMGAR(1) no subsistema Nordeste	73
5.46	Teste de Levene para o PAGAR(1) e PMGAR(1) no subsistema Norte	73
5.47	Cenário PAGAR(1) no Subsistema Sul	75
5.48	Desvio-Padrão PAGAR(1) no Subsistema Sul	76
5.49	Cenário PAGAR(1) no Subsistema Sudeste/Centro-Oeste	76
5.50	Desvio-Padrão PAGAR(1) no Subsistema Sudeste/Centro-Oeste	77
5.51	Cenário PAGAR(1) no Subsistema Nordeste	77
5.52	Desvio-Padrão PAGAR(1) no Subsistema Nordeste	78
5.53	Cenário PAGAR(1) no Subsistema Norte	78
5.54	Desvio-Padrão PAGAR(1) no Subsistema Norte	79

## Lista de tabelas

4.1	Teste de Kolmogorov-Smirnov para o Subsistema Norte	37
5.1	Teste de Kolmogorov-Smirnov por mês para cada Subsistema	46
5.2	Parâmetros estimados via Método de Momentos	48
5.3	Parâmetros estimados via Método de Máxima Verossimilhança	50
5.4	Intervalo de Confiança (99%) para a Assimetria via Bootstrap no Subsistema Sul	61
5.5	Intervalo de Confiança (99%) para a Assimetria via Bootstrap no Subsistema Sudeste/Centro-Oeste	62
5.6	Intervalo de Confiança (99%) para a Assimetria via Bootstrap no Subsistema Nordeste	62
5.7	Intervalo de Confiança (99%) para a Assimetria via Bootstrap no Subsistema Norte	63
5.8	Proporção de p-valores do Teste de Mann-Whitney superiores a 1% dos Subsistemas para o Modelo Aditivo e Multiplicativo	74

*"Insanidade é continuar fazendo sempre a  
mesma coisa e esperar resultados diferentes"*

**Albert Einstein, .**

# 1

## Introdução

O planejamento da operação energética de um sistema de geração de energia elétrica hidrotérmico é dado através das metas de geração para as usinas hidrelétricas e termelétricas para cada momento ao longo do horizonte de estudo, atendendo à demanda de energia elétrica, às restrições operativas das usinas e às restrições elétricas do sistema. Tal planejamento consiste em uma tarefa árdua por se tratar de um problema acoplado no tempo e espaço e por ser um problema estocástico.

Os reservatórios das usinas hidrelétricas possuem capacidade limitada em seu armazenamento e a decisão de geração em um estágio influencia no nível de armazenamento no final do mesmo estágio, que será o armazenamento inicial do próximo estágio. Logo, há uma dependência entre a decisão operativa de um determinado período com os períodos subsequentes, o que explica o problema de acoplamento no tempo.

O acoplamento no espaço se dá pelo fato do parque hidrelétrico estar disponível em diversas bacias hidrográficas - com várias usinas hidrelétricas dispostas em uma mesma cascata - em que a operação das usinas a montante afeta a disponibilidade de água e, portanto, a operação das usinas a jusante.

O problema estocástico é caracterizado pela existência da incerteza em relação às afluições futuras, que variam sazonalmente e regionalmente e podem possuir períodos secos de longa duração. Além disso, existem ainda as incertezas com relação à demanda de energia elétrica.

A alocação dos recursos hídricos e térmicos é baseada em alguns critérios, dentre eles, a minimização dos custos ótimos globais de produção de energia elétrica. Na figura (Figura 1.1) ilustra-se a relação existente entre uma decisão de despacho tomada no presente e possíveis consequências no futuro. Por exemplo, ao tomar uma decisão de utilizar energia hidrelétrica para atender a demanda presente e, no futuro, ocorrer um baixo nível de afluição, pode ter como consequência o uso da geração térmica de custo elevado ou a suspensão do fornecimento de energia. Por outro lado, se utilizar a geração térmica para atender a demanda presente e no futuro acontecer uma cheia (nível alto de afluição), caracterizará um vertimento, ocorrendo desperdício de energia e

consequentemente um aumento desnecessário do custo de operação.

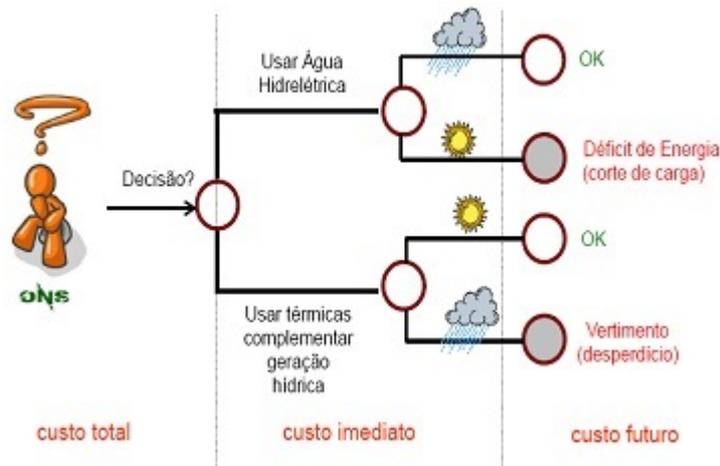


Figura 1.1: Problema de Decisão da Operação - Fonte: (ONS, 2015)

Deste modo, é necessário encontrar um equilíbrio da operação do sistema. Segundo (Kelman, 2001), o sistema é equilibrado quando é possível gerar energia suficiente para atender a demanda, com um risco pequeno de falha. Logo, o equilíbrio está relacionado com a redução do uso da água e da não redução dos reservatórios. A variável de decisão será o volume de água armazenado no final do período da operação, ou seja, o volume final.

Durante a tomada de decisão, é necessário comparar as vantagens entre o uso imediato da água e de seu armazenamento no presente para uso futuro. O uso imediato da água pode ser representado através de uma Função de Custo Imediato (FCI), enquanto o armazenamento no presente para uso futuro pode ser representado pela Função de Custo Futuro (FCF). De acordo com (Pereira et al.,1998) o custo total será a soma destes custos e a decisão ótima é obtida igualando a derivada do custo total em relação ao volume final armazenado a zero, isto é, a decisão será baseada no ponto que minimiza a soma de ambos os custos (Figura 1.2).

$$\frac{\partial(FCI + FCF)}{\partial V} = \frac{\partial FCI}{\partial V} + \frac{\partial FCF}{\partial V} = 0 \Rightarrow \frac{\partial FCI}{\partial V} = -\frac{\partial FCF}{\partial V} \quad (1-1)$$

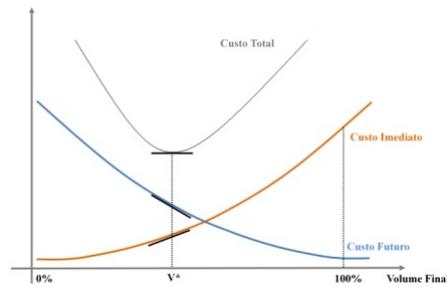


Figura 1.2: Problema de Decisão da Operação - Fonte:(Pereira, et al., 1998)

Com as mudanças do marco regulatório em 1999, o Operador Nacional do Sistema Elétrico (ONS) se tornou o responsável pelo planejamento da operação energética do Sistema Interligado Brasileiro (SIN). A ONS utiliza de uma cadeia de modelos matemáticos desenvolvida pelo Centro de Pesquisas de Energia Elétrica (CEPEL) onde os principais modelos são: DESSEM-PAT, DECOMP e NEWAVE. Esses modelos são responsáveis pela política ótima de operação para os problemas de programação diária (DESEM-PAT), curto prazo (DECOMP) e médio/longo prazo (NEWAVE) baseada no critério de minimização dos custos de operação ao longo do horizonte de planejamento.

Neste contexto, o modelo NEWAVE é adotado pelo Sistema Elétrico Brasileiro (SEB) para o planejamento de horizonte de 5 a 10 anos. Tal modelo baseia-se na técnica PDDE (Programação Dinâmica Dual Estocástica) e tem por objetivo traçar uma estratégia de geração hidráulica e térmica para a minimização do custo total de operação ao longo do horizonte. Para reduzir a dimensionalidade do problema utiliza-se uma representação simplificada do parque gerador, por meio de sistemas equivalentes de energia. Além disso, (Penna, 2009) também descreve que a estocasticidade das aflúências é dada por um vasto número de cenários hidrológicos gerados sinteticamente do modelo adotado pela SEB, conhecido como Modelo Periódico Autorregressivo (PAR(p)).

O presente trabalho tem por objetivo explorar a metodologia descrita por (Fernandez, 1984) e (Fernandez & Salas, 1986), denominado Modelo Periódico Autorregressivo Gama, dentro do contexto de séries de vazões hidrológicas, viabilizando a geração de novas séries sintéticas hidrológicas tão eficientes quanto as encontradas na modelagem consolidada pelo SEB, além de contornar alguns problemas encontrados na modelagem do setor através desta metodologia. Para isto, a dissertação se organizará do seguinte modo:

No capítulo 2 é apresentada ao leitor a metodologia Box & Jenkins, ilustrando o seu mecanismo de funcionamento através de detalhes matemáticos, além de explicitar sua utilidade dentro do SEB por meio de sua modelagem periódica.

O capítulo 3 trata da revisão de literatura da Modelagem Periódica Gama proposta por (Fernandez, 1984) e (Fernandez & Salas, 1986) detalhando através da matemática a sua construção.

Fruto do objetivo do trabalho e do capítulo 3, o capítulo 4 trabalha com simulação destes modelos, detalhando a sua programação para fins acadêmicos.

O capítulo 5 avalia, detalha e compara os resultados obtidos dos modelos propostos para todos os subsistemas por meio de testes estatísticos e visualizações gráficas.

Por fim, o capítulo 6 apresenta as considerações finais e trabalhos futuros com sugestões de possíveis implementações no intuito de otimizar os resultados obtidos.

## 2

### A Metodologia Box e Jenkins

Como o SEB adota o modelo autorregressivo periódico no sistema NEWAVE, neste capítulo é apresentado ao leitor uma breve descrição sobre o modelo em questão, apresentando a sua estrutura matemática, bem como o seu papel dentro do setor.

Suponha que exista um mecanismo que atue como um filtro, o qual é estimulado por uma série de ruídos brancos, resultantes de um processo de geração de números aleatórios, e que com esse estímulo seja gerada pelo sistema uma sequência de valores observados seguindo um padrão, que corresponde à uma série temporal  $Y_t$  (Figura 2.1).



Figura 2.1: Filtro no contexto de Séries Temporais - Fonte: (Cavalcanti Netto, M. A., 2015)

Embora, na prática, o caminho inverso retrate a realidade, ou seja, conhece-se o conjunto de observações sequenciais  $Y_t$  geradas pelo mecanismo em questão, onde busca-se associar um modelo que simule os processos internos ao mecanismo que as gerou. Uma vez que se estabeleça um modelo operacional para essa representação, a série aleatória  $e_t$  corresponde uma sequência aleatória de valores denominados resíduos resultantes da extração entre os valores reais observados de  $Y_t$  com os valores do modelo ajustado do mecanismo em questão (Figura 2.2).

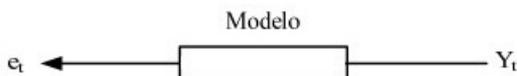


Figura 2.2: Filtro para o modelo e resíduos - - Fonte: (Adaptado de Cavalcanti Netto, M. A., 2015)

A modelagem através de Box & Jenkins tem por objetivo prever valores futuros de uma série temporal por meio de seus valores presentes e passados.

No entanto, esta metodologia só é aplicável para processos estacionários de segunda ordem, isto é, processos que atendam as seguintes condições: média e variância constantes e função de autocovariância dependendo apenas do lag-k.

A metodologia de Box & Jenkins baseia-se no ajuste sequencial de modelos denominados ARIMA(p,d,q) (autoregressivos integrados e de médias móveis) a séries temporais de valores observados de forma que a diferença entre os valores gerados pelo ajuste do modelo com os valores reais observados resulte em uma sequência de resíduos com comportamento aleatório em torno de zero. O referente modelo representa a junção da parte autorregressiva (parâmetro p) com médias móveis (parâmetro q) e adota o parâmetro d como sucessivas diferenciações da série para obtenção de estacionaridade. É importante ressaltar a existência do modelo ARMA, que também adota a junção autorregressiva com médias móveis, no entanto, tal modelo não possui estrutura para diferenciação da série e parte do princípio que o pressuposto de estacionaridade foi atendido.

Os modelos ARIMA pertencem a uma classe de modelos robustos capazes de descrever os processos de geração de uma variedade de séries temporais para os previsores (que correspondem aos filtros) sem considerar a existência de outras séries explicativas que geraram as séries.

É possível transferir a ideia deste modelo para o mecanismo mencionado anteriormente atuante como filtro. Como o modelo ARIMA atua com três parâmetros (ARIMA(p,d,q)), suponha que existam sucessivos filtros (Figura 2.3) aos quais cada um deles associado a um resíduo, representando o seu sistema: o filtro de médias móveis (parâmetro q), o filtro autorregressivo (parâmetro p) e o filtro de integração não-estacionário (parâmetro d).

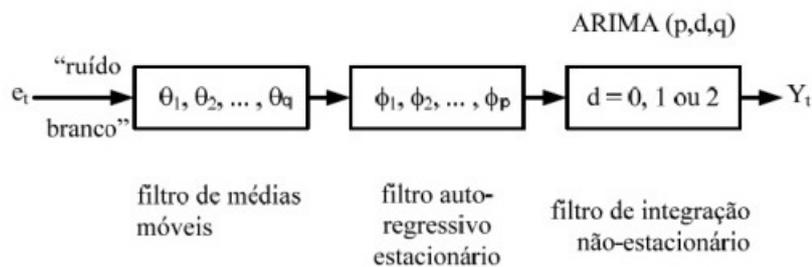


Figura 2.3: Filtros do modelo ARIMA(p,d,q) - Fonte: (Cavalcanti Netto, M. A., 2015)

Portanto, é possível representar o modelo ARIMA(p,d,q) matematicamente do seguinte modo:

$$(1 + \phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p)(1 - B)^d Y_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) e_t \quad (2-1)$$

Onde B representa o operador de retardo, isto é,  $B^j Y_t = Y_{t-j}$ .

A metodologia de Box e Jenkins segue três etapas (Figura 2.4):

Etapa 1. Identificação de possíveis modelos e seus parâmetros

Etapa 2. Estimação dos parâmetros p e q e testes de diagnósticos

Etapa 3. Previsão, simulação ou controle do mecanismo de geração dos valores observados  $Y_t$ .

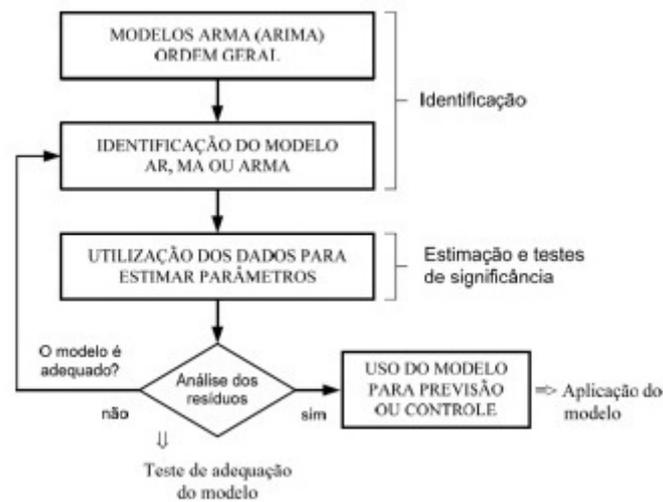


Figura 2.4: Etapas da Metodologia Box & Jenkins - Fonte: Oliveira(2010)

É importante ressaltar que uma maneira de identificar os parâmetros do modelo é através de uma inspeção gráfica das funções de autocorrelação (FAC) e autocorrelação parcial (FACP). Estas funções permitem que se meça a força relativa de interação entre as variáveis  $Y_t$  defasadas.

Esta metodologia serve de arcabouço para que modelos periódicos possam ser estudados e empregados em suas devidas particularidades. (Salas & Obeysekera, 1982) enfatizam que estes modelos de séries temporais aplicados às séries hidrológicas desempenham um importante papel para planejamento e operação de sistemas hidrotérmicos. (Ferreira, 2013) lista diversos modelos que se enquadram neste contexto, entre eles, encontram-se os seguintes modelos: Periódico Autorregressivo de ordem p (PAR(p)), Periódico ARMA (PARMA(p,q)), Periódico Autorregressivo Aditivo Gama de ordem 1 (PAGAR(1)), Periódico Autorregressivo Multiplicativo Gama de Ordem 1 (PMGAR(1)) e, por fim, o Periódico Autorregressivo Gama de Ordem 1 (PGAR(1)), que é a forma híbrida do modelo aditivo e multiplicativo.

O modelo  $PAR(p)$ , desenvolvido por (Thomas & Fiering, 1962) e com contribuições significativas na sua formalização por (Jones & Brelsford, 1967) e (Hipel & McLeod, 1994), ajusta para cada período da série um modelo autorregressivo de ordem  $p$  ( $AR(p)$ ). De modo análogo, o  $PARMA(p,q)$  ajusta um  $ARMA(p,q)$  para cada período de tempo. Entende-se como período de tempo séries disponíveis, por exemplo, em dia, semana ou mês.

Estes modelos são abundantemente estudados na literatura para aperfeiçoamento e aplicações em diversas séries periódicas. Em termos de aplicações, no Brasil, (Maceira, et al., 1999) trabalharam com previsões de vazões semanais baseados em modelos  $ARMA$ , associado parâmetros periódicos e não periódicos.

O Modelo Periódico Autorregressivo de ordem  $p$  ( $PAR(p)$ ) é adotado como ferramenta pelo SEB(NEWAVE) objetivando a geração de cenários de afluências que, posteriormente, serão utilizadas na PDDE para dar origem às funções de custo futuro ((Maceira et al., 2005); (CEPEL, 2011a)). O histórico de vazões e/ou ENAs consiste em uma série temporal cujo interesse do SEB é aproximar seu comportamento estocástico após o ajuste do  $PAR(p)$  para que possa-se gerar novas séries, diferentes da original, mas igualmente possível do ponto de vista estatístico. No entanto, existem problemas nesta abordagem ao tratar de vazões incrementais pequenas, pois corre-se o risco da geração de valores negativos para as vazões. Para contornar este problema, adota-se o uso da distribuição log-normal aos resíduos presentes no modelo (Maceira et al., 2005), uma vez que a mesma assegura-se que apenas valores positivos sejam gerados no ajuste de séries.

Na década de 1980, (Fernandez, 1984) e (Fernandez & Salas, 1986) propuseram modelos de estruturas autorregressivas com parâmetros periódicos que admitem distribuição gama para auxiliar na modelagem e geração de séries sintéticas hidrológicas e evitar que transformações sejam empregadas para cumprir o requisito de normalidade. Esta modelagem será estudada de forma mais detalhada no capítulo a seguir.

O modelo em questão também foi abordado por outros e pelos próprios autores em diferentes anos. Em (Fernandez & Salas, 1990) o modelo gama autorregressivo de ordem 1 ( $GAR(1)$ ) foi sugerido com correções no viés baseado em estudos computacionais aplicados na estimação de parâmetros, enquanto que em (Chebaane et. al, 1995) o  $PGAR$  foi proposto em conjunto com outro modelo conhecido como periódico discreto autorregressivo ( $PDAR$ ) o qual foi denominado  $PGAR$ - $PDAR$ , cujo intuito era modelar séries hidrológicas para regiões secas. Já em (Şarлак & Şorman, 2007) o modelo  $GAR(1)$  foi estudado em séries hidrológicas e suas estimações foram construídas através do método

de máxima verossimilhança modificada.

Existem estudos recentes referentes a modelagem PGAR para o SEB. Em (Ferreira, 2013) o modelo híbrido é abordado como estudo alternativo ao periódico autoregressivo (PAR(p)) em séries de interconfigurações do SEB para a geração de séries sintéticas. No entanto, o autor evidenciou restrições quanto ao uso desta modelagem. Já em (Braga & Calmon, 2016) tem como proposta a comparação dos resultados do PGAR com a modelagens periódicas Normal e Lognormal adotadas pelo setor. Os autores relataram bons resultados do PGAR quando comparado com o modelo periódico Normal. Além disso, também destacaram esta modelagem como uma boa alternativa para o modelo Lognormal.

### 3

## O Modelo Periódico Autorregressivo Gama

Segundo (Ferreira, 2013) séries sintéticas de vazão/ENA são utilizadas com frequência em estudos de planejamento e operação do sistema hidroelétrico. Por isso, uma grande quantidade de modelos são estudados usando como base o modelo periódico autorregressivo.

De acordo (Fernandez & Salas, 1986) séries de vazões anuais podem ser aproximadas para distribuições normais, porém em períodos de tempo curtos como diário, semanal e mensal esta característica não é observada, especialmente pelo problema de assimetria. Para contornar esta dificuldade, diferentes métodos são abordados na literatura na tentativa de uma aproximação para a distribuição normal, mas nenhum deles apresenta sucesso. Grande parte desses modelos possuem estruturas autorregressivas de ordem 1 constante para toda série ou com parâmetros periódicos e assumem distribuição normal para os resíduos.

(Fernandez & Salas, 1986) propuseram uma modelagem capaz de tratar as limitações acima mencionadas adotando a distribuição Gama e estrutura autorregressiva. Tanto o modelo linear quanto o não linear são adotados como extensão para acomodar o comportamento sazonal da distribuição marginal e da função de autocorrelação. Basicamente o procedimento para a modelagem se resumirá na estimação dos parâmetros periódicos e geração de resíduos.

Os modelos envolvendo distribuição gama lineares e não lineares possuem restrições, mas são complementares no tratamento do coeficiente periódico de assimetria. O modelo proposto apresentará estrutura autoregressiva, além de ser periódico e não linear. A vantagem deste método é que ele pode ser aplicado diretamente em séries periódicas assimétricas, como é o caso das séries de energia.

A seguir serão apresentados os modelos propostos por (Fernandez, 1984) e (Fernandez & Salas, 1986), que serão utilizados neste trabalho na tentativa de propor um método de modelagem alternativa para geração de cenários do SEB. Os modelos são denominados por Periódico Gama Aditivo Autorregressivo (PAGAR), Periódico Gama Multiplicativo Autorregressivo (PMGAR) e Periódico Gama Autorregressivo (PGAR). Todos são processos autorregressi-

vos de ordem 1.

### 3.1

#### O Modelo Periódico Autorregressivo Aditivo Gama de Ordem 1

Considere o processo periódico Autorregressivo de ordem 1:

$$X_{v,t} = \phi_t X_{v,t-1} + \varepsilon_{v,t} \quad (3-1)$$

Onde:

- $X_{v,t}$  : Série temporal do período  $t$  no ano  $v$  com distribuição  $Gama(\alpha_t, \beta_t, \lambda_t)$
- $\varepsilon_{v,t}$  : Série de resíduo do período  $t$  no ano  $v$ .

Partindo do pressuposto de estacionariedade cíclica (estacionariedade periódica) com distribuição gama, o modelo proposto em (3-1) representa o PAGAR(1) e sua função de densidade é dada por:

$$f_x(x) = \{\alpha_t \Gamma(\beta_t)\}^{-1} \{(x - \lambda_t)/\alpha_t\}^{\beta_t-1} \exp\{-(x - \lambda_t)/\alpha_t\} \quad (3-2)$$

Os parâmetros  $\alpha_t$ ,  $\beta_t$  e  $\lambda_t$  são denominados, respectivamente, por escala, forma e deslocamento, onde  $\alpha_t > 0$ ,  $\beta_t > 0$  e  $x > \lambda_t$ .

Segundo (Fernandez & Salas, 1986), para definir o modelo (3-1) é necessário encontrar a distribuição para o resíduo  $\varepsilon_{v,t}$ , condicionado que  $X_{v,t}$  representa uma distribuição gama. Para isso, o autor recorre a transformação matemática de Laplace-Stieltjes (Lawrence & Lewis, 1981):

$$LS_X(s) = E[\exp\{-sx\}] = \int_0^\infty e^{-sx} f_X(x) dx \quad (3-3)$$

Admitindo independência dos resíduos  $\varepsilon_{v,t}$  para todo o procedimento  $X_{v,t}$  e supondo que o processo é periodicamente ergótico, (Fernandez & Salas, 1986) simplificaram a notação de  $\varepsilon_{v,t}$  omitindo ano  $v$ .

Utilizando a definição (3-3) para a série de ruídos  $\varepsilon_t$  e fazendo manipulações matemáticas, chega-se ao resultado (3-4). Para mais detalhes matemáticos, ver (Fernandez, 1984).

$$LS_{\varepsilon_t}(s) = e^{(\lambda_t - \phi_t \lambda_{t-1})s} \left( \frac{1}{1 + \alpha_t s} \right)^{\beta_t - \beta_{t-1}} \left( \phi_t \frac{\alpha_{t-1}}{\alpha_t} \right) \quad (3-4)$$

$$+ [1 - \phi_t \frac{\alpha_{t-1}}{\alpha_t}] \frac{1}{1 + \alpha_t s})^{\beta_{t-1}}$$

Para obter a inversa de (3-4), (Fernandez & Salas, 1986) baseiam-se nas seguintes restrições:

$$\beta_t \geq \beta_{t-1} > 0 \tag{3-5a}$$

$$0 \leq \phi_t \frac{\alpha_{t-1}}{\alpha_t} \leq 1 \tag{3-5b}$$

Nota-se, portanto, que o modelo aditivo adota como restrição o aumento do parâmetro  $\beta_t$  ao longo do período, enquanto o parâmetro periódico,  $\phi_t$ , limita-se ao intervalo de valores positivos, mas inferiores à razão do parâmetro  $\alpha_t$  em relação ao estágio anterior.

Aplicando as propriedades da transformação de Laplace-Stieltjes detalhadas em (Fernandez, 1984),  $\varepsilon_t$  pode ser representado pela soma de duas variáveis aleatórias independentes:

$$\varepsilon_t = \varepsilon_t(0) + \varepsilon_t(1) \tag{3-6}$$

Neste caso,  $\varepsilon_t(0)$  segue uma distribuição gama de 3 parâmetros, com parâmetro deslocamento  $\lambda_t - \phi_t \lambda_{t-1}$ , escalar  $\alpha_t$  e forma  $\beta_t - \beta_{t-1}$ , enquanto  $\varepsilon_t(1)$  é denominada como variável inovação e é definida por:

$$\varepsilon_t(1) = \begin{cases} 0, & \text{se } N = 0, \\ \sum_{m=1}^N Y_m (\phi_t \frac{\alpha_{t-1}}{\alpha_t})^{U_m}, & \text{se } N > 0. \end{cases} \tag{3-7}$$

As seqüências  $U_1, \dots, U_N$  e  $Y_1, \dots, Y_N$  são, respectivamente, variáveis aleatórias independentes e identicamente distribuídas com distribuição uniforme(0,1) e exponencial( $\alpha_t$ ). Além disso, N representa uma variável aleatória Poisson de valor esperado  $\theta$  dado por:

$$\theta = -\beta_{t-1} \ln \left\{ \phi_t \frac{\alpha_{t-1}}{\alpha_t} \right\} \tag{3-8}$$

A estrutura de autocorrelação do modelo PAGAR(1) é dada por:

$$\rho_{k,t} = \frac{\sigma_{t-k}}{\sigma_t} \prod_{j=0}^{k-1} \phi_{t-j} \quad k = 1, 2, \dots \quad (3-9)$$

onde  $\rho_{k,t}$  representa a autocorrelação entre  $X_{v,t}$  e  $X_{v,t-k}$ , e  $\sigma_t^2$  a variância de  $X_{v,t}$ . Em particular, para  $k=1$ :

$$\rho_{1,t} = \phi_t \frac{\sigma_{t-1}}{\sigma_t} \quad (3-10)$$

É possível notar que o coeficiente de autocorrelação aparece como função do parâmetro periódico do modelo, tal como o seu desvio-padrão. Dispondo-se da variância e coeficiente de assimetria do processo em relação aos parâmetros do modelo, as restrições dadas em (3-5a) e (3-5b) são reduzidas para:

$$\rho_{1,t}^2 \leq \frac{\beta_{t-1}}{\beta_t} \leq 1 \quad (3-11a)$$

$$\rho_{1,t} \leq \frac{\gamma_t}{\gamma_{t-1}} \leq 1 \quad (3-11b)$$

Onde  $\gamma_t$  representa o coeficiente de assimetria.

As equações (3-11a) e (3-11b) representam a região admissível do modelo em função do coeficiente de correlação periódica de lag-1. Analisando a restrição imposta em (3-5a) e (3-5b), pode-se interpretar que o parâmetro  $\beta_t$  segue dois caminhos distintos: ser constante ou aumentar ao longo do período. Para o caso de  $\beta_t$  constante, é possível mostrar que  $\varepsilon_t(0)$  é igual a  $\lambda_t - \phi\lambda_{t-1}$ . Neste caso, as restrições em (3-10) e (3-11b) conduzem para  $0 \leq \rho_{1,t} \leq 1$  viabilizando que modelo PAGAR(1) tenha média, variância e autocorrelação periódica de lag-1, mas assimetria constante.

Além do mais, um modelo mais simples surge ao assumir que  $\beta_t = 1$  para todo  $t$ , onde a distribuição marginal é exponencial com parâmetro de deslocamento  $\lambda_t$  e escalar  $\alpha_t$ . O modelo PAGAR(1) por si só admite distribuição marginal gama ou exponencial e assume que todos os parâmetros são periódicos, com exceção de  $\beta$ . Em (Fernandez, 1984) encontra-se o uso desta modelagem dentro do contexto de séries hidrológicas ao supor que a assimetria não varia de mês para mês, isto é, são iguais

durante o ano.

Por outro lado, ao considerar que  $\beta_t$  é variante no tempo, por exemplo, ao longo dos meses, é possível notar que a restrição (3-5a) implica em  $\beta_\omega > \beta_{\omega-1} > \dots > \beta_2 > \beta_1$ , sendo que a condição  $\beta_1 > \beta_\omega$  não é estabelecida, acarretando na perda da correlação entre  $X_{v,1}$  e  $X_{v-1,\omega}$ .

Ademais, o modelo PAGAR(1), sob a restrição (3-5a), apresenta uma sequência gama de média, variância e correlação periódica com assimetria decrescente. Como a correlação entre o último mês de um determinado ano com o primeiro mês do ano seguinte é perdida, as sequências anuais derivadas de cada um dos modelos são descorrelatadas.

Desse modo, o modelo PAGAR(1) é somente aplicável para um processo gama de parâmetro  $\beta$  constante ao longo dos anos, implicando em assimetria constante, ao passo que os demais momentos amostrais são constantes ou periódicos.

### 3.1.1

#### O Modelo PAGAR(1) com Marginais Exponenciais

Um caso particular pode ser concebido ao assumir que a distribuição marginal de  $X_{t,v}$  segue uma exponencial de parâmetros  $\lambda_t$  e  $\alpha_t$ . Como a distribuição exponencial pode ser reescrita como distribuição gama, o parâmetro  $\beta$  será constante e igual a 1. Neste caso, o inverso da transformação (3-4) será dado por:

$$\varepsilon_t = \lambda_t - \phi_t \lambda_{t-1} + \varepsilon_t(1) \quad (3-12)$$

Aqui, a variável  $\varepsilon_t(1)$  é uma combinação de exponenciais dada por:

$$\varepsilon_t(1) = \begin{cases} 0, & \text{com probabilidade } \phi_t \frac{\sigma_{t-1}}{\sigma_t}, \\ Ex(\alpha_t), & \text{com probabilidade } 1 - \phi_t \frac{\sigma_{t-1}}{\sigma_t}. \end{cases} \quad (3-13)$$

Onde  $Ex(\alpha)$  representa uma variável aleatória exponencial de média  $\frac{1}{\alpha}$ .

Este tipo de modelagem pode ser útil para todo o instante de tempo com a única restrição do coeficiente de correlação ser positivo.

## 3.2

### Modelo Periódico Autorregressivo Multiplicativo Gama de Ordem 1

Apesar do processo autorregressivo multiplicativo ser não linear, este tipo de modelagem é utilizado de maneira implícita em hidrologia quando uma transformação logarítmica é feita na série temporal, ou seja, ao usar este tipo de transformação em um modelo linear gaussiano as variáveis assumem comportamento nor-

mal, então, implicitamente, um processo autorregressivo multiplicativo pode ser pensado como um processo linear de domínio log (Ferreira, 2013). No entanto, (Fernandez, 1984) ressalta que esse processo é limitado, uma vez que não há representação matemática formal do processo de dependência, principalmente 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 de ordem 1 (PM-GAR(1)) é dado por:

$$X_{v,t} = X_{v,t-1}^{\delta_t} V_{v,t} \quad t = 1, 2, \dots, \omega \quad (3-14)$$

Onde:

- $X_{v,t}$  representa a variável no mês  $t$  e ano  $v$ ;
- $\delta_{v,t}$  representa o parâmetro periódico do processo autorregressivo;
- $V_{t,v}$  representa a série de resíduos no mês  $t$  e ano  $v$ ;

Sem perda de generalidade, a distribuição marginal de  $X_{v,t}$  é Gama de parâmetros periódicos  $\alpha_t$  e  $\beta_t$ .

Com a finalidade de encontrar uma distribuição para os resíduos do modelo (3-14), a transformação de Mellin torna-se uma ferramenta bastante útil (Griffin, 1975):

$$M_X(s) = E[X^{\{s-1\}}] = \int_0^\infty x^{s-1} f(x) dx \quad (3-15)$$

Adotando a transformação de Mellin no contexto da série de resíduos  $V_{v,t}$ , a expressão (3-15) é obtida em (3-16). Todos os detalhes matemáticos podem ser encontrados em (Fernandez, 1984).

$$M_{V_t}(s) = E(V_t^{s-1}) = \left[ \frac{\alpha_t}{(\alpha_{t-1})^{\delta_t}} \right]^{s-1} \frac{\Gamma(\beta_{t-1})\Gamma(\beta_t + s - 1)}{\Gamma(\beta_t)\Gamma(\beta_{t-1} + \delta_t s - \delta_t)} \quad (3-16)$$

Simplificando a notação de  $V_{v,t}$  por  $V_t$ . Para encontrar a inversa de (3-16) é necessário assumir a seguinte restrição:

$$0 < \beta_t \leq \beta_{t-1} \quad (3-17a)$$

$$0 \leq \delta_t \leq 1 \quad (3-17b)$$

Portanto, o ruído  $V_t$  pode ser escrito como:

$$V_t = \frac{\alpha_t}{(\alpha_{t-1})^{\delta_{t-1}}} Z_t Y_t^{\delta_t} \quad (3-18)$$

onde  $Y_t$  é uma variável aleatória de distribuição Beta com parâmetros  $\beta_t$  e  $\beta_{t-1} - \beta_t$  e  $Z_t$  corresponde ao ruído do processo multiplicativo gama estacionário

definido por (MacKenzie, 1982). A variável  $Z_t$  pode ser simulada usando as propriedades de decomposição do logaritmo de uma variável gama conforme relatado por (Shanbhag e Srechari, 1977). Através destas propriedades, (Fernandez, 1984) mostrou que  $Z_t$  pode ser simulado por:

$$Z_t = \exp\{V'_t + \lim_{K \rightarrow \infty} \sum_1^K V_{t,k}\} \quad (3-19)$$

Onde:

$$V'_t = -\psi(1 - \delta_t) \quad \text{com probabilidade } \delta_t \quad (3-20a)$$

$$V'_t = -\psi(1 - \delta_t) - Ex(\beta_t) \quad \text{com probabilidade } (1 - \delta_t) \quad (3-20b)$$

$$V_{t,k} = \frac{(1 - \delta_t)}{k} \quad \text{com probabilidade } \delta_t \quad (3-20c)$$

$$V_{t,k} = \frac{(1 - \delta_t)}{k} - Ex(\beta_t + k) \quad \text{com probabilidade } (1 - \delta_t) \quad (3-20d)$$

Onde  $Ex(a)$  representa uma distribuição exponencial de valor esperado  $\frac{1}{a}$  e  $\psi$  a constante de Euler. De acordo com (Fernandez, 1984) a série representada por (3-20a) a (3-20d) converge quase certamente e, na prática, não é necessário um valor de  $K$  tão grande para obtenção de uma boa representação para  $Z_t$ .

A função de autocorrelação do modelo PMGAR(1), demonstrada por (Fernandez e Salas, 1986), pode ser escrita como:

$$\rho_{k,t} = \left(\frac{\beta_t}{\beta_{t-k}}\right)^{1/2} \prod_{j=0}^{k-1} \delta_{t-j} \quad k=1,2,\dots \quad (3-21)$$

Em particular, para  $k=1$ :

$$\rho_{1,t} = \left(\frac{\beta_t}{\beta_{t-1}}\right)^{1/2} \delta_t \quad (3-22)$$

Então, a função de autocorrelação do modelo PMGAR(1) é similar ao processo autorregressivo gaussiano de ordem 1 e também ao que foi obtido ao PAGAR(1).

Assumindo que os parâmetros do modelo são escritos em termos do coeficiente de assimetria periódico e da função de autocorrelação de lag-1, as restrições (3-17a) e (3-17b) podem ser reescritas como:

$$1 \leq \frac{\beta_{t-1}}{\beta_t} \leq \frac{1}{\rho_{1,t}^2} \quad (3-23a)$$

$$1 \leq \frac{\gamma_t}{\gamma_{t-1}} \leq \frac{1}{\rho_{1,t}} \quad (3-23b)$$

Estas restrições definem o espaço paramétrico do modelo em função do coeficiente de autocorrelação de lag-1.

Segundo (Fernandez & Salas, 1986), há duas implicações relacionadas às restrições:  $\beta_t$  pode ser constante ou variante no tempo. Se  $\beta_t$  for constante, a transformação de Mellin (3-16) e a restrição (3-17a) conduzem para a restrição  $0 \leq \rho_{1,t} \leq 1$ , apresentando ao modelo PMGAR(1) média, variância e autocorrelação periódica, mas assimetria constante.

Caso  $\beta_t$  não seja constante, a restrição (3-17a) implica na condição  $\beta_\omega < \beta_{\omega-1} < \dots < \beta_2 < \beta_1$ , mas não em  $\beta_1 < \beta_\omega$ , fazendo com que a correlação entre  $X_{v,1}$  e  $X_{v-1,\omega}$  seja perdida, o que implica que as mesmas são decorrelatadas. Portanto, o modelo PMGAR(1) sob a restrição (3-17a), apresenta uma sequência gama de média, variância e correlação periódica e assimetria decrescente.

Ademais, ao considerar a variação de  $\beta_t$  e levando em consideração que as restrições impostas no PAGAR(1) e PMGAR(1) são opostas, (Fernandez & Salas, 1986) estabeleceram um modelo híbrido capaz de capturar as vantagens das duas metodologias, conhecido como Periódico Gama Autorregressivo de ordem 1 (PGAR(1)), que será detalhado posteriormente. Segundo os autores, o PGAR(1) aparenta ser o mais representativo para fenômenos geofísicos, uma vez que demonstra ser capaz de reproduzir um processo gama com média, variância, autocorrelação e assimetria periódica.

### 3.3

#### Casos particulares para a modelagem PMGAR(1)

Segundo (Fernandez, 1984), dois casos particulares para a modelagem multiplicativa podem ser sugeridos. O primeiro refere-se ao parâmetro  $\beta_t$  ser constante para todo período, enquanto o segundo especifica um processo periódico autorregressivo assumindo distribuição marginal exponencial. Como a segunda abordagem representa um caso particular da primeira, a sua estrutura matemática será apresentada mais detalhadamente.

#### 3.3.1

##### O Modelo PMGAR(1) com Marginais Exponenciais

Assumindo que a distribuição marginal da série  $X_{v,t}$  tenha distribuição exponencial de parâmetro  $\alpha_t$ , a transformação de Mellin, descrita em (3-16), pode ser simplificada para:

$$M_{V_t}(s) = \left[ \frac{\alpha_t}{(\alpha_{t-1})^{\delta_t}} \right]^{s-1} \frac{\Gamma(s)}{\Gamma(\delta_t s + 1 - \delta_t)} \quad (3-24)$$

Logo, o resíduo  $V_t$  é dado por:

$$V_t = \frac{\alpha_t}{\alpha_{t-1}^{\delta_t}} Z_t \quad (3-25)$$

(Fernandez, 1984) destaca que este caso particular pode representar uma versão periódica para o modelo original exponencial dado por (MaKenzie, 1982).

### 3.4

#### Modelo Periódico Autorregressivo Gama de Ordem 1

Seja  $Q_{v,t}$  um processo periódico autorregressivo assimétrico, com distribuição marginal Gama de parâmetros  $\alpha_t$ ,  $\beta_t$  e  $\lambda_t$  conforme apresentado em (3-2), então:

$$X_{v,t} = Q_{v,t} - \lambda_t \quad t = 1, 2, \dots, \omega \quad v = 1, 2, \dots, N \quad (3-26)$$

Tem distribuição Gama de parâmetros  $\alpha_t$  e  $\beta_t$ . Para representar o processo  $X_{t,v}$ , propõe-se o seguinte modelo:

$$Q_{v,t} = \lambda_t + \phi_t X_{v,t-1} + (X_{v,t-1})^{\delta_t} W_{v,t} \quad (3-27)$$

Onde  $X_{v,0} = X_{v-1,\omega}$  corresponde ao período  $\omega$  do ano  $v - 1$ ,  $\phi_t$  e  $\delta_t$  são coeficientes periódicos autorregressivos e  $W_{v,t}$  o resíduo do processo. O modelo (3-27) é não linear e não gaussiano e representa uma forma híbrida entre os modelos PAGAR(1) e PMGAR(1), denominada PGAR(1) por (Fernandez & Salas, 1986).

Para utilizar as propriedades de ambos os modelos, os parâmetros  $\phi_t$  e  $\delta_t$  são definidos em função dos parâmetros da distribuição marginal Gama,  $\alpha_t$  e  $\beta_t$ , e da autocorrelação de lag-1  $\rho_{1,t}$ :

$$\phi_t = 0 \quad \beta_t < \beta_{t-1} \quad (3-28a)$$

$$\phi_t = \rho_{1,t} \frac{\alpha_t}{\alpha_{t-1}} \left( \frac{\beta_t}{\beta_{t-1}} \right)^{1/2} \quad \beta_t \geq \beta_{t-1} \quad (3-28b)$$

$$\delta_t = \rho_{1,t} \left( \frac{\beta_{t-1}}{\beta_t} \right)^{1/2} \quad \beta_t < \beta_{t-1} \quad (3-28c)$$

$$\delta_t = 0 \quad \beta_t \geq \beta_{t-1} \quad (3-28d)$$

Utilizando o mesmo raciocínio dos modelos descritos nas seções anteriores,  $W_{v,t}$  representará o resíduo do modelo aditivo ou multiplicativo, dependendo apenas do valor relativo do parâmetro periódico  $\beta_t$ .

$$W_{v,t} = \begin{cases} V_t & \text{dado por (3-18)} & \beta_t < \beta_{t-1} \\ \varepsilon_t & \text{dado por (3-6)} & \beta_t \geq \beta_{t-1} \end{cases} \quad (3-29)$$

Quanto às restrições dos parâmetros do modelo, obtidas em (3-11) e (3-23), têm-se:

$$0 < \delta_t < 1 \quad \beta_t < \beta_{t-1} \quad (3-30a)$$

$$0 < \phi_t \frac{\alpha_{t-1}}{\alpha_t} < 1 \quad \beta_t \geq \beta_{t-1} \quad (3-30b)$$

Expressando os valores  $\phi_t$  e  $\delta_t$  baseado nas equações em (3-29) e (3-30), respectivamente, as restrições são simplificadas para as seguintes equações:

$$\rho_{1,t}^2 \leq \frac{\beta_{t-1}}{\beta_t} \leq \rho_{1,t}^{1/2} \quad (3-31a)$$

$$\rho_{1,t} \leq \frac{\gamma_t}{\gamma_{t-1}} \leq \frac{1}{\rho_{1,t}} \quad (3-31b)$$

Segundo (Fernandez e Salas, 1986), a expressão (3-31) define o espaço paramétrico do modelo PGAR(1) em função da autocorrelação  $\rho_{1,t}$  assumindo valores positivos. Não existem outras restrições aos parâmetros de deslocamento e escala além dos impostos da distribuição marginal Gama. Portanto, o modelo PGAR(1) mantém a estrutura autorregressiva, permitindo que a média, variância, assimetria e autocorrelação sejam periódicas.

### 3.4.1

#### Estimação dos parâmetros

O conjunto de parâmetros descritos no modelo PGAR(1) descrito acima é  $\{\lambda_t, \alpha_t, \beta_t, \phi_t, \delta_t\}$ , que envolve parâmetros relacionados à distribuição marginal e parâmetros autorregressivos aditivos e multiplicativos. Como (3-28) é utilizado para estimar os parâmetros  $\phi_t$  e  $\delta_t$ , então o conjunto de parâmetros reduz-se para  $\{\lambda_t, \alpha_t, \beta_t, \rho_{1,t}\}$ , uma vez que os parâmetros periódicos são descritos em função deste conjunto.

Os parâmetros  $\lambda_t, \alpha_t, \beta_t$  fazem parte da distribuição marginal Gama de  $Q_{v,t}$ . Uma vez que os valores da série  $Q_{v,t} = \{Q_{1,t}, Q_{2,t}, \dots, Q_{N,t}\}$  estão separados por ano, então supõe-se que são aproximadamente não correlacionados (Fernandez & Salas, 1986). Logo, é possível estimar os parâmetros periódicos pelo método de Máxima Verossimilhança(MV)(Casella & Berger, 2001). No entanto, (Kite 1977), (Choi & Wette, 1969), (Salas & Smith,1981) relatam que o método não tem solução quando o parâmetro  $\beta_t$  é menor do que 1, o que implica que o coeficiente de assimetria não pode ultrapassar 2.

De acordo com (Ferreira, 2013), ao tratar de séries de vazão/ENA não é incomum encontrar séries com assimetria maior que 2. Seguindo a mesma lógica dos modelos anteriores, o método dos momentos é adotado como solução para este problema, mesmo que, em geral, não tenham propriedades tão boas quanto o estimador de MV. Os estimadores dos parâmetros por método dos momentos, descrito em (Fernandez & Salas, 1986), são:

$$\hat{\lambda}_t = \bar{Q}_t - 2 \left( \frac{S_t}{g_t} \right) \quad (3-32a)$$

$$\hat{\alpha}_t = \frac{S_t g_t}{2} \quad (3-32b)$$

$$\hat{\beta}_t = \frac{4}{g_t^2} \quad (3-32c)$$

$$\hat{\rho}_{1,t} = r_{1,t} \quad (3-32d)$$

Onde  $\bar{Q}_t$ ,  $S_t$ ,  $g_t$  e  $r_{1,t}$  correspondem, respectivamente, a média amostral, desvio padrão, assimetria e coeficiente de autocorrelação de lag-1.

### 3.4.2 Previsão com o modelo PGAR(1)

Um conceito importante dentro de séries hidrológicas refere-se ao conhecimento dos fenômenos futuros baseado em seu passado. Para estimar um valor futuro, calcula-se o valor esperado da variável de interesse condicionado com o seu acontecimento passado, ou seja, acredita-se que informações passadas são capazes de descrever o futuro baseado no comportamento temporal padrão de uma série. Dadas as propriedades Markovianas de um modelo Gama, (Fernandez, 1984) descreve a seguinte expressão:

$$E(Q_{v,t+1} | Q_{v,t} = q, Q_{v,t-1} = q_1, \dots, Q_{v,t-n} = q_n, \dots) = E(Q_{v,t+1} | Q_{v,t} = q) \quad (3-33)$$

Assume-se que o modelo foi ajustado e todos seus parâmetros conhecidos. De (3-26) e (3-27), segue-se o seguinte processo de previsão:

$$Q_{v,t+1} = \lambda_{t+1} + \phi_{t+1}(Q_{v,t} - \lambda_t) + (Q_{v,t} - \lambda_t)^{\delta_{t+1}} V_{t+1} \quad (3-34)$$

Para propor um intervalo de confiança, necessita-se do conhecimento da distribuição do resíduo  $V_{t+1}$ , no entanto, tal distribuição não é desenvolvida para este modelo. Em vista disso, (Fernandez, 1984) enfatiza que o processo residual, descrito neste capítulo, torna-se útil para as estimativas dos valores futuros. Portanto, assume-se que uma função de distribuição pode ser atribuída ao condicionar o valor de interesse com o seu passado, isto é:

$$\begin{aligned} E(Q_{v,t+1} | Q_{v,t} = q, \delta_{t+1} = 0) &= \\ &= \lambda_{t+1} + \phi_{t+1} + \alpha_{t+1} \beta_{t+1} - \phi_{t+1} (\lambda_t + \alpha_t \beta_t - q) \end{aligned} \quad (3-35)$$

A expressão (3-35) refere-se à previsão um passo a frente do modelo aditivo. Partindo desta lógica, (Fernandez,1984) menciona que a variância e assimetria podem ser representadas, respectivamente, por:

$$\begin{aligned} Var(Q_{v,t+1}|Q_{v,t} = q, \delta_{t+1} = 0) &= Var(\lambda_{t+1} + \phi_{t+1}(q - \lambda) + V_{t+1}) \\ &= Var(V_{t+1}) = \sigma_{t+1}^2 \beta_{t+1} - \phi_{t+1}^2 \sigma_t^2 \beta_t \end{aligned} \quad (3-36)$$

$$\begin{aligned} E[(Q_{v,t+1}|Q_{v,t} = q, \delta_{t+1} = 0) - E[Q_{v,t+1}|Q_{v,t} = q, \delta_{t+1} = 0)]^3 \\ = E\{[V_{t+1} - E(V_{t+1})]^3\} = \gamma_{V_{t+1}} \sigma_{V_{t+1}}^{3/2} \end{aligned} \quad (3-37)$$

Onde:

$$\gamma_{\hat{Q}_{v,t+1}} = \frac{2(\sigma_{t+1}^3 \beta_{t+1} - \phi_{t+1}^3 \sigma_t^3 \beta_t)}{(\sigma_{t+1}^2 \beta_{t+1} - \phi_{t+1}^2 \sigma_t^2 \beta_t)^{3/2}}$$

Para encontrar as equações do modelo multiplicativo, basta assumir em (3-34) que  $\phi_{t+1}$  é igual a 0. Então a média, variância e assimetria são dadas, respectivamente, por:

$$\begin{aligned} E(Q_{v,t+1}|Q_{v,t} = q, \phi_{t+1} = 0) &= \\ = \lambda_{t+1} + \left(\frac{q - \lambda_t}{\alpha_t}\right)^{\delta_{t+1}} \frac{\alpha_{t+1} \beta_{t+1} \Gamma(\beta_t)}{\Gamma(\beta_t + \delta_{t+1})} \end{aligned} \quad (3-38)$$

$$\begin{aligned} Var(Q_{v,t+1}|Q_{v,t} = q, \phi_{t+1} = 0) &= Var(\lambda_{t+1} + (q - \lambda_t)^{\delta_{t+1}} V_{t+1}) \\ &= \left[\alpha_{t+1} \left(\frac{q - \lambda_t}{\alpha_t}\right)\right]^2 \beta_{t+1} \Gamma(\beta_t) \left[\frac{1 + \beta_{t+1}}{\Gamma(\beta_t + 2\delta_{t+1})} - \frac{\beta_{t+1} \Gamma(\beta_t)}{\Gamma(\beta_t + \delta_{t+1})}\right] \end{aligned} \quad (3-39)$$

$$E[(Q_{v,t+1}|Q_{v,t} = q, \phi_{t+1} = 0) - E[Q_{v,t+1}|Q_{v,t} = q, \phi_{t+1} = 0)]^3$$

$$= (q - \lambda_t)^{3\delta_{t+1}} \gamma_{V_{t+1}} \quad (3-40)$$

Onde:

$$\gamma_{\hat{Q}_{v,t+1}} = \left[ \alpha_{t+1} \left( \frac{q - \lambda_t}{\alpha_t} \right) \right]^3 \beta_{t+1} \Gamma(\beta_t) \left[ \frac{(1 + \beta_{t+1})(2 + \beta_{t+1})}{\Gamma(\beta_t + 3\delta_{t+1})} - \frac{3\beta_{t+1}(1 + \beta_{t+1})(\Gamma(\beta_t))}{\Gamma(\beta_t + \delta_{t+1})\Gamma(\beta_t + 2\delta_{t+1})} + \frac{2\beta_{t+1}^2 \Gamma^2(\beta_t)}{\Gamma^3(\beta_t + \delta_{t+1})} \right] \sigma_{\hat{Q}_{v,t+1}}^{-3/2}$$

É importante ressaltar que as equações mencionadas nesta seção foram construídas por meio de manipulações matemáticas baseadas em (3-4) e (3-16). Para mais detalhes, ver (Fernandez, 1984).

## 4

### Simulação de Modelos Não Lineares Gama

Um dos grandes desafios da modelagem proposta por (Fernandez & Salas, 1986) encontra-se na criação do código de programação, uma vez que esta metodologia não está disponível em qualquer software atual. Devido a isso, faz-se necessário escolher uma linguagem para implementar um programa para suprir a falta de um código de programação para metodologia proposta pelos autores. Por se tratar de uma proposta com ênfase em estatística, optou-se pela linguagem de programação R. Além de ser uma linguagem bastante rotineira no meio acadêmico, o software conta com uma gama de pacotes auxiliares que facilitam e otimizam a programação. Além disso, os procedimentos encontrados neste capítulo facilitam a transição para que outras linguagens de programação sejam utilizadas.

Fazendo uma análise conjunta da metodologia destacada por (Fernandez & Salas, 1986) e do objetivo do presente trabalho, pode-se separar o modelo nas seguintes etapas:

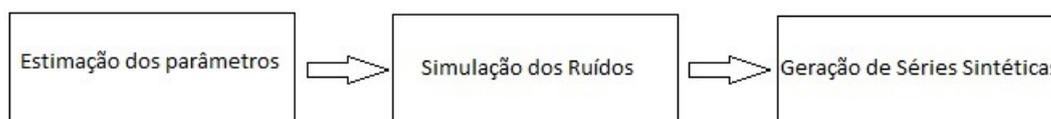


Figura 4.1: Etapas da Modelagem

Considerando os subsistemas do Setor Elétrico Brasileiro, decidiu-se adotar o subsistema Norte como forma de exemplificar o funcionamento da presente metodologia.

As séries históricas de energia estão disponibilizadas no site da ONS e compreendem o período de 1931 à 2013. No intuito de conhecer o comportamento da série temporal do subsistema Norte, optou-se, inicialmente, por uma análise mais exploratória (Apêndice, Bloco 1).

A função encarregada de gerar os boxplots periódicos provém do pacote "pear" (McLeod, 1994) e é denominada "peboxplot". Este pacote é bastante conhecido no contexto de séries periódicas.

Analisando as figuras (Figura 4.2) e (Figura 4.3) pode-se observar uma incerteza durante o período úmido de Dezembro à Maio, isto é, a variabilidade

da série de vazões nesse período é maior comparada com as estações de período seco. Além disso, cabe ressaltar a presença de diversos valores extremos (outliers) presentes em todos os períodos do subsistema, o que a priori comprova a ideia de assimetria para os períodos.

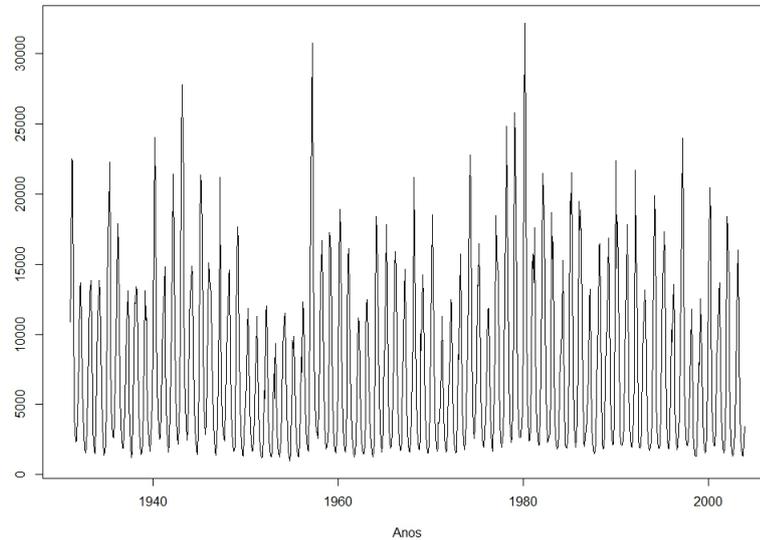


Figura 4.2: Série do Subsistema Norte

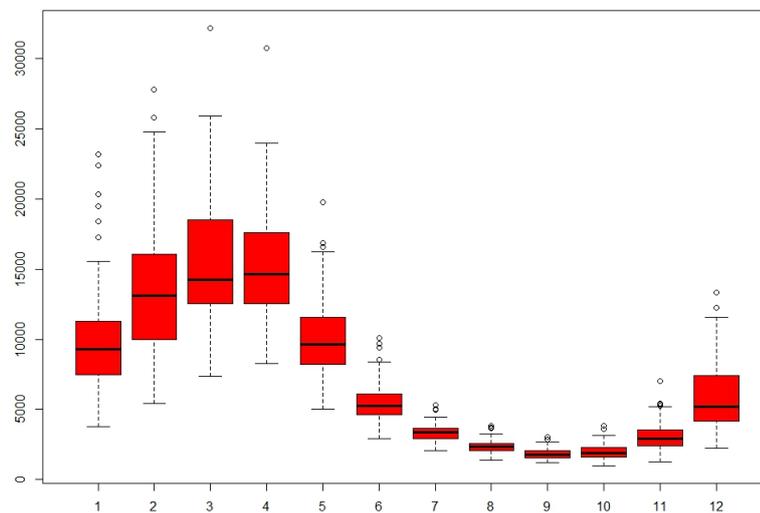


Figura 4.3: Boxplot Periódico para o Subsistema Norte

Ademais, conforme descrito por (Fernandez, 1984) e (Fernandez & Salas, 1986), é necessário que o histórico não tenha apenas evidência assimétrica, mas também que seja possível ajustar a distribuição gama a ele. Para se certificar desta hipótese, o pacote auxiliar "dgof" desenvolvido por (Arnold & Emerson, 2011) será

de suma importância. O pacote trabalha com funções que testam ajustes para quaisquer distribuições de interesse. Neste caso a função atribuída a esta tarefa chama-se "fitdist" (Apêndice, Bloco 2).

O histograma abaixo (Figura 4.4) representa a série histórica de energia ao longo dos anos em conjunto com a função de densidade da distribuição gama. Corroborando com as figuras anteriores, o subsistema Norte evidencia a assimetria do histórico. No que diz respeito a curva ajustada a ele, pode-se dizer preliminarmente que não houve um bom ajuste.

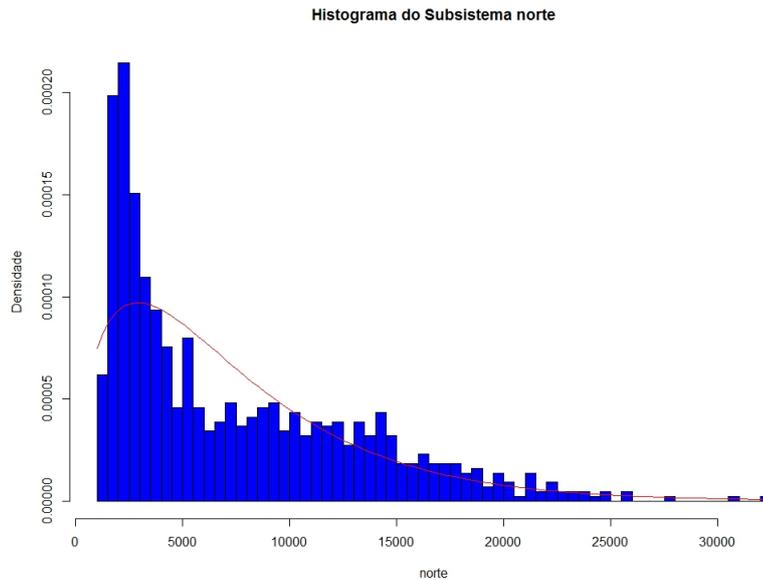


Figura 4.4: Histograma para o Subsistema Norte

Ao analisar o comportamento marginal da série temporal, isto é, mês a mês, necessita-se de um teste capaz de ajustar valores observados à uma distribuição de interesse. A literatura estatística fornece o teste de Kolmogorov-Smirnov como alternativa para solução deste problema (Apêndice, Bloco 2).

O mecanismo do teste de Kolmogorov-Smirnov se resume a comparação da diferença entre a distribuição empírica dos dados com a distribuição acumulada assumida por eles. Em seguida este resultado é comparado a um valor crítico originado a partir de um nível de significância.

Pela tabela (Tabela 4.1) pode-se dizer que todos os meses do subsistema Norte se ajustaram de maneira adequada a distribuição gama, ao nível de significância de 5%, tornando-se aptos a continuação da modelagem proposta por (Fernandez, 1984) e (Fernandez & Salas, 1986).

Subsistema Norte			
Período	P-valor	Período	P-valor
Janeiro	0.34	Julho	0.77
Fevereiro	0.97	Agosto	0.72
Março	0.50	Setembro	0.55
Abril	0.96	Outubro	0.92
Maiο	0.75	Novembro	0.51
Junho	0.78	Dezembro	0.51

Tabela 4.1: Teste de Kolmogorov-Smirnov para o Subsistema Norte

## 4.1

### Estimação dos Parâmetros

Conforme destacado por (Fernandez & Salas, 1986) e mencionado neste trabalho, utiliza-se o método de momentos como base para a estimação dos parâmetros do modelo (3-32). Como o cálculo da assimetria não faz parte dos pacotes padrões do software R, adotou-se o pacote complementar denominado "moments" (Komsta & Novomestky, 2015) como solução.

A função "apply" trabalha de forma que para cada mês seja calculado uma função de interesse. Portanto, "Q", "g" e "S" representam, respectivamente, a média, assimetria e desvio-padrão para os períodos do subsistema Norte. Além disso, a função seguinte implementa a estimação da autocorrelação de um mês corrente com o seu defasado (Apêndice, Bloco 3).

A partir das primeiras estimativas encontradas, o passo seguinte consiste na execução da estimação dos parâmetros periódicos. As funções encontradas em (Apêndice, Bloco 4) baseiam-se nas formulações apresentadas por (Fernandez & Salas, 1986) na estimação por método de momentos em (3-32) e pela estimação dos parâmetros periódicos descritos em (3-28).

#### 4.1.1

##### Procedimento alternativo para estimação

Na prática é comum encontrar problemas de estimação. O principal problema situa-se na violação do espaço paramétrico da distribuição Gama. Um modo alternativo para "corrigir" este problema é partir para métodos alternativos de estimação. Na literatura, é comum adotar o método de máxima verossimilhança como alternativa.

Através do pacote "MASS" (Venables & Ripley, 2002) e sua função "fitdistr", estima-se, por meio de "chutes iniciais" os parâmetros de interesse em cada período (Apêndice, Bloco 7).

## 4.2

### Geração dos resíduos do Modelo PGAR(1)

Esta é a etapa pilar para a programação do modelo. Conforme visto no capítulo anterior, os resíduos do modelo provém de uma complexa manipulação matemática resultante do mecanismo descrito em ((3-7) e (3-20)). Para programar esses resíduos torna-se necessário fornecer à função o momento de empregar o método aditivo e multiplicativo.

Inicia-se a função estabelecendo o comparativo entre o parâmetro  $\beta$  corrente com o seu defasado (Apêndice, Bloco 8). Para o modelo multiplicativo a programação contempla a parte descrita em (3-14) e adota-se K igual a 100.

Conforme mencionado no capítulo anterior, o modelo aditivo aborda a distribuição Gama de 3 parâmetros. Existem diversos pacotes disponíveis na biblioteca do R que trabalham com este tipo de distribuição. Sendo assim, adotou-se o uso do pacote "PearsonDS" (Becker & Klöbner, 2013) cuja função é denominada "rgamma3".

A parte aditiva é decorrente do valor encontrado para o parâmetro  $\theta$  (3-8) pertencente à distribuição Poisson. Na prática é comum encontrar problemas referentes a este parâmetro, uma vez que o espaço paramétrico da distribuição admite apenas valores positivos. Uma possível explicação para este fato pode estar ligada à influência de valores discrepantes. Para contornar este problema, criou-se a restrição dada em (Apêndice, Bloco 9) dentro da função geradora de resíduo.

#### 4.2.1

##### Geração dos resíduos em Modelos Aditivos e Multiplicativos

(Fernandez, 1984) destacou estruturas particulares para modelos PAGAR(1) e PMGAR(1). Admitindo que, marginalmente, os períodos assumam distribuição exponencial e, por consequência, assimetria constante, a simulação de modelos aditivos e multiplicativos se simplificará, respectivamente, para as funções encontradas em (Apêndice, Bloco 10) e (Apêndice, Bloco 11)

## 4.3

### Geração de Séries Sintéticas

Depois da estimação dos parâmetros e dos resíduos, a última etapa consiste em aproximar o comportamento estocástico e gerar novas séries sintéticas diferentes da original, mas possível do ponto de vista estatístico. Sendo assim, o presente trabalho adotou o seguinte método: omite-se 10 anos da série histórica (2004 à 2013) e gera-se novas séries para o mesmo intervalo de tempo.

Após isto, adotou-se o critério de gerar 500 novos resíduos compreendidos no referido período de tempo (Apêndice, Bloco 12), que participarão, posteriormente, da equação das séries sintéticas para o mesmo período de tempo.

Analisando o desempenho dos ruídos, nota-se um comportamento assimétrico do subsistema Norte (Figura 4.5) satisfazendo as condições impostas por (Fernandez

& Salas, 1986). Também pode ser observado a não geração de valores negativos, impossibilitando que erros matemáticos possam existir na geração das séries sintéticas do modelo (3-27).

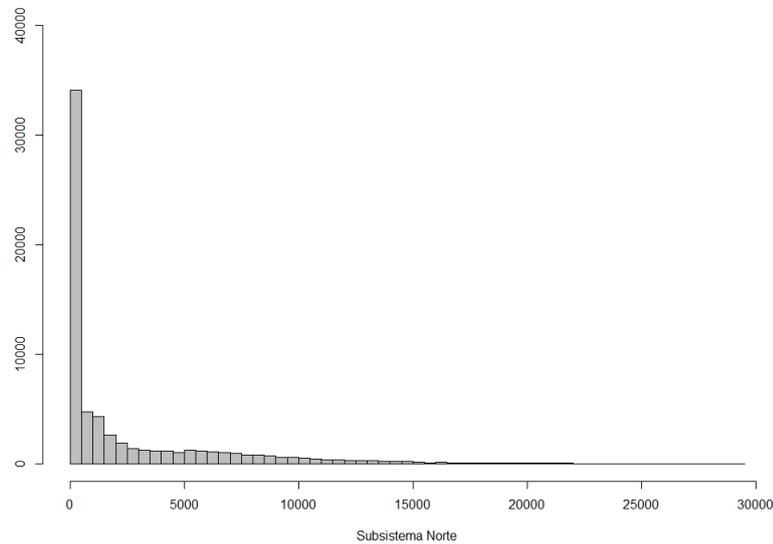


Figura 4.5: Ruído gerado para o Subsistema Norte

Por fim, gera-se o mesmo número de séries sintéticas (Apêndice, Bloco 13) anexando-as em uma mesma imagem (Figura 4.6).

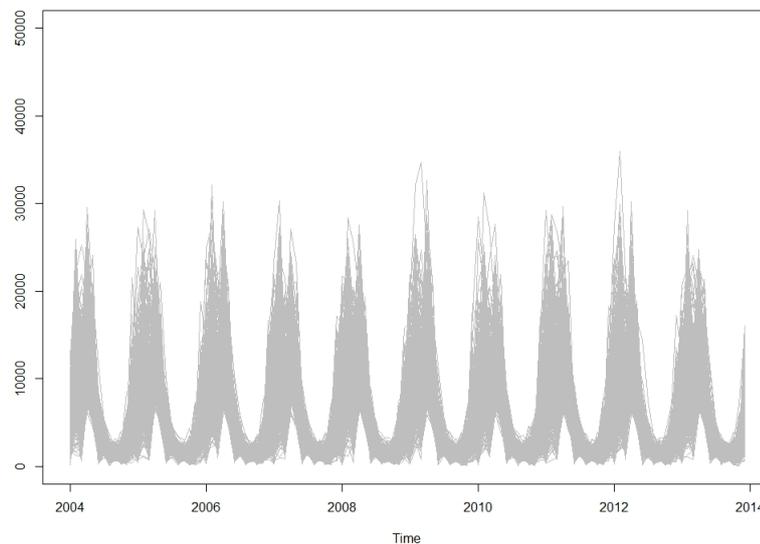


Figura 4.6: Simulação de 500 séries sintéticas para o Subsistema Norte

Conforme descrito anteriormente, as etapas da modelagem apresentadas neste capítulo foram baseadas através da teoria de (Fernandez, 1984) e (Fernandez & Salas, 1986) e é a partir dela que serão considerados os resultados desta dissertação.

A motivação para criação deste capítulo advém do fato de que outros pesquisadores poderão otimizá-lo e possivelmente adotá-lo em seus respectivos estudos, podendo publicá-los em revistas consolidadas no meio acadêmico. O código abre portas para novos horizontes dentro do software R tornando-se possível que pacotes futuros sejam implementados a fim de solucionar diversos problemas neste campo de pesquisa.

## 5 Resultados

Após a consolidação da teoria desenvolvida por (Fernandez, 1984) e (Fernandez & Salas, 1986) a presente seção apresenta os resultados da utilização da teoria sob o contexto dos subsistemas do Setor Elétrico Brasileiro, mais especificamente Sul, Sudeste/Centro-Oeste, Norte e Nordeste. A abordagem passará tanto pelo aspecto descritivo quanto pela geração de séries sintética.

Em sequência, testes de diagnósticos serão implementados no intuito de checar se a simulação possui validação tanto estatística, quanto estocástica.

### 5.1 Análise Preliminar das Séries ENAs dos Subsistemas

Dando continuidade ao que foi feito no Subsistema Norte no capítulo anterior, a seguir são plotados os boxplots periódicos e série histórica dos demais Subsistemas na intenção de observar seus respectivos desempenhos.

A partir de uma análise gráfica, nota-se diversos padrões para as vazões em cada região. No Subsistema Sul (Figura 5.2), por exemplo, observa-se uma enorme incerteza durante as estações presentes nos períodos intermediários, isto significa que sua variabilidade nesse período é maior quando comparado com as estações mais secas. Cabe ressaltar, assim como na pesquisa feita sobre o Norte, a presença de diversos valores extremos (outliers) presentes em todos os períodos do Subsistema Sul, em especial ao mês de Julho, apresentando um valor para vazão completamente atípico do histórico. Uma possível explicação para a ocorrência desse fenômeno pode estar ligado ao aumento significativo da precipitação de chuva na região Sul nos últimos 50 anos, conforme relata (Marengo et al., 2007 apud Marengo, 2008). Ademais (Groisman et al., 2005 apud Marengo, 2008) citam a tendência do aumento contínuo de chuva na região subtropical do Sul.

No Subsistema Sudeste/Centro-Oeste (Figura 5.4) observa-se um comportamento sazonal ao longo dos meses. Em estações mais quentes as vazões aumentam, assim como a sua variabilidade, enquanto no meses intermediários esse comportamento é reduzido. Conforme destacado no Sul, (Marengo et al., 2007 apud Marengo, 2008) relata um aumento na precipitação de chuva nos últimos anos para a região Sudeste. (Groisman et al., 2005 apud Marengo, 2008) descrevem um aumento sistemático de chuvas para a região. O autor também cita que fenômenos como o El Niño tem participação crucial para o aumento de precipitação de chuva na região.

No Subsistema Nordeste, representado pela figura(Figura 5.6), encontra-se um comportamento sazonal assim como no Sudeste/Centro-Oeste, isto é, há um aumento de vazão em períodos mais quentes, também conhecido como períodos úmidos. Este aumento está associado mais especificamente aos meses de Dezembro à Abril, enquanto que os demais meses, considerados como períodos secos, apresentam nitidamente uma queda nas vazões.

De modo geral, percebe-se que de todos os subsistemas, o Sul demonstrou o comportamento mais irregular, enquanto os demais demonstraram nitidamente um comportamento sazonal.

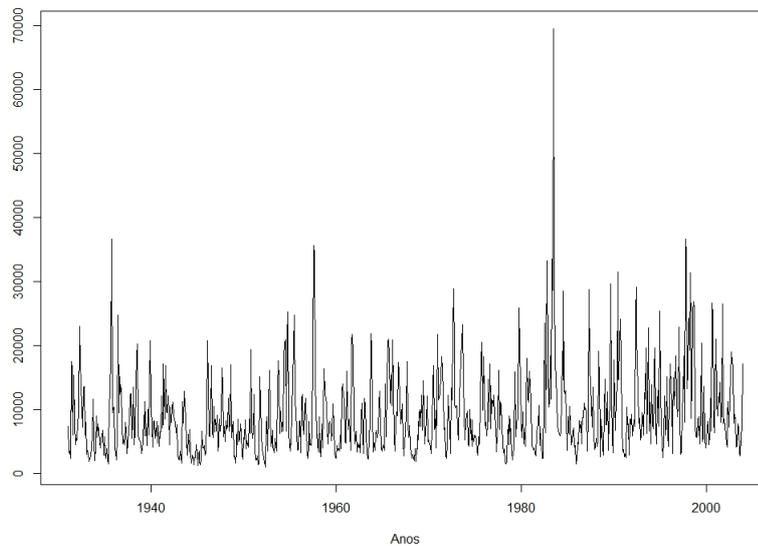


Figura 5.1: Série do Subsistema Sul

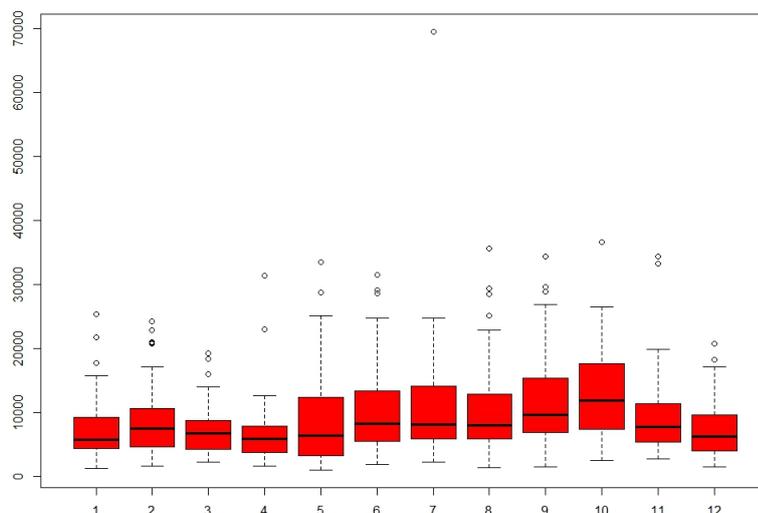


Figura 5.2: Boxplot Periódico para o Subsistema Sul

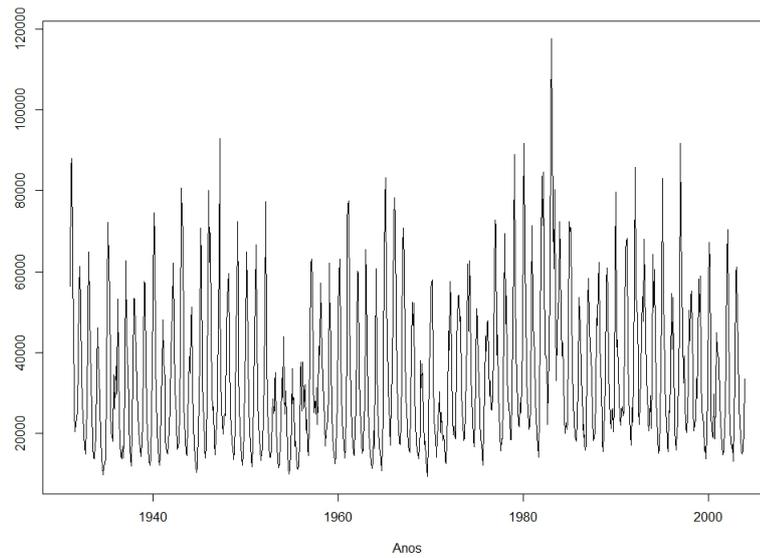


Figura 5.3: Série do Subsistema Sudeste/Centro-Oeste

PUC-Rio - Certificação Digital N° 1412776/CA

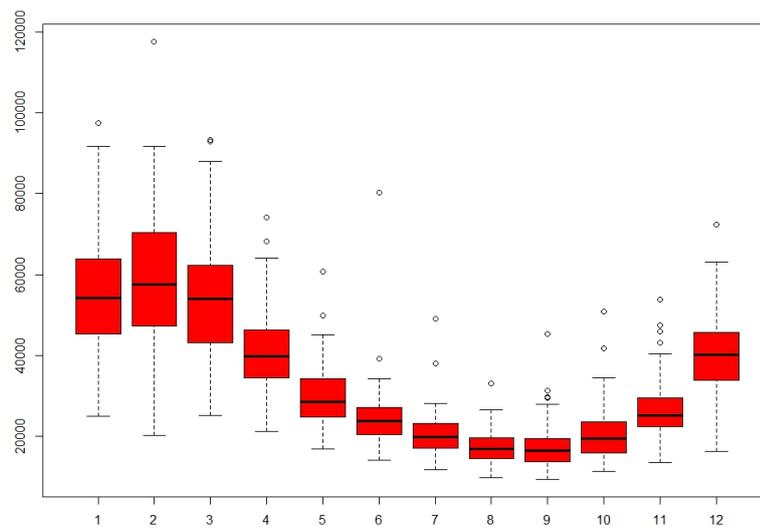


Figura 5.4: Boxplot Periódico para o Subsistema Sudeste/Centro-Oeste

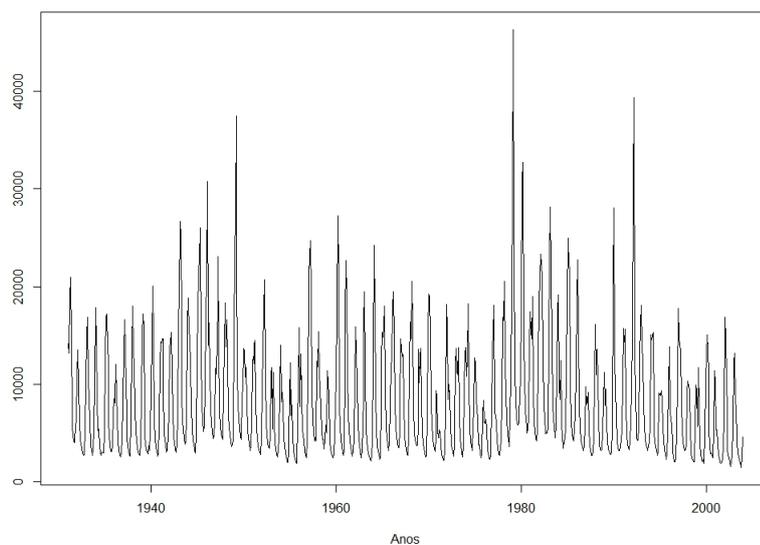


Figura 5.5: Série do Subsistema Nordeste

PUC-Rio - Certificação Digital Nº 1412776/CA

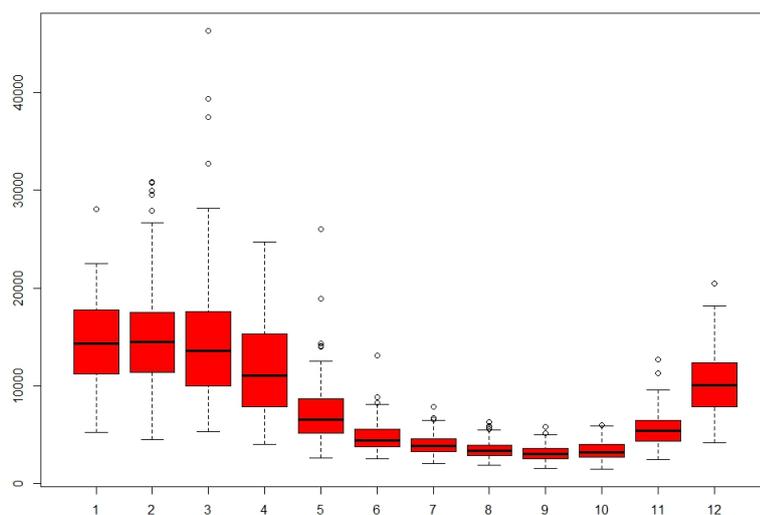


Figura 5.6: Boxplot Periódico para o Subsistema Nordeste

(Fernandez & Salas, 1986) relatam que para que sua modelagem seja válida, necessita-se de um comportamento gama da série temporal de estudo. Desse modo, as figuras (Figura 5.7), (Figura 5.8) e (Figura 5.9)) apresentam o comportamento da série histórica de energia ao longo dos anos em conjunto com a função de densidade da distribuição gama. Este fato contribui para a percepção de que em todos os subsistemas há um comportamento assimétrico. Visualmente o subsistema que aparenta melhor ajuste a distribuição gama é o Sul.

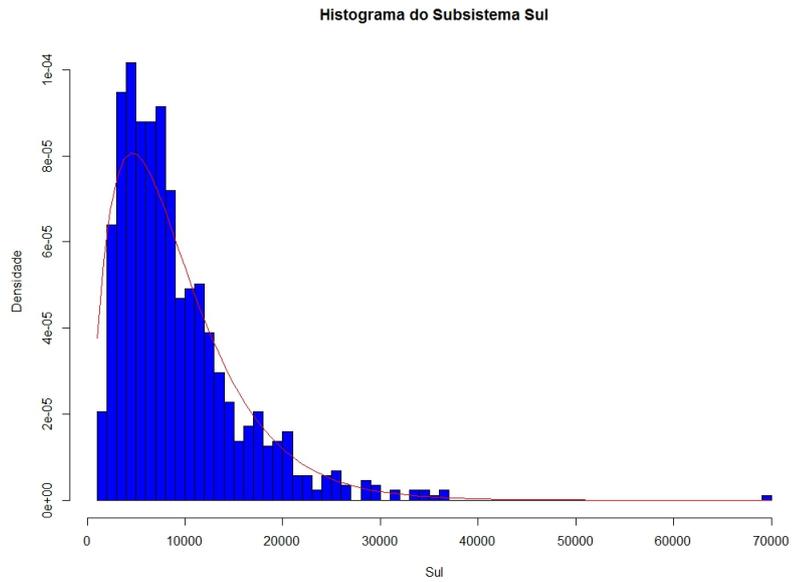


Figura 5.7: Histograma para o Subsistema Sul

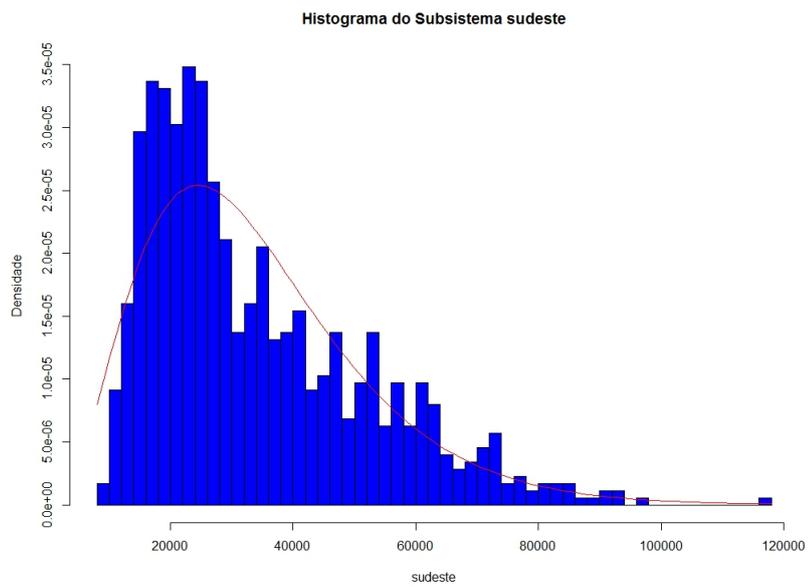


Figura 5.8: Histograma para o Subsistema Sudeste/Centro-Oeste

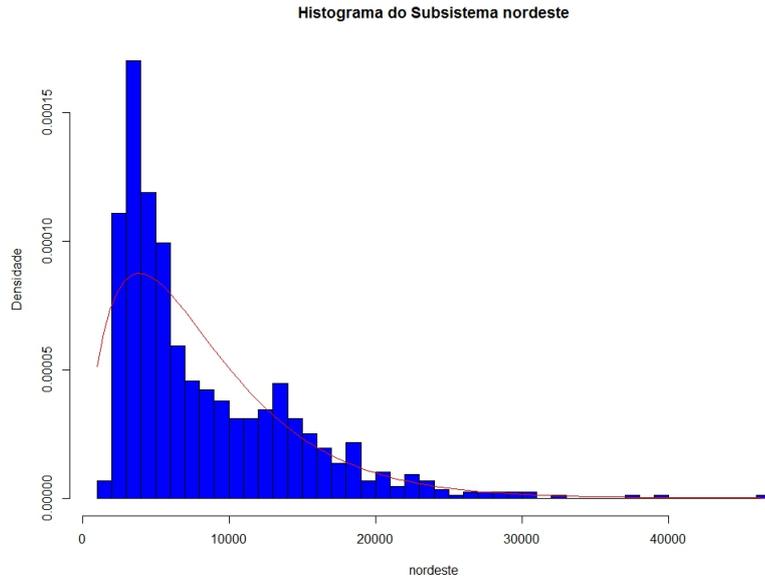


Figura 5.9: Histograma para o Subsistema Nordeste

Para analisar o comportamento marginal da série temporal, isto é, mês a mês, realizou-se o teste de hipótese de Kolmogorov-Smirnov com o intuito de averiguar o ajuste da distribuição gama. Ao nível de significância de 5%, nota-se a existência desta distribuição para todos os meses (Tabela 5.1) .

Período	Subsistemas		
	Sudeste/Centro-Oeste	Nordeste	Sul
	p-valor		
Janeiro	0.90	0.90	0.59
Fevereiro	0.62	0.57	0.91
Março	0.81	0.48	0.34
Abril	0.97	0.46	0.39
Mai	0.57	0.60	0.60
Junho	0.17	0.59	0.92
Julho	0.70	0.70	0.06
Agosto	0.94	0.90	0.84
Setembro	0.37	0.60	0.89
Outubro	0.69	0.75	0.55
Novembro	0.35	0.51	0.52
Dezembro	0.97	0.78	0.97

Tabela 5.1: Teste de Kolmogorov-Smirnov por mês para cada Subsistema

A partir do momento que os valores observados são comprovadamente assimétricos, segue-se a etapa que visa estimar os parâmetros que respeitem o espaço

paramétrico da distribuição especificada pelos autores em (3-2). Aceitar estimativas que violem o espaço paramétrico resultaria em erros matemáticos e no processo de simulação.

Na tabela (Tabela 5.2) estão listados as estimativas encontradas para cada subsistema. Como pode ser visto, o único subsistema a se enquadrar na metodologia de (Fernandez, 1984) e (Fernandez & Salas, 1986) é a região Norte, pois os demais subsistemas apresentam valores negativos para um dos parâmetros da distribuição gama, violando o seu espaço paramétrico. Uma possível explicação para a ocorrência deste fenômeno pode estar ligado aos valores discrepantes encontrados no histórico.

Período	Subsistemas											
	Sudeste/Centro-Oeste			Nordeste			Norte			Sul		
	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$
Janeiro	55088.5	2968.0	25.4	-10589.5	654.5	38.3	4248.7	2677.5	3.0	2276.3	3982.3	1.2
Fevereiro	57603.1	3732.9	20.3	-2833.0	2206.2	8.2	247.2	1729.6	5.5	465.1	3160.8	2.5
Março	53795.0	3141.6	21.9	6317.7	6920.2	1.3	6585.9	2280.3	4.0	900.8	2149.6	2.9
Abril	39865.4	4183.6	6.1	-3436.4	1934.3	8.0	7317.4	1990.3	4.0	3291.3	6539.4	0.5
Maiο	27367.1	4845.9	2.3	4298.9	4336.6	0.7	4268.1	1395.1	4.2	-669.9	5071.4	1.8
Junho	-46538.4	17158.7	0.2	2911.4	1553.7	1.3	3020.6	823.2	3.5	306.3	4522.6	2.1
Julho	12913.3	6193.9	0.8	1644.3	573.0	4.2	1572.3	271.9	5.1	6191.4	17945.7	0.3
Agosto	16894.4	2046.6	4.2	1076.4	388.4	6.4	1394.8	251.3	4.0	436.8	4966.2	2.0
Setembro	12890.1	5511.9	1.2	1199.7	368.8	5.4	937.4	172.1	4.6	-514.1	4070.4	3.0
Outubro	17132.1	5714.9	1.5	3.8	311.9	11.2	940.4	288.5	3.8	-776.3	3973.6	3.5
Novembro	25200.0	4629.7	2.6	1998.0	1329.1	2.80	1292.1	611.0	3.4	3691.1	6286.7	0.9
Dezembro	40372.2	2628.7	16.0	-10321.6	688.6	30.2	1273.0	1268.2	3.8	-1372.9	2214.0	4.0

Tabela 5.2: Parâmetros estimados via Método de Momentos

Para contornar o problema das estimativas, adotou-se como alternativa o método de Máxima Verossimilhança, inclusive para o Subsistema Norte. Muito embora (Fernandez, 1984) evidencie problemas práticos de assimetria para este método quando  $\beta_t$  for menor do que 1, optou-se por descartar esta hipótese devido a magnitude dos valores observados. O resultado deste método está ilustrado na tabela seguinte (Tabela 5.3).

Por consequência da não violação dos parâmetros para o referido método, estes resultados farão parte dos demais procedimentos de seus respectivos subsistemas.

É importante ressaltar que as estimativas encontradas pelo método de Máxima Verossimilhança forneceram valores completamente distintos quando do método de Momentos. Esta diferença pode estar relacionada com as propriedades distintas que os métodos possuem.

Período	Subsistemas											
	Sudeste/Centro-Oeste			Nordeste			Norte			Sul		
	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$
Janeiro	934.73	4107.77	13.22	2504.52	1508.05	7.94	3499.99	3299.98	2.19	952.92	2610.29	2.36
Fevereiro	901.63	5347.34	10.71	1897.12	3116.96	4.28	489.26	1695.03	7.65	589.01	2904.57	2.69
Março	1472.45	4328.39	12.23	3627.59	4227.94	2.73	6331.48	2437.88	3.83	2093.10	2909.47	1.70
Abril	971.16	2664.11	14.98	2429.82	3081.92	3.14	7000.00	2000.00	4.24	299.99	1999.97	3.01
Maior	2025.46	1800.17	15.35	1951.55	1858.11	2.97	4200.00	1300.00	4.70	199.97	3499.97	2.16
Junho	3072.64	2029.65	10.83	2425.69	1199.24	2.11	2473.16	657.28	4.62	999.98	3499.98	2.39
Julho	995.98	1246.19	15.85	1690.96	575.57	4.17	1395.90	244.99	8.09	99.91	2999.80	3.34
Agosto	984.86	970.40	16.95	1420.11	456.03	4.72	1067.62	180.16	7.32	1005.65	5106.09	1.78
Setembro	1576.16	1819.45	8.76	1188.74	357.33	5.61	1089.30	210.48	3.61	876.45	4429.77	2.44
Outubro	4499.98	2499.87	6.55	973.86	436.93	5.77	784.59	239.56	4.97	742.44	4168.43	2.95
Novembro	1768.63	1939.62	12.98	2191.44	1403.31	2.50	1018.35	523.80	4.01	999.88	2499.74	3.13
Dezembro	1000.00	3000.00	13.21	4113.15	4903.24	1.41	1442.84	1197.51	3.74	999.80	2999.76	2.12

Tabela 5.3: Parâmetros estimados via Método de Máxima Verossimilhança

## 5.2

**Simulação de Séries Sintéticas do Modelo PGAR(1) para os Subsistemas**

Uma vez obtidos os parâmetros da distribuição, o passo seguinte constitui-se nas estimações dos coeficientes periódicos da equação (3-27). Após esta estimação, a próxima etapa consiste na reprodução dos resíduos descritos pelas transformações de Merllin e Laplace-Siteltjes representadas, respectivamente, em (3-6) e (3-25).

A ideia básica da geração dos resíduos é fundamentada através de um mecanismo capaz de entender o momento adequado para reproduzir o PAGAR(1) e PMGAR(1). (Fernandez & Salas, 1986) demonstram que essa reprodução dependerá exclusivamente do comportamento de  $\beta_t$  (3-29). Para ilustrar a presente situação, na figura (Figura 5.10) encontra-se os  $\beta_t$  mensais dos subsistemas.

Analisando a figura (Figura 5.10) e fazendo um comparativo com o  $\beta_t$  e  $\beta_{t-1}$  observa-se, por exemplo, que para o mês de Fevereiro os subsistemas se comportam de maneiras distintas. Enquanto os subsistemas Norte e Sul adotariam o PAGAR(1) ( $\beta_t \geq \beta_{t-1}$ ), os subsistemas Nordeste e Sudeste/Centro-Oeste empregariam o PMGAR(1) ( $\beta_t \leq \beta_{t-1}$ ). Esta característica se estende para os meses subsequentes.

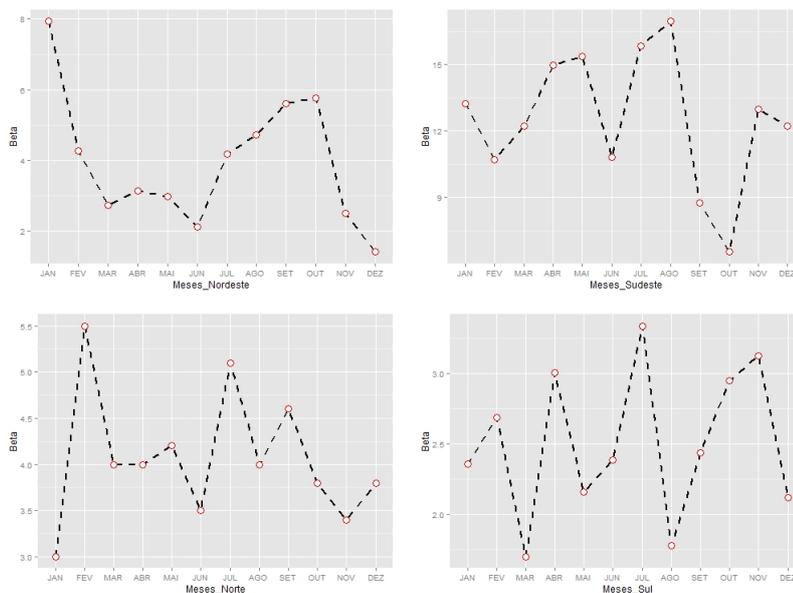


Figura 5.10: Geração dos "Betas" para os Subsistemas

A partir deste entendimento, os parâmetros estimados incorporam-se ao método da geração de resíduos descrito na seção anterior para, posteriormente, integrarem-se a equação do modelo PGAR(1) para simulação de séries hidrológicas sintéticas. Como é de praxe dentro do contexto de séries temporais, é importante verificar o comportamento residual de cada subsistema. Para checar este comportamento adotou-se o critério de simular 500 séries de ruídos ao longo de 10 anos, uma vez que o sistema NEWAVE trabalha com séries com este período de tempo. É importante ressaltar que, futuramente, esta simulação residual servirá como

arcabouço para a geração das séries sintéticas.

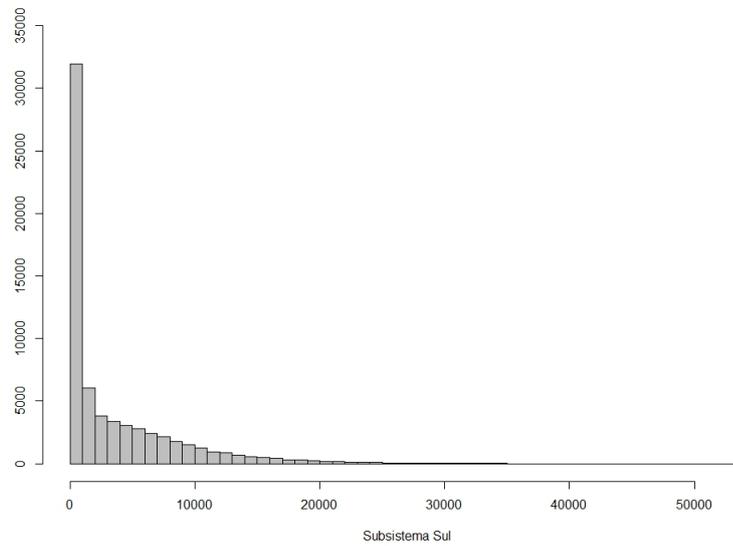


Figura 5.11: Ruído gerado para o Subsistema Sul

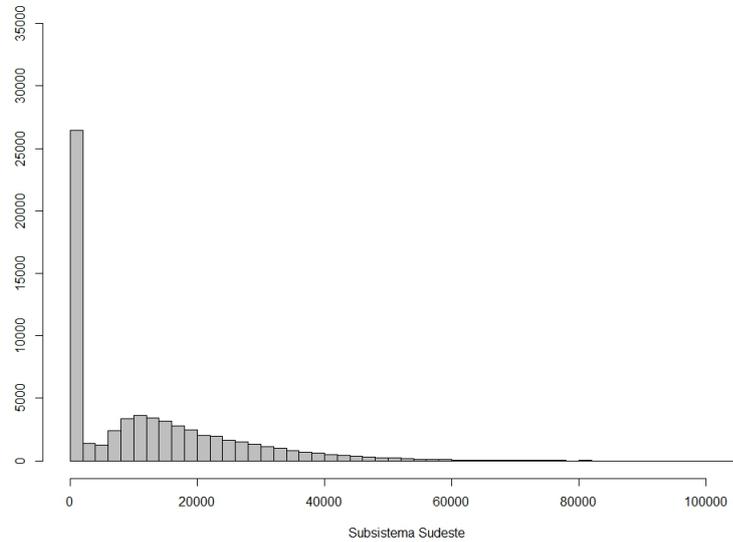


Figura 5.12: Ruído gerado para o Subsistema Sudeste/Centro-Oeste

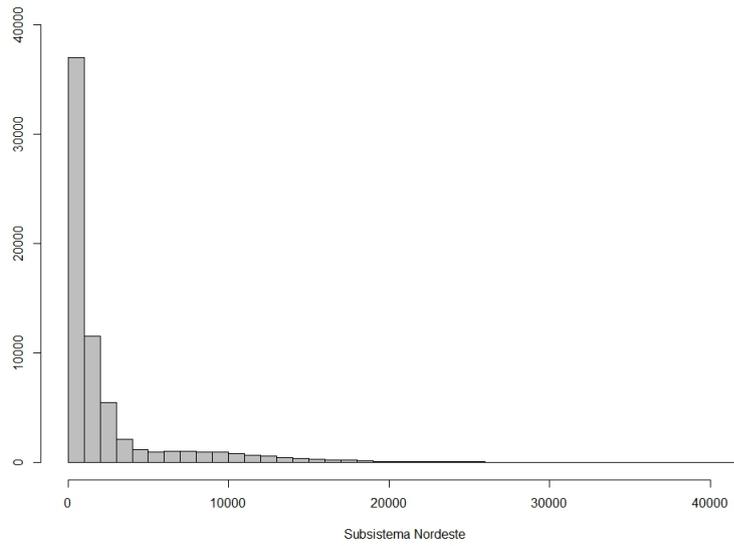


Figura 5.13: Ruído gerado para o Subsistema Nordeste

PUC-Rio - Certificação Digital N° 1412776/CA

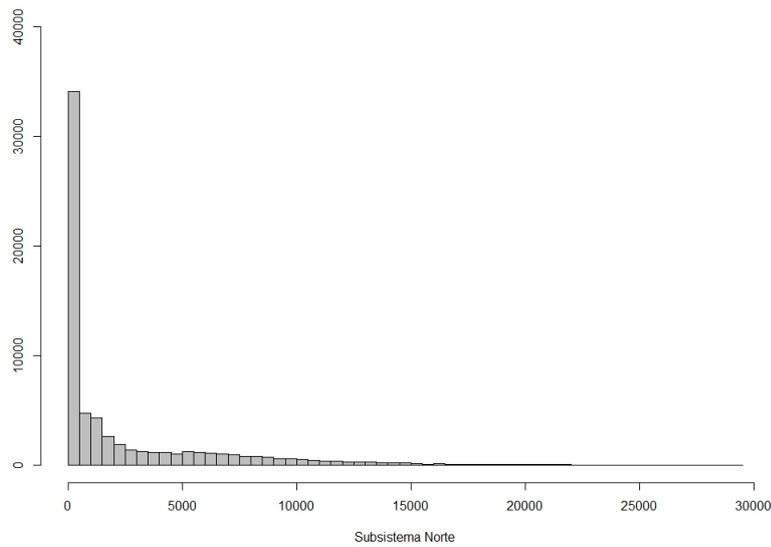


Figura 5.14: Ruído gerado para o Subsistema Norte

Pelas figuras (Figura 5.11), (Figura 5.12), (Figura 5.13) e (Figura 5.14) nota-se um comportamento assimétrico de todos os resíduos dos subsistemas, satisfazendo as condições impostas por (Fernandez, 1984). Também pode ser observado que estes resíduos não produziram valores negativos, o que era esperado, uma vez que toda a metodologia estudada neste trabalho parte do princípio da distribuição gama.

Considerando os valores obtidos para os resíduos, a próxima etapa tem por finalidade simular novas séries capazes de refletir realidades coerentes com a

original observada para os subsistemas, tanto do ponto de vista estatístico, quanto estocástico.

Conforme mencionado nesta seção, a série compreende o ano de 1931 à 2013. Para testar a simulação, optou-se por trabalhar com o método *out of the sample*. Formalizando, "corta-se" os 10 últimos anos da série temporal, constroi-se um bloco de dados contendo os anos de 1931 à 2003 e, em sequência, simula-se a partir do último ano observado novas séries hidrológicas para os próximos 10 anos. As figuras (Figura 5.15), (Figura 5.16), (Figura 5.17) e (Figura 5.18) apresentam a simulação de 500 novas séries seguindo este princípio.

Para fins comparativos, plotou-se o comportamento médio dos cenários (linha vermelha) bem como a média do período de 1931 à 2003 (linha amarela).

Dentre todos os subsistemas simulados, o Sul (Figura 5.15) configura a região de menor ENA e a série mais irregular. Esta irregularidade também é notada observando o histórico. Numa inspeção visual, percebe-se, a princípio, que a média dos cenários e a média do histórico não se sobrepõem. Para o subsistema Sudeste/Centro-Oeste (Figura 5.16) observa-se cenários mais compactos, mostrando um comportamento mais regular ao longo dos anos e altos valores ENA. Além disso, a média do histórico em comparação com à média dos cenários não demonstrou indícios de uma boa reprodução para o subsistema. O mesmo problema aparenta persistir para os subsistemas Nordeste (Figura 5.17) e Norte (Figura 5.18).

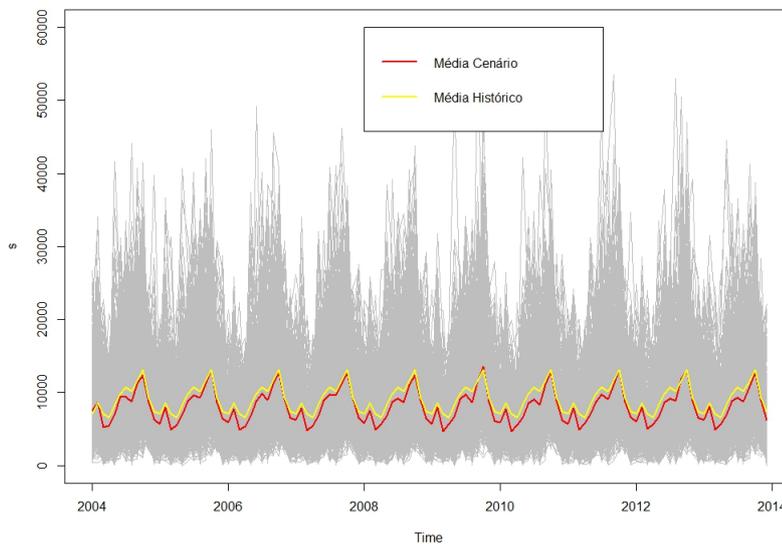


Figura 5.15: PGAR no Subsistema Sul

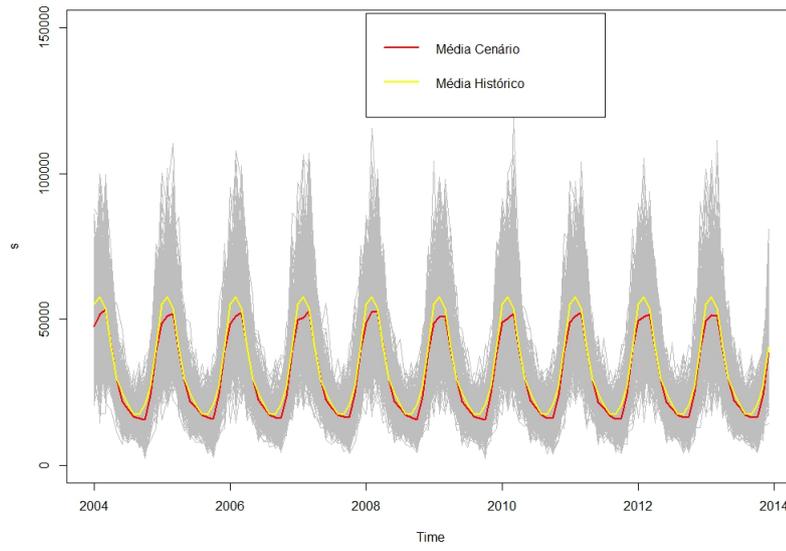


Figura 5.16: PGAR no Subsistema Sudeste/Centro-Oeste

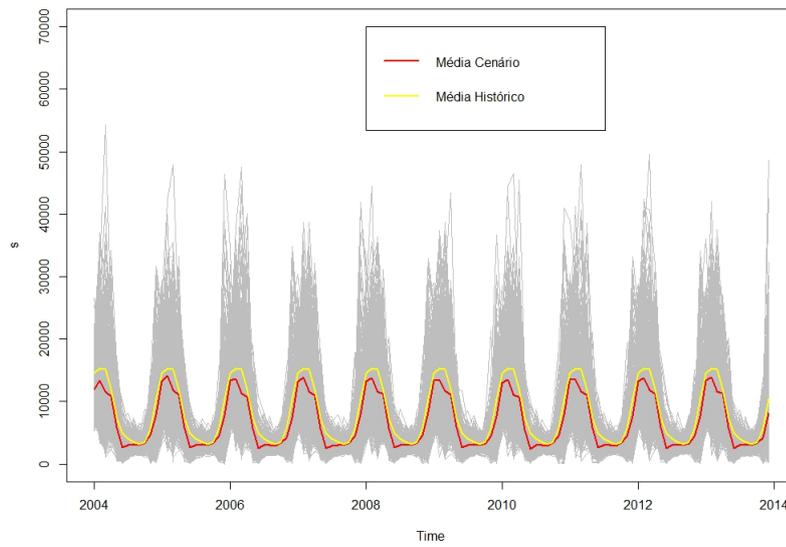


Figura 5.17: PGAR no Subsistema Nordeste

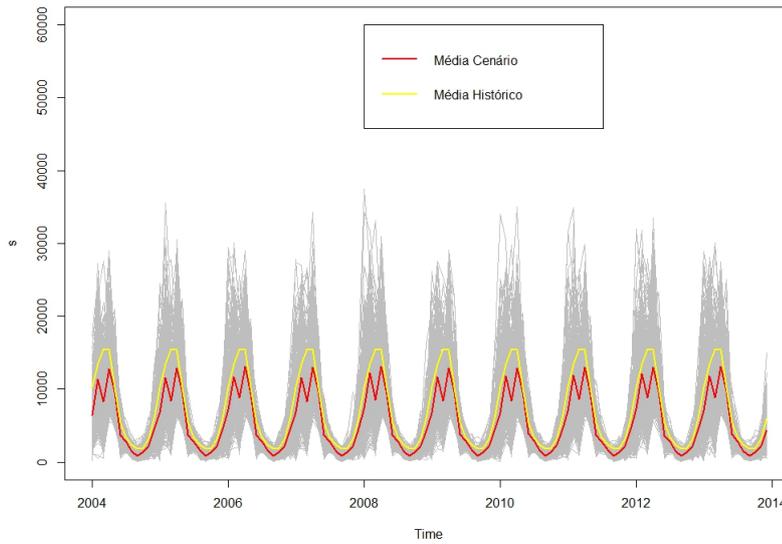


Figura 5.18: PGAR no Subistema Norte

Ainda que a simulação em conjunto com o histórico traga uma ideia inicial sobre o andamento da simulação, é interessante observar o comportamento da variabilidade em cada subsistema simulado. As figuras (Figura 5.19), (Figura 5.20), (Figura 5.21) e (Figura 5.22) ilustram esta situação.

Conforme se observa, os subsistemas Norte (Figura 5.22) e Nordeste (Figura 5.21) aparentam melhores reproduções ao modelo PGAR(1) em relação aos demais subsistemas, sobrepondo-se, inclusive, a alguns períodos estudados.

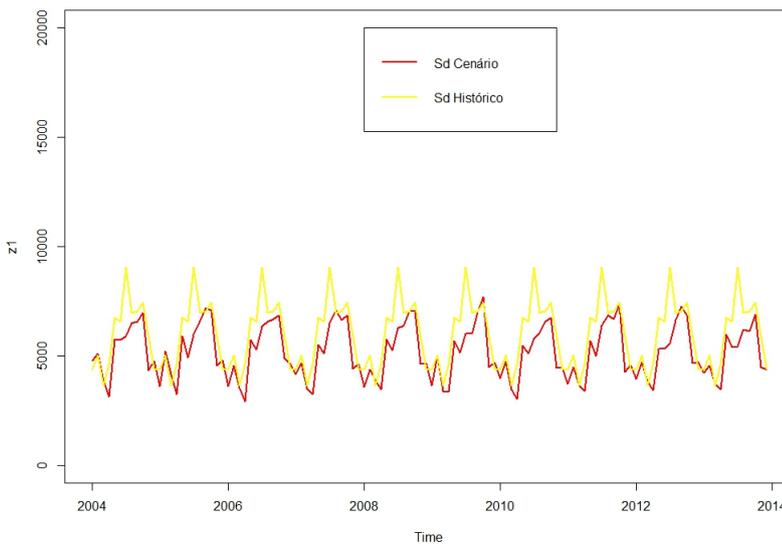


Figura 5.19: Desvio-Padrão PGAR no Subistema Sul

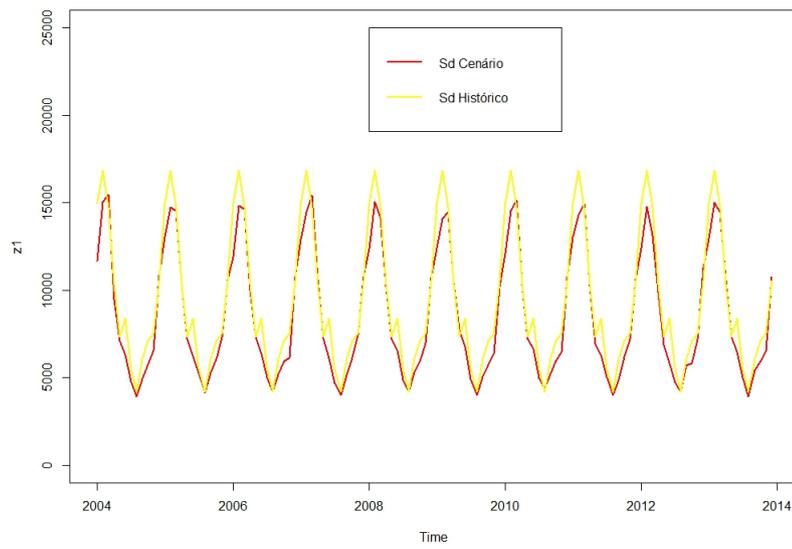


Figura 5.20: Desvio-Padrão PGAR no Subsistema Sudeste/Centro-Oeste

PUC-Rio - Certificação Digital N° 1412776/CA

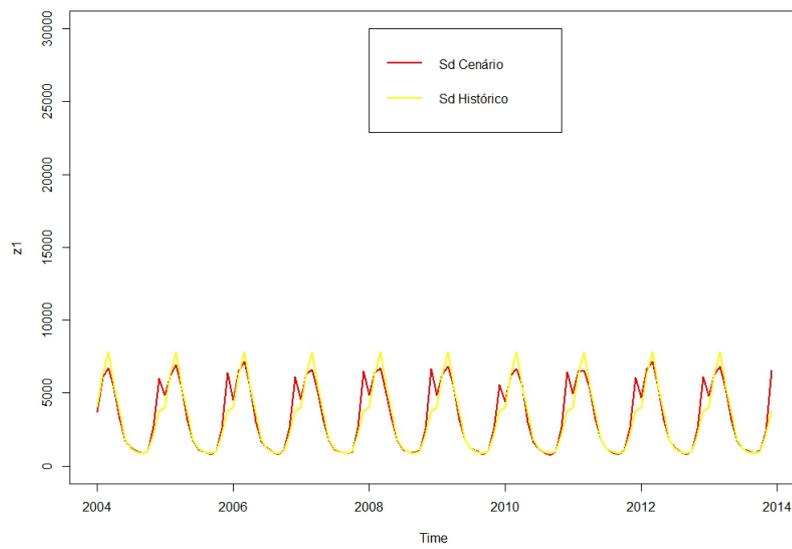


Figura 5.21: Desvio-Padrão PGAR no Subsistema Nordeste

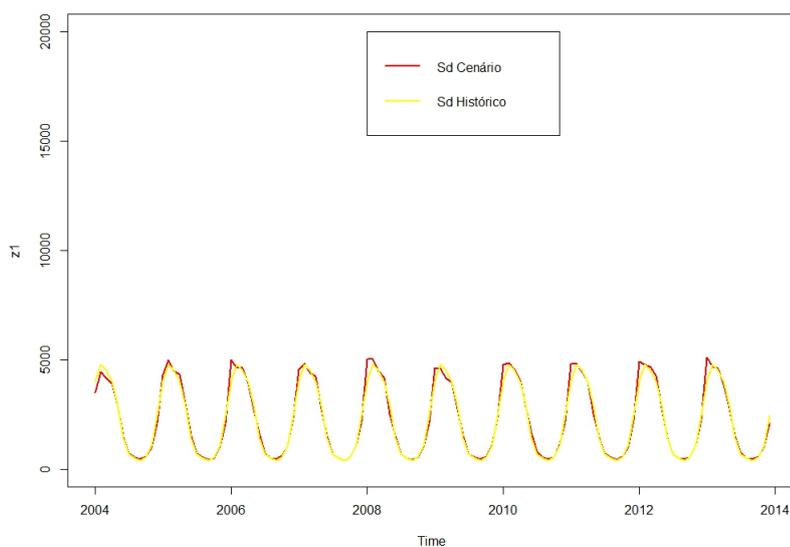


Figura 5.22: Desvio-Padrão PGAR no Subsistema Norte

As análises vistas até o momento permitem apenas uma avaliação visual para o tratamento do problema. Para dar continuidade, testes de hipóteses são necessários para aferir se há uma boa representação da simulação com a realidade.

### 5.3 Teste de Diagnóstico

Nesta seção as séries sintéticas são testadas para verificar seu comportamento estatístico quanto ao histórico. Para construir os testes estatísticos, adota-se o seguinte critério: para cada um dos 500 cenários simulados, testa-se, marginalmente, se a simulação reproduzem bem o histórico.

Na literatura existem dois métodos para testar hipóteses: paramétrico e não-paramétrico. Por se tratar de um problema não linear e com observações de tamanho pequeno, o presente trabalho aderiu ao método não-paramétrico.

#### 5.3.1 Teste de Variância e Assimetria

Como o tratamento é feito ao longo de 10 anos e o interesse é marginal, cada mês terá uma amostra de tamanho 10. Portanto, o primeiro teste realizado é conhecido como "Teste de Levene" e verifica a igualdade das variâncias entre a simulação e o histórico. As figuras (Figura 5.23), (Figura 5.24), (Figura 5.25) e (Figura 5.26) ilustram a presente situação.

Nota-se que as séries simuladas reproduziram de forma satisfatória o histórico ao nível de significância de 5% (linha vermelha). Logo, pode-se dizer que as variâncias são iguais.

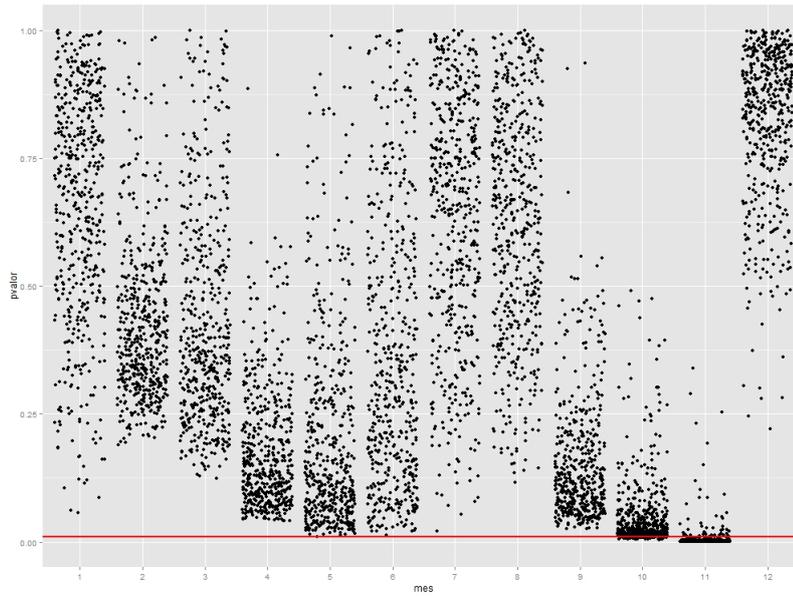


Figura 5.23: P-valor do Teste de Levene para o modelo no Subsystema Sul

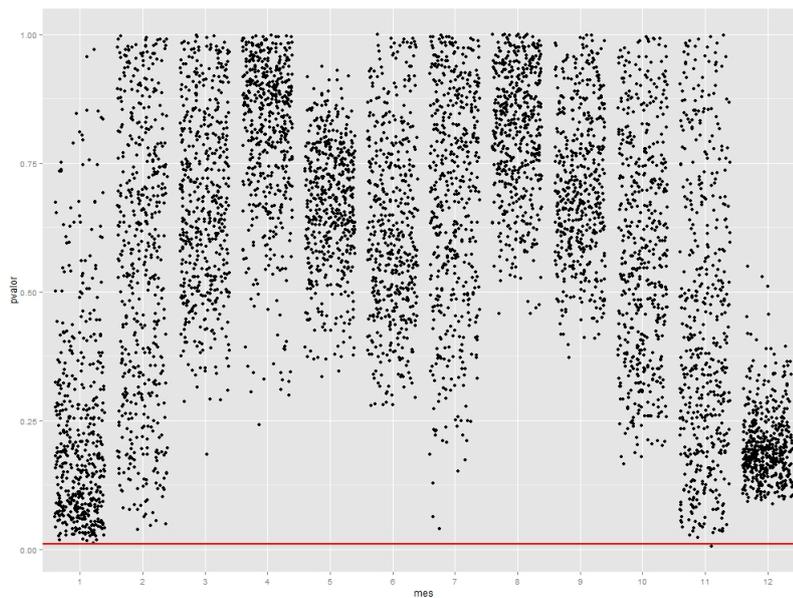


Figura 5.24: P-valor do Teste de Levene para o modelo PGAR no Subsystema Sudeste/Centro-Oeste

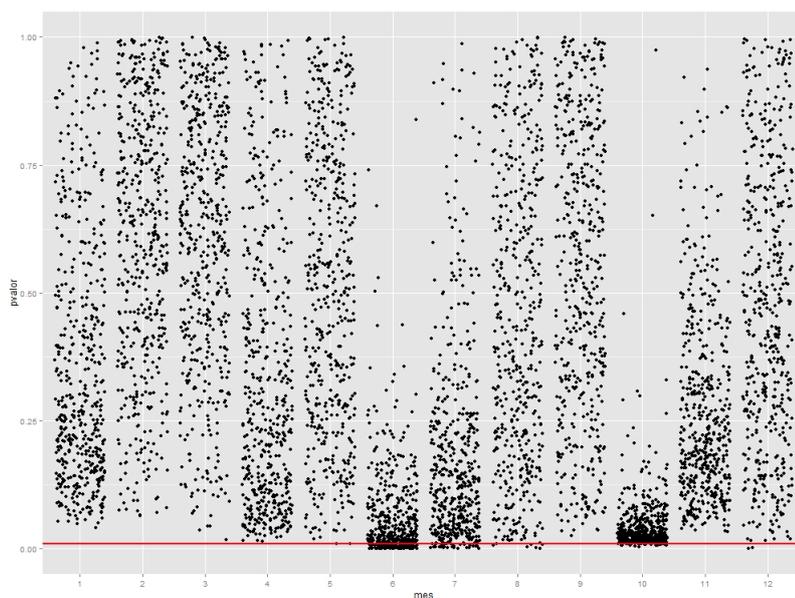


Figura 5.25: P-valor do Teste de Levene para o modelo PGAR no Subistema Nordeste

PUC-Rio - Certificação Digital N° 1412776/CA

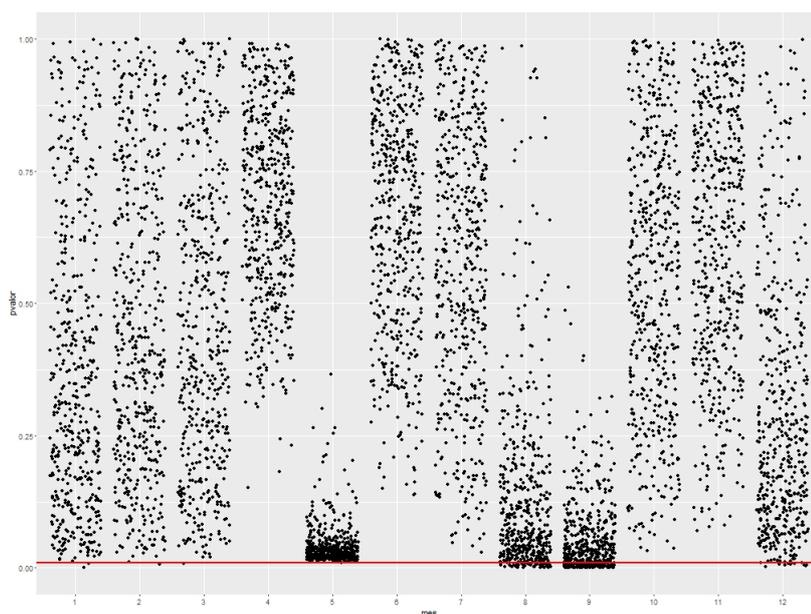


Figura 5.26: P-valor do Teste de Levene para o modelo PGAR no Subistema Norte

O desempenho da assimetria dos cenários é avaliado através do método *Bootstrap* (Efron & Tibshirani, 1994). De modo geral, o *Bootstrap* estima parâmetros através de reamostragem. Esta técnica é bastante empregada devido à sua eficiência e por não necessitar de muitas suposições na estimação de parâmetros. Utilizando esta ferramenta dentro do contexto do presente trabalho, reamostraram-se 1000 vezes cada mês do histórico pertencentes ao período de 1931 à 2003. Para cada

reamostragem, calcula-se a sua respectiva assimetria e, posteriormente, reproduz-se seu intervalo com 99% de confiança . Em seguida, computa-se a assimetria dos cenários e compara-se com os intervalos gerados.

Os resultados evidenciados nas tabelas (5.4) a (5.7) revelaram problemas (números em vermelho) para esta estimativa nos subsistemas Sul, Sudeste/Centro-Oeste e Nordeste, com destaque maior para o subsistema Sudeste/Centro-Oeste, apresentando frequência maior de meses fora do intervalo de confiança. Em contrapartida, pode-se observar tanto para a Assimetria Simulada quanto Assimetria Observada que o histórico e cenários reforçam a ideia da classificação de assimetria moderada a forte.

É importante ressaltar que esta etapa da análise representa uma maneira diferente de analisar a assimetria dos dados com a assimetria simulada, uma vez que não há nenhum teste de hipótese específico na literatura que trabalhe com esta situação.

Período	Subsistema Sul			
	Assimetria Simulada	Assimetria Observada	LI	LS
Janeiro	0.68	1.82	0.50	2.61
Fevereiro	0.58	1.26	0.44	1.73
Março	0.74	1.18	0.25	1.79
Abril	0.60	2.83	0.29	3.84
Maió	0.82	1.51	0.66	2.27
Junho	0.63	1.38	0.67	1.96
Julho	0.66	3.97	0.38	4.84
9 Agosto	0.72	1.43	0.63	2.14
Setembro	0.63	1.16	0.54	1.80
Outubro	0.57	1.07	0.25	1.68
Novembro	0.59	2.13	0.39	2.78
Dezembro	0.72	1.00	0.47	1.53

Tabela 5.4: Intervalo de Confiança (99%) para a Assimetria via Bootstrap no Subsistema Sul

Período	Subsistema Sudeste			
	Assimetria Simulada	Assimetria Observada	LI	LS
Janeiro	0.25	0.40	-0.20	0.94
Fevereiro	0.23	0.44	-0.60	1.38
Março	0.34	0.43	-0.25	0.90
Abril	0.29	0.81	-0.03	1.43
Maio	0.30	1.32	-0.22	2.11
Junho	0.33	4.08	0.03	5.06
Julho	0.23	2.24	-0.12	3.12
Agosto	0.22	0.98	0.05	1.74
Setembro	0.32	1.82	0.58	2.74
Outubro	0.34	1.61	0.47	2.40
Novembro	0.28	1.23	0.30	2.02
Dezembro	0.25	0.50	-0.19	1.21

Tabela 5.5: Intervalo de Confiança (99%) para a Assimetria via Bootstrap no Subsistema Sudeste/Centro-Oeste

Período	Subsistema Nordeste			
	Assimetria Simulada	Assimetria Observada	LI	LS
Janeiro	0.43	0.32	-0.56	1.16
Fevereiro	0.46	0.70	0.18	1.22
Março	0.57	1.77	0.66	2.58
Abril	0.54	0.71	0.25	1.25
Maio	0.55	2.36	0.42	3.35
Junho	0.66	1.76	0.36	2.79
Julho	0.49	0.97	0.39	1.62
Agosto	0.40	0.79	0.23	1.31
Setembro	0.39	0.86	0.25	1.42
Outubro	0.41	0.60	0.10	1.11
Novembro	0.53	1.20	0.42	1.82
Dezembro	0.79	0.36	-0.11	0.80

Tabela 5.6: Intervalo de Confiança (99%) para a Assimetria via Bootstrap no Subsistema Nordeste

Período	Subsistema Norte			
	Assimetria Simulada	Assimetria Observada	LI	LS
Janeiro	0.67	1.34	0.59	2.01
Fevereiro	0.40	0.72	-0.03	1.35
Março	0.53	1.01	0.08	1.80
Abril	0.54	0.99	0.08	1.87
Maiο	0.50	0.96	0.34	1.64
Junho	0.44	1.15	0.18	1.71
Julho	0.35	0.78	-0.10	1.24
Agosto	0.39	1.01	0.20	1.53
Setembro	0.43	0.87	0.26	1.45
Outubro	0.45	1.06	0.15	1.73
Novembro	0.56	1.16	0.31	1.95
Dezembro	0.59	1.04	0.37	1.62

Tabela 5.7: Intervalo de Confiança (99%) para a Assimetria via Bootstrap no Subsistema Norte

### 5.3.2

#### Teste de Média

Assumindo a igualdade das variâncias, o teste de Mann-Whitney foi empregado para verificar a existência de heterogeneidade dos dados simulados com o histórico. Este teste representa uma versão não paramétrica do teste t. O teste foi realizado para cada um dos 500 cenários simulados computando seu respectivo p-valor. Em seguida, este p-valor é comparado com o nível de significância de 5% para verificar a aceitação do teste. As figuras (5.27) a (5.30) representam a proporção de aceitação do teste em cada período.

Analisando o resultado do teste, é possível perceber que os subsistemas exibiram comportamentos distintos. Enquanto o Nordeste (Figura 5.29) e Norte (Figura 5.30) configuraram os piores resultados, os subsistemas Sul (Figura 5.27) e Sudeste/Centro-Oeste (Figura 5.28) apresentaram resultados mais satisfatórios. Muito embora no mês de Outubro para o subsistema Sudeste/Centro-Oeste a aceitação tenha sido abaixo de 50%, pode-se dizer que para este subsistema o modelo em questão teve uma boa reprodução. No entanto, o subsistema Sul se destaca como melhor reprodução do modelo de estudo, apresentando aceitação superior a 60% para todos os períodos.

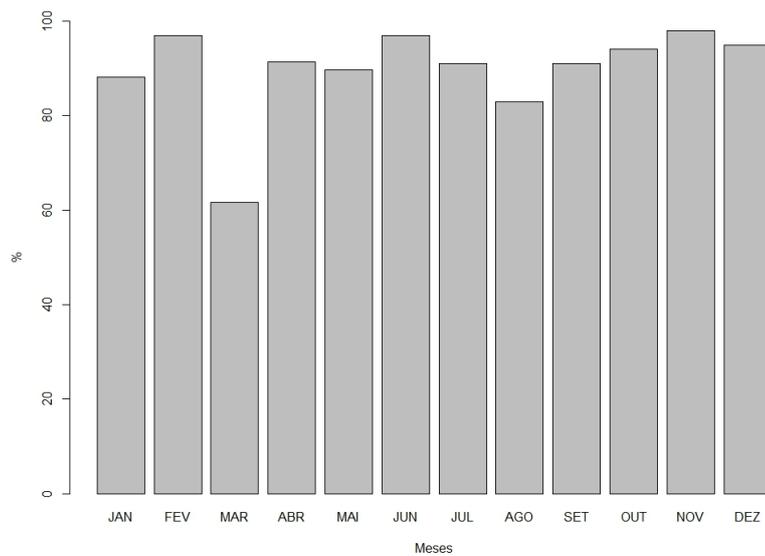


Figura 5.27: Proporção de aceitação do Teste de Mann-Whitney - Subsistema Sul

PUC-Rio - Certificação Digital N° 1412776/CA

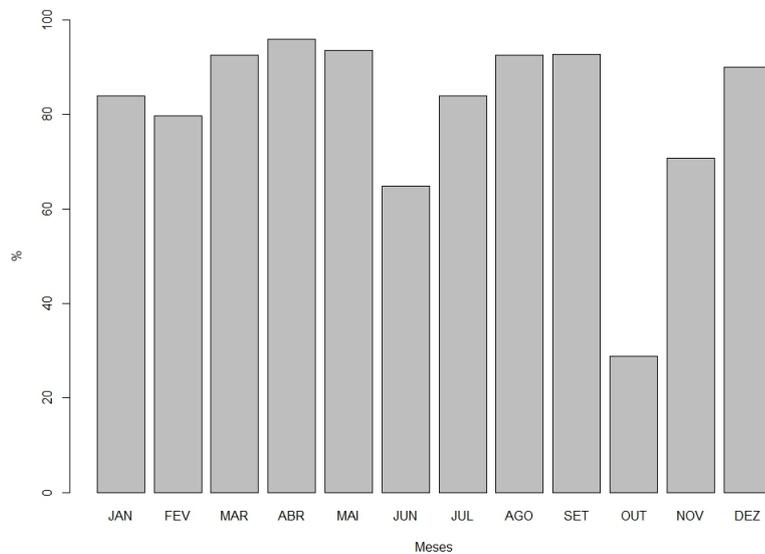


Figura 5.28: Proporção de aceitação do Teste de Mann-Whitney - Subsistema Sudeste/Centro-Oeste

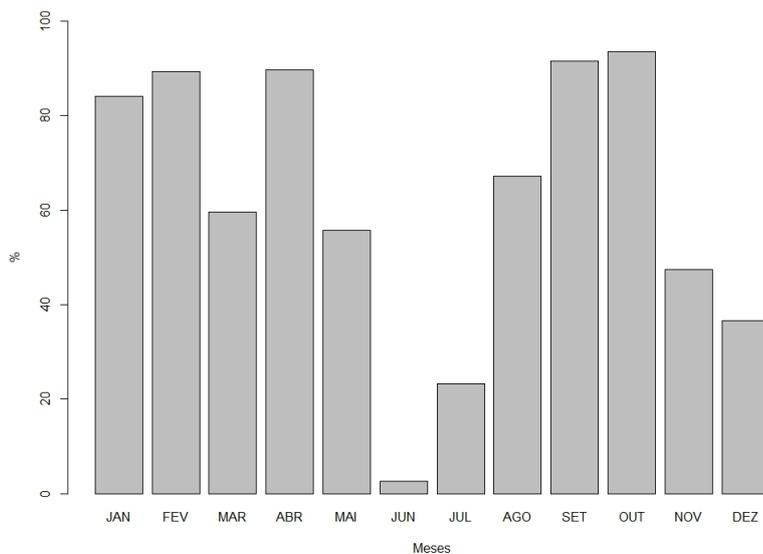


Figura 5.29: Proporção de aceitação do Teste de Mann-Whitney - Subsistema Nordeste

PUC-Rio - Certificação Digital N° 1412776/CA

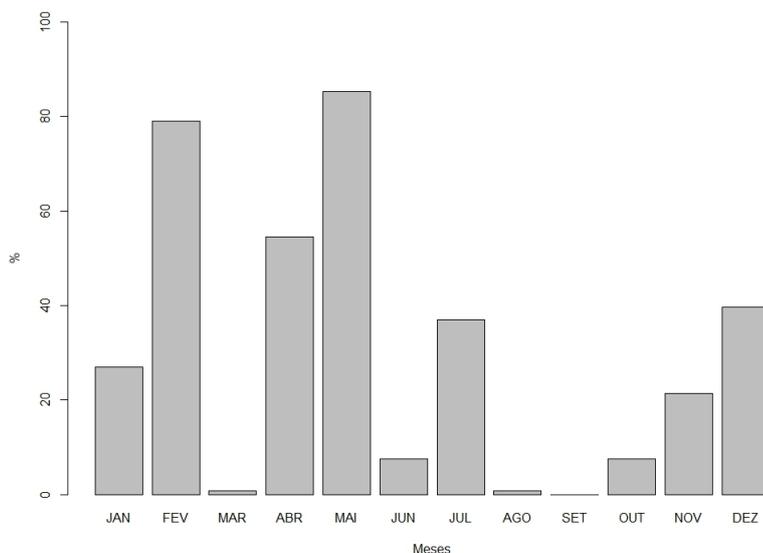


Figura 5.30: Proporção de aceitação do Teste de Mann-Whitney - Subsistema Norte

### 5.3.3 Restrições do Modelo PGAR(1)

Até o momento foram estudados testes para diagnosticar o comportamento da simulação. No entanto, é importante estudar a restrição paramétrica imposta pela modelagem descrita no capítulo 3. A presente seção tem o objetivo de checar o comportamento periódico do parâmetro  $\beta_t$ , descrito em (3-31), ao longo dos períodos.

Analisando as figuras (Figura 5.31), (Figura 5.32), (Figura 5.33) e (Figura 5.34), percebe-se que o subsistema Sul atendeu perfeitamente a restrição, enquanto no subsistema Sudeste/Centro-Oeste a restrição não foi atendida apenas no mês Setembro. No subsistema Nordeste o problema encontra-se nos meses de Junho e Outubro. Para o subsistema Norte a restrição foi violada nos meses de Março, Junho, Julho e Agosto.

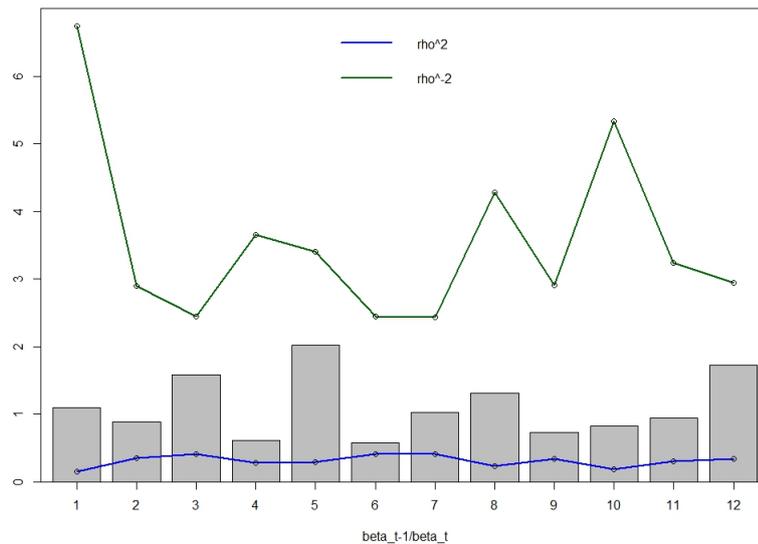


Figura 5.31: Avaliação da restrição do modelo PGAR no Subsistema Sul

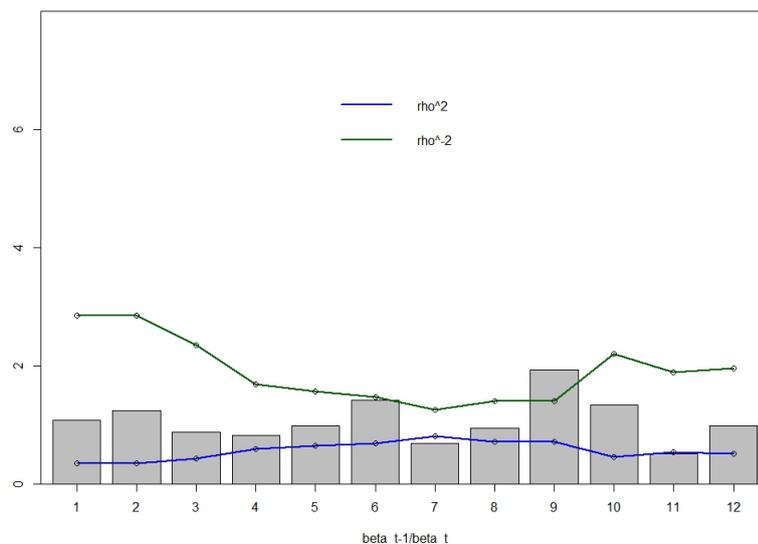


Figura 5.32: Avaliação da restrição do modelo PGAR no Subsistema Sudeste/Centro-Oeste

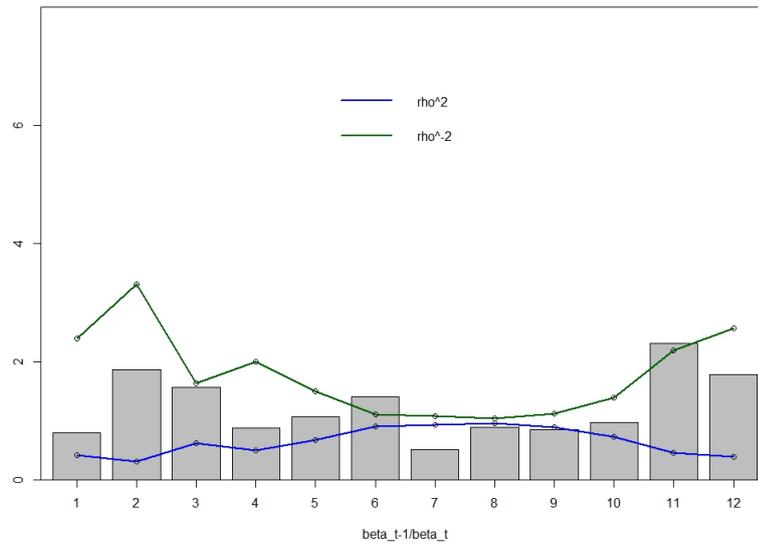


Figura 5.33: Avaliação da restrição do modelo PGAR no Subsistema Nordeste

PUC-Rio - Certificação Digital Nº 1412776/CA

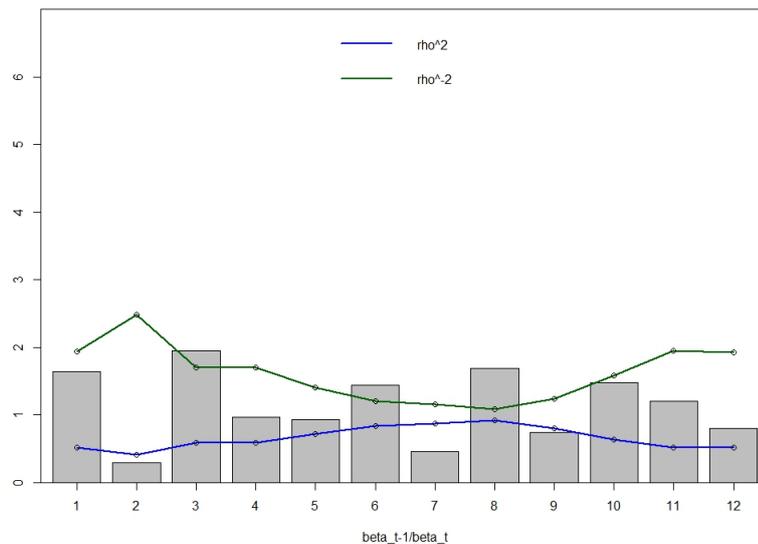


Figura 5.34: Avaliação da restrição do modelo PGAR no Subsistema Norte

Até o momento foram estudados métodos que diagnosticaram a eficácia da simulação. Para isto, testes e restrições do modelo foram impostos e comparados com o histórico observado.

Como análise alternativa, a seguir são plotados box-plots referentes a simulação em conjunto com o histórico de mesmo período (2004 à 2013) para visualizar seu comportamento com o intervalo interquartil (Primeiro e Terceiro Quartil). As figuras (Figura 5.35; Figura 5.36), (Figura 5.37; Figura 5.38), (Figura 5.39; Figura 5.40) e (Figura 5.41; Figura 5.42) ilustram esta situação.

Conforme se observa, o subsistema Sul apresenta a maior irregularidade, corroborando com o que foi visto na seção anterior. Além disso, os valores simulados para o subsistema Sudeste/Centro-Oeste não se mostram ajustados ao histórico, demonstrando em alguns períodos valores muito aquém do ideal. Contudo, pode-se dizer que o histórico do subsistema Nordeste e Norte acompanharam com êxito a simulação.

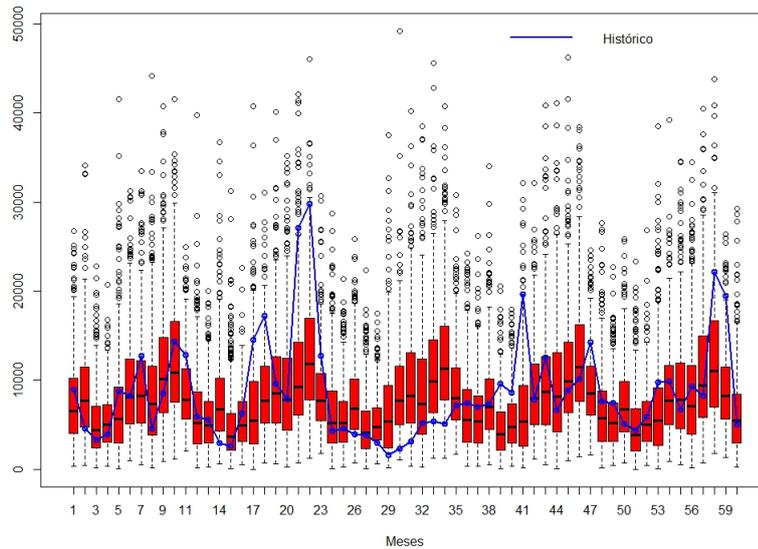


Figura 5.35: Box-plot da Simulação Sul dos meses 1 à 60

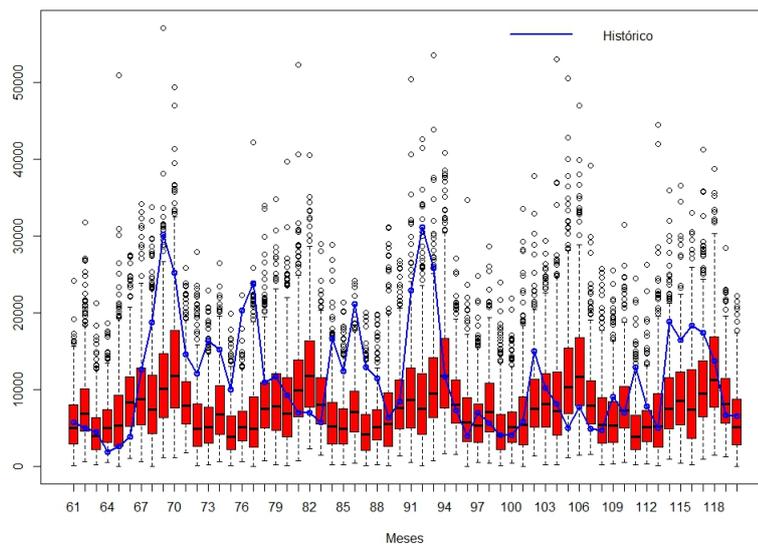


Figura 5.36: Box-plot da Simulação Sul dos meses 61 à 120

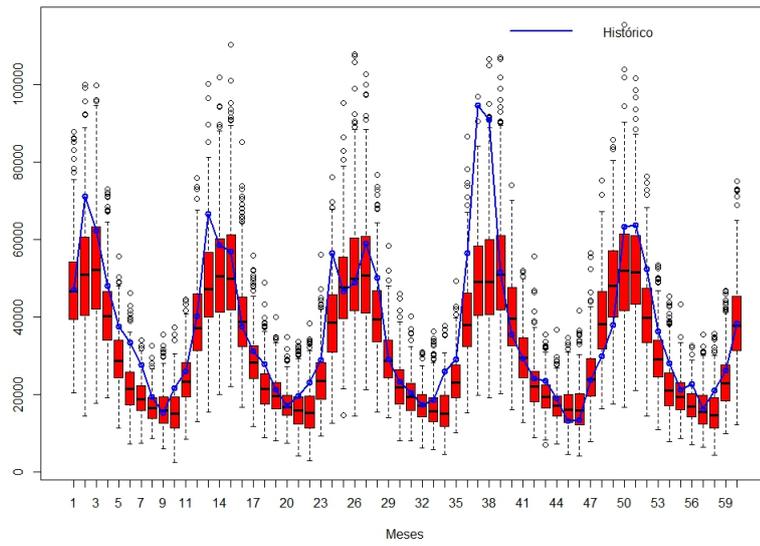


Figura 5.37: Box-plot da Simulação Sudeste/Centro-Oeste dos meses 1 à 60

PUC-Rio - Certificação Digital Nº 1412776/CA

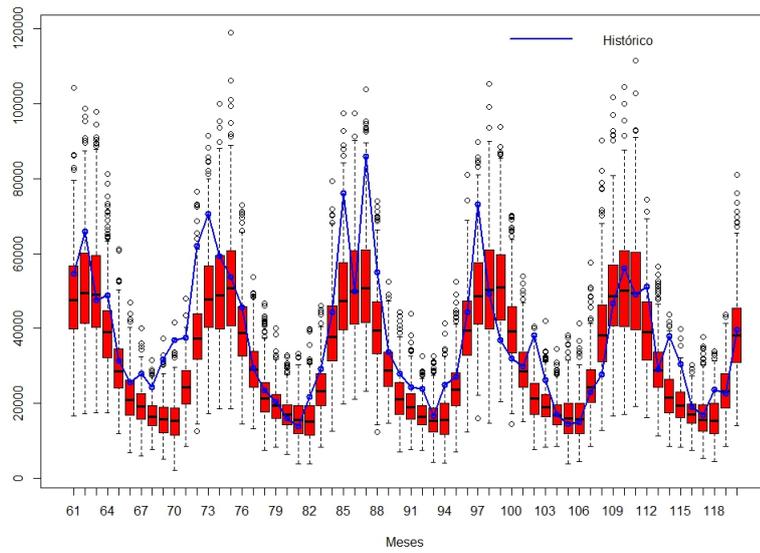


Figura 5.38: Box-plot da Simulação Sudeste/Centro-Oeste dos meses 61 à 120

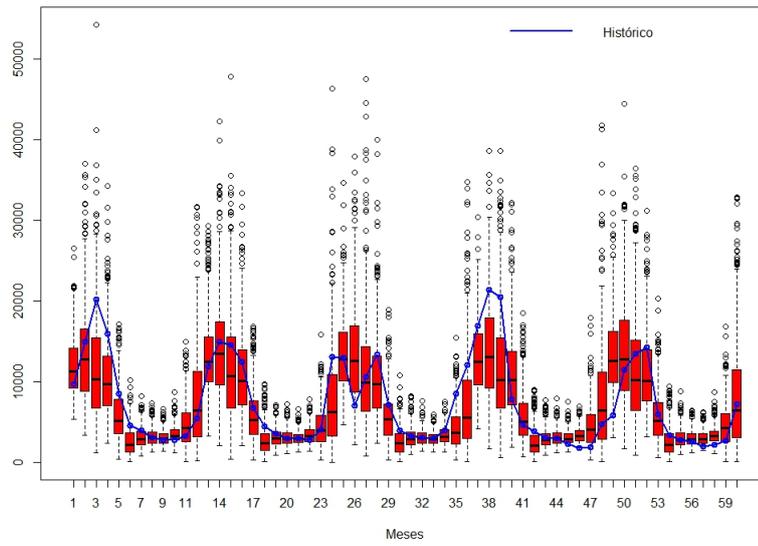


Figura 5.39: Box-plot da Simulação Nordeste dos meses 1 à 60

PUC-Rio - Certificação Digital Nº 1412776/CA

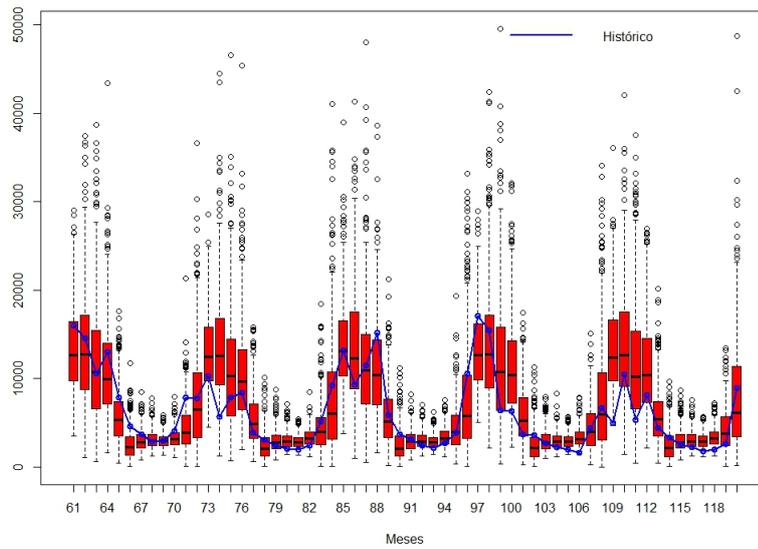


Figura 5.40: Box-plot da Simulação Nordeste dos meses 61 à 120

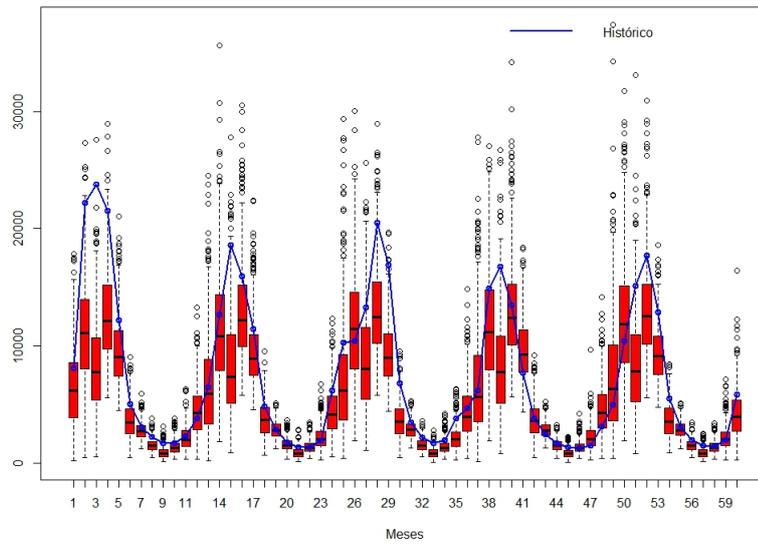


Figura 5.41: Box-plot da Simulação Norte dos meses 1 à 60

PUC-Rio - Certificação Digital N° 1412776/CA

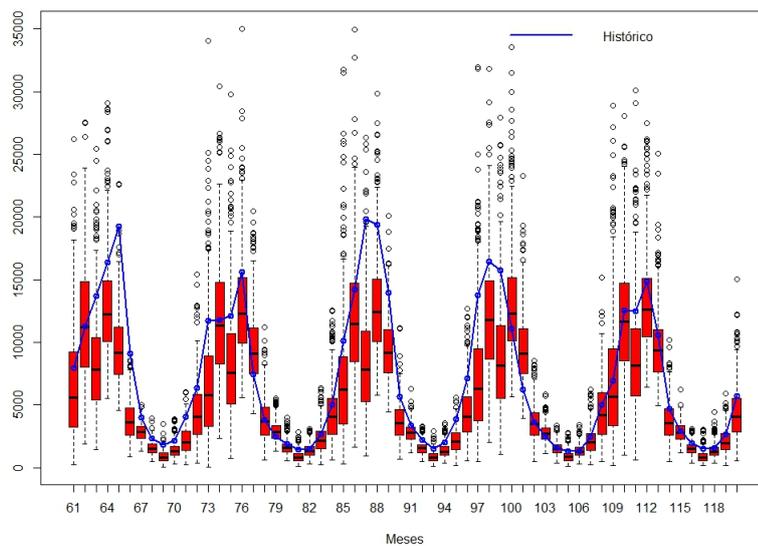


Figura 5.42: Box-plot da Simulação Norte dos meses 61 à 120

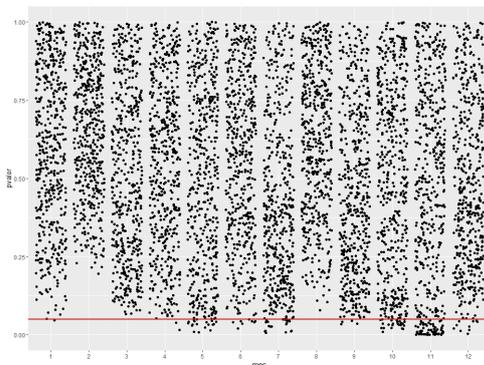
## 5.4 A Modelagem Aditiva e Multiplicativa

No contexto de séries temporais, é comum admitir métodos aditivos ou multiplicativos, exclusivamente, para ajustes de modelos. Conforme relatado nas seções anteriores, (Fernandez, 1984) propôs a integração destes dois métodos na busca por melhores ajustes para dados assimétricos em séries periódicas.

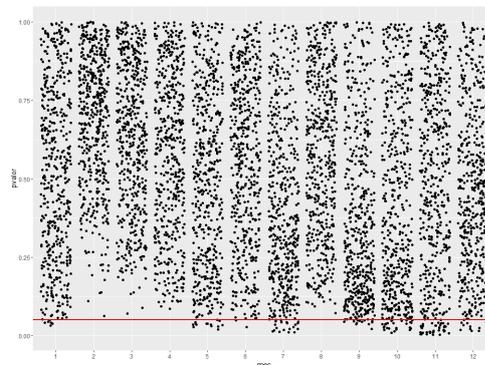
A presente seção tem por finalidade apresentar estes métodos como forma alternativa de análise, comparando-os entre si e com o PGAR(1). Além disso, esta modelagem admitirá a não variabilidade da assimetria, tornando o parâmetro  $\beta_t$  constante.

Seguindo de forma análoga as análises do PGAR(1), as figuras (Figura 5.43), (Figura 5.44), (Figura 5.45) e (Figura 5.46) representam o teste de Levene de cada subsistema para o método aditivo (PAGAR(1)) e também para o multiplicativo (PMGAR(1)).

Nota-se, ao nível de significância de 5% (linha vermelha), o bom comportamento das variâncias dos subsistemas, período a período, com o histórico. A partir disto, a etapa seguinte consiste no teste de Mann-Whitney para checar o comportamento dos cenários com o histórico.

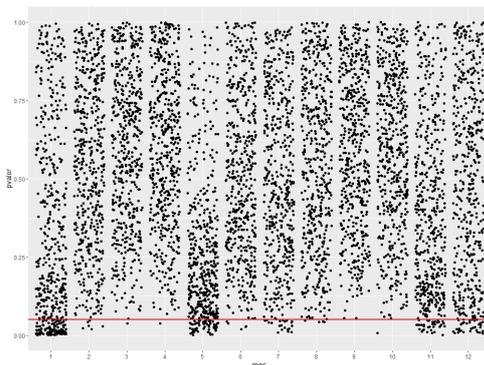


5.43(a): Modelo Aditivo

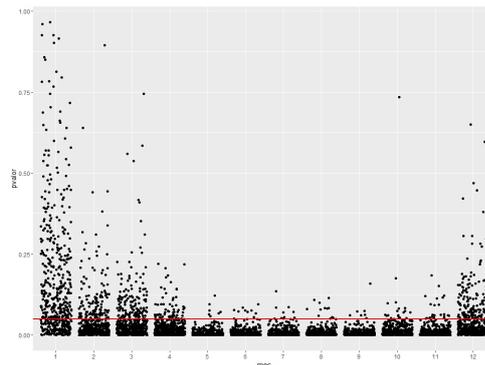


5.43(b): Modelo Multiplicativo

Figura 5.43: Teste de Levene para o PAGAR(1) e PMGAR(1) no subsistema Sul

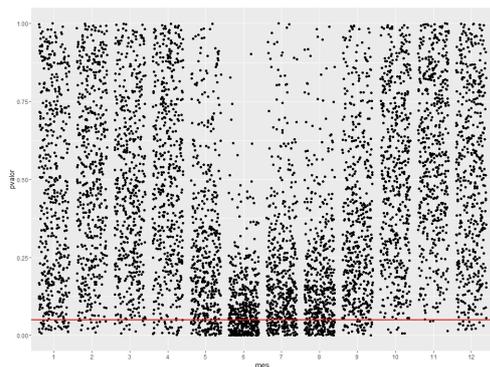


5.44(a): Modelo Aditivo

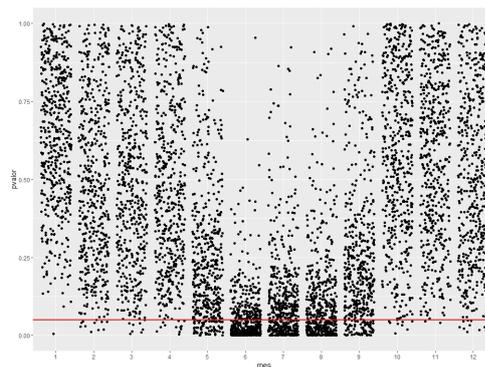


5.44(b): Modelo Multiplicativo

Figura 5.44: Teste de Levene para o PAGAR(1) e PMGAR(1) no subsistema Sudeste/Centro-Oeste

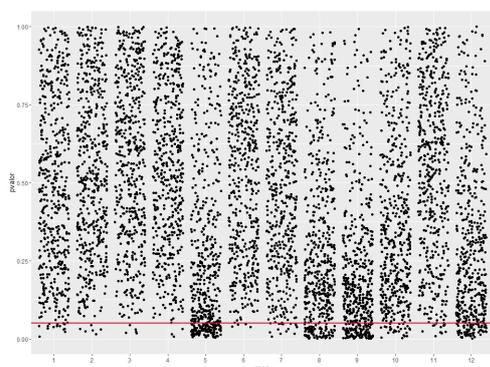


5.45(a): Modelo Aditivo

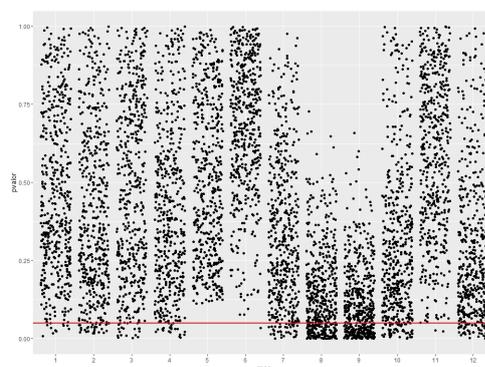


5.45(b): Modelo Multiplicativo

Figura 5.45: Teste de Levene para o PAGAR(1) e PMGAR(1) no subsistema Nordeste



5.46(a): Modelo Aditivo



5.46(b): Modelo Multiplicativo

Figura 5.46: Teste de Levene para o PAGAR(1) e PMGAR(1) no subsistema Norte

Analisando os resultados da tabela (Tabela 5.8) nota-se que o modelo PAGAR(1) ajustou-se melhor ao histórico, período a período, em todos os subsistemas. Os resultados tornam-se ainda mais evidentes ao compará-los com o PMGAR(1). Estes resultados também evidenciam que períodos considerados secos mostraram uma maior "resistência" ao teste para o subsistema Sudeste/Centro-Oeste. Além disso, em uma comparação entre o PAGAR(1) e PGAR(1), os resultados indicaram melhores ajustes para o modelo aditivo.

Período	Subsistemas									
	Sul		Sudeste/Centro-Oeste		Nordeste		Norte			
	Aditivo	Multiplicativo	Aditivo	Multiplicativo	Aditivo	Multiplicativo	Aditivo	Multiplicativo		
Janeiro	100%	78%	100%	87%	100%	20%	100%	43%		
Fevereiro	100%	97%	98%	74%	100%	64%	97%	17%		
Março	100%	90%	99%	85%	100%	83%	99%	4%		
Abril	100%	95%	94%	61%	100%	37%	99%	1%		
Maior	100%	97%	79%	47%	100%	67%	100%	1%		
Junho	100%	86%	76%	40%	99%	24%	100%	3%		
Julho	98%	88%	57%	27%	95%	13%	100%	0%		
Agosto	100%	79%	74%	34%	82%	17%	94%	1%		
Setembro	100%	69%	100%	7%	87%	16%	88%	1%		
Outubro	100%	63%	100%	59%	100%	15%	100%	7%		
Novembro	100%	75%	97%	59%	100%	52%	100%	24%		
Dezembro	100%	91%	100%	67%	100%	20%	100%	27%		

Tabela 5.8: Proporção de p-valores do Teste de Mann-Whitney superiores a 1% dos Subsistemas para o Modelo Aditivo e Multiplicativo

A partir destes primeiros resultados, o próximo passo representa a simulação de cada subsistema do modelo de melhor resultado. Partindo da mesma proposta feita na seção anterior, as figuras a seguir, seguem o comportamento sintético de cada subsistema em conjunto com sua respectiva variabilidade.

De modo geral, pode-se dizer que os cenários apresentam um bom ajuste para todos os subsistemas. Além do mais, os cenários revelaram valores superiores ao modelo PGAR(1), podendo assim, capturar melhor o histórico. No entanto, é importante destacar que uma pequena parte do histórico no subsistema Nordeste não capturou a simulação.

Em relação ao desvio-padrão, nota-se que os valores gerados sobrepõem-se aos valores observados na maioria dos períodos. Muito embora as variâncias sejam estatisticamente iguais, é importante relatar que, visualmente, devido a alta irregularidade do subsistema Sul, seu desvio-padrão simulado gera certo grau de "dificuldade" quanto ao acompanhamento do histórico em comparação com os demais subsistemas.

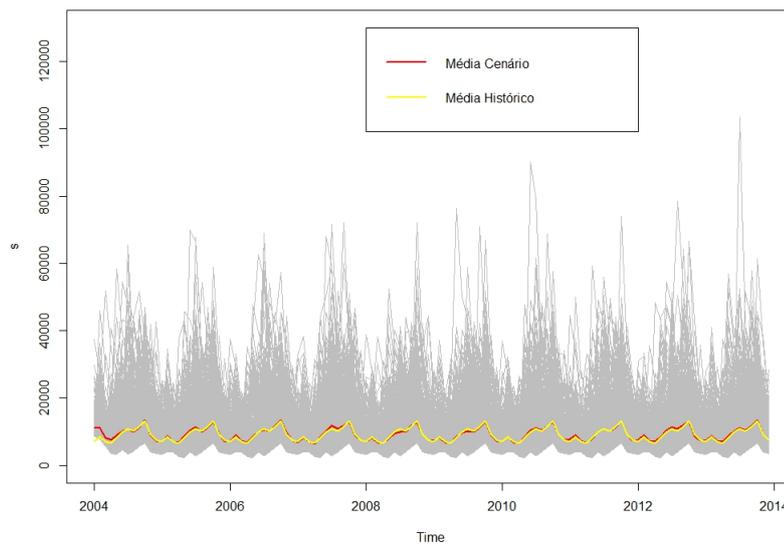


Figura 5.47: Cenário PAGAR(1) no Subsistema Sul

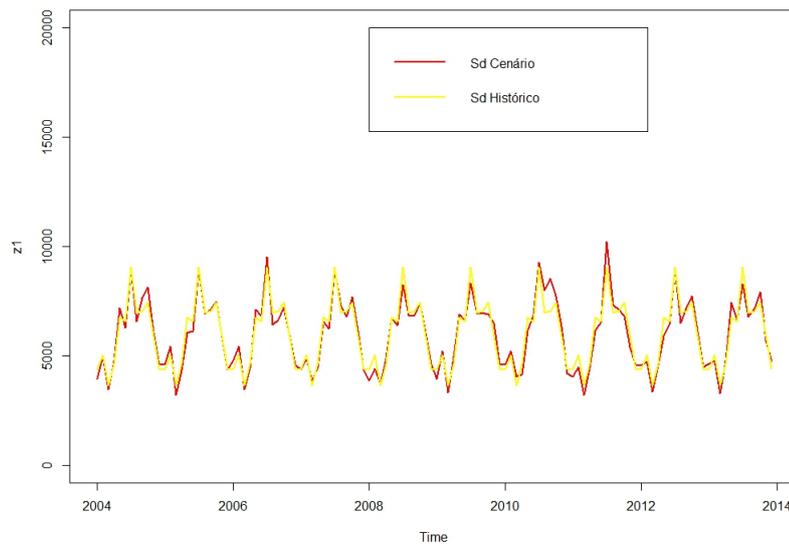


Figura 5.48: Desvio-Padrão PAGAR(1) no Subsistema Sul

PUC-Rio - Certificação Digital N° 1412776/CA

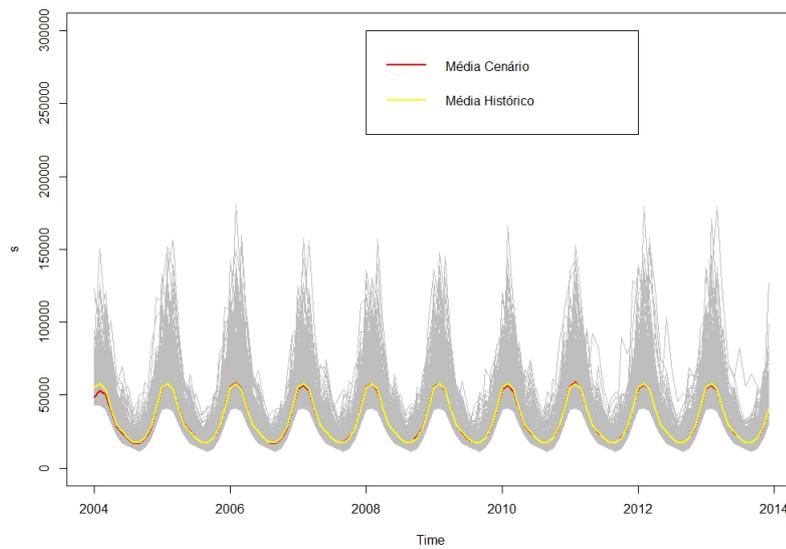


Figura 5.49: Cenário PAGAR(1) no Subsistema Sudeste/Centro-Oeste

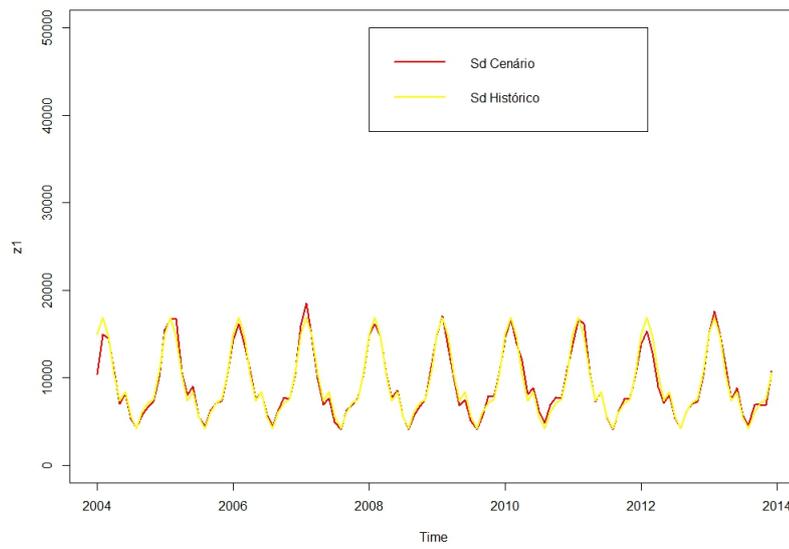


Figura 5.50: Desvio-Padrão PAGAR(1) no Subsistema Sudeste/Centro-Oeste

PUC-Rio - Certificação Digital N° 1412776/CA

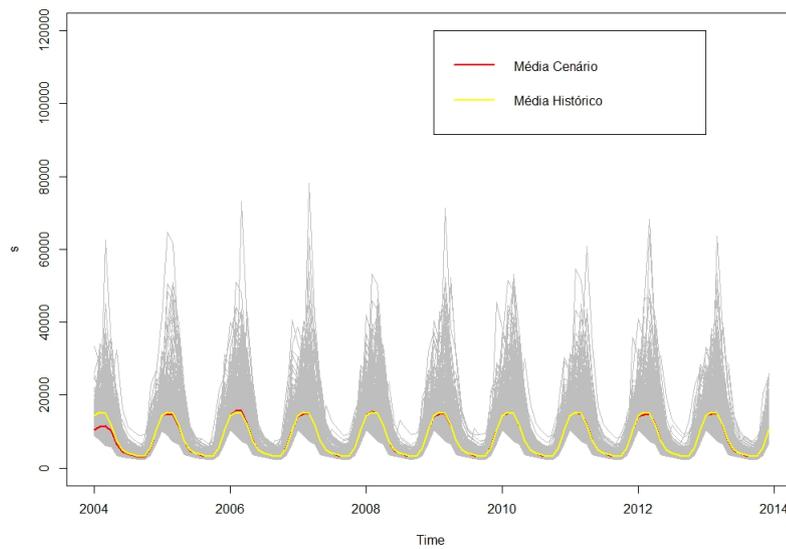


Figura 5.51: Cenário PAGAR(1) no Subsistema Nordeste

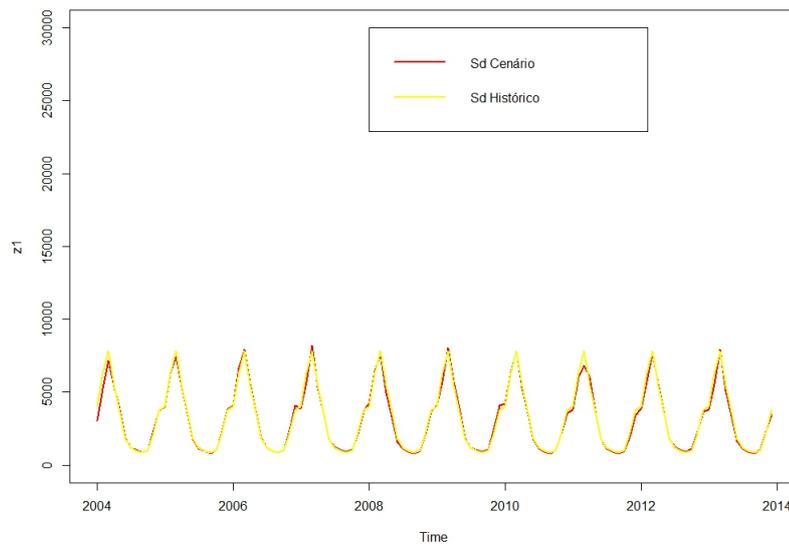


Figura 5.52: Desvio-Padrão PAGAR(1) no Subsistema Nordeste

PUC-Rio - Certificação Digital Nº 1412776/CA

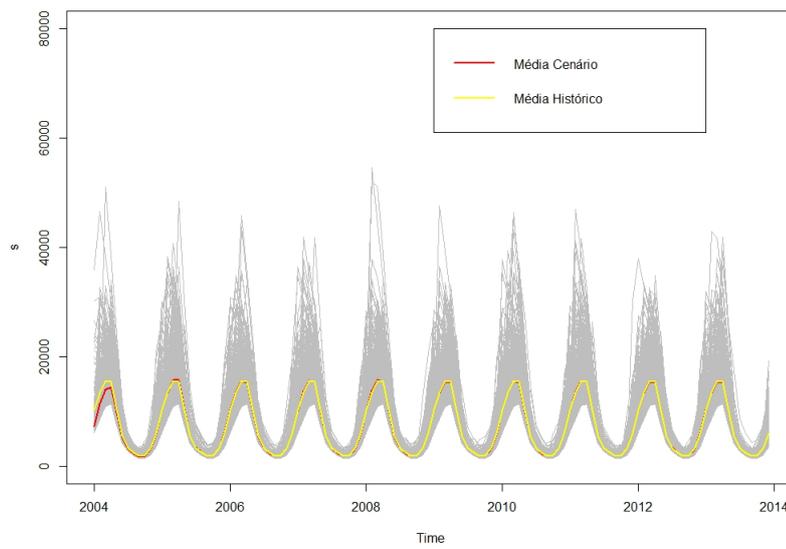


Figura 5.53: Cenário PAGAR(1) no Subsistema Norte

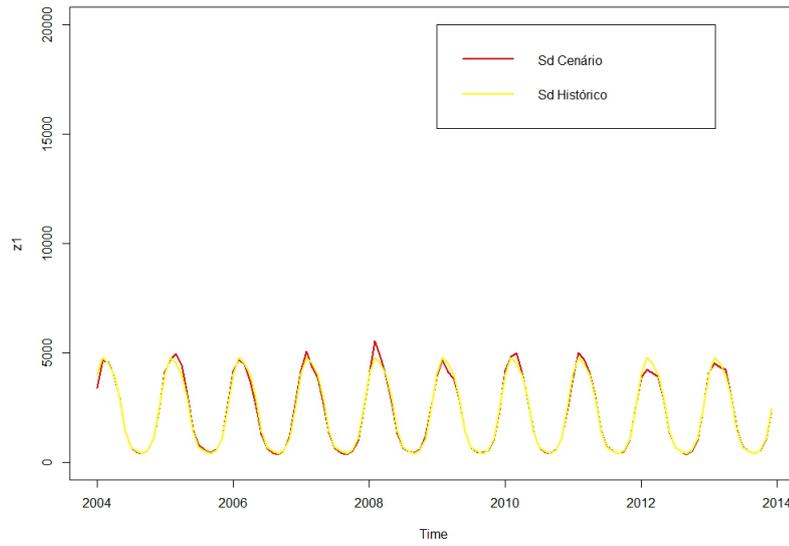


Figura 5.54: Desvio-Padrão PAGAR(1) no Subsistema Norte

## 6

### Considerações Finais e Trabalhos futuros

Ao longo deste trabalho foi estudada uma metodologia, desenvolvida por (Fernandez, 1984) e (Fernandez & Salas, 1986), cujo intuito era ser tão eficiente quanto a consolidada pelo SEB. Os autores propuseram um modelo capaz de incorporar métodos aditivos e multiplicativos em uma única estrutura, denominado PGAR(1). Estes mesmos autores consideraram a possibilidade de se aplicar o uso destes métodos de modo isolado, ou seja, o modelo simplesmente aditivo (PAGAR(1)) ou simplesmente multiplicativo (PMGAR(1)).

A modelagem PGAR(1) demonstrou, a priori, certa fragilidade quando aplicada em séries de vazões e/ou ENAs. Estes problemas podem estar interligados à natureza dos dados, uma vez que diversos outliers foram encontrados ao longo dos períodos. Sendo assim, os subsistemas não corresponderam de forma satisfatória ao que foi proposto por (Fernandez, 1984) e (Fernandez & Salas, 1986).

De início, a modelagem encontrou problemas para estimação dos parâmetros. Tal problema foi contornado empregando outro método consolidado na literatura, conhecido como método de máxima verossimilhança. A partir disto, deu-se continuidade ao processo de simulação. Um dos pontos positivos desta metodologia, comprovado neste trabalho, refere-se a simulação de resíduos não só assimétricos, mas também não negativo.

Com relação à simulação dos cenários, foi observado que o PGAR(1) não reproduziu bem os valores observados nos subsistemas. Como medida alternativa, o trabalho partiu para métodos exclusivamente aditivos e multiplicativos, denominados PAGAR(1) e PMGAR(1), utilizando distribuições marginais exponenciais. Tais métodos demonstraram serem mais robustos do que os resultados do PGAR(1). Comparando os dois métodos, o PAGAR(1) apresentou resultados mais eficientes.

O modelo PAGAR(1) foi capaz de reproduzir satisfatoriamente o histórico. No que diz respeito a parte estatística, o modelo reproduziu bem o histórico levando a uma aceitação da proporção de p-valores a 100% para a maioria dos períodos nos subsistemas.

Toda a metodologia deste trabalho foi devidamente programada em um ambiente estatístico e possibilita que outras séries possam ser trabalhadas neste contexto. Além disso, outro ponto positivo deixado pelo trabalho refere-se a possibilidade de contribuição para a implementação de pacotes acadêmicos sobre o tema.

Novas ordens dos modelos propostos no trabalho também podem ser estudadas

para verificar a eficácia do modelo quanto ao histórico das séries hidrológicas de vazões e/ou ENAs.

Uma outra possibilidade é checar a existência de distintas transformações matemáticas além das utilizadas pelos autores na construção do modelo.

Além disso, diferentes valores para o parâmetro  $\beta_t$  também podem ser aferidos na finalidade de melhorar os resultados encontrados para o PAGAR(1).

Embora o SEB adote modelagens com propostas Gaussianas e trabalhe com séries de interconfigurações, isto é, com o histórico sofrendo mudanças ao longo do tempo devido a entrada e retirada de usinas, o presente trabalho abre um leque de possibilidades para estudo de Modelos Não-Lineares devido a sua eficácia comprovada, podendo, inclusive, ser estendida para os métodos atuais do SEB. Além disso, outros trabalhos futuros podem ser empregados nesta direção, abordando novas distribuições assimétricas.

## Referências Bibliográficas

- [1] Arnold T. B. , Emerson J. W. (2011). Nonparametric Goodness-of-Fit Tests for Discrete Null Distributions. *The R Journal*, Vol. 3/2, December 2011, URL:[http : //journal.r – project.org/archive/2011 – 2/RJournal2011 – 2Arnold+Emerson.pdf](http://journal.r-project.org/archive/2011-2/RJournal2011-2Arnold+Emerson.pdf)
- [2] Becker M. & Klößner S. (2013). PearsonDS: Pearson Distribution System. R package version 0.98. URL <http://CRAN.R-project.org/package=PearsonDS>.
- [3] Braga, Diogo & Calmon, W. (2016). Periodic gamma autoregressive model: an application to the brazilian hydroelectric system *RAIRO Operations Research*, DOI: 10.1051/ro/2016035. No prelo.
- [4] Casella, G. & Berger, R.L. (2001). Statistical Inference. *Duxbury Press*.
- [5] Cavalcanti Netto, M. A. (2015). A Previsão com a Metodologia de Box-Jenkins. Disponível em: <http://www.ie.ufrj.br/download/APrevisaoComMetodologiadeBox-Jenkins.pdf>. Acesso em Setembro de 2015.
- [6] CEPEL. (2011) Manual de referência do modelo NEWAVE. Rio de Janeiro, Brazil: Centro de Pesquisas em Energia Elétrica.
- [7] Chebaane, M., Salas J.D. & Boes D.C., (1995), Product Periodic Autoregressive Processes for Modeling Intermittent Monthly Streamflows, *Water Resour. Res.*, 31(6):1513-1518.
- [8] Choi, S. C. & R. Wette, (1969), Maximum likelihood estimation of parameters of the gamma distribution and their bias, *Technometrics*, 11(14), 683-690.
- [9] Efron, B. & Tibshirani R.J., (1994), An Introduction to the Bootstrap, *Chapman and Hall/CRC*.
- [10] Fernandez, B. (1984). Stochastic modeling of periodic streamflow series with gamma distributions. Ph.D. dissertation *Colorado State University*, Fort Collins.
- [11] Fernandez, B., & Salas, J. (1990). Gamma-Autoregressive Models for StreamFlow Simulation *Journal of Hydraulic Engineering* 116.
- [12] Fernandez, B., & Salas, J. D. (1986). Periodic Gamma Autoregressive Processes for Operational Hydrology *Water Resources Research*,22, pp. 1385-1396.

- [13] Ferreira, P.G.C. (2013) A estocasticidade associada ao Setor Elétrico Brasileiro e uma nova abordagem para a geração de aflúências via modelos periódicos Gama. *Tese (Doutorado) - DEE, PUC-Rio, Rio de Janeiro.*
- [14] Griffin, W.C. (1975). *Transform Techniques for Probability Modeling*, Academic, New York, Orlando, Fla.
- [15] Groisman, P. Y.; Knight, R. W.; Easterling, D. R.; Karl, T. R.; Hegerl, Gabi & Razuvaev, V. A. N. Trends in intense precipitation in the climate record, *Journal of Climate*, n.18, p.1326-50, 2005.
- [16] Hipel, K. W., & McLeod, A. I. (1994). *Time series modelling of water resources and environmental systems*. Amsterdam, The Netherlands: Elsevier
- [17] Jones, R. H., & Brelsford, W. M. (1967). Time Series with periodic structure. *Biometrika* 54, pp. 403-408.
- [18] Kelman, J. (2001). Relatório da Comissão de Análise do Sistema Hidrotérmico de Energia Elétrica, Brasília, Brazil.
- [19] Kite, G. W. (1977). Frequency and Risk analysis in hydrology. *Water Resources Publications*
- [20] Komsta L. & Novomestky F. (2015). moments: Moments, cumulants, skewness, kurtosis and related tests. R package version 0.14. <http://CRAN.R-project.org/package=moments>.
- [21] Lawrance, A. J., & Lewis, P. A. (1981). A new autoregressive time series model in exponential variables (NEAR(1)). *Adv. Appl. Probab.*, pp. 826-845.
- [22] 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*, Budapeste, Hungary.
- [23] 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*.
- [24] Marengo, J. A., Alves, L. M., Valverde, M. C., Rocha, R. P. & Laborbe, R.(2007). Eventos extremos em cenários regionalizados de clima no Brasile América do Sul para o Século XXI: Projeções de clima futuro usando três modelos regionais.Eventos extremos em cenários regionalizados de clima no Brasile América do Sul para o Século XXI: Projeções de clima futuro usando três modelos regionais ? Subprojeto: Caracterização do clima atual e definição das alterações climáticas para o território brasileiro ao longo do século XXI. Brasília, fevereiro 2007.
- [25] Marengo, J. A.(2008). Água e mudanças climáticas. *Estud. av.*, São Paulo, v. 22, n. 63, p. 83-96.

- [26] McKenzie, E. (1982). Product Autoregression: A Time-Series characterization of the Gama distribution. *J. Appl. Prob.*, 19, pp. 463-468.
- [27] McLeod, A. I. (1994). Diagnostic Checking of Periodic Autoregression *Journal of Time Series Analysis*, Vol. 15, No. 2 pp.221–233.
- [28] Oliveira, F.L.C. (2010), Nova abordagem para geração de cenários de afluências no planejamento da operação energética de médio prazo, *Dissertação (Mestrado) - DEE, PUC-Rio, Rio de Janeiro*.
- [29] ONS. (2015). *www.ons.org.br*. Acesso em Outubro de 2015.
- [30] 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.
- [31] Pereira, M. V., Campodónico, N., & Kelman, R. (1998). Long-term Hydro Scheduling based on Sthocastic Models *EPSOM'98*.
- [32] Salas, J. D., & Smith, R. A. (1981). Computer Programs of Probability Distribution in Hydrology. *Department of Civil Engineering, Colorado State University, Fort Collins*.
- [33] Salas, J. D., & Obeysekera, J. T. (1982). ARMA Model Identification of Hydrologic Time Series. *Water Resources Research*, 18, pp. 1011-1021.
- [34] Şarlak N. & Şorman A.Ü., (2007), Gamma Autoregressive Models and Application on the K?z?l?rmak Basin, *Published in Teknik Dergi*, Vol. 18, pp: 4219-4227.
- [35] Shanbhag, D.N., & Srechari, M. (1977). On certain self-decomposable distribution. *Z. Wahr. Verw. Geb.* , 38, 217 ? 222.
- [36] 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.
- [37] Venables, W. N. & Ripley, B. D. (2002). Modern Applied Statistics with S.Fourth Edition. *Springer*, New York. ISBN 0-387-95457-0.

## 7 Apêndice

### Bloco 1

```
> dados_norte<-read.csv2("norte.csv",header=T)
> fgam<-fitdist(dados_norte,"gamma",method="mme")
> hist(dados_norte,breaks=50,prob=T,col='blue',
+      main="Histograma do Subistema norte" ,
+      xlab="norte",ylab="Densidade")
> curve(dgamma(x,fgam$estimate[1],fgam$estimate[2])
+       ,add=T,col="red")
> dados_norte<-ts(dados_norte,start=1931,frequency=12)
> plot(dados_norte)
> peboxplot(dados_norte,col="red")
```

### Bloco 2

```
> fjan<-fitdist(subset(dados_norte, cycle(dados_norte) == 1)
+              ,"gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 1)
+         ,"pgamma",fjan$estimate[1],fjan$estimate[2])

> ffev<-fitdist(subset(dados_norte, cycle(dados_norte) == 2)
+              ,"gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 2)
+         ,"pgamma",ffev$estimate[1],ffev$estimate[2])

> fmar<-fitdist(subset(dados_norte, cycle(dados_norte) == 3)
+              ,"gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 3),
+         "pgamma",fmar$estimate[1],fmar$estimate[2])

> fabr<-fitdist(subset(dados_norte, cycle(dados_norte) == 4)
+              ,"gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 4),
+         "pgamma",fabr$estimate[1],fabr$estimate[2])
```

```
> fmai<-fitdist(subset(dados_norte, cycle(dados_norte) == 5),
+              "gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 5),
+         "pgamma",fmai$estimate[1],fmai$estimate[2])

> fjun<-fitdist(subset(dados_norte, cycle(dados_norte) == 6),
+              "gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 6),
+         "pgamma",fjun$estimate[1],fjun$estimate[2])

> fjul<-fitdist(subset(dados_norte, cycle(dados_norte) == 7),
+              "gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 7),
+         "pgamma",fjul$estimate[1],fjul$estimate[2])

> fago<-fitdist(subset(dados_norte, cycle(dados_norte) == 8),
+              "gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 8),
+         "pgamma",fago$estimate[1],fago$estimate[2])

> fset<-fitdist(subset(dados_norte, cycle(dados_norte) == 9),
+              "gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 9),
+         "pgamma",fset$estimate[1],fset$estimate[2])

> fout<-fitdist(subset(dados_norte, cycle(dados_norte) == 10),
+              "gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 10),
+         "pgamma",fout$estimate[1],fout$estimate[2])

> fnov<-fitdist(subset(dados_norte, cycle(dados_norte) == 11),
+              "gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 11),
+         "pgamma",fnov$estimate[1],fnov$estimate[2])

> fdez<-fitdist(subset(dados_norte, cycle(dados_norte) == 12),
+              "gamma",method="mme")
> ks.test(subset(dados_norte, cycle(dados_norte) == 12),
+         "pgamma",fdez$estimate[1],fdez$estimate[2])
```

```

> Q<-apply(meses,2,mean)
> g<-apply(meses,2,skewness)
> S<-apply(meses,2,sd)
> autocor_lag1<-function(dados,dado_defasado){
+   r<-NULL
+   for( i in 1:12){
+     if(i==1){
+       r[i]<-cor(meses[,i],dado_defasado)
+     }else{
+       r[i]<-cor(meses[,i],meses[i-1])
+     }
+   }
+   return(r)
+ }

rho<-autocor_lag1(meses,dado_defasado)

```

#### Bloco 4

```

> param_fx<-function(parametros){
+   lambda<-NULL
+   alpha<-NULL
+   beta<-NULL
+   for(i in 1:dim(meses)[2]){
+     lambda[i]<-parametros$Q[i] -
+       2*(parametros$S[i]/parametros$g[i])
+     alpha[i]<-(parametros$S[i]*parametros$g[i])/2
+     beta[i]<- 4/(parametros$g[i]^2)
+   }
+   return(list(lambda,alpha,beta))
+ }

> lambda<-param_fx(parametros)[[1]]
> alpha<-param_fx(parametros)[[2]]
> beta<-param_fx(parametros)[[3]]

```

#### Bloco 5

```

> phi_delta_fx<-function(betatal,alphatal,rhotal){
+   phi<-NULL
+   delta<-NULL
+   for(i in 2:13){

```

```

+   if(betatal[i] >= betatal[i-1]){
+     phi[i-1]<-rhotal[i]*(alphatal[i]/alphatal[i-1])*
+       ((betatal[i]/betatal[i-1])^1/2)
+     delta[i-1]<-0
+   }else{
+     phi[i-1]<-0
+     delta[i-1]<-rhotal[i]*((betatal[i-1]/betatal[i])^1/2)
+   }
+ }
+ return(list(phi,delta))
+ }

phital<-phi_delta_fx(betatal,alphatal,rhotal)[[1]]
deltatal<-phi_delta_fx(betatal,alphatal,rhotal)[[2]]

```

### Bloco 6

```

> distribuicao<-function(x,lambda,alpha,beta){
+   ((alpha*gamma(beta))^-1)*(((x-lambda)/alpha)^
+     (beta-1))*exp(-(x-lambda)/alpha)
+ }

```

### Bloco 7

```

param_est<-matrix(,nrow=12,ncol=3)

param_est[1,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 1)
  +     , distribuicao, list(beta = chute_beta
  +     , alpha = chute_alpha, lambda= chute_lambda)
  +     , lower = 0.001)$estimate)

param_est[2,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 2)
  +     , distribuicao, list(beta = chute_beta
  +     , alpha = chute_alpha, lambda= chute_lambda)
  +     , lower = 0.001)$estimate)

param_est[3,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 3)
  +     , distribuicao, list(beta = chute_beta

```

```
      +, alpha = chute_alpha, lambda= chute_lambda)
      +, lower = 0.001)$estimate)

param_est[4,]<-as.vector(fitdistr(subset(
  + dados, cycle(dados) == 4)
  +, distribuicao, list(beta = chute_beta
  +, alpha = chute_alpha, lambda= chute_lambda)
  +, lower = 0.001)$estimate)

param_est[5,]<-as.vector(fitdistr(subset(
  + dados, cycle(dados) == 5)
  +, distribuicao, list(beta = chute_beta
  +, alpha = chute_alpha, lambda= chute_lambda)
  +, lower = 0.001)$estimate)

param_est[6,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 6)
  +, distribuicao, list(beta = chute_beta
  +, alpha = chute_alpha, lambda= chute_lambda)
  +, lower = 0.001)$estimate)

param_est[7,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 7)
  +, distribuicao, list(beta = chute_beta
  +, alpha = chute_alpha, lambda= chute_lambda)
  +, lower = 0.001)$estimate)

param_est[8,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 8)
  +, distribuicao, list(beta = chute_beta
  +, alpha = chute_alpha, lambda= chute_lambda)
  +, lower = 0.001)$estimate)

param_est[9,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 9)
  +, distribuicao, list(beta = chute_beta
  +, alpha = chute_alpha, lambda= chute_lambda)
  +, lower = 0.001)$estimate)

param_est[10,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 10)
```

```

      +, distribuicao, list(beta = chute_beta
      +, alpha = chute_alpha, lambda= chute_lambda)
      +, lower = 0.001)$estimate)

param_est[11,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 11)
      +, distribuicao, list(beta = chute_beta
      +, alpha = chute_alpha, lambda= chute_lambda)
      +, lower = 0.001)$estimate)

param_est[12,]<-as.vector(fitdistr(subset(
dados, cycle(dados) == 12)
      +, distribuicao, list(beta = chute_beta
      +, alpha = chute_alpha, lambda= chute_lambda)
      +, lower = 0.001)$estimate)

colnames(param_est)<-c('betatal', 'alphatal', 'lambdatal')

```

### Bloco 8

```

> funcao_erro<-function(betatal,phital,alphatal,lambdatal){
+   ztal<-NULL
+   wtal<-NULL
+   vtal<-NULL
+   etal<-NULL
+
+   for(i in 2:13){
+     if (betatal[i]<betatal[i-1]){
+       ytal<-rbeta(1,betatal[i], betatal[i-1]-betatal[i])
+       e<- -digamma(1)
+       prob<-runif(1)
+
+       if(prob<=deltatal[i-1]){
+         vlinhatal<- -e*(1-deltatal[i-1])
+       }else{
+         vlinhatal<- -e*(1-deltatal[i-1]) - rexp(1,betatal[i])
+       }
+       somavtalk<-0
+       for (k in 1:100){
+         prob<-runif(1)
+         if (prob<=deltatal[i-1]){
+           vtalk<- (1-deltatal[i-1])/k

```

```

+     }else{
+         vtalk<- ((1-deltatal[i-1])/k)
+             - (rexp(1,(betatal[i]+k)))
+     }
+     somavtalk<-somavtalk+vtalk
+ }
+
+ ztal<- exp(vlinhatal+somavtalk)
+ vtal<-(alphatal[i]/(alphatal[i-1]^deltatal[i-1]))*
+     ztal*(ytal^deltatal[i-1])
+ wtal[i-1]<-vtal
+
+ }else{
+
+     etalzero<-rgamma3(1,shape=betatal[i]-betatal[i-1],
+         scale =alphatal[i]
+         ,thres=lambdatal[i]
+         -phital[i-1]*lambdatal[i-1])
+     theta <- -betatal[i-1]*log(phital[i-1]
+         *(alphatal[i-1]/alphatal[i]))
+     if(theta<0){
+         theta<-0
+     }
+     N <- rpois(1,theta)
+     if(N != 0){
+         etalum<-0
+         for(j in 1:N){
+             ym <- rexp(1,alphatal[i])
+             um <- runif(1)
+             etalum <- etalum + ym*(phital[i-1]*
+                 (alphatal[i-1]/alphatal[i]))^um
+         }
+     }
+
+     }else{
+         etalum<-0
+     }
+     wtal[i-1] <- etalzero + etalum
+
+ }
+ }
+ return(wtal)

```

```
+ }
```

### Bloco 9

```
> if(theta<0){
+   theta<-0
+ }
```

### Bloco 10

```
erro_aditivo<-function(phital,alphatal,lambdata){
+   etal<-NULL
+   erro<-NULL
+   for(i in 2:13){ # Modelo Aditivo
+     erro<- lambdata[i] - phital[i-1]*lambdata[i-1]
+     prob<-runif(1)
+     if(prob<= phital[i-1]*(alphatal[i-1]/alphatal[i])){
+       erro<-erro + 0
+     }else{
+       erro<-erro + rexp(1,1/alphatal[i])
+     }
+     etal[i-1]<-erro
+   }
+   return(etal)
+ }
```

### Bloco 11

```
erro_multiplicativo<-function(phital,alphatal){
+   ztal<-NULL
+   wtal<-NULL
+   vtal<-NULL
+   etal<-NULL
+   for(i in 2:13){
+     e<- -digamma(1)
+
+     prob<-runif(1)
+     if(prob<=deltatal[i-1]){
+       vlinhatal<- -e*(1-deltatal[i-1])
+     }else{
+       vlinhatal<- -e*(1-deltatal[i-1]) - rexp(1,1)
+     }
+   }
+ }
```

```

+   }
+   somavtalk<-0
+   for (k in 1:100){
+
+     prob<-runif(1)
+     if (prob<=deltatal[i-1]){
+       vtalk<- (1-deltatal[i-1])/k
+     }else{
+       vtalk<- ((1-deltatal[i-1])/k
+               - (rexp(1,(1+k))))
+     }
+     somavtalk<-somavtalk+vtalk
+   }
+
+   ztal<- exp(vlinhatal+somavtalk)
+   vtal<-(alphatal[i]/(alphatal[i-1]
+     ^deltatal[i-1]))*ztal
+   wtal[i-1]<-vtal
+ }
+ return(wtal)
+ }

```

### Bloco 12

```

> ruidotal<-funcao_erro(betatal,phital,
+                       alphatal,lambdatal)
> ruido_array<-array(,dim=c(11,12,500))

> for(j in 1:500){
+   for(i in 1:11){
+     ruido_array[i,,j]<-matriz_erro(betatal,
+                                   phital,alphatal,lambdatal)
+   }
+ }

```

### Bloco 13

```

> serie<-array(,dim=c(11,12,500))
> serie[1,,]<-as.numeric(meses[72,])

> for(k in 1:500){
+   for(j in 2:11){

```

```
+   for(i in 1:12){  
+  
+     if(i==1){  
+       serie[j,i,k]<-phital[i]*(serie[j-1,12,k])  
+       +((serie[j-1,12,k])^deltatal[i])*ruído_array[j,i,k]  
+     }else{  
+       serie[j,i,k]<-phital[i]*(serie[j,i-1,k])  
+       +((serie[j,i-1,k])^deltatal[i])*ruído_array[j,i,k]  
+     }  
+   }  
+ }  
+ }
```