

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E
AMBIENTAL**

**PREVISÃO HIDROLÓGICA PARA A BACIA DO RIO
MADEIRA UTILIZANDO PREDITORES CLIMÁTICOS**

GABRIEL DIAS ROGAE

ORIENTADOR: CARLOS HENRIQUE RIBEIRO LIMA

COORIENTADOR: MARCUS SUASSUNA SANTOS

**MONOGRAFIA DE PROJETO FINAL II EM ENGENHARIA
AMBIENTAL**

BRASÍLIA/DF: DEZEMBRO/2020

UNIVERSIDADE DE BRASÍLIA

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E
AMBIENTAL**

**PREVISÃO HIDROLÓGICA PARA A BACIA DO RIO
MADEIRA UTILIZANDO PREDITORES CLIMÁTICOS**

GABRIEL DIAS ROGAE

MONOGRAFIA DE PROJETO FINAL SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL DA UNIVERSIDADE DE BRASÍLIA COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM ENGENHARIA AMBIENTAL.

APROVADA POR:

**CARLOS H. R. LIMA, PhD (Universidade de Brasília)
(ORIENTADOR)**

**DIRCEU S. R. JÚNIOR, PhD (Universidade de Brasília)
(EXAMINADOR INTERNO)**

**SAULO AIRES DE SOUZA, MSc (Universidade de Brasília)
(EXAMINADOR EXTERNO)**

DATA: BRASÍLIA/DF, 09 DE DEZEMBRO DE 2020.

FICHA CATALOGRÁFICA

ROGAE, GABRIEL DIAS

Previsão Hidrológica para a Bacia do Rio Madeira Utilizando Preditores Climáticos [Distrito Federal] 2020.

48 p., 297 mm (ENC/FT/UnB, Bacharel, Engenharia Ambiental, 2020)

Monografia de Projeto Final - Universidade de Brasília. Faculdade de Tecnologia. Departamento de Engenharia Civil e Ambiental.

1.Previsão de cheias 2.Teleconexões climáticas 3.Funções densidade de probabilidade

I. ENC/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

ROGAE, G.D. (2020). Previsão Hidrológica para a Bacia do Rio Madeira Utilizando Preditores Climáticos. Monografia de Projeto Final, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 38 p.

CESSÃO DE DIREITOS

NOME DA AUTOR: Gabriel Dias Rogae

TÍTULO DA MONOGRAFIA DE PROJETO FINAL: Previsão Hidrológica para a Bacia do Rio Madeira Utilizando Preditores Climáticos

GRAU / ANO: Bacharel em Engenharia Ambiental / 2020

É concedida à Universidade de Brasília a permissão para reproduzir cópias desta monografia de Projeto Final e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. A autora reserva outros direitos de publicação e nenhuma parte desta monografia de Projeto Final pode ser reproduzida sem a autorização por escrito da autora.

Gabriel Dias Rogae

gabrielrogae@gmail.com

Brasília/DF – Brasil

AGRADECIMENTOS

Agradeço aos meus pais, Célio e Conceição, que me incentivaram e foram meu suporte durante todo o período acadêmico, e que me ajudaram a permanecer focado na minha graduação.

Ao meu orientador Carlos Henrique Ribeiro Lima, que esteve sempre presente e disposto a me apoiar durante este estudo, como também em trabalhos passados, juntamente com meu coorientador Marcus Suassuna Santos.

Aos meus amigos, que tive o privilégio de conhecer na Universidade de Brasília: Hiarque, Caio, Miguel, Júlia, Camilla, Marcella, Letícia, e outros que foram companheiros nesta jornada acadêmica, oferecendo todo o apoio nos momentos de tensão de prova e trabalho.

Aos amigos que conheço antes da faculdade e que considero como irmãos. A Família Orgutal: Moreira, Thulio e Soares, que são parceiros de longa data e que foram essenciais para tornar esse período mais equilibrado. Aos chapas Arthur e Yan, que da mesma forma foram fundamentais neste tempo de UnB.

Aos meus amigos de trabalho na empresa Double Eco: Thaysa, Marcela, Francisco e Adriana, que me proporcionaram a experiência profissional para me tornar um engenheiro ambiental na prática.

À Universidade de Brasília, que no semestre de execução deste trabalho, sofreu com a paralização das aulas presenciais, mas que se mantém como uma instituição focada na pesquisa e inovação, incentivando seus alunos ao sucesso.

RESUMO

Este estudo teve por objetivo principal desenvolver um modelo estatístico, baseado em regressão linear com preditores climáticos, para previsão do hidrograma anual de vazões do rio Madeira, sendo este um local com curso d'água distribuído para atividades que passa pela necessidade do poder público monitorar seu comportamento. A série histórica de vazão levantada da estação fluviométrica de Porto Velho foi dividida em hidrogramas anuais de cheia, e feita a separação entre seus componentes, a fim de avaliar a porção de escoamento com resposta direta. Dessa forma, os dados calculados foram submetidos a uma padronização, se tornando vazões adimensionais. Com o objetivo de identificar uma distribuição que melhor se ajuste ao formato da nova série, foram feitas modelagens de quatro diferentes distribuições estatísticas: Beta, Gama, Weibull e Log-Normal. Observando o melhor ajuste representado pela função Beta, foram tomados seus dois parâmetros, o volume e a duração da cheia, como objetos de correlação entre 14 índices climáticos, a fim de identificar as teleconexões climáticas com maior interferência na bacia do rio Madeira. Adotando assim, mediante a correlação, os índices Niño 3 e *Pacific Decadal Oscillation* (PDO) foram utilizados como preditores no modelo de regressão. Os resultados encontrados de previsão foram pouco precisos neste estudo, muito ligado ao fato da adoção de uma análise de regressão individual à cada parâmetro, onde havia correlação significativa entre estes, sendo recomendável o uso de um modelo com variáveis agrupadas.

Palavras-chave: previsão de cheias, teleconexões climáticas, funções densidade de probabilidade

SUMÁRIO

1. INTRODUÇÃO	1
2. OBJETIVOS	3
2.1. OBJETIVO GERAL	3
2.2. OBJETIVOS ESPECÍFICOS.....	3
3. FUNDAMENTAÇÃO TEÓRICA E REVISÃO	
BIBLIOGRÁFICA	4
3.1. CARACTERIZAÇÃO FÍSICA DA BACIA DO RIO MADEIRA	4
3.1.1. Geomorfologia e Hidrologia	4
3.1.2. Clima	5
3.2. PREVISÃO DE HIDROGRAMAS ANUAIS	6
3.2.1. Análise do escoamento de base	7
3.2.2. Modelagem do hidrograma anual com funções densidade de probabilidade... 8	
3.3. TELECONEXÕES CLIMÁTICAS	11
3.3.1. Índices Climáticos	13
3.3.2. Índices SST	14
3.3.3. Índices de Temperatura	15
3.4. ANÁLISE DE REGRESSÃO COM PREDITORES CLIMÁTICOS	16
4. METODOLOGIA	18
4.1. DADOS CLIMÁTICOS E HIDROLÓGICOS	18
4.2. SEPARAÇÃO DO ESCOAMENTO DE BASE	19
4.3. AJUSTE DAS FUNÇÕES DE DENSIDADE PROBABILIDADE.....	20
4.4. CORRELAÇÃO ENTRE OS DADOS HIDROCLIMÁTICOS E O	
DESENVOLVIMENTO DO MODELO DE PREVISÃO	21
5. RESULTADOS E DISCUSSÕES	23
5.1. DISTRIBUIÇÕES DE PROBABILIDADE AJUSTADAS AO	
HIDROGRAMA DE CHEIA ANUAL	23
5.2. CORRELAÇÃO ENTRE OS DADOS CLIMÁTICOS E O MODELO DE	
PREVISÃO	29
6. CONCLUSÕES E RECOMENDAÇÕES.....	33
REFERÊNCIAS BIBLIOGRÁFICAS	34

APÊNDICE	35
 APÊNDICE A - CÓDIGO-FONTE EM LINGUAGEM PYTHON UTILIZADO NO ESTUDO	36

LISTA DE FIGURAS

Figura 3.1 Localização da bacia do rio Madeira.	4
Figura 3.2 Componentes de resposta de um hidrograma.....	7
Figura 3.3 Resumo das regiões usadas pelos índices Niño. 4 corresponde ao Niño 4; 3.4 ao Niño 3.4; 3 ao Niño 3 e 1+2 ao Niño 1+2.	15
Figura 3.4 Funcionamento do método de validação cruzada <i>leave-one-out</i>	17
Figura 4.1 Série temporal de vazões médias diárias da estação de Porto Velho.....	18
Figura 5.1 Vazões médias diárias e escoamento de base em vermelho (a) e escoamento superficial rápido (b).....	23
Figura 5.2 Vazões médias diárias agrupadas em hidrogramas anuais de cheia..	24
Figura 5.3 Hidrogramas médios adimensionais.	24
Figura 5.4 Ajuste da distribuição Beta aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b)..	25
Figura 5.5 Ajuste da distribuição Gama aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b).....	25
Figura 5.6 Ajuste da distribuição Log-Normal aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b).....	26
Figura 5.7 Ajuste da distribuição Weibull aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b).....	26
Figura 5.8 Ajuste das distribuições Beta, Gama, Log-Normal e Weibull aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b).....	27
Figura 5.9 RMSE do ajuste das funções de densidade probabilidade.	28
Figura 5.10 R ² do ajuste das funções densidade probabilidade.....	28
Figura 5.11 Correlação cruzada entre I9 e Beta (a) e I13 e a Duração (b).....	29
Figura 5.12 Validação cruzada entre os dados previstos e observados de alfa (a), beta (b), volume (c) e duração (d).....	31
Figura 5.12 Dados observados e previstos do hidrograma de cheia de 2013 a 2014 (a) e de 1996 a 1997 (b).....	31
Figura 5.14 Representação gráfica da matriz de correlação dos parâmetros.....	32

LISTA DE TABELAS

Tabela 3.1 Função densidade probabilidade, tempo de pico e vazão de pico para a distribuição Beta.	9
Tabela 3.2 Função densidade probabilidade, tempo de pico e vazão de pico para a distribuição Weibull.	9
Tabela 3.3 Função densidade probabilidade, tempo de pico e vazão de pico para a distribuição Gama.	10
Tabela 3.4 Função densidade probabilidade, tempo de pico e vazão de pico para a distribuição Log-Normal.	11
Tabela 4.1 Índices climáticos utilizados.	19
Tabela 5.1 Parâmetros estimados das funções Beta, Gama, Log-Normal e Weibull para o hidrograma médio de 2013 a 2014.	27
Tabela 5.2 Parâmetros estimados das funções Beta, Gama, Log-Normal e Weibull para o hidrograma médio de 1996 a 1997.	27
Tabela 5.3 Matriz de valores máximos absolutos de correlação cruzada entre as regiões e os índices.	30

LISTA DE SÍMBOLOS, NOMENCLATURAS E ABREVIACÕES

AB	- Alta da Bolívia
ANA	- Agência Nacional de Águas
a, b	- Parâmetros da distribuição Beta
ENSO	- El Niño/Oscilação Sul
HARDCRUT4	- Banco de dados de temperatura global
MJO	- Oscilação de Madden e Julian
MMQ	- Método dos Mínimos Quadrados
k, λ	- Parâmetros da distribuição Weibull
μ, σ	- Parâmetros da distribuição Log-Normal
NAO	- Oscilação do Atlântico Norte
ρ	- Coeficiente de correlação de Pearson
PSA	- Pacífico da América do Sul
q_p	- Vazão de pico
RMSE	- Raiz do erro médio quadrático
R^2	- Coeficiente de determinação
SST	- Temperatura da Superfície do Mar
SAM	- Modo Anular Sul
t_p	- Tempo de pico
τ, θ	- Parâmetros da distribuição Gama
ZCIT	- Zona de Convergência Intertropical

1- INTRODUÇÃO

A relação entre clima e o ciclo hidrológico pode ser observada em uma escala local, regional ou global. É perceptível que as cheias são correspondentes aos períodos de estação chuvosa, e nos momentos de diminuição do nível dos rios, com os períodos de seca. Essa relação entre precipitação e vazão ou nível d'água, se torna uma relação direta, sendo objeto de estudo via modelagem hidrológica do sistema. A partir disso, é possível o desenvolvimento de metodologias, como sistemas de alerta e previsão, que possam ser capazes de amenizar os impactos negativos das cheias e inundações, como a destruição de bens, ao meio ambiente e a vida humana.

Porém a precipitação em uma região não é o suficiente para representar o escoamento de um rio. A existência desta grande variabilidade que a hidrologia e o clima possuem, incentiva o estudo na tentativa de identificar os principais fatores e respostas para esta indefinição (Diniz, 2008). Alguns desses componentes climáticos, que se faz como objeto de estudo para este trabalho, são elementos para padrões ocorrentes em uma região, gerando anomalias climáticas entre áreas diferentes do globo, que são conhecidos como teleconexões climáticas. Estes fenômenos estão relacionados com características de um local, como a diferença de pressão ou entre a diferença de temperatura em uma tal época do ano, que se tornam um outro instrumento de estudo de correlação para a hidrologia (Cavalcanti, 2009).

Este trabalho traz essa proposta aplicado diretamente à bacia do rio Madeira, região esta que é composta por inúmeras atividades que faz jus a necessidade do entendimento e da implantação de previsões hidrológicas, como a pesca, o sistema hidroviário e o complexo hidrelétrico composto por duas grandes usinas hidrelétricas, a de Jirau e a de Santo Antônio, que é responsável pela geração de energia para parte da região Centro Oeste e Sudeste do Brasil (Cemaden, 2018).

Para o desenvolvimento estatístico deste modelo foi feita a separação do escoamento de base dos hidrogramas, a fim de analisar a porção de escoamento superficial rápido. Com o atual hidrograma separado, foi realizada a sua modelagem utilizando quatro funções densidade de probabilidade: Beta, Gama, Log-Normal e Weibull. Tomando a distribuição com melhor ajuste, são identificados quatro parâmetros para a correlação entre os índices climáticos: os dois parâmetros da função, o volume da cheia e a duração da cheia. Realizando a correlação entre os índices climáticos e os quatro parâmetros hidrológicos, foi feita a seleção dos índices

com maior valor de coeficiente de Pearson, partindo para o desenvolvimento do modelo, baseado na regressão linear de cada parâmetro.

Após essa introdução, o trabalho é dividido em seis capítulos. O primeiro faz uma breve introdução da importância do trabalho. O segundo capítulo apresenta os objetivos a serem atingidos com o desenvolvimento do trabalho. O terceiro, traz as referências bibliográficas. O quarto capítulo apresenta uma proposta de metodologia a ser seguida para que possa se atingir os resultados desejados. No quinto capítulo são apresentados os resultados atingidos e um breve discussão, onde serão interpretados e avaliados no sexto capítulo. E por fim são exibidas as referências bibliográficas utilizadas no trabalho, seguido do apêndice.

2- OBJETIVOS

2.1- OBJETIVO GERAL

Desenvolver um modelo estatístico com preditores climáticos para previsão do hidrograma anual de cheia do rio Madeira.

2.2- OBJETIVOS ESPECÍFICOS

- Avaliar o ajuste e adequação de diversas funções ao hidrograma anual de escoamento rápido do rio Madeira;
- Identificar como os parâmetros da função escolhida se correlacionam com os elementos principais do hidrograma (duração e volume);
- Identificar a influência das teleconexões climáticas na bacia do rio Madeira, em particular nos parâmetros da função que modela o hidrograma anual de vazões.

3- FUNDAMENTAÇÃO TEÓRICA E REVISÃO BIBLIOGRÁFICA

Nesta seção é feita uma caracterização da bacia de interesse e alguns exemplos de padrões e anomalias climáticas que impactam os padrões de ocorrência de chuva e de vazão no território brasileiro. Destaca-se ainda a parte estatística que será utilizada no desenvolvimento do trabalho, como a introdução de distribuições estatísticas e os métodos de correlação.

3.1- CARACTERIZAÇÃO FÍSICA DA BACIA DO RIO MADEIRA

3.1.1- Geomorfologia e Hidrologia

A bacia do rio Madeira está localizada na maior bacia hidrográfica do mundo, a bacia Amazônica, representa 23% de sua área. Ela apresenta 1.370.000 km², se estendendo pela Bolívia (51%), Brasil (42%) e Peru (7%) (GOYOT, 1993).

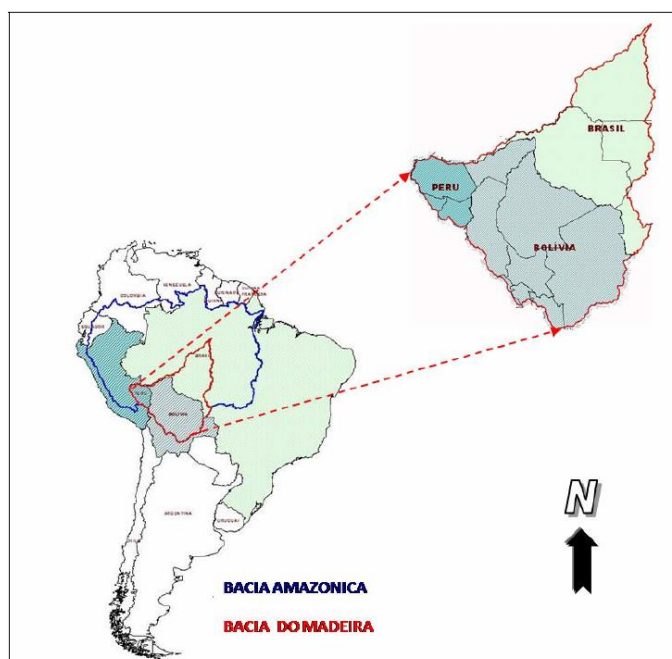


Figura 3.1 – Localização da bacia do rio Madeira
Fonte: Daza (2011)

A bacia apresenta uma geomorfologia dividida em três grandes unidades morfo-estruturais, sendo elas: a cordilheira dos Andes, a planície Amazônica e o escudo brasileiro. Estes diferentes relevos distribuídos, segundo Daza (2011) fazem interferência direta tanto no clima quanto na hidrografia da bacia. O rio Madeira também apresenta a característica de ser um rio novo, pois apresenta um extenso comprimento de margens erodidas e de sedimentos sendo depositados ao longo do curso.

Quanto ao regime hidrológico, em estudos realizados por Jean Rodier (1964), a bacia do rio Madeira fora classificada com a característica de possuir rios de regime tropical austral, com um pico de cheia, normalmente apresentado no primeiro semestre do ano calendário. Já Guyot (1993), com a intenção de analisar a fração boliviana da bacia, também exibe a mesma classificação quanto ao regime, porém adicionando a informação da maior regularidade dos rios das planícies, em comparação com os rios andinos, sendo melhor o comportamento entre o nível d'água e a vazão.

3.1.2- Clima

O clima da bacia do rio Madeira se caracteriza por ser tropical, com um período de seca curto, sendo classificada, predominantemente, em Am, pela classificação climática de Köppen, que analisa os valores médios anuais e mensais de temperatura do ar e precipitação (Hayakawa *et al.* 2015). Ou seja, caracteriza-se por apresentar temperatura média do mês mais frio superior a 18°C, com uma estação seca de pequena duração que é compensada pelos totais elevados de precipitação (Golfari *et al.*, 1978). As combinações de fatores como a temperatura, radiação solar, umidade do ar, ventos, cobertura vegetal, pressão atmosférica e precipitação determinam o clima atual da área. De uma forma geral, as características básicas da bacia predominam pela alta umidade relativa do ar, elevado grau de nebulosidade, precipitação abundante e temperatura média alta.

Entretanto, se faz de grande influência para a região, os controles climáticos mais dinâmicos, relacionados com os sistemas de circulação atmosférica.

3.1.2.1- Sistemas de circulação atmosférica

A circulação geral da atmosfera é resultado de uma distribuição desigual de energia na superfície do globo, onde em regiões polares recebem menos energia solar que regiões equatoriais e tropicais (Longo *et al.*, 2013). Este resultado se dá por trocas de calor sensível ou latente, no deslocamento de massas de ar, por processos de condensação ou pelas correntes oceânicas. Estes sistemas para a bacia de estudo, impõe para a região as características gerais do clima regional, podendo estar a nível global, regional ou local (Proclira, 2007). Segundo Daza (2011), dentre esses sistemas de circulação atmosférica que mais condicionam a área, podem-se citar os:

- Sistemas de norte, relacionadas com a Zona de Convergência Intertropical (ZCIT);

- Sistemas de sul, correspondente ao anticiclone polar com influência de frentes frias, levando a baixas temperaturas e chuvas intermitentes;
- Anticiclone da Alta da Bolívia (AB), estando relacionado com a forte convecção da região da Amazônia;
- Anticiclones subtropicais do Atlântico Sul e os anticiclones dos Açores, sendo associados a tempos estáveis em período de menor precipitação;
- Sistemas de depressões barométricas no interior de massas de ar equatorianas, devido à forte incidência de radiação solar e a baixa latitude, estando relacionadas com as gerações de zonas de instabilidade, com chuvas mais intensas, caracterizadas chuvas de verão.

3.2- PREVISÃO DE HIDROGRAMAS ANUAIS

Com objetivo do trabalho definido em poder prever hidrogramas de cheias anuais, ou seja, poder estimar o escoamento com antecedência no tempo, os benefícios que envolvem a previsão de vazão pode ser ligada a diversas áreas distintas. Desde os usos no abastecimento de água, na irrigação, navegação, hidrelétricas, inundações e na conservação ambiental, dependem do entendimento da quantidade de água disponível nos rios. Esta previsão de vazão, pode ser realizada tanto em curto prazo, com antecedência de poucas horas até cerca de 14 dias, ou em longo prazo, com antecedência de até nove meses (Georgakakos e Kryzstofowicz, 2001), onde este estudo melhor se enquadra.

Nos anos recentes houve uma ampliação do entendimento das conexões entre anomalias climáticas de grande escala e os eventos hidrológicos, o que favorece a muitos trabalhos poder desenvolver modelos que façam compreender essa conexão, ampliando também as possibilidades do uso de previsões hidroclimáticas de longo prazo (Silva *et al.* 2006). Esta compreensão para as bacias brasileiras mostram grande potencial, visto que a matriz energética brasileira é predominantemente baseada em energia hidráulica. A previsão a longo prazo permite a estimativa de afluências aos reservatórios hidrelétricos, possibilitando planejar a melhor operação (Tucci, *et al.* 2003).

Para este trabalho, na organização dos dados hidrológicos e climáticos para o desenvolvimento deste modelo, são necessários seguir algumas etapas em sua preparação, a fim de obter os elementos que melhor se correlacionarão entre as duas componentes.

3.2.1- Análise do escoamento de base

Um hidrograma é um registro de série temporal das condições de escoamento, como o nível ou da vazão de água, em um local de medição. Esta série pode ser um agregado de diferentes fontes de água que contribuem para o escoamento. Segundo Brodie *et al.* (2005), esses componentes podem ser subdivididos em escoamento superficial rápido e escoamento de base, como visto na figura 3.2.

- Escoamento superficial rápido (Quickflow): onde se obtém a resposta direta a um evento de chuva, como o próprio escoamento superficial de um rio, o movimento lateral do perfil do solo e a precipitação direta sobre a superfície.
- Escoamento de base (Baseflow): resposta dada de um longo prazo devido a armazenamentos naturais, oriundos também de fontes subterrâneas.

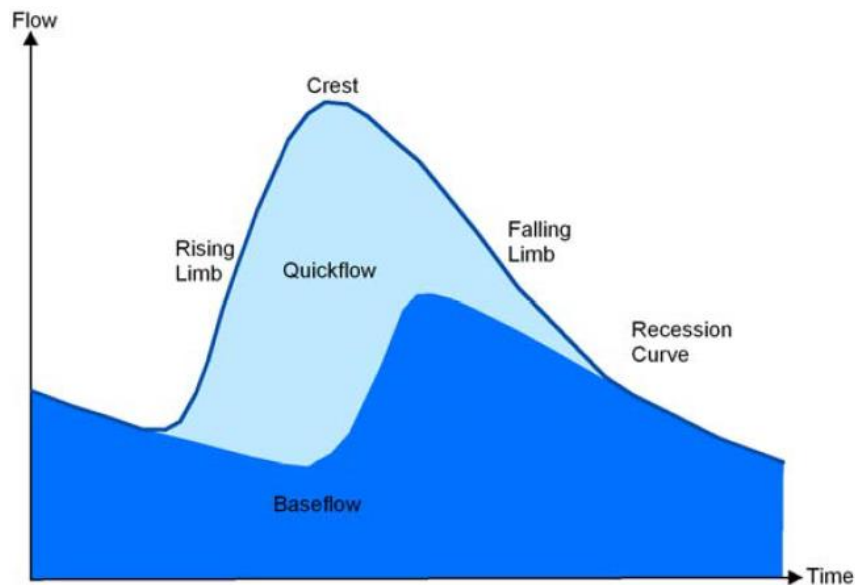


Figura 3.2: Componentes de resposta de um hidrograma
Fonte: Brodie *et al.*, 2005.

Com o objetivo de analisar apenas a porção de resposta rápida do hidrograma, sendo esta a parcela utilizada como dado de entrada para modelos de previsão de eventos de enchentes e inundações, existem técnicas de separação entre os componentes, podendo ser de forma gráfica ou por filtros digitais (Pramanik *et al.*, 2010).

As técnicas gráficas de separação são utilizadas para observar o componente do escoamento de base de um evento de hidrograma de inundação, diretamente no gráfico, incluindo um ponto onde se presume que o escoamento subsequente seja totalmente de base, e a outra parte seja a de escoamento superficial rápida. Por outro lado, as técnicas baseadas em filtros

recursivos processam todo o hidrograma do escoamento para derivar o hidrograma de escoamento básico. Estes filtros são ferramentas de análise de sinais que removem o sinal de escoamento rápido de alta frequência para derivar um sinal de escoamento básico de baixa frequência, sendo necessário a um processo de calibração de parâmetro do filtro, para que os resultados possam ser considerados numericamente válidos (Brodie *et al.*, 2005). Lyne *et al.* (1979), em seus estudos aplicou um destes filtros digitais, desenvolvido por Lynn em 1973, que foi utilizado como referência neste trabalho.

3.2.2- Modelagem do hidrograma anual com funções densidade de probabilidade

Com o intuito de representar matematicamente uma série temporal, os modelos de distribuição de probabilidades são utilizados para sintetizar o comportamento de variáveis aleatórias. Estes modelos são uma forma abstrata matemática, com características intrínsecas de variabilidade e conformação, que representam as variações possíveis de uma variável aleatória. De uma forma paramétrica, o modelo é descrito por parâmetros que definem, por valores numéricos, uma certa amostra de observações. Uma vez estimados estes valores, pode-se constituir uma síntese plausível do comportamento da variável aleatória, para posteriormente, manipular e gerar correlações com outras variáveis (Naghetini e Pinto, 2007).

Porém, de forma paralela, o formato das funções densidade de probabilidade, segundo Bhunya *et al.* (2004), é semelhante a de um hidrograma convencional, podendo assim alterar seus parâmetros, a fim de ajustar as dimensões dos hidrogramas do estudo. Neste trabalho, serão abordadas quatro distribuições estatísticas: Distribuição Beta, Distribuição Weibull, Distribuição Gama e Distribuição Log-Normal.

A distribuição Beta foi desenvolvida por Johnson e Kotz em 1970. Ela foi representada por duas variações, uma com dois parâmetros e outra com três parâmetros. Para fins deste estudo será abordado a distribuição Beta dois parâmetros. A expressão matemática que a define está representada na tabela 3.1, juntamente com as representações do tempo do pico (t_p) e da vazão de pico (q_p). A função Beta, $B(a, b)$, representada na equação 3.1, é descrita pelos parâmetros a e b , no qual sempre são positivos, e pelos termos t e q , que representam o tempo e vazão, respectivamente (Pramanik *et al.*, 2009). Quando ambos os parâmetros são iguais, a densidade Beta é simétrica, contrariamente, a distribuição é assimétrica. Se os parâmetros forem maiores que 1, a densidade Beta é unimodal, e para suas diferentes variações, faz com

que gere modelações onde se $a > b$, ela é negativamente assimétrica e o contrário se $a < b$. (Naghattini e Pinto, 2007). O mérito da distribuição Beta para o ajuste com os hidrograma unitários é que ela fornece todas as formas possíveis, dependendo da magnitude de seus parâmetros (Bhunya *et al.*, 2004).

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt \quad (3.1)$$

Tabela 3.1: Função densidade probabilidade, tempo de pico e vazão de pico para a distribuição Beta.

Função de distribuição de probabilidade	$q(t, a, b) = \frac{1}{B(a, b)} t^{a-1}(1-t)^{b-1}$
Vazão de pico (q_p)	$q_p = \frac{1}{B(a, b)} \left(\frac{a-1}{a+b-2}\right)^{a-1} \left(\frac{b-1}{a+b-2}\right)^{b-1}$
Tempo de pico (t_p)	$t_p = \frac{a-1}{a+b-2}$

A distribuição Weibull foi aplicado pela primeira vez por Rosin e Rammler em 1993. Ela é representada pelos parâmetros k e λ , que são números positivos que representam a forma e a escala do hidrograma, respectivamente. Sob certos valores, a função Weibull representa a distribuição normal e exponencial. O parâmetro de escala (λ), é responsável na alteração das dimensões do hidrograma, ampliando ou reduzindo, enquanto a assimetria da função é obtida na combinação entre os dois parâmetros. Dado um valor fixo para λ , quanto maior for k , maior será o hidrograma. A tabela 3.2 mostra as expressões matemáticas para esta função (Pramanik *et al.*, 2009).

Tabela 3.2: Função densidade probabilidade, tempo de pico e vazão de pico para a distribuição Weibull.

Função de distribuição de probabilidade	$q(t, k, \lambda) = \frac{k}{\lambda} \left(\frac{t}{\lambda}\right)^{k-1} \exp - \left(\frac{t}{\lambda}\right)^k$
Vazão de pico (q_p)	$q_p = \frac{k}{\lambda} \left(1 - \frac{1}{k}\right)^{1-\frac{1}{k}} \exp\left[-\left(1 - \frac{1}{k}\right)\right]$
Tempo de pico (t_p)	$t_p = \lambda \left(\frac{k-1}{k}\right)^{1/k}$

A distribuição Gama apresenta dois parâmetros τ e θ que representam a forma e a escala, respectivamente, e é amplamente usada para modelar hidrograma, muito devido a sua versatilidade de formas, ao coeficiente de assimetria variável e positivo, junto com o fato das variáveis aleatórias não serem definidas para valores negativos. Abaixo está a representação da função Gama, $\Gamma(\tau)$ para valores inteiros de τ !, juntamente com a tabela 3.3 que mostra as expressões matemáticas para o hidrograma (Pramanik *et al.*, 2009).

$$\Gamma(\tau) = (\tau - 1)! \quad (3.2)$$

Tabela 3.3: Função densidade probabilidade, tempo de pico e vazão de pico para a distribuição Gama.

Função de distribuição de probabilidade	$q(t, \tau, \theta) = t^{\tau-1} \frac{\exp(-\frac{t}{\theta})}{\theta^\tau \Gamma(\tau)}$
Vazão de pico (q_p)	$q_p = \frac{(\tau - 1)^\tau \exp(1 - \tau)}{\Gamma(\tau) \theta (\tau - 1)}$
Tempo de pico (t_p)	$t_p = \theta(\tau - 1)$

Por fim, a distribuição Log-Normal trata-se de uma função de densidade de probabilidade unimodal de qualquer variável aleatória que tenha um número grande de componentes aleatórios independentes, em decorrência do limite central seja normalmente distribuído. É representado por dois parâmetros, μ e σ que representam a média e o desvio padrão, respectivamente, e momentos de $\ln(t)$. É utilizada para valores de t maiores que 0, e com μ variando entre $-\infty$ e $+\infty$. A tabela 3.4 apresenta as expressões matemáticas da distribuição Log-Normal para serem aplicadas ao hidrograma (Naghattini e Pinto, 2007).

Tabela 3.4: Função densidade probabilidade, tempo de pico e vazão de pico para a distribuição Log-Normal.

Função de distribuição de probabilidade	$q(t, \mu, \sigma) = \frac{1}{t\sigma\sqrt{2\pi}} \exp^{-(\ln(t)-\mu)^2/2\sigma^2}$
Vazão de pico (q_p)	$q_p = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{\sigma^2}{2} - \mu\right)$
Tempo de pico (t_p)	$t_p = \exp(\mu - \sigma^2)$

3.3- TELECONEXÕES CLIMÁTICAS

O termo teleconexão está relacionado com a ligação entre anomalias climáticas locais e forçantes localizadas geralmente a grandes distâncias (Reboita, 2015). A circulação atmosférica é considerada um processo dinâmico, apresentando grande variabilidade, observado em diversas escalas de tempo. Os padrões gerais de circulação atmosférica comportam grande variabilidade, podendo ser condicionada por padrões de teleconexões climáticas, resultado de diferentes padrões de temperatura e pressão nos oceanos. A análise de associação desses fenômenos pode ser explorada por técnicas estatísticas, por meio de correlações espaciais, entre anomalias de uma variável em diferentes pontos (Cavalcanti, 2009). Um exemplo de teleconexão está associado com o fenômeno El Niño, que é provocado pelo aquecimento anormal das águas do Oceano Pacífico, que propicia condições para o aumento da probabilidade de ocorrência de seca no Nordeste brasileiro e maior intensidade de chuva no Sul do Brasil (Daza, 2011). A partir destes efeitos observados pelas diferentes teleconexões, como entre o El Niño e a precipitação de uma região, seria possível utilizar esse indicador como preditor de parâmetros de modelos hidrológicos em escala local.

O estudo realizado¹ sobre os principais teleconexões que afetam as regiões do Brasil são: o El Niño-Southern Oscillation (ENSO), Madden and Julian Oscillation (MJO), a North Atlantic Oscillation (NAO), o Southern Annular Mode (SAM) e o Pacific South America (PSA) pattern. Segundo Mântua (1997), o Pacific Decadal Oscillation (PDO) também exerce influência.

¹ Estudo fornecido por Iracema F.A. Cavalcanti no Instituto Nacional de Pesquisas Espaciais (INPE), em março de 2018. Disponível: <https://www.ufsm.br/app/uploads/sites/453/2018/05/palestra_stmaria_mar2018_F.pdf>

O ENSO (El Niño-Southern Oscillation) é uma mudança periódica do sistema oceano-atmosfera no Pacífico tropical que afeta o clima em todo o mundo. A variação irregular deste fenômeno gera duas faces opostas. Um dos extremos é resultado do aquecimento das águas simultaneamente com a diminuição da pressão atmosférica no Pacífico Leste, chamado de El Niño. A outra face, com o resfriamento das águas e o aumento na pressão atmosférica na região leste do Pacífico, nomeado de La Niña. Para Daza (2011), este é considerado o fenômeno que causa maior influência em relação as precipitações na bacia do rio Madeira, podendo em períodos de maior intensidade desta oscilação, apresentar secas e grandes inundações.

Quanto ao MJO (Madden and Julian Oscillation) trata-se de uma oscilação caracterizada pelo distúrbio de nuvens, chuvas, ventos e pressão que se move para leste, atravessando pelos trópicos do planeta e retornando ao seu ponto inicial, em períodos de 30 a 60 dias. Esta perturbação atmosférica difere do ENSO, que mostra características mais resistentes, devido a maior duração de influência. Logo, o MJO é melhor descrito por ter uma variabilidade climática tropical intrassazonal, ou seja, que varia em uma base de semana para semana. (Kayano *et al.*,2009).

O NAO (North Atlantic Oscillation) é associado a uma gangorra de pressão atmosférica em larga escala na região do Atlântico Norte, estando relacionado pelas mudanças de padrões climáticos locais, como a temperatura, precipitação, força e direção do vento, principalmente voltado para o hemisfério norte, mas podendo ser analisado seus efeitos na parte sul. As características de pressão observadas no Oceano Atlântico Norte diferem em regiões centrada nas ilhas dos Açores, apresentando pressão alta, e baixa centrada na Islândia. As mudanças relativas de pressão entre essas regiões que regem o efeito climático do NAO (NOAA, 2019).

O SAM (Southern Annular Mode) descreve um movimento norte-sul de fortes ventos vindos do oeste, que atua continuamente em latitudes médias e altas do hemisfério sul. Esse cinturão de ventos, ao mudar de posição, influencia na força e na direção de frentes frias e de sistemas de tempestades, sendo este um impulsor na variabilidade de chuvas, com maior atuação na Austrália. Similar com outros demais sistemas de oscilação, a atuação do SAM varia quando positivo e negativo. Quando positivo, o cinturão de ventos se contrai na Antártida, levando a ventos mais fracos, maiores pressões na região sul da Austrália e a restrição de

frentes frias no interior. Já em SAM negativo, ocorre a expansão destes ventos, resultando em tempestades mais fortes e sistemas de baixa pressão na região australiana (BoM, 2019).

Quanto o PSA (Pacific South America), resumidamente, apresenta dois modos distintos de variabilidade atmosférica no Hemisfério Sul, que influencia as anomalias de circulação e precipitação na América do Sul. Estes modos são chamados de PSA e PSA2. O PSA1 está relacionado com anomalias a Temperatura da Superfície do Mar (SST) sobre o Pacífico central e oriental, suprimindo a convecção sobre o Oceano Índico, sendo considerado uma resposta do ENSO, resultando em déficit de chuvas no Nordeste e chuvas intensas no Sudeste (Kingtse *et al.*, 2001). Já o PSA2, está associado com a intensificação da convecção no Pacífico Central, próximo ao sul do Equador, e a supressão no Pacífico Oeste (Reboita *et al.*, 2014).

O PDO (Pacific Decadal Oscillation) é descrito como um padrão de longa duração semelhante ao El Niño, de variabilidade climática do Pacífico. Os extremos do PDO são marcados por variações generalizadas na Bacia do Pacífico e no clima da América do Norte. Paralelo ao fenômeno ENSO, as fases extremas do PDO são classificadas como quentes ou frias, conforme definido pelas anomalias da temperatura do oceano no nordeste e no oceano Pacífico tropical. Quando as SSTs são anormalmente frias no interior do Pacífico Norte e quentes ao longo da costa do Pacífico, e quando as pressões ao nível do mar estão abaixo da média no Pacífico Norte, o PDO tem um valor positivo (Mântua *et al.*, 1997).

Apesar da atuação de algumas dessas teleconexões climáticas comentadas não atuarem diretamente na bacia do rio Madeira, a influência de suas respostas pode gerar anomalias de baixo ou de alto impacto no regime hidrológico da região. Para este estudo, são utilizados Índices Climáticos, que quantificam essas variáveis climáticas, como a pressão superficial e a temperatura da superfície do mar, por séries históricas temporais. Isso possibilita a correlação entre as variáveis hidrológicas, podendo assim analisar a atividade dessas teleconexões na bacia de interesse.

3.3.1- Índices Climáticos

Um índice é um valor calculado, com o objetivo de descrever o estado e as mudanças no sistema climático. Estes índices permitem um estudo estatístico das variações dos aspectos climatológicos dependentes, analisando e comparando as séries temporais, as médias, os extremos e as tendências. São baseados em parâmetros que descrevem os aspectos do clima,

existindo assim uma grande variedade de índices climáticos que são criados e examinados na literatura. Cada índice climático porta uma equação que são chamados de elementos climáticos, que são parâmetros mensuráveis que influenciam outros parâmetros atmosféricos, como a temperatura do ar, precipitação e radiação solar, mas também parâmetros não atmosféricos, como a temperatura do mar ou cobertura do gelo (UHH, 2017).

Como exemplo destes índices climáticos, se pode citar alguns importantes para o estudo, definido pela Administração Oceânica e Atmosfera Nacional (NOAA) dos Estados Unidos:

- Southern Oscillation Index (SOI): é definido pela diferença de pressão normalizada entre o Taiti e Darwin.
- North Atlantic Oscillation (NÃO): é definido pela diferença de pressão normalizada entre uma estação nos Açores e na Islândia.
- Arctic Oscillation (AO): oscilação típica de região polar, que influencia particularmente o Hemisfério Norte, com maior potencialidade. É calculado projetando anomalias diárias no padrão de carregamento.
- Southern Annular Mode (SAM): é definido pela variabilidade de baixa frequência dos extratropicais do Hemisfério Sul, sendo calculada pela diferença de pressão média zonal ao nível do mar entre 40°S e 65°S.
- Trans Polar Index (TPI): é definido como a diferença de pressão normalizada entre Hobart e Stanley.
- Pacific Decadal Oscillation (PDO): é definido pelo cálculo da média do SST com base na covariância temporal na bacia do Pacífico.
- Tripole Index for the Pacific Interdecadal Oscillation (TPI IPO): é baseado na diferença entre o SST calculado sobre o Pacífico equatorial central e a média do SST no noroeste e no sudoeste do Pacífico.
- Dipole Mode Index (DMI): é um gradiente anômalo de SST entre o Oceano Índico equatorial ocidental e o Oceano Índico equatorial sudeste.

3.3.2- Índices SST

Estes são índices que utilizam do gradiente de temperatura da superfície do mar, em análise entre duas estações, e não para apenas uma. Com isso, se obtém padrões atmosféricos e séries temporais para estes índices. Alguns destes são:

- Niño 3: é calculado a partir de um banco de dados de observações realizadas da temperatura da superfície do mar, HadISST1, em uma área média de 5S-5N e 150W-90W.
- Niño 3.4: calculado pelo banco de dados HadISST1, em uma área média área média SST de 5S-5N e 170-120W.
- Niño 4: Calculado a partir do HadISST1 e em uma área média de 5S-5N e 160E-150W.
- Niño 1+2: Calculado a partir do HadISST. É a área SST com média de 0-10S e 90W-80W.

As faixas atribuídas a cada um destes índices de El Niño variam ao longo do Oceano Pacífico, tendo cada um seu uso para observar e monitorar o fenômeno climático. A figura 3.3 mostra uma esquematização da área de atuação de cada índice.

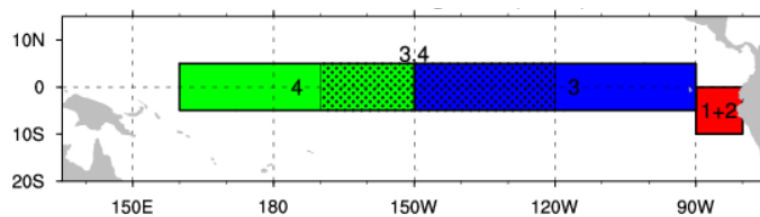


Figura 3.3: Resumo das regiões usadas pelos índices Niño. **4** corresponde ao Niño 4; **3.4** ao Niño 3.4; **3** ao Niño 3 e **1+2** ao Niño 1+2
Adaptado de Shea, 2018.

3.3.3- Índices de Temperatura

Quanto aos índices de temperatura em geral acabam por incluir a temperatura do ar como um parâmetro atmosférico, sendo registrados valores médios, mínimos e máximos que são calculados pelo índice. Além do componente atmosférico, parâmetros de SST podem ser combinados para a geração de um índice que envolva o meio Terra-Oceano. São exemplos destes índices:

- Anomalias médias globais de temperatura: é feito uma média global da temperatura da combinação de registros da temperatura da superfície do mar com a temperatura do ar, pelo banco de dados HADCRUT4.
- Índice de manchas solares: é um registro de números mensais de manchas solares, podendo associar o efeito do Sol com clima da região.

3.4- ANÁLISE DE REGRESSÃO COM PREDITORES CLIMÁTICOS

Obtendo os dados hidroclimáticos a serem analisados, primeiramente como forma de identificar os possíveis efeitos dos fenômenos climáticos quantificados pelos índices, a correlação junto aos parâmetros do hidrograma, com uso do coeficiente de Pearson, indica a força de relação linear entre as variáveis aleatórias (Benesty et al., 2009). A equação 3.3, representa matematicamente esse coeficiente:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.3)$$

onde, x são as variáveis aleatórias climáticas e y são as variáveis aleatórias dos parâmetros do hidrograma, i é o índice e n é n -ésimo termo. Quanto mais próximo o valor de ρ for de 1, mais forte será a correlação entre as variáveis, porém o inverso não é verdadeiro, pois o coeficiente detecta dependências lineares. Porém em casos em que as séries analisadas são conjuntamente normais, a não-correlação equivale a dados independentes (Benesty et al., 2009).

Com a seleção dos índices climáticos que apresentaram maior correlação entre os parâmetros estimados, como forma de encontrar uma relação matemática entre estes, é possível uma abordagem por regressão linear, em que uma série de dados observados y_i é aproximada por uma variável dependente x_i e do conjunto de coeficientes β_i (Zeileis et al., 2003). Essa formulação pode ser representada pela equação 3.4:

$$y_i = x_i \beta_i + \varepsilon_i \quad (3.4)$$

onde ε_i são os erros do modelo. Podendo estar ser uma forma de modelar uma série de dados por meio de um conjunto de variáveis dependentes. Para que este modelo preveja dados observados o mais próximo possível, é necessário que haja sua calibração, sendo capaz de dividir os dados para o treinamento e para o teste deste modelo.

A abordagem realizada neste estudo foi por *leave-one-out*. Esse método pode ser descrito da seguinte forma: (i) para o primeiro ano, considerando-se n anos de observações, o primeiro ano é removido e o modelo é ajustado utilizando-se os $n-1$ dados remanescentes; (ii) o modelo é utilizado para estimar o parâmetro de posição para o ano que ficou de fora do ajuste do modelo; (iii) o procedimento é repetido para os n dados da série histórica (Santos, 2019). A figura 3.4 ilustra o funcionamento desta metodologia.

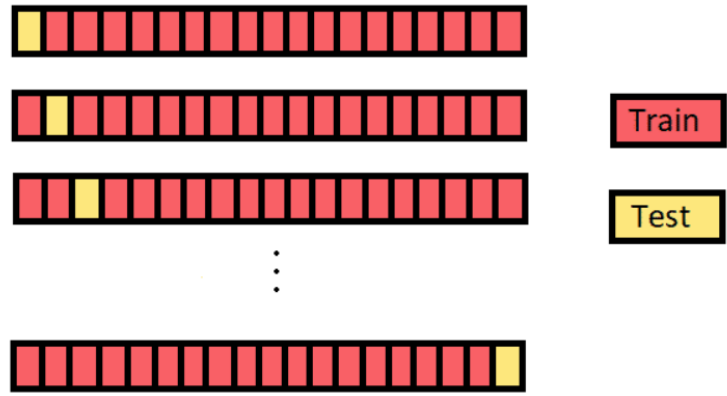


Figura 3.4: Funcionamento do método de validação cruzada *leave-one-out*
Fonte: Soneji, 2017.

4- METODOLOGIA

Para esta etapa do trabalho será descrita a metodologia adotada, desde a obtenção dos dados climáticos e hidrológicos e sua análise estatística, até a identificação da influência de teleconexões climáticas na região da bacia do Rio Madeira, por meio de índices climáticos.

Toda parte estatística do trabalho foi realizada em linguagem Python, estando no apêndice os códigos principais do desenvolvimento da pesquisa.

4.1- DADOS CLIMÁTICOS E HIDROLÓGICOS

Foram utilizados dados de vazão médias diárias, de 1967 a 2019, proveniente da estação fluviométrica Porto Velho, com código 15400000 (Hidroweb) da Agência Nacional de Águas. Com os dados obtidos, foi feita análise de consistência com análise da curva-chave.

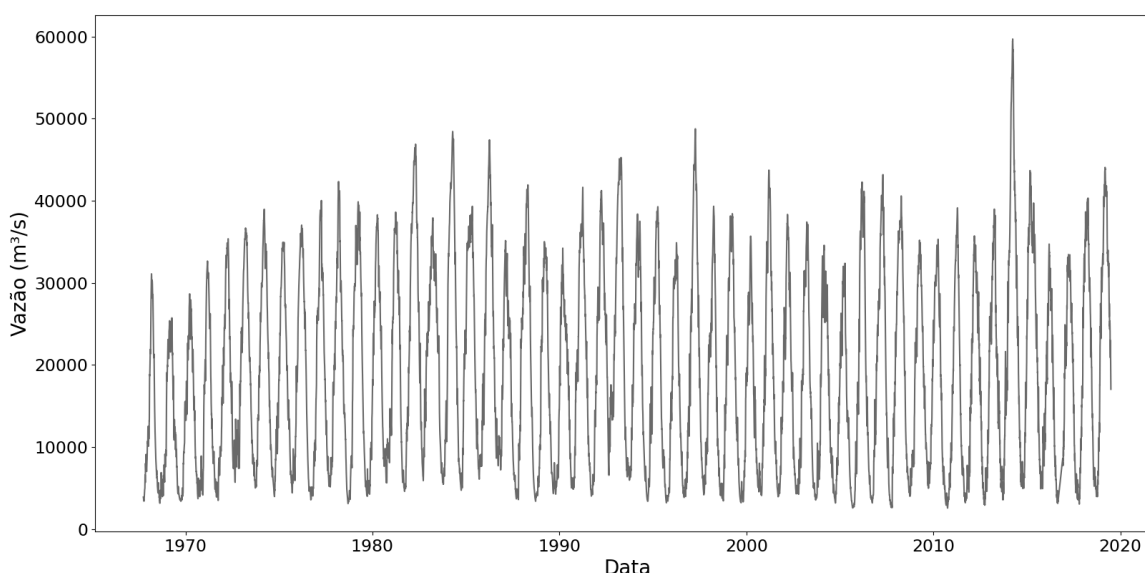


Figura 4.1: Série temporal de vazões médias diárias da estação de Porto Velho.

Foram também utilizados 14 índices climáticos, com dados no intervalo de tempo mensal, e que podem ser consultados pelo banco de dados do Laboratório de Pesquisa do Sistema Terrestre NOAA, pelo Sistema Global de Observação do Clima (GCOS). O sistema pode ser acessado pelo endereço, disponível em: https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/. Para melhor identificação durante o trabalho, a tabela 4.1 apresenta os índices climáticos com suas respectivas abreviações.

Tabela 4.1: Índices climáticos utilizados.

Índice	Nome	Abreviação
1	Southern Oscillation Index (SOI)	I1
2	North Atlantic Oscillation (NAO)	I2
3	Arctic Oscillation (AO)	I3
4	Southern Annular Mode (SAM)	I4
5	Trans Polar Index (TPI)	I5
6	Pacific Decadal Oscillation (PDO)	I6
7	Tripole Index for the Pacific Interdecadal Oscillation (TPI IPO)	I7
8	Dipole Mode Index (DMI)	I8
9	Niño 3	I9
10	Niño 3.4	I10
11	Niño 4	I11
12	Niño 1+2	I12
13	Anomalias médias globais de temperatura	I13
14	Índice de manchas solares	I14

4.2- SEPARAÇÃO DO ESCOAMENTO DE BASE

Com o objetivo de definir os componentes de resposta do escoamento para o hidrograma, foi utilizado a técnica desenvolvida por Lynn (1973), no qual obteve resultados satisfatórios. A técnica de filtragem digital segue a seguinte expressão matemática:

$$f_k = a \cdot f_{k-1} + \frac{(1+a)}{2} \cdot y_k - y_{k-1} \quad (4.1)$$

Onde f é o escoamento superficial rápido, y é o escoamento total, k é o índice temporal, e a é o parâmetro do filtro. O valor do parâmetro adotado foi considerado por análise gráfica,

se aproximando mais das curvas de escoamento de base vistas na literatura, onde ao diminuir o valor de a , o escoamento de base aumenta, e o contrário quando se aumenta a .

Ao separar a série de vazão de interesse para o trabalho, a próxima etapa consistiu em calcular os períodos de cheia para cada ano da bacia. Para isso, foi adotado como hidrograma de cheia, valores de escoamento rápido acima de 0 m³/s em um período mínimo de 100 dias consecutivos. Assim, foi possível retirar os hidrogramas anuais de cheia da série, e conseqüentemente identificar os primeiros dois parâmetros, que serão elementos de correlação com os índices climáticos, a duração da cheia e o volume da cheia. A duração é determinada entre extensão do início do hidrograma até o seu final, e o volume consiste na área abaixo da curva do hidrograma.

4.3- AJUSTE DAS FUNÇÕES DE DENSIDADE PROBABILIDADE

Uma vez obtido os hidrogramas anuais de cheia, para o ajuste da curva de função de probabilidade ao seu formato, o hidrograma é transformado em adimensional. Para este fim, é feita a multiplicação da vazão [m³/s] pela duração do hidrograma [s], e depois é feita a divisão pelo volume [m³], obtendo assim uma série com valores adimensionais.

$$q_{iAd.} = \frac{q_i \left(\frac{m^3}{s} \right) * Duração (s)}{Volume (m^3)} \quad (4.2)$$

As funções de probabilidade utilizadas foram descritas na seção 3.4, no qual o termo comum t encontrado na equação das quatro distribuições, representa os valores de vazão da série, enquanto os outros dois termos distintos são os parâmetros de cada função. Estes dois últimos termos foram objetos para definir o formato da curva de cada função, no qual seus valores são estimados a partir da minimização do erro quadrático médio (RMSE), visto na equação 4.2.

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^n (q_{(o)i} - q_{(fdp)i})^2}{n_j}} \quad (4.3)$$

onde $q_{(o)}$ é vazão observada, $q_{(fdp)}$ é a vazão calculada por cada distribuição, j é o índice para os hidrogramas anuais, i é o índice para as vazões diárias e n é a quantidade de dias.

Ao definir os parâmetros de cada função e ajustando sua curva com os hidrogramas, foi utilizado novamente como critério de qualidade, porém agora para distinguir a distribuição que melhor se ajustou aos dados observados, o RMSE e o coeficiente de determinação (R^2), tendo este valor mais próximo de 1 com a função melhor se ajustando. A representação matemática do R^2 é mostrada na equação 4.3.

$$R^2_j = \left\{ \frac{\sum_{i=1}^n (q_{(o)i} - \overline{q_{(o)}})(q_{(fdp)i} - \overline{q_{(fdp)}})}{\sqrt{\sum_{i=1}^n ((q_{(o)i} - \overline{q_{(o)}})^2) \times \sum_{i=1}^n ((q_{(fdp)i} - \overline{q_{(fdp)}})^2)}} \right\}^2 \quad (4.4)$$

dentre as distribuições trabalhadas, o objetivo é encontrar a função que apresenta o menor desvio em comparação com a série observada e utilizar seus dois parâmetros para as próximas etapas do trabalho.

4.4- CORRELAÇÃO ENTRE OS DADOS HIDROCLIMÁTICOS E O DESENVOLVIMENTO DO MODELO DE PREVISÃO

Após identificar a função de densidade com melhor ajuste ao formato do hidrograma, foram utilizados seus dois parâmetros como objetos de correlação com os índices climáticos, totalizando em 4 elementos a serem correlacionados: volume de cheia, duração da cheia, e os dois parâmetros da distribuição.

Dispondo de 14 índices climáticos, com área de atuação em diferentes regiões, foi necessário desenvolver uma correlação cruzada com cada um dos parâmetros. Para esta etapa, foi utilizado como referência o mês de agosto para o início das cheias na bacia do rio Madeira. Com os índices postos em intervalos mensais, foi usado em primeiro momento, na correlação entre os índices climáticos e os 4 parâmetros, o valor dos índices no mês 8 para o mesmo ano. Este mesmo processo foi repetido, porém utilizados lags negativos, a fim de identificar os valores máximos absolutos de correlação entre cada par índice-parâmetro, bem como o tempo de atraso no qual ocorriam. Para quantificar a correlação entre os elementos, foi utilizado o coeficiente de Pearson, visto na equação 3.3.

Obtendo os valores máximos do coeficiente de Pearson para cada correlação, foi utilizado de referência como tempo de atraso do índice climático, o maior valor de coeficiente dentre os quatro parâmetros. Ao final, foi selecionado os dois índices climáticos com maior grau de correlação para a construção do modelo.

Portando os melhores índices a se representar os 4 parâmetros, foi feita uma regressão linear individual, adotando para este estudo que as variáveis são independentes entre si, possibilitando esta abordagem. O desenvolvimento do modelo foi utilizado a equação 3.4 e o método de *leave-one-out* para a sua validação. O uso dos critérios de qualidade RMSE e R^2 também foram abordados para avaliar o seu desempenho aos dados previstos. Sendo assim, foi tida a tentativa de identificar uma relação matemática em que seja possível prever os parâmetros que descrevem o hidrograma, utilizando apenas os índices climáticos.

5- RESULTADOS E DISCUSSÕES

Nesta seção serão apresentados os resultados encontrados a partir da metodologia adotada.

5.1- DISTRIBUIÇÕES DE PROBABILIDADE AJUSTADAS AO HIDROGRAMA DE CHEIA ANUAL

Primeiramente, lidando com a necessidade da separação da componente de resposta rápida do hidrograma, foi utilizado o filtro digital desenvolvido por Lynn, adotando um valor de 0.997 para o parâmetro a por análise gráfica, obtendo o resultado da figura 5.1, onde a linha em vermelho no primeiro gráfico representa o escoamento de base calculado, e no segundo a porção do escoamento rápido, que foi utilizada para as próximas etapas.

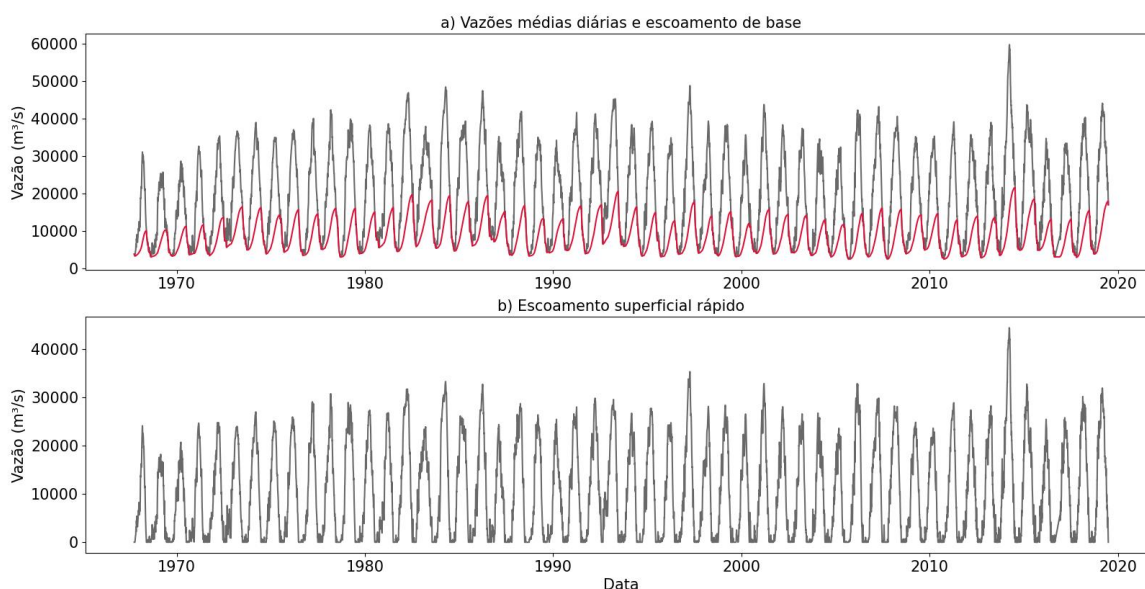


Figura 5.1: Vazões médias diárias e escoamento de base em vermelho (a) e escoamento superficial rápido (b).

Para ser feita a avaliação de forma individual, foi feita a separação em 52 hidrogramas de acordo com sua duração no período de cheia, como indicado na figura 5.2. Foi utilizado como referência para a apresentação dos resultados, os períodos que houve as maiores vazões de pico, entre os anos de 2013 a 2014 e de 1996 a 1997, como destacado em vermelho no gráfico.

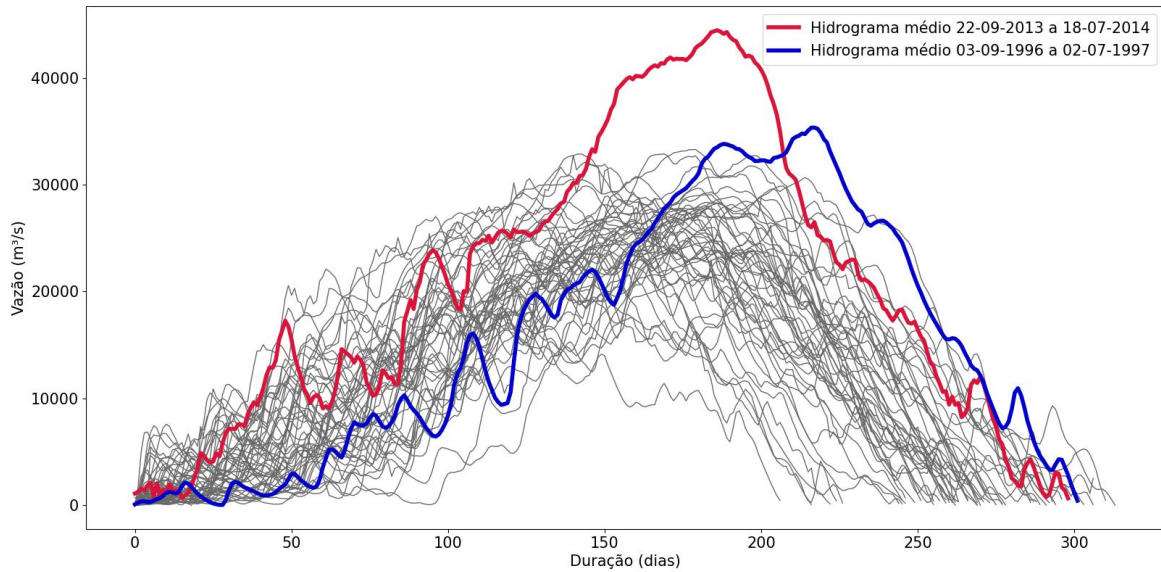


Figura 5.2: Vazões médias diárias agrupadas em hidrogramas anuais de cheia.

Com o intuito de padronizar os dados calculados, para assim seguir com os ajustes das funções de densidade probabilidade, foi necessário tornar as vazões em adimensionais, de forma em que seja possível a comparação entre os elementos, em função do volume e da duração de cada hidrograma. A figura 5.3 mostra o resultado deste processo.

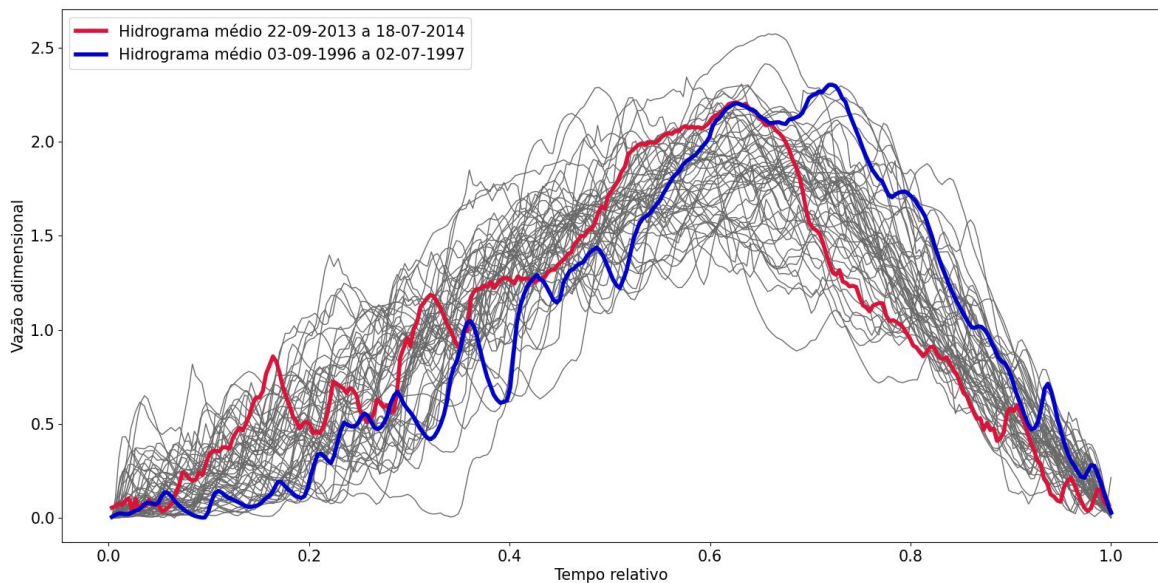


Figura 5.3: Hidrogramas médios adimensionais.

Com as dados de vazões já prontos para os ajustes das distribuições, os hidrogramas foram analisados separadamente para cada distribuição, aplicando as equações das tabelas 3.1 a 3.4, estimando seus parâmetros, minimizando o erro quadrático médio (RMSE). O resultado do ajuste de cada função foi calculado para toda a série, e pode ser observado nas figuras 5.4 a 5.7.

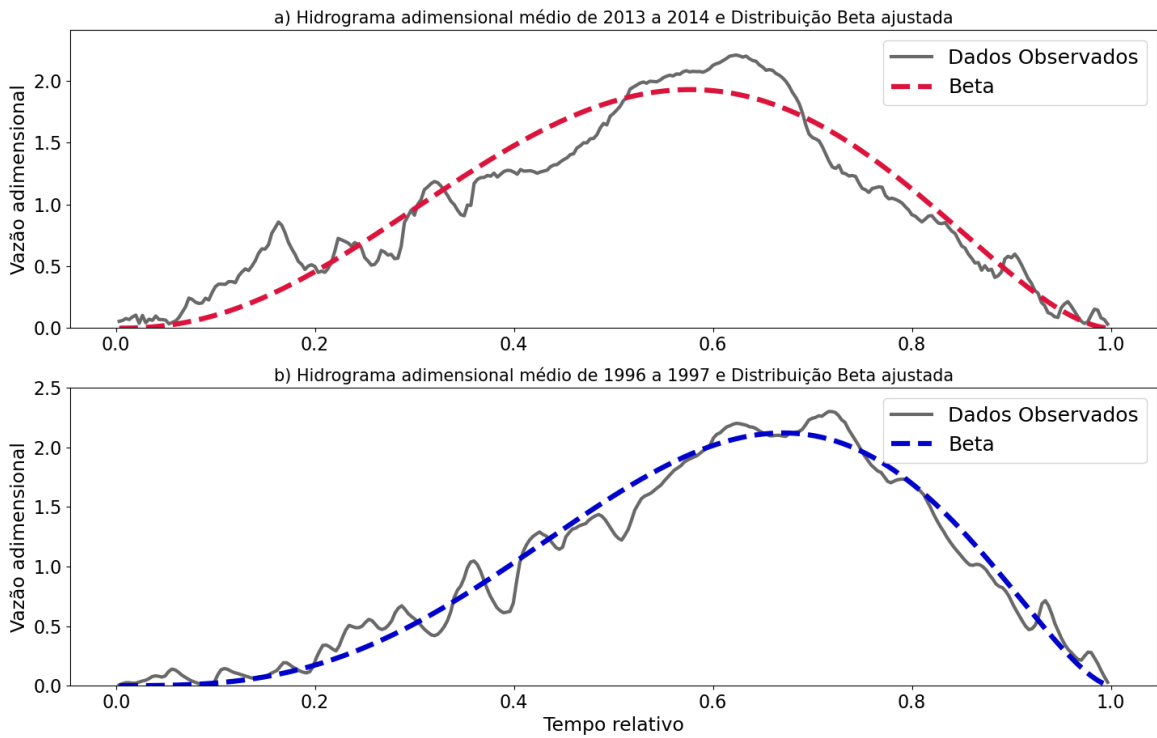


Figura 5.4: Ajuste da distribuição Beta aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b).

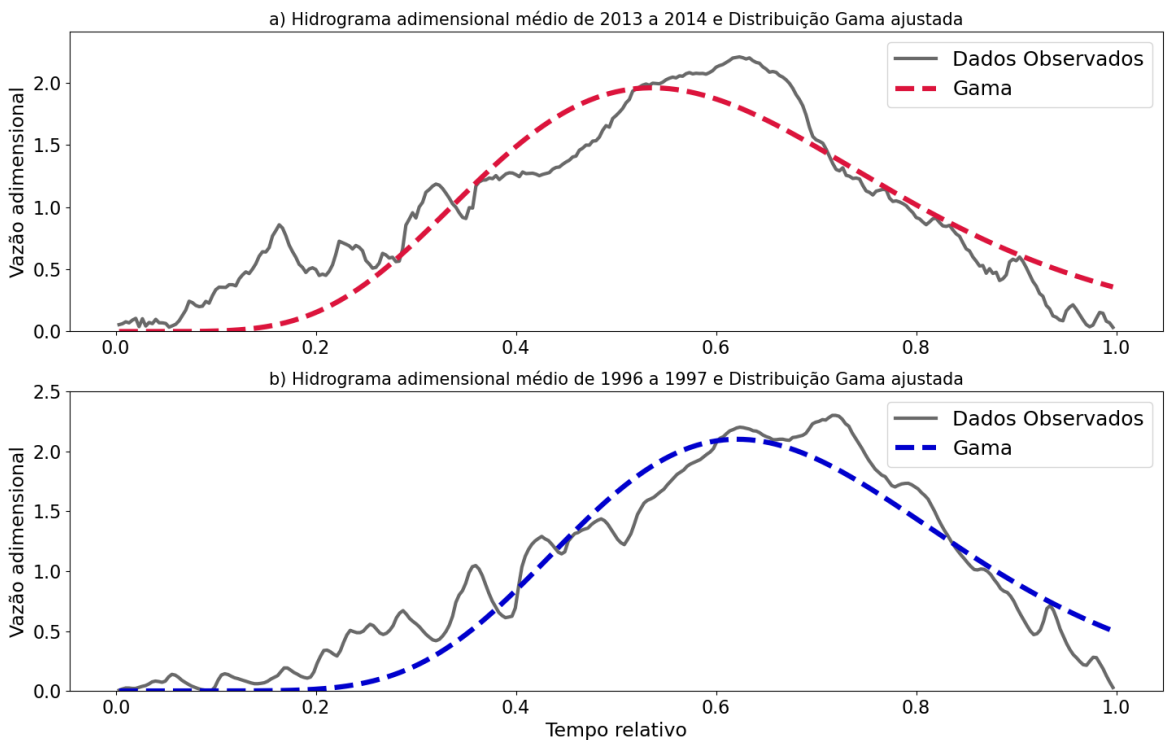


Figura 5.5: Ajuste da distribuição Gama aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b).

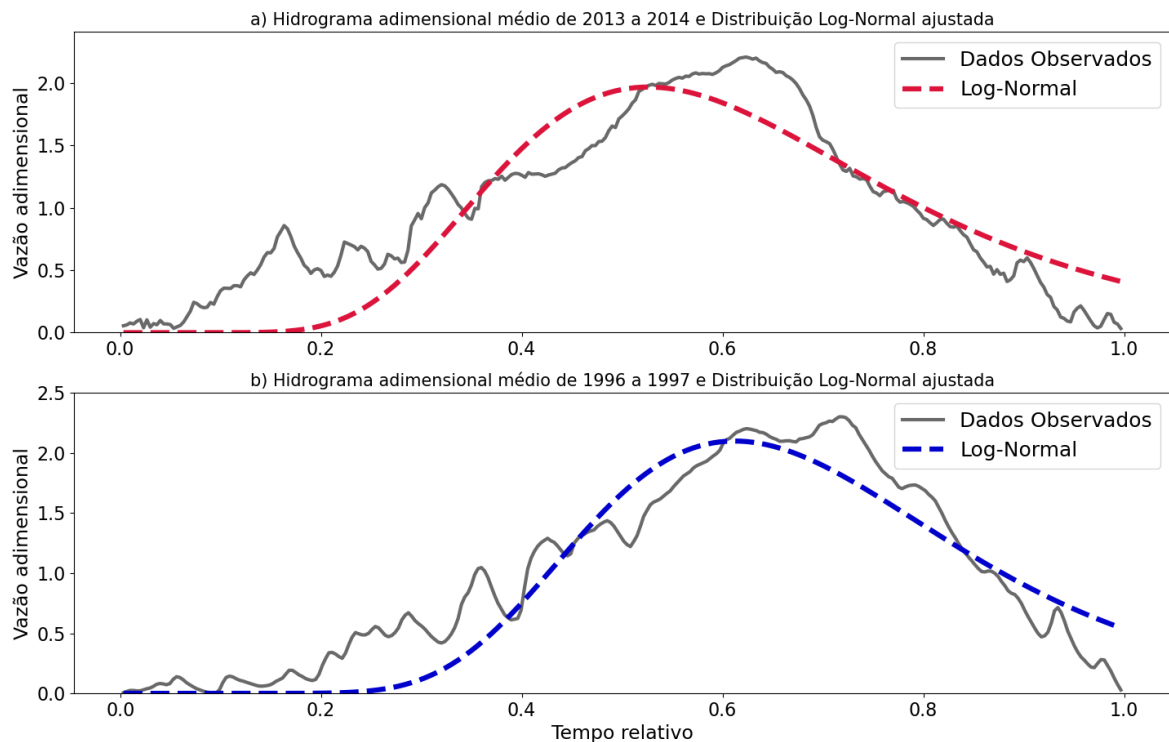


Figura 5.6: Ajuste da distribuição Log-Normal aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b).

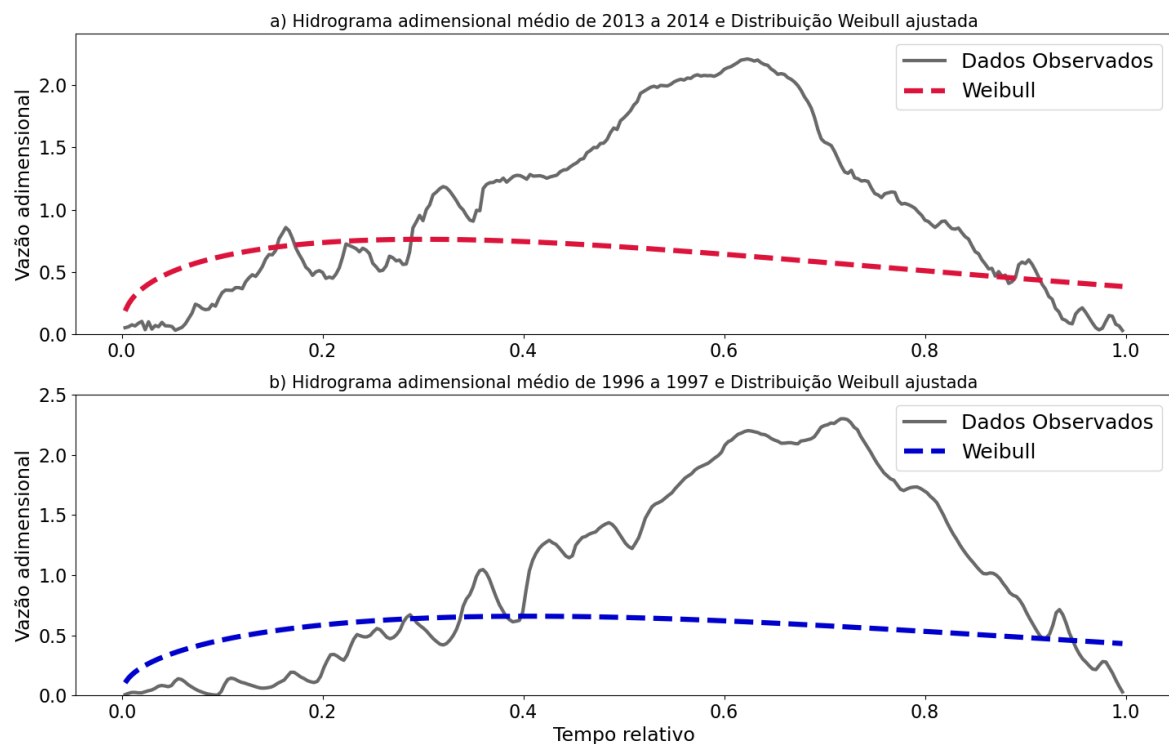


Figura 5.7: Ajuste da distribuição Weibull aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b).

Como forma de avaliar juntamente cada distribuição ajustada, a figura 5.8 mostra os mesmos hidrogramas e funções ajustadas em um mesmo gráfico. Em seguida, a tabela 5.1 e 5.2

mostram os valores dos parâmetros de cada distribuição, para os hidrogramas de cheia dos anos de referência.

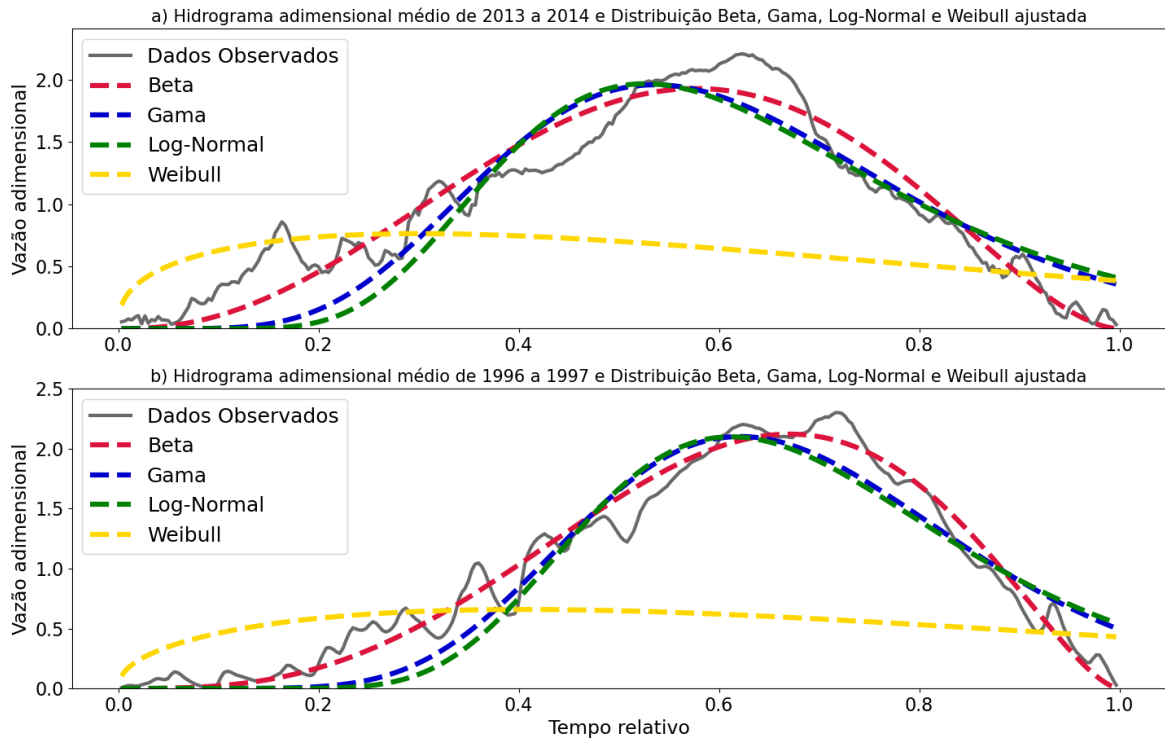


Figura 5.8: Ajuste das distribuições Beta, Gama, Log-Normal e Weibull aos dados observados adimensionais de vazão no ano de 2013 a 2014 (a) e de 1996 a 1997 (b).

Tabela 5.1: Parâmetros estimados das funções Beta, Gama, Log-Normal e Weibull para o hidrograma médio de 2013 a 2014.

Beta		Gama		Log-Normal		Weibull	
a	b	τ	θ	μ	σ	k	λ
3,430	2,781	8,099	0,075	-0,512	0,361	1,369	0,781

Tabela 5.2: Parâmetros estimados das funções Beta, Gama, Log-Normal e Weibull para o hidrograma médio de 1996 a 1997.

Beta		Gama		Log-Normal		Weibull	
a	b	τ	θ	k	λ	μ	σ
4,248	2,597	11,886	0,057	-0,403	0,297	1,446	0,914

Para avaliar a distribuição que melhor se ajustou aos dados observados, foi feita a análise de critério de qualidade do ajuste. As figura 5.9 e 5.10 representam graficamente esta solução, onde primeiramente é calculado o a raiz da soma dos desvio dos pontos da função com as vazões (RMSE), e quanto menor o seu valor, mais próxima esta função foi aproximada dos dados reais. De maneira similar, foi medido o ajuste do modelo estatístico pelo coeficiente de determinação, com avaliação de quanto mais próximo o valor for de 1, mais ajustado está a função.

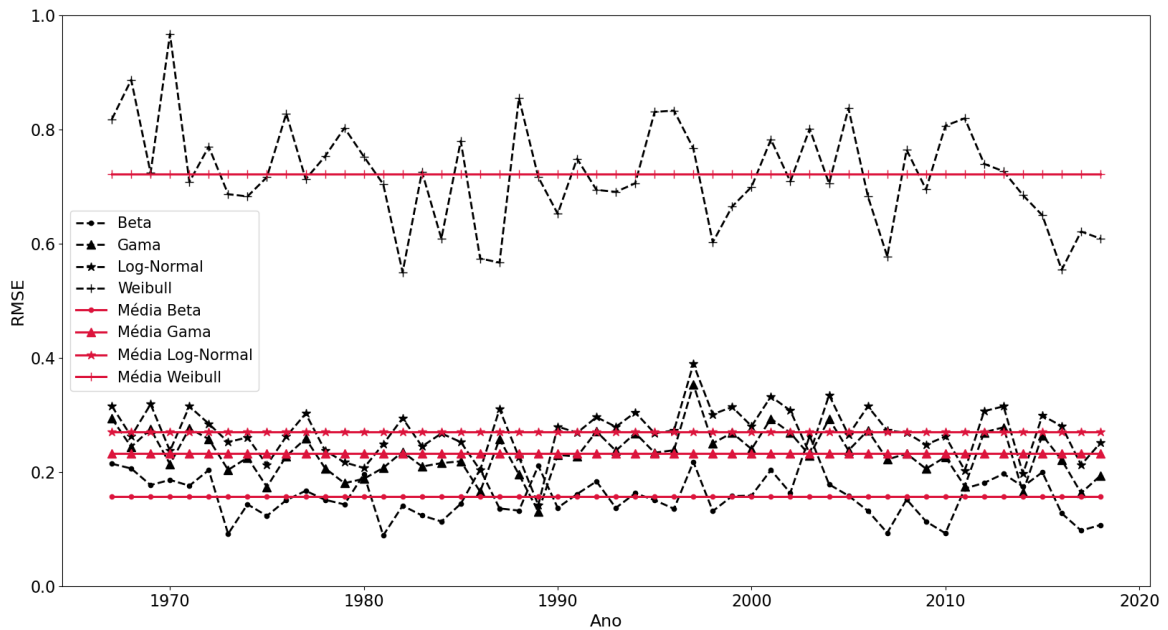


Figura 5.9: RMSE do ajuste das funções de densidade probabilidade.

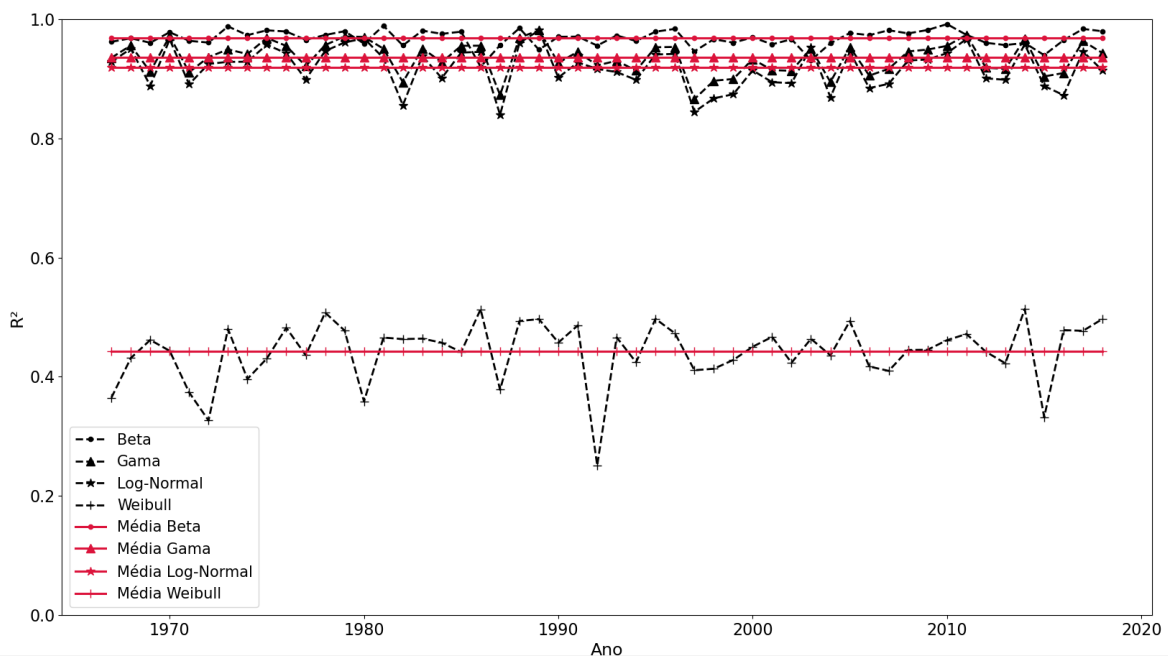


Figura 5.10: R^2 do ajuste das funções de densidade probabilidade.

Feita a análise dos valores de RMSE e de R^2 para o ajuste das distribuições, nota-se que a distribuição Beta foi a que melhor se ajustou ao formato dos hidrogramas. seguindo assim com os valores de seus parâmetros para as etapas de correlação com os índices climáticos.

5.2- CORRELAÇÃO ENTRE OS DADOS CLIMÁTICOS E O MODELO DE PREVISÃO

Seguindo com os valores dos parâmetros dos hidrogramas para as etapas de correlação com os índices climáticos, foi feita a correlação cruzada em função do tempo de atraso dos índices, estando em lag 0, o mês de referência para o início das cheias. A figura 5.11 exemplifica o processo feito, onde o maior valor absoluto no gráfico mostra a melhor correlação entre o índice e o parâmetro. As barras vermelhas indicam as regiões críticas para um teste de hipótese para a correlação, assim valores acima ou abaixo delas são estatisticamente significantes ao nível de significância de 5%.

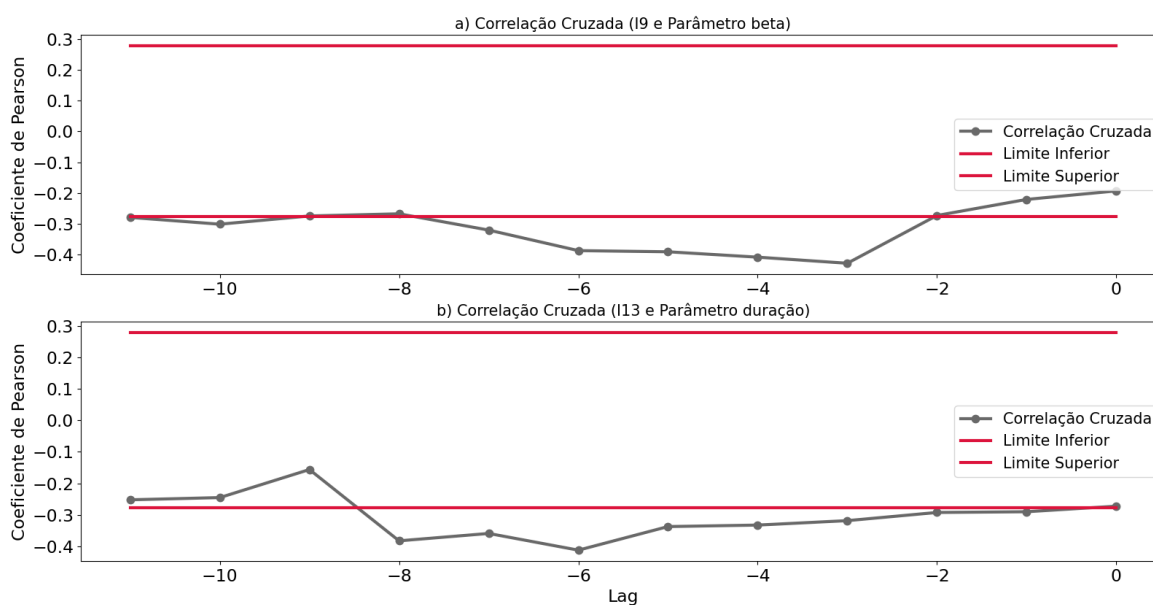


Figura 5.11: Correlação cruzada entre I9 (Niño 3) e Beta (a) e I13 (Anomalias de temperaturas globais) e a Duração (b).

Após este método ser executado, foi adotado o tempo de atraso com maior correlação de cada índice dentre os 4 parâmetros, para ser denominado como o lag necessário a se ter maior influência no hidrograma. A tabela 5.3 mostra o resultado das correlações.

Tabela 5.3: Matriz de valores máximos absolutos de correlação cruzada entre as regiões e os índices

ÍNDICE	LAG	PARÂMETRO			
		<i>a</i>	<i>b</i>	Volume	Duração
I1	-5	0,249	0,378	-0,188	0,111
I2	-5	-0,235	-0,194	-0,248	-0,324
I3	-5	-0,112	0,009	-0,361	-0,202
I4	-4	-0,184	-0,126	0,018	-0,305
I5	-3	-0,247	-0,465	0,146	-0,088
I6 (PDO)	-4	-0,300	-0,390	0,244	-0,316
I7	-3	-0,247	-0,465	0,147	-0,087
I8	-5	0,098	0,174	-0,278	-0,094
I9 (Niño 3)	-3	-0,314	-0,428	0,141	-0,100
I10	-3	-0,278	-0,432	0,128	-0,099
I11	-3	-0,184	-0,345	0,039	-0,118
I12	-11	-0,241	-0,342	0,108	-0,227
I13	-6	-0,348	-0,264	0,142	-0,412
I14	-2	0,178	0,305	-0,132	0,198

Para seguir para próxima etapa, foi tomado como critério a seleção dos 2 índices que melhor se correlacionaram com os parâmetros, tendo como referência física, as teleconexões comentadas na seção 3.3. Sendo estes o I6 (Pacific Decadal Oscillation) e o I9 (Niño 3).

Partindo de uma série de 2 índices climáticos e de 4 parâmetros, foi feita a regressão linear individual de forma a representar matematicamente cada parâmetro. Com uso do método *leave-one-out*, a figura 5.12 apresenta a validação cruzada entre os dados previstos e observados, resultantes do modelo.

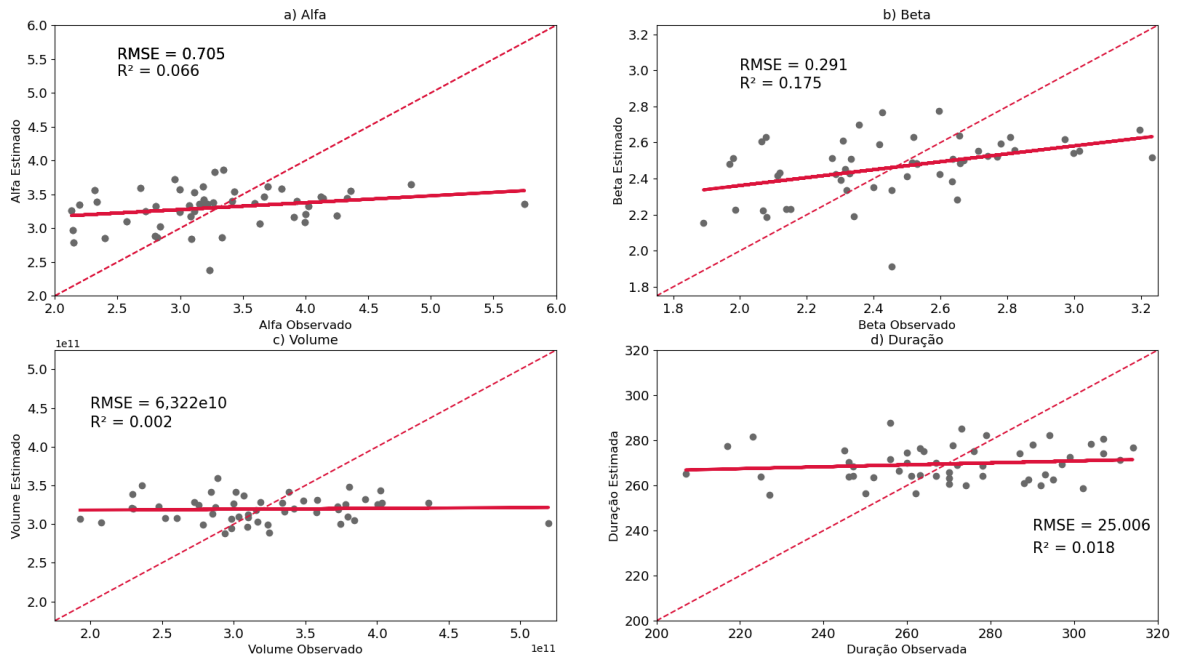


Figura 5.12: Validação cruzada entre os dados previstos e observados de alfa (a), beta (b), volume (c) e duração (d).

A fim de apresentar de outra forma os dados previstos em relação aos observados, a figura 5.13 põe os hidrogramas de referência observados, vista na figura 5.2, com suas respectivas previsões.

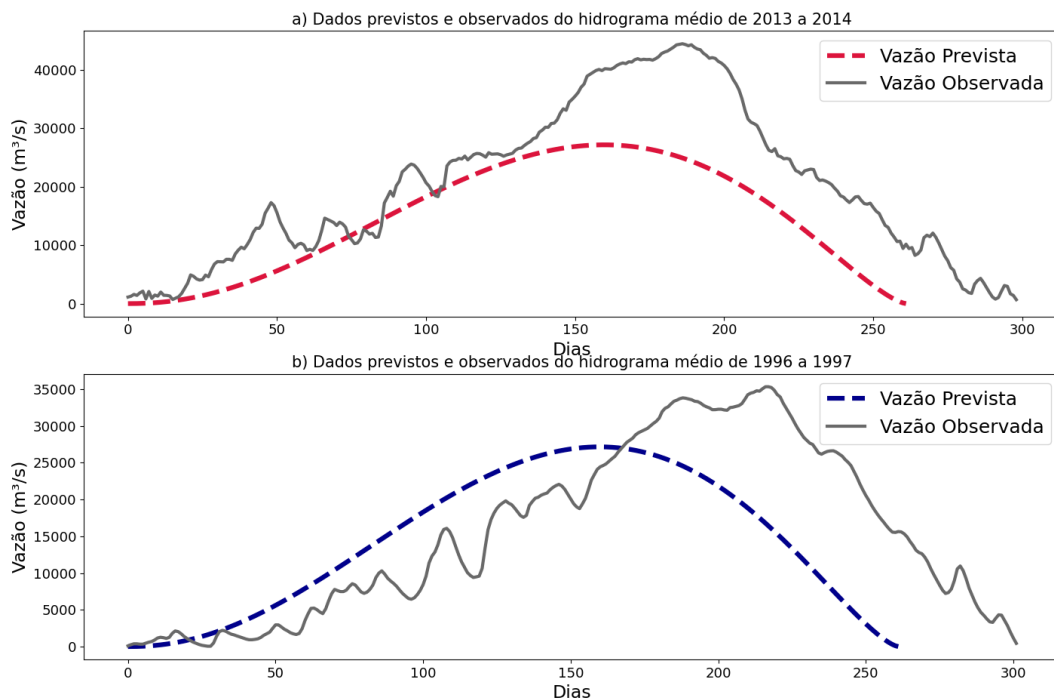


Figura 5.13: Dados observados e previstos do hidrograma de cheia de 2013 a 2014 (a) e de 1996 a 1997 (b).

Nota-se que pouca precisão do modelo ao prever os parâmetros, pode estar relacionada pela forma adotada em criar uma relação entre as variáveis de modo independente. Para testar a correlação entre os parâmetros foi feita uma correlação entre essas variáveis de forma a avaliar sua dependência, vista na figura 5.14. Nela é possível notar a significativa correlação, incentivando assim, para trabalhos futuros, uma nova aplicação no modelo, avaliando alternativas de implementar uma análise multivariada.

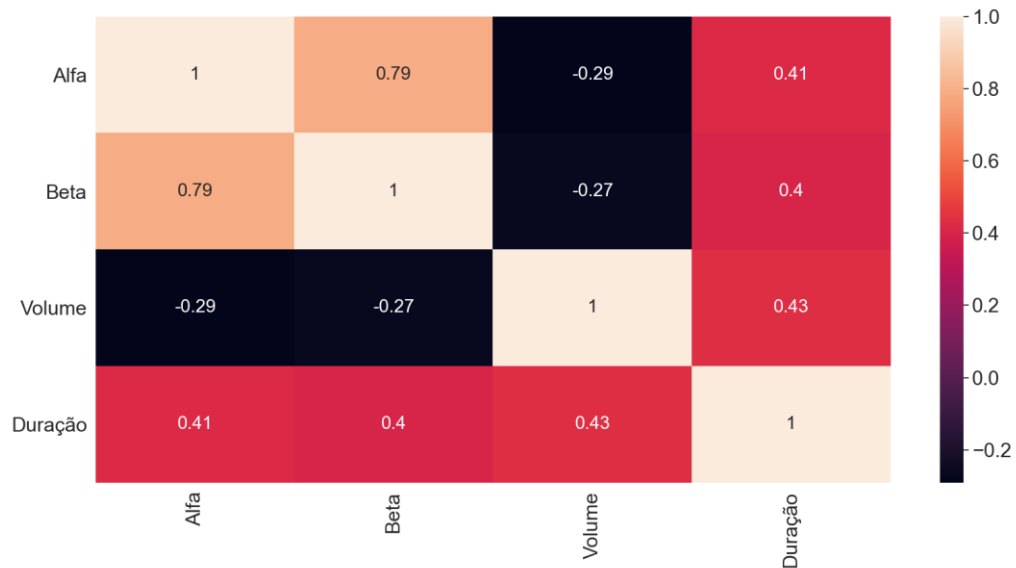


Figura 5.14: Representação gráfica da matriz de correlação dos parâmetros.

6- CONCLUSÕES E RECOMENDAÇÕES

Neste trabalho buscou-se implementar um modelo que fosse capaz de simular e prever uma cheia na região da bacia do rio Madeira, utilizando preditores climáticos que houvesse relação com a hidrologia da área. Para tal, foi feito a descrição deste modelo em quatro parâmetros, sendo necessário a análise dos componentes do hidrograma, em escoamento de base e de fluxo rápido, como também do ajuste do seu formato, por meio de funções de densidade probabilidade. Com isso, os índices climáticos foram utilizados como forma de representação quantitativa dos fenômenos climáticos, de modo em que pudesse correlacionar estas séries com os parâmetros do hidrograma. Na tentativa de identificar a relação matemática entre estes componentes, foi elaborado uma regressão linear individual para cada parâmetro.

Pode-se concluir dos resultados atingidos que houve um entrave na elaboração do modelo, onde os dados previstos da relação matemática foram pouco precisos, visto nos gráficos de validação cruzada. Sugere-se tal erro, na tentativa de avaliar individualmente cada variável, onde é recomendado por Souza Filho e Lall (2003), em que há relações dependentes entre os parâmetros, que se faça uma regressão agrupada para a previsão destas variáveis.

Porém outras informações foram concluídas do trabalho:

- Na aproximação dos hidrogramas de cheia da bacia do rio Madeira, a função Beta de densidade probabilidade se mostrou adequada, por possibilitar um melhor ajuste ao hidrograma médio.
- Os índices climáticos que mostraram melhores correlações, utilizando estudos passados, foram relacionados ao fenômeno ENSO e ao PDO.
- Além da aplicação de um regressão agrupada das variáveis, outra forma de poder tornar o modelo mais preciso, pode estar relacionado com a necessidade de implementar preditores climáticos que sejam mais identificados com região. Sugere-se até no estudo de outros índices, podendo haver uma combinação entre esses, ou também na inclusão dos índices climáticos do Atlântico.
- Recomenda-se também, em trabalhos futuros, a inclusão de dados hidrológicos ao modelo, não privando apenas às séries climáticas. O uso da correlação entre os dados no tempo ou utilizando informações de outras estações fluviométricas, pertencentes a região da bacia do rio Madeira, pode vim a aperfeiçoar o modelo.

REFERÊNCIAS BIBLIOGRÁFICAS

- Benesty, J., Chen, J., Huang, Y., Cohen, I. (2009). *“Noise Reduction in Speech Processing.”* Springer Publishing Company. 229p.
- Bessani, A. S., Souza, E. A., Rocha, E. S., Silva, R. L. S., Ramires, T. V. A. e Amorim, V. S. (2018). *“Fisiografia do rio Madeira: uma análise sobre a formação, estrutura e processos fluviais”*. *Revista do Grupo PET e Acadêmicos de Geografia da Ufac*, 1(1), 118-127.
- Brodie., R. S. e Hostetler. S. (2005). *“A review of techniques for analysing baseflow from stream hydrographs”* Bureau of Rural Sciences, Australia.
- Brunya, M. K., Kumar, S., Jena, J. e Brunya, P. K (2015). *Flood hydrograph with synthetic unit hydrograph routing”* *Water Resour Manage*, 29(15), 5765-5782.
- Brunya, P. K., Berndtsson, R., Ojha, C. S. P. e Mishra, S. K. (2004). *“Parameter estimation of Beta distribution for unit hydrograph derivation”*. *Journal of Hydrology*, 325-332.
- Brunya, P. K., Berndtsson, R., Ojha, C. S. P. e Mishra, S. K. (2007). *“Suitability of Gamma, Chi-square, Weibull, and Beta distributions as synthetic unit hydrographs”*. *Journal of Hydrology*, 28-38.
- Brunya, P. K., Berndtsson, R., Ojha, C. S. P. e Mishra, S. K. (2004). *“Parameter estimation of Beta distribution for unit hydrograph derivation”*. *Journal of Hydrology*, 325-332.
- Cavalcanti, Iracema F. A. et al. (2009). *Tempo e Clima no Brasil*. Oficina de Textos. São Paulo, Brasil.
- Cemaden - Centro Nacional de Monitoramento e Alertas de Desastres Naturais (2018). *“28/03/2018 – Previsão de vazão para bacia do Rio Madeira”*. Disponível: <<http://www.cemaden.gov.br/28032018-previsao-de-vazao-para-bacia-do-rio-madeira/>> Acesso em: outubro de 2019.
- Climate Data Guide. *“Nino SST Indices (Nino 1+2, 3, 3.4, 4; ONI and TNI)”*. Disponível em: <<https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni>>. Acesso em: 17 de novembro de 2019.
- Daza. C. A. (2008). *Análise pluviométrica da bacia transfronteiriça do rio Madeira*. Dissertação de Mestrado, Universidade Federal do Rio de Janeiro, Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia, Rio de Janeiro, RJ, 128p.

- Diniz, L. S. (2008). “*Regionalização de parâmetros de modelo chuva-vazão usando redes neurais*”. Tese de Doutorado. Universidade Federal do Rio Grande do Sul, Porto Alegre -RS, 213p.
- Eckhardt, K. (2008). “*A comparison of baseflow indices, which were calculated with seven different baseflow separation methods*” *Journal of Hydrology*, 168-173.
- Filho, F.A.S e Lall, U. (2003). “*Seasonal to interannual ensemble streamflow forecasts for Ceara, Brazil: applications of a multivariate, semiparametric algorithm*”. *Water Resources Research*, **39**(11).
- Georgakakos, K.P.; Krzysztofowicz, R. (2001). “*Probabilistic and ensemble forecasting.*” *Journal of Hydrology*. 249, p.1.
- Golfari, L.; Caser, R. L.; Moura, V. P. G. (1978) “*Zoneamento ecológico esquemático para reflorestamento no Brasil*”. Belo Horizonte: 66 p.
- Grimm, A. M. (2010). “*Interannual climate variability in South America: impacts on seasonal precipitation, extreme events, and possible effects of climate change*” *Stochastic Environmental Research and Risk Assessment*, **25**(4), 537-554.
- Grimm, A. M., Doyle, V. R. e Barros, V. R. (2000). “*Climate variability in Southern South America associated with El Niño and La Niña events*”. *Journal of Climate*, **13**(1), 35-58.
- Guyot, J. L. (1993). *Hydrogéochimie des fleuves de L'Amazone Bolivienne*. Études et Thèses. Paris, França.
- Hayakawa, E. H., Rossetti, D. F. (2015). “*Late quaternary dynamics in the Madeira river basin, southern Amazonia (Brazil), as revealed by paleomorphological analysis.*” *An. Acad. Bras. Ciênc.*, Rio de Janeiro , v. 87, n. 1, p. 29-49.
- Kingse, C. e Peagle, J. N. (2001). “*The pacific-south american modes and their downstream effects*”. *Journal of Climatology*, **21**(10), 1211-1229.
- Nobre, C. A. D., Obregón G. O. e Marnego, J. A. (2009). “*Características do clima amazônico: aspectos principais*”. *American Monograph Series*, 149-162.
- Labadessa, A. S. e Oliveira C. C. (2013). “*Navegação no rio madeira com auxílio da ferrovia transcontinental no trecho de Vilhena x Porto Velho*”. *Revista Labirinto*, **19**(1), 146-167.
- Leila, L. e Silva, M. E. S. (2012). “*Precipitação e vazão mensal na Amazônia*”. *Revista Geonorte*, **1**(5), 719-728.
- Lyne, V e Hollick, M. (1979). “*Stochastic time-variable rainfall-runoff modelling*”. *Australia National Conference Publication*, 89-93.

- Mântua N, Hare S, Zhang Y, Wallace J, Francis R. A. (1997). “*Pacific interdecadal climate oscillation with impacts on salmon production.*” *B. Am. Meteorol. Soc.* 78; p1069-1079.
- Muniz, L. S. (2013). *Análise dos padrões fluviométricos da bacia do rio Madeira*. Dissertação de Mestrado, Universidade Federal do Amazonas, Departamento de Geografia, Manaus, AM, 144p.
- Naghetini, M. e Pinto, E. J. A. (2007). *Hidrologia Estatística*. CPRM. Belo Horizonte, Brasil.
- Pramanik. N., Panda, R. K. e Sen. D (2010). “*Development of design flood hydrographs using probability density functions*”. *Hydrological Processes*, **24**(4), 415-428.
- Proclira, Clima e Ambiente. Saúde e ambiente. Universidade de Évora, Évora, Portugal. (2007). Disponível: <<http://www.proclira.uevora.pt/modulos/modulo06.pdf>>. Acesso em novembro de 2019.
- Reboita, M. S. e Santos, I. A. (2015). “*Influência de alguns padrões de teleconexão na precipitação no Norte e Nordeste do Brasil*”. *Revista Brasileira de Climatologia*, **15**(1), 28-48.
- Reef Resilience Network. “*El Niño Southern Oscillation*”. Disponível em: <<https://reefresilience.org/climate-and-ocean-change/el-nino-southern-oscillation>> Acesso em: 17 de novembro de 2019.
- SANTOS, M. S. (2019). “*Estudo do comportamento não estacionário das vazões no rio Paraguai: detecção, atribuição e previsão de cheias.*” Tese de Doutorado em Tecnologia Ambiental e Recursos Hídricos, Publicação PTARH.TD-024/2019, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 199p.
- Tucci, C. E. M. (2001). *Hidrologia: ciência e aplicação*. Editora da Universidade, Rio Grande do Sul, Brasil, 943p.
- Tucci, C.E.M., Clarke, R.T., Collischonn W.; Dias, P.L.S.; Sampaio, G.O. (2003). “*Long term flow forecast based on climate and hydrological modeling: Uruguay river basin.*” *Water Resources Research*. 39(7), p1-11.
- Silva, B. C., Tucci, C. E. M., Collischonn, W. (2006). “*Previsão de Vazão com Modelos Hidroclimáticos*”. *Revista Brasileira de Recursos Hídricos*, **11**(3), p15-29.
- Zeileis, A., Kleiber, C., Walter, K., Hornik, K., 2003. “*Testing and dating of structural changes in practice.*” *Comput. Stat. Data Anal.* 44, 109–123.

APÊNDICE

APÊNDICE A – CÓDIGO-FONTE EM LINGUAGEM PYTHON UTILIZADO NO ESTUDO

```
SEPARAÇÃO DO ESCOAMENTO DE BASE
def BaseFlow(Q_Table, Beta):
    n = len(Q_Table)
    Q = Q_Table['Vazoes']
    q = [np.nan]*n
    q[0] = 0

    for t in range(1, n):
        q[t] = Beta * q[t-1] + (1 + Beta) / 2 * (Q[t] - Q[t-1])
        if(Q[t] - q[t] > Q[t]):
            q[t] = 0
        else:
            q[t] = q[t]
    Qbase = Q - q
    BaseSup = pd.DataFrame({'Data': Q_Table['Data'], 'Qsup':Q,
'Qbase': Qbase, 'Qrunoff':q})
    BaseSup = BaseSup.merge(Q_Table, how = 'right')

    return(BaseSup)

QBase = BaseFlow(vazoes, 0.997)
```

```
CALCULO PARÂMETROS: DURAÇÃO E VOLUME
DURAÇÃO = []
for i in range(0, len(Anos)):
    D.append(len(SeriesRunoff[i]))

VOLUME = []
for i in range(0, len(SeriesRunoff)):
    y = SeriesRunoff[i].reset_index(drop = True)
    Vp = y[0] * (24 * 3600) / 2
    for j in range(0, len(y) - 2):
        Vp = Vp + ((y[j] + y[j+1]) / 2 * (24*3600))
    V.append(Vp)
```

```
FUNÇÃO BETA
def HydroBeta(x, alfa, beta):
    return((1/(math.gamma(alfa) * math.gamma(beta) /
math.gamma(alfa+beta))) * (x**(alfa-1)) * ((1-x)**(beta-1)))
```

```
FUNÇÃO GAMA
def HydroGamma(x, tau, omega):
    return(x ** (tau-1) * (math.exp(-x/omega)) / (omega**tau *
math.gamma(tau)))
```

```
FUNÇÃO LOG-NORMAL
def HydroLogNorm(x, mi, sigma):
    return(1 / (x * sigma * math.sqrt(2*math.pi)) * math.exp((-
(math.log(x) - mi) ** 2) / (2 * sigma ** 2)))
```


FUNÇÃO WEIBULL

```
def HydroWeibull(x, kapa, lamb):
    return((kapa / lamb) * (x / kapa) ** (kapa - 1) * math.exp(- (x
/ lamb) ** kapa))
```

ÍNDICES CLIMÁTICOS: NINO 3 E PDO

```
i6 = pd.read_excel(r'D:\UnB\Projeto Final\Dados\indices\I6.xlsx',
delimiter = ';', decimal = ',')
i6_01 = i6.iloc[:,1:].columns
i6a = pd.melt(i6, id_vars = ['Ano'], value_vars = i6_01, var_name =
'Mês', value_name = 'Indice')
i6a['Ano'] = i6a['Ano'].astype(str)
i6a['Mês'] = i6a['Mês'].astype(str)
i6a['Data'] = i6a[['Mês', 'Ano']].agg('-'.join, axis=1)
i6a['Data'] = pd.to_datetime(i6a['Data'], format = '%m-%Y', errors
= 'coerce')
i6 = pd.DataFrame()
i6['Data'] = i6a['Data']
i6['Indice'] = i6a['Indice']
i6 = i6.sort_values(by = 'Data').reset_index(drop = True)

i9 = pd.read_excel(r'D:\UnB\Projeto Final\Dados\indices\I9.xlsx',
delimiter = ';', decimal = ',')
i9_01 = i9.iloc[:,1:].columns
i9a = pd.melt(i9, id_vars = ['Ano'], value_vars = i9_01, var_name =
'Mês', value_name = 'Indice')
i9a['Ano'] = i9a['Ano'].astype(str)
i9a['Mês'] = i9a['Mês'].astype(str)
i9a['Data'] = i9a[['Mês', 'Ano']].agg('-'.join, axis=1)
i9a['Data'] = pd.to_datetime(i9a['Data'], format = '%m-%Y', errors
= 'coerce')
i9 = pd.DataFrame()
i9['Data'] = i9a['Data']
i9['Indice'] = i9a['Indice']
i9 = i9.sort values(by = 'Data').reset_index(drop = True)
```

MODELO: REGRESSÃO LINEAR

```
y_pred = list()
for train_ix, test_ix in cv.split(X):
    # split data
    X_train, X_test = X.iloc[train_ix, :], X.iloc[test_ix, :]
    y_train, y_test = y[train_ix], y[test_ix]
    # fit model
    model = LinearRegression(fit_intercept = False)
    model.fit(X_train, y_train)
    # evaluate model
    yhat = model.predict(X_test)
    # store
    y_pred_Alfa.append(yhat[0])

y_coef = [model.coef [0], model.coef [1]]
```