

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MODELAGEM DO POTENCIAL DE GERAÇÃO DE ENERGIA HIDRELÉTRICA: UMA
CONTRIBUIÇÃO PARA O PLANEJAMENTO ENERGÉTICO NA AMAZÔNIA

MÁRCIO NIRLANDO GOMES LOPES

TD 08/2019

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém – Pará – Brasil
2019

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MÁRCIO NIRLANDO GOMES LOPES

MODELAGEM DO POTENCIAL DE GERAÇÃO DE ENERGIA HIDRELÉTRICA: UMA
CONTRIBUIÇÃO PARA O PLANEJAMENTO ENERGÉTICO NA AMAZÔNIA

TD 08/2019

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém – Pará – Brasil
2019

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MÁRCIO NIRLANDO GOMES LOPES

MODELAGEM DO POTENCIAL DE GERAÇÃO DE ENERGIA HIDRELÉTRICA: UMA
CONTRIBUIÇÃO PARA O PLANEJAMENTO ENERGÉTICO NA AMAZÔNIA

Tese submetida à Banca
Examinadora do Programa de Pós-
Graduação em Engenharia Elétrica
da UFPA para obtenção do grau de
Doutor em Engenharia Elétrica na
área de Sistemas de Energia Elétrica

UFPA/ITEC/PPGEE
Campus Universitário do Guamá
Belém – Pará – Brasil
2019

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Biblioteca do ITEC/UFPA-Belém-PA

L864 Lopes, Márcio Nirlando Gomes, 1972-
Modelagem do potencial de geração de energia hidrelétrica: uma contribuição para o planejamento energético na Amazônia/ Márcio Nirlando Gomes Lopes. - 2019.

Orientadora: Brígida Ramati Pereira da Rocha

Tese (Doutorado) – Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Belém, 2019.

1. Energia hidrelétrica-Simulação por computador. 2. Inteligência artificial. 3. Climatologia. 4. Usinas hidrelétricas – aspectos econômicos. I. Título.

CDD 23. ed. – 6 2 1 . 3 1 2 1 3 4

Elaborado por Kelren Cecília dos Santos Lima da Mota – CRB-2/1461

UNIVERSIDADE FEDERAL DO PARÁ
 INSTITUTO DE TECNOLOGIA
 PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**“MODELAGEM DO POTENCIAL DE GERAÇÃO DE ENERGIA HIDRELÉTRICA: UMA
 CONTRIBUIÇÃO PARA O PLANEJAMENTO ENERGÉTICO NA AMAZÔNIA”**

AUTOR: MARCIO NIRLANDO GOMES LOPES


TESE DE DOUTORADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO
 COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO
 JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE DOUTOR EM ENGENHARIA
 ELÉTRICA NA ÁREA DE SISTEMAS DE ENERGIA ELÉTRICA.

APROVADA EM: 08/05/2019

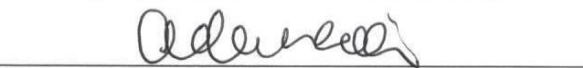
BANCA EXAMINADORA:




Prof.ª Dr.ª Brigida Ramati Pereira da Rocha
 (Orientadora – PPGEE/UFPA)




Prof. Dr. Carlos Tavares da Costa Júnior
 (Avaliador Interno – PPGEE/UFPA)




Prof. Dr. Arthur da Costa Almeida
 (Avaliador Externo ao Programa – CCAST/UFPA)




Prof. Dr. Everaldo Barreiros de Souza
 (Avaliador Externo ao Programa – IG/UFPA)



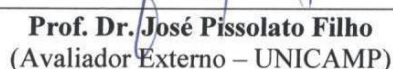
Prof.ª Dr.ª Valquíria Gusmão Macêdo
 (Avaliadora Externa ao Programa – FEEB/UFPA)



Prof.ª Dr.ª Ana Cristina Santos Strava Correa
 (Avaliadora Externa – CENSIPAM)



Prof. Dr. José Alberto Silva de Sá
 (Avaliador Externo – UEPA)



Prof. Dr. José Pissolato Filho
 (Avaliador Externo – UNICAMP)

VISTO:

Prof.ª Dr.ª Maria Emília de Lima Tostes
 (Coordenadora do PPGEE/ITEC/UFPA)

DEDICATÓRIA

*À minha esposa, Tatiana Rescinho Amador Lopes,
e aos meus filhos Vinícius Amador Lopes
e Thiago Amador Lopes.*

AGRADECIMENTOS

À Prof^a. Brígida Rocha pelo incentivo, pelos ensinamentos, pela paciência, pela orientação e, sobretudo, pela amizade sincera.

Aos Professores membros da banca, Valquíria Macedo, Ana Cristina Correa, Everaldo de Souza, Carlos Tavares, José Pissolato, José Alberto Sá e Arthur Almeida, pelas relevantes contribuições que levaram ao aprimoramento deste trabalho.

Aos amigos Alen Vieira, Pedro Rolim e Arilson Silva pela atenção dedicada e apoio irrestrito para elaboração desta pesquisa.

À UFPA e ao PPGEE, em especial ao corpo docente, os quais contribuíram para minha formação e aprendizado e abriram o meu horizonte para novos desafios e perspectivas.

Ao CENSIPAM pelas oportunidades profissionais que me proporcionaram a experiência e o conhecimento necessários para perseguir o doutorado.

A todos o meu humilde e sincero obrigado.

“Não basta conquistar a sabedoria, é preciso usá-la.”

Marco Túlio Cícero

SUMÁRIO

1. INTRODUÇÃO	26
1.1. RELEVÂNCIA E JUSTIFICATIVA DA TESE.....	26
1.2. LOCALIZAÇÃO DO SÍTIO DE ESTUDO.....	31
1.3. PRINCÍPIO FÍSICO TEÓRICO DA TESE	35
1.4. DEFINIÇÃO DE OBJETIVOS	36
1.4.1. Objetivo Geral	36
1.4.2. Objetivos Específicos.....	36
1.5. ESTRUTURA DA TESE	36
2. CONCEITOS BÁSICOS	38
2.1. CONCEITOS BÁSICOS SOBRE ENERGIA	38
2.1.1. Hidroeletricidade	40
2.1.2. Potencial de Geração de Energia Hidrelétrica.....	45
2.2. CONCEITOS BÁSICOS SOBRE HIDROLOGIA	46
2.2.1. Medição de Vazão em Cursos D'água	49
2.2.2. Determinação da Precipitação Média em Bacias Hidrográficas.....	53
2.3. CONCEITOS BÁSICOS SOBRE CLIMATOLOGIA	56
2.3.1. Variabilidade Climática	57
2.3.2. Mudanças Climáticas	58
2.3.2.1. <i>Mudanças Climáticas e Impactos na Amazônia</i>	60
2.3.3. Tratamento Estatístico de Dados Climatológicos.....	62
2.4. CONCEITOS BÁSICOS SOBRE INTELIGÊNCIA COMPUTACIONAL	64
2.4.1. Tópicos sobre <i>Machine Learning</i>.....	65
2.4.1.1. <i>Redes Neurais Artificiais (RNA)</i>	65
2.4.1.1.1. <i>Multilayer Perceptron (MPL) / Algoritmo Back-propagation (BPA)...</i>	69
2.4.1.1.2. <i>Algoritmo de Levenberg-Marquardt.....</i>	71
2.4.1.1.3. <i>Algoritmo de Regulação Bayesiana.....</i>	72
2.4.1.2. <i>Deep Learning e Trabalhos Relacionados.....</i>	76
2.4.1.2.1. <i>Group Method of Data Handling (GMDH).....</i>	78
3. DADOS E METODOLOGIA.....	82
3.1. DADOS PRIMÁRIOS DE RESUMO DE DESCARGA.....	82
3.2. DADOS PRIMÁRIOS DE NÍVEL FLUVIOMÉTRICO.....	82

3.3.	DADOS PRIMÁRIOS DE PRECIPITAÇÃO OBSERVADA	83
3.4.	DADOS PRIMÁRIOS DE PRECIPITAÇÃO SIMULADA	83
3.5.	FLUXO DE PROCESSOS	83
3.6.	SÉRIE TEMPORAL DE VAZÃO	84
3.7.	PRECIPITAÇÃO MÉDIA POR SUB-BACIA	85
3.8.	ANÁLISE DE TENDÊNCIA	86
3.9.	SÉRIE DE POTENCIAL DE GERAÇÃO DE ENERGIA	86
3.10.	MODELAGEM	87
	3.10.1. Previsão do Potencial de Geração de Energia.....	87
	3.10.1.1. Modelagem usando GMDH:	88
	3.10.1.2. Modelagem usando RNA.....	89
	3.10.1.3. Métricas de Avaliação dos Modelos	89
	3.10.1.3.1. Coeficiente de Correlação (R).....	89
	3.10.1.3.2. Erro Médio Absoluto (MAE).....	90
	3.10.1.3.3. Erro Médio Absoluto Percentual (MAPE).....	90
3.11.	CENÁRIOS DE PRECIPITAÇÃO	90
	3.11.1. Cenário 1: RCP4.5	90
	3.11.2. Cenário 2: Tendência de Precipitação.....	91
4.	RESULTADOS.....	93
4.1.	CURVA-CHAVE	93
4.2.	CLIMATOLOGIA	96
	4.2.1. Climatologia da Precipitação.....	96
	4.2.2. Análise Exploratória da Precipitação	98
	4.2.3. Climatologia do Potencial de Geração de Energia	111
	4.2.4. Análise Estatística de Tendência	112
4.3.	MODELAGEM DO POTENCIAL DE GERAÇÃO DE ENERGIA.....	126
4.4.	CENÁRIOS FUTUROS	136
	4.4.1. Cenário 1: RCP4.5	137
	4.4.2. Cenário 2: Tendência de Precipitação	144
5.	ANÁLISE DOS RESULTADOS.....	145
5.1.	RELAÇÃO DA PRECIPITAÇÃO COM O POTENCIAL DE GERAÇÃO DE ENERGIA.....	145

5.2. MODELAGEM PREDITIVA DO POTENCIAL DE GERAÇÃO DE ENERGIA	147
6. CONSIDERAÇÕES FINAIS	149
6.1. PROPOSIÇÃO DE TRABALHOS FUTUROS.....	151
REFERÊNCIAS	152

LISTA DE ILUSTRAÇÕES

Figura 1 - Consumo residencial anual de eletricidade: (a) no estado do Pará, (b) na região Norte, e (c) no Brasil. (d) Aumento percentual acumulado do consumo de energia com base no ano de 2008.	26
Figura 2 - Bacia do rio Tapajós. (a) Localização dos AHE do projeto CHT. (b) Estações hidrológicas na bacia do rio Tapajós com respectiva duração das séries temporais.	29
Figura 3 - Classes de cobertura do solo para o ano base 2018.	33
Figura 4 - Área de estudo: bacia do rio Tapajós.	34
Figura 5 - Oferta (%) de energia elétrica por fonte.	39
Figura 6 - Esquema de uma usina hidrelétrica.	40
Figura 7 - Usina hidrelétrica com reservatório no município de Tucuruí-PA.	42
Figura 8 - Usina hidrelétrica do tipo fio d'água no município de Estreito-MA.	44
Figura 9 - Ciclo hidrológico global. Os números representam uma estimativa da quantidade total de água (milhares de km ³) em cada processo por ano. E é a evaporação, P é a precipitação, Q_G é o escoamento subterrâneo e Q é o escoamento superficial.	47
Figura 10 - Processos do ciclo hidrológico operando na escala de bacia hidrográfica. Q é o <i>runoff</i> , Q_G é o escoamento subterrâneo, Q_{TF} é o escoamento sub-superficial em solo não saturado, I é a interceptação, E é a evaporação e P é a precipitação.	48
Figura 11 - Lances de régua linimétrica no rio Xingu, Altamira-PA.	51
Figura 12 - Molinete para medição de velocidade da água.	52
Figura 13 - Perfil de velocidade em uma seção de rio.	52
Figura 14 - Traçado da mediatriz para definição dos polígonos de Thiessen.	55
Figura 15 - Neurônio biológico.	66
Figura 16 - Rede <i>feedforward</i> com uma única camada de neurônios.	67

Figura 17 - Rede recorrente sem <i>loops</i> de auto- <i>feedback</i> e sem neurônios escondidos.....	68
Figura 18 - Rede recorrente com neurônios escondidos.....	69
Figura 19 - Esquema MLP com BPA.....	70
Figura 20 - Arquitetura geral para uma rede polinomial do tipo GMDH.....	80
Figura 21 - Processo para previsão do potencial de geração de energia hidrelétrica.	84
Figura 22 - Cálculo da área para cada ponto de grade (em km ²).....	85
Figura 23 - Diagrama de dispersão Cota × Vazão para a seção do rio Tapajós em na estação Jatobá.....	93
Figura 24 - Curvas-chave para a seção do rio Tapajós na estação Jatobá usando as seguintes equações: (a) exponencial, (b) potência, (c) polinômio de ordem 2 e (d) polinômio de ordem 3.	95
Figura 25 - Climatologia mensal da precipitação média na (a) sub-bacia Jat, (b) sub-bacia 444, (c) sub-bacia 445, (d) sub-bacia 446. Período base 1984-2018.	96
Figura 26 - Climatologia mensal da precipitação média na (e) sub-bacia 447, (f) sub-bacia 448 e (g) sub-bacia 449 para o período base 1984-2018. (h) Distribuição anual da precipitação por sub-bacia.....	97
Figura 27 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia Jat.	103
Figura 28 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 444.	104
Figura 29 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 445.	105
Figura 30 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 446.	106
Figura 31 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 447.	107

Figura 32 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 448.	109
Figura 33 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 449.	110
Figura 34 - Distribuição mensal do potencial teórico de geração de energia hidrelétrica para o AHE Jatobá no período base 1984-2013.	111
Figura 35 - Curva de permanência para o potencial de geração mensal de energia hidrelétrica no AHE Jatobá no período base 1984-2013. A linha vermelha destaca o valor de geração no período crítico (1.282 MW), conforme projeto.	112
Figura 36 - Série mensal de precipitação na sub-bacia Jat (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.	113
Figura 37 - Série mensal de precipitação na sub-bacia 444 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.	113
Figura 38 - Série mensal de precipitação na sub-bacia 445 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.	114
Figura 39 - Série mensal de precipitação na sub-bacia 446 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.	115
Figura 40 - Série mensal de precipitação na sub-bacia 447 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.	116
Figura 41 - Série mensal de precipitação na sub-bacia 448 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.	117

Figura 42 - Série mensal de precipitação na sub-bacia 449 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.	118
Figura 43 - Série anual de precipitação por sub-bacia (período base-1979-2018). A reta tracejada representa a tendência linear da série. (a) Sub-bacia Jat, (b) sub-bacia 444, (c) sub-bacia 445, (d) sub-bacia 446, (e) sub-bacia 447, (f) sub-bacia 448 e (g) sub-bacia 449.	125
Figura 44 - Desempenho dos principais modelos com base no coeficiente de correlação (R).	128
Figura 45 - Desempenho dos principais modelos com base no MAE	128
Figura 46 - Desempenho dos principais modelos com base no $MAPE$	129
Figura 47 - Resultados de desempenho durante a fase de aprendizado para RNA-RB e RNA-LM sobre 100 repetições de simulação para cada modelo. As linhas horizontais tracejada e contínua dentro da na caixa representam a média e mediana, respectivamente.	130
Figura 48 - Comparativo entre os valores observados (CPC) e previstos pelos modelos (GMDH, RNA-LM e RNA-RB) durante o período de treinamento/validação. As áreas sombreadas nas curvas inferior e do centro representam as 100 curvas de repetição dos modelos com inicialização randômica dos pesos para cada uma das rodadas.	131
Figura 49 - Distribuição do erro relativo por faixa de energia para o conjunto de dados de aprendizagem. (a) GMDH, (b) RNA-RB e (c) RNA-LM.	132
Figura 50 - Resultados de desempenho durante a fase de teste para RNA-RB e RNA-LM sobre 100 repetições de simulação para cada modelo. As linhas horizontais tracejada e contínua dentro da na caixa representam a média e mediana, respectivamente.	134
Figura 51 - Comparativo entre os valores observados (CPC) e previstos (GMDH, RNA-LM e RNA-RB) durante o período de teste. As áreas sombreadas nas curvas inferior e do centro representam as 100 curvas de repetição dos modelos com inicialização randômica dos pesos	135
Figura 52 - Distribuição do erro relativo por faixa de energia para o subconjunto de dados de teste. (a) GMDH, (b) RNA-RB e (c) RNA-LM.	136

Figura 53 - Erro absoluto do modelo climático RegCM4 para simulação da precipitação média mensal por sub-bacia, considerando o período base 1981-2005. (a) Sub-bacia Jat e (b) sub-bacia 444.	137
Figura 54 - Erro absoluto do modelo climático RegCM4 para simulação da precipitação média mensal por sub-bacia, considerando o período base 1981-2005. (c) sub-bacia 445, (d) sub-bacia 446 e (e) sub-bacia 447.	138
Figura 55 - Erro absoluto do modelo climático RegCM4 para simulação da precipitação média mensal por sub-bacia, considerando o período base 1981-2005. (f) sub-bacia 448 e (g) sub-bacia 449.	139
Figura 56 - Diagrama de dispersão entre os dados observados do CPC e a simulação usando o modelo climático RegCM4, por sub-bacia considerando o período base 1981-2005. (a) Sub-bacia Jat, (b) sub-bacia 444, (c) sub-bacia 445, (d) sub-bacia 446, (e) sub-bacia 447, (f) sub-bacia 448 e (g) sub-bacia 449.	140
Figura 57 - Comparativo entre os dados observados do CPC e a simulação usando o modelo climático RegCM4, para a climatologia da precipitação por sub-bacia considerando o período base 1981-2005. (a) Sub-bacia Jat, (b) sub-bacia 444, (c) sub-bacia 445 e (d) sub-bacia 446.	141
Figura 58 - Comparativo entre os dados observados do CPC e a simulação usando o modelo climático RegCM4, para a climatologia da precipitação por sub-bacia considerando o período base 1981-2005. (e) Sub-bacia 447, (f) sub-bacia 448 e (g) sub-bacia 449.	142
Figura 59 - Curva de permanência para o potencial de geração mensal de energia hidrelétrica sob cenário 1 (RCP4.5) para o período base 2019-2038. A linha vermelha destaca o valor de geração no período crítico (1.282 MW), conforme projeto.	143
Figura 60 - Curva de permanência para o potencial de geração mensal de energia hidrelétrica sob cenário 2 (tendência de precipitação) para o período base 2019-2038. A linha vermelha destaca o valor de geração no período crítico (1.282 MW), conforme projeto.	144
Figura 61 - Média mensal do potencial de geração de energia (vermelho) e da precipitação média na sub-bacia Jatobá.	145

Figura 62 - Média mensal do potencial de geração de energia (vermelho) e da precipitação média por sub-bacia..... 146

Figura 63 - Curvas de permanência para o potencial de geração de energia para os dados observados e para os cenários previstos. 147

LISTA DE TABELAS

Tabela 1 - Potencial hidrelétrico brasileiro em cada estágio por bacia hidrográfica (valores em MW).	28
Tabela 2 - Tendência linear calculada da precipitação média na sub-bacia Jat.	92
Tabela 3 - Série temporal mensal de precipitação média simulada pelo modelo climático RegCM4.	92
Tabela 4 - Série sintética da precipitação média corrigida com dados de tendência estatística.	92
Tabela 5 - Dados de resumo de descarga para a estação Jatobá.....	94
Tabela 6 - Equações para determinação da melhor curva de ajuste aos dados de cota \times vazão na estação Jatobá.	95
Tabela 7 - Estatística descritiva da precipitação na sub-bacia Jat.	98
Tabela 8 - Estatística descritiva da precipitação na sub-bacia 444.....	98
Tabela 9 - Estatística descritiva da precipitação na sub-bacia 445.....	99
Tabela 10 - Estatística descritiva da precipitação na sub-bacia 446.....	99
Tabela 11 - Estatística descritiva da precipitação na sub-bacia 447.....	99
Tabela 12 - Estatística descritiva da precipitação na sub-bacia 448.....	100
Tabela 13 - Estatística descritiva da precipitação na sub-bacia 449.....	100
Tabela 14 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia Jat, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$	101
Tabela 15 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 444, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$	101

Tabela 16 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 445, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.	101
Tabela 17 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 446, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.	101
Tabela 18 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 447, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.	102
Tabela 19 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 448, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.	102
Tabela 20 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 449, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.	102
Tabela 21 - Estatística descritiva do potencial mensal para geração de energia no AHE Jatobá.	111
Tabela 22 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia Jat no rio Tapajós.	121
Tabela 23 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 444 no rio Tapajós.	121
Tabela 24 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 445 no rio Tapajós.	122
Tabela 25 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 446 no rio Tapajós. Fonte: autor.	122
Tabela 26 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 447 no rio Tapajós.	123

Tabela 27 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 448 no rio Tapajós.	123
Tabela 28 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 449 no rio Tapajós.	124
Tabela 29 - Quadro resumo da tendência estatística observada para a precipitação média mensal nas sub-bacias do rio Tapajós.	124
Tabela 30 - Estatística do teste de Mann-Kendall para a precipitação média anual.	126
Tabela 31 - Configurações dos hiperparâmetros do modelo usando método GMDH.	126
Tabela 32 - Configurações dos hiperparâmetros do modelo usando RNA.	127
Tabela 33 - Resultados para a avaliação dos modelos durante a fase de treinamento/validação. DP refere-se ao desvio padrão da métrica de avaliação no índice.	130
Tabela 34 - Resultados para avaliação dos modelos no período de teste. DP refere-se ao desvio padrão da métrica de avaliação no índice.	133
Tabela 35 - Teste Mann-Whitney para comparação de duas amostras sobre os dados de precipitação, considerando como hipótese nula H_0 que a diferença de posição entre as amostras é igual a zero, para $\alpha = 0,05$	142

LISTA DE ABREVIATURAS

ADCP	<i>Acoustic Doppler Current Profiler</i>
AHE	Aproveitamento Hidrelétrico
ANA	Agência Nacional de Águas
ANEEL	Agência Nacional de Energia Elétrica
AR5	<i>Fifth Assessment Report</i>
BEN	Balanco Energético Nacional
BPA	<i>Back Propagation Algorithm</i>
CCEE	Câmara de Comercialização de Energia Elétrica
CHT	Complexo Hidrelétrico do Tapajós
CMIP5	<i>Coupled Model Intercomparison Project</i>
CNPE	Conselho Nacional de Política Energética
CNRH	Conselho Nacional de Recursos Hídricos
CPC	<i>Climate Prediction Center</i>
CPU	<i>Central Processing Unit</i>
DP	<i>Deep Learning</i>
ELETROBRAS	Centrais Elétricas Brasileiras
EPE	Empresa de Pesquisa Energética
GEE	Gases de Efeito Estufa
GTE	Grupo de Estudos Tapajós
GMDH	<i>Group Method of Data Handling</i>
GPU	<i>Graphics Processing Unit</i>
GRU	<i>Gated Recurrent Unit</i>
HadGEM2	<i>Hadley Centre's Global Environmental Model version 2</i>
IA	Inteligência Artificial
IBAMA	Inst. Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis
ICTP	<i>International Centre for Theoretical Physics</i>
INPE	Instituto Nacional de Pesquisas Espaciais
IPCC	<i>Intergovernmental Panel on Climate Change</i>
IRENA	<i>International Renewable Energy Agency</i>
LSTM	<i>Long-Short Term Memory</i>
MAE	<i>Mean Absolute Error (Erro Médio Absoluto)</i>

MAPE	<i>Mean Absolute Percentage Error</i> (Erro Médio Percentual Absoluto)
MDH	Modelo Hidrológico Distribuído
MIA	<i>Multilayered Iteration Algorithm</i>
MK	Teste de Mann-Kendall
MLP	<i>Multilayer Perceptron</i>
MMA	Ministério do Meio Ambiente
MME	Ministério de Minas e Energia
MSE	<i>Mean Squared Error</i> (Erro Quadrático Médio)
NCAR	<i>National Center for Atmospheric Research</i>
NCEP	<i>National Centers for Environmental Prediction</i>
PBMC	Painel Brasileiro de Mudanças Climáticas
PDE	Plano Decenal de Expansão de Energia
RAN1	Primeiro Relatório de Avaliação Anual
RCP	<i>Representative Concentration Pathways</i>
RegCM	<i>Regional Climate Model</i>
RNA	Redes Neurais Artificiais
RNA-LM	Rede Neural Artificial com algoritmo Levenberg-Marquardt
RNA-RB	Rede Neural Artificial com algoritmo regularização bayesiana
RNC	Redes Neurais Convolucionais
UHE	Usina Hidrelétrica
WCRP	<i>World Climate Research Programme</i>
WG1	<i>Working Group 1</i>
WGCM	<i>Working Group on Coupled Modeling</i>
XGB	<i>Extreme Gradient Booster</i>

RESUMO

Os estudos de viabilidade para aproveitamentos hidrelétricos (AHE) utilizam dados observados de vazão para estimar o potencial de geração de energia. De acordo com o regramento brasileiro, os registros necessitam abranger um período mínimo de setenta anos de observação. No entanto, em algumas regiões os registros de vazão compreendem curtos períodos, de modo que a série temporal não captura a variabilidade natural do clima e isto pode levar a erros na interpretação das informações, de tal forma que a geração de energia da planta pode ser inferior à esperada. Para tratar tal problema, esta tese propõe uma metodologia inovadora para estimar o potencial de geração de energia. O princípio consiste em utilizar dados de precipitação, que possui séries temporais longas, para estimar o potencial de geração de energia através de uma modelagem que emprega técnicas de inteligência artificial. Dois métodos distintos de aprendizado de máquina foram testados. O primeiro é uma técnica de *deep learning* denominada *Group Method of Data Handling* – GMDH. O segundo utiliza redes neurais artificiais com opções distintas para algoritmos de otimização, Levenberg-Marquardt e regularização bayesiana. A metodologia foi aplicada para o aproveitamento hidrelétrico de Jatobá, na bacia do rio Tapajós, no Pará. Os indicadores de avaliação mostraram que os modelos apresentam habilidade para simular a geração de energia, com melhor desempenho para o GMDH que atingiu 95% de correlação com os dados reais e apenas 12% de erro percentual médio. As simulações obtiveram melhor desempenho durante a estação seca, o que é fundamental, visto que este período crítico para a geração de energia hidrelétrica define a energia firme do empreendimento. Análises estatísticas sobre os dados observados detectaram tendência significativa de redução da precipitação em algumas sub-bacias do rio Tapajós. Simulações incorporando um cenário de mudanças climáticas, proposto pelo Painel Intergovernamental sobre Mudança do Clima, assim como um cenário de tendência estatística de longo prazo, ambos indicaram redução da capacidade de produção energética para os próximos vinte anos, sugerindo que o AHE Jatobá poderá não atender a demanda de energia firme do projeto no longo prazo, sob tais condições. O sucesso desta abordagem pode contribuir para reduzir incertezas e subsidiar os estudos preliminares para implementação de usinas hidrelétricas, assim como simular cenários para apoio ao planejamento, reduzir custos e gerar dados sintéticos para séries temporais de geração de energia cobrindo períodos sem dados observacionais de campo.

Palavras-chave: geração de energia, Amazônia, GMDH, RNA, mudanças climáticas.

ABSTRACT

The feasibility studies for hydroelectric power plants (HPP) use observed flow data to estimate the potential for energy generation. According to the Brazilian rule, the records need to cover a minimum period of seventy years of observation. However, in some regions the flow records comprise short periods, so that the time series does not capture the natural variability of the climate and this can lead to errors in the interpretation of the information, such that the power generation of the plant can be lower than expected. To address this problem, this thesis proposes an innovative methodology to estimate the potential of power generation. The principle is to use precipitation data, which has long time series, to estimate the potential of power generation through a modeling that employs artificial intelligence techniques. Two distinct methods of machine learning were tested. The first is a deep learning technique called Group Method Data Handling (GMDH). The second uses artificial neural networks with distinct options for optimization algorithms, Levenberg-Marquardt and Bayesian regularization. The methodology was applied to the Jatobá hydroelectric project in the Tapajós river basin in Pará. The evaluation indicators showed that the models have the skill to simulate the power generation, with a better performance for the GMDH, which reached 95% correlation with the actual data and only 12% average error percentage. The simulations obtained better performance during the dry season, which is fundamental, since this critical period for the generation of hydropower defines the firm energy of the enterprise. Statistical analysis on the observed data detected a significant tendency of precipitation reduction in some sub-basins of the Tapajós River. Simulations incorporating a climate change scenario proposed by the Intergovernmental Panel on Climate Change, as well as a scenario of long-term statistical trend, both indicated a reduction in power production capacity for the next twenty years, suggesting that the HPP Jatobá may not meet the project's firm energy demand for long-term under such conditions. The success of this approach can contribute to reducing uncertainties and subsidizing preliminary studies for the implementation of hydropower plants, as well as simulating scenarios to support planning, reduce costs and generate synthetic data for time series of power generation covering periods without observational field data.

Keywords: power generation, Amazon, GMDH, ANN, climate change.

CAPÍTULO I

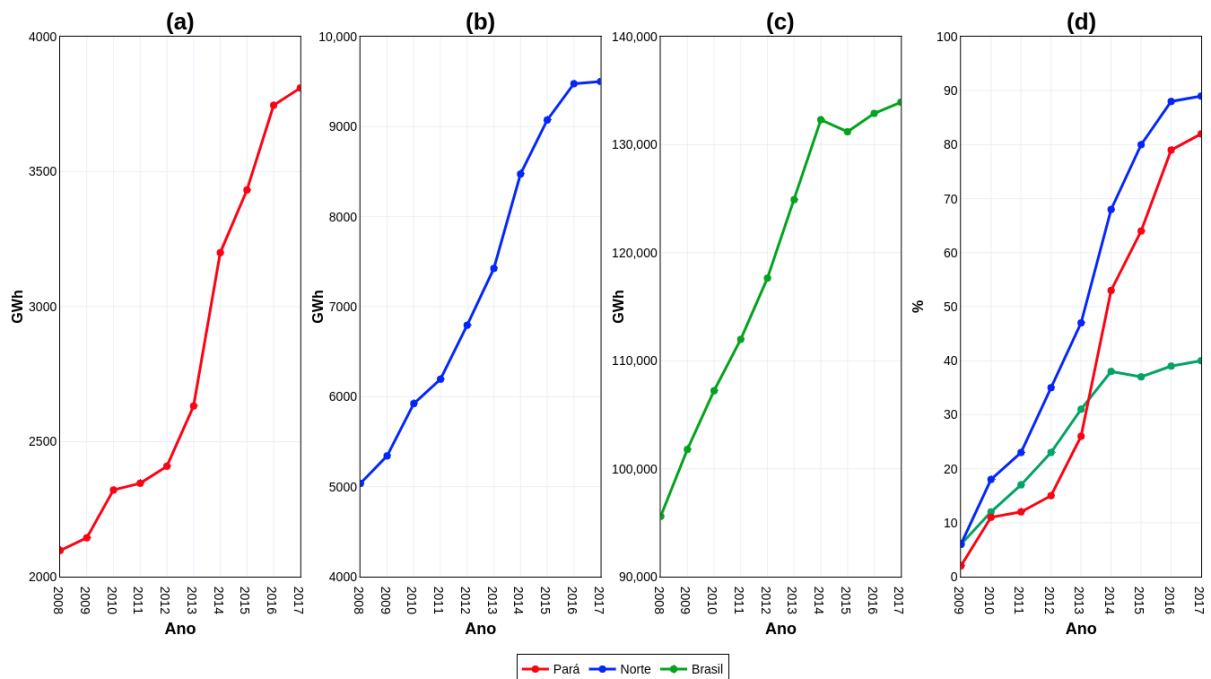
1. INTRODUÇÃO

Esta tese apresenta uma metodologia inovadora para estimar o potencial de geração de energia em aproveitamentos hidrelétricos com foco na Amazônia, através do reconhecimento de padrões de precipitação nas bacias hidrográficas por meio de técnicas de inteligência computacional.

1.1. RELEVÂNCIA E JUSTIFICATIVA DA TESE

A importância deste trabalho é caracterizada através da necessidade de uma metodologia que possa tratar um problema real e complexo, como a modelagem do potencial de geração de energia, utilizando uma quantidade reduzida de dados e alcançando resultados confiáveis estatisticamente, de forma que possam auxiliar tomadores de decisão em escolhas assertivas envolvendo a melhor relação custo/benefício, considerando o potencial de geração energética e os impactos sociais e ambientais associados ao empreendimento.

Figura 1 - Consumo residencial anual de eletricidade: (a) no estado do Pará, (b) na região Norte, e (c) no Brasil. (d) Aumento percentual acumulado do consumo de energia com base no ano de 2008.



Fonte: EPE, 2018.

A energia é seguramente fundamental para o desenvolvimento de qualquer nação. Para manter o seu crescimento, é natural que um país faça uso de seus recursos naturais para geração de energia. Atualmente, há uma tendência mundial e cada vez maior pela utilização de fontes renováveis de energia (IEA, 2018), tais como vento, sol e água. O mundo moderno, por sua vez, rendeu-se às facilidades proporcionadas por uma das formas de energia mais comumente utilizada, a eletricidade. Como exemplo, apenas nos últimos dez anos, o consumo de energia elétrica residencial cresceu em todo o Brasil cerca de 40%, sendo que a região Norte apresentou aumento de 89%, com especial destaque para o estado do Pará, que alcançou 82% de crescimento, conforme exibido na Figura 1.

A crescente demanda por energia elétrica torna a implementação de usinas hidrelétricas (UHE) necessária em muitas situações; entretanto este modelo pode causar impactos ambientais e sociais bastante significativos (FEARNSIDE, 2014, 2006). Na segunda metade do século passado, o Brasil implantou diversas usinas com reservatórios de acumulação. Entretanto, por razões ambientais, atualmente os novos projetos de UHE de grande porte consideram apenas a construção de barragens sem reservatórios ou com reservatórios muito pequenos, de maneira que o armazenamento de água é significativamente reduzido (MME, 2007; SHARMA; SINGH, 2013; TOLMASQUIM, 2016). Este novo tipo de usina é denominado de fio d'água e, em razão de seu pequeno reservatório, a produção de energia é mais suscetível à variabilidade do clima (IRENA, 2015; KOCH et al., 2011). Algumas usinas fio d'água nem mesmo possuem reservatório, e por esta razão tornam-se extremamente dependentes das vazões naturais dos rios, que apresentam comportamento sazonal e, portanto, durante a estação seca sofrem uma redução significativa da capacidade de geração de energia. Esta característica torna imperativa a determinação correta do período crítico, ou seja, aquele momento em que a geração é a mais baixa, o que define a “energia firme” da usina. O termo energia firme é usualmente empregado para estabelecer a máxima capacidade de produção contínua de energia por uma usina hidrelétrica. Tal determinação é de extrema relevância visto que faz parte dos contratos de energia, além de estabelecer penalidades para as geradoras em caso de não atendimento.

Considerando os estudos preliminares de viabilidade para um aproveitamento hidrelétrico, que é uma componente essencial para a implementação de projetos de usinas hidrelétricas (MME, 2007; TOLMASQUIM, 2016), uma abordagem importante deste estudo é a estimativa do potencial de geração de energia. Os principais parâmetros empregados para determinar a quantidade de energia hidrelétrica que um aproveitamento hidrelétrico (AHE) pode gerar são a vazão do rio e a queda hidráulica (KOCH et al., 2011; STICKLER et al., 2013; WALI, 2013). A queda hidráulica é um parâmetro de engenharia dependente do relevo e do projeto da usina,

cujo método oficialmente empregado para avaliar o potencial de um aproveitamento hidrelétrico considera a série temporal de longo prazo de medição de vazão realizada utilizando dados observacionais de estações hidrológicas.

No Brasil, o maior potencial ainda não explorado (estimado e inventariado) de energia hidrelétrica está na região Amazônica (Tabela 1), onde estima-se que mais de 70% do potencial brasileiro para hidroeletricidade seja encontrado na bacia do rio Amazonas e mais de 50% já foi inventariado (ANEEL, 2002; ELETROBRAS, 2017).

Tabela 1 - Potencial hidrelétrico brasileiro em cada estágio por bacia hidrográfica (valores em MW).

Bacia	Estimado	(%)	Inventário	(%)	Viabilidade	Projeto Básico	Construção	Operação	Total Geral
Rio Amazonas	30.595	70,1	38.580	53,0	774	948	794	23.213	94.905
Rio Tocantins	1.875	4,2	8.077	11,1	3.738	120	-	13.250	27.060
Atlântico Norte e Nordeste	707	1,6	871	1,2	466	50	-	812	2.905
Rio São Francisco	1.561	3,6	3.895	5,3	6.140	234	-	10.786	22.616
Atlântico Leste	1.423	3,3	5.796	8,0	665	811	54	5.424	14.172
Rio Paraná	5.107	11,7	9.563	13,1	1.889	2.206	400	43.538	62.704
Rio Uruguai	342	0,8	4.097	5,6	292	432	152	6.406	11.720
Atlântico Sudeste	2.031	4,7	1.857	2,6	2.218	326	5	3.793	10.231
Totais	43.640	100	72.737	100	16.181	5.127	1.405	107.223	246.313

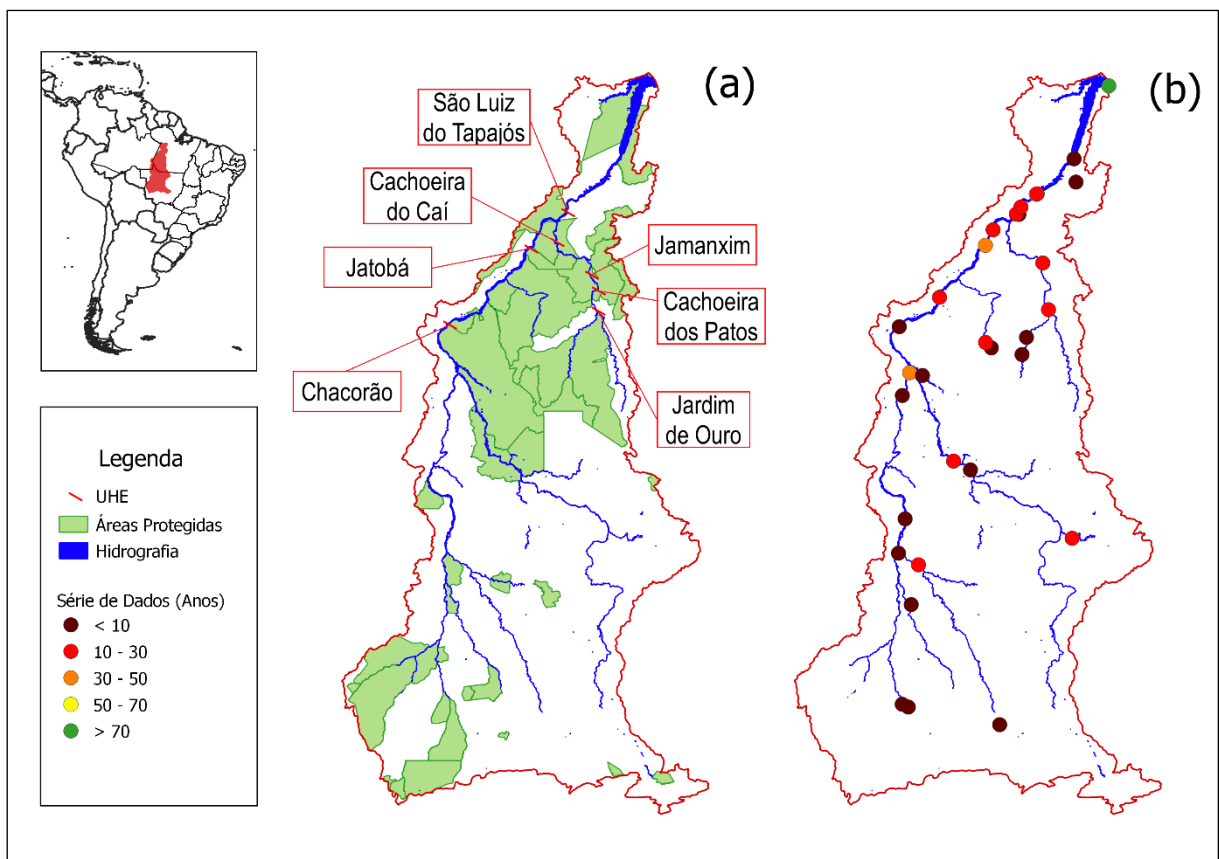
Fonte: Adaptado de ELETROBRAS, 2017 (<https://eletrobras.com/pt/Paginas/Potencial-Hidreletrico-Brasileiro.aspx>).

Para os próximos anos, há várias usinas hidrelétricas planejadas para implementação na região Amazônica. Dentre as sub-bacias do rio Amazonas, a sub-bacia do rio Tapajós é a que possui maior potencial hidrelétrico. Entretanto a maior parte deste potencial possui fortes restrições ambientais, visto que os principais aproveitamentos hidrelétricos estão situados em áreas protegidas, sejam terras indígenas ou unidades de conservação (Figura 2a). Ainda assim, um projeto denominado Complexo Hidrelétrico do Tapajós (CHT) foi elaborado, o qual prevê a instalação de um complexo de usinas.

O projeto CHT considera os seguintes AHE, com seus respectivos potenciais de geração: São Luiz do Tapajós (6.133 MW), Jatobá (2.338 MW), Chacorão (3.336 MW), Cachoeira dos Patos (528 MW), Jamanxim (881 MW) e Cachoeira do Caí (802 MW). O Conselho Nacional de Política Energética (CNPE), através da resolução CNPE nº 3, publicada em maio de 2011, destaca que as futuras usinas São Luiz do Tapajós e Jatobá são estratégicas de interesse público, estruturantes e prioritárias para implantação (MME; EPE, 2014). Dentre essas, a maior usina é

a de São Luiz do Tapajós, que teria potência inferior apenas à Itaipu, Belo Monte e Tucuruí, com previsão de até 6.133 MW de potência total a partir da construção de uma barragem de 3.483 metros de comprimento atravessando áreas protegidas. Os estudos para o AHE São Luiz do Tapajós eram os mais avançados, mas o Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) arquivou o seu processo de licenciamento sob argumento de que o projeto apresentado e o seu respectivo Estudo de Impacto Ambiental (EIA) não continham informações necessárias para a análise de viabilidade socioambiental, o que levou à extrapolação do prazo, conforme previsto na resolução do Conselho Nacional do Meio Ambiente (CONAMA). Jatobá é outro importante AHE, situado no município de Itaituba-PA, e que deverá gerar em média 1.282 MW, com potência instalada de 2.338 MW (MME; EPE, 2014).

Figura 2 - Bacia do rio Tapajós. (a) Localização dos AHE do projeto CHT. (b) Estações hidrológicas na bacia do rio Tapajós com respectiva duração das séries temporais.



Fonte: Elaborado pelo autor do trabalho.

Como forma de minimizar os impactos ambientais na região, a modalidade de exploração do potencial hidroelétrico adotado pelo governo será de usina-plataforma do tipo

fio d'água. As usinas-plataformas adotam esta denominação em referência às plataformas usadas para exploração de petróleo. Isto implica que na fase de construção as instalações levantadas sejam temporárias, diferentemente do modelo tradicional de vilas operárias, que depois transformam-se em cidades. As instalações ficarão, portanto, restritas ao entorno da usina, sem necessariamente ocorrer a abertura de rodovias de acesso, o que segundo a EPE contribuiria para reduzir o potencial de devastação inerente às ocupações do modo usual. Por outro lado, a característica mais limitante das usinas fio d'água é sua vulnerabilidade às flutuações do regime de precipitação.

O regramento brasileiro requer que os estudos prévios para implantação de uma UHE empreguem dados de medição de descarga (vazão) coletados por um período de setenta anos, de maneira que possa capturar com segurança estatística, os períodos secos que definem a energia firme (ANEEL, 2005). Entretanto, é bastante comum em inúmeras ocasiões não haver um longo registro de dados observacionais para os rios da Amazônia. Por exemplo, os dados disponíveis pelas estações hidrológicas na bacia do rio Tapajós apresentam duração das séries temporais de vazão ou cota inferiores a trinta anos, em média (Figura 2b). Esta condição pode ser um fator limitante para a estimativa do potencial de geração de energia hidrelétrica com o grau de confiança necessário. Além disso, também pode comprometer o sucesso do empreendimento que envolve elevadas quantidades de recursos financeiros e cujos impactos sociais e ambientais são extremamente importantes, demandando o máximo rigor para o planejamento do projeto (AWOJOBI; JENKINS, 2016).

Os sistemas de recursos hídricos, por sua vez, são projetados a partir da suposição de que a sequência de dados hidrológicos, de vazão ou de precipitação, seja estacionária. Em outras palavras, supõe-se que as características essenciais dos processos hidrológicos, tais como suas medidas de posição e variabilidade permaneçam constantes ao longo do tempo. Entretanto, tal suposição pode não ser verificada quando são considerados aspectos como a mudança de cobertura vegetal, ou a variabilidade e mudança climática. Nesses casos, projetos de médio e longo prazos baseados na estacionariedade dos eventos hidrológicos devem ser criteriosamente avaliados (ALEXANDRE; BAPTISTA; NAGHETTINI, 2010).

Segundo o *Fifth Assessment Report (AR5) - The Physical Science Basis* (2013), último relatório de avaliação divulgado pelo *Intergovernmental Panel on Climate Change (IPCC)*, é esperado um aumento progressivo da temperatura global de até 4,5°C para o final deste século, o que ocasionaria alterações significativas no regime pluviométrico em todo o globo. Esta afirmação está em consonância com os resultados publicados pelo Primeiro Relatório de Avaliação Anual (RAN1) (PBMC, 2014), do Painel Brasileiro de Mudanças Climáticas

(PBM). Todavia, particularmente no Brasil e, mais especificamente na Amazônia, mesmo com todo o avanço já alcançado, as simulações dos modelos numéricos de previsão climática não apresentam consenso absoluto, sugerindo diferentes resultados que apontam tanto para déficits acentuados de chuva, como para aumento no volume de precipitação (MARENGO, 2007). Muitos destes resultados contraditórios estão associados também às falhas nas séries temporais de dados para a Amazônia e aos poucos dados existentes. De acordo com Nobre (2008), há grandes chances de aumento na frequência de eventos extremos de secas e cheias decorrentes das mudanças climáticas, semelhante ao que se observou na Amazônia nos anos de 2005/2010 e 2009/2012, sendo anos de secas e inundações acentuadas, respectivamente.

A sua vasta rede de drenagem qualifica a Amazônia como região de elevado potencial para produção de energia hidrelétrica e pode ser considerada como alternativa vital para suprir as demandas energéticas do país, sobretudo, por meio de uma fonte renovável. A instalação de usinas hidrelétricas, especialmente em áreas sensíveis como as unidades de conservação existentes na Amazônia, inevitavelmente causa impactos ambientais de grande relevância, intensificando a pressão sobre a região e, por isso, a execução de um projeto desta envergadura exige uma avaliação criteriosa do potencial de geração de energia para verificar se o empreendimento será capaz de manter de forma tecnicamente sustentada a sua capacidade de geração no médio e longo prazos.

Com base no cenário que se desenha para a Amazônia, a inegável importância estratégica que possui para a geração de energia no Brasil, considerando-se as dificuldades frequentemente encontradas na região acerca dos registros hidrológicos, e diante das incertezas sobre a variabilidade e mudanças potenciais do clima, este trabalho busca apresentar uma metodologia aplicada à modelagem do potencial mensal de geração de energia hidrelétrica usando séries temporais de precipitação para ajudar na tomada de decisão antes da implantação dos projetos de UHE.

Visto que fenômenos naturais como a chuva, por exemplo, tem comportamento altamente não-linear e influencia diretamente a vazão e, conseqüentemente, o potencial de geração de energia, esta pesquisa adotou a hipótese de que este problema pode ser tratado com aplicação de técnicas de inteligência artificial, visto que tem sido demonstrado em inúmeras aplicações a sua capacidade para resolver problemas reais e complexos.

1.2. LOCALIZAÇÃO DO SÍTIO DE ESTUDO

A metodologia proposta para o potencial de geração de energia hidrelétrica foi aplicada para o AHE Jatobá, que está situado na bacia do rio Tapajós. A bacia hidrográfica do Tapajós

compreende os estados do Pará, Mato Grosso, Amazonas e Rondônia, e pertence à bacia do rio Amazonas. O rio Tapajós tem aproximadamente 2.000 km de extensão, nasce no Mato Grosso, da confluência dos rios Juruena e São Manuel, também conhecido como Teles Pires, e flui para o rio Amazonas pelo estado do Pará (SOUSA JÚNIOR, 2014). A área total da bacia do rio Tapajós tem 764.183 km², o relevo é constituído de planícies, regiões montanhosas e platôs, cuja altitude varia de 51 metros a cerca de 900 metros (ELETRONORTE; CAMARGO CORREA; CNEC, 2008).

Parte do relevo da bacia do rio Tapajós é resposta ao soerguimento do escudo brasileiro associado à formação arco magmático da cordilheira dos Andes. Grande parte da bacia do rio Tapajós apresenta baixa amplitude de relevo, cujos delimitadores geomorfológicos são escassos, predominando os rios e seus afluentes como marcos geográficos locais. A drenagem dos principais afluentes do rio Tapajós apresenta evolução geomorfológica de longo prazo (GET, 2014).

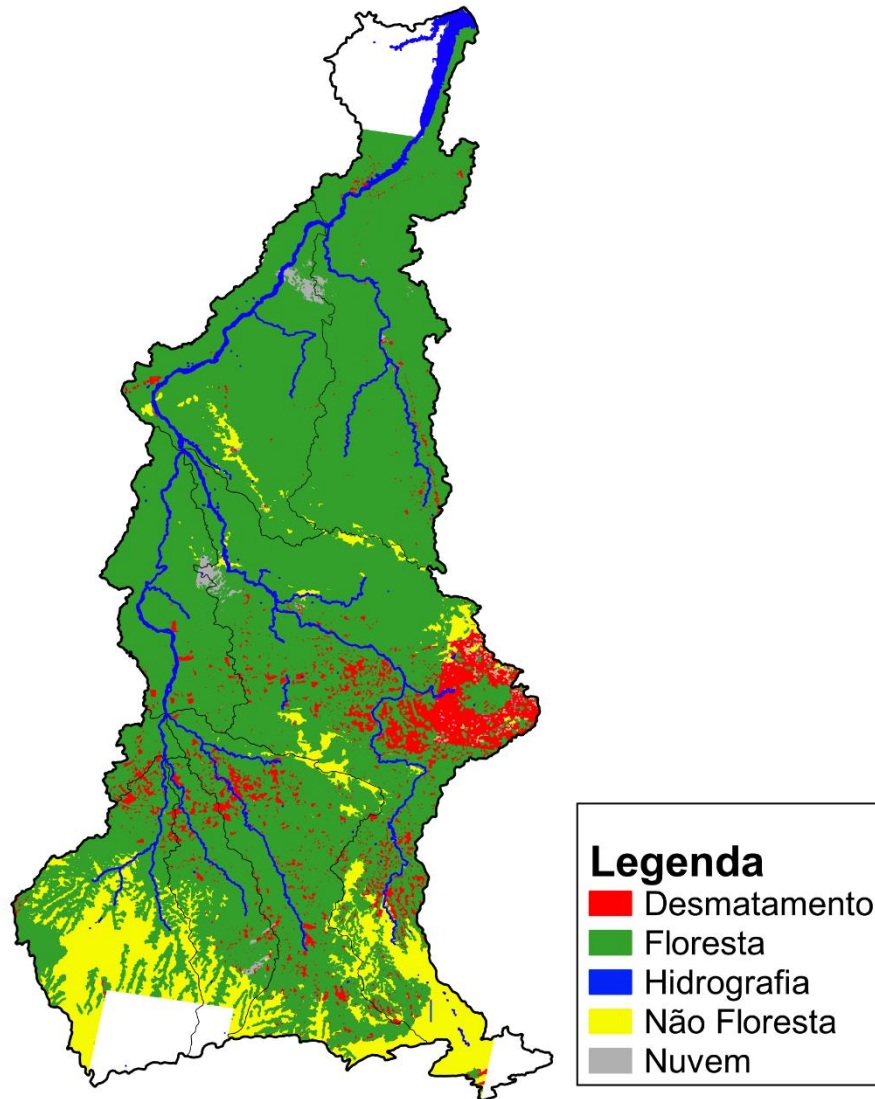
A maior sub-bacia do Tapajós é a do rio Teles Pires (sub-bacia 444), que cobre a maior parte do território. Na bacia do Tapajós há uma dinâmica de sedimentação fluvial caracterizada pela presença de ilhas, bancos de areia e extensas planícies fluviais.

Considerando as características pedológicas da bacia do Tapajós, predomina a classe Latossolo Vermelho-Amarelo (GET, 2014). Esta classe de solo ocorre principalmente sob condições de relevo ondulado, com variações entre suavemente ondulado e fortemente ondulado. Também apresenta elevada susceptibilidade aos processos erosivos, especialmente em áreas de morros residuais, onde os solos são pouco profundos e a declividade mais acentuada pode favorecer o movimento de massas. Tal condição pode ainda ser agravada pelas condições climáticas, em especial o regime pluviométrico, que será melhor detalhado no capítulo 4.

Sob o ponto de vista da fertilidade, os solos na bacia do Tapajós são predominantemente pobres, ou seja, baixa fertilidade natural, característicos de solos lixiviados (GET, 2014). Ainda assim, a fitofisionomia predominante sobre a região é de floresta de terra firme ou floresta ombrófila. A relação entre a floresta vigorosa em meio a um solo pobre é decorrente da ciclagem dos nutrientes, que consiste em um processo contínuo de troca de elementos químicos essenciais entre as frações viva e morta da biomassa. A bacia do Tapajós ainda mantém a maior parte da sua área com vegetação nativa preservada, com pequeno grau de desmatamento, condição favorecida pela existência de áreas protegidas em grande parte do território. Segundo registros do Projeto de Monitoramento do Desflorestamento da Amazônia (PRODES), realizado pelo Instituto Nacional de Pesquisas Espaciais (INPE), percebe-se que o maior grau de

desmatamento ocorre em uma pequena área no setor Oeste da bacia do Tapajós (Figura 3). A referida área pertence à sub-bacia 444 (Figura 4), cujo principal curso d'água é o rio Teles Pires.

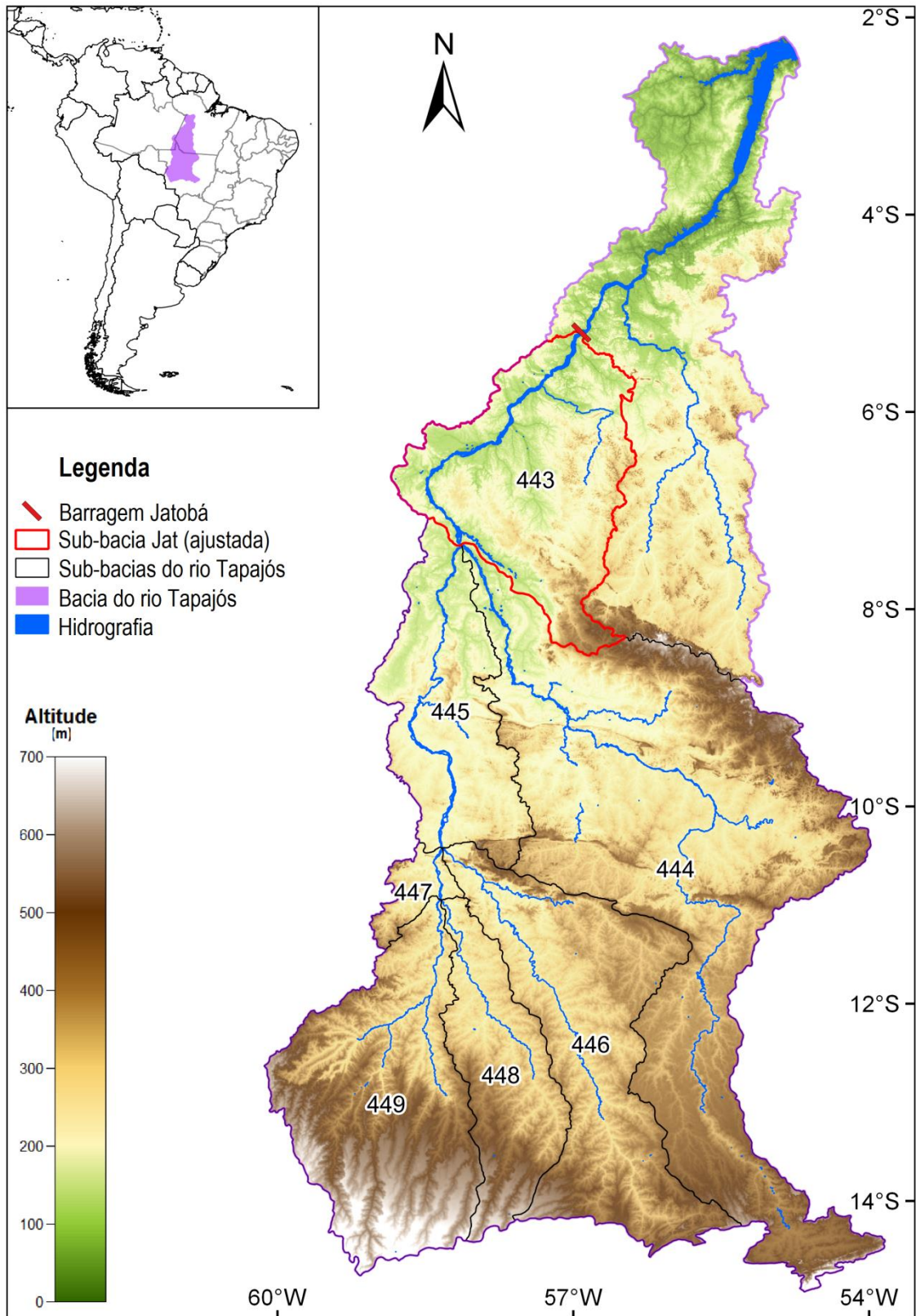
Figura 3 - Classes de cobertura do solo para o ano base 2018.



Fonte: INPE, 2018.

A instalação e operação da UHE Jatobá deverá impactar diretamente os municípios de Itaituba e Jacareacanga, com uma área do reservatório de 646,3 km². Porém a área inundada será de 204 km², descontando-se a área do rio Tapajós. A barragem deverá apresentar uma queda de 16 metros entre a montante e a jusante, com comprimento de 1.287 metros e altura máxima da barragem de 35 metros.

Figura 4 - Área de estudo: bacia do rio Tapajós.



Fonte: Elaborado pelo autor do trabalho.

A bacia do rio Tapajós pode ser subdividida em sete sub-bacias menores, cuja área de drenagem tem influência direta no AHE Jatobá (Figura 4). Os códigos numéricos apresentados nas áreas delimitadas são os mesmos adotados pela Base Hidrográfica Ottocodificada da Agência Nacional de Águas (ANA), conforme estabelecido pela Resolução nº30/2002 do CNRH (MMA, 2002) e representam as sub-bacias que compõem a bacia principal do Tapajós, segundo a hierarquia estabelecida pela metodologia de Otto Pfafstetter. O método Otto Pfafstetter de codificação considera inicialmente o curso d'água principal da bacia para estabelecer sua codificação. Para determinar qual o curso d'água principal, parte-se da foz da bacia e, a cada confluência com outros rios tributários decide-se qual o trecho de maior área de contribuição. Esse processo agregando trechos ao curso principal é repetido a cada confluência até a montante. A metodologia de codificação pode ser consultada em detalhes no “Manual de Construção da Base Hidrográfica Ottocodificada da ANA” (ANA, 2007).

As sub-bacias 443, 444, 445, 446, 447, 448 e 449 constituem a área de drenagem do AHE Jatobá (Figura 4). A barragem da UHE deverá ficar situada aproximadamente 1 km a jusante da localidade de Jatobá (ELETRONORTE; CAMARGO CORREA; CNEC, 2008).

1.3. PRINCÍPIO FÍSICO TEÓRICO DA TESE

A metodologia proposta é baseada na hipótese de que para produzir hidroeletricidade, além dos parâmetros de engenharia do projeto, a geração de energia propriamente dita depende principalmente da vazão do rio. Por sua vez, para que haja naturalmente vazão suficiente é necessário que ocorra chuva na bacia hidrográfica. A água precipitada na bacia pode percorrer os seguintes caminhos: infiltração no solo, retorno à atmosfera por evaporação, e escoamento (*runoff*). Assim, de forma simplificada, o conceito físico empregado nesta metodologia é transformar a entrada do sistema (precipitação) em potencial de geração de energia, que é a saída. Para isso, foram utilizadas técnicas de inteligência artificial porque são capazes de reconhecer padrões através de um processo iterativo de aprendizagem.

Esta metodologia reserva algumas condições para sua aplicação efetiva, como restrição para barramentos em cascata a montante do AHE, cujo sítio se busca realizar a modelagem do potencial de geração de energia. Todavia, esta condição pode ser viável desde que tais barragens também sejam do tipo fio d'água e o fluxo natural do rio não seja subitamente alterado por ações de controle da(s) UHE a montante. As oscilações no fluxo de vazão devem ser restritas à variabilidade climática natural, sob pena de introduzir forte ruído ao aprendizado de máquina.

1.4. DEFINIÇÃO DE OBJETIVOS

De modo a tratar o problema apresentado nos tópicos anteriores acerca da baixa disponibilidade de dados para estimar o potencial de geração de energia na Amazônia, a tese elencou um objetivo geral e seu desdobramento em objetivos específicos para abordar de forma apropriada a temática em foco.

1.4.1. Objetivo Geral

Desenvolver uma metodologia capaz de simular e prever, por meio de técnicas de inteligência computacional, o potencial de geração de energia em aproveitamentos hidrelétricos, sobretudo em regiões como a Amazônia, onde os registros de dados observacionais de vazão são escassos, com séries temporais curtas e/ou com muitas falhas.

1.4.2. Objetivos Específicos

- (a) Desenvolver e avaliar modelos preditivos com base em inteligência artificial para estimar o potencial de geração de energia em aproveitamentos hidrelétricos;
- (b) Realizar análises estatísticas e climatológicas sobre os dados observados de precipitação na bacia do rio Tapajós e potencial de geração energética no AHE Jatobá;
- (c) Avaliar o impacto na geração de energia do AHE Jatobá de acordo com cenários simulados para um período de até vinte anos futuros.

1.5. ESTRUTURA DA TESE

A tese está organizada em capítulos, conforme se apresenta a seguir:

Capítulo 1: contextualiza o ambiente no qual o problema central da tese se desenvolve, apresentando os objetivos da pesquisa e a sua justificativa.

Capítulo 2: realiza uma abordagem multidisciplinar sobre os principais conceitos e aspectos de relevância nas áreas de energia, hidrologia, climatologia e aprendizado de máquina para embasar o tratamento da problemática da tese.

Capítulo 3: caracteriza os dados utilizados e apresenta os procedimentos metodológicos para o tratamento dos dados utilizados e desenvolvimento dos modelos preditivos empregados.

Capítulo 4: apresenta os principais resultados obtidos para atender os objetivos geral e específicos propostos na tese.

Capítulo 5: faz a discussão sobre os resultados encontrados no âmbito da climatologia estatística e modelagem preditiva do potencial de geração de energia.

Capítulo 6: apresenta as considerações finais sobre a pesquisa e proposições de continuidade da pesquisa com trabalhos futuros.

CAPÍTULO II

2. CONCEITOS BÁSICOS

Visto que esta tese tem um caráter multidisciplinar, considerando diferentes áreas do conhecimento técnico-científico, este capítulo apresenta uma visão geral sobre alguns pontos importantes abrangendo os temas abordados na pesquisa para sua melhor compreensão.

2.1. CONCEITOS BÁSICOS SOBRE ENERGIA

Segundo estabelece a mecânica Newtoniana, o trabalho pode ser definido como o produto escalar entre uma força e o deslocamento realizado durante um intervalo de tempo (MICHAELIDES, 2012). A potência é o trabalho por unidade de tempo. Quando a força é variável, o trabalho é dado por uma integral que descreve o movimento da força entre dois pontos a e b . Portanto, o trabalho e sua derivada no tempo, a potência instantânea, são definidas da seguinte forma:

$$W = \int_a^b \vec{F} \cdot d\vec{x} \quad (1)$$

onde W é o trabalho, a e b são, respectivamente, ponto de partida e chegada para o deslocamento $d\vec{x}$, e \vec{F} é a força aplicada. A potência, por sua vez, é o produto escalar da força pela velocidade.

$$\dot{W} = \vec{F} \cdot \frac{d\vec{x}}{dt} = \vec{F} \cdot \vec{V} \quad (2)$$

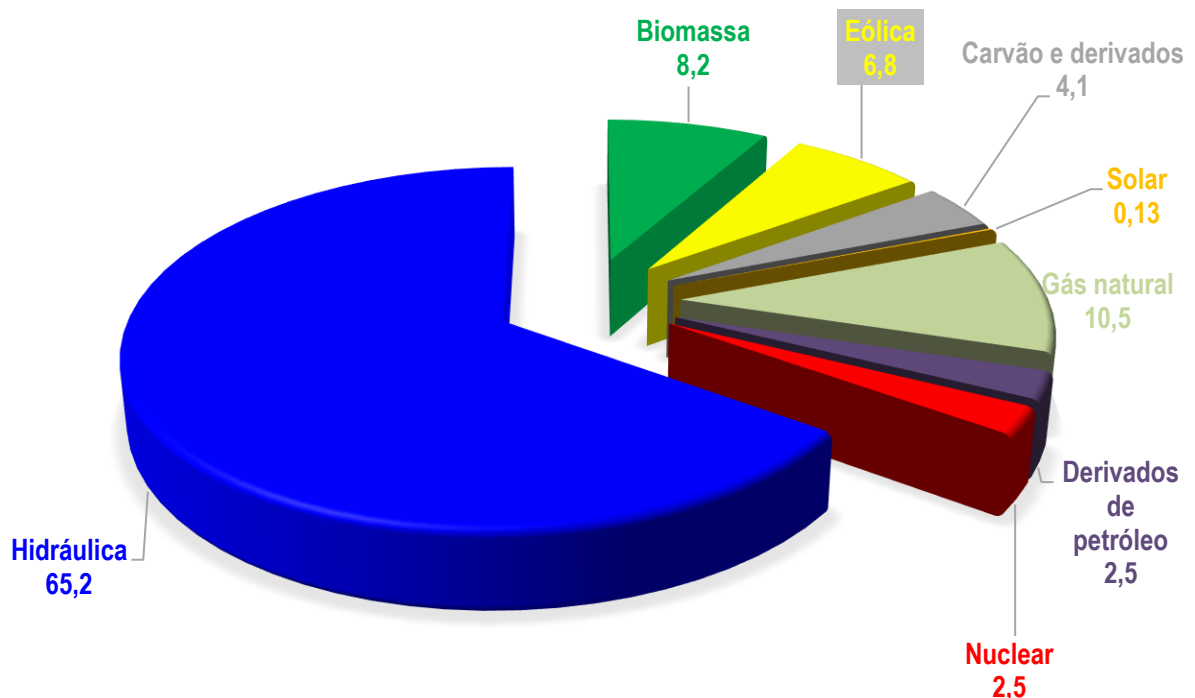
onde \dot{W} é a potência instantânea, dt é a variação no tempo e \vec{V} é a velocidade.

A energia é um conceito mais amplo e pode ser definida como o potencial dos materiais e sistemas para realizar algum trabalho útil. Pode ser mensurada na mesma unidade que o trabalho e vir em diversas formas. Por exemplo, energia calorífica ou termal é uma forma de energia que é transferida de materiais em temperaturas mais altas para aqueles de temperaturas mais baixas; sendo usualmente produzida pela combustão de certos materiais, que são considerados como fontes de energia. Existem diversos sistemas de engenharia, como por exemplo, usinas e máquinas de combustão interna, que convertem calor em trabalho. Outras formas de energia que são comumente encontradas em aplicações práticas no dia a dia da engenharia são: a energia potencial, que é uma forma de energia que fica armazenada em um

corpo; a energia cinética que está relacionada com o movimento de um corpo; a energia química que está associada às substâncias e é liberada quando ocorre a quebra das ligações químicas entre os átomos que compõem a matéria, como carvão e hidrocarbonetos, por exemplo; a energia nuclear, que é resultante da transformação de núcleos atômicos; e energia elétrica que é uma forma de energia que surge da diferença de potencial elétrico entre dois pontos, o que permite estabelecer uma corrente elétrica entre eles. A energia elétrica é uma forma de energia muito importante, especialmente para o modo de vida moderno que utiliza inúmeros equipamentos eletro-eletrônicos, tanto para facilitar as atividades diárias da rotina doméstica, como para o lazer, assim como para atividades da indústria e demais setores produtivos.

Na matriz de energia elétrica do Brasil predominam as fontes renováveis de energia (EPE, 2018), ou seja, aquelas que são geradas a partir de fontes inesgotáveis, ou de outra forma, que podem ser recicladas, sendo tipicamente de fontes naturais como a energia solar, a energia eólica e a energia hidráulica. Dentre estas, as fontes hídricas são responsáveis por 65% da oferta de energia elétrica no Brasil (Figura 5).

Figura 5 - Oferta (%) de energia elétrica por fonte.

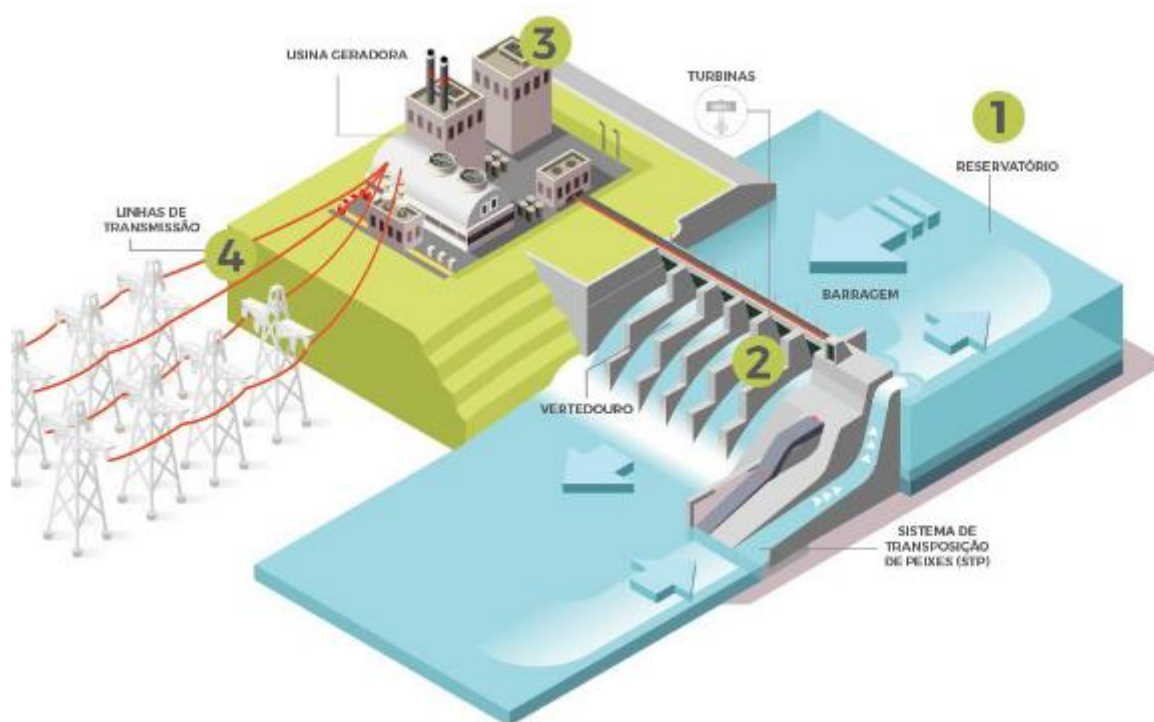


Fonte: EPE, 2018.

2.1.1. Hidroeletricidade

A hidroeletricidade é o termo que se refere a energia elétrica gerada pelo movimento das águas, seja pela força gravitacional da queda ou pelo próprio fluxo da água. Este é um conceito já bastante sedimentado pela sociedade moderna. A hidroeletricidade tem sido utilizada por mais de cem anos na maioria dos países que possuem alguma fonte hídrica disponível para geração de energia elétrica. Basicamente, a hidroeletricidade é produzida pelo processo e controle do fluxo de água através de uma barragem. Na Figura 6 pode-se observar esquematicamente os processos para a geração de energia hidrelétrica. [1] Os sistemas de captação e adução levam a água até a casa de força, estrutura na qual são instaladas as turbinas. [2] As turbinas são equipamentos cujo movimento giratório provocado pelo fluxo de água faz girar o rotor do gerador, fazendo com que o deslocamento do campo magnético produza energia elétrica [3]. O vertedouro, por sua vez, permite a saída do excesso de água do reservatório, quando o nível ultrapassa determinados limites. Outros aspectos e outros equipamentos também são importantes, entretanto, em última análise, trata-se de uma busca por “queda” e “vazão”, onde a primeira é fixa, mas a segunda é variável. [4] E por fim, a energia elétrica produzida é transmitida para uma ou mais linhas de transmissão, as quais são interligadas à rede de distribuição.

Figura 6 - Esquema de uma usina hidrelétrica.



Fonte: <https://www.todamateria.com.br/usina-hidreletrica/>

A hidroeletricidade tem sido a forma mais comum no mundo de geração de energia elétrica em grande escala e provavelmente, ainda será assim por alguns anos. A construção de barragens em locais estratégicos importantes, conforme decidido por engenheiros, significa que a geração de energia pode ser aumentada ou diminuída dependendo das necessidades da comunidade que a utiliza. Durante os períodos de baixa utilização a saída pode ser reduzida, e aumentada durante os períodos de alta demanda de produção (DIVYA; JOSEPH, 2014); essas mudanças podem ser feitas rapidamente em comparação com a produção de petróleo, por exemplo, que tem um atraso devido à necessidade de refinar o produto bruto. A velocidade com que a produção de energia hidrelétrica pode ser alterada é uma grande vantagem para nossas crescentes necessidades energéticas.

Considerando apenas o custo financeiro e a longevidade do empreendimento, a hidroeletricidade é uma das formas de energia elétrica de menor custo, visto que não requer combustível, que por sua vez significa ausência de mineração, sem processamento e mesmo sem custo de transporte (ANEEL, 2008; MICHAELIDES, 2012; DIVYA; JOSEPH, 2014). Pode ser considerada como uma das formas mais limpas de energia, mesmo levando-se em conta que o processo de construção da barragem, assim como sua manutenção, possa significar emissões de carbono, sobretudo por conta das áreas alagadas, cujas emissões decorrentes são inferiores comparadas às emissões por combustíveis fósseis, sobretudo no longo prazo.

Atualmente, as barragens apresentam usos diversificados, não sendo destinadas exclusivamente à geração de energia. A regularização de vazão exercida pelo barramento pode contribuir para amenizar inundações em períodos extremamente chuvosos, assim como manter o abastecimento de água durante estações mais secas (HASSAN, 2007). Um caso clássico ocorreu nos anos 1980, quando países como Sudão e Etiópia foram arrasados com secas severas, enquanto a barragem de Assuã manteve o Egito em condições bem mais confortáveis que seus vizinhos.

Embora a hidroeletricidade seja bastante utilizada no Brasil e no mundo, a sua implantação em geral acarreta algum custo social e ambiental, o que faz com que engenheiros e tomadores de decisão devam ter em mente um planejamento muito bem realizado de modo que a facilidade da energia gerada e seus benefícios associados sejam superiores às desvantagens de sua implantação. A construção de barragens sofre muitas críticas especialmente de ambientalistas. A cobrança é maior sobre aquelas UHE de grande porte e com reservatório de acumulação, ou seja, as que formam um lago extenso e volumoso, utilizado para armazenar água no período chuvoso e manter a geração de energia elétrica no período seco.

Este processo é conhecido como regularização de vazão e permite o controle e planejamento mensal ou anual, como é o caso, por exemplo, da usina de Tucuruí, no Pará (Figura 7).

Figura 7 - Usina hidrelétrica com reservatório no município de Tucuruí-PA.



Fonte: http://www3.ana.gov.br/porta/ANA/imagens/imagens-divisao-hidrografica/imagens-regiao-hidrografica-tocantins-araguaia/toca_143-8040444-usina-hidreletrica-de-tucurui-pa-rui-faquini.jpg/image_view_fullscreen.

Um dos principais pontos de questionamento acerca das barragens são os danos ambientais e sociais inerentes à implantação de uma usina hidrelétrica, os quais são considerados extremamente elevados devido às extensas áreas alagadas, prejudicando a fauna e a flora da região. A inundação de grandes áreas para o represamento do curso d'água de maneira a formar o reservatório de acumulação requer o deslocamento de comunidades inteiras, com abandono de cidades e realocação dos seus habitantes e empreendimentos de negócios. As áreas inundadas também provocam mudanças ecológicas profundas, prejudicando a fauna e a flora local, além da inundação de sítios e monumentos arqueológicos (HASSAN, 2007; MICHAELIDES, 2012).

As barragens também afetam a fauna aquática, em especial a migração de peixes, além de restringir o caminho natural das águas. Para mitigar tal impacto, são criados mecanismos de transposição dos peixes através de corredores em forma de escadarias laterais. A geração de

energia elétrica altera o ambiente aquático corrente abaixo, devido à sedimentação e passagem de água através de peneiras.

Alguns estudos associam às usinas a elevação dos níveis de gás metano sobre os reservatórios e área do entorno, sobretudo em regiões tropicais, como a Amazônia, em razão de processos químicos anaeróbicos por conta das árvores que ficaram submersas após o enchimento do lago (DEEMER et al., 2016; FEARNSSIDE, 2016). O metano é um dos gases de efeito estufa e embora seu tempo de vida na atmosfera seja menor que o do gás carbônico, seu potencial de aquecimento pode ser até 60 vezes maior do que o do CO₂. Outra crítica severa é o risco de rompimento da barragem, cujo dano potencial é absolutamente catastrófico. Em tais circunstâncias, qualquer assentamento no vale abaixo seria inundado, levando possivelmente à perda de muitas vidas humanas, destruindo casas, interrompendo o fornecimento de energia a todas as residências afetadas e possivelmente inundando a paisagem, além dos danos ecológicos.

Como forma de minimizar os aspectos negativos de grandes aproveitamentos hidrelétricos, a legislação incorporou medidas para reduzir os impactos ambientais e sociais, passando a construir usinas do tipo fio d'água, prática que o governo federal adotou em todas as licitações para aproveitamentos hidrelétricos de grande porte desde 2005 até os dias atuais. As usinas do tipo fio d'água não possuem reservatórios para armazenamento de água ou possuem reservatórios com dimensões bastante reduzidas, apenas para manter as condições mínimas de funcionamento das turbinas, o que permite tão somente a regularização do fluxo diário ou semanal. Este é um conceito que tem sido aplicado especialmente na Amazônia, como as usinas de Santo Antônio e Jirau, no rio Madeira, em Rondônia; UHE Belo Monte, no rio Xingu, no Pará; e UHE Estreito (Figura 8), no rio Tocantins, no Maranhão.

Figura 8 - Usina hidrelétrica do tipo fio d'água no município de Estreito-MA.



Fonte: <http://www.oas.com.br/oas-com/oas-engenharia/realizacoes/energia/#>

A barragem em si interrompe o curso do rio e represa a água formando um reservatório que permite regular a vazão. De tal forma, esse reservatório pode ser diretamente manejado pelo gestor da usina para controlar a vazão. Em uma usina a fio d'água, este controle é comprometido pela ausência de reservatório, ficando mais suscetível às alternâncias do comportamento climático. Em geral, usinas a fio d'água têm baixos “fatores de capacidade”. O fator de capacidade é uma grandeza adimensional obtida pela divisão da energia efetivamente gerada ao longo do ano – em geral, medida em MWh/ano – pela energia máxima que poderia ser gerada no sistema (ABBUD; TANCREDI, 2010). Trata-se, portanto, de uma medida da limitação da usina no que diz respeito à sua capacidade de gerar energia.

Os projetos de engenharia para usinas hidrelétricas são complexos e exigem várias fases antes da energia chegar ao consumidor final. Após o leilão do aproveitamento hidrelétrico, seguem as fases para elaboração do projeto básico e dos programas e projetos sócio-ambientais, que devem estar em consonância com os estudos de viabilidade e sócio-ambientais aprovados. A Agência Nacional de Energia Elétrica (ANEEL) é a responsável por aprovar o projeto básico e autorizar a construção da usina. O órgão de licenciamento ambiental, por sua vez, é responsável pela emissão da Licença de Instalação, que é o documento que permite o início da

construção da usina sob fiscalização da ANEEL. E antes da usina entrar em operação, é necessária uma Licença de Operação, condição para o enchimento do reservatório. Todo este processo é relativamente demorado, podendo levar entre cinco e quinze anos.

2.1.2. Potencial de Geração de Energia Hidrelétrica

De forma mais geral, podemos considerar como principais variáveis utilizadas na classificação de uma usina hidrelétrica: a altura da queda d'água (ou queda hidráulica), a vazão natural do curso d'água, a capacidade ou potência instalada, o tipo de turbina empregada, o sítio de localização, o tipo de barragem e o reservatório (ANEEL, 2008). Todos são fatores interdependentes. A altura da queda d'água e a vazão dependem do local de construção e determinarão qual será a capacidade instalada - que, por sua vez, determina o tipo de turbina, barragem e reservatório. De maneira mais direta, pode-se dizer que a energia elétrica gerada em uma usina é proporcional ao produto entre a vazão natural do curso d'água e a queda hidráulica. Com base nesta premissa, o potencial de geração de energia hidrelétrica de um AHE é estimado usando a seguinte expressão (KOCH et al., 2011; SHARMA; SINGH, 2013; STICKLER et al., 2013; BUTERA; BALESTRA, 2015):

$$PH_i = \eta \times \rho \times Q_i \times g \times \Delta h \quad (3)$$

onde PH_i é o potencial hidrelétrico médio mensal (em megawatts) no mês i ; η é o fator de eficiência, que corresponde basicamente ao rendimento do conjunto turbina-gerador; ρ é a densidade da água (em quilogramas por metro cúbico); Q_i é a vazão média mensal (em metros cúbicos por segundo); g é a aceleração gravitacional (em metros por segundo quadrado); Δh é a queda hidráulica (em metros). Assim, substituindo os valores de Q_i em (3) para cada mês da série de dados, pode-se obter a série temporal mensal do potencial de geração de energia.

O potencial hidrelétrico é a quantidade de energia possível de ser tecnicamente e economicamente aproveitada empregando as tecnologias atualmente disponíveis (EPE, 2008). Usualmente, o potencial hidrelétrico pode ser mensurado em termos de energia firme, ou seja, a quantidade máxima de geração de forma contínua mesmo em condições de regime hidrológico crítico.

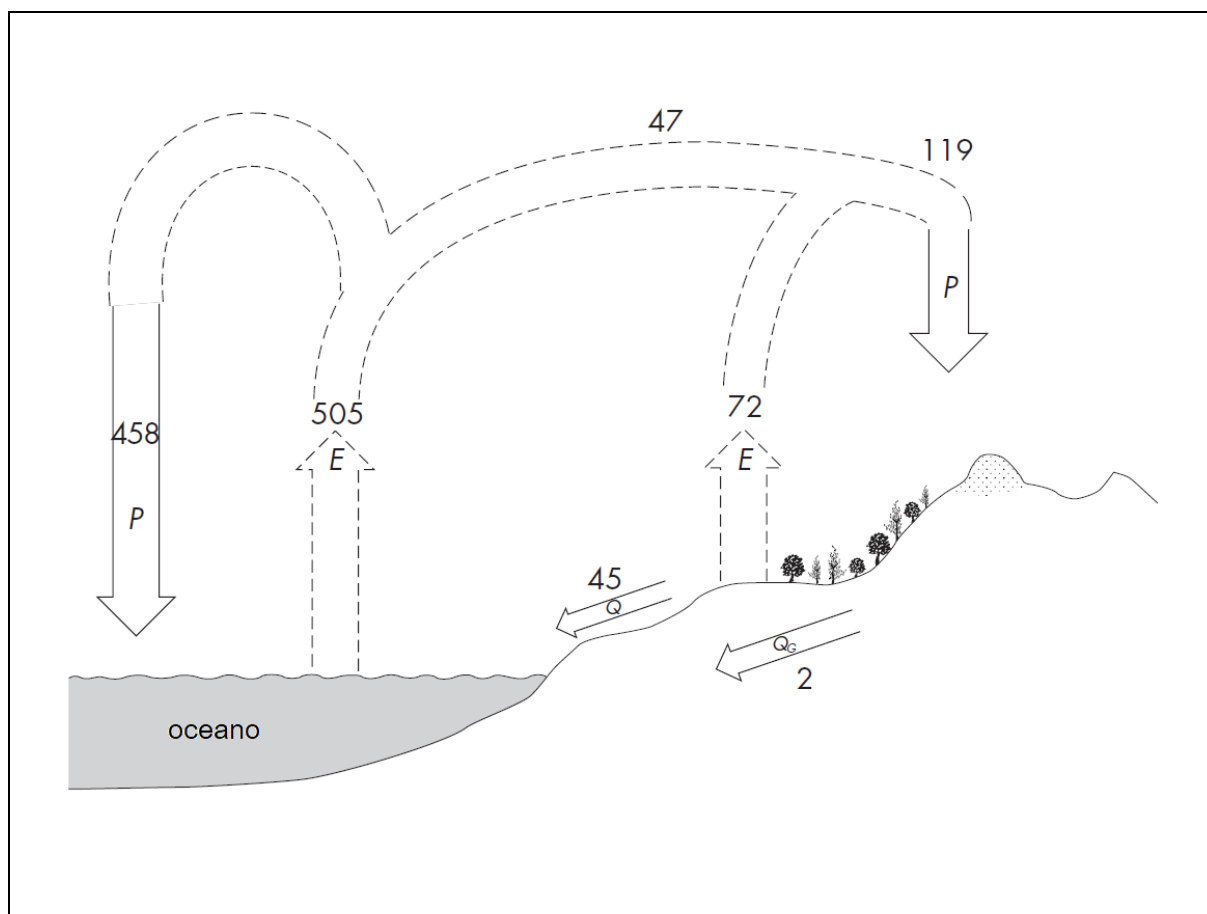
2.2. CONCEITOS BÁSICOS SOBRE HIDROLOGIA

Os recursos hídricos têm importância direta na vida dos seres vivos, pois todos dependem de água para manter suas atividades fisiológicas. Além disso, o homem faz uso e gestão dos recursos hídricos para o abastecimento e realização de diversas atividades, seja lazer, irrigação de culturas agrícolas, criação de animais terrestres e aquáticos, navegação, geração de energia, entre outras. Para continuar utilizando um bem tão necessário e precioso, é fundamental conhecer o seu comportamento para realizar melhor e adequadamente o seu gerenciamento. É a hidrologia a ciência que engloba o estudo da água na superfície e abaixo da superfície da Terra, a ocorrência e o movimento da água, as suas propriedades físicas e químicas, e sua relação com os componentes vivos e materiais do meio ambiente (BALES, 2015).

A unidade espacial mais comum sob consideração da hidrologia é a bacia hidrográfica, que consiste de uma área de terra a partir da qual os fluxos d'água correm em direção a um rio e, então, do rio para o mar. Tal terminologia sugere que a área seja análoga a uma bacia onde todas as águas se movem em direção a um ponto central. O denominador comum em qualquer ponto de uma bacia é que onde quer que a chuva caia, ela finalizará no mesmo lugar, ou seja, onde o rio encontra o mar. Naturalmente que este conceito está centrado no escoamento da água, desprezando algumas perdas, seja por evaporação ou por infiltração. As bacias hidrográficas podem apresentar tamanhos variados, desde poucos hectares a milhões de quilômetros quadrados, como é o caso da bacia Amazônica que possui cerca de 7 milhões de quilômetros quadrados (MMA, 2006).

Como ponto de partida sobre qualquer abordagem hidrológica, é sempre útil considerar o ciclo hidrológico. Este é na verdade um modelo conceitual de como a água se move entre a Terra e a atmosfera nos seus diferentes estados, seja gasoso, líquido ou sólido. Como qualquer outro modelo conceitual, este pode conter algumas simplificações. Na Figura 9 é apresentado um esquema do movimento da água em torno do sistema Terra-atmosfera e é uma representação do ciclo hidrológico global.

Figura 9 - Ciclo hidrológico global. Os números representam uma estimativa da quantidade total de água (milhares de km^3) em cada processo por ano. E é a evaporação, P é a precipitação, Q_G é o escoamento subterrâneo e Q é o escoamento superficial.



Fonte: Adaptado de Shiklomanov, 1993.

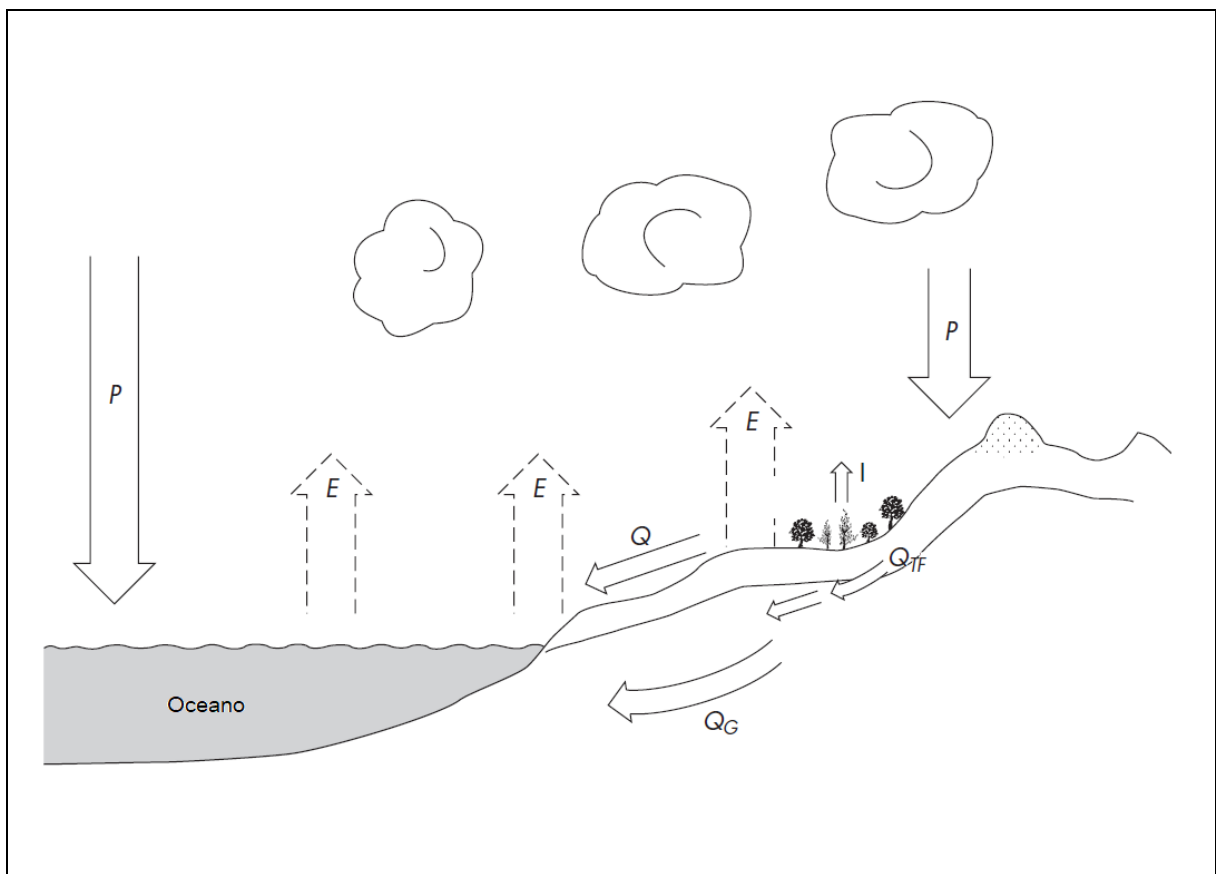
O ciclo consiste de evaporação da água líquida em vapor d'água que é movido através da atmosfera. Em algum estágio, o vapor d'água é condensado em líquido (ou sólido) novamente e cai na superfície como precipitação. Os oceanos perdem mais água por evaporação do que recebem por precipitação. Por outro lado, os continentes recebem mais água precipitada do que perdem para a atmosfera por evaporação. Essa diferença entre precipitação e evaporação na zona terrestre é o escoamento superficial e sub-superficial das águas (ou *runoff*) em direção aos oceanos, completando assim o ciclo hidrológico (RAGHUNATHY, 2006). A partir da Figura 9, percebe-se que as maiores quantidades de evaporação e precipitação ocorrem sobre os oceanos. Os valores no diagrama são totais globais, mas variam enormemente ao redor do globo. A precipitação total é particionada em diversos processos hidrológicos em diferentes quantidades dependendo do clima.

O termo precipitação abrange a queda de qualquer hidrometeoro, seja no estado líquido ou sólido. Hidrometeoro é um termo empregado para representar um fenômeno atmosférico

formado por um conjunto de partículas aquosas, na forma líquida ou sólida (por ex.: chuva, neve, granizo), em queda livre ou suspensão na atmosfera, ou levantada da superfície pela ação do vento, ou depositada em superfícies (por exemplo: orvalho).

Sob a ótica de uma escala menor o ciclo hidrológico de uma bacia hidrográfica, o modelo conceitual pode ser observado com mais profundidade acerca das operações dos processos hidrológicos. Adaptando a Figura 9 para a Figura 10, ainda há essencialmente três processos predominando (evaporação, precipitação e *runoff*), mas agora pode-se subdividir cada um em diferentes subprocessos. A evaporação é uma mistura de evaporação da água livre (ou seja, rios e lagos), evaporação do solo, evaporação da superfície das plantas, interceptação e transpiração das plantas.

Figura 10 - Processos do ciclo hidrológico operando na escala de bacia hidrográfica. Q é o *runoff*, Q_G é o escoamento subterrâneo, Q_{TF} é o escoamento sub-superficial em solo não saturado, I é a interceptação, E é a evaporação e P é a precipitação.



Fonte: Adaptado de Shiklomanov, 1993.

A interceptação da precipitação pelas plantas torna a água disponível novamente para evaporação antes mesmo que ela alcance a superfície do solo. O termo *runoff* de forma mais

abrangente engloba o movimento da água líquida tanto acima, quanto abaixo da superfície terrestre. Os processos do ciclo hidrológico são bastante complexos e ocorrem em escalas de tempo dentro da bacia hidrográfica. Matematicamente, tais processos do ciclo hidrológico podem ser representados por meio de uma equação, denominada equação do balanço hídrico, conforme segue:

$$P + E + \Delta S + Q = 0 \quad (4)$$

onde P é a precipitação, E é a evaporação, ΔS é a mudança no armazenamento e Q é o *runoff*. (4) incorpora os princípios de continuidade de massa e energia, ou seja, o ciclo hidrológico é considerado um sistema fechado, onde nenhuma massa ou energia é criada ou perdida. Neste caso, a massa refere-se à água. Os termos da equação do balanço hídrico podem ser interpretados como uma série de fluxos e armazenamentos. A equação do balanço hídrico, portanto, avalia o fluxo relativo de água para a superfície e a partir da superfície com o termo de armazenamento também incorporado. A precipitação na equação do balanço hídrico representa a principal entrada de água em uma superfície, por exemplo, na bacia hidrográfica. Uma grande parte da hidrologia envolve medições e estimativas da quantidade de água envolvida nestes fluxos de transferência e armazenamento de água.

2.2.1. Medição de Vazão em Cursos D'água

Para entender o comportamento dos rios, suas variabilidades sazonais, as vazões e os regimes de precipitação que os influenciam, levando em consideração a sua distribuição tanto no espaço, como no tempo, torna-se necessária a coleta e interpretação dos dados. É a hidrometria, um ramo da hidrologia, que é responsável pela medição e análise das características físicas e químicas da água, o que inclui os métodos, suas técnicas e instrumentos empregados (CHOW et al., 1994).

Um segmento importante dentro da hidrometria é a fluviometria, que abrange as medições de cotas e vazões dos rios. As cotas linimétricas, ou simplesmente cotas, ou ainda níveis fluviométricos são a altura da superfície da água tomando como referência uma escala arbitrária. Para observação sistemática das características de um curso d'água faz-se necessária a instalação de instrumentos que permitam efetuar as medições que se deseja. Para tal, são instaladas estações hidrométricas que consistem em postos situados em uma determinada seção de um rio e que possuem um dispositivo de medição do nível das águas (régua linimétrica ou

linígrafos), devidamente referidos a uma cota conhecida ou materializada no terreno, chamada de referencial de nível (RN), eventualmente dotadas de facilidades (botes, pontes, outros) para medição de vazão e, quando necessário, providas de estruturas artificiais de controle (CHOW et al., 1994).

A vazão ou descarga de um rio é o volume de água que passa através de uma seção transversal do rio por unidade de tempo. Sua medição é extremamente importante, não há dúvida, porém, a obtenção diária da vazão por processos de medição direta é, além de bastante oneroso, operacionalmente complicado, sobretudo para rios com grande volume, correnteza, turbulência e margens muito distantes, particularmente como ocorrem nos rios da Amazônia. Por esta razão, operacionalmente são adotados processos indiretos para medição de vazão.

O processo mais usual para obtenção da vazão de forma indireta é através da curva-chave, ou seja, uma relação estabelecida entre a cota e a vazão por meio de uma equação matemática ajustada aos dados. Esta relação pode ser unívoca ou não, constante ou variável com o tempo, dependendo das condições locais (PINTO et al., 1973). Em linhas gerais, a curva-chave é nada mais que uma equação ajustada aos dados de medição de vazão, com a seguinte forma:

$$Q = a (h - h_0)^b \quad (5)$$

onde Q é a vazão, h é a cota, h_0 é a cota quando a vazão é zero e, a e b são parâmetros da equação de potência. Os parâmetros h_0 , a e b são normalmente ajustados usando-se um método de minimização de desvios entre valores previstos pela equação e valores medidos, como por exemplo o método de mínimos desvios quadrados. Outras formulações têm sido propostas, sobretudo com o avanço dos computadores, o que facilita o cálculo. Herschy (2009) propôs uma representação geral da vazão por meio de um polinômio de grau M :

$$Q = a_0 + a_1 h + a_2 h^2 + \dots + a_M h^M = \sum_{m=0}^M a_m h^m \quad (6)$$

onde a_0 , a_1 , ..., a_M são coeficientes e h é a cota. Outras possibilidades de representação matemática da vazão podem encontrar bons resultados (FENTON, 2018).

A cota é observada com auxílio da régua linimétrica (Figura 11) ou registrada por meio de linígrafos. A régua linimétrica é uma escala graduada usada para medir a altura do nível fluviométrico, sendo usualmente registradas pelo observador duas leituras diárias, a primeira às 07:00 h e a segunda às 17:00 h.

Figura 11 - Lances de régua linimétrica no rio Xingu, Altamira-PA.



Fonte: autor.

Para obtenção da curva-chave, é necessário que seja realizada a medição da vazão através de campanhas de campo, oportunidade em que são coletados, além da vazão, diversos outros dados, como largura da seção, área molhada, velocidade, profundidade. A esse conjunto de informações de campo chama-se resumo de descarga. As campanhas de campo para obtenção do resumo de descarga ocorrem periodicamente.

A obtenção da vazão é efetuada por meio de um processo de medição e integração do campo de velocidades na seção transversal do rio. Para medição da velocidade da água em qualquer profundidade de um curso d'água, um dos instrumentos frequentemente utilizados é o molinete hidráulico (Figura 12), o qual consiste de pequenas hélices que giram impulsionadas pela passagem da água. A relação entre velocidade da água e velocidade de rotação do molinete é estabelecida por uma equação fornecida pelo fabricante do equipamento, assim denominada de equação do molinete (COLLISCHONN; TASSI, 2011). Mais recentemente, o molinete tem sido progressivamente substituído por modernos equipamentos (*Acoustic Doppler Current Profiler* - ADCP), que utilizam o efeito doppler em ondas acústicas para obtenção da velocidade por diferença de frequência entre os sons emitidos pelo equipamento e refletidos pela água.

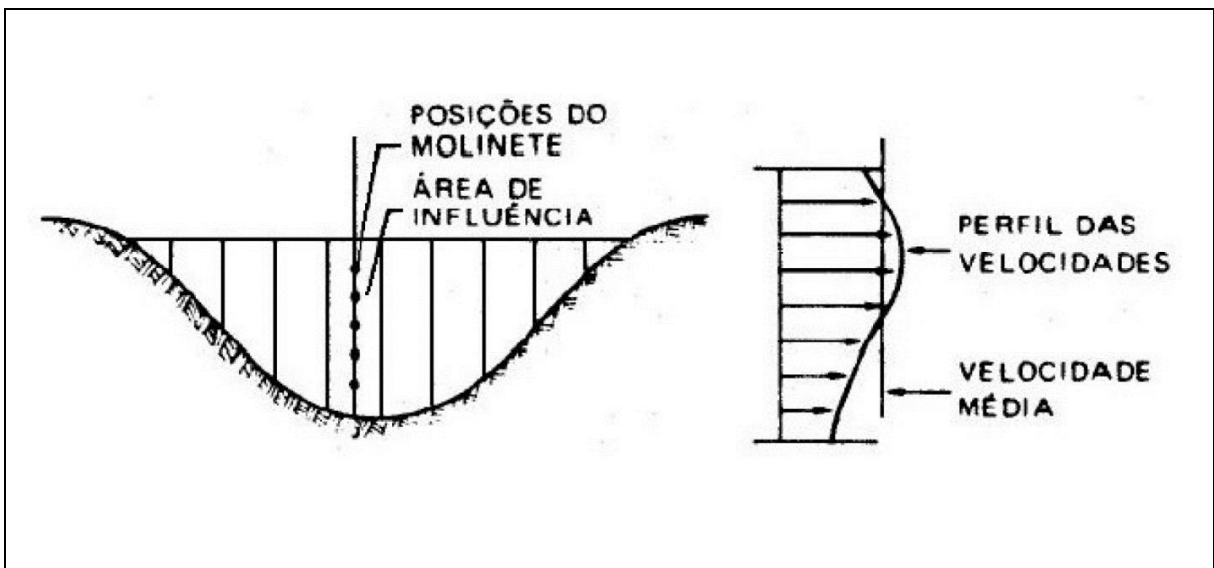
Figura 12 - Molinete para medição de velocidade da água.



Fonte: <http://www.aparelhosmedir.com/>.

A velocidade da corrente de um fluxo é, geralmente, maior na região central do rio comparado àquela medida em suas margens. Em razão de tal variação em diferentes pontos da seção transversal, as medidas de velocidade devem ser tomadas tanto na superfície da seção transversal, como também em diferentes profundidades (Figura 13).

Figura 13 - Perfil de velocidade em uma seção de rio.



Fonte: <https://capacitacao.ead.unesp.br/dspace/bitstream/ana/112/1/apostila.pdf>

Atualmente, vários postos de observação hidrológica contam com plataformas de coleta de dados, as quais podem efetuar as medições e registros dos dados hidrológicos automaticamente sem necessidade de um observador. Entretanto, para se estabelecer a relação cota-vazão, ainda é necessária a realização de campanhas de campo para obtenção do resumo de descarga.

Dentre as análises estatísticas com os dados hidrológicos, uma das mais importantes devido à sua aplicabilidade é a curva de permanência, a qual consiste em uma distribuição de frequências acumuladas relativas das vazões de um determinado rio para um ponto específico e serve para fornecer informações mais completas sobre a variabilidade do escoamento e a avaliação da transformação chuva-vazão (CASTELLARIN et al., 2013), sendo passível de utilização para identificar as alterações em decorrência das mudanças ambientais de natureza regional e global (PEÑA-ARANCIBIA et al., 2012; RODRIGUEZ et al., 2010; SIQUEIRA JÚNIOR et al., 2015).

A curva de permanência estabelece uma relação entre a vazão e a frequência com que determinado valor de vazão pode ser superado ou igualado. Para aplicações em empreendimentos de geração de energia hidrelétrica, utiliza-se como base para definir a energia assegurada de uma planta o referencial Q_{95} , ou seja, o valor de vazão que é superado em 95% do tempo (MME, 2007; EPE, 2008; CCEE, 2010).

2.2.2. Determinação da Precipitação Média em Bacias Hidrográficas

As medições e estimativas de precipitação (volume e intensidade) são críticas para qualquer estudo ou modelagem envolvendo o ciclo hidrológico. Os pluviômetros são os dispositivos primários para observação da precipitação líquida (chuva), cuja quantidade é usualmente medida em milímetros, considerando que esta medida é por unidade de área. Assim, por exemplo, 10 mm de precipitação corresponde a uma lâmina d'água de 10 mm de altura ou 10 litros por metro quadrado. As redes de pluviômetros, particularmente na Amazônia, frequentemente têm suas distribuições esparsas e outras limitações podem não fornecer a resolução espacial e temporal necessária para vários esforços da pesquisa científica e sobretudo da modelagem. Embora os pluviômetros obtenham uma medida de certa forma precisa, é uma informação pontual. Neste sentido, os avanços do sensoriamento remoto orbital ou de radares meteorológicos têm contribuído para aperfeiçoar as informações de precipitação em áreas com poucos ou nenhum sensor de superfície, visto que sua informação é referente a uma área e não simplesmente a um único ponto. Entretanto, as informações de satélites e radares são estimativas e necessitam dos pluviômetros para sua devida calibração.

A precipitação sobre grandes áreas não é uniforme e, portanto, usar o volume de precipitação de um único ponto não é apropriado para representar o volume precipitado na bacia. Por essa razão, é necessário determinar a precipitação média na bacia (RAGHUNATH, 2006). Para tal, há pelo menos três métodos consagrados na literatura especializada: (a) método de média aritmética, (b) método do polígono de Thiessen, e (c) método das isoietas.

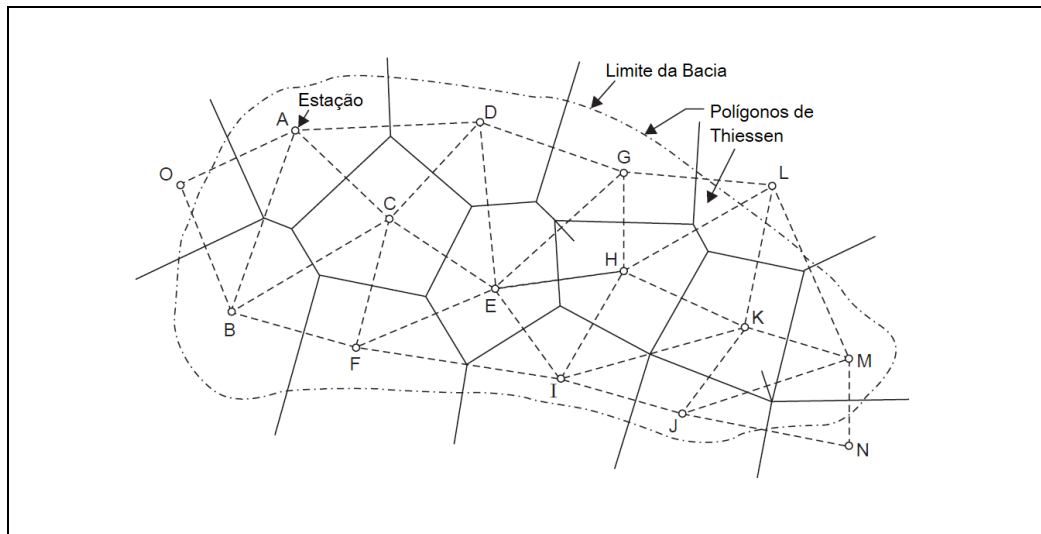
(a) Método de média aritmética: consiste em aplicar uma média aritmética simples e direta sobre os volumes de chuva registrados pelos pluviômetros que compõem a área da bacia, conforme equação a seguir. Embora este método seja simples e rápido de ser aplicado, somente apresenta bons resultados quando a área é relativamente plana, os pluviômetros estão uniformemente distribuídos e a precipitação registrada não apresenta grandes variações em torno da média.

$$\bar{P} = \frac{\sum_{i=1}^n P_i}{n} \quad (7)$$

onde \bar{P} é a precipitação média na bacia, P_i é a precipitação no pluviômetro i , n é o número de pluviômetros na bacia.

(b) Método do polígono de Thiessen: este método admite que cada área do polígono é influenciada pelo pluviômetro dentro desta área. Para determinação dos polígonos, plota-se os pluviômetros em um mapa da área de interesse, conectando-os entre si por meio de linhas retas (THIESSEN, 1911). A partir deste traçado, são desenhadas bissetrizes perpendiculares às retas anteriormente traçadas, juntando-se os pluviômetros adjacentes para formar polígonos (Figura 14).

Figura 14 - Traçado da mediatriz para definição dos polígonos de Thiessen.



Fonte: Adaptado de Raghunathy (2006).

Esta abordagem permite utilizar uma distribuição irregular de pluviômetros, aplicando-se um de fator de ponderação para cada ponto de medição da precipitação, conforme determinado pela expressão a seguir:

$$\bar{P} = \frac{\sum_{i=1}^n A_i P_i}{A} \quad (8)$$

onde \bar{P} é a precipitação média na bacia; P_i é a precipitação no pluviômetro i ; A é a área total da bacia; e A_i é a área parcial da bacia sob influência do pluviômetro i . Os resultados obtidos pelo método de Thiessen são usualmente mais acurados do que o método anterior (RAGHUNATH, 2006).

(c) Método das isoietas: neste método os pontos de precipitação são plotados em um mapa de base e as linhas de mesmo valor de precipitação, denominadas de isoietas, são traçadas levando-se em consideração a orografia e a morfologia das nuvens de tempestade. O valor médio entre duas isoietas sucessivas é tomado como a média das duas isoietas, sendo ponderado com a área entre as isoietas. O somatório de cada produto é dividido pela área total da bacia para obtenção da precipitação média, conforme segue:

$$\bar{P} = \frac{\sum \left(\frac{P_i + P_{i+1}}{2} \right) A_i}{\sum A_i} \quad (9)$$

onde A é a área total, A_i é a área entre as duas isoietas sucessivas P_i e P_{i+1} . Este método pode obter resultados melhores que os anteriores quando apropriadamente analisados e levando em consideração os fatores de fisiografia da região e características das nuvens de precipitação (RAGHUNATH, 2006). Todavia, sem tais considerações e para uma grade regular, tende a aproximar-se do método proposto por Thiessen.

2.3. CONCEITOS BÁSICOS SOBRE CLIMATOLOGIA

A atmosfera afeta a todos os seres no planeta e nós reagimos às condições do tempo atmosférico por meio de nossa habilidade em tomar decisões. O tempo e o clima são fatores importantes na determinação das atividades do nosso dia a dia e também de longo prazo. O tempo refere-se às condições atmosféricas em um dado momento. Já o clima é frequentemente definido como uma média das condições de tempo. Embora esta definição não esteja errada, também não se deve entender o clima como algo constante, pelo contrário, o sistema climático está em constante mudança.

A climatologia é o ramo da meteorologia que estuda o clima. Enquanto a climatologia aplicada pode ser entendida como a análise científica de dados climáticos à luz de uma aplicação útil para um propósito operacional (LANDSBERG; JACOBS, 1951). Entretanto, mesmo quando há disponibilidade de grandes quantidades de dados climáticos, pode não ser trivial obter respostas definitivas. Algumas das incertezas resultam do fato de nos dias atuais ainda não termos um entendimento completo do sistema climático, alguns dados são inadequados ou de baixa qualidade, e particularmente pela falta de habilidade para resolver todas as escalas, espacial e temporal, relevantes de variabilidade climática. Sob este aspecto, a análise estatística pode frequentemente não apresentar respostas definitivas, entretanto pode ser de grande ajuda para estabelecer relações e quantificar as incertezas dentro dos dados climáticos.

Há três propriedades dos dados climáticos a considerar: normais, extremos e frequências. Essas propriedades são usadas para avaliar o estado da atmosfera sobre um período de tempo particular quando comparadas com as condições atmosféricas sobre um período de tempo similar no passado. As “Normais” se referem à média das condições de tempo em um lugar. As “Normais Climatológicas” são tipicamente calculadas para períodos de trinta anos e fornecem uma visão das condições de tempo esperadas para um determinado local ao longo do ano. Por exemplo, o outono em Belém-PA é bastante chuvoso.

Os “Extremos” são usados para descrever as medições de máximo e mínimo para uma determinada variável atmosférica que pode ser esperada ocorrer em um certo local em um dado

tempo, baseado em um longo período de observações. Por exemplo, uma temperatura mínima de 5°C não deve ocorrer em Belém-PA qualquer que seja a época do ano.

E por fim, as “Frequências” que se referem à taxa de incidência de um fenômeno em um local particular sobre um longo período de tempo. Os dados de frequência são especialmente importantes para avaliação de risco, aplicações de engenharia ou agricultura, entre outros. Por exemplo, caso um engenheiro projete uma rede de drenagem pluvial para suportar 30 mm de precipitação em uma hora e esta frequência é excedida em média cinco vezes no ano, talvez essa taxa de falha não seja aceitável para as pessoas que terão suas casas invadidas pela água.

2.3.1. Variabilidade Climática

As variações climáticas resultam na modificação da média, da variância e da frequência de valores extremos de temperatura, precipitação, pressão atmosférica e outras variáveis meteorológicas. Tais flutuações, que podem ocorrer nas escalas de tempo e espaço, além daquelas de eventos individuais de tempo, são definidas pelo *Intergovernmental Panel on Climate Change* – IPCC como variabilidade climática. As mudanças com significância estatística que ocorrem no estado médio do clima ou na sua variabilidade esperada e que persistem por um período prolongado, da ordem de décadas ou mais, são referidas como mudanças climáticas (IPCC, 2013).

A análise climática requer dados homogêneos e isto é particularmente importante para análises de variabilidade e mudança do clima. Uma série temporal climática homogênea é definida como aquela onde as variações são causadas exclusivamente por variações no clima. Esta colocação se faz importante porque há muitos registros de séries climatológicas de longo prazo que foram afetadas por muitos fatores não climáticos e que tornaram tais dados não representativos da variação climática real que ocorreu ao longo do tempo. Esses fatores podem incluir mudanças: nos instrumentos de medição, nas práticas de observação, no ambiente entorno da estação, o que pode provocar um viés gradual nos dados. Todas essas não homogeneidades podem inserir um viés na série temporal conduzindo a interpretações equivocadas.

As causas de variabilidade do clima e/ou mudança do clima estão relacionadas com: processos internos e naturais como as interações dinâmicas e termodinâmicas entre a atmosfera e os oceanos, erupções vulcânicas; e/ou com forçantes externas como mudanças no fluxo de radiação solar, ou mudanças antropogênicas persistentes na composição da atmosfera, ou ainda, mudanças no uso e cobertura da terra. Tais fatores podem afetar diretamente o clima alterando o balanço de energia local e/ou indiretamente podem modificar a circulação planetária de ventos

e correntes oceânicas pelos quais o clima desde a escala local até grandes extensões são regulados. Os efeitos da variação climática são numerosos e frequentemente afetam os sistemas naturais e socioeconômicos.

2.3.2. Mudanças Climáticas

O tema “mudanças climáticas” foi discutido pela primeira vez entre governantes e cientistas em reunião ocorrida em Toronto, Canadá, no ano de 1988. A motivação para tal discussão foi a preocupação da comunidade científica com o aquecimento global e suas consequências para o clima do planeta. Este evento originou o IPCC e, dois anos depois, em 1990, foi divulgado o seu primeiro relatório no qual os cientistas já advertiam para a necessidade de estabilizar os níveis crescentes de dióxido de carbono (CO_2), visto que este constitui um importante gás do efeito estufa (GEE), cujas elevadas emissões antrópicas representariam uma ameaça ao sistema climático. Muito embora o efeito estufa seja um processo natural em que diversos gases presentes na atmosfera como vapor d’água, metano (CH_4), dióxido de carbono e outros aprisionam grande parte da radiação de onda longa emitida pela superfície terrestre, elevando a temperatura do ar na troposfera (camada da atmosfera mais próxima da superfície), o aumento na concentração desses gases, em especial do CO_2 e CH_4 , os quais têm ciclo de vida de até 100 anos, atuaria para intensificar o efeito estufa, ocasionando impactos desastrosos para o planeta.

Segundo anunciado através dos relatórios produzidos pelo IPCC, o aquecimento do sistema climático é inequívoco, com alterações observadas desde a década de 1950, como aquecimento da atmosfera e oceanos, redução nas quantidades de neve e gelo, elevação do nível dos mares e, principalmente, o aumento crescente na concentração de GEE.

Diante de tais evidências, o quinto relatório do IPCC produzido pelo Grupo de Trabalho 1 (*Working Group 1 – WG1*), divulgado no segundo semestre de 2013, definiu um conjunto de quatro novos cenários que estão relacionados com as emissões ou concentrações de gases de efeito estufa, aerossóis e outros fatores climáticos. Tais cenários são baseados em emissões de caráter antropogênico e não incluem mudanças produzidas por forçantes naturais dos fatores climáticos, como radiação solar e erupções vulcânicas. Os cenários estabelecidos, denominados por Caminhos da Concentração Representativa (*Representative Concentration Pathways – RCP*), são identificados de acordo com a forçante radiativa total (resultante) aproximada para o ano de 2100, em relação ao ano de 1750, sendo assim especificadas: RCP2.6 para forçante de $2,6 \text{ Wm}^{-2}$; RCP4.5 para forçante de $4,5 \text{ Wm}^{-2}$; RCP6.0 para forçante de $6,0 \text{ Wm}^{-2}$; RCP8.5 para forçante de $8,5 \text{ Wm}^{-2}$. O RCP2.6 considera um cenário de mitigação que conduzirá o nível da

forçante radiativa a um valor muito baixo. O RCP4.5 e o RCP6.0 representam cenários de estabilização das emissões, enquanto o RCP8.5 consiste em um cenário de emissões muito elevadas para os gases de efeito estufa.

Em setembro de 2008 ocorreu um encontro envolvendo 20 grupos de modelagem climática de todo o mundo, o Grupo de Trabalho sobre Modelagem Acoplada (*Working Group on Coupled Modeling – WGCM*) do Programa Mundial de Pesquisa Climática (*World Climate Research Programme – WCRP*). Na oportunidade, os participantes decidiram promover um novo conjunto de experimentos coordenados de modelagem climática. Estes experimentos passaram a constituir a quinta fase do Projeto de Intercomparação de Modelos Acoplados (*Coupled Model Intercomparison Project - CMIP5*), a qual tem por objetivo fornecer um contexto multi-modelo para: (i) avaliar os mecanismos responsáveis pelas diferenças de modelo quanto aos *feedbacks* pouco compreendidos associados ao ciclo do carbono e às nuvens; (ii) examinar a “previsibilidade” do clima e explorar a capacidade dos modelos para prever o clima em escalas de tempo decadais; e (iii) determinar por que os modelos forçados da mesma forma produzem uma gama de diferentes respostas.

O CMIP iniciou em 1995 sob a égide do WGCM, tendo como objetivo entender melhor as mudanças climáticas passadas, presentes e futuras decorrentes de causas naturais, ou em resposta a mudanças na forçante radiativa em um contexto multi-modelo. Esse entendimento inclui avaliações de desempenho do modelo durante o período histórico e quantificações das causas da propagação nas projeções futuras. Também são usados experimentos idealizados para ampliar a compreensão das respostas dos modelos. Além dessas respostas na escala de longo prazo, são realizados experimentos para investigar a previsibilidade do sistema climático em várias escalas de tempo e espaço, bem como fazer previsões a partir de condições climáticas observadas. Os resultados dos modelos utilizados no CMIP5 são disponibilizados através de centros de pesquisa localizados em diversos países, como Inglaterra, Alemanha e Estados Unidos.

Um dos modelos climáticos de circulação global empregados pelo CMIP5 e projetados para rodar os principais cenários para o IPCC é o HadGEM2 (*Hadley Centre’s Global Environmental Model version 2*) do *United Kingdom Meteorological Office*. O HadGEM2 é um modelo de sistema terrestre acoplado para simulações centenárias do CMIP5, e compreende uma gama de configurações de modelos específicos, incorporando diferentes níveis de complexidade, mas com uma estrutura física comum. A família HadGEM2 inclui uma configuração atmosfera-oceano acoplada, com ou sem uma extensão vertical na atmosfera para incluir uma estratosfera bem resolvida, e uma configuração do Sistema Terra que inclui

vegetação dinâmica, biologia oceânica e química atmosférica (COLLINS et al., 2008). Como os modelos climáticos globais possuem resolução mais grosseira, aproximadamente 100 km, é usual fazer um *downscaling* com modelos regionais ou de mesoescala.

O *downscaling* é um conceito geral que abrange vários métodos para aumentar a resolução espacial em áreas de interesse, objetivando melhorar a usabilidade dos cenários climáticos dos modelos climáticos globais (MCGUFFIE; HENDERSON-SELLERS, 2005; TUCKER et al., 2018). No *downscaling* dinâmico ocorre o acoplamento entre as diferentes grades dos modelos regional e global, aumentando a resolução em uma área limitada, fazendo uso pelo modelo regional das condições iniciais e de contorno a partir do modelo global.

Um dos modelos regionais também empregado no CMIP5 é o *Regional Climate Model* (RegCM) que foi desenvolvido inicialmente pelo *National Center for Atmospheric Research* (NCAR) (GIORGI; BATES, 1989; DICKINSON et al., 1989) e, atualmente é mantido pelo *International Centre for Theoretical Physics* (ICTP), onde passa frequentemente por modificações e aprimoramentos (GIORGI et al., 1993a, 1993b, GIORGI; MEARNNS, 1999; PAL et al., 2007). O RegCM4 é a sua versão atual para simulação regional de climas futuros, podendo ser aplicado para qualquer região do globo, cuja resolução espacial pode chegar até 10 km. O RegCM4 é um modelo de área limitada, com as equações primitivas para fluido compressível e hidrostático, em coordenada vertical sigma e grade horizontal de Arakawa-Lamb B, a qual estabelece que as variáveis escalares (temperatura, pressão, etc.) são definidas no centro da ponto de grade, enquanto as variáveis vetoriais, como o vento, são colocadas nos vértices do ponto de grade. O esquema de superfície, o processo de transferência radiativa, os esquemas de parametrização convectiva e outras características que detalham toda a física do modelo são descritas por Giorgi et al. (2012).

2.3.2.1. Mudanças Climáticas e Impactos na Amazônia

A região Amazônica há muito é tema de pesquisa nas mais diferentes áreas. Mesmo diante da escassez de dados na região e de séries temporais curtas e com muitas falhas, alguns estudos apontam impactos potenciais significativos que a região estará sujeita até o fim deste século em razão das mudanças do clima (MARENGO et al., 2011a, c; JOETZJER et al., 2013). Dentre as mudanças esperadas, acredita-se na elevação nas temperaturas do ar e redução no volume de chuvas, em especial devido à maior duração da estação seca. Situações como esta certamente ocasionarão impactos significativos no regime hidrológico (COX et al., 2004, 2008; LI et al., 2006; SALAZAR et al., 2007). Contudo, ainda há certo grau de incerteza em tais afirmações acerca das avaliações sobre os impactos das mudanças climáticas, principalmente

devido ao desempenho dos modelos climáticos e aos erros de propagação inerentes aos modelos (JONES, 2000).

Ainda que as projeções dos modelos climáticos sejam circundadas por incertezas (DESSAI; HULME, 2004), as pesquisas que abordam os impactos são baseadas na hipótese de que os sinais das mudanças climáticas projetadas são, até certo ponto, confiáveis quando obtidas a partir da climatologia dos modelos, e não em relação à climatologia sobre dados observados (WOOD et al., 2002). Em outras palavras, é mais importante a avaliação realizada sobre a climatologia dos modelos para períodos históricos e cenários futuros (ALLEN; INGRAM, 2002; BRAVO et al., 2013), naturalmente que considerando os ajustes estatísticos para minimizar qualquer viés dos modelos (BATES et al., 2008; BÁRDOSSY; PEGRAM, 2011; DEMARIA et al., 2013), do que usar valores absolutos dos cenários previstos.

Para a região Amazônica, alguns estudos não apresentam consenso sobre as alterações na vazão para os rios da bacia quando são consideradas projeções de distintos modelos climáticos (ARORA; BOER, 2001; MILLY et al., 2005; SALATI et al., 2009; LAVADO CASIMIRO et al., 2011; GUIMBERTEAU et al., 2013; SIQUEIRA JÚNIOR et al., 2015; DE SOUZA et al., 2016). Basicamente, tais diferenças surgem em razão das incertezas não projetadas sobre as mudanças de precipitação, mesmo sobre simulações que consideram igual cenário climático. De certa forma, isto caracteriza a complexidade do problema encontrado na modelagem numérica.

De Souza et al. (2016) realizaram simulações sazonais de precipitação sobre a Amazônia Legal para um período de 25 anos utilizando o modelo regional RegCM4 aninhado ao modelo global HadGEM2, sob o cenário de forçante radiativa RCP4.5 projetado pelo IPCC. Sobre a bacia do rio Tapajós, o modelo capturou de forma satisfatória os principais aspectos do comportamento da precipitação, quando comparados os resultados simulados contra os dados observados. Houve tendência de subestimativa do volume de chuvas apenas na Calha Norte da região do Baixo Amazonas durante o trimestre dezembro-janeiro-fevereiro e também março-abril-maio.

Um aspecto relevante a ser considerado sobre as alterações no regime pluviométrico regional na Amazônia devido às mudanças climáticas consiste na variabilidade da vazão no período de vazante, condição particularmente importante para a produção energética das plantas hidrelétricas, em especial aquelas cujo reservatório não permite grande armazenamento de água, como é o caso das usinas a fio d'água, impossibilitando assim o armazenamento durante a estação chuvosa e a dificuldade de regulação ao longo da estação seca. Mohor et al. (2015) estudaram o impacto das mudanças do clima em algumas estações da sub-bacia do rio Tapajós

utilizando o modelo MHD-INPE, o qual simulou os dados históricos de vazão, apresentando subestimativa para os picos de vazão e algumas limitações para simular o período de recessão. Comportamentos semelhantes também foram verificados em outros estudos (COLLISCHONN et al., 2007; BRAVO et al., 2012; DEMARIA et al., 2013; LAVADO CASIMIRO et al., 2011; GUIMBERTEAU et al., 2012).

Abordando particularmente a sub-bacia do rio Tapajós, há certo consenso dos modelos em apontar elevação da temperatura do ar, porém alguns membros de determinados modelos projetam aumento de precipitação, enquanto outros sugerem redução (MOHOR et al., 2015). Por outro lado, Marengo et al. (2011b) sugerem que a duração da estação seca na bacia Amazônica deverá aumentar gradualmente, mesmo havendo maiores volumes de chuva durante a estação chuvosa. Mohor et al. (2015) concluíram que há fortes indícios de que as mudanças climáticas afetarão sensivelmente a capacidade de geração de energia hidrelétrica na sub-bacia do Tapajós a médio e longo prazos.

2.3.3. Tratamento Estatístico de Dados Climatológicos

Como a climatologia trabalha com agregados de condições de tempo atmosférico, a estatística é usada para reduzir uma grande quantidade de registros em uma única informação condensada ou em poucos números de maneira que as informações sejam mais facilmente compreensíveis. Por exemplo, quando, a partir de registros diários de precipitação ao longo de trinta anos, é calculada a precipitação média para se obter a normal climatológica.

Praticamente todos os dados climatológicos podem ser ordenados como séries temporais. Como os processos climatológicos operam sobre uma ampla gama de escalas de tempo (por exemplo, diária, mensal, sazonal, anual, etc.), as análises de séries temporais constituem uma abordagem oportuna para detectar, descrever e modelar a variabilidade climática e seus impactos. Uma propriedade importante das séries temporais é a estacionariedade ou não-estacionariedade. Uma série não estacionária tem propriedades estatísticas (por exemplo, média e/ou variância) que mudam ao longo do tempo. Há muitos tipos diferentes de variabilidade temporal que ocorrem em séries temporais de dados climatológicos. Pode haver uma tendência não linear, uma variabilidade cíclica, comportamento quase periódico, além de descontinuidades (ROBESON, 2005).

Para analisar a presença ou não de tendência em séries temporais cuja distribuição não possui uma distribuição normal, condição frequentemente observada para dados de precipitação, uma abordagem usualmente empregada aos dados climáticos é o teste de Mann-Kendall, pois é um teste não-paramétrico robusto para casos em que a tendência pode ser assumida como

monotônica e, portanto, nenhum ciclo sazonal ou outro está presente nos dados. Outra vantagem do método é não ser influenciado por séries não homogêneas (ZHANG et al., 2009). A estatística Mann-Kendall é calculada segundo a equação:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{Sgn}(x_j - x_i) \quad (10)$$

onde x_j e x_i são os valores anuais nos anos j e i , $j > i$, respectivamente, e

$$\text{Sgn}(x_j - x_i) = \begin{cases} \mathbf{1} & \text{se } x_j - x_i > \mathbf{0} \\ \mathbf{0} & \text{se } x_j - x_i = \mathbf{0} \\ -\mathbf{1} & \text{se } x_j - x_i < \mathbf{0} \end{cases} \quad (11)$$

Se n for igual ou menor que 9, o valor absoluto de S é comparado à distribuição teórica de S derivada por Mann e Kendall. O teste bicaudal é usado para quatro níveis de significância α : 0,1; 0,05; 0,01 e 0,001. Em certo nível de probabilidade, h_0 é rejeitado em favor de h_1 se o valor absoluto de S for igual ou superior a um valor especificado $S_{\alpha/2}$, onde $S_{\alpha/2}$ é o menor S que tem a probabilidade inferior a $\alpha/2$ de aparecer no caso de não haver tendência. Um valor positivo (negativo) de S indica uma tendência crescente (decrecente).

Quando uma tendência monotônica é demonstrada usando o teste de Mann-Kendall (MK) e a tendência parece ser linear, pode-se usar uma reta de Theil-Sen para estimar a inclinação da tendência. A reta Theil-Sen é uma alternativa não-paramétrica à reta de regressão de mínimos quadrados ordinários paramétricos. Enquanto uma reta de regressão de mínimos quadrados ordena como a concentração média muda linearmente com o tempo; uma reta de Theil-Sen modela como a concentração mediana (percentil 50) muda linearmente com o tempo (SEN, 1968).

O método de Sen utiliza um modelo linear para estimar a inclinação da tendência e a variância dos resíduos que devem ser constantes no tempo, e não é muito afetado por erros de dados individuais ou valores atípicos (*outliers*), o que o caracteriza como uma ferramenta estatística bastante robusta (GOCIC; TRAJKOVIC, 2013). Assim, se há presença de tendência na série temporal, a inclinação pode ser estimada ajustando-se um modelo linear, como segue:

$$f(t) = \alpha (t - t_0) + \beta \quad (12)$$

onde α é a inclinação (coeficiente angular – inclinação de Sen) e β é uma constante (coeficiente linear). Para deduzir uma estimativa da inclinação α , as inclinações de todos os pares de dados são calculadas:

$$\alpha_i = \frac{x_j - x_k}{j - k} \quad \text{para } i = 1, 2, \dots, N \quad (13)$$

onde x_j e x_k são os valores dos dados nos tempos j e k , respectivamente, com $j > k$. Se houver n valores x_j na série temporal, serão obtidos tantos quanto $N = n(n - 1) / 2$ estimativas de inclinação α_i . O estimador da inclinação de Sen é a mediana desses valores N de α_i . Os valores N de α_i são classificados do menor para o maior e o estimador do Sen é

$$\alpha = \begin{cases} \alpha_{\frac{N+1}{2}} & \text{se } N \text{ for ímpar} \\ \frac{1}{2} \left(\alpha_{\frac{N}{2}} + \alpha_{\frac{N+2}{2}} \right) & \text{se } N \text{ for par} \end{cases} \quad (14)$$

Para obter uma estimativa de β em (12), são calculados os n valores das diferenças $x_i - \alpha t_i$. A mediana desses valores fornece uma estimativa de β (SIROIS, 1998).

2.4. CONCEITOS BÁSICOS SOBRE INTELIGÊNCIA COMPUTACIONAL

Inteligência Computacional, também denominada Inteligência Artificial, ou IA, é um dos campos relativamente mais recentes da ciência e das engenharias. Os trabalhos com IA começaram mais ativamente após a segunda guerra mundial e o termo IA surgiu em 1956. Atualmente abrange uma enorme variedade de subcampos, que vão do geral (aprendizagem e percepção) ao específico, como jogar xadrez, provar teoremas matemáticos, escrever poesia, dirigir um carro em uma rua movimentada e diagnosticar doenças. A inteligência artificial é constituída por um conjunto de três fatores principais que dão às máquinas a capacidade de “pensarem” como seres humanos. Esses fatores são:

- Uma gama de dados numerosos e complexos (*big data*) que possam ser utilizados;
- Modelos (algoritmos) de aprendizagem de máquina (*machine learning*) que sejam capazes de resolver problemas de classificação, processamento e análise dos dados; e
- Computação (*hardware*) avançada com processamento paralelo rápido e eficiente.

2.4.1. Tópicos sobre *Machine Learning*

Na última década houve um aumento significativo do interesse pelo aprendizado de máquina, com interações cada vez maiores entre homem e o computador, como o reconhecimento facial utilizado frequentemente nos aparelhos do tipo *smartphone*; os automóveis sem condutores, chamados carros autônomos; os atendentes virtuais, bastante empregados em bancos e outros serviços de *telemarketing*, além de inúmeras outras aplicações. O aprendizado de máquina pode ser entendido como o emprego de algoritmos para análise de dados brutos a partir da construção de modelos matemáticos, de maneira automática. Muitos destes algoritmos podem realizar esta tarefa, porém durante bom tempo, as redes neurais estiveram em evidência.

2.4.1.1. *Redes Neurais Artificiais (RNA)*

Muito embora as redes neurais remontem desde a década de 1950, foi o volume cada vez mais crescente de dados, o *big data*, juntamente com a evolução da programação paralela em GPU (*Graphics Processing Unit*) que permitiram uma maior popularização e aplicação das redes neurais para problemas reais e complexos. A utilização de uma GPU em associação com a CPU (*Central Process Unit*) permite acelerar aplicativos de uso geral, mas em especial para as áreas de ciências e engenharia, porque transfere parte do processamento intensivo dos aplicativos para a GPU, enquanto o restante do código segue sendo executado pela CPU. E assim, as Rede Neurais Artificiais (RNA) ganharam cada vez mais evidência.

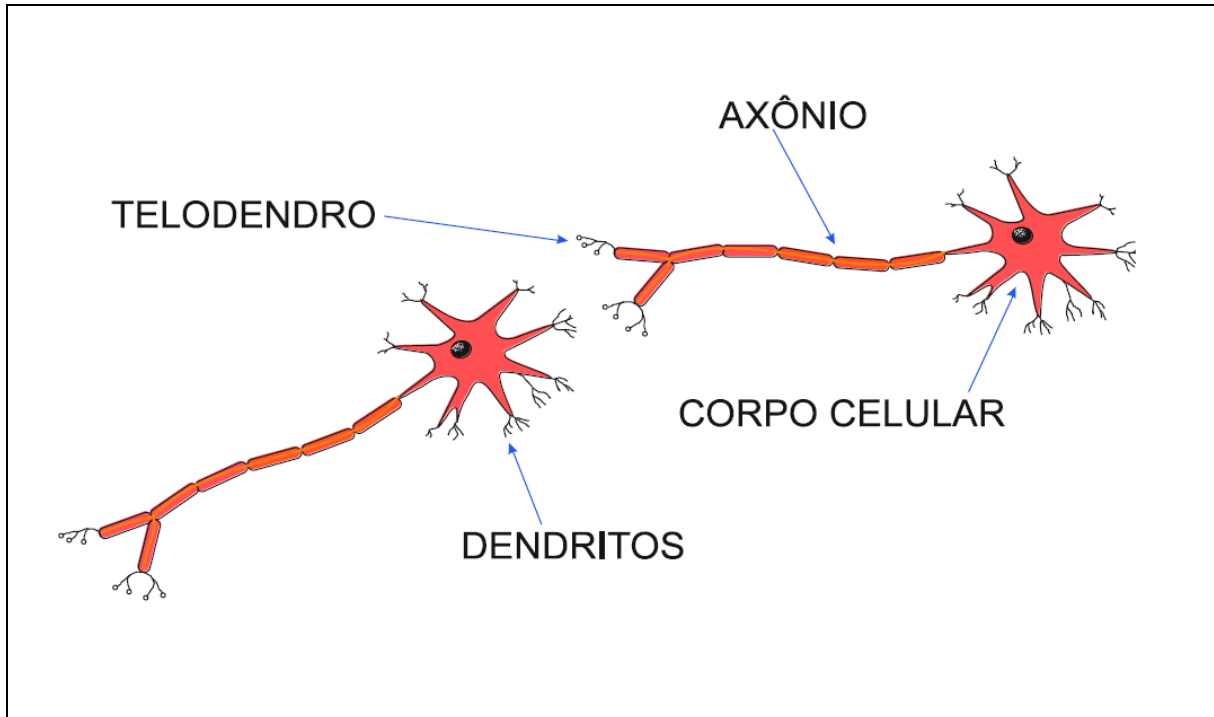
Uma RNA em termos simples é um modelo computacional inspirado biologicamente, que consiste em processar elementos (chamados neurônios), e conexões entre eles com coeficientes (denominados pesos) ligados às conexões. Essas conexões constituem a estrutura neuronal e, anexadas a essa estrutura, estão os algoritmos de treinamento e recuperação. Em razão das conexões encontradas entre os neurônios, as RNA são denominadas de modelos conexionistas (KASABOV, 1995).

Kohonen e Deboeck (1998) descreveram redes neurais como um conjunto de técnicas matemáticas que podem ser usadas para processamento de sinais, previsão e agrupamento, denominando-as como técnicas de regressão paralela não-lineares e multicamadas.

A primeira definição descreve a RNA de acordo com suas semelhanças com o funcionamento do cérebro humano (Figura 15) e o segundo (Kohonen) em uma perspectiva de aplicação. É realmente aceito, inclusive, as recentes iniciativas de pesquisa do cérebro de que o cérebro humano é muito mais complexo, pois muitas de suas funções cognitivas ainda são desconhecidas. No entanto, as principais características consideradas e descritas como funções

comuns em redes reais e artificiais são: aprendizado e adaptação, generalização, paralelismo massivo, robustez, armazenamento associativo de informação, processamento de informação espaço-temporal.

Figura 15 - Neurônio biológico.



Fonte: Valença (2010).

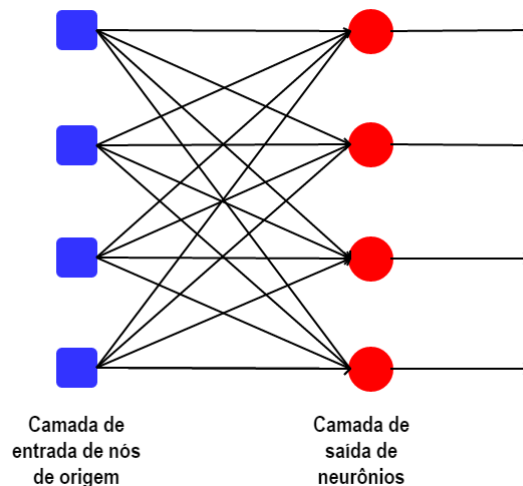
A maneira pela qual os neurônios de uma rede neural são estruturados está intimamente ligada ao algoritmo de aprendizado usado para treinar a rede. Portanto, podemos falar de algoritmos de aprendizado (regras) usados no projeto de redes neurais como sendo estruturados. Em geral, pode-se identificar três classes fundamentalmente diferentes de arquiteturas de rede:

I. Rede feedforward de camada única

Em uma rede neural em camadas, os neurônios são organizados na forma de camadas. Na forma mais simples de uma rede em camadas, temos uma camada de entrada de nós de origem que se projeta diretamente em uma camada de saída de neurônios (nós de computação), mas não vice-versa. Em outras palavras, essa rede é estritamente do tipo *feedforward*. Um exemplo esquemático para o caso de quatro nós nas camadas de entrada e saída é apresentado na Figura 16. Essa rede é chamada de rede de camada única, com a designação de “camada

única” referente à camada de saída dos nós de computação (neurônios). Não é contada a camada de entrada dos nós de origem porque nenhuma computação é executada lá.

Figura 16 - Rede *feedforward* com uma única camada de neurônios.



Fonte: Elaborado pelo autor do trabalho.

II. Rede *feedforward* multicamadas

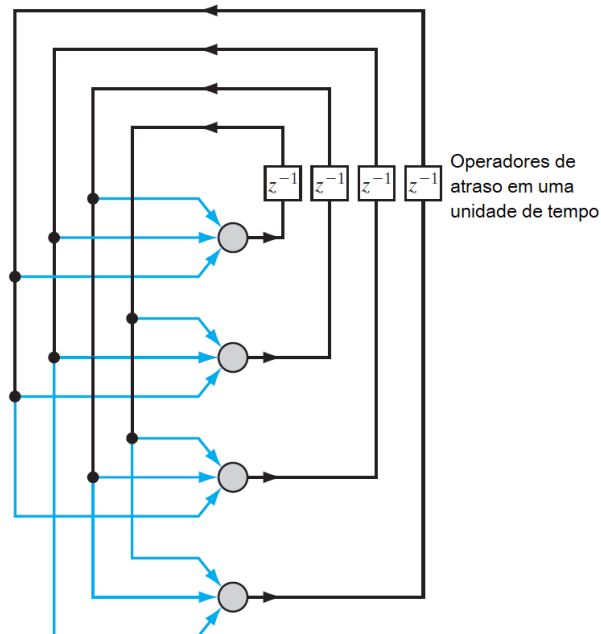
A rede neural *feedforward* multicamadas se distingue pela presença de uma ou mais camadas ocultas, cujos nós de computação são correspondentemente chamados de neurônios ocultos ou unidades ocultas; o termo “oculto” refere-se ao fato de que esta parte da rede neural não é vista diretamente da entrada ou saída da rede. A função dos neurônios ocultos é intervir entre a entrada externa e a saída da rede de alguma maneira útil. Ao adicionar uma ou mais camadas ocultas, a rede é ativada para extrair estatísticas de ordem superior de sua entrada. Num sentido mais amplo, a rede adquire uma perspectiva global, apesar de sua conectividade local, devido ao conjunto extra de conexões sinápticas e à dimensão extra de interações neurais (CHURCHLAND; SEJNOWSKI, 1992).

III. Redes recorrentes

Uma rede neural recorrente distingue-se de uma rede neural *feedforward* na medida em que tem pelo menos um *loop* de *feedback*. Por exemplo, uma rede recorrente pode consistir em uma única camada de neurônios, com cada neurônio alimentando seu sinal de saída de volta para as entradas de todos os outros neurônios (Figura 17). Na estrutura representada não há *loops* de *auto-feedback* na rede; *auto-feedback* refere-se a uma situação em que a saída de um

neurônio é realimentada em sua própria entrada. A rede recorrente também não possui neurônios ocultos.

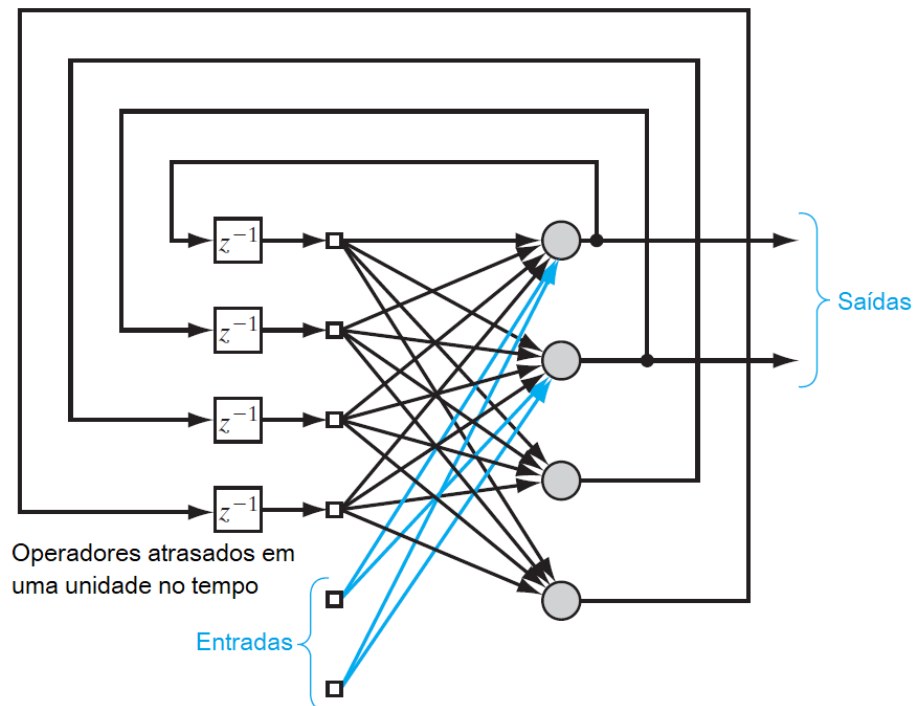
Figura 17 - Rede recorrente sem *loops* de auto-*feedback* e sem neurônios escondidos.



Fonte: Adaptado de Haykin (2009).

Um outro tipo de rede recorrente, agora com neurônios ocultos, é mostrado na Figura 18. As conexões de realimentação se originam dos neurônios ocultos, bem como dos neurônios de saída. A presença de *loops* de realimentação, seja na estrutura recorrente da Figura 17 ou na da Figura 18, tem um impacto profundo na capacidade de aprendizagem da rede e no seu desempenho. Além disso, os circuitos de realimentação envolvem o uso de ramificações específicas compostas de elementos de atraso de tempo unitário (representado por z^{-1}), que resultam em um comportamento dinâmico não-linear, assumindo que a rede neural contém unidades não-lineares.

Figura 18 - Rede recorrente com neurônios escondidos.



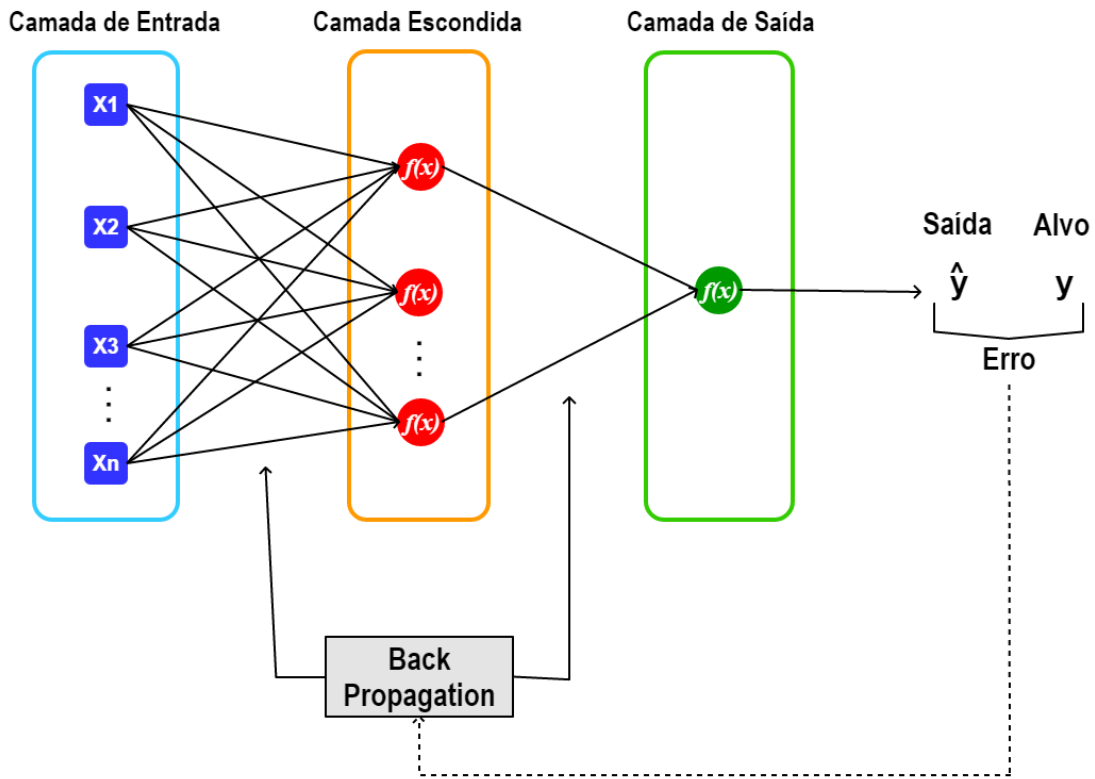
Fonte: Adaptado de Haykin (2009).

2.4.1.1.1. Multilayer Perceptron (MPL) / Algoritmo Back-propagation (BPA)

O algoritmo *Multilayer Perceptron* (MLP) consiste de um tipo de rede neural progressiva composta por uma camada de entrada, uma camada de saída, além de pelo menos, uma camada de neurônios intermediários, denominada camada escondida, que é responsável pela não linearidade das redes MLP permitindo que elas sejam capazes de resolver problemas complexos (DOAN; LIONG, 2004).

O treinamento das redes MLP requer o uso de um algoritmo que permita estabelecer um conjunto ótimo de pesos para a rede. O algoritmo de aprendizagem mais amplamente utilizado para tarefas separáveis de não linearidade é o algoritmo de retro propagação (*back-propagation*) do erro (RUMELHART et al., 1986; JANI et al., 2017). O problema encontrado durante o treinamento de uma MLP é que com a inclusão de uma ou mais camadas intermediárias, o erro dessas camadas é desconhecido, mas mesmo assim é necessário ajustar os pesos. Mas esse problema pode ser tratado com o BPA (Figura 19), que realiza uma propagação recursiva dos erros.

Figura 19 - Esquema MLP com BPA.



Fonte: Elaborado pelo autor do trabalho.

Visto que o aprendizado é supervisionado, o objetivo do BPA é ajustar os pesos tal que o erro entre a saída desejada e a calculada seja minimizado. Este processo ocorre a partir do erro entre os pares de amostras de entrada e saída dos dados da rede, consistindo de duas fases: progressiva (*forward*) e recursiva (*backward*). A fase progressiva é usada para encontrar uma saída a partir dos valores de entrada de um determinado padrão, ou seja, os sinais são propagados no sentido progressivo (da camada de entrada para a camada de saída), encontrando a saída do sinal e o erro, mas mantendo os pesos fixos. A fase recursiva compara esta saída com a saída desejada e a atualiza recursivamente (da camada de saída para a camada de entrada) os valores dos pesos das conexões dos neurônios da estrutura (RUMELHART et al., 1986; HAYKIN, 2009). O erro é determinado de acordo com a regra de ajuste dos pesos (regra delta generalizada) (HINTON et al., 2006; XIE et al., 2016):

$$E = 1/2 \sum_{i=1}^p \sum_{j=1}^n (d_j^i - y_j^i)^2 \quad (15)$$

onde p é o número de padrões usados no treinamento, n é o número de saídas, d é a saída desejada, e y é a saída estimada. A atualização do peso após o cálculo do erro E é dada pela seguinte expressão:

$$W = W - \alpha \frac{\partial E}{\partial W} \quad (16)$$

onde α representa a taxa de treinamento.

Os pesos da rede são configurados aleatoriamente antes do início do aprendizado e esta característica importante deve ser considerada visto que ela muda a maneira como o modelo aprende (MULLER; GUIDO, 2016), ou seja, a mesma arquitetura pode gerar modelos bastante diferentes entre si usando os mesmos parâmetros, uma vez que empregam diferentes sementes randômicas.

O BPA pode ser aplicado a qualquer rede que usa uma função de ativação diferenciável e aprendizagem supervisionada, sendo baseada no gradiente descendente. O processo de treinamento é realizado baseado em um método de tentativa e erro, de tal sorte que um dos problemas do algoritmo é seu tempo de treinamento. A taxa de atualização dos pesos é o que influencia este tempo. Se a taxa configurada é muito pequena, a rede consome muito tempo para o treinamento. Por outro lado, se a taxa é alta, a rede pode convergir em um intervalo de tempo curto, mas quando uma outra entrada é apresentada, a rede pode se tornar instável, obtendo resultados não confiáveis. Por esta razão, este trabalho experimentou dois diferentes algoritmos para otimização do BPA: Levenberg-Marquardt e regulação Bayesiana.

2.4.1.1.2. Algoritmo de Levenberg-Marquardt

Enquanto o BPA padrão usa o gradiente descendente como método de aproximação do mínimo da função de erro, o algoritmo Levenberg-Marquardt (LM) usa a aproximação de Newton (HAGAN; MENHAJ, 1994). Esta aproximação é obtida a partir da modificação do método de Gauss-Newton, introduzindo-se o parâmetro μ , de acordo com a equação a seguir:

$$W_{n+1} = W_n - (J_n^T J_n + \mu I)^{-1} J_n^T E_n \quad (17)$$

onde W é o peso, I é uma matriz identidade, E é o erro e J é a matriz Jacobiana. O parâmetro μ é um escalar não negativo multiplicado por um fator (β) sempre que um passo resultar em um

aumento na função erro, o qual deve ser minimizado. Quando um passo resulta em uma diminuição na função de desempenho, μ é dividido pelo fator β . Isto significa que se há uma convergência para o mínimo da função, μ é pequeno e o algoritmo se aproxima do método de Newton (passo $1/\mu$); no caso de não haver convergência, o método se aproxima do gradiente descendente. Assim, o algoritmo Levenberg-Marquardt combina as melhores características de dois métodos numéricos. O parâmetro μ atua como um fator de estabilização do treinamento, ajustando a aproximação para usar a rápida convergência do método de Newton e evitar passos muito grandes que possam levar a um erro de convergência (HAGAN; MENHAJ, 1994). O método de Newton é mais rápido e mais acurado para alcançar o erro mínimo, tal que o objetivo do algoritmo Levenberg-Marquardt passar para o método de Newton o mais rápido possível (DEMUTH; BEALE, 2009).

2.4.1.1.3. Algoritmo de Regulação Bayesiana

A regularização é um termo genérico para métodos envolvendo a modificação da função de desempenho de redes neurais, que é usualmente a soma dos quadrados dos erros de treinamento, e objetiva aperfeiçoar sua capacidade de generalização (MACKAY, 1992). O principal problema com a implementação da regularização é definir os valores corretos para os parâmetros da função objetivo. A estrutura bayesiana para redes neurais é baseada na interpretação probabilística dos parâmetros da rede. Em outras palavras, enquanto no treinamento de uma rede convencional o conjunto ideal de pesos é escolhido pela minimização de uma função de erro, a abordagem bayesiana envolve uma distribuição de probabilidade dos pesos da rede. Como resultado, as previsões da rede também são uma distribuição de probabilidade (XU et al., 2006; KAYRI, 2016).

A regularização inicializa uma distribuição anterior sobre o conjunto de parâmetros do vetor peso w , incentivando os pesos a serem tão pequenos quanto possível até que sejam suportados pelos dados. Este método efetivamente proíbe valores de parâmetro muito distantes da origem e, portanto, reduz a variância (DEMYANOV, 2015).

O objetivo de treinar uma rede neural é obter um conjunto de pesos e vieses que minimize o erro entre a saída desejada e a saída estimada pela rede. A função objetivo F mais usual é o erro médio quadrático (E_D):

$$F = E_D(\mathbf{D}|\mathbf{w}, \mathbf{M}) = \frac{1}{n} \sum_{i=1}^n (\mathbf{d}_i - \mathbf{y}_i)^2 \quad (18)$$

onde D representa o conjunto de dados de treinamento do par entrada-saída, w é o vetor dos pesos da rede, M é a arquitetura do modelo usado da rede neural, n é o número de amostras de treinamento, d_i representa a saída desejada e y_i é a saída da rede. Entretanto, esta função objetivo clássica pode frequentemente superajustar os dados de treinamento, condição conhecida como *overfitting*. Se ocorrer o *overfitting* da rede para os dados de treinamento, a rede certamente terá um fraco desempenho sobre os dados de teste (desconhecidos). Para melhorar a habilidade de generalização da rede neural e fazê-la ter um bom desempenho sobre dados desconhecidos, adiciona-se um termo extra à função objetivo, denominado termo de decaimento do peso. Com esta função de decaimento do peso os pesos da rede podem ser menores, o que melhora sua habilidade de generalização. A nova função objetivo torna-se, então:

$$F = \alpha E_w(D|w, M) + \beta E_D(w|M) \quad (19)$$

onde E_w é a função de decaimento do peso e α é o coeficiente de decaimento do peso. α e β são denominados como parâmetros de regularização ou hiperparâmetros, e representam a importância relativa do termo de decaimento do peso. Se $\alpha \ll \beta$, então o algoritmo de treinamento será conduzido para os menores erros. Se $\alpha \gg \beta$, o treinamento enfatizará a redução do tamanho do peso em detrimento de um bom ajuste, produzindo assim uma resposta mais suave da rede. Se as estimativas Bayes de α são grandes, as densidades posteriores dos pesos são altamente concentradas em torno de zero, de modo que os pesos efetivamente desaparecem e o modelo desconta conexões na rede. Portanto, α e β são adaptativamente previstos para lidar com a complexidade do modelo e a qualidade do ajuste (OKUT, 2016). Há muitos tipos de termos de decaimento do peso, sendo o mais usual a função quadrática de decaimento do peso:

$$E_w(w|M) = \frac{1}{m} \sum_{j=1}^m [w]_j^2 \quad (20)$$

onde $[\cdot]_j$ significa o j -ésimo componente do vetor peso e m é o número de todos os pesos da rede.

Os parâmetros ótimos de regularização α e β podem ser determinados pela técnica bayesiana. Na teoria bayesiana os pesos da rede podem ser considerados como variáveis aleatórias. Para um determinado conjunto de dados, α e β , a probabilidade posterior é dada por:

$$P(\mathbf{w}|\mathbf{D}, \alpha, \beta, \mathbf{M}) = \frac{P(\mathbf{D}|\mathbf{w}, \beta, \mathbf{M}) P(\mathbf{w}|\alpha, \mathbf{M})}{P(\mathbf{D}|\alpha, \beta, \mathbf{M})} \quad (21)$$

onde $P(\mathbf{D}|\mathbf{w}, \beta, \mathbf{M})$ é a função de verossimilhança de \mathbf{w} , que corresponde à probabilidade do dado ocorrer para determinados pesos \mathbf{w} . $P(\mathbf{w}|\alpha, \mathbf{M})$ é a densidade a priori (distribuição prévia dos pesos). $P(\mathbf{D}|\alpha, \beta, \mathbf{M})$ é o fator de normalização, também conhecido como evidência, e que garante que a probabilidade total seja 1. O fator de normalização não depende de \mathbf{w} (KUMAR, 2004)

$$P(\mathbf{D}|\alpha, \beta, \mathbf{M}) = \int P(\mathbf{D}|\mathbf{w}, \beta, \mathbf{M}) P(\mathbf{w}|\alpha, \mathbf{M}) d\mathbf{w} \quad (22)$$

Os pesos \mathbf{w} foram assumidos serem identicamente distribuídos, cada um seguindo uma distribuição gaussiana ($\mathbf{w}|\alpha, \mathbf{M}) \sim N(0, \alpha^{-1})$, onde N denota uma distribuição normal. Se a distribuição de probabilidade dos dados e pesos for considerada gaussiana, então as densidades de probabilidade podem ser escritas como:

$$P(\mathbf{D}|\mathbf{w}, \beta, \mathbf{M}) = \frac{1}{Z_D(\beta)} \exp(-\beta E_D) \quad (23)$$

$$P(\mathbf{w}|\alpha, \mathbf{M}) = \frac{1}{Z_w(\alpha)} \exp(-\alpha E_w) \quad (24)$$

onde $Z_D(\beta) = (\pi/\beta)^{n/2}$ e $Z_w(\alpha) = (\pi/\alpha)^{m/2}$. Z_D e Z_w são fatores de normalização. Substituindo (23) e (24) em (21), obtemos:

$$P(\mathbf{w}|\mathbf{D}, \alpha, \beta, \mathbf{M}) = \frac{\frac{1}{Z_w(\alpha)Z_D(\beta)} \exp(-(\alpha E_w + \beta E_D))}{P(\mathbf{D}|\alpha, \beta, \mathbf{M})} = \frac{1}{Z_F(\alpha, \beta)} \exp(-F(\mathbf{w})) \quad (25)$$

onde $Z_F(\alpha, \beta) = Z_D(\beta) Z_w(\alpha) P(\mathbf{D}|\alpha, \beta, \mathbf{M})$ é um fator de normalização.

Na teoria bayesiana maximizar a probabilidade posterior $P(\mathbf{w}|\mathbf{D}, \alpha, \beta, \mathbf{M})$ é equivalente a minimizar a função objetivo regularizada $F(\mathbf{w})$. Então os parâmetros de regularização podem ser otimizados através da regra de Bayes:

$$P(\alpha, \beta | D, M) = \frac{P(D|\alpha, \beta, M)P(\alpha, \beta | M)}{P(D|M)} \quad (26)$$

Considerando a densidade a priori, $P(\alpha, \beta | M)$, como sendo uniforme para os parâmetros de regularização α e β , então a maximização do termo posterior pode ser alcançada maximizando-se a função de probabilidade $P(D|\alpha, \beta, M)$, que é o fator de normalização em (21). Como foi considerado que todas as probabilidades adotam a forma gaussiana, a forma da densidade posterior de (18) é mostrada em (20). Assim, (18) pode ser resolvida para o fator de normalização (FORESEE, HAGAN, 1997).

$$\begin{aligned} P(D|\alpha, \beta, M) &= \frac{P(D|w, \beta, M)P(w|\alpha, M)}{P(w|D, \alpha, \beta, M)} = \frac{\left[\frac{1}{Z_D(\beta)} \exp(-\beta E_D) \right] \left[\frac{1}{Z_w(\alpha)} \exp(-\alpha E_w) \right]}{\frac{1}{Z_F(\alpha, \beta)} \exp(-F(w))} \\ &= \frac{Z_F(\alpha, \beta)}{Z_D(\beta)Z_w(\alpha)} \cdot \frac{\exp(-\beta E_D - \alpha E_w)}{\exp(-F(w))} = \frac{Z_F(\alpha, \beta)}{Z_D(\beta)Z_w(\alpha)} \end{aligned} \quad (27)$$

Se α é pequeno, a grande liberdade no intervalo a priori de valores possíveis de w é automaticamente penalizada pelo grande valor de Z_w ; modelos que se ajustam bem aos dados alcançam um grande valor de Z_F . O valor ótimo de α atinge um compromisso entre ajustar bem os dados e ser um modelo simples (MACKAY, 1992). As constantes $Z_D(\beta)$ e $Z_w(\alpha)$ são conhecidas de (19). A única parte não conhecida é $Z_F(\alpha, \beta)$, mas que pode ser estimada através da expansão em séries de Taylor. Visto que a função objetivo tem a forma quadrática em uma área pequena em torno do valor mais provável (MP), então $F(w)$ pode ser expandida em torno do valor mais provável da densidade posterior de w^{MP} , onde o gradiente é zero (MACKAY, 1992). Resolvendo para a constante de normalização, obtém-se

$$Z_F \approx (2\pi)^{m/2} (\det(\mathbf{H}^{\text{MP}}))^{-1/2} \exp(-F(w^{\text{MP}})) \quad (28)$$

onde \mathbf{H} é a matriz hessiana da função objetivo, sendo definida como:

$$\mathbf{H} = \alpha \nabla^2 E_w + \beta \nabla^2 E_D \quad (29)$$

Substituindo este resultado em (27), pode-se obter os valores ótimos para α e β tomando a derivada com relação a cada um dos logaritmos de (27) e igualando a zero. Durante o processo a matriz hessiana precisa ser calculada. O cálculo da matriz hessiana pode ser realizado através da aproximação de Gauss-Newton, igual como utilizado no algoritmo de Levenberg-Marquardt na seção anterior. Então obtém-se:

$$\alpha^{\text{MP}} = \frac{\gamma}{2E_w(\mathbf{w}^{\text{MP}})}, \quad \beta^{\text{MP}} = \frac{n-\gamma}{2E_D(\mathbf{w}^{\text{MP}})} \quad (30)$$

onde $\gamma = N - 2\alpha \text{Traço}(\mathbf{H}^{\text{MP}})^{-1}$ (FORESEE; HAGAN, 1997). O termo “Traço” refere-se à diagonal principal de uma matriz. O parâmetro γ pode variar entre 0 e N , sendo denominado de número efetivo de parâmetros; e consiste em uma medida de quantos parâmetros na rede neural são efetivamente usados na redução da função de erro. N é o número total de parâmetros da rede.

2.4.1.2. Deep Learning e Trabalhos Relacionados

No início dos anos 2000 o poder computacional aumentou vertiginosamente, o que permitiu uma verdadeira explosão de novas técnicas computacionais que até então eram praticamente impossíveis de serem utilizadas. Um dos problemas das redes neurais tradicionais até pouco tempo atrás era o custo computacional para treinar modelos grandes, o qual foi resolvido quando Hinton et al. (2006) mostraram que a rede *deep belief* pode ser eficientemente treinada utilizando um algoritmo denominado *greedy layer-wise pretraining*. A *deep belief* é uma rede neural profunda, composta de múltiplas camadas de unidades escondidas, com conexões entre as camadas, mas não entre unidades dentro de cada camada. Assim, nos desenvolvimentos seguintes relacionados com o tema, os pesquisadores passaram a ser capazes de treinar de forma eficiente redes neurais complexas, que não mais se limitavam a uma única camada. Essas novas estruturas mostraram sistematicamente melhores resultados, assim como aperfeiçoamento da capacidade de generalização. Neste contexto, surgiu o que passou a ser chamado *deep learning* (DL) ou aprendizado profundo que nada mais é do que um ramo do aprendizado de máquina baseado em um conjunto de algoritmos que tentam modelar abstrações de alto nível de dados usando uma estrutura do tipo grafo profundo com inúmeras camadas de processamento, as quais são constituídas de várias transformações lineares e, principalmente, não lineares (GOODFELLOW et al., 2016). Grafo é um termo usualmente empregado na ciência da computação para se referir a um conjunto de nós conectados dois a dois. Enquanto

as abstrações de alto nível de dados referem-se ao nível de visão dos dados em um banco de dados, em que neste caso, apenas parte dos dados é descrita.

A medida em que os modelos de DL foram sendo amplamente utilizados pelas aplicações das ciências da computação, como reconhecimento de imagens (KRIZHEVSKY et al., 2012), nos últimos anos os benefícios da DL têm sido amplamente utilizados nas mais diferentes áreas, inclusive para aplicações relacionadas à energia (WANG et al., 2016; COELHO et al., 2017; FAN et al., 2017; WANG et al., 2017; KONG et al., 2017; FENG et al., 2017). Com relação a este campo, a previsão de energia eólica é indiscutivelmente uma das aplicações que mais se utilizou desta ferramenta. Wang et al. (2016) mostraram que o uso de uma rede do tipo *deep belief* e regressão quantílica, podem melhorar a previsão probabilística da velocidade do vento. Feng et al. (2017) propuseram um algoritmo de seleção profunda de características que, em combinação com uma estrutura de multi-modelo, melhorou a acurácia de previsão do vento em 30%. Ainda nesta mesma linha de pesquisa, Wang et al. (2017) propuseram um *ensemble* de redes neurais convolucionais para obter uma probabilidade mais acurada da previsão de geração de energia eólica.

Além das aplicações de DL relacionadas com energia eólica, a DL também se mostrou exitosa em diversos casos relacionados a outros campos de energia. No contexto de previsão de carga, por exemplo, Fan et al. (2017) propuseram um *autoencoder* profundo em combinação com modelo *extreme gradient booster* (XGB) que mostrou como eles puderam prever de forma mais acurada a carga de resfriamento do prédio do que utilizando técnicas alternativas. Um *autoencoder* é um tipo de rede neural artificial usada para aprender códigos de dados eficientes de maneira não supervisionada. O XGB é um algoritmo que consiste da implementação de árvores de decisão aprimoradas por gradientes projetadas para velocidade e desempenho. Ainda neste mesmo trabalho foi proposta uma rede neural profunda para previsão mais acurada da carga de resfriamento de um prédio, para o qual realizaram a pesquisa. Coelho et al. (2017) propuseram uma estratégia com DL para previsão de séries temporais e mostraram como esta técnica pode ser usada com sucesso para a previsão do consumo de eletricidade em residências.

O modelo básico de DL é uma extensão natural do tradicional MLP que usa múltiplas camadas escondidas (GOODFELLOW et al., 2016). Quando comparada com a MLP padrão, a DL requer mudanças específicas do modelo para que seja eficientemente treinado, ou seja, funções de ativação diferentes da sigmoide padrão.

Há muitas arquiteturas de DL que podem extrair automaticamente características complexas a partir de dados de entrada sem intervenção humana. As redes neurais recorrentes (RNR) são um tipo de rede ligeiramente mais complexas do que as DL básicas, e que constroem

mapeamentos adicionais para manter as informações relevantes a partir das entradas do passado e que são confiáveis para modelar dados de séries temporais (GOODFELLOW et al., 2016). Os dois principais tipos de redes recorrentes são a LSTM (*Long-Short Term Memory*) (HOCHREITER; SCHMIDHUBER, 1997) e a GRU (*Gated Recurrent Unit*) (CHUNG et al., 2014). Diferentemente das RNA padrões, a LSTM e a GRU são capazes de modelar um comportamento seletivo de “esquecimento-lembrança”. Ambas as estruturas são muito semelhantes, mas a GRU tem uma estrutura mais simples e é mais rápida para treinar.

Um tipo diferente de DL são as redes neurais convolucionais (RNC), as quais são modeladas usando três blocos de construção: uma operação de convolução, uma operação de agrupamento, e uma camada totalmente conectada. Para uma determinada matriz de dados, a operação de convolução passa um filtro através da matriz de dados e calcula o produto cruzado local por elemento entre o filtro e os dados. Quando filtros diferentes capturam propriedades diferentes, as RNC tipicamente usam vários filtros para obter diferentes matrizes de dados conhecidas como mapas de características. Em um passo subsequente, a operação de agrupamento reduz o tamanho desses mapas de características, reduzindo grandes áreas em valores únicos. Após diversas operações de convolução e agrupamento, os valores dos últimos mapas de características são usados como entradas para uma rede totalmente conectada (LAGO et al., 2018).

Uma outra estrutura de DL são as redes neurais polinomiais do tipo GMDH (*Group Method of Data Handling*). Visto que este método não requer conhecimento prévio do especialista para selecionar as variáveis mais relevantes para uso na camada de entrada do modelo, ele é bastante apropriado para aplicações práticas. Os algoritmos multicamadas GMDH podem ser utilizados para resolver problemas de modelagem quando o número de observações é inferior ao número de variáveis (ONWUBOLU, 2015). A descrição mais detalhada do método será apropriadamente apresentada na seção seguinte.

2.4.1.2.1. *Group Method of Data Handling (GMDH)*

Na última década, o uso de técnicas de inteligência computacional para modelagem e previsão de carga em sistemas de potência tem aumentado bastante (AHMAD et al., 2014) e um dos métodos mais promissores cujos resultados têm mostrado animadores é o GMDH. Esta abordagem é geralmente aplicada para previsão de carga em curto-prazo (ABDEL-AAL, 2005, 2006; ELATTAR et al., 2012; ZJAVKA; SNÁSEL, 2016) ou modelagem e previsão de consumo de eletricidade (SRINIVASAN, 2008; XIAO et al., 2015). A literatura especializada também traz muitas outras aplicações para diagnóstico e previsão de faltas em sistemas elétricos

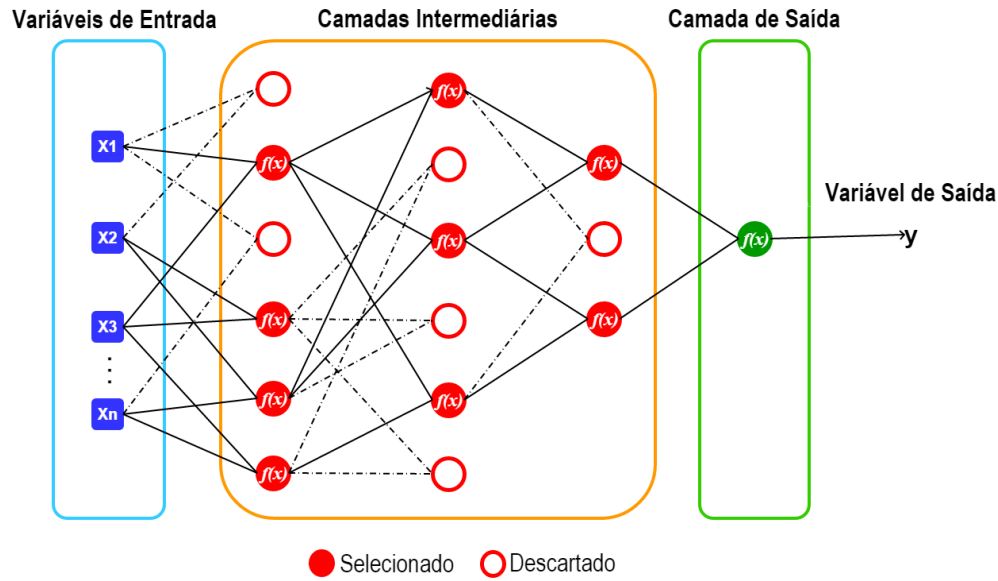
(LI et al., 2009; WITCZAK et al., 2006), assim como previsão de chuva (CHENG; YANGJUN-HE, 2016) e vazão (IKEDA et al., 1976), mas este método ainda não foi suficientemente explorado e aplicado para a modelagem preditiva do potencial de geração de energia hidrelétrica, especialmente na região Amazônica.

O pioneiro da técnica GMDH foi Alexey. G. Ivakhnenko (IVAKHNENKO, 1968), mas muitos outros pesquisadores realizaram contribuições subsequentes e posteriores desenvolvimentos igualmente significativos. O GMDH consiste de uma abordagem auto organizada e indutiva para modelagem matemática de sistemas complexos (FARLOW, 1984). Este método é uma técnica versátil e robusta, e também é considerada para resolver problemas de inteligência artificial como identificação, previsão de curto e longo prazo de processos aleatórios, e reconhecimento de padrões em sistemas complexos (FARLOW, 1981; IVAKHNENKO; IVAKHNENKO, 1995; YANGJUN-HE, 2016).

As características apresentadas pelo GMDH e sua diversidade de usos são aspectos que levaram ao emprego deste algoritmo para a sua aplicação nesta metodologia. Além disso, a seleção das variáveis de entrada mais relevantes, graças a sua habilidade de evitar problemas com *overfitting*, e a escolha objetiva do modelo mais ajustado, são as vantagens que favoreceram a sua escolha.

O conceito fundamental da abordagem GMDH é estruturar uma rede neural multicamadas *feedforward* usando modelos parciais conectados hierarquicamente, tal que os melhores modelos possam ser selecionados através do método apropriado, e os modelos de baixo desempenho sejam descartados (DAG; YOZGATLIGIL, 2016; MRUGALSKI, 2013). A Figura 20 mostra a estrutura esquemática de uma rede genérica do tipo GMDH, onde cada camada consiste de pares de variáveis de entrada e uma variável de saída.

Figura 20 - Arquitetura geral para uma rede polinomial do tipo GMDH.



Fonte: Elaborado pelo autor do trabalho.

A formulação matemática que resolve as conexões entre as variáveis de entrada e saída pode ser explicada pela teoria da função polinomial, conhecida por polinômios de Kolmogorov-Gabor (IVAKHNENKO, 1971; XU et al., 2012):

$$y = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n \sum_{j=1}^n a_i a_j x_i x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n a_i a_j a_k x_i x_j x_k + \dots \quad (31)$$

onde $X = (x_1, x_2, \dots, x_n)$ é o vetor de entrada e y é a variável de saída. Desta maneira, uma descrição completa pode ser simplificada por meio de um sistema de polinômios parciais quadráticos com apenas duas variáveis, da seguinte forma:

$$\hat{y} = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1^2 + w_4 x_1 x_2 + w_5 x_2^2 \quad (32)$$

onde x_1 e x_2 são as variáveis de entrada, \hat{y} é a variável de saída (alvo), e w_0, \dots, w_5 são os parâmetros (pesos).

Para cada conjunto de dados de aprendizagem que inclui a variável dependente y , assim como as variáveis independentes x_1, x_2, \dots, x_n , a amostra de dados é separada em subconjuntos

de treinamento e teste. A partir de n variáveis de entrada, são geradas combinações usando duas a duas variáveis para todas as variáveis de entrada, de modo a formar pares que constituem as unidades que integrarão a primeira camada. Para todas as unidades são estimados alguns pesos usando o subconjunto de treinamento.

Um critério é definido como métrica de comparação entre os modelos parciais e o resultado esperado de maneira a selecionar os melhores modelos parciais, descartando aqueles que não alcançaram o critério de escolha. Os modelos parciais selecionados tornam-se parte da primeira camada e as previsões das unidades da primeira camada são configuradas como novas variáveis de entrada que alimentarão a camada seguinte e assim se constrói uma estrutura multicamadas aplicando-se os passos anteriores. Quando o critério de seleção dos modelos parciais não é mais alcançado, a adição de novas camadas é finalizada e a unidade de menor erro na camada de ordem mais alta é escolhida como modelo final.

A estrutura do método GMDH nesta pesquisa utilizou o algoritmo de iteração multicamadas (*Multilayered Iteration Algorithm – MIA*) com funções de transferência polinomiais. O MIA foi considerado apropriado para tratar o problema porque permite trabalhar com um grande número de variáveis que por sua vez requer a análise de um número elevado de combinações de modelos parciais (KORDÍK, 2009).

O MIA pode combinar quatro tipos de funções polinomiais de transferência, como mostrado abaixo:

$$\text{- linear: } \mathbf{y} = \mathbf{w}_0 + \mathbf{w}_1\mathbf{x}_1 + \mathbf{w}_2\mathbf{x}_2 \quad (33)$$

$$\text{- linear-covariante: } \mathbf{y} = \mathbf{w}_0 + \mathbf{x}_1(\mathbf{w}_1 + \mathbf{w}_3\mathbf{x}_2) + \mathbf{w}_2\mathbf{x}_2 \quad (34)$$

$$\text{- quadrática: } \mathbf{y} = \mathbf{w}_0 + \mathbf{x}_1(\mathbf{w}_1 + \mathbf{w}_3\mathbf{x}_2 + \mathbf{w}_4\mathbf{x}_1) + \mathbf{x}_2(\mathbf{w}_2 + \mathbf{w}_5\mathbf{x}_2) \quad (35)$$

$$\begin{aligned} \text{- cúbica: } \mathbf{y} = & \mathbf{w}_0 + \mathbf{w}_1\mathbf{x}_1 + \mathbf{w}_2\mathbf{x}_2 + \mathbf{w}_3\mathbf{x}_1\mathbf{x}_2 + \mathbf{w}_4\mathbf{x}_1^2 + \mathbf{w}_5\mathbf{x}_2^2 + \mathbf{w}_6\mathbf{x}_1\mathbf{x}_1^2 + \\ & + \mathbf{w}_7\mathbf{x}_1^2\mathbf{x}_2 + \mathbf{w}_8\mathbf{x}_1\mathbf{x}_2^2 + \mathbf{w}_9\mathbf{x}_2\mathbf{x}_2^2 \end{aligned} \quad (36)$$

CAPÍTULO III

3. DADOS E METODOLOGIA

O presente capítulo descreve os dados primários utilizados na tese, o tratamento e processamento realizado sobre eles e a metodologia desenvolvida.

3.1. DADOS PRIMÁRIOS DE RESUMO DE DESCARGA

Para determinação sistemática da vazão na estação hidrológica de Jatobá, cujo código adotado pela Agência Nacional de Águas (ANA) é 17650000, foram empregados dados de resumo de descarga que estão disponíveis em formato tabular e podem ser obtidos através do sistema Hidroweb (<http://hidroweb.ana.gov.br/>) da ANA. O resumo de descarga da estação Jatobá compreende medições realizadas em diferentes períodos do ano, entre 1972 e 1999. As informações disponíveis continham dados tanto de consistência 1, quanto de consistência 2. Os dados de consistência 1 referem-se aos dados brutos, sem qualquer tipo de análise ou correção. Os dados de consistência 2 sofreram análise de consistência, conforme preconiza a ANA através das diretrizes recomendadas para consistência de dados fluviométricos no âmbito da Resolução Conjunta ANA e ANEEL nº 3, de 10 de agosto de 2010.

O resumo de descarga traz as seguintes informações: código da estação, nível de consistência do dado, data e hora da medição, número ordinário da medição, cota (em metros), vazão (em metros cúbicos por segundo), área molhada (em quilômetros quadrados), largura da seção (em metros), velocidade média (em metros por segundo) e profundidade (em metros).

Para utilização na tese foram empregados apenas os dados de consistência 2.

3.2. DADOS PRIMÁRIOS DE NÍVEL FLUVIOMÉTRICO

Para processamento da série de vazão, também foram utilizados dados diários de cotas, ou seja, níveis fluviométricos na estação Jatobá, também disponíveis através do portal Hidroweb. Os dados diários de nível fluviométrico apresentam duas observações, a primeira realizada às 07:00 h e a segunda às 17:00 h. As observações de nível fluviométrico em Jatobá foram realizadas através de régua linimétrica. Os registros fluviométricos da estação de Jatobá são disponibilizados em formato tabular e contêm as seguintes informações: código da estação;

nível de consistência; data e hora; média diária da cota; tipo da medição (escala, linígrafo ou *data logger*); cotas máxima, mínima e média (mensais); dia de ocorrência das cotas máxima e mínima mensais; cota média anual; cotas diárias.

O período de cobertura dos registros de cota para Jatobá iniciou em 1972 e foi encerrado no ano de 2013, com a desativação da estação. Visto que os dados de consistência 2 cobriam somente até o mês de fevereiro de 2007, foram utilizados os dados de consistência 1 por cobrir maior período de informação.

3.3. DADOS PRIMÁRIOS DE PRECIPITAÇÃO OBSERVADA

Os registros diários de precipitação usados na tese são provenientes do *Climate Prediction Center / National Centers for Environmental Prediction* (CPC / NCEP), e estão disponíveis em ftp://ftp.cpc.ncep.noaa.gov/precip/CPC_UNI_PRCP/GAUGE_GLB/. Os dados de precipitação têm cobertura global para uma grade regular de $0,5^\circ$ latitude \times $0,5^\circ$ longitude (XIE et al., 2007; CHEN et al., 2008). Esta base de dados se apresenta no formato binário e consiste de uma combinação de dois tipos de dados: registros diários por pluviômetros (observações de superfície) e estimativas de precipitação por satélites usando canais em micro-ondas e infravermelho. A cobertura dos dados pluviométricos vai de 01 de janeiro de 1979 aos dias atuais.

3.4. DADOS PRIMÁRIOS DE PRECIPITAÇÃO SIMULADA

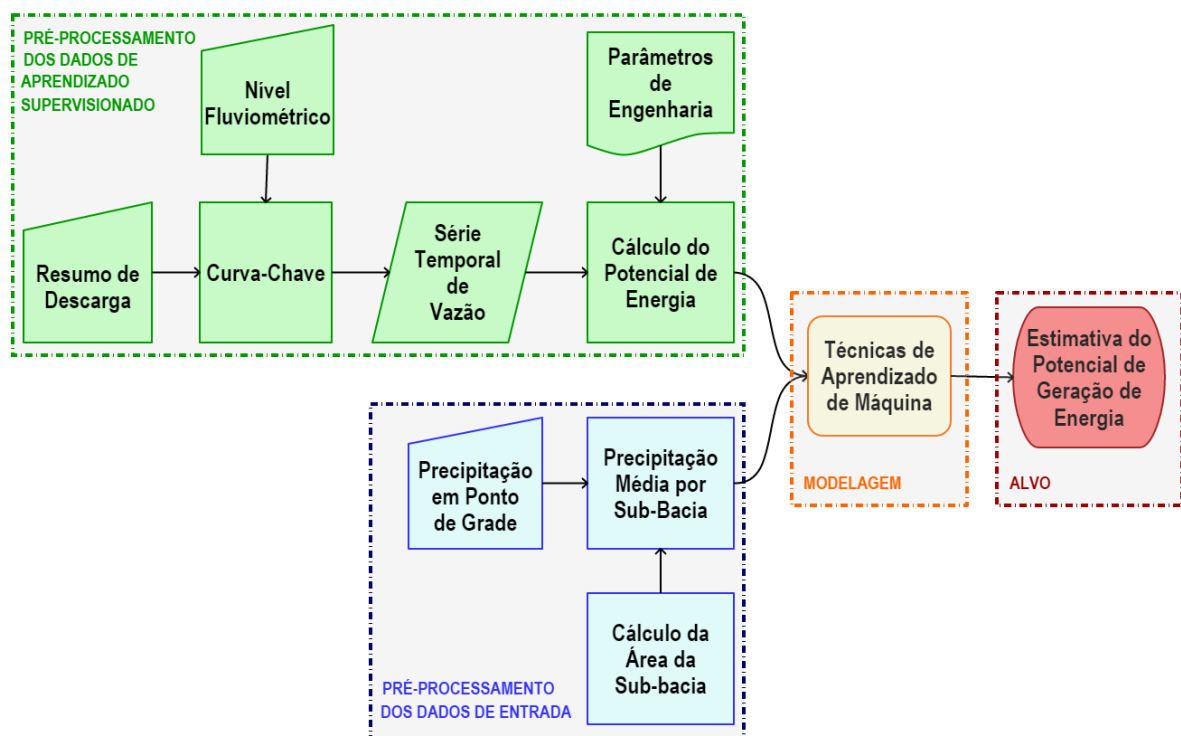
A tese também utilizou dados sintéticos de precipitação mensal através de modelagem climática com o RegCM4 (*Regional Climate Model version 4*). Os dados foram disponibilizados pelo Laboratório de Modelagem Climática da Faculdade de Meteorologia da Universidade Federal do Pará, que realizou simulações do clima para o período de janeiro de 1984 a dezembro de 2038. Os dados de precipitação mensal se apresentam em formato binário e estão disponíveis para uma grade regular de $0,5^\circ$ latitude \times $0,5^\circ$ longitude, cobrindo o domínio da Amazônia Legal.

3.5. FLUXO DE PROCESSOS

Na Figura 21 pode ser observado o fluxo geral dos processos para modelagem preditiva do potencial de geração de energia hidrelétrica. Inicialmente os dados de resumo de descarga e nível fluviométrico são processados para gerar uma curva-chave que fornece a série temporal de vazão. Os dados de vazão são convertidos em potencial de geração hidrelétrica por meio de equações matemáticas que empregam parâmetros do projeto da UHE. Os volumes de

precipitação em cada ponto de grade são ponderados pela área para se obter uma média de cada sub-bacia, gerando assim a série temporal mensal de precipitação por sub-bacia. E, finalmente, os dados de precipitação por sub-bacia juntamente com o potencial mensal de geração de energia hidrelétrica real são introduzidos em uma rede neural artificial (polinomial e tradicional) supervisionada para modelar o potencial mensal de geração de energia elétrica no AHE. A descrição em detalhe de cada passo individual do processo será apresentada nas seções seguintes.

Figura 21 - Processo para previsão do potencial de geração de energia hidrelétrica.



Fonte: Elaborado pelo autor do trabalho.

3.6. SÉRIE TEMPORAL DE VAZÃO

Geralmente, a ANA fornece juntamente com os dados de resumo de descarga, a curva-chave da seção do rio. Entretanto, neste caso particular, por não haver curva-chave oficial publicada para a estação Jatobá, foi necessário desenvolver uma equação para determinação da série temporal de vazão.

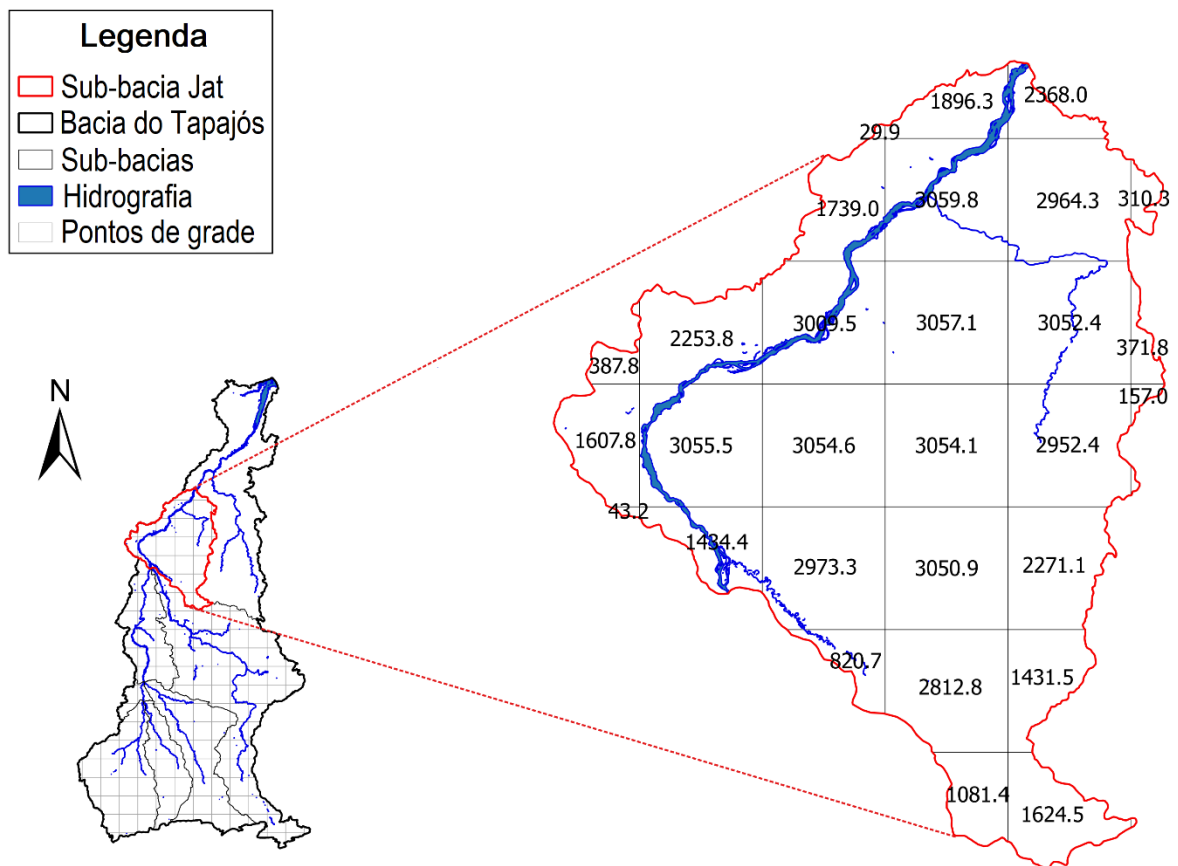
A elaboração da curva-chave para obter a série temporal de vazão foi realizada lançando-se os dados de resumo de descarga em uma planilha eletrônica de cálculo e, em seguida, o ajuste dos dados foi definido por meio de funções polinomiais e exponenciais, entre

outras. O coeficiente de determinação (R^2) foi adotado como critério para selecionar a melhor função de ajuste adaptada ao comportamento do rio na seção de medição. Assim, usando a curva-chave e os dados mensais de níveis fluviométricos, foi possível gerar a série temporal mensal de vazão média em Jatobá.

3.7. PRECIPITAÇÃO MÉDIA POR SUB-BACIA

Para o cálculo da precipitação média por sub-bacia, foi necessário determinar as áreas de cada sub-bacia hidrográfica e dos pontos de grade que as compõem. Este procedimento foi executado utilizando o *software* de geoprocessamento Quantum GIS (Figura 22). A área da sub-bacia 443 foi reduzida do tamanho original e ajustada para o ponto de instalação da barragem, descartando assim uma pequena área à jusante da barragem e cuja precipitação não contribui para a geração de energia. Assim, a nova forma adotada pela sub-bacia 443 considerou a topografia do terreno (divisores de água) e a rede de drenagem do sítio, passando a ser referida como sub-bacia JAT (443 ajustada).

Figura 22 - Cálculo da área para cada ponto de grade (em km²).



Fonte: Elaborado pelo autor do trabalho.

Em seguida, foram totalizados os dados diários de precipitação em ponto de grade para cada mês. E por fim, os dados mensais de precipitação em ponto de grade e as suas respectivas áreas foram aplicados em (8) para gerar a série temporal de precipitação mensal média para cada sub-bacia do rio Tapajós com influência no AHE Jatobá.

3.8. ANÁLISE DE TENDÊNCIA

Os dados de precipitação média mensal por sub-bacia foram submetidos à estatística de Mann-Kendall (MANN, 1945; KENDALL, 1975; LOPES et al., 2012) para um nível de confiança de 95%, considerando como hipótese h_0 testada a inexistência de tendência na série de dados. A formulação para a hipótese alternativa h_1 foi haver tendência monotônica, seja de aumento ou de diminuição da precipitação média para o mês investigado em uma determinada sub-bacia do rio Tapajós.

Os dois principais resultados obtidos com este teste referem-se, primeiramente, ao nível de significância da tendência porventura detectada, obtido através da comparação com o valor p . Caso o valor p seja maior que o nível de significância $\alpha = 0.05$, aceita-se a hipótese h_0 . Do contrário, a hipótese nula é rejeitada e admite-se que há tendência de aumento ou redução. Em segundo, o valor τ que representa a medida da força da tendência. Assim, τ assume um valor entre -1 e $+1$, onde o sinal indica a direção da tendência.

O teste MK foi aplicado à série temporal de precipitação mensal abrangendo o período de 1979 a 2018 para cada sub-bacia no rio Tapajós com influência no AHE Jatobá, de forma a avaliar a presença de tendência sistemática no comportamento da precipitação ao longo do tempo. E usando o teste de inclinação de Sen foram determinados a magnitude da tendência e os coeficientes (angular e linear) para as equações da reta de tendência linear nos casos em que houve redução estatisticamente significativa, conforme encontrado no teste MK.

3.9. SÉRIE DE POTENCIAL DE GERAÇÃO DE ENERGIA

Para geração da série de potencial mensal de geração de energia, necessária para o treinamento e validação das redes neurais, foram utilizados os dados da série de vazão mensal média e os parâmetros específicos de engenharia, obtidos a partir do projeto da usina, conforme disponibilizados pela Empresa de Pesquisas Energéticas (EPE). A queda hidráulica considerou 35 metros e a eficiência adotada do conjunto turbina-gerador foi de 95%. Tais valores de cada

mês foram substituídos em (3) para gerar a série temporal do potencial mensal médio de geração hidrelétrica abrangendo o período de 1984 a 2013.

3.10. MODELAGEM

A modelagem do potencial de geração de energia hidrelétrica foi realizada empregando duas técnicas de inteligência computacional: (i) *deep learning*, através do método GMDH, e (ii) redes neurais artificiais. As seções seguintes trazem as configurações adotadas para cada técnica empregada.

3.10.1. Previsão do Potencial de Geração de Energia

Como dados de entrada para os modelos preditivos desenvolvidos foram considerados os totais mensais de precipitação para as sete sub-bacias do rio Tapajós que influenciam o sítio de previsão, assim como o potencial mensal de geração de energia. Esta última variável foi empregada no aprendizado supervisionado das redes.

Visto que a bacia hidrográfica do Tapajós apresenta uma forma longitudinal, cujo fluxo dos rios tributários segue aproximadamente de Sul para Norte; e considerando que os volumes de água de pontos a diferentes distâncias percorrerão suas trajetórias em intervalos diferentes de tempo até o barramento; e portanto a precipitação que cai na bacia no ponto mais distante do barramento pode levar alguns meses para chegar no AHE, as variáveis de entrada do modelo, ou seja, a precipitação mensal média em cada sub-bacia, foram defasadas em até 12 meses. Assim, o número de variáveis de entrada do modelo totalizou 91 informações de precipitação [7 sub-bacias \times (mês corrente + 12 meses passados)].

As séries temporais foram divididas em dois subconjuntos adotando-se a técnica de validação *hold out* (HAYKIN, 2009). A primeira massa de dados correspondeu ao subconjunto de treinamento e validação, ficando a segunda para a realização do teste. Desta maneira, o subconjunto de treinamento correspondeu ao período de janeiro de 1984 a agosto de 1990, totalizando 80 pares de entrada-saída para otimizar os valores dos hiperparâmetros do modelo. O subconjunto de validação empregou 40 observações abrangendo o período de setembro de 1990 a dezembro de 1993 para otimizar a arquitetura do modelo. E por fim, o subconjunto de teste utilizou 240 observações considerando o período de janeiro de 1994 a dezembro de 2013 e foi usado para mensurar o desempenho do modelo usando um subconjunto de dados totalmente independente e desconhecido durante o processo de treinamento e validação.

3.10.1.1. Modelagem usando GMDH:

Dentre os principais hiperparâmetros de configuração da rede polinomial GMDH que foram testados tem-se: o critério de seleção dos modelos parciais; a técnica de validação empregada e o parâmetro alfa.

Há quatro tipos de configuração para os critérios de seleção: teste, viés, teste com viés, e viés com retreinamento. Na opção “teste” os modelos são comparados com base no erro do conjunto de teste. Na opção “viés” os modelos parciais são comparados com base no erro do viés do conjunto de teste. Na opção “teste com viés” a seleção utiliza um critério combinado através do erro do teste e viés. Na opção “viés com retreinamento” primeiramente os modelos são comparados com base no erro do viés, sendo então retreinados depois sobre o conjunto de treinamento e teste.

A técnica de validação consiste no método de separação do conjunto de dados para treinamento e teste, apresentando seis modos opcionais, conforme descritos a seguir:

- Modo 1: o conjunto de dados é particionado em vários subconjuntos de treinamento e teste, de tal maneira que o último conjunto de dados pertence ao subconjunto de teste.
- Modo 2: é semelhante ao anterior, porém o último subconjunto é de treinamento.
- Modo 3.1: os dados são divididos em subconjuntos de treinamento e teste na proporção de 1 subconjunto de teste para cada 3 de treinamento, sendo que a partição inicia com o subconjunto de treinamento e termina com teste.
- Modo 3.2: é análogo ao anterior, porém a proporção fica em 1 subconjunto de teste para 2 de treinamento, sendo que a partição inicia e termina com o subconjunto de treinamento.
- Modo 4.1: os dados de treinamento e teste são divididos na proporção de 4 de treinamento para cada 1 de teste, onde o modo inicia com o subconjunto de treinamento e finaliza com o subconjunto de teste.
- Modo 4.2: a proporção da distribuição dos subconjuntos é de 3 teste para 1 treinamento, iniciando e finalizando com dados do subconjunto de treinamento.

Quanto ao parâmetro alfa, este é um valor positivo que corresponde à força da regularização e reduz a variância das estimativas, conforme apresentado em (30). Quanto maior o valor, maior a regularização.

Os modelos parciais foram comparados baseado no viés e, então retreinados com base nos dados totais de treinamento. A decisão de cessar o treinamento foi adotada após comparar o erro relativo da camada de treinamento com o erro mínimo da camada. Quando a diferença

alcançou um valor inferior ao epsilon, então o treinamento terminou. O epsilon é o limiar de referência da condição de parada do treinamento, cujo valor adotado foi 0,001.

3.10.1.2. Modelagem usando RNA

Os modelos de redes neurais com BPA também foram estruturados utilizando o mesmo conjunto de 91 variáveis de entrada como descrito anteriormente. A função de ativação adotada foi a tangente hiperbólica. As arquiteturas testadas para a rede mostraram que foi suficiente para modelar o potencial de geração de energia hidrelétrica (BASHEER, 2000). Dois algoritmos foram testados para otimização do BPA: Levenberg-Marquardt e regulação bayesiana. As configurações foram testadas usando de 1 até 90 neurônios na camada escondida; com taxas de aprendizagem variando de 0,01 a 0,5; momentum de 0 a 0,9. Para evitar *overfitting* a fase de aprendizagem foi interrompida em 100 iterações. A aprendizagem foi parada quando o MSE passou a aumentar durante cinco iterações seguidas.

Todas as arquiteturas foram treinadas 20 vezes. A arquitetura que alcançou melhor desempenho baseado no MSE para o subconjunto de validação foi a escolhida. Após a escolha da melhor arquitetura, as simulações foram repetidas mais 80 vezes, totalizando 100 repetições para os subconjuntos de treinamento e validação. Os subconjuntos de teste também foram repetidos por 100 vezes, conforme metodologia proposta por Brownlee (2017).

3.10.1.3. Métricas de Avaliação dos Modelos

Avaliar a habilidade do modelo é fundamental para definir sua aplicação potencial. Para isto, as seguintes métricas foram usadas para quantificar o desempenho do modelo: coeficiente de correlação (R), erro médio absoluto (MAE) e erro médio absoluto percentual ($MAPE$).

3.10.1.3.1. Coeficiente de Correlação (R)

É uma medida que indica o grau de associação entre duas ou mais variáveis, cujo intervalo de ocorrência é limitado entre os valores -1 e $+1$ (CHARTTERJEE; HADI, 2012). A correlação é considerada fraca quando ela alcança valores abaixo de 0,6 e pode ser obtida por meio da formulação abaixo.

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \quad (24)$$

onde $x_i - \bar{x}$ é o desvio de cada observação x_i a partir da média da variável resposta; $y_i - \bar{y}$ é o desvio de cada observação y_i em relação à média da variável preditora; e n é o número de observações.

3.10.1.3.2. Erro Médio Absoluto (MAE)

Consiste de uma média aritmética dos valores absolutos dos desvios entre os membros de cada par. Esta é uma medida comum para avaliar a acurácia de predições para preditores contínuos (WILKS, 2006) e pode ser determinada usando a seguinte equação:

$$MAE = \frac{1}{n} \sum_{i=0}^{n-1} |x_i - y_i| \quad (25)$$

onde y_i é o valor estimado da i -ésima observação; x_i é o valor real correspondente; e n é o número de observações.

3.10.1.3.3. Erro Médio Absoluto Percentual (MAPE)

Consiste de uma medida do tamanho do erro em termos de percentagem e é usado como uma estimativa da acurácia de uma previsão.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{x_i} \right| \quad (26)$$

onde as variáveis são semelhantes as já descritas.

3.11. CENÁRIOS DE PRECIPITAÇÃO

Foram construídos dois cenários com horizonte de vinte anos futuros para avaliar o potencial de geração de energia hidrelétrica no AHE Jatobá ao longo do tempo, utilizando para tal o modelo GMDH. Os processos para gerar os dados de precipitação mensal que constituem a entrada do modelo GMDH são descritos nas seções seguintes.

3.11.1. Cenário 1: RCP4.5

Para construção do cenário RCP4.5 do IPCC, foram utilizadas as saídas do modelo RegCM4. Para rodar a simulação do clima foi realizado um *downscaling* dinâmico sob o domínio da Amazônia Legal, aninhando a grade do RegCM4 ao modelo global HadGEM2, cuja configuração detalhada para simulação encontra-se descrita em De Souza et al. (2016). Para

avaliar a habilidade do modelo climático em simular a precipitação mensal do clima observado para a Amazônia no período de janeiro de 1984 a dezembro de 2005, foram calculados os valores de precipitação média mensal por sub-bacia do rio Tapajós, conforme descrito na seção 3.7.

A partir dos dados de precipitação mensal por sub-bacia, foi estabelecida a climatologia dos dados simulados. A climatologia dos dados simulados foi comparada à climatologia mensal dos dados reais observados de precipitação para avaliar a consistência dos dados e sua aplicabilidade na previsão de cenários futuros.

Por outro lado, as previsões climáticas do modelo RegCM4 até o mês de dezembro de 2038 foram utilizadas como dado de entrada do modelo GMDH para previsão de potencial de geração de energia hidrelétrica para avaliação de cenário RCP4.5.

3.11.2. Cenário 2: Tendência de Precipitação.

Para construção deste cenário, foram utilizadas as equações de regressão linear para cada um dos meses que apresentaram tendência estatística significativa de redução da precipitação média para uma determinada sub-bacia, conforme descrito na seção 2.3.3. Dessa forma, foram gerados valores de precipitação média mensal para o período de 2019 a 2038. Em seguida, foi realizada uma composição com os dados simulados pelo modelo climático RegCM4, cujos dados foram exclusivamente substituídos pelos dados da tendência apenas para os meses de referência. A partir da série composta de dados de precipitação, estes passaram a ser utilizados na entrada do modelo GMDH para realizar a previsão do potencial de geração de energia para os próximos vinte anos. A seguir é apresentado um exemplo para o mês de janeiro na sub-bacia Jat referente ao processo de composição da série de precipitação média mensal com a tendência.

1º) Determina-se a equação de tendência $f(t) = \alpha (t - t_0) + \beta$, onde α e β são os coeficientes angular e linear, respectivamente, determinados conforme apresentados em (14). Por exemplo, considerando que a série iniciou no ano de 1979, para o mês de janeiro na sub-bacia Jat, o cálculo para o valor da precipitação no ano de 2019 fica:

$$f(2019) = -4,814 (2019 - 1979) + 385,7 \quad (37)$$

onde t é o ano cujo valor da precipitação média mensal se deseja calcular.

2º) Após a determinação da equação, obtém-se a série, conforme exibido na Tabela 2:

Tabela 2 - Tendência linear calculada da precipitação média na sub-bacia Jat.

ANO	JAN
2019	193
2020	188
⋮	⋮
2036	111
2037	106
2038	102

Fonte: Elaborado pelo autor do trabalho.

3º) Da série de precipitação simulada com o RegCM4 abaixo:

Tabela 3 - Série temporal mensal de precipitação média simulada pelo modelo climático RegCM4.

ANO	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
2019	153	173	217	216	156	66	10	12	31	147	206	265
2020	143	142	204	250	75	50	36	62	104	167	197	254
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2036	231	277	187	248	86	16	19	39	55	95	203	318
2037	221	183	194	127	78	18	57	73	58	168	211	291
2038	217	190	199	289	103	24	25	16	29	137	285	290

Fonte: Elaborado pelo autor do trabalho.

faz-se a substituição pelos dados gerados no passo 2. A nova série composta por tendência passa a ser:

Tabela 4 - Série sintética da precipitação média corrigida com dados de tendência estatística.

ANO	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
2019	193	173	217	216	156	66	10	12	31	147	206	265
2020	188	142	204	250	75	50	36	62	104	167	197	254
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2036	111	277	187	248	86	16	19	39	55	95	203	318
2037	106	183	194	127	78	18	57	73	58	168	211	291
2038	102	190	199	289	103	24	25	16	29	137	285	290

Fonte: Elaborado pelo autor do trabalho.

4º) Repete-se, então, o mesmo procedimento para os demais meses que também apresentaram alguma tendência estatística significativa sobre os dados de precipitação média em cada sub-bacia.

CAPÍTULO IV

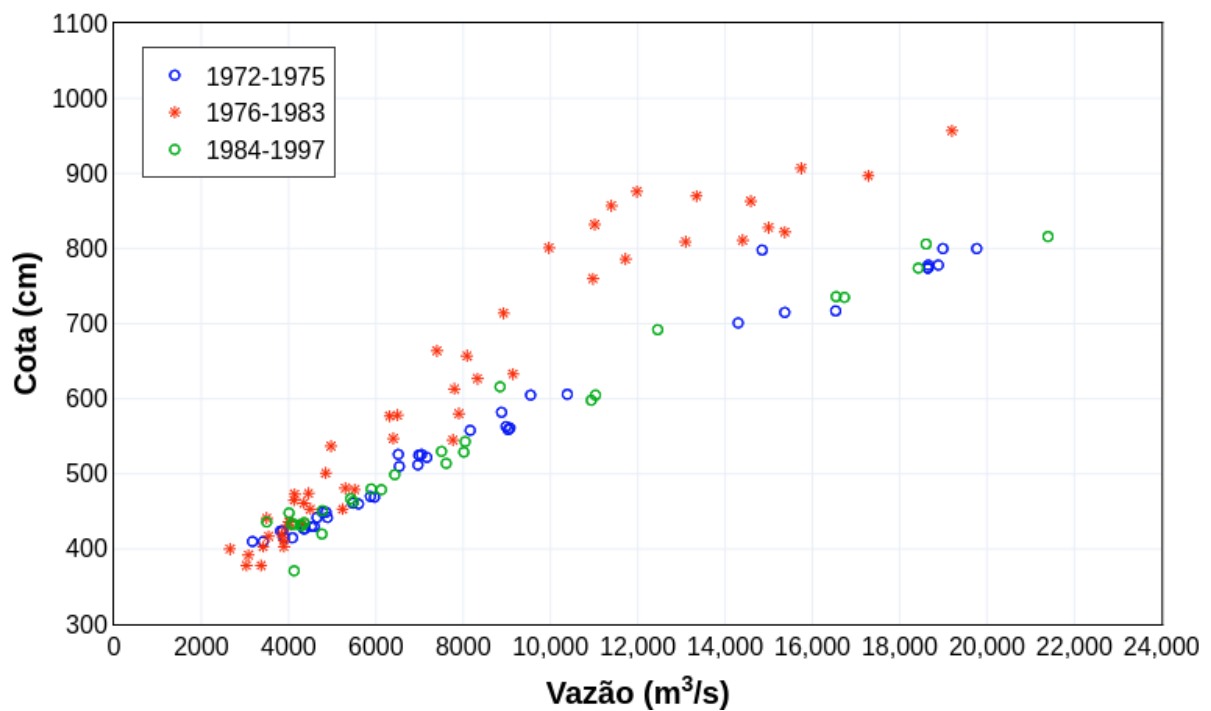
4. RESULTADOS

Este capítulo apresenta os resultados encontrados na tese.

4.1. CURVA-CHAVE

Como mencionado em capítulos anteriores, a geração da série de vazão utiliza a curva-chave da seção do rio para sua determinação sistemática. Na Figura 23 foram plotados os dados provenientes do resumo de descarga da estação Jatobá. Claramente fica denotado que há mais de uma curva para ligar os pontos deste diagrama de dispersão.

Figura 23 - Diagrama de dispersão Cota \times Vazão para a seção do rio Tapajós em na estação Jatobá.



Fonte: Elaborado pelo autor do trabalho.

A Tabela apresenta os dados de cota, vazão, profundidade, largura da seção e áreas molhada, coletados em campo (resumo de descarga) em três datas distintas. Cada uma das datas selecionadas pertence a um dos períodos identificados na Figura 23. Ao aumentar qualquer uma das referidas grandezas (cota, profundidade, largura da seção e áreas molhada), espera-se uma reposta no mesmo sentido para a vazão. Entretanto, os dados coletados no dia 26/02/1976 mostram que embora todas as variáveis tenham apresentado maior valor para este dia comparado aos outros dois dias analisados, o valor da vazão no dia 26/02/1976 reduziu. Como este não é um caso isolado, não se trata de um erro de anotação. Outro ponto importante é que embora as magnitudes dos valores atribuídos às variáveis sejam próximas umas das outras, isto não ocorre com a cota. Para os dados de cota, houve uma variação de cerca de 2 metros. Essa diferença de altura de lâmina d'água é demasiada para considerar que ambos os pontos pertençam à mesma curva-chave. Considerando os 2 metros de altura da lâmina d'água vezes a largura em torno de 1000 metros do rio na seção de medição, seriam pelo menos 2.000 m³/s de vazão a mais para a cota de 8 metros de altura. Portanto, para o dia 26/02/1976 seria razoável uma vazão próxima de 12.000 m³/s, levando-se em conta os dados dos dias 17/12/1973 e 14/06/1995.

Tabela 5 - Dados de resumo de descarga para a estação Jatobá.

Data	Cota (m)	Vazão (m³/s)	Profundidade (m)	Largura da Seção (m)	Área Molhada (km²)
17/12/1973	6,06	10.385	13,59	1005	13.434
26/02/1976	8,01	9.963	14,14	1036	14.649
14/06/1995	5,98	10.934	14,00	1019	14.224

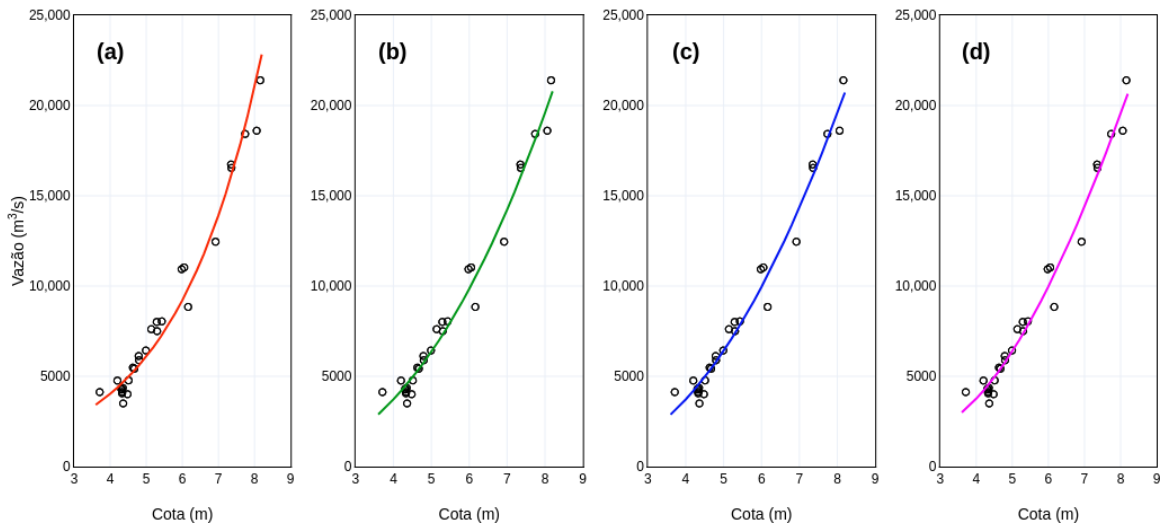
Fonte: Elaborado pelo autor do trabalho.

Poderia supor tratar-se este de um caso de remanso, ou seja, quando nível da água do curso d'água se eleva por influência de algum barramento que pode ser natural, devido o encontro com um rio de maior volume, ou por alguma estrutura de engenharia, como as barragens. Todavia, tais circunstâncias não se repetem nas medições ocorridas após 1984 que, inclusive, compreendem um período de medições bastante superior, sobretudo quando acrescido do período dos primeiros três anos de medição (1972 – 1975). Uma razão plausível, portanto, para tal comportamento seria uma mudança da seção de medição do rio. Visto que nas anotações oficiais das cadernetas de campo e nos dados coletados não há subsídios para afirmar se houve mudança de seção ou definir de forma categórica as razões para as inconformidades

nos dados de resumo de descarga, foram utilizados os dados de resumo de descarga somente após o ano de 1984 para calcular as informações derivadas necessárias para a continuidade da pesquisa.

Considerando, portanto, apenas os dados de descarga a partir de 1984, foram traçadas curvas para determinar a melhor equação de ajuste da relação cota \times vazão usando diferentes equações, conforme apresentado na Figura 24 e Tabela 6. O critério adotado para seleção da equação de ajuste foi a que apresentou maior R^2 . Como as equações polinomiais de ordem 2 e 3 apresentaram melhor resultado, adotou-se a equação mais simples, ou seja, o polinômio de ordem 2.

Figura 24 - Curvas-chave para a seção do rio Tapajós na estação Jatobá usando as seguintes equações: (a) exponencial, (b) potência, (c) polinômio de ordem 2 e (d) polinômio de ordem 3.



Fonte: Elaborado pelo autor do trabalho.

Tabela 6 - Equações para determinação da melhor curva de ajuste aos dados de cota \times vazão na estação Jatobá.

Curva	Equação	R^2
Exponencial	$y = 774,57e^{0,4126x}$	0,9534
Potência	$y = 134,69x^{2,3949}$	0,9595
Polinômio ordem 2	$y = 423,3x^2 - 1116,8x + 1415,5$	0,9792
Polinômio ordem 3	$y = -19,747x^3 + 774,91x^2 - 3142x + 5193,5$	0,9792

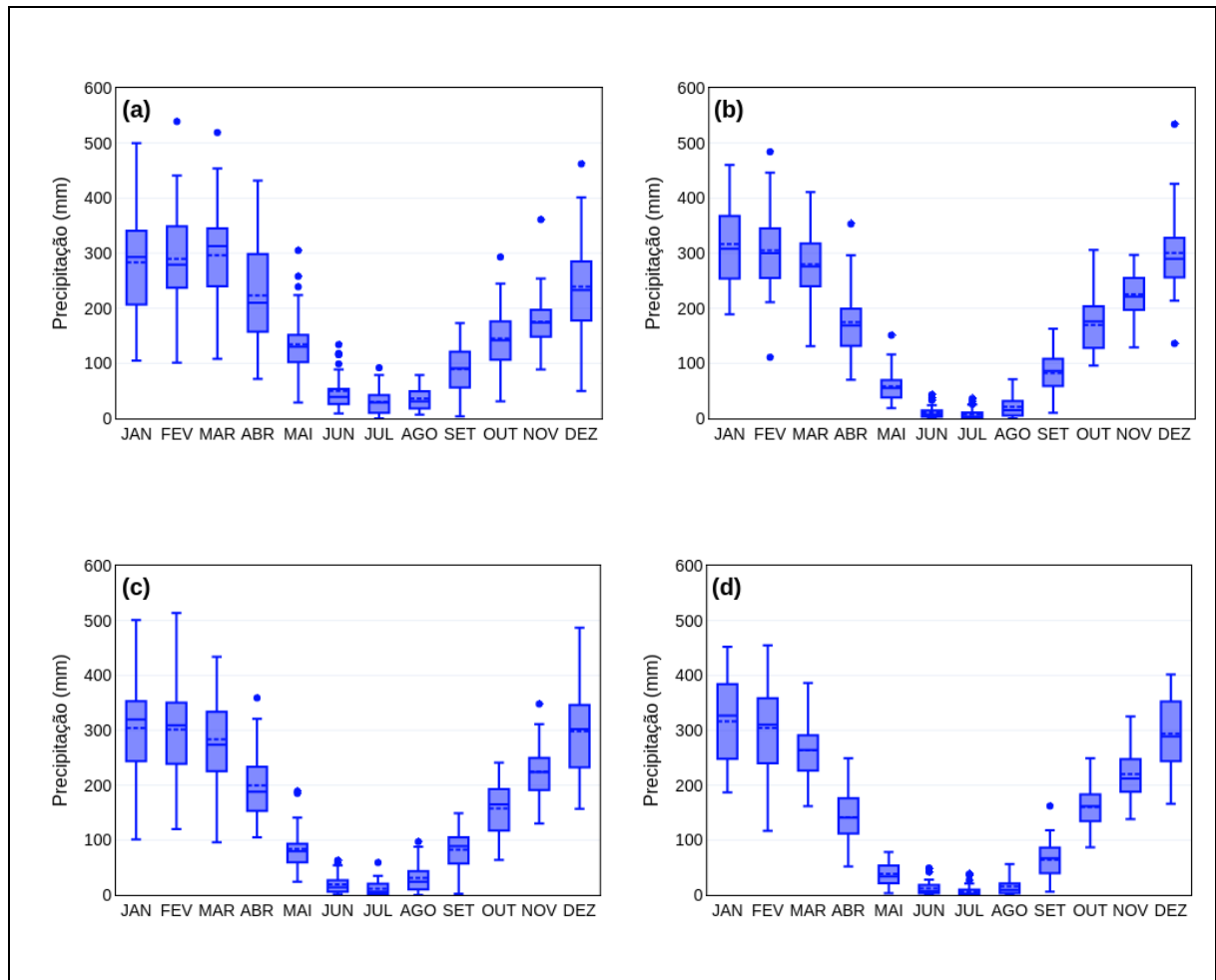
Fonte: Elaborado pelo autor do trabalho.

4.2. CLIMATOLOGIA

4.2.1. Climatologia da Precipitação

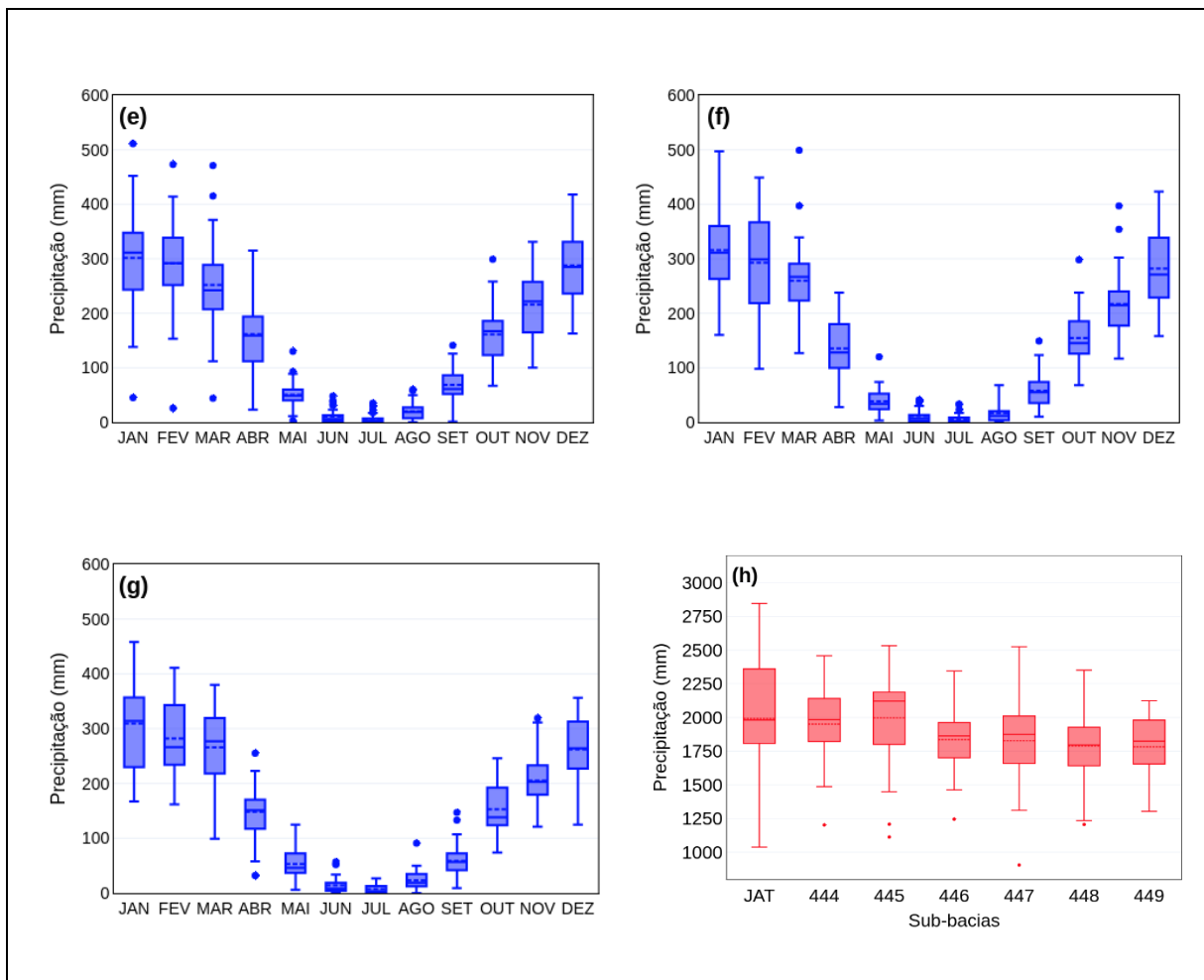
A climatologia mensal da precipitação e o comportamento anual para cada sub-bacia que influencia o AHE Jatobá podem ser observados nas Figuras 25 e 26. Todas as sub-bacias estudadas apresentaram comportamento climatológico semelhante. O auge do período chuvoso vai de janeiro a março na bacia do rio Tapajós. Nos meses mais chuvosos, cada sub-bacia registra cerca de 300 mm mensais. A estação chuvosa também é caracterizada por apresentar grande variabilidade.

Figura 25 - Climatologia mensal da precipitação média na (a) sub-bacia Jat, (b) sub-bacia 444, (c) sub-bacia 445, (d) sub-bacia 446. Período base 1984-2018.



Fonte: Elaborado pelo autor do trabalho.

Figura 26 - Climatologia mensal da precipitação média na (e) sub-bacia 447, (f) sub-bacia 448 e (g) sub-bacia 449 para o período base 1984-2018. (h) Distribuição anual da precipitação por sub-bacia.



Fonte: Elaborado pelo autor do trabalho.

O trimestre mais seco ocorre nos meses de junho, julho e agosto. Em média, em qualquer que seja a sub-bacia, o mês mais seco recebe menos de 30 mm, o que representa apenas 10% ou menos do volume de precipitação do mês mais chuvoso. A estação seca pode, inclusive, não registrar precipitação em um dos meses.

Ao longo do ano as sub-bacias situadas mais ao norte da bacia (JAT, 444 e 445) acumulam cerca de 2.000 mm de precipitação, enquanto as demais recebem aproximadamente

1.850 mm. Cerca de 75% dessa diferença no volume de precipitação acumulada a favor das sub-bacias ao norte ocorrem basicamente nos meses de abril e maio, período de transição da estação chuvosa para a seca.

A sub-bacia Jat é a que apresentou maior variabilidade anual de precipitação (Figura 26h). A sub-bacia 446 destacou-se por apresentar comportamento interanual mais regular, com menor intervalo interquartil e dados simétricos, com pequena dispersão. A sub-bacia 445 apresentou forte assimetria em seus dados com mediana distante da média e maior variância comparada às outras sub-bacias.

4.2.2. Análise Exploratória da Precipitação

A seguir são apresentados os resultados das sínteses dos dados de precipitação média mensal para cada sub-bacia, abrangendo medidas de posição e dispersão (Tabelas 7 a 13). Observa-se que o trimestre junho, julho e agosto, que corresponde à estação seca, pode não registrar precipitação neste período. Os valores máximos registrados na estação chuvosa são importantes, sobretudo, para adequado dimensionamento da barragem. O maior volume mensal registrado em uma das sub-bacias foi de 620 mm, no mês de janeiro, na sub-bacia 449. Esse valor é praticamente o dobro da sua média, o que caracteriza a grande variabilidade da precipitação nesta sub-bacia, mas esta característica também se estende às demais sub-bacias do rio Tapajós. O segundo maior volume mensal de precipitação média registrado (539 mm) foi justamente na sub-bacia Jat, local do futuro barramento.

Tabela 7 - Estatística descritiva da precipitação na sub-bacia Jat.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	ANO
Mínimo	105	101	108	72	29	9	0	7	4	31	89	50	1038
Máximo	500	539	519	432	305	134	92	79	173	293	361	462	2846
Média	283	290	296	223	134	50	30	36	90	145	175	239	1993
Mediana	293	279	313	210	130	39	29	31	91	142	174	233	1981
Desv Pad	97	85	94	85	59	33	22	21	40	61	52	91	464
Variância	9390	7247	8877	7141	3518	1109	479	424	1588	3663	2711	8217	215199

Fonte: Elaborado pelo autor do trabalho.

Tabela 8 - Estatística descritiva da precipitação na sub-bacia 444.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	ANO
------------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Mínimo	189	111	131	70	19	0	0	0	10	96	129	136	1202
Máximo	460	484	411	353	151	43	36	71	163	306	297	534	2458
Média	317	305	280	175	58	11	7	21	83	170	225	300	1951
Mediana	308	300	276	169	55	7	2	15	86	176	221	290	1984
Desv Pad	79	77	58	62	31	11	10	20	35	46	40	72	267
Variância	6224	5853	3329	3855	940	114	100	404	1200	2071	1586	5247	71068

Fonte: Elaborado pelo autor do trabalho.

Tabela 9 - Estatística descritiva da precipitação na sub-bacia 445.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	ANO
Mínimo	101	120	96	105	24	1	0	0	2	64	130	157	1113
Máximo	501	514	434	359	189	63	59	97	149	241	348	487	2533
Média	304	301	284	200	84	19	11	31	83	158	224	298	1997
Mediana	320	309	274	188	80	14	6	24	89	165	224	302	2122
Desv Pad	86	83	82	63	38	17	13	26	33	48	50	83	345
Variância	7449	6853	6766	3960	1441	292	163	698	1096	2261	2520	6944	118865

Fonte: Elaborado pelo autor do trabalho.

Tabela 10 - Estatística descritiva da precipitação na sub-bacia 446.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	ANO
Mínimo	187	117	162	52	3	0	0	0	6	87	138	166	1245
Máximo	452	455	386	249	78	49	38	56	162	249	325	402	2345
Média	316	304	264	141	39	12	6	16	64	160	220	294	1836
Mediana	327	310	264	141	34	7	2	9	67	162	212	289	1863
Desv Pad	74	78	54	49	21	13	10	18	34	37	47	64	228
Variância	5509	6128	2917	2386	432	170	99	319	1168	1398	2238	4058	52083

Fonte: Elaborado pelo autor do trabalho.

Tabela 11 - Estatística descritiva da precipitação na sub-bacia 447.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	ANO
Mínimo	45	26	44	23	3	0	0	0	1	67	100	163	905
Máximo	511	473	471	315	130	48	35	60	141	299	331	418	2525
Média	301	292	252	162	50	11	6	20	69	161	216	288	1827
Mediana	311	292	242	159	49	6	2	19	61	167	222	285	1874
Desv Pad	89	84	82	62	25	13	9	16	31	51	59	68	322
Variância	7923	7038	6756	3878	626	160	79	268	949	2565	3447	4646	103972

Fonte: Elaborado pelo autor do trabalho.

Tabela 12 - Estatística descritiva da precipitação na sub-bacia 448.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	ANO
Mínimo	160	98	127	28	3	0	0	0	10	68	117	158	1206
Máximo	497	449	499	238	120	41	33	68	149	298	397	423	2351
Média	316	293	260	135	38	11	6	17	58	155	217	282	1788
Mediana	311	299	267	128	34	7	1	12	55	145	215	271	1797
Desv Pad	79	88	73	49	23	13	9	18	31	51	58	68	270
Variância	6293	7671	5379	2420	532	159	79	334	974	2555	3376	4676	72808

Fonte: Elaborado pelo autor do trabalho.

Tabela 13 - Estatística descritiva da precipitação na sub-bacia 449.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	ANO
Mínimo	167	162	99	32	6	0	0	0	9	74	121	125	1303
Máximo	620	411	380	255	125	57	27	91	147	246	319	356	2124
Média	309	282	266	148	53	14	7	23	59	153	205	262	1782
Mediana	314	266	277	151	46	8	3	19	57	138	203	264	1824
Desv Pad	91	67	63	44	27	15	9	19	29	44	46	59	243
Variância	8300	4470	3977	1941	731	231	77	350	863	1933	2127	3471	59145

Fonte: Elaborado pelo autor do trabalho.

Nas Tabelas 14 a 20 são apresentados resultados do teste de normalidade de Shapiro-Wilk aplicado sobre os dados de precipitação média mensal para cada sub-bacia. De acordo com Leotti et al. (2005), este é o melhor teste de aderência à normalidade. Os testes de normalidade são usados para determinar se um conjunto de dados de uma dada variável aleatória, é bem modelada por uma distribuição normal ou não, ou para calcular a probabilidade da variável aleatória subjacente estar normalmente distribuída. A aplicação deste tipo de teste é importante para um tratamento estatístico adequado ao conjunto de dados. Para a comparação de médias e variâncias entre duas amostras quando os dados não apresentam distribuição

normal, é necessário lançar mão de testes não paramétricos, cuja base teórica não é tão robusta, porém estes testes podem ser aplicados independentemente do tipo de distribuição. Os resultados apresentados indicaram que a hipótese H_0 foi rejeitada em vários momentos, denotando que a precipitação média não possui distribuição normal.

Tabela 14 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia Jat, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
Valor-p	0,525	0,093	0,546	0,988	0,069	0,079	0,049	0,180	0,427	0,772	0,715	0,208
H_0	Aceita	Aceita	Aceita	Aceita	Aceita	Aceita	Rejeita	Aceita	Aceita	Aceita	Aceita	Aceita
H_a	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita

Fonte: Elaborado pelo autor do trabalho.

Tabela 15 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 444, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
Valor-p	0,097	0,131	0,885	0,275	0,005	0,0001	0,0001	0,001	0,985	0,088	0,225	0,025
H_0	Aceita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita	Rejeita
H_a	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita	Aceita

Fonte: Elaborado pelo autor do trabalho.

Tabela 16 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 445, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
Valor-p	0,377	0,990	0,821	0,068	0,014	0,000	0,0001	0,005	0,751	0,235	0,892	0,625
H_0	Aceita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita	Aceita
H_a	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita	Rejeita

Fonte: Elaborado pelo autor do trabalho.

Tabela 17 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 446, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
Valor-p	0,160	0,741	0,305	0,804	0,311	0,0001	0,0001	0,0001	0,561	0,935	0,326	0,318
H_0	Aceita	Aceita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita	Aceita
H_a	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita	Rejeita

Fonte: Elaborado pelo autor do trabalho.

Tabela 18 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 447, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
Valor-p	0,527	0,337	0,583	0,772	0,100	0,0001	0,0001	0,015	0,894	0,434	0,680	0,636
H_0	Aceita	Aceita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita	Aceita
H_a	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita	Rejeita

Fonte: Elaborado pelo autor do trabalho.

Tabela 19 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 448, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
Valor-p	0,902	0,721	0,039	0,136	0,014	0,0001	0,0001	0,0001	0,074	0,324	0,032	0,298
H_0	Aceita	Aceita	Rejeita	Aceita	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Rejeita	Aceita
H_a	Rejeita	Rejeita	Aceita	Rejeita	Aceita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Aceita	Rejeita

Fonte: Elaborado pelo autor do trabalho.

Tabela 20 - Teste de Shapiro-Wilk para normalidade sobre os dados de precipitação na sub-bacia 449, considerando como hipótese nula H_0 que a variável apresenta distribuição normal, para $\alpha = 0,05$.

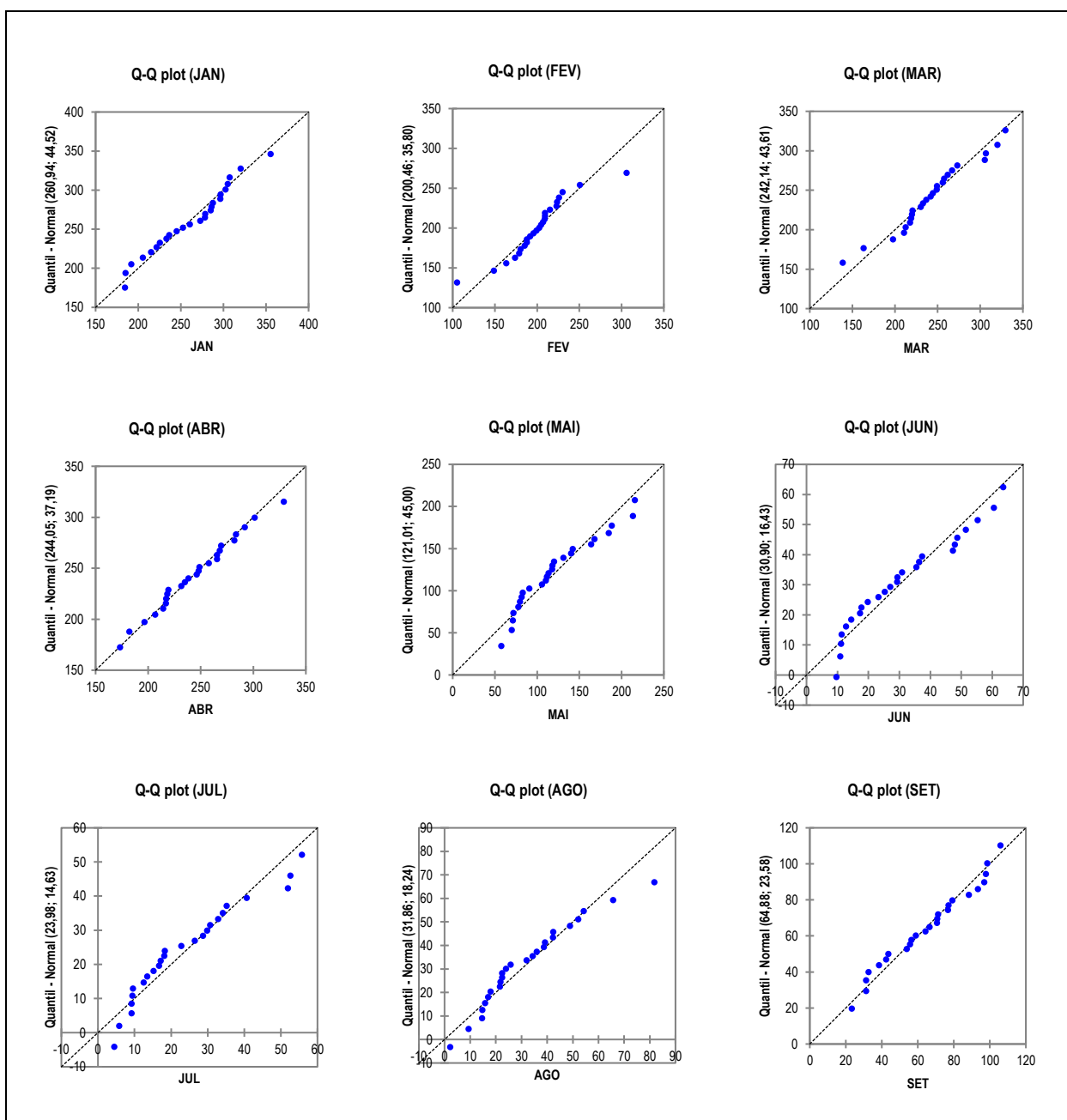
Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
Valor-p	0,025	0,211	0,338	0,757	0,409	0,0001	0,0001	0,002	0,023	0,488	0,451	0,393
H_0	Rejeita	Aceita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita
H_a	Aceita	Rejeita	Rejeita	Rejeita	Rejeita	Aceita	Aceita	Aceita	Aceita	Rejeita	Rejeita	Rejeita

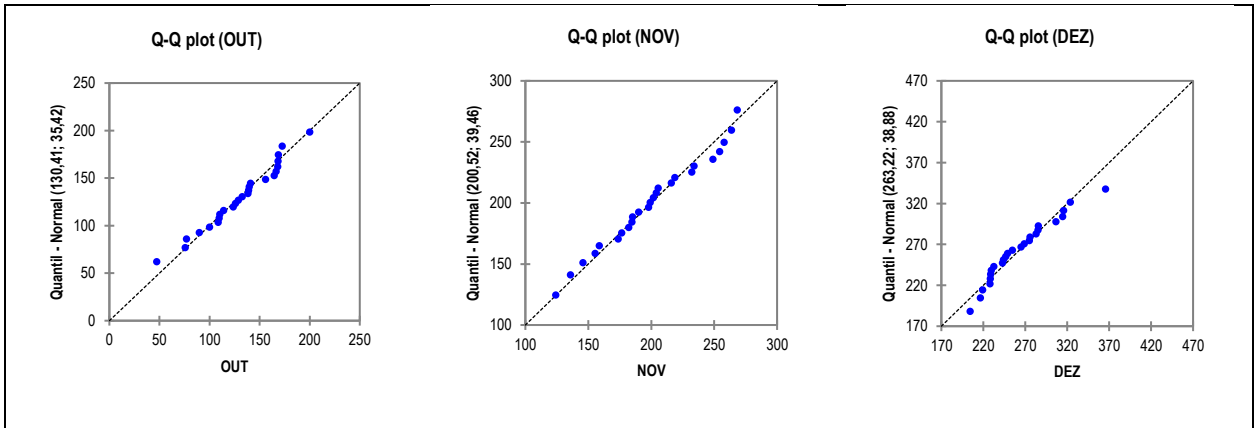
Fonte: Elaborado pelo autor do trabalho.

Os gráficos Normal Q-Q são mostrados nas Figuras 27 a 33. O gráfico Q-Q (quantil-quantil) é uma ferramenta para avaliar se um conjunto de dados possui uma distribuição teórica, como por exemplo, a distribuição normal. O Normal Q-Q permite uma verificar de forma visual e direta se os dados obedecem a uma distribuição normal. Um gráfico Q-Q é um gráfico de dispersão criado pela plotagem de dois conjuntos de quantis, um contra o outro. Se os dois

conjuntos de quantis forem da mesma distribuição, os pontos formam uma linha que é praticamente reta. Os resultados encontrados evidenciam que na estação seca os dados não apresentam distribuição normal, coerente com o teste de Shapiro-Wilk.

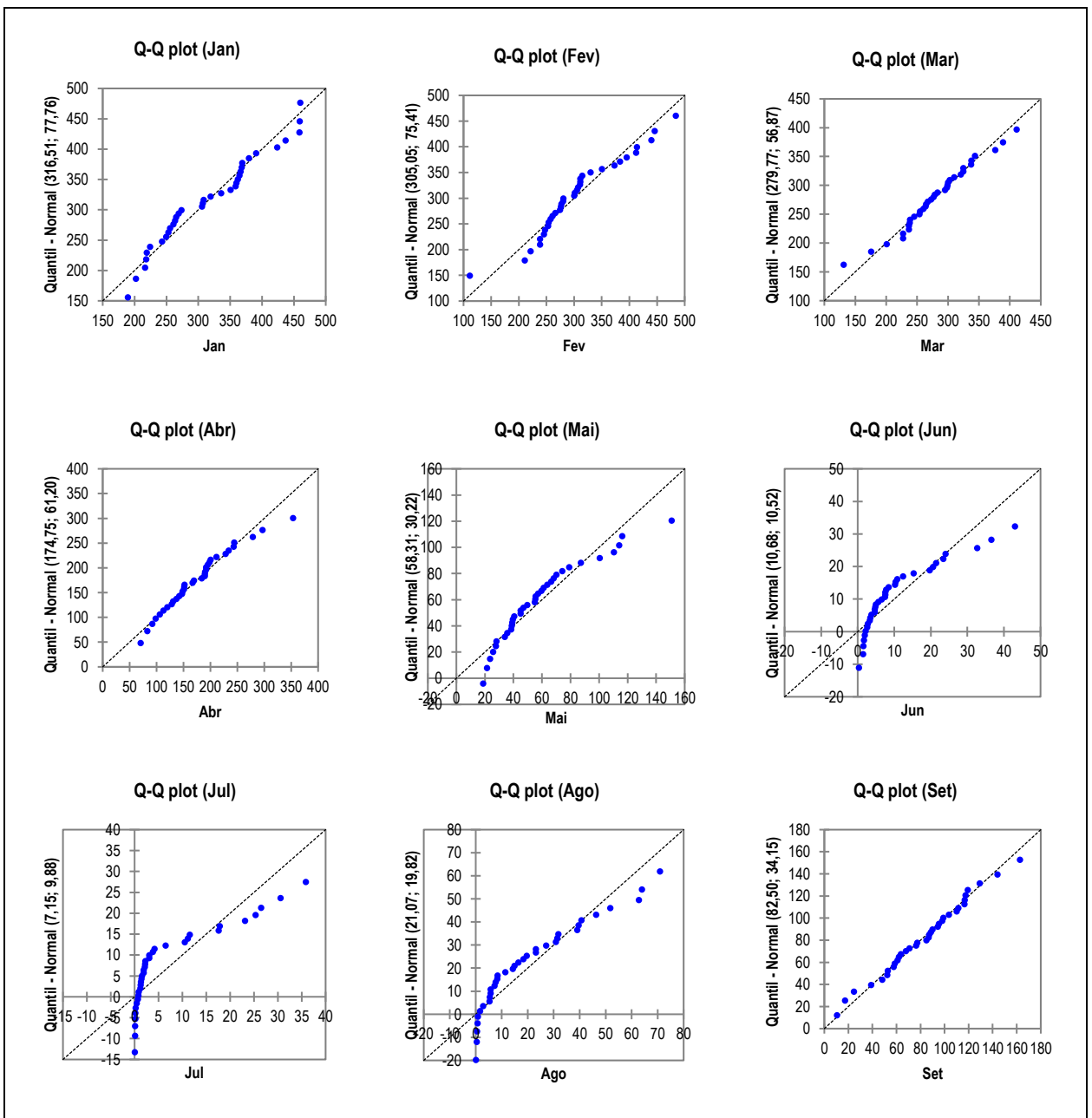
Figura 27 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia Jat.

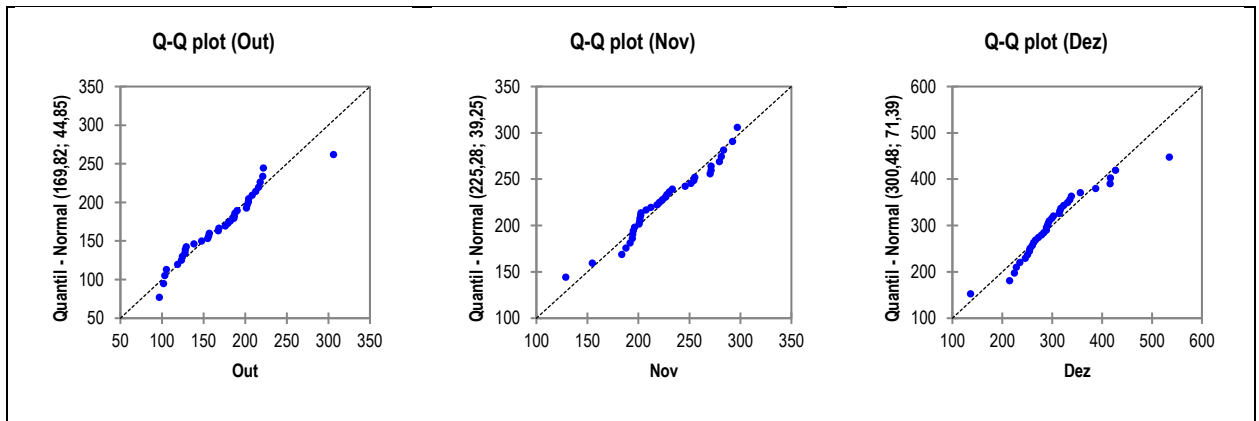




Fonte: Elaborado pelo autor do trabalho.

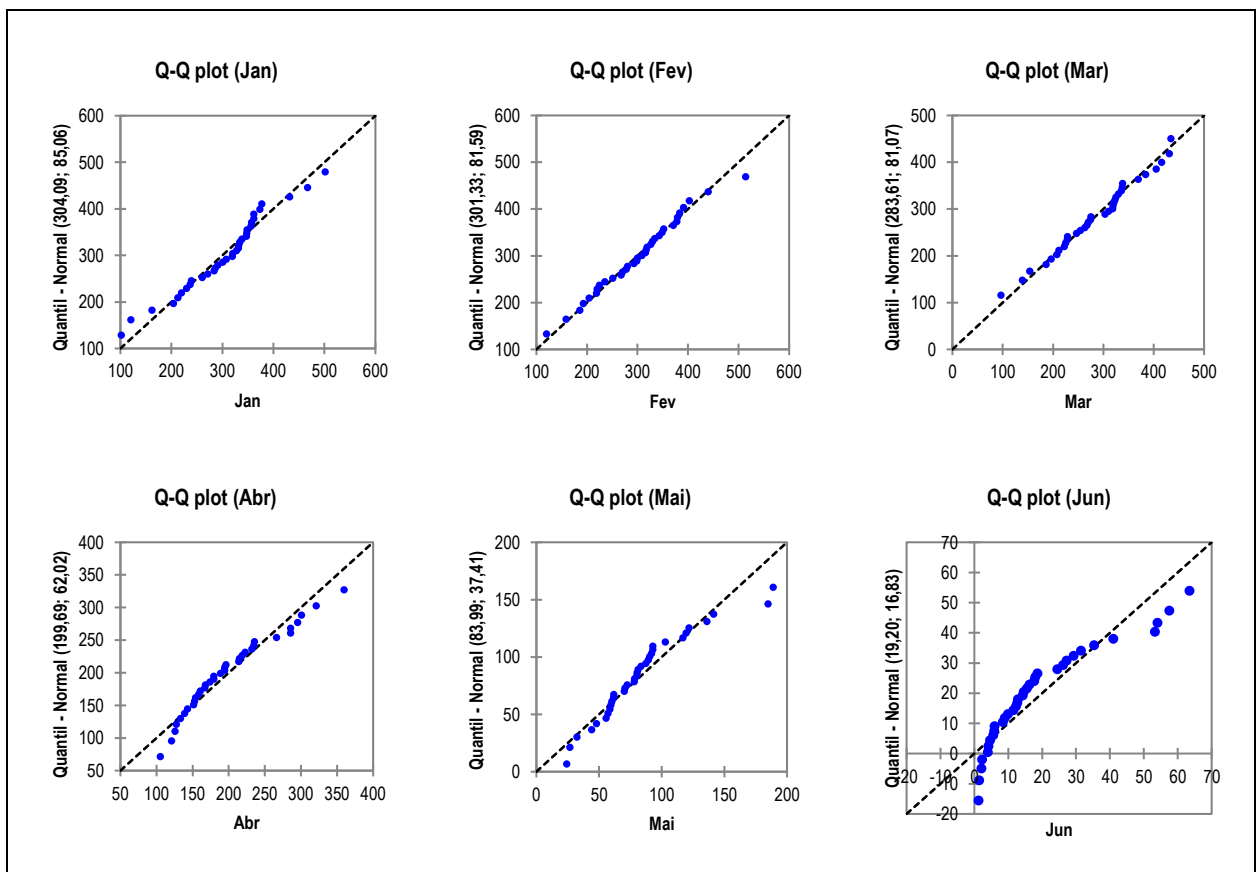
Figura 28 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 444.

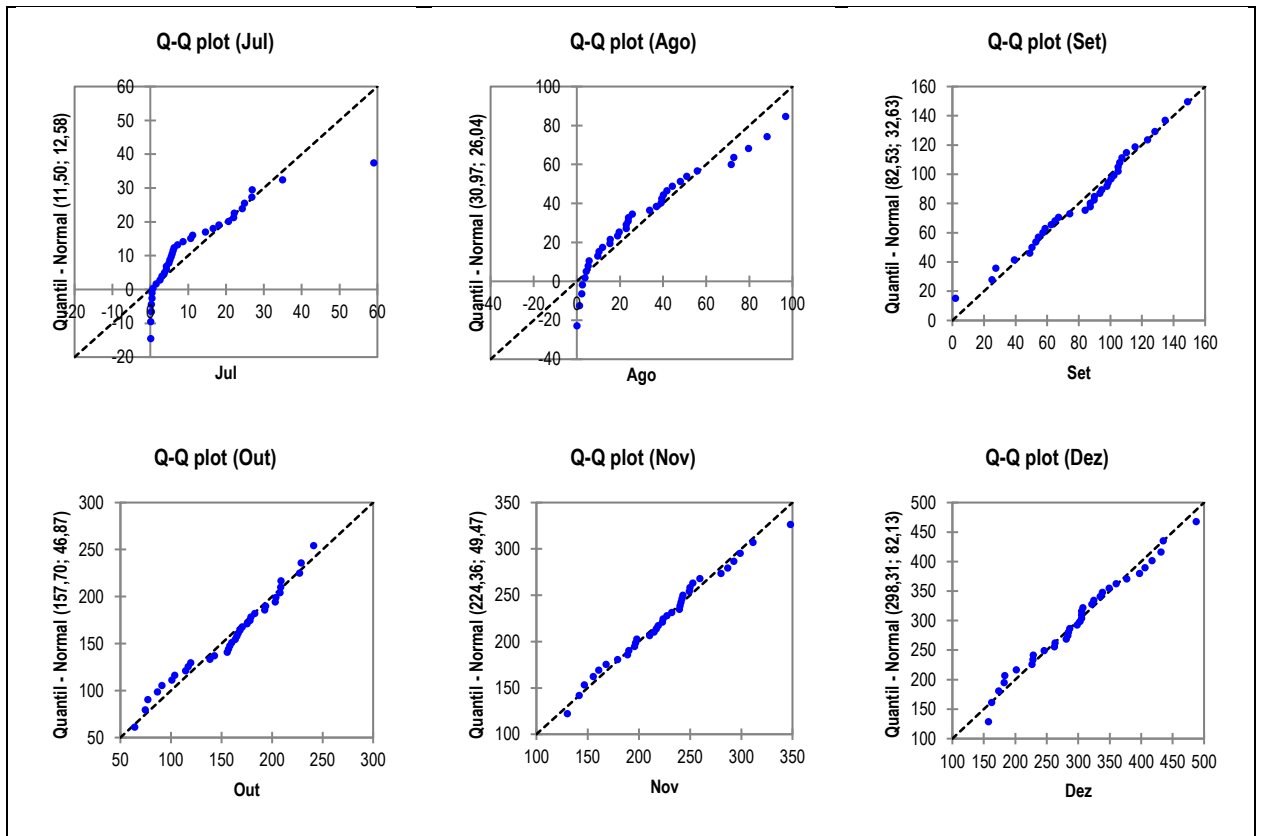




Fonte: Elaborado pelo autor do trabalho.

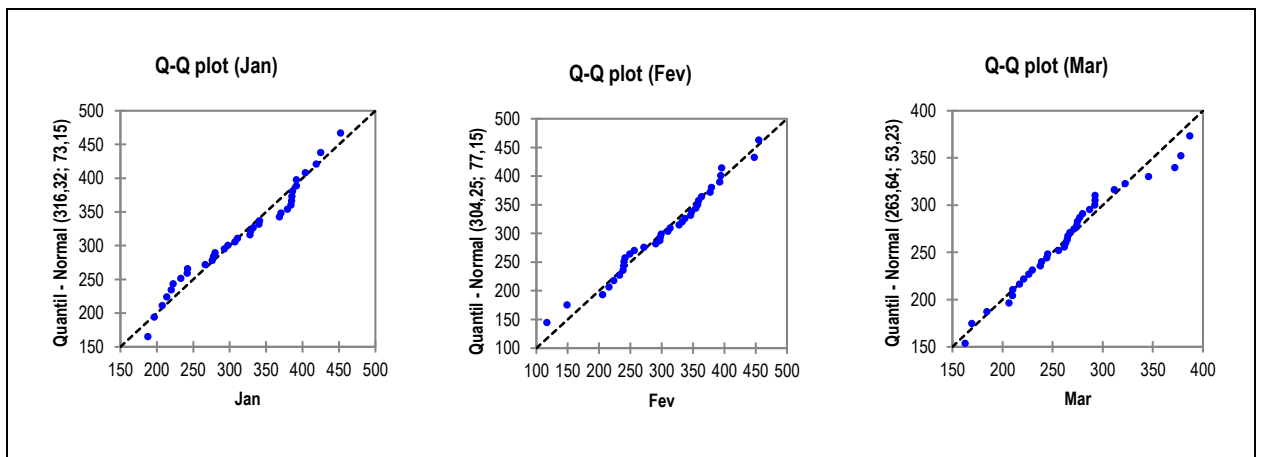
Figura 29 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 445.

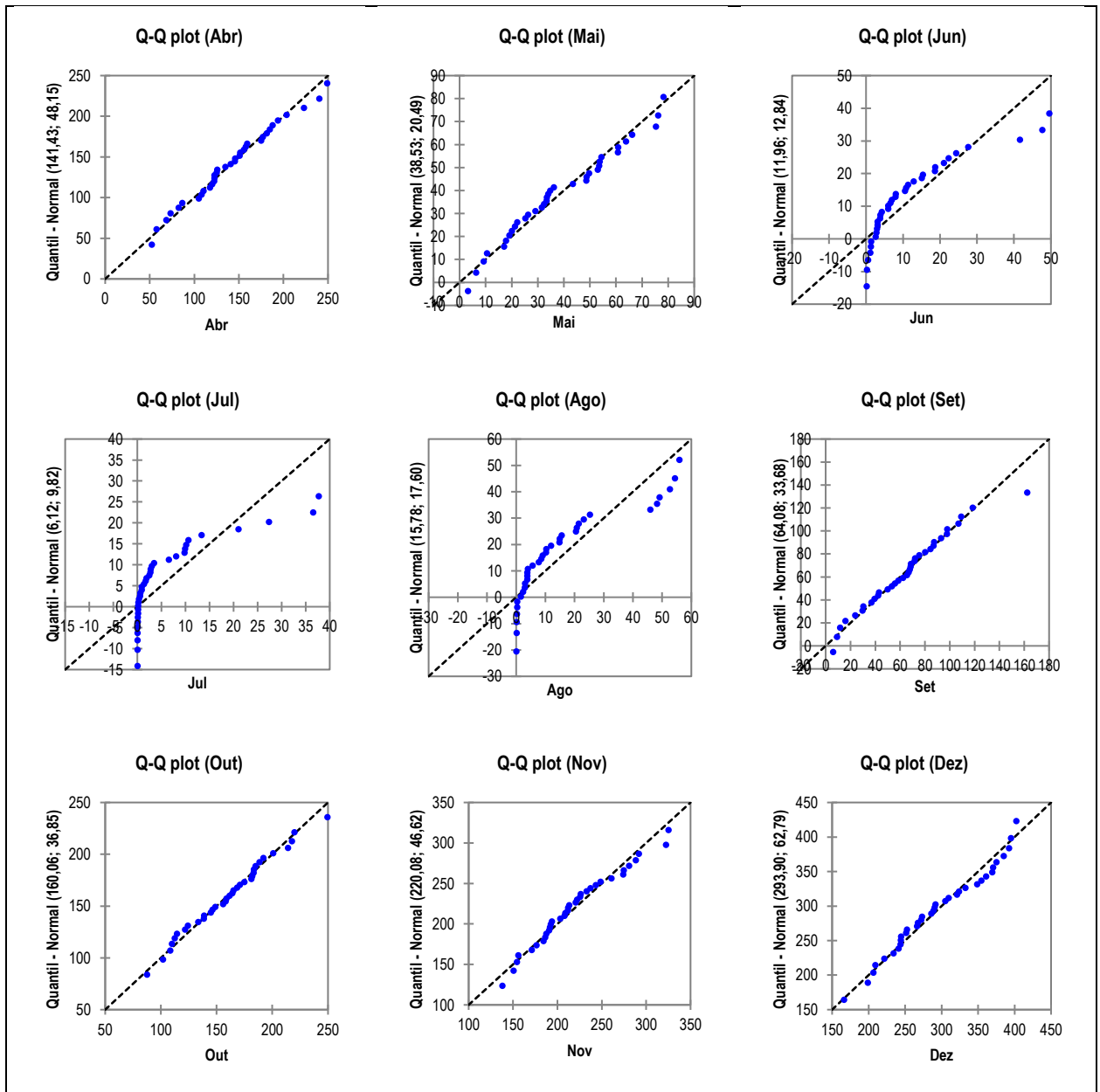




Fonte: Elaborado pelo autor do trabalho.

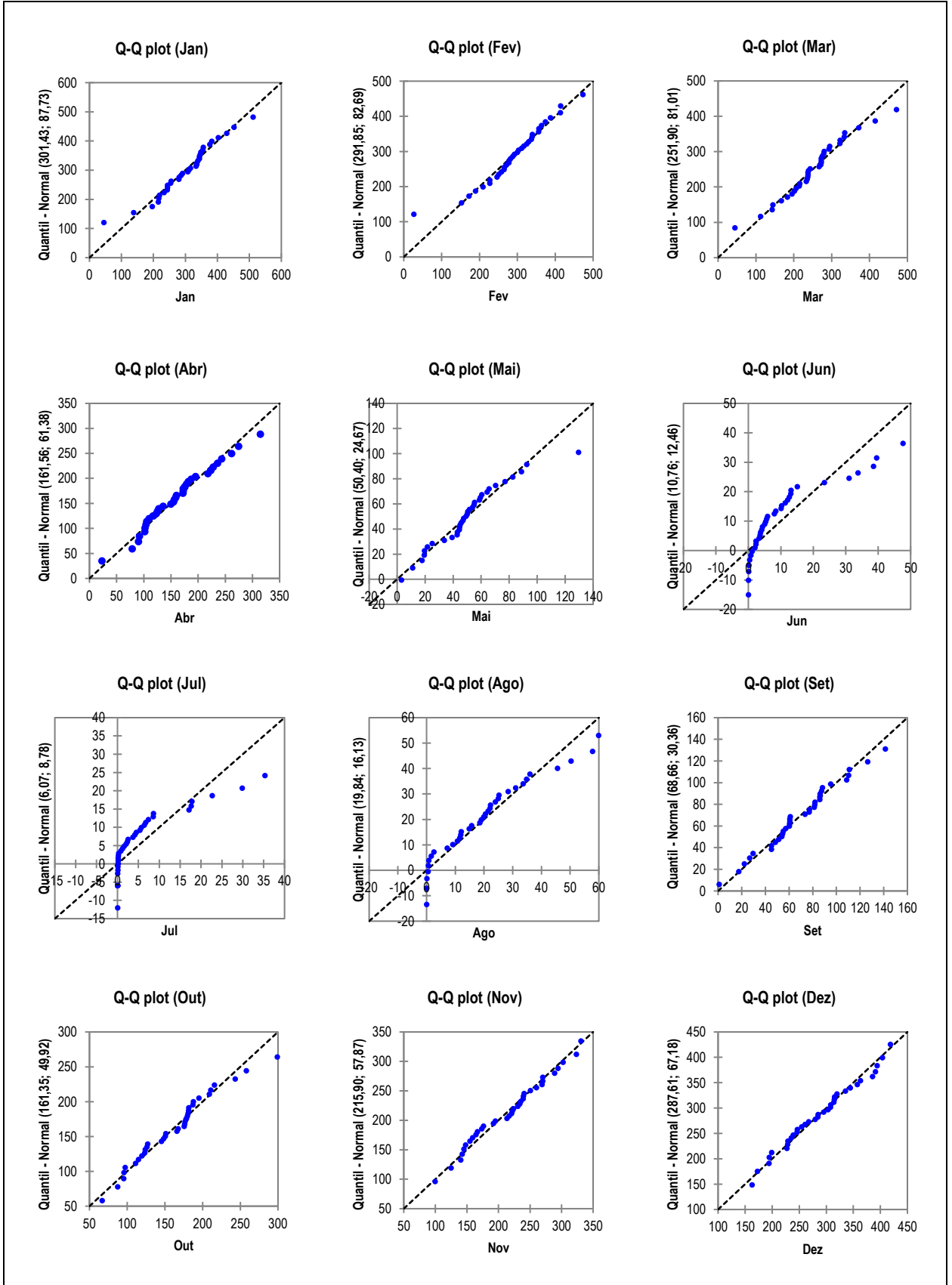
Figura 30 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 446.





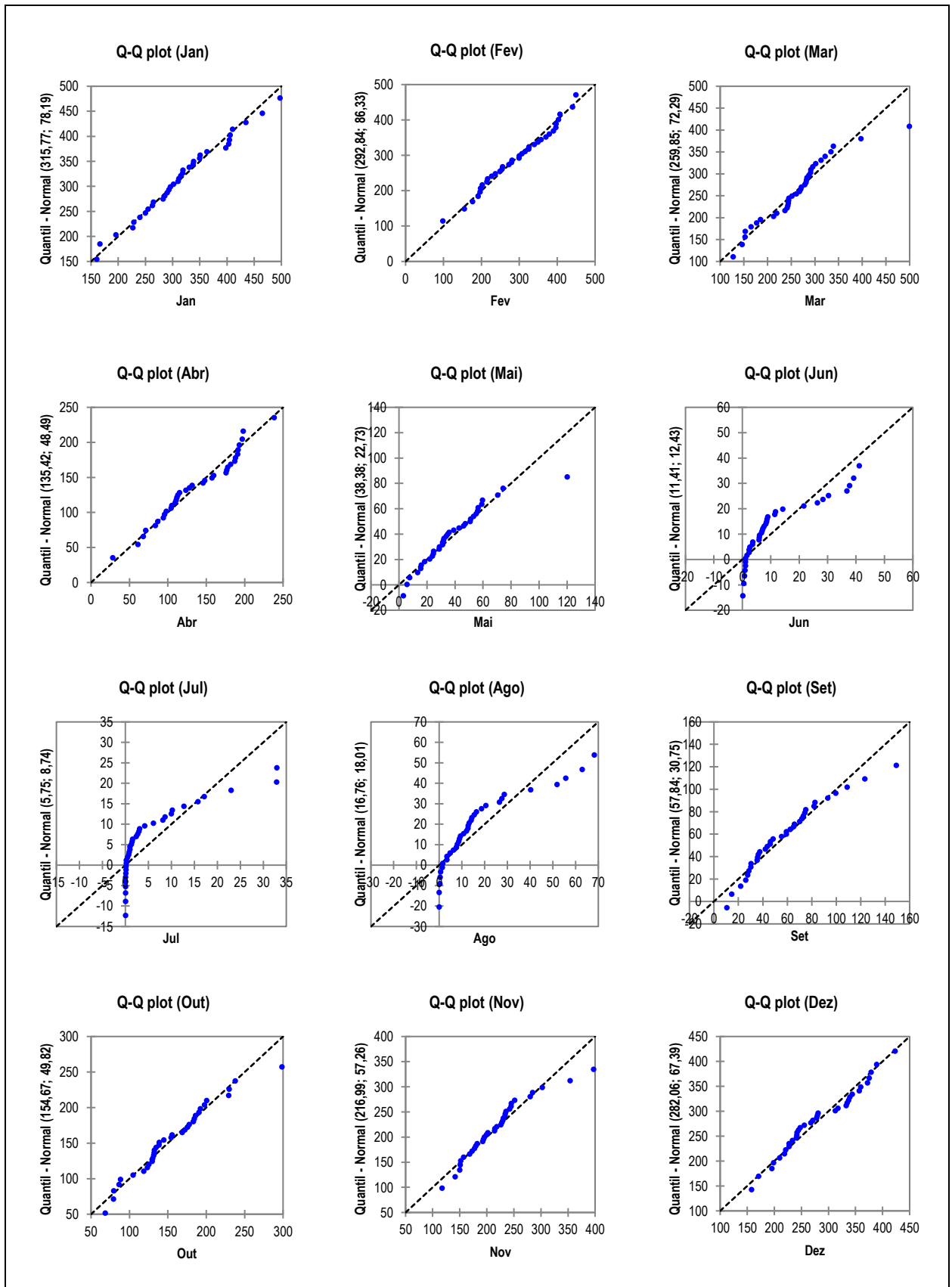
Fonte: Elaborado pelo autor do trabalho.

Figura 31 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 447.



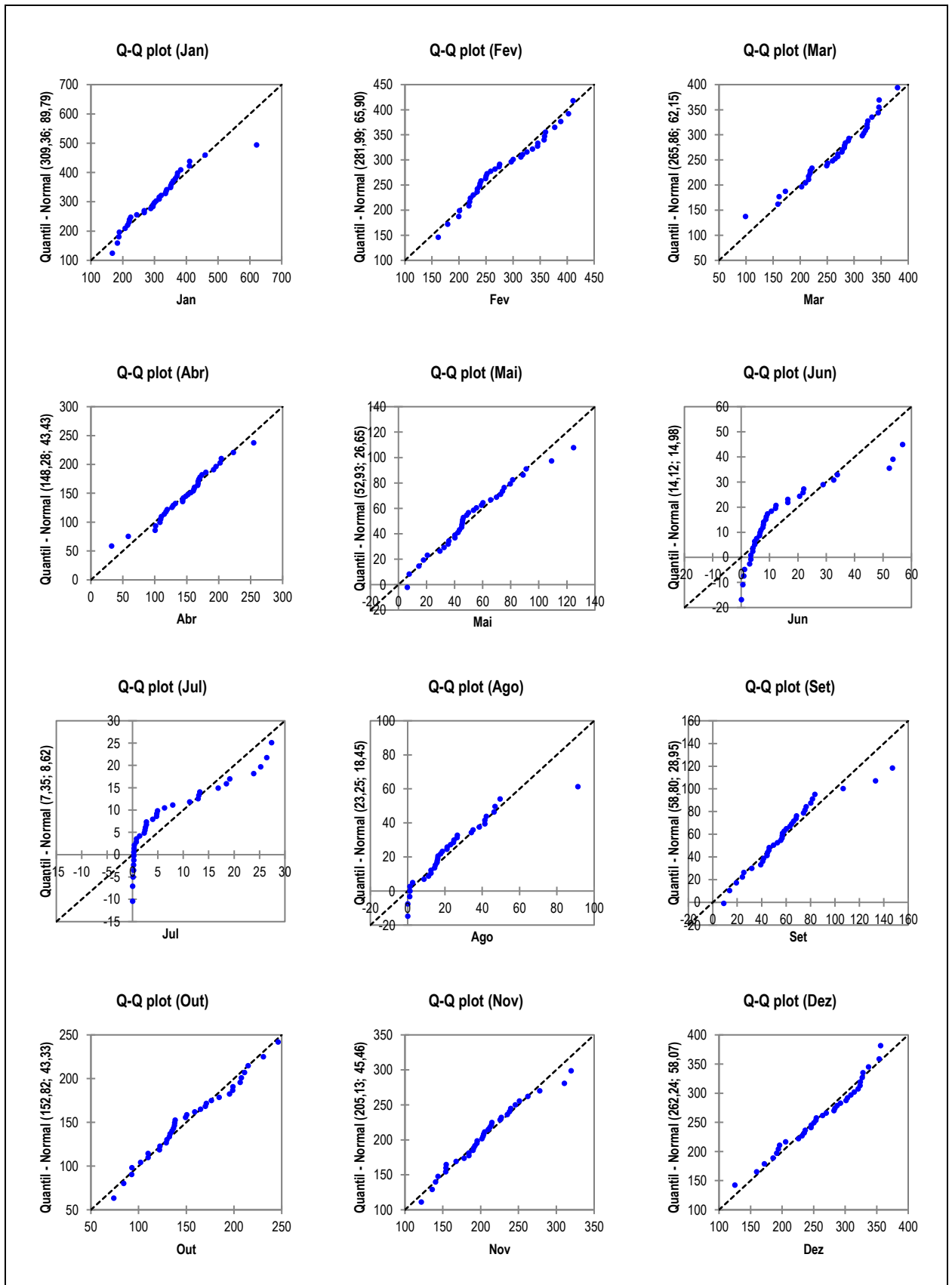
Fonte: Elaborado pelo autor do trabalho.

Figura 32 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 448.



Fonte: Elaborado pelo autor do trabalho.

Figura 33 - Teste de normalidade Q-Q para os dados de precipitação mensal na sub-bacia 449.

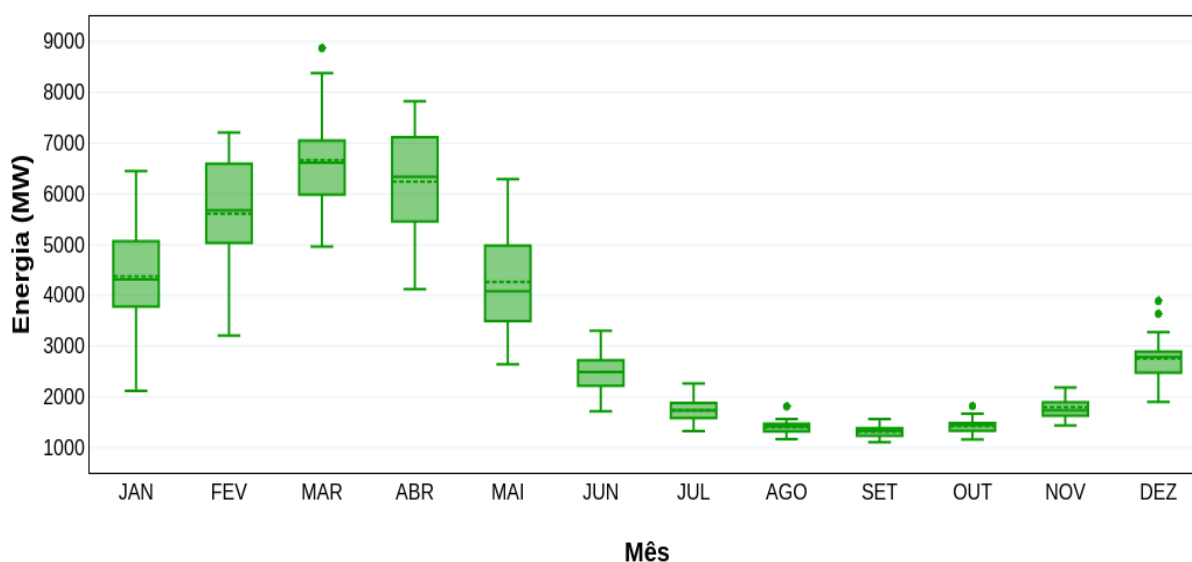


Fonte: Elaborado pelo autor do trabalho.

4.2.3. Climatologia do Potencial de Geração de Energia

A climatologia do potencial médio mensal de geração de energia hidrelétrica para o AHE Jatobá é apresentada na Figura 34 e Tabela 21. Pode ser observada a simetria dos dados e sua dispersão. O trimestre agosto, setembro e outubro apresenta o período mais crítico, com menor potencial para geração de hidreletricidade. No mês de setembro, que apresenta menor capacidade de geração de energia, seu potencial para geração energética varia entre 1.229 MW e 1.381 MW (correspondente ao intervalo interquartil), com média de 1.307 MW. Enquanto o trimestre fevereiro, março e abril é caracterizado como o período de maior potencial teórico, cuja capacidade de geração mensal média é superior a 5.000 MW, ultrapassando ligeiramente os 6.000 MW no mês de março, que corresponde ao auge da capacidade de geração de energia.

Figura 34 - Distribuição mensal do potencial teórico de geração de energia hidrelétrica para o AHE Jatobá no período base 1984-2013.



Fonte: Elaborado pelo autor do trabalho.

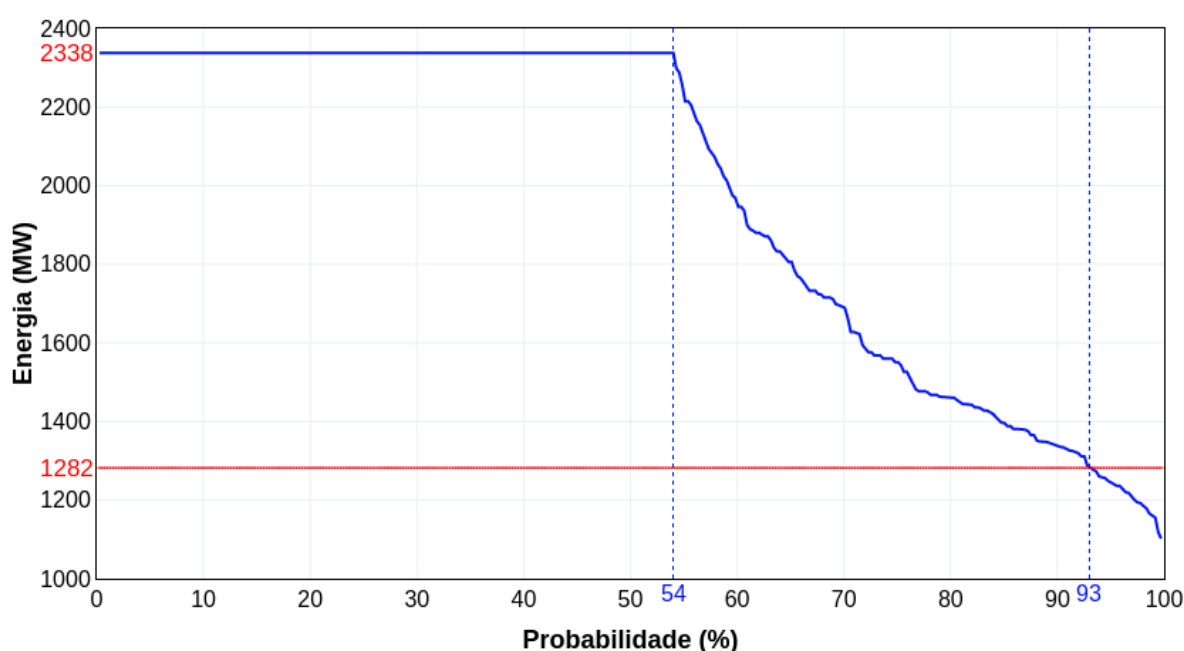
Tabela 21 - Estatística descritiva do potencial mensal para geração de energia no AHE Jatobá.

Parâmetros	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
Mínimo	2112	3209	4962	4120	2638	1711	1323	1160	1102	1155	1436	1899
Máximo	6448	7209	8873	7826	6291	3299	2257	1806	1560	1815	2184	3886
Média	4373	5606	6665	6240	4263	2488	1730	1394	1307	1419	1792	2751
Mediana	4314	5676	6618	6339	4082	2482	1733	1410	1328	1443	1733	2784
Desv Pad	955	1034	913	1029	999	393	205	131	106	142	209	435
Variância	912907	1069368	833488	1059841	997347	154838	41920	17277	11308	20284	43697	189596

Fonte: Elaborado pelo autor do trabalho.

Na Figura 35 observa-se a curva de permanência adaptada para o potencial de geração de energia no AHE Jatobá. A geração de energia prevista em projeto considerando os dados observados e eficiência do conjunto turbina-gerador de 95%, mostram que em apenas 93% do período as condições planejadas para geração de energia serão atendidas, ou seja, o mínimo de 1.282 MW. Em outros 7% a geração ficará abaixo do previsto. A UHE poderá gerar energia em sua capacidade máxima (2.338 MW) durante 54% do período.

Figura 35 - Curva de permanência para o potencial de geração mensal de energia hidrelétrica no AHE Jatobá no período base 1984-2013. A linha vermelha destaca o valor de geração no período crítico (1.282 MW), conforme projeto.

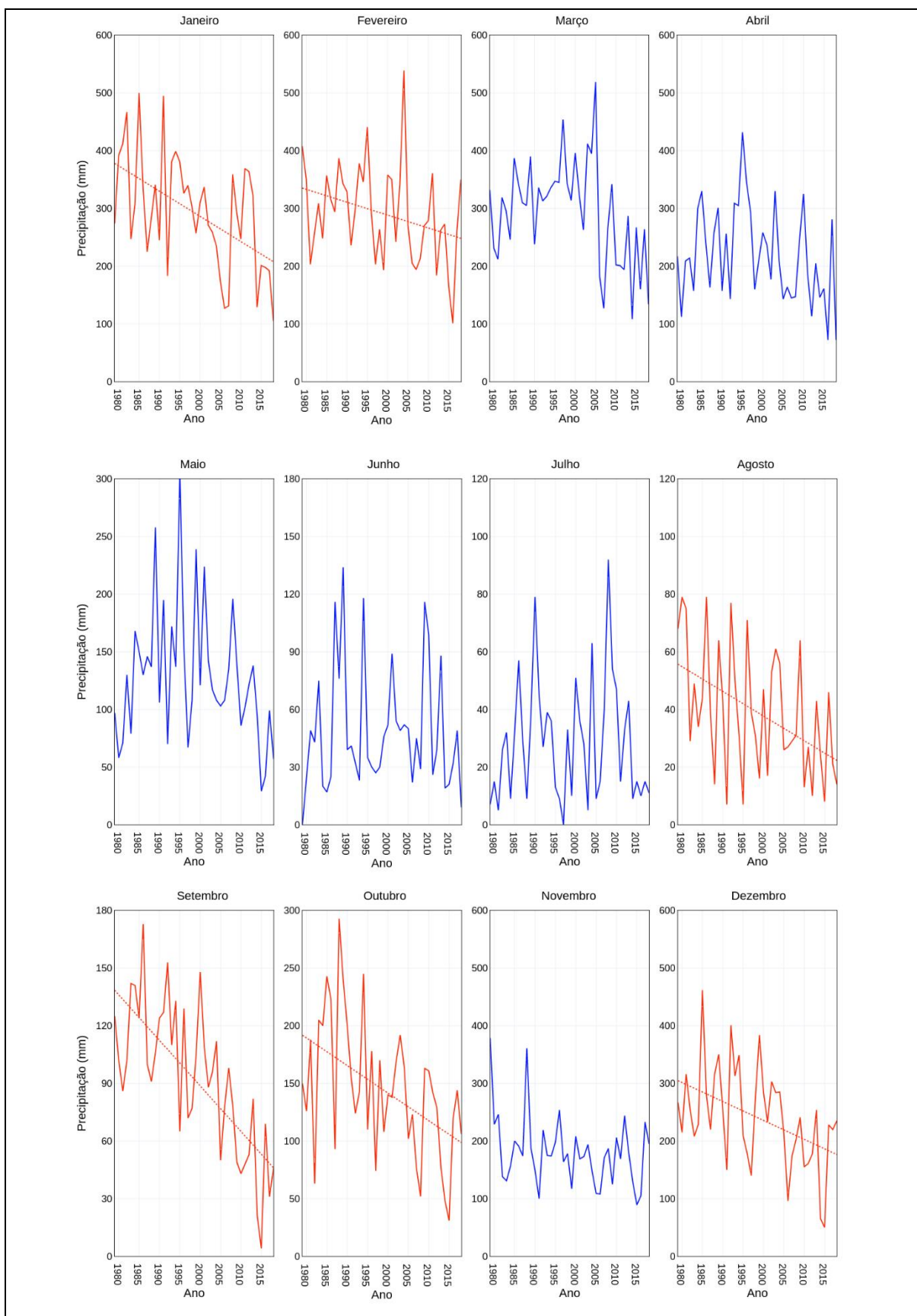


Fonte: Elaborado pelo autor do trabalho.

4.2.4. Análise Estatística de Tendência

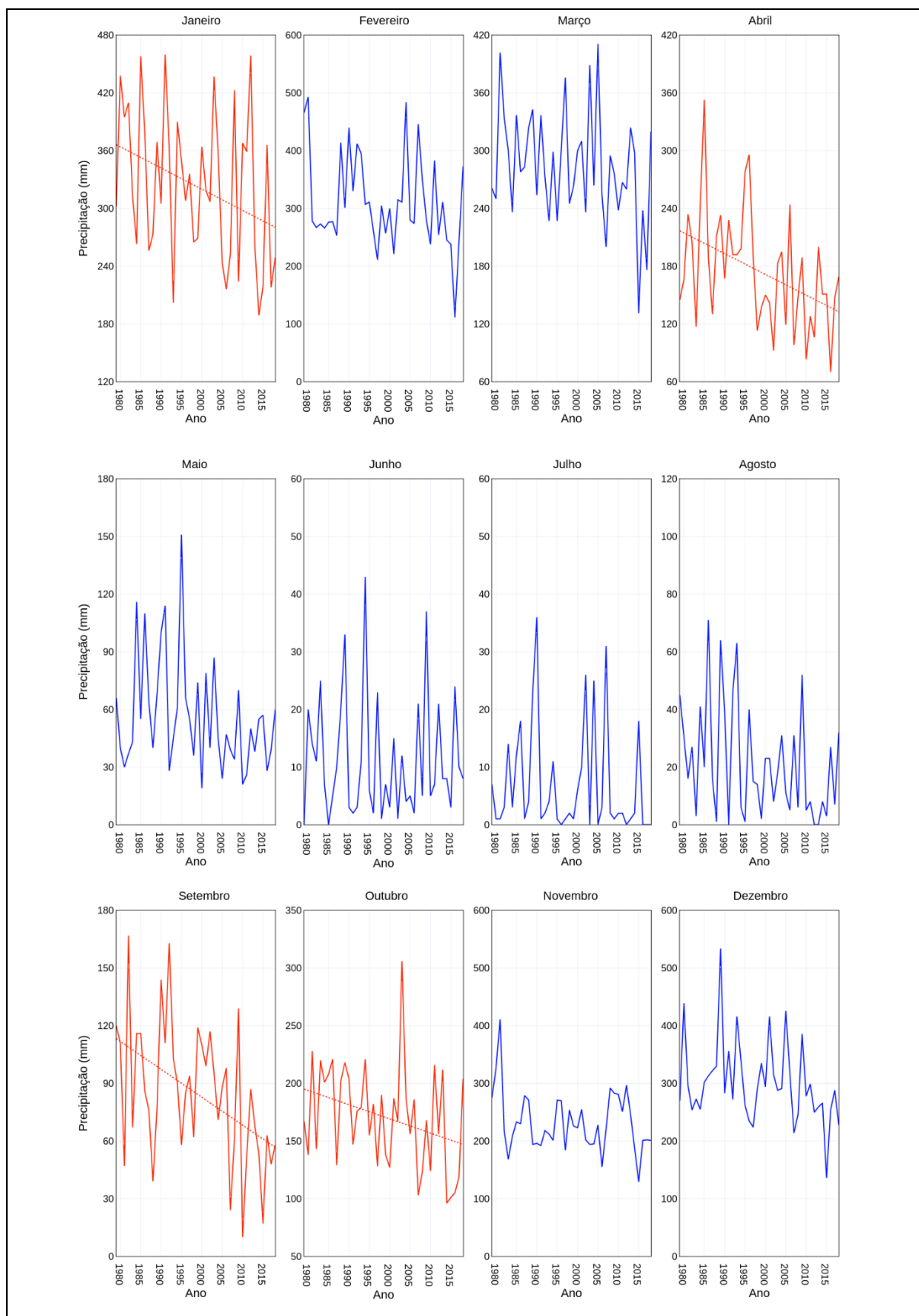
O teste MK foi aplicado para as séries de dados mensais de precipitação em cada sub-bacia e também para o total anual de precipitação, conforme exibido nas Figuras 36 a 43. Os meses em que foi identificada tendência significativa de redução da precipitação mensal ou anual foram traçados em vermelho. Para os meses cuja tendência monotônica foi detectada, a reta de tendência linear foi traçada.

Figura 36 - Série mensal de precipitação na sub-bacia Jat (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.



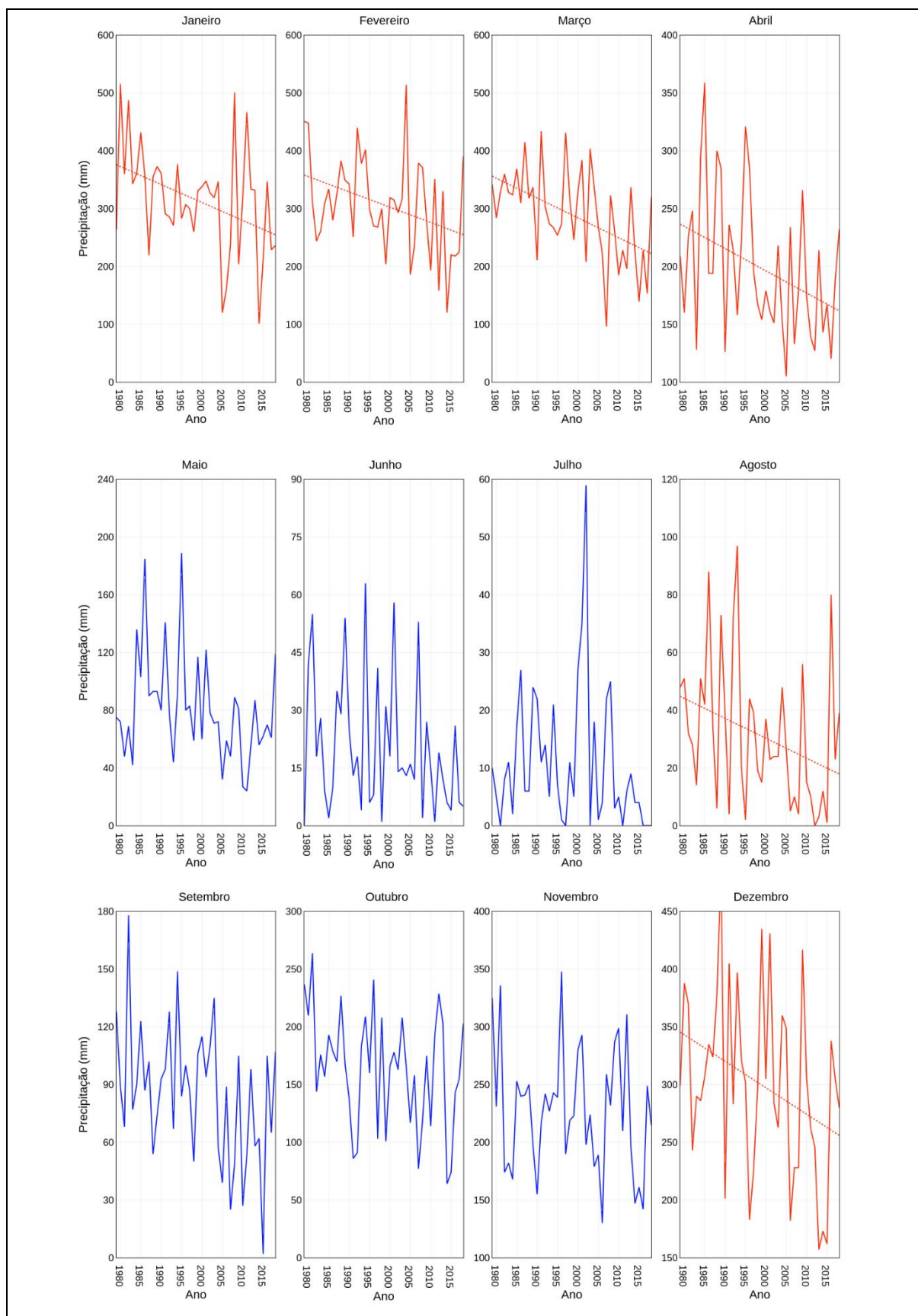
Fonte: Elaborado pelo autor do trabalho.

Figura 37 - Série mensal de precipitação na sub-bacia 444 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.



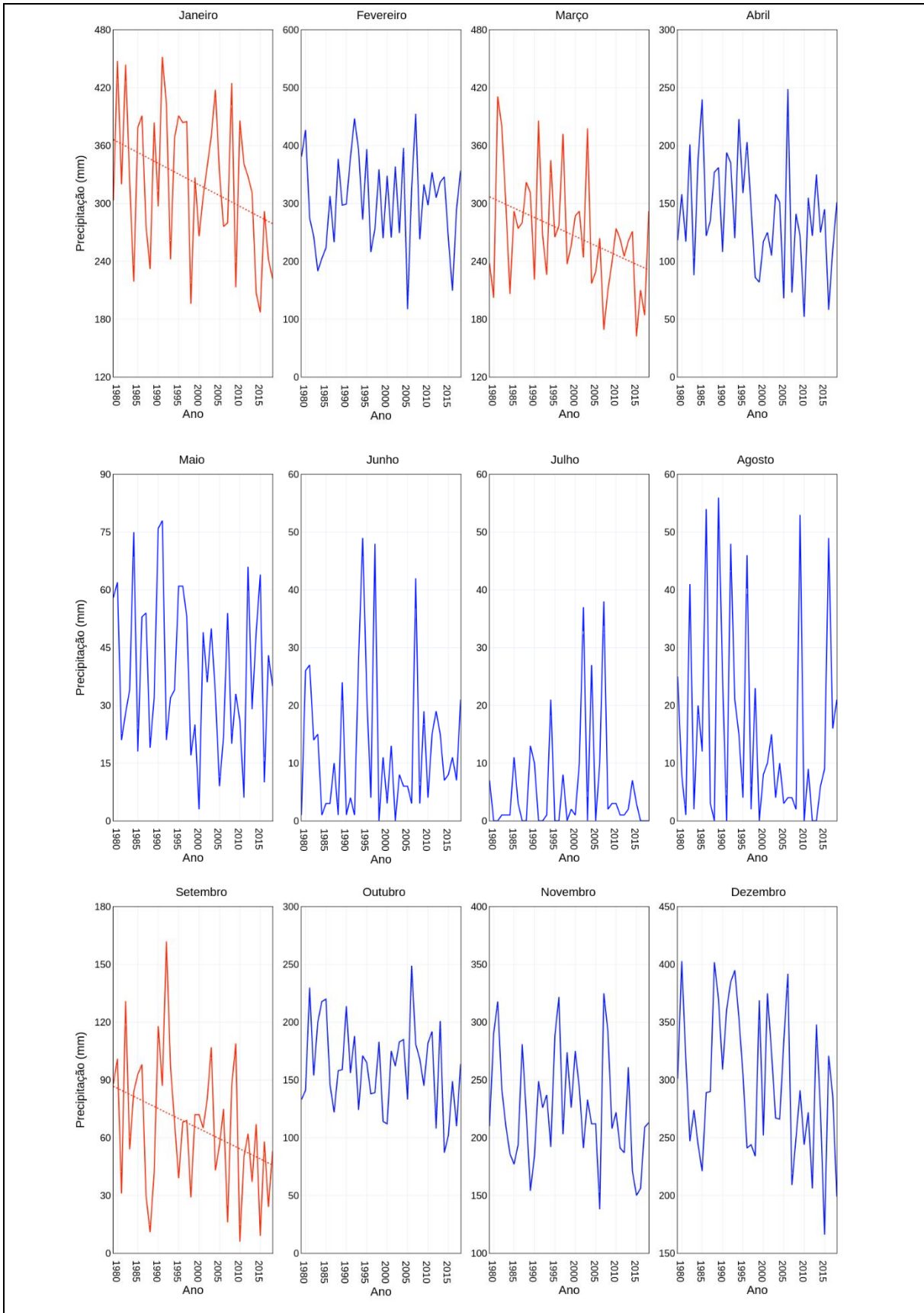
Fonte: Elaborado pelo autor do trabalho.

Figura 38 - Série mensal de precipitação na sub-bacia 445 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.



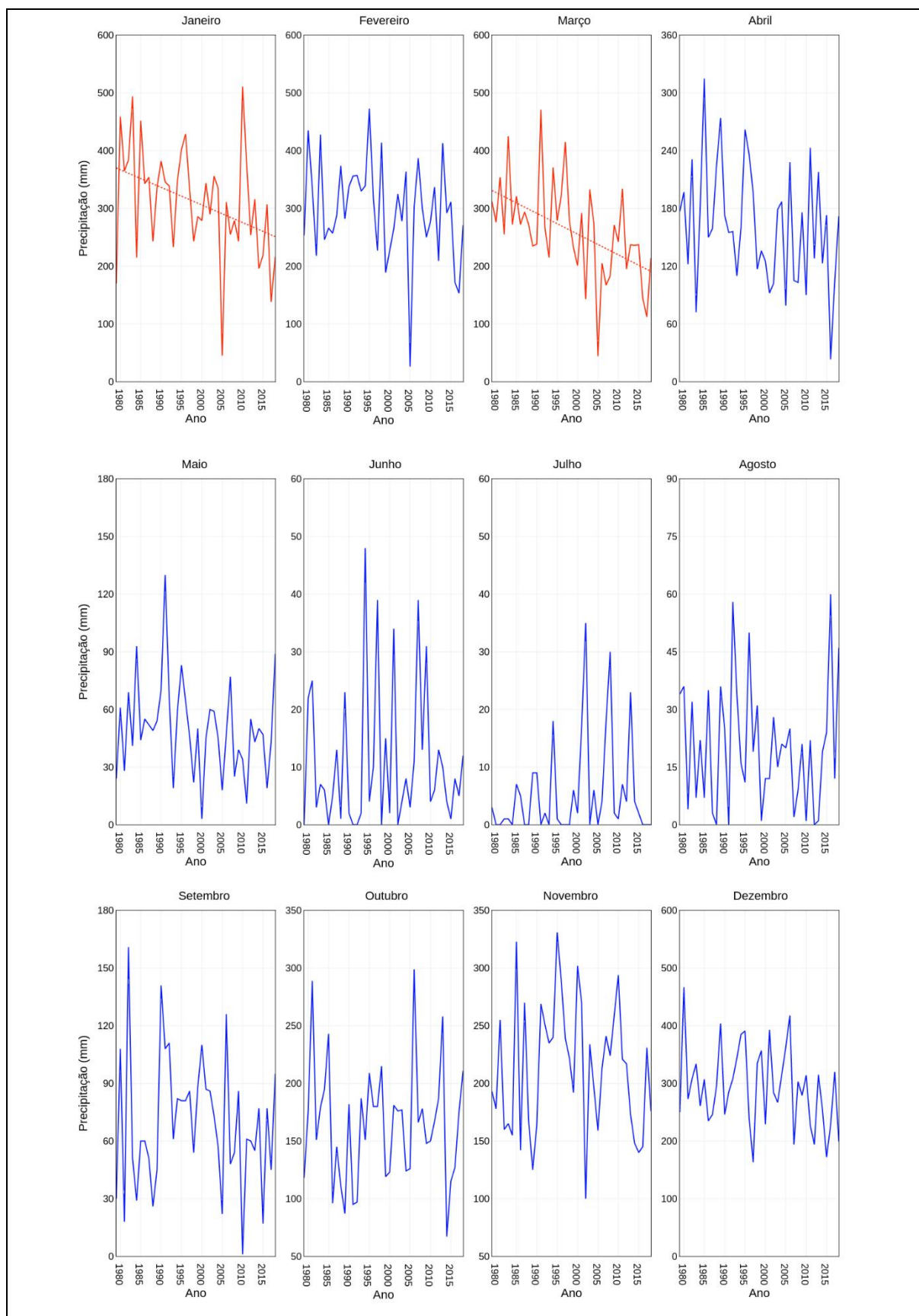
Fonte: Elaborado pelo autor do trabalho.

Figura 39 - Série mensal de precipitação na sub-bacia 446 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.



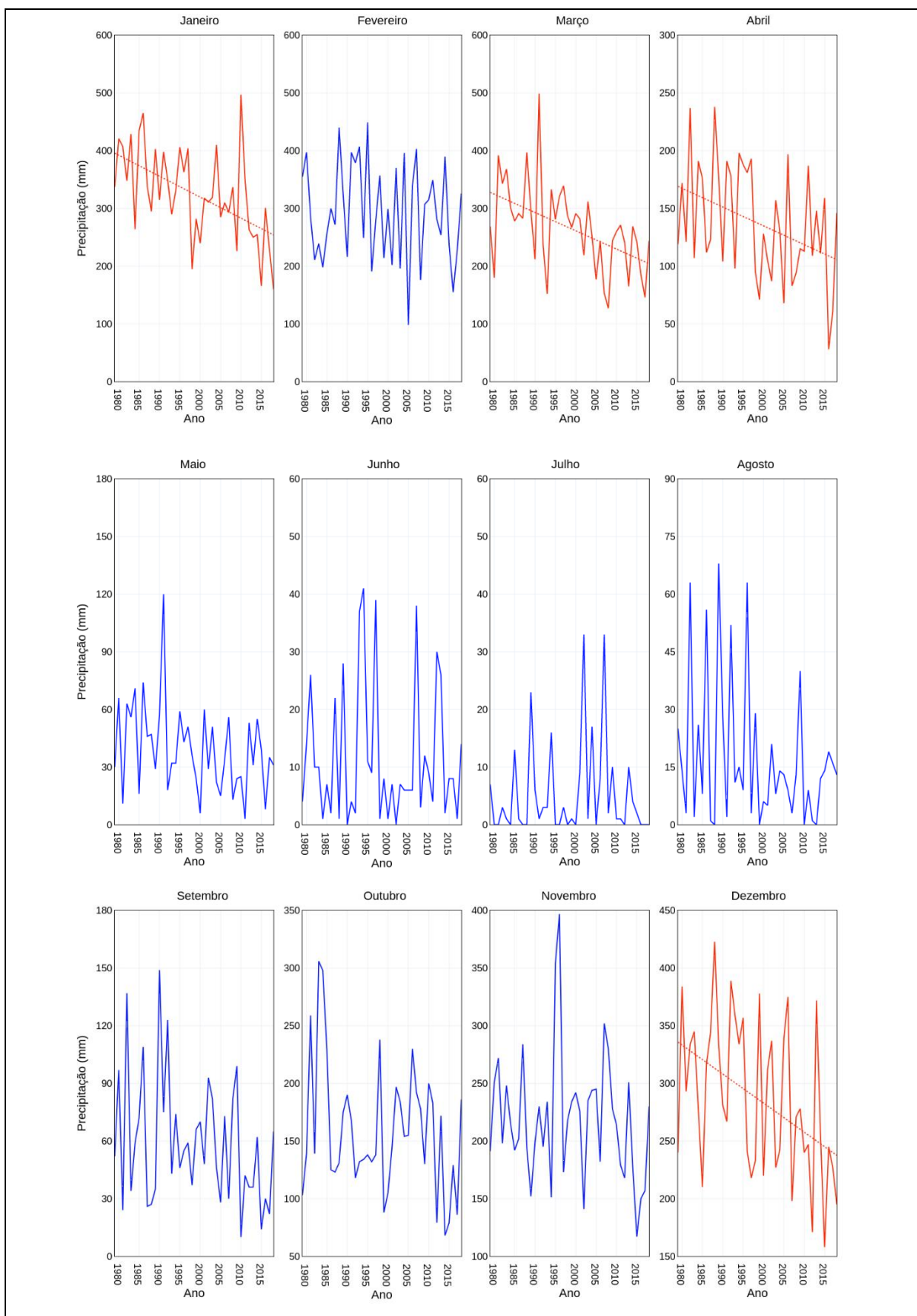
Fonte: Elaborado pelo autor do trabalho.

Figura 40 - Série mensal de precipitação na sub-bacia 447 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.



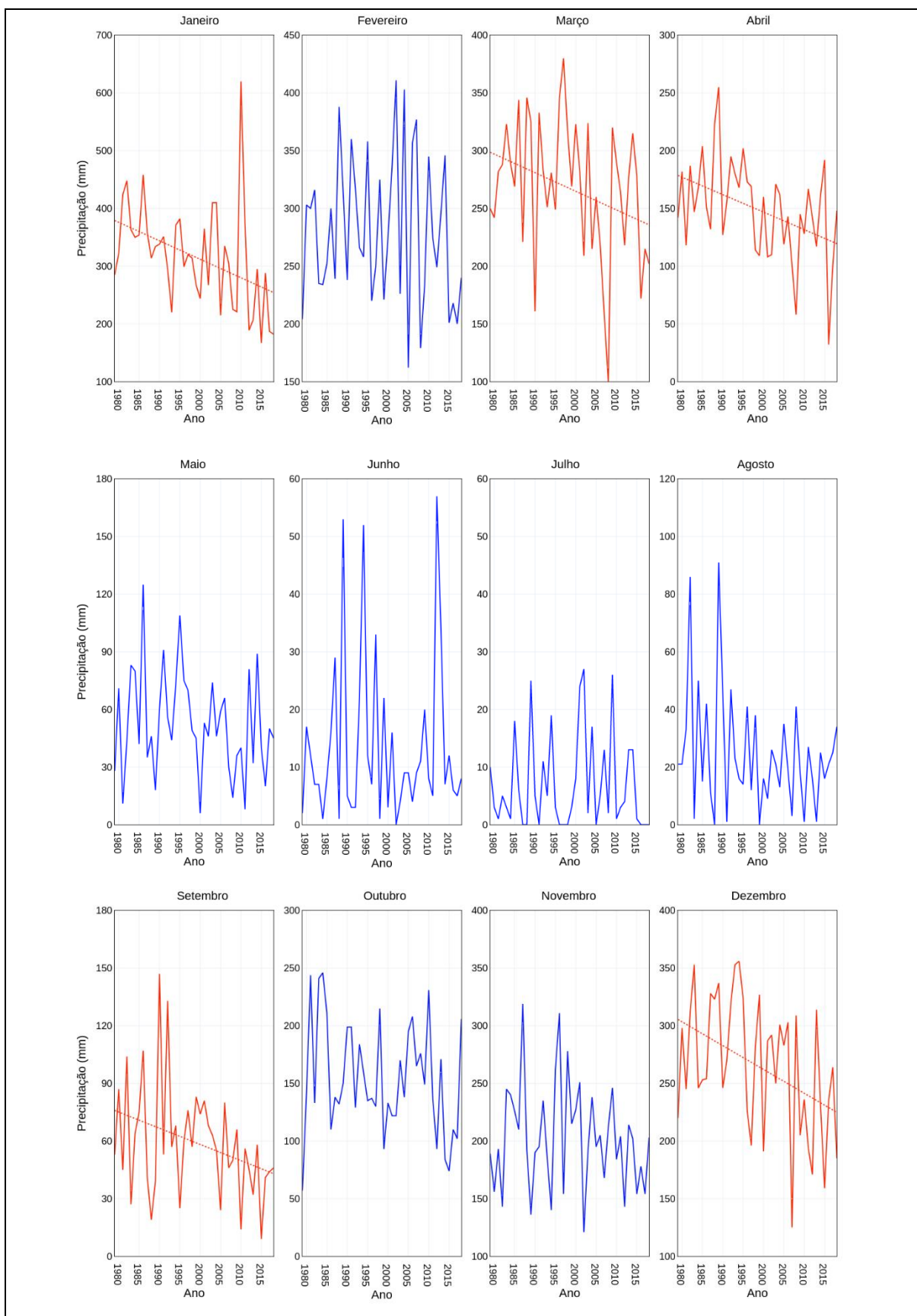
Fonte: Elaborado pelo autor do trabalho.

Figura 41 - Série mensal de precipitação na sub-bacia 448 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.



Fonte: Elaborado pelo autor do trabalho.

Figura 42 - Série mensal de precipitação na sub-bacia 449 (período base-1979-2018). As curvas em vermelho apresentaram tendência significativa de redução da precipitação. A reta tracejada representa a tendência linear da série.



Fonte: Elaborado pelo autor do trabalho.

Os resultados das tendências sistemáticas identificadas usando a estatística MK para a precipitação média mensal em cada sub-bacia do rio Tapajós são apresentados nas Tabelas de 22 a 28. A sequência mostrada nas tabelas refere-se às seguintes informações: mês, mínimo volume de chuva registrado, média do volume de chuva, máximo volume de chuva registrado, variância, tau de Kendall, estatística S de MK, valor p , inclinação de Sen e a tendência.

Em todas as sub-bacias houve tendência de redução de precipitação mensal. As sub-bacias Jat e 445 foram as que apresentaram maior quantidade de meses com tendência de redução da precipitação, ambas com seis meses (Tabela 29). Enquanto a sub-bacia 447 foi a que registrou menor número de meses com tendência à redução da precipitação, apenas dois. A sub-bacia Jat apresentou as tendências mais acentuadas, com inclinação de Sen de até $-4,814$ registrado para o mês de janeiro (Tabela 22).

Os meses com maior volume de precipitação foram os que apresentaram tendência de redução com maior frequência. O mês de janeiro apresentou as maiores magnitudes de tendência de redução, além de ocorrer tal condição em todas as sete sub-bacias. O mês de março registrou redução da precipitação em cinco das sete sub-bacias. Apenas os meses de maio, junho, julho e novembro não registraram qualquer tendência significativa.

Tabela 22 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia Jat no rio Tapajós.

Mês	Mínimo	Média	Máximo	σ	τ	S	p	Sen	Tendência
Jan	105	293	500	99	-0,359	-280	0,001	-4,814	Redução
Fev	101	292	539	84	-0,223	-174	0,044	-2,374	Redução
Mar	108	294	519	90	-0,185	-144	0,096	-2,205	Nula
Abr	72	218	432	81	-0,187	-146	0,091	-2,004	Nula
Mai	29	128	305	58	-0,159	-124	0,152	-1,021	Nula
Jun	0	48	134	33	0,003	2,0	0,991	0,012	Nula
Jul	0	29	92	21	0,062	48,0	0,584	0,067	Nula
Ago	7	39	79	22	-0,315	-246,0	0,004	-0,843	Redução
Set	4	92	173	39	-0,508	-396,0	< 0,0001	-2,291	Redução
Out	31	145	293	59	-0,341	-266,0	0,002	-2,491	Redução
Nov	89	182	379	61	-0,177	-138,0	0,110	-1,151	Nula
Dez	50	241	462	86	-0,279	-218,0	0,011	-3,028	Redução

Fonte: Elaborado pelo autor do trabalho.

Tabela 23 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 444 no rio Tapajós.

Mês	Mínimo	Média	Máximo	σ	τ	S	p	Sen	Tendência
Jan	189	323	460	78	-0,244	-190	0,028	-2,472	Redução
Fev	111	311	493	82	-0,169	-132	0,127	-1,296	Nula
Mar	131	283	411	58	-0,162	-126	0,145	-1,177	Nula
Abr	70	175	353	60	-0,269	-210	0,015	-1,993	Redução
Mai	19	56	151	29	-0,151	-118	0,173	-0,425	Nula
Jun	0	11	43	10	0,044	34	0,701	0,042	Nula
Jul	0	7	36	10	-0,185	-144	0,096	-0,064	Nula
Ago	0	21	71	19	-0,215	-168	0,052	-0,468	Nula
Set	10	85	167	36	-0,323	-252	0,003	-1,458	Redução
Out	96	171	306	45	-0,264	-206	0,017	-1,437	Redução
Nov	129	232	411	51	-0,121	-94	0,279	-0,720	Nula
Dez	136	301	534	72	-0,197	-154	0,075	-1,425	Nula

Fonte: Elaborado pelo autor do trabalho.

Tabela 24 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 445 no rio Tapajós.

Mês	Mínimo	Média	Máximo	σ	τ	S	p	Sen	Tendência
Jan	101	315	516	92	-0,287	-224	0,009	-2,586	Redução
Fev	120	307	514	85	-0,226	-176	0,041	-2,624	Redução
Mar	96	289	434	79	-0,362	-282	0,001	-3,606	Redução
Abr	105	199	359	61	-0,254	-198	0,022	-1,902	Redução
Mai	24	81	189	37	-0,192	-150	0,083	-0,707	Nula
Jun	0	20	63	18	-0,179	-140	0,105	-0,332	Nula
Jul	0	11	59	12	-0,179	-140	0,105	-0,156	Nula
Ago	0	31	97	25	-0,272	-212	0,014	-0,824	Redução
Set	2	86	178	35	-0,190	-148	0,087	-0,945	Nula
Out	64	164	264	50	-0,213	-166	0,055	-1,195	Nula
Nov	130	227	348	54	-0,105	-82	0,345	-0,835	Nula
Dez	157	301	487	80	-0,226	-176	0,041	-2,413	Redução

Fonte: Elaborado pelo autor do trabalho.

Tabela 25 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 446 no rio Tapajós. Fonte: autor.

Mês	Mínimo	Média	Máximo	σ	τ	S	p	Sen	Tendência
Jan	187	323	452	75	-0,221	-172	0,046	-2,656	Redução
Fev	117	304	455	80	0,005	4	0,972	0,200	Nula
Mar	162	269	411	60	-0,249	-194	0,025	-1,772	Redução
Abr	52	141	249	48	-0,174	-136	0,116	-1,171	Nula
Mai	3	39	78	20	-0,105	-82	0,345	-0,305	Nula
Jun	0	13	49	13	0,069	54	0,537	0,085	Nula
Jul	0	6	38	9	-0,038	-30	0,735	-0,004	Nula
Ago	0	16	56	18	-0,092	-72	0,408	-0,112	Nula
Set	6	66	162	35	-0,249	-194	0,025	-1,074	Redução
Out	87	161	249	38	-0,115	-90	0,300	-0,678	Nula
Nov	138	224	325	48	-0,149	-116	0,180	-0,900	Nula
Dez	166	296	403	63	-0,192	-150	0,083	-1,574	Nula

Fonte: Elaborado pelo autor do trabalho.

Tabela 26 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 447 no rio Tapajós.

Mês	Mínimo	Média	Máximo	σ	τ	S	p	Sen	Tendência
Jan	45	311	511	96	-0,318	-248	0,004	-3,806	Redução
Fev	26	297	473	86	-0,144	-112	0,196	-1,538	Nula
Mar	44	261	471	83	-0,369	-288	0,001	-3,298	Redução
Abr	23	161	315	62	-0,203	-158	0,067	-1,726	Nula
Mai	3	50	130	24	-0,149	-116	0,180	-0,377	Nula
Jun	0	11	48	12	0,090	70	0,421	0,090	Nula
Jul	0	5	35	9	0,089	69	0,427	0,013	Nula
Ago	0	20	60	16	-0,064	-50	0,568	-0,130	Nula
Set	1	69	161	35	-0,023	-18	0,843	-0,087	Nula
Out	67	164	299	52	0,026	20	0,825	0,099	Nula
Nov	100	213	331	57	-0,072	-56	0,522	-0,528	Nula
Dez	163	292	467	70	-0,146	-114	0,188	-1,536	Nula

Fonte: Elaborado pelo autor do trabalho.

Tabela 27 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 448 no rio Tapajós.

Mês	Mínimo	Média	Máximo	σ	τ	S	p	Sen	Tendência
Jan	160	325	497	79	-0,400	-312	0,000	-3,983	Redução
Fev	98	293	449	86	-0,085	-66	0,449	-0,936	Nula
Mar	127	266	499	76	-0,387	-302	0,000	-3,166	Redução
Abr	28	137	238	49	-0,226	-176	0,041	-1,474	Redução
Mai	3	39	120	23	-0,213	-166	0,055	-0,581	Nula
Jun	0	12	41	12	-0,018	-14	0,880	-0,016	Nula
Jul	0	5	33	8	-0,009	-7	0,944	-5,313E-4	Nula
Ago	0	17	68	19	-0,079	-62	0,477	-0,104	Nula
Set	10	59	149	33	-0,162	-126	0,145	-0,653	Nula
Out	68	159	306	56	-0,110	-86	0,322	-0,714	Nula
Nov	117	219	397	56	-0,118	-92	0,289	-0,728	Nula
Dez	158	287	423	67	-0,285	-222	0,010	-2,763	Redução

Fonte: Elaborado pelo autor do trabalho.

Tabela 28 - Estatística do teste de Mann-Kendall para a precipitação média mensal na sub-bacia 449 no rio Tapajós.

Mês	Mínimo	Média	Máximo	σ	τ	S	p	Sen	Tendência
Jan	167	317	620	90	-0,372	-290	0,001	-4,295	Redução
Fev	162	281	411	64	-0,067	-52	0,552	-0,514	Nula
Mar	99	267	380	60	-0,221	-172	0,046	-1,391	Redução
Abr	32	149	255	42	-0,272	-212	0,014	-1,375	Redução
Mai	6	52	125	27	-0,110	-86	0,322	-0,450	Nula
Jun	0	13	57	14	0,046	36	0,683	0,044	Nula
Jul	0	7	27	8	-0,072	-56	0,522	-0,015	Nula
Ago	0	24	91	20	-0,090	-70	0,421	-0,232	Nula
Set	9	59	147	29	-0,228	-178	0,039	-0,790	Redução
Out	57	154	246	48	-0,146	-114	0,188	-0,964	Nula
Nov	121	203	319	45	-0,085	-66	0,449	-0,425	Nula
Dez	125	265	356	58	-0,249	-194	0,025	-1,937	Redução

Fonte: Elaborado pelo autor do trabalho.

Tabela 29 - Quadro resumo da tendência estatística observada para a precipitação média mensal nas sub-bacias do rio Tapajós.

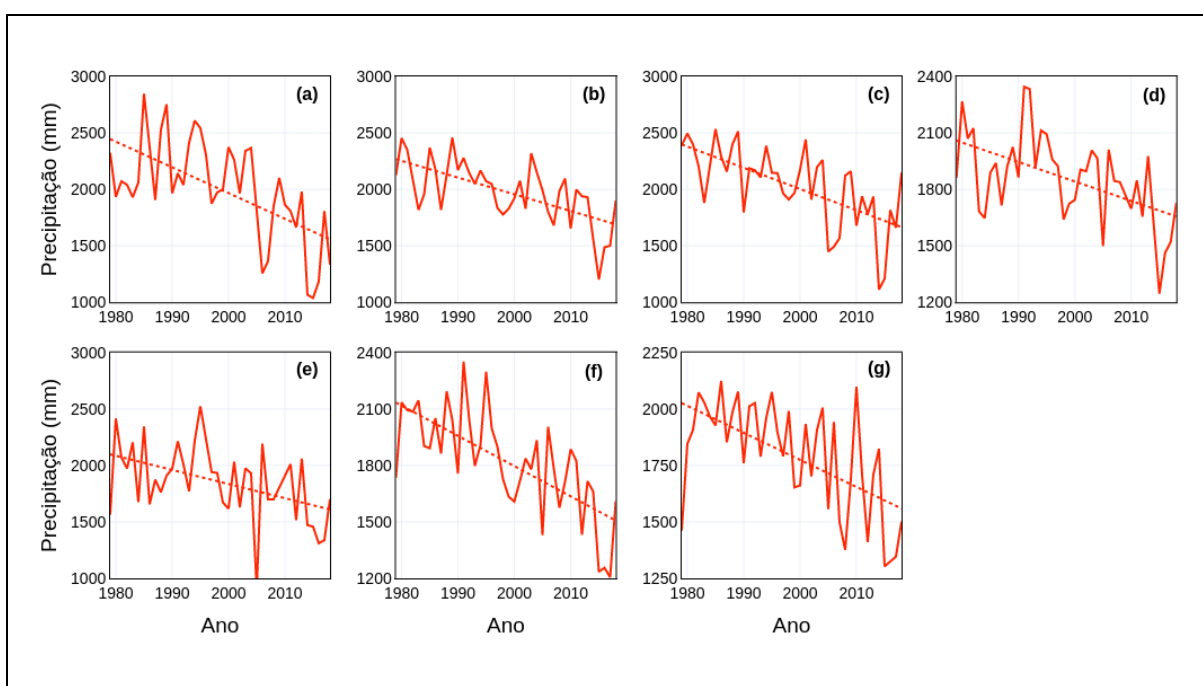
SUB-BACIA	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
JAT	Redução	Redução	-	-	-	-	-	Redução	Redução	Redução	-	Redução
444	Redução	-	-	Redução	-	-	-	-	Redução	Redução	-	-
445	Redução	Redução	Redução	Redução	-	-	-	Redução	-	-	-	Redução
446	Redução	-	Redução	-	-	-	-	-	Redução	-	-	-
447	Redução	-	Redução	-	-	-	-	-	-	-	-	-
448	Redução	-	Redução	Redução	-	-	-	-	-	-	-	Redução
449	Redução	-	Redução	Redução	-	-	-	-	Redução	-	-	Redução

Fonte: Elaborado pelo autor do trabalho.

Dos meses do período chuvoso, o mês de janeiro destaca-se por apresentar tendência de redução em cinco das sete sub-bacias analisadas, como apresentado no quadro resumo da Tabela 29. A sub-bacia 445 apresentou redução no volume de chuvas em quase todos os meses do período chuvoso. Outro fator relevante é que algumas sub-bacias também apresentaram redução no volume de chuvas no período de transição do final da estação seca para o início da estação chuvosa. Considerando o período seco compreendido entre maio e setembro, quando os volumes de precipitação são, em média, inferiores a 100 mm mensais, o mês de setembro configurou-se como o que mais apresentou redução entre os meses da estação seca.

A análise das séries anuais de precipitação para as sete sub-bacias denotou que em todas o volume de precipitação tende a reduzir significativamente ano após ano (Figura 43). As sub-bacias observadas com tendência mais significativa de redução por ordem de magnitude foram a Jat, 445, 448, 444, 449, 447 e 446 (Tabela 29), com a inclinação de Sen variando de $-21,506$ a $-10,067$.

Figura 43 - Série anual de precipitação por sub-bacia (período base-1979-2018). A reta tracejada representa a tendência linear da série. (a) Sub-bacia Jat, (b) sub-bacia 444, (c) sub-bacia 445, (d) sub-bacia 446, (e) sub-bacia 447, (f) sub-bacia 448 e (g) sub-bacia 449.



Fonte: Elaborado pelo autor do trabalho.

Tabela 30 - Estatística do teste de Mann-Kendall para a precipitação média anual.

Mês	Mínimo	Média	Máximo	σ	τ	S	p	Sen	Tendência
JAT	1038	2001	2846	437	-0,423	-330	0,000	-21,506	Redução
444	1202	1978	2458	271	-0,459	-358	< 0,0001	-14,594	Redução
445	1113	2032	2533	344	-0,472	-368	< 0,0001	-17,140	Redução
446	1245	1857	2345	232	-0,344	-268	0,002	-10,067	Redução
447	905	1855	2525	326	-0,285	-222	0,010	-12,247	Redução
448	1206	1819	2351	271	-0,526	-410	< 0,0001	-14,778	Redução
449	1303	1792	2124	242	-0,385	-300	0,000	-12,859	Redução

Fonte: Elaborado pelo autor do trabalho.

4.3. MODELAGEM DO POTENCIAL DE GERAÇÃO DE ENERGIA

Na Tabela 31 são apresentadas algumas das configurações com melhor desempenho testadas usando o método GMDH.

Tabela 31 - Configurações dos hiperparâmetros do modelo usando método GMDH.

Modelo ^(a)	Variáveis entrada ^(b)	Critério	Técnica Validação	Alfa
01	8	Viés + retreinamento	Modo 4.1	1,5
02	10	Teste	Modo 4.1	1,5
03	22	Teste	Modo 4.2	1,0
04	18	Teste	Modo 3.2	1,0
05	20	Teste	Modo 3.1	1,0
06	13	Teste	Modo 2	1,0
07	16	Teste	Modo 1	1,0
08	16	Teste + viés	Modo 3.1	1,0
09	13	Viés	Modo 3.1	1,0
10	16	Teste + viés	Modo 3.1	1,0

Fonte: Elaborado pelo autor do trabalho.

(a) “Modelo” se refere ao número ordinal de apresentação.

(b) “Variáveis de entrada” corresponde ao número de variáveis que o método selecionou dentro da gama disponível de 91 e utilizou no modelo.

As configurações de melhor desempenho usando a técnica de redes neurais artificiais da forma tradicional é apresentada na Tabela 32.

Tabela 32 - Configurações dos hiperparâmetros do modelo usando RNA.

Modelo	Algoritmo	Neurônios ^(c)	Função de Ativação	Taxa Aprend. ^(d)
11	RNA-RB	1	Tangente Hiperbólica	0,3
12	RNA-RB	2	Sigmoide	0,3
13	RNA-RB	3	Tangente Hiperbólica	0,2
14	RNA-RB	4	Sigmoide	0,2
15	RNA-RB	5	Tangente Hiperbólica	0,1
16	RNA-RB	6	Sigmoide	0,1
17	RNA-RB	7	Tangente Hiperbólica	0,1
18	RNA-RB	8	Sigmoide	0,1
19	RNA-RB	9	Tangente Hiperbólica	0,1
20	RNA-RB	10	Sigmoide	0,1
21	RNA-RB	11	Tangente Hiperbólica	0,1
22	RNA-RB	12	Sigmoide	0,1
23	RNA-RB	13	Tangente Hiperbólica	0,1
24	RNA-RB	15	Sigmoide	0,1
25	RNA-RB	20	Tangente Hiperbólica	0,1
26	RNA-RB	30	Tangente Hiperbólica	0,1
27	RNA-LM	1	Sigmoide	0,3
28	RNA-LM	2	Tangente Hiperbólica	0,2
29	RNA-LM	3	Sigmoide	0,1
30	RNA-LM	4	Tangente Hiperbólica	0,1
31	RNA-LM	5	Sigmoide	0,1
32	RNA-LM	6	Tangente Hiperbólica	0,1
33	RNA-LM	7	Sigmoide	0,1
34	RNA-LM	8	Tangente Hiperbólica	0,1
35	RNA-LM	9	Sigmoide	0,1
36	RNA-LM	10	Tangente Hiperbólica	0,1
37	RNA-LM	11	Sigmoide	0,1
38	RNA-LM	12	Sigmoide	0,1
39	RNA-LM	13	Sigmoide	0,1
40	RNA-LM	20	Tangente Hiperbólica	0,1
41	RNA-LM	30	Sigmoide	0,1

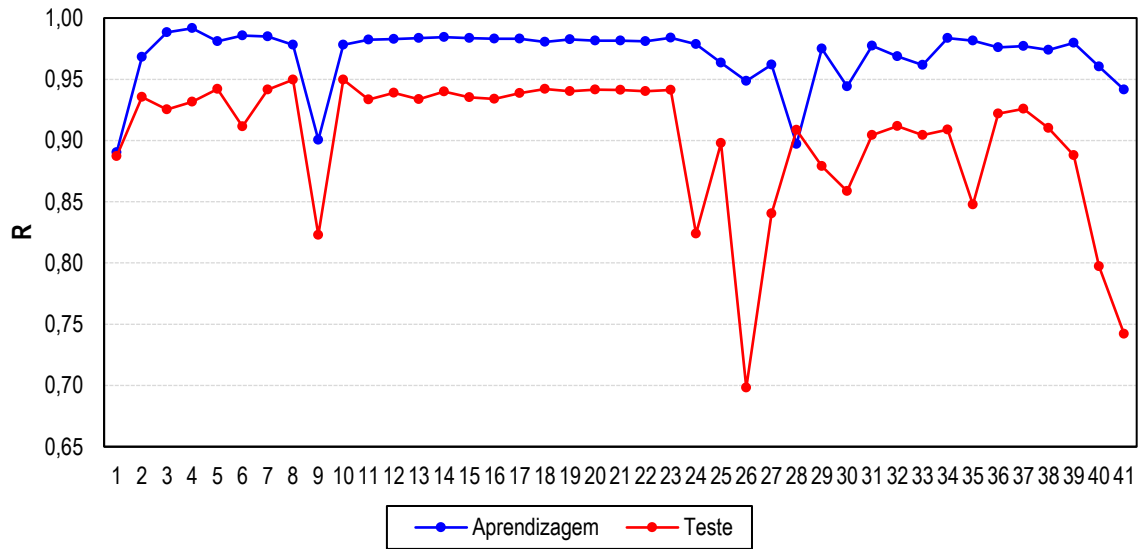
Fonte: Elaborado pelo autor do trabalho.

(c) Refere-se ao número de neurônios na camada escondida.

(d) Taxa de aprendizado é um parâmetro que está ligado à mudança dos pesos sinápticos.

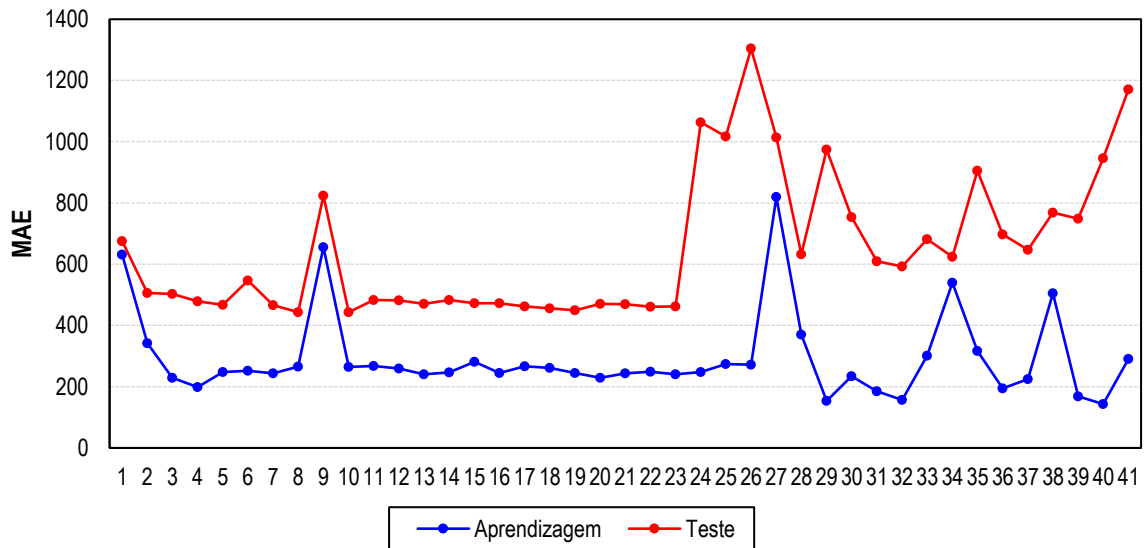
Nas Figuras 44 a 46 são traçados os indicadores (coeficientes de correlação, *MAE* e *MAPE*) para as diferentes arquiteturas, destacando-se que para os modelos do tipo RNA foram executadas 20 repetições da simulação e obtida a sua média.

Figura 44 - Desempenho dos principais modelos com base no coeficiente de correlação (*R*).

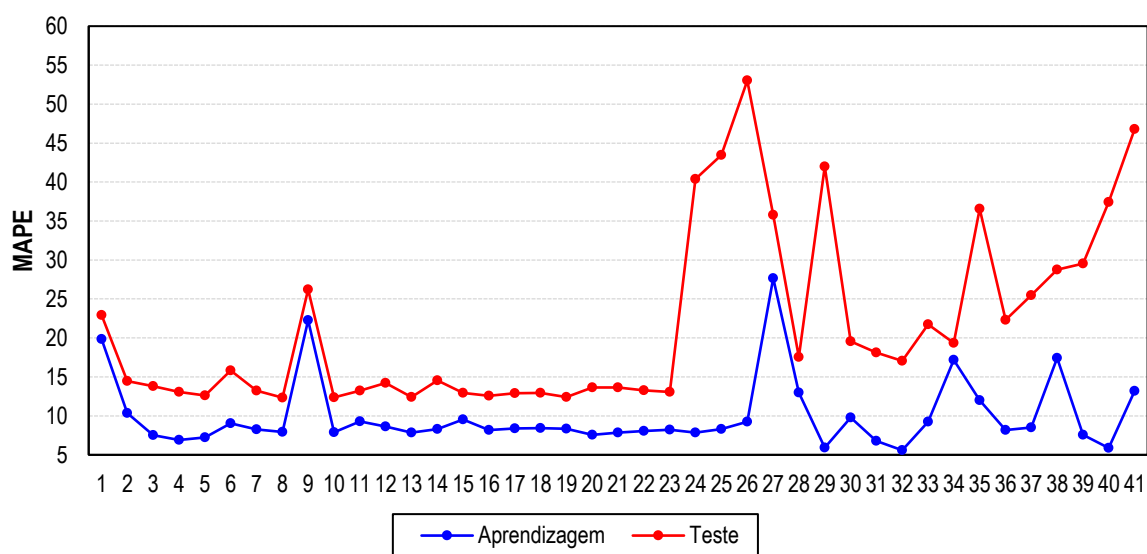


Fonte: Elaborado pelo autor do trabalho.

Figura 45 - Desempenho dos principais modelos com base no *MAE*.



Fonte: Elaborado pelo autor do trabalho.

Figura 46 - Desempenho dos principais modelos com base no *MAPE*.

Fonte: Elaborado pelo autor do trabalho.

A literatura recomenda selecionar a melhor arquitetura de modelagem utilizando os indicadores da fase de treinamento/validação. Entretanto, nem todas as configurações que mostraram bom desempenho durante a fase de aprendizagem repetiram o comportamento na fase de teste. Por isso a escolha do modelo de melhor desempenho levou em consideração também a sua regularidade/estabilidade, comparando-se as diferenças obtidas entre a fase de treinamento e teste.

A melhor arquitetura obtida para a rede polinomial do tipo GMDH selecionou 16 variáveis para construir o modelo a partir de um total de 91 variáveis possíveis, conforme as combinações a seguir: P-JAT₍₁₎, P-JAT₍₈₎, P-JAT₍₁₀₎, P-444, P-444₍₁₎, P-444₍₃₎, P-444₍₁₀₎, P-445, P-445₍₂₎, P-445₍₉₎, P-447, P-447₍₉₎, P-448₍₂₎, P-449, P-449₍₂₎, P-449₍₉₎, onde os índices entre parênteses mostram o tempo de defasagem mensal utilizado, os quais foram selecionados pelo próprio algoritmo. As funções de ativação empregadas foram: linear, linear-covariante e quadrática. Quanto aos modelos de RNA convencionais, o algoritmo Levenberg-Marquardt, aqui denominado de RNA-LM, obteve melhor desempenho usando 6 neurônios na camada escondida. Enquanto o algoritmo usando a regulação bayesiana, denominado RNA-RB, obteve melhor resultado empregando 9 neurônios na camada escondida.

Na Tabela 33 são apresentados os resultados dos modelos para os subconjuntos de treinamento e validação (com 100 repetições para as simulações de RNA) usando algumas métricas objetivas de avaliação de desempenho. Considerando o coeficiente de correlação (*R*) entre o conjunto de dados observados e simulados, todos os modelos mostraram boa destreza,

com ligeira vantagem para RNA-RB, que alcançou correlação média de 98,48%. A RNA-LM mostrou o mais baixo erro absoluto médio (*MAE*) entre os modelos, com média de 243 MW. E considerando o critério do erro percentual médio (*MAPE*), o GMDH obteve melhor desempenho, com 7,92 %.

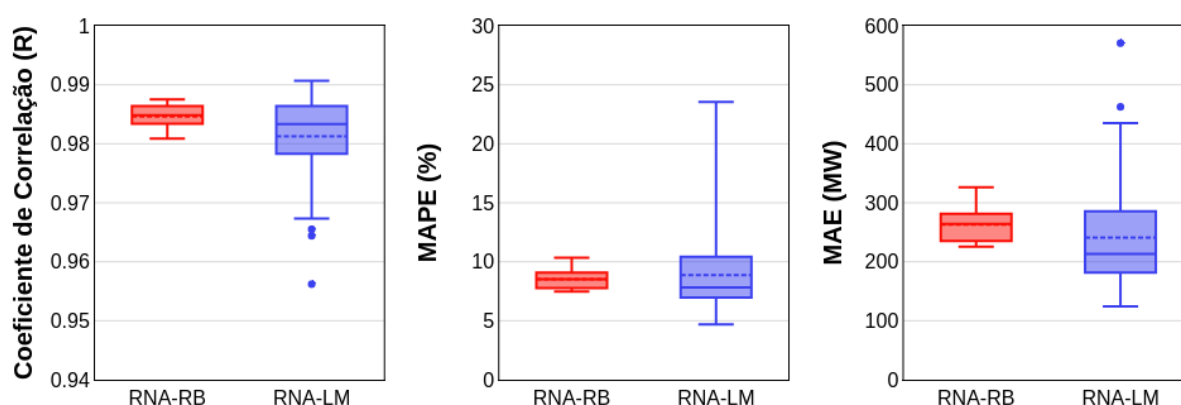
Tabela 33 - Resultados para a avaliação dos modelos durante a fase de treinamento/validação. DP refere-se ao desvio padrão da métrica de avaliação no índice.

Modelo	<i>R</i>	DP_R	<i>MAE</i>	DP_{MAE}	<i>MAPE</i>	DP_{MAPE}
GMDH	0,9782	-	265	-	7,92	-
RNA-RB	0,9848	0,0017	262	24,83	8,50	0,67
RNA-LM	0,9811	0,0070	243	91,93	8,86	3,25

Fonte: Elaborado pelo autor do trabalho.

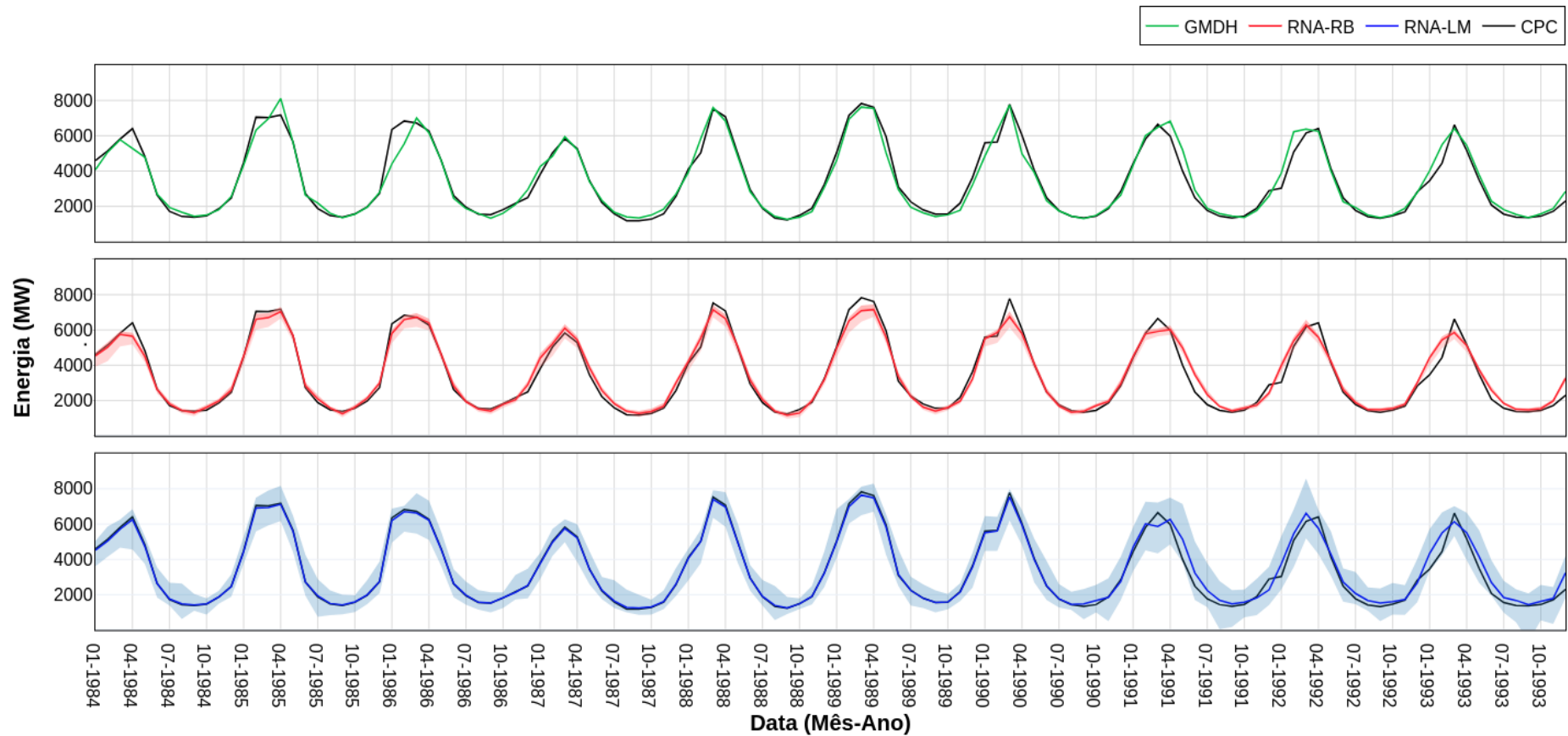
Para melhor compreensão da Tabela 33 diante das diversas simulações realizadas pelas RNA, a Figura 47 mostra um *box plot* para caracterizar o desempenho dos modelos RNA-RB e RNA-LM acerca dos três critérios de avaliação de desempenho. Claramente o modelo RNA-LM apresentou maior dispersão, com vários *outliers*, corroborando com as medidas de desvio padrão apresentadas anteriormente, fazendo com que a diferença entre a média e mediana nos resultados denote grande assimetria dos dados. A simulação da série temporal para o potencial mensal médio de geração energia de hidrelétrica para o conjunto de aprendizado (treinamento e validação) é mostrado na Figura 48.

Figura 47 - Resultados de desempenho durante a fase de aprendizado para RNA-RB e RNA-LM sobre 100 repetições de simulação para cada modelo. As linhas horizontais tracejada e contínua dentro da na caixa representam a média e mediana, respectivamente.



Fonte: Elaborado pelo autor do trabalho.

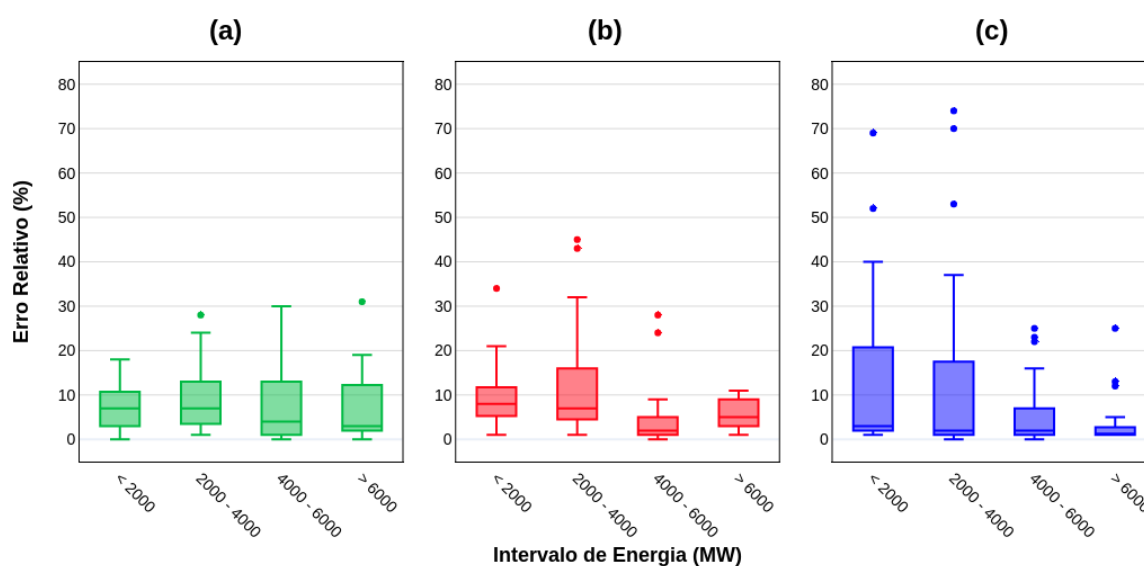
Figura 48 - Comparativo entre os valores observados (CPC) e previstos pelos modelos (GMDH, RNA-LM e RNA-RB) durante o período de treinamento/validação. As áreas sombreadas nas curvas inferior e do centro representam as 100 curvas de repetição dos modelos com inicialização randômica dos pesos para cada uma das rodadas.



Fonte: Elaborado pelo autor do trabalho.

Como os dados simulados podem apresentar pequena dispersão em um intervalo de valores para a geração de energia, mas grande dispersão em outra faixa de intervalo, na Figura 49 são apresentados dados estatísticos que mostram os limites e distribuição de erro dos modelos para quatro gamas distintas de valores: inferior a 2000, entre 2000 e 4000, entre 4000 e 6000, e superior a 6000. Visto que o erro relativo é fortemente dependente do dado real, ele fornece uma ideia da acurácia da previsão.

Figura 49 - Distribuição do erro relativo por faixa de energia para o conjunto de dados de aprendizagem. (a) GMDH, (b) RNA-RB e (c) RNA-LM.



Fonte: Elaborado pelo autor do trabalho.

O primeiro intervalo (< 2000) é o mais importante para a previsão porque está associado com os valores mínimos, que deveriam definir a energia firme da usina. Portanto, é fundamental que os modelos apresentem comportamento consistente neste intervalo. O GMDH mostrou distribuição simétrica para o erro relativo no intervalo inferior a 2000, centrado em 7% (mediana) e com máximo de 18%, e uma dispersão pequena, condições desejáveis para o modelo. Para a segunda categoria (2000 – 4000), o erro permaneceu centrado em 7%, embora a dispersão tenha sido um pouco maior. No intervalo seguinte (4000 – 6000) a distribuição do erro relativo foi bastante assimétrica, com erro alcançando até 30%. E para o último intervalo que corresponde aos máximos valores, houve um aumento da assimetria e, exceto pelos *outliers*, o erro relativo máximo ficou abaixo de 20%.

Considerando o modelo RNA-RB, a distribuição do erro foi bastante simétrica para o intervalo inferior, com erro centrado em 8% e *outside* alcançando 21%, embora tenha ocorrido

um valor considerado *outlier*. Para o intervalo seguinte (2000 – 4000) houve marcada assimetria e grande dispersão, com *outside* acima de 30%. No intervalo entre 4000 e 6000, a RNA-BR mostrou seu melhor desempenho, com um erro máximo inferior a 10% e baixa dispersão, embora ainda tenham sido observados dois valores considerados *outliers*. E para os valores mais altos, o erro relativo foi em torno de 5%, com baixa dispersão e variações de até 11%.

O terceiro modelo, RNA-LM, embora tenha apresentado erros muito pequenos durante a fase de treinamento, a validação influenciou claramente o desempenho geral do modelo durante o aprendizado, fazendo o erro apresentar grande variância, especialmente nas primeiras duas classes de intervalo, alcançando até 40% de erro. Enquanto nas últimas duas classes de intervalo, mesmo com uma redução do erro, especialmente para os valores máximos, ainda ocorreram muitos *outliers*.

Em relação ao conjunto de dados de teste, que foi usado para uma avaliação sem viés do modelo final ajustado, a Tabela 34 apresenta os resultados da acurácia dos modelos desenvolvidos. O modelo GMDH apresentou o melhor desempenho comparado aos outros, embora seus indicadores tenham apresentado pequena diferença com relação ao RNA-BR. Esses indicadores mostram que o modelo GMDH apresentou melhor capacidade de generalização, porque seu desempenho durante a fase de teste obteve uma menor variação comparada a fase de treinamento, mantendo uma boa acurácia, com coeficiente de correção de 0,9496, erro médio absoluto de 443 MW e somente 12,34 % de erro médio percentual.

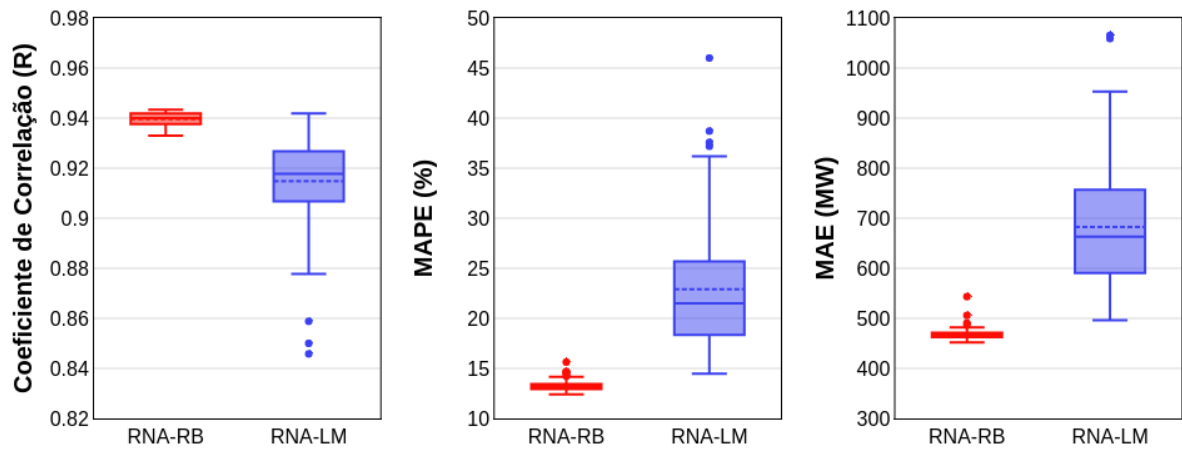
Tabela 34 - Resultados para avaliação dos modelos no período de teste. DP refere-se ao desvio padrão da métrica de avaliação no índice.

Modelo	R	DP_R	MAE	DP_{MAE}	$MAPE$	DP_{MAPE}
GMDH	0,9496	-	443	-	12,34	-
RNA-RB	0,9397	0,0024	468	11,31	13,28	0,53
RNA-LM	0,9148	0,0182	683	117,89	22,91	6,02

Fonte: Elaborado pelo autor do trabalho.

A Figura 50 caracteriza o comportamento das RNA na fase de teste, considerando 100 repetições para as simulações mantendo as arquiteturas utilizadas em RNA-LM e RNA-RB. O modelo RNA-LM mostrou grande variância em seu comportamento, com expressiva dispersão para qualquer que seja a métrica de avaliação e frequentes *outliers*, denotando a sua dificuldade para obter uma generalização adequada. Por outro lado, a RNA-RB apresentou desempenho claramente superior, capturando melhor o comportamento da variável, resultando em baixa variância nas simulações e, conseqüentemente, uma generalização mais apropriada do modelo.

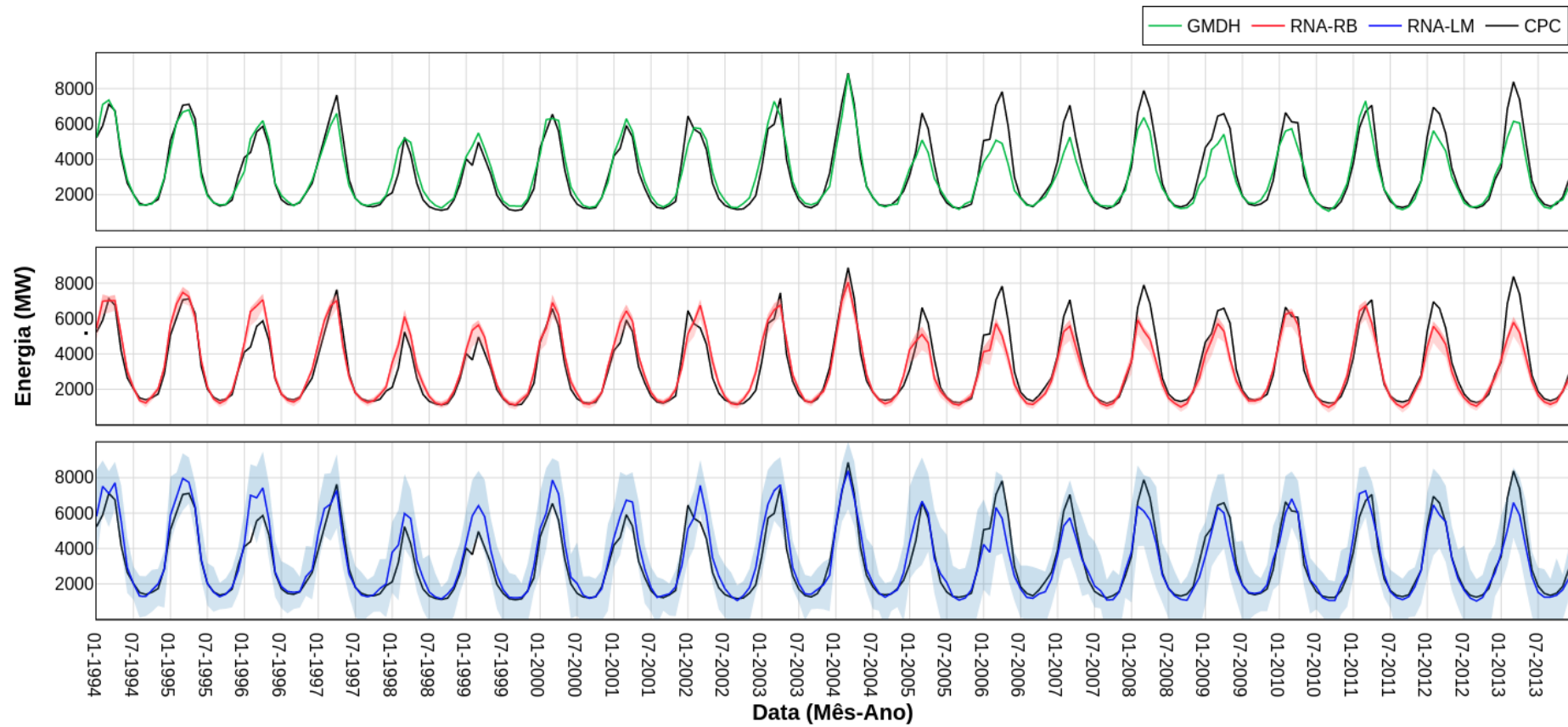
Figura 50 - Resultados de desempenho durante a fase de teste para RNA-RB e RNA-LM sobre 100 repetições de simulação para cada modelo. As linhas horizontais tracejada e contínua dentro da na caixa representam a média e mediana, respectivamente.



Fonte: Elaborado pelo autor do trabalho.

A habilidade dos modelos pode ser observada na Figura 51, que se refere ao subconjunto de teste.

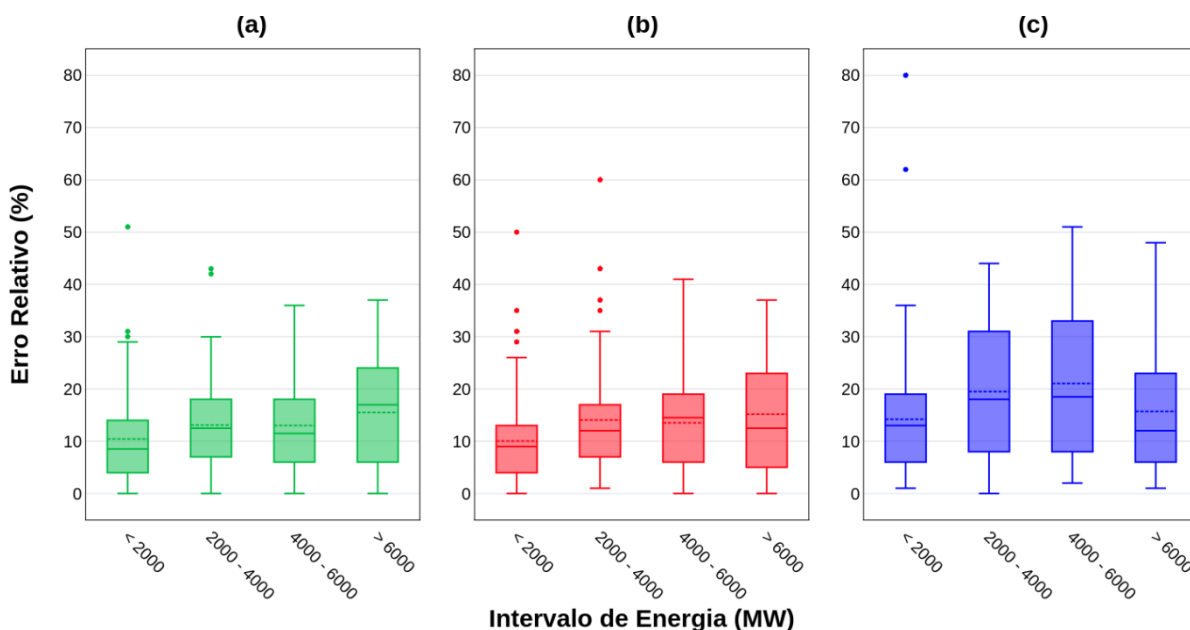
Figura 51 - Comparativo entre os valores observados (CPC) e previstos (GMDH, RNA-LM e RNA-RB) durante o período de teste. As áreas sombreadas nas curvas inferior e do centro representam as 100 curvas de repetição dos modelos com inicialização randômica dos pesos



Fonte: Elaborado pelo autor do trabalho.

A partir do subconjunto de dados de teste, na Figura 52 é mostrada a variação do erro relativo por intervalo de valores. Para o intervalo inferior, o modelo GMDH apresentou melhor desempenho do que os outros, com mediana próxima de 8%, mas seguido de muito próximo pelo RNA-RB. Em relação às classes seguintes, a mediana variou entre 11% e 17% para o modelo GMDH, enquanto para o modelo RNA-RB mostrou menor oscilação, variando entre 12% e 14%. Para a segunda e terceira categoria, houve alternância de melhor desempenho entre RNA-RB e GMDH. Enquanto para a última categoria ambos os modelos apresentaram desempenhos semelhantes. O modelo RNA-LM apresentou comportamento inferior, com grande intervalo interquartil e pior desempenho comparativo para todas as classes, especialmente para a primeira (< 2000), que é a mais crítica.

Figura 52 - Distribuição do erro relativo por faixa de energia para o subconjunto de dados de teste. (a) GMDH, (b) RNA-RB e (c) RNA-LM.



Fonte: Elaborado pelo autor do trabalho.

4.4. CENÁRIOS FUTUROS

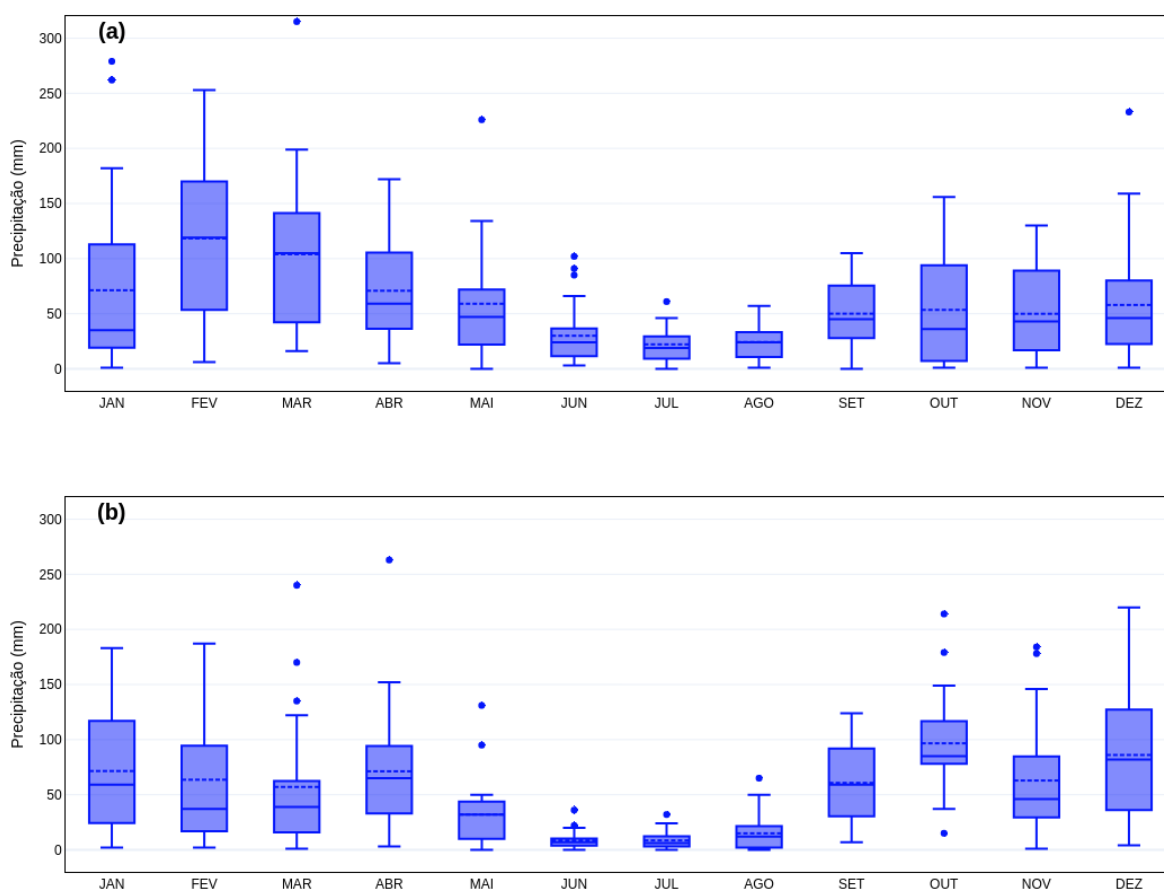
Para simulação do potencial de geração de energia para cenários futuros, foi adotado o modelo GMDH devido ter apresentado melhor índice de desempenho que os demais. As simulações foram realizadas para dois tipos de cenários, no qual o primeiro considera a forçante de emissões de gases de efeito estufa RCP4.5, e o segundo leva em consideração as tendências estatísticas detectadas na seção 5.2.3.

4.4.1. Cenário 1: RCP4.5

O Cenário RCP4.5 utilizou a precipitação simulada pelo modelo climático RegCM4 como dado de entrada para o modelo GMDH, fins gerar a previsão do potencial para geração de energia hidrelétrica para as próximas décadas. Porém, antes da simulação preditiva foi realizada uma análise comparativa entre os dados reais de precipitação observada e os dados sintéticos de precipitação gerada pelo modelo climático RegCM4 para avaliar a sua destreza em simular o clima passado (precipitação).

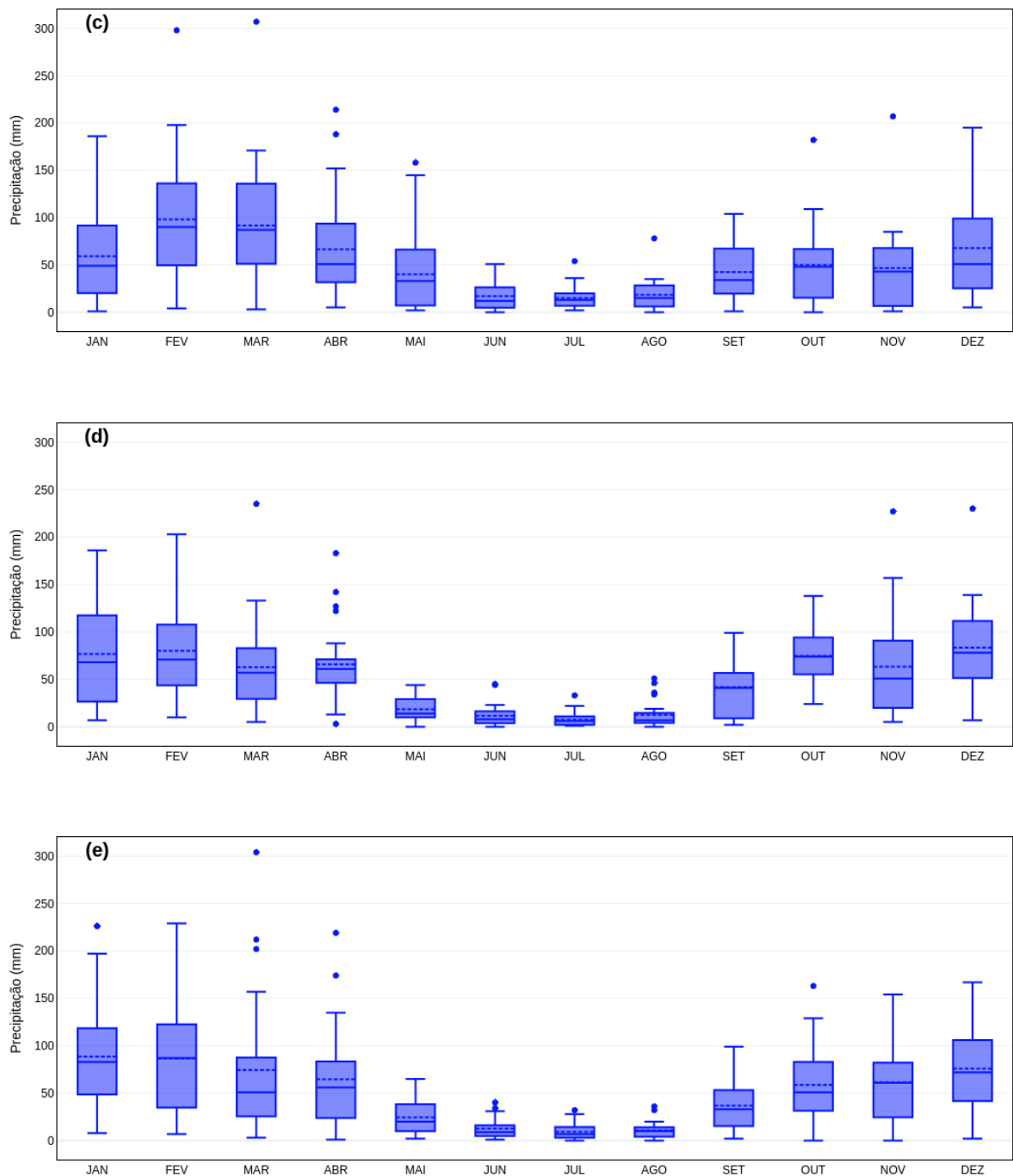
Inicialmente foi calculado o erro absoluto do modelo para cada sub-bacia de estudo, conforme apresentado nas Figuras 53 a 55. Os maiores erros ocorrem durante o período chuvoso, naturalmente quando há maior variabilidade e, também maior dispersão nos dados simulados.

Figura 53 - Erro absoluto do modelo climático RegCM4 para simulação da precipitação média mensal por sub-bacia, considerando o período base 1981-2005. (a) Sub-bacia Jat e (b) sub-bacia 444.



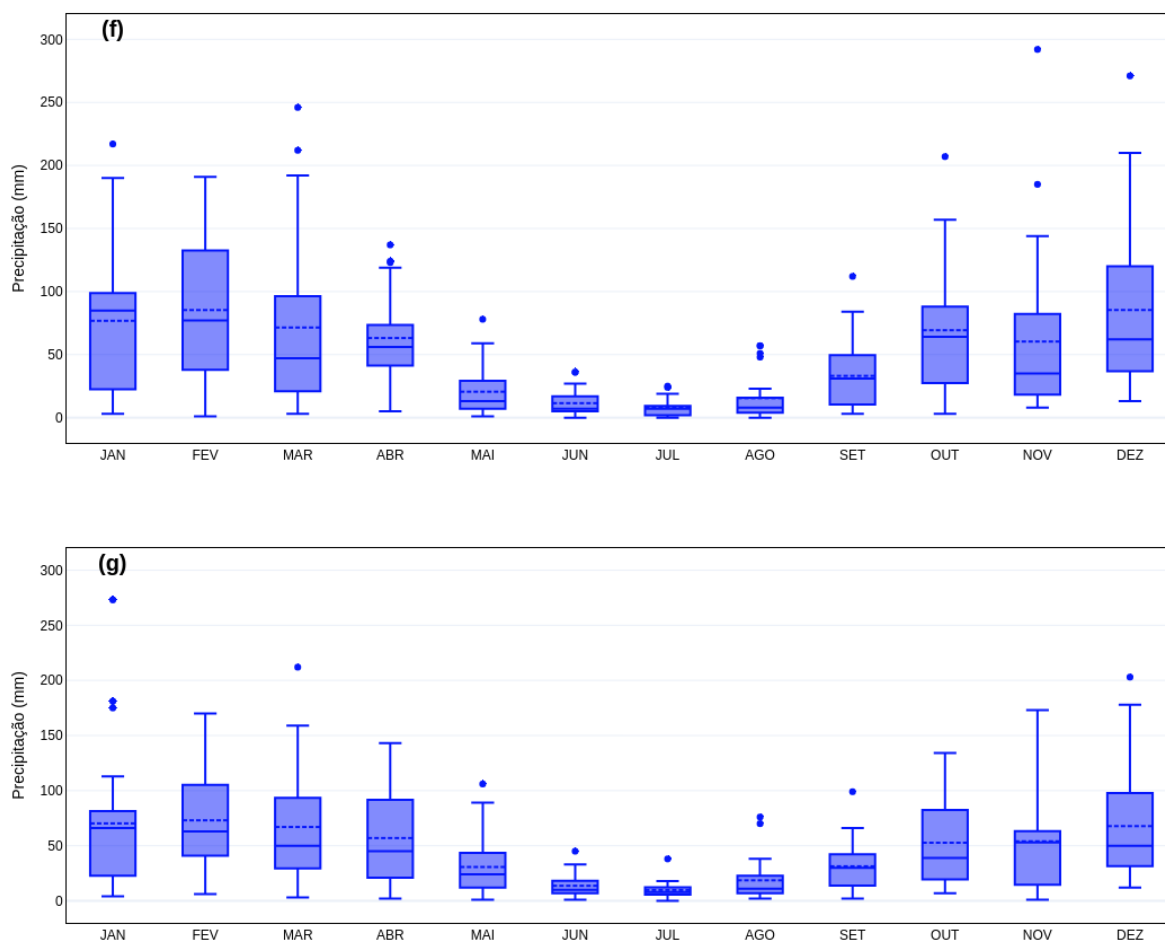
Fonte: Elaborado pelo autor do trabalho.

Figura 54 - Erro absoluto do modelo climático RegCM4 para simulação da precipitação média mensal por sub-bacia, considerando o período base 1981-2005. (c) sub-bacia 445, (d) sub-bacia 446 e (e) sub-bacia 447.



Fonte: Elaborado pelo autor do trabalho.

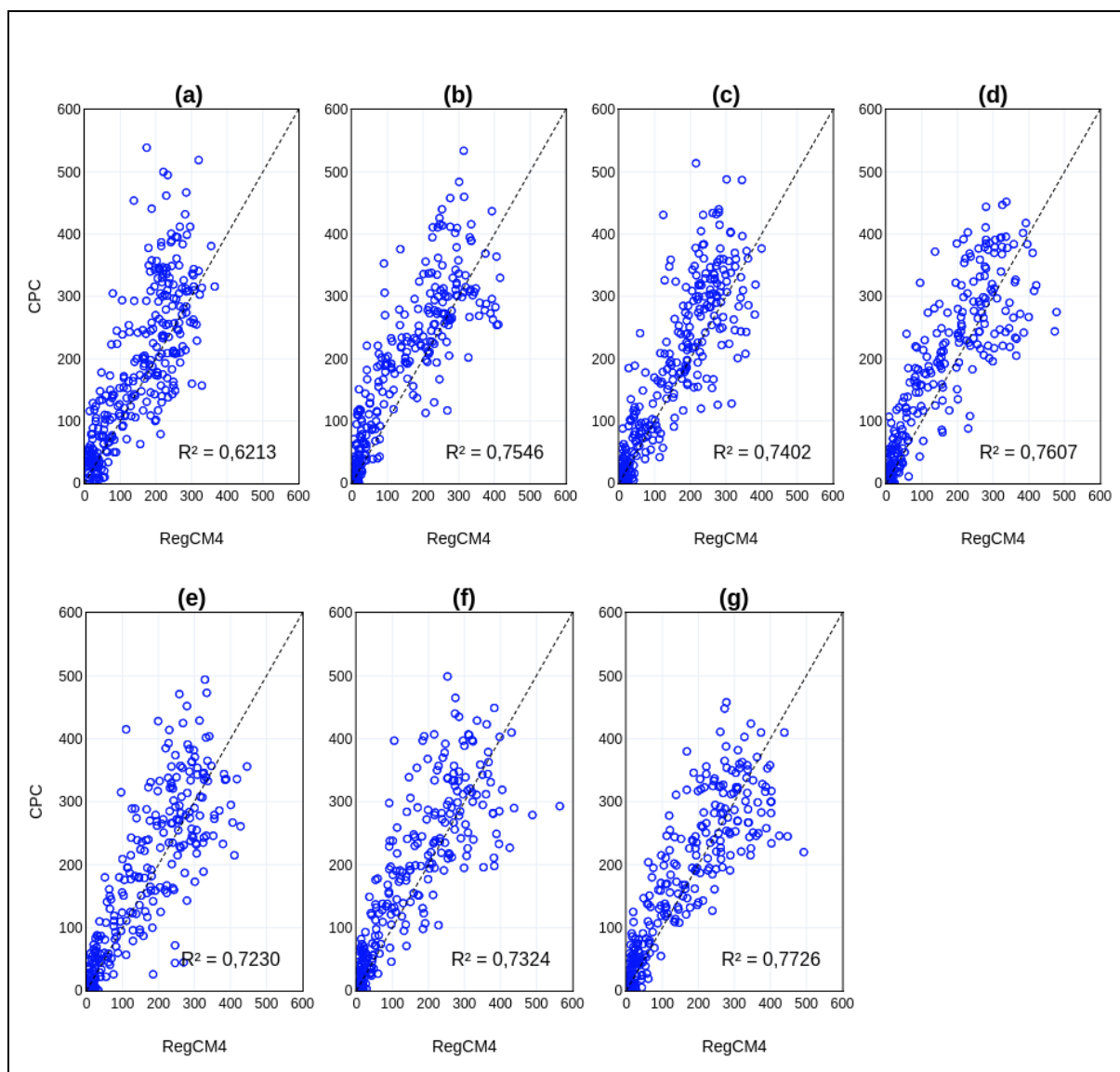
Figura 55 - Erro absoluto do modelo climático RegCM4 para simulação da precipitação média mensal por sub-bacia, considerando o período base 1981-2005. (f) sub-bacia 448 e (g) sub-bacia 449.



Fonte: Elaborado pelo autor do trabalho.

Em seguida foi avaliada a correlação entre as séries através de um diagrama de dispersão (Figura 56), o qual mostrou que há correspondência entre os dados observados (CPC) e simulados (RegCM4), com coeficientes de determinação variando entre as sub-bacias de 0,6213 até 0,7726, com melhor desempenho para a simulação na sub-bacia 449. Também pode ser observado um ligeiro viés negativo na simulação climática.

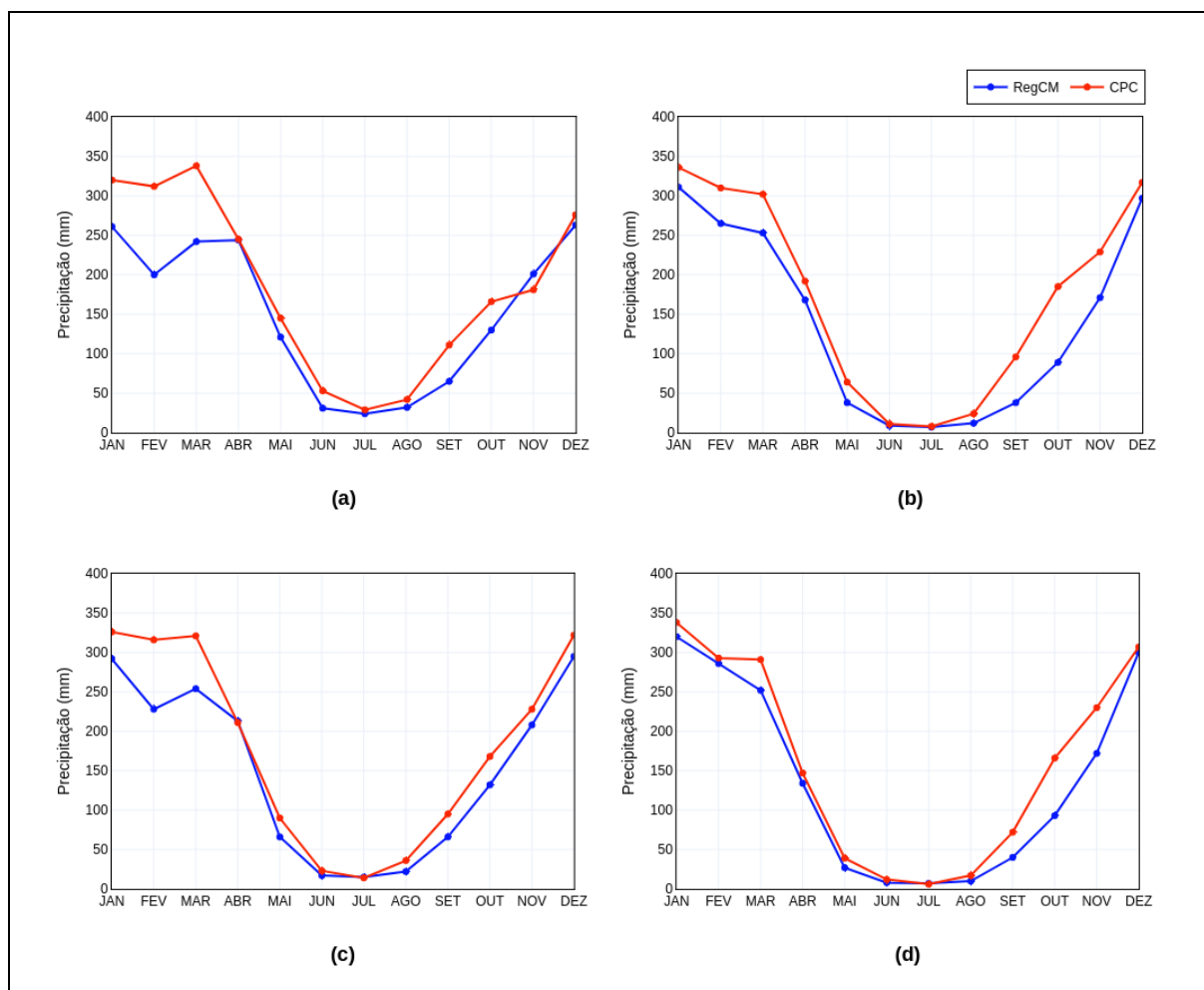
Figura 56 - Diagrama de dispersão entre os dados observados do CPC e a simulação usando o modelo climático RegCM4, por sub-bacia considerando o período base 1981-2005. (a) Sub-bacia Jat, (b) sub-bacia 444, (c) sub-bacia 445, (d) sub-bacia 446, (e) sub-bacia 447, (f) sub-bacia 448 e (g) sub-bacia 449.



Fonte: Elaborado pelo autor do trabalho.

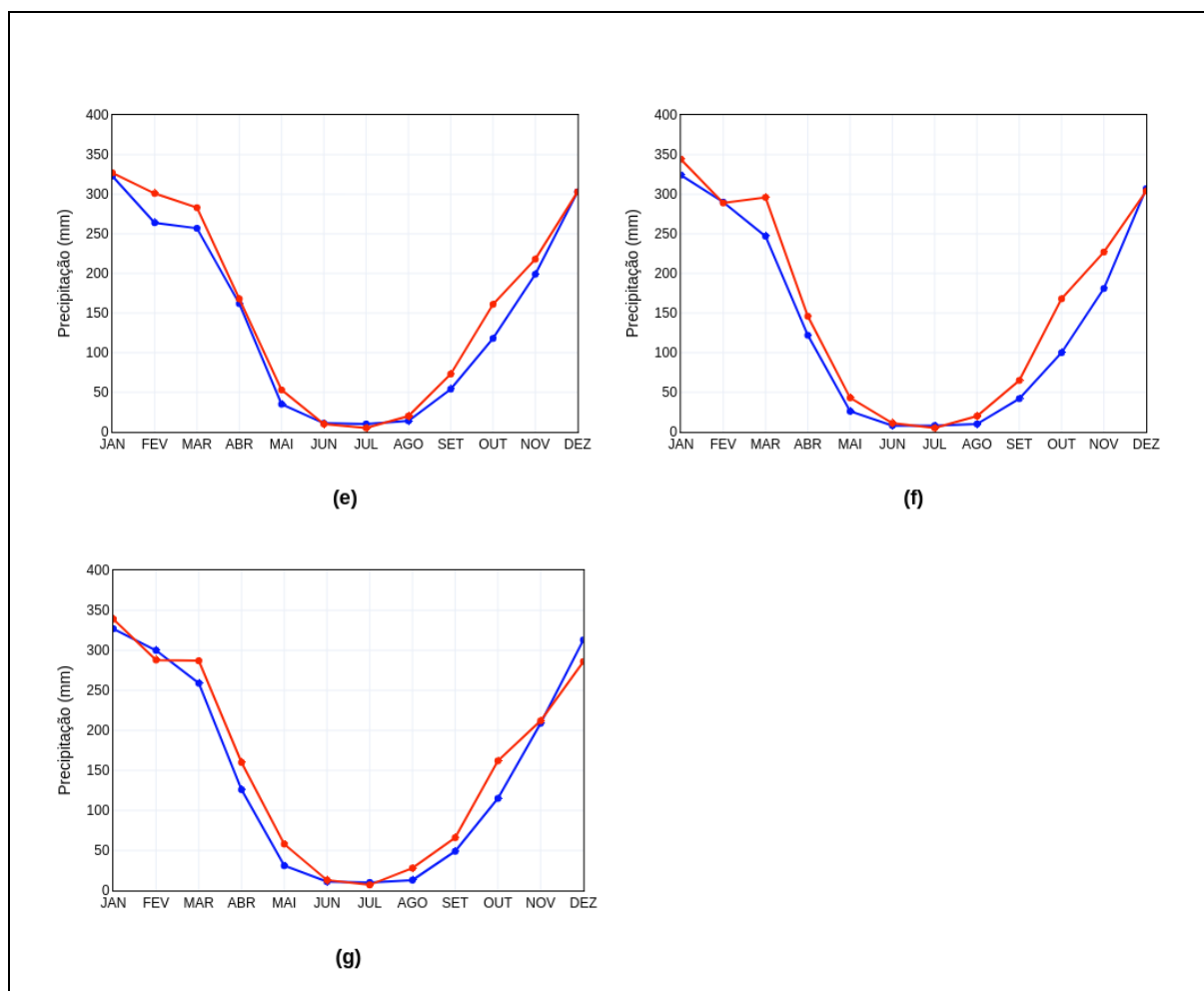
A climatologia da precipitação observada (CPC) e simulada (RegCM4) por sub-bacia, considerando o período base 1981 a 2005 é apresentada nas Figuras 57 e 58. A avaliação comparativa entre os dados observacionais e a simulação obteve resultado positivo, ou seja, o comportamento da precipitação foi capturado pelo modelo climático. Entretanto, visualmente denota-se subestimativas mais pronunciadas para os três primeiros meses do ano que corresponde ao período chuvoso. Esta discrepância ocorreu particularmente nas sub-bacias Jat, 444 e 445, que estão situadas mais próximas do AHE Jatobá. Para comparar as médias entre os dados observacionais e os simulados, foi empregado o teste não-paramétrico Mann-Whitney em razão da condição de não normalidade da série de precipitação.

Figura 57 - Comparativo entre os dados observados do CPC e a simulação usando o modelo climático RegCM4, para a climatologia da precipitação por sub-bacia considerando o período base 1981-2005. (a) Sub-bacia Jat, (b) sub-bacia 444, (c) sub-bacia 445 e (d) sub-bacia 446.



Fonte: Elaborado pelo autor do trabalho.

Figura 58 - Comparativo entre os dados observados do CPC e a simulação usando o modelo climático RegCM4, para a climatologia da precipitação por sub-bacia considerando o período base 1981-2005. (e) Sub-bacia 447, (f) sub-bacia 448 e (g) sub-bacia 449.



Fonte: Elaborado pelo autor do trabalho.

Aplicando-se o teste de Mann-Whitney, não foi identificada diferença estatisticamente significativa para um nível de confiança de 95% entre as duas bases de dados de precipitação (RegCM e CPC), conforme apresentado na Tabela 34.

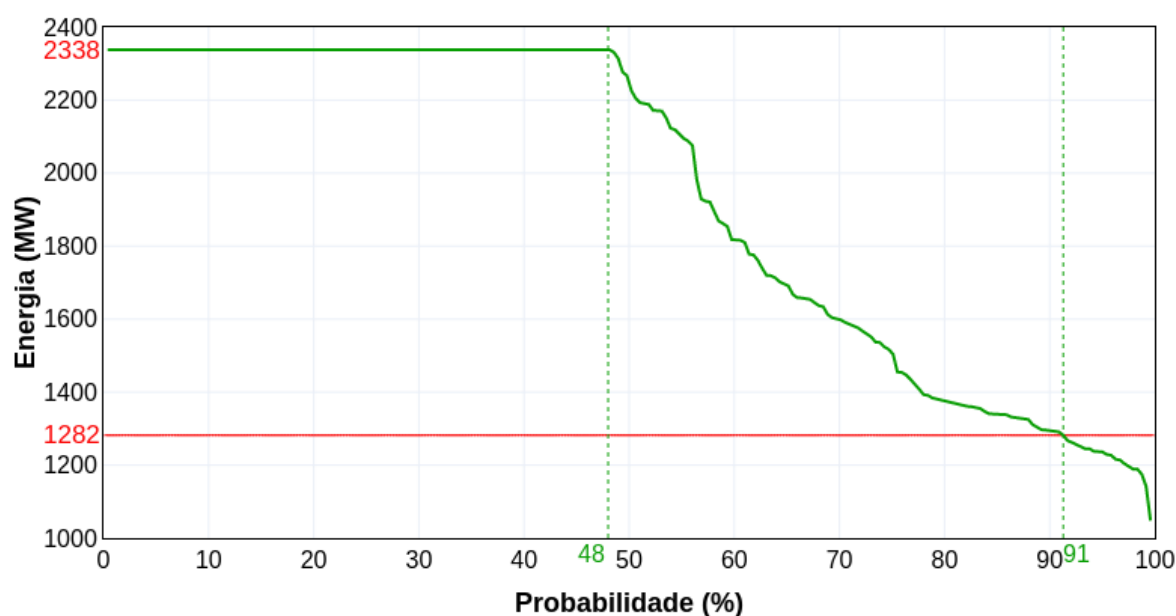
Tabela 35 - Teste Mann-Whitney para comparação de duas amostras sobre os dados de precipitação, considerando como hipótese nula H_0 que a diferença de posição entre as amostras é igual a zero, para $\alpha = 0,05$.

Parâmetros	Jat	444	445	446	447	448	449
Valor-p	0,403	0,403	0,436	0,665	0,840	0,707	0,795
H_0	Aceita	Aceita	Aceita	Aceita	Aceita	Aceita	Aceita
H_a	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita	Rejeita

Fonte: Elaborado pelo autor do trabalho.

A curva de permanência para o potencial de geração de energia no AHE Jatobá utilizando os dados simulados para os próximos vinte anos (2019-2038) é mostrada na Figura 59. De acordo com o esperado para o cenário 1, a probabilidade de atender a capacidade assegurada de geração de energia cai para 91%, quando deveria ser minimamente de 95%. Sob tais condições, a probabilidade de atender a capacidade máxima de geração seria reduzida para 48%.

Figura 59 - Curva de permanência para o potencial de geração mensal de energia hidrelétrica sob cenário 1 (RCP4.5) para o período base 2019-2038. A linha vermelha destaca o valor de geração no período crítico (1.282 MW), conforme projeto.

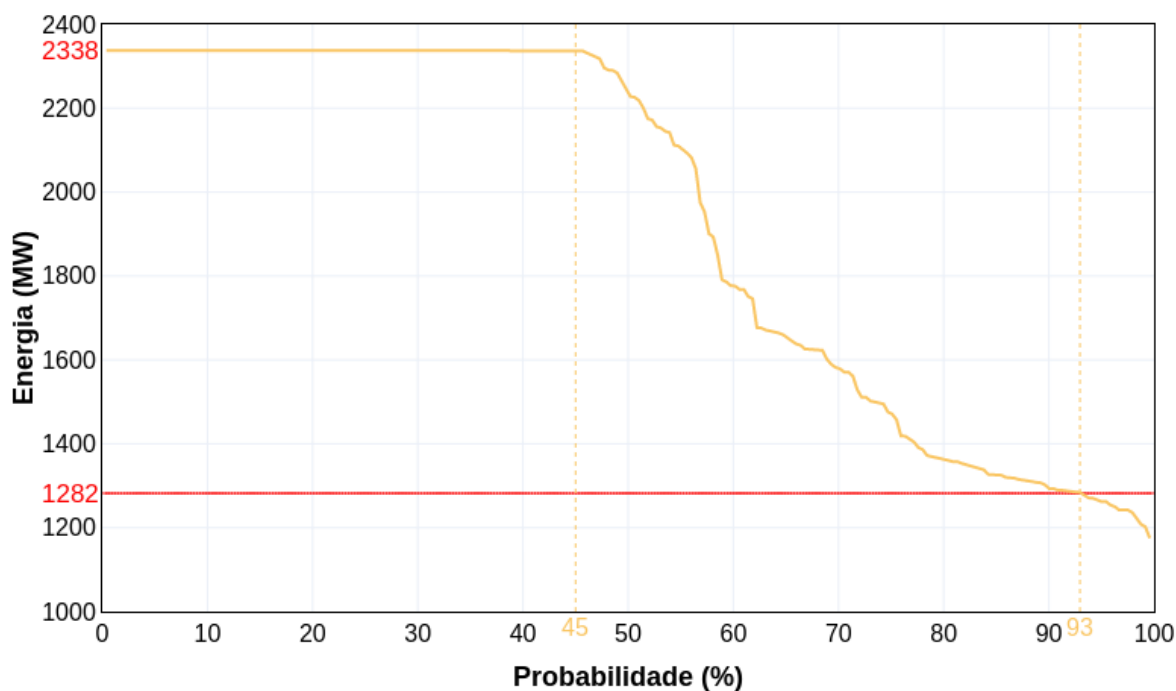


Fonte: Elaborado pelo autor do trabalho.

4.4.2. Cenário 2: Tendência de Precipitação

Sob as condições apresentadas para o cenário 2 (Figura 60), a probabilidade de manter o mínimo da capacidade de geração de energia seria ligeiramente reduzida, alcançando o índice de 93%. Enquanto a probabilidade de disponibilidade de geração na capacidade máxima cairia drasticamente para 45% quando comparada aos dados observados (Figura 35).

Figura 60 - Curva de permanência para o potencial de geração mensal de energia hidrelétrica sob cenário 2 (tendência de precipitação) para o período base 2019-2038. A linha vermelha destaca o valor de geração no período crítico (1.282 MW), conforme projeto.



Fonte: Elaborado pelo autor do trabalho.

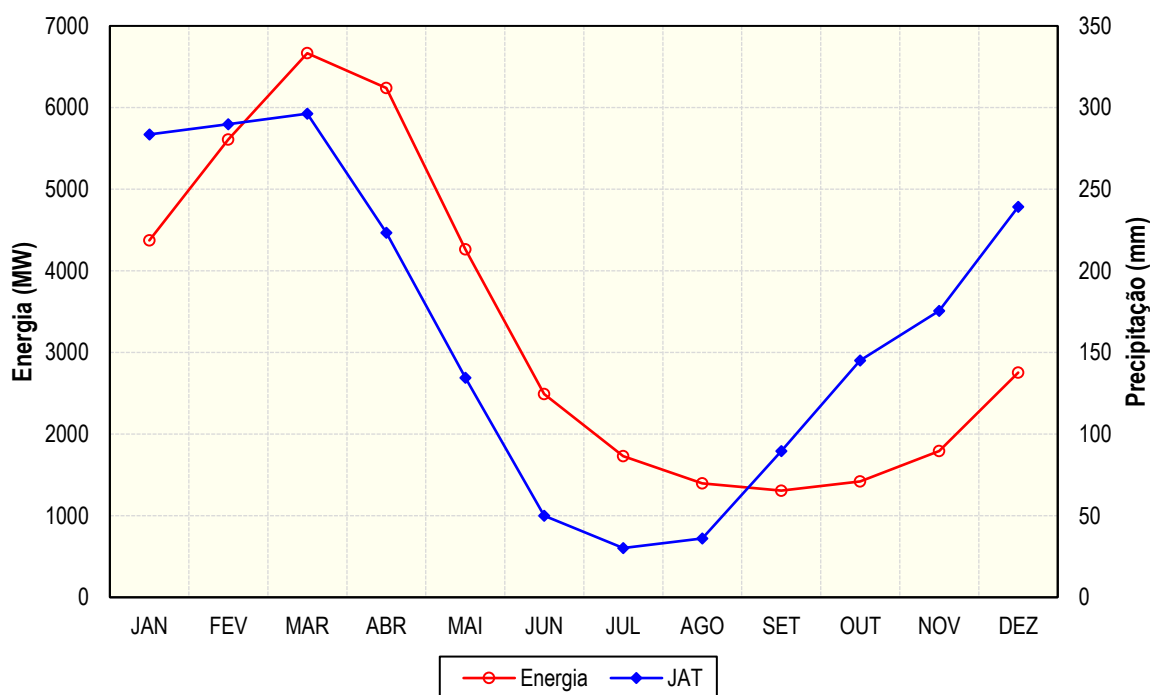
CAPÍTULO V

5. ANÁLISE DOS RESULTADOS

5.1. RELAÇÃO DA PRECIPITAÇÃO COM O POTENCIAL DE GERAÇÃO DE ENERGIA

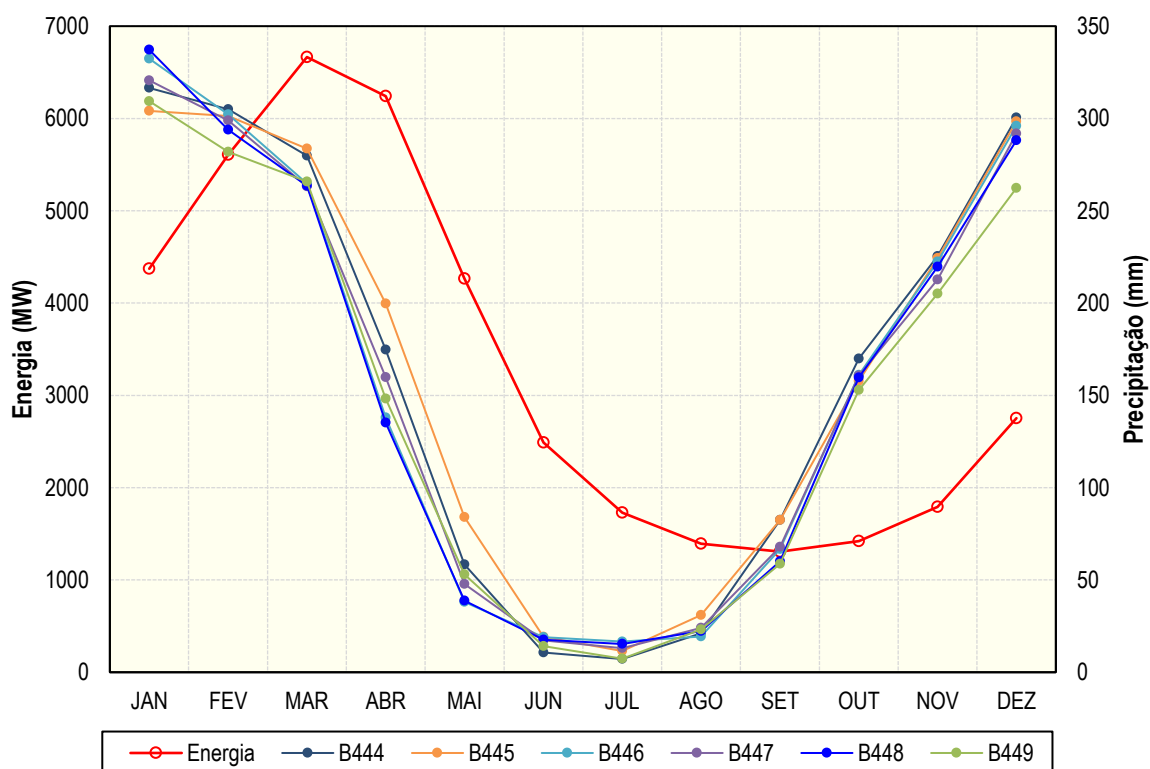
O regime pluviométrico na bacia do rio Tapajós é marcado por grande variabilidade mensal, sazonal e anual, o que, por conseguinte, torna a vazão igualmente variável ao longo do tempo. A consequente variabilidade natural da vazão também influencia a capacidade de geração de energia hidrelétrica no AHE Jatobá. O auge da ocorrência dos períodos mais chuvosos e mais secos é defasado dos máximos e mínimos do potencial para geração de energia denotando que a onda de vazão leva cerca de dois a três meses para alcançar o AHE Jatobá, dependendo da sub-bacia considerada (Figuras 61 e 62). Por estar localizada na sub-bacia Jat, o AHE Jatobá apresenta resposta direta ao potencial de geração de energia em decorrência da precipitação nesta sub-bacia.

Figura 61 - Média mensal do potencial de geração de energia (vermelho) e da precipitação média na sub-bacia Jatobá.



Fonte: Elaborado pelo autor do trabalho.

Figura 62 - Média mensal do potencial de geração de energia (vermelho) e da precipitação média por sub-bacia.

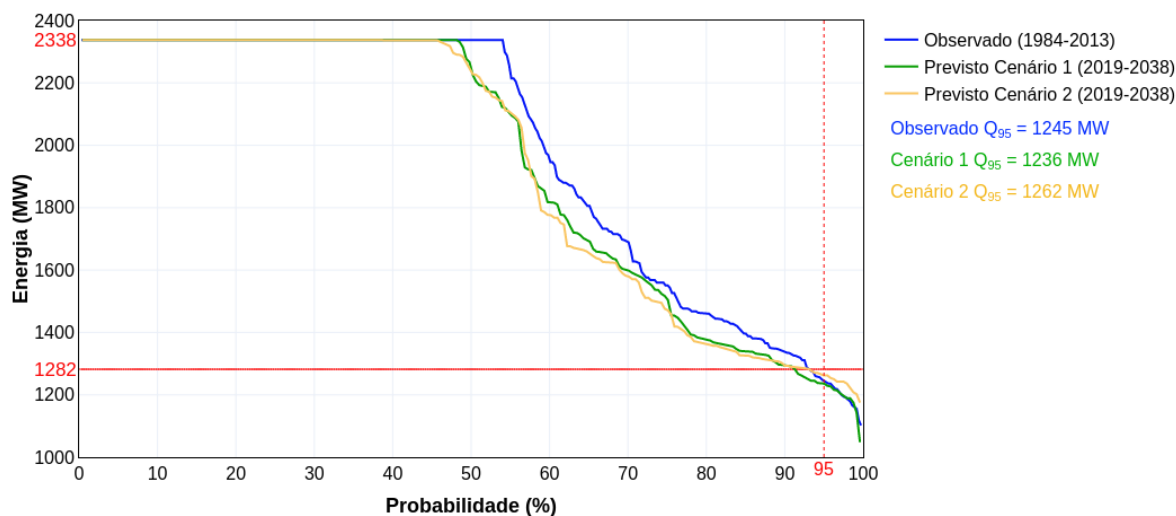


Fonte: Elaborado pelo autor do trabalho.

Os dados observacionais apresentaram significância estatística para uma tendência sistemática de redução da precipitação em alguns meses do ano, com ênfase no período chuvoso. Tal comportamento foi detectado para diferentes sub-bacias do rio Tapajós, o que certamente pode levar a comprometer o fator de capacidade de geração de energia hidrelétrica no AHE Jatobá no médio e longo prazos. Naturalmente que há parâmetros de projeto, como a eficiência do conjunto turbina/gerador, que podem ser ajustados para compensar a perda natural da capacidade de geração por redução da precipitação na bacia hidrográfica. Ainda assim, há indícios de que a probabilidade de manter a energia firme constante no projeto não seria alcançada (Figura 63).

As simulações realizadas para cenários futuros utilizando as saídas do modelo climático também corroboram para o comprometimento da viabilidade do empreendimento nos próximos 20 anos (Figura 63). Entretanto, deve-se levar em consideração também que o modelo climático apresentou viés ligeiramente negativo, ou seja, tendência de subestimar o volume de precipitação, especialmente para os meses mais chuvosos.

Figura 63 - Curvas de permanência para o potencial de geração de energia para os dados observados e para os cenários previstos.



Fonte: Elaborado pelo autor do trabalho.

Também deve ser considerado que pode estar ocorrendo uma variabilidade climática natural em uma escala temporal que supera o período de registros observacionais de precipitação da série trabalhada. E sob tais condições, o viés de redução da precipitação ora detectado em determinados meses para algumas sub-bacias poderia mudar de sentido em outra fase do evento, ou seja, passar para um viés de alta. Mas embora essa seja uma hipótese possível, não parece provável a partir dos dados analisados.

5.2. MODELAGEM PREDITIVA DO POTENCIAL DE GERAÇÃO DE ENERGIA

De uma forma mais geral, os três modelos desenvolvidos capturaram bem o comportamento sazonal das variáveis, mostrando bom desempenho durante a fase de treinamento. Embora os valores extremos da série tenham apresentado grande variância, os erros encontrados foram reduzidos, caracterizando a habilidade dos algoritmos durante o processo de aprendizado.

O modelo RNA-LM destacou-se por apresentar um ajuste quase perfeito para o subconjunto de dados de treinamento (primeiras 80 observações) e grande variância na fase de validação (40 observações seguintes), caracterizando a dificuldade para generalização do modelo e uma tendência de *overfitting*, com superestimação dos dados. Este comportamento pode ser explicado pelo pequeno número de observações para o subconjunto de treinamento e grande número de variáveis na camada de entrada. Para tais condições, quanto menor o número de neurônios na camada escondida, qualquer que seja a arquitetura adotada pela RNA-LM, o

número de parâmetros a ser estimado será sempre muito maior do que o número de observações. Muitos parâmetros contra poucas observações para o treinamento reduzem a habilidade de generalização da rede por não capturar as complexidades nos padrões entre as variáveis de entrada e alvo, frequentemente conduzindo ao *overffiting* (OKUT, 2016). Isto explica o desempenho inferior do modelo RNA-LM comparado aos demais.

O modelo RNA-RB apresentou regularidade em seu desempenho ao longo de toda a fase de aprendizado, denotando melhor capacidade de generalização. Este comportamento difere do modelo anterior porque a técnica de regularização conduz os vieses dos parâmetros para uma direção na qual os valores são mais prováveis. Portanto, esta técnica tende a reduzir a variância das estimativas do custo de introduzir vieses (OKUT, 2016).

Quanto ao modelo GMDH, este também apresentou comportamento regular durante a fase de aprendizado, sendo favorecido neste caso pela escolha das variáveis de maior relevância para o modelo, reduzindo número de parâmetros para otimizar o treinamento e evitar o *overffiting*.

Para a fase de teste e considerando o foco nos valores mínimos de potencial de geração de energia, o desempenho dos modelos alcançou uma boa destreza, especialmente para o GMDH e RNA-RB. Claramente houve subestimativas dos picos ocorridos para os anos de 2005 até 2009, e de 2012 a 2013. Provavelmente este comportamento está associado com o conjunto de dados de precipitação do CPC. Em geral, para regiões com pobre cobertura de redes de pluviômetros, tal qual ocorre na Amazônia, os dados em ponto de grade do CPC tendem a subestimar os eventos de chuva volumosa devido à precipitação convectiva (BECK et al., 2017). Portanto, é muito provável que um ruído nos dados de entrada tenha levado a erros maiores durante a simulação dos picos de máxima geração de energia. Entretanto, essa possível subestimativa dos picos, mesmo não sendo desejável, não compromete os estudos de viabilidade do AHE, pois estes são baseados na energia firme, ou seja, a máxima produção contínua de energia calculada a partir do período crítico.

CAPÍTULO VI

6. CONSIDERAÇÕES FINAIS

Na região Amazônica a existência de séries de dados observacionais de vazão e/ou níveis fluviométricos com duração inferior a trinta anos é majoritária. Esta característica marcante dificulta a realização dos estudos de viabilidade para aproveitamentos hidrelétricos na região, que possui o maior potencial ainda não explorado para geração de energia hidrelétrica no país. A presente tese desenvolveu uma metodologia inovadora para tratar esta questão empregando técnicas de inteligência computacional. Sob esta abordagem três modelos preditivos do potencial de geração de energia hidrelétrica foram desenvolvidos: GMDH, RNA-RB e RNA-LM. O grande diferencial para esta modelagem é a utilização de dados de precipitação mensal média nas sub-bacias hidrográficas como variável de entrada dos modelos. Os dados de precipitação são mais abundantes, disponíveis em diversas fontes e cobrem toda a Amazônia.

Os modelos desenvolvidos foram testados para o AHE Jatobá, situado na bacia do rio Tapajós. A modelagem do potencial de geração de energia mostrou-se exitosa para aplicações em usinas do tipo fio d'água. Mesmo utilizando apenas 120 amostras (dez anos de dados mensais) para o treinamento das redes, o que pode ser considerado como uma base de dados pequena, os modelos apresentaram elevada acurácia.

A técnica de *deep learning* aplicada através do modelo GMDH se destacou sobre as demais redes neurais, alcançando acurácia de 12,34% para o erro percentual médio. Este desempenho se deu graças à sua habilidade em selecionar as variáveis de maior relevância para a construção do modelo, reduzindo de 91 para 16 as variáveis de entrada. Esta característica do algoritmo permite diminuir a quantidade de parâmetros, reduzindo a complexidade do modelo e favorecendo a obtenção de melhor generalização da rede.

A rede neural artificial tradicional, particularmente utilizando o algoritmo de regularização bayesiana, representado pelo modelo RNA-RB, alcançou índices muito próximos do GMDH, com MAPE médio de 13,28%. Como a regularização utiliza uma função de decaimento dos pesos, isso mantém pequena a magnitude dos pesos e melhora a sua capacidade de generalização, mesmo diante de uma pequena amostra para treinamento. Outra característica do RNA-RB que o aproxima do resultado do GMDH é que, para reduzir a complexidade do modelo, o método despreza alguns pesos considerados menos significantes.

Por outro lado, o modelo RNA-LM mesmo tendo apresentado indicadores numericamente aceitáveis, foi o que obteve pior desempenho, não sendo apropriado para sua aplicação no AHE Jatobá por apresentar grande variância nos seus resultados, típico de uma condição de *overfitting*. Ficou claro que o modelo não conseguiu generalizar diante de uma pequena amostra de treinamento e de um grande número de parâmetros, o que o conduziu a um ajuste muito preciso no treinamento e erros dilatados para o conjunto de dados desconhecidos utilizado na fase de teste.

De uma forma geral, os indicadores de avaliação denotaram que os modelos desenvolvidos, GMDH e RNA-BR, são confiáveis, com vantagem para o GMDH, podendo ser utilizados para a estimativa e demais simulações sob cenários particulares visando a determinação do potencial de geração de energia em aproveitamentos hidrelétricos. Nesse sentido, as simulações realizadas com o GMDH para cenários climáticos futuros, considerando os próximos vinte anos, indicaram que tanto para condições de emissão sob a forçante radiativa RCP4.5 proposta pelo IPCC, como para condição de manutenção das tendências climáticas atuais de precipitação observadas na bacia do rio Tapajós, em ambos os cenários há grande possibilidade do AHE não conseguir manter sua capacidade de geração conforme parâmetros de projeto.

Mesmo com todas as limitações e incertezas que a modelagem pode apresentar, os dados observacionais dos últimos quarenta anos (1979-2018) deixam claro que as sub-bacias hidrográficas que contribuem para a vazão e conseqüente geração de energia no aproveitamento Jatobá já estão sofrendo redução significativa da precipitação em determinados meses do ano, especialmente nos períodos de maior volume de precipitação, cujo reflexo produz efeitos no total anual de precipitação das principais sub-bacias do rio Tapajós.

A análise dos dados observacionais e simulados sugerem uma revisão cautelosa e apropriada do projeto para a UHE Jatobá visto que o impacto social e ambiental para implantação de uma usina hidrelétrica, sobretudo na região Amazônica, é significativo e há risco considerável do projeto não conseguir alcançar a capacidade de geração projetada. Os resultados encontrados qualificam a metodologia proposta como um método alternativo que pode auxiliar técnicos e tomadores de decisão a fazerem escolhas assertivas, com redução das incertezas nos estudos prévios envolvendo empreendimento hidrelétricos.

6.1. PROPOSIÇÃO DE TRABALHOS FUTUROS

Esta tese alcançou os objetivos propostos ao desenvolver uma metodologia inédita para a Região Amazônica e constitui um ponto de partida para encorajar outros pesquisadores a contribuir com novos *insights* que poderão conduzir à generalização do método e/ou à aplicação de diferentes técnicas de aprendizado de máquina para aperfeiçoar a acurácia dos modelos e aumentar a segurança das previsões em projetos de usinas hidrelétricas a fio d'água.

Certamente ainda há muito para avançar na utilização desta metodologia. Como descrito na própria tese, há inúmeras técnicas de inteligência computacional com potencial para aplicações em previsão de séries temporais. Como pesquisas futuras, pretende-se explorar outras técnicas de inteligência computacional como redes convolucionais, máquina vetor de suporte, sistemas de inferência fuzzy, neurofuzzy, métodos híbridos que combinam diferentes abordagens. Realizar comparativos entre métodos estatísticos e algoritmos de inteligência computacional, a exemplo do ARIMA ou SARIMA, também é objeto de interesse futuro para investigação.

Como estabelecido nesta pesquisa, os dados observacionais de vazão e/ou nível fluviométrico são relativamente escassos e muitas vezes de qualidade duvidosa devido o elevado número de falhas das séries e as dificuldades de operação inerentes da região Amazônica. Por isso esta metodologia se mostra relevante, visto que os dados de precipitação podem advir de muitas fontes com cobertura na região Amazônica e possuem, usualmente, séries temporais mais longas. Por exemplo, pode-se utilizar dados de precipitação estimados por satélites, dados de plataformas de coleta em superfície, além de outras bases que combinam dados de superfície com estimativas por satélite. Desta forma, utilizar diferentes fontes de dados de precipitação com propósito de avaliar a robustez e habilidade dos modelos também é uma linha de pesquisa que se pretende explorar futuramente.

E por fim, há inúmeros outros cenários que podem ajudar a investigar os impactos na geração de energia que também podem ser mais amplamente explorados, por exemplo, utilizando outras condições de emissões de gases de efeito estufa, conforme proposto pelo IPCC. Ou ainda, avaliar o impacto para a geração de energia exercido por eventos extremos de clima, como os fenômenos El Niño-Oscilação Sul (ENOS), gradiente inter-hemisférico de anomalia de temperatura da superfície do mar no oceano Atlântico tropical, oscilação decadal do Pacífico (ODP).

REFERÊNCIAS

- ABDEL-AAL, R. E. Modeling and forecasting electric daily peak using abductive networks. **International Journal Electrical Power & Energy Systems**, v. 28, p. 133-141, 2006.
- ABDEL-AAL, R. E. Improving electric load forecasts using network committees. **Electric Power Systems Research**, v. 74, p. 83-94, 2005.
- ABBUD, O.; TANCREDI, M. **Transformações recentes da matriz brasileira de geração de energia elétrica – causas e impactos principais**. Centro de Estudos da Consultoria do Senado, 64 p., 2010.
- AHMAD, A. S.; HASSAN, M. Y.; ABDULLAH, M. P.; RAHMAN, H. A.; HUSSIN, F.; ABDULLAH, H.; SAIDUR, R. A review on applications of ANN and SVM for building electrical energy consumption forecasting. **Renewable Sustainable Energy Review**, v. 33, p. 102-109, 2014.
- ALLEN, M. R.; INGRAM, W. J. Constraints on future changes in climate and the hydrologic cycle. **Nature**, v. 419, 224-232, 2002.
- ALEXANDRE, G. R.; BAPTISTA, M. B.; NAGHETTINI, M. Estudo para identificação de tendências do regime pluvial na região metropolitana de Belo Horizonte a partir de métodos estatísticos. **Revista Brasileira de Recursos Hídricos**, v. 15 (2), p. 115-126, 2010.
- ANEEL (Agência Nacional de Energia Elétrica). **Atlas de energia elétrica do Brasil. Parte II – Fontes renováveis – Cap. 3 Energia Hidráulica**. ANEEL, Brasília-DF, 13 p., 2002.
- ANEEL (Agência Nacional de Energia Elétrica). **Cadernos Temáticos ANEEL 3 – Energia Assegurada**. ANEEL, Brasília-DF, 18 p., 2005.
- ANEEL (Agência Nacional de Energia Elétrica). **Atlas de energia elétrica do Brasil. 3ª Edição**. ANEEL, Brasília-DF, 236 p., 2008.
- ANA (AGÊNCIA NACIONAL DE ÁGUAS). **Manual de Construção da Base Hidrográfica Ottocodificada da ANA**. Brasília-DF, 150 p., 2007.
- ARORA, V.; BOER, G. Effects of simulated climate change on the hydrology of major river basins. **Journal of Geophysical Research**, v. 106, p. 3335-3348, 2001.

AWOJOBI, O.; JENKINS, G. P. Managing the cost overrun risks of hydroelectric dams: an application of reference class forecasting techniques. **Renewable Sustainable Energy Reviews**, v. 63, p. 19-32, 2016. <https://doi.org/10.1016/j.rser.2016.05.006>

BALES, R. C. **Hydrology, floods and droughts – overview**. Encyclopedia of Atmospheric Sciences. Second Edition. p.180-184, 2015. <https://doi.org/10.1016/B978-0-12-382225-3.00166-3>

BASHEER, I. A. Selection of methodology for neural network modeling of constitutive hysteresis behavior of soils. **Computer-Aided Civil and Infrastructure Engineering**, v. 15, p. 440-458, 2000. <https://doi.org/10.1111/0885-9507.00206>

BÁRDOSSY, A.; PEGRAM, G. Downscaling precipitation using regional climate models and circulation patterns toward hydrology. **Water Resource Research**, v. 47(4), W04505, 2011.

BATES, B. C.; KUNDZEWICZ, Z. W.; WU, S.; PALUTIKOF, J. P. Climate Change and Water. **Technical Paper of the Intergovernmental Panel on Climate Change**. IPCC Secretariat, Geneva, 210 p., 2008.

BECK, H. E.; VERGOPOLAN, N.; PAN, M.; LEVIZZANI, V.; VAN DIJK, A. I. J. M.; WEEDON, G. P.; BROCCA, L.; PAPPENBERGER, F.; HUFFMAN, G. J.; WOOD, E. F. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. **Hydrology and Earth System Sciences**, v. 21, p. 6201-6217, 2017. <https://doi.org/10.5194/hess-21-6201-2017>

BRAVO, J.; ALLASIA, D.; PAZ, A., COLLISCHONN, W.; TUCCI, C. Coupled hydrologic–hydraulic modeling of the Upper Paraguay River Basin. **Journal of Hydrologic Engineering**, v. 17(5), p. 635-646, 2012.

BRAVO, J. M.; COLLISCHONN, W.; DA PAZ, A.R.; ALLASIA, D.; DOMECCQ, F. Impact of projected climate change on hydrologic regime of the Upper Paraguay River basin. **Climatic Change**, v. 127(1), p. 27-41, 2013.

BROWNLEE, J. **Statistical methods for machine learning**. 291 p., 2017.

BUTERA, I.; BALESTRA, R. Estimation of the hydropower potential of irrigation networks. **Renewable Sustainable Energy Reviews**, v. 48, p. 140-151, 2015. <https://doi.org/10.1016/j.rser.2015.03.046>

CASTELLARIN, A.; BOTTER, G.; HUGHES, D. A.; LIU, S.; OUARDA, T. B. M. J.; PARAJKA, J.; POST, D. A.; SIVAPALAN, M.; SPENCE, C.; VIGLIONE, A.; VOGEL, R. M. Prediction of flow duration curves in ungauged basins. Cambridge University Press, p. 135-162, 2013. <https://doi.org/10.1017/CBO9781139235761.010>

CCEE – Câmara de Comercialização de Energia Elétrica. **Regras de Comercialização – Módulo 4: Garantias Físicas**. 2010.

CHARTTERJEE, S.; HADI, A. S. **Regression analysis by example**. 5th Ed. John Wiley & Sons Inc, New Jersey, 393 p., 2012.

CHEN, M.; SHI, W.; XIE, P.; SILVA, V. B. S.; KOUSKY, V. E.; HIGGINS, R. W.; JANOWIAK, J. E. Assessing objective techniques for gauge-based analyses of global daily precipitation. **Journal of Geophysical Research – Atmosphere**, v. 112, D4, 2008. <https://doi.org/10.1029/2007JD009132>

CHENG, C. –H.; JUN-HE, Y. A novel rainfall forecast model based on the integrated non-linear attribute selection method and support vector regression. **Journal of Intelligent & Fuzzy Systems**, v. 31, n. 2, p. 915-925, 2016. DOI: 10.3233/JIFS-169021

CHOW, V. T.; MAIDMENT, D. R.; MAYS, L. W. **Hidrología aplicada**. McGraw-Hill Interamericana, Bogotá, 584 p., 1994.

CHUNG, J.; GULCEHRE, C.; CHO, K.; BENGIO, Y. Empirical evaluation of gated recurrent neural networks on sequence modelling. 2014. arXiv:1412.3555.

CHURCHLAND, P. S.; SEJNOWSKI, T.J. **The Computational Brain**. Cambridge, MA. MIT Press, 558 p., 1992.

COELHO, I.; COELHO, V.; LUZ, E.; OCHI, L.; GUIMARÃES, F.; RIOS, E. A GPU deep learning metaheuristic based model for time series forecasting. **Applied Energy**, v. 201, p. 412-418, 2017. <https://doi.org/10.1016/j.apenergy.2017.01.003>

- COLLINS, W.; BELLOUIN, N.; DOUTRIAUX-BOUCHER, M.; GEDNEY, N.; HALLORAN, P.; HINTON, T.; HUGHES, J.; JONES, C.; JOSHI, M.; LIDDICOAT, S.; MARTIN, G.; O'CONNOR, F.; RAE, J.; SENIOR, C.; SITCH, S.; TOTTERDELL, I.; WILTSHIRE, A.; WOODWARD, S. Development and evaluation of an Earth-System model – HadGEM2. **Geoscientific Model Development**, v. 4, p. 1051-1075, 2011. doi:10.5194/gmd-4-1051-2011
- COLLISCHONN, W.; ALLASIA, D. G.; SILVA, B. C., TUCCI, C. E. M. The MGB-IPH model for large-scale rainfall-runoff modeling. **Hydrological Sciences Journal**, v. 52, p. 878-895, 2007.
- COLLISCHONN, W.; TASSI, R. **Introduzindo Hidrologia**. IPH UFRGS. 8º Ed, 151 p., 2011.
- COX, P. M.; BETTS, R. A.; COLLINS, M.; et al. Amazonian forest dieback under climate-carbon cycle projections for the 21st century. **Theoretical and Applied Climatology**, v. 78(1), p. 137-156, 2004.
- COX, P. M.; HARRIS, P. P.; HUNTINGFORD, C.; et al. Increasing risk of Amazonian drought due to decreasing aerosol pollution. **Nature**, v. 453, p. 212-215, 2008.
- DAG, O.; YOZGATLIGIL, C. GMDH: an R package for short term forecasting via GMDH-type neural network algorithms. **The R Journal**, v. 8, p. 379-386, 2016. <https://journal.r-project.org/archive/2016/RJ-2016-028/RJ-2016-028.pdf>
- DE SOUZA, E. CARMO, A.; MORAES, B.; NACIG, A.; FERREIRA, D.; ROCHA, E.; SOUZA, P. Sazonalidade da precipitação sobre a Amazônia Legal brasileira: clima atual e projeções futuras usando o modelo RegCM4. **Revista Brasileira de Climatologia**, v. 18, p. 293-306, 2016.
- DEEMER, B. R.; HARRISON, J. A.; LI, S. et al. Greenhouse gas emissions from reservoir water surfaces: a new global synthesis. **BioScience**, v. 66(11), p. 949-964, 2016. <https://doi.org/10.1093/biosci/biw117>
- DEMARIA, E. M. C.; MAURER, E. P.; THRASHER, B.; VICUÑA, S.; MEZA, F. J. Climate change impacts on an alpine watershed in Chile: do new model projections change the story? **Journal of Hydrology**, v. 502, p.128-138, 2013.

DEMUTH, H.; BEALE, M. **Neural network toolbox user's guide for use with Matlab**. Mathworks, 951 p., 2009.

DEMYANOV, S. **Regularization methods for neural networks and related models**. Tese de doutorado. University of Melbourne, 182 p., 2015.

DESSAI, S.; HULME, M. Does climate adaptation policy need probabilities? **Climate Policy**, v. 4, p. 107-128, 2004.

DICKINSON, R. E.; ERRICO, R.M; GIORGI, F.; BATES, G. A regional climate model for the western United States. **Climate Change**, v. 15, p. 383-422, 1989.

DIVYA, S.; JOSEPH, J. A brief review on recent trends in alternative sources of energy. **International Journal of Electrical and Computer Engineering**, v. 8 (2), p. 389-395, 2014.

DOAN, C. D.; LIONG, S. Generalization for multilayer neural network: Bayesian regularization or early stopping. In: **2nd Asia Pacific Association of Hydrology and Water Resources**. Suntec-Singapore, p. 5-8, 2004.

ELATTAR, E. E.; GOULERMAS, J. Y.; WU, Q. H. Generalized locally weighted GMDH for short term load forecasting. **IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews**, v. 42, p. 345-356, 2012.

ELETRORBRAS (Centrais Elétricas Brasileiras). SIPOT (Sistema de Informações do Potencial Hidrelétrico Brasileiro), 2017. [Online] Acessado em 15 maio de 2018. Disponível em <https://eletrobras.com/pt/Paginas/Potencial-Hidreletrico-Brasileiro.aspx>

ELETRONORTE, CAMARGO CORREA, CNEC. **Estudos de inventário hidrelétrico das bacias dos rios Tapajós e Jamanxim**. Rio de Janeiro, 117 p., 2008.

EPE (EMPRESA DE PESQUISA ENERGÉTICA). **Metodologia de Cálculo da Garantia Física das Usinas**. EPE, Rio de Janeiro, 15 p., 2008.

EPE (EMPRESA DE PESQUISA ENERGÉTICA). **Balanco Energético Brasileiro 2018 Ano 2017**. EPE, Rio de Janeiro, 294 p., 2018.

FAN, C.; XIAO, F; ZHAO, Y. A short-term building cooling load prediction method using deep learning algorithms. **Applied Energy**, v. 195, p. 222-233, 2017. <https://doi.org/10.1016/j.apenergy.2017.03.064>

FARLOW, S. J. The GMDH algorithm of Ivakhnenko. **The American Statistician**, v. 35, p. 210-215, 1981.

FARLOW, S. J. **Self-organizing methods in modelling: GMDH type algorithms**. Ed. Marcel Dekker, New York, 350 p., 1984.

FEARNSIDE, P. M. Greenhouse gas emissions from hydroelectric dams in tropical forests. **Alternative Energy and Shale Gas Encyclopedia**, Chapter 42, p.426-436, 2016. <https://doi.org/10.1002/9781119066354.ch42>

FEARNSIDE, P. M. Impacts of Brazil's Madeira River dams: unlearned lessons for hydroelectric development in Amazonia. **Environmental Science & Policy**, v. 38, p. 164-172, 2014. <https://doi.org/10.1016/j.envsci.2013.11.004>

FEARNSIDE, P. M. Dams in the Amazon: Belo Monte and Brazil's hydroelectric development of the Xingu River Basin. **Environmental Management**, v. 38, p. 16-27, 2006. <https://doi.org/10.1007/s00267-005-0113-6>

FENG, C.; CUI, M.; HODGE, B.; ZHANG, J. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. **Applied Energy**, v. 190, p.1245-1257, 2017. <https://doi.org/10.1016/j.apenergy.2017.01.043>

FENTON, J. D. On the generation of stream rating curves. **Journal of Hydrology**, v. 564, p. 748-757, 2018. <https://doi.org/10.1016/j.jhydrol.2018.07.025>

FORESEE, F. D.; HAGAN, M. T. Gauss-Newton Approximation to Bayesian learning. In: **Proceedings of the International Conference on Neural Networks**. Houston-Texas, p. 1930-1935, 1997.

GIORGI, F.; BATES, G. T. The climatological skill of a regional climate model over complex terrain. **Monthly Weather Review**, v. 117 (11), p. 2325-2347, 1989.

GIORGI, F.; MARINUCCI, M. R.; BATES, G. T. Development of a Second-Generation Regional Climate Model (RegCM2). Part I: Boundary-layer and radiative transfer processes. **Monthly Weather Review**, v. 121, p. 2794-2813, 1993a.

GIORGI, F.; MARINUCCI, M. R.; BATES, G. T.; DE CANIO, G. Development of a second-generation regional climate model (RegCM2). Part II: Convective processes and assimilation of lateral boundary conditions. **Monthly Weather Review**, v. 121, p. 2814-2832, 1993b.

GIORGI, F.; MEARNS, L. O. Introduction to special section: regional climate modeling revisited. **Journal of Geophysical Research**, v. 104, p. 6335–6352, 1999.

GIORGI, F.; COPPOLA, E.; SOLMON, F.; MARIOTTI, L.; SYLLA, M.; BI, X.; ELGUINDI, N.; DIRO, G.; NAIR, V.; GIULIANI, G.; TURUNCOGLU, U.; COZZINI, S.; GÜTTLER, I.; O'BRIEN, T.; TAWFIK, A.; SHALABY, A.; ZAKEY, A.; STEINER, A.; STORDAL, F.; SLOAN, L.; BRANKOVIC, C. RegCM4: model description and preliminary tests over multiple CORDEX domains. **Climate Research**, v. 52, p. 7-29, 2012.

GOCIC, M.; TRAJKOVIC, S. Analysis of changes in meteorological variables using Mann-Kendall and Sen's slope estimator tests in Serbia. **Global and Planetary Change**, v. 100, p. 172-182, 2013. <http://dx.doi.org/10.1016/j.gloplacha.2012.10.014>

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. MIT Press, 2016.
<http://www.deeplearningbook.org/>

GTE (Grupo de Estudos Tapajós). Avaliação Ambiental integrada da bacia do Tapajós. Síntese do diagnóstico ambiental. Relatório Técnico, 46 p., 2014.

GUIMBERTEAU, M.; DRAPEAU, G.; RONCHAIL, J.; SULTAN, B.; POLCHER, J.; MARTINEZ, J. -M.; PRIGENT, C.; GUYOT, J. -L.; COCHONNEAU, G.; ESPINOZA, J. C.; FILIZOLA, N.; FRAIZY, P.; LAVADO, W.; DE OLIVEIRA, E.; POMBOSA, R.; NOBRIEGA, L.; VAUCHEL, P. Discharge simulation in the sub-basins of the Amazon using ORCHIDEE forced by new datasets. **Hydrology and Earth System Sciences**, v. 16, p. 911-935, 2012.

GUIMBERTEAU, M.; RONCHAIL, J.; ESPINOZA-VILLAR, J. C.; LENGAINNE, M.; SULTAN, B.; POLCHER, J.; DRAPEAU, G.; GUYOT, J. -L.; DUCHARNE, A.; CIAIS, P.

Future changes in precipitation and impacts on extreme streamflow over Amazonian sub-basins. **Environmental Research Letters**, v. 8(1), 014035, 2013.

HAGAN, M. T.; MENHAJ, M. B. Training feedforward networks with the Marquardt algorithm. **IEEE Transactions on Neural Networks**, v. 5 (6), p. 989-993, 1994. <https://doi.org/10.1109/72.329697>

HASSAN, F. The Aswan high dam and the international rescue Nubia campaign. **The African Archaeological Review**, v. 24, p. 73-94, 2007. <https://www.jstor.org/stable/40743449>

HAYKIN, S. **Neural networks and learning machines**. Third Ed. Pearson Prentice Hall. New Jersey, 938 p., 2009.

HERSCHY, R. D. **Streamflow measurement**. Third edition, Taylor & Francis, London, 534 p., 2009.

HINTON, G. E.; OSINDERO, S.; TEH, Y. -W. A fast learning algorithm for deep belief nets. **Journal Neural Computation**, v. 18, p. 1527-1554, 2006. <https://doi.org/10.1162/neco.2006.18.7.1527>

HOCHREITER, S.; SCHMIDHUBER, J. **Long short-term memory**. **Neural Computation**, v. 9 (8), p. 1735-1780, 1997.10.1162/neco.1997.9.8.1735

IKEDA, S.; SAWARAGI, Y.; OCHIAI, M. Sequential GMDH algorithm and its application to river flow prediction. **IEEE Transactions on Systems, Man and Cybernetics**, v. 6, p. 473-479, 1976. <https://doi.org/10.1109/TSMC.1976.4309532>

IEA (International Energy Agency). **World Energy Outlook**. Technical report, 12 p., 2018.

INPE (Instituto Nacional de Pesquisas Espaciais). Projeto de Monitoramento do Desflorestamento da Amazônia (PRODES). Dados PRODES (mapas e imagens) 2018. Acesso em 12 de maio de 2019. Disponível em <http://www.dpi.inpe.br/prodesdigital/dadosn/>.

IPCC (Intergovernmental Panel on Climate Change). **Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change** [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (editors.)].

Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 p., 2013.

IRENA (International Renewable Energy Agency). **Hydropower Technology Brief**. IEA-ETSAP & IRENA, Abu Dhabi, 20 p., 2015.

IVAKHNENKO, A. G. The group method of data handling – a rival of the method of stochastic approximation. *Soviet Automatic Control*, v. 1 (3), p. 43-55, 1968.

IVAKHNENKO, A. G. Polynomial theory of complex systems. **IEEE Transactions on Systems, Man and Cybernetics**, v. SMC-1 (4), p. 364–378, 1971. <https://doi.org/10.1109/TSMC.1971.4308320>

IVAKHNENKO, A. G.; IVAKHNENKO, G. A. The review of problems solvable by algorithms of the group method of data handling (GMDH). **Pattern Recognition and Image Analysis**, v. 5 (4), p. 527–535, 1995.

JANI, D. B.; MISHRA, M.; SAHOO, P. K. Application of artificial neural network for predicting performance of solid desiccant cooling systems – A review. **Renewable Sustainable Energy Reviews**, v. 80, p. 352-366, 2017. <https://doi.org/10.1016/j.rser.2017.05.169>

JOETZJER, E.; DOUVILLE, H.; DELIRE, C.; CIAIS, P. Present-day and future Amazonian precipitation in global climate models: CMIP5 versus CMIP3. **Climate Dynamics**, v. 41(11), p. 2921-2936, 2013.

JONES, R. N. Managing uncertainty in climate change projections: issues for impact assessment. **Climate Change**, v. 45, p. 403-419, 2000.

KAYRI, M. Predictive abilities of Bayesian regularization and Levenberg-Marquardt algorithms in artificial neural networks: a comparative empirical study on social data. **Mathematical and Computational Applications**, v. 21(2), p. 20, 2016.

KASABOV, N. K. **Foundations of neural networks, fuzzy systems and knowledge engineering**. A Bradford Book, The MIT Press, London, 581 p., 1995.

KENDALL, M. **Multivariate Analysis**. Charles Griffin & Company, London, 210 p., 1975.

KOCH, F.; PRASCH, M.; BACH, H.; MAUSER, W.; APPEL, F.; WEBER, M. How will hydroelectric power generation develop under climate change scenarios? A case study in the upper Danube basin. **Energies**, v. 4, p. 1508–1541, 2011. <https://doi.org/10.3390/en4101508>

KOHONEN, T.; DEBOECK, G. **Visual explorations in finance with self-organizing maps**. Springer, London, 258 p., 1998.

KONG, X; XU, X.; YAN, Z.; CHEN, S.; YANG, H; HAN, D. Deep learning hybrid method for islanding detection in distributed generation. **Applied Energy**, v. 210, p. 776-785, 2017. <https://doi.org/10.1016/j.apenergy.2017.08.014>

KORDÍK, P. GAME - Hybrid self-organizing modeling system based on GMDH. In: **Hybrid Self-Organizing Modeling Systems – Studies in Computational Intelligence**, v. 211, p. 233–280, 2009. https://doi.org/10.1007/978-3-642-01530-4_6

KRIZHEVSKY, A; SUTSKEVER, I.; HINTON, G. ImageNet classification with deep convolutional neural networks. In: **Proceedings of the 25th International Conference on Neural Information Processing Systems, NIPS'12**. USA: Curran Associates Inc., p.1097-1105, 2012. <http://dx.doi.org/10.1145/3065386>.

KUMAR, P.; MERCHANT, S. N.; DESAI, U. B. Improving performance in pulse radar detection using Bayesian regularization for neural network training. **Digital Signal Processing**, v. 14 , p. 438-448, 2004. doi:10.1016/j.dsp.2004.06.002

LAGO, J.; RIDDER, F. D.; SCHUTTER, B. D. Forecasting spot electricity prices: deep learning approaches and empirical comparison of traditional algorithms. **Applied Energy**, v. 221, p. 386-405, 2018. <https://doi.org/10.1016/j.apenergy.2018.02.069>

LANDSBERG, H.; JACOBS, W. C. **Applied climatology**. In Malone, T.F., ed., *Compendium of Meteorology*. Boston, MA: American Meteorological Society, p. 976 – 992, 1951.

LAVADO CASIMIRO, W. S.; LABAT, D.; GUYOT, J. L.; ARDOIN-BARDIN, S. Assessment of climate change impacts on the hydrology of the Peruvian Amazon – Andes basin. **Hydrological Processes**, v. 25(24), p. 3721-3734, 2011.

LEOTTI, V.; BIRCK, A.; RIBOLDI, J. Comparação dos testes de aderência à normalidade Kolmogorov-Smirnov, Anderson-Darling, Cramer–Von Mises e Shapiro-Wilk por simulação.

In: **Simpósio de Estatística Aplicada à Experimentação Agronômica**, Londrina/PR, v. 11, p. 192, 2005.

LI, W. H.; FU, R.; DICKINSON, R. E. Rainfall and its seasonality over the Amazon in the 21st century as assessed by the coupled models for the IPCC AR4. **Journal of Geophysical Research**, v. 111, D02111, 2006.

LI, F.; UPADHYAYA, B. R.; COFFEY, L. A. Model-based monitoring and fault diagnosis of fossil power plant process units using Group Method of Data Handling. **ISA Trans**, v. 48, p. 213–219, 2009. [https://doi.org/S0019-0578\(08\)00075-X](https://doi.org/S0019-0578(08)00075-X) [pii]r10.1016/j.isatra.2008.10.014

LOPES, M.; ROLIM, P.; BOUÇÃO, G.; CASTRO, B.; LOPES, J. Análise de tendência da precipitação em Belém-PA. In: XVII Congresso Brasileiro de Meteorologia, Gramado-RS, 2012.

MACKAY, D. J. C. Bayesian Interpolation. **Neural Computation**, v. 4, p. 415–447, 1992. <https://doi.org/10.1162/neco.1992.4.3.415>

MANN, H. B. Nonparametric tests against trend. **Econometrica**, v. 13, p. 245-259, 1945.

MARENGO, J. A. Cenários de Mudanças Climáticas para o Brasil em 2100. **Ciência & Ambiente**, v. 34, p. 100-125, 2007.

MARENGO, J. A.; BETTS, R.; NOBRE, C.; KAY, G.; CHOU, S.; TOMASELLA, J.; SAMPAIO, G.; ALVES, L.; OBREGON, G.; SOAREAS, W. **Riscos das Mudanças Climáticas no Brasil: Análise Conjunta Brasil-Reino Unido sobre os Impactos das Mudanças Climáticas e do Desmatamento na Amazônia**. 56 p., 2011a.

MARENGO, J. A.; TOMASELLA, J.; ALVES, L. M.; SOARES, W. R.; RODRIGUEZ, D. A. The drought of 2010 in the context of historical droughts in the Amazon region. **Geophysical Research Letters**, v. 38, L12703, 2011b.

MARENGO, J. A.; CHOU, S. C.; KAY, G.; ALVES, L. M.; et al. Development of regional future climate change scenarios in South America using the Eta CPTEC/HadCM3 climate change projections: climatology and regional analyses for the Amazon, São Francisco and the Paraná River basins. **Climate Dynamics**, v. 38(9), p. 1829-1848, 2011c.

MCGUFFIE, K.; HENDERSON-SELLERS, A. **A Climate Modelling Primer**. Third Edition. John Wiley & Sons. London, 296 p., 2005.

MICHAELIDES, E. **Alternative energy sources – Green energy and technology**. Springer, London, 478 p., 2012.

MILLY, P.; DUNNE, K.; VECCHIA, A. Global pattern of trends in streamflow and water availability in a changing climate. **Nature**, v. 438, p. 347-350, 2005.

MMA (Ministério do Meio Ambiente). **Resolução nº 30**. MMA/Conselho nacional de Recursos Hídricos, D.O.U., 2002.

MMA (Ministério do Meio Ambiente). **Caderno da região Amazônica**. MMA/Secretaria de Recursos Hídricos, 126 p., 2006.

MME (Ministério de Minas e Energia). **Manual for hydropower studies of river basins**. MME, Rio de Janeiro, 678 p., 2007.

MOHOR, G. S.; RODRIGUEZ, D. A.; TOMASELLA, J.; SIQUEIRA JÚNIOR, J. L. Exploratory analyses for the assessment of climate change on the energy production in an Amazon run-of-river hydropower plant. **Journal of Hydrology: Regional Studies**, v. 4, p. 41-59, 2015.

MRUGALSKI, M. An unscented Kalman filter in designing dynamic GMDH neural networks for robust fault detection. **International Journal of Applied Mathematics and Computer Science**, v. 23, p. 157–169, 2013. <https://doi.org/10.2478/amcs-2013-0013>

MULLER, A. C.; GUIDO, S. **Introduction to Machine Learning with Python**. O'Reilly Media Inc., Sebastopol, 340 p., 2016.

NOBRE, C. **Mudanças climáticas e o Brasil - Contextualização**. Parcerias Estratégicas (Brasília), v. 27, p. 7-17, 2008.

OKUT, H. Bayesian Regularized Neural Networks for Small n Big p Data, in: **Artificial Neural Networks**. INTECH, p. 27–48, 2016. <https://doi.org/10.5772/63256>

ONWUBOLU, G. **GMDH-Methodology and implementation in C**. Imperial College Press. London, p. 304, 2015.

PAL, J. S.; GIORGI, F.; BI, X.; et al. The ICTP RegCM3 and RegCNET: regional climate modeling for the developing World. **Bulletin of the American Meteorological Society**, p. 1395-1409, 2007.

PBMC (PAINEL BRASILEIRO DE MUDANÇAS CLIMÁTICAS). Base científica das mudanças climáticas – Volume 1 – Primeiro relatório de avaliação nacional. [Ambrizzi, T.; Araújo, M. (editores)]. COPPE/UFRJ, Rio de Janeiro, 464 p., 2014.

PEÑA-ARANCIBIA, J. L.; VAN DIJK, A. I. J. M.; GUERSCHMAN, J. P., et al. Detecting changes in streamflow after partial woodland clearing in two large catchments in the seasonal tropics. **Journal of Hydrology**., v. 416-417, p. 60-71, 2012.

PINTO, N. L. S.; HOLTZ, A. C. T.; MARTINS, J. A. **Hidrologia de Superfície**. 2ª Ed. Edgard Blücher. Curitiba-PR, Brasil. Centro de Estudos e Pesquisas de Hidráulica da Universidade Federal do Paraná. 180 p., 1973.

RAGHUNATH, H. M. **Hydrology – principles, analysis, design**. 2nd Ed. New Age, New Delhi, 477 p., 2006.

ROBESON, S. Statistical climatology. In: **Encyclopedia of World Climatology**, [Oliver, J. Ed.] Springer, 874 p., 2005

RODRIGUEZ, D. A.; TOMASELLA, J.; LINHARES, C. Is the forest conversion to pasture affecting the hydrological response of Amazonian catchments? Signals in the Ji-Paraná Basin. **Hydrological Processes**, v. 24 (10), p.1254-1269, 2010.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v 323, p. 533–536, 1986. <https://doi.org/10.1038/323533a0>

SALATI, E.; SCHINDLER, W.; VICTORIA, D.; SALATI, E.; SOUZA, J.; VILLA NOVA, N. **The economics of climate change in Brazil: costs and opportunities**. Vox Editora e Gráfica Ltda., São Paulo, 2009.

SALAZAR, L. F.; NOBRE, C. A.; OYAMA, M. D. Climate change consequences on the biome distribution in tropical South America. **Geophysical Research Letters**, v. 34, L09708, 2007.

SEN, P. K. Estimates of the regression coefficient based on Kendall's Tau. **Journal of the American Statistical Association**, v. 63 (324), p. 1379–1389, 1968.

SHARMA, H.; SINGH, J. Run off river plant: status and prospects. **International Journal of Innovative Technology and Exploring Engineering**, v. 3 (2), p. 210–213, 2013.

SHIKLOMANOV, I. A. World fresh water resources. In: P.H. Gleick (ed.) **Water in crisis: a guide to the world's fresh water resources**. Oxford University Press, New York, p. 13–24, 1993.

SIQUEIRA JÚNIOR, J. L.; TOMASELLA, J.; RODRIGUEZ, D. A. Impacts of global and regional land-use and land-cover changes in the hydrological regime of the Madeira River basin. **Climatic Change**, v. 129(1), p. 117-129, 2015.

SIROIS, A. A Brief and Biased Overview of Time Series Analysis or How to Find that Evasive Trend. In: **WMO report No. 133 WMO/EMEP workshop on Advanced Statistical methods and their Application to Air Quality Data sets**. Helsinki, p. 14–18, 1998,

SOUSA JÚNIOR, W. **Tapajós: hidrelétricas, infraestrutura e caos. Elementos para a governança em uma região singular**. 1ª Ed. ITA/CTA, São José dos Campos, 192 p., 2014.

SRINIVASAN, D. Energy demand prediction using GMDH networks. **Neurocomputing**, v. 72, p. 625–629, 2008. <https://doi.org/10.1016/j.neucom.2008.08.006>

STICKLER, C. M.; COE, M. T.; COSTA, M. H.; NEPSTAD, D. C.; MCGRATH, D. G.; DIAS, L. C. P.; RODRIGUES, H. O.; SOARES-FILHO, B. S. Dependence of hydropower energy generation on forests in the Amazon Basin at local and regional scales. **Proceedings of the National Academy of Sciences-PNAS**, v. 110 (23), p. 9601–9606, 2013. <https://doi.org/10.1073/pnas.1215331110>

THIESSEN, A. Precipitation averages for large areas. **Monthly Weather Review**, v. 39, p. 1082–1089, 1911.

TOLMASQUIM, M. T. **Energia renovável: hidráulica, biomassa, eólica, solar, oceânica**. EPE, Rio de Janeiro, 452 p., 2016.

TUCKER, S.; JONES, R. G.; BUONOMO, E.; BURGİN, L.; GALLO, F. Dynamical downscaling of GloSea5 over Ethiopia. **Climate Services**, v.9, p. 57-71, 2018.

WALI, U. G. Estimating Hydropower Potential of an Ungauged Stream. **International Journal of Emerging Technology and Advanced Engineering**, v. 3 (11), p. 592–600, 2013.

WANG, D.; LUO, H.; GRUNDER, O.; LIN, Y.; GUO, H. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. **Applied Energy**, v.190, p. 390-407, 2017. 10.1016/j.apenergy.2016.12.134

WANG, H.; WANG, G.; LI, G.; PENG, J.; LIU, Y. Deep belief network based deterministic and probabilistic wind speed forecasting approach. **Applied Energy**, v. 182, p.80-93, 2016. 10.1016/j.apenergy.2016.08.108

WILKS, D. S. **Statistical methods in the atmospheric sciences**. 2nd Ed. Academic Press, New York, 649 p., 2006.

WITCZAK, M.; KORBICZ, J.; MRUGALSKI, M.; PATTON, R. J. A GMDH neural network-based approach to robust fault diagnosis: application to the DAMADICS benchmark problem. **Control Engineering Practice**, v. 14 (6), p. 671–683, 2006. <https://doi.org/10.1016/j.conengprac.2005.04.007>

WOOD, A. W.; MAURER, E. P.; KUMAR, A.; LETTENMAIER, D. P. Long-range experimental hydrologic forecasting for the eastern United States. **Journal of Geophysical Research Atmosphere**, v. 107, D20, 2002.

VALENÇA, M. **Fundamentos das redes neurais: exemplos em Java**. 2^a Edição, Livro Rápido, Olinda, 386 p., 2010.

XIAO, J.; SUN, H.; HU, Y.; XIAO, Y. GMDH based auto-regressive model for China's energy consumption prediction. In: **2015 International Conference on Logistics, Informatics and Service Science - LISS**, 2015. <https://doi.org/10.1109/LISS.2015.7369754>

XIE, K.; ZHANG, H.; SINGH, C. Reliability forecasting models for electrical distribution systems considering component failures and planned outages. **International Journal of Electrical Power & Energy System**, v. 79, p. 228-234, 2016. <https://doi.org/10.1016/j.ijepes.2016.01.020>

XIE, P.; CHEN, M.; YANG, S.; YATAGAI, A.; HAYASAKA, T.; FUKUSHIMA, Y.; LIU, C. A gauge-based analysis of daily precipitation over East Asia. **Journal of Hydrometeorology**, v. 8, p. 607-626, 2007. <https://doi.org/10.1175/JHM583.1>

XU, H.; DONG, Y.; WU, J.; ZHAO, W. Application of GMDH to Short-Term Load Forecasting In: **Advances in Intelligent and Soft Computing**, v. 138, p. 27–32, 2012. https://doi.org/10.1007/978-3-642-27869-3_4

XU, M.; ZENG, G.; XU, X., HUANG, G.; JIANG, R.; SUN, W. Application of bayesian regularized BP neural network model for trend analysis, acidity and chemical composition of precipitation in North Carolina. **Water, Air, and Soil Pollution**, v. 172, p. 167-184, 2006.

ZHANG, W.; YAN, Y.; ZHENG, J.; LI, L.; DONG, X; CAI, H. Temporal and spatial variability of annual extreme water level in the Pearl River Delta region, China. **Global and Planetary Change**, v. 69, p. 35-47, 2009. <https://doi.org/10.1016/j.gloplacha.2009.07.003>

ZJAVKA, L.; SNÁŠEL; V. Short-term power load forecasting with ordinary differential equation substitutions of polynomial networks. **Electric Power System Research**, v. 137, p. 113–123, 2016. <https://doi.org/10.1016/j.epsr.2016.04.003>