



GERAÇÃO DE CENÁRIOS SINTÉTICOS MENSIS DE VAZÕES
CONDICIONADAS AO EL NIÑO/OSCILAÇÃO SUL - APLICAÇÃO AO
PLANEJAMENTO DA OPERAÇÃO DE SISTEMAS HIDROTÉRMICOS

Felipe Treistman

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Civil, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Civil.

Orientadores: Otto Corrêa Rotunno Filho
Maria Elvira Piñeiro Maceira

Rio de Janeiro
Dezembro de 2018

GERAÇÃO DE CENÁRIOS SINTÉTICOS MENSAIS DE VAZÕES
CONDICIONADAS AO EL NIÑO/OSCILAÇÃO SUL - APLICAÇÃO AO
PLANEJAMENTO DA OPERAÇÃO DE SISTEMAS HIDROTÉRMICOS

Felipe Treistman

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA
CIVIL.

Examinada por:

Prof. Otto Corrêa Rotunno Filho, Ph.D.

Prof^a. Maria Elvira Piñeiro Maceira, D.Sc.

Prof. Afonso Augusto Magalhães de Araujo, D.Sc.

Prof^a. Claudine Pereira Dereczynski, D.Sc.

Dr^a. Débora Dias Jardim Penna, D.Sc.

Prof. Jerson Kelman, Ph.D.

Prof. Jorge Machado Damazio, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
DEZEMBRO DE 2018

Treistman, Felipe

Geração de Cenários Sintéticos Mensais de Vazões Condicionadas ao El Niño/Oscilação Sul - Aplicação ao Planejamento da Operação de Sistemas Hidrotérmicos/Felipe Treistman. – Rio de Janeiro: UFRJ/COPPE, 2018.

XVII, 241 p.: il.; 29, 7cm.

Orientadores: Otto Corrêa Rotunno Filho

Maria Elvira Piñeiro Maceira

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Civil, 2018.

Referências Bibliográficas: p. 118 – 129.

1. Geração de Cenários Sintéticos de Vazões. 2. El Niño - Oscilação Sul. 3. Planejamento da Operação Energética. I. Rotunno Filho, Otto Corrêa *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Civil. III. Título.

Agradecimentos

Aos meus pais, Ronaldo Treistman e Laís Sampaio Treistman, e toda minha família, por me apoiarem em todas as etapas de minha vida.

À minha amiga e companheira da vida, Marcela Coitinho de Aquino e Castro, por todo amor dedicado ao longo dos anos, que me fazem querer ser melhor a cada dia.

À minha orientadora Dra. Maria Elvira Piñeiro Maceira, pelo constante apoio à minha formação profissional, pelo incentivo e auxílio durante o desenvolvimento deste trabalho.

Ao meu orientador professor Otto Corrêa Rotunno Filho, pelos ensinamentos acadêmicos durante o mestrado e por aceitar participar deste trabalho, contribuindo com suas opiniões e revisões, tornando possível sua conclusão.

À amiga Dra. Débora Dias Jardim Penna, um muito obrigado por todo apoio e pelos ensinamentos diários. Ainda no CEPEL, agradeço especialmente ao Dr. Jorge Machado Damazio, pelas diversas dúvidas tiradas, sugestões dadas e histórias contadas. Agradeço também ao amigo Dr. Wanderson Luiz Silva pelo seu bom humor diário, mas principalmente pelo essencial auxílio em assuntos climáticos.

Agradeço também aos amigos e pesquisadores do CEPEL, Diego Maia, Hugo Santarém, Juan Colonese, Lilian Brandão, Luis Fernando Cerqueira, Priscilla Chan e Victor Almeida, pelas risadas e cafés compartilhados, que tornam os dias de trabalho menos árduos.

Ao CEPEL, pelas oportunidades ofertadas, possibilitando o meu desenvolvimento profissional e pessoal. Por todos modelos desenvolvidos anteriormente nesta instituição, especificamente no departamento de Otimização Energética e Meio Ambiente, que tornaram possível a dissertação proposta.

Ao programa de Engenharia Civil da COPPE/UFRJ por contribuir de forma essencial para continuidade de minha formação acadêmica; a todos os professores e técnicos administrativos, obrigado. Aos funcionários técnico-administrativos do Programa de Engenharia Civil da COPPE/UFRJ, em especial à Secretaria Executiva, nas pessoas de Sra. Marcilia Mascarenhas e Sr. Roberto Mallet, e à Secretaria Acadêmica, nas pessoas de Sra. Elisabeth Cornélio, Sr. Jairo A. Leite e Sra. Marcia Reis, bem como aos funcionários do Laboratório de Informática que foram gentis

e prestativos quando precisei de sua orientação e suporte em todos os trâmites requeridos nesta etapa de minha vida. Dedico reconhecimento ao Laboratório de Recursos Hídricos e Meio Ambiente (LABH2O) pela infraestrutura oferecida.

Registra-se, também, que o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Complementarmente, em nome do Laboratório de Recursos Hídricos e Meio Ambiente (LABH2O) do Programa de Engenharia Civil da COPPE/UFRJ, agradeço, ainda, o suporte da FAPERJ pelo apoio financeiro através dos projetos PEC/COPPE - FAPERJ - Pensa Rio - Edital 34/2014 (2014-2020) -E-26/010.002980/2014, projeto FAPERJ No. E-12/2015, projeto FAPERJ No. E-22/2016, projeto CNPq Edital Universal No. 14/2013 - Processo 485136/2013-9, projeto CNPq Edital No. 12/2016 - Processo 306944/2016, pelo contínuo apoio à pesquisa científica no Brasil.

Agradeço ainda às seguintes instituições pelo fornecimento dos dados necessários ao desenvolvimento desta dissertação: Operador Nacional do Sistema Elétrico (ONS), Agência Nacional de Águas (ANA), *National Oceanic and Atmospheric Administration* (NOAA) e *International Research Institute for Climate and Society* (IRI).

A todos que de forma, direta ou indireta, contribuíram para a realização do trabalho.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

GERAÇÃO DE CENÁRIOS SINTÉTICOS MENSAIS DE VAZÕES
CONDICIONADAS AO EL NIÑO/OSCILAÇÃO SUL - APLICAÇÃO AO
PLANEJAMENTO DA OPERAÇÃO DE SISTEMAS HIDROTÉRMICOS

Felipe Treistman

Dezembro/2018

Orientadores: Otto Corrêa Rotunno Filho

Maria Elvira Piñeiro Maceira

Programa: Engenharia Civil

O planejamento da operação energética do Sistema Interligado Nacional brasileiro (SIN) é um complexo problema de otimização de recursos, devido às suas proporções continentais e seu sistema de geração de energia hidrotérmico. Atualmente, utiliza-se um modelo periódico auto-regressivo (PAR(p)) para a geração de séries sintéticas mensais de vazão e energia natural afluyente como insumo para os modelos de otimização de curto e médio/longo prazo. O fenômeno climático El Niño - Oscilação Sul (ENOS) é capaz de alterar os regimes hidrológicos brasileiros, através da redução ou aumento da precipitação em diversas regiões, e assim, impactando a geração de energia. Esta dissertação tem como objetivo propor um aprimoramento metodológico para a geração de cenários de vazões de forma a incluir informações sobre o ENOS. Propõe-se um novo modelo estocástico, chamado de modelo autorregressivo periódico com chaveamento markoviano (MS-PAR(p)), no qual os parâmetros do modelo passam a ser segmentados e ajustados entre diferentes estados no mesmo mês. Para a geração de cenários dos estados de ENOS, são ajustadas cadeias de Markov mensais. Um procedimento para inclusão de previsões probabilísticas do fenômeno climático é desenvolvido. Os modelos original e proposto são comparados através de diversas simulações realizadas com a configuração do SIN. O modelo MS-PAR(p) apresenta melhor desempenho nas regiões sul e norte, onde há uma maior influência do fenômeno climático, sendo capaz de gerar cenários sintéticos cuja distribuição se assemelha mais à observada. Esse fato ocorre, principalmente, durante períodos onde as vazões afluentes estão abaixo da média, o que pode representar um ganho para o planejamento da operação energética do Brasil.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

SYNTHETIC SCENARIO GENERATION OF MONTHLY STREAMFLOWS
CONDITIONED TO THE EL NIÑO/SOUTHERN OSCILLATION -
APPLICATION TO OPERATION PLANNING OF HYDROTHERMAL
SYSTEMS

Felipe Treistman

December/2018

Advisors: Otto Corrêa Rotunno Filho
Maria Elvira Piñeiro Maceira

Department: Civil Engineering

The energy operation planning of the Brazilian National Interconnected System (NIS) is a complex problem of resources optimization, due to its continental proportions and its hydrothermal energy generation system. Currently, a periodic auto-regressive (PAR(p)) model is used to generate synthetic scenarios of monthly streamflow and energy for the mid-term and long-term optimization models. It is known that the climate phenomenon El Niño - Southern Oscillation (ENSO) is capable of changing the Brazilian hydrological regimes through the reduction or increase of rainfall in several regions. This dissertation aims to propose a methodological improvement on the generation of synthetic scenarios of streamflow, to include information from the ENSO. A new stochastic model called Markov-switching periodic auto-regressive (MS-PAR(p)) is proposed, in which the model parameters are segmented and estimated between different states in the same month. For the scenario generation of the ENSO states, monthly Markov chains are adjusted. A procedure for inclusion of probabilistic forecasts of the climatic phenomenon is developed. The original and proposed models are compared through several simulations performed with the configuration of the NIS. The MS-PAR(p) model presents better performance in the south and north regions, where there is a more significant influence of the ENSO. It is demonstrated that the proposed methodology is able to generate synthetic scenarios which distribution better resembles the observed one. This fact occurs, mainly, during periods when the streamflows are below-average, which may represent an improvement for the planning of the Brazilian energy operation.

Sumário

Lista de Figuras	xi
Lista de Tabelas	xvi
1 Introdução	1
1.1 Considerações Iniciais	1
1.2 Objetivos	2
1.3 Estrutura Textual	2
2 Revisão Bibliográfica	4
2.1 Planejamento da Operação Energética	4
2.2 El Niño - Oscilação Sul	11
2.2.1 Impactos Relacionados ao El Niño - Oscilação Sul nos Padrões de Precipitação	14
2.2.2 Impactos Relacionados ao El Niño - Oscilação Sul nos Padrões de Vazões Fluviais	20
2.2.3 Previsões para o El Niño - Oscilação Sul	23
2.3 Modelos Estocásticos em Hidrologia	26
2.3.1 Modelos de Geração de Cenários Sintéticos de Vazões com uso de Informações Climáticas	27
3 Metodologia	33
3.1 Dados de Vazão Natural Afluyente	33
3.2 Dados do Fenômeno El Niño - Oscilação Sul	33
3.2.1 Histórico do Fenômeno El Niño - Oscilação Sul	33
3.2.2 Previsões do Fenômeno El Niño - Oscilação Sul	36
3.3 Modelo Autorregressivo Periódico	37
3.3.1 Identificação do Modelo	38
3.3.2 Estimação do Modelo	40
3.3.3 Correlações Espaciais	40
3.3.4 Vazões Incrementais Negativas	41
3.3.5 Geração de Séries Sintéticas com o Modelo PAR(p)	42

3.4	Modelo Autorregressivo Periódico com Chaveamento Markoviano - MS-PAR(p)	44
3.4.1	Modelagem do Processo Markoviano	47
3.4.2	Identificação da Influência do Fenômeno El Niño - Oscilação Sul nas Vazões	53
3.5	Métricas de Avaliação	55
3.5.1	CrITÉrios de avaliação para Casos Não-Condicionados	55
3.5.2	CrITÉrios de avaliação para Casos Condicionados	57
4	Resultados e Discussões	59
4.1	Estudos de caso	59
4.1.1	Casos Analisados	59
4.2	Influência do Fenômeno El Niño - Oscilação Sul nas Séries de Vazões Naturais	63
4.3	Caso Não-Condicionado	78
4.4	Caso Condicionado - período pré-ajuste	86
4.5	Caso Condicionado - período pós-ajuste	101
5	Conclusões e Desenvolvimentos Futuros	115
	Referências Bibliográficas	118
A	<i>Synthetic Scenario Generation of Monthly Streamflows Conditioned to the El Niño/Southern Oscillation - Application to Operation Planning of Hydrothermal Systems</i>	130
A.1	Introduction	130
A.1.1	Initial Considerations	130
A.1.2	Objectives	131
A.1.3	Textual Structure	131
A.2	Literature Review	132
A.2.1	Energy Operation Planning	132
A.2.2	El Niño - Southern Oscillation	139
A.2.3	Stochastic Models in Hydrology	154
A.3	Methodology	159
A.3.1	Natural Streamflow Data	159
A.3.2	El Niño - Southern Oscillation Data	159
A.3.3	Periodic Autoregressive Model	162
A.3.4	Markov-Switching Periodic Autoregressive - MS-PAR(p)	169
A.3.5	Evaluation Metrics	181
A.4	Results and Discussions	184

A.4.1	Case Studies	184
A.4.2	Influence of El Niño - Southern Oscillation on the Streamflows	187
A.4.3	Unconditioned Case	201
A.4.4	Conditioned cases - in-sample period	210
A.4.5	Conditioned cases - out-of-sample period	225
A.5	Conclusions and Future Improvements	238

Lista de Figuras

2.1	Participação de fontes de geração de energia na potência instalada do SIN em 31 de dezembro de 2017.	5
2.2	Processo de tomada de decisão operativa e suas possíveis consequências em um sistema hidrotérmico.	7
2.3	Exemplo de acoplamento hidráulico entre usinas hidroelétricas na bacia do rio Grande.	8
2.4	Comportamento dos custos da operação de um sistema hidrotérmico em função do volume final armazenado.	9
2.5	Cadeia de modelos do planejamento da operação e expansão energética.	11
2.6	Regiões de acompanhamento da TSM do Pacífico para identificação de ENOS	12
2.7	Percentis de precipitação média esperada no Brasil durante episódios de El Niño	17
2.8	Percentis de precipitação média esperada no Brasil durante episódios de La Niña	18
2.9	Comportamento das células de circulação de Walker e Hadley durante eventos de El Niño (a) e La Niña (b) para os impactos canônicos (I) e reversos (II)	20
2.10	Anomalias de precipitação e vazão natural afluente anuais para a UHE Itá entre 1961 e 2004	22
2.11	Pluma de previsão da anomalia da TSM NINO 3.4 obtida em novembro de 2017	24
2.12	Categorias observadas e previstas de ENOS.	26
3.1	Fluxograma metodológico - etapas do trabalho.	34
3.2	Histórico do índice ONI. Áreas em vermelho (azul) representam episódios de EN (LN).	35
3.3	Geração de cenários sintéticos condicionados ao passado recente.	43
3.4	Geração de cenários sintéticos não-condicionados ao passado recente.	43
3.5	Procedimento de identificação de influência do ENOS nas vazões naturais afluentes às UHEs - parte I	54

3.6	Procedimento de identificação de influência do ENOS nas vazões naturais afluentes às UHEs - parte II	55
3.7	Exemplo de uma sequência negativa hipotética. A linha cheia corresponde a série temporal observada e a linha pontilhada ao limite pré-estabelecido.	56
4.1	Distribuição geográfica das UHEs selecionadas para análise.	61
4.2	Teste KS para a região sul do Brasil.	63
4.3	Estatísticas históricas da UHE Itá segundo o estado ENOS.	65
4.4	Estatísticas históricas da UHE Itaipu segundo o estado ENOS.	65
4.5	Estatísticas históricas da UHE Salto Caxias segundo o estado ENOS.	66
4.6	Teste KS para a região sudeste e centro-oeste do Brasil.	68
4.7	Estatísticas históricas da UHE Capivara segundo o estado ENOS.	70
4.8	Estatísticas históricas da UHE Furnas segundo o estado ENOS.	70
4.9	Estatísticas históricas da UHE Porto Primavera segundo o estado ENOS.	71
4.10	Teste KS para a região nordeste do Brasil.	71
4.11	Estatísticas históricas da UHE Sobradinho segundo o estado ENOS.	72
4.12	Teste KS para a região norte do Brasil.	73
4.13	Estatísticas históricas da UHE Belo Monte segundo o estado ENOS.	74
4.14	Estatísticas históricas da UHE Santo Antônio Jari segundo o estado ENOS.	74
4.15	Distribuições de frequências histórica e sintética, gerada através da cadeia de Markov.	78
4.16	Matrizes de transição mensais históricas (a) e o módulo da diferença entre as matrizes de transições mensais sintéticas com relação às históricas (b), em valores percentuais.	79
4.17	Estatísticas obtidas para a UHE Itá para o caso não condicionado.	81
4.18	Estatísticas obtidas para a UHE Furnas para o caso não condicionado.	81
4.19	Estatísticas obtidas para a UHE Sobradinho para o caso não condicionado.	82
4.20	Estatísticas obtidas para a UHE Belo Monte para o caso não condicionado.	82
4.21	Assimetrias histórica e sintética da UHE Furnas.	84
4.22	Correlações espaciais anuais histórica (a) e geradas com os modelos PAR(p) (b), MS-PAR(p)1 (c) e MS-PAR(p)2 (d)	85
4.23	Cenários sintéticos de vazões gerados para o período entre 1953 a 1956 na UHE Santo Antônio Jari com os modelos PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) e sua distribuição histórica (d).	98

4.24	Cenários sintéticos de vazões gerados para o período entre 1953 a 1956 na UHE Salto Caxias com os modelos PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) e sua distribuição histórica (d).	100
4.25	Percentual de cenários de ENOS gerados até doze meses à frente para os anos entre 2010 e 2017.	102
4.26	RPSS em função do período observado (eixo horizontal) e início da previsão (eixo vertical) para os anos entre 2010 e 2017.	103
4.27	Mediana do RPSS mensal com relação ao horizonte de previsão para os anos entre 2010 e 2017.	104
4.28	Cenários sintéticos gerados para o período entre 2010 a 2013 na UHE Belo Monte com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).	107
4.29	Cenários sintéticos gerados para o período entre 2014 a 2017 na UHE Belo Monte com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).	108
4.30	Cenários sintéticos gerados para o período entre 2010 a 2013 na UHE Itá com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).	109
4.31	Cenários sintéticos gerados para o período entre 2014 a 2017 na UHE Itá com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).	110
4.32	Cenários sintéticos gerados de forma sequencial para o período entre 2015 a 2017 na UHE Estreito Tocantins com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).	113
A.1	Participation of sources of power generation in the installed capacity of the NIS on December 31, 2017.	133
A.2	Operational decision-making process and its possible consequences in a hydrothermal system.	135
A.3	Example of hydraulic coupling between HPPs in the Rio Grande basin.	136
A.4	Behavior of operating costs of a hydrothermal system as a function of the final volume stored.	137
A.5	Chain of models for the expansion and operation energy planning.	138
A.6	Pacific SST monitoring regions for ENSO identification	141
A.7	Mean precipitation percentile expected in Brazil during El Niño episodes	145
A.8	Percentis of average precipitation expected in Brazil during episodes of La Niña	146
A.9	Walker and Hadley circulation cell behavior during El Niño (a) and La Niña (b) events for canonical (I) and reverse (II) impacts	148

A.10 Annual rainfall and natural streamflow for Itá HPP between 1961 and 2004	149
A.11 Prediction plume of SST anomaly in NINO 3.4 obtained in November 2017.	152
A.12 Observed and predicted ENSO categories	154
A.13 Methodological flowchart.	160
A.14 ONI History.	161
A.15 Sythentic scenarios generation conditioned to the recent past.	168
A.16 Sythentic scenarios unconditioned generation.	169
A.17 Procedure to identify the influence of ENSO on the natural flows affluent to the HPPs - part I	180
A.18 Procedure to identify the influence of ENSO on natural inflows to HPPs - Part II	180
A.19 Example of a hypothetical negative sequence. The full line corresponds to the observed time series and the line dotted to the pre-established limit.	182
A.20 Geographic distributribution of the selected HPPs.	186
A.21 KS Test for the southern region of Brazil.	189
A.22 Historical Statistics of Itá HPP according to ENSO state.	190
A.23 Historical statistics of the Itaipu HPP according to ENSO state.	191
A.24 Historical statistics of the Salto Caxias HPP according to ENSO state.	191
A.25 KS Test for the southeastern and central-west region of Brazil.	193
A.26 Historical statistics of the Capivara HPP according to ENSO state.	195
A.27 Historical statistics of the Furnas HPP according to the state of ENSO.	195
A.28 Historical statistics of the Porto Primavera HPP according to the state of ENSO.	196
A.29 KS Test for the northeast region of Brazil.	196
A.30 Historical statistics of Sobradinho HPP according to ENSO state.	197
A.31 KS Test for the north region of Brazil.	198
A.32 Historical statistics of Belo Monte Hydroelectric Plant according to ENSO state.	199
A.33 Historical statistics of Santo Antônio Jari HPP according to ENSO state.	199
A.34 Distribution of historical and synthetic probability generated through the Markov chain.	201
A.35 Historical monthly transition matrices (a) and the absolute value of the difference between the synthetic monthly transitions matrices compared to the historical ones (b), in percentage values.	203
A.36 Statistics obtained for Itá HPP for the unconditioned case.	205

A.37	Statistics obtained for Furnas HPP for the unconditioned case.	205
A.38	Statistics obtained for Sobradinho HPP for the unconditioned case.	206
A.39	Statistics obtained for Belo Monte HPP for the unconditioned case.	206
A.40	Historical and synthetic asymmetry of Furnas HPP.	208
A.41	Hystorical annual cross-correlations (a) ans synthetic ones obtained with the models PAR(p) (b), MS-PAR(p)1 (c) and MS-PAR(p)2 (d)	209
A.42	Synthetic scenarios generated for the period between 1953 and 1956 at the Santo Antônio Jari HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).	222
A.43	Synthetic scenarios generated for the period 1953 to 1956 in the Salto Caxias HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).	224
A.44	Percentage of ENSO scenarios generated up to twelve months ahead for the period between 2010 and 2017.	226
A.45	RPSS depending on observed period (horizontal axis) and lead time (vertical axis), for the period between 2010 and 2017.	227
A.46	Median of the monthly RPSS with respect to the forecast horizon.	228
A.47	Synthetic scenarios generated for the period between 2010 and 2013 in the Belo Monte HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS- PAR(p)2 (c) models, and the historical distribution (d).	231
A.48	Synthetic scenarios generated for the period between 2014 and 2017 in the Belo Monte HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS- PAR(p)2 (c) models, and the historical distribution (d).	232
A.49	Synthetic scenarios generated for the period between 2010 and 2013 in the Itá HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).	233
A.50	Synthetic scenarios generated for the period between 2014 and 2017 in the Itá HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).	234
A.51	Synthetic scenarios generated sequentially for the period 2015 to 2017 for the Estreito Tocantins HPP with models PAR(p) (a), MS- PAR(p)1 (b) and MS-PAR(p)2 (c) and its historical distribution (d).	237

Lista de Tabelas

2.1	Previsão probabilística de ENOS obtida em novembro de 2017	24
3.1	Quantidade de meses classificados entre os estados de ENOS no período de 1950 a 2017	36
4.1	Relação de UHEs representantes para a apresentação dos resultados	60
4.2	Resumo dos estudos de caso realizados para a avaliação dos modelos	62
4.3	Efeitos do ENOS nas vazões afluentes às UHEs brasileiras	77
4.4	Estatísticas histórica e sintética referentes ao comprimento de eventos de ENOS	80
4.5	Resumo do teste KS para os cenários simulados	83
4.6	Resumo dos testes de sequências negativas para os cenários simulados de vazões	86
4.7	Estatísticas históricas e sintéticas do comprimento de sequências negativas de vazões	87
4.8	Estatísticas históricas e sintéticas da intensidade de sequências negativas de vazões	88
4.9	Estatísticas históricas e sintéticas da soma de sequências negativas de vazões	89
4.10	Avaliação da geração de cenários de vazões para o período de ajuste dos modelos entre os anos de 1951 e 2009	91
4.11	EMPA mensal do modelo PAR(p) para o período de ajuste dos modelos entre os anos de 1951 e 2009.	93
4.12	EMPA mensal do modelo MS-PAR(p)1 para o período de ajuste dos modelos entre os anos de 1951 e 2009.	93
4.13	EMPA mensal do modelo MS-PAR(p)2 para o período de ajuste dos modelos entre os anos de 1951 e 2009.	94
4.14	EMPA dividido entre os estados de ENOS para a geração de cenários no período de ajuste dos modelos	95
4.15	Avaliação da geração de cenários no período pós-ajuste para os anos entre 2010 e 2017	105

4.16	Avaliação da simulação no período pós-ajuste acumulados em 6 meses para os anos entre 2010 e 2017	112
A.1	Probability forecast of ENSO obtained in November 2017	151
A.2	Number of months classified between ENSO states between 1950 and 2017	162
A.3	Representing HPPs	185
A.4	Case Studies Summary	188
A.5	Effects of ENSO on streamflows to Brazilian HPPs	202
A.6	Historical and Synthetic statistics of the length of ENSO states	204
A.7	KS test summary	207
A.8	Negative sequence tests summary	210
A.9	Historic and synthetic statistics of negative sequence length	211
A.10	Historic and synthetic statistics of negative sequence intensity	212
A.11	Historic and synthetic statistics of negative sequence sum	213
A.12	In-sample Evaluation	215
A.13	Monthly MAPE of the PAR(p) model	216
A.14	Monthly MAPE of the MS-PAR(p)1 model	217
A.15	Monthly MAPE of the MS-PAR(p)2 model.	217
A.16	MAPE of the in-sample period divided among the states of ENSO	219
A.17	Evaluation of the out-of-sample generated scenarios for the period between 2010 and 2017	229
A.18	Evaluation of six months accumulated out-of-sample errors	236

Capítulo 1

Introdução

1.1 Considerações Iniciais

A geração de energia elétrica no Brasil é realizada majoritariamente através de suas usinas hidroelétricas (UHEs), dispersas em doze principais bacias hidrográficas, perfazendo um total de 114GW. No ano de 2017, cerca de 70% da geração verificada foi disponibilizada via energia hidroelétrica, 7% através das usinas eólicas, restando um pouco mais de 23% para a geração térmica, caracterizando-se como uma matriz energética verde (ONS, 2018a). Ao contrário de sistemas puramente térmicos, matrizes energéticas com a composição brasileira possuem uma complexidade maior, devido ao desconhecimento do suprimento futuro de sua maior fonte de energia, hidroelétrica. Sendo assim, pode-se afirmar que as vazões afluentes futuras às UHEs são a principal fonte de incerteza para o planejamento da operação energética do setor elétrico brasileiro e devem ser foco de estudos de forma a aprimorar constantemente a sua representação.

Atualmente, para o segundo mês do planejamento da operação de curto prazo e para todo o horizonte do planejamento da operação de médio/longo prazo, são gerados cenários sintéticos de aflúências através de um modelo estocástico autorregressivo periódico de ordem p . Esse tipo de modelo estocástico descreve as realizações futuras como uma regressão linear dos valores passados, acrescido de um termo aleatório. Ou seja, utilizam-se as próprias vazões passadas como insumo ao modelo. Um grande esforço científico vem sendo realizado com o intuito de incluir informações climáticas provenientes de outras variáveis relevantes ao ciclo hidrológico em modelos estocásticos de vazões fluviais. Dentre os vários fenômenos climáticos influentes nos ciclos hidrológicos brasileiros, o fenômeno El Niño - Oscilação Sul (ENOS) vem sendo apontado, ao longo de décadas de pesquisa, como um dos mais importantes. Sua ocorrência impacta de diferentes formas e intensidades a precipitação e, conseqüentemente, as vazões fluviais, podendo-se afirmar que a geração

de energia no Brasil é amplamente influenciada pelo ENOS.

1.2 Objetivos

Considerando o exposto acima, esta dissertação tem como objetivo principal propor um aprimoramento ao modelo utilizado na geração de séries sintéticas de afluências às UHEs nos estudos oficiais do planejamento da operação energética do setor elétrico brasileiro. É apresentado um novo modelo estocástico, capaz de incluir informações sobre o fenômeno climático El Niño - Oscilação Sul (ENOS), visando determinar, com uma precisão maior, as distribuições multivariadas de probabilidades condicionadas para as afluências futuras.

Dado o objetivo principal deste trabalho, uma série de desafios impostos à sua obtenção surgem. Assim, caracterizam-se os seguintes objetivos secundários necessários:

- identificar a melhor forma de se incluir a informação sobre o ENOS no modelo;
- avaliar a reprodução de estatísticas históricas relevantes por parte do modelo tradicional e com o aprimoramento sugerido e
- avaliar a geração de cenários de vazões afluentes em estudos de casos que simulem um caso real.

1.3 Estrutura Textual

Este primeiro capítulo tem como intuito contextualizar o problema dentro do qual se insere este trabalho bem como descrever seus objetivos. O Capítulo 2 tratará da revisão bibliográfica realizada, subdividindo-se em três principais temáticas: a Seção 2.1 irá se aprofundar no planejamento da operação energética com foco no caso brasileiro; a Seção 2.2 discorrerá sobre o fenômeno climático ENOS, iniciando com sua caracterização, abordando seus impactos tanto em escala global quanto no Brasil e finalizando com seus modelos de previsão; por fim, a Seção 2.3 apresentará modelos estocásticos utilizados em hidrologia, com foco em modelos para a geração de séries sintéticas, desde os modelos mais tradicionais passando pelos que incluem informações climáticas na sua modelagem.

No Capítulo 3, primeiramente, são apresentados os dados obtidos de vazão natural afluente e os dados relativos ao ENOS. Em seguida, é apresentada toda a metodologia referente ao modelo de geração de séries sintéticas mensais de vazão e energia, GEVAZP, e o aprimoramento proposto nesta dissertação. No final do capítulo, expõem-se os estudos de casos utilizados na avaliação dos modelos e as

métricas utilizadas. A apresentação dos resultados é realizada no Capítulo 4, onde o potencial do modelo MS-PAR(p) frente ao modelo PAR(p) proposto é verificado para a geração de cenários sintéticos de vazões afluentes. No Capítulo 5 são feitas as ponderações e conclusões referentes a este trabalho sobretudo com relação aos resultados obtidos ao longo dele, além de indicar possíveis aprimoramentos futuros. Apresenta-se por fim as referências bibliográfica que ajudaram a inspirar e compor esta dissertação. Cabe destacar que, nesta dissertação, optou-se por incluir a versão em inglês do texto no Apêndice A.

Capítulo 2

Revisão Bibliográfica

Este capítulo apresentará a revisão bibliográfica realizada para os três temas centrais desta dissertação. A primeira seção tratará do planejamento da operação energética, apresentando as características do Sistema Interligado Nacional (SIN) e as metodologias utilizadas na solução do problema do planejamento de sua operação. A segunda seção discorrerá sobre o fenômeno climático ENOS, com sua caracterização, impactos associados a ele e o estado da arte de suas previsões. Por fim, será feita uma revisão sobre o processo de geração de cenários sintéticos de afluições fluviais aos reservatórios de usinas hidroelétricas (UHEs), apresentando a abordagem tradicional de sua modelagem estocástica, e, em seguida, abordando os modelos estocásticos que consideram outras variáveis climáticas.

2.1 Planejamento da Operação Energética

O SIN é o sistema brasileiro de geração e transmissão de energia elétrica, caracterizado por ser um sistema hidrotérmico de grande porte. Devido às suas características singulares como dimensões continentais, diferentes regimes hidrológicos e eólicos, múltiplos proprietários e usuários, pode-se afirmar que o SIN é um dos sistemas mais complexos do mundo no que diz respeito ao planejamento de sua expansão e operação. Segundo dados obtidos do Plano da Operação Energética 2018/2022 (ONS, 2018a), pode ser observado na Figura 2.1 que as UHEs compõe a maior parcela da matriz de geração de energia elétrica brasileira, com 67,8% do total de potência instalada, ou 114 GW. As usinas eólicas representam 7,9% da matriz de energia elétrica brasileira até o final do ano de 2016, com 9,6 GW de potência instalada, a quarta maior fonte de geração de energia, atrás das usinas térmicas a gás.

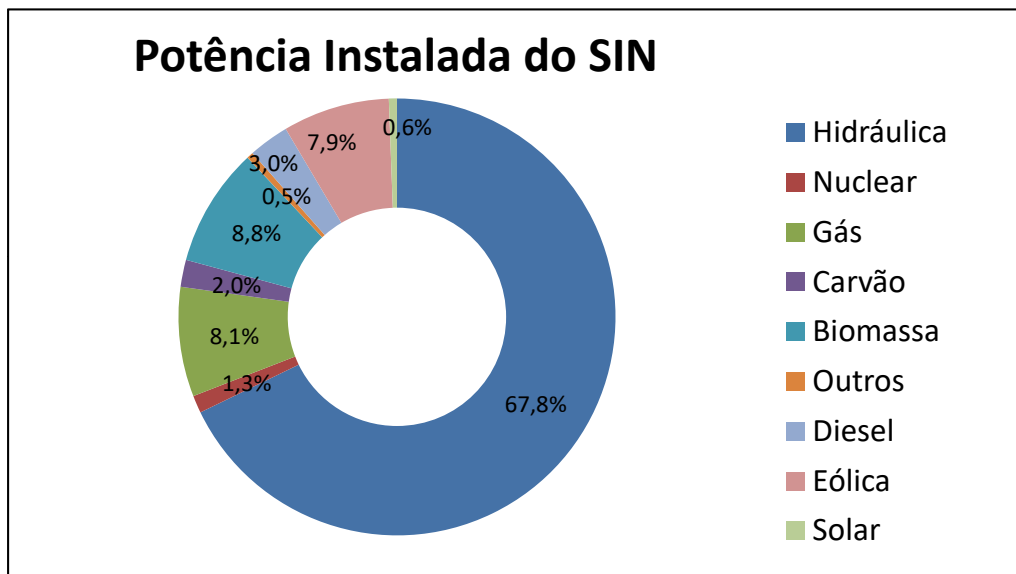


Figura 2.1: Participação de fontes de geração de energia na potência instalada do SIN em 31 de dezembro de 2017. Fonte: ONS (2018a).

Para o atendimento da evolução da carga, a expansão da oferta de geração até o ano de 2022 apresenta um incremento de 18.728 MW de potência instalada, o que representa aproximadamente 12% da oferta de geração atual. Desse montante, as hidroelétricas continuam sendo a principal fonte de oferta na geração, apresentando o maior crescimento bruto da matriz energética, com um incremento de 8.989 MW, 8,5% de sua oferta atual, porém diminuindo sua participação para 65,6% do total em 2021. É interessante observar que, em sua maioria, as usinas projetadas são desprovidas de reservatório de regulação entre elas, não apresentando ganho de energia armazenável ao SIN. Dessa forma, o SIN continua perdendo gradativamente sua “inércia hidroenergética”, podendo ser cada vez mais frequente o deplecionamento acentuado dos grandes reservatórios ao final de cada estação seca. Além disso, o sistema se encontrará cada vez mais exposto às secas prolongadas.

Apesar das usinas eólicas representarem um baixo percentual de potência instalada total, sua importância é cada vez maior para o SIN, principalmente na região nordeste, onde há a maior concentração desse tipo de usina. Outro fator que tem contribuído seguidamente para o alto aproveitamento dessa fonte está no fato da região nordeste ter passado por condições hidrológicas críticas nos últimos anos, com as vazões afluentes muito abaixo da média de longo termo (MLT), atingindo fortemente a produção hidrelétrica. Segundo dados apurados pelo Operador Nacional do Sistema Elétrico (ONS), a geração eólica foi responsável pelo atendimento da carga em cerca de 40% em base mensal, chegando até o atendimento de 71% da carga em base horária do nordeste. Atualmente, a demanda energética da região nordeste pode ser completamente atendida sem a necessidade de geração hidroelétrica, apenas

com o acionamento das usinas termoelétricas (UTES), eólica e intercâmbios provenientes de outras regiões. As usinas eólicas com conclusão das obras já projetadas para 2022 acrescentarão 3.064 MW no SIN, um considerável incremento de 24,9%, alcançando 15.373 MW instalados, representando 8,8% do total; o que colocará essa fonte como a segunda maior em potência instalada. Já as usinas solares passarão de apenas 952 MW para 3.638 MW. Para efeito de comparação, as usinas nucleares de Angra I e Angra II somam 1.990 MW instalados. Apesar dos benefícios, a inclusão de fontes intermitentes, como a eólica e a solar, traz uma série de dificuldades para a operação energética. É, neste ponto, que a geração hidroelétrica assume um novo papel, de firmar a geração eólica e solar, já que hidroelétricas possuem partida rápida, dando uma estabilidade para a geração proveniente de fontes intermitentes.

Como pode ser observado, as características do SIN encontram-se em um processo de grandes alterações. A matriz de geração de energia elétrica que antes era predominantemente hidráulica, com grandes reservatórios de regulação pluri-anuais, e apenas complementado por usinas térmicas, está se encaminhando para uma matriz com alta penetração de fontes renováveis intermitentes e diminuição da capacidade de armazenamento de energia por parte dos reservatórios. Apesar de todos esses fatores, a energia hidroelétrica continuará a ser a principal fonte de geração. Com essa perspectiva, torna-se importante o aprimoramento das metodologias empregadas na sua representação nos diferentes horizontes do planejamento energético.

Ao contrário de sistemas puramente térmicos, em que, de forma simplificada, o planejamento da operação pode ser solucionado através da hierarquização das UTES por custo de produção, sistemas hidrotérmicos, incluindo sua extensão com a inclusão de usinas eólicas, possuem um planejamento mais complexo (FORTUNATO, 1990). Para sistemas hidrotérmicos, com predominância da energia hidráulica, a principal fonte de incertezas é representada pelas afluições hidrológicas futuras, que pressupõe acoplamentos temporais de decisões operativas. Existem vínculos entre a tomada de decisão em instantes de tempo consecutivos induzidos pela capacidade de armazenamento de energia nos reservatórios das usinas hidroelétricas. A Figura 2.2 traz um resumo das possíveis decisões operativas em um reservatório. Caso seja tomada a decisão de armazenar água em um reservatório e a condição hidrológica futura seja favorável, essa política apresenta-se como inadequada, representando um aumento desnecessário do custo presente da operação, sendo a decisão operativa correta a geração hidroelétrica no presente. Se as afluições futuras forem baixas, e toma-se a decisão no instante presente de turbinar a água de um reservatório, essa também é uma operação inadequada, causando o acionamento de térmicas mais caras nos passos à frente, ou, em casos mais extremos, corte de carga (déficit) – representando um incremento do custo futuro. Nesse

caso, a operação correta seria o armazenamento de água no instante presente. Ou seja, ações no instante presente têm impactos tanto na operação do reservatório em questão, quanto nos custos futuros de geração do sistema.

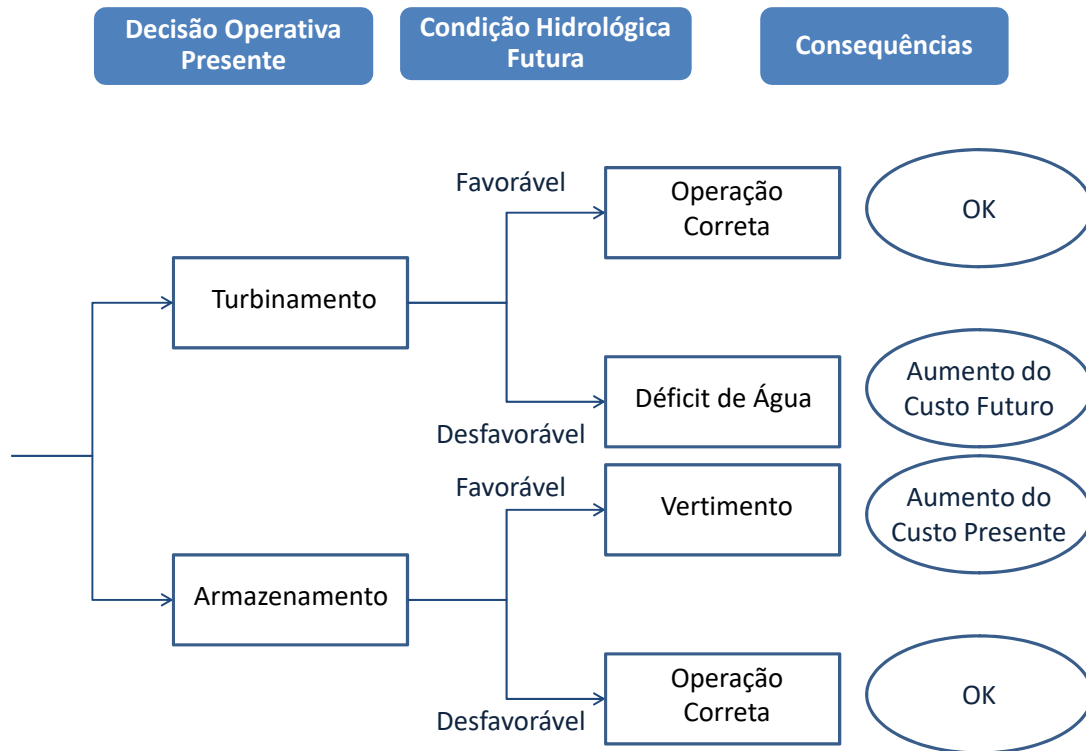


Figura 2.2: Processo de tomada de decisão operativa e suas possíveis consequências em um sistema hidrotérmico.

Além disso, o planejamento da operação hidrotérmica também se defronta com acoplamentos hidráulicos entre UHEs, exemplificado pela Figura 2.3, que traz uma representação esquemática das UHEs existentes na bacia do rio Grande. Nessa imagem, UHEs com reservatório são representadas através de triângulos e UHEs a fio d'água por círculos. Ações operativas de UHEs localizadas a montante (por exemplo, UHE Camargos) acabam afetando a operação de todas as outras UHEs localizadas a jusante, representando um vínculo espacial entre as decisões.

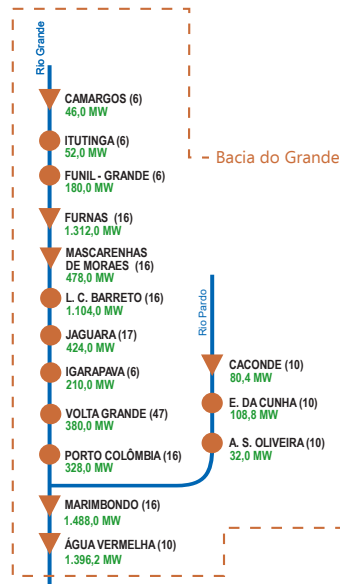


Figura 2.3: Exemplo de acoplamento hidráulico entre usinas hidroelétricas na bacia do rio Grande. Fonte: adaptado de ONS (2018b)

De maneira simplificada, uma boa coordenação de todos os recursos disponíveis no SIN deve levar em conta critérios de segurança, garantindo o suprimento constante de energia a todas as regiões do Brasil, tanto no presente quanto nos estágios futuros, além de garantir a modicidade tarifária. A garantia do menor custo total da operação deve levar em conta tanto o custo presente (ou custo imediato) quanto o custo futuro acarretado pela operação realizada.

A Figura 2.4 apresenta o comportamento dos custos da operação de um sistema hidrotérmico em função do volume final armazenado. A função de custo imediato (FCI) representa o custo presente associado pelo acionamento de UTEs para a produção de energia. Como pode ser observado, é uma função que cresce com o volume final armazenado. Quanto maior a quantidade de recurso hídrico estocado mais UTEs são acionadas para a produção de energia no instante de tempo t , tornando a operação nesse instante mais cara. A função de custo futuro (FCF) possui um comportamento oposto a FCI. Ao armazenar recursos, o custo futuro da operação – do instante $t+1$ até o final do período de estudo - tende a ser menor, devido a uma disponibilidade maior de geração por UHEs. A FCF além de considerar o custo de produção pelas UTEs leva em conta o custo futuro relacionado aos impactos causados por um déficit de energia no sistema. Existe, então, um custo de oportunidade relacionado à água armazenada, medido pela derivada da FCF, chamado de “valor da água”. Dessa forma, o custo total da operação (CT) é o somatório dos custos presente e futuro. A operação ótima de um sistema hidrotérmico ocorre com a minimização do CT. A linha tracejada em vermelho, representa o ponto que possui o menor CT, ou seja, o ponto de operação ótima.

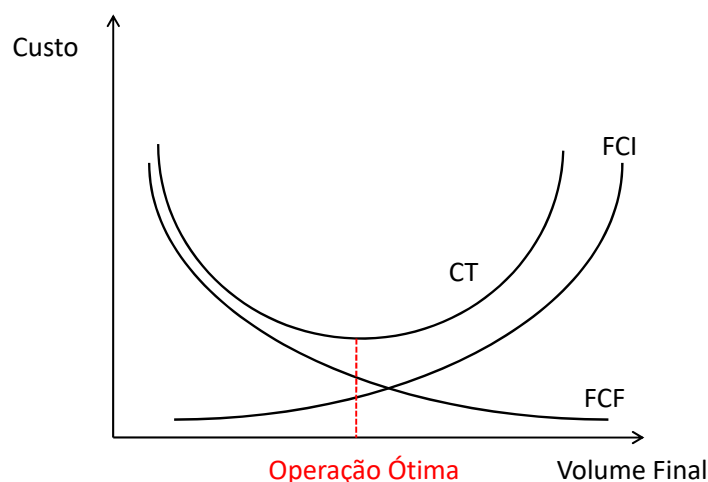


Figura 2.4: Comportamento dos custos da operação de um sistema hidrotérmico em função do volume final armazenado.

Pode-se então afirmar que a operação do SIN é um complexo problema de otimização de recursos com múltiplos objetivos e possíveis soluções. Devido a sua alta complexidade, o planejamento da operação do SIN é dividido em diferentes etapas (programação da operação, curto prazo e médio/longo prazo) que consideram distintos horizontes de tempo e graus de detalhamento da representação do SIN e das incertezas relacionadas. Atualmente, utiliza-se a cadeia de modelos desenvolvida pelo Centro de Pesquisas de Energia Elétrica – CEPTEL (MACEIRA *et al.*, 2002) para a solução do problema supracitado.

A Figura 2.5 apresenta uma perspectiva da cadeia de modelos citada no parágrafo anterior. O planejamento da operação energética do SIN é dividido em três etapas distintas. Quanto maior o horizonte de planejamento, menor é o grau de detalhamento do SIN considerado no problema e maiores as incertezas sobre o futuro.

Na etapa de planejamento de médio/longo prazo o modelo NEWAVE (MACEIRA *et al.*, 2018b) agrupa as UHEs em reservatórios equivalentes de energia (REEs), a fim de representar as incertezas detalhadamente além de manter o tempo computacional de execução do modelo em níveis adequados. Atualmente a topologia considerada no NEWAVE representa o parque hidrelétrico em doze REEs. O modelo NEWAVE utiliza a Programação Dinâmica Dual Estocástica (PEREIRA e PINTO, 1991) para definir qual será a participação do parque térmico e de cada REE no atendimento da carga, além dos intercâmbios do sistema, até trinta anos à frente, com discretização mensal. Em outras palavras, o modelo NEWAVE é responsável pelo cálculo da política de operação do sistema interligado, representado pelas FCFs de cada mês. Essas funções fornecem o valor da energia armazenada para cada estado do sistema (armazenamento no início de cada mês) e condição

hidrológica.

A incerteza hidrológica durante o horizonte de planejamento é considerada através de cenários multivariados sintéticos de energia natural afluyente ¹(ENA) que são gerados estocasticamente pelo modelo GEVAZP (JARDIM *et al.*, 2001; MACEIRA *et al.*, 2006; MACEIRA e DAMÁZIO, 2006; MACEIRA e BEZERRA, 1997). O modelo GEVAZP, objeto de aprimoramento dessa dissertação, será apresentado de forma detalhada no Capítulo 3. A partir dos resultados da operação, o modelo NEWAVE constrói a FCF, que servirá como dado de entrada para o modelo DECOMP (DINIZ *et al.*, 2018). Ainda na etapa de planejamento da operação de médio prazo, o modelo SUSHI (MACEIRA *et al.*, 2018a) realiza a simulação da operação hidrotérmica para usinas individualizadas, seguindo a política de operação e o conjunto de FCFs, determinadas no modelo NEWAVE.

Na etapa de curto prazo, o modelo DECOMP resolve o problema do planejamento da operação do SIN, considerando-o, de forma mais detalhada, com a representação individualizada do parque hidroelétrico. Nessa etapa, é definido o despacho de cada uma das usinas térmicas e hidráulicas, de forma a minimizar, ao final do horizonte, o custo total da operação, levando em conta a FCF construída pelo NEWAVE. Para o primeiro mês, dividido em etapas semanais, o modelo DECOMP utiliza previsões de vazões de forma determinística, fornecidas em grande parte pelo modelo estocástico PREVIVAZ (MACEIRA *et al.*, 1999). Para os demais meses, o modelo GEVAZP pode fornecer uma árvore de cenários multivariados sintéticos de vazões fluviais. Atualmente utiliza-se o modelo DECOMP para um horizonte de dois meses.

Na última etapa da cadeia de modelos, a programação diária da operação pode ser resolvida através do modelo DESSEM (DOS SANTOS e DINIZ, 2009), que representa o SIN com restrições elétricas e hidráulicas mais detalhadas, designando o despacho de cada unidade da usina a cada meia hora para o primeiro dia e de forma horária para os demais dias, até duas semanas a frente. Para essa etapa, as previsões de vazões diárias podem ser realizadas pelo modelo estocástico PREVIVAZH (DE CARVALHO, 2001), que realiza uma desagregação não paramétrica das previsões semanais fornecidas pelo modelo PREVIVAZ.

¹ENA é a transformação da vazão natural afluyente de uma UHE em energia dada a sua produtividade acumulada.

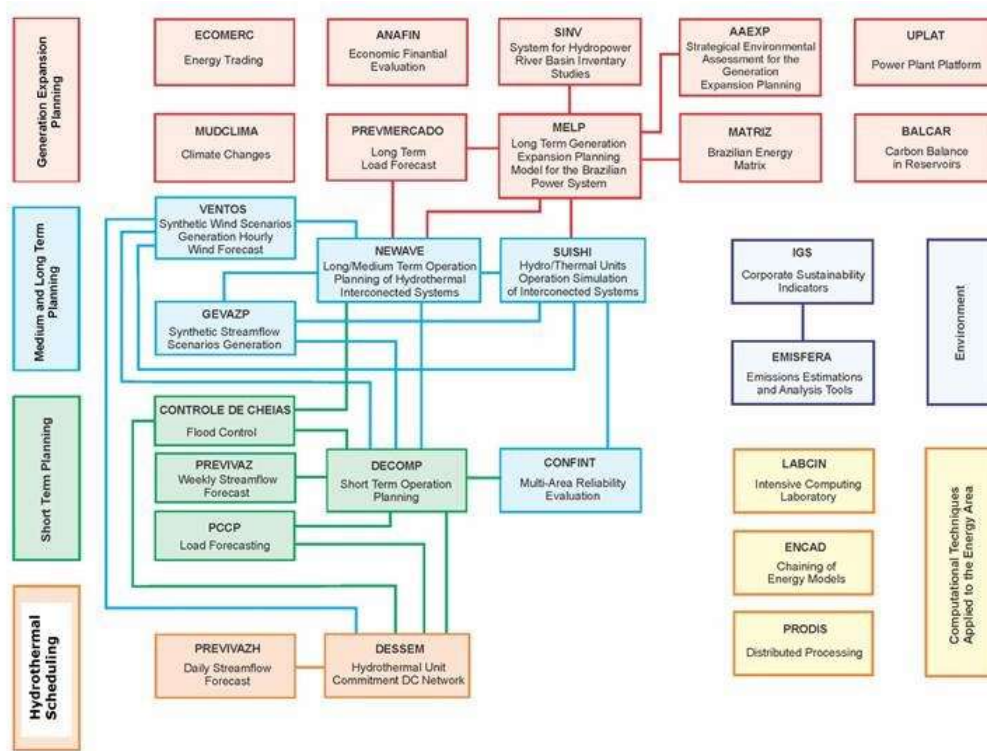


Figura 2.5: Cadeia de modelos do planejamento da operação e expansão energética.

2.2 El Niño - Oscilação Sul

O fenômeno El Niño – Oscilação Sul (ENOS) é um dos fenômenos climáticos naturais mais importantes, tendo diversos impactos nos padrões de circulação global da atmosfera e de variações climáticas, incluindo especialmente precipitação, por todo o mundo (RASMUSSEN e WALLACE, 1983). O fenômeno é classificado em três fases distintas: La Niña (LN), Neutro (N), e El Niño (EN). A primeira é caracterizada por temperatura da superfície do mar (TSM) abaixo da média na região equatorial do Oceano Pacífico. Essa área engloba as porções central e leste do Pacífico. A fase conhecida como El Niño é caracterizada por ser o oposto da La Niña, com TSM acima da média nessa área do Pacífico. Essas características sobre a TSM em ambas as fases precisam ser acopladas com mudanças atmosféricas de pressão e ventos para que um episódio seja deflagrado. A fase neutra caracteriza-se por ser uma etapa de transição entre os eventos, com TSMs próximas da média no Pacífico. Em alguns casos, a TSM apresenta condições de EN ou LN, porém, a atmosfera não apresenta as outras mudanças necessárias para a classificação das outras fases, conforme se verá a seguir.

Ainda que a causa exata para o desencadeamento das fases quente e fria do fenômeno ENOS ainda não seja completamente conhecida, devido a sua alta complexidade relacionada à interação oceano-atmosfera (TIMMERMANN *et al.*,

2018), ambas as fases são conectadas com a Oscilação Sul e a circulação de Walker (WANG, 2002). A célula de circulação de Walker é causada por uma força do gradiente de pressão que ocorre entre um sistema de baixa pressão sobre a Indonésia e um de alta pressão que se encontra sobre o leste do Pacífico. Durante um episódio de EN, a célula de circulação de Walker é enfraquecida, ou até mesmo revertida, causando um aquecimento nas TSM do Pacífico, especialmente da região equatorial. Já em eventos de LN, a Circulação de Walker é fortalecida, resultando em TSMs mais frias na região equatorial do Pacífico (WANG, 2004).

Usualmente, episódios do fenômeno ENOS não possuem uma periodicidade bem definida, podendo se repetir a cada dois ou sete anos, e tendem a persistir por nove ou doze meses, porém, em alguns casos de eventos extremos, eles perduram por mais de dois anos. Pesquisas recentes vêm fazendo uma distinção entre o ENOS Leste, também conhecido como canônico, e o ENOS Central, chamado Modoki, que ocorre com anomalias de TSM na região mais centralizada do Pacífico (SAHU *et al.*, 2014; TEDESCHI, 2013; YUAN e YANG, 2012). Esta dissertação não discutirá sobre essas subclassificações do ENOS.

Para auxiliar na sua classificação e previsão, alguns índices foram desenvolvidos, como o Índice de Oscilação Sul (IOS), que é baseado na diferença de pressão entre duas regiões específicas: Taiti e Darwin (CHEN, 1982). Todavia, o monitoramento do estado ENOS é baseado, principalmente, em medições de TSMs de regiões geográficas específicas do Pacífico (RASMUSSEN e CARPENTER, 1982). As mais usadas são a região Nino 1+2, Nino 3, Nino 4 e a região conhecida como Nino 3.4, uma região que compreende partes das regiões NINO 3 e NINO 4 entre as longitudes 170°O e 120°O, como apresentado na Figura 2.6.

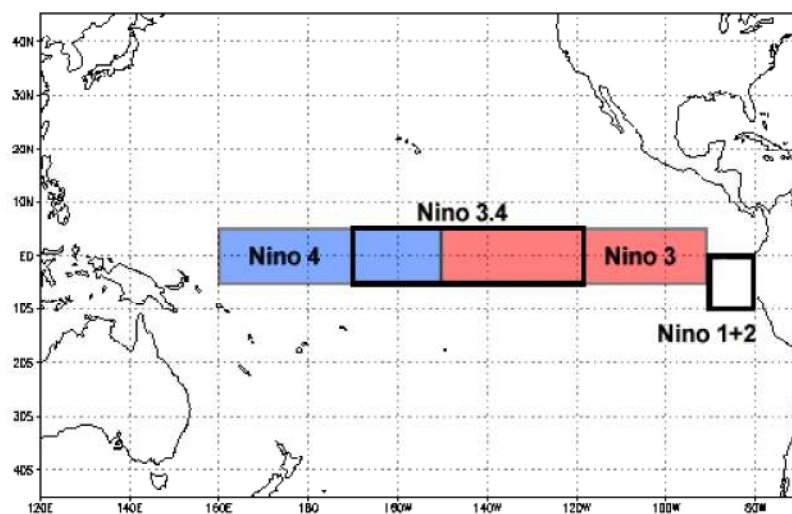


Figura 2.6: Regiões de acompanhamento da TSM do Pacífico para identificação de ENOS. Fonte: NOAA (2018).

A região NINO 3.4 é considerada, de fato, a mais importante para a medição, classificação e previsão de episódios de ENOS (KOUSKY e HIGGINS, 2007). Como essa área engloba porções tanto na região oeste quanto na região leste do Pacífico, características importantes relacionadas com os padrões de circulação atmosférica em eventos de ENOS, como mudanças na TSM e em gradientes de TSM, podem ser propriamente medidas (BARNSTON *et al.*, 1997). É com base na TSM dessa região que o índice chamado de Oceanic Niño Index (ONI) é construído. O ONI é definido como uma média móvel de três meses sobre a TSM da região NINO 3.4. O seu histórico, que tem início no ano de 1950, é baseado na série de TSM estendida e reconstruída versão 5 (Ersstv5), que foi atualizada no ano de 2017 (HUANG *et al.*, 2017).

Devido à tendência significativa de aumento da TSM no Pacífico por conta do aquecimento global, o órgão responsável pela produção do ONI, *Climate Prediction Center* (CPC), adota uma estratégia para a remoção dessa tendência. O desvio em relação à TSM média, ou anomalia, é calculado com base em um período centralizado de trinta anos, que é estimado a cada janela de cinco anos. Por exemplo, o histórico do ONI entre os anos de 1950 e 1955 é calculado com base na TSM média de 1936-1965. O próximo período de cinco anos (1956-1960) é baseado na TSM média de 1941-1970. Esse procedimento é repetido seguindo a mesma lógica para todo o histórico. O último período base (1986-2015) é utilizado para calcular as anomalias mais recentes. Porém, no ano de 2021, quando haverá um novo período base (1991-2021), os valores da última década serão ligeiramente diferentes por conta da atualização realizada com base no período base mais recente.

Embora haja diversos índices que conseguem capturar algumas causas e efeitos do ENOS, como o IOS e as diversas TSM localizadas na região equatorial do Pacífico, o *National Oceanic and Atmospheric Administration* (NOAA) utiliza como um de seus critérios para o monitoramento operativo o índice ONI. Quando o ONI apresenta anomalias quentes (maior ou igual que $+0,5$ °C) ou frias (menor ou igual que $-0,5$ °C), chamados de condições de El Niño ou de La Niña, por pelo menos cinco períodos seguidos, é então declarado a existência de um evento de EN ou LN. Qualquer outra sequência de valores é classificada como Neutro. Nesta dissertação, a classificação de eventos ENOS será realizada apenas com base no mesmo critério adotado pelo NOAA, devido, principalmente, a sua grande disponibilidade de previsões.

2.2.1 Impactos Relacionados ao El Niño - Oscilação Sul nos Padrões de Precipitação

Sendo o ENOS um fenômeno climático com interações entre a circulação dos oceanos e atmosfera, suas consequências são percebidas por todo o mundo. Eventos extremos como furacões (ELSNER *et al.*, 2001; SAUNDERS *et al.*, 2000), secas e cheias (ANDREWS *et al.*, 2004; TONG *et al.*, 2006), aumento de doenças (CARDENAS *et al.*, 2006; POVEDA, 2001) além de impactos na economia (SOLOW *et al.*, 1998) já foram relacionados com seu acontecimento. Na maior parte dos casos, episódios de EN possuem efeitos contrários quando comparados a eventos de LN.

Um dos impactos mais proeminentes do ENOS pode ser observado na mudança nos padrões de precipitação e de vazões fluviais. Na Europa, um estudo encontrou conexões entre a fase quente do ENOS e o aumento da precipitação no inverno, enquanto episódios de LN foram relacionados a uma tendência de redução no inverno (ZANCHETTIN *et al.*, 2008). Esse impacto aparenta ser mais significativo na região mais ocidental do continente, que é considerada a área mais sensível a flutuações climáticas. Nesse mesmo estudo, também, é apontada a possibilidade da existência de uma relação entre o ENOS e outro padrão de teleconexão, conhecido como Oscilação do Atlântico Norte (OAN), porém o seu mecanismo de interação pode na verdade ser espúrio, ou não tão relevante. Em outro estudo para a Europa, BRÖNNIMANN (2007) revisou os impactos e evidências do ENOS no clima europeu. Na Austrália, uma pesquisa concluiu pela existência da relação entre o ENOS e o fenômeno Oscilação Inter-Decadal do Pacífico (OIP) (POWER *et al.*, 1999). É demonstrada uma grande relação entre o ENOS e a variabilidade interanual da precipitação. Esse fato, porém, não ocorre em conjunto com valores positivos do OIP, somente em valores negativos. Uma ligação entre eventos de EN e o aumento da precipitação no sudoeste asiático, região que compreende países como Iraque, Irã, Afeganistão, Cazaquistão, entre outros, além de parte do leste europeu, foi encontrada em MARIOTTI (2007). Episódios de ENOS também foram altamente relacionados ao sistema de monções do leste asiático (TORRENCE e WEBSTER, 1999; WANG *et al.*, 2000).

As conexões existentes entre o ENOS e as variações na precipitação e nas vazões dos rios são amplamente estudadas no continente americano. Nos Estados Unidos da América (EUA), foi encontrada uma relação entre eventos de EN com o aumento na frequência de dias com ocorrência de eventos extremos de chuva e afluência na região Sudoeste, enquanto que a região Noroeste sofre o efeito inverso (CAYAN *et al.*, 1999). Já em episódios de LN ocorre o oposto. Além disso, observou-se que as alterações nos padrões de vazões fluviais são maiores do que as que ocorrem nas precipitações, podendo ser dez vezes mais frequente o acontecimento de um

evento extremo na primeira variável. PIECHOTA e DRACUP (1996) estudaram a relação entre ENOS e eventos de seca nos EUA, encontrando, em várias regiões, respostas consistentes.

Na América do Sul, a maior parte da precipitação ocorre durante o verão austral (dezembro a fevereiro) e outono (março a maio). A influência do fenômeno ENOS nas chuvas do continente é objeto de estudo há décadas (ACEITUNO, 1988). Para o caso brasileiro, especificamente, uma análise conjunta entre os impactos de eventos de EN na precipitação, temperatura, estrutura de circulação e termodinâmica da atmosfera foi realizada em GRIMM (2003). Esse estudo foi realizado com dados obtidos através de uma grande rede de estações de precipitação (1175) cobrindo o período entre 1956 e 1992, sendo nove desses anos considerados episódios de EN, em conjunto com dados de reanálise atmosférica. As análises do sistema de monção de verão foram feitas tanto em base mensal, quanto sazonal entre novembro e fevereiro. Os resultados apresentados mostram uma variação subsazonal substancial, mostrando que estudos com base mensal são capazes de capturar padrões relevantes. Eventos de EN impactam aumentando ou reprimindo os mecanismos que produzem a precipitação. Em novembro e dezembro, anticiclones são favorecidos nas regiões norte e centro-leste, causando chuvas abaixo da média, enquanto que o fluxo de umidade do oceano Atlântico é transportado para as regiões nordeste e sul, causando anomalias positivas de chuva. Em janeiro, porém, o fluxo de umidade é divergido da região sul para o centro-leste, revertendo o déficit de precipitação nessa região, enquanto que o sul do Brasil passa a ter chuvas com valores médios. Em fevereiro, esses padrões são novamente revertidos, e a região norte do Brasil e a Zona de Convergência do Atlântico Sul (ZCAS) passam a sofrer com chuvas abaixo da média. Esses resultados demonstrados, que não foram detectados em estudos com base sazonal, corroboram a importância da análise em base mensal dos efeitos do ENOS nas precipitações. Ainda que os resultados sejam bastante consistentes, é apontado que deve-se ter cautela ao inferir sobre os efeitos de eventos de EN para novembro e dezembro, dado que eles podem variar de um ano para o outro.

Um estudo similar avaliou os impactos da LN no sistema de monção do verão no Brasil (GRIMM, 2004), considerando ao todo oito eventos de LN. Em novembro, foram identificadas fortes anomalias positivas para a precipitação na região nordeste da Amazônia, e ligeiramente acima nas regiões norte e centro-leste do Brasil, enquanto, na região sul, o oposto foi observado. Já em dezembro, a anomalia positiva na região nordeste da Amazônia é mantida, enquanto, na região centro-leste é enfraquecida, e a região sul passa a ter precipitações próximas dos valores médios. Em janeiro, alterações bruscas ocorrem nas anomalias encontradas anteriormente. No nordeste da Amazônia, a anomalia positiva é enfraquecida, enquanto que a região sul do Brasil passa a ter precipitações acima da média na sua área mais ao norte,

e próximos da média na área mais ao sul. A região centro-oeste passa a ter chuvas acima da média enquanto que no centro-leste, incluindo a ZCAS, anomalias negativas são observadas. Em fevereiro, a anomalia negativa na região Centro-Oeste é revertida e tornam-se ligeiramente positivas em algumas áreas. No extremo sul do Brasil, anomalias negativas voltam a aparecer, assim como as anomalias positivas na região nordeste da Amazônia. Apesar dos resultados mostrarem alguns padrões fortes com efeitos inversos em comparação com eventos de EN, é apontado que as mudanças no estado atmosférico, durante episódios de LN, encontradas entre os meses, dificilmente seria a única causa das alterações encontradas.

As Figuras 2.7 e 2.8 mostram o resumo dos resultados encontrados para os percentis de precipitação média esperados no Brasil durante eventos de EN e LN, respectivamente, em novembro (lado esquerdo superior), dezembro (direito superior), janeiro (esquerdo inferior) e fevereiro (direito inferior). O índice '(0)' que aparece ao lado dos meses de novembro e dezembro, indica o ano em que se iniciou o evento, e o índice '(+)' indica o ano subsequente. As cores vermelhas representam percentis baixos, enquanto as cores azuis estão relacionadas a valores altos. As regiões em cinza claro e escuro representam os resultados com 85% e 90% de nível de significância. Como pode ser observado, em muitos casos, os efeitos de episódios de EN e LN são opostos.

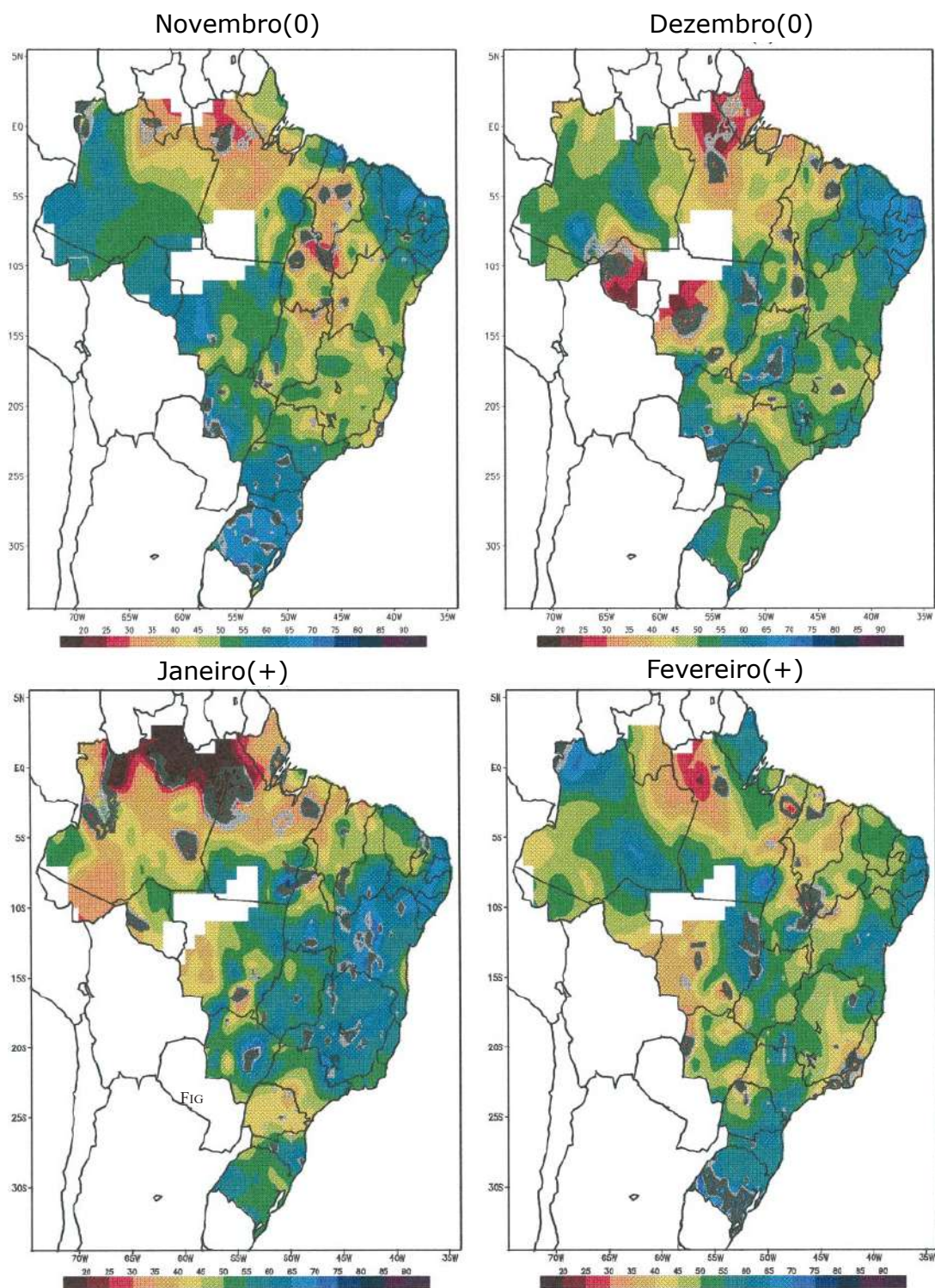


Figura 2.7: Percentis de precipitação média esperada no Brasil durante episódios de El Niño. Fonte: adaptado de GRIMM (2003)

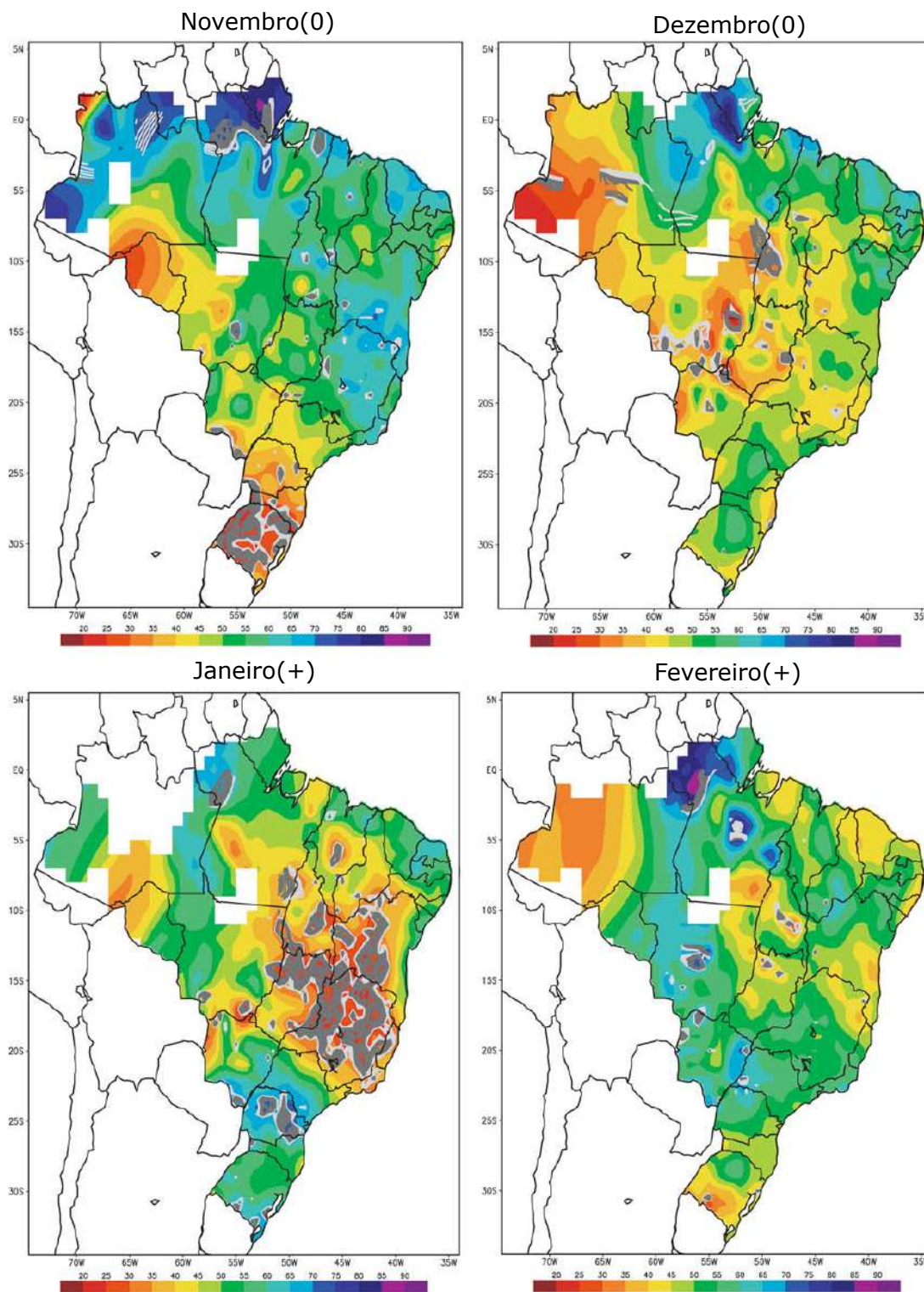


Figura 2.8: Percentis de precipitação média esperada no Brasil durante episódios de La Niña. Fonte: adaptado de GRIMM (2004)

Em UVO *et al.* (1998), foi feita uma análise da relação entre os padrões de precipitação mensal na região nordeste do Brasil e a TSM dos oceanos Atlântico e Pacífico. Demonstra-se que a TSM do Pacífico não é a única forçante climática capaz de modular as anomalias de precipitação da região nordeste, sendo a TSM do

Atlântico um importante fator. Um estudo numérico realizado em PEZZI e CAVALCANTI (2001) corrobora essa relação entre as TSMs do Atlântico e do Pacífico e os padrões de precipitação nessa mesma região. Além dessa conclusão, destaca-se que, durante eventos de EN, quando a TSM do Pacífico é maior que a média, o resto da América do Sul não é afetada pela TSM do Atlântico. Porém, em episódios de LN, a TSM do Atlântico passa a influenciar nas precipitações do continente, especialmente na região central e sul do Brasil.

Uma avaliação da influência do fenômeno ENOS em eventos de precipitação extrema na América do Sul é apresentada em GRIMM e TEDESCHI (2009). Utilizou-se nesse estudo uma grande base de dados, contendo mais de dez mil estações cobrindo o período entre 1956 e 2002. Eventos extremos são definidos como períodos em que a média móvel de três dias é superior ao percentil 90 após o ajuste de uma distribuição do tipo Gamma para cada dia do ano. Foi encontrada uma influência significativa de episódios de ENOS na frequência e intensidade de eventos de chuva intensa em diversas regiões da América do Sul. As alterações na frequência em que os eventos de chuva intensa ocorrem são quase simétricos entre os eventos LN e EN, podendo essa mudança ser observada mais claramente durante o período úmido (verão austral). Esse comportamento também é consistente com a variação da precipitação mensal total durante eventos de ENOS. Quando uma região apresenta um aumento na chuva mensal, a frequência de eventos extremos tende a crescer. O oposto, porém, nem sempre é verdadeiro. Conclui-se que o impacto nos eventos extremos é observado de forma bastante dispersa por todo o continente, indicando que o fenômeno ENOS possui uma influência ainda maior nas chuvas diárias em comparação com mensais.

Um estudo interessante mostra que é necessário cuidado ao se fazer previsões da precipitação e de seus impactos diretos, como as vazões fluviais, condicionados a eventos de ENOS na América do Sul (AMBRIZZI *et al.*, 2004). Nesse trabalho, episódios de ENOS são divididos em dois grupos: o primeiro é composto por anos em que os impactos do ENOS na precipitação se comportam da forma esperada conforme estudos anteriores; já o segundo grupo é composto por eventos de EN e LN que tiveram padrões de atuação diferentes do primeiro. AMBRIZZI *et al.* (2004) investigou as interações atmosféricas e o comportamento das células de Walker e Hadley durante os eventos. O segundo grupo apresentou, para as duas fases extremas do ENOS, padrões opostos ao primeiro grupo quando analisadas a circulação atmosférica e a precipitação. Conforme esteja presente a influência da TSM do Atlântico, uma intensificação da ZCAS e da Zona de Convergência Intercontinental (ZCIT) pode ocorrer, desenvolvendo os padrões encontrados no segundo grupo. A Figura 2.9 apresenta a diferença encontrada nas células de Hadley e Walker entre os dois grupos analisados, sendo o grupo (I) correspondente aos impactos canônicos

que ocorrem durante eventos de ENOS e o grupo (II) associado com os impactos reversos. Enquanto os movimentos de subsidência são relacionados a chuvas abaixo da média (setas orientadas de cima para baixo), movimentos ascendentes (setas orientadas de baixo para cima) representam o oposto. As áreas sombreadas em azul e em vermelho no continente Sul-Americano representam anomalias negativas e positivas, respectivamente, de precipitação no continente. Já as áreas em azul e em vermelho no oceano representam anomalias positivas e negativas, respectivamente, da TSM.

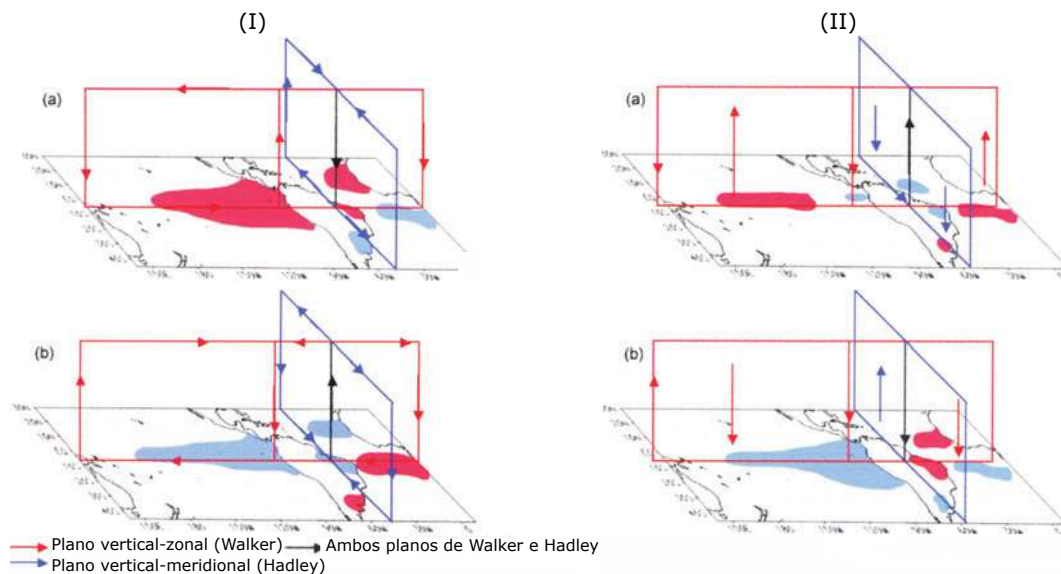


Figura 2.9: Células de circulação de Walker Walker e Hadley para eventos de El Niño (a) e La Niña (b) para os impactos canônicos (I) e reversos (II). Fonte: adaptado de AMBRIZZI *et al.* (2004)

Parte dos estudos revisados neste capítulo apresentam uma relação sólida entre eventos de ENOS e alterações nos padrões de precipitação. Ainda que não seja uma relação direta de causa e efeito, as evidências decorrentes do levantamento bibliográfico realizado acima indicam que o ENOS é um dos fenômenos climáticos mais importantes, com capacidade de alterar os padrões de precipitação no Brasil. Sendo a relação chuva-vazão uma das bases do ciclo hidrológico, espera-se que também haja impactos nas afluições hidrológicas às UHEs.

2.2.2 Impactos Relacionados ao El Niño - Oscilação Sul nos Padrões de Vazões Fluviais

Sob o ponto de vista da sinergia entre chuva e vazão, TREISTMAN (2015) quantificou a correlação existente entre os eventos de precipitação e a vazão natural afluente mensais em cinco UHEs, representativas de diferentes regiões hidrográficas do Brasil. Dados consistidos de 814 estações pluviométricas, com pelo menos trinta

anos de histórico, foram analisados em conjunto com o histórico de vazões naturais^{2,3} durante o mesmo período para as cinco UHEs. Uma alta correlação foi encontrada entre as chuvas e as vazões naturais mensais contemporâneas para quase todas as UHEs analisadas durante o período úmido que ocorre na maior parte das bacias hidrográficas (outubro a março). A UHE Sobradinho representa uma exceção, tendo o maior valor de correlação entre a vazão natural afluyente com a precipitação com um mês de atraso. Como a maior parte da chuva ocorre durante o período úmido, as vazões afluentes do período seco (abril a setembro) são, em sua maioria, correlacionadas com as chuvas dos meses anteriores.

Outra exceção é apontada para a UHE Itá, que se localiza na região sul. Essa área do Brasil apresenta pouca sazonalidade para as duas variáveis, com chuvas distribuídas ao longo do ano. Consequentemente, todos os meses apresentam valores altos de correlação entre as duas variáveis sem atraso, demonstrando que as bacias hidrográficas, nessa região, têm uma rápida resposta a eventos de precipitação. Para essa mesma UHE, a correlação anual entre a precipitação anual e a vazão natural anual é de 0,97, um valor extremamente alto. A Figura 2.10 apresenta as anomalias anuais das duas variáveis para a UHE Itá, onde pode ser observado com clareza a forte relação entre elas. Destacam-se o ano de 1962, em que ocorreu uma LN, apresentando uma das menores anomalias, e o ano de 1983, caracterizado por um episódio de EN, possuindo o maior valor de anomalia nas duas séries. É possível observar, também, que a variação das anomalias de vazão natural afluyente é maior que nas precipitações durante os eventos de ENOS.

²Vazões naturais são aquelas que ocorreriam em uma seção do rio caso não houvesse ações antrópicas na sua bacia contribuinte – tais como regularizações de vazões realizadas por reservatórios, desvios de água, evaporações em reservatórios e usos consuntivos (irrigação, criação de animal e abastecimento urbano, rural e industrial).

³O processo de consistência dos dados, e naturalização das vazões é realizado pelo ONS. Maiores informações sobre esses dados podem ser obtidos no <http://ons.org.br/>

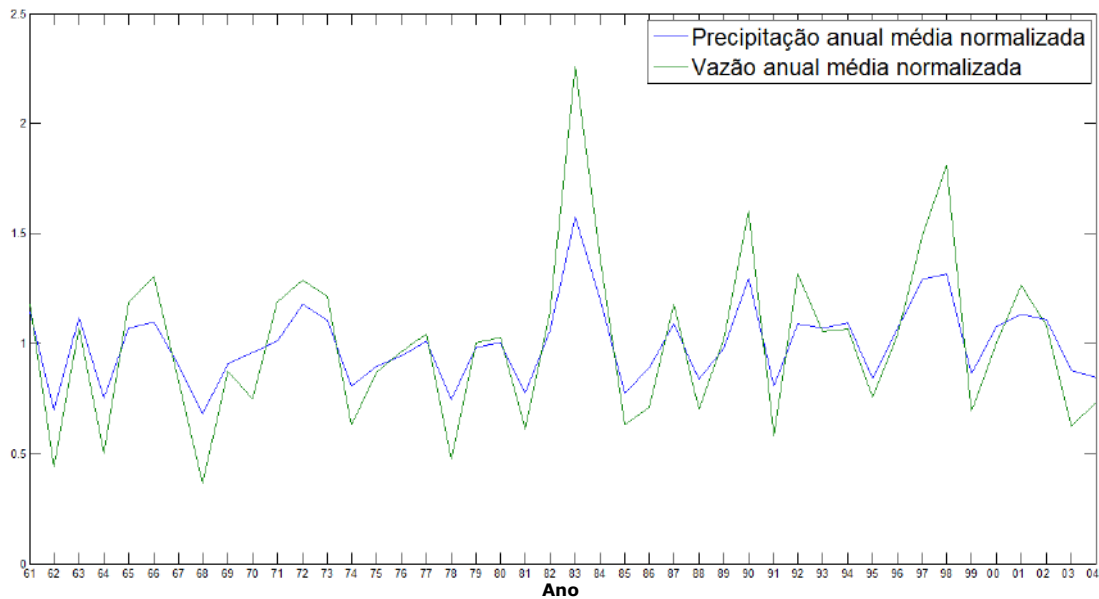


Figura 2.10: Anomalias de precipitação e vazão natural af luente anuais para a UHE Itá entre 1961 e 2004. Fonte: TREISTMAN (2015).

Como o fenômeno ENOS pode modular a intensidade e a frequência das precipitações no Brasil, e conhecida a forte relação existente entre as chuvas e as vazões afluentes, é esperado que as af luências nos rios brasileiros sejam afetadas em algum grau, por esse fenômeno climático. Diversos estudos sobre a associação do ENOS com alterações nos padrões de af luências fluviais brasileiras vêm sendo realizados ao longo dos anos.

DETTINGER *et al.* (2000) buscou identificar a relação entre o ENOS e as vazões fluviais que ocorrem em rios por todo o mundo, com foco no continente americano. Utilizando-se de uma extensa base de dados, através do uso de técnicas estatísticas, como análise de correlação e análise de componentes principais, foi possível identificar as relações existentes entre as vazões fluviais e o fenômeno ENOS em uma escala global. Para o Brasil, especificamente, identificaram-se alterações nos padrões de vazão semelhantes aos padrões citados anteriormente nesta dissertação para as precipitações. Segundo os resultados encontrados, episódios de EN tendem a ser mais secos que eventos de LN na bacia Amazônica, principalmente em sua porção superior, ocorrendo o oposto no sul do continente. Conclusões semelhantes para o território brasileiro também foram relatados em MARENGO *et al.* (1998), CAMILLONI e BARROS (2000) e GRIMM e TEDESCHI (2009).

CAPOZZOLI *et al.* (2017) avaliou a relação entre diferentes padrões climáticos com as vazões naturais em 163 UHEs no SIN, utilizando a análise de componentes principais. Analisando os picos espectrais e os padrões dos fenômenos climáticos, apontou-se que as duas primeiras componentes principais explicam mais de 50% da variabilidade, podendo estar associadas aos fenômenos ENOS, às os-

cilações decadais, como a Oscilação Decadal do Pacífico (PDO) e a Oscilação Multi-decadal do Atlântico (AMO), bem como a anomalia da TSM no Atlântico Tropical Sul (TSA). São destacadas, também, no segundo componente principal as variabilidades opostas entre a região sul do Brasil, e a porção restante do país, principalmente sobre as bacias do São Francisco e Tocantins.

2.2.3 Previsões para o El Niño - Oscilação Sul

As previsões de episódios do ENOS são baseadas, em sua maioria, em previsões da TSM na região NINO 3.4. O *International Research Institute for Climate and Society* (IRI, 2018) faz uma coletânea de previsões realizadas para as anomalias nessa região por modelos desenvolvidos por diversos institutos de pesquisa (vinte e cinco modelos até o relatório de novembro de 2017) e disponibiliza publicamente em seu endereço eletrônico ⁴. Nesse relatório, são incluídos apenas modelos com base mensal, onde a previsão é realizada para dez períodos à frente. Essas informações são disponibilizadas a cada mês de duas formas: através de uma pluma (conjunto) de previsões, como apresentado na Figura 2.11; e através de uma tabela de previsões probabilísticas dividida entre as três condições de ENOS. É também disponibilizada um valor médio das anomalias considerando apenas os modelos de previsão de origem dinâmica, apenas modelos estatísticos, para todo o conjunto de previsões, e para um valor de consenso do *Climate Prediction Center* - CPC, não considerado no valor médio do conjunto de previsões. Modelos dinâmicos de previsão são baseados em equações que buscam representar as relações físicas entre o oceano e a atmosfera. Já modelos estatísticos baseiam-se em modelos de regressão, de inteligência artificial, e em outras técnicas análogas que utilizem dados da série histórica para prever o futuro. O valor esperado pelo consenso do CPC para o índice baseia-se na informação disponível por todos os modelos, segundo o estado atual da atmosfera, consoante o julgamento humano dos especialistas do IRI e CPC. Desse diagnóstico, também é produzida uma outra tabela de previsão probabilística, que leva em conta as percepções dos pesquisadores.

⁴ <http://iri.columbia.edu/our-expertise/climate/forecasts/enso/current>

Tabela 2.1: Previsão probabilística de ENOS obtida em novembro de 2017. Fontes: adaptado de IRI (2018).

Trimestre	La Niña	Neutro	El Niño
NDJ 2017	75%	25%	0%
DJF 2018	71%	29%	0%
JFM 2018	60%	40%	0%
FMA 2018	45%	54%	1%
MAM 2018	26%	72%	2%
AMJ 2018	16%	76%	8%
MJJ 2018	17%	61%	22%
JJA 2018	17%	51%	32%
JAS 2018	16%	46%	38%

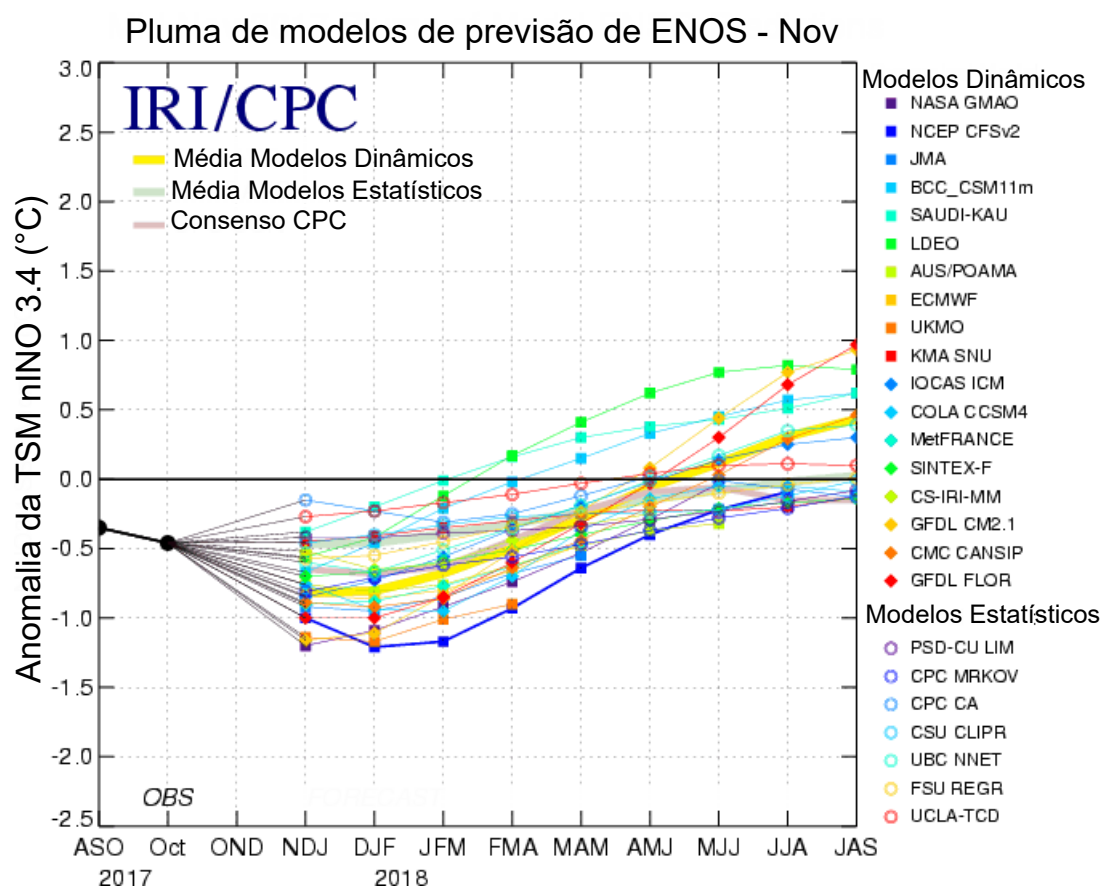


Figura 2.11: Pluma de previsão da anomalia da TSM NINO 3.4 obtida em novembro de 2017. Fonte: adaptado de IRI (2018).

As previsões probabilísticas são baseadas em todos os modelos disponíveis, ponderados igualmente, em conjunto com uma função de desvio padrão centrada na média. Essa função é construída como uma curva Gaussiana, levando em conta a

precisão global esperada para cada período do ano e para o horizonte em que está sendo feita a previsão. Quanto mais próximo do mês em que se iniciam as previsões, menores são os erros, e, assim, mais estreita é a distribuição. De forma geral, a precisão das previsões é maior durante o inverno austral (junho a agosto), quando os eventos de ENOS tendem a se iniciar, enquanto que, no verão austral (dezembro a fevereiro), as previsões são mais precisas.

BARNSTON *et al.* (2012) analisaram a capacidade preditiva da pluma de previsões em tempo real do IRI durante os anos de 2002 e 2011 em comparação com as previsões retrospectivas fornecidas por alguns dos modelos incluídos na pluma. Apesar da precisão da previsão em tempo real ser ligeiramente menor do que os modelos menos avançados das décadas passadas, os resultados podem ter sido influenciados pelo período de estudo escolhido, que se mostrou particularmente mais desafiador que o período 1981-2010. Considerando apenas os últimos nove anos das previsões retrospectivas, a pluma de previsões apresenta resultados ligeiramente superiores. Além disso, verificou-se que ambas as previsões possuem uma tendência de iniciar os eventos de ENOS mais tarde que o observado, ou com uma intensidade menor, sendo as previsões mais longas piores. Esse tipo de erro é intensificado durante o período de transição do ENOS.

Em um estudo anterior, TIPPETT e BARNSTON (2008) mostraram alguns dos benefícios das previsões probabilísticas de ENOS. Foram coletadas as previsões do projeto *Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction* (DEMETER). Para produzir as previsões probabilísticas de ENOS, diferentes métodos de calibração foram utilizados, desde métodos simples sem qualquer tipo de ponderação, em que cada membro da previsão era classificado entre os três distintos estados de ENOS, até métodos mais complexos de calibração conjunta, onde cada membro foi ponderado de forma ótima. As análises basearam-se nas previsões produzidas de 1980 a 2001. Através das métricas utilizadas, concluiu-se que as previsões produzidas pelo conjunto de todos os membros era melhor que a utilização de um único membro que possuía a melhor performance. Uma interessante conclusão desse estudo mostra que, apesar do período de transição do ENOS ser quando os modelos de previsão tem a menor previsibilidade, é exatamente nele que existe a maior vantagem frente a utilização única dos valores históricos.

TIPPETT *et al.* (2012) também analisaram as previsões probabilísticas produzidas através da pluma de previsões do IRI. Dois tipos de previsão probabilística foram utilizados: uma versão simplificada usando apenas a frequência relativa obtida entre as previsões da pluma enquadrada entre os três estados de ENOS; e outra utilizando as previsões da pluma e um ajuste da distribuição de probabilidade através de um modelo gaussiano. Não foi verificado um ganho na previsão do segundo modelo em relação ao primeiro. A Figura 2.12 apresenta a categoria prevista para o

ENOS em cada trimestre entre os anos de 2002 a 2011, e a categoria observada. O eixo horizontal corresponde ao tempo em que foi realizada a previsão, e o eixo vertical representa o número de passos à frente em relação ao período inicial. Os valores 1, 0 e -1 representam os estados de EN, N e LN respectivamente. Os valores positivos ou negativos são apresentados quando as categorias de EN e LN possuem a maior probabilidade, enquanto que valores de zero representam eventos N. Pode ser observado um declínio do poder de previsão em relação ao tempo inicial da previsão, além da tendência em iniciar o início dos eventos um pouco mais tarde.

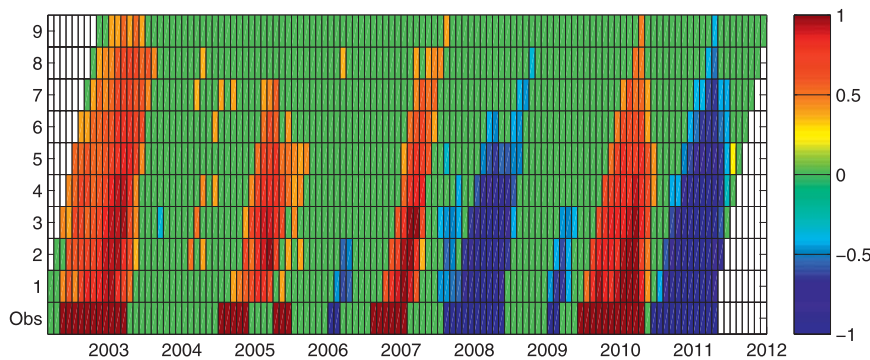


Figura 2.12: Categorias observadas e previstas de ENOS. Fonte: adaptado de: TIPPETT *et al.* (2012)

2.3 Modelos Estocásticos em Hidrologia

Em grande parte das situações, é muito difícil prever deterministicamente o que de fato ocorrerá em um instante futuro para fenômenos naturais, como precipitação em uma bacia hidrográfica ou a vazão em determinada seção de um rio. Apesar de existirem modelos que produzam previsões determinísticas, sempre haverá um erro, pois modelos são apenas formas aproximadas para representar a natureza.

O conjunto de observações de um processo estocástico ordenadas ao longo do tempo (em dias, semanas, anos) compõe uma série temporal. A evolução da série temporal de um fenômeno natural, como a vazão de um rio, desenvolve-se seguindo leis probabilísticas que podem ser aproximadas por expressões matemáticas. Pode-se interpretar que a série histórica observada é apenas uma possível realização do processo estocástico que a produziu, delimitando, portanto, uma série temporal. O mesmo processo estocástico que gerou a série histórica pode gerar também outras séries temporais. Essas séries, chamadas de sintéticas, também possuem uma chance de ocorrência. Um conjunto de séries sintéticas, ou cenários sintéticos, guarda as mesmas características estatísticas observadas na série histórica.

A ideia de se gerar cenários sintéticos através de um modelo estocástico tem como um dos seus objetivos inferir sobre os passos futuros de uma série temporal,

e utilizar essa informação em algum tipo de tomada de decisão, como dimensionamento de reservatórios (KOTTEGODA, 1980), ou para o planejamento da operação energética, que é onde se enquadra essa dissertação. Atualmente, para o SIN, o modelo GEVAZP gera uma árvore de cenários para o segundo mês do planejamento da operação de curto prazo; já para o planejamento da operação de médio/longo prazo são gerados cenários sintéticos com o horizonte de dez anos. Ambas as etapas do planejamento energético requerem possíveis realizações de vazões futuras, o que torna útil a modelagem estocástica das aflúências. Desta forma, o SIN passa a estar preparado para operar de forma eficiente os seus recursos, inclusive para eventos extremos.

Modelos estocásticos tradicionais utilizados na geração de cenários sintéticos de vazões, usualmente derivados da família de modelos BOX e JENKINS (1970), descrevem as realizações futuras como uma regressão das observações passadas (HIPPEL e MCLEOD, 1994; SALAS *et al.*, 1980). Dessa forma, são utilizadas apenas variáveis endógenas na representação do processo estocástico de aflúências. Dada as conexões existentes entre as vazões afluentes de rios e outras variáveis do ciclo hidrológico (evaporação, precipitação), e as teleconexões com outros fenômenos climáticos, como ENOS, PDO, que podem ser identificados através da TSM de algumas regiões específicas, ou ventos, alguns trabalhos tem buscado incluir essas informações exógenas em modelos estocásticos hidrológicos. Com isso, espera-se determinar com uma precisão maior as distribuições multivariadas de probabilidades condicionadas para as aflúências futuras dada todas as informações relevantes disponíveis em um certo instante.

2.3.1 Modelos de Geração de Cenários Sintéticos de Vazões com uso de Informações Climáticas

Recentemente, diversas metodologias foram desenvolvidas de forma a incluir informações climáticas nos modelos estocásticos utilizados na previsão e geração de séries sintéticas de aflúências. Apesar de conceitualmente ser fácil deduzir que as vazões afluentes são fruto de diversos processos físicos conectados, a consideração de informações climáticas é um desafio em constante desenvolvimento na hidrologia estocástica.

COLONESE (2016) investigou a utilização de previsão quantitativa de precipitação fornecida pelo modelo ETA (MESINGER *et al.*, 2012) com resolução de 15 km em modelos de previsão de vazões. Uma minuciosa análise sobre os dados do modelo atmosférico foi realizada, apontando possíveis caminhos de sua inserção visando aprimorar as previsões de aflúências diárias e semanais obtidas para quatro UHEs localizadas na região sul do Brasil através de dois modelos distintos: um modelo

estocástico - PREVIVAZH (DE CARVALHO, 2001); e um modelo determinístico - SMAP II em sua versão suavizada (SILVA, 2012).

O trabalho desenvolvido por VIEIRA (1997) teve como um de seus objetivos considerar a influência do ENOS no modelo auto-regressivo periódico de ordem p - PAR(p) - através de uma técnica que não alterasse a estrutura do modelo. O primeiro passo consistiu na investigação da teleconexão entre as séries de vazões (mensais e anuais) e a série do IOS (Índice de Oscilação Sul) através das funções de correlação cruzada. O ano hidrológico considerado se inicia em novembro e termina em outubro do ano subsequente. Detectou-se um relacionamento contemporâneo (*lag* 0) significativo em UHEs pertencentes às bacias hidrográficas dos rios Grande, Paranaíba, Paraná, Tietê e São Francisco. Já para as bacias do rio Paraíba do Sul, Ribeirão das Lajes (região sudeste) e Iguaçu e Jacuí (região sul) não se mostraram correlações significativas. Outro fato relevante desta análise resultou na conclusão que as séries anuais apresentaram um relacionamento mais forte com a série do IOS que as séries mensais.

A proposta para o modelo PAR(p) de VIEIRA (1997) inclui um algoritmo alternativo para estimação dos parâmetros através de um critério de similaridade para classificar os anos do histórico segundo sua similitude com o estado corrente da variável IOS, considerando os últimos 12 meses. Para isso, calcula-se o desvio médio absoluto dado pela razão entre o estado corrente e os estados passados. Durante a estimação dos parâmetros do PAR(p), cada ano do histórico é ponderado através de um peso estimado com relação ao desvio calculado no passo anterior. A aplicação dos pesos é feita seis meses à frente do período em que ele foi calculado. Realizaram-se testes utilizando o algoritmo proposto para a previsão de vazões da UHE Furnas até doze meses a frente, em três anos hidrológicos descritos como secos e três úmidos. Segundo as métricas utilizadas, mostrou-se que há um ganho, apesar de modesto, na maioria dos anos analisados quando considerada a variável macroclimática IOS na previsão de vazões a essa usina. A modelagem proposta permite a utilização da informação climática até doze meses a frente apenas. Com isso, destaca-se que a geração de variáveis macroclimáticas poderia prolongar a utilização deste algoritmo para estudos de longo prazo. Outra aplicação desse mesmo modelo no caso do sistema de UHEs na Colômbia pode ser encontrado em KELMAN *et al.* (2000), onde se mostraram ganhos nos cenários gerados com informação climática em comparação ao modelo tradicional.

Uma metodologia para a consideração do índice IOS na geração de cenários de vazões diárias é descrita em COSTA *et al.* (2014). Nessa aplicação, voltada para o controle de cheias, o registro histórico de vazões é classificado, utilizando o índice IOS, entre anos secos, normais e úmidos. Dado essa classificação, três conjuntos de parâmetros são estimados para a geração de cenários sintéticos de vazões diárias

condicionadas ao fenômeno ENOS, por meio do modelo DIANA (KELMAN *et al.*, 1983).

ANDERSON *et al.* (2015) incorporaram a informação disponível no conjunto de previsões do ONI fornecida pelo IRI em um modelo estocástico não paramétrico utilizado no gerenciamento dos reservatórios da bacia hidrográfica do rio Baixo Colorado no Texas (ANDERSON e WALKER, 2011). Cada previsão existente na pluma foi classificada entre as três fases do fenômeno ENOS (La Niña, Neutro, El Niño) com o intuito de se calcular uma matriz de probabilidade de transição entre estados para os nove meses de previsão fornecida. Utilizando o histórico de vazões, classificado em meses secos, moderados e úmidos, e sua matriz de transição entre esses estados em conjunto com a matriz de transição entre as condições de ENOS, o modelo proposto faz uma reamostragem do histórico, simulando através de uma cadeia de Markov, dois mil cenários para os próximos sessenta meses. Segundo os resultados, há um ganho na geração dos cenários futuros, tendo sido estes particularmente úteis para as projeções do ano de 2015 no gerenciamento da bacia hidrográfica.

Um dos modelos estocásticos mais utilizados em estudos recentes para a previsão e geração de cenários sintéticos de vazões com a incorporação de informações climáticas é o modelo auto-regressivo periódico com variáveis exógenas (PARX). Esse tipo de modelagem utiliza, além das afluições passadas, outras variáveis externas ao processo que está sendo modelado que possuam alguma relação. Assim, espera-se que haja ganhos na modelagem da variável de interesse. O trabalho descrito em SANGY (2016) adicionou a previsão de precipitação mensal fornecida pelo modelo ETA40 do Instituto Nacional de Pesquisas Espaciais (INPE) no processo de geração de cenários sintéticos de vazões em cinco UHEs do SIN, através do ajuste de modelos PARX. Apesar de ser demonstrado que a utilização de dados de previsão de chuva pode trazer ganhos significativos para a geração de cenários de vazões, principalmente durante o período úmido, a qualidade da previsão de precipitação para um horizonte mensal ainda é baixa, o que acaba limitando o aprimoramento dos resultados da geração de cenários de vazão. Neste sentido, em alguns casos, demonstra-se que a inserção da variável exógena pode inclusive levar a resultados piores quando comparados ao modelo tradicional.

Uma abordagem similar foi realizada por DE FARIAS (2012) para a geração de cenários de afluições na UHE Furnas. Nesse trabalho, foram ajustados diversos modelos do tipo auto-regressivo (AR) com ordens diferentes, e auto-regressivos com variáveis exógenas (ARX), sendo essas variáveis a precipitação observada na bacia hidrográfica referente a UHE e a precipitação prevista, fornecida pelo modelo de circulação global ECHAM 4.5. Cada modelo do tipo AR foi responsável por gerar um único cenário (uma previsão), enquanto que o modelo ARX ajustado gerou vinte e quatro cenários distintos, ambos para o período de teste de janeiro de 2002 até

dezembro de 2007. Segundo as métricas avaliadas, chegou-se à conclusão que o modelo ARX foi capaz de capturar de forma mais adequada variações bruscas na série de vazões, obtendo um desempenho superior.

Em LIMA e LALL (2008), foram comparados os resultados de previsão de afluências feitas com até seis meses de antecedência para 57 UHEs do SIN através de quatro modelos: PAR de ordem 1; PARX, sendo a variável exógena a TSM entre as latitudes 25°N e 45°S dos oceanos Pacífico e Atlântico; e outras duas variantes com a inclusão da correlação espacial. A partir dos resultados obtidos, analisando-se o coeficiente de eficiência médio, o modelo PAR obteve, em média, uma melhor performance, apesar dessa conclusão não ser uma realidade para todas as UHEs e meses analisados. Porém, observou-se que, nos meses com maiores afluências, em UHEs com influência de forçantes remotas (variáveis climáticas não diretamente relacionadas), o modelo PARX apresentou um desempenho melhor de forma geral.

Uma análise similar foi realizada em LIMA e LALL (2010), onde foram comparadas as previsões para um horizonte de seis meses. Foram utilizados os modelos PAR, PARX com a possibilidade de inclusão da TSM da região NIÑO 3, TSM do Atlântico Sul e índices de vento zonal de baixo-nível na região sudeste do Brasil, com atraso de um até seis meses, e outro modelo PARX onde os parâmetros da equação de regressão são ajustados com uma metodologia diferente. Os resultados demonstraram que a inclusão das variáveis climáticas aprimorou as previsões para até três meses à frente para a maioria das UHEs durante o período de teste. Observou-se que a utilização de variáveis climáticas incluídas com um atraso maior pode melhorar as previsões em horizontes maiores, quando comparado com a inclusão de variáveis com atraso menor.

Em MAÇAIRA *et al.* (2017), utilizou-se um modelo PARX com até oito variáveis exógenas para geração de cenários de afluências para oito UHEs em diferentes regiões, apresentando apenas uma ligeira vantagem sobre a modelagem PAR(p). Em SILVEIRA *et al.* (2017), foi feita uma comparação entre um modelo PARX com até quatro variáveis exógenas e o modelo PAR(p) para a previsão de afluências em 88 UHEs brasileiras. Os resultados foram analisados separadamente por mês, mostrando que, para a maior parte do ano, o modelo PAR(p) é capaz de produzir previsões melhores, sendo superado apenas no período entre setembro e novembro pelo modelo PARX.

Buscando uma nova abordagem para previsão e geração de cenários de vazões com a inclusão de informações climáticas, outra classe de modelo estocástico é apresentada em GELATI *et al.* (2010). Define-se nesse artigo, o modelo autorregressivo com variáveis exógenas modulado por cadeia de Markov (MARX), onde os parâmetros do modelo ARX são alternados de acordo com um regime climático modelado por cadeia oculta de Markov, tendo a probabilidade de transição entre os

estados como função de informações climáticas. Esse modelo é uma extensão natural do trabalho apresentado em HAMILTON (1989), que introduziu os modelos da família *Markov-Switching Autorregresive* (MS-AR) para modelar o produto interno bruto, onde os estados ocultos representavam as fases de crescimento econômico e de recessão. Após esse trabalho, novos caminhos para modelagem estocástica de séries temporais foram abertos. Outra referência para essa família de modelos pode ser encontrada em KROLZIG (1997), onde a modelagem autorregressiva foi estendida para uma versão vetorial.

O trabalho desenvolvido por GELATI *et al.* (2010) utilizou o modelo MARX para a previsão e geração de cenários de vazão em uma UHE localizada na região oeste do Equador. Nesse modelo, as anomalias das afluições quadrimestrais foram assumidas como função de seus valores passados (para previsão um passo à frente), das variáveis climáticas (previsão e simulação de longo prazo), e são moduladas por um processo climático oculto que também é influenciado pela variável climática. Após um processo de testes com diversas variáveis climáticas relacionadas ao fenômeno ENOS, identificou-se através do critério de informação Bayesiana (BIC), a anomalia da TSM na região NINO 1+2 e o índice Trans-Nino como sendo os mais adequados ao modelo, bem como a utilização de apenas dois estados para a cadeia de Markov oculta, um relacionado a eventos de El Niño e outro a eventos neutros e a La Niña. Para este, trabalho foram utilizadas as observações das variáveis climáticas. Observou-se que, tanto no período de calibração, quanto no de validação, as afluições estimadas obtiveram um bom desempenho, ainda que, para vazões baixas, o modelo, de forma geral, produzisse superestimativas, tendo um melhor ajuste às vazões altas. Esse fato pode ser explicado pela alta correlação de eventos de El Niño com anomalias altas, enquanto eventos mais secos não parecem estar correlacionados. Aplicações dessa modelagem na otimização de reservatórios podem ser encontradas em GELATI *et al.* (2011) e em GELATI *et al.* (2014), onde a previsão das afluições foi realizada de forma multivariada para dois reservatórios.

Os trabalhos apresentados nesta seção buscaram incluir variáveis climáticas na previsão ou geração de cenários, de forma a aprimorar os modelos estocásticos tradicionais. Os modelos que incluem variáveis exógenas, como ARX ou PARX, apesar de apresentarem alguns bons resultados, acabam dependentes da qualidade das previsões para as variáveis climáticas para a geração de cenários futuros. Além disso, esse tipo de modelagem acaba considerando que as variáveis climáticas influenciam somente no nível médio das afluições, ignorando possíveis alterações na variabilidade associadas. Outras modelagens propostas acabam incorporando essa questão, porém ainda assim acabam por ter limitações no horizonte de estudo devido à inclusão da variável climática, ou ainda não se adequam à restrição de linearidade que envolve o modelo NEWAVE. Dessa forma, a metodologia que será apresentada

no próximo capítulo busca inspirar-se nessas pesquisas, porém sendo aplicável ao planejamento da operação SIN.

Capítulo 3

Metodologia

Neste capítulo, apresentam-se os dados utilizados ao longo desta dissertação. Em seguida, detalham-se a formulação matemática e estatística do modelo de geração de séries sintéticas mensais de energia e vazão - GEVAZP - e de um modelo alternativo nele baseado, porém capaz de capturar informações climáticas referentes ao fenômeno ENSO. Por fim, discriminam-se os testes empreendidos na análise dos modelos. A abordagem do problema delineado engloba um conjunto de procedimentos e dados dispostos na Figura 3.1 constituindo a metodologia desta dissertação.

3.1 Dados de Vazão Natural Afluente

O histórico de vazão natural afluente para as 141 UHEs utilizadas nos estudos de casos tratados nesta dissertação foram obtidos no conjunto de dados de entrada do planejamento mensal da operação de janeiro de 2018. Nesta dissertação, serão utilizados os dados entre 1950 e 2017, mesmo período do histórico do índice ONI.

3.2 Dados do Fenômeno El Niño - Oscilação Sul

3.2.1 Histórico do Fenômeno El Niño - Oscilação Sul

A classificação do histórico de ENOS nos seus três possíveis estados (LN, N e EN) utilizará o índice ONI (HUANG *et al.*, 2017), construído como uma média móvel de três meses sobre a TSM da região NINO 3.4. O histórico do ONI, apresentado na Figura 3.2, abrange os anos entre 1950 e 2017. Nesta dissertação, o critério de classificação de estados de ENOS será o mesmo empregado no monitoramento operativo realizado pela *National Oceanic and Atmospheric Administration* (NOAA), ou seja, baseado no ONI:

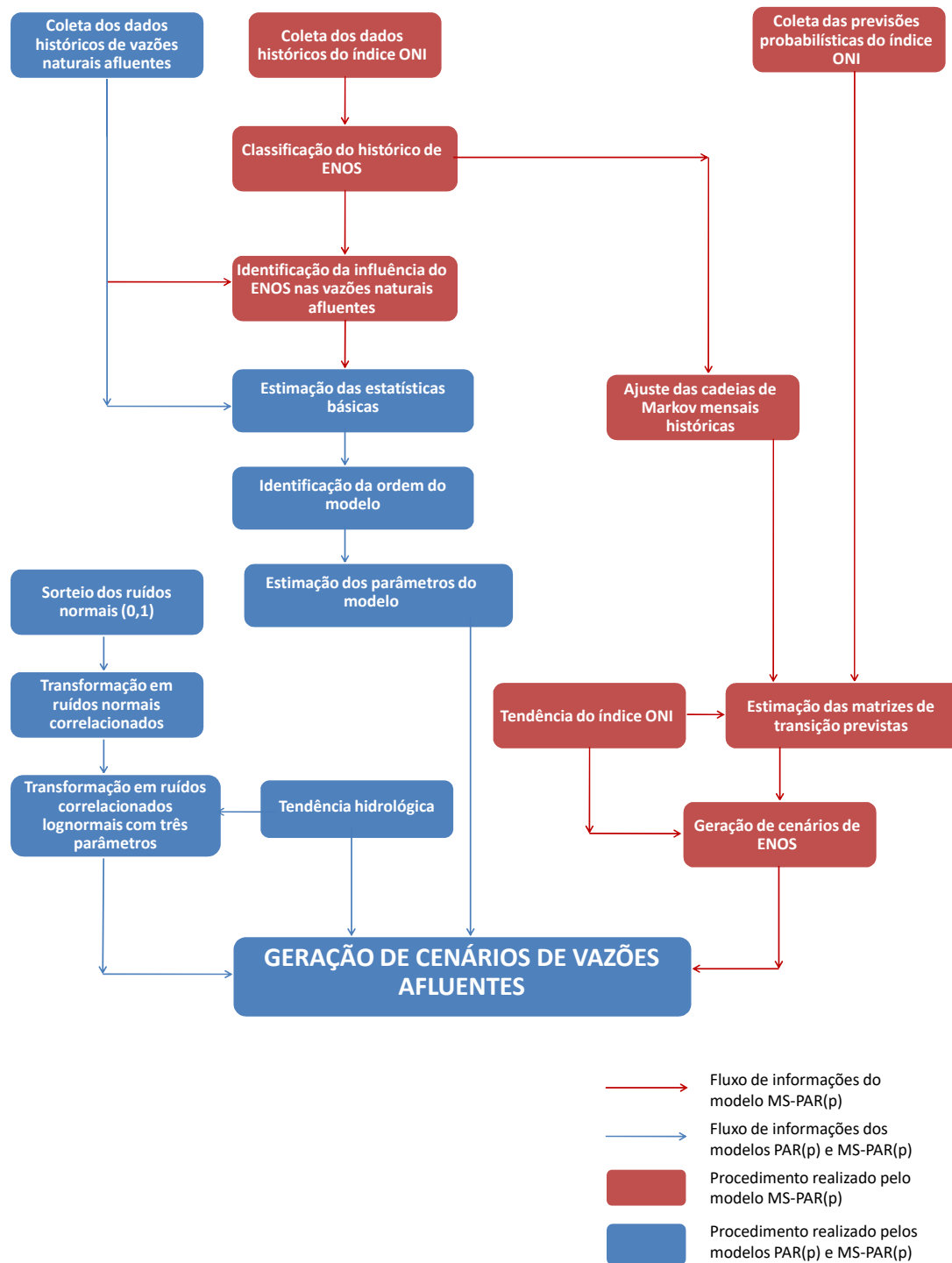


Figura 3.1: Fluxograma metodológico - etapas do trabalho.

1. caso ocorram cinco trimestres seguidos em que o ONI seja menor ou igual a $-0,5$ (condição de LN), o estado desse período é classificado como LN;
2. caso ocorram cinco trimestres seguidos em que o ONI seja maior ou igual a $+0,5$ (condição de EN), o estado desse período é classificado como EN;
3. qualquer outra combinação resulta na classificação de estado N.

Na Figura 3.2, áreas em vermelho (azul) representam episódios, ou estados, de EN (LN), enquanto a linha em vermelho (azul) define o limite das condição de EN (LN). Deve-se observar que o critério de classificação utilizado é dependente tanto dos valores passados de ONI quanto dos valores futuros. Por exemplo, até o final do ano de 2017, os últimos trimestres estão classificados como neutro, apesar do ONI apresentar valores menores que $-0,5$. Após o início do ano de 2018, manteve-se a tendência de condições de LN, caracterizando um episódio de LN.

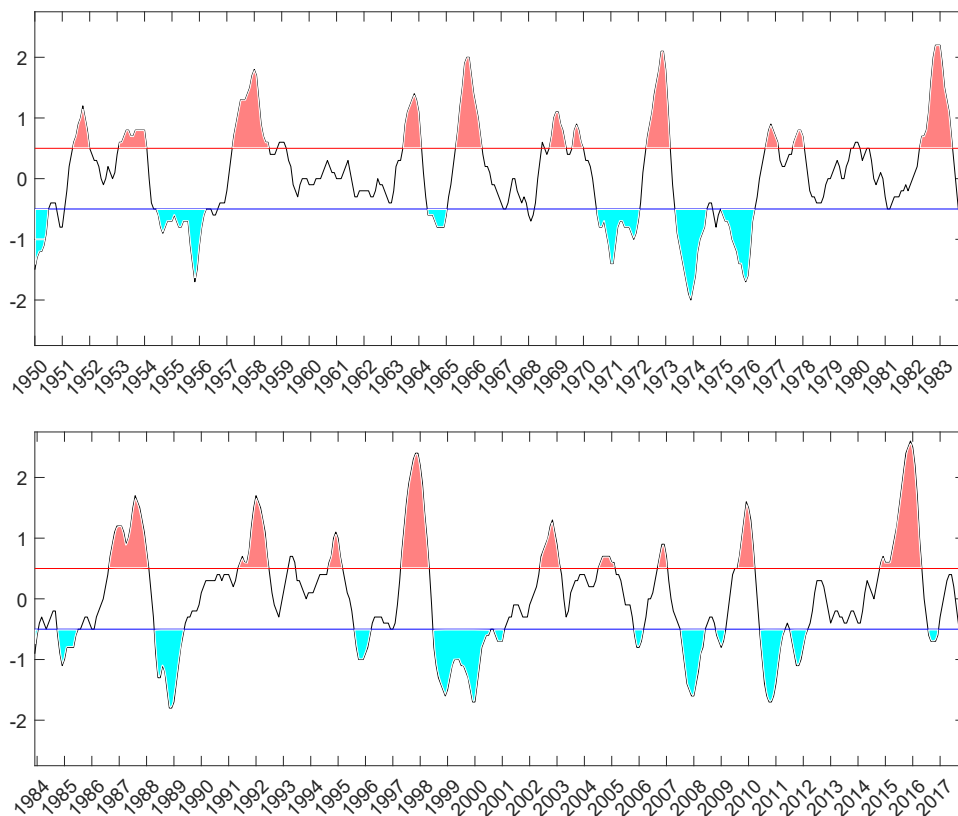


Figura 3.2: Histórico do índice ONI. Áreas em vermelho (azul) representam episódios de EN (LN)

Visando sua utilização no modelo GEVAZP, que gera cenários sintéticos com discretização mensal, é realizada uma adaptação do critério de classificação do ENOS. Para a classificação dos meses entre os três estados de ENOS, será considerado o mês central de cada trimestre como o representativo. Por exemplo, o

índice do trimestre janeiro/fevereiro/março (JFM) será utilizado para classificar o mês de fevereiro; o mês de março será classificado segundo o índice do período fevereiro/março/abril (FMA); e assim sucessivamente. O número de meses classificados entre cada estado é apresentado na Tabela 3.1.

Tabela 3.1: Quantidade de meses classificados entre os estados de ENOS no período de 1950 a 2017

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
La Niña	21	18	17	13	16	15	18	18	18	19	21	21
Neutro	23	29	37	43	37	38	35	35	30	27	23	23
El Niño	24	21	14	12	15	15	15	15	20	22	24	24

Deve-se observar que, ao se utilizar esse procedimento em simulações de estudos oficiais, a classificação do mês anterior ao caso sempre será desconhecida. Exemplificando, supondo que esteja sendo realizada a geração de cenários sintéticos a partir de fevereiro, como o mês de janeiro é classificado pelo índice ONI do trimestre DJF, que depende da TSM observada em fevereiro (ainda não observada), o valor para DJF é inexistente. Então, para o mês anterior ao estudo, também será prevista a classificação de seu estado, usando as previsões probabilísticas fornecidas pelo IRI.

3.2.2 Previsões do Fenômeno El Niño - Oscilação Sul

O modelo proposto nesta dissertação utilizará, como dado de entrada, as previsões probabilísticas fornecidas pelo IRI, que são baseadas nas previsões realizadas por diversos modelos dinâmicos e estatísticos coletados. São fornecidas previsões para nove trimestres à frente. É importante ressaltar que essas previsões dizem respeito à classificação do índice ONI entre as três condições de ENOS. Também deve-se destacar que as previsões fornecidas são feitas a partir do trimestre seguinte ao mês em que foram realizadas. Por exemplo, a primeira previsão fornecida ao final do mês de março diz respeito ao trimestre MAM. Segundo o critério de classificação utilizado, a previsão seria enquadrada para a classificação de abril. Sendo assim, para março, é necessário obter a previsão do trimestre FMA, que foi disponibilizada em fevereiro. Por fim, para a estimação do estado do mês anterior ao caso (fevereiro), coleta-se a previsão disponibilizada em janeiro para o primeiro trimestre (JFM).

3.3 Modelo Autorregressivo Periódico

O modelo GEVAZP é responsável pela geração de cenários multivariados sintéticos de ENA e vazões mensais para os modelos NEWAVE e DECOMP. Os cenários gerados pelo GEVAZP representam a principal incerteza para o planejamento do SIN e, por essa razão, são essenciais para o cálculo da política ótima da operação. Em geral, séries hidrológicas de intervalo de tempo menor que o ano, tais como séries mensais, possuem como principal característica o comportamento periódico das suas propriedades probabilísticas, como, por exemplo, a média, a variância, a assimetria e a estrutura de autocorrelação. O modelo GEVAZP baseia-se no modelo estocástico conhecido como autorregressivo periódico de ordem ‘p’ (HIPEL e MCLEOD, 1994; SALAS *et al.*, 1980). Esses modelos são referenciados por modelos PAR(p), onde p é a ordem do modelo, ou seja, o número de termos autorregressivos do modelo. As estatísticas básicas amostrais de uma série de vazão fluvial, que são utilizadas na modelagem PAR(p), são determinadas mensalmente conforme as Equações (3.1), (3.2), (3.3) abaixo:

$$\mu_m = \sum_{t=1}^N \frac{Z_t}{N} \quad (3.1)$$

$$\sigma_m = \sqrt{\sum_{t=1}^N \frac{(Z_t - \mu_m)^2}{N}} \quad (3.2)$$

$$\rho_m(k) = \frac{1}{N} \sum_{t=1}^N \left(\frac{Z_t - \mu_m}{\sigma_m} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \quad (3.3)$$

onde:

Z_t é a série temporal sazonal de período s ;

s é o número de períodos ($s=12$ para séries mensais);

N é o número de anos;

t é o índice do tempo, $t = 1, 2, \dots, sN$, função do ano T ($T = 1, 2, \dots, N$) e do período m ($m = 1, 2, \dots, s$);

μ_m é a média sazonal de período s ;

σ_m é o desvio-padrão sazonal de período s ;

$\rho_m(k)$ é a autocorrelação sazonal lag k dada pela correlação entre Z_t e Z_{t-k} de período s .

O modelo PAR(p) pode ser descrito matematicamente por (3.4):

$$\left(\frac{Z_t - \mu_m}{\sigma_m} \right) = \phi_m^1 \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_m^{p_m} \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + a_t \quad (3.4)$$

onde:

p_m é a ordem do modelo;

$\phi_m^{p_m}$ é o coeficiente autorregressivo de ordem p_m ;

a_t é a série de ruídos independentes com média zero e variância $\sigma_a^{2(m)}$.

Sendo assim, os cenários sintéticos futuros, gerados conforme o modelo PAR(p) na Equação (3.4), podem ser interpretados como uma combinação linear de observações passadas do histórico, acrescido de um termo aleatório. A geração de séries sintéticas pode ser, então, realizada através de sucessivos sorteios de a_t , sendo cada um desses cenários uma possível realização futura do processo estocástico. Com vista a reduzir o tempo computacional nos modelos de otimização energética, são aplicadas técnicas de agregação do vetor de ruídos a_t , de forma a diminuir sua cardinalidade, porém mantendo-se a qualidade da geração, conforme mostrado em JARDIM *et al.* (2001).

3.3.1 Identificação do Modelo

A primeira etapa para o ajuste de um modelo PAR(p) é a identificação da ordem p_m mais apropriada para os operadores autorregressivos de cada período. O critério de identificação da ordem do modelo através da função de autocorrelação parcial (FACP) foi sugerido, inicialmente, por BOX e JENKINS (1970) para a família de modelos estocásticos autorregressivos integrados de médias móveis (ARIMA). Essa mesma abordagem pode ser facilmente estendida para modelos da família PAR(p). Dada a Equação (3.4), multiplicando-se ambos os lados por $\left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}}\right)$ e tomando-se o valor esperado, obtém-se, para cada período, a seguinte expressão:

$$E \left[\left(\frac{Z_t - \mu_m}{\sigma_m} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] = \phi_m^1 E \left[\left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] + \dots + \phi_m^{p_m} E \left[\left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] + E \left[a_t \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] \quad (3.5)$$

Exemplificando, para $k = 1$, a expressão anterior reduz-se a

$$\rho_m(1) = \phi_m^1 + \phi_m^2 \rho_{m-1}(1) + \dots + \phi_m^{p_m} \rho_{m-1}(p_m - 1) \quad (3.6)$$

Fixando-se m e variando k de 1 a p_m na Equação (3.5), define-se, para cada período um conjunto de equações comumente denominado de equações de Yule-

Walker. Para um período m qualquer, tem-se:

$$\begin{bmatrix} 1 & \rho_{m-1}(1) & \rho_{m-1}(2) & \dots & \rho_{m-1}(p_m - 1) \\ \rho_{m-1}(1) & 1 & \rho_{m-2}(1) & \dots & \rho_{m-2}(p_m - 2) \\ \rho_{m-1}(2) & \rho_{m-2}(1) & 1 & \dots & \rho_{m-3}(p_m - 3) \\ \vdots & \dots & \vdots & \ddots & \vdots \\ \rho_{m-1}(p_m - 1) & \rho_{m-2}(p_m - 2) & \rho_{m-3}(p_m - 3) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_m^1 \\ \phi_m^2 \\ \phi_m^3 \\ \vdots \\ \phi_m^{p_m} \end{bmatrix} = \begin{bmatrix} \rho_m(1) \\ \rho_m(2) \\ \rho_m(3) \\ \vdots \\ \rho_m(p_m) \end{bmatrix} \quad (3.7)$$

Denomina-se ϕ_m^{kk} o último parâmetro deste processo auto-regressivo de ordem k . As equações de Yule-Walker para cada período m podem ser reescritas da seguinte forma:

$$\begin{bmatrix} 1 & \rho_{m-1}(1) & \rho_{m-1}(2) & \dots & \rho_{m-1}(k-1) \\ \rho_{m-1}(1) & 1 & \rho_{m-2}(1) & \dots & \rho_{m-2}(k-2) \\ \rho_{m-1}(2) & \rho_{m-2}(1) & 1 & \dots & \rho_{m-3}(k-3) \\ \vdots & \dots & \vdots & \ddots & \vdots \\ \rho_{m-1}(k-1) & \rho_{m-2}(k-2) & \rho_{m-3}(k-3) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_m^{k1} \\ \phi_m^{k2} \\ \phi_m^{k3} \\ \vdots \\ \phi_m^{kk} \end{bmatrix} = \begin{bmatrix} \rho_m(1) \\ \rho_m(2) \\ \rho_m(3) \\ \vdots \\ \rho_m(k) \end{bmatrix} \quad (3.8)$$

Cada um dos m sistemas resultantes pode ser resolvido através da decomposição de Cholesky. Denomina-se de função autocorrelação parcial do período m o conjunto de valores ϕ_m^{kk} . O conjunto de parâmetros ϕ_m^{kk} , $m = 1, 2, \dots, s$, é uma outra forma de representar a estrutura de dependência do processo estocástico ao longo do tempo, podendo ser compreendido como uma medida de correlação de *lag* k , retirando a influência dos lags menores. Em um processo auto-regressivo de ordem p_m , a função de autocorrelação parcial ϕ_m^{kk} será diferente de zero para k menor ou igual a p_m e zero para k maior que p_m .

A identificação da ordem p_m pela FACP é realizada comparando-se os valores de ϕ_m^{kk} a um determinado intervalo de confiança (95% por exemplo), podendo ser realizada de duas formas:

1. identifica-se o primeiro ϕ_m^{kk} significativo, fazendo a recursão da maior ordem possível até a menor;
2. identifica-se o último ϕ_m^{kk} significativo anterior a uma ordem não significativa.

Ambas as formas de identificação são corretas do ponto de vista teórico. Alguns pesquisadores defendem o uso da primeira, com a justificativa de aproveitar ao máximo as informações contidas na série histórica, que poderiam ser perdidas ao se utilizar o segundo critério. Os pesquisadores que sugerem o uso do segundo processo acreditam que, por ser mais parcimonioso, ele pode produzir resultados melhores, considerando espúrias as autocorrelações parciais significativas de ordem superior. O

modelo GEVAZP segue a primeira linha de identificação, e será esse o procedimento adotado nesta dissertação. Sendo o GEVAZP um gerador de cenários sintéticos de afluências, utilizado também para o planejamento de médio e longo prazo, espera-se que seja capaz de reproduzir os períodos secos multivariados observados ao longo do histórico. Sendo assim, é mais plausível supor que modelos com ordens maiores tenham uma maior aderência à autocorrelação anual histórica, com capacidade de mimetizar os períodos críticos do histórico.

3.3.2 Estimação do Modelo

Após a etapa de identificação da ordem p_m do modelo, deve-se obter as estimativas dos parâmetros da Equação (3.4). Para essa etapa, utilizam-se os estimadores obtidos pelo método dos momentos. As médias e os desvios padrões mensais são obtidos através das Equações (3.1) e (3.2) respectivamente. Os parâmetros ϕ_m^i , $i = 1, \dots, p_m$ podem ser estimados substituindo-se, na Equação (3.6), as estimativas de $\rho_{m-j}(k)$, $j = 0, \dots, (p_m - 1)$ obtidas na Equação (3.3).

Por fim, o último parâmetro a ser estimado é a variância $\sigma_a^{2(m)}$ dos ruídos independentes a_t . Dada a equação 3.5, para o lag $k = 0$, tem-se:

$$1 = \phi_m^1 \rho_m(1) + \phi_m^2 \rho_m(2) + \dots + \phi_m^{p_m} \rho_m(p_m) + E \left[a_t \left(\frac{Z_t - \mu_m}{\sigma_m} \right) \right] \quad (3.9)$$

Multiplicando-se ambos os lados da Equação (3.4) por a_t e tomando o valor esperado, obtém-se:

$$E \left[a_t \left(\frac{Z_t - \mu_m}{\sigma_m} \right) \right] = \sigma_a^{2(m)} \quad (3.10)$$

Substituindo (3.10) em (3.9), chega-se a:

$$\sigma_a^{2(m)} = 1 - \phi_m^1 \rho_m(1) - \phi_m^2 \rho_m(2) - \dots - \phi_m^{p_m} \rho_m(p_m) \quad (3.11)$$

sendo essa expressão válida para qualquer período m .

3.3.3 Correlações Espaciais

Uma das principais vantagens do SIN reside na complementariedade hídrica existente entre as regiões hidrográficas que o compõe. Nesse sentido, essa característica deve ser propriamente incorporada no modelo de geração de cenários sintéticos de forma que as séries sejam geradas de forma multivariada, levando em conta as correlações espaciais existentes entre as diversas UHEs e REEs. A princípio, o modelo PAR(p) gera séries de forma univariada. Porém pode-se realizar uma

transformação nos resíduos espacialmente não correlacionados b_t de forma que eles passem a considerar a correlação espacial, resultando em resíduos W_t :

$$W_t = Db_t \quad (3.12)$$

onde D é a matriz quadrada de dimensão igual ao número de UHEs (ou REEs) que correlaciona espacialmente os resíduos, usualmente chamada de matriz de carga. A matriz de carga D , por sua vez, é uma transformação obtida da matriz de correlações espaciais lag 0 (mensais ou anuais) U , entre cada uma das UHEs. Para obtê-la, pode-se, por exemplo, utilizar uma técnica conhecida como método de Jacobi (PRESS *et al.*, 1992) para realizar uma decomposição espectral, em que a matriz de correlações espaciais U (anual ou mensal) é decomposta em uma matriz diagonal contendo seus autovalores (A) e uma matriz simétrica de autovetores (V), conforme as Equações (3.13) e (3.14):

$$U = AVA' \quad (3.13)$$

$$D = A\sqrt{V}A' \quad (3.14)$$

3.3.4 Vazões Incrementais Negativas

Séries temporais como as vazões naturais mensais, usualmente, não possuem valores negativos. Durante o processo de geração de cenários sintéticos, deve-se impor as condições necessárias para obter-se um Z_t positivo:

$$Z_t = \mu_m + \phi_m^1 \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_m^{p_m} \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + a_t \sigma_m > 0 \quad (3.15)$$

ou,

$$a_t > - \left(\frac{\mu_m}{\sigma_m} \right) - \phi_m^1 \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) - \dots - \phi_m^{p_m} \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \quad (3.16)$$

Séries de vazões naturais e, conseqüentemente, de ENA, usualmente, possuem uma distribuição de probabilidades não-normal. Uma possível solução para o problema de geração de valores negativos, e para lidar com a não-normalidade dos resíduos da série temporal Z_t , seria a utilização da transformação BOX e COX (1964). Sendo um dos objetivos do modelo GEVAZP a geração de cenários sintéticos mensais de ENA para o modelo NEWAVE, é necessário observar o requisito de linearidade dos dados de entrada para esse modelo. Dessa forma, a utilização da transformação (não-linear) BOX e COX (1964) da série temporal Z_t deve ser descartada.

A solução adotada para tratar as fortes assimetrias dos resíduos a_t e impedir

a geração de cenários negativos é o ajuste de uma distribuição lognormal com três parâmetros (CHARBENEAU, 1978) aos resíduos mensais, onde o terceiro parâmetro Δ_t , denominado como deslocamento, é calculado de forma a se evitar a geração de cenários negativos, conforme apresentado na Equação 3.16. Isto posto, tem-se que, para cada mês, a variável ζ_t possui distribuição normal, com média μ_{ζ_t} , variância $\sigma_{\zeta_t}^2$.

$$\zeta_t = \ln(a_t - \Delta_t) \quad (3.17)$$

Os parâmetros da distribuição lognormal são estimados de forma a preservar os momentos dos resíduos a_t , segundo as equações

$$\mu_{\zeta_t} = \frac{1}{2} \ln \left(\frac{\sigma_a^2}{\theta^2 - \theta} \right) \quad (3.18)$$

$$\sigma_{\zeta_t}^2 = \ln(\theta) \quad (3.19)$$

$$\Delta_t = \frac{-\sigma_a}{\sqrt{(\theta - 1)}} \quad (3.20)$$

Igualando-se a Equação (3.16) ao deslocamento Δ_t , apresentado na equação 3.20, obtém-se:

$$\theta = 1 + \frac{\sigma_a^2}{(\mu_a - \Delta_t)^2} \quad (3.21)$$

3.3.5 Geração de Séries Sintéticas com o Modelo PAR(p)

Ajustados todos os parâmetros do modelo, é possível realizar a geração de séries sintéticas. Cada cenário gerado pode ser compreendido como uma possível realização do processo estocásticos modelado. Pode-se reescrever a Equação (3.4), tornando-a mais conveniente para a geração sequencial de cenários Z_t , obtidos como uma função dos valores passados, acrescido de um termo aleatório, conforme a Equação 3.22:

$$Z_t = \mu_m + \phi_m^1 \sigma_m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_m^{p_m} \sigma_m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + a_t \quad (3.22)$$

onde a série de ruídos E_t é independente com média zero e variância unitária.

Para dar início ao processo de geração de cenários, é necessário o conhecimento dos valores passados de $Z_{t-\tau}$, $\tau = 1, \dots, p_m$. Os valores passados $Z_{t-\tau}$ podem ser obtidos através de valores previamente observados da série, usualmente chamados de tendência hidrológica quando o processo estocástico é uma série hidrológica, dando início ao que se chama de geração condicionada ao passado recente. O processo de geração condicionada é ilustrado na Figura 3.3. Nesse tipo de geração, todos

os cenários gerados no instante $t = 1$ partem de um mesmo passado em comum.

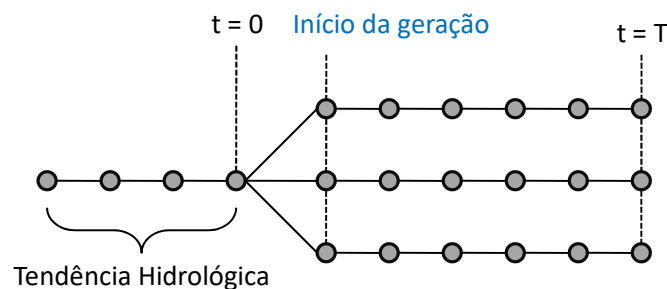


Figura 3.3: Geração de cenários sintéticos condicionados ao passado recente.

Os valores passados $Z_{t-\tau}$ podem, também, ser obtidos de forma a que cada cenário gerado no instante $t = 1$ venha de um passado distinto, caracterizando, assim, uma geração não-condicionada. A geração não condicionada é realizada em duas etapas. Na primeira etapa, arbitram-se os valores de $Z_{t-\tau}$ como sendo iguais a média mensal da série histórica, comumente chamada de MLT. A partir desses valores, geram-se séries sintéticas para um determinado horizonte, por exemplo, cinco anos. Após esse tempo, cada série sintética segue uma trajetória distinta, possuindo passados diversos. Os valores gerados para os últimos períodos da primeira etapa servirão como valores iniciais para a segunda etapa. O processo de geração não-condicionado é ilustrado na Figura 3.4.

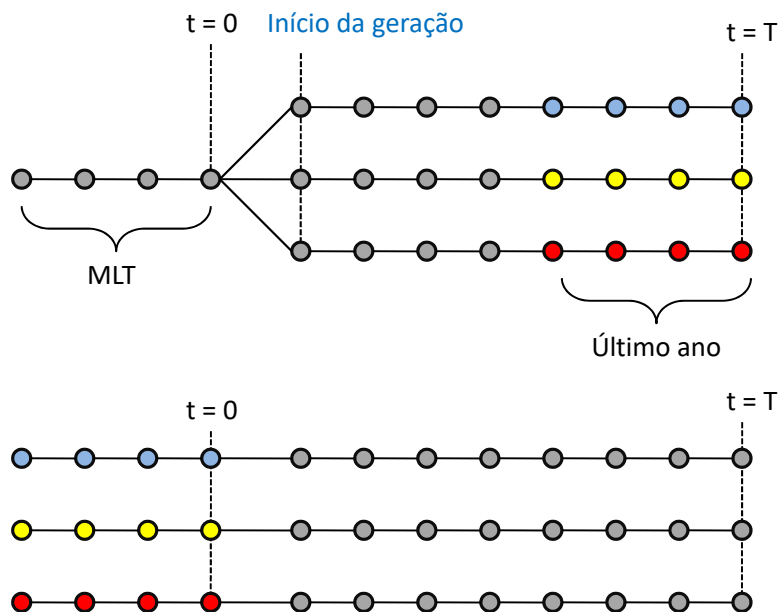


Figura 3.4: Geração de cenários sintéticos não-condicionados ao passado recente.

3.4 Modelo Autorregressivo Periódico com Chaveamento Markoviano - MS-PAR(p)

A nova modelagem concebida nesta dissertação visa aprimorar o modelo GEVAZP por meio da consideração de informações sobre o fenômeno climático ENOS. Para tanto é proposto um novo tipo de modelo estocástico: autorregressivo periódico com chaveamento Markoviano - MS-PAR(p) (*Markov-Switching Periodic Autoregressive*). Ele pode ser compreendido como uma combinação entre a modelagem PAR(p) em que seus parâmetros são chaveados segundo alguns estados que seguem uma cadeia de Markov (WILKS, 2011). O MS-PAR(p) é uma evolução natural da família de modelos iniciada em HAMILTON (1989). A ideia principal aqui apresentada, é através da identificação dos diferentes estados de ENOS no histórico, fazer uma estimação segmentada dos parâmetros utilizados na parte autorregressiva periódica do modelo. Além dos parâmetros serem estimados de acordo com o mês em questão, eles também irão variar de acordo com os estados de LN, N e EN. Deve-se observar que ao contrário dos modelos propostos por GELATI *et al.* (2011); HAMILTON (1989); KROLZIG (1997), o modelo MS-PAR(p), apresentado nesta dissertação, apesar de fazer parte da mesma família de modelos, não segue uma cadeia de Markov oculta para seus estados, sendo esse processo estocástico identificado entre os estados do ENOS. Mais detalhes sobre modelagem da cadeia de Markov, serão vistos, posteriormente, na sub-seção 3.4.1.

Assim sendo, o modelo MS-PAR(p) proposto nesta dissertação possui a seguinte formulação matemática:

$$\left(\frac{Z_t - \mu_m^{\varepsilon_t}}{\sigma_m^{\varepsilon_t}} \right) = \phi_m^1 \left(\frac{Z_{t-1} - \mu_{m-1}^{\varepsilon_{t-1}}}{\sigma_{m-1}^{\varepsilon_{t-1}}} \right) + \dots + \phi_m^{p_m} \left(\frac{Z_{t-p_m} - \mu_{m-p_m}^{\varepsilon_{t-p_m}}}{\sigma_{m-p_m}^{\varepsilon_{t-p_m}}} \right) + a_t \quad (3.23)$$

onde:

Z_t é a série temporal sazonal de período s ;

ε_t é a série temporal discreta sazonal de período s com espaço de estados $r = 1, 2, \dots, d$ ($d=3$ para a representação do ENOS), que segue uma cadeia de Markov homogênea;

s é o número de períodos ($s=12$ para séries mensais);

N é o número de anos;

t é o índice do tempo, $t = 1, 2, \dots, sN$, função do ano T ($T = 1, 2, \dots, N$) e do período m ($m = 1, 2, \dots, s$);

$\mu_m^{\varepsilon_t}$ é a média sazonal do período m para o estado de ε_t ;

$\sigma_m^{\varepsilon_t}$ é o desvio-padrão sazonal de período m para o estado de ε_t ;

p_m é a ordem do modelo para o período m ;

$\phi_m^{p_m}$ é o coeficiente autorregressivo de ordem p_m para o período m ;
 a_t é a série de ruídos independentes com média zero e variância $\sigma_a^{2(m)}$.

Sobre o modelo MS-PAR(p) descrito na Equação (3.23), deve-se notar que os parâmetros média e desvio padrão alternam de acordo com o período m e com o estado de ε_t , ao contrário do parâmetro autorregressivo $\phi_m^{p_m}$, que somente varia com o período m . Como esse modelo será utilizado para a geração de cenários sintéticos de vazões afluentes, optou-se por apresentar aqui uma formulação mais próxima da realidade do processo estocástico em questão. Enquanto a média e desvio padrão podem sofrer alterações de acordo com o estado do fenômeno ENOS, devido ao aumento ou decréscimo da precipitação, que se transformará futuramente em vazão fluvial, o parâmetro autorregressivo $\phi_m^{p_m}$ está relacionado à resposta da área de drenagem da UHE modelada oferece a eventos passados. Em outras palavras, o parâmetro $\phi_m^{p_m}$ por representar estatisticamente uma característica física da bacia de drenagem, não deve sofrer modificações por conta do fenômeno ENOS e, com isso, não é chaveado segundo o estado de ε_t . Aqui, define-se o espaço de estados de ε_t como $r = \{1, 2, 3\}$, representando, respectivamente, os estados LN, N e EN.

Para o cálculo da média e do desvio padrão amostrais, calculados para cada mês e estado, e da autocorrelação amostral, variável mensalmente apenas, apresentam-se as Equações (3.24), (3.25) e (3.26). É interessante observar que a formulação do modelo MS-PAR(p) traz alguns conceitos de estatística Bayesiana para a modelagem estatística clássica, utilizada nos modelos Box-Jenkins, como o modelo PAR(p). O cálculo da média e do desvio padrão é realizado com informações *a priori* sobre o espaço de estados.

$$\mu_m^r = E(Z_t | \varepsilon_t = r) = \sum_{t=1}^N \mathbf{1}_{(\varepsilon_t=r)} \frac{Z_t}{N_m^r} \quad r = 1, 2, 3 \quad (3.24)$$

$$\sigma_m^r = E(Z_t | \varepsilon_t = r) = \sqrt{\sum_{t=1}^N \mathbf{1}_{(\varepsilon_t=r)} \frac{(Z_t - \mu_m^r)^2}{N_m^r}} \quad r = 1, 2, 3 \quad (3.25)$$

$$\rho_m(k) = \frac{1}{N} \sum_{t=1}^N \left(\frac{Z_t - \mu_m^r}{\sigma_m^r} \right) \left(\frac{Z_{t-k} - \mu_{m-k}^r}{\sigma_{m-k}^r} \right) \quad (3.26)$$

onde $\mathbf{1}_{(\varepsilon_t=r)}$ denota a função indicadora, assumindo o valor unitário quando $\varepsilon_t = r$ e zero para qualquer outro valor e N_m^r é o número de ocorrências do estado r no mês m .

Definida a formulação matemática do modelo MS-PAR(p), o processo de identificação da ordem e estimação dos parâmetros, bem como o processo de transformações dos ruídos em lognormal três parâmetros correlacionados espacialmente, seguirão exatamente os mesmos passos do modelo apresentado na seção 3.3. Sendo

assim, o modelo proposto MS-PAR(p) é um aprimoramento metodológico do modelo GEVAZP. Com o advento de sua implementação, informações sobre o fenômeno climático ENOS serão incorporadas na geração de cenários sintéticos de vazões e energias mensais voltados para o SIN. Espera-se, assim, poder determinar com uma precisão maior as distribuições multivariadas de probabilidades para as afluências futuras. Esse novo modelo, altera a concepção da modelagem autorregressiva de olhar somente para os valores passados da série temporal para prever os passos futuros. Dado o estado r esperado para o mês m no instante t em que se deseja fazer a geração de cenários, um determinado conjunto de parâmetros é utilizado para gerar os cenários, capturando efetivamente a sua função de distribuição histórica, que é diferente dos outros estados.

A modelagem proposta possui algumas vantagens em relação aos modelos PARX que incluem informações climáticas. Primeiramente, nos modelos PARX, a informação climática somente altera o nível do valor esperado. Conforme será visto no próximo capítulo, estados de ENOS diferentes produzem variabilidades distintas para as vazões, alterando a função de distribuição de probabilidades das vazões. Sendo o modelo GEVAZP um modelo de geração de séries sintéticas de afluências, e não um modelo voltado a previsões, torna-se atrativo conseguir identificar funções de distribuição de probabilidades distintas. Da mesma forma que um modelo único do tipo AR consegue capturar menos informações sobre o processo estocástico de variáveis aleatórias que possuem alguma sazonalidade do que um modelo do tipo PAR(p), espera-se que um modelo MS-PAR(p) traga benefícios semelhantes. Identificando-se funções de distribuição de probabilidades distintas dentro de um mesmo período, o processo estocástico poderá ser melhor explorado. O segundo ponto, e o mais importante, diz respeito ao horizonte de geração possível. Como as variáveis exógenas são inseridas diretamente nos modelos PARX, muitas vezes através de previsões determinísticas, a geração de cenários fica restrita ao horizonte de previsão da variável exógena. No modelo MS-PAR(p) aqui proposto, são gerados cenários de afluências através da parte autorregressiva do modelo, em conjunto com cenários de estados de ENOS, por meio da cadeia de Markov. Essa alternativa flexibiliza o modelo, podendo este ser utilizado durante todo o horizonte desejado, constituindo-se em candidato natural para geração de cenários de curto prazo até o longo prazo. Outro ponto de destaque do modelo MS-PAR(p) está na manutenção da modelagem linear das vazões, sendo compatível a sua utilização com o modelo NEWAVE.

O desenvolvimento apresentado nesta dissertação ocorreu simultaneamente com o artigo de ALIAT e HAMDI (2018), os primeiros a introduzir a família de modelos MS-PARMA, uma formulação ainda mais generalizada com a inclusão de termos relacionados à médias móveis. Nesse artigo, também são discutidos aspec-

tos relevantes sobre a estacionariedade e a ergodicidade dos modelos MS-PARMA, essenciais para o correto funcionamento de um modelo linear estocástico, além de pequenos exemplos práticos que mostram o potencial do modelo. Assim mesmo, pode-se afirmar que o trabalho aqui desenvolvido possui um caráter inovador, ao se construir um modelo MS-PAR(p) aplicado a um caso real e de grande complexidade como a modelagem das afluências do SIN. Além disso, há algumas diferenças metodológicas importantes devido à especificidade do problema aqui analisado, onde os estados são observáveis no histórico (estados de ENOS), sendo modelados através de uma cadeia de Markov (não oculta).

3.4.1 Modelagem do Processo Markoviano

Sendo o objetivo final do modelo MS-PAR(p) proposto gerar cenários de afluências mensais com informações sobre o fenômeno ENOS para o horizonte de curto, médio e longo prazo, é necessário que a geração de cenários seja feita também para os eventos de ENOS. Os cenários gerados para o fenômeno ENOS devem possuir a mesma discretização temporal dos cenários de vazão, ou seja, em base mensal.

A série temporal de estados do fenômeno ENOS pode ser compreendida como uma variável discreta ε_t com espaço de estados r relacionados aos seus três estados possíveis (LN, N e EN). A classe de modelo mais comumente utilizada para representar a série temporal de uma variável discreta é conhecida como cadeia de Markov (WILKS, 2011). O comportamento de uma cadeia de Markov é conduzido por um conjunto de probabilidades de transição entre seus estados, \mathbf{P} , sendo a forma mais simples a cadeia de Markov de ordem um (ou um modelo auto-regressivo de ordem um). Ou seja, o estado seguinte depende apenas do estado mais recente, independente da sequência de estados anteriores, conforme apresentado na Equação (3.27):

$$P(\varepsilon_t = r \mid \varepsilon_{(t-1)}, \varepsilon_{(t-2)}, \dots, \varepsilon_1) = P(\varepsilon_t = r \mid \varepsilon_{(t-1)}) \quad (3.27)$$

As probabilidades de transição são probabilidades condicionadas com relação ao estado $t - 1$ mais recente. Como o fenômeno ENOS será dividido entre seus três possíveis estados, a matriz de transição \mathbf{P}_m será de dimensão 3x3.

$$\mathbf{P}_m = \begin{bmatrix} p_m^{1,1} & p_m^{1,2} & p_m^{1,3} \\ p_m^{2,1} & p_m^{2,2} & p_m^{2,3} \\ p_m^{3,1} & p_m^{3,2} & p_m^{3,3} \end{bmatrix} \quad (3.28)$$

Dado que não existe transição direta entre estados de EN para LN, assim como para as condições, os elementos $p_m^{1,3}$ e $p_m^{3,1}$ serão sempre nulos.

Para cada mês m será estimada a matriz de transição para os estados de

ENOS e para as condições de ENOS, ambas utilizando a seguinte expressão matemática:

$$p_m^{i,j} = P(\varepsilon(t) = j \mid \varepsilon(t-1) = i)_m = \frac{N_m^{i,j}}{N_{m-1}^i} \quad (3.29)$$

onde $p_m^{i,j}$ é a probabilidade condicionada a se transitar para o estado j no mês m dado que o mês anterior $m - 1$ encontrava-se no estado i ; $N_m^{i,j}$ é o número de vezes em que ocorreu a transição do estado i para o estado j no mês m no histórico; e N_{m-1}^i é o número de ocorrências no histórico do estado i no mês $m - 1$. A matriz de transição \mathbf{P}_m pode ser estimada segundo a Equação (3.28) variando-se os índices dos estados.

Inclusão de Previsões Probabilísticas do Fenômeno El Niño - Oscilação Sul

Mantendo-se em perspectiva a melhor utilização do modelo MS-PAR(p) proposto em estudos de planejamento da operação energética do SIN, apresenta-se uma metodologia de inclusão de previsões probabilísticas de ENOS, produzindo matrizes de transição previstas entre os estados. Em casos reais, até mesmo o estado de meses anteriores pode não ser conhecido, dado que a classificação do ENOS depende tanto dos índices ONI passados quanto futuros, sendo, assim, necessário estimar a probabilidade para os valores passados. Ilustrando se os últimos quatro índices foram enquadrados como condições de EN, caso o próximo mês também assim seja, serão efetivamente classificados como estados de EN; porém, caso o próximo índice ONI seja um valor menor que +0,5, esse período como um todo será classificado como neutro. Nesse contexto, a utilização de previsões probabilísticas apresenta-se vantajosa.

Conforme apresentado na seção 3.2, o IRI fornece previsões probabilísticas para as condições do índice ONI até nove trimestres à frente. A princípio, dadas as previsões disponibilizadas de condições de LN, N e EN, denominadas por $P_c(LN)_m^{prev}$, $P_c(N)_m^{prev}$ e $P_c(EN)_m^{prev}$, respectivamente, não se pode inferir sobre as probabilidades condicionadas de transição entre as condições de ENOS. Porém, pode-se calcular uma matriz de transição prevista mais próxima possível das matrizes históricas de transição entre as condições, respeitando a distribuição de probabilidade prevista pelo IRI. Para diferenciar os valores referentes às condições de ENOS, o sub-índice "c" será adicionado nas notações, evitando-se possíveis confusões entre os conceitos de condição e estado de ENOS. O cálculo das matrizes de transição previstas mensais entre as condições de ENOS pode ser compreendido como um problema de otimização cujo objetivo é minimizar as diferenças (Δ_i) entre a matriz histórica e a

prevista:

$$\min \sum_{i=1}^{14} \Delta_i \quad (3.30)$$

s.a.

$$\begin{aligned} P_c(LN | LN)_m^{prev} &= P_c(LN | LN)_m^{hist} + \Delta_1^+ - \Delta_2^- \\ P_c(LN | N)_m^{prev} &= P_c(LN | N)_m^{hist} + \Delta_3^+ - \Delta_4^- \\ P_c(N | LN)_m^{prev} &= P_c(N | LN)_m^{hist} + \Delta_5^+ - \Delta_6^- \\ P_c(N | N)_m^{prev} &= P_c(N | N)_m^{hist} + \Delta_7^+ - \Delta_8^- \\ P_c(N | EN)_m^{prev} &= P_c(N | EN)_m^{hist} + \Delta_9^+ - \Delta_{10}^- \\ P_c(EN | N)_m^{prev} &= P_c(EN | N)_m^{hist} + \Delta_{11}^+ - \Delta_{12}^- \\ P_c(EN | EN)_m^{prev} &= P_c(EN | EN)_m^{hist} + \Delta_{13}^+ - \Delta_{14}^- \end{aligned}$$

Restrições referentes ao somatório das probabilidades condicionadas

$$\begin{aligned} P_c(LN | LN)_m^{prev} + P_c(N | LN)_m^{prev} &= 1 \\ P_c(LN | N)_m^{prev} + P_c(N | N)_m^{prev} + P_c(EN | N)_m^{prev} &= 1 \\ P_c(N | EN)_m^{prev} + P_c(EN | EN)_m^{prev} &= 1 \end{aligned}$$

Restrições referentes aos limites das probabilidades condicionadas

$$\begin{aligned} 0 &\leq P_c(LN | LN)_m^{prev} \leq 1 \\ 0 &\leq P_c(LN | N)_m^{prev} \leq 1 \\ 0 &\leq P_c(N | LN)_m^{prev} \leq 1 \\ 0 &\leq P_c(N | N)_m^{prev} \leq 1 \\ 0 &\leq P_c(N | EN)_m^{prev} \leq 1 \\ 0 &\leq P_c(EN | N)_m^{prev} \leq 1 \\ 0 &\leq P_c(EN | EN)_m^{prev} \leq 1 \end{aligned}$$

Restrições referentes à reprodução das previsões disponibilizadas pelo IRI

$$\begin{aligned} P_c(LN)_m^{prev} &= P_c(LN)_{m-1}^{prev} \times P_c(LN | LN)_m^{prev} + P_c(N)_{m-1}^{prev} \times P_c(LN | N)_m^{prev} \\ P_c(N)_m^{prev} &= P_c(LN)_{m-1}^{prev} \times P_c(N | LN)_m^{prev} + P_c(N)_{m-1}^{prev} \times P_c(N | N)_m^{prev} + \\ &\quad P_c(EN)_{m-1}^{prev} \times P_c(N | EN)_m^{prev} \\ P_c(EN)_m^{prev} &= P_c(EN)_{m-1}^{prev} \times P_c(EN | EN)_m^{prev} + P_c(N)_{m-1}^{prev} \times P_c(EN | N)_m^{prev} \end{aligned}$$

O problema de otimização acima possui uma função objetivo linear, assim como

todas as suas restrições, caracterizando-se como um problema de programação linear. Esse problema pode ser solucionado através de um algoritmo simplex, por exemplo.

Solucionado o problema de otimização acima para todos os meses em que existam previsões de condição de ENOS, pode-se, então, estimar a matriz de transição prevista entre os estados ENOS. O cálculo da matriz de transição de estados de ENOS segue o mesmo critério utilizado para a classificação do histórico, conforme abaixo disposto.

1. Caso o estado anterior seja classificado como LN:
 - (a) a probabilidade de persistir como LN será igual a probabilidade de persistir na condição LN;
 - (b) a probabilidade de transitar ao estado N será igual a probabilidade de transitar para a condição N dado que estava em uma LN.
2. Caso o estado anterior seja classificado como N:
 - (a) a probabilidade de transitar a um estado LN será igual ao produto de se transitar à condição LN no mês m e sua persistência na condição LN por mais quatro meses;
 - (b) a probabilidade de transitar a um estado LN será igual ao produto de se transitar à condição LN no mês m e sua persistência na condição LN por mais quatro meses;
 - (c) a probabilidade de persistir como neutro é dada pelo complemento das duas opções anteriores.
3. Caso o estado anterior seja classificado como EN:
 - (a) a probabilidade de persistir como EN será igual a probabilidade de persistir na condição EN;
 - (b) a probabilidade de transitar ao estado N será igual a probabilidade de transitar para a condição N dado que estava em uma EN.

Matematicamente, pode-se resumir os critérios anteriores segundo as seguintes ex-

pressões:

$$\begin{aligned}
P(LN | LN)_m^{prev} &= P_c(LN | LN)_m^{prev} \\
P(N | LN)_m^{prev} &= P_c(N | LN)_m^{prev} \\
P(LN | N)_m^{prev} &= P_c(LN | N)_m^{prev} \times \prod_{i=1}^4 P_c(LN | LN)_{m+i}^{prev} \\
P(EN | N)_m^{prev} &= P_c(EN | N)_m^{prev} \times \prod_{i=1}^4 P_c(EN | EN)_{m+i}^{prev} \\
P(N | N)_m^{prev} &= 1 - P(EN | N)_m^{prev} - P(LN | N)_m^{prev} \\
P(EN | EN)_m^{prev} &= P_c(EN | EN)_m^{prev} \\
P(N | EN)_m^{prev} &= P_c(N | EN)_m^{prev}
\end{aligned} \tag{3.31}$$

Enquanto houver previsões disponíveis, o modelo calculará as matrizes \mathbf{P}_m^{prev} . Para o cálculo de $P(LN | N)_m^{prev}$ e $P(EN | N)_m^{prev}$, quando não existirem previsões para os termos dos meses $m + 1$, $m + 2$, $m + 3$ e $m + 4$, esses serão substituídos pelos valores históricos. Com essa implementação, pretende-se aproveitar ao máximo a previsão probabilística disponibilizada.

Geração de Cenários de ENOS

Estimadas as matrizes de transição mensais para os estados e condições do fenômeno ENOS, a sua geração de cenários pode ser realizada de maneira direta para todo o horizonte de estudo. Dada a matriz de transição de estados, histórica ou prevista, sorteia-se qual será o estado do período seguinte. Para manter coerência com o critério de classificação dos estados, caso um cenário transite de um estado N para LN ou EN, o mesmo permanecerá nesse estado por mais quatro períodos no mínimo, contabilizando-se assim uma persistência de no mínimo cinco períodos. Assim como ocorre na geração de cenários de afluências, a geração de cenários de ENOS poderá ser realizada de forma não condicionada, ou condicionada ao passado recente.

Geração Não Condicionada

O processo de geração não condicionada de cenários de ENOS segue o mesmo princípio utilizado na geração não condicionada de afluências, dividindo o processo em duas etapas. Primeiramente, sorteia-se, da distribuição de probabilidades histórica dos estados, os valores para o mês $m - 1$. Após esse passo, é feita a geração de cenários de ENOS utilizando as cadeias de Markov históricas para definir as probabilidades de transição ou persistência dos estados, por cinco anos. Finda essa etapa, os cenários gerados para os últimos 11 meses (valor máximo da ordem do

modelo MS-PAR(p)) são guardados, servindo de passado para a geração de cenários para a segunda etapa de geração.

Geração Condicionada

Para a geração condicionada dos estados, o modelo MS-PAR(p) aplicado a outros tipos de problemas poderia gerar cenários de estados de forma direta, partindo do último estado observado. Porém, a aplicação nesta dissertação visa a geração de cenários de estados de ENOS. Conforme descrito na seção 3.2, o estado de ε_{t-1} nunca será conhecido, dado que ele depende de um valor do ONI que leva em consideração a TSM do período t . Além disso, o próprio critério de classificação de ENOS é dependente dos valores futuros. De forma a estimar o estado de ENOS probabilisticamente em ε_{t-1} serão utilizadas as previsões fornecidas pelo IRI.

Para estimar o estado ε_{t-1} o seguinte procedimento é realizado:

1. Verifica-se recursivamente as condições de ENOS nos períodos $t - \tau$, $\tau=2,3,4,5$, para identificar uma tendência de formação de EN ou LN;
2. Contabiliza-se o número de períodos passados a partir de $\tau - 2$ (chamado de δ) em que condições de EN ou LN foram observadas seguidamente, depois da ocorrência de uma condição N;
3. Caso δ seja igual a zero, a condição em $t - 2$ foi classificada como neutra. Com isso:
 - (a) a probabilidade do estado ε_{t-1} ser EN será igual a probabilidade ao produto de se transitar à condição EN no período $t - 1$ e sua persistência na condição LN por mais quatro períodos, de t a $t + 3$;
 - (b) a probabilidade do estado ε_{t-1} ser LN será igual a probabilidade ao produto de se transitar à condição LN no período $t - 1$ e sua persistência na condição LN por mais quatro períodos, de t a $t + 3$;
 - (c) a probabilidade de persistir como neutro é dada pelo complemento das duas opções anteriores;
4. Caso δ seja maior que zero, as condições $t - \tau, \tau=2, \dots, \delta+1$ foram classificado como LN (ou EN). Assumindo-se:
 - (a) a probabilidade do estado ε_{t-1} ser LN (ou EN) será igual ao produto de sua persistência por mais $5-\delta$ períodos;
 - (b) a probabilidade do estado ε_{t-1} ser N é dada pelo complemento da opção anteriores;

Resumindo matematicamente o procedimento anterior, para a condição anterior N, tem-se:

$$\begin{aligned}
P(LN)_{t-1}^{prev} &= P_c(LN | N)_{t-1}^{prev} \times \prod_{i=0}^3 P_c(LN | LN)_{t+i}^{prev} \\
P(EN)_{t-1}^{prev} &= P_c(EN | N)_{t-1}^{prev} \times \prod_{i=0}^3 P_c(EN | EN)_{t+i}^{prev} \\
P(N)_{t-1}^{prev} &= 1 - P(LN)_{t-1}^{prev} - P(EN)_{t-1}^{prev}
\end{aligned} \tag{3.32}$$

Para casos em que a condição anterior seja LN ou EN, e dado δ , o número de condições de mesma classificação imediatamente anteriores ao período $t - 1$, tem-se:

$$\begin{aligned}
P(r)_{t-1}^{prev} &= \prod_{i=0}^{4-\delta} P_c(r | r)_{t+i}^{prev} \\
P(N)_{t-1}^{prev} &= 1 - P(r)_{t-1}^{prev}
\end{aligned} \tag{3.33}$$

sendo $r = 1$ para o estado LN e $r = 3$ para o estado EN.

Calculada a distribuição de probabilidades dos estados em ε_{t-1} , realiza-se um sorteio para definir o estado passado de cada série sintética. O procedimento para a geração de cenários dos períodos posteriores será feito através de sorteios aleatórios seguindo as matrizes de transição prevista \mathbf{P}_m^{prev} , enquanto houver dados previstos disponíveis, e, posteriormente, se necessário, as matrizes de transição histórica \mathbf{P}_m^{hist} .

3.4.2 Identificação da Influência do Fenômeno El Niño - Oscilação Sul nas Vazões

Com vistas a produzir um modelo mais parcimonioso, a primeira avaliação a ser realizada diz respeito à identificação da influência do fenômeno climático ENOS nas séries históricas de vazões afluentes mensais de cada UHE. Durante a revisão bibliográfica realizada no item 2.2.1, foram apresentadas diversos procedimentos para identificar a influência do ENOS nas vazões. Como esta dissertação visa o desenvolvimento de uma metodologia de geração de cenários sintéticos de afluições condicionados ao fenômeno ENOS, a identificação da influência do ENOS será verificada nas distribuições de frequências acumuladas obtidas através da segmentação do histórico de vazões entre seus diferentes estados. Para tanto, será realizado o teste não-paramétrico de aderência de Kolmogorov-Smirnov (KS) (PINTO *et al.*, 1976). Sob a hipótese nula da igualdade entre duas distribuições, verifica-se a máxima diferença absoluta $W = \text{Max}_x |S_1(x) - S_2(x)|$, onde $S_1(x)$ e $S_2(x)$ são as distribuições de frequências acumuladas. Para amostras de tamanho suficientemente grande, o

valor crítico do teste no nível de significância de 10% é aproximadamente dado por:

$$W_{critico} = 1,22 \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \quad (3.34)$$

Quando $W > W_{critico}$, a hipótese de igualdade entre as distribuições é rejeitada. O teste de aderência KS é sensível à diferenças na média e no desvio padrão entre as distribuições, o que o torna interessante na contexto da geração de cenários. O teste será realizado em cada mês para todas as combinações possíveis de estados, e em todas as UHEs, através do procedimento exemplificado nas Figuras 3.5 e 3.6.

Primeiramente, inicializa-se o teste KS entre a amostra da série de EN e a amostra dos outros dois estados em conjunto (LN e N). Caso a hipótese nula não seja rejeitada (as duas distribuições são iguais), testa-se a amostra da série de LN contra os outros dois estados em conjunto (EN e N); caso a primeira hipótese nula seja rejeitada (as duas distribuições são diferentes) testa-se a separação da amostra restante entre LN e N. Após essa primeira sequência de testes, o procedimento é novamente realizado, porém tendo, para o primeiro teste KS, a amostra da série LN contra a amostra dos outros dois estados em conjunto (EN e N). Caso o resultado do primeiro procedimento indique apenas a influência do EN e o segundo procedimento sinalize apenas a influência da LN, não se rejeita a hipótese que ambas as fases do ENOS são influentes.

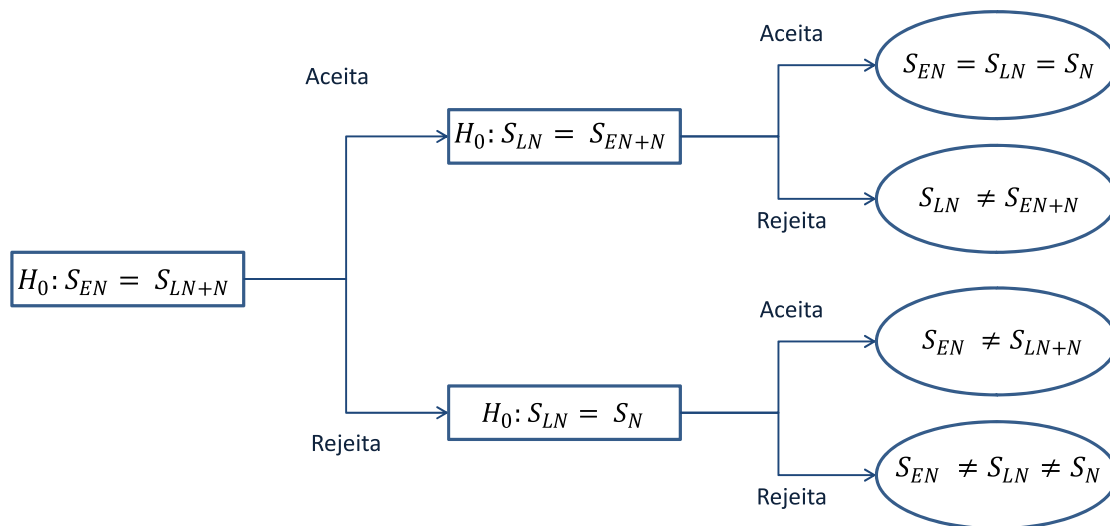


Figura 3.5: Procedimento de identificação de influência do ENOS nas vazões naturais afluentes às UHEs - parte I

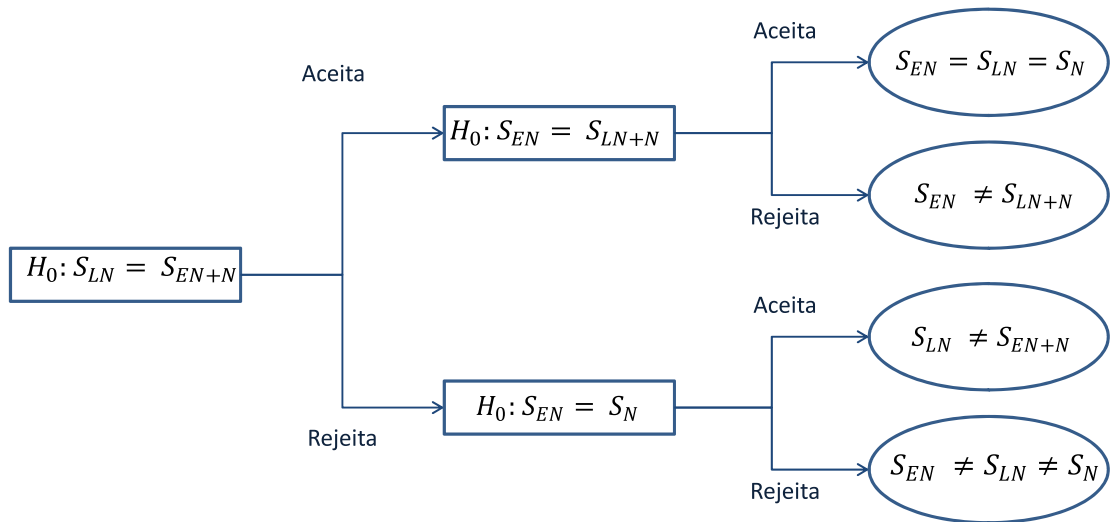


Figura 3.6: Procedimento de identificação de influência do ENOS nas vazões naturais afluentes às UHEs - parte II

Caso em determinado mês nenhuma das fases do ENOS seja relevante para a geração de cenários, o modelo MS-PAR(p), para esse mês, comportar-se-á como o PAR(p), possuindo a mesma média e desvio padrão em todos os estados de ENOS. Caso apenas uma das fases seja influente, as outras duas fases comportar-se-ão como uma. Por exemplo, se em janeiro identificou-se que o EN possui relevância sobre as vazões, porém a LN se comporta de forma parecida ao estado N, os dois últimos estados terão as mesmas estatísticas históricas. Acredita-se que essa identificação prévia da influência do ENOS nas vazões naturais afluentes produza modelos mais parcimoniosos do que o MS-PAR(p) sem esse processo, e, assim, traga benefícios para a geração de cenários. Ao longo desta dissertação, serão comparadas as três modelagens descritas neste capítulo: PAR(p), MS-PAR(p) sem identificação prévia da influência do ENOS (chamado de MS-PAR(p)1) e MS-PAR(p) com identificação prévia da influência do ENOS (chamado de MS-PAR(p)2).

3.5 Métricas de Avaliação

3.5.1 Critérios de avaliação para Casos Não-Condicionados

Um modelo de geração de séries sintéticas deve preservar as principais características da série observada. A adequação dos modelos utilizados nesta dissertação será avaliada através da reprodução das estatísticas históricas. Serão comparadas as estatísticas sintéticas de média, desvio padrão e correlações espaciais anuais com seus valores históricos. Para cada período simulado, serão testadas as distribuições de probabilidades sintéticas de cada UHE através do teste de aderência KS aplicado no nível de significância de 5%. Adicionalmente, avalia-se, para o modelo MS-PAR(p),

a reprodução das médias e desvios padrões condicionados pelos estados de ENOS. Ainda para o modelo MS-PAR(p), realiza-se uma análise dos cenários sintéticos de ENOS, comparando-os com as distribuições históricas de probabilidades, e com relação às matrizes de transição entre os estados.

No caso do planejamento da operação de médio/longo prazo, a representação de seus períodos críticos também são extremamente relevantes. Nesse contexto, faz-se necessário avaliar a reprodução das sequências negativas, definidas como um período em que as vazões afluentes estão ininterruptamente abaixo de um valor pré-determinado, por exemplo, a MLT. A Figura 3.7 ilustra o conceito de sequência negativa e suas variáveis associadas. A linha contínua representa afluências e a linha pontilhada um limite pré-estabelecido. Os intervalos $(t_1 - t_2)$ e $(t_3 - t_4)$ correspondem às sequências negativas, isto é, vazões abaixo da MLT.

Aqui, definem-se as variáveis aleatórias relacionadas às sequências negativas:

$$C = t_2 - t_1 \quad (3.35)$$

$$S = \sum_{i=t_1}^{t_2} (Z_i - \mu_i) \quad (3.36)$$

$$I = \frac{S}{C} \quad (3.37)$$

sendo C o comprimento de sequência correspondente ao comprimento dos intervalos $(t_2 - t_1)$, $(t_4 - t_3)$ na Figura 3.7; S é a soma de sequência correspondente à área abaixo do limite durante a sequência (representadas na Figura 3.7 como as áreas A1 e A2); e I a intensidade de sequência correspondente ao valor médio abaixo do limite, isto é, a soma de sequência dividida pelo respectivo comprimento de sequência.

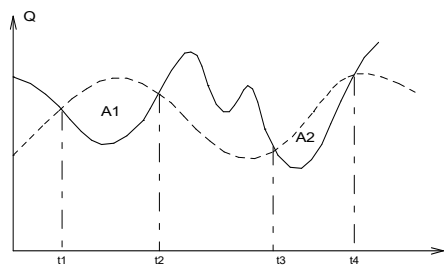


Figura 3.7: Exemplo de uma sequência negativa hipotética. A linha cheia corresponde a série temporal observada e a linha pontilhada ao limite pré-estabelecido.

O mesmo procedimento é realizado tanto para a série histórica quanto para as séries sintéticas geradas pelos três modelos. Para cada variável aleatória são comparados os valores de média, máximo, mínimo, e desvio padrão, além de se testar a hipótese nula de igualdade das distribuições de frequências das duas amostras. Seja n_1 o número de observações da variável aleatória X , obtida a partir da série histórica

e n_2 o número de observações da variável aleatória Y , obtida a partir da série gerada. Para o comprimento da sequência, que é uma variável discreta, utiliza-se o teste de igualdade de duas distribuições multinomiais (MOOD, 1950). O teste baseia-se na divisão do espaço amostral em $(k+1)$ subconjuntos, e a hipótese nula estabelece que:

$$H_0 : p_{1,j} = p_{2,j} \quad j = 1, \dots, k + 1 \quad (3.38)$$

onde $p_{1,j}$ é a probabilidade de que uma observação da primeira população pertença ao j -ésimo subconjunto, e $p_{2,j}$ é a correspondente probabilidade para a segunda população. Demonstra-se que:

$$\chi_k^2 = \sum_{i=1}^2 \sum_{j=1}^{k+1} \frac{\left(N_{ij} - n_i \frac{N_{1j} + N_{2j}}{n_1 + n_2} \right)^2}{n_i \frac{N_{1j} + N_{2j}}{n_1 + n_2}} \quad (3.39)$$

tem assintoticamente distribuição Qui-quadrado com k graus de liberdade, onde N_{ij} é o número de observações do j -ésimo subconjunto da i -ésima população.

Para variáveis aleatórias contínuas, como, por exemplo, a soma e a intensidade de sequência, o teste de aderência de Kolmogorov-Smirnov (PINTO *et al.*, 1976) é utilizado.

Os períodos secos multivariados são uma característica extremamente relevante no contexto do planejamento energético brasileiro. Assim, além da análise das variáveis aleatórias de forma univariada, será conduzida a avaliação de forma multivariada, observando a concomitância de sequências negativas entre diferentes UHEs, conforme apresentado em SANTARÉM *et al.* (2017).

Para a verificação dos cenários de ENOS, também pode-se analisar o comprimento de cada evento de ENOS nas séries sintéticas geradas. Será utilizado o teste multinomial, descrito anteriormente, para comparar a distribuição histórica de comprimentos de eventos de ENOS em seus três estados, com a distribuição sintética obtida nos cenários gerados. Dessas distribuições, também serão avaliados os comprimentos médios e máximos, além do desvio padrão das distribuições.

3.5.2 Critérios de avaliação para Casos Condicionados

Sendo os cenários sintéticos utilizados nos modelos de planejamento da operação, uma forma de comparação entre os modelos de geração de cenários seria através da comparação dos resultados da operação em tempo real. Porém, nesta dissertação, a avaliação dos casos condicionados será baseada em critérios comumente utilizados para a verificação do poder de previsão médio das séries sintéticas. O erro médio absoluto (EMA), erro médio percentual absoluto (EMPA) e a raiz do erro

quadrático médio (REQM) são apresentados, respectivamente, nas Equações (3.40), (3.41) e (3.42):

$$EMA = \frac{1}{n} \sum_{i=1}^n |Z_t - Cen_t^i| \quad (3.40)$$

$$EMPA = \frac{1}{n} \sum_{i=1}^n \left| \frac{Z_t - Cen_t^i}{Z_t} \right| \quad (3.41)$$

$$REQM = \frac{1}{n} \sum_{i=1}^n \sqrt{(Z_t - Cen_t^i)^2} \quad (3.42)$$

onde Cen_t^i é o i -ésimo cenário sintético simulado de um total de n cenários.

Um objetivo correlato desta dissertação é a avaliação da qualidade dos cenários gerados para os estados de ENOS, dado que eles possuem influência direta nas vazões sintéticas. Primeiramente, calcula-se o *ranked probability score* (RPS) (WILKS, 2011), que é a soma das diferenças quadráticas entre as previsões probabilísticas acumuladas com relação ao evento observado. A probabilidade acumulada para cada categoria é dada por:

$$P_i = \sum_{j=1}^i p_j \quad (3.43)$$

onde p_j é a previsão probabilística da ocorrência do j -ésimo estado. O RPS é calculado segundo a expressão

$$RPS = \sum_{i=1}^C (P_i - O_i)^2 \quad (3.44)$$

onde O_i é 1 quando o estado é observado e zero para os outros estados. Após o cálculo do RPS, é possível calcular o *ranked probability skill score* (RPSS), onde a confiabilidade das previsões probabilísticas geradas são comparadas com relação a utilização da frequência climatológica¹ em cada período. Quanto mais próximo de 1, melhor é a previsão. Define-se o RPSS através da Equação (3.45):

$$RPSS = 1 - \frac{RPS}{RPS_{clim}} \quad (3.45)$$

onde RPS_{clim} é o valor do RPS utilizando a frequência climatológica.

¹Frequência (ou probabilidade) climatológica corresponde à frequência de ocorrência de um fenômeno climático ao longo do histórico

Capítulo 4

Resultados e Discussões

Este capítulo apresentará os resultados obtidos ao longo dos estudos de caso apresentados nesta dissertação. Nesta seção, os nomes dos modelos PAR(p), MS-PAR(p)1 (sem identificação prévia da influência do ENOS) e MS-PAR(p)2 (com identificação prévia da influência do ENOS) serão abreviados como P, MSP1 e MSP2 nas tabelas e gráficos. Primeiramente, será abordada a influência do fenômeno ENOS nas séries de vazões naturais afluentes às UHEs através do procedimento proposto. Em seguida, serão realizadas as análises nos estudos de casos não condicionados, condicionados no período de ajuste do modelo e, por fim, condicionados no período pós ajuste do modelo.

4.1 Estudos de caso

Visando avaliar o aprimoramento proposto nesta dissertação para o modelo GEVAZP, através da adoção da modelagem estocástica MS-PAR(p), com consideração de informações sobre o fenômeno climático ENOS, será feita uma comparação da performance do modelo GEVAZP original com relação a nova modelagem proposta. A seguir, apresentam-se os casos a serem analisados.

4.1.1 Casos Analisados

A configuração hidráulica será a mesma para todos os casos analisados, sendo esta a configuração encontrada no PMO de janeiro de 2018, que conta com 141 UHEs. Foram excluídas da configuração UHEs que possuem vazão incremental nula. Para facilitar a avaliação, uma UHE de cada um dos doze REEs constantes na configuração atual do SIN foi escolhida para a apresentação dos resultados. A Tabela 4.1 apresenta as UHEs eleitas e a potência instalada de cada uma. A distribuição geográfica dessas UHEs é apresentada na Figura 4.1. Pode-se observar que grande parte das principais bacias hidrográficas brasileiras estão contempladas.

Para avaliar a adequação do modelo PAR(p) e dos modelos MS-PAR(p) com e sem identificação prévia da influência do ENOS, será realizado um caso não-condicionado. Nesse tipo de caso, busca-se, através da geração de cenários sintéticos, a reprodução das estatísticas históricas. Serão geradas 2000 séries sintéticas, com um horizonte de 67 anos (igual ao número de anos utilizados no ajuste dos modelos).

Já para a comparação do desempenho de ambos os modelos, serão realizados diversos casos condicionados ao passado recente. As rodadas condicionadas serão divididas em dois tipo:

1. geração de cenários um passo à frente para cada mês dentro do período de ajuste do modelo de 1951 a 2009;
2. geração de cenários para cinco anos à frente para cada mês entre os anos de 2010 e 2017.

No primeiro tipo de caso condicionado, para a geração de cenários do modelo MS-PAR(p), será considerado o conhecimento completo dos estados de ENOS, ou seja, previsões perfeitas de ENOS. Já para os casos condicionados do tipo 2, utilizam-se as previsões probabilísticas disponibilizadas pelo IRI. Em ambos os casos, serão gerados 2000 cenários sintéticos. A Tabela 4.2 traz um resumo dos casos que serão analisados nesta dissertação.

Tabela 4.1: Relação de UHEs representantes para a apresentação dos resultados

UHE	Bacia Hidrográfica	Potência Instalada (MW)
Itá	Uruguai	1.450
Salto Caxias	Iguaçu	1.240
Itaipu	Paraná	14.000
Capivara	Paranapanema	619
Porto Primavera	Paraná	1.540
Furnas	Grande	1.312
Sobradinho	São Francisco	1.050
Santo Antônio Jari	Amazonas	390
Estreito Tocantins	Tocantins	1.087
Belo Monte	Xingu	6.111
Teles Pires	Teles Pires	1.820
Jirau	Madeira	3.750

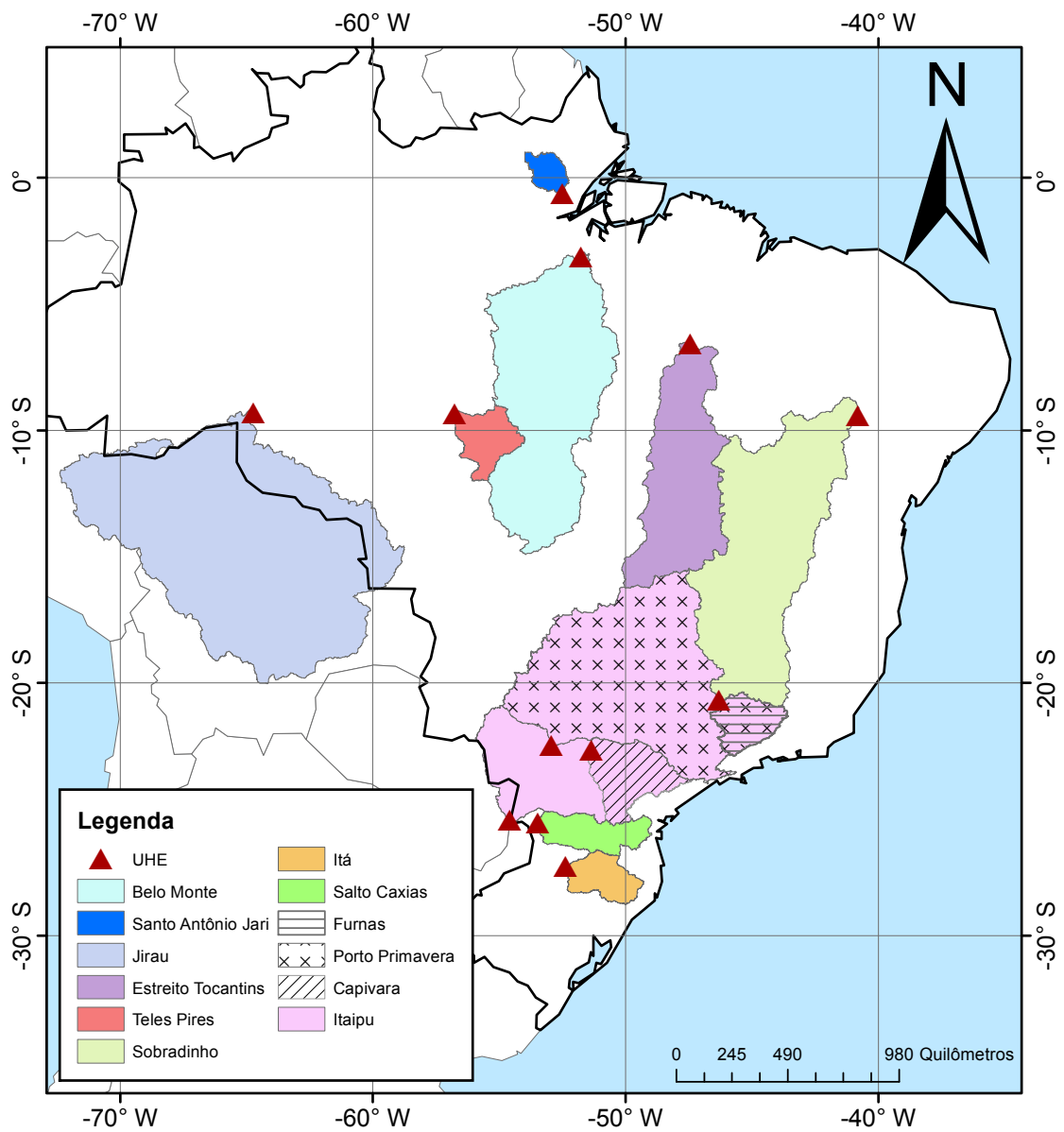


Figura 4.1: Distribuição geográfica das UHEs selecionadas para análise.

Tabela 4.2: Resumo dos estudos de caso realizados para a avaliação dos modelos

Tipo de caso	Quantidade	Nº de Períodos	Período de Teste	Tipo de Avaliação	Métricas de Avaliação
Não Condicionado	1	67 x 12	-	Adequação do modelo	- Momentos gerados - Correlação espacial - Sequências negativas - EMA
Condicionado	59 x 12	1	1951 a 2009	Performance no período de ajuste	- EMPA - REQM - EMA
Condicionado	8 x 12	5 x 12	2010 a 2017	Performance no período pós-ajuste	- EMPA - REQM - RPSS

4.2 Influência do Fenômeno El Niño - Oscilação Sul nas Séries de Vazões Naturais

Os primeiros resultados a serem apresentados fazem parte de um diagnóstico sobre a influência do fenômeno ENOS nas séries de vazões naturais às UHEs. Utilizou-se o teste de aderência KS, descrito no item 3.5, para o período entre 1950 e 2016. Para facilitar a visualização dos resultados, optou-se por dividir as UHEs entre as bacias hidrográficas as quais pertencem e, posteriormente, apresentá-los separadamente nos quatro submercados do SIN (sul, sudeste/centro-oeste, nordeste, norte). A Figura 4.2 apresenta o resultado do teste de aderência KS realizado para todos os meses em UHEs localizadas na região sul do Brasil. As células em vermelho indicam a influência do EN, azul da LN, lilás de ambos os estados, e branca sem influência. Observa-se um padrão bem definido de identificação dos efeitos do EN entre novembro e março para a maioria das UHEs. Destaca-se que, para a bacia do rio Iguaçu, que se encontra na parte mais ao norte da região sul, em geral, apenas em novembro e em dezembro o EN é identificado. Nas regiões hidrográficas em latitudes mais elevadas, ambas as fases do ENOS produzem alterações em novembro e dezembro, e somente o EN entre janeiro a março. Em algumas UHEs das outras bacias hidrográficas, é identificada, também, a influência da LN entre os meses de maio e julho. Destaca-se que a maior parte dos meses que sofrem influência do ENOS (de novembro a março) são meses em que as vazões médias na região sul se encontram mais baixas.

UHE	REE	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
FUNDÃO	IGUAÇU												
G.B. MUNHOZ	IGUAÇU												
JORDÃO	IGUAÇU												
MAUÁ	IGUAÇU												
SALTO CAXIAS	IGUAÇU												
SALTO OSÓRIO	IGUAÇU												
SEGREDO	IGUAÇU												
SLT.SANTIAGO	IGUAÇU												
STA CLARA PR	IGUAÇU												
ITAIPU	ITAIPU												
14 DE JULHO	SUL												
CASTRO ALVES	SUL												
D. FRANCISCA	SUL												
ERNESTINA	SUL												
G.P. SOUZA	SUL												
ITAUBA	SUL												
JACUÍ	SUL												
MONTE CLARO	SUL												
PASSO REAL	SUL												
QUEBRA QUEIXO	SUL												
SALTO PILÃO	SUL												
BARRA GRANDE	URUGUAI												
CAMPOS NOVOS	URUGUAI												
FOZ CHAPECÓ	URUGUAI												
GARIBALDI	URUGUAI												
ITA	URUGUAI												
MACHADINHO	URUGUAI												
MONJOLINHO	URUGUAI												
PASSO FUNDO	URUGUAI												
PASSO S JOÃO	URUGUAI												
SÃO JOSÉ	URUGUAI												

Figura 4.2: Teste KS para a região sul do Brasil. Células em branco, vermelho, azul e lilás indicam a influência de nenhum dos estados, EN, LN e N, respectivamente.

Apresentam-se, em seguida, as estatísticas históricas de média e desvio padrão

calculadas segundo as Equações (3.24) e (3.25), e o coeficiente de variação ($CV_m^r = \frac{\sigma_m^r}{\mu_m^r}$), para cada um dos estados de ENOS e para o histórico completo, em algumas das UHEs selecionadas na região sul. Os resultados, a seguir, foram obtidos sem levar em consideração o teste KS, ou seja, calculam-se as estatísticas para todos os estados de ENOS em todos os meses.

A Figura 4.3 apresenta os resultados obtidos para a UHE Itá, localizada no rio Uruguai. A região sul do Brasil é caracterizada por não ter uma sazonalidade bem definida de suas vazões fluviais, conforme pode ser observado. Em relação às médias, observa-se que o fenômeno EN produz um aumento da vazão afluyente em relação aos outros estados e ao histórico completo, e o inverso ocorre no estado de LN (quando os estados são influentes). É interessante ressaltar que, em geral, o aumento da média é também acompanhado pelo acréscimo do desvio padrão, e o contrário ocorre nos períodos de LN. Ou seja, há uma variabilidade maior associada quando se encontra em um EN, porém menor em LN ou N. Maio é único mês que foge ao padrão descrito, tendo a LN uma média maior, porém um desvio padrão menor. Para a UHE Itá o teste KS indica a influência de ambas as fases do ENOS em novembro e dezembro, do EN em janeiro e março e da LN em maio e outubro.

A Figura 4.4 traz os resultados obtidos para a UHE Itaipu, também localizada na região sul do Brasil, no rio Paraná. Apesar de se localizar na região sul, a UHE Itaipu possui uma extensa bacia de contribuição incremental que também abrange parte da região sudeste. Para a UHE Itaipu, identificou-se, apenas em abril, maio e dezembro, a influência do EN nas suas vazões afluentes. Assim como ocorre em grande parte dos meses na UHE Itá, o EN causa um aumento da vazão média, assim como do desvio padrão, o que representa uma incerteza maior nesse período. No caso da UHE Itaipu, em maio há um acréscimo de 1000 m³/s na vazão média, o que representa 50% a mais quando comparado com o histórico sem classificação. O desvio padrão, que representa a incerteza do processo estocástico de afluências, também cresce cerca de 60%. Nesse mesmo mês, para os estados N e LN, que são considerados como um único estado para a UHE Itaipu dado o resultado do procedimento de identificação da influência, ocorre um efeito contrário, com vazões médias menores e um desvio padrão também inferior.

A UHE Salto Caxias, última usina da região sul a ser analisada, que se encontra na bacia do rio Iguaçu, é apresentada na Figura 4.5. Conforme a análise anterior, o teste KS apresenta apenas o estado de EN influente em novembro e março. Assim como ocorre nas UHEs destacadas anteriormente, o EN provoca um aumento da média e do desvio padrão em relação ao histórico completo, enquanto os outros dois estados acarretam uma redução das estatísticas, principalmente entre os meses de novembro e maio.

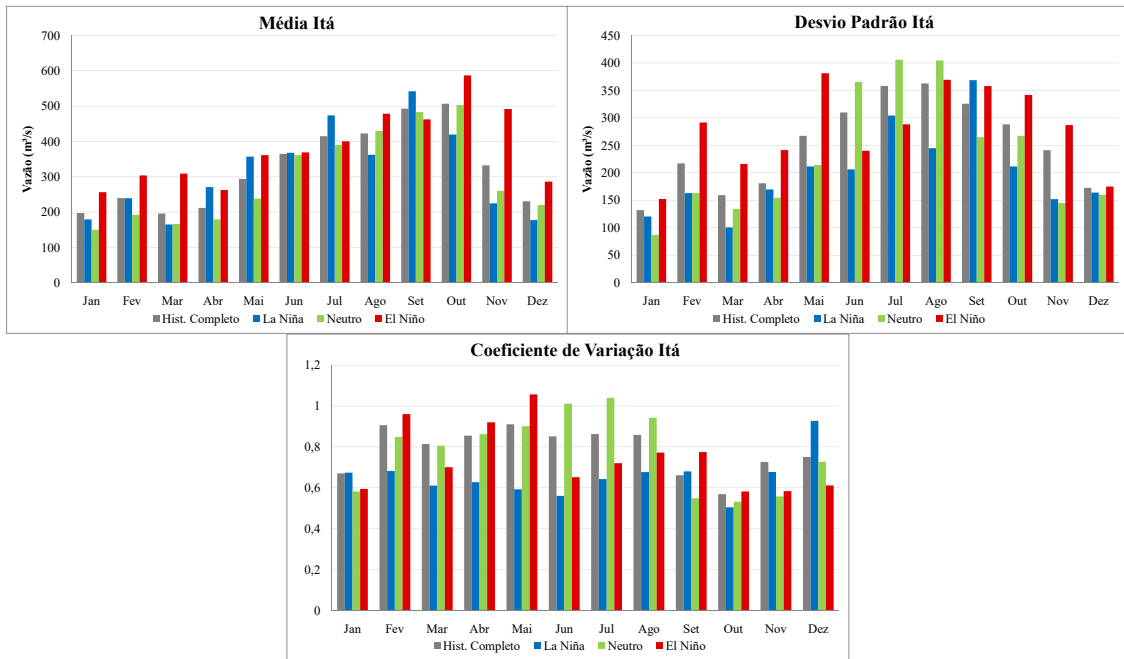


Figura 4.3: Estatísticas históricas da UHE Itá segundo o estado ENOS. As estatísticas obtidas para o histórico completo, e segmentado para o estado LN, N e EN, são representadas pelas barras de cor cinza, azul, verde e vermelho, respectivamente. Dados correspondentes ao período entre 1950 e 2017.

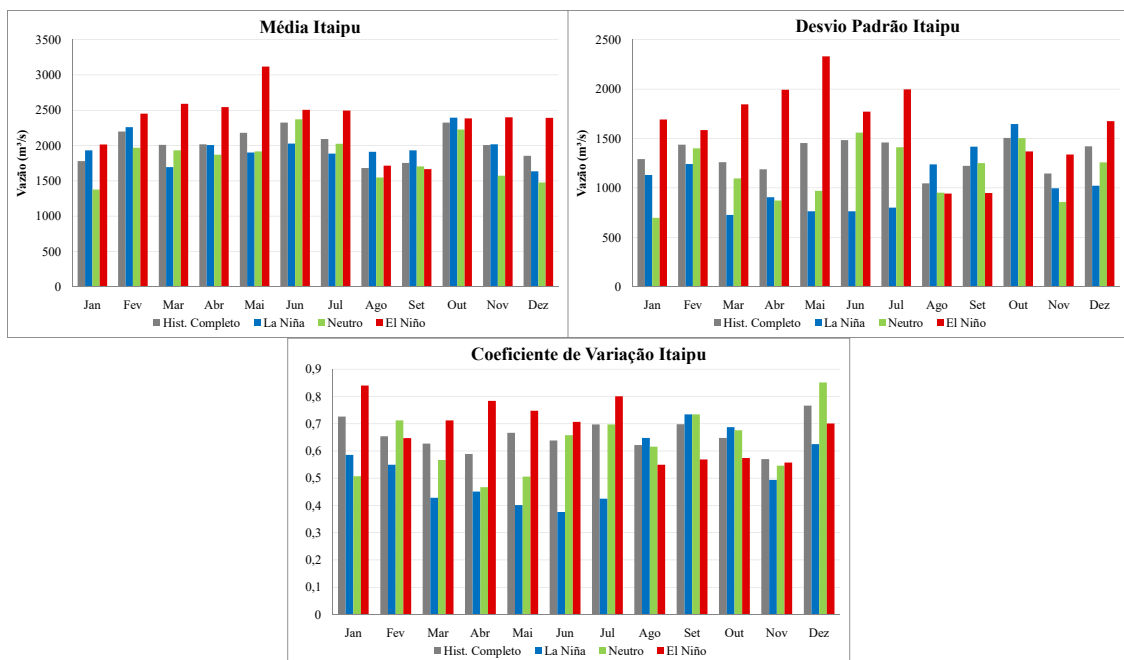


Figura 4.4: Estatísticas históricas da UHE Itaipu segundo o estado ENOS. Dados correspondentes ao período entre 1950 e 2017.

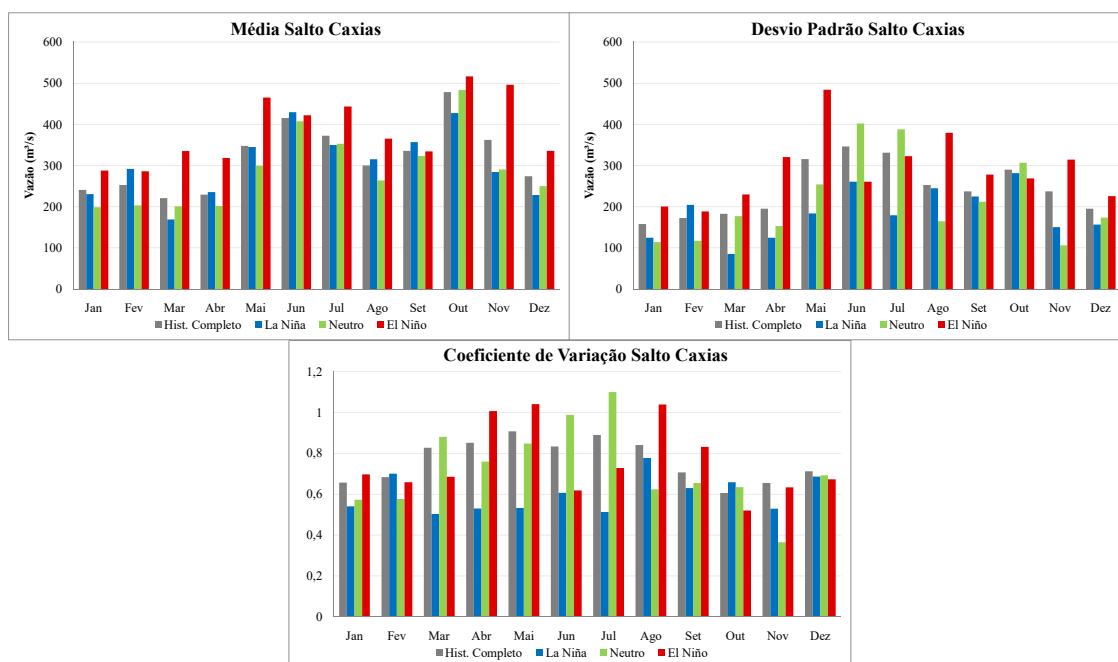


Figura 4.5: Estatísticas históricas da UHE Salto Caxias segundo o estado ENOS. Dados correspondentes ao período entre 1950 e 2017.

Esses padrões são observados na maioria das UHEs localizadas na região sul do Brasil, e vão ao encontro da revisão bibliográfica realizada no item 2.2.1. Quanto mais ao sul, maior é a influência do ENOS, sendo o estado EN responsável, muitas vezes, por um aumento na vazão afluyente, e a LN por uma diminuição. Porém, observa-se, também, que não há uma relação direta entre o fenômeno ENOS e as vazões, o que explica a maior incerteza nos eventos de EN, representada pelo crescimento do desvio padrão. Existe uma tendência de ocorrência de mais eventos de cheias durante esse período, porém esse fato não é verificado sempre. Todavia, durante a LN, uma vazão média abaixo da MLT é mais provável.

A seguir, serão investigadas UHEs localizadas em bacias hidrográficas da região sudeste e centro-oeste. Primeiramente, apresentam-se os resultados obtidos para o teste KS nessas UHEs. Para a bacia hidrográfica do rio Paranapanema, que se localiza na porção mais ao sul no estrato das bacias analisadas na Figura 4.6, o teste indica, de forma geral, uma influência marcante do EN, principalmente em novembro e em dezembro, e entre maio e agosto. As vazões afluentes de suas UHEs, possuem um comportamento de transição entre as regiões sul e sudeste, evidenciado pela similaridade de seu resultado no teste KS comparado ao de UHEs na região sul. Os resultados obtidos para a bacia do rio Tietê não apresentam um padrão bem definido, sendo identificado, entre os meses de abril a julho, em algumas de suas UHEs, a influência de um dos estados de ENOS. Para a bacia do rio Grande e do Atlântico sudeste, que englobam as bacias hidrográficas costeiras, em grande parte das UHEs, é identificado um padrão consistente de influência do EN entre

maio e julho (em alguns casos, a LN também é relevante) e da LN entre agosto e setembro. Nas UHEs localizadas a montante da bacia hidrográfica do rio Paranaíba, é identificada a distinção da LN entre janeiro e setembro.

UHE	BACIA	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
AIMORES	ATL. SUDEST												
BAGUARÍ	ATL. SUDEST												
BILLINGS	ATL. SUDEST												
CANDONGA	ATL. SUDEST												
FUNIL	ATL. SUDEST												
GUILMAN-AMOR	ATL. SUDEST												
ILHA DOS POMBOS	ATL. SUDEST												
LAJES	ATL. SUDEST												
MASCARENHAS	ATL. SUDEST												
P. ESTRELA	ATL. SUDEST												
PARAIBUNA	ATL. SUDEST												
PICADA	ATL. SUDEST												
ROSAL	ATL. SUDEST												
S. A. CARVALHO	ATL. SUDEST												
SALTO GRANDE	ATL. SUDEST												
SANTA BRANCA	ATL. SUDEST												
SANTANA	ATL. SUDEST												
SIMPLÍCIO	ATL. SUDEST												
SOBRAGÍ	ATL. SUDEST												
STA. CECÍLIA	ATL. SUDEST												
TOCOS	ATL. SUDEST												
A. VERMELHA	GRANDE												
A. S. OLIVEIRA	GRANDE												
CACONDE	GRANDE												
CAMARGOS	GRANDE												
E. DACUNHA	GRANDE												
ESTREITO	GRANDE												
FUNIL-GRANDE	GRANDE												
FURNAS	GRANDE												
IGARAPAVA	GRANDE												
JAGUARA	GRANDE												
JAGUARÍ	GRANDE												
M. DE MORAES	GRANDE												
MARIMBONDO	GRANDE												
P. COLOMBIA	GRANDE												
VOLTA GRANDE	GRANDE												
ITIQUEIRA 1	PARAGUAI												
JAUÚ	PARAGUAI												
MANSO	PARAGUAI												
PONTE PEDRA	PARAGUAI												
I. SOLT. EQV	PARANÁ												
JUPIÁ	PARANÁ												
P. PRIMAVERA	PARANÁ												
B. COQUEIROS	PARANAÍBA												
BATALHA	PARANAÍBA												
CACH. DOURADA	PARANAÍBA												
CAÇU	PARANAÍBA												
CAPIM BRANCO 1	PARANAÍBA												
CAPIM BRANCO2	PARANAÍBA												
CORUMBÁ I	PARANAÍBA												
CORUMBÁ III	PARANAÍBA												
CORUMBÁ IV	PARANAÍBA												
EMBORCAÇÃO	PARANAÍBA												
ESPORA	PARANAÍBA												
FOZ R. CLARO	PARANAÍBA												
ITUMBIARA	PARANAÍBA												
MIRANDA	PARANAÍBA												
NOVA PONTE	PARANAÍBA												
SALTO	PARANAÍBA												
SÃO SIMÃO	PARANAÍBA												
SERRA FACÃO	PARANAÍBA												
SLT. VERDINHO	PARANAÍBA												
A. A. LAYDNER	PARANAPANEMA												
CANOAS I	PARANAPANEMA												
CANOAS II	PARANAPANEMA												
CAPIVARA	PARANAPANEMA												
CHAVANTES	PARANAPANEMA												
L. N. GARCEZ	PARANAPANEMA												
OURINHOS	PARANAPANEMA												
PIRAJU	PARANAPANEMA												
ROSANA	PARANAPANEMA												
TAQUARUCU	PARANAPANEMA												
A. S. LIMA	TIETÊ												
BARRA BONITA	TIETÊ												
EDGARD SOUZA	TIETÊ												
GUARAPIRANGA	TIETÊ												
IBITINGA	TIETÊ												
N. AVANHANDAVA	TIETÊ												
PROMISSÃO	TIETÊ												

Figura 4.6: Teste KS para a região sudeste e centro-oeste do Brasil.

A Figura 4.7 apresenta as estatísticas para a UHE Capivara, localizada no rio Paranapanema. É possível observar que essa UHE não possui uma sazonalidade definida e que, assim como em Itaipu e Itá, suas vazões, de forma geral, possuem médias maiores que a MLT durante eventos de EN. Esse efeito, assim como nas outras usinas analisadas, vem acompanhado de uma maior variabilidade também. Destaca-se que, segundo o critério do teste KS, o estado EN possui uma influência em março, maio, julho, novembro e dezembro, enquanto a LN entre junho e agosto e em novembro.

São apresentados, na Figura 4.8, os resultados para a UHE Furnas, que pertence a bacia hidrográfica do rio Grande em Minas Gerais. A UHE Furnas é notadamente uma das UHEs mais importantes no SIN devido ao seu potencial de regularização, e por se localizar na região a montante de toda a bacia hidrográfica do rio Paraná. É possível verificar que, entre janeiro e julho, excetuando-se março, as afluições médias no período de EN são maiores que nas outras fases. O desvio padrão de janeiro a junho é maior para a fase quente do ENOS, enquanto que, em setembro e outubro, são apresentados valores maiores em períodos de LN, com destaque para o mês de setembro, quando o coeficiente de variação passa a ser maior que o valor unitário nesse estado. Segundo o teste KS, a LN é influente em fevereiro e entre maio e setembro (maior parte do período seco nessa região) e o estado EN responsável por variações entre maio e julho. Esse padrão pode ser estendido para grande parte das UHEs na bacia do Grande.

A UHE Porto Primavera, localizada no rio Paraná, tem suas estatísticas apresentadas na Figura 4.9. Pode-se observar que, com exceção de junho e dezembro, não há uma variação perceptível nas afluições devido aos diferentes estados de ENOS, como ocorre para a UHE Furnas. O procedimento de identificação de influência do ENOS indica apenas o estado EN como significativo para o mês de dezembro, demonstrando a pouca influência sofrida pela UHE Porto Primavera. Como ela se encontra na região mais a jusante do rio Paraná, ela possui uma extensa área de drenagem e contribuições de diversos outros rios e regiões, o que pode explicar a falta de uma influência mais forte do ENOS.

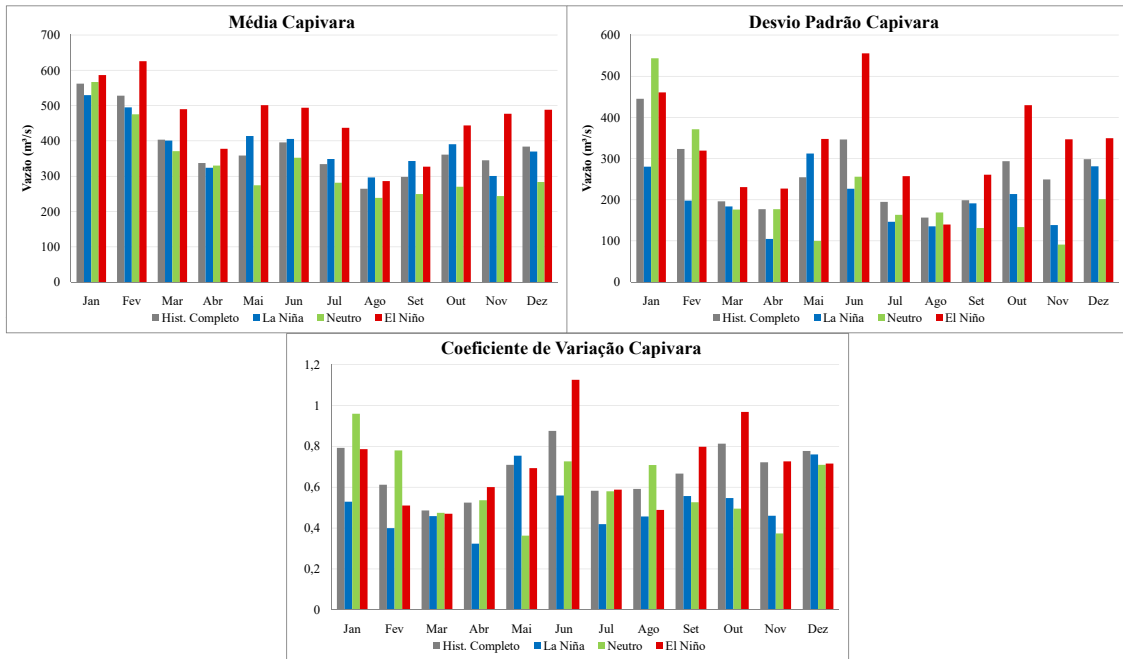


Figura 4.7: Estatísticas históricas da UHE Capivara segundo o estado ENOS.

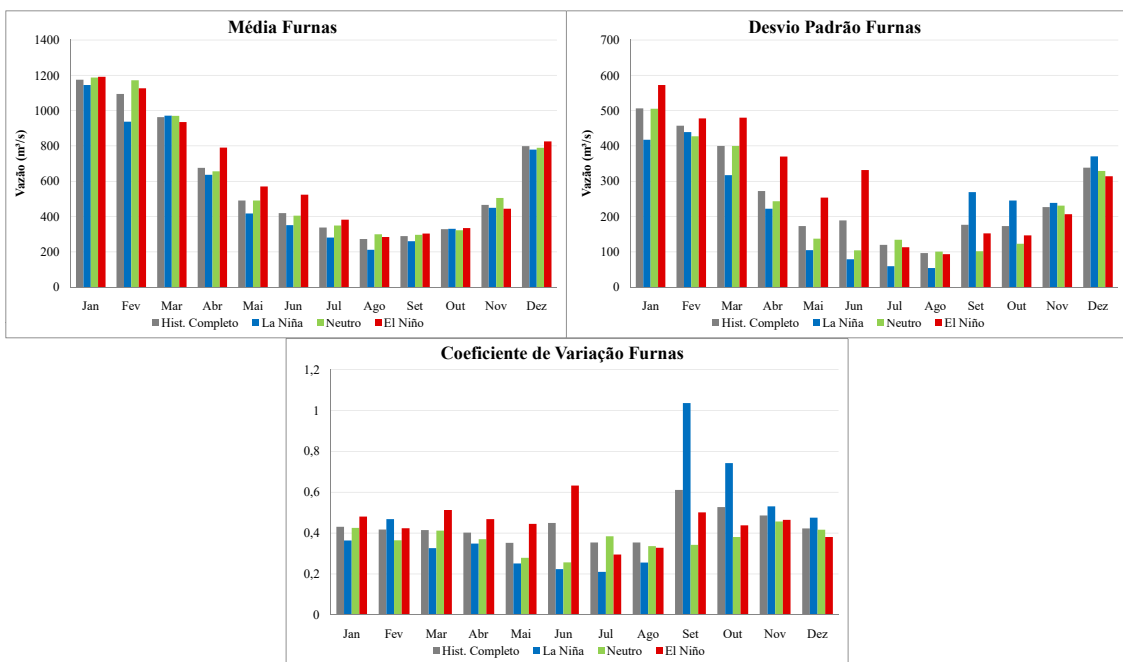


Figura 4.8: Estatísticas históricas da UHE Furnas segundo o estado ENOS.

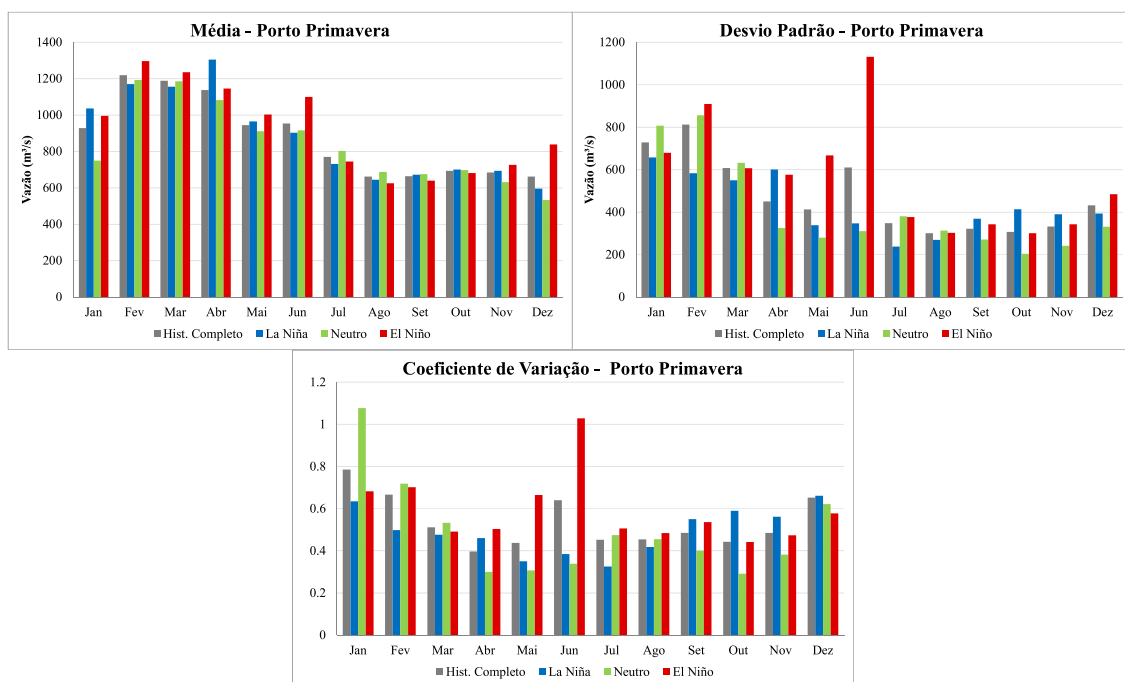


Figura 4.9: Estatísticas históricas da UHE Porto Primavera segundo o estado ENOS.

O resultado do teste KS para as bacias hidrográficas pertencentes à região nordeste é apresentado na Figura 4.10. Observa-se que, em sua maioria, é identificada a LN como uma influência importante em janeiro e fevereiro e em agosto e setembro. As UHEs localizadas nas bacias do Atlântico nordeste e leste apresentam alguma influência em novembro e dezembro.

A Figura 4.11 traz os resultados para a UHE Sobradinho, localizada na região nordeste do Brasil, pertencente à bacia do rio São Francisco. Nota-se que as vazões médias são menores em quase todos os meses do ano durante eventos de LN, excetuando-se os meses de novembro e dezembro. Em fevereiro e março, meses com as maiores afluências, é possível observar que o efeito da LN é mais intenso e possui uma variabilidade menor, principalmente em março. Essa inversão no padrão de influência observada nas fases do ENOS, entre os meses de novembro e fevereiro, também foi destacada na análise das precipitações nessa região em GRIMM (2003) e em GRIMM (2004). A LN foi identificada como influente em fevereiro e março e entre julho e setembro.

UHE	BACIA	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
IRAPÉ	ATL LESTE												
PEDRA DO CAVALO	ATL LESTE												
ITAPEBI	ATL LESTE												
SANTA CLARA MG	ATL LESTE												
BOA ESPERANÇA	PARNAÍBA												
QUEIMADO	SÃO FRANCISCO												
RETIRO BAIXO	SÃO FRANCISCO												
SOBRADINHO	SÃO FRANCISCO												
TRÊS MARIAS	SÃO FRANCISCO												

Figura 4.10: Teste KS para a região nordeste do Brasil.

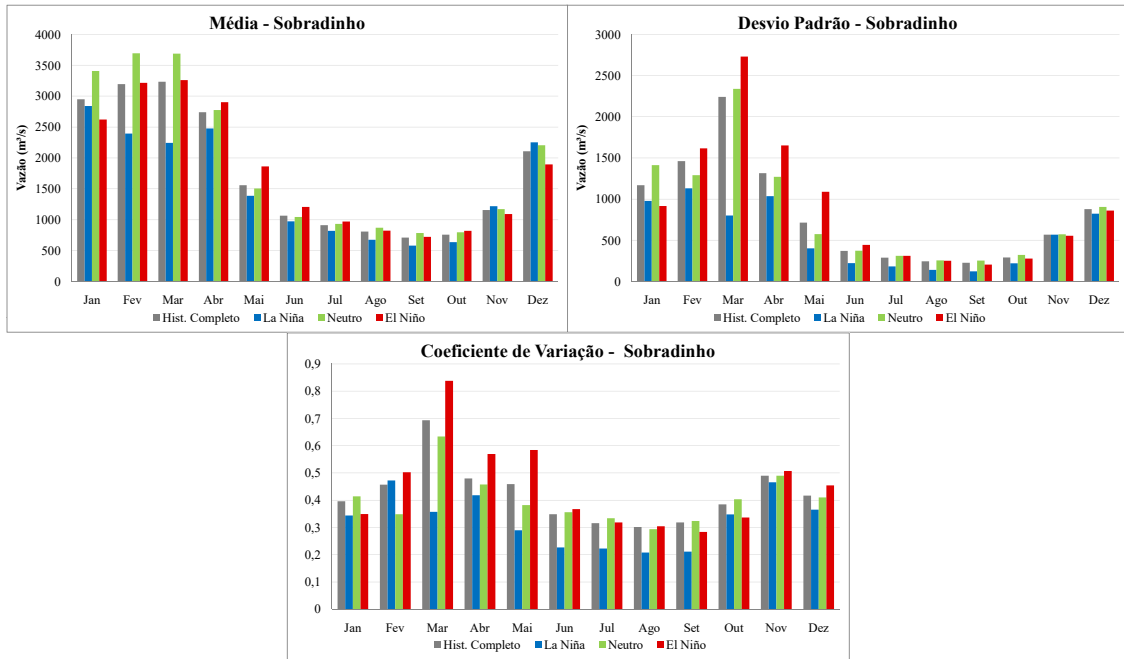


Figura 4.11: Estatísticas históricas da UHE Sobradinho segundo o estado ENOS.

Os resultados para o teste KS das UHEs da região norte do Brasil são expostos na Figura 4.12. Pode-se observar alguns padrões bastante distintos entre as diversas bacias que a compõe. Na bacia hidrográfica do rio Teles Pires e do rio Madeira, as UHEs não possuem uma influência bem demarcada do ENOS. Já nas outras, o contrário ocorre, principalmente da bacia do Xingu e nas UHEs pertencentes à margem esquerda do rio Amazonas, denotadas como Amazonas. Assim como destacado em diversas pesquisas citadas anteriormente, a região norte do Brasil sofre grande influência do fenômeno ENOS, e, quanto maiores as latitudes, maiores são os efeitos. De maneira geral, as duas fases do ENOS são identificadas como influentes na bacia do Amazonas em quase todos os meses, o que configura essa bacia como uma das que sofrem efeitos do ENOS de maneira mais bem marcada.

A Figura 4.13 ilustra os resultados encontrados para a UHE Belo Monte, que se localiza na bacia do rio Xingu no estado do Pará. É possível observar que durante os períodos de EN os meses de janeiro a março apresentam médias menores que a MLT. O mês de abril, apesar de apresentar médias próximas, possui um desvio padrão maior em períodos de EN. Por outro lado, em maio e junho, a média em períodos de LN, apresenta-se com valores superiores, sem grandes alterações no desvio padrão. Pelo fato da UHE Belo Monte se localizar na região norte do Brasil, o ENOS causa uma alteração no padrão do regime hidrológico bastante distinto do resto do Brasil. Na UHE Belo Monte, e, de maneira geral, nas outras UHEs da região norte, a LN provoca anomalias positivas, enquanto que o EN causa anomalias negativas nas afluências. Destaca-se que, para a UHE Belo Monte, o teste KS indica a influência do EN entre janeiro e março e no mês de novembro, e a LN denota ser

influyente em maio, setembro e novembro.

Na Figura 4.14 são apresentadas as estatísticas calculadas para UHE Santo Antônio Jari, que se encontra região norte do Brasil, especificamente na margem esquerda do rio Amazonas. De janeiro a julho, pode ser observado que há uma alteração importante no comportamento das afluências entre os diferentes estados, principalmente nos meses mais úmidos. Os períodos de LN apresentam média e desvio padrão maiores que os demais estados, enquanto o inverso ocorre nos períodos de EN. Já os períodos N apresentam médias próximas aos valores encontrados no histórico sem a classificação. Entre fevereiro e julho, e em novembro, ambos os estados de ENOS são identificados como influentes, enquanto que em dezembro apenas o EN.

É interessante notar que esses resultados vão ao encontro dos padrões caracterizados nos estudos apresentados anteriormente sobre a influência do fenômeno ENOS nas precipitações que ocorrem no Brasil. A região norte, de forma geral, sofre com chuvas abaixo da média em períodos de EN, e o contrário ocorre em períodos de LN.

UHE	BACIA	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
BELO MONTE	XINGU												
CURUÁ-UNA	XINGU												
DARDANELOS	MADEIRA												
GUAPORÉ	MADEIRA												
JIRAU	MADEIRA												
RONDON II	MADEIRA												
SAMUEL	MADEIRA												
STO. ANTÔNIO	MADEIRA												
BALBINA	AMAZONAS												
CACH. CALDEIRÃO	AMAZONAS												
COARACY NUNES	AMAZONAS												
FERREIRA GOMES	AMAZONAS												
STO. ANT. JARI	AMAZONAS												
CANA BRAVA	TOCANTINS												
ESTREITO TOC.	TOCANTINS												
LAJEADO	TOCANTINS												
PEIXE ANGICAL	TOCANTINS												
SÃO SALVADOR	TOCANTINS												
SERRA DA MESA	TOCANTINS												
TUCURUÍ	TOCANTINS												
SÃO MANOEL	TELES PIRES												
TELES PIRES	TELES PIRES												

Figura 4.12: Teste KS para a região norte do Brasil.

Outro importante ponto que deve ser destacado diz respeito às alterações no desvio padrão. Apesar das vazões na UHE Santo Antônio Jari, em média, serem maiores em períodos classificados como LN do que em outros períodos, observa-se que isso sucede em conjunto com uma maior variabilidade. Ou seja, há uma incerteza maior sobre o que ocorrerá em períodos LN, enquanto que os períodos de EN possuem médias menores e menor incerteza. Além disso, ao se dividir o histórico nos três estados, dois deles passaram a apresentar um desvio padrão menor do que no histórico sem classificação.

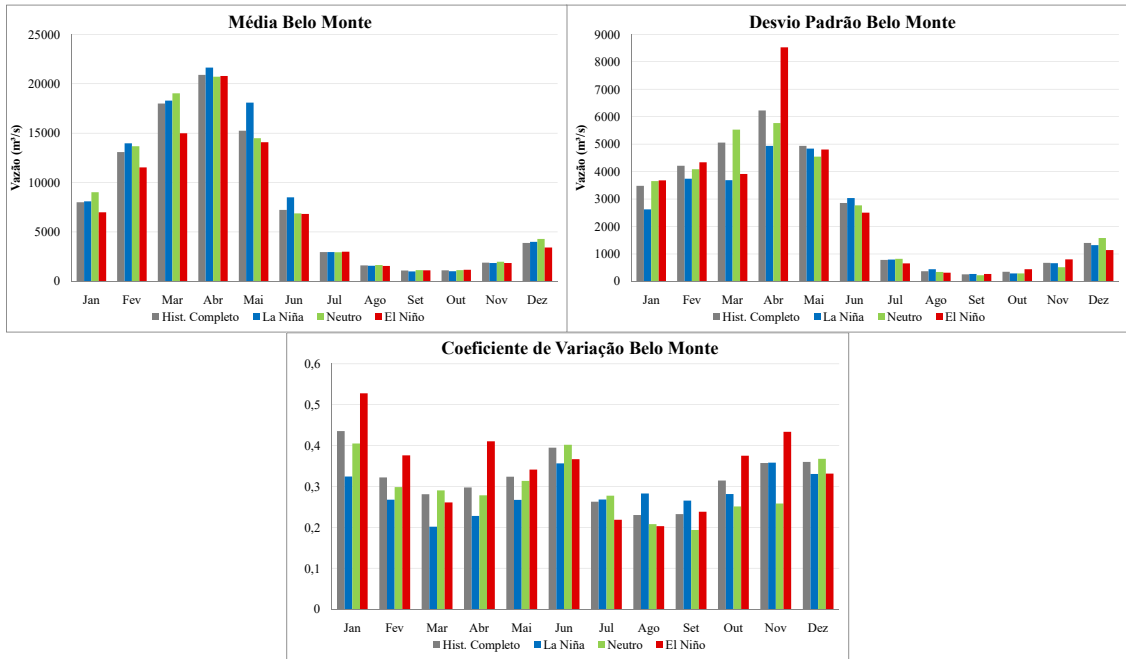


Figura 4.13: Estatísticas históricas da UHE Belo Monte segundo o estado ENOS.

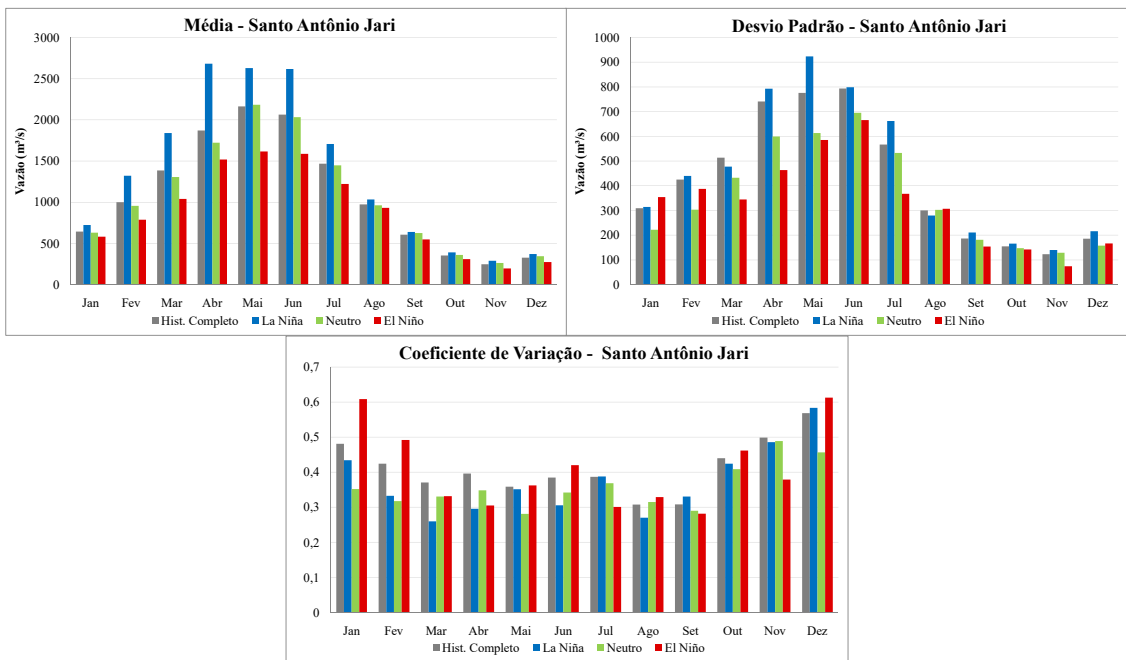


Figura 4.14: Estatísticas históricas da UHE Santo Antônio Jari segundo o estado ENOS.

O fenômeno ENOS, assim como a maioria dos fenômenos climáticos que atuam em escala global, produz efeitos de forma não-linear. O modelo MS-PAR(p), apresentado nesta dissertação, aproxima linearmente as consequências causadas pelo ENOS através da divisão do histórico de vazões naturais mensais das UHEs brasileiras entre os três estados de ENOS, segundo a classificação obtida através do índice

ONI. Conforme exposto nesta seção, a aproximação realizada apresenta-se adequada, conseguindo capturar as diversas alterações produzidas no ciclo hidrológico em escala mensal ao longo de todas as bacias hidrográficas brasileiras. Além do processo de classificação do histórico, propôs-se uma segunda etapa de identificação da influência do ENOS através da aplicação de um procedimento envolvendo o teste de aderência KS.

Essa segunda etapa busca deixar o modelo MS-PAR(p)2 mais parcimonioso, descartando o uso de divisão do histórico quando esse procedimento não traz qualquer informação relevante ao processo. Por muitas vezes, identifica-se que apenas um dos estados do ENOS traz alguma informação relevante. Nesses casos, a fase oposta e o estado N possuem estatísticas bastante próximas, e, assim, são considerados como apenas um estado. Do total de 141 UHEs analisadas, apenas 19 não apresentam diferenças entre os estados de ENOS em pelo menos um mês. Conforme exposto, o ENOS não provoca alterações apenas nas vazões médias, mas também no desvio padrão. Essa constatação indica que, determinados estados de ENOS, há uma previsibilidade maior do que ocorrerá, enquanto outros possuem uma variabilidade associada muito grande. No contexto da geração de cenários sintéticos, esse tipo de informação é também de extrema relevância.

Muitos dos resultados expostos nesta seção vão ao encontro de diversas pesquisas citadas na revisão bibliográfica. De maneira geral, o estado EN produz anomalias positivas na região sul entre os meses de novembro e março, com o efeito inverso no estado LN. Como as vazões nessa região não possuem uma correlação mensal forte, esse efeito não é propagado para os outros meses do ano. Na região sudeste/centro-oeste existe uma heterogeneidade grande de resultados nas diferentes bacias hidrográficas que a compõe, devido ao fato dela se situar em uma região de transição dos efeitos do ENOS, onde, segundo a revisão bibliográfica realizada, não há um sinal bem definido. Apresenta-se um ligeiro aumento das médias e desvios padrão das vazões afluentes no estado EN e diminuição na LN e N, quando comparados ao histórico completo. A bacia do Paranapanema, por estar mais ao sul, sofre uma influência mais forte do fenômeno ENOS.

Em todas as outras bacias hidrográficas que compõe a região sudeste, o ENOS produz efeitos de magnitude menor, porém sua atuação não se restringe apenas ao período úmido. Devido à forte correlação mensal existente nas vazões das UHEs presentes na região sudeste/centro-oeste, os impactos do ENOS, quando presentes, são propagados ao longo do ano. Na região nordeste, existe uma forte atuação do ENOS nas aflúências entre janeiro e março, onde ocorre a maior parte das chuvas na região, sendo o efeito observado também nos meses subsequentes. Assim como nas regiões sul e sudeste, o EN provoca anomalias positivas, e o oposto ocorre na LN. É interessante ressaltar que, em novembro e em dezembro, os efeitos são inversos

na região nordeste, com a LN apresentando anomalias positivas e o EN negativas, porém com menor intensidade. Essa variabilidade em escala mensal já havia sido observada nas referências citadas anteriormente. Por último, na região norte, a influência do fenômeno ENOS é observada com bastante intensidade, principalmente nas UHEs localizadas na margem esquerda do rio Amazonas. Ao contrário do que ocorre na maior parte do Brasil, a LN provoca anomalias positivas, e oposto ocorre com o EN. A Tabela 4.3 traz um resumo dos efeitos do fenômeno ENOS nas quatro macro-regiões brasileiras.

Tabela 4.3: Efeitos do ENOS nas vazões afluentes às UHEs brasileiras

Região	La Niña	El Niño
Sul	Anomalias negativas, restritas aos meses de novembro a março	Anomalias positivas, restritas aos meses de novembro a março
Sudeste / Centro-Oeste	Heterogeneidade de resultados; Anomalias negativas de pequena intensidade propagadas por todo o ano	Heterogeneidade de resultados; Anomalias positivas de pequena intensidade propagadas por todo o ano
Nordeste	Anomalias negativas de grande intensidade entre janeiro e março; Anomalias positivas de pequena intensidade em novembro e dezembro	Anomalias positivas de grande intensidade entre janeiro e março; Anomalias negativas de pequena intensidade em novembro e dezembro
Norte	Anomalias positivas de grande intensidade, propagadas por todo o ano	Anomalias negativas de grande intensidade propagadas por todo o ano

4.3 Caso Não-Condicionado

Esta seção apresentará os resultados obtidos através da geração de cenários de vazões afluentes realizada de forma não-condicionada. Um modelo estocástico adequado para a geração de cenários de vazões aplicado ao planejamento da operação energética do SIN deve ser capaz de reproduzir diversas características encontradas nos registros históricos de afluências, como os momentos e correlações espaciais entre as UHEs. Visando o planejamento de médio/longo prazo, também é interessante que o modelo esteja apto a reproduzir as sequências negativas (períodos consecutivos abaixo da MLT) tanto de forma univariada quanto multivariada. Para essa avaliação, foi feita a geração multivariada de 2.000 cenários sintéticos por 67 anos (mesmo comprimento do registro histórico utilizado no ajuste do modelo).

Além da geração de cenários de afluências, o modelo MS-PAR(p) realiza, em conjunto, a geração de cenários de ENOS. Essa geração é realizada através de uma cadeia de Markov de ordem 1 ajustada a cada mês do histórico, conforme apresentado no item 3.4.1. Sendo assim, deve-se avaliar se o modelo proposto consegue reproduzir, nas séries sintéticas, as características encontradas no histórico, ou seja, investiga-se a reprodução da distribuição histórica de probabilidade de ENOS com relação ao que é gerado através do modelo MS-PAR(p). Na Figura 4.15, são apresentadas a distribuição histórica do ENOS ao longo dos meses (gráfico à esquerda) e a distribuição obtida através dos 2.000 cenários sintéticos ao longo de 67 anos (gráfico à direita). As barras vermelhas, verdes e azuis representam o percentual de EN, N e LN. Pode-se observar que a distribuição sintética reproduz corretamente os valores históricos.

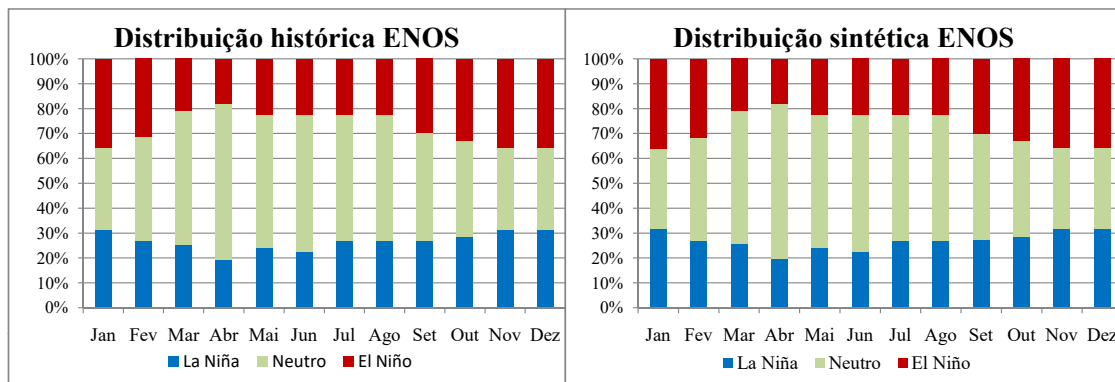


Figura 4.15: Distribuições de frequências histórica e sintética, gerada através da cadeia de Markov.

Dado que a evolução dos eventos de ENOS é condicionado ao estado passado, e, especificamente, no modelo MS-PAR(p) é adotada uma cadeia de Markov de ordem 1 de forma a representar o desenvolvimento dos eventos, avalia-se, também, a reprodução das transições históricas entre os três estados de ENOS para cada

mês. Dessa forma, apresentam-se na Figura 4.16 as matrizes de transições mensais históricas (a) e módulo da diferença entre as matrizes de transições mensais sintéticas com relação às históricas (b), em valores percentuais. Quanto mais verde, maior é probabilidade de ocorrência da transição de um estado para o outro indicado, e valores em branco representam probabilidade nula. É possível observar que em alguns meses, os estados de ENOS persistem onde estiverem, ou seja, não há transição. Nota-se que os resultados obtidos na simulação reproduzem adequadamente o comportamento dessa variável ao longo do tempo.

		Mês seguinte		
		La Niña	Neutro	El Niño
Jan	La Niña	100	0	0
	Neutro	0	100	0
	El Niño	0	0	100
Fev	La Niña	85.714	14.286	0
	Neutro	0	95.455	4.545
	El Niño	0	16.667	83.333
Mar	La Niña	94.444	5.556	0
	Neutro	0	100	0
	El Niño	0	33.333	66.667
Abr	La Niña	76.471	23.529	0
	Neutro	0	94.444	5.556
	El Niño	0	28.571	71.429
Mai	La Niña	92.308	7.692	0
	Neutro	9.524	80.952	9.524
	El Niño	0	8.333	91.667
Jun	La Niña	87.5	12.5	0
	Neutro	2.778	88.889	8.333
	El Niño	0	20	80
Jul	La Niña	93.333	6.667	0
	Neutro	10.811	83.784	5.405
	El Niño	0	13.333	86.667
Ago	La Niña	88.889	11.111	0
	Neutro	5.882	91.176	2.941
	El Niño	0	6.667	93.333
Set	La Niña	94.444	5.556	0
	Neutro	2.941	82.353	14.706
	El Niño	0	0	100
Out	La Niña	94.444	5.556	0
	Neutro	6.897	86.207	6.897
	El Niño	0	0	100
Nov	La Niña	100	0	0
	Neutro	7.692	84.615	7.692
	El Niño	0	0	100
Dez	La Niña	100	0	0
	Neutro	0	100	0
	El Niño	0	0	100

(a)

		Mês seguinte		
		La Niña	Neutro	El Niño
La Niña	0	0	0	
Neutro	0	0	0	
El Niño	0	0	0	
La Niña	0.24	0.24	0	
Neutro	0	0.159	0.159	
El Niño	0	0.268	0.268	
La Niña	0.156	0.156	0	
Neutro	0	0	0	
El Niño	0	0.007	0.007	
La Niña	0.459	0.459	0	
Neutro	0	0.017	0.017	
El Niño	0	0.416	0.416	
La Niña	0.184	0.184	0	
Neutro	0.014	0.002	0.016	
El Niño	0	0.198	0.198	
La Niña	0.469	0.469	0	
Neutro	0.051	0.15	0.201	
El Niño	0	0.194	0.194	
La Niña	0.202	0.202	0	
Neutro	0.04	0.104	0.064	
El Niño	0	0.298	0.298	
La Niña	0.098	0.098	0	
Neutro	0.003	0.045	0.044	
El Niño	0	0.135	0.135	
La Niña	0.299	0.299	0	
Neutro	0.011	0.033	0.044	
El Niño	0	0	0	
La Niña	0.122	0.122	0	
Neutro	0.013	0.019	0.006	
El Niño	0	0	0	
La Niña	0	0	0	
Neutro	0.028	0.045	0.016	
El Niño	0	0	0	
La Niña	0	0	0	
Neutro	0	0	0	
El Niño	0	0	0	

(b)

Figura 4.16: Matrizes de transição mensais históricas (a) e o módulo da diferença entre as matrizes de transições mensais sintéticas com relação às históricas (b), em valores percentuais.

De forma a analisar mais profundamente as séries sintéticas de ENOS, avalia-se a seguir se essas são capazes de reproduzir corretamente o comprimento (ou duração) típico de eventos de ENOS. Sendo assim, contabilizou-se o comprimento

de cada sequência de estado de ENOS, formando-se uma distribuição de comprimentos entre os três estados de ENOS. O mesmo procedimento foi realizado nas séries sintéticas, obtendo-se uma distribuição sintética de comprimentos para os três estados. Através da aplicação do teste multinomial apresentado no Capítulo 3.5, aceitou-se a hipótese nula de que as distribuições histórica e sintética, para cada estados de ENOS, são iguais. A Tabela 4.4 traz a média, o máximo e o desvio padrão das distribuições de frequência histórica e sintéticas de comprimento para as sequências dos três estados. É possível observar que os cenários sintéticos conseguem reproduzir a duração dos eventos de ENOS de forma adequada.

Tabela 4.4: Estatísticas histórica e sintética referentes ao comprimento de eventos de ENOS

	Média		Máximo		Desvio Padrão	
	Histórico	Sintético	Histórico	Sintético	Histórico	Sintético
La Niña	13,1	13,6	32,0	97,0	8,1	8,9
Neutro	9,7	9,6	50,0	100,0	10,9	9,8
El Niño	10,0	10,0	19,0	58,0	4,4	5,0

Observada a geração correta de cenários de ENOS, serão realizadas, a seguir, as avaliações para os cenários de vazões. Primeiramente, apresentam-se os resultados referentes à reprodução dos momentos de primeira e segunda ordem para algumas das UHEs selecionadas na Tabela 4.1. Como esses parâmetros são utilizados na definição do modelo, é esperada sua reprodução nos cenários sintéticos após uma simulação por um período suficientemente longo, como é o caso não condicionado realizado. As Figuras 4.17, 4.18, 4.19 e 4.20 trazem as médias mensais, os desvios padrões mensais e os coeficientes de variação mensais obtidos nas simulações realizada para as UHEs Itá, Furnas Sobradinho e Belo Monte. As barras em cinza, azul, verde e vermelho representam os valores históricos e aqueles obtidos pelo modelo PAR(p), MS-PAR(p)1 e MS-PAR(p)2 para cada um dos meses. É possível observar que todos os três modelos conseguem reproduzir as estatísticas históricas de forma adequada em todas as UHEs, conforme exposto nas figuras supracitadas.

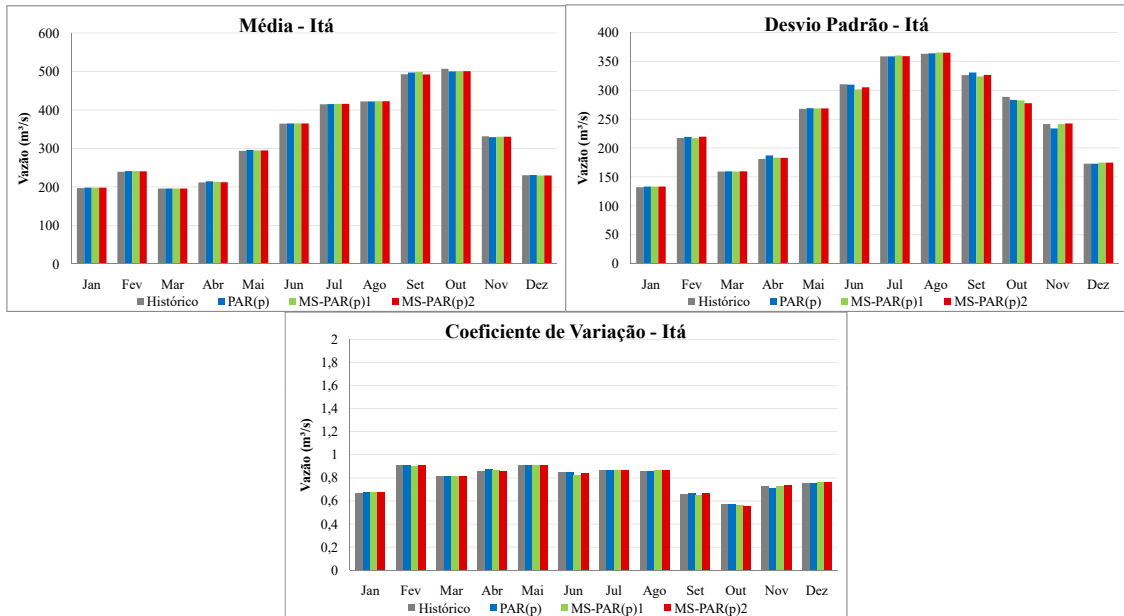


Figura 4.17: Estatísticas obtidas para a UHE Itá para o caso não condicionado.

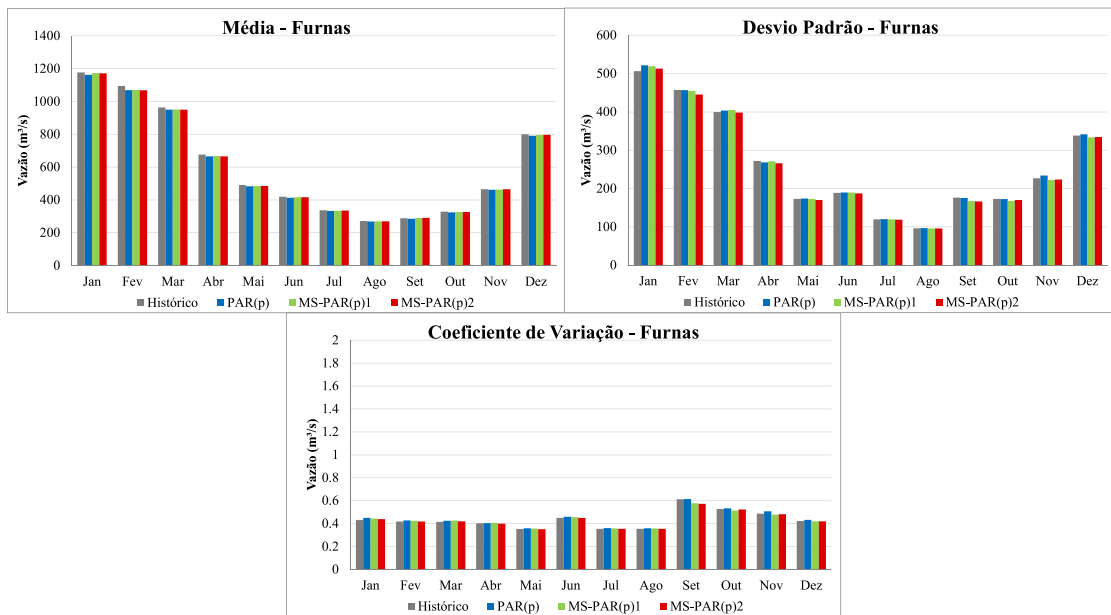


Figura 4.18: Estatísticas obtidas para a UHE Furnas para o caso não condicionado.

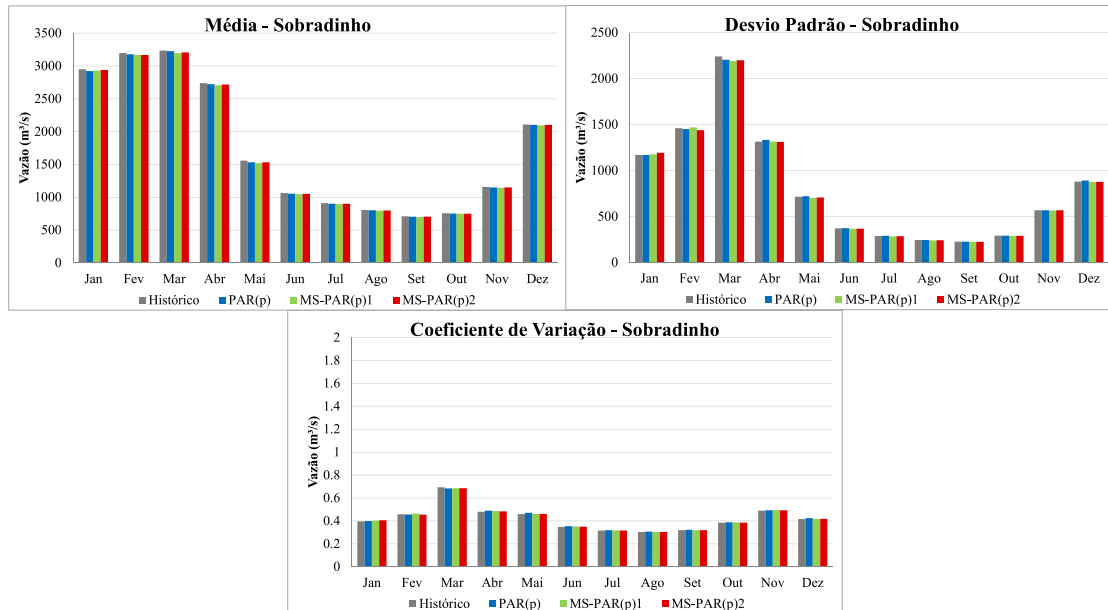


Figura 4.19: Estatísticas obtidas para a UHE Sobradinho para o caso não condicionado.

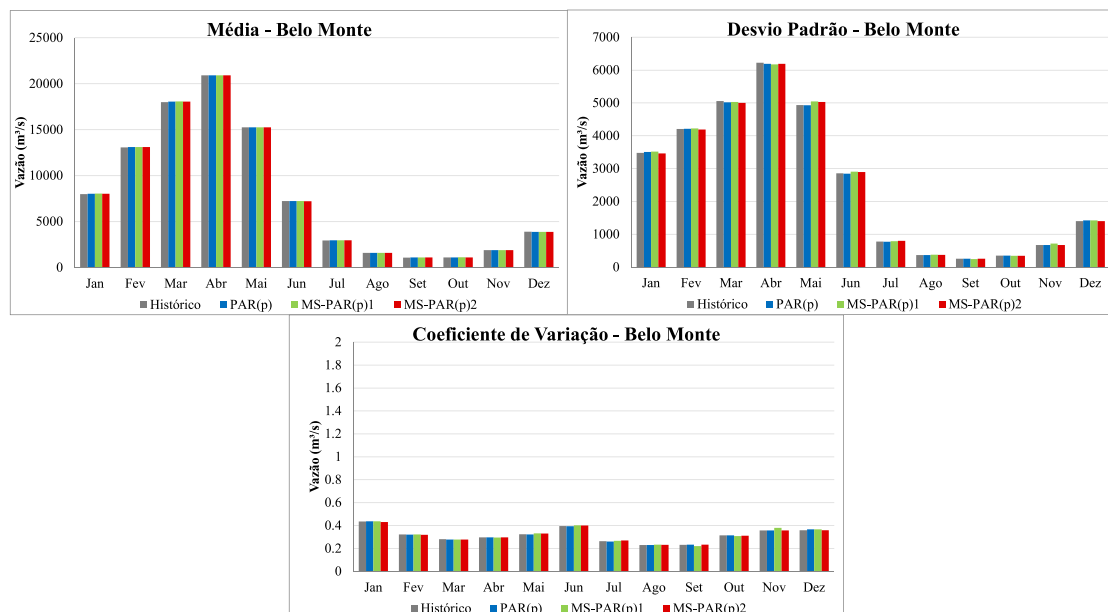


Figura 4.20: Estatísticas obtidas para a UHE Belo Monte para o caso não condicionado.

Conforme demonstrado, os momentos de primeira e segunda ordem são reproduzidos corretamente, conforme esperado. Após a análise inicial do caso não-condicionado, a Tabela 4.5 traz um resumo dos resultados dos testes de aderência KS obtidos para cada uma das 141 UHEs em todos os períodos simulados (804 períodos), totalizando 113364 testes realizados. Nesse teste, para cada período simulado, compara-se a distribuição de frequência sintética dos 2000 cenários gerados à distribuição de frequências histórica, através do teste KS no nível de significância

de 5%. Como pode ser observado, o modelo MS-PAR(p)1 obteve os melhores resultados, sendo o MS-PAR(p)2 o segundo melhor modelo e por último o PAR(p). O total de rejeições manteve-se baixo em todos os três modelos analisados sendo a diferença do modelo MS-PAR(p)1 para o PAR(p) de 4% a menos em termos de rejeição. Destaca-se também que grande parte das rejeições ocorrem em UHEs que possuem vazões incrementais muito pequenas.

Analisando-se individualmente cada UHE, buscou-se identificar aquelas em que houve uma taxa de rejeição maior que 5% e 1% do total de 804 períodos simulados. Nessa análise, o modelo MS-PAR(p)1 possui cerca de metade do percentual de rejeição do modelo PAR(p), ao passo que o modelo MS-PAR(p)2 apresenta uma diminuição do percentual de rejeições nessas duas análises, porém menor que o modelo MS-PAR(p)1. O resultado indica que há uma melhora sensível na reprodução de função de distribuição de frequências de algumas UHEs quando utilizado os modelos MS-PAR(p), com um destaque maior para a opção sem identificação prévia da influência do ENOS.

Tabela 4.5: Resumo do teste KS para os cenários simulados

	PAR(p)	MS-PAR(p)1	MS-PAR(p)2
Rejeição total teste KS (%)	11,04%	7,07%	9,08%
UHEs com pelo menos 5% dos períodos rejeitados (%)	46,8%	23,4%	41,8%
UHEs com pelo menos 1% dos períodos rejeitados (%)	57,5%	31,9%	53,2%

Como foi explicitado na Tabela 4.5, a geração de cenários sintéticos para algumas UHEs não consegue reproduzir corretamente a função de distribuição de frequências. Após uma exploração detalhada, identificou-se que as rejeições quando ocorrem se concentram em determinados meses para algumas UHEs. Para a UHE Furnas, o modelo PAR(p) apresenta, em junho, cerca de 23% de rejeições, enquanto que, em setembro e outubro, 100% dos períodos foram rejeitados. Por outro lado, os modelos MS-PAR(p)1 e MS-PAR(p)2 possuem rejeições apenas no mês de outubro, de 70% e 95%, respectivamente. De forma a identificar qual o fator que leva com que esses meses possuam uma taxa de rejeição elevada, analisa-se, na Figura 4.21, o terceiro momento, assimetria, para a UHE Furnas. Percebe-se que os meses entre junho e outubro possuem uma assimetria bastante elevada, a qual não é reproduzida pelo modelo PAR(p), enquanto que, nos modelos MS-PAR(p), os cenários sintéticos conseguem se aproximar mais dos valores históricos. Pode-se observar que, em outubro, o modelo MS-PAR(p)2 possui assimetria próxima ao modelo PAR(p), tendo

assim uma alta taxa de rejeição. Segundo os resultados apresentados no item 4.2, o mês de outubro não sofre efeito segundo o estado de ENOS, o que explica a proximidade entre os resultados do PAR(p) com o MS-PAR(p)2 para esse mês.

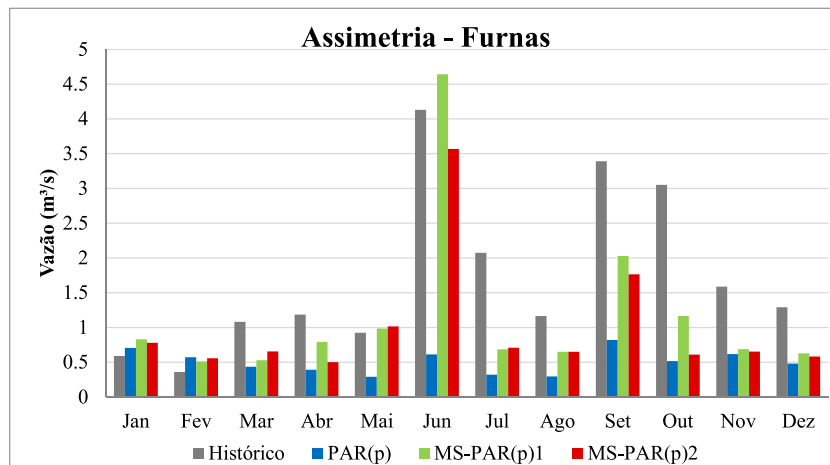


Figura 4.21: Assimetria histórica e sintética da UHE Furnas.

Analisando-se os meses em que houve uma taxa de rejeição alta do teste KS entre as distribuições de frequências sintética e histórica, observa-se que, em sua maioria, as rejeições ocorrem em distribuições altamente assimétricas, que não conseguem ser reproduzidas em sua totalidade nos cenários sintéticos. Deve-se ressaltar que essa observação não é algo que ocorre com frequência. O tratamento de distribuições asssimétricas no modelo GEVAZP é realizado na geração de ruídos lognormais com três parâmetros, conforme descrito no item 3.3.4. Porém, observa-se que o modelo MS-PAR(p)1 possui uma menor taxa de rejeição total do teste KS justamente por reproduzir melhor a assimetria quando comparado ao modelo PAR(p). O MS-PAR(p)2 consegue reproduzir melhor essa característica quando é indicado que o mês em questão possui influência do ENOS. Caso contrário, ele se comporta como o modelo PAR(p), o que explica o fato dele possuir uma rejeição intermediária entre os três modelos analisados. Sendo assim, pode-se inferir que há ganhos na modelagem MS-PAR(p) quanto à reprodução da função de distribuição acumulada de frequências em comparação com o modelo PAR(p), principalmente para distribuições altamente assimétricas.

Uma importante característica do SIN está na complementaridade hídrica que existe entre as diversas regiões hidrográficas que o compõe. Sendo assim, um bom modelo de geração de cenários sintéticos de vazões deve ser capaz de reproduzir as correlações espaciais entre as UHEs observadas no registro histórico. A Figura 4.22 apresenta um mapa de calor das correlações cruzadas anuais entre o histórico (a), e os valores simulados com os modelos PAR(p) (b), MS-PAR(p)1 (c) e MS-PAR(p)2 (d) para as 141 UHEs presentes no estudo de caso. As cores em vermelho correspondem

a correlações positivas e, em azul, a correlações negativas. É possível observar que ambos os três modelos conseguem representar corretamente essa importante característica do histórico.

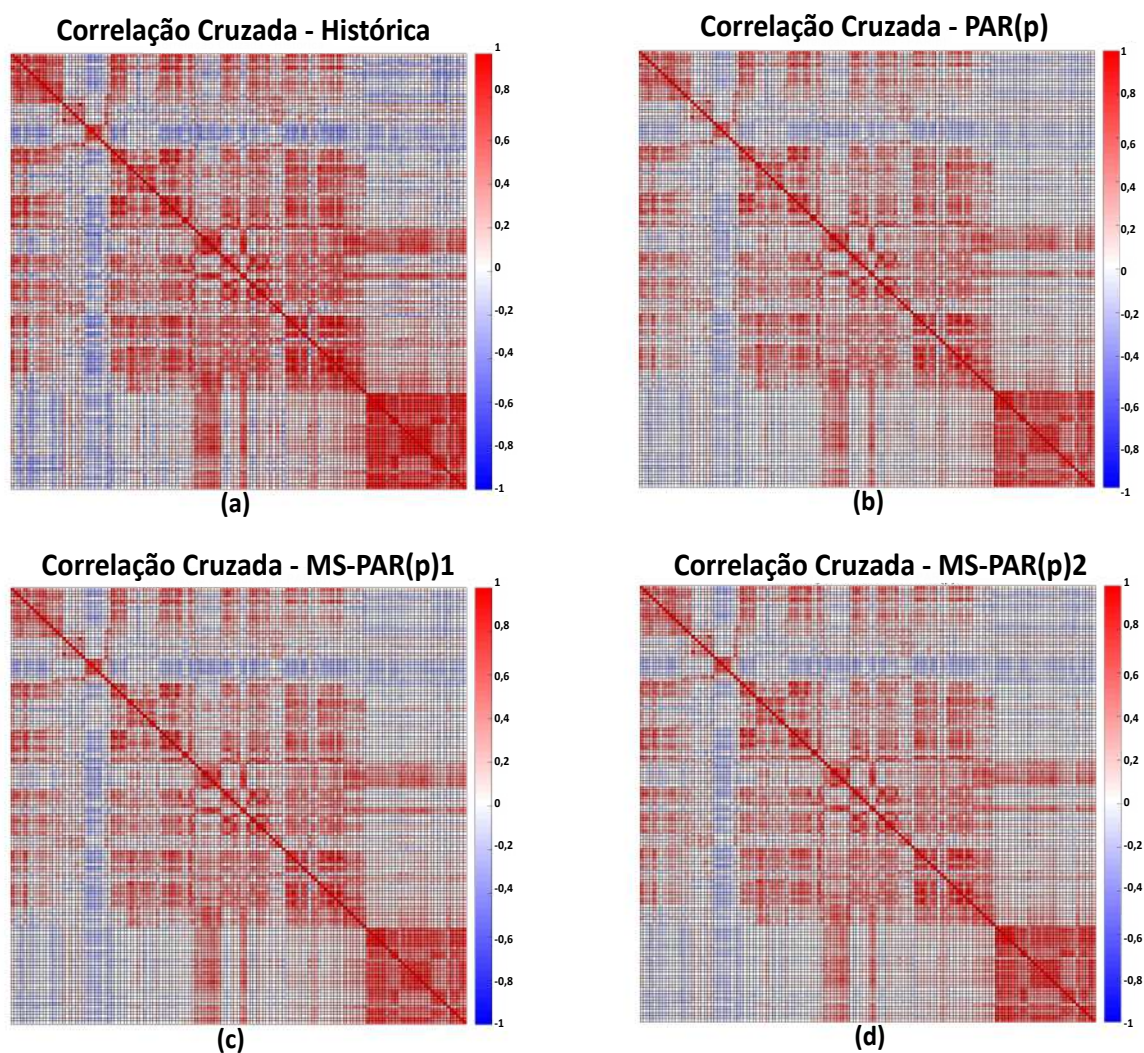


Figura 4.22: Correlações espaciais anuais histórica (a) e geradas com os modelos PAR(p) (b), MS-PAR(p)1 (c) e MS-PAR(p)2 (d)

A geração de cenários de sequências negativas (cenários que possuem vazões abaixo da MLT) é de extrema importância para o planejamento da operação do SIN. Um bom modelo estocástico de geração de séries sintéticas deve ser capaz de reproduzir o histórico de sequências negativas, inclusive as secas que ocorreram de forma multivariada, além de estar apto a gerar cenários ainda mais severos. Sendo assim, apresenta-se, em seguida, uma avaliação das métricas de soma, intensidade e comprimento das sequências negativas, conforme definidas no item 3.5, geradas para este caso não-condicionado. Além da análise univariada para cada UHE, será apresentada uma análise multivariada de sequências negativas que ocorreram simultaneamente nas UHEs Furnas, Sobradinho, Belo Monte e Itá. Primeiramente, apresenta-se, na Tabela 4.6, um resumo das métricas para todas as UHEs do SIN. É

possível observar que o modelo MS-PAR(p)1 se sobressai, tendo as menores taxas de rejeição dos testes, principalmente para o teste de intensidade, onde a rejeição passa de 30,5% no modelo PAR(p) para 15,6%. O modelo MS-PAR(p)2 é ligeiramente melhor nos testes para as variáveis soma e intensidade, enquanto o contrário ocorre para o comprimento, quando comparado com o modelo PAR(p).

Tabela 4.6: Resumo dos testes de sequências negativas para os cenários simulados de vazões

	P	MSP1	MSP2
Rejeição teste KS			
soma da sequência negativa (%)	10,6%	7,8%	7,8%
Rejeição teste KS			
intensidade da sequência negativa (%)	30,5%	15,6%	27,7%
Rejeição teste multinomial			
comprimento sequência negativa (%)	24,1%	24,1%	27,7%

As Tabelas 4.7, 4.8 e 4.9 trazem os resultados obtidos para as variáveis comprimento, intensidade e soma das treze UHEs avaliadas individualmente. Além das análises univariadas, faz-se também uma avaliação das sequências negativas de forma multivariada. Para tanto, elegeram-se dois pares de UHEs de regiões hidrográficas distintas: Furnas e Belo Monte; Itaipu e Sobradinho. Além disso, uma análise foi realizada levando-se em conta quatro UHEs Itá, Porto Primavera, Sobradinho e Belo Monte. Para cada uma das variáveis examinadas de sequência negativa, são apresentadas a média, desvio padrão e valor máximo, tanto do histórico, quanto dos valores simulados por cada um dos modelos. É possível observar que, excetuando-se alguns casos, as médias dos valores históricos e sintéticos das três variáveis de sequência negativa são próximos, inclusive para as análises multivariadas. Além disso, todos os modelos foram capazes de produzir valores máximos acima do histórico. Essa característica demonstra a capacidade desses em simular cenários ainda mais críticos que os observados no histórico, que persistem por uma quantidade de tempo maior, e são ainda mais intensos.

4.4 Caso Condicionado - período pré-ajuste

Nesta seção, serão apresentados os resultados obtidos através da geração de 2.000 cenários sintéticos para o período de 1951 a 2009 (período em que os modelos foram ajustados nestes casos). A cada período, a tendência hidrológica é atualizada com o passado recente. Para os modelos MS-PAR(p), o estado ENOS ao longo de

Tabela 4.7: Estatísticas históricas e sintéticas do comprimento de sequências negativas de vazões

	Média												Máximo			Desvio Padrão			
	Histórico			MSP1			MSP2			Histórico			MSP1			MSP2			
	P	MSP1	MSP2	P	MSP1	MSP2	Histórico	P	MSP1	MSP2	Histórico	P	MSP1	MSP2	Histórico	P	MSP1	MSP2	
Sobradinho	4,9	4,9	4,4	4,6	4,6	4,4	25	87	71	77	5,5	5,9	5,2	5,0					
Belo monte	4,0	4,1	3,9	3,9	3,9	3,9	26	62	57	49	4,5	4,3	4,1	4,0					
Estreito Toc.	4,1	4,5	4,3	4,3	4,3	4,3	26	75	80	75	4,4	4,7	4,8	4,5					
Jirau	3,7	3,6	3,4	3,5	3,4	3,4	35	81	89	81	6,2	5,6	5,5	5,5					
Sto. Ant. Jari	4,2	4,1	4,0	4,0	4,0	4,0	27	74	61	60	5,4	4,9	4,5	4,3					
Teles Pires	4,1	4,6	4,3	4,3	4,3	4,3	24	83	65	77	4,3	5,2	4,5	4,7					
Capivara	5,0	4,1	4,1	4,1	4,0	4,0	37	63	60	74	5,8	4,7	4,5	4,4					
Furnas	4,2	4,8	4,5	4,5	4,5	4,5	39	134	83	71	4,8	6,9	5,6	5,6					
P. Primavera	4,0	3,8	3,7	3,8	3,7	3,7	30	90	83	86	5,5	5,1	4,5	5,0					
Itá	4,3	3,8	3,7	3,7	3,7	3,7	16	44	42	45	3,7	3,6	3,5	3,6					
Itaipu	4,8	4,4	4,2	4,2	4,1	4,1	41	128	128	125	7,0	6,1	5,8	6,1					
Salto Caxias	4,1	3,7	3,6	3,7	3,7	3,6	32	55	49	45	4,6	3,7	3,5	3,6					
Furnas + Belo Monte	2,9	2,6	2,4	2,4	2,4	2,4	11	54	33	32	2,6	2,6	2,3	2,3					
Itaipu + Sobradinho	2,3	2,4	2,2	2,3	2,2	2,2	12	39	37	32	2,0	2,4	2,2	2,1					
Itá + P. Primavera																			
+ Sobradinho	1,4	2,3	2,4	2,4	2,3	2,3	4	34	33	38	0,8	2,1	2,1	2,1					
+ Belo Monte																			

Tabela 4.8: Estatísticas históricas e sintéticas da intensidade de sequências negativas de vazões

	Média												Máximo						Desvio Padrão					
	Histórico			P			MSP1			MSP2			Histórico			P			MSP1			MSP2		
	Histórico	P	MSP1	MSP2	Histórico	P	MSP1	MSP2	Histórico	P	MSP1	MSP2	Histórico	P	MSP1	MSP2	Histórico	P	MSP1	MSP2	Histórico	P	MSP1	MSP2
Sobradinho	440	500	498	502	1736	3065	3058	3081	377	402	403	407	440	500	498	502	1736	3065	3058	3081	377	402	403	407
Belo monte	1269	1551	1536	1577	3750	12738	12847	12725	1067	1401	1423	1453	1269	1551	1536	1577	3750	12738	12847	12725	1067	1401	1423	1453
Estreito Toc.	355	401	380	403	1264	2818	2589	2781	254	318	308	322	355	401	380	403	1264	2818	2589	2781	254	318	308	322
Jirau	978	1238	1238	1184	4792	8909	10028	9564	1237	1391	1413	1396	978	1238	1238	1184	4792	8909	10028	9564	1237	1391	1413	1396
Sto. Ant. Jari	239	239	241	243	798	1574	1785	1777	191	187	190	188	239	239	241	243	798	1574	1785	1777	191	187	190	188
Teles Pires	316	358	353	354	1087	2339	2296	2225	274	264	264	265	316	358	353	354	1087	2339	2296	2225	274	264	264	265
Capivara	114	133	124	130	299	512	499	508	63	75	70	71	114	133	124	130	299	512	499	508	63	75	70	71
Furnas	138	136	138	138	573	1146	1009	963	105	112	112	109	138	136	138	138	573	1146	1009	963	105	112	112	109
P. Primavera	322	258	252	258	1390	1182	1198	1172	282	174	169	175	322	258	252	258	1390	1182	1198	1172	282	174	169	175
Itá	130	132	127	129	300	492	491	492	73	71	68	69	130	132	127	129	300	492	491	492	73	71	68	69
Itaipu	575	631	592	618	1570	2158	2127	2174	380	371	359	373	575	631	592	618	1570	2158	2127	2174	380	371	359	373
Salto Caxias	125	128	123	126	260	439	408	401	56	65	63	63	125	128	123	126	260	439	408	401	56	65	63	63
Furnas + Belo Monte	1691	1970	1929	1953	9374	13525	14121	13301	1648	1850	1883	1872	1691	1970	1929	1953	9374	13525	14121	13301	1648	1850	1883	1872
Itaipu + Sobradinho	1270	1256	1195	1257	3365	4716.1	4584	4770	678	664	630	663	1270	1256	1195	1257	3365	4716.1	4584	4770	678	664	630	663
Itá + P. Primavera																								
+ Sobradinho	502	451	432	446	1217	1365	1393	1380	262	205	198	205	502	451	432	446	1217	1365	1393	1380	262	205	198	205
+ Belo Monte																								

Tabela 4.9: Estatísticas históricas e sintéticas da soma de sequências negativas de vazões

		Média						Máximo						Desvio Padrão									
		Histórico		MSP1		MSP2		Histórico		P		MSP1		MSP2		Histórico		P		MSP1		MSP2	
		P	Histórico	P	Histórico	P	Histórico	P	Histórico	P	Histórico	P	Histórico	P	Histórico	P	Histórico	P	Histórico	P	Histórico	P	Histórico
Sobradinho	Nordeste	2932	2733	2647	2563	27245	72491	60105	64951	4680	4677	3896	3818										
Belo monte	Norte	7802	7092	7443	7381	47349	152356	151282	136546	9547	11410	11053	10832										
Estreito Toc.	Norte	2034	1905	1930	1944	13692	52532	55455	53872	2422	2824	2919	2724										
Jirau	Norte	10039	13614	9678	9637	158546	395763	403514	439209	23850	21407	21452	21659										
Sto. Ant. Jari	Norte	1303	1188	1233	1213	11286	36087	27171	25633	1780	2192	1924	1810										
Teles Pires	Norte	2070	2279	1873	1932	14667	60635	40290	53693	2842	3380	2808	3035										
Capivara	Sudeste	690	805	642	647	8345	14671	13709	14393	1234	1011	911	902										
Furnas	Sudeste	972	1046	838	849	10695	40964	24163	23481	1990	2013	1467	1482										
P. Primavera	Sudeste	1231	1532	1138	1205	10978	48561	34904	48145	2262	2296	1820	2259										
Itá	Sul	574	500	541	545	2167	8390	7562	7791	517	662	615	639										
Itaipu	Sul	3956	3649	3532	3616	35996	176358	168307	180286	5443	7680	6829	7746										
Salto Caxias	Sul	551	572	529	525	3216	8903	7451	7538	609	681	600	635										
Furnas + Belo Monte		5550	6194	5110	5169	48098	142278	102237	104671	9359	7873	7293	7262										
Itaipu + Sobradinho		3292	2870	2983	3038	17743	75092	78175	73826	2893	4394	3729	3909										
Itá + P. Primavera																							
+ Sobradinho		1167	700	1120	1142	3259	19193	20312	26379	524	1388	1257	1369										
+ Belo Monte																							

todo período a ser simulado é um dado previamente conhecido (previsão perfeita do estado do ENOS). Para a análise dos resultados nesta seção, serão utilizadas as métricas de avaliação EMA, EMPA e REQM, apresentados no item 3.5, calculadas para cada UHE em todos os períodos simulados.

A Tabela 4.10 apresenta a avaliação dos casos condicionado para o período de pré-ajuste nas UHEs selecionadas ao longo de todos os períodos, obtendo-se um valor médio para cada métrica de avaliação. Os valores destacados em negrito correspondem ao menor valor da métrica entre os três modelos analisados, e consequentemente, indicam o melhor deles segundo a métrica em questão. Observa-se que, na maior parte das UHEs analisadas, o modelo MS-PAR(p)1 foi indicado como o melhor, principalmente na região sul, obtendo uma redução do EMPA em relação ao modelo PAR(p) de quase 10% na UHE Itá. O modelo MS-PAR(p)2, em muitos casos, acaba reduzindo o erro em relação ao modelo PAR(p), porém é pior quando comparado ao modelo MS-PAR(p)1. Ainda utilizando o EMPA como balizador, o modelo MS-PAR(p)1 tem erros menores em 72% das UHEs em relação ao modelo PAR(p), destacando-se, principalmente, nas regiões sul, nordeste e norte do Brasil. Já o modelo MS-PAR(p)2, é apontado como superior em 46,7% das UHEs.

Tabela 4.10: Avaliação da geração de cenários de vazões para o período de ajuste dos modelos entre os anos de 1951 e 2009

UHE	Região	EMA						EMPA						REQM																																																																																												
		P		MSP1		MSP2		P		MSP1		MSP2		P		MSP1		MSP2																																																																																								
		P	MSP1	MSP2	P	MSP1	MSP2	P	MSP1	MSP2	P	MSP1	MSP2	P	MSP1	MSP2	P	MSP1	MSP2																																																																																							
Sobradinho	Nordeste	622,04	599,64	610,14	31,70	30,82	31,24	773,03	745,07	757,76	2023,05	1976,18	2000,41	25,49	25,03	25,34	2470,60	2414,20	2444,85	480,92	459,43	470,68	24,78	23,99	24,30	588,62	561,61	576,31	1965,58	1982,24	1955,44	13,81	13,76	13,65	2373,14	2394,92	2361,09	36,26	36,20	37,24	380,74	384,93	387,77	424,77	431,63	439,16	19,46	20,12	20,38	517,38	525,43	535,04	187,06	181,53	186,88	60,89	58,90	61,18	247,12	237,67	244,84	179,72	184,44	190,08	30,53	32,18	32,93	220,26	227,86	234,00	350,42	358,25	351,48	67,02	66,87	66,30	439,25	452,78	441,38	203,02	196,09	201,04	101,56	93,17	96,79	269,62	261,57	268,26	855,36	811,25	825,12	71,18	68,61	67,99	1090,64	1033,14	1053,43	187,01	180,94	188,24	93,90	88,94	93,05	248,90	240,45	250,98

Como o modelo GEVAZP gera cenários sintéticos mensais, há, também, a necessidade de avaliar os resultados mês a mês, identificando, assim, em quais períodos há um maior potencial de ganho com o modelo MS-PAR(p). As Tabelas 4.11, 4.12 e 4.13 trazem o EMPA mensal para cada uma das UHEs analisada. Os resultados em negrito correspondem ao menor valor entre os três modelos. É possível que o modelo PAR(p) leve vantagem sobre os modelos MS-PAR(p), principalmente nos meses de recessão das vazões afluentes, que são meses mais secos na maior parte do Brasil, com exceção da região sul. As UHEs Furnas, Porto Primavera e Teles Pires, assim como exposto na Tabela 4.10, apresentam melhores resultados com o modelo PAR(p) na maior parte dos meses. Porém, nota-se que, para a maior parte das UHEs e, em especial, durante o período úmido, há uma ampla vantagem dos modelos MS-PAR(p) sobre o PAR(p).

Uma possível explicação para esses resultados pode estar no fato de que o fenômeno ENOS possui uma maior influência na precipitação e, conseqüentemente, nas vazões afluentes que ocorrem durante o período úmido. Como no período de recessão as afluições ocorrem mais como um produto das vazões passadas, o modelo MS-PAR(p) deixa de possuir vantagem sobre o PAR(p). A melhor informação que um modelo pode utilizar para gerar cenários futuros de vazões afluentes durante o período seco é a magnitude da vazão observada ao longo do período úmido, e qualquer outro tipo de informação pode acabar introduzindo mais ruído ao modelo. Destacado esse ponto, deve-se observar que as vazões afluentes ocorrem em uma proporção muito maior durante o período úmido, e que o fato dos modelos MS-PAR(p) apresentarem um menor EMPA nesses períodos pode ser considerado uma vantagem em relação ao modelo PAR(p).

Tabela 4.11: EMPA mensal do modelo PAR(p) para o período de ajuste dos modelos entre os anos de 1951 e 2009.

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Sobradinho	39	49	53	47	21	16	10	9	13	24	47	51
Belo Monte	45	29	21	32	27	21	20	12	15	25	29	30
Estreito Toc.	31	43	30	33	28	12	12	8	14	22	30	34
Jirau	11	8	9	9	9	9	13	13	22	22	27	15
Sto. Ant. Jari	50	37	34	35	31	31	28	24	28	31	42	62
Teles Pires	23	19	23	19	14	7	6	5	10	35	41	31
Capivara	88	67	44	50	69	54	52	50	59	76	56	65
Furnas	41	40	33	27	19	27	15	18	39	35	36	36
P. Primavera	168	101	93	43	35	44	25	29	37	56	53	122
Itá	82	100	96	99	171	119	86	134	96	73	80	82
Itaipu	101	85	75	51	73	49	42	59	57	65	50	147
Salto Caxias	72	95	104	96	170	107	92	99	81	75	62	74

Tabela 4.12: EMPA mensal do modelo MS-PAR(p)1 para o período de ajuste dos modelos entre os anos de 1951 e 2009.

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Sobradinho	35	43	52	46	18	16	10	11	14	27	46	51
Belo Monte	40	28	20	34	25	21	21	13	15	25	30	27
Estreito Toc.	28	44	28	29	29	11	12	12	14	23	27	31
Jirau	10	7	9	10	11	10	13	14	19	22	26	14
Sto. Ant. Jari	49	39	34	34	30	31	32	26	28	32	41	59
Teles Pires	24	19	24	18	14	9	8	7	13	35	40	30
Capivara	84	63	51	50	61	54	49	50	60	69	50	66
Furnas	41	37	33	28	20	22	21	20	46	41	39	38
P. Primavera	157	110	96	39	35	43	35	29	40	56	52	110
Itá	74	98	84	95	132	100	93	135	93	69	64	80
Itaipu	93	78	85	45	70	47	42	67	53	60	46	137
Salto Caxias	70	81	89	101	165	105	91	96	79	70	54	67

Tabela 4.13: EMPA mensal do modelo MS-PAR(p)2 para o período de ajuste dos modelos entre os anos de 1951 e 2009.

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Sobradinho	36	46	50	47	21	16	12	11	14	27	43	50
Belo Monte	42	28	20	33	26	24	21	12	15	25	30	28
Estreito Toc.	28	43	29	31	30	12	12	8	14	22	30	32
Jirau	10	8	9	9	9	9	13	13	20	22	26	15
Sto. Ant. Jari	51	39	34	34	30	31	32	27	27	31	44	67
Teles Pires	23	19	23	22	15	8	10	7	11	34	42	31
Capivara	87	67	51	52	72	56	49	50	59	77	49	65
Furnas	43	38	34	30	23	21	21	20	46	40	41	37
P. Primavera	167	103	94	41	35	44	25	29	37	56	53	111
Itá	76	100	92	101	159	103	88	134	97	70	61	79
Itaipu	100	72	75	44	73	49	42	64	57	61	50	128
Salto Caxias	79	95	94	102	170	109	91	95	82	73	52	76

Um outro ponto de destaque em ambos os modelos do tipo MS-PAR(p) está nos períodos onde se sobressaem em relação ao modelo PAR(p). Conforme apresentado na Tabela 4.3, eventos de LN causam anomalias negativas nas vazões naturais afluentes principalmente na região sul, e de menor intensidade em parte das regiões sudeste/centro-oeste e em parte da região nordeste. A região norte, principalmente em sua porção superior, sofre uma redução nas aflúências durante eventos de EN. A Tabela 4.14 apresenta a mesma avaliação feita anteriormente para o EMPA, porém dividindo-se entre os estados de ENOS. Como pode ser observado, a UHE Santo Antônio Jari possui melhor desempenho durante eventos de EN com os modelos MS-PAR(p)1 e MS-PAR(p)2, apresentando uma redução de aproximadamente 4% e 3%, respectivamente, com relação ao modelo PAR(p). O mesmo comportamento ocorre na UHE Estreito Tocantins. No que concerne as UHEs localizadas no restante do Brasil, principalmente as que se localizam na região sul, os modelos MS-PAR(p)1 e MS-PAR(p)2 produzem, em sua maioria, melhores resultados em períodos de LN. Em alguns casos específicos, essa redução chega a mais de 10% em Itá e Salto Caxias. As UHEs das regiões sul e sudeste também apresentam melhor desempenho em estados neutros.

Tabela 4.14: EMPA dividido entre os estados de ENOS para a geração de cenários no período de ajuste dos modelos

	La Niña						Neutro			El Niño		
	P		MSP1	MSP2	P		MSP1	MSP2	P		MSP1	MSP2
Sobradinho	35,82	29,04	31,74	28,43	29,78	29,62	33,32	34,16	33,45			
Belo Monte	23,99	22,61	24,26	24,26	24,13	23,90	28,88	28,72	28,68			
Estreito Toc.	24,16	21,64	23,64	24,31	25,22	24,65	26,12	24,10	24,31			
Jirau	15,11	13,13	13,88	12,74	13,55	12,91	14,39	14,67	14,64			
Sto. Ant. Jari	34,34	37,50	38,99	32,55	33,05	33,76	44,10	40,17	41,37			
Teles Pires	22,04	21,66	22,76	17,38	18,64	18,03	20,51	21,14	22,05			
Capivara	56,72	53,95	54,40	63,64	54,82	61,67	60,20	70,13	66,55			
Furnas	34,06	37,03	37,19	27,78	27,39	29,06	31,84	35,62	35,41			
P.Primavera	61,14	60,41	60,57	73,36	68,32	71,33	61,98	70,39	63,28			
Itá	86,88	74,90	77,02	113,40	100,86	109,73	95,54	97,23	93,62			
Itaipu	62,49	57,82	60,03	81,10	75,51	74,11	62,82	67,13	65,23			
Salto Caxias	81,29	69,45	79,86	107,02	93,16	104,40	83,88	99,82	86,48			

Para tornar a análise dos resultados apresentados na Tabela 4.14 mais prática do ponto de vista da geração de cenários sintéticos, expõe-se, a seguir, algumas figuras que trazem a distribuição dos cenários gerados por cada modelo em alguns dos anos simulados. O gráfico exibido é do tipo *fanchart*, onde cada faixa de cor representa uma concentração percentual das séries sintéticas. Quanto mais forte a cor, menor é a faixa que representa as séries geradas. A linha tracejada em preto simboliza a mediana da distribuição, e a linha cheia com marcadores em preto os valores históricos. Para facilitar a visualização dos gráfico, de forma a diminuir sua escala, serão apresentados 90% dos cenários, excluindo-se, assim, os mais extremos. No contexto da geração de cenários, espera-se que os valores observados estejam pelo menos contidos em alguma das faixas.

Selecionou-se, para ser analisado primeiramente, o período entre janeiro de 1953 e dezembro de 1956. Esses anos correspondem à parte central do período crítico brasileiro, período onde o SIN passou pela pior seca multivariada que se tem registro (a partir do ano de 1931). O ano de 1953 iniciou com um estado de EN que persistiu até fevereiro de 1954, tendo, em seguida, a formação de uma LN a partir de maio, que perdurou até setembro de 1956. Para esse primeiro subconjunto de períodos, serão avaliadas as gerações de cenários da UHE Santo Antônio Jari, localizada na região norte do Brasil, e da UHE Salto Caxias, localizada na região sul, observando-se, assim, a variação dos regimes hidrológicos durante o período crítico.

A Figura 4.23 apresenta os cenários gerados no período supracitado para a UHE Santo Antônio Jari com os modelos PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) e sua distribuição histórica (d). De maneira geral, é possível afirmar que as três gerações realizadas conseguiram capturar corretamente o comportamento observado na série histórica, porém há diferenças importantes que devem ser destacadas. Observa-se que, no primeiro ano, a série histórica da UHE Santo Antônio Jari possui valores bem abaixo dos registrados nos anos subsequentes devido à presença do fenômeno EN. Analisando-se os cenários gerados pelos três modelos, é possível perceber que, para o ano de 1953, o modelo PAR(p) gera uma nuvem de cenários mais úmidos no mês de maior vazão em comparação com os modelos MS-PAR(p). Além disso, os dois últimos modelos conseguem capturar o comportamento desse ano, mesmo com uma variabilidade menor das séries sintéticas geradas. Conforme apontado na Figura 4.14, os períodos de EN tendem a ser mais secos, e possuem uma incerteza menor. O contrário ocorre na LN, conforme verificado nos anos subsequentes entre 1954 e 1956. Em 1955, observa-se um decréscimo brusco da vazão afluente de abril para maio que não conseguiu estar presente em qualquer dos cenários mostrados do modelo PAR(p), porém capturados pelos modelo MS-PAR(p) devido a maior variabilidade dos cenários gerados. Para essa UHE, não existem grandes diferenças entre a geração de cenários realizada pelos modelos MS-PAR(p) entre si,

pois, de acordo com o teste de influência do ENOS apresentado em 4.12, a maior parte dos meses sofrem efeitos de ambas as fases.

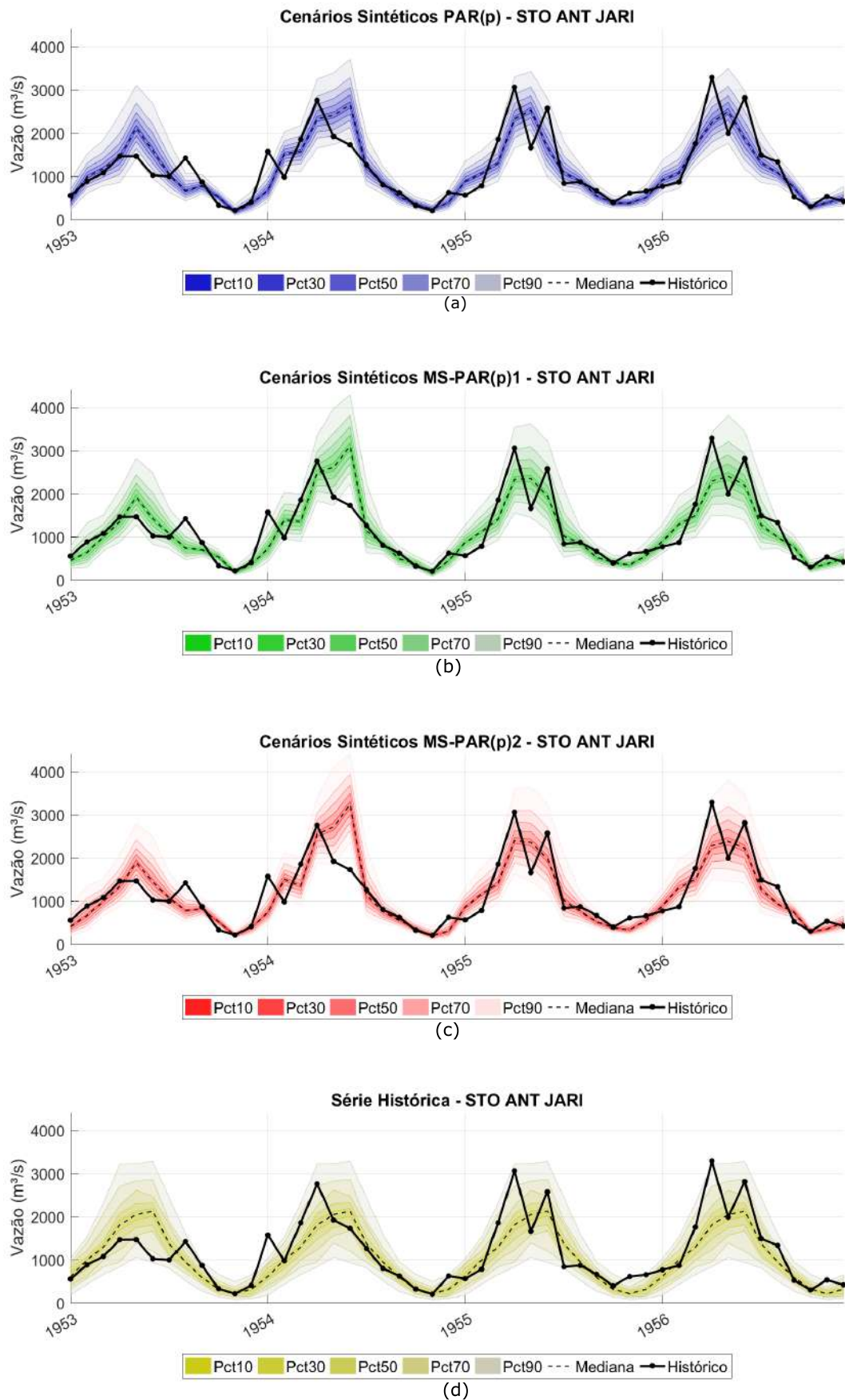


Figura 4.23: Cenários sintéticos de vazões gerados para o período entre 1953 a 1956 na UHE Santo Antônio Jari com os modelos PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) e sua distribuição histórica (d).

São apresentadas, na Figura 4.24 (a), (b) e (c), as gerações de cenários sintéticos de vazões afluentes a UHE Salto Caxias realizadas pelos modelos PAR(p), MS-PAR(p)1 e MS-PAR(p)2, respectivamente. Primeiramente, é possível observar que há uma variabilidade grande na série observada em Salto Caxias, característica marcante das UHEs da região sul. Mais uma vez, nota-se que os três modelos são capazes de gerar cenários de maneira aderente ao histórico, com algumas diferenças entre si. Ao contrário do que ocorre na UHE Santo Antônio Jari, na UHE Salto Caxias, a geração do modelo MS-PAR(p)2 é mais próxima ao modelo PAR(p). Conforme observou-se na Figura 4.2, apenas o estado de EN em março e novembro é identificado como influente. Analisando-se o ano de 1953, é possível observar que o modelo MS-PAR(p) gera uma nuvem de cenários com uma variabilidade maior que nos outros dois modelos por conta do estado de EN. A partir de 1954, já no estado de LN, sua geração passa a possuir uma dispersão menor, sem comprometer a aderência aos valores observados.

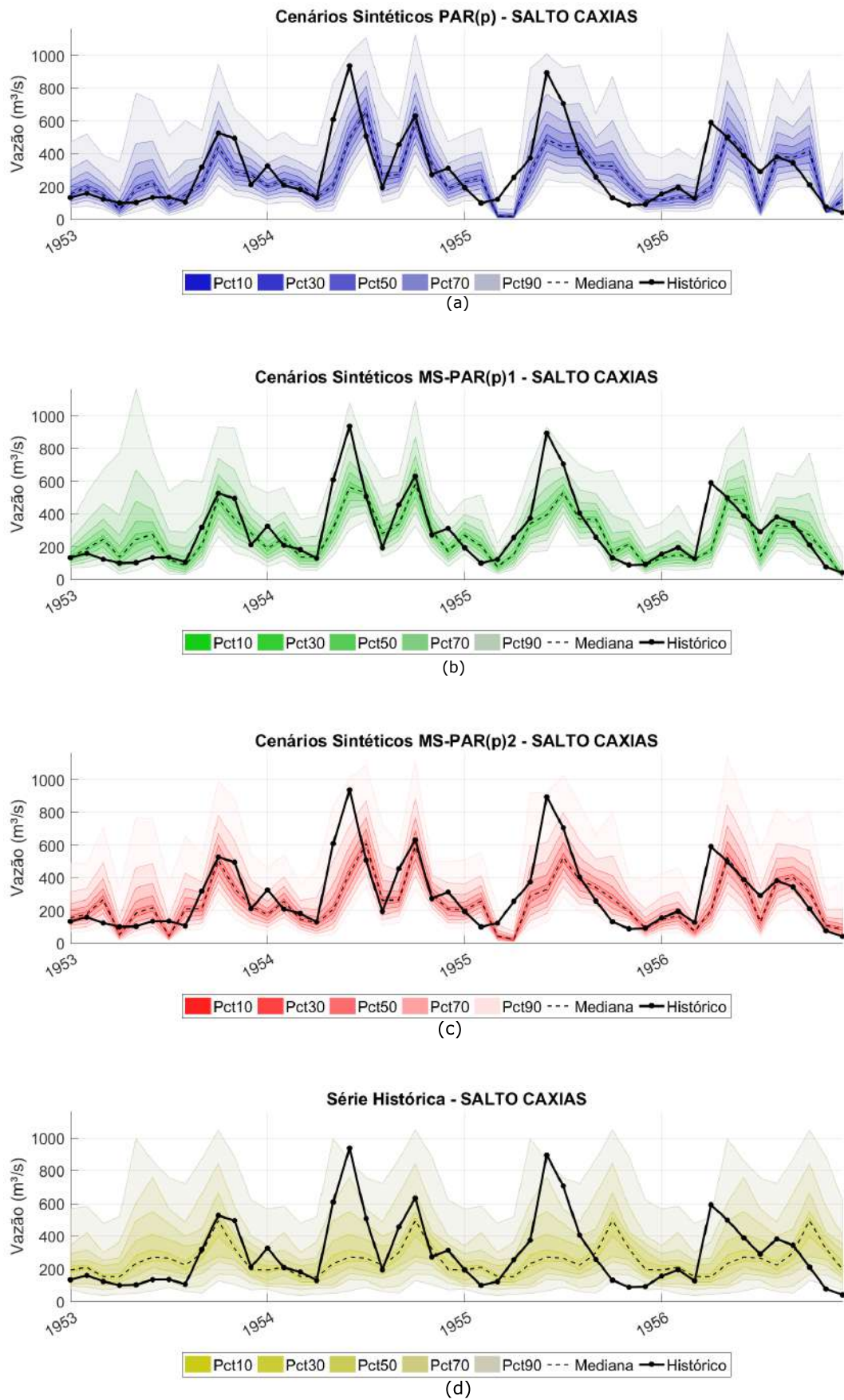


Figura 4.24: Cenários sintéticos de vazões gerados para o período entre 1953 a 1956 na UHE Salto Caxias com os modelos PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) e sua distribuição histórica (d).

Os resultados exibidos, nesta seção, demonstram que os modelos MS-PAR(p) são melhores quando avaliados segundo as métricas EMA, EMPA e REQM em comparação com o modelo PAR(p). A diferença entre os resultados é maior durante o período úmido e durante os períodos onde as vazões afluentes estão abaixo da MLT, em sua maior parte períodos de EN na região para UHEs da região norte e LN no restante do Brasil. As UHEs que sofrem uma influência maior do fenômeno ENOS são aquelas em que o modelo MS-PAR(p) se destaca em relação ao modelo PAR(p). Pode-se afirmar, então, que as distribuições de cenários multivariados foram gerados de forma mais aderente aos valores observados no histórico, tendo essa distribuição um desvio padrão menor ou maior, dependendo do estado de ENOS. Esses resultados, porém, devem ser analisados com cautela. Dado que o período simulado estava contido no período de ajuste dos modelos, é esperado que modelos com mais parâmetros ajustados (que é o caso dos modelos MS-PAR(p)) possuam uma performance melhor do que modelos com menos parâmetros. Será apresentada, na seção seguinte, uma avaliação da geração de cenários para o período pós-ajuste (2010 a 2017), de modo a identificar melhor o potencial da metodologia desenvolvida nesta dissertação.

4.5 Caso Condicionado - período pós-ajuste

Esta seção tratará dos casos condicionados realizados no período pós-ajuste dos modelos. O período selecionado para a avaliação corresponde aos anos entre 2010 e 2017. Para cada ano simulado, os modelos serão ajustados até o ano observado anterior. Ou seja, para o ano de 2010, os modelos foram ajustados com o histórico de vazões até 2009; para o ano de 2011, os modelos foram ajustados com os dados até 2010, e assim sucessivamente. Para cada período simulado, a geração será realizada até cinco anos à frente, com o intuito de se avaliar a geração de cenários sintéticos tanto para o curto prazo quanto para o médio/longo prazo. Com relação ao fenômeno ENOS, serão utilizadas as previsões probabilísticas fornecidas pelo IRI, conforme descrito no item 3.4.1, para a geração dos cenários de ENOS.

Sendo assim, antes de se avaliar a geração de cenários de vazões afluentes, dado que os modelos MS-PAR(p) dependem dos estados de ENOS, primeiramente será feita uma avaliação da qualidade dos cenários de ENOS gerados pelos modelos. Durante o horizonte analisado, houve dois eventos de EN: o primeiro, já em sua fase final no começo de 2010; e um forte EN entre o final de 2014 e meados de 2016. Ocorreram, também, quatro eventos de LN: o primeiro entre meados de 2010 até maio de 2011; um segundo evento que pode ser considerado como a continuação do anterior, a partir de julho de 2011 até o princípio de 2012; o terceiro evento uma LN fraca que perdurou apenas nos últimos cinco meses de 2016; e o último evento de

LN que teve formação em outubro de 2017. No restante dos períodos, manteve-se o estado N. Ainda sobre o período analisado, destaca-se que os anos de 2014 e 2015 estão entre os cinco piores anos hidrológicos do histórico.

A Figura 4.25 traz o histórico de ENOS comparado aos cenários gerados pelo modelo MS-PAR(p). Valores positivos (em vermelho) representam a probabilidade de formação de EN quando este valor é maior que sua probabilidade climatológica; valores negativos (em azul) representam (em módulo) a probabilidade de formação de LN quando este valor é maior que sua probabilidade climatológica; valor zero (em verde) representa uma maior probabilidade de ocorrência de estado N. Dados não existentes são representados em branco. A linha mais abaixo do gráfico representa o estado de ENOS observado, enquanto que as linhas subsequentes representam as previsões probabilísticas de ENOS realizadas em diferentes instantes de tempo. Pode-se observar que os cenários gerados de ENOS são bastante aderentes ao que ocorreu de fato. De maneira geral, quanto menor a diferença entre o tempo de previsão e o de observação, maior é o percentual de cenários gerados na categoria correta, assim como o contrário também é verdade. Durante o horizonte analisado, houve dois casos de falso alarme, quando o modelo previu, para a maioria dos cenários, a ocorrência de um EN no final de 2012 e em meados de 2017, fato que acabou não se confirmando. Porém, destaca-se que, mesmo para geração de cenários realizadas para um horizonte de até doze meses à frente, a qualidade das previsões é satisfatória em muitos casos. Por exemplo, observa-se que a maior parte dos cenários gerados em junho de 2010 indicavam a formação de uma LN que persistiria até maio de 2011. O mesmo pode ser dito para o forte EN ocorrido entre o final de 2014 e meados de 2016.

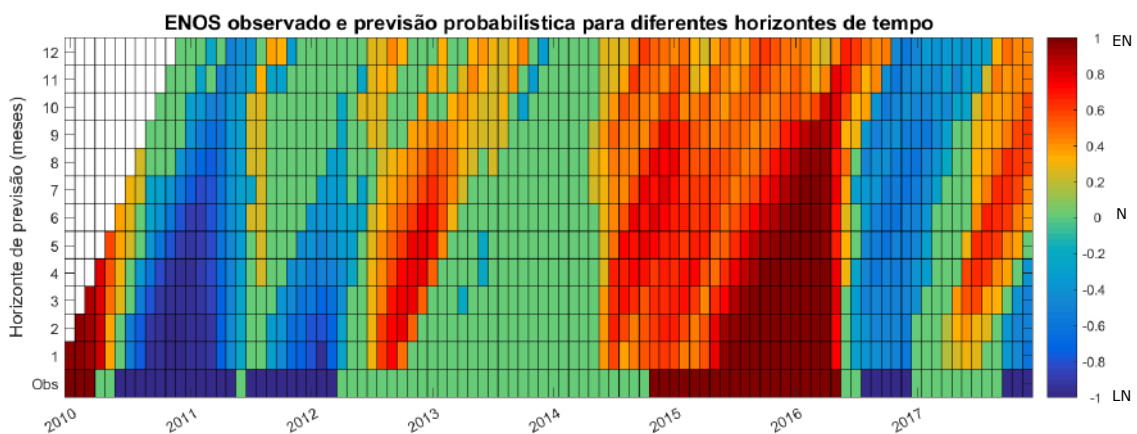


Figura 4.25: Percentual de cenários de ENOS gerados até doze meses à frente para os anos entre 2010 e 2017.

Para quantificar o ganho de se utilizar as previsões probabilísticas do IRI em relação a probabilidade climatológica, foi calculado, para o mesmo horizonte

de geração, a métrica RPSS, conforme apresentada na Equação (3.45). Para essa métrica, valores positivos indicam que a previsão probabilística é melhor do que a utilização da probabilidade climatológica, sendo o valor 1 atribuído para uma previsão perfeita. Apresenta-se, na Figura 4.26, o RPSS para a geração de cenários de ENOS com relação ao início da simulação. Índices menores que zero, representados pela cor branca, indicam que a utilização da climatologia é melhor do que a consideração da previsão probabilística. Valores maiores que zeros, representados pela escala de cores, indicam que é mais vantajoso utilizar as previsões probabilísticas. Mais uma vez, é possível observar a qualidade dos cenários de ENOS gerados, até mesmo para longos períodos de previsão em alguns casos. Nota-se que a qualidade dos cenários, após iniciado um evento de ENOS, é maior do que nos períodos iniciais, o que confere às figuras apresentadas um aspecto de rampa.

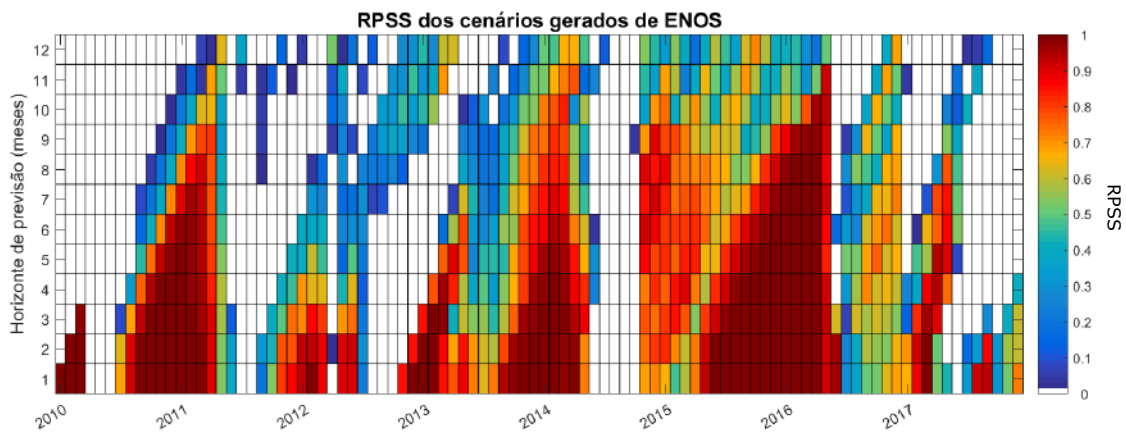


Figura 4.26: RPSS em função do período observado (eixo horizontal) e início da previsão (eixo vertical) para os anos entre 2010 e 2017.

De forma a sintetizar o resultado obtido, apresenta-se a mediana do RPSS para cada mês em relação ao horizonte de previsão na Figura 4.27. Optou-se por apresentar a mediana, pois, em alguns poucos períodos ocorreu, um valor de RPSS demasiadamente negativo, o que acabaria levando a um entendimento errado da qualidade da geração de cenários de ENOS caso fosse apresentada a média. Observa-se que, para a maioria dos meses, a previsão probabilística é bastante superior nos primeiros meses, decaindo conforme a geração se afasta do tempo de observação. Além disso, os meses entre junho e agosto apresentam-se com os menores valores de RPSS para a geração um mês à frente. Esse resultado vai ao encontro de outras pesquisas citadas nesta dissertação sobre a dificuldade de se prever os eventos de ENOS durante esse período. Nos demais meses, o RPSS oscila de valores próximos a 1,0 até 0,5 no horizonte de oito meses, o que confirma a qualidade dos cenários gerados.

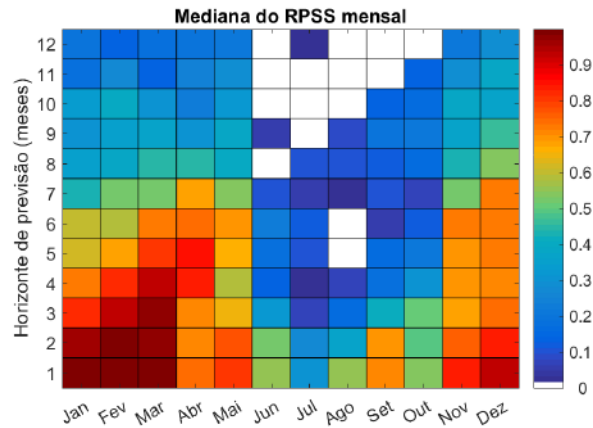


Figura 4.27: Mediana do RPSS mensal com relação ao horizonte de previsão para os anos entre 2010 e 2017.

Passada a verificação da geração de cenários de ENOS, será apresentada uma análise da geração de cenários de vazões afluentes dos três modelos avaliados nesta dissertação. Para cada mês entre janeiro de 2010 a dezembro de 2017, foram gerados 2000 cenários sintéticos com cinco anos de horizonte. As métricas EMA, EMPA e REQM foram calculadas para até um horizonte de seis meses à frente em cada simulação, visando avaliar a geração além do curto prazo. Apresentam-se, na Tabela 4.15, os resultados médios obtidos para a geração de cenários um mês à frente, para as treze UHEs selecionadas. É possível observar que, de maneira geral, os modelos MS-PAR(p) possuem resultados melhores nas regiões norte e sul nas três métricas analisadas. A única exceção é a UHE Itaipu, que apresenta resultados melhores no modelo PAR(p). De maneira geral, observa-se uma predominância do modelo MS-PAR(p)2 sobre o modelo MS-PAR(p)1.

Esse resultado é exatamente o oposto do obtido quando foi feita a avaliação da geração no período pré-ajuste, conforme demonstrado na Tabela 4.10. Esse fato pode ser uma indicação de que o modelo MS-PAR(p)1, em muitos casos, realiza uma parametrização excessiva no ajuste para algumas UHEs, o que acaba introduzindo mais ruído no modelo. Porém, em alguns casos, esse tipo de ajuste apresenta resultados melhores, principalmente na região norte. O modelo MS-PAR(p)2, através da identificação de influência do ENOS, consegue, por muitas vezes, produzir resultados melhores, principalmente, para as UHEs da região sul. A região sul possui melhores resultados com a geração de cenários realizada pelo modelo MS-PAR(p), diminuindo em 0,7% o EMPA em relação ao modelo PAR(p), havendo melhoria em exatamente metade das UHEs que a compõe. Apesar da UHE Capivara possuir melhores resultados com o modelo PAR(p), 80% das UHEs (de um total de dez) pertencentes à bacia hidrográfica do rio Paranapanema possuem melhores resultados com os modelo MS-PAR(p). No restante da região sudeste, o modelo PAR(p) possui resultados superiores.

Tabela 4.15: Avaliação da geração de cenários no período pós-ajuste para os anos entre 2010 e 2017

UHE	Região	EMA						EMPA			REQM		
		P		MSP1		MSP2		P		MSP1		MSP2	
		P	MSP1	MSP2	MSP1	MSP2	P	MSP1	MSP2	P	MSP1	MSP2	
Sobradinho	Nordeste	545,87	604,09	584,62	59,20	57,34	53,32	59,20	57,34	720,49	788,93	765,53	
Belo Monte	Norte	1958,09	1851,90	1905,20	26,29	26,98	27,22	26,29	26,98	2472,78	2387,70	2416,12	
Estreito Toc.	Norte	428,75	422,14	426,81	27,34	27,14	27,15	27,34	27,14	554,00	549,15	552,43	
Jirau	Norte	2667,69	2710,61	2659,78	19,49	19,16	19,49	19,55	19,16	3075,92	3146,19	3072,86	
Sto. Ant. Jari	Norte	335,43	319,08	322,31	34,07	34,85	34,78	34,07	34,85	402,17	389,04	392,77	
Teles Pires	Norte	450,96	432,70	444,94	20,19	19,68	20,19	19,49	19,68	538,53	524,70	535,94	
Capivara	Sudeste	241,45	249,06	244,62	49,12	48,70	47,13	49,12	48,70	298,76	309,51	301,97	
Furnas	Sudeste	184,30	200,61	182,13	58,10	52,95	51,89	58,10	52,95	230,17	252,48	236,66	
P. Primavera	Sudeste	358,86	357,61	352,74	34,29	33,40	34,29	33,94	33,40	436,08	442,68	429,76	
Itá	Sul	276,06	277,84	270,97	121,83	119,39	121,83	124,50	119,39	339,01	344,25	336,48	
Itaipu	Sul	1309,84	1351,54	1407,98	39,60	42,19	39,60	41,75	42,19	1507,54	1565,08	1619,13	
Salto Caxias	Sul	261,51	255,47	255,11	86,29	83,59	86,29	83,37	83,59	318,94	315,75	314,38	

Será analisada, a seguir, a distribuição dos cenários gerados para algumas das UHEs de forma a compreender melhor as diferenças entre a geração realizada dos três modelos. As Figuras 4.28 e 4.29 expõem as séries sintéticas geradas um mês à frente para a UHE Belo Monte entre 2010 e 2013, e 2014 a 2017, respectivamente. De maneira geral, ambos os três modelos geraram nuvens de cenários adequados para o a UHE Belo Monte. Destacam-se, desse período, os anos de 2010, 2015 e 2016, que foram os de ocorrência dos piores períodos úmidos para a UHE Belo Monte entre 2010 e 2017. Não por coincidência, em ambos os períodos estava em atividade o estado EN. Observa-se que, também, nos três períodos, o modelo MS-PAR(p)1 consegue gerar cenários mais próximos do valor observado, principalmente no mês de abril, que é quando, em geral, ocorrem as maiores vazões afluentes na UHE Belo Monte. Em abril de 2015, apenas o modelo MS-PAR(p)1 consegue gerar cenários de mesma magnitude da vazão observada, que foi extremamente baixa em relação à MLT.

O motivo pelo qual o modelo MS-PAR(p)2 não conseguiu capturar o mesmo comportamento está no fato de não ter sido selecionado, para o mês de abril, qualquer estado de ENOS como influente na UHE Belo Monte, conforme foi apresentado na Tabela 4.12. Caso houvesse sido selecionado, provavelmente o modelo MS-PAR(p)2 teria gerado cenários mais aderentes no mês de abril, demonstrando que, para alguns casos, o processo de identificação de influência sugerido pode não ser o ideal. É importante destacar que, para abril, os três modelos identificam um AR(1) com um coeficiente autorregressivo de aproximadamente 0,15. Ou seja, as vazões afluentes do mês de abril são fracamente correlacionadas com a vazão do mês anterior. O uso da informação climática de ENOS, nesse caso, trouxe um claro benefício à geração de cenários, trazendo relevante informação ao modelo.

As Figuras 4.30 e 4.31 trazem a geração de cenários um passo à frente realizada para a UHE Itá. É possível notar que a série observada durante o período simulado possui um comportamento bastante errático, principalmente no ano de 2015, quando ocorreu um forte EN. Ainda assim, pode-se afirmar que os três modelos conseguem gerar cenários próximos aos valores observados.

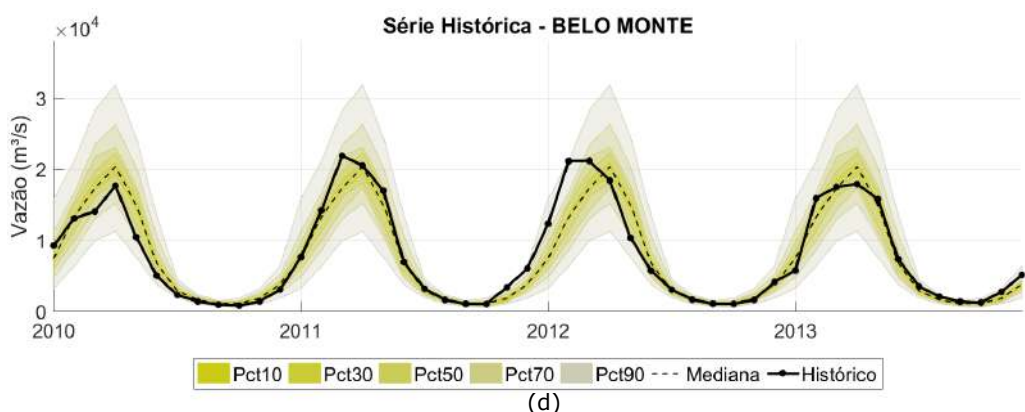
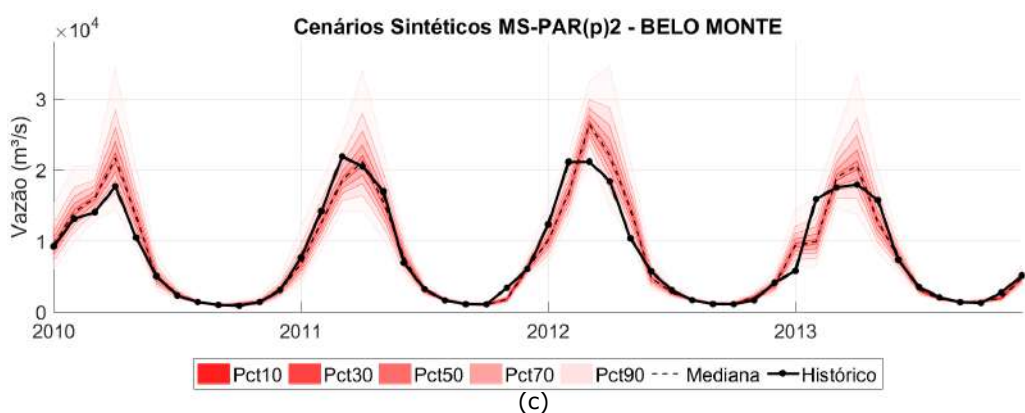
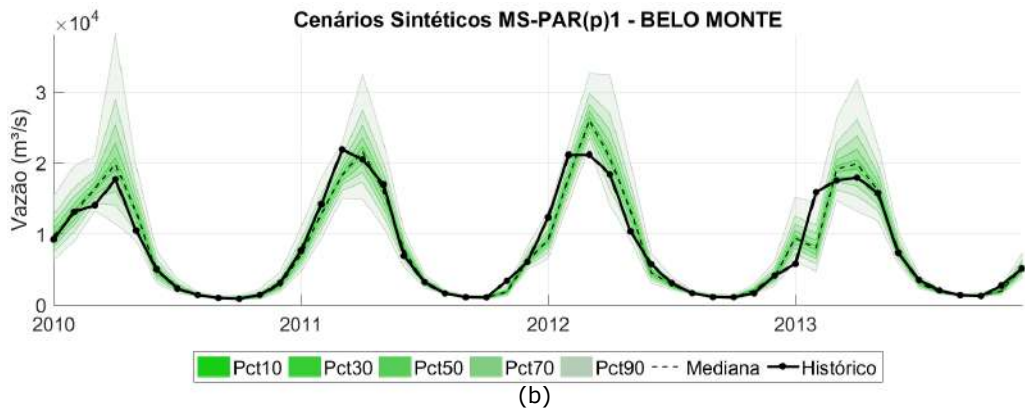
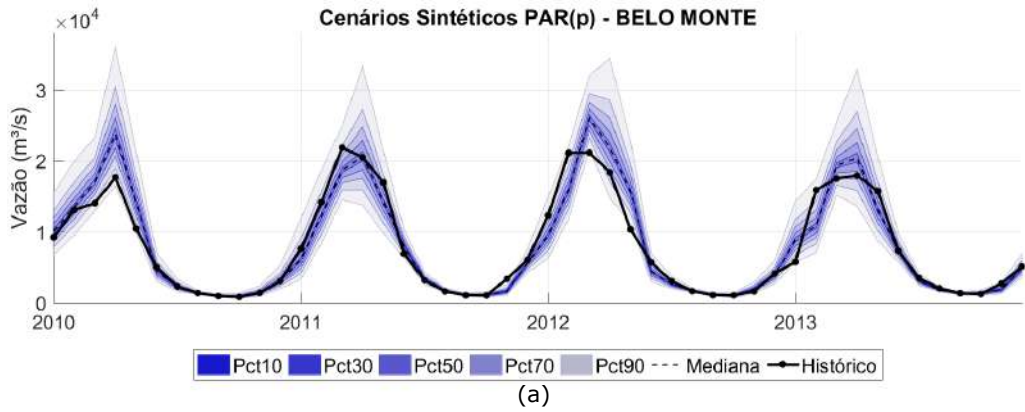


Figura 4.28: Cenários sintéticos gerados para o período entre 2010 a 2013 na UHE Belo Monte com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).

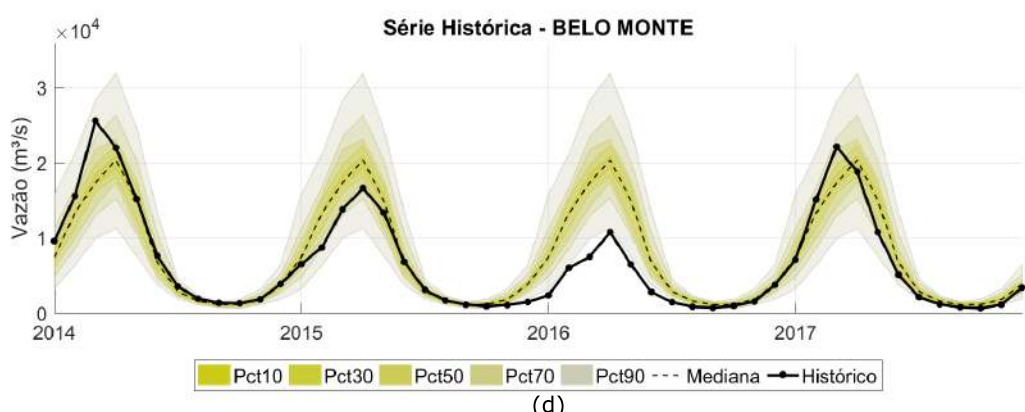
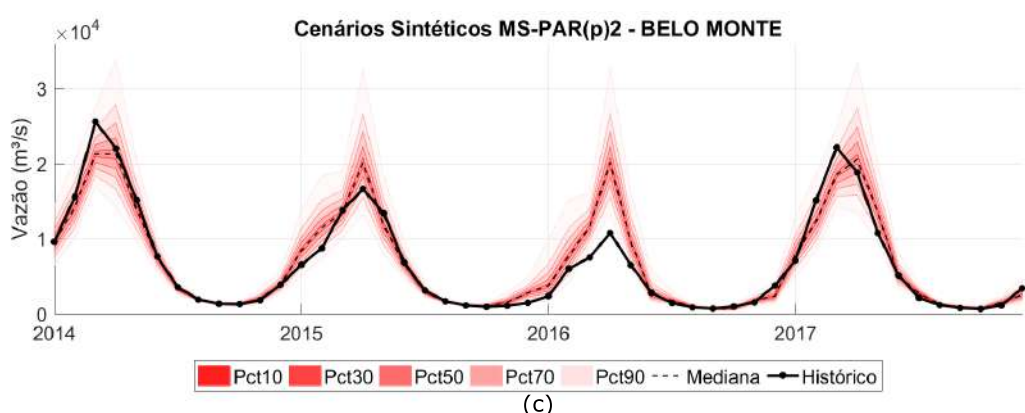
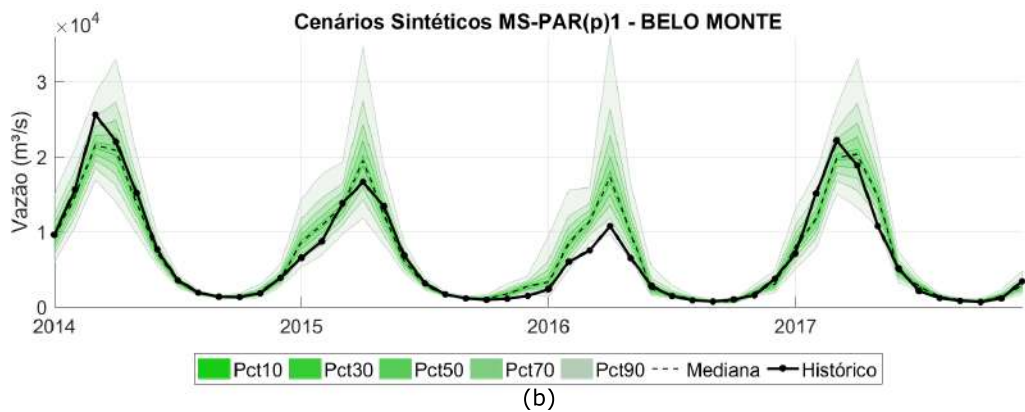
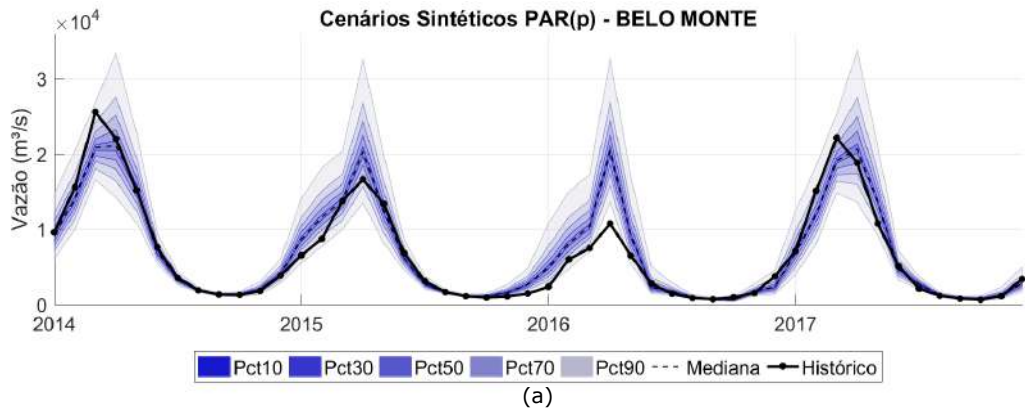


Figura 4.29: Cenários sintéticos gerados para o período entre 2014 a 2017 na UHE Belo Monte com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).

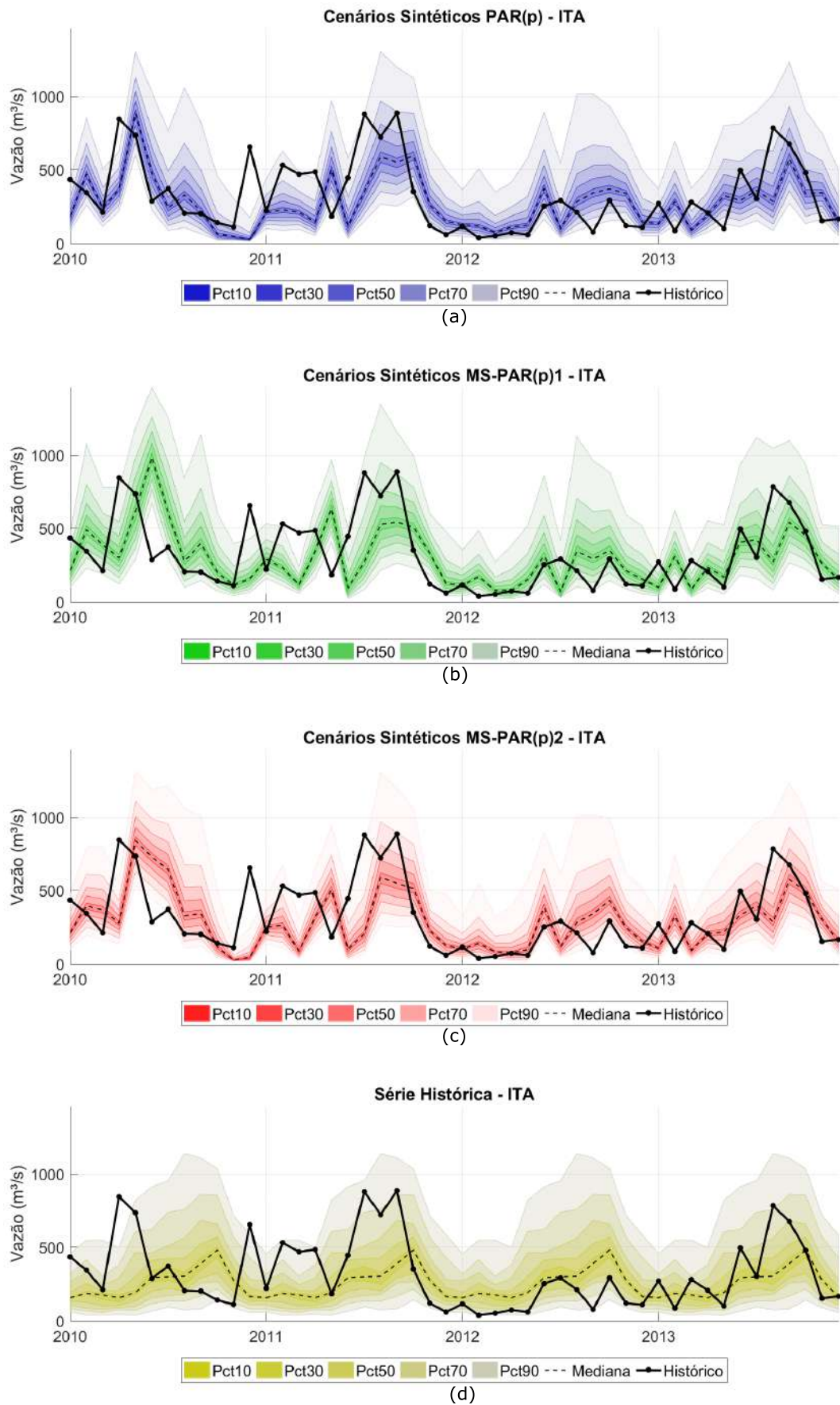


Figura 4.30: Cenários sintéticos gerados para o período entre 2010 a 2013 na UHE Itá com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).

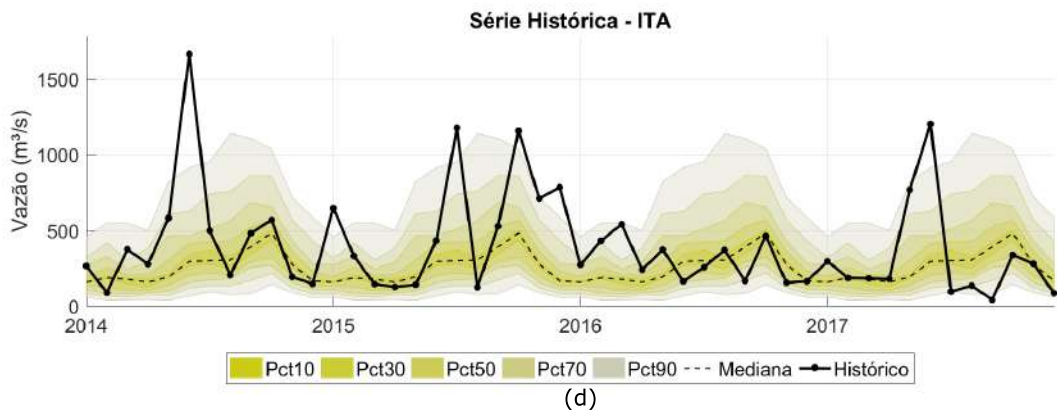
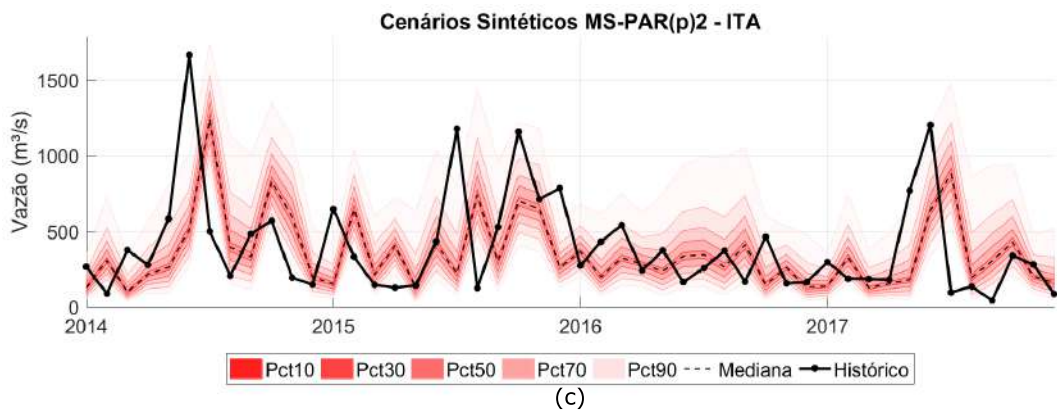
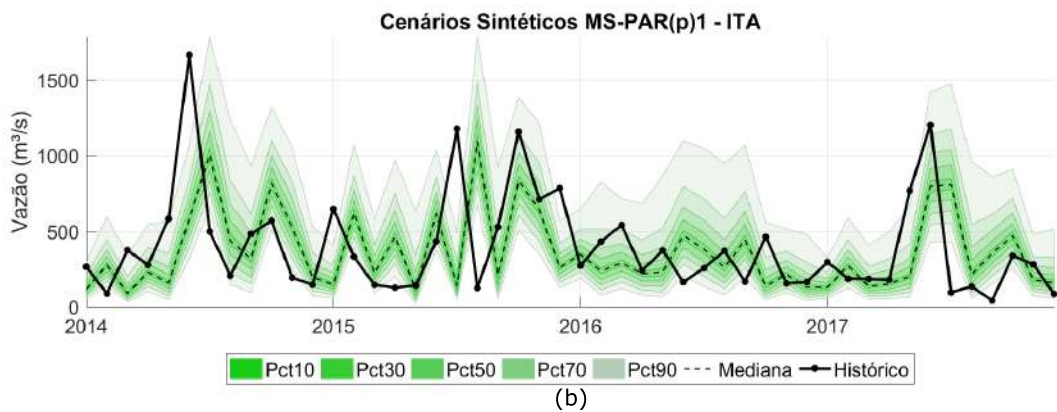
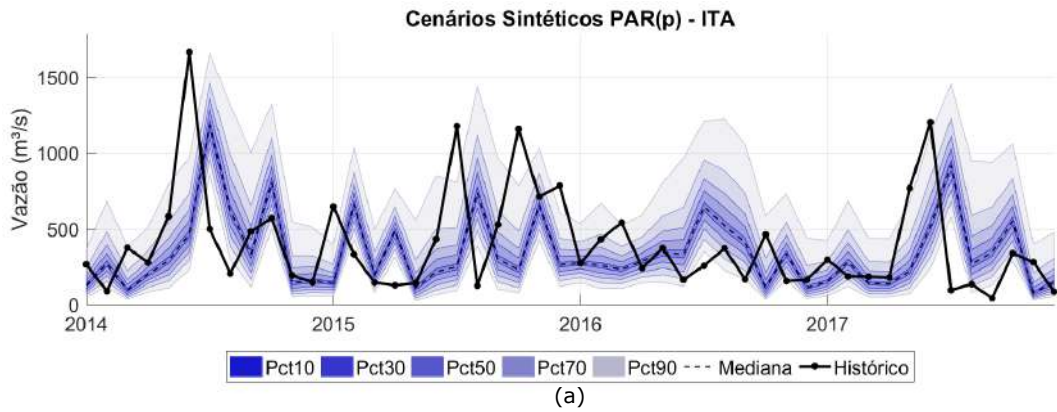


Figura 4.31: Cenários sintéticos gerados para o período entre 2014 a 2017 na UHE Itá com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).

Como para cada caso simulado foi feita uma geração de cenários sintéticos por cinco anos, de forma a avaliar a geração para o médio prazo, será feita uma avaliação dos erros médios acumulados por seis meses consecutivos, apresentada na Tabela 4.16. Como pode ser observado, os resultados são próximos aos observados na Tabela 4.15, indicando uma superioridade dos modelos MS-PAR(p) nas regiões sul e norte, e o modelo PAR(p) se sobressaindo na região sudeste, e nas UHEs Sobradinho e Itaipu. Comparando-se os resultados individualmente para cada UHE, o modelo MS-PAR(p)2 apresenta EMA médio menor que o modelo PAR(p) em 53% das UHEs, ao passo que para o modelo MS-PAR(p)1 apenas 31% das UHEs tiveram um EMA médio menor que o modelo PAR(p).

Na região sul, de um total de 31 UHEs, o aprimoramento ocorreu em 77,5% e 58% no modelos MS-PAR(p)2 e MS-PAR(p)1, respectivamente, demonstrando, mais uma vez, a sua superioridade nessa localidade. Destacam-se, também, os resultados obtidos para a bacia do rio Paranapanema, cujo aprimoramento ocorreu em 70% das UHEs com o modelo MS-PAR(p)2 e 60% com o modelo MS-PAR(p)1 de um total de 10 UHEs. Já na região norte, esses índices foram de 50% e 25%, respectivamente. Porém, é necessário salientar que ocorreu um aprimoramento para as UHEs Belo Monte e Jirau, que são duas das maiores produtoras de energia elétrica do Brasil.

Para exemplificar a análise da geração de médio e longo prazo, a Figura 4.32 apresenta os três anos de cenários gerados em janeiro de 2015 para os três modelos e sua distribuição histórica para a UHE Estreito Tocantins. O ano de 2015 caracterizou-se por um forte EN, que, por sua vez, provoca anomalias negativas em UHEs localizadas na região norte do Brasil. É possível observar que, principalmente nos quatro primeiros meses de 2015, os cenários gerados com os modelo MS-PAR(p) estão mais próximo dos valores observados. Com o passar do tempo, os modelos acabam perdendo a memória do passado recente e passam a produzir cenários próximos à distribuição histórica.

Tabela 4.16: Avaliação da simulação no período pós-ajuste acumulados em 6 meses para os anos entre 2010 e 2017

UHE	Região	EMA						EMPA						REQM						
		P		MSP1		MSP2		P		MSP1		MSP2		P		MSP1		MSP2		
Sobradinho	Nordeste	593,15	653,85	636,06	66,36	73,88	71,30	787,46	862,14	835,76										
Belo Monte	Norte	2269,76	2197,11	2226,30	35,16	33,94	34,79	2807,11	2752,03	2755,91										
Estreito Toc.	Norte	521,12	507,02	505,53	35,09	34,39	34,44	650,71	636,12	633,75										
Jirau	Norte	4038,58	4095,93	4024,35	28,95	29,08	28,76	4632,84	4703,44	4618,43										
Sto. Ant. Jari	Norte	404,68	396,62	397,07	43,42	43,06	43,87	480,99	475,81	477,07										
Teles Pires	Norte	529,65	527,53	528,89	23,90	24,11	23,89	634,98	634,98	636,29										
Capivara	Sudeste	266,28	275,57	273,94	54,47	56,32	56,11	337,14	337,14	331,31										
Furnas	Sudeste	221,03	239,43	232,65	62,50	72,03	70,16	297,68	297,68	292,48										
P. Primavera	Sudeste	383,00	408,79	377,03	37,26	38,91	36,41	474,39	519,80	468,20										
Itá	Sul	283,27	281,54	278,98	131,34	132,12	130,84	352,54	352,31	349,79										
Itaipu	Sul	1431,01	1474,48	1502,23	45,60	47,62	47,78	1659,97	1717,33	1740,61										
Salto Caxias	Sul	279,65	275,37	273,75	103,69	104,06	101,55	343,82	343,04	339,49										

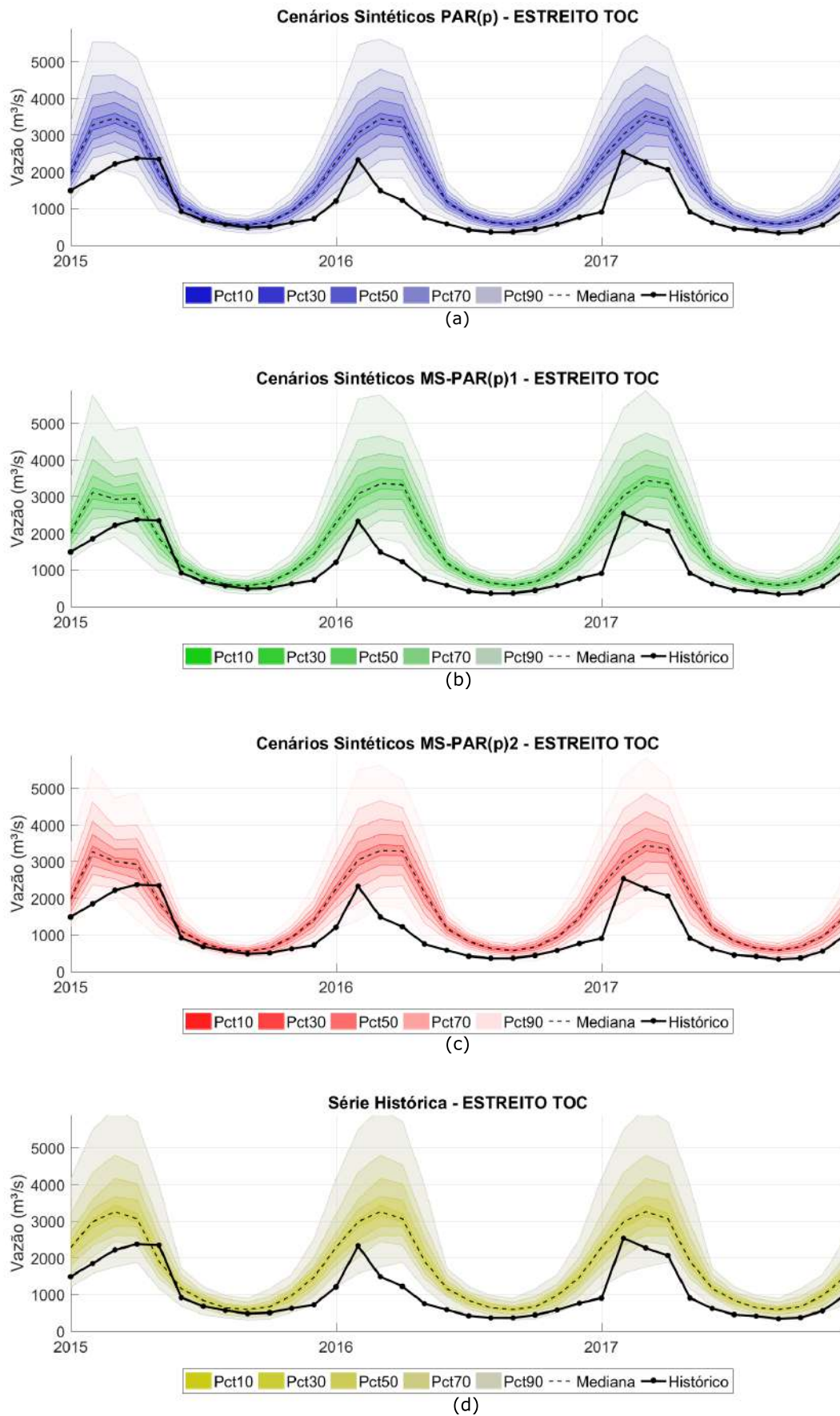


Figura 4.32: Cenários sintéticos gerados de forma sequencial para o período entre 2015 a 2017 na UHE Estreito Tocantins com os modelos PAR(p) (a), MS-PAR(p)1 (b) e MS-PAR(p)2 (c) e sua distribuição histórica (d).

Os resultados expostos pelos casos condicionados indicam alguns pontos importantes que devem ser destacados. Nas regiões mais afetadas pelo fenômeno ENOS, a geração de cenários realizada pelos modelos MS-PAR(p) ocorre de forma mais aderente aos valores históricos observados em comparação ao modelo PAR(p). Já a região sudeste, que, em diversas pesquisas citadas nesta dissertação, apontam para um comportamento de transição entre o dipolo norte/sul dos efeitos causados pelo ENOS, não apresentam um padrão claro de aprimoramento ao se utilizar a modelagem proposta. Outro ponto que deve ser destacado diz respeito as UHEs Sobradinho e Itaipu. Ambas UHEs possuem uma das maiores áreas de drenagem do SIN, e também não produziram resultados melhores com os modelos MS-PAR(p). Os efeitos ao longo de suas bacias de drenagem tendem a ser difusos, o que poderia ser uma explicação para o fato de UHEs que possuem áreas de drenagem extensas não apresentarem melhores resultados com os modelos MS-PAR(p).

Os modelos MS-PAR(p) foram capazes de aprimorar a geração de cenários tanto no curto prazo (um mês à frente) quanto na análise de médio prazo (até seis meses à frente), principalmente nas regiões citadas anteriormente. É importante ressaltar que houve inclusive uma melhora dos resultados quando se avaliou um horizonte mais longo. Pode-se afirmar que uma das vantagens do modelo MS-PAR(p) é conseguir identificar estados diferentes em um determinado mês, indicando períodos com médias e incertezas que podem ser maiores ou menores do que o histórico, o que acaba beneficiando a geração de cenários no médio prazo. Como os eventos de EN e LN tendem a persistir por pelo menos cinco meses, o modelo MS-PAR(p) acaba se beneficiando dessa característica no médio / longo prazos.

Capítulo 5

Conclusões e Desenvolvimentos Futuros

O Brasil possui um sistema de geração de energia elétrica hidrotérmico de proporções continentais, no qual a energia hidroelétrica é predominante, e, assim, representa a principal fonte de incerteza para o planejamento do SIN. A tomada de decisão na operação dos reservatórios acaba por influenciar os instantes futuros, representando um acoplamento temporal das decisões. Existem ainda acoplamentos espaciais da operação, pelo fato de ações operativas de UHEs a montante influenciarem na disponibilidade de recursos de todas as UHEs a jusante da bacia hidrográfica.

Durante a operação do SIN, deve-se levar em conta critérios de segurança energética e modicidade tarifária, dois objetivos opostos. Sendo assim, o planejamento da operação do SIN é um complexo problema de otimização de recursos com múltiplos objetivos, que deve ser solucionado em tempo real, todavia observando-se o longo prazo. De forma a lidar com esse problema, o planejamento da operação é dividido em diferentes etapas, com distintos horizontes e detalhamentos do SIN e das incertezas associadas. Para o tratamento da principal fonte de incerteza do planejamento da operação de curto, médio e longo prazo, as vazões afluentes às UHEs, utiliza-se o modelo estocástico PAR(p) para a geração de séries sintéticas adotado pelo programa GEVAZP. Esta dissertação teve como objetivo propor um aprimoramento metodológico a esse modelo, de forma a incluir informações climáticas referentes ao fenômeno El Niño - Oscilação Sul (ENOS).

O ENOS é um dos principais fenômenos climáticos naturais, produzindo diversos impactos nos padrões de circulação global da atmosfera. Especificamente no Brasil, sua ocorrência impacta de distintas formas e intensidades a precipitação, e, conseqüentemente, as vazões fluviais. Sua consideração na modelagem estocástica das aflúências pode trazer ganhos para o planejamento da operação do SIN. Na metodologia autorregressiva periódica (PAR(p)) as séries sintéticas de aflúências futuras são representadas como uma combinação linear das vazões passadas acresci-

das de um termo aleatório. Uma nova modelagem estocástica foi apresentada nesta dissertação, mais especificamente, o modelo autorregressivo periódico com chaveamento markoviano - MS-PAR(p). Neste modelo, através da subdivisão de cada período considerado (cada mês) em distintos estados, faz-se uma estimação segmentada dos parâmetros utilizados no modelo PAR(p). Para esta dissertação, os estados são representações dos estados de ENOS (La Niña, Neutro e El Niño - LN, N e EL). A média e o desvio padrão, além de variarem com o mês, possuem valores distintos para cada estado de ENOS. A evolução do estado de ENOS é modelada através de uma cadeia de Markov que é ajustada para cada mês. Com o intuito de ajustar um modelo mais parcimonioso, propôs-se um procedimento de identificação de influência do ENOS nas vazões afluentes por meio da aplicação do teste de aderência de Kolmogorov-Smirnov (KS).

Nesta dissertação propôs-se utilizar previsões probabilísticas de ENOS para a geração de cenários de ENOS. Dada as matrizes de transição históricas, os valores do índice ONI anteriores, e as previsões provenientes do IRI, é possível obter matrizes de transição previstas semelhantes às matrizes de transição históricas, reproduzindo a distribuição de probabilidade das previsões probabilísticas, por meio da resolução de um problema de programação linear.

Para avaliar a metodologia proposta, foram comparados, através de estudos de caso, as três metodologias descritas: o modelo PAR(p) tradicional, o modelo MS-PAR(p) sem e com a identificação prévia de influência do ENOS. Utilizou-se a configuração do SIN de janeiro de 2018, possuindo 141 UHEs, nessas avaliações. Primeiramente, realizou-se uma longa simulação não-condicionada ao passado recente, para observar a reprodução de estatísticas históricas e sequências negativas. Ambos os três modelos conseguem reproduzir as estatísticas, observando-se uma ampla vantagem dos modelos MS-PAR(p) em relação ao modelo PAR(p) tradicional. Analisou-se o desempenho dos modelos em casos reais através de diversas simulações condicionadas, tanto no período de ajuste do modelo, quanto no período pós-ajuste. Nos casos realizados no período de pré-ajuste do modelo, observou-se uma clara vantagem do modelo MS-PAR(p)₁, seguido pelo modelo MS-PAR(p)₂ e o PAR(p) em último. A região sul apresentou as maiores reduções dos erros com os modelos MS-PAR(p). Destaca-se que, durante a análise mensal, foi possível observar que os ganhos eram maiores durante os meses do período úmido. Outro relevante ponto apresentado diz respeito aos períodos em que os modelos MS-PAR(p) se sobressaíram. Durante eventos de EN, a região norte possui as menores vazões afluentes, e é também quando o modelo apresenta a maior diferença entre os modelos MS-PAR(p) e o modelo PAR(p).

Nos casos condicionados pós-ajuste, primeiramente foi avaliada a geração de cenários de estados de ENOS, demonstrando-se que os mesmos possuem boa

qualidade até mesmo para horizontes mais extensos. Na análise dos erros obtidos nos cenários de afluência um mês a frente, em média, o modelo PAR(p) obteve o melhor desempenho, seguido pelo modelo MS-PAR(p)2 e, por último, o modelo MS-PAR(p)1. Esse resultado é exatamente o oposto do observado no período pré-ajuste. Fazendo-se uma análise mais segmentada, pode-se observar que o modelo MS-PAR(p)2 obteve resultados melhores nas regiões norte e sul e nas UHEs localizadas na bacia do rio Paranapanema. Já a região sudeste, de maneira geral, obteve os melhores resultados com o modelo PAR(p).

Por fim, avaliaram-se os erros médios acumulados por seis meses com o intuito de se ter uma análise da geração para o médio/longo prazo. Os resultados nessa avaliação corroboraram a potencialidade do modelo MS-PAR(p) para a geração de cenários nas regiões sul e norte. Cerca de 53% das UHEs do SIN apresentaram erros menores na comparação do modelo MS-PAR(p)2 com o modelo PAR(p). Esse índice é superior ao encontrado na avaliação realizada um mês a frente.

O processo de identificação de influência do ENOS, a despeito de ajustar modelos mais parcimoniosos, e com isso trazer alguns benefícios, por algumas vezes, obteve resultados contra-intuitivos ou deixou de incluir a informação para algumas UHEs quando deveria. Aperfeiçoamentos podem ser realizados nessa questão através do desenvolvimento de outros testes estatísticos, ou até mesmo com inclusão de julgamento humano por parte de especialistas.

A comparação entre modelos de geração de cenários sintéticos deve ser realizada com um enfoque diferente da análise de modelos de previsão, onde se visa chegar mais próximo do valor observado. A avaliação realizada nos casos condicionado serve como um indicativo, demonstrando algumas possíveis vantagens do modelo MS-PAR(p) em relação ao PAR(p). Por exemplo, a identificação de períodos mais incertos que outros e a geração de cenários mais aderentes em períodos mais secos que o normal, que são justamente os momentos em que o SIN encontra-se em uma condição hidrológica mais desfavorável, podem representar um avanço para o planejamento da operação energética. Porém, esse tipo de aprimoramento só pode ser verificado através da aplicação da modelagem proposta no modelo de planejamento da operação. Como desenvolvimento futuro, sugere-se a inclusão e a realização de testes dessa nova metodologia nos modelos de planejamento da operação energética.

Referências Bibliográficas

- ACEITUNO, P., 1988, “On the Functioning of the Southern Oscillation in the South American Sector. Part I: Surface Climate”, *Monthly Weather Review*, v. 116, n. 3 (mar), pp. 505–524. ISSN: 0027-0644. doi: 10.1175/1520-0493(1988)116<0505:OTFOTS>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281988%29116%3C0505%3A0TFOTS%3E2.0.CO%3B2>>.
- ALIAT, B., HAMDI, F., 2018, “On Markov-switching periodic ARMA models”, *Communications in Statistics - Theory and Methods*, v. 47, n. 2 (jan), pp. 344–364. ISSN: 0361-0926. doi: 10.1080/03610926.2017.1303734. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/03610926.2017.1303734>>.
- AMBRIZZI, T., DE SOUZA, E. B., PULWARTY, R. S., 2004, “The Hadley and Walker Regional Circulations and Associated ENSO Impacts on South American Seasonal Rainfall”. In: Diaz, H., Bradley, R. (Eds.), *The Hadley circulation: Present, past and future*, Springer, pp. 203–235, Dordrecht. ISBN: 978-1-4020-2944-8. doi: 10.1007/978-1-4020-2944-8_8. Disponível em: <http://link.springer.com/10.1007/978-1-4020-2944-8_8>.
- ANDERSON, R., ROSE, B., OLIVER, L., 2015, “Use of IRI Ensembles to Characterize ENSO Uncertainty in Water Supply Forecasting for the Lower Colorado River in Texas”. In: *Watershed Management 2015*, pp. 79–90, Reston, VA, jul. American Society of Civil Engineers. ISBN: 9780784479322. doi: 10.1061/9780784479322.008. Disponível em: <<http://ascelibrary.org/doi/10.1061/9780784479322.008>>.
- ANDERSON, R., WALKER, D., 2011, “Use of Stochastic Modeling during the 2008 and 2009 Drought on the Lower Colorado River in Texas”. In: *World Environmental and Water Resources Congress 2011*, n. 2010, pp. 3113–3123, Reston, VA, may. American Society of Civil Engineers. ISBN: 9780784411735. doi: 10.1061/41173(414)325. Disponível em: <<http://ascelibrary.org/doi/10.1061/41173%28414%29325>>.

- ANDREWS, E. D., ANTWEILER, R. C., NEIMAN, P. J., et al., 2004, “Influence of ENSO on Flood Frequency along the California Coast”, *Journal of Climate*, v. 17, n. 2 (jan), pp. 337–348. ISSN: 0894-8755. doi: 10.1175/1520-0442(2004)017<0337:IOEOFF>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%282004%29017%3C0337%3AIOEOFF%3E2.0.CO%3B2>>.
- BARNSTON, A. G., CHELLIAH, M., GOLDENBERG, S. B., 1997, “Documentation of a highly ENSO-related sst region in the equatorial pacific: Research note”, *Atmosphere-Ocean*, v. 35, n. 3 (sep), pp. 367–383. ISSN: 0705-5900. doi: 10.1080/07055900.1997.9649597. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/07055900.1997.9649597>>.
- BARNSTON, A. G., TIPPETT, M. K., L’HEUREUX, M. L., et al., 2012, “Skill of Real-Time Seasonal ENSO Model Predictions during 2002–11: Is Our Capability Increasing?” *Bulletin of the American Meteorological Society*, v. 93, n. 5 (may), pp. 631–651. ISSN: 0003-0007. doi: 10.1175/BAMS-D-11-00111.1. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-11-00111.1>>.
- BOX, G. E., COX, D. R., 1964, “An analysis of transformations”, *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 211–252.
- BOX, G. E., JENKINS, G. M., 1970, *Time series analysis: forecasting and control*. Holden-Day.
- BRÖNNIMANN, S., 2007, “Impact of El Niño–southern oscillation on European climate”, *Reviews of Geophysics*, v. 45, n. 3. doi: 10.1029/2006RG000199.
- CAMILLONI, I., BARROS, V., 2000, “The Parana River Response to El Nino 1982 – 83 and 1997 – 98 Events”, *Journal of Hydrometeorology*, v. 1, pp. 412–430.
- CAPOZZOLI, C. R., CARDOSO, A. D. O., FERRAZ, S. E. T., 2017, “Padrões de Variabilidade de Vazão de Rios nas Principais Bacias Brasileiras e Associação com Índices Climáticos”, *Revista Brasileira de Meteorologia*, v. 32, n. 2 (jun), pp. 243–254. ISSN: 0102-7786. doi: 10.1590/0102-77863220006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_{_}arttext{&}pid=S0102-77862017000200243{&}lng=pt{&}tlng=pt>.
- CARDENAS, R., SANDOVAL, C. M., RODRÍGUEZ-MORALES, A. J., et al., 2006, “Impact of climate variability in the occurrence of leishmaniasis

in Northeastern Colombia”, *American Journal of Tropical Medicine and Hygiene*, v. 75, n. 2, pp. 273–277. ISSN: 00029637. doi: 75/2/273[pii].

CAYAN, D. R., REDMOND, K. T., RIDDLE, L. G., 1999, “ENSO and Hydrologic Extremes in the Western United States*”, *Journal of Climate*, v. 12, n. 9 (sep), pp. 2881–2893. ISSN: 0894-8755. doi: 10.1175/1520-0442(1999)012<2881:EAHEIT>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%281999%29012%3C2881%3AEAHEIT%3E2.0.CO%3B2>>.

CHARBENEAU, R. J., 1978, “Comparison of the two- and three-parameter log normal distributions used in streamflow synthesis”, *Water Resources Research*, v. 14, n. 1 (feb), pp. 149–150. ISSN: 00431397. doi: 10.1029/WR014i001p00149. Disponível em: <<http://doi.wiley.com/10.1029/WR014i001p00149>>.

CHEN, W. Y., 1982, “Assessment of Southern Oscillation Sea-Level Pressure Indices”, *Monthly Weather Review*, v. 110, n. 7 (jul), pp. 800–807. ISSN: 0027-0644. doi: 10.1175/1520-0493(1982)110<0800:AOSOSL>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281982%29110%3C0800%3AAOSOSL%3E2.0.CO%3B2>>.

COLONESE, J. P., 2016. “Uso de Dados de Precipitação em Modelos de Vazões Diárias Utilizados no Planejamento da Operação do Sistema Interligado Nacional”. Dissertação (Engenharia Civil), Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

COSTA, F., RAUPP, I., DAMÁZIO, J., et al., 2014, “The methodologies for the flood control planning using hydropower reservoirs in Brazil”. In: *6th Internacional Conference on Flood Management*, 09.

DE CARVALHO, A. R. L., 2001. “Previsão de Afluências Diárias para Horizontes de 1 a 7 dias através de uma abordagem estocástica - Modelo Previvazh”. Dissertação (Engenharia Civil), Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

DE FARIAS, W. C. M., 2012. “Aplicação de um modelo autorregressivo multivariado para feação de cenários de afluências incorporando informações climáticas”. Monografia (Engenharia Ambiental), Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

DETTINGER, M., CAYAN, D., MCCABE, G., et al., 2000, “Multiscale streamflow variability associated with El Niño/Southern Oscillation”. In: Diaz, H.,

Markgraf, V. (Eds.), *El Niño and the Southern Oscillation—Multiscale Variability and Global and Regional Impacts*, Cambridge University Press, pp. 113–146.

DINIZ, A., SANTOS, T., CABRAL, R., et al., 2018, “Short/mid-term hydrothermal dispatch and spot pricing for large-scale systems - the case of Brazil”. In: *20th Power Systems Computation Conference - PSCC*, Dublin, Ireland, June.

DOS SANTOS, T. N., DINIZ, A. L., 2009, “A new multiperiod stage definition for the multistage benders decomposition approach applied to hydrothermal scheduling”, *IEEE Transactions on Power Systems*, v. 24, n. 3, pp. 1383–1392. ISSN: 08858950. doi: 10.1109/TPWRS.2009.2023265.

ELSNER, J. B., BOSSAK, B. H., NIU, X. F., 2001, “Secular changes to the ENSO-U.S. hurricane relationship”, *Geophysical Research Letters*, v. 28, n. 21 (nov), pp. 4123–4126. ISSN: 00948276. doi: 10.1029/2001GL013669. Disponível em: <<http://doi.wiley.com/10.1029/2001GL013669>>.

FORTUNATO, L. A. M., 1990, *Introdução ao planejamento da expansão e operação de sistemas de produção de energia elétrica*. Universidade Fluminense, EDUFF. ISBN: 9788522801039. Disponível em: <<https://books.google.com.br/books?id=w40QkQEACAAJ>>.

GELATI, E., MADSEN, H., ROSBJERG, D., 2010, “Markov-switching model for nonstationary runoff conditioned on El Niño information”, *Water Resources Research*, v. 46, n. 2 (feb), pp. W02517. ISSN: 00431397. doi: 10.1029/2009WR007736. Disponível em: <<https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009WR007736>>.

GELATI, E., MADSEN, H., ROSBJERG, D., 2014, “Reservoir operation using El Niño forecasts—case study of Daule Peripa and Baba, Ecuador”, *Hydrological Sciences Journal*, v. 59, n. 8 (aug), pp. 1559–1581. ISSN: 0262-6667. doi: 10.1080/02626667.2013.831978. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/02626667.2013.831978>>.

GELATI, E., MADSEN, H., ROSBJERG, D., 2011, “Stochastic reservoir optimization using El Niño information: case study of Daule Peripa, Ecuador”, *Hydrology Research*, v. 42, n. 5 (jun), pp. 413. ISSN: 0029-1277. doi: 10.2166/nh.2011.009. Disponível em: <<http://hr.iwaponline.com/cgi/doi/10.2166/nh.2011.009>>.

- GRIMM, A. M., 2004, “How do La Niña events disturb the summer monsoon system in Brazil?” *Climate Dynamics*, v. 22, n. 2-3, pp. 123–138. ISSN: 0930-7575. doi: 10.1007/s00382-003-0368-7. Disponível em: <<http://link.springer.com/10.1007/s00382-003-0368-7>>.
- GRIMM, A. M., 2003, “The El Niño Impact on the Summer Monsoon in Brazil: Regional Processes versus Remote Influences”, *Journal of Climate*, v. 16, n. 2 (jan), pp. 263–280. ISSN: 0894-8755. doi: 10.1175/1520-0442(2003)016<0263:TENIOT>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%282003%29016%3C0263%3ATENIOT%3E2.0.CO%3B2>>.
- GRIMM, A. M., TEDESCHI, R. G., 2009, “ENSO and Extreme Rainfall Events in South America”, *Journal of Climate*, v. 22, n. 7 (apr), pp. 1589–1609. ISSN: 0894-8755. doi: 10.1175/2008JCLI2429.1. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/2008JCLI2429.1>>.
- HAMILTON, J. D., 1989, “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle”, *Econometrica*, v. 57, n. 2 (mar), pp. 357. ISSN: 00129682. doi: 10.2307/1912559. Disponível em: <<http://www.jstor.org/stable/1912559>>.
- HIPEL, K. W., MCLEOD, A. I., 1994, *Time series modelling of water resources and environmental systems*, v. 45. Elsevier.
- HUANG, B., THORNE, P. W., BANZON, V. F., et al., 2017, “Extended Reconstructed Sea Surface Temperature version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons”, *Journal of Climate*, v. 5, pp. JCLI-D-16-0836.1. ISSN: 0894-8755. doi: 10.1175/JCLI-D-16-0836.1. Disponível em: <<http://journals.ametsoc.org/doi/10.1175/JCLI-D-16-0836.1>>.
- IRI, 2018. “IRI”. https://iri.columbia.edu/our-expertise/climate/forecasts/enso/2017-November-quick-look/?enso_tab=enso-sst-table. Accessed: 2018-02-10.
- JARDIM, D., MACEIRA, M., FALCAO, D., 2001, “Stochastic streamflow model for hydroelectric systems using clustering techniques”. In: *Power Tech Proceedings, 2001 IEEE Porto*, v. 3, pp. 6–pp. IEEE, sep. doi: 10.1109/PTC.2001.964916.
- KELMAN, J., DE M. VIEIRA, A., RODRIGUEZ-AMAYA, J. E., 2000, “El Niño influence on streamflow forecasting”, *Stochastic Environmental Research and Risk Assessment*, v. 14, n. 2 (jun), pp. 123–138. ISSN: 1436-3240.

doi: 10.1007/PL00009776. Disponível em: <<http://link.springer.com/10.1007/PL00009776>>.

KELMAN, J., DAMÁZIO, J. M., COSTA, J. P. D., 1983, “Geração de séries sintéticas de vazões diárias–Modelo DIANA”, *Revista Brasileira de Engenharia*, v. 1, n. 2, pp. 5–22.

KOTTEGODA, N. T., 1980, *Stochastic water resources technology*. Springer. ISBN: 9781349034673.

KOUSKY, V. E., HIGGINS, R. W., 2007, “An Alert Classification System for Monitoring and Assessing the ENSO Cycle”, *Weather and Forecasting*, v. 22, n. 2, pp. 353–371. ISSN: 0882-8156. doi: 10.1175/WAF987.1. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/WAF987.1>>.

KROLZIG, H.-M., 1997, *Markov-Switching Vector Autoregressions*, v. 454, *Lecture Notes in Economics and Mathematical Systems*. Berlin, Heidelberg, Springer Berlin Heidelberg. ISBN: 978-3-540-63073-9. doi: 10.1007/978-3-642-51684-9. Disponível em: <<http://link.springer.com/10.1007/978-3-642-51684-9>>.

LIMA, C. H. R., LALL, U., 2010, “Climate informed monthly streamflow forecasts for the Brazilian hydropower network using a periodic ridge regression model”, *Journal of Hydrology*, v. 380, n. 3-4, pp. 438–449. ISSN: 00221694. doi: 10.1016/j.jhydrol.2009.11.016. Disponível em: <<http://dx.doi.org/10.1016/j.jhydrol.2009.11.016>>.

LIMA, C. H. R., LALL, U., 2008, “Inclusão de informação climática e variabilidade espacial em modelos estatísticos de previsão de afluências”. In: *II Simpósio de Recursos Hídricos do Sul-Sudeste*, pp. 1–18.

MAÇAIRA, P. M., OLIVEIRA, F. L. C., FERREIRA, P. G. C., et al., 2017, “Introducing a Causal PAR(p) Model to Evaluate the influence of Climate Variables in Reservoir Inflows: a Brazilian Case”, *Pesquisa Operacional*, v. 37, n. 1 (jan), pp. 107–128. ISSN: 0101-7438. doi: 10.1590/0101-7438.2017.037.01.0107. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-74382017000100107&lng=en&tlng=en>.

MACEIRA, E., PENNA, D., DAMÁZIO, J., 2006, “Geração de Cenários Sintéticos de Energia e Vazão para o Planejamento da Operação Energética”, *Cadernos do IME - Série Estatística*, v. 21, pp. 25. doi: 10.12957.

- MACEIRA, M. E. P., DAMÁZIO, J. M., 2006, “Use of PAR(p) model in the stochastic dual dynamic programming optimization scheme used in the operation planning of the Brazilian hydropower system”, *Probability in the Engineering and Informational Sciences*, v. 20, n. 01 (jan), pp. 143–156. ISSN: 0269-9648. doi: 10.1017/S0269964806060098. Disponível em: <http://www.journals.cambridge.org/abstract_S0269964806060098>.
- MACEIRA, M. E., DAMAZIO, J. M., GHIRARDI, A. O., et al., 1999, “Periodic ARMA models applied to weekly streamflow forecasts”, *International Conference on Electric Power Engineering, PowerTech Budapest 1999*, p. 1. doi: 10.1109/PTC.1999.826517.
- MACEIRA, M. E., BATISTA, F., MELO, A., et al., 2018a, “A Probabilistic Approach to Define the Amount of Energy to be Traded in Hydro Dominated Interconnected Systems”. In: *20th Power Systems Computation Conference - PSCC*, Dublin, Ireland, Junea.
- MACEIRA, M. E., PENNA, D., DINIZ, A., et al., 2018b, “Twenty Years of Application of Stochastic Dual Dynamic Programming in Official and Agent Studies in Brazil – Main Features and Improvements on the Newave Model”. In: *20th Power Systems Computation Conference - PSCC*, Dublin, Ireland, Juneb.
- MACEIRA, M., BEZERRA, C., 1997, “Stochastic Streamflow Model for Hydroelectric Systems”. In: *5th Conference on Probabilistic Methods Applied to Power Systems - PMAPS*, Vancouver, Canada.
- MACEIRA, M., TERRY, L., F.S.COSTA, et al., 2002, “Chain of optimization models for setting the energy dispatch and spot price in the Brazilian system”, *Power System Computation Conference*, v. 2, n. June, pp. 24–28. Disponível em: <http://www.psc-central.org/uploads/tx_ethpublications/s43p01.pdf>.
- MARENGO, J. A., TOMASELLA, J., UVO, C. R., 1998, “Trends in streamflow and rainfall in tropical South America: Amazonia, eastern Brazil, and northwestern Peru”, *Journal of Geophysical Research: Atmospheres*, v. 103, n. D2 (jan), pp. 1775–1783. ISSN: 01480227. doi: 10.1029/97JD02551. Disponível em: <<http://doi.wiley.com/10.1029/97JD02551>>.
- MARIOTTI, A., 2007, “How ENSO impacts precipitation in southwest central Asia”, *Geophysical Research Letters*, v. 34, n. 16 (aug), pp. 2–6. ISSN:

00948276. doi: 10.1029/2007GL030078. Disponível em: <<http://doi.wiley.com/10.1029/2007GL030078>>.

MESINGER, F., CHOU, S. C., GOMES, J. L., et al., 2012, “An upgraded version of the Eta model”, *Meteorology and Atmospheric Physics*, v. 116, n. 3-4, pp. 63–79.

MOOD, A. M., 1950, *Introduction to the Theory of Statistics*. McGraw-hill.

NOAA, 2018. “NOAA”. <https://www.ncdc.noaa.gov/teleconnections/enso/indicators/sst.php>. Accessed: 2018-02-10.

ONS, 2018a. “PEN 2018”. Planejamento da Operação Energética 2018/2022. Operador Nacional do Sistema Elétrico. Rio de Janeiro. 2018, a.

ONS, 2018b. “ONS”. <https://www.ons.org.br>, b. Accessed: 2018-06-10.

PEREIRA, M. V., PINTO, L. M., 1991, “Multi-stage stochastic optimization applied to energy planning”, *Mathematical programming*, v. 52, n. 1-3, pp. 359–375.

PEZZI, L. P., CAVALCANTI, I. F. A., 2001, “The relative importance of ENSO and tropical Atlantic sea surface temperature anomalies for seasonal precipitation over South America: a numerical study”, *Climate Dynamics*, v. 17, pp. 205–212. ISSN: 0930-7575. doi: 10.1007/s003820000104. Disponível em: <<http://link.springer.com/10.1007/s003820000104>>.

PIECHOTA, T. C., DRACUP, J. A., 1996, “Drought and regional hydrologic variation in the United States: Associations with the El Niño-Southern Oscillation”, *Water Resources Research*, v. 32, n. 5, pp. 1359–1373.

PINTO, N. L. D. S., HOLTZ, A. C. T., MARTINS, J. A., 1976, “Hidrologia básica”. In: *Hidrologia básica*, Edgard Blucher.

POVEDA, G., 2001, “Coupling between Annual and ENSO Timescales in the Malaria-Climate Association of Columbia.” *Environmental Health Perspectives*, v. 109, n. 5, pp. 489. ISSN: 00916765. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1240308/>>.

POWER, S., CASEY, T., FOLLAND, C., et al., 1999, “Inter-decadal modulation of the impact of ENSO on Australia”, *Climate Dynamics*, v. 15, n. 5 (may), pp. 319–324. ISSN: 0930-7575. doi: 10.1007/s003820050284. Disponível em: <<http://link.springer.com/10.1007/s003820050284>>.

- PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., et al., 1992. “Numerical recipes in FORTRAN” . .
- RASMUSSEN, E. M., WALLACE, J. M., 1983, “Meteorological Aspects of the El Niño/Southern Oscillation”, *Science*, v. 222, n. 4629 (dec), pp. 1195–1202. ISSN: 0036-8075. doi: 10.1126/science.222.4629.1195. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17806710><http://www.sciencemag.org/cgi/doi/10.1126/science.222.4629.1195>>.
- RASMUSSEN, E. M., CARPENTER, T. H., 1982, “Variations in Tropical Sea Surface Temperature and Surface Wind Fields Associated with the Southern Oscillation/El Niño”, *Monthly Weather Review*, v. 110, n. 5 (may), pp. 354–384. ISSN: 0027-0644. doi: 10.1175/1520-0493(1982)110<0354:VITSST>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281982%29110%3C0354%3AVITSST%3E2.0.CO%3B2>>.
- SAHU, N., BEHERA, S. K., RATNAM, J. V., et al., 2014, “El Niño Modoki connection to extremely-low streamflow of the Paranaíba River in Brazil”, *Climate Dynamics*, v. 42, n. 5-6, pp. 1509–1516. ISSN: 09307575. doi: 10.1007/s00382-013-2006-3.
- SALAS, J. D., DELLEUR, J., YEVJEVICH, V., et al., 1980, *Applied modeling of hydrologic time series*. Water Resources Publication.
- SANGY, P. E. R., 2016, *Inclusão Da Precipitação Na Geração De Cenários De Vazão*. Tese de Doutorado, Universidade Federal do Rio de Janeiro. Disponível em: <<http://pee.ufrj.br/teses/textocompleto/2016033008.pdf>>.
- SANTARÉM, H., PENNA, D. D. J., TREISTMAN, F., 2017, “Aplicação das Correlações Espaciais Históricas na Geração de Cenários Sintéticos Multivariados para o Gerenciamento do Sistema Interligado Nacional”. In: *XXII Simpósio Brasileiro de Recursos Hídricos*, pp. 1–8, Florianópolis, nov.
- SAUNDERS, M. A., CHANDLER, R. E., MERCHANT, C. J., et al., 2000, “Atlantic hurricanes and NW Pacific typhoons: ENSO spatial impacts on occurrence and landfall”, *Geophysical Research Letters*, v. 27, n. 8 (apr), pp. 1147–1150. ISSN: 00948276. doi: 10.1029/1999GL010948. Disponível em: <<http://doi.wiley.com/10.1029/1999GL010948>>.
- SILVA, L., 2012. “Calibração Automática de Modelos Tipo Chuva-Vazão Utilizando Técnicas de Suavização – Uma Aplicação ao Modelo SMAPII”.

Dissertação (Engenharia Civil), Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

SILVEIRA, C. D. S., ALEXANDRE, A. M. B., DE SOUZA FILHO, F. D. A., et al., 2017, “Monthly streamflow forecast for National Interconnected System (NIS) using Periodic Auto-regressive Endogenous Models (PAR) and Exogenous (PARX) with climate information”, *RBRH*, v. 22. ISSN: 2318-0331. doi: 10.1590/2318-0331.011715186. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2318-03312017000100233&lng=en&tlng=en>.

SOLOW, A. R., ADAMS, R. F., BRYANT, K. J., et al., 1998, “No Title”, *Climatic Change*, v. 39, n. 1, pp. 47–60. ISSN: 01650009. doi: 10.1023/A:1005342500057. Disponível em: <<http://link.springer.com/10.1023/A:1005342500057>>.

TEDESCHI, R. G., 2013, *As Influências de Tipos Diferentes de ENOS na Precipitação e nos seus Eventos Extremos sobre a América do Sul - Observações, Simulações e Projeções*. Doutorado em meteorologia, Instituto Nacional de Pesquisas Espaciais - INPE.

TIMMERMANN, A., AN, S.-I., KUG, J.-S., et al., 2018, “El Niño–Southern Oscillation complexity”, *Nature*, v. 559, n. 7715 (jul), pp. 535–545. ISSN: 0028-0836. doi: 10.1038/s41586-018-0252-6. Disponível em: <<http://www.nature.com/articles/s41586-018-0252-6>>.

TIPPETT, M. K., BARNSTON, A. G., 2008, “Skill of Multimodel ENSO Probability Forecasts”, *Monthly Weather Review*, v. 136, n. 10 (oct), pp. 3933–3946. ISSN: 0027-0644. doi: 10.1175/2008MWR2431.1. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/2008MWR2431.1>>.

TIPPETT, M. K., BARNSTON, A. G., LI, S., 2012, “Performance of Recent Multimodel ENSO Forecasts”, *Journal of Applied Meteorology and Climatology*, v. 51, n. 3 (mar), pp. 637–654. ISSN: 1558-8424. doi: 10.1175/JAMC-D-11-093.1. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-11-093.1>>.

TONG, J., QIANG, Z., DEMING, Z., et al., 2006, “Yangtze floods and droughts (China) and teleconnections with ENSO activities (1470–2003)”, *Quarterly International*, v. 144, n. 1 (feb), pp. 29–37. ISSN: 10406182. doi: 10.1016/j.quaint.2005.05.010. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1040618205000947>>.

- TORRENCE, C., WEBSTER, P. J., 1999, “Interdecadal Changes in the ENSO–Monsoon System”, *Journal of Climate*, v. 12, n. 8 (aug), pp. 2679–2690. ISSN: 0894-8755. doi: 10.1175/1520-0442(1999)012<2679:ICITEM>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%281999%29012%3C2679%3AICITEM%3E2.0.CO%3B2>>.
- TREISTMAN, F., 2015. “Análise das correlações existentes entre precipitação e vazão em usinas hidrelétricas brasileiras”. Disponível em: <<http://monografias.poli.ufrj.br/monografias/monopoli10014907.pdf>>. Monografia (Engenharia Ambiental), Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.
- UVO, C. B., REPELLI, C. A., ZEBIAK, S. E., et al., 1998, “The Relationships between Tropical Pacific and Atlantic SST and Northeast Brazil Monthly Precipitation”, *Journal of Climate*, v. 11, n. 4 (apr), pp. 551–562. ISSN: 0894-8755. doi: 10.1175/1520-0442(1998)011<0551:TRBTPA>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%281998%29011%3C0551%3ATRBTPA%3E2.0.CO%3B2>>.
- VIEIRA, A. D. M., 1997, *Hidrologia estocástica e operação de reservatórios*. Tese de Doutorado, Universidade Federal do Rio de Janeiro.
- WANG, B., WU, R., FU, X., 2000, “Pacific–East Asian Teleconnection: How Does ENSO Affect East Asian Climate?” *Journal of Climate*, v. 13, n. 9 (may), pp. 1517–1536. ISSN: 0894-8755. doi: 10.1175/1520-0442(2000)013<1517:PEATHD>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%282000%29013%3C1517%3APEATHD%3E2.0.CO%3B2>>.
- WANG, C., 2002, “Atmospheric Circulation Cells Associated with the El Niño–Southern Oscillation”, *Journal of Climate*, v. 15, n. 4 (feb), pp. 399–419. ISSN: 0894-8755. doi: 10.1175/1520-0442(2002)015<0399:ACCAWT>2.0.CO;2. Disponível em: <<http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%282002%29015%3C0399%3AACCAWT%3E2.0.CO%3B2>>.
- WANG, C., 2004, “ENSO, Atlantic Climate Variability, and the Walker and Hadley Circulations”. In: *The Hadley Circulation: Present, Past and Future*, Springer Berlin Heidelberg, cap. 6, pp. 173–202. ISBN: 1402029438. doi: 10.1007/978-1-4020-2944-8_7. Disponível em: <http://link.springer.com/10.1007/978-1-4020-2944-8_7>.

- WILKS, D. S., 2011, *Statistical methods in the atmospheric sciences*, v. 100. Academic press.
- YUAN, Y., YANG, S., 2012, “Impacts of Different Types of El Niño on the East Asian Climate: Focus on ENSO Cycles”, *Journal of Climate*, v. 25, n. 21, pp. 7702–7722. doi: 10.1175/JCLI-D-11-00576.1. Disponível em: <<https://doi.org/10.1175/JCLI-D-11-00576.1>>.
- ZANCHETTIN, D., FRANKS, S. W., TRAVERSO, P., et al., 2008, “On ENSO impacts on European wintertime rainfalls and their modulation by the NAO and the Pacific multi-decadal variability described through the PDO index”, *International Journal of Climatology*, v. 28, n. 8 (jun), pp. 995–1006. ISSN: 08998418. doi: 10.1002/joc.1601. Disponível em: <<https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.1601>>.

Apêndice A

Synthetic Scenario Generation of Monthly Streamflows Conditioned to the El Niño/Southern Oscillation - Application to Operation Planning of Hydrothermal Systems

A.1 Introduction

A.1.1 Initial Considerations

The generation of electric power in Brazil is carried out mainly through its hydroelectric power plants (HPPs), dispersed in twelve main hydrographic basins, with a total of 114GW. In 2017, about 70% of the generation verified was available via hydroelectric power, 7% through wind power plants, leaving a little more than 20% for thermal generation, characterizing as a green energy matrix (ONS, 2018a). In contrast to purely thermal systems, energetic matrices with the Brazilian composition have greater complexity, due to the lack of knowledge of the future supply of their largest source of energy, hydroelectric. Thus, it can be affirmed that future streamflows are the main source of uncertainty for the Brazilian energy planning and should be a focus of studies in order to constantly improve their representation. Currently, for the second month of mid-term planning and for the whole horizon of long-term planning, synthetic scenarios are generated for streamflows through a stochastic periodic autoregressive model.

This type of stochastic model describes future realizations as a linear regression of past values, plus a random term. In other words, only the past streamflows are used as inputs to the model. Recently, a great scientific effort has been made to include climatic information from other variables relevant to the hydrological cycle in stochastic models. Among the several influential climatic phenomena in the Brazilian hydrological cycles, the El Niño - Southern Oscillation (ENSO) has been pointed out over decades of research as one of the most important. Its occurrence impacts of different forms and intensities the precipitation, and, consequently, the streamflows. Therefore, the energy generation in Brazil is widely influenced by the ENSO.

A.1.2 Objectives

Considering what is exposed above, this dissertation has as main objective to propose an improvement to the model used in the generation of synthetic series of inflows in the official studies of the energy planning. A new stochastic model is presented, capable of including information on the ENSO climatic phenomenon, aiming at determining with greater precision the multivariate distributions of probabilities conditioned for future streamflows.

Given the main objective of this work, a series of challenges to its attainment arise, and thus are characterized as necessary secondary objectives:

- to identify the best way of including information on ENSO in the model;
- to evaluate the reproduction of relevant historical statistics by the traditional model and with the suggested improvement;
- to evaluate the generation of streamflow scenarios in case studies that simulate a real case.

A.1.3 Textual Structure

This first chapter aim to contextualize the problem to which this work belongs and to describe its objectives. The Chapter A.2 will deal with the literature review, subdividing into three main themes: Section A.2.1 will deepen in the planning of energy operation with a focus on the Brazilian case; the Section A.2.2 will discuss the ENSO climate phenomenon, initializing its characterization, its impacts both on a global scale and in Brazil, ending with its forecasting models; Finally, the section A.2.3 will present stochastic models used in hydrology, focusing on models for the generation of synthetic series, from the more traditional models going through which include climatic information in their modeling.

In the chapter A.3, the data obtained from the natural streamflow and the El Niño - Southern Oscillation are presented first. Then the methodology for the generation of monthly synthetic series of streamflow and energy is presented, GEVAZP, and the improvement proposed in this dissertation. At the end of the chapter we present the case studies used in the evaluation of the models, and the metrics used. The presentation of the results is done in Chapter A.4, where the potential of the MS-PAR(p) model compared to the proposed PAR (p) model is verified for the generation of synthetic scenarios of streamflows. In Chapter A.5 the discussion and conclusions regarding this work and the results obtained along it are made, besides indicating possible future improvements . Finally, the bibliographical references that helped to inspire and compose this work are presented. Appendices A, B, C, D and E present the Portuguese version on this dissertation.

A.2 Literature Review

This chapter will present the literature review for the three central themes of this dissertation. The first section will deal with energy operation planning, characterizing the Nation Interconnected System and the methodologies used to solve the problem of it. The second section will discuss the climate phenomenon El Niño – Southern Oscillation, with its characterization, impacts associated with it and the state of the art of its predictions. Finally, a review on the stochastic modeling of synthetic streamflow scenario generation to the hydroelectric power plant (HPP) presenting the traditional approach of the stochastic modeling, and then addressing the stochastic models that consider other climatic variables.

A.2.1 Energy Operation Planning

The National Interconnected System (NIS) is the Brazilian electric energy system of generation and transmission, characterized by being a sizeable hydrothermal. Due to its unique characteristics such as continental dimensions, different hydrological and wind regimes, multiple owners and users of its hydroelectric power plants (HPPs), it can be said that the NIS is one of the most complex systems in the world regarding planning and operation. According to data obtained from the Energy Operation Plan 2018/2022 (ONS, 2018a), it can be observed in Figure A.1 that HPPs are still part of the most significant portion of the Brazilian electricity generation matrix, with 67.8% of total installed power, or 114 GW. Wind power plants represented 7.9% of the Brazilian electric power matrix by the end of 2017, with 9.6 GW of installed capacity, the third largest source of power generation, behind gas-fired thermal plants.

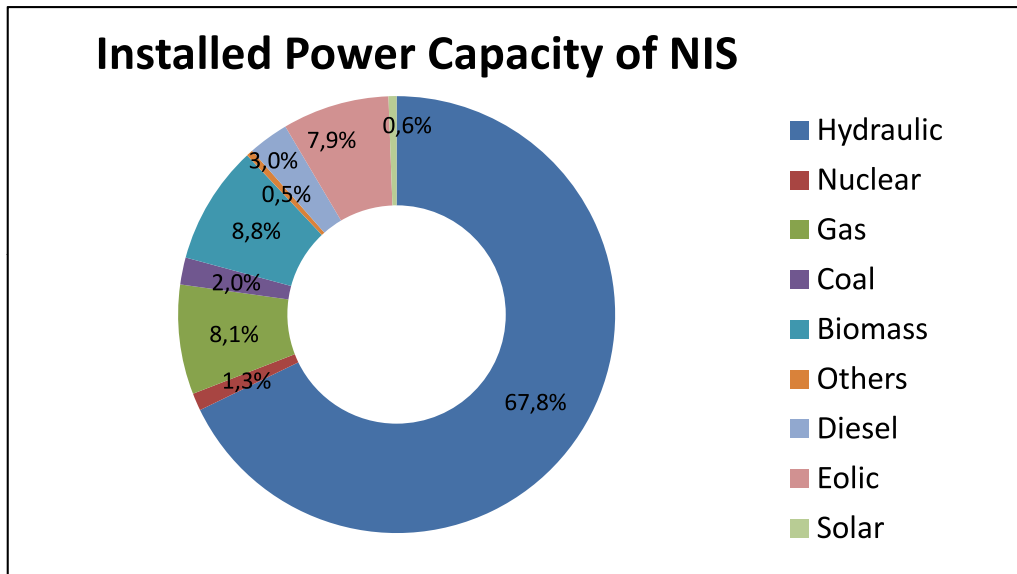


Figure A.1: Participation of sources of power generation in the installed capacity of NIS on December 31, 2017. Source: ONS (2018a).

The expansion of the generation supply up to 2022 presents an increase of 18,728 MW of installed capacity, which represents approximately 12% of the current generation supply. Of this amount, hydroelectric plants continue to be the primary source of supply in generation, presenting the highest gross growth in the energy matrix, with an increase of 8,989 MW, 8.5% of its current supply, but decreasing its share to 65.6% total in 2022. It is interesting to note that the majority of the projected plants are devoid of a regulation reservoir, with no gain of storable energy to the NIS. In this way, the NIS continues to lose its "hydro energetic inertia" gradually, and the massive depletion of large reservoirs can be more frequent at the end of each dry season. Also, the system will be increasingly exposed to prolonged droughts.

Although wind farms represent a low percentage of total installed power, their importance is increasing for the NIS, mainly in the northeast, where there is the highest concentration of this type of power plant. Another factor that has contributed to the high utilization of this source is the fact that the northeast region has undergone critical hydrological conditions in recent years, with inflow well below the long-term mean (LTM), hitting hydropower production strongly. According to data obtained by the National Electric System Operator (ONS), wind generation was responsible for handling the load by about 40 % on a monthly basis, reaching 71% of the load on an hourly basis in the northeast. Currently, the energy demand of the northeast region can be adequately met without the need for hydroelectric generation, only with the activation of thermal power plants (TPPs), wind power and exchanges from other regions. Wind farms planned for 2022 will add 3,064

MW to the NIS, a considerable increase of 24.9%, reaching 15,373 MW installed, representing 8.8% of the total; which places this source as the second largest in installed power. Solar power plants will increase from just 952 MW to 3,638 MW. For comparison purposes, the Angra I and Angra II nuclear power plants have 1,990 MW installed. Despite the benefits, the inclusion of intermittent sources, like solar and wind power, poses some difficulties for energy operation. This leads the hydroelectric generation to assume a new role in the electric sector. Its reservoirs can give stability to wind and solar generation.

As can be observed, the characteristics of the NIS are in the process of significant changes. The electricity generation matrix, which was previously predominantly hydraulic, with large multi-annual regulation reservoirs, and only complemented by thermal plants, is moving towards a matrix with high penetration of intermittent renewable sources and a decrease in the energy storage capacity of the reservoirs. Despite all these factors, hydroelectric power will continue to be the main source of generation. Thus, it is essential to improve the methodologies used in its representation in the different horizons of energy planning.

In contrast to purely thermal systems, in which, in a simplified way, the planning of its energy operation can be solved through the hierarchization of the dispatched TPPs by the cost of production, hydrothermal systems (and their extension with the inclusion of wind power plants) has a far more complex planning (FORTUNATO, 1990). For hydrothermal systems, the main source of uncertainties is represented by future hydrologic inflows faced with temporal couplings of operational decisions. There are links among decision making in consecutive time periods induced by the energy storage capacity in hydroelectric power plant reservoirs. Figure A.2 gives a summary of possible operational decisions in a reservoir. If the decision to store water in a reservoir is made and the future hydrological condition is favorable, this policy is inadequate, representing an unnecessary increase of the present cost of the operation, being the correct operative decision the hydroelectric generation in the present. If future inflows are low, and the decision is made to deplete water from a reservoir at the present moment, this is also an inadequate operation, causing more expensive thermal plants to be driven in the next steps, or, in more extreme cases, load cut (deficit) - representing an increment of future cost. In this case, the appropriated operation would be the storage of water at the present moment. As one can see, actions made at the present moment impact both the operation of the reservoir in question and the future generation costs of the system.

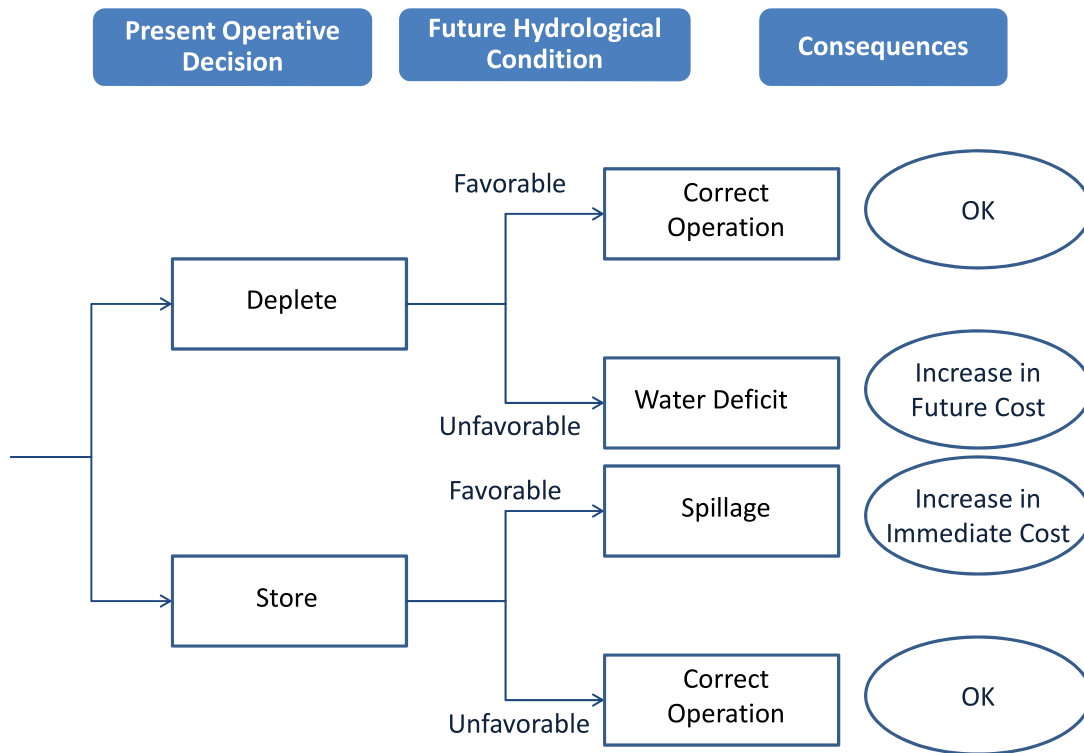


Figure A.2: Operational decision-making process and its possible consequences in a hydrothermal system.

Also, hydrothermal operation planning is confronted with hydraulic couplings between HPPs, exemplified by Figure A.3, which presents a schematic representation of the existing HPPs in the Rio Grande basin. In this image, HPPs with a reservoir are represented by triangles and HPPs without a reservoir by circles. Operational actions of HPPs located upstream (for example, Camargos HPP) end up affecting the operation of all other HPPs located downstream, representing a spatial link among decisions.

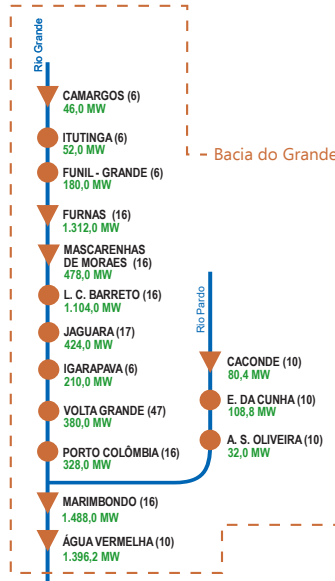


Figure A.3: Example of hydraulic coupling between HPPs in the Rio Grande basin. Adapted from: ONS (2018b)

In a simplified way, proper coordination of all the available resources in the NIS must take into account security criteria, guaranteeing the constant supply of energy to all the regions of Brazil in the present as in the future stages, besides guaranteeing the tariff affordability. The guarantee of the lowest total cost of the operation must take into account both the present cost (or immediate cost) and the future cost incurred by the operation.

Figure A.4 shows the behavior of the operating costs of a hydrothermal system as a function of the final stored volume. The Immediate Cost Function (ICF) represents the current cost associated with supplying TPPs for energy production. As can be seen, it is a function that grows with the final stored volume. The higher the amount of water resources stored, the more TPPs are activated for the production of energy at time t , rendering the immediate operation more expensive. The Future Cost Function (FCF) has an opposite behavior to that of the ICF. When storing resources, the future cost of the operation - from time $t + 1$ to the end of the study period - tends to be lower due to greater availability of generation by the HPPs. The FCF in addition to considering the cost of production of the TPPs takes into account the future cost related to the impacts caused by an energy deficit in the system. There is then an opportunity cost related to stored water as measured by the FCF derivative called the "water value." Thus, Total Cost of Operation (TC) is the sum of present and future costs. The optimal operation of a hydrothermal system occurs with the minimization of the TC. The dashed line in red represents the point where it has the lowest TC, i.e., the ideal operating point.

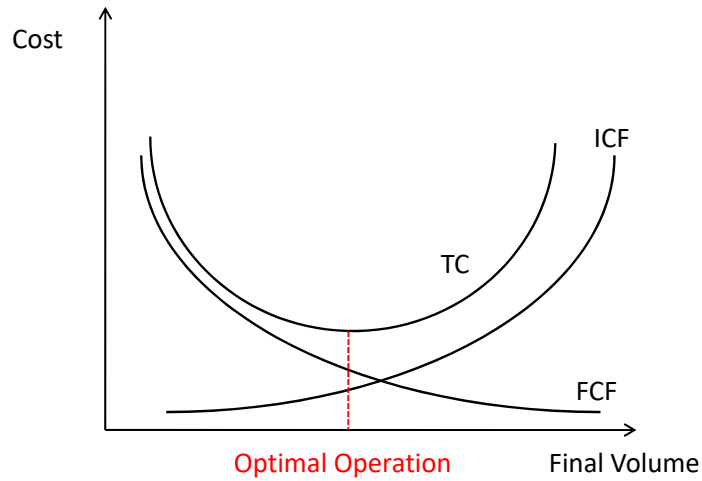


Figure A.4: Behavior of the costs of operating a hydrothermal system as a function of the final volume stored.

It can then be said that the NIS operation is a complex resource optimization problem with multiple goals and possible solutions. Due to its high complexity, the NIS operation planning is divided into different stages (long-term, mid-term, and operational programming) that consider different time horizons and detailed degrees of the NIS representation and its related uncertainties. Currently, the chain of models developed by the Electric Energy Research Center - CEPTEL (MACEIRA *et al.*, 2002) is used to solve the above problem.

Figure A.5 presents a perspective from the chain of models mentioned in the previous paragraph. The longer the planning horizon, the lower the degree of detail of the NIS considered in the problem and higher are the uncertainties about the future.

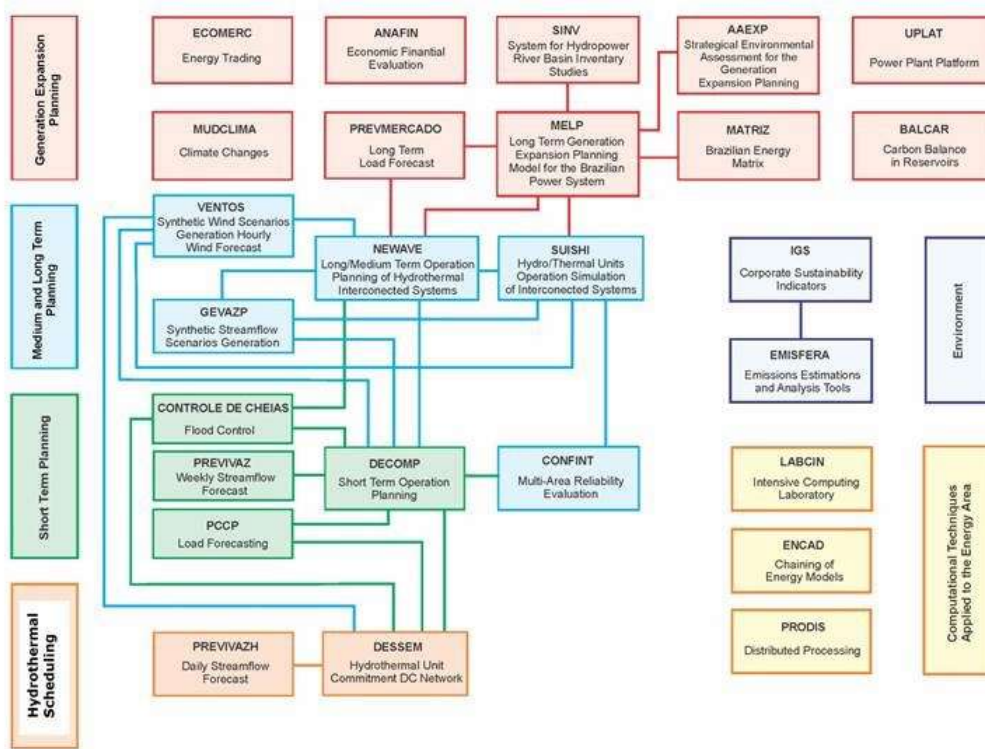


Figure A.5: Chain of models for the expansion and operation energy planning.

In the long-term planning stage, the NEWAVE model (MACEIRA *et al.*, 2018b) groups the HPPs in energy equivalent reservoirs (EER) in order to adequately represent the hydrological uncertainty, in addition to maintaining the computational time of execution of the model at appropriate levels. Currently, the topology considered in NEWAVE represents the hydroelectric park in twelve EERs. The NEWAVE model uses Stochastic Dual Dynamic Programming (PEREIRA e PINTO, 1991) to define the participation of the thermal park and each EER in the load service, in addition to the system interchanges, up to thirty years ahead, with monthly discretization. In other words, the NEWAVE model is responsible for calculating the operation policy of the interconnected system, represented by the FCFs of each month. These functions provide for each system state, storage at the beginning of each month and hydrological condition of the stored energy value.

The hydrological uncertainty during the planning horizon is considered through multivariate synthetic scenarios of affluent natural energy (ANE)¹ that is stochastically generated by GEVAZP (JARDIM *et al.*, 2001; MACEIRA *et al.*, 2006; MACEIRA e DAMÁZIO, 2006; MACEIRA e BEZERRA, 1997). The GEVAZP model, which is the subject of improvement of this dissertation, will be presented in detail in Chapter A.3. From the results of the operation, the NEWAVE model constructs the FCF, which will serve as data input for the DECOMP model (DINIZ

¹ANE is the transformation of the natural streamflow of a HPPs into energy given its accumulated productivity through the basin

et al., 2018). Also, in the long-term operation planning phase, the SUIISHI model (MACEIRA *et al.*, 2018a) performs the simulation of the hydrothermal operation in individual plants, following the operation policy and the set of FCFs, determined in the NEWAVE model.

Regarding a more detailed representation, the DECOMP model solves the operation planning of the NIS in the mid-term stage, with the individualized representation of the hydroelectric park. In this phase, the dispatch of each of the thermal and hydraulic plants is defined, in order to minimize the total cost of the operation at the end of the horizon, taking into account the FCFs built by NEWAVE. For the first month, divided into weekly steps, the DECOMP model uses deterministic streamflow forecasts, provided in large part by the stochastic model PREVIVAZ (MACEIRA *et al.*, 1999). For the other months, the GEVAZP model provides a tree of multivariate synthetic scenarios of streamflows. Currently, the DECOMP model is used for a two-month horizon.

The last stage of the model chain, the daily operation schedule, can be solved through the DESSEM (DOS SANTOS e DINIZ, 2009) model, that represents the NIS with more electrical and hydraulic restrictions details, designating the dispatch of each unit of all powerplants every half hour for the first day and from time to time for the remaining days, up to two weeks ahead. For this stage, daily predictions can be performed by the stochastic model PREVIVAZH (DE CARVALHO, 2001), which performs a non-parametric breakdown of the weekly forecasts provided by the PREVIVAZ model.

A.2.2 El Niño - Southern Oscillation

The El Niño - Southern Oscillation (ENSO) phenomenon is one of the most important natural climatic phenomena, having several impacts on the global circulation patterns of the atmosphere, and climatic variations (including precipitation) around the world (RASMUSSON e WALLACE, 1983). The phenomenon is classified in three distinct phases: La Niña (LN), Neutral (N), and El Niño (EN). The first is characterized by below-average sea surface temperature (SST) in the equatorial region of the Pacific Ocean. This area encompasses the central and eastern portions of the Pacific. The phase known as El Niño is characterized by being the opposite of La Niña, with SST above average in this area of the Pacific. These characteristics about SST in both phases need to be coupled with atmospheric pressure changes and winds for which an episode is triggered. The Neutral phase is characterized by being a transition stage between the other events, with SSTs close to the Pacific average. In some cases, the SST presents EN or LN conditions, but the atmosphere does not present the other changes necessary to classify the other phases, as will be

seen below.

Although the exact cause for the triggering of the hot and cold phases of the ENSO phenomenon is not yet fully known due to their high complexity related to the ocean-atmosphere interaction (TIMMERMANN *et al.*, 2018), both phases are connected with the Southern Oscillation and the Walker Circulation (WANG, 2002). Walker's circulation cell is caused by a pressure gradient that occurs between a low-pressure system on Indonesia and a high-pressure system that lies in the eastern Pacific. During an EN episode, the Walker Circulation cell is weakened, or even reversed, causing warming in the Pacific SST, especially in the equatorial region. At LN events, Walker Circulation is strengthened, resulting in colder SSTs in the Pacific equatorial region (WANG, 2004).

Episodes of the ENSO phenomenon do not usually have a well-defined periodicity, which can be repeated every two or seven years. Also, they tend to persist for nine or twelve months, but in some cases, extreme events can last for more than two years. Currently, there is a distinction between the East El-Niño (also known as canonical) and the Central El Niño (called Modoki), which occurs with SST anomalies in the more centralized Pacific region. Some recent studies have shown the differences between their impacts (SAHU *et al.*, 2014; TEDESCHI, 2013; YUAN e YANG, 2012). This dissertation will not discuss about these sub-classifications of ENSO.

To the ENSO classification and prediction, some indices have been developed, such as the Southern Oscillation Index (SOI), which is based on the pressure difference between two specific regions: Tahiti and Darwin (CHEN, 1982). However, ENSO status monitoring is based primarily on SST measurements from Pacific-specific geographic regions (RASMUSSON e CARPENTER, 1982). The most commonly used are the Nino 1 + 2, Nino 3, Nino 4 and the region known as Nino 3.4 (a region comprising parts of the NINO 3 and NINO 4 regions between the lengths 170°0 and 120°0), as shown in Figure A.6.

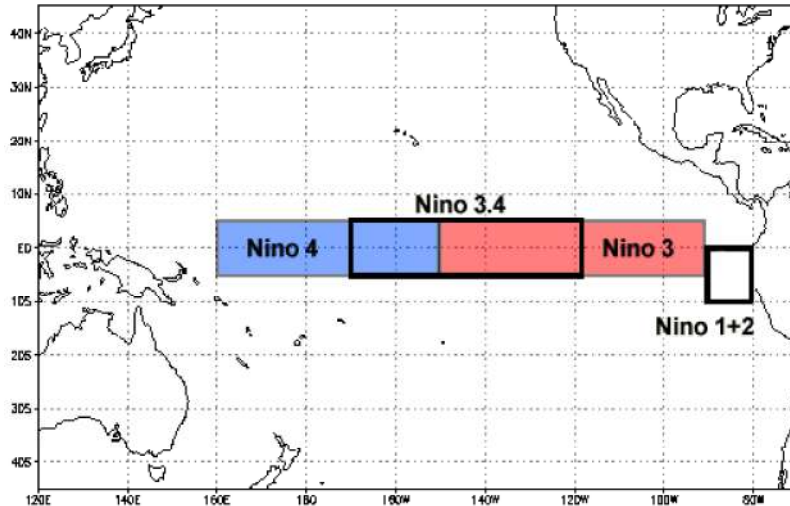


Figure A.6: Pacific SST monitoring regions for ENSO identification. Source: NOAA (2018).

The NINO 3.4 region is considered the most important for the measurement, classification, and prediction of ENSO episodes (BARNSTON *et al.*, 1997; KOUSKY e HIGGINS, 2007). Since this area comprises portions both in the western and eastern Pacific, important features related to atmospheric circulation patterns in ENSO events, such as changes in SST and SST gradients can be appropriately measured (BARNSTON *et al.*, 1997). It is based on the SST of that region that the Oceanic Niño Index (ONI) is constructed. The ONI is defined as a three-month moving average over the SST of the NINO 3.4 region. Its historical data, which begins in the year 1950, is based on the extended and rebuilt SST series version 5 (Ersstv5), which was updated in the year 2017 (HUANG *et al.*, 2017).

Climate Prediction Center (CPC), the agency responsible for producing the ONI, adopts a strategy to remove a significant upward trend in SST in the Pacific due to global warming. The deviation from the mean SST, or anomaly, is calculated based on a centralized period of thirty years, which is estimated at each five-year window. For example, ONI’s historical data between the years 1950 and 1955 is calculated based on the average SST of 1936-1965. The next five-year period (1956-1960) is based on the averages SST of 1941-1970. This procedure is repeated following the same logic for the whole history. The last base period (1986-2015) is used to calculate the most recent anomalies. However, in the year 2021, when there will be a new base period (1991-2021), the values of the last decade will be slightly different because of the update made based on the most recent base period.

Although there are several indices that capture some of the causes and effects of ENSO, such as SOI and various SSTs located in the equatorial Pacific region, and that no single index can confirm the occurrence of an ENSO phenomenon, the National Oceanic and Atmospheric Administration (NOAA) uses the ONI as

the criterion for real-time operational monitoring. When ONI has warm anomalies (greater than or equal to $+0.5$ °C) or cold ones (less than or equal to -0.5 °C), called El Niño or La Niña conditions, for at least five consecutive periods, it is then declared the existence of an EN or LN event. Any other sequence of values is classified as Neutral. In this dissertation, the classification of ENSO events will be performed by the same NOAA's criteria, due to its high acceptance among scientists, simplicity, and the high availability of forecasts.

Related Impacts of El Niño - Southern Oscillation in Precipitation Patterns

Since ENSO is a climatic phenomenon with interactions between the circulation of the oceans and atmosphere, its consequences are perceived all over the world. Extreme events such as hurricanes (ELSNER *et al.*, 2001; SAUNDERS *et al.*, 2000), droughts and floods (ANDREWS *et al.*, 2004; TONG *et al.*, 2006), increase in diseases (CARDENAS *et al.*, 2006; POVEDA, 2001) and economy impacts (SOLOW *et al.*, 1998) have already been related to its occurrence. In most cases, EN episodes have opposite effects when compared to LN events.

One of the most prominent impacts of ENSO can be observed in the change in precipitation and streamflows patterns. In Europe, a study found connections between the ENSO warm phase and an increase in winter's precipitation, while LN episodes were related to a downward precipitation trend in winter (ZANCHETTIN *et al.*, 2008). This impact appears to be more significant in the westernmost region of the continent, which is considered the most sensitive area to climate fluctuations. In this same study, it is also pointed out the possibility of a relationship between ENSO and another teleconnection pattern known as the North-Atlantic Oscillation (NAO), but its mechanism of interaction may actually be spurious, or not so relevant. Stil, in Europe, BRÖNNIMANN (2007) reviewed the impacts and evidences of ENSO in the European climate. In Australia, a survey concluded the existence of the relationship between ENSO and the phenomenon Pacific Inter-Decadal Oscillation (OIP) (POWER *et al.*, 1999). A significant relationship between the ENSO and the interannual rainfall variability is demonstrated. This fact, however, does not occur in conjunction with positive OIP values, only in negative values. A link between EN events and increased precipitation in south-east Asia (a region comprising countries such as Iraq, Iran, Afghanistan, Kazakhstan, among others, and parts of eastern Europe) was found in MARIOTTI (2007). Episodes of ENSO were also highly related to the East Asian monsoon system (TORRENCE e WEBSTER, 1999; WANG *et al.*, 2000).

The connections between ENSO and variations in precipitation and stream-flow are widely studied in the American continent. In the United States of America

(USA), a relationship between EN events was found with the increase in the frequency of days with extreme rainfall and streamflow in the southwest region, whereas the northwest region suffers the opposite effect (CAYAN *et al.*, 1999). However, the opposite occurs, LN episodes. Also, it has been observed that changes in the patterns of inflows are greater than those occurring in precipitation, and the occurrence of an extreme event in the first variable may be tenfold. PIECHOTA e DRACUP (1996) studied the relationship between ENSO and drought in USA, finding consistent response en several regions.

In South America, most of the precipitation occurs during the austral summer (December to February) and autumn (March to May). The influence of the ENSO phenomenon on the South American's rain has been studied for decades (ACEITUNO, 1988). For the Brazilian case, specifically, a joint analysis between the impacts of EN events on precipitation, temperature, circulation structure and thermodynamics of the atmosphere was performed in GRIMM (2003). This study was carried out with data obtained through an extensive network of precipitation stations (1175) covering the period between 1956 and 1992, nine of these years being considered EN episodes, together with atmospheric reanalysis data. The analyzes of the summer monsoon system were done on both monthly and seasonal basis for the months between November and February. The results presented show substantial sub-seasonal variation, showing that studies on a monthly basis can capture relevant patterns. EN events impact by increasing or repressing the mechanisms that produce precipitation. In November and December, anticyclones are favored in the north and center-east regions, causing below-average rainfall, while the Atlantic Ocean moisture flow is transported to the northeast and south regions, causing positive rainfall anomalies. However, in January, the moisture flow is diverged from the south to the east-central region, reversing the precipitation deficit in that region, while the south of Brazil experiences average values of rainfall. In February, these patterns are reversed again, and the northern region of Brazil and the South Atlantic Convergence Zone (SACZ) suffer from below-average rainfall. These results, which were not detected in seasonally based studies, corroborate the importance of the monthly analysis of the effects of ENSO on precipitation. Although the results are quite consistent, it is pointed out that caution should be exercised in inferring the effects of EN events for November and December, since they may vary from year to year.

A similar study evaluated the impacts of LN on the summer monsoon system in Brazil (GRIMM, 2004), considering all eight LN events. In November, strong positive precipitation anomalies were identified in the northeastern region of *Amazônia*, and slightly higher anomalies in the north and Center-East regions of Brazil, while the opposite was observed south region. The positive anomaly in the northeastern

region of *Amazônia* is maintained in December, while in the Center-East region it is weakened and the south region has precipitations close to the average values. In January, abrupt changes occur in the anomalies previously found. In the north-east of the *Amazônia*, the positive anomaly is weakened, while the southern region of Brazil experiences above-average rainfall in its northernmost area, and near the average in the southernmost area. The Central-West region is experiencing above-average rainfall while in the central east, including the SACZ, negative anomalies are observed. In February, the negative anomaly in the Midwest region is reversed and become slightly positive in some areas. In the extreme south of Brazil, negative anomalies reappear, as well as positive anomalies in the northeastern region of the *Amazônia*. Although the results show some strong patterns with inverse effects compared to EN events, it is pointed out that changes in the atmospheric state during episodes of LN found between months would hardly be the only cause of the alterations found.

Figures A.7 and A.8 show a summary of the results found for the mean precipitation percentiles expected in Brazil during EN and LN events, respectively, in November (upper left), December (upper right), January (lower left) and February (lower right). The '(0)' index that appears next to November and December indicates the year in which the event started, and the '(+)' index for the subsequent year. Red colors represent low percentiles, while blue colors are related to high values. Light and dark gray regions represent the results with 85% and 90% significance level. As can be seen, in many cases the effects of episodes of EN and LN are opposite.

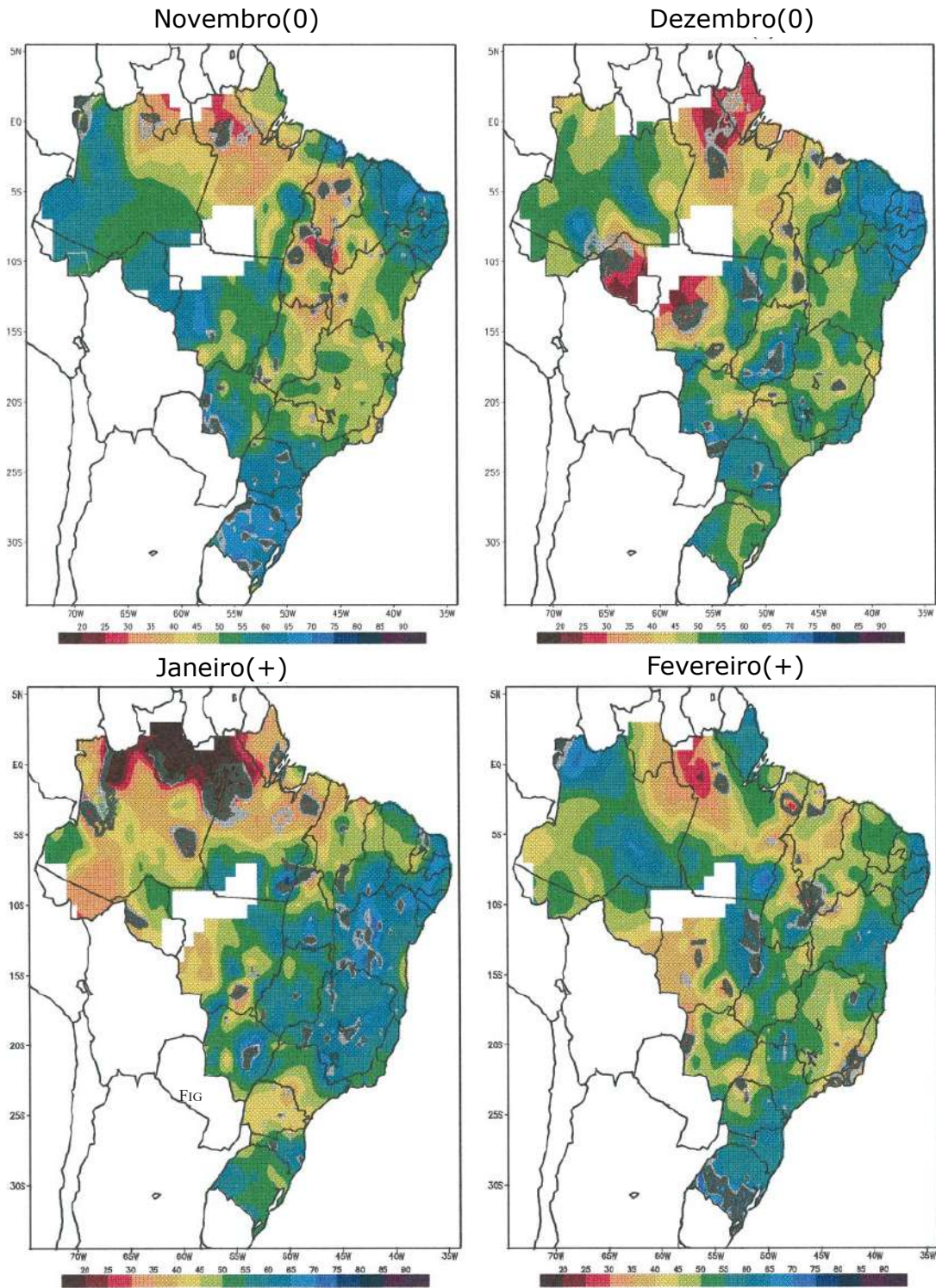


Figure A.7: Percentis of expected mean precipitation in Brazil during El Niño episodes. Source: GRIMM (2003)

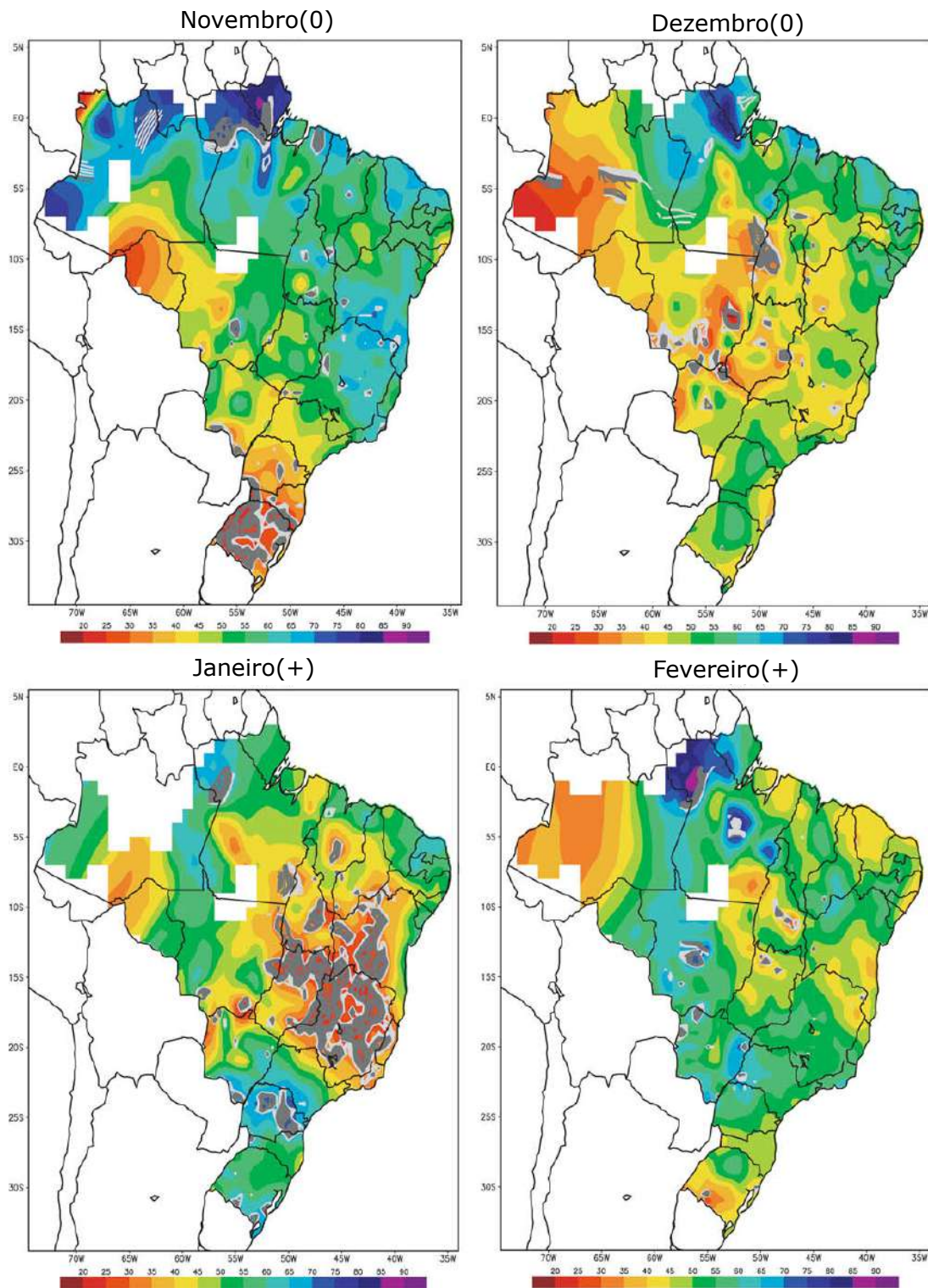


Figure A.8: Percentis of average precipitation expected in Brazil during episodes of La Niña. Source: GRIMM (2004)

In UVO *et al.* (1998) an analysis was made of the relationship between monthly precipitation patterns in the northeast region of Brazil and the SST of the Atlantic and Pacific Oceans. It is shown that the Pacific SST is not the only climatic force capable of modulating precipitation anomalies in the northeast region,

with Atlantic SST being an important factor. A numerical study in PEZZI e CAV-ALCANTI (2001) corroborates this relationship between the Atlantic and Pacific SSTs and precipitation patterns in the same region. In addition to this conclusion, it is highlighted that during EN events (when the Pacific SST is higher than the average) the rest of South America is not affected by the Atlantic SST. However, in LN episodes, the Atlantic SST starts to influence the precipitation of the continent, especially in central and southern Brazil.

An evaluation of the ENSO's influence on extreme precipitation events in South America is presented in GRIMM e TEDESCHI (2009). An extensive database containing more than ten thousand stations covering the period between 1956 and 2002 was used in this study. Extreme events are defined as periods in which the three-day moving average is higher than the 90th percentile after adjusting a Gamma distribution for each day of the year. A significant influence of ENSO episodes on the frequency and intensity of intense rainfall events was found in several regions of South America. Changes in the frequency of intense rainfall events are almost symmetrical between LN and EN events, observed more clearly during the wet season (austral summer). This behavior is also consistent with the variation of the total monthly precipitation during ENSO events. When a region shows an increase in monthly rainfall, the frequency of extreme events tends to increase. Still, the opposite, is not always true. It is concluded that the impact on extreme events is observed in a very dispersed way throughout the continent, indicating that the ENSO phenomenon has an even more significant influence on daily rainfall compared to monthly rainfall.

An interesting study shows that caution is needed when precipitation and their direct impacts (such as river flows), forecasts are made conditioned to ENSO events in South America AMBRIZZI *et al.* (2004). In this study, ENSO episodes are divided into two groups: the first is composed of years in which the impacts of ENSO on precipitation behave as expected according to previous studies; the second group is composed of EN and LN events that had different performance patterns from the first one. AMBRIZZI *et al.* (2004) investigated the atmospheric interactions and behavior of Walker and Hadley cells during these events. For the two extreme phases of the ENSO, the second group presented different patterns from the first group, regarding atmospheric circulation and precipitation. According to the Atlantic SST, an intensification of the SACZ and the Intercontinental Convergence Zone (ITCZ) may occur, developing the patterns found in the second group. Figure A.9 presents the difference found in the Hadley and Walker cells between the two analyzed groups being group (I) the canonical impacts that occur during ENSO events and group (II) with the reverse impacts. While subsidence movements are related to below-average rains (arrows oriented from top to bottom), upward motions (arrows oriented from

the bottom to top) represent the opposite. The areas shaded in blue and red on the South American continent represent negative and positive anomalies, respectively, of precipitation on the continent. The blue and red areas in the ocean represent positive and negative anomalies of SST, respectively.

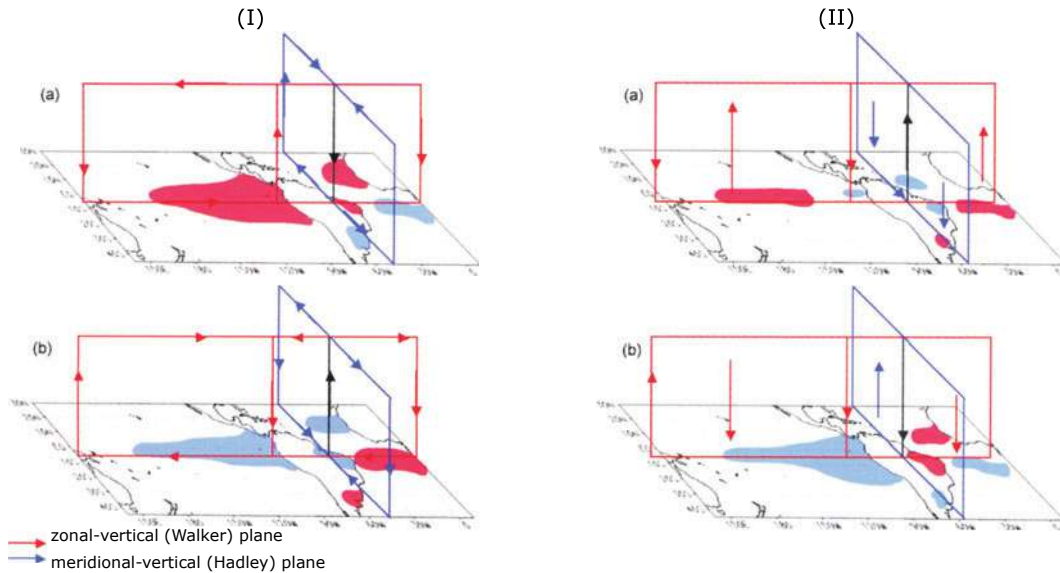


Figure A.9: Walker Walker and Hadley circulation cells for events of El Niño (a) and La Niña (b) for canonical (I) and reverse (II) impacts. Source: AMBRIZZI *et al.* (2004)

Related Impacts of El Niño - Southern Oscillation in Streamflow Patterns

Part of the studies reviewed in this chapter present a strong relationship between ENSO events and some change in precipitation patterns. Although it is not a direct cause and effect relationship, the evidence from the bibliographic survey carried out above indicates that ENSO is one of the most important climatic phenomena, with the capacity to change precipitation patterns in Brazil. Since the rain-streamflow relationship is one of the bases of the hydrological cycle, it is expected that there will also be impacts on hydrological flows to the HPPs.

From the point of view of the synergy between rainfall and streamflow, TREISTMAN (2015) quantified the correlation between monthly rainfall and streamflow in five HPPs, representatives of different hydrographic regions of Brazil. Consisted data of 814 rainfall stations, with at least thirty years of history, were analyzed together with the historical data of natural inflows^{2 3} during the same period

²Natural flows are those that would occur in a section of the river if there were no anthropogenic actions in its contributing basin - such as regularization of flows by reservoirs, water diversions, evaporation in reservoirs and consumptive uses (irrigation, animal husbandry, and urban supply, rural and industrial).

³The process of data consistency and the naturalization of streamflows is carried out by ONS. Further information on these data can be obtained at <http://ons.org.br/>

for the five UHEs. A high correlation was found between rainfall and contemporary monthly natural flows for almost all UHEs analyzed during the wet period (October to March). The Sobradinho HPP is an exception, having the highest correlation value between the natural streamflow and the precipitation one month late. As most of the rainfall occurs during the wet season, the streamflows of the dry period (April to September) are mostly correlated with the rains of the previous months.

Another exception is pointed to the Itá HPP, which is located in the south region. This area of Brazil presents little seasonality for the two variables, with rainfall distributed throughout the year. Consequently, each month shows high values of correlation between the two variables without delay, demonstrating that the watersheds in this region have a rapid response to precipitation events. For this same HPP, the annual correlation between the annual precipitation and the natural annual flow is 0.97, an extremely high value. Figure A.10 presents the annual anomalies of these two variables for the Itá HPP, where a strong relationship between them can be observed. It is worth noting the year 1962, when there was an LN episode, presenting one of the smallest anomalies, and the year 1983, characterized by an episode of EN, with the highest anomaly value in both series. It is also possible to observe that the variation of the anomalies of affluent natural flow is greater than in the precipitations during the ENSO events.

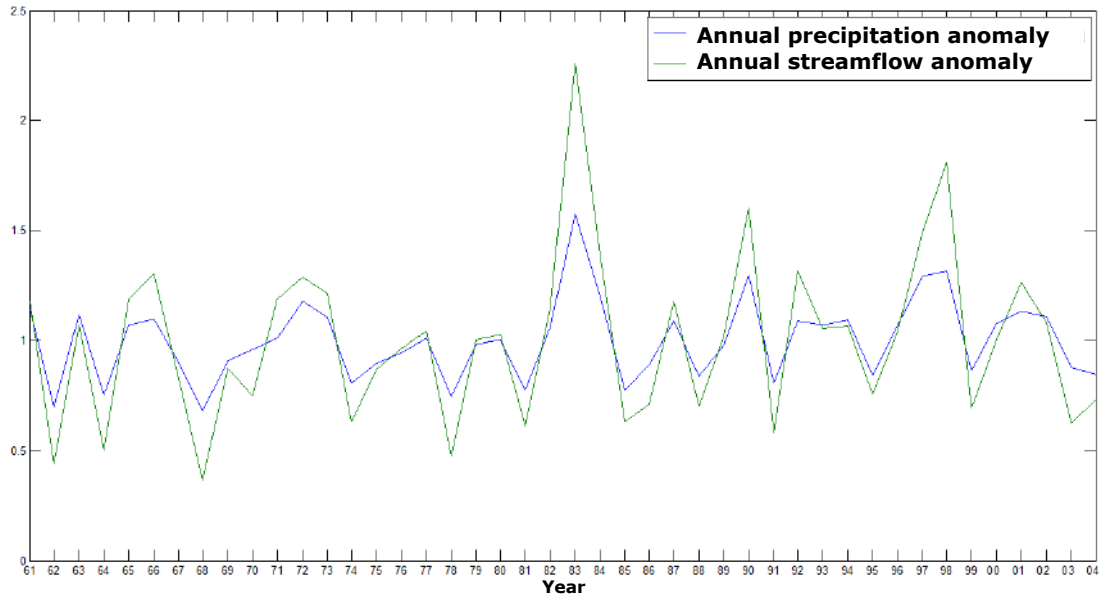


Figure A.10: Annual rainfall anomalies and natural streamflow for Itá HPP between 1961 and 2004. Source: TREISTMAN (2015).

As the ENSO phenomenon can modulate the intensity and frequency of rainfall in Brazil, and the strong relationship between rains and tributaries is known, it is expected that the influxes in the Brazilian rivers will be affected to some degree by this climatic phenomenon. Several studies on the association of ENSO with changes

in patterns of Brazilian streamflows have been carried out over the years.

DETTINGER *et al.* (2000) sought to identify the relationship between ENSO and streamflows that occur in rivers around the world, with a focus on the American continent. Using an extensive database, through the use of statistical techniques such as Correlation Analysis and Principal Component Analysis, it was possible to identify the relationships between river flows and the ENSO phenomenon on a global scale. For Brazil, specifically, changes were identified in flow patterns similar to the standards cited earlier in this dissertation for rainfall. According to the results, EN episodes tend to be drier than LN events in the Amazon basin, mainly in its upper portion, occurring the opposite in the south of the continent. Similar conclusions for the Brazilian territory were also reported in MARENGO *et al.* (1998), GRIMM e TEDESCHI (2009) and CAMILLONI e BARROS (2000).

CAPOZZOLI *et al.* (2017) evaluated the relationship between different climatic patterns with the natural streamflows in 163 HPPs of the NIS, using Principal Components Analysis. Analyzing the spectral peaks and patterns of climatic phenomena, it was pointed out that the first two main components, which account for more than 50% of the variability, may be associated with the ENSO phenomenon and to decade oscillations, such as the Pacific Decadal Oscillation (PDO), Atlantic Multi-decadal Oscillation (AMO), and with the SST anomaly in the South Tropical Atlantic (STA). It is also highlighted that the second main component represents the opposing variabilities between the south region of Brazil, with the remaining portion of the country, mainly on the São Francisco and Tocantins basins.

Predictions for El Niño - Southern Oscillation

ENSO episode predictions are mostly based on SST forecasts in the NINO 3.4 region. The International Research Institute for Climate and Society (IRI) makes a collection of forecasts for the anomalies in this region by models developed by several research institutes (twenty-five models by the November 2017 report) and let publicly available on their site ⁴. In this report only models with monthly basis are included, where the forecasts are carried out for ten periods ahead. This information is made available each month in two ways: through a prediction plume, as shown in Figure A.11; and through a probability table divided by the three ENSO conditions. An average value of the anomalies is given by considering only the dynamic origin prediction models, only statistical models, for the whole set of forecasts, and for a consensus of the Climate Prediction Center - CPC (not considered in the average value of the set of forecasts). Dynamic forecasting models are based on equations that seek to represent the physical relationships between the ocean and the atmo-

⁴<http://iri.columbia.edu/our-expertise/climate/forecasts/enso/current>

sphere. Statistical models are based on regression models, artificial intelligence, and other similar techniques that use data from the historical series to predict the future. The expected CPC consensus value for the index is based on information available to all models, the current state of the atmosphere, and the human judgment of IRI and CPC experts. From this diagnosis, another probabilistic prediction table is also produced, which takes into account the researchers' perceptions.

Table A.1: Probability forecast of ENSO obtained in November 2017. Source: IRI (2018).

Season	La Niña	Neutral	El Niño
NDJ 2017	75%	25%	0%
DJF 2018	71%	29%	0%
JFM 2018	60%	40%	0%
FMA 2018	45%	54%	1%
MAM 2018	26%	72%	2%
AMJ 2018	16%	76%	8%
MJJ 2018	17%	61%	22%
JJA 2018	17%	51%	32%
JAS 2018	16%	46%	38%

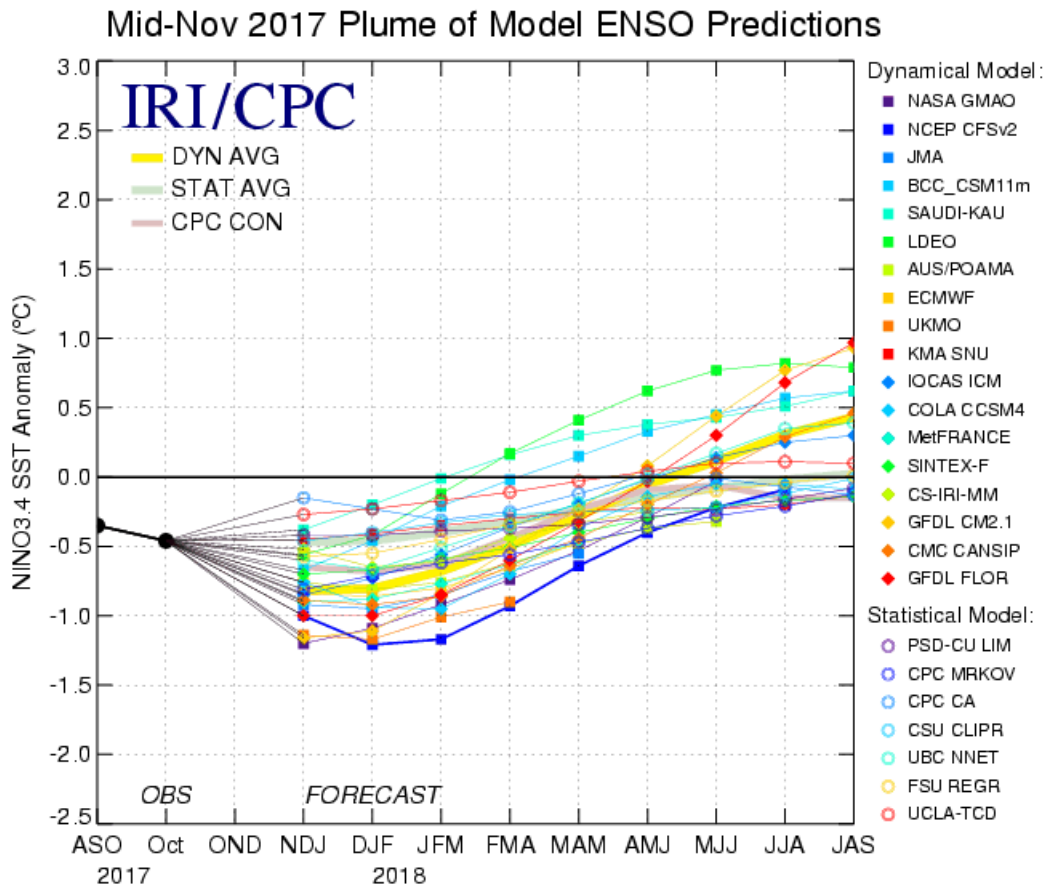


Figure A.11: Prediction plume of SST anomaly in NINO 3.4 obtained in November 2017. Source: IRI (2018).

The probabilistic predictions are based on all available models, weighted equally, together with an average-centered standard deviation function. This function is constructed as a Gaussian curve, taking into account the expected global precision for each period of the year and the horizon in which the forecast is being made. The closer to the month in which the prediction starts, the smaller the errors, and thus the narrower the distribution. Overall forecast accuracy is highest during the austral winter (June to August), when ENSO events tend to start, whereas in the austral summer (December to February) forecasts are better.

BARNSTON *et al.* (2012) analyzed the predictive capacity of the IRI real-time forecast plume during the years of 2002 and 2011 compared to the retrospective predictions provided by some of the models included in the plume. Although the precision of the real-time forecast is slightly lower than the less advanced models of the past decades, the results may have been influenced by the chosen study period, which proved to be particularly more challenging than the 1981-2010 period. Considering only the last nine years of retrospective forecasts, the forecast plume shows slightly better results. Also, it was found that both predictions tend to start the ENSO events later than observed, or with a lower intensity, been longer

predictions worse. This type of error is intensified during the ENSO transition period.

In an earlier study, TIPPETT e BARNSTON (2008) showed some of the benefits of the probabilistic forecasts of ENSO. The predictions of the project called Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER) were collected. In order to produce the probabilistic predictions of ENSO, different calibration methods were used, from simple methods without any weighting, in which each member of the prediction was classified between the three distinct states of ENSO, to more complex methods of joint calibration, where each was weighted optimally. The analyzes were based on the forecasts produced for 1980 and 2001. Through the metrics used, it was concluded that the forecasts produced by the set of all members were better than the use of a single member that had the best performance. An interesting conclusion of this study shows that although the transition period of ENSO is when forecasting models have the least predictability, it has the most significant advantage over the use of historical values.

TIPPETT *et al.* (2012) also analyzed the probabilistic forecasts produced through the IRI forecast plume. Two types of probabilistic prediction were used: a simplified version using only the relative frequency obtained between predictions of the plume framed between the three states of ENSO, and another using the plume predictions and an adjustment of the probability distribution through a Gaussian model. There was no gain in the prediction of the second model about the first one. Figure A.12 shows the predicted category for ENSO in each season between the years 2002 and 2011, and the category observed. The horizontal axis corresponds to the time that the prediction was made, and the vertical axis the number of steps ahead about the initial period. Values 1, 0 and -1 represent the states of EN, N, and LN respectively. Positive or negative values are presented when the categories of EN and LN are most likely, while values of zero represent N events. A decline in forecasting power can be observed about the initial forecast time, in addition to the tendency to start the beginning of events a little later.

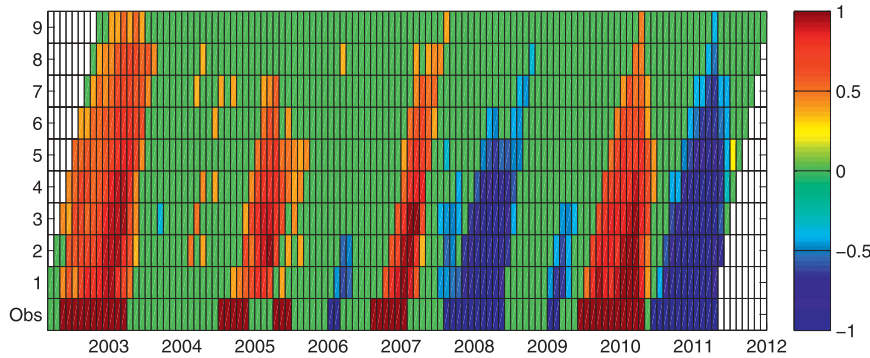


Figure A.12: Observed and predicted ENSO categories. Source: TIPPETT *et al.* (2012)

A.2.3 Stochastic Models in Hydrology

In most situations, it is challenging to predict deterministically what will occur in a future instant for natural phenomena, such as precipitation in a river basin or the streamflow in a particular section of a river. Although there are models that produce deterministic predictions, there will always be an error, because models are only approximated ways to represent nature.

The set of observations ordered over time (in days, weeks, years) of a stochastic process composes a time series. The evolution of the time series of a statistical phenomenon, like the flow of a river, develops following probabilistic laws that can be approximated by mathematical expressions. It can be interpreted that the observed historical series is only a possible realization of the stochastic process that produced it, and, therefore, it is also a time series. The same stochastic process that generated the historical series can also generate other series. These series, called synthetic, also have a chance of occurring. A set of synthetic series, synthetic scenarios, keeps the same statistical characteristics observed in the historical series.

The idea of generating synthetic scenarios through a stochastic model has as one of its objectives to infer about the future steps of a time series, and to use this information in some type of decision making, such as reservoir dimensioning (KOTTEGODA, 1980), or for the energy operation planning, which is where this dissertation fits. Currently, for the NIS, the GEVAZP model generates a scenario tree for the second month of mid-term operation planning and for the long-term operation planning are generated synthetic scenarios with the horizon of ten years. Both phases of planning require possible realizations of future flows, which makes it necessary the stochastic modeling of streamflows. In this way, NIS is prepared to operate its resources efficiently, including for extreme events.

Traditional stochastic models used in the generation of synthetic streamflow scenarios, usually derived from the BOX e JENKINS (1970) family of models, de-

scribe future realizations as a combination of past observations (HIPEL e MCLEOD, 1994; SALAS *et al.*, 1980). In this way, only endogenous variables are used to represent the stochastic process of inflows. Given the connections between river flows and other variables of the hydrological cycle (evaporation, precipitation), and the teleconnections with other climatic phenomena such as ENSO and PDO, some studies have sought to include this exogenous information in hydrological stochastic models. With this, it is expected to determine with greater accuracy the the conditioned multivariate frequency distributions for future streamflow given all the relevant information available at a given time.

Synthetic Scenario Generation Models using Climate Information

Recently, several methodologies have been developed to include climatic information in the stochastic models used in the forecast and generation of synthetic series of inflows. Although conceptually it is easy to deduce that the streamflows are the result of several connected physical processes, the consideration of climatic information is a continually developing challenge in stochastic hydrology.

COLONESE (2016) investigated the use of quantitative precipitation forecast provided by the ETA model (MESINGER *et al.* (2012)) with a resolution of 15 km in flow prediction models. A detailed analysis of the data of the atmospheric model was made, pointing out possible paths of its insertion in order to improve the predictions of daily and weekly inflows obtained for four UHEs located in the Southern region of Brazil through two distinct models: a stochastic model - PREVIVAZH (DE CARVALHO, 2001); and a deterministic model - SMAP II in its smoothed version (SILVA, 2012).

The work developed by VIEIRA (1997) had as one of its objectives to consider the influence of ENSO on the periodic autoregressive model - PAR(p) - through a technique that did not alter the structure of the model. The first step consisted in the investigation of the teleconnection between the streamflow series (monthly and annual) and the SOI series through the functions of cross-correlation. The hydrological year considered starts in November and ends in October of the subsequent year. A significant contemporary (lag 0) relationship was detected in HPPs belonging to the hydrographic basins of the Grande, Paranaíba, Paraná, Tietê and São Francisco rivers. For the basins of the Paranaíba do Sul, Ribeirão das Lajes (southeastern region) and Iguaçú and Jacuí (southern region), there were no significant correlations. Another relevant fact of this analysis was the verification that the annual series presented a stronger relationship with the SOI series than the monthly series.

The proposal for the PAR(p) model of VIEIRA (1997) includes an alternative algorithm for parameter estimation using a similarity criterion to classify the years according to their similarity to the current state of the SOI variable, considering

the last 12 months. The absolute mean deviation given by the ratio between the current state and the past states is calculated. During the estimation of the PAR(p) parameters, each year of the historical is weighted by an estimated weight concerning the deviation calculated in the previous step. The application of the weights is made six months ahead of the period in which it was calculated. Tests were carried out using the proposed algorithm for predicting the streamflows of the Furnas HPP up to twelve months ahead, in three hydrological years described as dry and three wet. According to the metrics used, it was shown that there is an improvement, although modest, in most of the years analyzed when considering the macroclimatic variable SOI in the prediction of streamflows to this plant. The proposed model allows the use of climate information up to twelve months ahead only. Therefore, it should be noted that the generation of macroclimatic variables could prolong the use of this algorithm for long-term studies. Another application of this same model in the a Colombian case can be found in KELMAN *et al.* (2000), where it was shown an improvement in the scenarios generated with climate information compared to the traditional model.

A methodology for considering the SOI in the generation of daily streamflow scenarios is described in COSTA *et al.* (2014). In this application, directed to flood control, the historical record of streamflows is classified, using the SOI, between dry, normal and wet years. Given this classification, three sets of parameters are estimated for the generation of synthetic scenarios of daily streamflows conditioned to the ENSO phenomenon, through the DIANA model (KELMAN *et al.*, 1983).

ANDERSON *et al.* (2015) incorporated the information available in the ONI prediction set provided by IRI in a non-parametric stochastic model used to manage the reservoirs of the Lower Colorado river basin in Texas (ANDERSON e WALKER, 2011). Each forecast in the plume was classified among the three phases of the ENSO (La Niña, Neutral, El Niño) phenomenon in order to calculate a transition probability matrix between states for the nine months forecast provided. Using the streamflow history, classified in dry, moderate and wet months, and its transition matrix between these states in conjunction with the transition matrix between the ENSO conditions, the proposed model resamples the history, simulating through a chain of Markov two thousand scenarios for the next sixty months. According to the results, there is an improvement in the generation of future scenarios, and these were particularly useful for the projections of the year 2015 in the management of the river basin.

One of the most used stochastic models in recent studies for prediction and generation of synthetic scenarios of flows with the incorporation of climatic information is the periodic autoregressive model with exogenous variables (PARX). This type of modeling uses, in addition to the past inflows, other variables external to the

process being modeled that have some relation. Thus, it is expected that there will be gains in the modeling of the variable of interest. The work described in SANGY (2016) added the monthly precipitation forecast provided by the ETA40 model of the National Institute of Space Research (INPE) in the process of generation of synthetic streamflow scenarios in five HPPs of the NIS, through the adjustment of PARX models. Although it is demonstrated that the use of rainfall data can bring significant gains in the prediction of streamflows, especially in wet periods, the quality of the precipitation forecast for a monthly horizon is still low, therefore, limiting the enhancement in the streamflow scenarios. In some cases, it is shown that the insertion of the exogenous variable may even lead to worse results when compared to the traditional model.

A similar approach was taken by DE FARIAS (2012) for the generation of inflow scenarios in the Furnas HPP. In this work, several autoregressive (AR) models with different orders and autoregressive with exogenous variables (ARX) were adjusted, these variables being the rainfall observed in the hydrographic basin related to HPU, and predicted precipitation, provided by the global circulation model ECHAM 4.5. Each AR model was responsible for generating a single scenario (one forecast), while the adjusted ARX model generated twenty-four distinct scenarios, both for the test period from January 2002 to December 2007. According to the metrics evaluated, it was concluded that the ARX model was able to capture more accurately abrupt variations in the flow series, obtaining a superior performance.

In LIMA e LALL (2008), it was compared the prediction results of inflows made up to six months in advance to 57 HPPs of the NIS through four models: PAR of order 1; PARX being the exogenous variable SST between latitudes 25°N and 45°S of the Pacific and Atlantic oceans; and two other variants with the inclusion of spatial correlation from the HPPs. From the obtained results, analyzing the average efficiency coefficient, the PAR model obtained the best performance, although the individual results vary widely between HPP and the month analyzed. However, it was observed that in the months with higher inflows, in HPPs with the influence of remote forcings (climate variables not directly related to streamflow), the PARX model presented a better overall performance.

A similar analysis was performed in LIMA e LALL (2010), where the forecasts were compared for a six-month horizon. It was utilized the model: PAR, PARX models with the possibility of inclusion of SST of the NINO3 region, SST of the South Atlantic and low-level zonal wind indexes in the southeast region of Brazil, with a delay of up to six months, and another PARX model were used the parameters of the regression equation are adjusted through a second methodology. The results showed that the inclusion of climatic variables improved predictions for up to three months ahead for most of the HPPs during the test period. It was observed that the

use of climatic variables included with a longer delay could improve the forecasts in larger horizons when compared with the inclusion of variables with less delay.

MAÇAIRA *et al.* (2017) tested a PARX model with up to eight exogenous variables to generate inflow scenarios for eight HPPs in different regions, presenting only a slight advantage over PAR(p) modeling. In SILVEIRA *et al.* (2017) a comparison was made between a PARX model with up to four exogenous variables and the PAR(p) model for the prediction of inflows in 88 Brazilian HPPs. The results were analyzed separately by month, showing that for most of the year the PAR(p) model can produce better predictions, being only surpassed in the period between September and November by the PARX model.

Looking for a new approach to predicting and generating streamflow scenarios with the inclusion of climate information, another class of stochastic model is presented in GELATI *et al.* (2010). In this paper, it is defined the Markov-modulated autoregressive model with exogenous input (MARX), where the parameters of the ARX model are alternated according to a climatic regime modeled by hidden Markov chain, with the probability of transition between states as a function of climate information. This model is a natural extension of the work presented in HAMILTON (1989), which introduced the models of the Markov-Switching Autoregressive family (MS-AR) to model the Gross Domestic Product where the hidden states represented the growth and recession phases. After this work, new paths for stochastic modeling of time series were opened. Another reference for this family of models can be found in KROLZIG (1997), where autoregressive modeling has been extended to a vector version.

The work developed by GELATI *et al.* (2010) used the MARX model for the prediction and generation of streamflow scenarios in a HPP located in the western region of Ecuador. In this model, the anomalies of the inflows were assumed as a function of their past values (for a one-step forecast), as well as of climate variables (for long-term prediction and simulation), and also modulated by a hidden climate process that is influenced by the same climate variable. After a process of tests with several climatic variables related to the ENSO phenomenon, the SST anomaly of the NINO 1 + 2 region and the Trans-Nino Index were identified through the Bayesian Information Criterion (BIC) as being the most adequate to the model. Also, only two states for the hidden Markov chain were adjusted, one related to El Niño events and another to neutral events and La Niña. For this work, the observations of climatic variables were used. It was noted that in the calibration and validation periods, the estimated inflows obtained an excellent performance, although at low streamflow the model, in general, overestimated, having a better adjustment to the high streamflow. This fact can be explained by the high correlation of El Niño events with high anomalies, whereas drier events do not seem to be correlated.

Applications of this modeling in reservoir optimization can be found in GELATI *et al.* (2011) and GELATI *et al.* (2014), where the prediction of the inflows was carried out in a multivariate manner for two reservoirs.

The papers presented in this section sought to include climatic variables in the prediction or generation of scenarios, in order to improve the traditional stochastic models. The reviewed models that include exogenous variables such as ARX or PARX, although they present some good results, end up depending on the quality of the predictions for the climatic variables for the generation of future scenarios. Also, this type of modeling consider that the climatic variables can influence only the average level of the inflows, ignoring possible changes in the associated uncertainties. Other proposed modeling do incorporate this question, but they end up having limitations in the study horizon due to the inclusion of the climatic variable, or they do not conform to the restriction of linearity that involves the NEWAVE model. Thus, the methodology that will be presented in the next chapter seeks to be inspired by these researches but applied to the operation planning of the NIS.

A.3 Methodology

In this chapter, firstly it is presented all the data used in this dissertation. Then, the mathematical and statistical formulation is detailed for the generation model of monthly synthetic series of energy and streamflow - GEVAZP - and an alternative model based on it, but capable of capturing climatic information regarding the ENSO phenomenon. Finally, the tests carried out in the scenarios generation are discriminated. The approach of the problem outlined encompasses a set of procedures and data arranged in Figure A.13 constituting the methodology of this dissertation

A.3.1 Natural Streamflow Data

The natural streamflow historical data for the 141 HPPs used in the case studies in this dissertation was obtained in the input data set of the monthly operation planning of the January 2018. It is used all record from 1950 to 2017, which is the same period covered by the ONI.

A.3.2 El Niño - Southern Oscillation Data

El Niño - Southern Oscillation Historical Data

The classification of the ENSO history in its three possible states (LN, N, and EN) will use the ONI (HUANG *et al.*, 2017), constructed as a three-month moving average over the SST of the NINO 3.4 region. The history of the ONI, shown in

Figure A.14, covers the years between 1950 and 2017. In this dissertation, the ENSO status classification criterion will be the same as that used in NOAA’s operational monitoring, i.e., based in ONI:

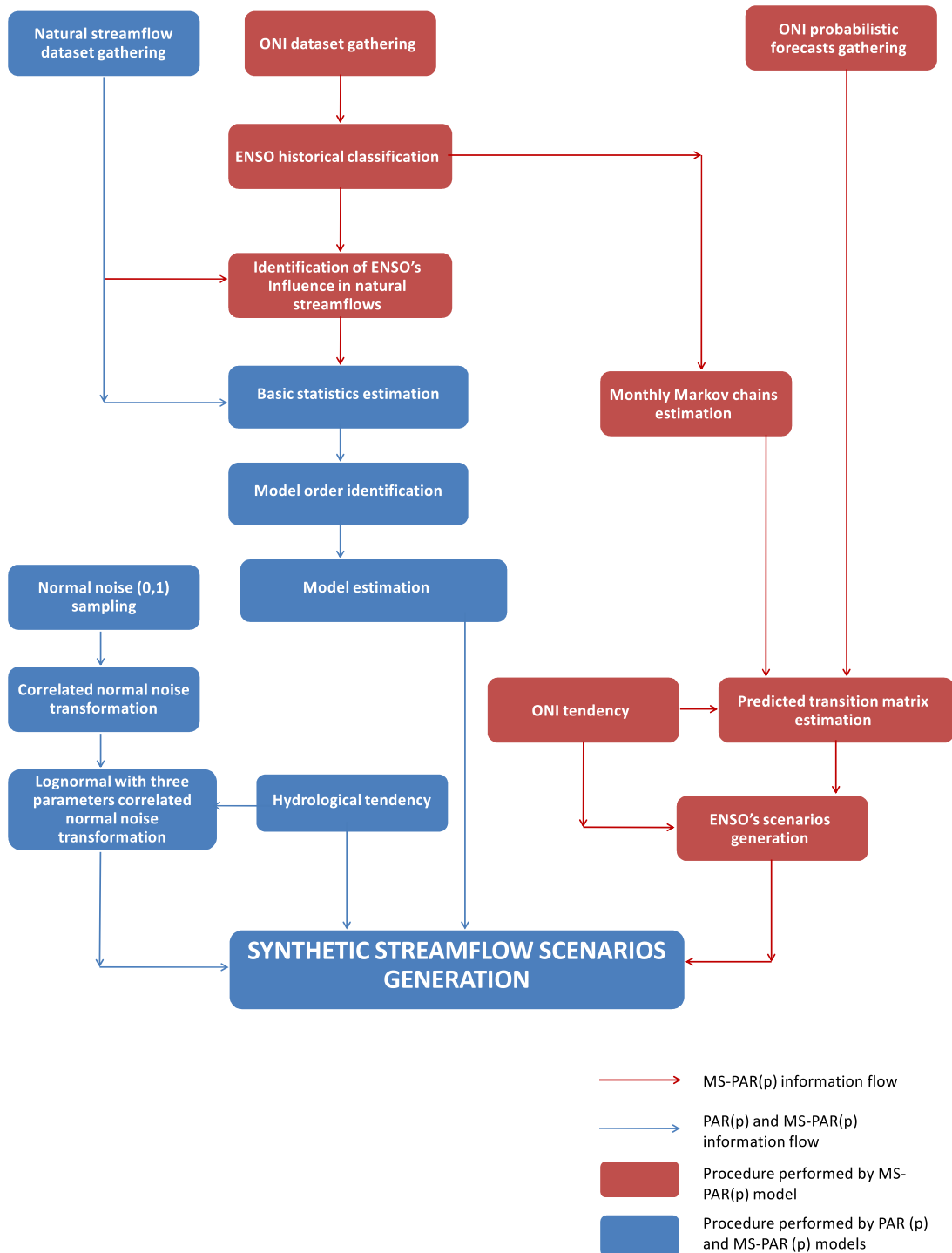


Figure A.13: Methodological flowchart.

1. if five consecutive seasons occur where the ONI is less than or equal to -0.5 (LN condition), the state of that period is classified as LN;
2. if five consecutive seasons occur where the ONI is greater than or equal to $+0.5$ (EN condition), the state of that period is classified as EN;
3. any other combination results in the N status classification.

In Figure A.14 areas in red (blue) represent episodes - or states - of EN (LN), while the line in red (blue) defines the boundary conditions of EN (LN). It should be noted that the classification criterion used is dependent on both the ONI past and future values. For example, until the end of 2017, the last seasons are classified as N, although the ONI presents values lower than -0.5 . After the beginning of 2018, the trend of LN conditions was maintained, characterizing an episode of LN.

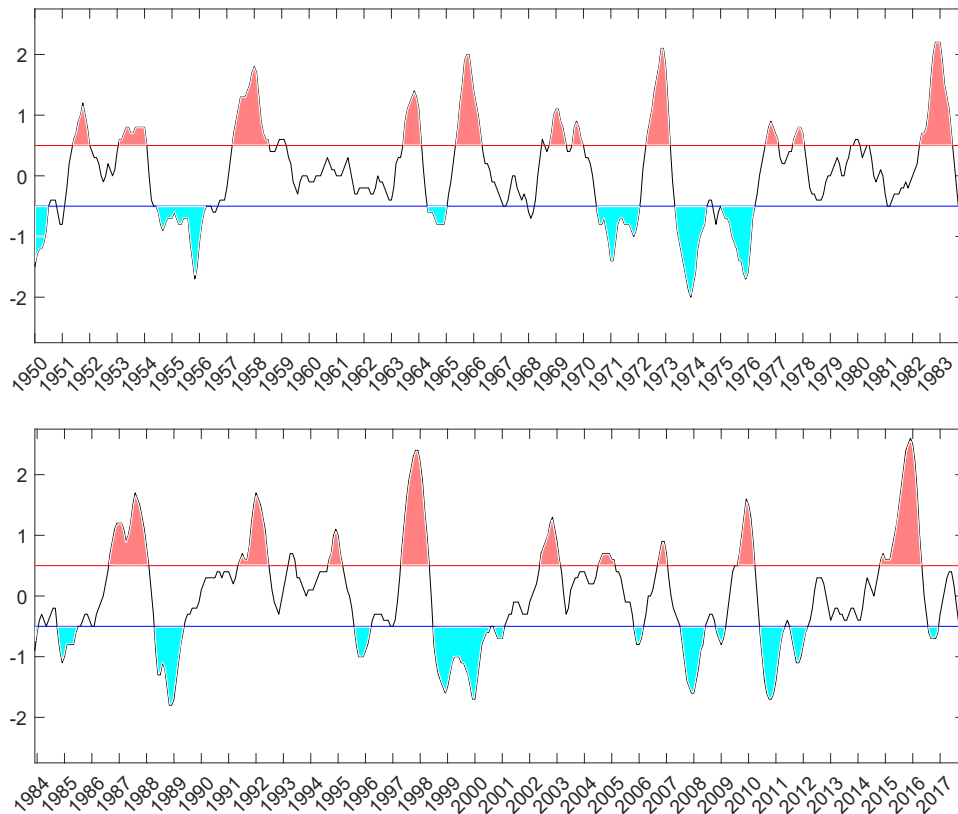


Figure A.14: ONI History. Areas in red (blue) represent episodes of EN (LN).

Aiming at its use in the GEVAZP model, which generates synthetic scenarios with monthly discretization, an adaptation of the ENSO classification criterion is performed. For the classification of the months between the three states of ENSO will be considered the middle month of each season as the representative. For example, the January/February/March quarter index (JFM) will be used to

rank February; March will be classified according to the index of the period February/March/April (FMA); and so on. The number of months classified between each state is shown in Table A.2.

Table A.2: Number of months classified between ENSO states between 1950 and 2017

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
La Niña	21	18	17	13	16	15	18	18	18	19	21	21
Neutro	23	29	37	43	37	38	35	35	30	27	23	23
El Niño	24	21	14	12	15	15	15	15	20	22	24	24

It should be noted that when using this procedure in simulations of official studies, the classification of the month before the case will always be unknown. For example, assuming that synthetic scenarios are being generated for February since January is classified by the ONI index of the DJF season, which depends on the SST in February (not yet observed), the value for DJF is non-existent. Then, for the month before the study, it will also be predicted its state, using the probabilistic forecasts provided by IRI.

El Niño - Southern Oscillation Predictions

The model proposed in this dissertation will use as input data the probabilistic forecasts provided by IRI, which are based on the predictions made by several dynamic and statistical models collected together. Forecasts are provided for ten seasons ahead. It is important to emphasize that these forecasts concern the classification of the ONI index among the three ENSO conditions. It should also be noted that the forecasts provided are made from the following season of the month in which they were made. For example, the first forecast provided at the end of March relates to the MAM season. According to the classification criterion used, this forecast would fit the classification of April. Therefore, for March, it is necessary to obtain the FMA quarter forecast, which was made available in February. Finally, for the estimation of the state of the month before the case (February), the forecast made available in January for the first quarter (JFM) is collected.

A.3.3 Periodic Autoregressive Model

The GEVAZP model is responsible for the generation of synthetic multivariate scenarios of energy and monthly streamflows for the NEWAVE and DECOMP models. The scenarios generated by GEVAZP represent the main uncertainty for

the planning of the NIS, and, therefore, are essential for the calculation of the optimum operation policy. In general, hydrological series of less than a year, such as monthly series, have as main characteristic the periodic behavior of their probabilistic properties, such as mean, variance, skewness and autocorrelation structure. The GEVAZP model is based on the stochastic model known as periodic autoregressive of order 'p' (HIPEL e MCLEOD, 1994; SALAS *et al.*, 1980). These models are referenced by PAR(p) models, where p is the order of the model, in other words, the number of auto-regressive terms of the model. The basic sample statistics of a streamflow series, which are used in PAR(p) modeling, are determined monthly according to the Equations (A.1), (A.2), (A.3) below:

$$\mu_m = \sum_{i=1}^N \frac{Z_t}{N} \quad (\text{A.1})$$

$$\sigma_m = \sqrt{\sum_{i=1}^N \frac{(Z_t - \mu_m)^2}{N}} \quad (\text{A.2})$$

$$\rho_m(k) = \frac{1}{N} \sum_{t=1}^N \left(\frac{Z_t - \mu_m}{\sigma_m} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \quad (\text{A.3})$$

where:

Z_t is the seasonal time series of period s ;

s is the number of periods ($s = 12$ for monthly series);

N is the number of years;

t is the time index, $t = 1, 2, \dots, sN$, function of year T ($T = 1, 2, \dots, N$) and period m ($m = 1, 2, \dots, s$);

μ_m is the seasonal mean of period s ;

σ_m is the seasonal standard deviation of period s ;

$\rho_m(k)$ is the correlation between Z_t and Z_{t-k} such that t corresponds to the period m .

The PAR(p) model can be mathematically described by A.4:

$$\left(\frac{Z_t - \mu_m}{\sigma_m} \right) = \phi_m^1 \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_m^{p_m} \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + a_t \quad (\text{A.4})$$

at where:

p_m is the order of the model;

$\phi_m^{p_m}$ is the autoregressive coefficient of order p_m ;

a_t is the series of independent noises with mean zero and variance $\sigma_a^{2(m)}$.

Thus, future synthetic scenarios, generated according to the PAR(p) model in the Equation (A.4), can be interpreted as a linear combination of past histori-

cal observations, plus a random term. The generation of synthetic series can then be performed through a successive sampling of a_t , each of these scenarios being a possible future realization of the stochastic process. In order to reduce the computational time in the energy optimization models, clustering techniques are applied to the noise vector a_t , reducing its cardinality, while maintaining the generation quality, as demonstrated in JARDIM *et al.* (2001).

Model Identification

The first step for fitting a PAR(p) model is the identification of the order p_m more appropriate for the autoregressive operators of each period. The criterion for identifying the order of the model through the partial autocorrelation function (PACF) was initially suggested by BOX e JENKINS (1970) for the autoregressive integrated moving average (ARIMA) family of stochastic models. This same approach can easily be extended to PAR(p) models. Given the Equation A.4, multiplying both sides by $\left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}}\right)$ and taking the expected value, we obtain for each period the expression (A.5):

$$E \left[\left(\frac{Z_t - \mu_m}{\sigma_m} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] = \phi_m^1 E \left[\left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] + \dots + \phi_m^{p_m} E \left[\left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] + E \left[a_t \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] \quad (\text{A.5})$$

Exemplifying, for $k = 1$, the above expression reduces to:

$$\rho_m(1) = \phi_m^1 + \phi_m^2 \rho_{m-1}(1) + \dots + \phi_m^{p_m} \rho_{m-1}(p_m - 1) \quad (\text{A.6})$$

By setting m and varying k from 1 to p_m in A.5 a set of equations commonly called Yule-Walker equations are defined for each period. For a period m :

$$\begin{bmatrix} 1 & \rho_{m-1}(1) & \rho_{m-1}(2) & \dots & \rho_{m-1}(p_m - 1) \\ \rho_{m-1}(1) & 1 & \rho_{m-2}(1) & \dots & \rho_{m-2}(p_m - 2) \\ \rho_{m-1}(2) & \rho_{m-2}(1) & 1 & \dots & \rho_{m-3}(p_m - 3) \\ \vdots & \dots & \vdots & \ddots & \vdots \\ \rho_{m-1}(p_m - 1) & \rho_{m-2}(p_m - 2) & \rho_{m-3}(p_m - 3) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_m^1 \\ \phi_m^2 \\ \phi_m^3 \\ \vdots \\ \phi_m^{p_m} \end{bmatrix} = \begin{bmatrix} \rho_m(1) \\ \rho_m(2) \\ \rho_m(3) \\ \vdots \\ \rho_m(p_m) \end{bmatrix} \quad (\text{A.7})$$

It is called ϕ_m^{kk} the last parameter of this autoregressive process of order k .

The Yule-Walker equations for each kk period m can be rewritten as follows:

$$\begin{bmatrix} 1 & \rho_{m-1}(1) & \rho_{m-1}(2) & \dots & \rho_{m-1}(k-1) \\ \rho_{m-1}(1) & 1 & \rho_{m-2}(1) & \dots & \rho_{m-2}(k-2) \\ \rho_{m-1}(2) & \rho_{m-2}(1) & 1 & \dots & \rho_{m-3}(k-3) \\ \vdots & \dots & \vdots & \ddots & \vdots \\ \rho_{m-1}(k-1) & \rho_{m-2}(k-2) & \rho_{m-3}(k-3) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_m^{k1} \\ \phi_m^{k2} \\ \phi_m^{k3} \\ \vdots \\ \phi_m^{kk} \end{bmatrix} = \begin{bmatrix} \rho_m(1) \\ \rho_m(2) \\ \rho_m(3) \\ \vdots \\ \rho_m(k) \end{bmatrix} \quad (\text{A.8})$$

Each of the resulting systems can be solved through the Cholesky Decomposition. It is called the partial autocorrelation function of the period m the set of values ϕ_m^{kk} . The set of parameters ϕ_m^{kk} , $m = 1, 2, \dots, s$, is another way of representing the structure of dependence of the stochastic process over time, and can be understood as a measure of lag k correlation, disregarding the influence of the other orders. In an autoregressive process of order p_m , the partial autocorrelation function ϕ_m^{kk} will be nonzero for k less than or equal to p_m and zero for k greater than p_m .

The identification of the order p_m by the autocorrelation function is performed by comparing the values of ϕ_m^{kk} to a given confidence interval (95 % for example). performed in two ways:

1. It is identified the first significant ϕ_m^{kk} , making the recursion of the largest possible order to the smallest;
2. It is identified last significant ϕ_m^{kk} before a non-significant order.

Both ways of identification are correct from the theoretical point of view. Some researchers defend the use of the former, with the justification for making the most of the information contained in the historical series, which could be lost by using the second criterion. Researchers who suggest the use of the second process, believe that because it is more parsimonious, it can produce better results, considering spurious the significant partial autocorrelations of a higher order. The GEVAZP model follows the first line of identification, and this will be the adopted procedure in this dissertation. Since GEVAZP is a generator of synthetic inflow scenarios, also used for medium and long-term planning, it is expected that it will be able to reproduce the multivariate dry periods observed throughout history. Thus, it is more plausible to suppose that models with larger orders have greater adherence to historical annual autocorrelation, capable of mimic critical periods of history.

Model Estimation

After the step of identifying the order p_m of the model, one must obtain the estimates of the parameters of the Equation (A.4). For this step, the estimators

obtained by the moments method are used. The means and the monthly standard deviations are obtained through the Equations (A.1) and (A.2) respectively. The parameters ϕ_m^i , $i = 1, \dots, p_m$ can be obtained by substituting in (A.6) the estimatives $\rho_{m-j}(k)$, $j = 0, \dots, (p_m - 1)$ obtained in (A.3).

Finally, the last parameter to be estimated is the variance $\sigma_a^{2(m)}$ of the independent noises a_t . Given the Equation (A.5), for lag $k = 0$:

$$1 = \phi_m^1 \rho_m(1) + \phi_m^2 \rho_m(2) + \dots + \phi_m^{p_m} \rho_m(p_m) + E \left[a_t \left(\frac{Z_t - \mu_m}{\sigma_m} \right) \right] \quad (\text{A.9})$$

Multiplying both sides of the Equation (A.4) by a_t and taking the expected value gives:

$$E \left[a_t \left(\frac{Z_t - \mu_m}{\sigma_m} \right) \right] = \sigma_a^{2(m)} \quad (\text{A.10})$$

Substituting (A.10) in (A.9):

$$\sigma_a^{2(m)} = 1 - \phi_m^1 \rho_m(1) - \phi_m^2 \rho_m(2) - \dots - \phi_m^{p_m} \rho_m(p_m) \quad (\text{A.11})$$

valid for any period m .

Spatial Correlations

One of the main advantages of the NIS lies in the hydrological complementary between the hydrographic regions that compose it. Thus, this characteristic must be properly incorporated in the synthetic scenario generation model, so that the series are generated in a multivariate manner, taking into account the spatial correlations between the various HPPs or EERs. At first, the PAR(p) model generates series univariate. However, a spatially uncorrelated residue b_t can be transformed so that they begin to consider the spatial correlation, resulting in residues W_t :

$$W_t = D b_t \quad (\text{A.12})$$

where D is the square matrix of dimension equal to the number of HPPs (or EERs) that spatially correlates the residues, usually called the load matrix. The load matrix D , in turn, is a transformation obtained from the spatial correlation matrix U . In order to obtain this, one can use, for example, a technique known as the Jacobi Method (PRESS *et al.*, 1992) to perform a spectral decomposition, in which the (annual or monthly) spatial correlation matrix (U) is decomposed in a diagonal matrix containing its eigenvalues (A) and a symmetric matrix of eigenvectors (V),

according to the expressions (A.13) and (A.14) :

$$U = AVA' \quad (\text{A.13})$$

$$D = A\sqrt{V}A' \quad (\text{A.14})$$

Negative Incremental Streamflows

Time series such as monthly natural flows usually do not have negative values. Thus, during the process of generating synthetic scenarios one must impose the necessary conditions to obtain a positive Z_t :

$$Z_t = \mu_m + \phi_m^1 \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_m^{p_m} \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + a_t \sigma_m > 0 \quad (\text{A.15})$$

or,

$$a_t > - \left(\frac{\mu_m}{\sigma_m} \right) - \phi_m^1 \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) - \dots - \phi_m^{p_m} \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \quad (\text{A.16})$$

Natural streamflow series, and hence natural energy, usually have a non-normal probability distribution. A possible solution to the problem of generation of negative values, and to deal with the non-normality of the residues of the time series Z_t , would be the use of the transformation BOX e COX (1964). As one of the objectives of the GEVAZP model is to generate monthly synthetic scenarios of natural energy for the NEWAVE model, it is necessary to observe the linearity requirement of the input data for this model. Thus, the use of the non-linear transformation BOX e COX (1964) of the time series Z_t must be discarded.

The solution adopted to treat the strong asymmetries of residuals a_t , and to prevent the generation of negative scenarios, is the adjustment of a Lognormal distribution with three parameters (CHARBENEAU, 1978) to the monthly residuals, where the third parameter Δ_t , called the offset, is calculated in order to avoid the generation of negative scenarios, as presented in Equation (A.16). For each month, the variable ζ_t has Normal distribution, with mean μ_{ζ_t} , variance $\sigma_{\zeta_t}^2$, and offset Δ_t .

$$\zeta_t = \ln(a_t - \Delta_t) \quad (\text{A.17})$$

The parameters of the Lognormal distribution are estimated in order to preserve the moments of the residues a_t , according to the equations:

$$\mu_{\zeta_t} = \frac{1}{2} \ln \left(\frac{\sigma_a^2}{\theta^2 - \theta} \right) \quad (\text{A.18})$$

$$\sigma_{\zeta_t}^2 = \ln(\theta) \quad (\text{A.19})$$

$$\Delta_t = \frac{-\sigma_a}{\sqrt{(\theta - 1)}} \quad (\text{A.20})$$

Equating the Equation (A.16) to the offset Δ_t , shown in the previous expression, one can obtain:

$$\theta = 1 + \frac{\sigma_a^2}{(\mu_a - \Delta_t)^2} \quad (\text{A.21})$$

Generation of Synthetic Series with the PAR(p) Model

After adjusting all the parameters of the model, it is possible to generate the synthetic series. Each generated scenario can be understood as a possible realization of the stochastic modeling process. We can rewrite the Equation (A.4), making it more convenient for the sequential generation of scenarios Z_t , obtained as a function of past values, plus a random term, A.22:

$$Z_t = \mu_m + \phi_m^1 \sigma_m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_m^{p_m} \sigma_m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + \sigma_m \sigma_{\zeta_t}^{(m)} E_t \quad (\text{A.22})$$

where the noise series E_t is independent with mean zero and unit variance.

To begin the process of generating scenarios, it is necessary to know the past values of $Z_{t-\tau}$, $\tau = 1, \dots, p_m$. The past values $Z_{t-\tau}$ can be obtained through previously observed values of the series, usually called the hydrological tendency when the stochastic process is a hydrological series, initiating what is called a conditioned (to the recent past) generation. The conditioned generation process is illustrated in Figure A.15. In this type of generation, all the scenarios generated at time $t = 1$ start from the same common past.

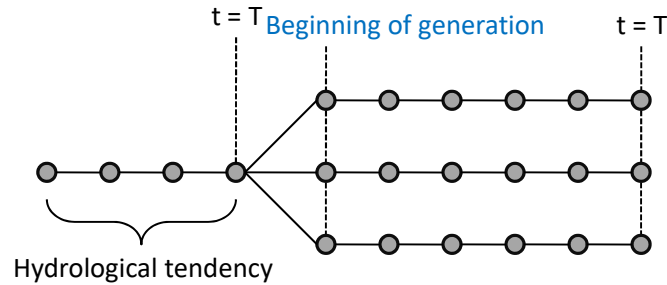


Figure A.15: Synthetic scenarios generation conditioned to the recent past.

Synthetic scenarios generation conditioned to the recent past.

The past values $Z_{t-\tau}$ can also be obtained so that each scenario generated at time $t = 1$ comes from a distinct past, thus characterizing an unconditioned

generation. Unconditioned generation is performed in two stages. In the first step, the values of $Z_{t-\tau}$ are arbitrated as being the monthly average of the historical series, commonly called the long-term mean (LTM). From these values, synthetic series are generated for a given horizon, for example, five years. After this time, each synthetic series follow a distinct trajectory, possessing different passages. The values generated for the last periods of the first step will serve as initial values for the second step. The unconditioned generation process is illustrated in Figure A.16.

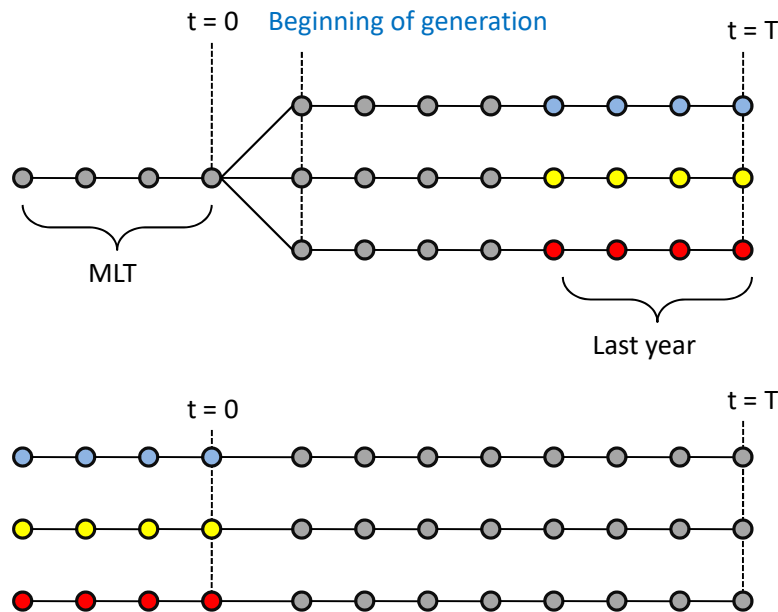


Figure A.16: Synthetic scenarios unconditioned generation.

A.3.4 Markov-Switching Periodic Autoregressive - MS-PAR(p)

The proposed modeling in this dissertation aims to improve the GEVAZP model through the consideration of information about the ENSO phenomenon. To achieve this objective, a new type of stochastic model is proposed: Markov-Switching Periodic Auto-Regressive - MS-PAR(p). It can be understood as a combination of PAR(p) modeling in which its parameters are switched according to some states that follow a Markov chain (WILKS, 2011). The MS-PAR(p) model is a natural evolution of the model family started in HAMILTON (1989). The main idea presented here is to identify the different ENSO states in historical data, and make a segmented estimation of the parameters used in the periodic autoregressive part of the model. Thus, in addition to the parameters being estimated according to the month in question, they will also vary according to the states of LN, N, and EN. It should be noted that unlike the models proposed by HAMILTON (1989), KROLZIG (1997),

and GELATI *et al.* (2011) the MS-PAR(p) model presented in this dissertation, despite being part of the same family of models, does not follow a hidden Markov chain for its states. The states are considered an observed stochastic process of the ENSO states series. More details on Markov chain modeling will be seen later in the A.3.4 subsection.

Thus, the MS-PAR(p) model proposed in this dissertation has the following mathematical formulation:

$$\left(\frac{Z_t - \mu_m^{\varepsilon_t}}{\sigma_m^{\varepsilon_t}} \right) = \phi_m^1 \left(\frac{Z_{t-1} - \mu_{m-1}^{\varepsilon_{t-1}}}{\sigma_{m-1}^{\varepsilon_{t-1}}} \right) + \dots + \phi_m^{p_m} \left(\frac{Z_{t-p_m} - \mu_{m-p_m}^{\varepsilon_{t-p_m}}}{\sigma_{m-p_m}^{\varepsilon_{t-p_m}}} \right) + a_t \quad (\text{A.23})$$

where:

- Z_t is the seasonal time series of period s ;
- ε_t is the seasonal discrete time series of period s with state-space $r = 1, 2, \dots, d$ ($d = 3$ for the ENSO representation), which follows a homogeneous Markov chain;
- s is the number of periods ($s = 12$ for monthly series);
- N is the number of years;
- t is the time index, $t = 1, 2, \dots, sN$, function of year T ($T = 1, 2, \dots, N$) and period m ($m = 1, 2, \dots, s$);
- $\mu_m^{\varepsilon_t}$ is the seasonal mean of period m for the state of ε_t ;
- $\sigma_m^{\varepsilon_t}$ is the seasonal standard deviation of period m for the state of ε_t ;
- p_m is the order of the model for period m ;
- $\phi_m^{p_m}$ is the autoregressive coefficient of order p_m for period m ;
- a_t is the series of independent noises with mean zero and variance $\sigma_a^{2(m)}$.

For the MS-PAR(p) model described in the expression (A.23), it should be noted that the mean and standard deviation parameters alternate according to the period m and the state of ε_t , unlike the autoregressive parameter $\phi_m^{p_m}$, which only varies with period m . It was chosen to present here a formulation closer to the reality of the stochastic process of the streamflow. While the mean and standard deviation may change according to the state of the ENSO phenomenon, due to the increase or decrease of precipitation, which will eventually become streamflow, the autoregressive parameter $\phi_m^{p_m}$ is related to the response of the drainage area of the modeled HPP offers to past events. In other words, the parameter $\phi_m^{p_m}$ because it statistically represents a physical characteristic of the drainage basin, should not be modified due to the ENSO phenomenon, and, therefore, it is not switched according to the state of ε_t . Here, we define the state space of ε_t as $r = \{1, 2, 3\}$, representing, respectively, the states LN, N and EN.

For the estimation of the mean and standard deviation of the sample, calculated for each month and state, and for the sample autocorrelation, monthly variable

only, we present the Equations (A.24), (A.25) and (A.26). It is interesting to note that the formulation of the MS-PAR(p) model brings some Bayesian statistical concepts for the traditional statistical modeling, used in Box-Jenkins models, such as the PAR(p) model. The calculation of the mean and standard deviation is performed with prior information on the state space.

$$\mu_m^r = E(Z_t | \varepsilon_t = r) = \sum_{t=1}^N \mathbf{1}_{(\varepsilon_t=r)} \frac{Z_t}{N_m^r} \quad r = 1, 2, 3 \quad (\text{A.24})$$

$$\sigma_m^r = E(Z_t | \varepsilon_t = r) = \sqrt{\sum_{t=1}^N \mathbf{1}_{(\varepsilon_t=r)} \frac{(Z_t - \mu_m^r)^2}{N_m^r}} \quad r = 1, 2, 3 \quad (\text{A.25})$$

$$\rho_m(k) = \frac{1}{N} \sum_{t=1}^N \left(\frac{Z_t - \mu_m^r}{\sigma_m^r} \right) \left(\frac{Z_{t-k} - \mu_{m-k}^r}{\sigma_{m-k}^r} \right) \quad (\text{A.26})$$

where $\mathbf{1}_{(\varepsilon_t=r)}$ denotes the indicator function, assuming the unit value when $\varepsilon_t = r$ and zero for any other value and N_m^r is the number of occurrences of state r in the month m .

Defined the mathematical formulation of the MS-PAR model (p) the process of order identification and parameter estimation, as well as the process of noise transformations in Lognormal three parameters and spatially correlated, will follow the same steps of the model presented in section A.3.3. Thus, the proposed model MS-PAR(p) is a methodological improvement of the GEVAZP model. With the advent of its implementation, information on the ENSO phenomenon will be incorporated into the generation of synthetic scenarios of streamflows and monthly energies focused on the NIS. It is expected, therefore, to be able to determine with greater accuracy the multivariate distributions of probabilities for future inflows. This new model alters the concept of autoregressive modeling that takes into account only the past values of the time series to predict future steps. Given the expected state r for the month m at time t , a given set of parameters is used to generate the scenarios, effectively capturing its historical distribution function, which is different from the other states.

The proposed model has some advantages over PARX models that include weather information. First, in PARX models, climate information only changes the level of the expected value. As will be seen in the next chapter, different ENSO states produce distinct variabilities for streamflow, changing the frequency distribution function of the streamflo. Since the GEVAZP model is a model for the generation of synthetic series of inflows and not a predictive model, it is attractive to be able to identify different probability distribution functions. In the same way that a single AR model can capture less information about the seasonal stochastic process

of random variables than a PAR(p) model, it is expected that an MS-PAR(p) bring similar benefits. In identifying distinct probability distribution functions within the same period, the stochastic process can be better explored. The second, and most important point, concerns the horizon of possible generation. As the exogenous variables are inserted directly into the PARX models, many times through deterministic predictions, scenario generation is restricted to the forecast horizon of the exogenous variable. In the MS-PAR(p) model proposed here, streamflow scenarios are generated with the autoregressive part of the model, together with ENSO state scenarios, through the Markov chain. This alternative makes the model more flexible, being able to be used throughout the desired horizon. So it is a natural candidate for the generation of mid-term scenarios and also for the long-term horizon. Another essential point of the MS-PAR(p) model is the maintenance of the linear modeling of the flows, being compatible its use with the NEWAVE model.

The development presented in this dissertation coincided with the article by ALIAT e HAMDI (2018), the first to introduce the MS-PARMA model family, an even more generalized formulation with the inclusion of terms related to moving averages. In this article, it is also discussed relevant aspects about the stationarity and ergodicity of the MS-PARMA models, essential for the correct operation of a linear stochastic model, as well as small practical examples that show the potential of the model. Likewise, it can be affirmed that the work developed here has an innovative character of building an MS-PAR(p) model applied to a real case of high complexity such as the modeling of NIS inflows. Also, there are some significant methodological differences due to the specificity of the problem analyzed here, where the states are observable in the history (ENSO states), being modeled through a Markov chain (not hidden).

Markovian Process Modeling

As the final objective of the proposed MS-PAR(p) model is to generate scenarios of monthly streamflows with information about the ENSO phenomenon for the mid and long-term horizon, it is necessary that the scenario generation is also done for ENSO events. The scenarios generated for the ENSO phenomenon must have the same temporal discretization of the flow scenarios (monthly basis).

The time series of states of the ENSO phenomenon can be understood as a discrete variable ε_t with state space r , related to its three possible states (LN, N and EN). The model class most commonly used to represent the time series of a discrete variable is known as Markov chain (WILKS, 2011). The behavior of a Markov chain is driven by a set of transition probabilities between its states, \mathbf{P} , the simplest form being the Markov chain of order one (which is also na autoregressive of order one). In other words, the next state depends only from the most recent state, independent

of the sequence of previous states, as presented in Equation (A.27):

$$P(\varepsilon(t) = r \mid \varepsilon_{(t-1)}, \varepsilon_{(t-2)}, \dots, \varepsilon_1) = P(\varepsilon(t) = r \mid \varepsilon_{(t-1)}) \quad (\text{A.27})$$

The transition probabilities are probabilities conditioned with respect to the most recent state $t-1$. As the ENSO phenomenon will be divided among its three possible states the transition matrix \mathbf{P}_m will be of 3x3 dimension.

$$\mathbf{P}_m = \begin{bmatrix} p_m^{1,1} & p_m^{1,2} & p_m^{1,3} \\ p_m^{2,1} & p_m^{2,2} & p_m^{2,3} \\ p_m^{3,1} & p_m^{3,2} & p_m^{3,3} \end{bmatrix} \quad (\text{A.28})$$

Given that there is no direct transition between states from EN to LN (as well as conditions), the elements $p_m^{1,3}$ and $p_m^{3,1}$ will always be null .

For each month m it will be estimated the transition matrix for the ENSO states, as well as the transition matrix of ENSO conditions, both using the following mathematical expression:

$$p_m^{i,j} = P(\varepsilon(t) = j \mid \varepsilon_{(t-1)} = i)_m = \frac{N_m^{i,j}}{N_{m-1}^i} \quad (\text{A.29})$$

where $p_m^{i,j}$ is the conditional probability of moving to state j in month m given that the previous month $m-1$ was in state i ; $N_m^{1,2}$ is the number of times the transition from state i to state j occurred in the month m in the history; and N_{m-1}^i is the number of occurrences in the history of state i in month $m-1$. The transition matrix \mathbf{P}_m can be estimated according to the Equation (A.28) by varying the state indices.

Inclusion of Probabilistic El Niño - Southern Oscillation Forecast

Aiming at the best use of the MS-PAR(p) model proposed in the energy operation planning studies of the NIS, it is presented a methodology for the inclusion of probabilistic forecasts of ENSO in the model, that produces predicted transition matrices between states. In real cases, even the state of previous months may not be known, since the ENSO classification depends on both past and future ONI values, so it is necessary to estimate the probability for past values. For example, if the last four indices were classified as EN conditions, if the next month were also so, they will be effectively classified as EN states; however, if the next ONI value is less than +0.5, this period as a whole will be classified as neutral. In this context, the use of probabilistic forecasts is advantageous.

As presented in section A.3.2, IRI provides probabilistic forecasts for ONI

index conditions up to nine quarters ahead. At first, given the predictions made available of LN, N and EN conditions, denoted by $P_c(LN)_m^{prev}$, $P_c(N)_m^{prev}$, and $P_c(EN)_m^{prev}$, respectively, one can not infer on the conditional transition probabilities between ENSO conditions. However, it is possible to calculate a predicted transition matrix as close as possible to the historical transition matrices between the conditions, respecting the probability distribution predicted by IRI. In order to differentiate the values referring to the ENSO conditions, the sub-index "c" will be added in the notations, avoiding possible confusions between the concepts of condition and state of ENSO. Thus, the calculation of the monthly predicted transition matrices between the ENSO conditions can be understood as an optimization problem whose objective is to minimize the differences (Δ_i) between the historical and the expected matrix:

$$\min \sum_{i=1}^{14} \Delta_i \quad (\text{A.30})$$

s.t.

$$\begin{aligned} P_c(LN | LN)_m^{prev} &= P_c(LN | LN)_m^{hist} + \Delta_1^+ - \Delta_2^- \\ P_c(LN | N)_m^{prev} &= P_c(LN | N)_m^{hist} + \Delta_3^+ - \Delta_4^- \\ P_c(N | LN)_m^{prev} &= P_c(N | LN)_m^{hist} + \Delta_5^+ - \Delta_6^- \\ P_c(N | N)_m^{prev} &= P_c(N | N)_m^{hist} + \Delta_7^+ - \Delta_8^- \\ P_c(N | EN)_m^{prev} &= P_c(N | EN)_m^{hist} + \Delta_9^+ - \Delta_{10}^- \\ P_c(EN | N)_m^{prev} &= P_c(EN | N)_m^{hist} + \Delta_{11}^+ - \Delta_{12}^- \\ P_c(EN | EN)_m^{prev} &= P_c(EN | EN)_m^{hist} + \Delta_{13}^+ - \Delta_{14}^- \end{aligned}$$

Restrictions on the sum of conditional probabilities

$$\begin{aligned} P_c(LN | LN)_m^{prev} + P_c(N | LN)_m^{prev} &= 1 \\ P_c(LN | N)_m^{prev} + P_c(N | N)_m^{prev} + P_c(EN | N)_m^{prev} &= 1 \\ P_c(N | EN)_m^{prev} + P_c(EN | EN)_m^{prev} &= 1 \end{aligned}$$

Restrictions on the limits of conditional probabilities

$$\begin{aligned}
0 &\leq P_c(LN | LN)_m^{prev} \leq 1 \\
0 &\leq P_c(LN | N)_m^{prev} \leq 1 \\
0 &\leq P_c(N | LN)_m^{prev} \leq 1 \\
0 &\leq P_c(N | N)_m^{prev} \leq 1 \\
0 &\leq P_c(N | EN)_m^{prev} \leq 1 \\
0 &\leq P_c(EN | N)_m^{prev} \leq 1 \\
0 &\leq P_c(EN | EN)_m^{prev} \leq 1
\end{aligned}$$

Restrictions on the reproduction of forecasts made available by IRI

$$\begin{aligned}
P_c(LN)_m^{prev} &= P_c(LN)_{m-1}^{prev} \times P_c(LN | LN)_m^{prev} + P_c(N)_{m-1}^{prev} \times P_c(LN | N)_m^{prev} \\
P_c(N)_m^{prev} &= P_c(LN)_{m-1}^{prev} \times P_c(N | LN)_m^{prev} + P_c(N)_{m-1}^{prev} \times P_c(N | N)_m^{prev} + \\
&\quad P_c(EN)_{m-1}^{prev} \times P_c(N | EN)_m^{prev} \\
P_c(EN)_m^{prev} &= P_c(EN)_{m-1}^{prev} \times P_c(EN | EN)_m^{prev} + P_c(N)_{m-1}^{prev} \times P_c(EN | N)_m^{prev}
\end{aligned}$$

The above optimization problem has a linear objective function, as well as all its constraints, characterizing itself as a linear programming problem. This problem can be solved through a simplex algorithm, for example.

Solved the above optimization problem for all months in which ENSO condition predictions exist, one can then estimate the predicted transition matrix between ENSO states. The calculation of the ENSO state transition matrix follows the same criteria used for the historical classification, i.e.:

1. If the previous state is classified as LN:
 - (a) the probability of persisting as LN will be equal to the probability of persisting the LN condition;
 - (b) the probability of transiting to the N state will equal the probability of passing the condition N since it was in a LN.
2. If the previous state is classified as N:
 - (a) the probability of transit to a state LN will be equal to the product of going to the condition LN in month m and its persistence in the condition LN for another four months;
 - (b) the probability of transit to a state LN will be equal to the product of going to the condition LN in month m and its persistence in the condition LN for another four months;

- (c) the probability of persisting as neutral is given by the complement of the two previous options.
3. If the previous state is classified as EN:
- (a) the probability of persisting as EN will be equal to the probability of persisting the EN condition;
- (b) the probability of transiting to the N state will equal the probability of passing the condition N given that it was in an EN.

Mathematically, one can summarize the previous criteria according to the following expressions:

$$\begin{aligned}
P(LN | LN)_m^{prev} &= P_c(LN | LN)_m^{prev} \\
P(N | LN)_m^{prev} &= P_c(N | LN)_m^{prev} \\
P(LN | N)_m^{prev} &= P_c(LN | N)_m^{prev} \times \prod_{i=1}^4 P_c(LN | LN)_{m+i}^{prev} \\
P(EN | N)_m^{prev} &= P_c(EN | N)_m^{prev} \times \prod_{i=1}^4 P_c(EN | EN)_{m+i}^{prev} \\
P(N | N)_m^{prev} &= 1 - P(EN | N)_m^{prev} - P(LN | N)_m^{prev} \\
P(EN | EN)_m^{prev} &= P_c(EN | EN)_m^{prev} \\
P(N | EN)_m^{prev} &= P_c(N | EN)_m^{prev}
\end{aligned} \tag{A.31}$$

As long as there are predictions available, the model will calculate the predicted matrices \mathbf{P}_m^{prev} . For the calculation of $P(LN | N)_m^{prev}$ and $P(EN | N)_m^{prev}$, when there is no forecast for the terms of months $m + 1$, $m + 2$, $m + 3$ and $m + 4$, these will be replaced by historical values. With this, it is intended to make the most of the probabilistic forecast available.

Generating ENSO Scenarios

After estimating the monthly transition matrices for the states and conditions of the ENSO phenomenon, its generation of scenarios can be performed directly for the entire study horizon. Given the state or historical transition matrix, the state of the next period can be easily sampled. In order to maintain consistency with the classification criterion of the states, if a scenario moves from a state N to LN or EN, it will remain in that state for another four periods at least, thus accounting for the persistence of at least five periods. As in the generation of inflow scenarios, the generation of ENSO scenarios can be performed in an unconditioned way, or conditioned to the recent past.

Unconditioned Generation

The process of unconditioned generation of ENSO scenarios follows the same principle used in the unconditioned generation of inflows, dividing the process into two stages. First, we draw from the historical probability distribution of the states the values for the month $m - 1$. After this step, ENSO scenarios are generated using the historical Markov chain to define the transition or persistence probabilities of the states, for five years. After this stage, the scenarios generated for the last 11 months (maximum value of the order of the MS-PAR(p) model) are stored, serving as a background for the generation of scenarios for the second generation stage.

Conditioned Generation

For the conditional generation of states, the MS-PAR(p) model, applied to other types of problems, could generate scenarios of states directly, starting from the last observed state. However, the application in this dissertation aims at generating scenarios of ENSO states. As described in the A.3.2 section, the state of ε_{t-1} will never be known, since it depends on the ONI value that takes into account the SST of period t . Also, the ENSO classification criterion itself is dependent on future values. However, the state in ε_{t-1} can be estimated probabilistically, using the predictions provided by IRI.

To estimate the state $t - 1$, the following procedure is performed:

1. It is verified the occurrence of ENSO conditions in the period $t - \tau$, $\tau = 2,3,4,5$ to identify a trend of EN or LN formation;
2. it is summed the number of past periods from $\tau - 2$ (called δ) in which conditions of EN or LN were observed next, after occurrence of a condition N, ;
3. If δ is equal to zero, the condition in $t - 2$ has been classified as neutral.

Thereby:

- (a) The probability of the state ε_{t-1} will be equal to the probability of the commodity to move to the condition EN in the period $t - 1$ and its persistence in the condition LN for four more periods, from t to $t + 3$;
- (b) The probability of the state ε_{t-1} will be equal to the probability that the product will transition to the condition LN in the period $t - 1$ and its persistence in the condition LN for four more periods , from t to $t + 3$;
- (c) The probability of persisting as neutral is given by the complement of the two previous options;

4. δ is greater than zero, the conditions $t - \tau, \tau=2, \dots, \delta+1$ have been classified as LN (or EN). Thereby:

- (a) The probability of the state ε_{t-1} being LN (or EN) will be equal to the output of its persistence for more $5 - \delta$ periods;
- (b) The probability of the state ε_{t-1} is N is given by the complement of the previous options;

Summarizing the previous procedure mathematically, for the previous condition N:

$$\begin{aligned}
 P(LN)_{t-1}^{prev} &= P_c(LN | N)_{t-1}^{prev} \times \prod_{i=0}^3 P_c(LN | LN)_{t+i}^{prev} \\
 P(EN)_{t-1}^{prev} &= P_c(EN | N)_{t-1}^{prev} \times \prod_{i=0}^3 P_c(EN | EN)_{t+i}^{prev} \\
 P(N)_{t-1}^{prev} &= 1 - P(LN)_{t-1}^{prev} - P(EN)_{t-1}^{prev}
 \end{aligned} \tag{A.32}$$

For cases where the previous condition is LN or EN, and given δ , the number of conditions of the same classification immediately preceding the period $t - 1$, we have:

$$\begin{aligned}
 P(r)_{t-1}^{prev} &= \prod_{i=0}^{4-\delta} P_c(r | r)_{t+i}^{prev} \\
 P(N)_{t-1}^{prev} &= 1 - P(r)_{t-1}^{prev}
 \end{aligned} \tag{A.33}$$

being $r = 1$ for the state LN and $r = 3$ for the state EN.

Calculating the probability distribution of the states in ε_{t-1} , a random sampling of the predicted distribution is made to define the past state of each synthetic series. The procedure for generating scenarios from later periods will be done by random draws following the expected transition matrices \mathbf{P}_m^{prev} as long as there are predicted data available, and later if necessary, matrices of historical transition \mathbf{P}_m^{hist} .

Identification of El Niño - Southern Oscillation Influence in the Streamflows

Aiming to provide a more parsimonious model, the first evaluation to be carried out concerns the identification of the influence of the ENSO phenomenon on the historical series of monthly inflows of each HPP. During the literature review carried out in chapter A.2.2, several methodologies were presented to identify the influence of ENSO on streamflow. As this dissertation aims to develop a methodology for the

generation of synthetic scenarios of inflows conditioned to the ENSO phenomenon, the identification of the influence of ENSO will be verified in the cumulative distribution functions obtained through the segmentation of the flow history between its different states. For this, the non-parametric adherence test called Kolmogorov-Smirnov (KS) (PINTO *et al.*, 1976) will be performed. Under the null hypothesis of equality between two distributions, we find the maximum absolute difference $W = \text{Max}_x |S_1(x) - S_2(x)|$, where $S_1(x)$ and $S_2(x)$ are the cumulative frequency distributions. For samples of sufficiently large size, the critical value of the test at a significance of 10% is approximately given by:

$$W_{critical} = 1.22 \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \quad (\text{A.34})$$

When $W > W_{critical}$ the null hypothesis of equality between the distributions is rejected. The KS test is sensitive to differences in mean and standard deviation between distributions, which makes it interesting in the context of scenario generation. The test will be performed every month, for all possible combinations of states, and in all HPPs, through the procedure exemplified in Figures A.17 and A.18.

First, the KS test is used to compare the distributions between the sample of the EN series and the sample of the other two states together (LN and N). If the null hypothesis is not rejected (the two distributions are equal), the sample of the LN series is tested against the other two states together (EN and N); if the first null hypothesis is rejected (the two distributions are different) the separation of the remaining sample between LN and N is tested. After this first sequence of tests, the procedure is again performed, but having as the first KS test made for the LN sample against the sample of the other two states together (EN and N). If the result of the first procedure indicates only the influence of the EN and the second procedure only signals the influence of the LN, it is not rejected the null hypothesis that both phases of the ENSO are influential.

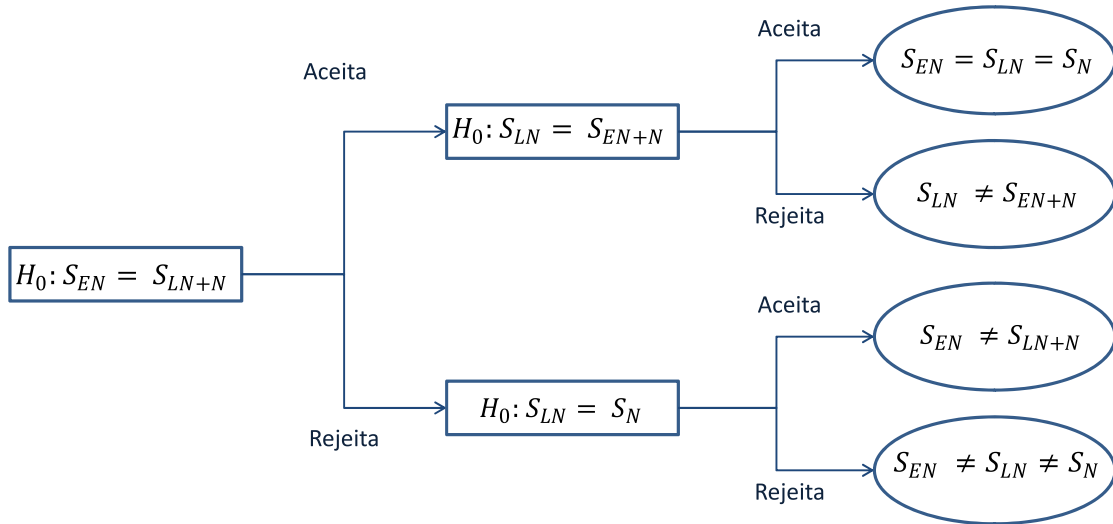


Figure A.17: Procedure to identify the influence of ENSO on the natural flows affluent to the HPPs - part I

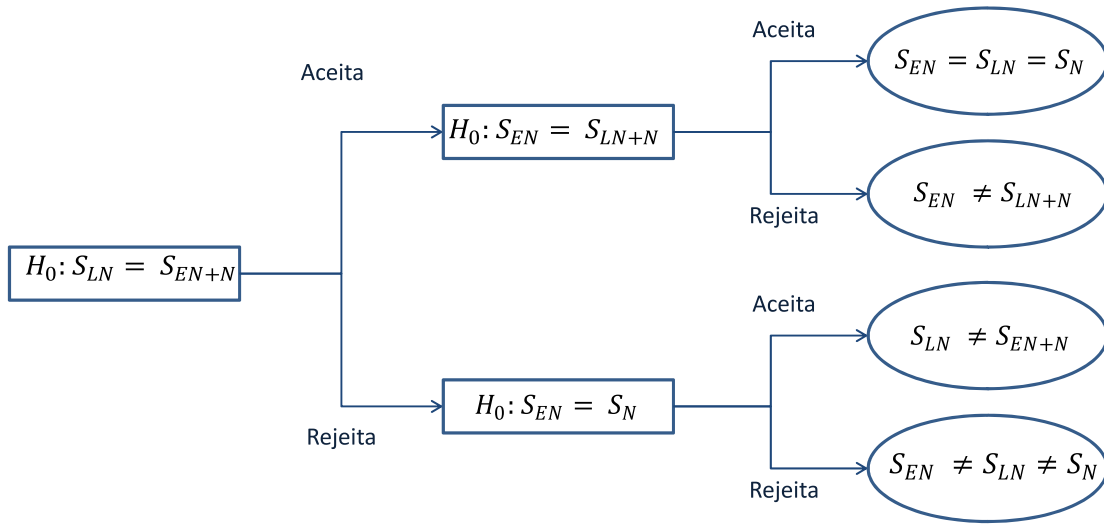


Figure A.18: Procedure to identify the influence of ENSO on natural inflows to HPPs - Part II

Thus, if in any given month none of the ENSO phases are relevant to the scenario generation, the MS-PAR(p) model for that month will behave as the PAR(p), having the same mean and standard deviation in all the ENSO states. If only one of the phases is influential, the other two phases will behave as one. For example, if in January it was identified that the EN has relevance over the streamflows, but the LN behaves similarly to the N state, the last two states will have the same historical statistics. It is believed that this prior identification of the influence of ENSO on natural streamflows produces more parsimonious models than MS-PAR(p) without this process, and thus benefits the generation of scenarios. In Chapter A.4 it will be compared the three models described in this chapter: PAR(p), MS-PAR(p) without

previous identification of the influence of ENSO (called MS-PAR(p)1) and MS-PAR(p) with previous identification of the influence of ENSO (called MS-PAR(p)2)

A.3.5 Evaluation Metrics

Evaluation Criteria for Non-Conditioned Cases

A synthetic series generation model should preserve the main statistical characteristics of the observed series. The adequacy of the models used in this dissertation will be evaluated through the reproduction of historical statistics. Thus, it will be compared the synthetic statistics of mean, standard deviation and annual spatial correlations with their historical values. For each simulated period, the synthetic probability distributions of each HPP will be tested through the KS test, with a significance level of 5%. Also, for the MS-PAR(p) model evaluates the reproduction of the means and standard deviations conditioned by the ENSO states. Also for the MS-PAR(p) model, it is performed an analysis of the ENSO synthetic scenarios, comparing them with the historical distributions of probability, and their transition matrices between the states.

For medium/long-term operation planning, the representation of its critical periods is also extremely relevant. In this context, it is necessary to evaluate the reproduction of the negative sequences, defined as a period in which the streamflows are uninterruptedly below a predetermined value, for example, the LTM. Figure A.19 illustrates the concept of negative sequence and its associated variables. The continuous line represents the streamflow and the dotted line a pre-established limit. The intervals (t1 - t2) and (t3 - t4) correspond to the negative sequences, that is, the flows are below the limits.

Here, we define the random variables related to negative sequences:

$$C = t_2 - t_1 \quad (\text{A.35})$$

$$S = \sum_{i=t_1}^{t_2} (Z_i - \mu_i) \quad (\text{A.36})$$

$$I = \frac{S}{C} \quad (\text{A.37})$$

being C the sequence length corresponds to the length of the intervals (t2 - t1), (t4 - t3), in Figure A.19; S the sequence sum corresponds to the area below the boundary during the sequence (represented in Figure A.19 as the areas A1 and A2); and I the sequence intensity corresponds to the mean value below the limit, i.e., the sequence sum divided by the respective sequence length.

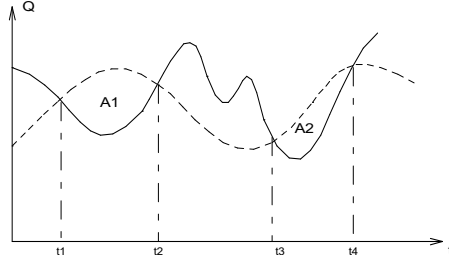


Figure A.19: Example of a hypothetical negative sequence. The full line corresponds to the observed time series and the line dotted to the pre-established limit.

The same procedure is performed for both the historical series and the synthetic series generated by the three models. For each random variable, the mean, maximum, minimum, and standard deviation values will be compared, in addition to testing the null hypothesis of equality of the distribution of the two samples. Let n_1 be the number of observations of the random variable X , obtained from the historical series and n_2 the number of observations of the random variable Y , obtained from the series generated. For the length of the sequence, which is a discrete variable, the equality test of two multinomial distributions (MOOD, 1950) is used. The test is based on the division of the sample space into $(k + 1)$ subsets, and the null hypothesis states that:

$$H_0 : p_{1,j} = p_{2,j} \quad j = 1, \dots, k + 1 \quad (\text{A.38})$$

where $p_{1,j}$ is the probability that an observation of the first population belongs to the j th subset, and $p_{2,j}$ is the corresponding probability for the second population. It is demonstrated that:

$$\chi_k^2 = \sum_{i=1}^2 \sum_{j=1}^{k+1} \frac{\left(N_{ij} - n_i \frac{N_{1j} + N_{2j}}{n_1 + n_2} \right)^2}{n_i \frac{N_{1j} + N_{2j}}{n_1 + n_2}} \quad (\text{A.39})$$

has asymptotically Chi-square distribution with k degrees of freedom, where N_{ij} is the number of observations of the j th subset of the i th population.

For continuous random variables, such as sum and sequence intensity, the Kolmogorov-Smirnov test is used.

The multivariate dry periods are an extremely relevant feature in the context of Brazilian energy planning. Thus, in addition to the univariate analysis of the random variables, it is conducted the evaluation in a multivariate manner, observing the concomitance of negative sequences between different HPPs, as presented in SANTARÉM *et al.* (2017).

In order to verify the ENSO scenarios, one can also analyze the length of each

ENSO event in the synthetic series generated. The multinomial test described above will be used to compare the historical distribution of the lengths from ENSO states sequences in its three states, with the synthetic distribution obtained in the scenarios generated. Also, from these distributions, the average and maximum lengths will be evaluated, as well as the standard deviation of the distributions.

Evaluation Criteria for Conditioned Cases

As the synthetic scenarios are used in the operation planning models, a way of comparing scenarios generation models would be by comparing the results of the real-time operation. However, in this dissertation, the assessment of conditional cases will be based on commonly used criteria for verifying the average forecast power of synthetic series. The Mean Absolute Error (MAE), Mean Absolute Percent Error (MAPE) and Root Mean Square Error (RMSE) are presented, respectively, in the Equations (A.40), (A.41) and (A.42)

$$MAE = \frac{1}{n} \sum_{i=1}^n |Z_t - Cen_t^i| \quad (A.40)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Z_t - Cen_t^i}{Z_t} \right| \quad (A.41)$$

$$RMSE = \frac{1}{n} \sum_{i=1}^n \sqrt{(Z_t - Cen_t^i)^2} \quad (A.42)$$

where Cen_t^i is the i -th simulated synthetic scenario of a total of n scenarios.

A related objective of this dissertation is the evaluation of the quality of the scenarios generated for the ENSO states since they have a direct influence on the synthetic streamflows. For this, first, the Ranked Probability Score (RPS) is calculated, which is the sum of the quadratic differences between the probabilistic forecasts accumulated in relation to the observed event. The cumulative probability for each category is given by:

$$P_i = \sum_{j=1}^i p_j \quad (A.43)$$

where p_j is the probabilistic prediction of the occurrence of the j -th state. The RPS is calculated according to the expression:

$$RPS = \sum_{i=1}^C (P_i - O_i)^2 \quad (A.44)$$

where O_i is 1 when the state is observed and zero for the other states. After calculating the RPS, it is possible to calculate the Ranked Probability Skill Score (RPSS),

where the reliability of the probabilistic forecasts generated is compared with the quality of the use of the climatological frequency ⁵ in each period. The closer to 1, the better the forecast. The RPSS is defined by the Equation (A.45):

$$RPSS = 1 - \frac{RPS}{RPS_{clim}} \quad (\text{A.45})$$

where RPS_{clim} is the RPS value using the climatological frequency.

A.4 Results and Discussions

This chapter will present the results obtained in the case studies presented in this dissertation. Here, the names of the models PAR(p), MS-PAR(p)1 (without prior identification of the influence of ENSO) and MS-PAR(p)2 (with prior identification of the influence of ENSO) will be abbreviated as P, MSP1 and MSP2 in the tables and graphs. Firstly, the case studies are presented. In the next section, the influence of the ENSO phenomenon on natural streamflows to the HPPs will be evaluated through the proposed procedure. In the last section, the results of the case studies are analyzed, in the unconditioned case studies, conditioned in the in-sample periods, and, at last, conditioned in the out-of-sample periods.

A.4.1 Case Studies

The hydraulic configuration will be the same for all analyzed cases, this being the configuration found in the monthly operation program of January 2018, which has 141 HPPs. HPPs with zero incremental streamflow were excluded from the configuration. In order to facilitate the evaluation, one HPP of each of the twelve EERs in the current configuration of the NIS was chosen for the presentation of the results. The table A.3 presents the chosen HPPs and the installed power of each one. The geographical distribution of these HPPs is shown in Figure A.20. It can be observed that most of the main Brazilian river basins are contemplated.

⁵Climatological frequency (or probability) corresponds to the frequency of occurrence of a climatic phenomenon throughout the history

Table A.3: Representing HPPs

HPP	Hydrographic basin	Installed Potency (MW)
Itá	Uruguai	1,450
Salto Caxias	Iguaçu	1,240
Itaipu	Paraná	14,000
Capivara	Paranapanema	619
Porto Primavera	Paraná	1,540
Furnas	Grande	1,312
Sobradinho	São Francisco	1,050
Santo Antônio Jari	Amazonas	390
Estreito Tocantins	Tocantins	1,087
Belo Monte	Xingu	6,111
Teles Pires	Teles Pires	1,820
Jirau	Madeira	3,750

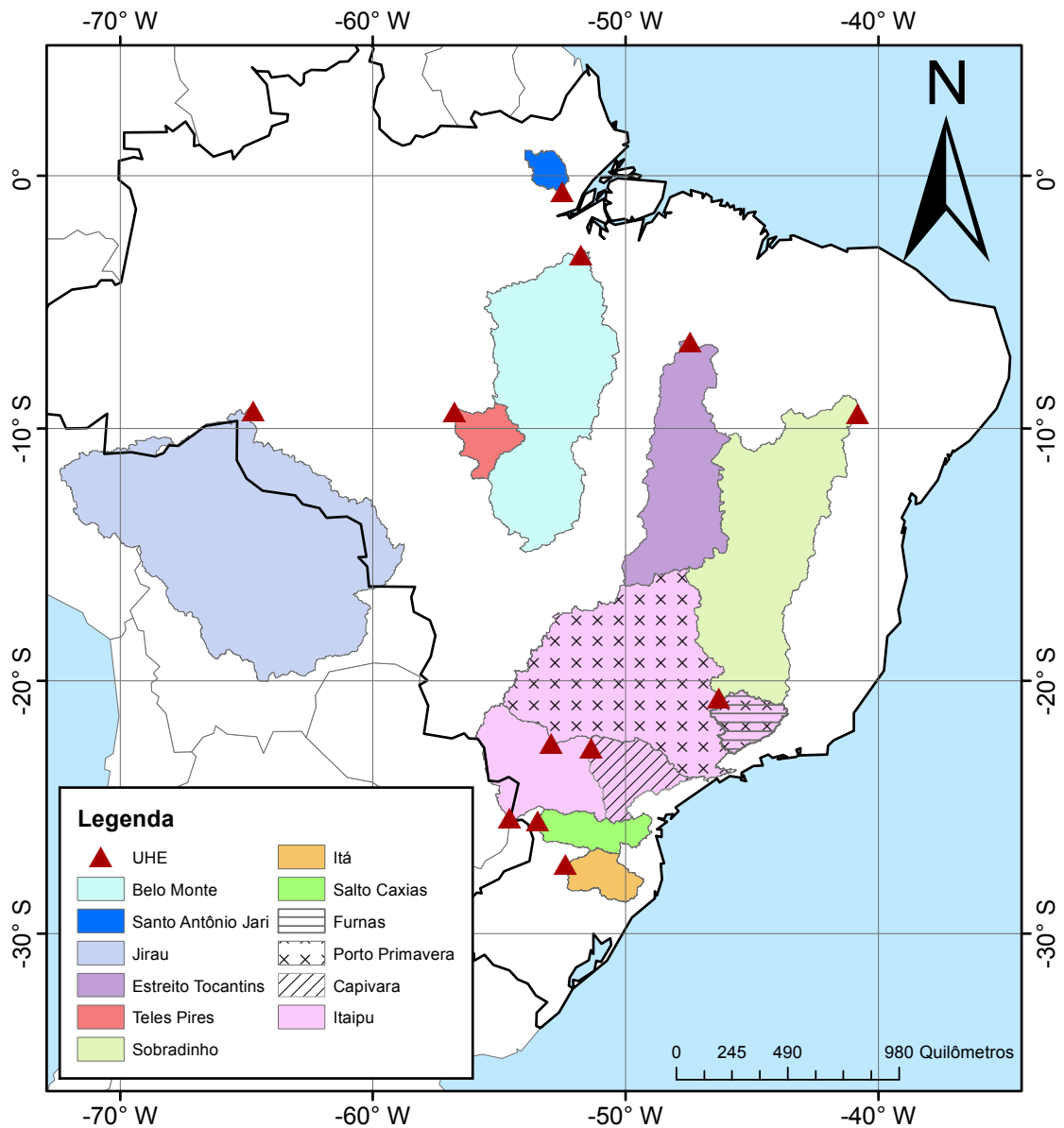


Figure A.20: Geographic distributrition of the selected HPPs.

To evaluate the adequacy of the PAR(p) model and the MS-PAR(p) models, with and without prior identification of ENSO's influence, an unconditioned case will be performed. In this type of case, the reproduction of historical statistics is sought through the generation of synthetic scenarios. There will be generated 2000 synthetic series, with a horizon of 67 years (equal to the number of years used in the adjustment of the models).

For the comparison of the performance of both models, it will be carried out several cases conditioned to the recent past. These cases will be divided into two types:

1. scenarios generation one step ahead, for each month within the in-sample period from 1950 to 2009;

2. scenario generation for five years ahead, for each month within the out-of-sample period from 2010 and 2017;

In the first type of conditioned case (in-sample cases), for the generation of scenarios of the MS-PAR(p) model, it will be considered the complete knowledge of the ENSO states, that is, perfect predictions of ENSO. For the out-of-sample cases, it will be used the probabilistic forecasts provided by IRI. In both cases, 2,000 synthetic scenarios will be generated.

The Table A.4 provides a summary of the cases that will be analyzed in this dissertation.

A.4.2 Influence of El Niño - Southern Oscillation on the Streamflows

The first results to be presented in this dissertation are part of diagnosis about the influence of the ENSO phenomenon in the series of natural streamflows of the HPPs. The KS adherence test, described in chapter A.3.5, was used for the period between 1950 and 2016. To facilitate the visualization of the results, it was decided to divide the HPPs between the hydrographic basins which they belong to and subsequently presented separately in the four NIS submarkets (south, southeast/central-west, northeast and north). Figure A.21 presents the result of the KS test performed for every month in HPPs located in the southern region of Brazil. The cells in red indicate the influence of EN, blue of the LN, lilac of both states, and white without influence. A distinct pattern of identification of the effects of EN between November and March is observed for most of the HPPs. It should be noted that for the Iguaçu river basin that is located in the northern part of the southern region, in general only in November and December the EN is identified. in hydrographic regions at higher latitudes, both phases of ENSO produce changes in November and December, and only EN in January to March. In some HPPs in the other river basins, the influence of LN between May and July is also identified. It is noteworthy that most of the months that are influenced by the ENSO (from November to March) are months in which the average flows in the south are lower.

Table A.4: Case Studies Summary

Case type	Quantity	Horizon	Test Period	Evaluation Type	Evaluation Metrics
Non-conditioned	1	67 x 12	-	Model adequacy	-moments reproduction - spacial correlation - negative sequences
In-sample conditioned	59 x 12	1	1951 a 2009	in-sample performance	- MAE - MAPE - RMSE
Out-of-sample conditioned	8 x 12	5 x 12	2010 a 2017	out-of-sample performance	- MAE - MAPE - RMSE - RPSS

HPP	BASIN	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
FUNDAO	IGUAÇU												
G.B. MUNHOZ	IGUAÇU												
JORDÃO	IGUAÇU												
MAUÁ	IGUAÇU												
SALTO CAXIAS	IGUAÇU												
SALTO OSÓRIO	IGUAÇU												
SEGREDO	IGUAÇU												
SLT.SANTIAGO	IGUAÇU												
STA CLARA PR	IGUAÇU												
ITAIPU	ITAIPU												
14 DE JULHO	SUL												
CASTRO ALVES	SUL												
D. FRANCISCA	SUL												
ERNESTINA	SUL												
G.P. SOUZA	SUL												
ITAUBA	SUL												
JACUÍ	SUL												
MONTE CLARO	SUL												
PASSO REAL	SUL												
QUEBRA QUEIXO	SUL												
SALTO PILÃO	SUL												
BARRA GRANDE	URUGUAI												
CAMPOS NOVOS	URUGUAI												
FOZ CHAPECÓ	URUGUAI												
GARIBALDI	URUGUAI												
ITA	URUGUAI												
MACHADINHO	URUGUAI												
MONJOLINHO	URUGUAI												
PASSO FUNDO	URUGUAI												
PASSO S JOÃO	URUGUAI												
SÃO JOSÉ	URUGUAI												

Figure A.21: KS Test for the southern region of Brazil. Cells in red, blue and lilac indicate the influence of any of the states, EN, LN and N, respectively.

It is presented the historical monthly mean, and standard deviation, calculated according to the Equations (A.24) and (A.25), and the monthly coefficient of variation $CV_m^r = \frac{\sigma_m^r}{\mu_m^r}$, for each of the ENSO states and for the complete historical data, for some of the selected HPPs in the south region. The following results were obtained without taking into account the KS test, that is, the statistics are calculated for all ENSO states in all months. Figure A.22 presents the results obtained for HPP Itá, located on the Uruguai river. The southern region of Brazil is characterized by not having a well-defined seasonality of its streamflows, as can be observed. About the averages, it is observed that the EN phenomenon produces an increase of the streamflow in relation to the other states and in the, and the inverse occurs in the state of LN (when the states are influential). It is interesting to note that, in general, the increase of the mean is also accompanied by the increase of the standard deviation, and the opposite occurs in the LN periods. To put it another way, there is greater variability associated in EN states, but lower in LN or N states. May is the only month that falls outside the described pattern, with LN a higher average, but a lower standard deviation. For the Itá HPP, the KS test indicates the influence of both ENSO phases in November and December, the EN in January and March and the LN in May and October.

Figure A.23 brings the results obtained to the Itaipu HPP, also located in the south region of Brazil, in the Paraná river. Despite being located in the south region, the Itaipu HPP has an extensive basin of contribution that also covers part of the southeast region. For the Itaipu HPP, only in April, May, and December, the influence of the EN on its streamflow was identified. As in most of the months in

the Itá HPP, the EN causes an increase in the average flow, as well as the standard deviation, which represents a greater uncertainty in this period. In the case of Itaipu HPP, in May, there is an increase of 1000 m³/s in the average flow, which represents a 50% increase when compared to the LTM. The standard deviation (which represents the uncertainty of the stochastic process of inflows) also grows about 60%. In May, for the N and LN states (which are considered as a single state for the Itaipu HPP given the result of the process of identifying the influence), there is an opposite effect, with smaller mean streamflows, and a lower standard deviation.

The statistical data of Salto Caxias HPP, the last plant in the south region to be analyzed, located in the Iguaçú river, is presented in Figure A.24. According to the previous analysis, the KS test only presents the influence of the EN state in November and March. As in the HPPs highlighted above, the EN causes an increase in the mean and the standard deviation, while the other two states result in a reduction of the statistics, mainly between November and May.

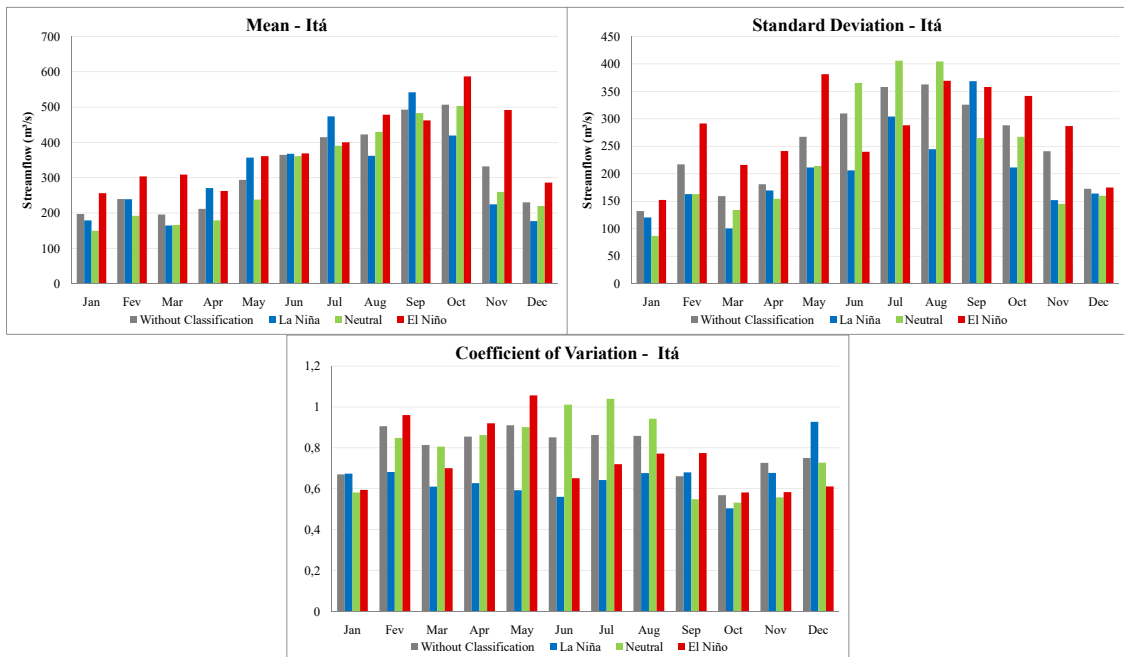


Figure A.22: Historical Statistics of Itá HPP according to the ENSO state. The statistics obtained for the complete history and segmented into the state LN, N and EN, are represented by the gray, blue, green and red color bars, respectively. Data of the period between 1950 and 2017

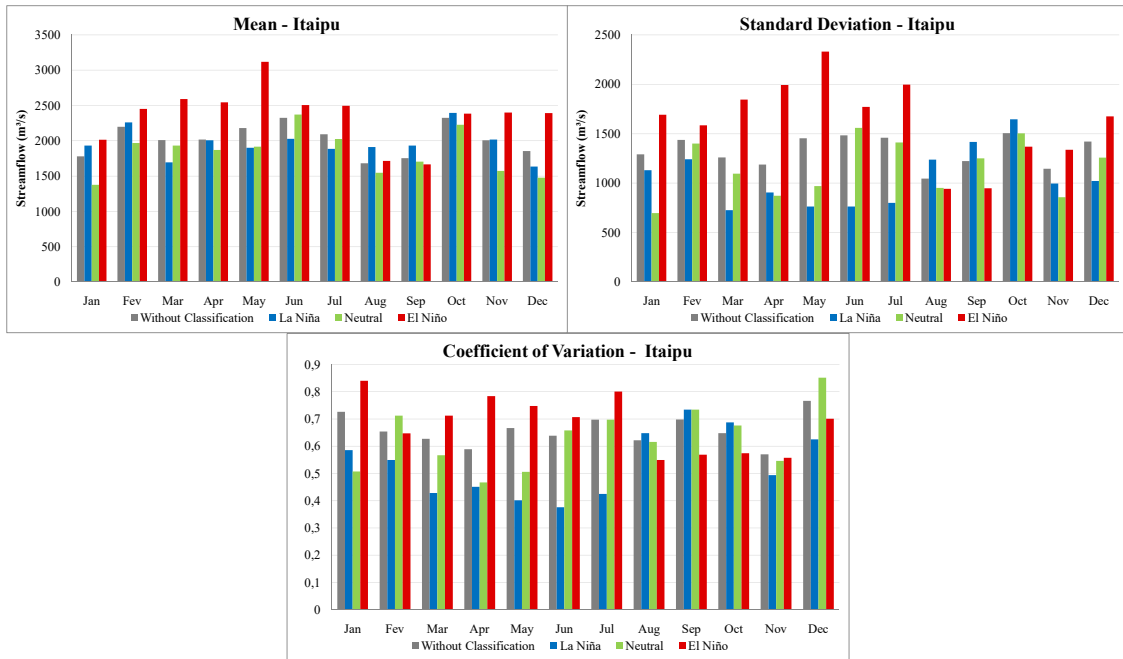


Figure A.23: Historical statistics of the Itaipu HPP according to ENSO state.

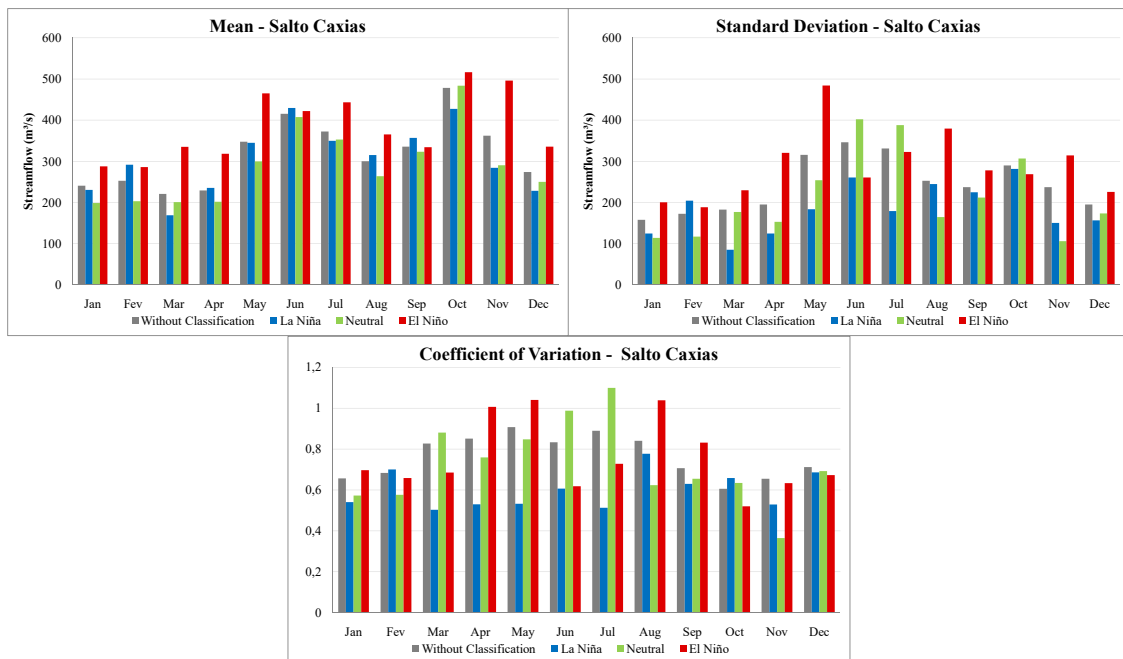


Figure A.24: Historical statistics of the Salto Caxias HPP according to ENSO state.

These patterns are observed in most of the HPPs located in the south region of Brazil and are in consonance with the literature review made in chapter A.2.2. The ENSO phenomenon has a great influence in the south region. Its warm phase is often responsible for an increase in the streamflow, while the LN by a decrease. However, it is also observed that it has not a direct relationship between the ENSO phenomenon and the flows, which explains the greater uncertainty in the EN events,

represented by the increase of the standard deviation. There is a tendency for more flood events to occur during this period, but this is not always verified. However, during LN, below average streamflow is more likely.

Next, it is investigated the HPPs located in hydrographic basins of the south-east and central-west. First, the results obtained for the KS test in these HPPs are presented. For the watershed of the Paranapanema river, which is located in the southernmost portion of the basin analyzed in Figure A.25, the test indicates, in general, a marked influence of the EN, mainly in November and December, and between May and August. The streamflows of their HPPs have a transition behavior between the south and southeast regions, evidenced by the similarity of their result in the KS test compared to the HPPs in the south region. The results obtained for the Tietê river basin do not present a well-defined pattern, being identified between April to July in some of its HPPs the influence of one of the ENSO states. For the Rio Grande basin and southeast Atlantic, which encompasses the coastal watersheds, in most of the HPPs, a consistent pattern of EN influence is identified between May and July (in some cases the LN is also relevant) and in August to September of the LN. In the HPPs located upstream of the Paranaíba river basin, the LN is identified influential between January and September.

UHE	BACIA	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
AIMORES	ATL. SUDEST												
BAGUARÍ	ATL. SUDEST												
BILLINGS	ATL. SUDEST												
CANDONGA	ATL. SUDEST												
FUNIL	ATL. SUDEST												
GUILMAN-AMOR	ATL. SUDEST												
ILHA DOS POMBOS	ATL. SUDEST												
LAJES	ATL. SUDEST												
MASCARENHAS	ATL. SUDEST												
P. ESTRELA	ATL. SUDEST												
PARAIBUNA	ATL. SUDEST												
PICADA	ATL. SUDEST												
ROSAL	ATL. SUDEST												
S. A. CARVALHO	ATL. SUDEST												
SALTO GRANDE	ATL. SUDEST												
SANTA BRANCA	ATL. SUDEST												
SANTANA	ATL. SUDEST												
SIMPLÍCIO	ATL. SUDEST												
SOBRAGÍ	ATL. SUDEST												
STA. CECÍLIA	ATL. SUDEST												
TOCOS	ATL. SUDEST												
A. VERMELHA	GRANDE												
A. S. OLIVEIRA	GRANDE												
CACONDE	GRANDE												
CAMARGOS	GRANDE												
E. DACUNHA	GRANDE												
ESTREITO	GRANDE												
FUNIL-GRANDE	GRANDE												
FURNAS	GRANDE												
IGARAPAVA	GRANDE												
JAGUARA	GRANDE												
JAGUARÍ	GRANDE												
M. DE MORAES	GRANDE												
MARIMBONDO	GRANDE												
P. COLOMBIA	GRANDE												
VOLTA GRANDE	GRANDE												
ITIQUEIRA 1	PARAGUAI												
JAUÚ	PARAGUAI												
MANSO	PARAGUAI												
PONTE PEDRA	PARAGUAI												
I. SOLT. EQV	PARANÁ												
JUPIÁ	PARANÁ												
P. PRIMAVERA	PARANÁ												
B. COQUEIROS	PARANAÍBA												
BATALHA	PARANAÍBA												
CACH. DOURADA	PARANAÍBA												
CAÇU	PARANAÍBA												
CAPIM BRANCO 1	PARANAÍBA												
CAPIM BRANCO2	PARANAÍBA												
CORUMBÁ I	PARANAÍBA												
CORUMBÁ III	PARANAÍBA												
CORUMBÁ IV	PARANAÍBA												
EMBORCAÇÃO	PARANAÍBA												
ESPORA	PARANAÍBA												
FOZ R. CLARO	PARANAÍBA												
ITUMBIARA	PARANAÍBA												
MIRANDA	PARANAÍBA												
NOVA PONTE	PARANAÍBA												
SALTO	PARANAÍBA												
SÃO SIMÃO	PARANAÍBA												
SERRA FACÃO	PARANAÍBA												
SLT. VERDINHO	PARANAÍBA												
A. A. LAYDNER	PARANAPANEMA												
CANOAS I	PARANAPANEMA												
CANOAS II	PARANAPANEMA												
CAPIVARA	PARANAPANEMA												
CHAVANTES	PARANAPANEMA												
L. N. GARCEZ	PARANAPANEMA												
OURINHOS	PARANAPANEMA												
PIRAJU	PARANAPANEMA												
ROSANA	PARANAPANEMA												
TAQUARUCU	PARANAPANEMA												
A. S. LIMA	TIETÊ												
BARRA BONITA	TIETÊ												
EDGARD SOUZA	TIETÊ												
GUARAPIRANGA	TIETÊ												
IBITINGA	TIETÊ												
N. AVANHANDAVA	TIETÊ												
PROMISSÃO	TIETÊ												

Figure A.25: KS Test for the southeastern and central-west region of Brazil.

Figure A.26 presents the historical statistics for the Capivara HPP, located on the Paranapanema river. It is possible to observe that this HPP does not have a defined seasonality and that as in Itaipu and Itá, their streamflow, in general, are above during EN events. This effect, as in the other plants analyzed, is accompanied by greater variability as well. It should be noted that according to the criterion of the KS test, the EN state influences in March, May, July, November and December, and the LN between June and August and November.

The results for the Furnas HPP, which belongs to the river basin of Grande in Minas Gerais, are presented in Figure A.27. The Furnas HPP is one of the most important in the NIS because of its regularization potential of the entire Paraná river basin. It is possible to verify that between January and July, except March, the average streamflows in the EN state are higher than in the others. The standard deviation from January to June is higher for the warm ENSO phase, whereas in September and October larger values are presented in LN periods, especially in September, when the coefficient of variation becomes higher than the unit value in that state. According to the KS test, LN is influential in February and between May and September (most of the dry period in this region) and the EN state responsible for variations in the months between May and July. This pattern can be extended to most HPPs in the Grande basin.

The Porto Primavera HPP, located on the Paraná river, has its statistics presented in Figure A.28. It can be observed that with the exception of June and December, there is no perceptible variation in the inflows due to the different states of ENSO, as occurs for the Furnas HPP. The ENSO influence identification procedure only indicates the EN status as significant for the month of December, demonstrating the little influence suffered by the Porto Primavera HPP. As it is in the region further downstream of the Paraná river, it has an extensive drainage area and contributions from several other rivers and regions, which may explain the lack of a stronger ENSO influence.

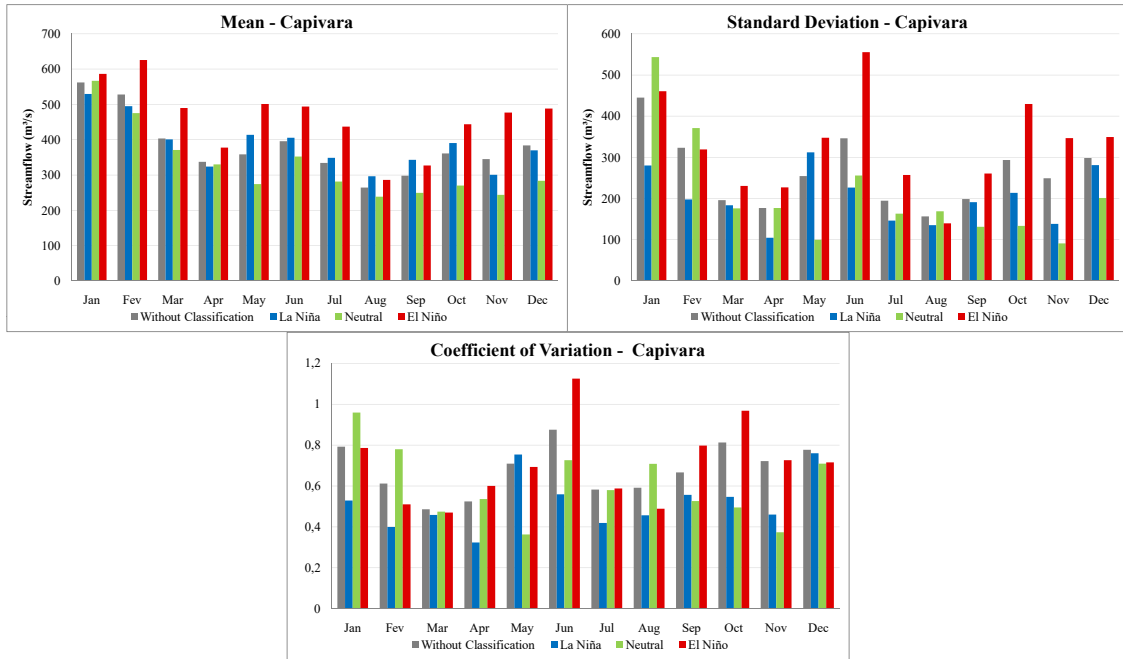


Figure A.26: Historical statistics of the Capivara HPP according to ENSO state.

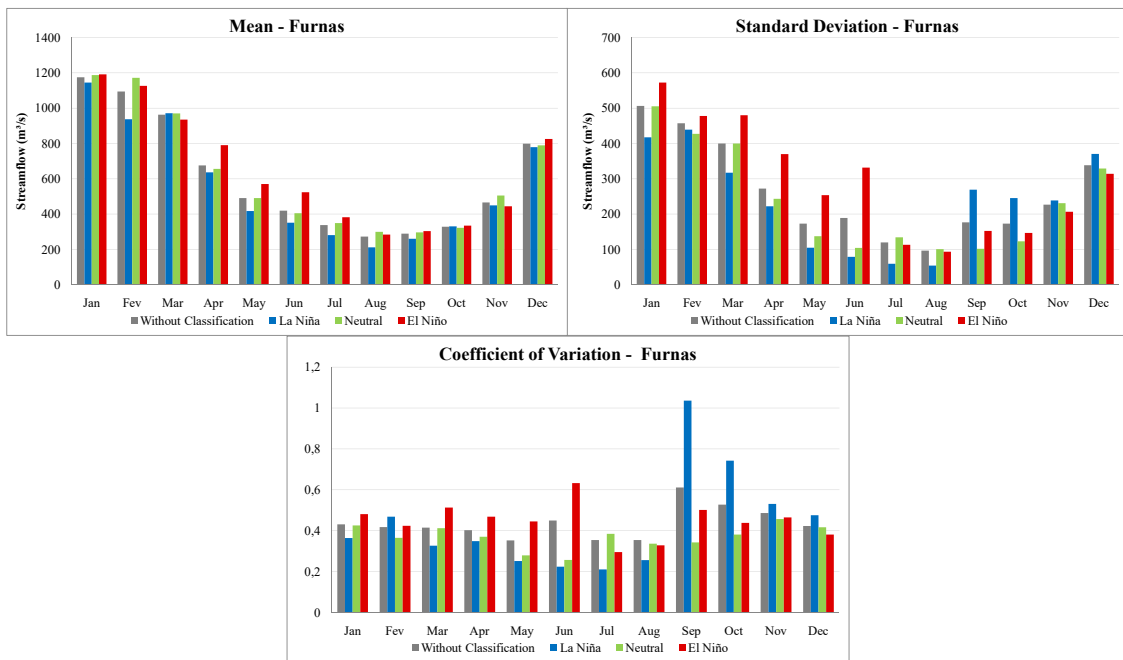


Figure A.27: Historical statistics of the Furnas HPP according to ENSO state.

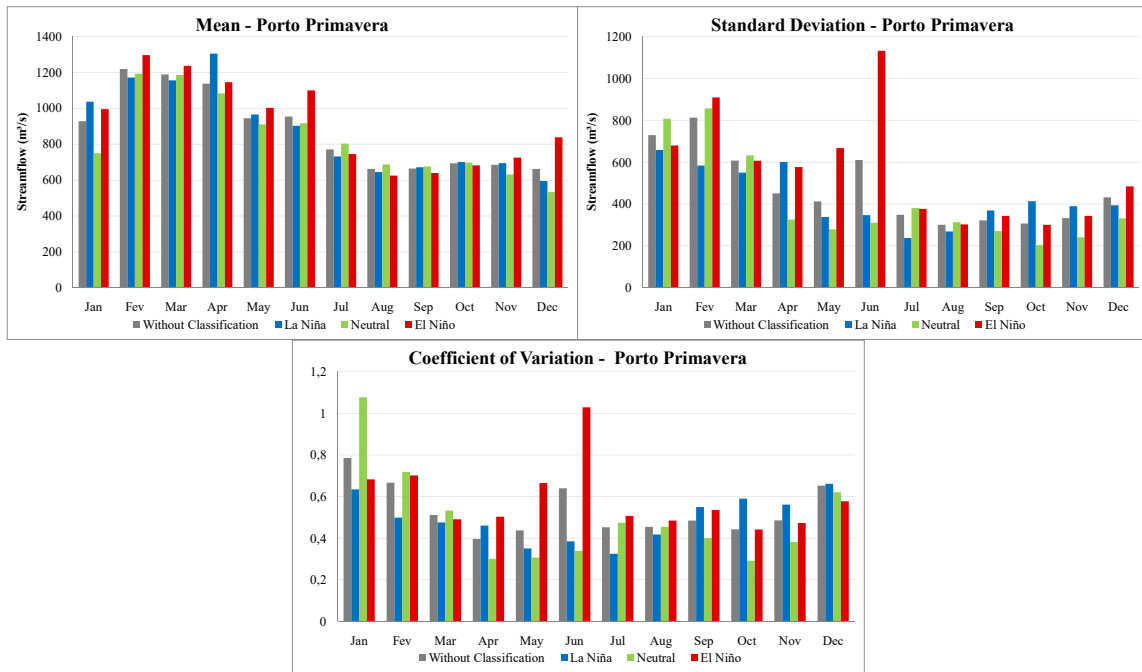


Figure A.28: Historical statistics of the Porto Primavera HPP according to ENSO state.

The result of the KS test for the watersheds belonging to the northeast region is shown in Figure A.29. It is observed that most of the LN are identified as an important influence in January and February and in August and September. HPPs located in the northeast and East Atlantic basins have some influence in November and December.

Figure A.30 brings the results to Sobradinho HPP, located in the northeastern region of Brazil, belonging to the São Francisco river basin. It is noticed that during events of LN the streamflows are below the average in almost all months, except for November and December. In February and March, the wettest months in this region, it is possible to observe that the effect of LN is more intense, and has a lower variability, especially in March. This inversion in the influence pattern observed between November and February was also highlighted in the analysis of rainfall in this region in GRIMM (2003) and GRIMM (2004). LN was identified as influential in February and March and between July and September.

HPP	BASIN	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
IRAPÉ	EAST ATL.												
PEDRA DO CAVALO	EAST ATL.												
ITAPEBI	EAST ATL.												
SANTA CLARA MG	EAST ATL.												
BOA ESPERANÇA	PARNAÍBA												
QUEIMADO	SÃO FRANCISCO												
RETIRO BAIXO	SÃO FRANCISCO												
SOBRADINHO	SÃO FRANCISCO												
TRÊS MARIAS	SÃO FRANCISCO												

Figure A.29: KS Test for the northeast region of Brazil.

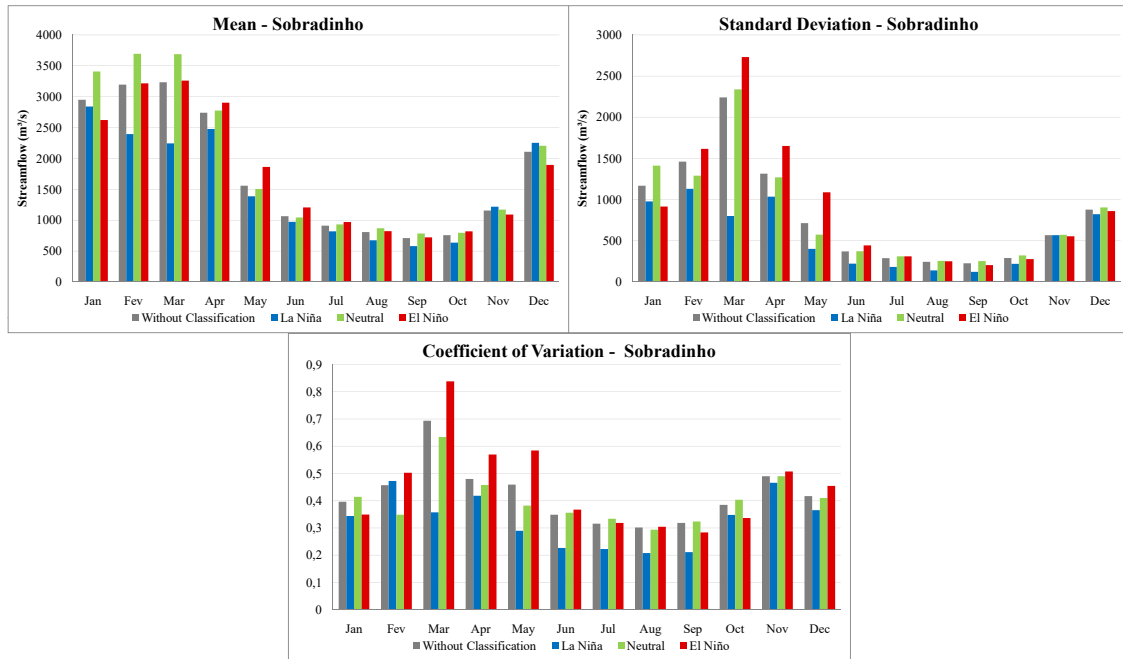


Figure A.30: Historical statistics of Sobradinho HPP according to the ENSO state.

The results for the KS test of the HPPs located in the north region of Brazil are shown in Figure A.31. It is possible to observe some quite distinct patterns between the several basins that compose it. In the hydrographic basin of the Teles Pires river and the Madeira river, the HPPs do not have a well-demarcated ENSO influence. In the others, the opposite occurs, mainly in the Xingu basin and in the HPPs belonging to the left bank of the Amazon river, denoted as Amazonas. As highlighted in several studies cited above, the north region of Brazil suffers a significant influence of the ENSO phenomenon, and the higher the latitudes, the greater the effects. In general, the two phases of the ENSO are identified as influential in the Amazon basin in almost every month, which constitutes this basin as one of those that suffer ENSO effects more markedly.

Figure A.32 shows the results found for the Belo Monte HPP, located in the Xingu basin in the state of Pará. It is possible to observe that during the EN periods, the streamflows are below the average from January to March. In April, despite presenting average values, it has a higher standard deviation in periods of EN. In May and June, above average values in LN periods were found, with no smaller changes in the standard deviation. Because Belo Monte HPP is located in the north region of Brazil, ENSO causes a change in the pattern of the hydrological regime quite distinct from the rest of Brazil. In the Belo Monte HPP, and in general in the other HPPs in the north region, LN causes positive anomalies, whereas the EN causes negative anomalies in the streamflows. It is noteworthy that for the Belo Monte HPP, the KS test indicates the influence of the EN between January and March and in November, and the LN indicated as influential in May, September,

and November.

In Figure A.33 the statistics for the Santo Antônio Jari HPP are presented, which is located in the north region of Brazil, specifically on the left bank of the Amazon river. From January to July it can be observed that there is a significant change in the behavior of the streamflows between the different states, especially in the wetter months. LN periods present above average mean and standard deviation, while the inverse occurs in periods of EN. The N periods have means close to the values found in the history without the classification. Between February and July, and in November, both ENSO states are identified as influential, while in December only the EN.

It is interesting to note that these results are in agreement with the patterns characterized in the researches presented previously on the influence of the ENSO phenomenon on the precipitations that occur in Brazil. The north region, in general, suffers from below-average rainfall in periods of EN, and the opposite occurs in periods of LN.

Another important point that should be highlighted concerns the changes in standard deviation. Although the streamflow in the Santo Antônio Jari HPP is higher in periods classified as LN than in other periods, it is observed that this happens together with greater variability. In other words, there is greater uncertainty about what will occur in LN periods, while the periods of EN have smaller means and lower uncertainty. Also, when dividing the historical into the three states, two of them started to present a smaller standard deviation than in the unclassified history.

HPP	BASIN	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
BELO MONTE	XINGU												
CURUÁ-UNA	XINGU												
DARDANELOS	MADEIRA												
GUAPORÉ	MADEIRA												
JIRAU	MADEIRA												
RONDON II	MADEIRA												
SAMUEL	MADEIRA												
STO. ANTÔNIO	MADEIRA												
BALBINA	AMAZONAS												
CACH. CALDEIRÃO	AMAZONAS												
COARACY NUNES	AMAZONAS												
FERREIRA GOMES	AMAZONAS												
STO. ANT. JARI	AMAZONAS												
CANA BRAVA	TOCANTINS												
ESTREITO TOC.	TOCANTINS												
LAJEADO	TOCANTINS												
PEIXE ANGICAL	TOCANTINS												
SÃO SALVADOR	TOCANTINS												
SERRA DA MESA	TOCANTINS												
TUCURUÍ	TOCANTINS												
SÃO MANOEL	TELES PIRES												
TELES PIRES	TELES PIRES												

Figure A.31: KS Test for the north region of Brazil.

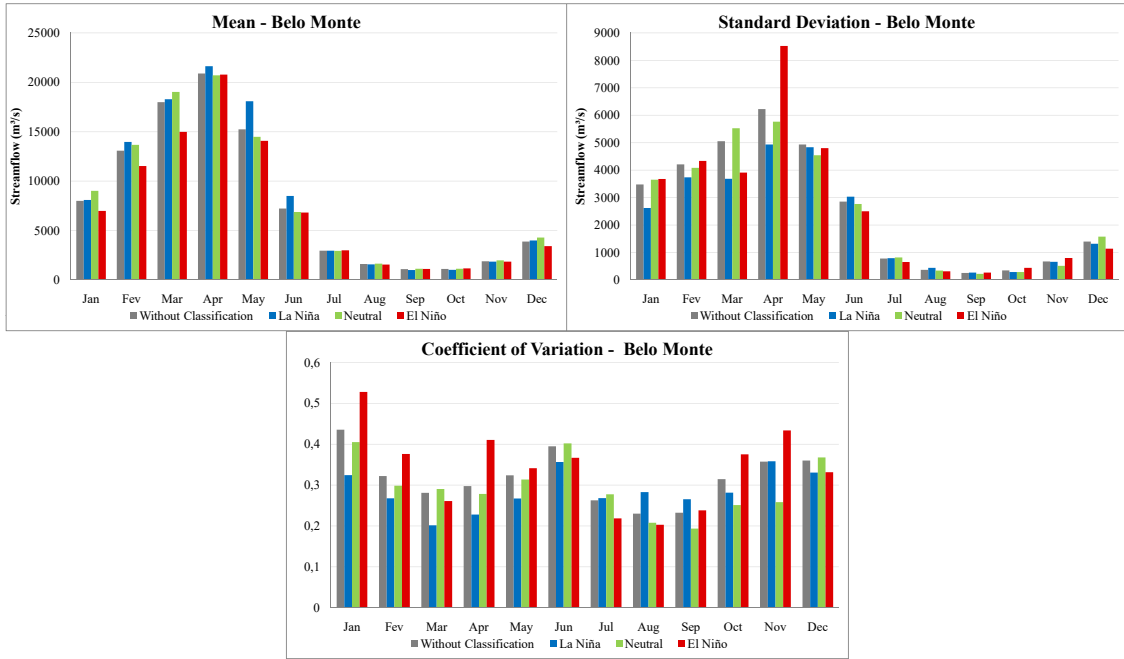


Figure A.32: Historical statistics of Belo Monte Hydroelectric Power Station according to ENSO state.



Figure A.33: Historical statistics of Santo Antônio Jari HPP according to ENSO state.

The ENSO phenomenon, like most climatic phenomena that act on a global scale, produces non-linear effects. The MS-PAR(p) model presented in this dissertation approximates linearly the consequences caused by ENSO, by classifying the monthly historical data of natural streamflows among the three ENSO states, given the classification obtained through the ONI. As shown in this section, the

approximation performed is adequate, capturing the various changes produced in the hydrological cycle on a monthly scale throughout all the Brazilian hydrographic basins. In addition to the historical classification process, a second step was proposed to identify the influence of ENSO through the application of a procedure involving the KS adherence test.

This second step seeks to make the MS-PAR(p)2 model more parsimonious, discarding the use of the classification when this procedure does not bring any relevant information to the generation process. Often, it is identified that only one of the ENSO phases produces important effects, with the opposite phase and the N state having close statistics to each other and thus considered as only one state. Of the total of 141 HPPs analyzed, only 19 did not present differences between the ENSO states for at least one month. As shown, the ENSO does not cause changes just in average streamflows, but also in the standard deviation. This finding indicates that in certain ENSO states there is greater predictability, while others have considerable uncertainty associated. In the context of the generation of synthetic scenarios, this kind of information is extremely relevant.

Many of the results presented in this section are in agreement with several studies cited in the literature review. In general, the EN state produces positive anomalies in the south region between November and March, with the opposite effect in the LN state. As the streamflows in this region do not have a strong monthly correlation, this effect is not propagated for the other months of the year. In the southeastern/central-west region there is great heterogeneity of results in the different river basins that compose it. This region, according to the literature review, is located in a transition region of the ENSO effects, and there is no well-defined effect pattern. It is presented a slight increase in the means and standard deviations of the streamflows in the EN state and a decrease in the NL and N. The Paranapanema basin, being further south, suffers a stronger influence of the ENSO phenomenon.

In all other hydrographic basins that make up the southeast region, ENSO produces effects of lesser magnitude, but its observation is not restricted to the wet period. Due to the strong monthly correlation in the streamflows, ENSO impacts, when they exist, are propagated throughout the year. In the northeast region, there is a strong impact of ENSO in the streamflows between January to March, when occurs most of the rainfall in the region. This effect is also observed in subsequent months due to correlation. As in the south and southeast regions, EN causes positive anomalies, and the opposite occurs in the LN. It is interesting to note that in November and December, the effects are inverted in the northeast region. In these months, the LN presents positive anomalies, while negative anomalies occur in EN, but with lower intensity. This variability in monthly scale had already been observed

in the references cited above. Finally, the north region suffer a strong influence of the ENSO phenomenon, especially in the HPPs located on the left bank of the Amazon river. Contrary to what occurs in most of Brazil, LN causes positive anomalies, and the opposite happens with EN. The Table A.5 provides a summary of the effects of the ENSO phenomenon on the four Brazilian macro-regions.

A.4.3 Unconditioned Case

This section will present the results obtained through the generation of inflow scenarios carried out in an unconditioned way. A stochastic model suitable for the generation of flow scenarios applied to the planning of the energetic operation of the NIS should be able to reproduce several characteristics found in the streamflow’s historical records, such as its moments, temporal correlation structure, and spatial correlations among HPPs. Aiming its use at medium/long-term planning, it is also interesting that the model can reproduce the negative sequences (consecutive periods below LTM) both in a univariate and multivariate form. For this evaluation, the multivariate generation of 2000 synthetic scenarios was done for 67 years (same length as the historical record used in the model adjustment).

In addition to the generation of streamflow scenarios, the MS-PAR(p) model also generates ENSO scenarios. This generation is performed through a Markov chain of order one adjusted to each month of the history, as presented in Chapter A.3.4. Therefore, it is necessary to evaluate if the proposed model can reproduce the historical characteristics in the synthetic series, i.e., the reproduction of the historical frequency distribution of ENSO. Figure A.34 the historical distribution of ENSO over the months (on the left) and the distribution obtained through the 2000 synthetic scenarios over 67 years are shown (on the right). The red, green and blue bars represent the percentage of EN, N, and LN. It can be observed that the synthetic distribution correctly reproduces historical values.

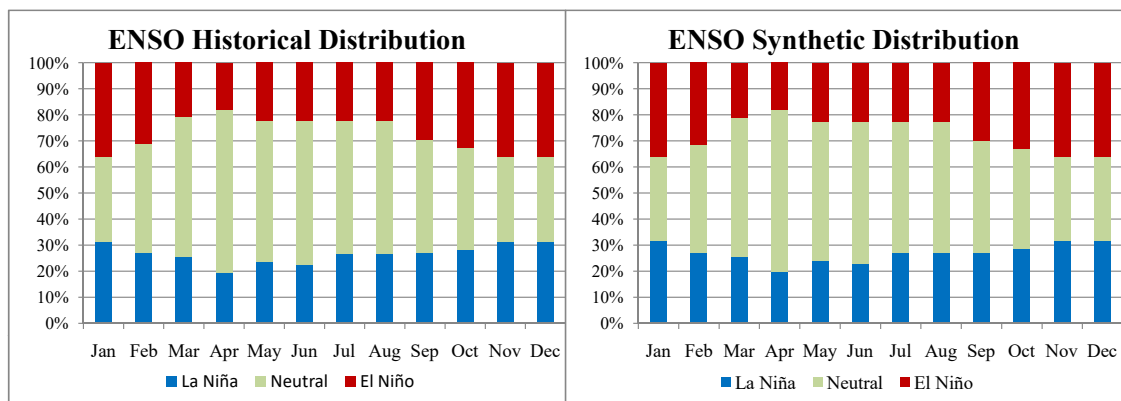


Figure A.34: Distribution of historical and synthetic probability generated through the Markov chain.

Table A.5: Effects of ENSO on streamflows to Brazilian HPPs

Region	La Niña	El Niño
South	negative anomalies, restricted from November to March	Positive anomalies, restricted from November to March
Southeast/Central-West	Heterogeneity of results; Low negative anomalies propagated all over the year	Heterogeneity of results; Low positive anomalies propagated all over the year
Northeast	High intensity negative anomalies, between January and March; Low intensity positive anomalies in November and December	High intensity positive anomalies, between January and March; Low intensity negative anomalies in November and December
North	High intensity positive anomalies, propagated all over the year	High intensity negative anomalies, propagated all over the year

As the evolution of ENSO events is conditioned to the past state, and specifically in the MS-PAR(p) model a Markov chain of order one is adopted to represent the development of the events, it is also evaluated the reproduction of the historical transitions between the three ENSO states for each month. Figure A.35 presents the historical monthly transition matrices (a) and the absolute value of the difference between the synthetic monthly transitions matrices compared to the historical ones (b), in percentage values. Green values show high probability of occurrence of the transition from one state to the other indicated and blank values represent zero probability. It is possible to observe in some months, the ENSO states always persist. Also, one can note that the results obtained in the simulation adequately reproduce the behavior of this transitions over time.

		Mês seguinte		
		La Niña	Neutro	El Niño
Jan	La Niña	100	0	0
	Neutro	0	100	0
	El Niño	0	0	100
Fev	La Niña	85.714	14.286	0
	Neutro	0	95.455	4.545
	El Niño	0	16.667	83.333
Mar	La Niña	94.444	5.556	0
	Neutro	0	100	0
	El Niño	0	33.333	66.667
Abr	La Niña	76.471	23.529	0
	Neutro	0	94.444	5.556
	El Niño	0	28.571	71.429
Mai	La Niña	92.308	7.692	0
	Neutro	9.524	80.952	9.524
	El Niño	0	8.333	91.667
Jun	La Niña	87.5	12.5	0
	Neutro	2.778	88.889	8.333
	El Niño	0	20	80
Jul	La Niña	93.333	6.667	0
	Neutro	10.811	83.784	5.405
	El Niño	0	13.333	86.667
Ago	La Niña	88.889	11.111	0
	Neutro	5.882	91.176	2.941
	El Niño	0	6.667	93.333
Set	La Niña	94.444	5.556	0
	Neutro	2.941	82.353	14.706
	El Niño	0	0	100
Out	La Niña	94.444	5.556	0
	Neutro	6.897	86.207	6.897
	El Niño	0	0	100
Nov	La Niña	100	0	0
	Neutro	7.692	84.615	7.692
	El Niño	0	0	100
Dez	La Niña	100	0	0
	Neutro	0	100	0
	El Niño	0	0	100

(a)

		Mês seguinte		
		La Niña	Neutro	El Niño
La Niña	0	0	0	
Neutro	0	0	0	
El Niño	0	0	0	
La Niña	0.24	0.24	0	
Neutro	0	0.159	0.159	
El Niño	0	0.268	0.268	
La Niña	0.156	0.156	0	
Neutro	0	0	0	
El Niño	0	0.007	0.007	
La Niña	0.459	0.459	0	
Neutro	0	0.017	0.017	
El Niño	0	0.416	0.416	
La Niña	0.184	0.184	0	
Neutro	0.014	0.002	0.016	
El Niño	0	0.198	0.198	
La Niña	0.469	0.469	0	
Neutro	0.051	0.15	0.201	
El Niño	0	0.194	0.194	
La Niña	0.202	0.202	0	
Neutro	0.04	0.104	0.064	
El Niño	0	0.298	0.298	
La Niña	0.098	0.098	0	
Neutro	0.003	0.045	0.044	
El Niño	0	0.135	0.135	
La Niña	0.299	0.299	0	
Neutro	0.011	0.033	0.044	
El Niño	0	0	0	
La Niña	0.122	0.122	0	
Neutro	0.013	0.019	0.006	
El Niño	0	0	0	
La Niña	0	0	0	
Neutro	0.028	0.045	0.016	
El Niño	0	0	0	
La Niña	0	0	0	
Neutro	0	0	0	
El Niño	0	0	0	

(b)

Figure A.35: Historical monthly transition matrices (a) and the absolute value of the difference between the synthetic monthly transitions matrices compared to the historical ones (b), in percentage values.

In order to further analyze the synthetic series of ENSO, it is evaluated if they are able to correctly reproduce the typical length (or duration) of ENSO events. Thus, the length of each ENSO state sequence was counted, forming a distribution of lengths between the three ENSO states. The same procedure was performed in the synthetic series, obtaining a synthetic distribution of lengths for the three states. By applying the multinomial test presented in Chapter A.3.5, the null hypothesis was accepted that the historical and synthetic distributions for each ENSO state are equal. Table A.6 gives the mean, maximum, and standard deviation of the historical and synthetic frequency distributions of length for the sequences of the three states. It is possible to observe that synthetic scenarios can reproduce the duration of ENSO events in an appropriate way.

Table A.6: Historical and Synthetic statistics of the length of ENSO states

	Mean		Maximum		Standard Deviation	
	Historical	Synthetic	Historical	Synthetic	Historical	Synthetic
La Niña	13.1	13.6	32.0	97.0	8.1	8.9
Neutral	9.7	9.6	50.0	100.0	10.9	9.8
El Niño	10.0	10.0	19.0	58.0	4.4	5.0

Observed the correct generation of ENSO scenarios will be carried out following the evaluations for the flow scenarios. Firstly, it will be presented the results concerning the reproduction of the first and second order moments will be presented for some of the HPPs selected in Table A.3. As these parameters are used in the definition of the model, its reproduction in synthetic scenarios is expected after a simulation for a sufficiently long period, as is the unconditioned case performed. Figures A.36, A.37, A.38 and A.39 show the mean, standard deviation and coefficient of variation obtained in the simulation performed for HPPs Itá, Furnas Sobradinho, and Belo Monte. The gray, blue, green and red bars represent the historical values, and the synthetic values obtained by the PAR(p), MS-PAR(p)1 and MS-PAR(p)2 models for each of the months. It is possible to observe that all three models can reproduce these historical statistics adequately in all the HPPs, as shown in aforementioned figures.

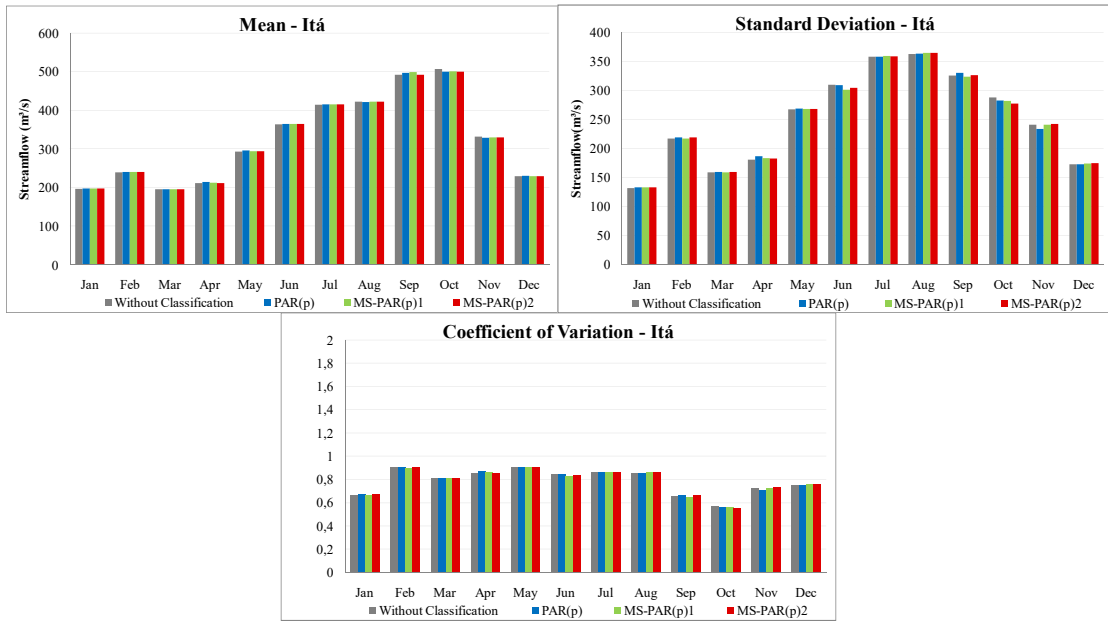


Figure A.36: Statistics obtained for Itá HPP for the unconditioned case.

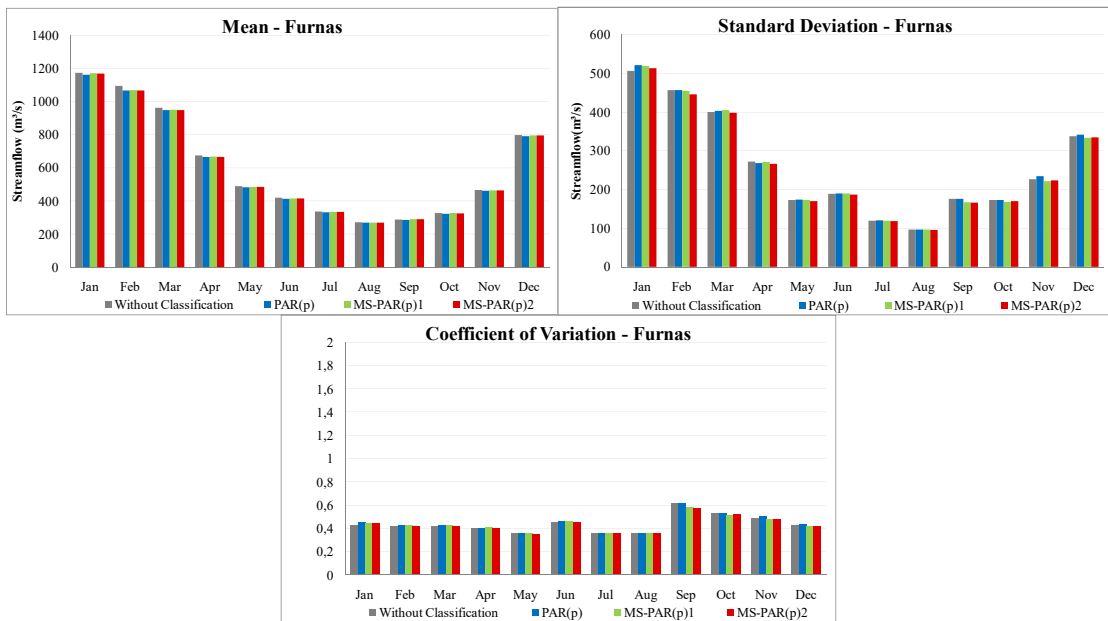


Figure A.37: Statistics obtained for Furnas HPP for the unconditioned case.

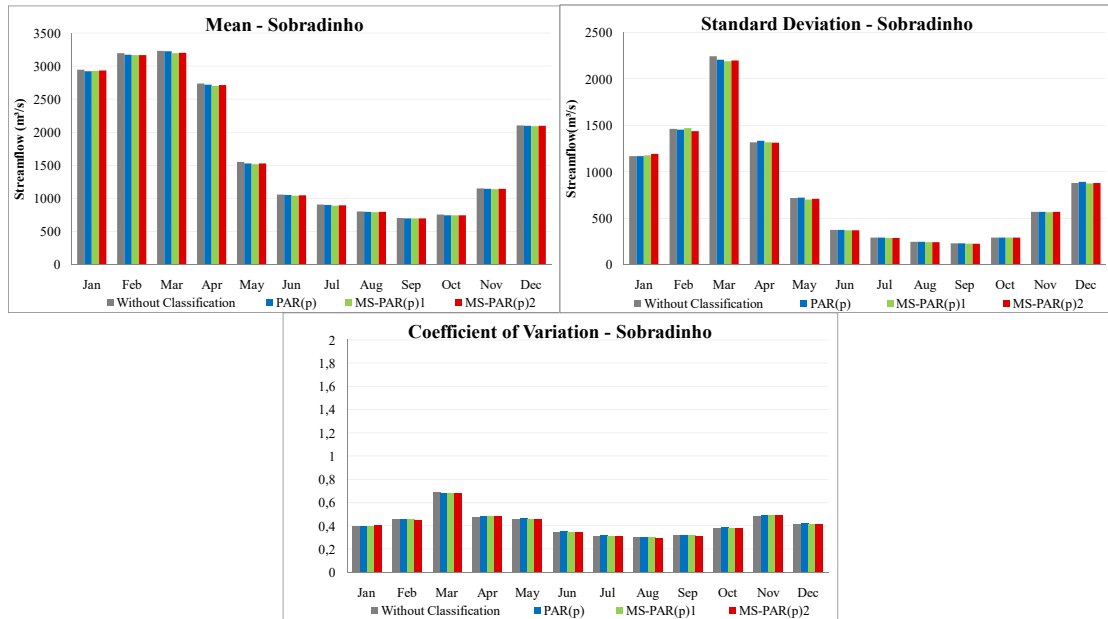


Figure A.38: Statistics obtained for Sobradinho HPP for the unconditioned case.

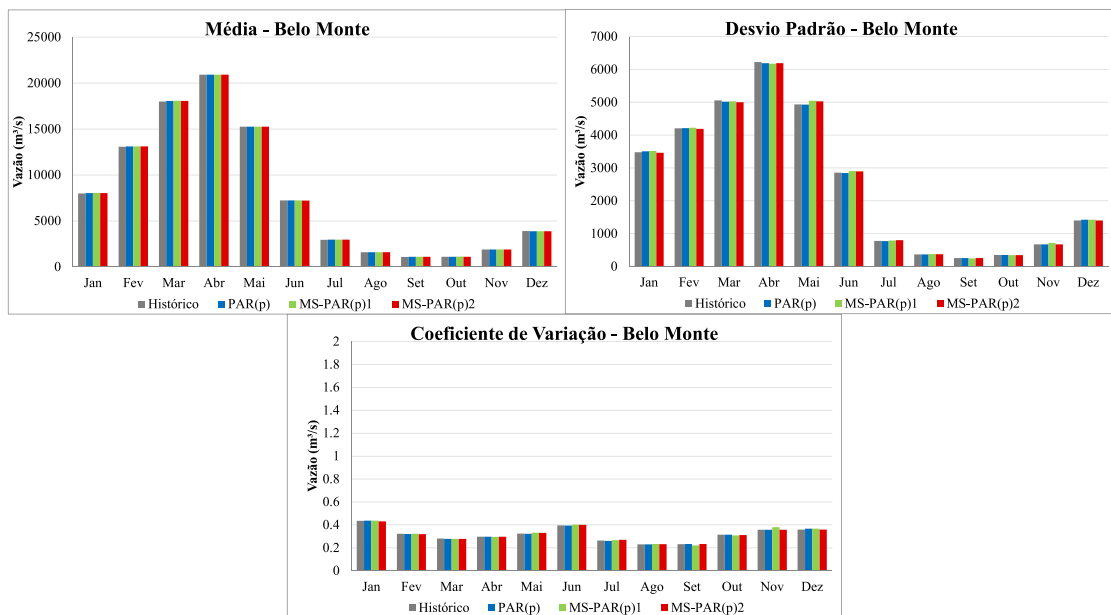


Figure A.39: Statistics obtained for Belo Monte HPP for the unconditioned case.

As shown, the first and second order moments are reproduced correctly, as expected. After the initial analysis of the unconditioned case, Table A.7 summarizes the results of the KS adherence tests performed for each of the 141 HPPs, in all simulated periods (804 periods), totaling 113364 tests that have been performed. In this test, for each simulated period, the synthetic frequency distribution of the 2000 generated scenarios is compared to the historical frequency distribution, using the KS test with a significance level of 5%. As can be observed, the MS-PAR(p)1 model obtained the best results, with MS-PAR(p)2 being the second best model

and finally PAR(p). The total rejections remained low in all three models analyzed, with the difference of the MS-PAR(p)1 model for the PAR(p) of 4% less rejection. It is also worth noting that most of the rejections occur in HPPs that have very small incremental streamflows.

Next, an analysis of the results for each HPP is made, to identify those in which there was a rejection rate greater than 5% and 1% of the 804 simulated periods. The MS-PAR(p)1 model has about half the percentage of rejection of the PAR(p) model, while the MS-PAR(p)2 model shows a decrease in the percentage of rejections in these two analyzes, but smaller than the MS-PAR(p)1 model. This result indicates that there is a significant improvement in the reproduction of the frequency distribution function of some HPP when using the MS-PAR(p) models, with a greater prominence on the model without prior identification of the ENSO's influence.

Table A.7: KS test summary

	PAR(p)	MS-PAR(p)1	MS-PAR(p)2
KS Test			
Total Rejection (%)	11.04%	7.07%	9.08%
HPPs with at least 5% rejected periods (%)	46.8%	23.4%	41.8%
HPPs with at least 1% dos rejected periods (%)	57.5%	31.9%	53.2%

As was explained in Table A.7, the generation of synthetic scenarios for some HPPs cannot reproduce correctly the frequency distribution. After a detailed exploration, it was identified that the rejections, when they occur, are concentrated in certain months for some HPPs. For the Furnas HPP, the PAR(p) model presented around 23% of rejections in June, while in September and October 100% of the periods were rejected. However, the MS-PAR(p)1 and MS- PAR(p)2 have rejections only in October, 70% and 95%, respectively. To identify the factor that causes these months to have a high rejection rate, the third moment, skewness, is analyzed in Figure A.40, for Furnas HPP. It can be seen that between June and October have a very high skewness, which is not reproduced by the PAR(p) model, whereas in the MS-PAR(p) models, the synthetic scenarios are closer to historical values. It can be observed that in October the MS-PAR(p)2 model has asymmetry close to the PAR(p) model, thus having a high rejection rate. According to the results presented in Section A.4.2, October is not affected by the ENSO, which explains the proximity between PAR(p) and MS-PAR(p)2 results for this month.

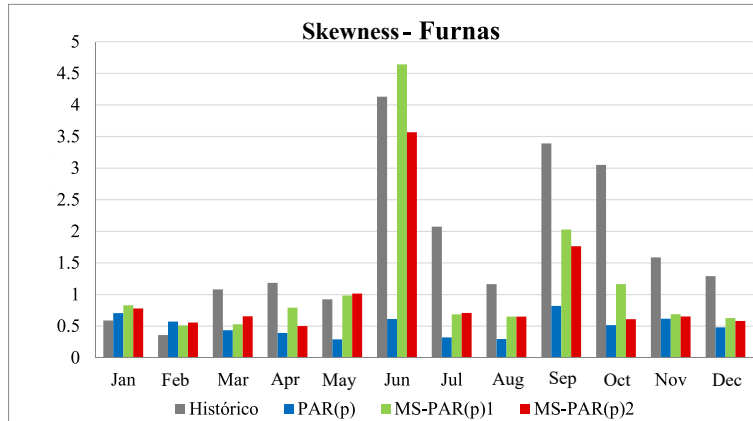


Figure A.40: Historical and synthetic asymmetry of Furnas HPP.

Analyzing the months in which there was a high rejection rate of the KS test between the synthetic and historical frequency distributions, it is observed that for the most part of them, the rejections occur in highly asymmetric distributions, which can not be reproduced in their entirety in the synthetic scenarios. It should be noted this is not something that frequently occurs, as both models are prepared to handle with the high skewed distributions of the streamflow. However, the MS-PAR(p)1 model has a lower total rejection rate of the KS test precisely because it replicates the skewness better when compared to the PAR(p) model. The MS-PAR(p)2 can better reproduce this characteristic when it is indicated that the month in question has ENSO influence. Otherwise, it behaves like the PAR(p) model, which explains the fact that it has an intermediate rejection between the three models analyzed. Thus, it can be inferred that there is an improvement in the MS-PAR(p) modeling regarding the reproduction of the cumulative frequency distribution function in comparison with the PAR(p) model.

An important feature of the NIS is the water complementarity that exists between the different hydrographic regions that compose it. Thus, a good model for the generation of synthetic flow scenarios must be able to reproduce the spatial correlations between the HPPs observed in the historical record. Figure A.41 shows a heatmap of the annual cross-correlations for the history (a), and the simulated values with the PAR(p) (b), MS-PAR(p)1 (c) and MS-PAR(p)2 (d) among the 141 HPPs present in the NIS. The colors in red correspond to positive correlations and blue colors to negative correlations. It is possible to observe that both three models can correctly represent this important historical characteristic, being the values.

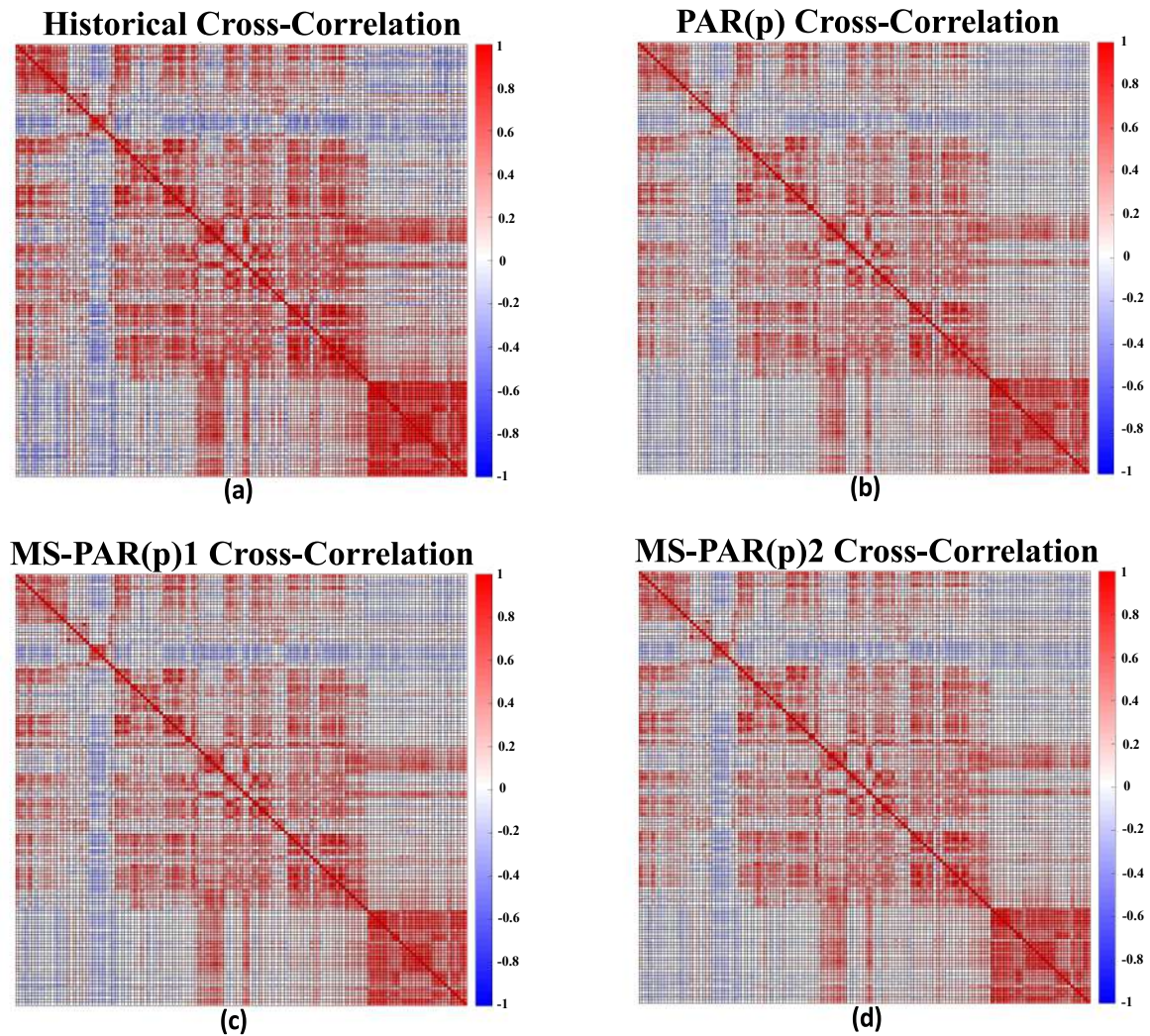


Figure A.41: Historical annual cross-correlations (a) and synthetic ones obtained with the models PAR(p) (b), MS-PAR(p)1 (c) and MS-PAR(p)2 (d)

The generation of scenarios of negative sequences (scenarios that have flows below the LTM) is extremely important for operation planning of the NIS. A good generation model of synthetic streamflow series must be able to reproduce the history of negative sequences, including droughts that occurred in a multivariate way, besides being able to generate even more severe scenarios. Thus, an evaluation of the sum, intensity and length metrics of the negative sequences, as defined in Chapter A.3.5 generated for this unconditioned case, is presented below. In addition to the univariate analysis for each HPP, a multivariate analysis of negative sequences that occurred simultaneously in the HPPs Furnas, Sobradinho, Belo Monte and Itá will be presented. First, a summary of the metrics for all the HPPs of the NIS is presented in the Table A.8. It is possible to observe that the MS-PAR(p)1 model stands out, having the lowest rejection rates of the tests, mainly for the intensity test, where the rejection goes from 30.5% in the PAR(p) model to 15.6%. The MS-PAR(p)2 model is slightly better in the tests for the sum and intensity variables,

and the opposite occurs for the length when compared to the PAR(p).

Table A.8: Negative sequence tests summary

	P	MSP1	MSP2
Rejection of negative sequence sum KS test (%)	10.6%	7.8%	7.8%
Rejection of negative sequence intensity KS test (%)	30.5%	15.6%	27.7%
Rejection of negative sequence length multinomial test (%)	24.1%	24.1%	27.7%

Tables A.9, A.10 and A.11 bring the obtained results for the variables length, intensity and sum of twelve HPPs evaluated individually. In addition to the univariate analysis, negative sequences are also multivariate evaluated. For this purpose, two pairs of HPPs were chosen from different hydrographic regions: Furnas and Belo Monte; Itaipu and Sobradinho; and one analysis carried out taking into account four HPPs Itá, Porto Primavera, Sobradinho and Belo Monte. For each of the negative sequence variables, the mean, standard deviation and maximum value of both the historic and the synthetic values for each model are presented. It is possible to observe that, except in some cases, the averages of the historical and synthetic values of the three negative sequence variables are close, including for multivariate analysis. In addition, all models were able to produce maximum values above historical. This feature demonstrates their ability to simulate even more critical scenarios than those observed in history, which persist for a greater amount of time, and are even more intense.

A.4.4 Conditioned cases - in-sample period

In this section, it is presented the results obtained through the generation of 2000 synthetic scenarios for the in-sample period from 1951 to 2009 . For each period, the hydrological tendency is updated with the recent past, and for the MS-PAR(p) model the ENSO state throughout the period to be simulated is a previously known data (perfect prediction of the ENSO state). For the analysis of the results in this section, the MAE, MAPE and RMSE metrics, presented in Chapter A.3.5, will be used and calculated for each HPP in all simulated periods.

Table A.12 presents the evaluation of the conditioned cases for the in-sample period in the selected HPPs, throughout all periods, obtaining an average value for each evaluation metrics. The values highlighted in bold correspond to the lowest metric value among the three models analyzed, and consequently, indicate the best

Table A.9: Historic and synthetic statistics of negative sequence length

	Mean						Maximum						Standard Deviation					
	Historic		MSP1		MSP2		Historic		MSP1		MSP2		Historic		MSP1		MSP2	
	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic
Sobradinho	4.9	4.9	4.6	4.4	25	87	71	77	5.5	5.9	5.2	5.0						
Belo monte	4.0	4.1	3.9	3.9	26	62	57	49	4.5	4.3	4.1	4.0						
Estreito Toc.	4.1	4.5	4.3	4.3	26	75	80	75	4.4	4.7	4.8	4.5						
Jirau	3.7	3.6	3.5	3.4	35	81	89	81	6.2	5.6	5.5	5.5						
Sto. Ant. Jari	4.2	4.1	4.0	4.0	27	74	61	60	5.4	4.9	4.5	4.3						
Teles Pires	4.1	4.6	4.3	4.3	24	83	65	77	4.3	5.2	4.5	4.7						
Capivara	5.0	4.1	4.1	4.0	37	63	60	74	5.8	4.7	4.5	4.4						
Furnas	4.2	4.8	4.5	4.5	39	134	83	71	4.8	6.9	5.6	5.6						
P. Primavera	4.0	3.8	3.8	3.7	30	90	83	86	5.5	5.1	4.5	5.0						
Itá	4.3	3.8	3.7	3.7	16	44	42	45	3.7	3.6	3.5	3.6						
Itaipu	4.8	4.4	4.2	4.1	41	128	128	125	7.0	6.1	5.8	6.1						
Salto Caxias	4.1	3.7	3.7	3.6	32	55	49	45	4.6	3.7	3.5	3.6						
Furnas + Belo Monte	2.9	2.6	2.4	2.4	11	54	33	32	2.6	2.6	2.3	2.3						
Itaipu + Sobradinho	2.3	2.4	2.3	2.2	12	39	37	32	2.0	2.4	2.2	2.1						
Itá + P. Primavera																		
+ Sobradinho	1.4	2.3	2.4	2.3	4	34	33	38	0.8	2.1	2.1	2.1						
+ Belo Monte																		

Table A.10: Historic and synthetic statistics of negative sequence intensity

	Mean						Maximum						Standard Deviation					
	Historic		MSP1		MSP2		Historic		MSP1		MSP2		Historic		MSP1		MSP2	
	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic	P	Historic
Sobradinho	440	500	498	502	1736	3065	3058	3081	377	402	403	407						
Belo monte	1269	1551	1536	1577	3750	12738	12847	12725	1067	1401	1423	1453						
Estreito Toc.	355	401	380	403	1264	2818	2589	2781	254	318	308	322						
Jirau	978	1238	1238	1184	4792	8909	10028	9564	1237	1391	1413	1396						
Sto. Ant. Jari	239	239	241	243	798	1574	1785	1777	191	187	190	188						
Teles Pires	316	358	353	354	1087	2339	2296	2225	274	264	264	265						
Capivara	114	133	124	130	299	512	499	508	63	75	70	71						
Furnas	138	136	138	138	573	1146	1009	963	105	112	112	109						
P. Primavera	322	258	252	258	1390	1182	1198	1172	282	174	169	175						
Itá	130	132	127	129	300	492	491	492	73	71	68	69						
Itaipu	575	631	592	618	1570	2158	2127	2174	380	371	359	373						
Salto Caxias	125	128	123	126	260	439	408	401	56	65	63	63						
Furnas + Belo Monte	1691	1970	1929	1953	9374	13525	14121	13301	1648	1850	1883	1872						
Itaipu + Sobradinho	1270	1256	1195	1257	3365	4716.1	4584	4770	678	664	630	663						
Itá + P. Primavera																		
+ Sobradinho	502	451	432	446	1217	1365	1393	1380	262	205	198	205						
+ Belo Monte																		

Table A.11: Historic and synthetic statistics of negative sequence sum

	Standard Deviation											
	Mean						Maximum					
	Historic	P	MSP1	MSP2	Historic	P	MSP1	MSP2	Historic	P	MSP1	MSP2
Sobradinho	2733	2932	2647	2563	27245	72491	60105	64951	4680	4677	3896	3818
Belo monte	7092	7802	7443	7381	47349	152356	151282	136546	9547	11410	11053	10832
Estreito Toc.	1905	2034	1930	1944	13692	52532	55455	53872	2422	2824	2919	2724
Jirau	13614	10039	9678	9637	158546	395763	403514	439209	23850	21407	21452	21659
Sto. Ant. Jari	1188	1303	1233	1213	11286	36087	27171	25633	1780	2192	1924	1810
Teles Pires	2279	2070	1873	1932	14667	60635	40290	53693	2842	3380	2808	3035
Capivara	805	690	642	647	8345	14671	13709	14393	1234	1011	911	902
Furnas	1046	972	838	849	10695	40964	24163	23481	1990	2013	1467	1482
P. Primavera	1532	1231	1138	1205	10978	48561	34904	48145	2262	2296	1820	2259
Itá	500	574	541	545	2167	8390	7562	7791	517	662	615	639
Itaipu	3649	3956	3532	3616	35996	176358	168307	180286	5443	7680	6829	7746
Salto Caxias	572	551	529	525	3216	8903	7451	7538	609	681	600	635
Furnas + Belo Monte	6194	5550	5110	5169	48098	142278	102237	104671	9359	7873	7293	7262
Itaipu + Sobradinho	2870	3292	2983	3038	17743	75092	78175	73826	2893	4394	3729	3909
Itá + P. Primavera												
+ Sobradinho	700	1167	1120	1142	3259	19193	20312	26379	524	1388	1257	1369
+ Belo Monte												

of them according to the metric in question. The MS-PAR(p)1 model was indicated as the best, mainly in the southern region, obtaining a MAPE reduction in relation to the PAR(p) model of almost 10% in the Itá HPP. The MS-PAR(p)2 model in many cases ends up reducing the error in relation to the PAR(p) model, but it is worse than the MS-PAR(p)1. The MS-PAR(p)1 model has smaller errors in 72% of the HPPs compared to the PAR(p) model, especially in the south, northeast and north regions of Brazil. The MS-PAR(p)2 model is indicated as superior in 46.7% of the HPPs, mainly in those regions previously highlighted.

Table A.12: In-sample Evaluation

HPP	region	EMA			EMPA			REMQ		
		P	MSP1	MSP2	P	MSP1	MSP2	P	MSP1	MSP2
Sobradinho	Northeast	622.04	599.64	610.14	31.70	30.82	31.24	773.03	745.07	757.76
Belo Monte	North	2023.05	1976.18	2000.41	25.49	25.03	25.34	2470.60	2414.20	2444.85
Estreito Toc.	North	480.92	459.43	470.68	24.78	23.99	24.30	588.62	561.61	576.31
Jirau	North	1965.58	1982.24	1955.44	13.81	13.76	13.65	2373.14	2394.92	2361.09
Sto. Ant. Jari	North	310.66	313.99	316.42	36.26	36.20	37.24	380.74	384.93	387.77
Teles Pires	North	424.77	431.63	439.16	19.46	20.12	20.38	517.38	525.43	535.04
Capivara	Southeast	187.06	181.53	186.88	60.89	58.90	61.18	247.12	237.67	244.84
Furnas	Southeast	179.72	184.44	190.08	30.53	32.18	32.93	220.26	227.86	234.00
P. Primavera	Southeast	350.42	358.25	351.48	67.02	66.87	66.30	439.25	452.78	441.38
Itá	South	203.02	196.09	201.04	101.56	93.17	96.79	269.62	261.57	268.26
Itaipu	South	855.36	811.25	825.12	71.18	68.61	67.99	1090.64	1033.14	1053.43
Salto Caxias	South	187.01	180.94	188.24	93.90	88.94	93.05	248.90	240.45	250.98

As the GEVAZP model generates monthly synthetic scenarios, there is also a need to evaluate the results month by month, thus identifying, in which periods there is a greater potential gain with the MS-PAR(p) model. Tables A.13, A.14 and A.15 bring the monthly MAPE to each of the analyzed HPPs. It is possible to note that the PAR(p) model is better than the MS-PAR(p) models mainly in the recession months, which are drier months in most of Brazil, except for the south region. The HPPs Furnas, Porto Primavera, and Teles Pires, as shown in Table A.12, present better results with the PAR(p) model in most months. However, it is noteworthy that for most HPPs, and especially during the wet period, there is a large advantage of the MS-PAR(p) models on PAR(p).

A possible explanation for these results may be the fact that the ENSO phenomenon has a greater influence on the precipitation, and consequently, on the streamflows that occur during the wet season. As in the recession period the streamflows occur more as a product of the past values, the model MS-PAR(p) no longer has an advantage over PAR(p). The best information that a model can use to generate future scenarios of streamflows during the dry period is the magnitude of flow observed during the wet period, and any other type of information may end up introducing more noise to the model. However, it should be noted that the streamflows occur in a much larger proportion during the wet period, and that the fact that the MS-PAR(p) models present a lower MAPE in these periods can be considered an advantage over the model PAR(p).

Table A.13: Monthly MAPE of the PAR(p) model

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Sobradinho	39	49	53	47	21	16	10	9	13	24	47	51
Belo Monte	45	29	21	32	27	21	20	12	15	25	29	30
Estreito Toc.	31	43	30	33	28	12	12	8	14	22	30	34
Jirau	11	8	9	9	9	9	13	13	22	22	27	15
Sto. Ant. Jari	50	37	34	35	31	31	28	24	28	31	42	62
Teles Pires	23	19	23	19	14	7	6	5	10	35	41	31
Capivara	88	67	44	50	69	54	52	50	59	76	56	65
Furnas	41	40	33	27	19	27	15	18	39	35	36	36
P. Primavera	168	101	93	43	35	44	25	29	37	56	53	122
Itá	82	100	96	99	171	119	86	134	96	73	80	82
Itaipu	101	85	75	51	73	49	42	59	57	65	50	147
Salto Caxias	72	95	104	96	170	107	92	99	81	75	62	74

Table A.14: Monthly MAPE of the MS-PAR(p)1 model

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Sobradinho	35	43	52	46	18	16	10	11	14	27	46	51
Belo Monte	40	28	20	34	25	21	21	13	15	25	30	27
Estreito Toc.	28	44	28	29	29	11	12	12	14	23	27	31
Jirau	10	7	9	10	11	10	13	14	19	22	26	14
Sto. Ant. Jari	49	39	34	34	30	31	32	26	28	32	41	59
Teles Pires	24	19	24	18	14	9	8	7	13	35	40	30
Capivara	84	63	51	50	61	54	49	50	60	69	50	66
Furnas	41	37	33	28	20	22	21	20	46	41	39	38
P. Primavera	157	110	96	39	35	43	35	29	40	56	52	110
Itá	74	98	84	95	132	100	93	135	93	69	64	80
Itaipu	93	78	85	45	70	47	42	67	53	60	46	137
Salto Caxias	70	81	89	101	165	105	91	96	79	70	54	67

Table A.15: Monthly MAPE of the MS-PAR(p)2 model.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Sobradinho	36	46	50	47	21	16	12	11	14	27	43	50
Belo Monte	42	28	20	33	26	24	21	12	15	25	30	28
Estreito Toc.	28	43	29	31	30	12	12	8	14	22	30	32
Jirau	10	8	9	9	9	9	13	13	20	22	26	15
Sto. Ant. Jari	51	39	34	34	30	31	32	27	27	31	44	67
Teles Pires	23	19	23	22	15	8	10	7	11	34	42	31
Capivara	87	67	51	52	72	56	49	50	59	77	49	65
Furnas	43	38	34	30	23	21	21	20	46	40	41	37
P. Primavera	167	103	94	41	35	44	25	29	37	56	53	111
Itá	76	100	92	101	159	103	88	134	97	70	61	79
Itaipu	100	72	75	44	73	49	42	64	57	61	50	128
Salto Caxias	79	95	94	102	170	109	91	95	82	73	52	76

Another point that should be highlighted in both models of the MS-PAR(p) type is the periods where they are better than the PAR(p) model. As shown in Table A.5, LN events cause negative anomalies in the streamflows, mainly in the southern region, and with lower intensity in part of the southeast/central-west and part of the northeast. The north region, mostly in its upper portion, suffers a reduction in inflows during EN events. Table A.16 presents the same evaluation done previously

for the MAPE but dividing it between the states of ENSO. As can be observed, the Santo Antônio Jari HPP has better performance during EN events with the MS-PAR(p)1 and MS-PAR(p)2 models, presenting a reduction of approximately 4% and 3% respectively with relation to the PAR(p) model. The same behavior occurs at the Estreito Tocantins HPP. On the other hand, the MS-PAR(p)1 and MS-PAR(p)2 models have the best results in LN periods, mainly in the rest of Brazil, especially those located in the south. In some specific cases, this reduction reaches more than 10% in Itá and Salto Caxias. The HPPs of the south and southeast regions also present better performance in N states.

Table A.16: MAPE of the in-sample period divided among the states of ENSO

	La Niña						Neutro			El Niño									
	P		MSP1		MSP2		P		MSP1		MSP2		P		MSP1		MSP2		
Sobradinho	35.82	29.04	31.74	28.43	29.78	29.62	33.32	34.16	33.45	29.78	29.62	33.32	34.16	33.45	29.78	29.62	33.32	34.16	33.45
Belo Monte	23.99	22.61	24.26	24.26	24.13	23.90	28.88	28.72	28.68	24.13	23.90	28.88	28.72	28.68	24.13	23.90	28.88	28.72	28.68
Estreito Toc.	24.16	21.64	23.64	24.31	25.22	24.65	26.12	24.10	24.31	25.22	24.65	26.12	24.10	24.31	25.22	24.65	26.12	24.10	24.31
Jirau	15.11	13.13	13.88	12.74	13.55	12.91	14.39	14.67	14.64	13.55	12.91	14.39	14.67	14.64	13.55	12.91	14.39	14.67	14.64
Sto. Ant. Jari	34.34	37.50	38.99	32.55	33.05	33.76	44.10	40.17	41.37	33.05	33.76	44.10	40.17	41.37	33.05	33.76	44.10	40.17	41.37
Teles Pires	22.04	21.66	22.76	17.38	18.64	18.03	20.51	21.14	22.05	18.64	18.03	20.51	21.14	22.05	18.64	18.03	20.51	21.14	22.05
Capivara	56.72	53.95	54.40	63.64	54.82	61.67	60.20	70.13	66.55	54.82	61.67	60.20	70.13	66.55	54.82	61.67	60.20	70.13	66.55
Furnas	34.06	37.03	37.19	27.78	27.39	29.06	31.84	35.62	35.41	27.39	29.06	31.84	35.62	35.41	27.39	29.06	31.84	35.62	35.41
P.Primavera	61.14	60.41	60.57	73.36	68.32	71.33	61.98	70.39	63.28	68.32	71.33	61.98	70.39	63.28	68.32	71.33	61.98	70.39	63.28
Itá	86.88	74.90	77.02	113.40	100.86	109.73	95.54	93.62	93.62	100.86	109.73	95.54	93.62	93.62	100.86	109.73	95.54	93.62	93.62
Itaipu	62.49	57.82	60.03	81.10	75.51	74.11	62.82	67.13	65.23	75.51	74.11	62.82	67.13	65.23	75.51	74.11	62.82	67.13	65.23
Salto Caxias	81.29	69.45	79.86	107.02	93.16	104.40	83.88	99.82	86.48	93.16	104.40	83.88	99.82	86.48	93.16	104.40	83.88	99.82	86.48

To analyze the results presented in the table A.16 in a more practical way for the generation of synthetic scenarios, it is present below some figures that bring the distribution of the generated scenarios by each model in some of the simulated years. The graphic shown is of type fanchart, where each color range represents a percentage concentration of the synthetic series. The stronger the color, the smaller the represented range generated. The dashed line in black symbolizes the median of the distribution and the full line in black with markers the historical values. To facilitate the visualization of the graphs, and to reduce their scale, 90% of the scenarios will be presented, thus excluding the most extreme ones. In the context of scenario generation model, it is expected that the observed values are at least contained in any of the bands.

The period between January 1953 and December 1956 was selected to be analyzed first. These years correspond to the central part of the Brazilian critical period, the period in which the NIS had the worst multivariate drought recorded (since 1931). The year 1953 began with a state of EN that persisted until February 1954, followed by the formation of an LN from May until September 1956. For this first subset of periods, it will be evaluated the generation of scenarios for the Santo Antônio Jari HPP, located in the north region of Brazil, and for the Salto Caxias HPP, located in the south region, thus observing the variation of hydrological regimes during the critical period.

Figure A.42 shows the scenarios generated in the period mentioned above for the Santo Antônio Jari HPP with the models PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) and its historical distribution (d). In general, it is possible to affirm that the three generations realized managed to capture correctly the behavior observed in the historical series. However, there are important differences that must be highlighted. It is observed that in the first year, the historical series of Santo Antônio Jari HPP has values well below those recorded in subsequent years, due to the presence of the EN phenomenon. By analyzing the scenarios generated by the three models, it is possible to see that for the year 1953 the PAR(p) model generates a spreader and wetter scenarios distribution, especially in May compared to the MS-PAR(p) models. Also, the last two models can capture the behavior of this year, even with lower variability of the synthetic series generated. As shown in Figure A.33, periods of EN tend to be drier and have a lower uncertainty. The opposite occurs in the LN, as observed in subsequent years between 1954 and 1956. In 1955, there was a sharp decrease in the streamflow from April to May that could not be present in any of the PAR(p) model scenarios but captured by the MS-PAR(p) model due to the greater variability of the scenarios generated. For this HPP, there are no big differences between the generation of scenarios performed by the MS-PAR(p) models among themselves, since according to the ENSO influence

test presented in A.31, most months are affected for both phases.

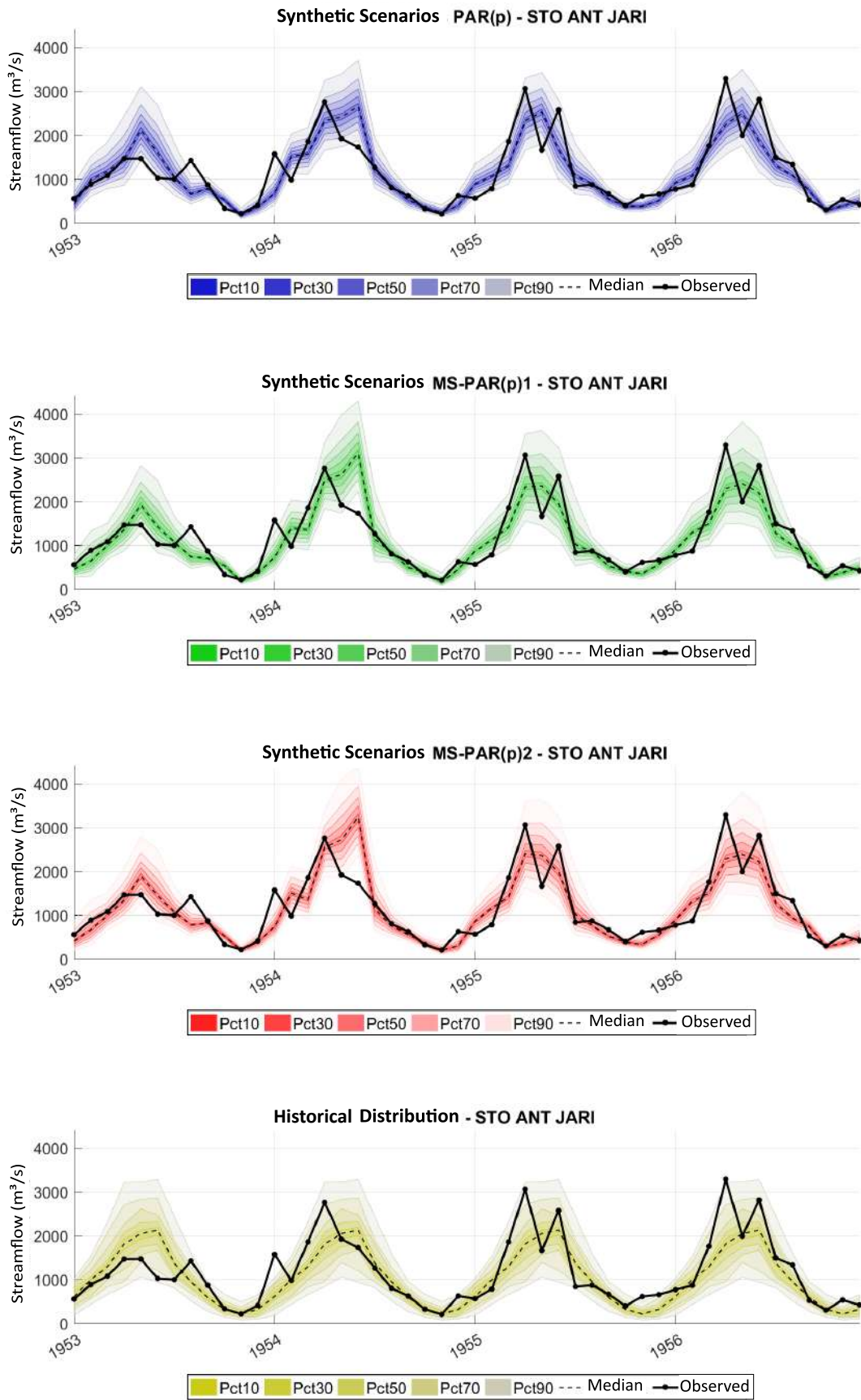


Figure A.42: Synthetic scenarios generated for the period between 1953 and 1956 at the Santo Antônio Jari HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).

Figure A.43 presents the synthetic scenarios generated to the Salto Caxias HPP, performed by the models PAR(p) (a), MS-PAR(p)1 (b) and MS-PAR(p)2 (c), respectively, and its historical distribution (d). Firstly, it is possible to observe that there is large variability in the series observed in Salto Caxias, a striking feature of the HPPs located in the south region. Once again, it can be seen that the three models can generate scenarios in a way that adheres to the observed values, with some differences from each other. Unlike in the Santo Antônio Jari HPP, in the Salto Caxias HPP the generation of the MS-PAR(p)2 model is closer to the PAR(p) model. As Figure A.21 shows, only the state of EN in March and November is identified as influential. Analyzing the year 1953, it is possible to observe that the model MS-PAR(p), generates scenarios with greater variability than in the other two models, due to the state of EN. From 1954, already in the state of LN, its generation has a smaller dispersion in the MS-PAR(p)1 model, without compromising the adherence to the observed values.

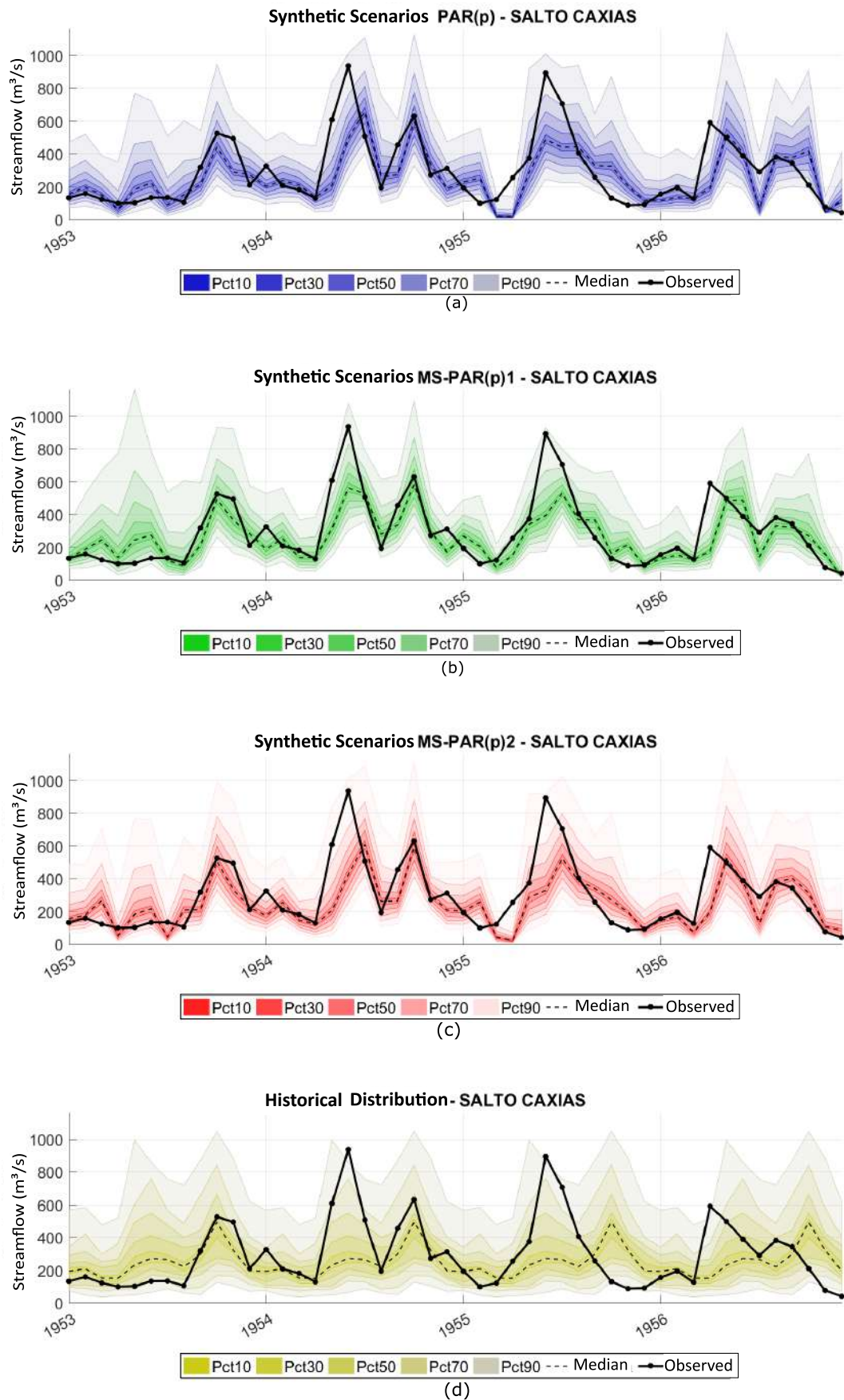


Figure A.43: Synthetic scenarios generated for the period 1953 to 1956 in the Salto Caxias HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).

The results shown in this section demonstrate that the MS-PAR(p) models are better when evaluated according to the MAE, MAPE and RMSE metrics, compared to the PAR(p) model, for most of the HPPs. The difference between the results is higher during the wet period, and during the periods where the streamflows are below the LTM for the most part of Brazil, during LN periods. The HPPs that suffer a greater influence of the ENSO phenomenon are those that the MS-PAR(p) model stands out about the PAR(p) model. It can be affirmed that the distributions of multivariate scenarios were generated more closely to the values observed in the history, having this distribution a smaller or larger standard deviation depending on the state of ENSO. These results, however, should be analyzed with caution. Given that the simulated period was contained in the adjustment period of the models, it is expected that models with more adjusted parameters (which is the case of the MS-PAR(p) models) have a better performance than models with fewer. An evaluation of the scenario generation for the out-of-sample period (2010 to 2017) will be presented in the following section, in order to better identify the potential of the methodology developed in this dissertation.

A.4.5 Conditioned cases - out-of-sample period

This section will address with the conditional cases performed in the out-of-sample period. The period selected for the evaluation corresponds to the years between 2010 and 2017. For each simulated year, the models will be adjusted until the previous observed year. In other words, for the year 2010 the models were adjusted with the streamflow history until 2009; for the year 2011, the models were adjusted with the data until 2010, and so on. For each simulated period, generation will be carried out up to five years ahead, to evaluate the generation of synthetic scenarios for both the mid-term and the long term. Regarding the ENSO phenomenon, the probabilistic forecasts provided by IRI will be used as described in chapter A.3.4, for the generation of the ENSO scenarios.

Therefore, before assessing the generation of streamflow scenarios, since the MS-PAR(p) models depend on the ENSO states, a first evaluation will be made of the quality of the ENSO scenarios generated by the models. During the analyzed horizon, there were two EN events: the first one, in its end, at the beginning of 2010; and a strong EN between the end of 2014 and mid-2016. There were also four LN events: the first between mid-2010 and May 2011; a second event that can be considered as the continuation of the previous one, from July 2011 to the beginning of 2012; the third event a weak LN that only lasted in the last five months of 2016; and the previous LN event that was formed in October 2017. In the remaining periods, state N was maintained. Still, over the analyzed period, it is highlighted

that the years of 2014 and 2015 are among the five worst hydrological years of history.

Figure A.44 brings the ENSO history compared to the scenarios generated by the MS-PAR(p) model for the evaluated period said before. Positive values (red colors) represent the probability of formation of EN when this value is greater than the climatological probability; negative values (blue colors) represent (in module) the probability of LN formation when this value is greater than the climatological probability; zeros (in green) represent a higher probability of occurrence of N state; non-existent data is presented in white. The line below the graph represents the observed ENSO state, while the subsequent lines represent the probabilistic predictions of ENSO performed at different time points. It can be observed that the generated scenarios of ENSO are quite adherent to what happened in fact.

In general, the smaller the difference between the prediction and the observation time, the greater the percentage of scenarios generated in the correct category, just as the opposite is also true. During the analyzed horizon, there were two cases of false alarm, when the model predicted for most scenarios the occurrence of an EN at the end of 2012 and in the middle of 2017, a fact that was not confirmed. However, it should be noted that even for the generation of scenarios carried out for a horizon up to twelve months ahead, the quality of forecasts is satisfactory in many cases. For example, a large part of the scenarios generated in June 2010 indicated the formation of an LN that would persist until May 2011. The same can be said for the strong EN between the end of 2014 and the middle of 2016.

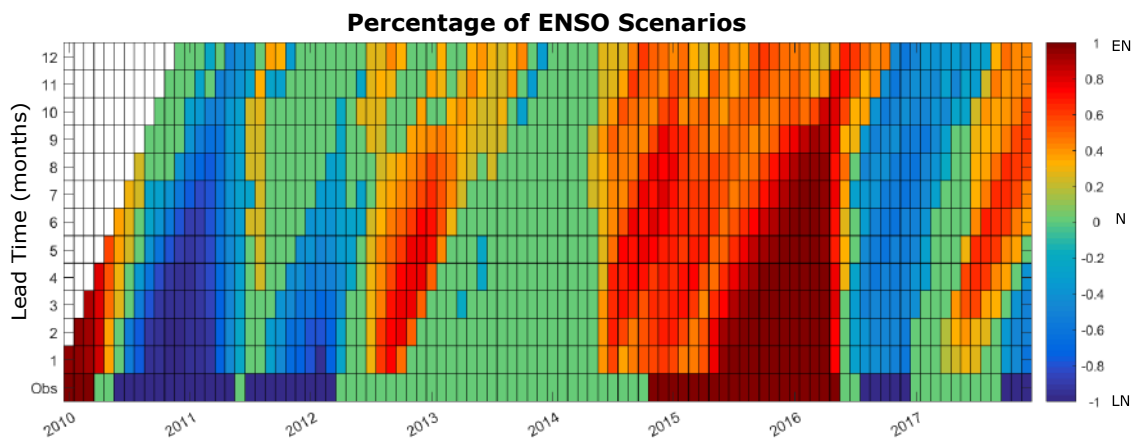


Figure A.44: Percentage of ENSO scenarios generated up to twelve months ahead for the period between 2010 and 2017.

To quantify the gain of using the probabilistic predictions of IRI about climatological probability, the RPSS metric was calculated for the same generation horizon, as presented in Equation A.45. For this metric, positive values indicate that the probabilistic forecast is better than the use of climatological probability,

with value 1 attributed to a perfect prediction. The RPSS for the generation of ENSO scenarios concerning the different time steps between the observation and forecast period is presented in Figure A.45. Indices less than zero, represented by white color, indicate that the use of climatology is better than the consideration of probabilistic forecasting. Once again it is possible to observe the quality of the generated ENSO scenarios, even for long forecast periods in some cases. It is noted that the quality of the scenarios after an ENSO event is started is higher than in the initial periods, which gives Figures a ramp aspect.

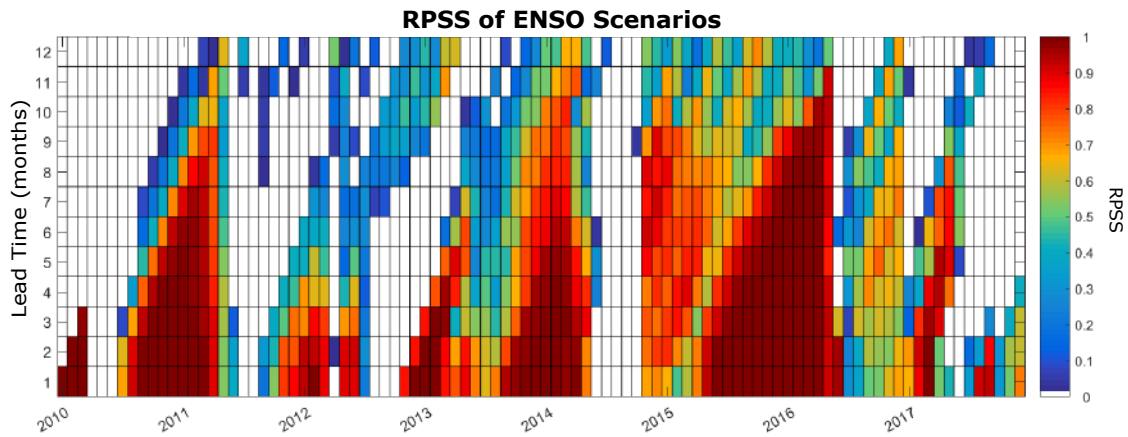


Figure A.45: RPSS depending on observed period (horizontal axis) and lead time (vertical axis), for the period between 2010 and 2017.

To synthesize the obtained results, the median RPSS for each month is presented about the forecast horizon in Figure A.46. It was decided to present the median, because in a few periods a very negative RPSS value occurred, which would lead to a misunderstanding of the quality of the ENSO scenario generation if the average was presented. It is observed that for most months the probabilistic forecast is much better than the use of the climatology in the first months, decreasing as the generation departs from the observation time. Between June and August, it is presented the lowest RPSS values for generation one month ahead. This result is in agreement with other studies cited in this dissertation about the difficulty of predicting during this period, when in many cases ENSO events begin to form. In the other months, the RPSS oscillates of values close to one to 0.5 in the horizon of eight months, which confirms the quality of the scenarios generated.

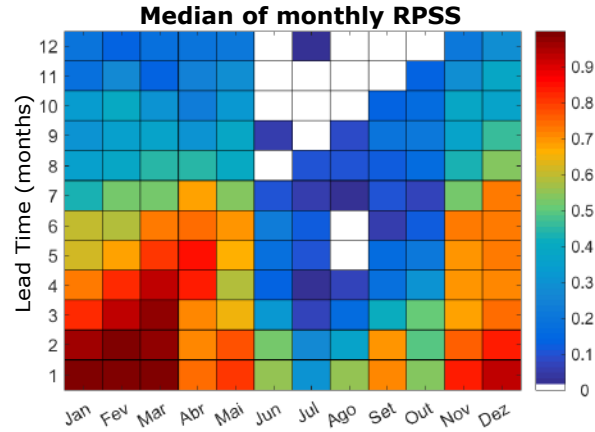


Figure A.46: Median of the monthly RPSS with respect to the forecast horizon.

After verifying the generation of ENSO scenarios, an analysis of the generation of streamflow scenarios of the three models evaluated in this dissertation will be presented. For each month between January 2010 and December 2017, were generated 2000 synthetic scenarios with five years horizon. The MAE, MAPE, and RMSE metrics were computed for up to six months ahead in each simulation, aiming to evaluate generation beyond the short term. The average results obtained for the generation of scenarios one month ahead, for the thirteen selected HPPs are presented in Table A.17. It is possible to observe that in general the MS-PAR(p) models have better results in the north and south regions in the three metrics analyzed. The only exception is the Itaipu HPP, which presents better results in the PAR(p) model. In general, a predominance of the MS-PAR (p) 2 model on the MS-PAR (p) 1 model is observed.

This result is precisely the opposite obtained when the generation evaluation was made in the in-sample period, as shown in Table A.12. This may be an indication that the MS-PAR(p)1 model in many cases performs an over-fitting in the parameters for some HPPs, which ends up introducing more noise into the model. However, in some cases, this type of adjustment presents better results, mainly in the north region. The MS-PAR(p)2 model, through the identification of the influence of ENSO, can often produce better results, mainly for the HPPs of the southern region. This region has better results with the scenarios generation of performed by the MS-PAR(p)2 model, decreasing the MAPE by 0.7% in relation to the PAR(p) model, with an improvement in exactly half of the HPPs composing it. Although the Capivara HPP has better results with the PAR(p) model, 80% of the HPPs (out of ten) belonging to the Paranapanema river basin have better results with the MS-PAR(p) model. In the rest of the southeast region, in general, the PAR(p) model has superior outcomes.

Table A.17: Evaluation of the out-of-sample generated scenarios for the period between 2010 and 2017

UHE	Região	EMA			EMPA			REMQ		
		P	MSP1	MSP2	P	MSP1	MSP2	P	MSP1	MSP2
Sobradinho	Northeast	545.87	604.09	584.62	53.32	59.20	57.34	720.49	788.93	765.53
Belo Monte	North	1958.09	1851.90	1905.20	27.22	26.29	26.98	2472.78	2387.70	2416.12
Estreito Toc.	North	428.75	422.14	426.81	27.15	27.34	27.14	554.00	549.15	552.43
Jirau	North	2667.69	2710.61	2659.78	19.49	19.55	19.16	3075.92	3146.19	3072.86
Sto. Ant. Jari	North	335.43	319.08	322.31	34.78	34.07	34.85	402.17	389.04	392.77
Teles Pires	North	450.96	432.70	444.94	20.19	19.49	19.68	538.53	524.70	535.94
Capivara	Southeast	241.45	249.06	244.62	47.13	49.12	48.70	298.76	309.51	301.97
Furnas	Southeast	184.30	200.61	182.13	51.89	58.10	52.95	230.17	252.48	236.66
P. Primavera	Southeast	358.86	357.61	352.74	34.29	33.94	33.40	436.08	442.68	429.76
Itá	South	276.06	277.84	270.97	121.83	124.50	119.39	339.01	344.25	336.48
Itaipu	South	1309.84	1351.54	1407.98	39.60	41.75	42.19	1507.54	1565.08	1619.13
Salto Caxias	South	261.51	255.47	255.11	86.29	83.37	83.59	318.94	315.75	314.38

The distribution of scenarios generated for some of the HPPs will be analyzed to better understand the differences between the generation of the three models. Figures A.47 and A.48 show the synthetic series generated one month ahead for the Belo Monte HPP between 2010 and 2013, and 2014 to 2017, respectively. In general, both three models generated a distribution of suitable scenarios for the Belo Monte HPP. It should be noted that 2010, 2015 and 2016 years were the worst wet periods for the Belo Monte HPP between 2010 and 2017. Not coincidentally, in both periods the EN was acting. It can be observed that in the three periods, the MS-PAR(p)1 model can generate scenarios closer to the observed value, especially in April, which is when the occurs the higher streamflows in the Belo Monte HPP. In April of 2015 only the model MS-PAR(p)1 can generate scenarios of the same magnitude of the observed flow, which was extremely above the LTM.

The reason why the MS-PAR(p)2 model failed to capture the same behavior is the fact that no ENSO state was selected for April as influential in the Belo Monte HPP, as presented in Table A.31. If it had been selected, probably the MS-PAR(p)2 model would have generated more adherent scenarios in April, demonstrating that for some cases the process of ENSO's influence identification proposed here may not be ideal. It is important to highlight that for April the three models identify an AR (1) with an autoregressive coefficient of approximately 0.15. In other words, the streamflows of April are weakly correlated with the streamflow of the previous month. The use of ENSO climatic information, in this case, brought a clear benefit to the generation of scenarios, bringing relevant information to the model.

Figures A.49 and A.50 bring the generation of scenarios a one step ahead for Itá HPP. It is possible to notice that the series observed during the simulated period has a very erratic behavior, especially in the year of 2015, when a strong EN occurred. Nevertheless, it can be affirmed that the three models can generate scenarios close to the observed values.

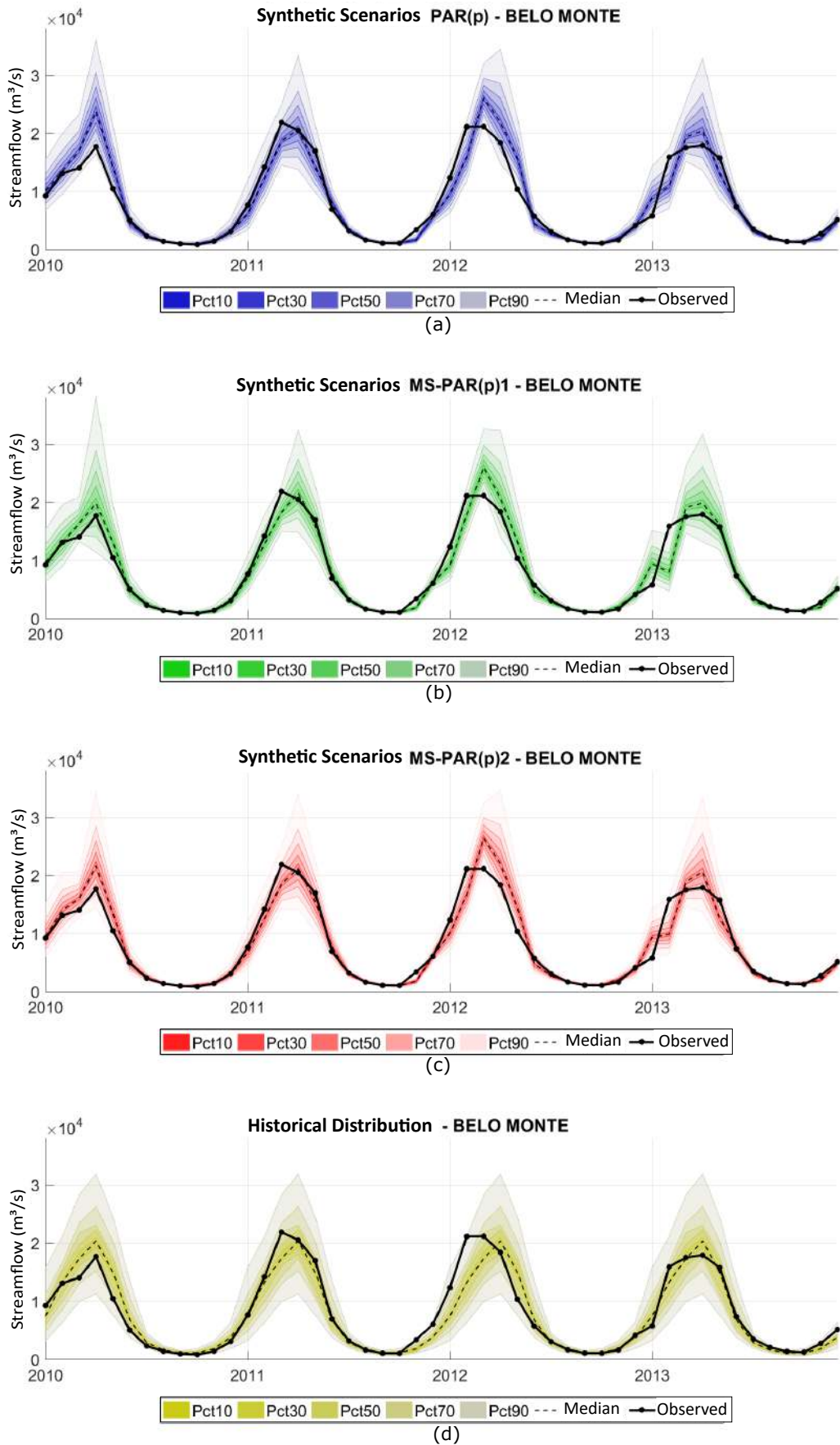


Figure A.47: Synthetic scenarios generated for the period between 2010 and 2013 in the Belo Monte HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).

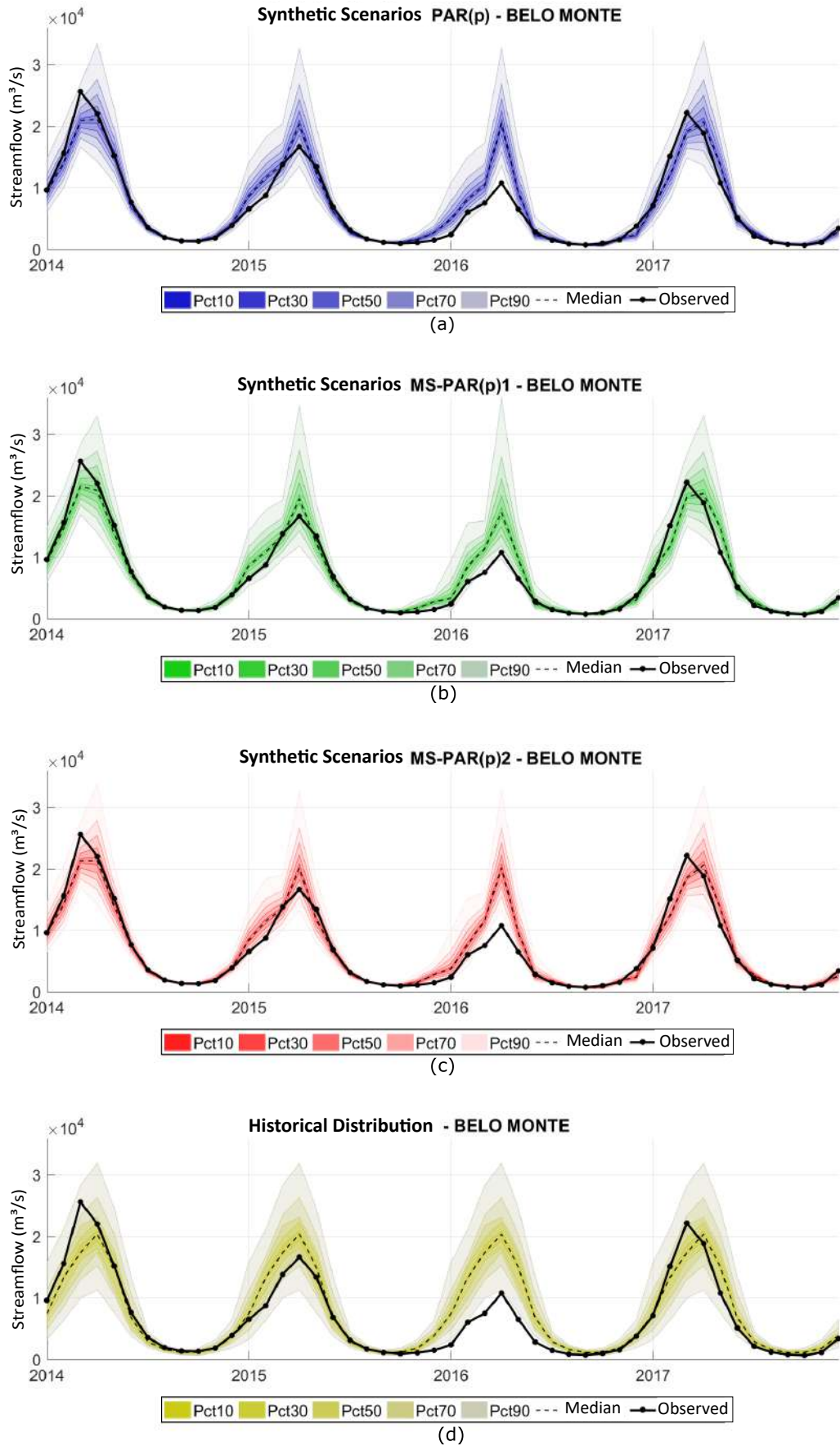


Figure A.48: Synthetic scenarios generated for the period between 2014 and 2017 in the Belo Monte HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).

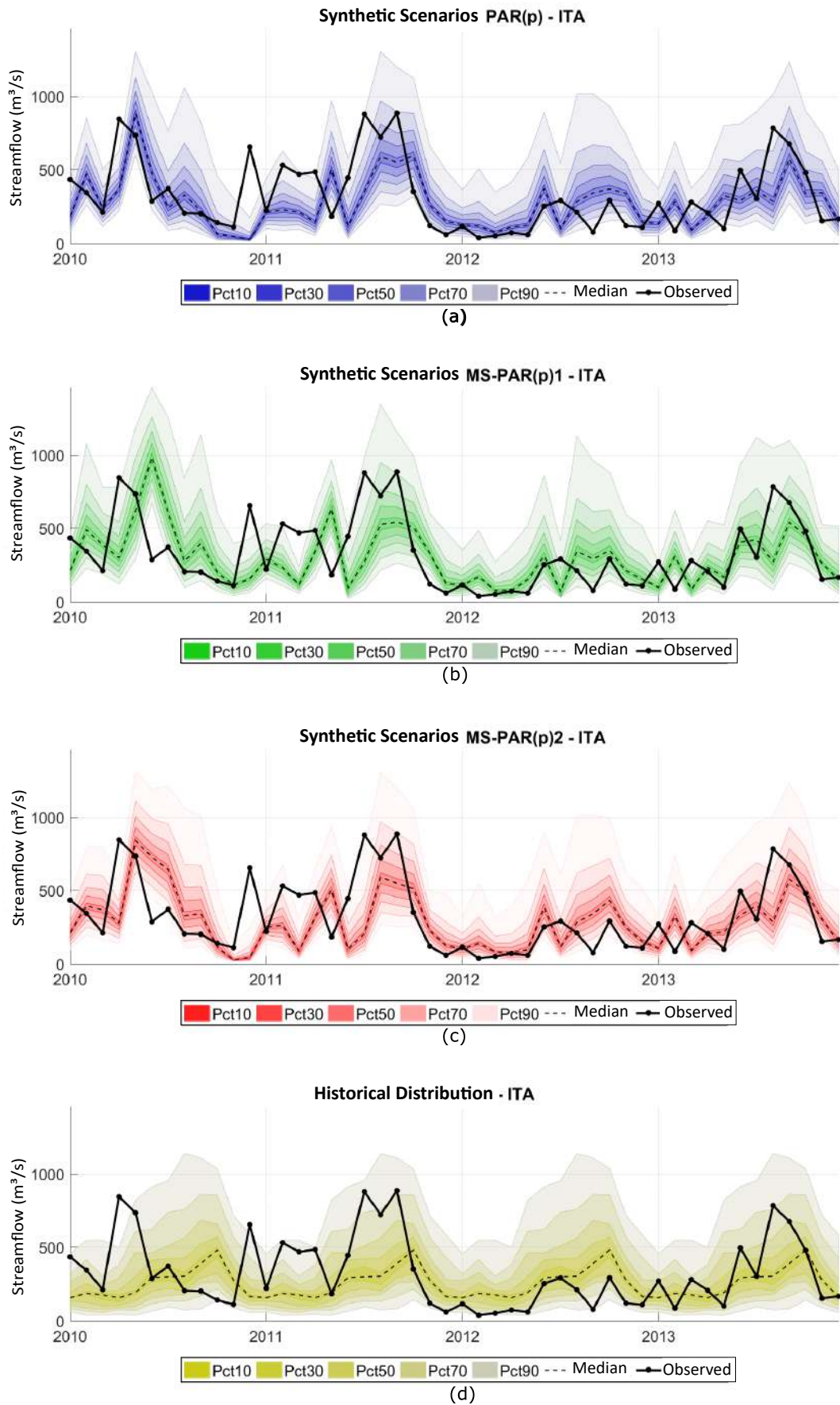


Figure A.49: Synthetic scenarios generated for the period between 2010 and 2013 in the Itá HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).

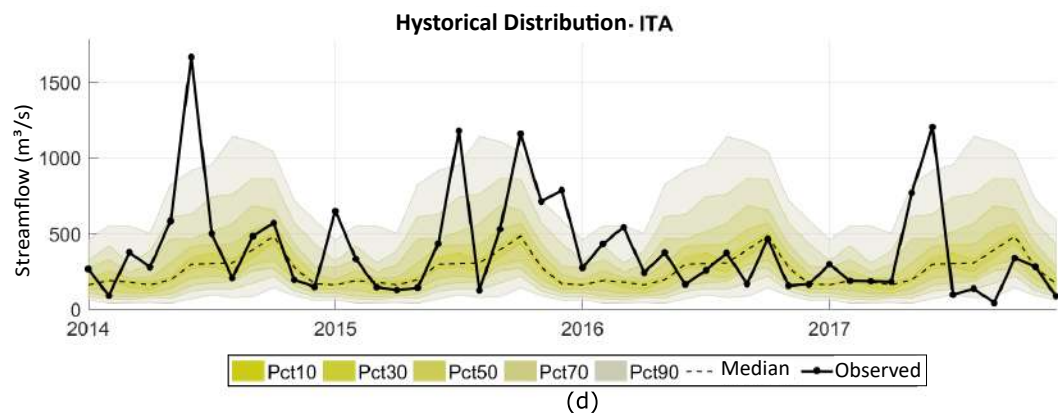
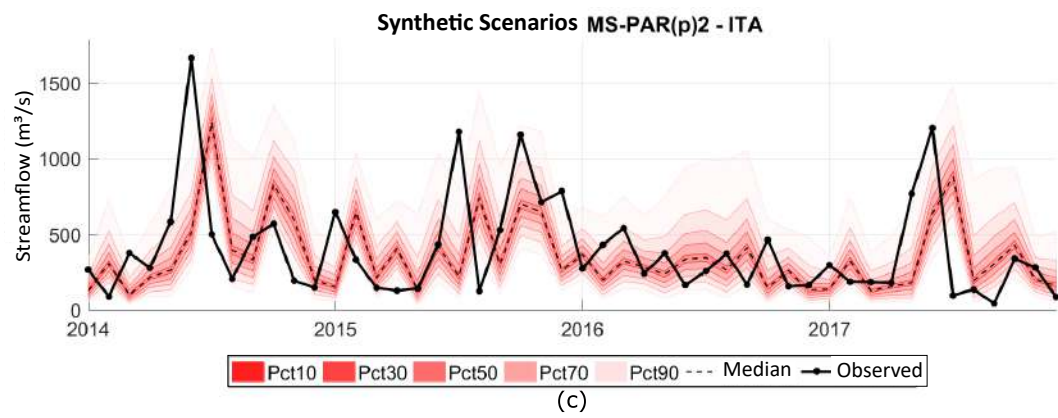
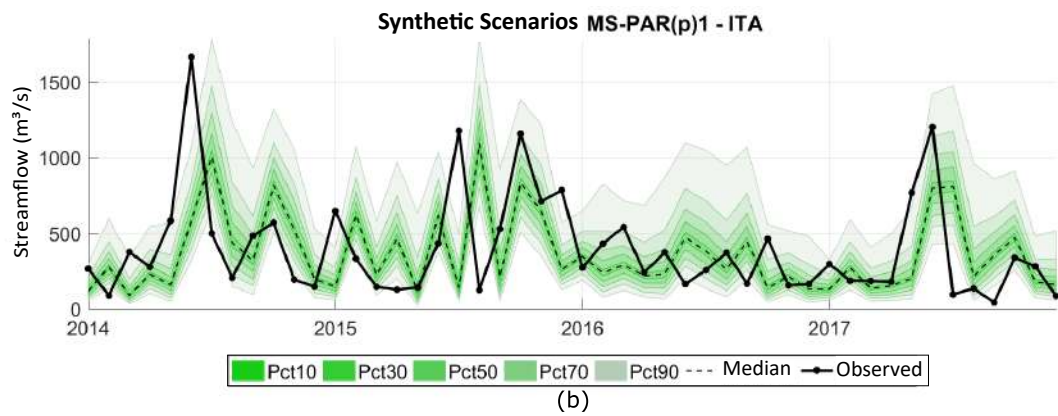
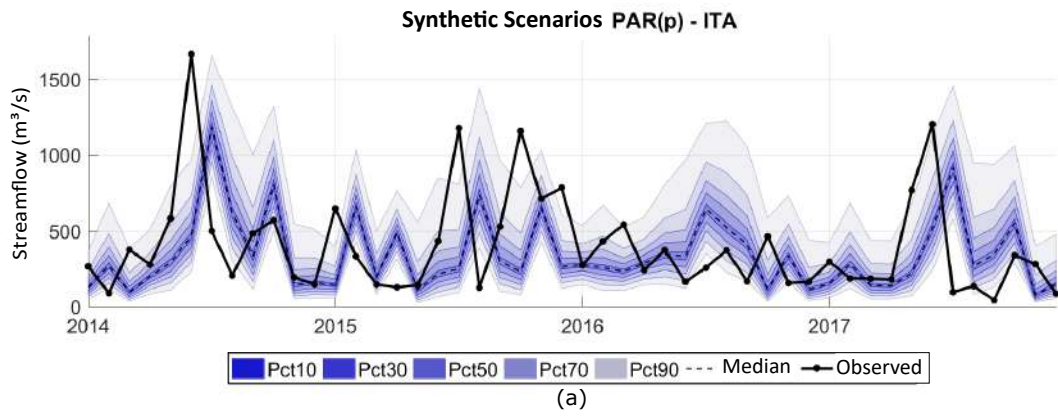


Figure A.50: Synthetic scenarios generated for the period between 2014 and 2017 in the Itá HPP with the PAR(p) (a), MS-PAR(p)1 (b), MS-PAR(p)2 (c) models, and the historical distribution (d).

In order to evaluate the generation for the medium term, an evaluation of the average errors accumulated for six consecutive months will be made, presented in Table A.18. As can be observed, the results are close to those observed in the table A.17, indicating a superiority of the MS-PAR(p) models in the southern and northern regions, and the PAR(p) standing out in the southeast, and in the So-bradinho and Itaipu HPPs. Comparing the results individually for each HPP, the MS-PAR(p)2 model presents a MAE lower than the PAR(p) model in 53% of the HPPs, whereas the MS-PAR(p)1 model only 31%.

In the south region, of a total of 31 HPPs, an improvement occurred in 77.5% and 58% of the HPPs with the MS-PAR(p)2 and MS-PAR(p)1 models respectively, demonstrating once again their superiority in this locality. Also noteworthy were the results obtained for the Paranapanema river basin, whose improvement occurred in 70% of the HPPs with the model MS-PAR(p)2 and 60% the MS-PAR(p)1 of a total of 10 HPPs. In the North, these rates were 50% and 25%, respectively. However, it should be noted that there has been an improvement for the Belo Monte and Jirau HPPs, which are two of the largest electricity producers in Brazil.

To illustrate the analysis of the medium and long-term generation, Figure A.51 presents the three years of scenarios generated in January 2015 with the three models and their historical distribution for the Estreito Tocantins HPP. The year 2015 was characterized by a strong EN, which in turn causes negative anomalies in HPPs located in the north region of Brazil. It is possible to observe that, especially in the first four months of 2015, the scenarios generated with the MS-PAR(p) model are closer to the observed values. Over time, the models end up losing the memory of the recent past and begin to produce scenarios close to the historical distribution.

Table A.18: Evaluation of six months accumulated out-of-sample errors

UHE	Região	MAE			MAPE			RMSE		
		P	MSP1	MSP2	P	MSP1	MSP2	P	MSP1	MSP2
Sobradinho	Northeast	593.15	653.85	636.06	66.36	73.88	71.30	787.46	862.14	835.76
Belo Monte	North	2269.76	2197.11	2226.30	35.16	33.94	34.79	2807.11	2752.03	2755.91
Estreito Toc.	North	521.12	507.02	505.53	35.09	34.39	34.44	650.71	636.12	633.75
Jirau	North	4038.58	4095.93	4024.35	28.95	29.08	28.76	4632.84	4703.44	4618.43
Sto. Ant. Jari	North	404.68	396.62	397.07	43.42	43.06	43.87	480.99	475.81	477.07
Teles Pires	North	529.65	527.53	528.89	23.90	24.11	23.89	634.77	634.98	636.29
Capivara	Southeast	266.28	275.57	273.94	54.47	56.32	56.11	322.46	337.14	331.31
Furnas	Southeast	221.03	239.43	232.65	62.50	72.03	70.16	273.15	297.68	292.48
P. Primavera	Southeast	383.00	408.79	377.03	37.26	38.91	36.41	474.39	519.80	468.20
Itá	South	283.27	281.54	278.98	131.34	132.12	130.84	352.54	352.31	349.79
Itaipu	South	1431.01	1474.48	1502.23	45.60	47.62	47.78	1659.97	1717.33	1740.61
Salto Caxias	South	279.65	275.37	273.75	103.69	104.06	101.55	343.82	343.04	339.49

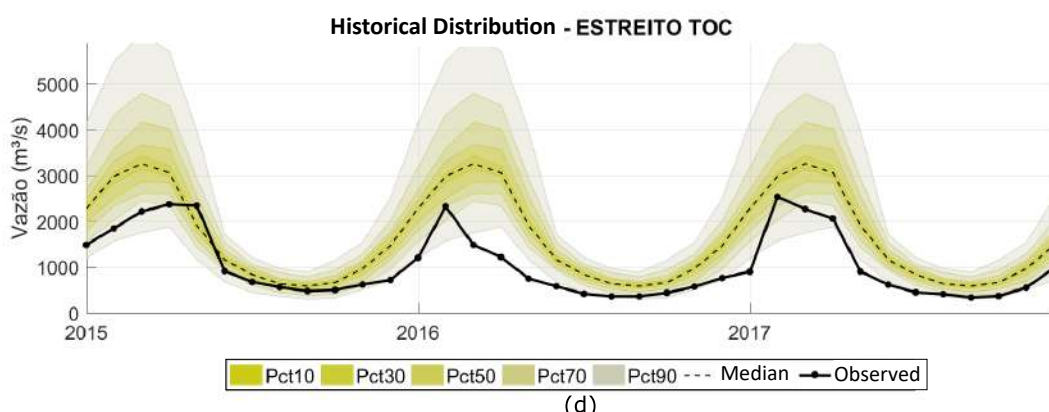
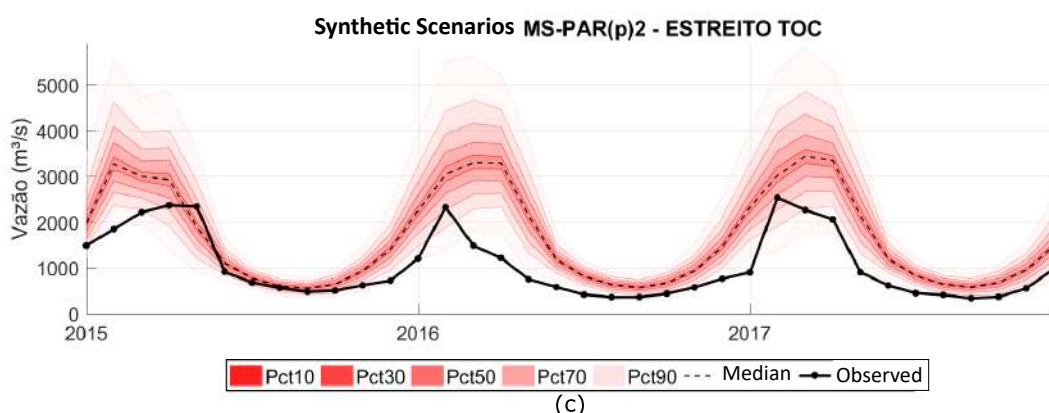
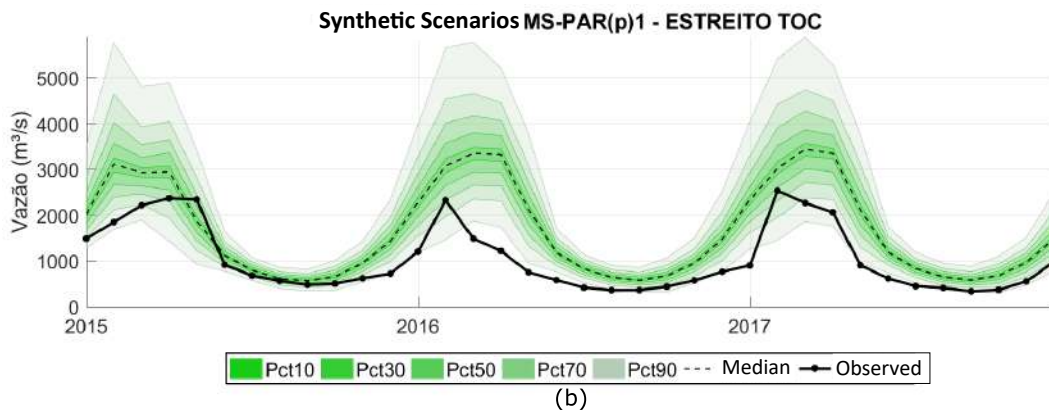
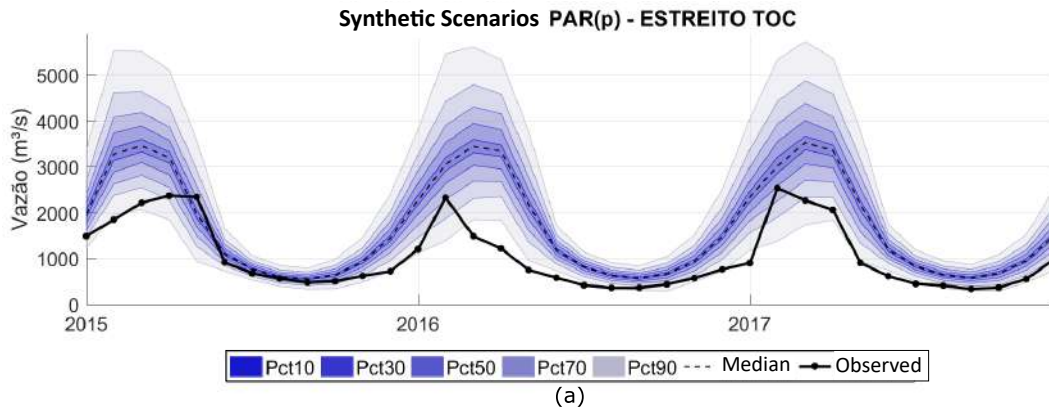


Figure A.51: Synthetic scenarios generated sequentially for the period 2015 to 2017 for the Estreito Tocantins HPP with models PAR(p) (a), MS-PAR(p)1 (b) and MS-PAR(p)2 (c) and its historical distribution (d).

The results presented by the out-of-sample cases indicate some important points that should be highlighted. In the regions most affected by the ENSO phenomenon, the generation of scenarios performed by the MS-PAR(p) models occurs more closely to the observed values, compared to the PAR(p) model. The south-east region, which in several studies cited in this dissertation point to a transition behavior between the north/south dipole of the effects caused by ENSO, do not present a clear pattern of improvement when using the proposed modeling. Another point that should be highlighted concerns the Sobradinho and Itaipu HPPs. Both HPPs have one of the largest NIS drainage areas, and also did not produce better results with the MS-PAR(p) models. The effects along their drainage basins could be diffuse, which could be an explanation their worse results with the MS-PAR(p) models.

The MS-PAR(p) models were able to improve the generation of scenarios both in the short term (one month ahead) and in the medium term analysis (up to six months ahead), especially in the regions previously mentioned. It is important to emphasize that the results in the medium term analysis were even better than the one obtained in the short-term evaluation. It can be stated that one of the advantages of the MS-PAR(p) model is to be able to identify different states in a given month, indicating periods with averages and uncertainties that may be above or below than the historical one, which ends up benefiting the generation of scenarios in the medium term. As the events of EN and LN tend to persist for at least five months, the MS-PAR(p) model profits from this feature in the medium/long term generation.

A.5 Conclusions and Future Improvements

Brazil has a hydrothermal power generation system of continental proportions, in which hydroelectric power is predominant, and thus represents the primary source of uncertainty for the planning of the NIS. Its planning is a complex problem of resource optimization with multiple objectives, which must be solved in real time, but observing the long term. In order to deal with this problem, the energy operation planning is divided into different stages, with different horizons and level of details of the NIS and associated uncertainties. To deal with the main source of uncertainty of the planning of the short, mid and long-term operation, a PAR(p) model is used in the GEVAZP program to generate synthetic series of streamflows. This dissertation aimed to propose a methodological improvement to this model, to include climatic information referring to the El Niño - Southern Oscillation (ENSO) phenomenon.

ENSO is one of the primary natural climatic phenomena and has several im-

pacts on the global circulation patterns of the atmosphere. Specifically, in Brazil, its occurrence impacts of different forms and intensities to the precipitation, and, consequently, the streamflows. Therefore, its consideration in the stochastic modeling of inflows could improve the operation planning of the NIS. In the PAR(p) model, the synthetic series of future inflows are represented as a linear combination of past flows plus a random term. A new stochastic model was presented in this dissertation, the Markov-Switching Periodic Autoregressive - MS-PAR(p). In this model, through the subdivision of each considered period (each month) into distinct states, a segmented estimation of the parameters used in the PAR(p) model is done. For this dissertation, in the GEVAZP model, the states are representations of the ENSO states (La Niña, Neutral and El Niño - LN, N, and EN). Thus, the mean and standard deviation in addition to varying with the month, also have different values within the same month according to the state of ENSO. The evolution of the ENSO's state is modeled through a Markov chain that is adjusted for each month. In order to fit a more parsimonious model, a procedure was proposed to identify the influence of ENSO on the streamflows through the application of the Kolmogorov-Smirnov (KS) adherence test.

In this dissertation, it was proposed to use these probabilistic forecasts for the generation of ENSO scenarios. Given the historical transition matrices, previous values of ONI, and the probabilistic predictions from IRI, it is possible to obtain, through the solution of a linear programming problem, predicted transition matrices similar to historical transition matrices, that also reproduces the frequency distribution of probabilistic forecasts.

In order to evaluate the proposed methodology, the three methodologies described were compared: the traditional PAR(p) model, the MS-PAR(p) model without (MS-PAR(p)1) and with previous identification of ENSO influence (MS-PAR(p)2). The NIS configuration of January 2018, having 141 UHEs, was used in these evaluations. First, a long simulation unconditioned to the recent past was performed to observe the reproduction of historical statistics and negative sequences. Both of these models can reproduce the most statistics, however, with a wide advantage of the MS-PAR(p) models.

The performance of the models in real cases was analyzed through several conditioned simulations, both in the in-sample period and in out-of-sample period. For this analysis, the MAE, MAPE and RMSE metrics were used. In the in-sample cases, a clear advantage of the MS-PAR(p)1 model was observed in all statistics, followed by the MS-PAR(p)2 model and the PAR(p) last. The south region presented the largest reduction of errors with the MS-PAR(p) models. It is noteworthy that during the monthly analysis, it was possible to observe that the improvements were higher during the wet period. Another relevant point is related to the periods

in which the MS-PAR(p) models excelled. During EN events, the north region has the smallest inflows, and it is also when the model presents the biggest difference between the MS-PAR(p) and the PAR(p) models.

In the out-of-sample conditioned cases, it was first evaluated if the ENSO states scenarios, demonstrating that they have good quality even for longer horizons. In the analysis of the errors obtained in the streamflow scenarios one month ahead, on average, the PAR(p) model obtained the best performance, followed by the MS-PAR(p)2 model and finally the MS-PAR(p)1 model. This result is precisely the opposite of that observed in the in-sample period. By making a more segmented analysis, it can be observed that the MS-PAR(p)2 model obtained better results in the north and south regions and the HPPs located in the Paranapanema river basin. It is interesting to note that for the Belo Monte HPP, the MS-PAR(p)1 model obtained the best results because it can reproduce better periods of low streamflows. During the ENSO influence identification process, no influential states was reported during April, which was shown to be inadequate given the subsequent results. The southeast region, in general, obtained the best results with the PAR(p) model.

It can be stated that the objective of this dissertation was successfully achieved, and might represents a possible improvement for the energy operation planning. Even though an application to stochastic hydrology has been made, the MS-PAR(p) model can be used for any time series, such as time series related to the economy or industrial production, as well as the family of models developed by HAMILTON (1989). The representation of the state space in this dissertation was made explicitly, being considered an observed variable, the ENSO time series. Other applications made with the MS-PAR(p) model could represent the states through hidden states.

The identification process of ENSO's influence despite leaving the adjusted models more parsimonious, sometimes obtained counterintuitive results or failed to include the information for some HPPs when it should. Enhancements can be made to this issue through the development of other statistical tests, or even with the inclusion of human judgment by experts.

The comparison between models of synthetic scenario generation should be performed with a different approach to the analysis of forecast models, where it is aimed to reach closer to the observed value. The evaluation performed in the conditioned cases serve as an indication of the potential of the model, demonstrating some possible advantages. For example, the identification of periods that are more uncertain than others, and the generation of more adherent scenarios in periods that are drier than normal, which are precisely the moments in which the NIS is more fragile, can represent a step forward in operation planning. However, this kind of improvement can only be verified through the application of the proposed modeling

in the operation planning model. Therefore, as future development, it is suggested the inclusion and testing of the proposed methodology in energy operation planning models.