

LEONARDO CAMPOS DE ASSIS

**SIMULAÇÃO DE PROCESSOS HIDROLÓGICOS NA BACIA
HIDROGRÁFICA DO RIO PIRANGA**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Civil, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS-BRASIL
2012

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

A848s
2012

Assis, Leonardo Campos de, 1980-
Simulação de processos hidrológicos na bacia hidrológica do
rio Piranga / Leonardo Campos de Assis. – Viçosa, MG,
2012.
xii, 111f. : il. ; (algumas color.) ; 29cm.

Inclui apêndices.

Orientador: Maria Lúcia Calijuri

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Hidrologia - Métodos de simulação. 2. Teoria bayesiana
de decisão estatística. 3. Processamento de imagens -
Técnicas digitais. 4. Sensoriamento remoto. I. Universidade
Federal de Viçosa. Departamento de Engenharia Civil.
Programa de Pós-Graduação em Engenharia Civil. II. Título.

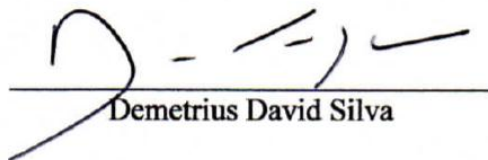
CDD 22. ed. 551.483

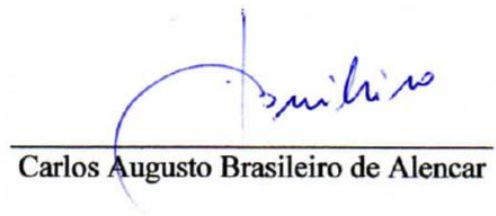
LEONARDO CAMPOS DE ASSIS

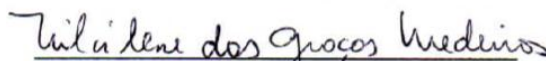
**SIMULAÇÃO DE PROCESSOS HIDROLÓGICOS NA BACIA
HIDROGRÁFICA DO RIO PIRANGA**

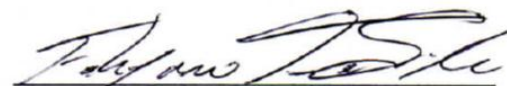
Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Civil, para obtenção do título de *Doctor Scientiae*.

APROVADA: 6 de agosto de 2012.


Demetrius David Silva


Carlos Augusto Brasileiro de Alencar


Nilcilene das Graças Medeiros
(Coorientadora)


Fabyano Fonseca e Silva
(Coorientador)


Maria Lúcia Calijuri
(Orientadora)

DEDICATÓRIA

*À memória de meu saudoso pai Jovem Campos de Assis,
dedico.*

“Quem alcança seu ideal, vai além dele.”

Friedrich W. Nietzsche.

AGRADECIMENTOS

À FAPEMIG pelo apoio financeiro ao projeto de pesquisa que suportou este trabalho, e pela concessão de uma bolsa de estudos, que me possibilitou dedicação integral e exclusiva à pesquisa.

À minha orientadora e cara amiga de longa data Profa. Lúcia Calijuri, pelo crédito, paciência, benevolência, incentivo constante, apoio incondicional e orientação na condução deste.

Aos professores Carlos Vieira, Nilcilene Medeiros, Fabyano Fonseca, Demetrius Silva, Sebastião Martins, Eduardo Marques e José Marinaldo Gleriani pelas importantes contribuições no exame de qualificação do doutorado e, ou co-orientação.

Aos colegas e amigos do programa de Pós-Graduação em Informações Espaciais, especialmente a Antônio Prata, Leila, Wellington, Marquinhos, Afonso, Inês e Graziela, também a todos do Núcleo de Pesquisas Ambientais Avançadas, especialmente a Marcos Dornelas, Madson, Ronan, Isabela, Everton, Aníbal, Luna, Mateus, Ivan, Rodrigo, Eduardo e Paula.

Aos meus familiares e amigos que sempre me apoiaram, especialmente meu irmão Rodrigo, aos primos Donizete, Nenzinha e Sônia, Raquel e Sueli, tias Nenzica e Aparecida, tios Walter e Juarez, também aos novos irmãos Frederico, Viviane, Renato, Simone e Larissa, aos meus sogros Euphranor e Elisabeth Martins, também ao Luís e Marlene e à cara Vó Urana.

À minha querida mãe Carmen e ao meu saudoso pai Jovem (*in memoriam*), pelas orações, incentivo, apoio e incondicional amor que sempre me ajudaram a enfrentar as adversidades do caminho.

Por fim à minha querida esposa Carminha. Meu sincero muito obrigado! Seu amor, paciência, apoio, incentivo, motivação, companheirismo e carinho foram fundamentais para conclusão desta importante etapa em minha vida.

SUMÁRIO

LISTA DE FIGURAS	ii
LISTA DE QUADROS	iv
LISTA DE TABELAS	v
RESUMO	vi
ABSTRACT	vii
I. INTRODUÇÃO GERAL	1
1 Objetivos	1
2 Hipótese de pesquisa	2
3 Justificativa	2
4 Caracterização da área de estudo	3
5 Problema	4
6 Organização do trabalho	7
7 Referências	7
II. Estudo de precipitações pluviiais máximas por meio da inferência Bayesiana: uma abordagem para eventos extremos	8
1 Introdução	8
1.1 Problema	9
1.2 Objetivo	9
1.3 Hipótese	9
2 Revisão de literatura	10
2.1 Distribuição de Valores Extremos	10
2.2 Abordagem Bayesiana	12
2.3 Verificação da convergência	14
2.4 Análise de Agrupamento Aglomerado Hierárquico	15
3 Método	16
3.1 Organização dos dados pluviométricos	17
3.2 Quantil da Distribuição GEV	18
3.3 Definição do Modelo Hierárquico Bayesiano	19
3.4 Análise de Agrupamento Baseado em Modelos	22
4 Resultados e discussões	22
4.1 Inferência multiparamétrica	23
4.2 Análise de Agrupamento Baseada em Modelos	29

4.3	Modelos de Representatividade Regional	32
4.4	Validação dos Modelos	34
5	Conclusões e Recomendações	38
6	Referências	39
III.	Classificação de imagens do sensoriamento remoto a partir de metodologias de representação da paisagem.	42
1	Introdução	42
1.1	Objetivo	43
1.2	Questão de pesquisa	44
2	Revisão de literatura	44
2.1	Regressão Logística	44
2.2	Tomada de Decisão por Agregação Multicritérios	47
2.3	Avaliação pela estatística ROC	50
3	Métodos	51
3.1	Base de dados acessória	51
3.1.1	Modelo Digital de Elevação	51
3.1.2	Imagem de satélite	52
3.2	Definição das classes informacionais	54
3.3	Amostragem das classes informacionais	55
3.4	Imagens de representação da paisagem	56
3.5	Classificação híbrida	58
4	Resultados e discussões	58
4.1	Definição das informações contextuais	58
4.2	IFVP obtidas por RLB	59
4.3	IFVP determinadas por MCE	61
4.4	Classificação híbrida de imagem via IFVP	65
5	Conclusões	69
6	Referências	70
IV.	Simulação de processos hidrológicos de evento extremo na bacia hidrográfica do rio Piranga.	73
1	Introdução	73
2	Objetivo	74
3	Revisão de literatura	74
3.1	Sistema de Simulação Hidrológica	74
3.2	Modelo Digital de Elevação Condicionado à Hidrografia	75
4	Métodos	77

4.1	Dados de entrada: hidrometeorológicos e espaciais	78
4.1.1	Dados hidrometeorológicos	78
4.1.2	Modelo Digital de Elevação Hidrograficamente Condicionado	79
4.1.3	Imagem de Uso e Cobertura do Solo	80
4.1.4	Mapa de Tipos de Solos	81
4.2	Sistemas de Informação Geográfica e de Simulação Hidrológica	83
4.3	Seleção de métodos e modelos hidrológicos	84
4.3.1	Precipitação pluvial	85
4.3.2	Método de determinação do escoamento superficial	85
4.3.3	Método de propagação de cheias em canais	87
4.3.4	Método de transformação chuva-deflúvio	88
4.3.5	Calibração dos modelos	89
4.3.6	Avaliação e validação do desempenho do modelo	90
5	Resultados e discussões	91
5.1	Calibração de modelos	92
5.2	Validação do modelo	95
5.3	Limitações do estudo	96
6	Conclusões e recomendações	97
7	Referências	98
V.	CONCLUSÕES GERAIS	101
	APÊNDICE II	103
	Apêndice (II) A – Série anual de eventos máximos diários de precipitação para cada estação pluviométrica em análise	103
	Apêndice (II) B – Teste diagnóstico de convergência de BGR aplicado ao par de cadeias paralelas dos modelos independentes.	106
	Apêndice (II) C – Teste diagnóstico de convergência de RL aplicado à cadeia unificada dos modelos independentes.	107
	Apêndice (II) D – Altura da precipitação máxima diária e seu respectivo período de retorno calculado por Análise de Frequência dos registros da série histórica das estações pluviométricas em análise.	108
	APÊNDICE III	109
	Apêndice (III) A – Parâmetros de avaliação da qualidade e seleção de modelos.	109
	Apêndice (III) B – Matriz de contingência resultante da avaliação da classificação híbrida via IFVP obtida por MCE	111

LISTA DE FIGURAS

- Figura (I) 1 – Localização das bacias hidrográficas dos Rios Piranga (polígono verde) e Doce (polígono azul), bem como sua posição no Estado de Minas Gerais no Território Brasileiro. 4
- Figura (I) 2 – Registros fotográficos de inundação em algumas cidades localizadas na BHRP e margeadas por esse rio. 6
- Figura (II) 1 – Fluxograma sequencial dos processos utilizados. 16
- Figura (II) 2 – Representação gráfica do modelo Bayesiano Hierárquico para estimar os parâmetros da distribuição GEV, aplicada a estudos de precipitações pluviiais máximas anuais. 20
- Figura (II) 3 – Inferência do parâmetro forma para cada estação com seus limites superior (Ls), inferior (Li) e média a posteriori (MP). 25
- Figura (II) 4 – Inferência do parâmetro escala para cada estação com seus limites superior (Ls), inferior (Li) e média a posteriori (MP). 26
- Figura (II) 5 – Inferência do parâmetro locação para cada estação com seus limites superior (Ls), inferior (Li) e média a posteriori (MP). 27
- Figura (II) 6 – Comparação de alturas máximas precipitadas diárias entre os valores preditos (terminação $\hat{\theta}$) pelos modelos ajustados e os registros observados (terminação θ_{obs}) da série histórica para as estações e1, e2, e3, e4, e5 e e6, com períodos de retorno de até 70 anos. 28
- Figura (II) 7 – Espaço paramétrico de θ resultante do ajuste da distribuição GEV para cada estação pluviométrica onde as dimensões locação (x) escala (y) e forma (z) associadas respectivamente aos parâmetros μ , σ e η . 29
- Figura (II) 8 – Dendrograma resultante da análise de agrupamento hierárquico dos parâmetros das estações pluviométricas com os p-valores calculados para cada agrupamento; em evidência pelos retângulos envoltórios grupos com nível de significância de 0.05. 30
- Figura (II) 9 – Áreas de representatividade geométrica dos grupos de estações G1 (em azul) e G2 (em verde), obtidas pela análise agrupamento de modelos. 32
- Figura (II) 10 – Inferência preditiva das alturas das precipitações pluviiais máximas diárias e seus respectivos períodos de retorno determinada a partir dos valores da média a posteriori, limite inferior e superior para o modelo G1. 36
- Figura (II) 11 – Inferência preditiva das alturas das precipitações pluviiais máximas diárias e seus respectivos períodos de retorno obtida a partir dos valores da média a posteriori, limite inferior e superior para o modelo G1 37
- Figura (III) 1 – Fotografias dos tipos de uso e ocupação do solo mais comuns encontrados na BHRP, os quais sejam: mata secundária (A); café (B); eucalipto (C); pastagem (D); solo exposto (E); e área agrícola (F). 55
- Figura (III) 2 – Espaço estratégico de decisão. 63

Figura (III) 3 – Imagem de uso e cobertura do solo de 2010 da bacia hidrográfica do rio Piranga.	67
Figura (IV) 1 – Principais processos metodológicos utilizados nesta pesquisa.	78
Figura (IV) 2 – Modelo digital de elevação condicionado à hidrografia, delimitação das áreas de drenagem de cada estação fluviométrica bem como sua identificação, estações pluviométricas, principais rios e seus respectivos canais evidenciados.	80
Figura (IV) 3 – Imagem de uso e cobertura do solo elaborada por processo de classificação de imagem do satélite Landsat 5 de agosto de 2010 da bacia hidrográfica do rio Piranga.	81
Figura (IV) 4 – Principais grupos tipológicos de solos ocorrentes na bacia hidrográfica do rio Piranga, delimitação das áreas de drenagem de cada estação fluviométrica em análise bem como sua identificação, estações pluviométricas, principais rios e seus respectivos canais evidenciados.	83
Figura (IV) 5 – Modelo de representação topológica da bacia do rio Piranga no HEC-HMS	92
Figura (IV) 6 – Resultado de simulação hidrológica da etapa de calibração de parâmetros dos modelos. A linha azul refere-se os valores de vazão observados enquanto a linha vermelha é referente aos simulados, barras em cor verde representam alturas precipitadas diárias na bacia durante o período do evento.	94
Figura (IV) 7 – Resultado de simulação hidrológica da etapa de validação dos modelos. A linha azul refere-se os valores de vazão observados, a linha vermelha é relativa aos valores de vazão simulados, e as barras em cor verde representam alturas precipitadas diárias na bacia.	96

LISTA DE QUADROS

Quadro (II) 1 – Estações pluviométricas selecionadas para análise.	18
Quadro (II) 2 – Relação de algoritmos utilizados na determinação dos parâmetros do modelo hierárquico Bayesiano pelo software OpenBUGS.	23
Quadro (III) 1 – Sumário de métodos e parâmetros utilizados no processo de ortorretificação.	53
Quadro (III) 2 – Vantagens e desvantagens das Imagens Fundamentadas em Valores de Probabilidade (IFVP).	57

LISTA DE TABELAS

Tabela (II) 1 – Teste diagnóstico de convergência de BGR aplicado às cadeias dos modelos dos grupos de estações, G1 e G2.	33
Tabela (II) 2 – Teste diagnóstico de convergência de RL aplicado à cadeia unificada dos modelos dos grupos de estações, G1 e G2.	33
Tabela (II) 3 – Valores das médias a posteriori (MP), limite inferior (Li) e limite superior (Ls) do intervalo HPD ao nível de significância de 0.05 dos parâmetros forma (η), locação (μ) e escala (σ) dos modelos ajustados para os grupos de estações, G1 e G2.	34
Tabela (III) 1 – Número de pixels utilizados nas etapas de treinamento e teste dos métodos para cada classe informacional.	56
Tabela (III) 2 – Parâmetros de avaliação de qualidade para seleção de modelos.	60
Tabela (III) 3 – Resultados OWA para o consórcio de classes “Vegetação de Porte Alto”.	63
Tabela (III) 4 – Resultados OWA para o consórcio de classes “Vegetação de Porte Baixo”.	64
Tabela (III) 5 – Resultados OWA para a classe superfície sombreada.	64
Tabela (III) 6 – Resultados OWA para a classe superfície aquosa.	65
Tabela (III) 7 – Resultados OWA para a classe Solo Exposto.	65
Tabela (III) 8 – Quantificação do uso e ocupação do solo na BHRP.	68
Tabela (IV) 1 – Parâmetros do Método de Propagação de cheias em canais Muskingum-Cunge	87
Tabela (IV) 2 – Identificação das áreas de drenagem que compõem a bacia do rio Piranga e valores iniciais e otimizados dos parâmetros.	93

RESUMO

ASSIS, Leonardo Campos de, D.Sc., Universidade Federal de Viçosa, agosto de 2012. **Simulação de processos hidrológicos na bacia hidrográfica do rio Piranga.** Orientadora: Maria Lúcia Calijuri. Coorientadores: José Marinaldo Gleriani, Fabyano Fonseca e Silva e Nilcilene das Graças Medeiros.

A região de nascente da bacia hidrográfica do rio Doce, enquanto este ainda é denominado rio Piranga, é suscetível a inundações recorrentes, quase sempre provocadas por grandes volumes de precipitação pluvial. Além do risco de perda de vidas, as inundações causam problemas socioeconômicos e ambientais para as comunidades locais. Diante do exposto, o intuito desta pesquisa foi apresentar métodos para estudo dos processos hidrológicos relacionados à ocorrência de inundações na região da Bacia Hidrográfica do rio Piranga (BHRP). Os métodos são referentes à inferência Bayesiana, à hidrologia e às ciências cartográficas e tiveram o propósito de subsidiar a obtenção dos seguintes produtos distintos: (i) um modelo de chuvas de representatividade regional para predição de eventos extremos de precipitação, notadamente alturas diárias máximas, obtido por análise de agrupamento baseado em modelos através de abordagem Bayesiana de modelos hierárquicos; (ii) uma imagem de uso e cobertura do solo referente a agosto de 2010, obtida por técnicas de classificação de imagens com suporte de modelos digitais para representação da paisagem como subsídios para identificação das classes temáticas; e (iii) um modelo de chuva-deflúvio para representação de cenário de inundação obtido por simulação hidrológica que permitiu reproduzir a vazão de enchente relativa a uma chuva histórica de longa duração, com boa acurácia. Conclui-se que a abordagem proposta oferece a possibilidade de compreensão dos processos hidrológicos mais relevantes no contexto de uma inundação, permite conhecer melhor o comportamento de resposta da bacia frente à ocorrência desses eventos e, assim, subsidiar entidades e organismos gestores dos recursos hídricos da região a avaliar medidas atenuantes dos efeitos dessas inundações.

ABSTRACT

ASSIS, Leonardo Campos de, D.Sc., Universidade Federal de Viçosa, August, 2012. **Hydrological processes simulation in the river Piranga watershed.** Adviser: Maria Lúcia Calijuri. Co-Advisers: José Marinaldo Gleriani, Fabyano Fonseca e Silva and Nilcilene das Graças Medeiros.

The highlands of the Doce river basin, while it still is called Piranga river is susceptible to recurring floods, often caused by expressive rainfall heights. Besides the risk of deaths, the floods cause severe socioeconomic and environmental problems for the local communities. Because of this, the purpose of this research was to present and apply methods to support the hydrological processes analysis related to flooding in the Piranga river watershed region. The methods are related to Bayesian inference, hydrology and cartographic sciences, purposing to support the generation of the following distinct products: (i) a rainfall model of regional representativeness applied to the prediction of extreme precipitation events, especially the maximum daily heights, obtained by model-based cluster analysis through Bayesian approach of hierarchical models; (ii) an image of the land use / land cover (LULC) from august, 2010, produced by image classification techniques supported by digital landscape models used to identify thematic categories; and (iii) a rainfall-runoff model obtained through hydrological simulation to represent a flooding scenario that has rebuilt with good accuracy, the observed discharge associated to a historic long term rainstorm. It was concluded that the proposed approach offers the possibility of knowing the most relevant hydrological process related to the flooding context, allows better understanding about the basin response behavior in the case of such events and so, supports the local organizations of water resources management on evaluate actions applied to attenuation of floods.

I. INTRODUÇÃO GERAL

Devido à gravidade dos transtornos ocasionados pelas enchentes, sobretudo em áreas urbanas, é fundamental conhecer com mais detalhe as principais áreas afetadas bem como quantificar a intensidade dos eventos hidrológicos extremos causadores das mesmas. Entretanto, observa-se severa escassez de dados espaciais e hidrometeorológicos (principalmente fluviométricos), comprometendo sobremaneira a análise hidrológica. Para reduzir os efeitos da escassez de informações, a simulação pode e deve ser empregada como um instrumento alternativo de análise, pois permite configurar cenários, tanto para verificar a situação atual quanto para avaliar alterações no ambiente. Dessa maneira, possibilita identificar as mais prováveis origens de problemas e verificar diferentes respostas das bacias frente a proposições de intervenções. A análise integrada entre uso e cobertura do solo com fenômenos hidrológicos permite considerar com maior eficiência os detalhes e a dinâmica dos processos naturais e decorrentes da atuação humana no meio, o que subsidia a tomada de decisão com maior chance de acerto por se aproximar da situação real. O estado da arte disponível na literatura técnica para arrostar o problema exposto é vasto e apresenta casos bem sucedidos de experimentação, entretanto na condição Brasileira há necessidade de desenvolvimento de métodos que considerem as limitações de informações disponíveis, que sejam adequados à averiguação das causas, investigação de possíveis soluções e, ainda, possam ser replicados.

1 Objetivos

Geral:

Integrar metodologias para o estudo de Bacias Hidrográficas, com foco nos eventos extremos, através de simulação de processos hidrológicos de forma a auxiliar na gestão dos recursos hídricos.

Específicos:

- Realizar estudo do tempo de retorno de precipitações pluviiais máximas para a área da nascente da bacia hidrográfica do rio Doce (na parte conhecida como rio Piranga) por meio de inferência estatística Bayesiana;
- Produzir base cartográfica digital atual do uso e cobertura do solo e da topografia do terreno no intuito de identificar as relações entre a paisagem e o comportamento hidrológico da bacia, bem como subsidiar a obtenção de parâmetros utilizados nos modelos hidrológicos e;
- Ajustar modelo de um cenário crítico de inundação através de simulação de processos hidrológicos na Bacia Hidrográfica do rio Piranga (BHrP) para a identificação de problemas relacionados à infiltração da água no solo e ao escoamento superficial na região.

2 Hipótese de pesquisa

Através da análise de eventos extremos de precipitação aliada ao diagnóstico espacial do uso e cobertura do solo e simulação de processos hidrológicos é possível representar o comportamento hidrológico de uma bacia hidrográfica frente a determinado evento de precipitação causador de inundação.

3 Justificativa

Atualmente observa-se grande dificuldade para obtenção de dados hidrológicos – principalmente fluviométricos – em mananciais superficiais localizados em ambiente urbanizado. Por vezes, quando estão disponíveis, as séries históricas são recentes e não abrangem período de tempo suficiente para análise hidrológica consistente, o que prejudica sobremaneira a investigação da frequência de eventos extremos nessas localidades. Contudo, o cenário urbano é certamente um dos mais vulneráveis a tais eventos, social e economicamente, inclusive com risco de vida para suas populações, notadamente em maior número. Em parte, a vulnerabilidade pode ser explicada pela

intensidade e quantidade das alterações no meio físico da bacia. A análise hidrológica integrada à simulação das condições do ambiente natural e sua variação espacial permite verificar os papéis dos vários elementos que atuam nos processos que caracterizam as inundações de maneira mais efetiva que a convencional. Deste modo, a busca por métodos simplificados de simulação que, com o mínimo de ajustes paramétricos e recursos informacionais possam ser replicados em diferentes localidades, é uma demanda de pesquisa para avaliar cenários impactantes decorrentes das inundações bem como possíveis soluções.

4 Caracterização da área de estudo

A área de estudo compreende a porção da nascente da Bacia Hidrográfica do rio Doce, onde esse recebe o nome de rio Piranga, portanto Bacia Hidrográfica do rio Piranga. A bacia hidrográfica do rio Doce representa grande importância socioeconômica e ambiental para o Brasil, por isso é objeto de vários estudos e projetos de pesquisa.

O rio Piranga nasce na Serra das Vertentes, no encontro das Serras da Mantiqueira e do Espinhaço, no município de Ressaquinha, e sua foz situa-se na confluência com o rio do Carmo, formando o rio Doce, no município de Santa Cruz do Escalvado (Figura (I) 1). Possui extensão aproximada de 245 km e área de drenagem de aproximadamente 6.600 km². Seus principais afluentes são os rios São Bernardo, Xopotó, Turvo Limpo e Oratórios. Segundo recenseamento demográfico do ano de 2010, a região possui população estimada de aproximadamente 300.000 habitantes, distribuídos em 28 municípios completamente inseridos na Bacia Hidrográfica, com predominância da população urbana sobre a rural (CAMARGO, 2012). A densidade demográfica da região está em torno de 10 hab./km² e o Índice de Desenvolvimento Humano (IDH) dos municípios da bacia hidrográfica do rio Piranga situa-se em torno de 0,5, com destaque para Viçosa e Ponte Nova, cujo IDH situa-se na faixa de 0,75 (CONSÓRCIO ECOPLAN-LUME, 2010).

A Bacia Hidrográfica do rio Piranga (BHrP) apresenta uso e cobertura do solo caracterizado predominantemente por pastagens que subsidiam atividade pecuária, áreas de cultivo agrícola com ocupação por lavouras temporárias e permanentes,

essas últimas com destaque para a cafeicultura, também há consideráveis áreas plantadas com eucalipto, as quais, em conjunto com as demais caracterizam exploração antropogênica da área. A cobertura vegetal natural da região é definida por remanescentes florestais.

O clima na região das nascentes do rio Doce, segundo classificação de Köppen-Geiger, enquadra-se como tropical de altitude com chuvas durante o verão e verões frescos. A temperatura média anual varia em torno de 18°C em Barbacena e a precipitação média anual é em torno de 1500 mm (CONSÓRCIO ECOPLANLUME, 2010).

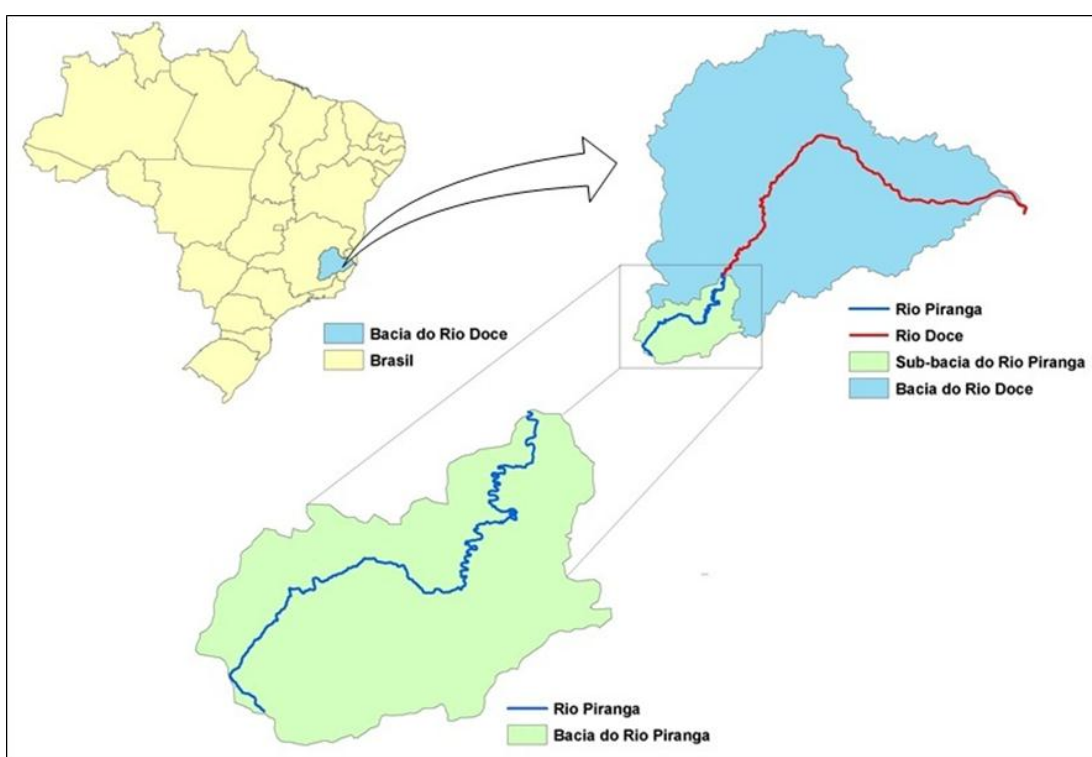


Figura (I) 1 – Localização das bacias hidrográficas dos Rios Piranga (polígono verde) e Doce (polígono azul), bem como sua posição no Estado de Minas Gerais no Território Brasileiro.

5 Problema

Várias cidades localizadas BHRP apresentam histórico de inundações (Figura (I) 2) provocadas por precipitações de longa duração e elevada intensidade. Barragens

instaladas no leito do rio Piranga podem contribuir para a magnitude dos eventos de inundação nas cidades banhadas pelo rio. Devido ao efeito da curva de remanso das barragens de jusante, a onda de cheia tem sua velocidade de propagação reduzida. Distante aproximadamente 34 km da cidade de Ponte Nova encontra-se o lago da barragem da Usina Hidrelétrica Risoleta Neves.

Na área de estudo observa-se os tipos de erosão laminar, sulcos e voçorocas, sendo os trechos mais críticos localizados ao norte da cidade de Viçosa, nas cabeceiras do rio Turvo Limpo e a sudeste da cidade de Alto Rio Doce. As colinas convexo-côncavas com vertentes ravinadas e escoamento concentrado favorecem o surgimento de sulcos e erosão laminar (CONSÓRCIO ECOPLAN-LUME, 2010).

O relevo bastante acidentado agrava a potencialidade do escoamento superficial com perda de solos devido em parte à velocidade de escoamento, fato este que justifica a ocorrência de voçorocas em vários locais na região. Outro condicionante é o pouco desenvolvimento econômico aliado à falta de recursos para o produtor agrícola, que dificultam o acesso a tecnologias que potencializam a produtividade e reduzem impactos ambientais.

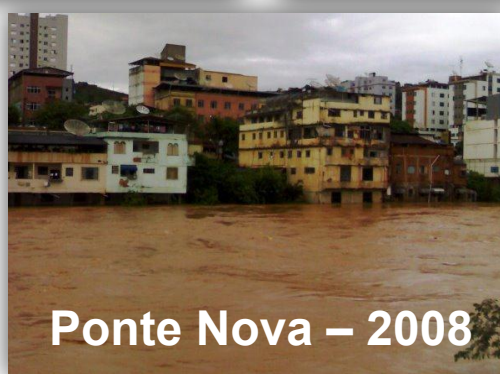


Figura (I) 2 – Registros fotográficos de inundação em algumas cidades localizadas na BHRP e margeadas por esse rio.

Diante do exposto é importante investigar os fatores relacionados à ocorrência de inundações na região.

6 Organização do trabalho

O trabalho está organizado em três estudos, o primeiro estudo versa sobre a análise estatística da série histórica de dados pluviométricos com vistas à obtenção de modelos de representação regional para predição de alturas máximas precipitadas associadas aos respectivos períodos de retorno utilizando abordagem Bayesiana. O segundo estudo apresenta métodos de distribuição espacial da informação especialista para incorporá-la ao processo de classificação de imagens digitais para melhorar seu resultado. O terceiro estudo apresenta a simulação de processos hidrológicos na BHrP como ferramenta de análise e avaliação das condições que ocasionam os problemas de inundação.

7 Referências

CAMARGO, R. D. A. Avaliação da suscetibilidade à erosão e proposta de zoneamento estratégico com vistas à garantia da sustentabilidade para a Bacia Hidrográfica do rio Piranga, MG. Universidade Federal de Viçosa. Viçosa, p. 104. 2012.

CONSÓRCIO ECOPLAN-LUME. Plano Integrado de Recursos Hídricos da Bacia Hidrográfica do Rio Doce e Planos de Ações para as Unidades de Planejamento e Gestão de Recursos Hídricos no Âmbito da Bacia do Rio Doce. [S.l.], p. 478. 2010.

II. Estudo de precipitações pluviiais máximas por meio da inferência Bayesiana: uma abordagem para eventos extremos

1 Introdução

Duzentos e onze dos duzentos e trinta municípios que se encontram na Bacia Hidrográfica do rio Doce (BHrD) possuem sede nela. A região tem população estimada em 3,7 milhões de habitantes (CONSÓRCIO ECOPLAN-LUME, 2010). Em sua região de nascente, onde o rio ainda é conhecido como Piranga, estão inseridos diversos municípios, entre os quais Ponte Nova, Guaraciaba e Porto Firme. Localizados às margens do rio Piranga, esses municípios são conhecidos por inundações recorrentes que em 1979, 1997, 2008 e 2010 tiveram consequências trágicas, de natureza humana e socioeconômica.

Eventos de origem natural como as precipitações intensas e de longa duração aliado a ações de ordem antropogênica como a elevação do potencial de assoreamento do rio e a construção de barragens no leito, são responsáveis pela ocorrência dessas inundações.

Tendo-se em vista não apenas o histórico dos eventos extremos de precipitação na região como também suas consequências e sabendo serem recorrentes, faz-se necessário conhecer o tempo de retorno associado aos volumes precipitados de maior magnitude, pois essas informações são necessárias para subsidiar a administração pública no planejamento de estratégias para atenuação dos seus efeitos. Modelos de distribuição teóricos de probabilidade são comumente ajustados aos dados da série histórica de observações pluviométricas para esse propósito.

Por meio da inferência Bayesiana é possível incorporar conhecimento prévio (BEIJO, VIVANCO e MUNIZ, 2009), o que se configura como estratégia útil ao lidar com modelos de probabilidade complexos, como é o caso da distribuição de Valores Extremos Generalizada – *GEV* (*Generalized Extreme Values*), indicada para estudos de eventos hidrológicos extremos (COLES e TAWN, 1996; KOUTSOYIANNIS, 2004; GAIONI, DEY e RUGGERI, 2010). A partir dos modelos *GEV* obtidos pretende-se aplicar análise de agrupamento baseada em

modelos (NASCIMENTO et al., 2012) como técnica para identificar similaridades entre eles.

1.1 Problema

Apesar de recorrentes, as precipitações de expressiva magnitude não são frequentes. Conhecer-las melhor bem como seus períodos de retorno é uma necessidade tanto para avaliar medidas direcionadas a atenuar os efeitos de inundações quanto para subsidiar projetos de obras de engenharia de grande porte. Como há, geralmente, poucos registros disponíveis desses eventos, utilizam-se métodos estatísticos de inferência para prever sua probabilidade de ocorrência. Contudo, os métodos estatísticos preditivos tradicionais estimam as maiores chuvas com incerteza elevada. Ainda é desconhecido um método adequado a eventos extremos que ofereça maior potencial de aplicação frente à escassez de dados e com maior acurácia.

1.2 Objetivo

Propor e avaliar um modelo Bayesiano hierárquico da distribuição Generalizada de Valores Extremos (*GEV – Generalized Extreme Values*), obtido pela técnica de agrupamento baseada em modelos, para predição de alturas máximas precipitadas e seus respectivos períodos de retorno e que represente uma determinada região.

1.3 Hipótese

É possível melhorar as estimativas de eventos extremos de precipitação através da obtenção de um modelo preditivo de representação regional obtido a partir de análise de agrupamento baseado em modelos de Função Densidade Probabilidade (FDP)

ajustados para dados das séries históricas das estações pluviométricas localizadas na região de uma bacia hidrográfica e suas proximidades.

Assume-se a premissa que, se n vetores paramétricos θ_i referentes a um mesmo tipo de FDP conhecida cujos respectivos modelos foram ajustados às amostras S_i de registros pluviométricos de n estações compartilham características semelhantes no espaço paramétrico e, portanto, são identificados em um mesmo agrupamento; Há então outro vetor paramétrico Φ que define uma FDP desse mesmo tipo que pode ser determinado a partir de um novo conjunto amostral S' formado pelo conjunto **união** das amostras S_i utilizadas na obtenção de cada θ_i que foram identificados na análise de agrupamento baseado em modelos.

2 Revisão de literatura

2.1 Distribuição de Valores Extremos

A teoria de valores extremos fornece um conjunto completo de ferramentas para analisar a distribuição estatística de valores extremos incluindo uma importante premissa que assume estacionariedade do modelo, isto é, implica em dizer que os seus parâmetros não alteram ao longo do tempo (BEGUERÍA et al., 2010). Com isso, apesar do modelo POT (*Peaks-Over-Threshold*) ser comum para hidrologia na estimativa de quantis extremos, suas premissas subordinadas ao grupo particular de leis paramétricas das Distribuições de Pareto Generalizadas são de independência e homogeneidade, o que conduz a subestimativas dos quantis mais elevados quando se considera dados heterogêneos (BACRO e CHAOUICHE, 2006).

Considerando que pelo método proposto as inferências são realizadas em dois momentos, no primeiro para determinação dos vetores paramétricos dos modelos (FDP) através dos registros das estações pluviométricas – um modelo para cada; e no segundo pelo novo ajuste da FDP aos registros agrupados (determinação de um novo vetor paramétrico) para que este modelo represente o grupo de estações, ou seja, uma região. Conclui-se que tal particularidade de ajuste em dois momentos não fere o

princípio teórico de independência dos eventos, mas fere o de homogeneidade no primeiro momento. Além disso, quando o tamanho da amostra aumenta, os únicos limites possíveis para distribuição da estatística de extremos independente da distribuição da população são os três tipos de distribuições de extremos Gumbel, Fréchet e Weibull, e esta é provavelmente a razão para o emprego difundido da distribuição Generalizada de Valores Extremos ou, do inglês, *GEV – Generalized Extreme Values* (PARK et al., 2011).

A FDP da distribuição *GEV* é dada pela Equação 1.

$$f(x) = \frac{1}{\sigma} \left(1 + \frac{\eta}{\sigma} (x - \mu) \right)^{-(1+\frac{1}{\eta})} \exp \left\{ - \left(1 + \frac{\eta}{\sigma} (x - \mu) \right)^{-\frac{1}{\eta}} \right\} \quad (1)$$

definida em $\frac{\eta}{\sigma} (x - \mu) \geq -1$, com parâmetros, η , σ e μ .

A Função de Distribuição Acumulada (FDA) da distribuição *GEV* é apresentada pela Equação 2.

$$P(X > x) = 1 - P(X \leq x)$$

$$P(X > x) - 1 = -P(X \leq x) * (-1)$$

$$1 - P(X > x) = P(X \leq x)$$

$$P(X \leq x) = \exp \left\{ - \left(1 + \frac{\eta}{\sigma} (x - \mu) \right)^{-\frac{1}{\eta}} \right\} \quad (2)$$

Em um estudo realizado no Canadá para escolher a melhor distribuição probabilística entre nove avaliadas (Beta-K, Beta-P, *GEV*, GNO, GP, Gumbel, Log-Pearson Tipo III, Pearson Tipo III e Wakeby) para representar precipitação máxima anual, apesar de resultados similares na avaliação da qualidade do ajustamento, os autores preferiram a *GEV* por requerer método mais simples para estimativa de seus parâmetros e ser baseada numa base teórica sólida para representar a distribuição de variáveis aleatórias de valores extremos (NGUYEN, TAO e BOURQUE, 2002).

Sua utilização para estudar vazões de inundação (GAIONI, DEY e RUGGERI, 2010), calcular o nível dos mares (HERRINGTON et al., 2008), monitorar a integridade de estruturas (PARK e SOHN, 2005), e até estimar rotações de uma turbina eólica (RAGAN e MANUEL, 2007) dão ideia da amplitude e diversidade de

suas possíveis aplicações. Além dessas, seu uso para análises de chuvas extremas é amplamente documentado na literatura (COLES e TAWN, 1996; CRISCI et al., 2002; NGUYEN, TAO e BOURQUE, 2002; BEIJO, VIVANCO e MUNIZ, 2009; PARK et al., 2011).

A distribuição *GEV* é um caso geral que combina outras três distribuições de valores extremos definidas por dois parâmetros, a de Gumbel dita do Tipo I; a de Fréchet dita do Tipo II; e a de Weibull dita do Tipo III; sendo essas últimas, casos especiais da *GEV* (HERRINGTON et al., 2008). A *GEV* é uma distribuição caracterizada por três parâmetros: locação (μ), escala (σ) e forma (η). Os dois primeiros representam na prática a média e desvio-padrão, respectivamente. O parâmetro forma depende da distribuição inicial dos dados e é responsável por determinar o peso da cauda da *GEV* bem como sua equivalência às distribuições de valores extremos do Tipo I ($\eta = 0$), do Tipo II ($\eta > 0$) ou do Tipo III ($\eta < 0$). É justamente o peso da cauda que torna a distribuição *GEV* útil para descrever fenômenos extremos (MICHELE e SALVADORI, 2005). Quando $\eta < 0$, o limite superior da distribuição é finito, portanto, a *GEV* caracteriza-se como uma distribuição de extremos mínimos. Porém, nos casos em que $\eta \geq 0$, o limite superior da distribuição *GEV* é infinito, assim, fica definida como uma distribuição de extremos máximos (COLES e TAWN, 1996).

2.2 Abordagem Bayesiana

Numa análise convencional a abordagem Bayesiana é empregada através do Teorema de Bayes (Equação 3), regra para atualização da informação sobre os parâmetros de uma distribuição a *posteriori* $\pi(\boldsymbol{\theta}|y)$, obtida pela combinação da função de verossimilhança $f(y|\boldsymbol{\theta})$ que contém a informação dos dados y em análise, com a distribuição a *priori* $\pi(\boldsymbol{\theta})$, que traz a informação prévia sobre o vetor paramétrico.

$$\pi(\boldsymbol{\theta}|y) = \frac{\pi(\boldsymbol{\theta})f(y|\boldsymbol{\theta})}{\int \pi(\boldsymbol{\theta})f(y|\boldsymbol{\theta})d(\boldsymbol{\theta})} \quad (3)$$

A forma proporcional da Equação 3 é suficiente para realizar estimativas dos parâmetros desconhecidos, contudo, para fazer inferências preditivas sobre uma observável desconhecida, é necessário $\int \pi(\boldsymbol{\theta})f(y|\boldsymbol{\theta})d(\boldsymbol{\theta})$. Conhecido o modelo e tendo em mente que a função de verossimilhança será obtida da FDP *GEV* nota-se que, para fazer a inferência *a posteriori* necessita-se conhecer a distribuição *a priori*. Visto que, a informação contida da distribuição *a priori* é responsável pela influência relativa de cada um dos componentes na informação *a posteriori*, alguns estudos Bayesianos se ocupam em incorporar expertise ao modelo através de consulta a especialistas (COLES e TAWN, 1996; BEIJO, VIVANCO e MUNIZ, 2009), o que fornece bons resultados uma vez que recorre a informações externas, além dos dados, para suportar as análises. Outros estudos se ocupam em avaliar o desempenho do modelo entre diferentes opções de distribuição *a priori* (GAIONI, DEY e RUGGERI, 2010), contudo essa abordagem pode gerar grande quantidade de informações, se apresentando adequada para uma estação, porém não a um conjunto delas, como é o caso. Adicionalmente, acredita-se que a consulta a especialista(s) sobre os parâmetros e quantis extremos de cada uma de quinze estações pluviométricas de uma mesma região seja algo aparentemente muito subjetivo.

Todo arcabouço conceitual que fundamenta o problema apresentado, com dados regionais (em várias estações), aponta como solução mais viável uma direção que toma outro sentido além do que tem sido utilizado. Conforme será atestado, essa direção é para utilização de modelos hierárquicos.

Permutabilidade é a principal característica dos modelos hierárquicos que confere a adequabilidade desses ao tipo de análise proposta. Se nenhuma informação além dos dados está disponível para distinguir os parâmetros entre si e nenhum ordenamento ou agrupamento pode ser feito (nesse momento), deve-se assumir simetria entre os parâmetros na sua distribuição *a priori*. Na prática, esse desconhecimento implica em permutabilidade (GELMAN et al., 1995) e sua expressão mais simples de distribuição permutável é apresentada na forma de distribuições iid (*independent and identically distributed*) para cada parâmetro θ como amostra de uma distribuição populacional governada por algum vetor paramétrico desconhecido ϕ , conforme a Equação 4.

$$p(\theta|\phi) = \int \left[\prod_{j=1}^j p(\theta_j|\phi) \right] p(\phi) d\theta \quad (4)$$

Nos modelos hierárquicos Bayesianos, como os hiperparâmetros da distribuição *a priori* θ são definidos em função de distribuições de probabilidade desconhecidas ϕ , essas são chamadas distribuições hiperprioris e tem seus próprios hiperparâmetros. Conforme a teoria Bayesiana, é necessário uma distribuição conjunta *a posteriori* para cada distribuição *a priori* especificada, quando se tem interesse em compará-las (SILVA et al., 2008). Portanto, para o tratamento completamente Bayesiano do modelo hierárquico, tem-se a Distribuição Conjunta *a Priori* dada por $P(\phi, \theta) = P(\phi)P(\theta|\phi)$ e sua Distribuição Conjunta *a Posteriori* apresentada pela Equação 5.

$$\begin{aligned} P(\phi, \theta|y) &\propto P(\phi, \theta)P(y|\phi, \theta) \\ &= P(\phi, \theta)P(y|\theta) \end{aligned} \quad (5)$$

Observa-se que ϕ afeta y somente através de θ . Apesar disso, deve-se ter muito cuidado com a ignorância ou desconhecimento em relação a ϕ . Informação suficientemente substancial a respeito dos parâmetros ϕ das hiperprioris deve ser empregada para, pelo menos restringir seus hiperparâmetros em uma região finita (GELMAN et al., 1995).

2.3 Verificação da convergência

Após um número elevado de iterações, pela teoria das Cadeias de Markov, espera-se que converjam para uma *distribuição estacionária*, entretanto nada garante que tal fato ocorra. Apesar disso, há testes estatísticos e de inspeção visual que verificam se uma cadeia aparentemente convergiu. Portanto, analisar a convergência das simulações é de crucial importância para a credibilidade dos resultados. Os testes de convergência mais comuns são os métodos diagnósticos de Raftery & Lewis (RL) para análise de cadeias individuais e o de Brooks, Gelman & Rubin (BGR),

apropriado para análise de duas ou mais cadeias paralelas. Para aplicar os testes utilizou-se o programa BOA (*Bayesian Output Analysis*) (SMITH, 2005).

2.4 Análise de Agrupamento Aglomerado Hierárquico

No contexto de análise de agrupamento, um grupo é um subconjunto de um conjunto de objetos, que estão, de certo modo, mais próximos uns dos outros, do que os demais objetos daquele conjunto. A análise de agrupamento é o processo de agregação de alguns objetos, de um dado conjunto de objetos, em grupos (BANERJEE e ROSENFELD, 1993).

Métodos hierárquicos de agrupamento procedem por estágios, produzindo uma sequência de partições nos dados, cada uma correspondente a um número diferente de grupos, que podem ser também aglomerados, no sentido de serem reunidos ou divididos em um ou mais grupos em cada estágio. Nos Agrupamentos Aglomerados Hierárquicos (*Hierarchical Agglomerative Clustering*), o número de estágios é limitado pela quantidade de grupos na partição inicial (FRALEY e RAFTERY, 1998), sendo essa, a análise de agrupamento mais difundida. Fundamenta-se no critério de dissimilaridade entre objetos, verificada através do cômputo de uma matriz de distância Euclidiana (Equação 6) dos mesmos.

$$d_{hk} = \alpha_i d_{hi} + \alpha_j d_{hj} + \beta d_{ij} + \lambda |d_{hi} - d_{hj}| \quad (6)$$

em que d_{ij} representa a distância Euclidiana entre os objetos i e j , combinados para formar um novo grupo k ; d_{hk} representa a distância Euclidiana entre um objeto restante h e o novo grupo k ; α_i , α_j , β e λ representam parâmetros que estão relacionados com o método de agrupamento escolhido (i.e., Média, Ward, Completo, Simples, entre outros) (SINGH, 2008). Para o método Completo, os valores para os parâmetros são $\alpha_i = 0.5$; $\alpha_j = 0.5$; $\beta = 0$; $\lambda = 0.5$. Recomenda-se calcular o nível de significância dos grupos identificados durante a análise de agrupamento, através dos respectivos p -value, como alternativa para verificar a medida de confiabilidade de cada um.

3 Método

O aspecto diferenciado desta abordagem está na análise de agrupamento se aplicar aos modelos ajustados aos registros das estações pluviométricas que definem a distribuição teórica de probabilidade por meio dos seus parâmetros e não nos registros de observações pluviométricas em si, como usualmente se verifica (COLES e TAWN, 1996; BEIJO, MUNIZ e CASTRO NETO, 2005; BEIJO, VIVANCO e MUNIZ, 2009; GAIONI, DEY e RUGGERI, 2010). A ideia fundamental está na possibilidade de elevar o número de observações empregadas no procedimento de determinação dos parâmetros de uma distribuição teórica de probabilidade. Desse modo, ao utilizar os registros de um determinado conjunto de estações, admite-se representatividade regional do modelo obtido.

A ilustração apresentada na Figura (II) 1 descreve resumidamente os processos sequenciais que foram seguidos para aplicação do método. Percebe-se que o escopo do método é definido de modo sequencial (*straight foward*), composto por cinco etapas principais.

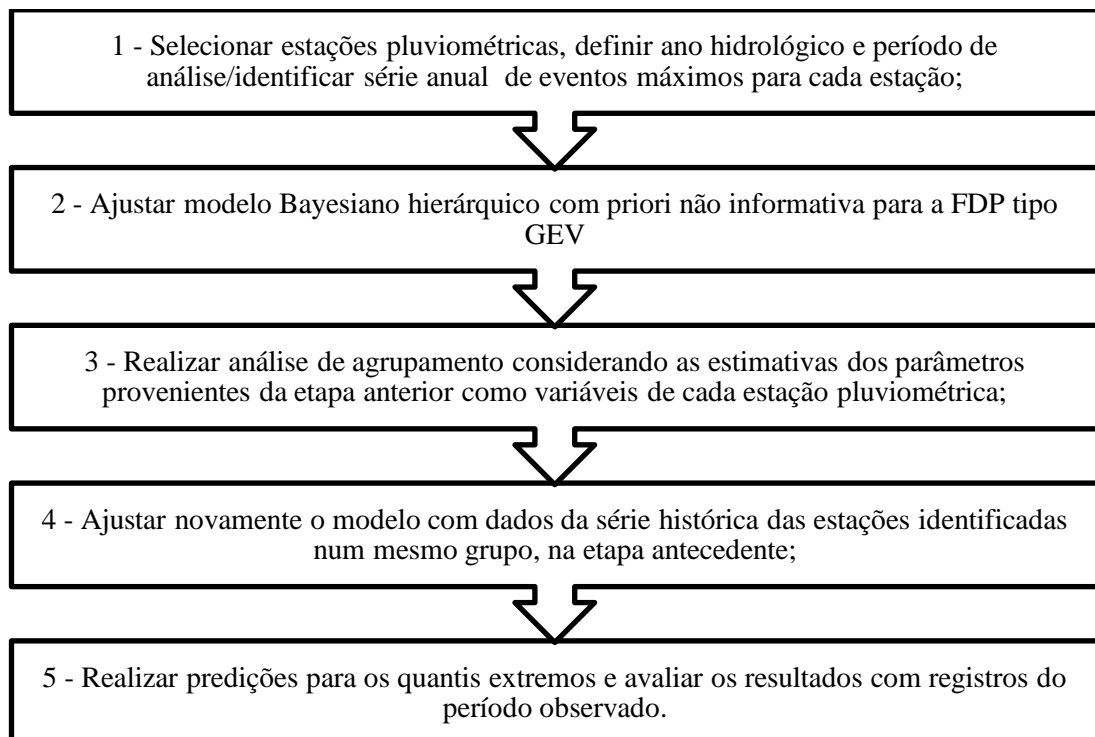


Figura (II) 1 – Fluxograma sequencial dos processos utilizados.

3.1 Organização dos dados pluviométricos

Após o levantamento e seleção das estações pluviométricas pertencentes à área de estudo que possuíam registros ininterruptos de dados, realizou-se uma análise exploratória das séries históricas com intuito de definir o ano hidrológico para, em seguida, selecionar os dados do período chuvoso. Em casos onde o enfoque do estudo seja o total mensal ou anual precipitado, pode-se lançar mão da técnica de preenchimento de falhas para aumentar a disponibilidade de estações à análise.

Quinze estações pluviométricas localizadas na Bacia Hidrográfica do rio Piranga (BHRP) e suas proximidades foram selecionadas para as análises (Quadro (II) 1).

O início da série histórica foi determinado pela amplitude de suas observações. Nesta pesquisa os maiores períodos são de 70 anos e os menores de 37 anos de dados. A série histórica anual de eventos máximos diários para cada estação em análise encontra-se disponível no Apêndice (II) A.

A partir dos registros de altura precipitada nas estações pluviométricas, calculou-se os valores médios mensais para determinar período chuvoso a partir da identificação do ano hidrológico. Definido o período chuvoso, selecionou-se os registros de altura máxima diária precipitada. O período-base foi fixado com base no menor.

Id.	Código	Nome	Sub-bacias	Coordenadas geográficas	Altitude (m)	Período de observação
e1	02043011	Fazenda Paraíso	Rio Doce	Lat.: -20°23'24" Lon.: -43°10'49"	477	1941-2011 / 70 anos
e2	02143003	Desterro do Melo	Rio Doce	Lat.: -21°08'57" Lon.: -43°31'12"	780	1941-2011 / 70 anos
e3	02043009	Acaiaca Jusante	Rio Doce	Lat.: -20°21'45" Lon.: -43°08'38"	423	1941-2011 / 70 anos
e4	02143005	Campolide	Rio Grande	Lat.: -21°16'44" Lon.: -43°49'12"	997	1941-2011 / 70 anos
e5	02043014	Porto Firme	Rio Doce	Lat.: -20°40'13" Lon.: -43°05'17"	598	1941-2011 / 70 anos
e6	02143007	Vargem do Engenho	Rio Grande	Lat.: 21°11'34" Lon.: -43°36'50"	1120	1941-2011 / 70 anos
e7	02043010	Piranga	Rio Doce	Lat.: -20°41'26" Lon.: -43°17'58"	620	1953-2011 / 58 anos

Id.	Código	Nome	Sub-bacias	Coordenadas geográficas	Altitude (m)	Período de observação
e8	02043025	Usina da Brecha	Rio Doce	Lat.: -20°31'00" Lon.: -43°01'00"	529	1959-2011 / 52 anos
e9	02043005	Conselheiro Lafaiete	Rio Paraopeba	Lat.: -20°39'54" Lon.: -43°46'40"	996	1965-2011 / 46 anos
e10	02042015	Seriquite	Rio Doce	Lat.: -20°43'34" Lon.: -42°55'02"	638	1966-2011 / 45 anos
e11	02043026	Braz Pires	Rio Doce	Lat.: -20°50'51" Lon.: -43°14'31"	632	1966-2011 / 45 anos
e12	02043027	Fazenda Ocidente	Rio Doce	Lat.: -20°17'08" Lon.: -43°05'56"	462	1967-2011 / 44 anos
e13	02042016	São Miguel do Anta	Rio Doce	Lat.: -20°40'57" Lon.: -42°48'24"	583	1966-2011 / 45 anos
e14	02042024	Viçosa	Rio Doce	Lat.: -20°45'00" Lon.: -42°51'00"	712,2	1967-2011 / 44 anos
e15	02042018	Ponte Nova Jusante	Rio Doce	Lat.: -20°23'05" Lon.: -42°54'10"	350	1974-2011 / 37 anos

Quadro (II) 1 – Estações pluviométricas selecionadas para análise.

3.2 Quantil da Distribuição *GEV*

A FDP adotada para a análise das precipitações anuais máximas foi a distribuição Generalizada de Valores Extremos – *GEV*, recomendada para estudos de frequência de inundações pelo *National Environmental Research Council* (NERC) do Reino Unido no Relatório de Estudos de Inundações - *Flood Studies Report* (SUTCLIFFE, 1978).

No contexto da hidrologia, o que se busca conhecer é a altura máxima precipitada associada a um determinado período de retorno em anos, isto é, o valor estimado para um determinado quantil extremo da distribuição. Essa estimativa pode ser obtida ao se desenvolver a Equação 2 para fornecer o quantil $(1 - p)$ da distribuição de máximas anuais demonstrada pela Equação 7.

$$1 - P(X > x) = P(X \leq x)$$

$$\begin{aligned}
1 - P(X > x) &= \exp \left\{ - \left(1 + \frac{\eta}{\sigma}(x - \mu) \right)^{\frac{1}{\eta}} \right\} \\
[\ln(1 - P(X > x))]^\eta &= \left[- \left(1 + \frac{\eta}{\sigma}(x - \mu) \right)^{\frac{1}{\eta}} \right]^\eta \\
[\ln(1 - P(X > x))]^{-\eta} &= \left(-1 - \frac{\eta}{\sigma}x + \frac{\eta}{\sigma}\mu \right)^{+\eta} \\
[\ln(1 - P(X > x))]^{-\eta} &= -1 - \frac{\eta}{\sigma}x + \frac{\eta}{\sigma}\mu \\
\frac{\eta}{\sigma}x &= -[\ln(1 - P(X > x))]^{-\eta} - 1 + \frac{\eta}{\sigma}\mu \\
x &= \frac{\sigma}{\eta} \left[-(\ln(1 - P(X > x)))^{-\eta} - 1 \right] + \frac{\sigma}{\eta} \cdot \frac{\eta}{\sigma} \mu \\
x &= \frac{\sigma}{\eta} \left[-(\ln(1 - P(X > x)))^{-\eta} - 1 \right] + \mu
\end{aligned} \tag{7}$$

onde $P(X > x)$ é o período de retorno em anos; e x é sua altura máxima precipitada associada. Adotou-se os períodos de retorno (T) de 5, 10, 20, 50, 70, 100, 200 e 500 anos. Como $T = \frac{1}{P}$, $P = T^{-1}$.

3.3 Definição do Modelo Hierárquico Bayesiano

Assumiu-se que a função de verossimilhança segue uma distribuição de valores extremos generalizada (*GEV*) definida por $\theta = (\mu, \sigma, \eta)$, ou seja $P(y|\theta) \sim GEV(\mu, \sigma, \eta)$. Os parâmetros do vetor θ , por sua vez, foram definidos em função de suas distribuições *a priori*, seja na forma de hiperparâmetros (no caso de η) ou na forma de hiperdistribuições (no caso de μ e σ), através de $P(\mu) \sim N(\mu_0, \tau)$; $P(\sigma) = \frac{1}{\sqrt{\tau}}$; $P(\eta) \sim U(a_0, b_0)$. As distribuições hiperprioris do vetor paramétrico ϕ foram definidas por $P(\mu_0) \sim U(a_1, b_1)$; $P(\tau) \sim \Gamma(\alpha, \beta)$. Certamente o modelo hierárquico (Figura (II) 2) é compreendido melhor através de sua representação gráfica.

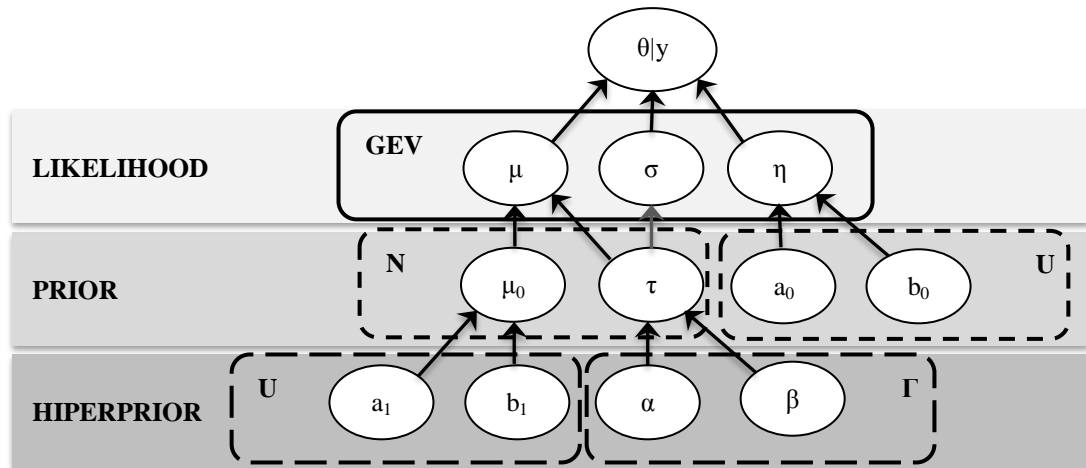


Figura (II) 2 – Representação gráfica do modelo Bayesiano Hierárquico para estimar os parâmetros da distribuição *GEV*, aplicada a estudos de precipitações pluviiais máximas anuais.

Importante mencionar que apesar da ideia de desconhecimento (ou não comprometimento) a respeito da definição das *prioris* para cada estação pluviométrica que fundamentou a escolha da abordagem do modelo hierárquico, adotou-se o critério de restringir a região limite de oscilação dos parâmetros forma (η) e locação (μ) com fundamento teórico da definição da *GEV* no primeiro caso e com base nos conhecimento das alturas máximas precipitadas no último. Desse modo, o parâmetro escala foi definido por uma distribuição Uniforme com hiperparâmetros no intervalo $[-1/2, 1/2]$, ou seja, $\eta \sim U(-0.5, 0.5)$ e o parâmetro locação foi definido através de distribuição *a priori* Normal com hiperparâmetros μ_0 e τ , isto é $\mu \sim N(\mu_0, \tau)$. Esses últimos, por sua vez, foram definidos por distribuições hiperprioris. Para a média μ_0 uma distribuição Uniforme com hiperparâmetros compreendidos no intervalo $[50, 100]$, isto é $\mu_0 \sim U(50, 100)$. Uma vez que o parâmetro precisão (τ) é definido como $\tau = \sigma^{-2}$, adotou-se uma distribuição hiperpriori Gamma com hiperparâmetros muito pequenos, ou seja, $\tau \sim \Gamma(0.001, 0.001)$. Desse modo, o parâmetro escala (σ) foi definido como sendo $\sigma = (\sqrt{\tau})^{-1}$.

O vetor Y corresponde aos valores de precipitação máxima anual da série histórica, para cada estação pluviométrica no primeiro momento, e da união dos dados do conjunto de estações agrupadas no segundo momento. Em resumo, a descrição de cada componente do modelo Bayesiano hierárquico ficou como segue:

$$Y|\mu, \sigma, \eta, \mu_0 \sim GEV(\mu, \sigma, \eta); \quad \mu \sim N(\mu_0, (1/\sigma)^2); \quad \mu_0 \sim U(50, 100); \\ (1/\sigma)^2 \sim \Gamma(0.001, 0.001); \text{ e } \eta \sim U(-0.5, 0.5).$$

As distribuições Normal, Uniforme e Gamma são apresentadas nas Equações 8, 9 e 10, respectivamente.

$$N = \sqrt{\frac{\tau}{2\pi}} \exp \left[-\frac{\tau}{2} (x - \mu)^2 \right]; -\infty < x < \infty \quad (8)$$

$$U = \frac{1}{b - a}; a < x < b \quad (9)$$

$$\Gamma = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}; x > 0 \quad (10)$$

Inferências sobre as quantidades de interesse na abordagem Bayesiana são fundamentadas nas suas distribuições marginais *a posteriori*, que podem ser obtidas integrando-se a distribuição conjunta *a posteriori* (SILVA e LEANDRO, 2009), resolvendo-se a Equação 5. Entretanto, ao desenvolver esse modelo hierárquico substituindo os parâmetros dos vetores pelas respectivas FDPs apresentadas nas equações 1, 8, 9 e 10 pode-se constatar que a solução da equação resultante não é tratável analiticamente (Equação 11).

$$\begin{aligned} P(\mu, \sigma, \eta, \mu_0) &\propto \\ P(Y|\mu, \sigma, \eta, \mu_0) \cdot P\left(\mu \middle| \mu_0, \frac{1}{\sigma^2}\right) \cdot P(\mu_0) \cdot P\left(\frac{1}{\sigma^2}\right) \cdot P(\eta) &\propto \\ \prod_{i=1}^n \frac{1}{\sigma} \left[1 + \frac{\eta}{\sigma} (x - \mu) \right]^{-\left(1 + \frac{1}{\eta}\right)} \exp \left\{ - \left[1 + \frac{\eta}{\sigma} (x - \mu) \right]^{-\frac{1}{\eta}} \right\} \cdot \sqrt{\frac{1}{2\pi\sigma^2}} \exp - \frac{1}{2\sigma^2} (\mu & \\ - \mu_0)^2 \cdot \frac{1}{(100 - 50)} \cdot \frac{(0.001)^{0.001}}{\Gamma(0.001)} \left(\frac{1}{\sigma^2} \right)^{0.001-1} e^{-0.001\left(\frac{1}{\sigma^2}\right)} \cdot \frac{1}{(0.5 + 0.5)} & \end{aligned} \quad (11)$$

Uma alternativa mais fácil para resolver o problema de determinar a distribuição *a posteriori* de modelos hierárquicos é a opção pelo método de Monte Carlo via Cadeias de Markov (MCMC) que, apesar de intrincado, quando usado com cautela apresenta resultados confiáveis (GELMAN et al., 1995). Portanto, para realizar as análises empregou-se o software livre OpenBUGS, um programa para análise Bayesiana de modelos estatísticos complexos.

3.4 Análise de Agrupamento Baseado em Modelos

A análise de agrupamento tem sentido apenas quando parte dos dados realmente constituem grupos (BANERJEE e ROSENFELD, 1993), sob a perspectiva de determinar a estrutura intrínseca de dados agrupados quando nenhuma outra informação está disponível além dos valores observados (FRALEY e RAFTERY, 1998). Assim sendo, cada estação pluviométrica em análise é uma candidata para o estudo das precipitações máximas, pois se localizam em uma mesma região, a bacia hidrográfica, e, por isso, supõe-se partilharem várias características ambientais e climáticas. Como a análise de agrupamento almeja identificar possíveis similaridades entre elas para então reuni-las em um grupo, esse procedimento reduz o esforço analítico uma vez que subsidia a escolha das estações com dados similares apenas.

A Análise de Agrupamento Baseado em Modelos (ou *Model-Based Cluster Analysis*) assume a premissa que a análise de agrupamento deveria ser baseada em um modelo para os dados, i.e., em suposições de como os dados foram gerados (BANERJEE e ROSENFELD, 1993). Para suportar essa hipótese, realizou-se uma análise de agrupamento baseado em modelo (*Model-Based Cluster Analysis*) com os parâmetros estimados pelo ajuste da distribuição *GEV* aos registros de precipitações diárias máximas anuais das estações. A Análise de Agrupamento foi realizada no software estatístico R (R Development Core Team, 2011), através do pacote *PVCLUST*, utilizado para calcular o *p*-valor aproximadamente não viesado (*Approximately Unbiased – AU p-value*) de cada grupo identificado. Essa abordagem permite determinar o nível de significância para cada agrupamento.

4 Resultados e discussões

Os resultados elementares do método de análise proposto dizem respeito aos algoritmos adotados no processo MCMC para se obter a estimativa de cada

parâmetro do modelo hierárquico. O relatório de informações sobre os nodos do modelo no software OpenBUGS forneceram a relação de algoritmo utilizado para cada parâmetro conforme apresentado no Quadro (II) 2.

Parâmetro	Algoritmo
η :	slice updater
μ :	adaptive metropolis 1D updater
μ_0 :	conjugate normal updater
τ :	slice updater

Quadro (II) 2 – Relação de algoritmos utilizados na determinação dos parâmetros do modelo hierárquico Bayesiano pelo software OpenBUGS.

4.1 Inferência multiparamétrica

Inicialmente se avaliou os resultados dos modelos ajustados para cada estação pluviométrica individualmente através da análise de convergência dos seus parâmetros, realizada em duas etapas. Como foram processadas duas cadeias paralelas para cada parâmetro, a avaliação de convergência foi realizada primeiramente para o par de cadeias. Em seguida o par de cadeias paralelas foi reunido em uma única cadeia para permitir avaliar sua convergência e obter estimativas dos parâmetros. Apesar da avaliação visual através de gráficos de métodos convergência (autocorrelação, densidade da *posteriori*, Gelman & Rubin, Geweke, entre outros) ser frequentemente empregada, devido à quantidade de parâmetros optou-se por apresentar os métodos diagnósticos descritivos.

Aplicou-se o teste de Brooks, Gelman & Rubin (BGR) para primeiramente avaliar a convergência do par de cadeias paralelas. Para esse teste, normalmente os valores dos parâmetros RF (*Reduction Factor*) e CSRF (*Corrected Scale Reduction Factor*) devem ser próximos a 1 para um melhor resultado, pois sinalizam que as amostras devem ser consideradas provenientes de uma distribuição estacionária. Todos os resultados foram favoráveis em relação à indicação de convergência dos parâmetros, com resultados próximos a unidade, e estão disponíveis no Apêndice. Os resultados do teste diagnóstico de convergência de BGR estão apresentados no Apêndice (II) B e os do teste diagnóstico de RL estão apresentados no Apêndice (II) C.

Deve-se analisar os números da referida tabela considerando que *Lower bound* é o número de iterações necessárias para estimar, a partir de amostras independentes, o quantil especificado com a acurácia desejada. *Total* é o número de iterações necessárias para estimar cada parâmetro. Em outras palavras, trata-se de uma medida indicadora da facilidade/dificuldade de convergência do modelo. *Burn-in* é o número de iterações iniciais descartadas no início de geração da cadeia e *Thinning* é o intervalo de amostragem que deve ser adotado. O *Dependence Factor* explora a relação entre *Total* e *Lower Bound*, na qual mede o incremento multiplicativo no número de iterações necessárias para alcançar convergência devido à correlação da cadeia. Valores próximos a unidade indicam bom resultados e maiores que 5.0 frequentemente indicam falha na convergência (SMITH, 2005). Uma vez mais, todos os resultados do *Dependence Factor*, o principal indicador de convergência, para todos os parâmetros da cadeia unificada encontram-se muito próximos do ideal, indicando que houve convergência dos parâmetros para uma distribuição estacionária. Em relação à medida *Total*, nota-se que os parâmetros apresentaram diferentes graus de dificuldade para atingir a convergência, entretanto com pouca variação. A partir da média de valores da medida *Total*, os parâmetros em ordem crescente de grau aparente de dificuldade para atingir a convergência foram: *forma*, *locação* e *escala*.

Uma vez atingida convergência das cadeias e, portanto, confirmada a condição de estacionariedade dos parâmetros, pode-se fazer uso deles nas etapas analíticas subsequentes. Nas ilustrações a seguir são apresentados os resultados obtidos dos parâmetros *forma* (Figura (II) 3), *locação* (Figura (II) 4) e *escala* (Figura (II) 5), com seus valores de limite inferior (*Li*) e limite superior (*Ls*), correspondentes ao intervalo de mais alta probabilidade (*HPD – High Posterior Density*) ao nível de significância de 0.05, bem com o valor da média *a posteriori* (*MP*).

Pelos valores obtidos em relação ao parâmetro *forma* (Figura (II) 3) nota-se que o modelo *GEV* não foi plenamente apropriado a todas as estações para ajuste de uma distribuição de extremos máximos, visto que alguns intervalos de valores são notadamente inferiores a zero, por exemplo, aqueles dos modelos *e1*, *e2*, *e11*, *e12* e *e13*, o que classifica os respectivos modelos como sendo Weibull - Tipo III, adequada a extremos mínimos. O restante dos modelos apresentaram faixas de variação no intervalo que admite valores próximos de zero ou superiores, o que os

classifica como sendo respectivamente de Gumbel – Tipo I ($e5, e6, e7, e8, e9$) ou Fréchet – Tipo III ($e3, e4, e10, e14$ e $e15$).

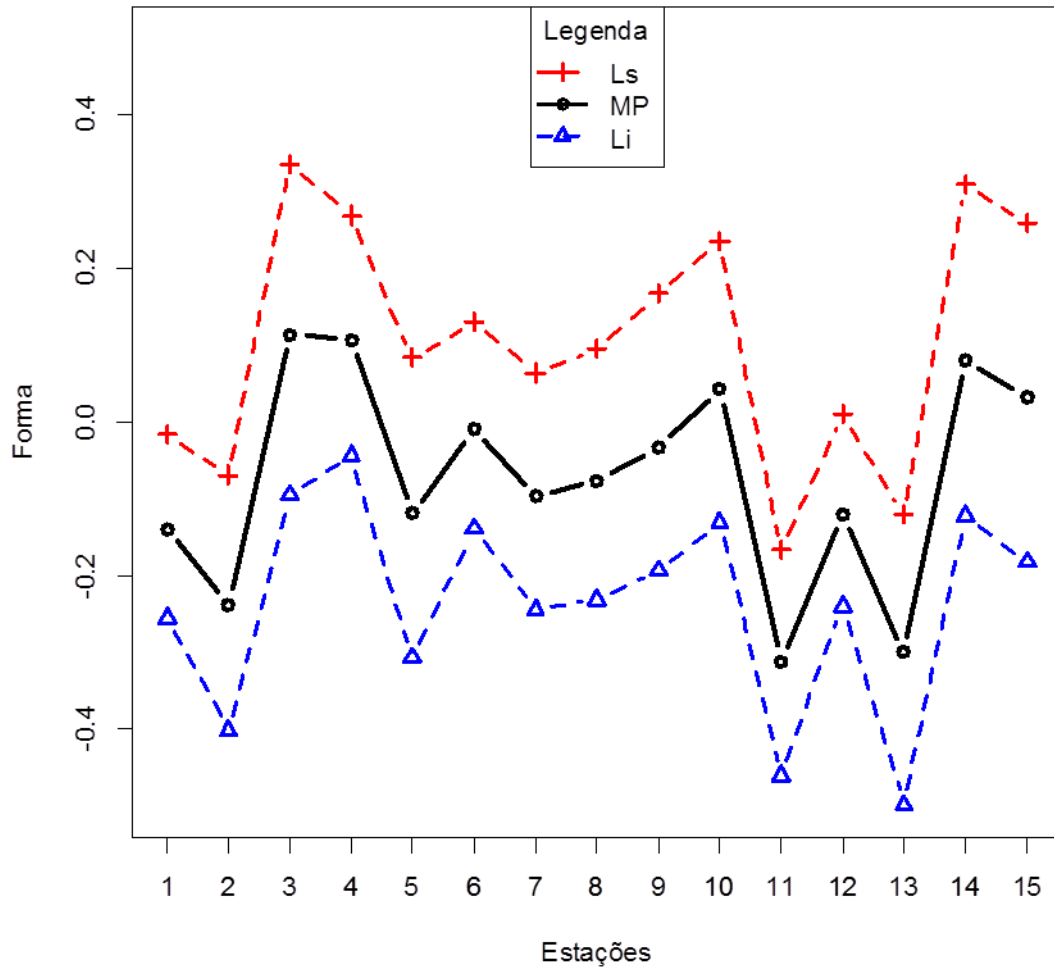


Figura (II) 3 – Inferência do parâmetro *forma* para cada estação com seus limites superior (*Ls*), inferior (*Li*) e média *a posteriori* (*MP*).

O parâmetro *escala* (Figura (II) 4) apresentou regularidade de variabilidade nas regiões entre 15 (*Li*) e 30 (*Ls*) unidades. Os menores valores de escala foram verificados para o modelo $e3$, com *Ls* abaixo de 20 unidades, enquanto os maiores foram para o modelo $e2$, com *Li* acima de 20 unidades.

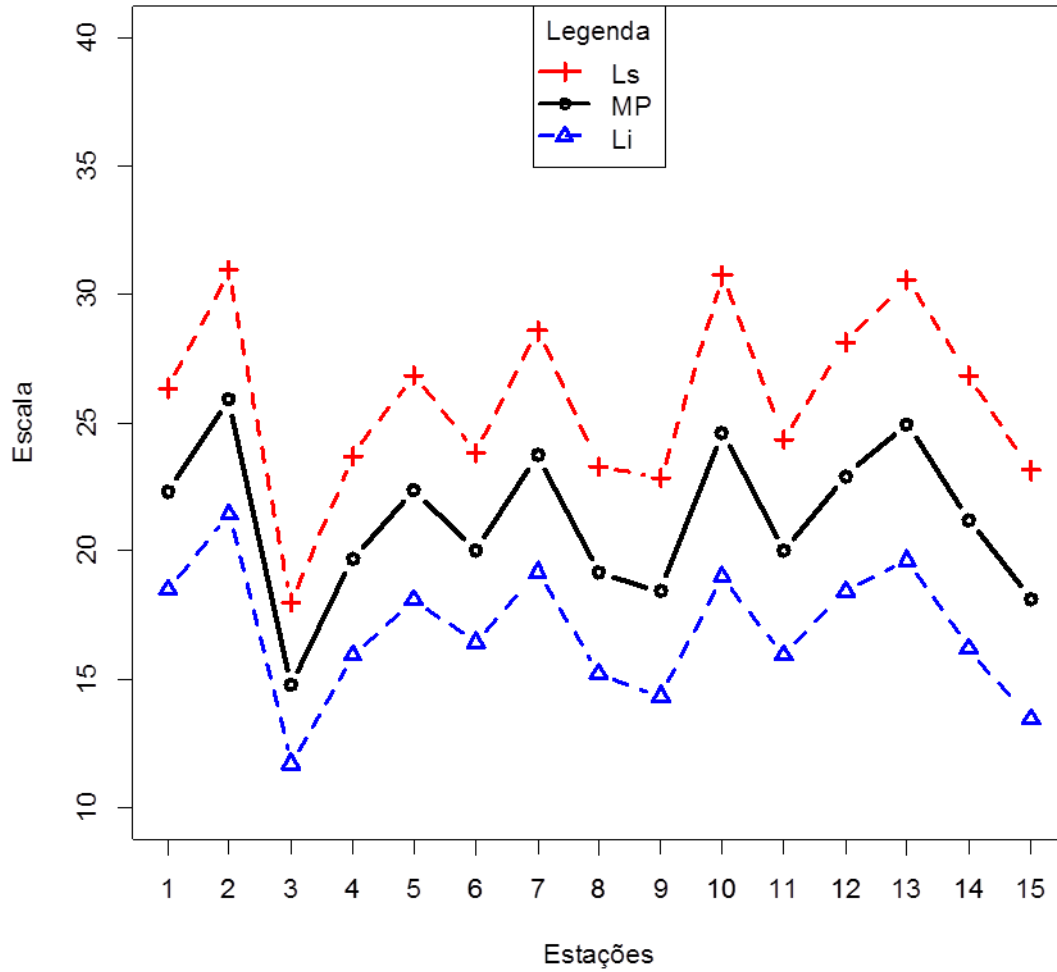


Figura (II) 4 – Inferência do parâmetro *escala* para cada estação com seus limites superior (*Ls*), inferior (*Li*) e média a *posteriori* (*MP*).

O parâmetro *locação* (Figura (II) 5) apresentou oscilação de maior magnitude entre modelos, bem como sua faixa de variação entre o intervalo *HPD*. Nota-se que os modelos *e2* e *e12* apresentaram respectivamente a maior e a menor expressão de valores, com fronteira na região próxima de 70 unidades. Em relação à amplitude de variação do parâmetro para um mesmo modelo, *e10* foi o maior e *e3* o menor.

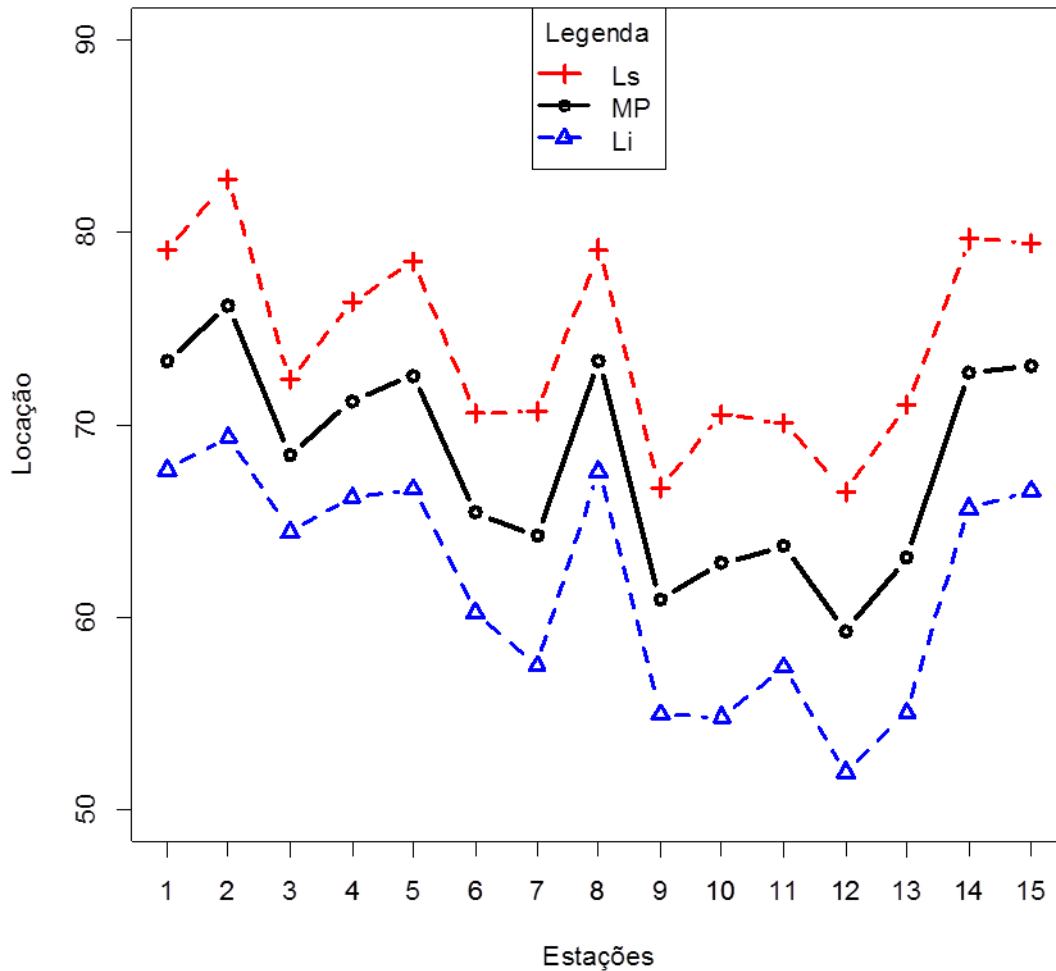


Figura (II) 5 – Inferência do parâmetro *locação* para cada estação com seus limites superior (*Ls*), inferior (*Li*) e média a *posteriori* (*MP*).

A partir dos valores da média a *posteriori* dos parâmetros dos modelos *e1*, *e2*, *e3*, *e4*, *e5* e *e6*, utilizou-se a Equação 3 para fazer inferência de altura máxima precipitada diária para os quantis cujas estações pluviométricas possuíam tempo de observação da série histórica correspondentes aos períodos de retorno (*T*) de até 70 anos (Figura (II) 6). Esta comparação teve o objetivo de verificar a aderência dos modelos ajustados aos dados observados das referidas estações. Utilizou-se os mesmos símbolos e nomes de identificação das estações/modelos para representar os valores dos registros dos períodos de retorno observados e os valores preditos pelo modelo, ambos acrescidos com terminação “_obs” ou “_hat”, respectivamente e para cada estação avaliada.

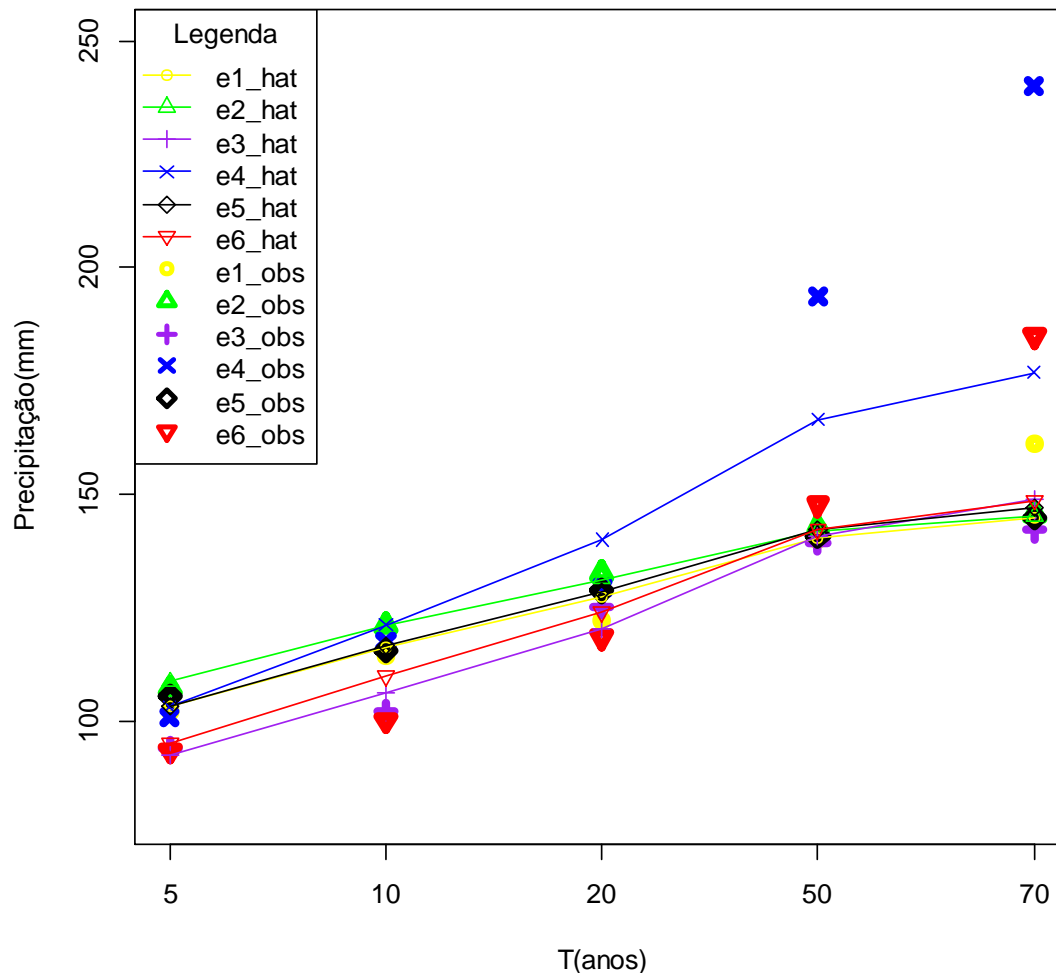


Figura (II) 6 – Comparação de alturas máximas precipitadas diárias entre os valores preditos (terminação _hat) pelos modelos ajustados e os registros observados (terminação _obs) da série histórica para as estações *e1*, *e2*, *e3*, *e4*, *e5* e *e6*, com períodos de retorno de até 70 anos.

Percebe-se que, apesar de alguma variação para os períodos de retorno até 20 anos, a maior parte dos modelos apresentaram discrepâncias em relação às alturas observadas, com exceção dos modelos *e2* e *e5*, que apresentaram ótima aderência para todos os períodos de retorno (Figura (II) 6). Com exceção de *e1*, *e4* e *e6*, os modelos apresentaram boa aderência para os períodos de 50 e 70 anos. Aliás, os modelos *e4* e *e6* não apresentaram aderência satisfatória para nenhum período de retorno, podendo ser classificados como os piores ajustes. Convém mencionar que o modelo *e1* apenas não foi aderente para o período de retorno de 70 anos, porém o foi para os períodos anteriores. O modelo *e3* apresentou desempenho apenas razoável. Optou-se por não apresentar a comparação para outras estações devido ao período de retorno ser inferior a 70, a maior parte menor que 50 anos.

4.2 Análise de Agrupamento Baseada em Modelos

Como se tratam de estações pluviométricas de uma mesma região, suspeita-se que as características dos modelos que as representem, compartilham características no espaço de parâmetros. Para verificar essas características optou-se por visualizar o espaço paramétrico através de abordagem que auxilia verificar melhor as semelhanças/diferenças dos vetores. Portanto, apresenta-se (Figura (II) 7) os vetores paramétricos plotados em uma figura de representação espacial tridimensional onde o eixo x assume os valores de *locação*, o eixo y assume os de *escala* e o eixo z os de *forma*, no espaço paramétrico.

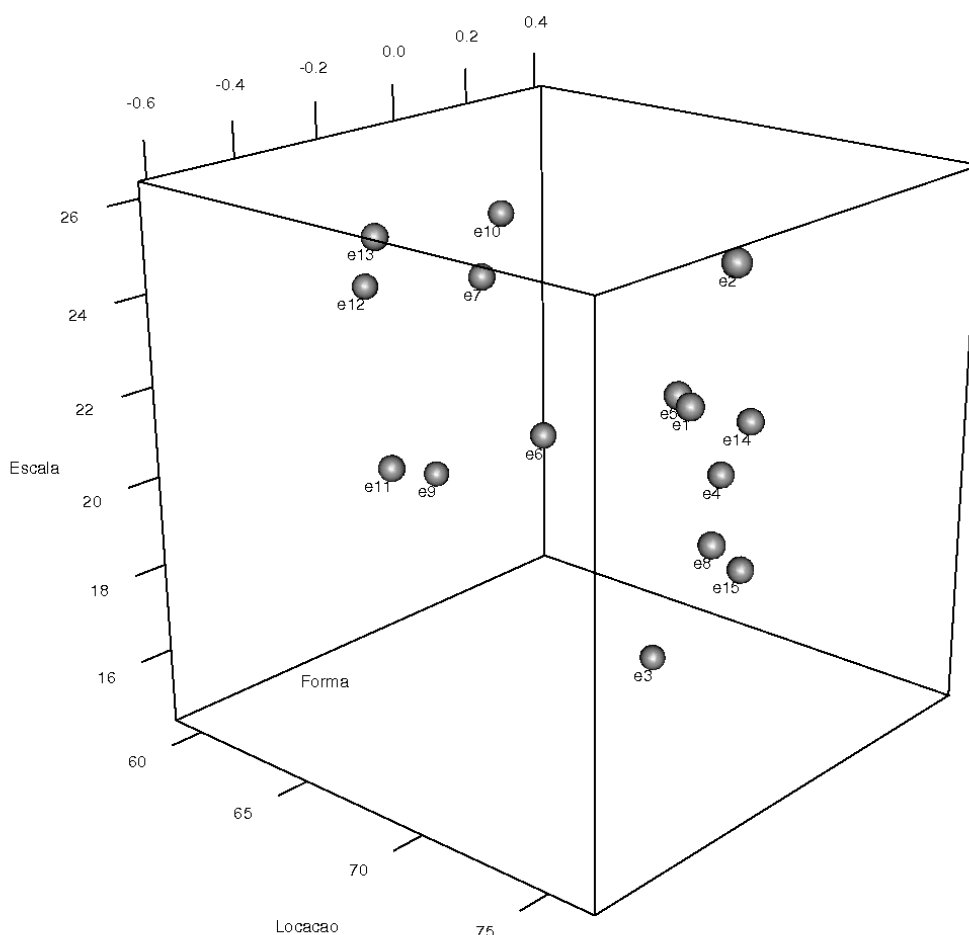


Figura (II) 7 – Espaço paramétrico de θ resultante do ajuste da distribuição GEV para cada estação pluviométrica onde as dimensões *locação* (x) *escala* (y) e *forma* (z) associadas respectivamente aos parâmetros μ , σ e η .

Ao observar a representação espacial dos parâmetros dos modelos das estações em análise (Figura (II) 7), realmente é possível constatar que alguns parâmetros ocupam posição muito próxima no espaço paramétrico, ou seja, compartilham uma mesma vizinhança.

Para comprovar as suspeitas, agora mais evidentes, realizou-se a análise de agrupamento baseada em modelos e obteve-se a identificação de grupos em diferentes níveis hierárquicos, conforme ilustra o dendrograma da Figura (II) 8. Nela os retângulos envolventes com bordas na cor azul evidenciam os grupos que não apenas parecem existir por erro de amostragem, mas de fato eles serão observados caso o número de observações aumente. Ou seja, a hipótese que o agrupamento não existe é rejeitada a determinado nível de significância.

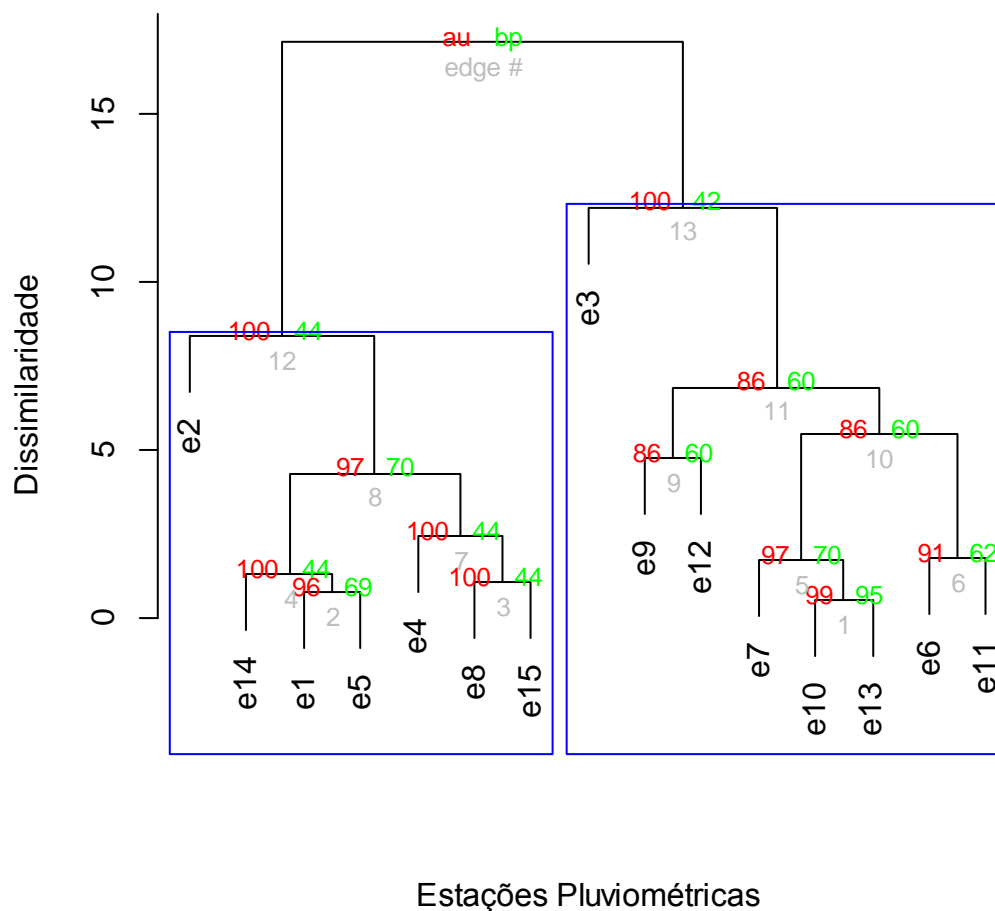


Figura (II) 8 – Dendrograma resultante da análise de agrupamento hierárquico dos parâmetros das estações pluviométricas com os *p*-valores calculados para cada agrupamento; em evidência pelos retângulos envoltórios grupos com nível de significância de 0.05.

Conforme a ilustração do dendrograma (Figura (II) 8), foram identificados dois grupos estações no primeiro nível da hierarquia, ao nível de significância *alpha* de 0.05, destacados pelos retângulos envolventes. Os grupos obtidos foram: grupo 1 pelo conjunto $G1 = \{e1, e2, e4, e5, e8, e14, e15\}$ com 7 estações e grupo 2 pelo conjunto $G2 = \{e3, e6, e7, e9, e10, e11, e12, e13\}$ com 8 estações.

Em seguida realizou-se uma análise espacial com a geometria representativa de cada grupo sobre a bacia, com intuito de verificar a representatividade de área dos modelos que serão ajustados. Essa análise foi governada pela produção de um mapa poligonal de áreas de influência de cada estação, obtido por polígonos de *Thiessen*. Cada polígono identificado por sua respectiva estação foi categorizado em um dos grupos, conforme resultado da análise de agrupamento, como *G1* ou *G2*. A ilustração da Figura (II) 9 mostra as áreas de influência dos grupos de estações. Nota-se que a área ocupada por *G1* encontra-se expressivamente na porção nordeste (NE) da bacia, na região à jusante, com alguma, porém menor influência na região sul (S) e leste (E). A área representada por *G2* ocupa majoritariamente a porção norte (N) e noroeste (NO), notadamente na região de nascentes do rio Piranga e se estende para a porção média e desta até os limites do seu divisor de águas, à sudeste (SE). A zona de transição de representatividade dos modelos na porção média da bacia pode ser explicada, em parte, pela interferência do terreno montanhoso nessa região, o que configura condições especiais para ocorrência de precipitações nesse local.

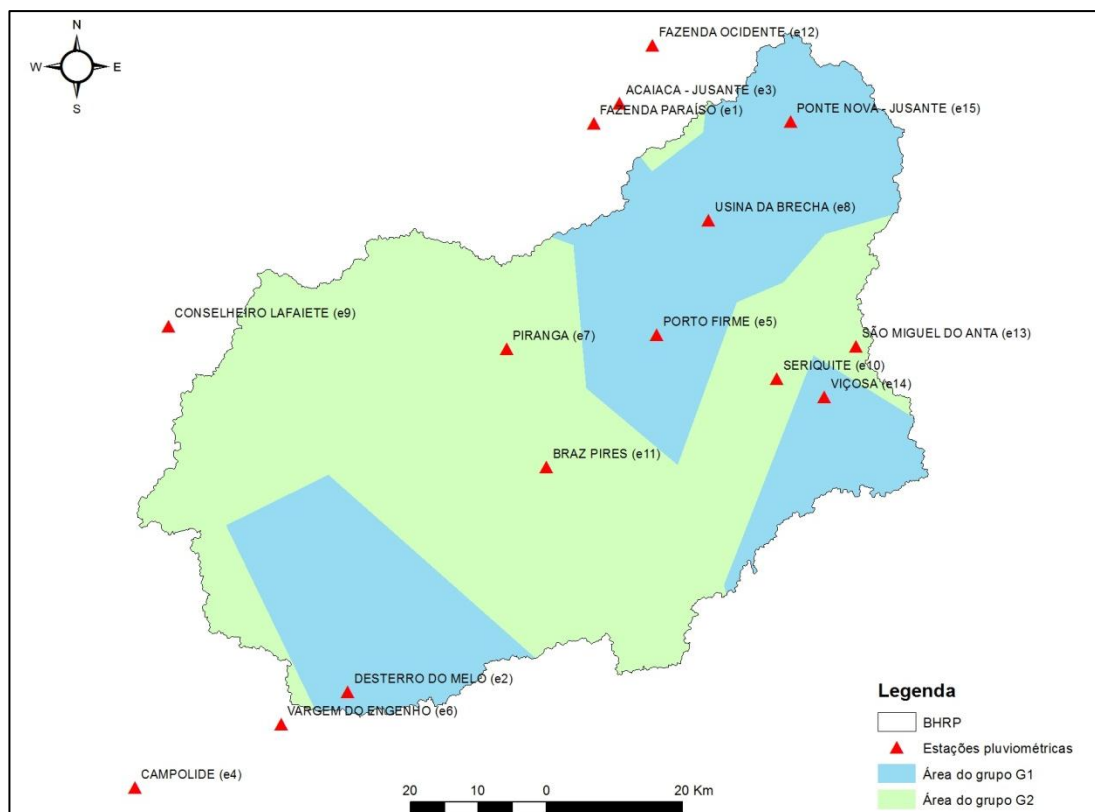


Figura (II) 9 – Áreas de representatividade geométrica dos grupos de estações *G1* (em azul) e *G2* (em verde), obtidas pela análise agrupamento de modelos.

A etapa seguinte consistiu da união dos dados da série histórica das estações de cada grupo, para geração de dois novos conjuntos de dados, referentes a *G1* e *G2*, os quais foram utilizados para novo procedimento de ajuste do modelo Bayesiano hierárquico proposto, conforme Equação 10.

4.3 Modelos de Representatividade Regional

Com subsídios da análise de agrupamento baseado em modelos que permitiu produzir novos conjuntos de dados, foram ajustados dois Modelos de Representatividade Regional (MRR), um para cada grupo de estações, *G1* e *G2*. Apesar da estrutura dos MRR ser a mesma do modelo hierárquico aplicado para ajuste dos dados das estações individuais (Equação 10), a característica diferenciada neste novo procedimento de ajuste foi quanto a dimensão do vetor Y , que agora contempla o número de anos da série histórica multiplicado pelo número de estações

pertencentes ao grupo. O resultado proporcionou aumento expressivo no tamanho da amostra, ou seja, o valor n . Durante análise pretérita individualizada, n era igual a 37 para cada estação, e na atual, seu valor foi de 259 para $G1$ e 293 para $G2$. Os valores das *prioris* não informativas foram mantidos.

Os testes diagnósticos de convergência de BGR (Tabela (II) 1) e RL (Tabela (II) 2) foram aplicados para avaliar os parâmetros *locação*, *escala* e *forma* da cada MRR, $G1$ e $G2$. Como pode ser verificado ao analisar os números dos testes de BGR e RL, em ambos os casos, todas as cadeias atingiram a condição de convergência e, portanto seus parâmetros e inferências de valores alcançaram a condição de distribuição estacionária.

Tabela (II) 1 – Teste diagnóstico de convergência de BGR aplicado às cadeias dos modelos dos grupos de estações, $G1$ e $G2$.

G1					G2				
Parâm.	CSRFB		PSRF	MPSRF	Parâm.	CSRFB		PSRF	MPSRF
	Estimado	Q.(0.975)				Estimado	Q.(0.975)		
η	1,000	1,000	1,000		η	1,000	1,000	1,000	
μ	0,999	1,000	0,999	1,000	μ	1,000	1,000	1,000	1,000
σ	0,999	0,999	0,999		σ	1,000	1,000	1,000	

Fonte: o autor.

Chama-se a atenção para uma particularidade nos resultados do teste RL (Tabela (II) 2), onde percebe-se pelos valores da medida *Total*, que foi necessário menor número de iterações para alcance da condição de convergência dos parâmetros *forma* e *locação* no modelo $G1$, cujo tamanho da amostra foi menor do que $G2$, contudo, em relação ao parâmetro *escala*, o inverso foi verificado.

Tabela (II) 2 – Teste diagnóstico de convergência de RL aplicado à cadeia unificada dos modelos dos grupos de estações, $G1$ e $G2$.

Parâm.	Thin (G1 e G2)	Lower Bound (G1 e G2)	G1			G2		
			Burn-in	Total	Dep. Factor	Burn-in	Total	Dep. Factor
η	1	3746	2	3727	0,994	2	3796	1,013
μ	1	3746	1	3774	1,007	2	3862	1,030
σ	1	3746	1	3822	1,020	2	3789	1,011

Fonte: o autor.

Os valores apresentados na Tabela (II) 3 são referentes aos parâmetros de média a *posteriori*, limites inferior e superior do intervalo de *HPD* dos parâmetros dos modelos dos dois grupos de estações, *G1* e *G2*. Verifica-se que houve redução da faixa de variação admissível para os valores dos parâmetros, em relação àqueles dos modelos individualizados das estações, provavelmente devido ao aumento da amostra. Os valores do parâmetro *forma* foram negativos e muito pequenos, contudo estão localizados em faixas de variação que admitem o valor zero, o que confere classificação equivalente à distribuição de Gumbel – Tipo I, apropriada para valores extremos máximos. Em relação ao parâmetro *locação*, o MRR do grupo *G1* apresentou valores significativamente superiores ao do grupo *G2*, indicando alturas precipitadas diárias maiores nessa região do médio e baixo rio Piranga do que nas suas áreas de nascente.

Tabela (II) 3 – Valores das médias a posteriori (MP), limite inferior (Li) e limite superior (Ls) do intervalo *HPD* ao nível de significância de 0.05 dos parâmetros *forma* (η), *locação* (μ) e *escala* (σ) dos modelos ajustados para os grupos de estações, *G1* e *G2*.

Parâmetros	G1			G2		
	MP	Li	Ls	MP	Li	Ls
η	-0,01679	-0,1026	0,07167	-0,03771	-0,08718	0,01473
μ	73,75	71,15	76,43	66,34	64,00	68,71
σ	19,34	17,41	21,25	19,20	17,61	20,86

Fonte: o autor.

4.4 Validação dos Modelos

De posse dos valores dos parâmetros obtidos para os grupos *G1* e *G2* (Tabela (II) 3), realizou-se procedimento de inferência dos quantis extremos (Equação 6) para os períodos de retorno de 5, 10, 20, 30, 40, 50, 70, 100, 200 e 500 anos, através da substituição dos valores de *forma*, *locação* e *escala*, por aqueles determinados pela *MP*, *Li* e *Ls*, respectivos. Utilizou-se os dados de série histórica as estações que formaram os grupos para verificação de aderência dos modelos aos dados observados.

Para avaliação do MRR do grupo *G1*, empregou-se as séries históricas das estações que o compuseram (*e1*, *e2*, *e4*, *e5*, *e8*, *e14*, *e15*), sob o critério de que deve ser capaz de fazer estimativas de alturas precipitadas para estações daquela região com boa acurácia, devido ao seu caráter espacial,. No gráfico da Figura (II) 10 pode-se verificar o desempenho apresentado pelo modelo do grupo *G1* por comparação dos seus valores preditos com os determinados pela análise de frequência a partir dos registros históricos das estações. As alturas máximas precipitadas associadas aos respectivos períodos de retorno, calculadas pela Análise de Frequência dos dados de série histórica de cada estação em análise, encontram-se disponíveis no Apêndice (II) D. À exceção de alguns registros das estações *e4*, *e8* e *e14*, que se posicionaram além das estimativas de *Li* e *Ls* do intervalo *HPD* (Figura (II) 10), todos os demais se localizaram no mesmo. Em relação a *e8*, apenas o valor correspondente ao período de retorno de 10 anos ficou abaixo de *Li*, os outros acompanharam muito próximos os valores de *MP*. Entretanto, *e4* e *e14* não apresentaram bom ajuste, o primeiro para os períodos de retorno de 40, 50 e 70 anos e o segundo para os períodos de retorno de 20, 30 e 40 anos. Em ambos os casos, os valores registrados superaram as estimativas de *Ls*, ou seja, o modelo subestimou as precipitações nas regiões de Campolide (*e4*) e Viçosa (*e14*). As estações *e1* e *e15* oscilaram entre a *MP* e o *Li*, *e2* e *e5* apresentaram ótima aderência aos valores da *MP*.

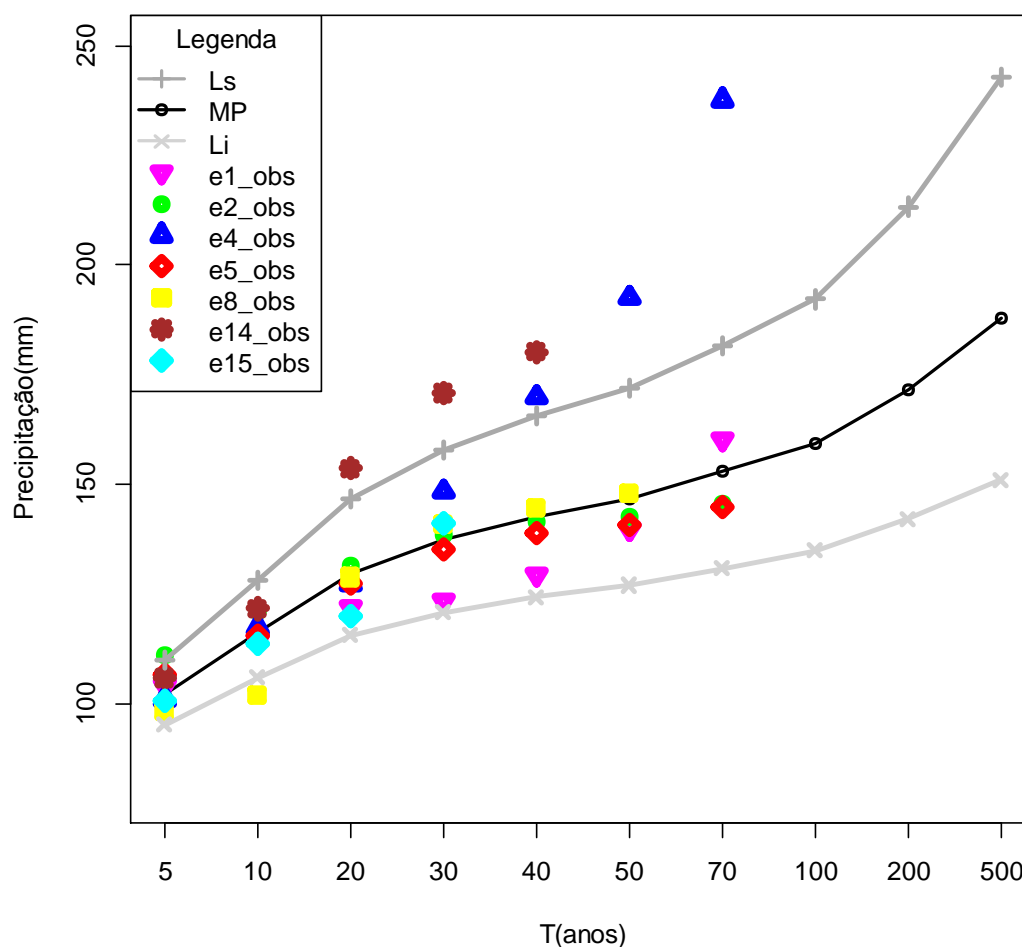


Figura (II) 10 – Inferência preditiva das alturas das precipitações pluviiais máximas diárias e seus respectivos períodos de retorno determinada a partir dos valores da média a *posteriori*, limite inferior e superior para o modelo *G1*.

Ao se analisar em conjunto, os valores de alturas precipitadas diárias máximas estimadas pelo modelo do grupo *G1* em comparação aos respectivos valores calculados pela análise de frequência para cada estação (Figura (II) 10), constata-se que aqueles determinados pela *MP* apresentaram ótima aderência para a maioria das estações na maior parte dos períodos de retorno.

Em relação ao modelo do grupo *G2*, cujos resultados obtidos estão exibidos no gráfico da Figura (II) 11, observa-se que das oito estações empregadas na sua determinação, apenas *e11* apareceu completamente fora, com todos os seus valores abaixo do intervalo *HPD*, portanto superestimados por *Li*. As estações *e9* e *e13* tiveram seus menores períodos de retorno, de 5 e 10 anos superestimados por *G2* e, no caso da primeira, também foram os períodos de retorno de 20 e 30 anos. Entretanto, em situação oposta, encontrou-se a estação *e10*, que apresentou seus

maiores períodos de retorno, de 30 e 40 anos, subestimados por L_s do intervalo HPD . Acompanhando $e10$, $e6$ teve seu maior período de retorno, de 70 anos, subestimado por L_s . Por outro lado, $e7$ foi subestimando em relação ao período de retorno de 10 anos. Todos os períodos de retorno das demais estações oscilaram entre os valores de L_i e L_s do intervalo HPD determinado pelo modelo do grupo $G2$.

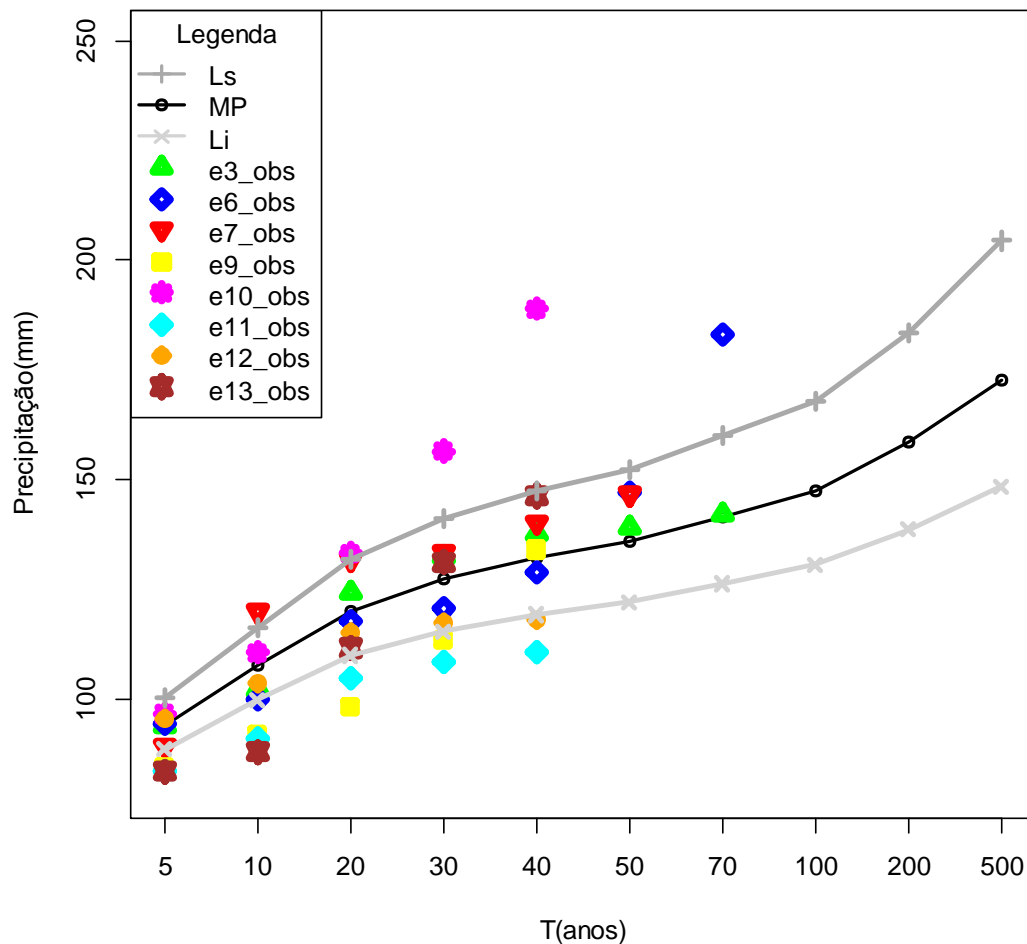


Figura (II) 11 – Inferência preditiva das alturas das precipitações pluviiais máximas diárias e seus respectivos períodos de retorno obtida a partir dos valores da média *a posteriori*, limite inferior e superior para o modelo $G1$

Apesar dos valores obtidos pela análise de frequência da série histórica das estações do grupo $G2$ não terem seguido majoritariamente uma linha de estimativas específica, calculada a partir dos parâmetros de L_i , L_s ou MP , considera-se que o resultado foi satisfatório, pois esses valores apresentaram pouca variação, ou seja, o intervalo HPD de $G2$ foi pequeno para cada período de retorno (Figura (II) 11). Em

outras palavras, acredita-se que a adoção dos parâmetros da *MP* do modelo do grupo *G2* para obtenção de estimativas para sua região representativa fornecerá resultados adequados, devido a pouca variação do intervalo *HPD* e da maior parte dos valores avaliados se enquadrar nessa faixa.

Ambos os modelos do grupo de estações *G1* e *G2* apresentaram resultados satisfatórios e podem ser empregados para obter boas estimativas a partir de suas *MPs* para cada respectiva região representativa, exceto em relação à localização específica das estações *e4* e *e14* no caso de *G1* e das estações *e9*, *e10* e *e11* no caso de *G2*.

5 Conclusões e Recomendações

É possível concluir que os objetivos foram alcançados satisfatoriamente uma vez que os resultados dos modelos hierárquicos da distribuição *GEV* ajustados por inferência Bayesiana em associação com análise de agrupamento baseada em modelos para determinação da altura máxima precipitada e seu respectivo período de retorno realmente apresentaram representatividade regional na bacia hidrográfica do rio Piranga. Além disso, pode-se concluir que:

- a análise de agrupamento baseado em modelos se mostrou eficiente uma vez que os grupos de estações identificados pelo método têm significado coerente na realidade;
- o método de análise Bayesiana através de modelos hierárquicos foi adequado ao ajuste da distribuição *GEV*, cujos parâmetros são geralmente complexos de se determinar;
- o modelo hierárquico mostrou-se uma forma recomendável de introdução de informação *a priori* de maneira “semi-informativa”, isto porque a região de variação dos parâmetros não foi completamente vaga (não informativa), porém não foi também efetivamente informativa;
- estimativas de precipitações diárias máximas realizadas a partir dos parâmetros dos dois MRR obtidos pelo método, *G1* e *G2*, nos seus respectivos intervalos de *HPD*, foram adequadas para os maiores períodos de

retorno na região da bacia hidrográfica do rio Piranga, pois compreendem a maior parte desses valores quando obtidos pela análise de frequência das series históricas das estações analisadas;

- embora os resultados da avaliação corroborem a recomendação do vetor paramétrico obtido pela *MP* para a maior parte das situações, para casos de aplicações que considerem o risco de vidas humanas, recomenda-se avaliar a possibilidade de adotar os valores obtidos com modelo parametrizado entre a *MP* e o *Ls* no intervalo *HPD*, por questões de segurança;
- apesar do método proposto poder ser replicado, teoricamente, em qualquer região, espera-se que seu desempenho seja melhor em bacias hidrográficas devido suas características hidrológicas e climatológicas.

Recomenda-se que para trabalhos futuros seja feita a avaliação de desempenho de diferentes configurações quanto à hierarquia do modelo utilizada, ou mesmo em relação à escolha de outros tipos de distribuições de valores extremos. Também pode ser avaliada qual seria a limitação espacial ou configuração adequada quanto ao número de estações para compor cada grupo identificado pela análise de agrupamento, e ainda, se há um número mínimo de anos da série histórica para que o ajuste tenha efeito. Além disso, pode-se investigar outras soluções para incorporação de conhecimento através das distribuições *a priori*, bem como avaliar o desempenho de diferentes *prioris* no modelo hierárquico, ambos em abordagem Bayesiana.

6 Referências

BACRO, J.-N.; CHAUCHE, A. Incertitude d'estimation des pluies extrêmes du pourtour méditerranéen: illustration par les données de Marseille. **hydrological Sciences Journal**, 51, n. 3, 2006. 389-405.

BANERJEE, S.; ROSENFELD, A. Model-Based Cluster Analysis. **Pattern Recognition**, 26, n. 6, 1993. 963-974.

BEGUERÍA, S. et al. Assessing trends in extreme precipitation events intensity and magnitude using non-stationary peaks-over-threshold analysis: a case study in northeast Spain from 1930 to 2006. **International Journal of Climatology**, 31, 2010. 2102-2114.

BEIJO, L. A.; MUNIZ, J. A.; CASTRO NETO, P. Tempo de retorno das precipitações máximas em Lavras (MG) pela distribuição de valores extremos do tipo I. **Ciência e Agrotecnologia**, Lavras, 29, n. 3, 2005. 657-667.

BEIJO, L. A.; VIVANCO, M. J. F.; MUNIZ, J. A. Análise Bayesiana no Estudo do Tempo de Retorno das Precipitações Pluviais Máximas em Jaboticabal (SP). **Ciência e Agrotecnologia**, Lavras, 33, n. 1, jan./fev. 2009. 261-270.

COLES, S. G.; TAWN, J. A. A Bayesian Analysis of Extreme Rainfall Data. **Applied Statistics**, 45, n. 4, 1996. 463-478.

CONSÓRCIO ECOPLAN-LUME. **Plano Integrado de Recursos Hídricos da Bacia Hidrográfica do Rio Doce e Planos de Ações para as Unidades de Planejamento e Gestão de Recursos Hídricos no Âmbito da Bacia do Rio Doce**. [S.l.], p. 478. 2010.

CRISCI, A. et al. Extreme rainfall in a changing climate: regional analysis and hydrological implications in Tuscany. **Hydrological Process**, 16, 2002. 1261-1274.
FRALEY, C.; RAFTERY, A. E. How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis. **The Computer Journal**, 41, n. 8, 1998. 578-588.

GAIONI, E.; DEY, D.; RUGGERI, F. Bayesian modeling of flash floods using generalized extreme value distribution with prior elicitation. **Chilean Journal of Statistics**, 1, n. 1, April 2010. 75-90.

GELMAN, A. et al. **Bayesian Data Analysis**. London: Chapman & Hall, 1995. 526 p. ISBN ISBN 0-412-03991-5.

HERRINGTON, T. O. et al. **A comparison of methods used to calculate extreme water levels**. Solutions to Coastal Desasters Congress. [S.l.]: ASCE. 2008. p. 198-209.

KOUTSOYIANNIS, D. Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. **Hydrological Sciences Journal**, 49, n. 4, 2004. 575-590.

MICHELE, C. D.; SALVADORI, G. Some hydrological applications of small sample estimators of generalized pareto and extreme value distributions. **Journal of Hydrology**, n. 301, 2005. 37-53.

NASCIMENTO, M. et al. Bayesian model-based clustering of temporal gene expression using autoregressive panel data approach. **Bioinformatics**, Oxford, 4, June 2012. 1 - 5.

NGUYEN, V.-T.-V.; TAO, D.; BOURQUE, A. **On selection of probability distributions for representing annual extreme rainfall series**. Global Solutions for Urban Drainage. [S.l.]: [s.n.]. 2002.

PARK, H. W.; SOHN, H. **Parameter estimation of the generalized extreme value distribution for structural health monitoring**. Structures Congress 2005. [S.l.]: ASCE. 2005.

PARK, J.-S. et al. Changes in the extreme daily rainfall in South Korea. **International Journal of Climatology**, 31, 2011. 2290-2299.

R DEVELOPMENT CORE TEAM. **A Language and Environment for Statistical Computing**. Vienna: [s.n.], 2011.

RAGAN, P.; MANUEL, L. **Statistical extrapolation methods for estimating wind turbine extreme loads**. 45th Aerospace Sciences Meeting and Exhibit. Reno: [s.n.]. 2007.

SILVA, F. F. et al. Comparação bayesiana de modelos de previsão de diferenças esperadas nas progênies no melhoramento genético de gado Nelore. **Pesquisa Agropecuária Brasileira**, Brasília, 43, 2008. 37-45.

SILVA, J. P.; LEANDRO, R. A. Uma abordagem bayesiana para o mapeamento de QTLs utilizando o método MCMC com saltos reversíveis. **Ciência e Agrotecnologia**, Lavras, 33, 2009. 1061-1070.

SINGH, W. **Robustness of three hierarchical agglomerative clustering techniques for ecological data**. University of Iceland. [S.l.], p. 100. 2008.

SMITH, B. J. **Bayesian Output Analysis Program (BOA) Version 1.1 User's Manual**. [S.l.], p. 43. 2005. Disponível em: Acesso em: jun. 2012.

SUTCLIFFE, J. V. **Methods of flood estimation: a guide to the flood studies report**. Natural Environment Research Council. [S.l.], p. 50. 1978. NERC Open Access Research Archive (NORA). Disponível em: Acesso em jun. 2012.

III. Classificação de imagens do sensoriamento remoto a partir de metodologias de representação da paisagem.

1 Introdução

Um mapa recente de uso do solo e cobertura vegetal é uma ferramenta útil tanto para fins de diagnóstico como de planejamento da ocupação, pois permite verificar em um dado instante, feições do desenvolvimento regional para com isso, avaliar suas características e, se caso for, adequá-lo aos seus modelos preconcebidos.

Com o advento das imagens multiespectrais obtidas pelo sensoriamento remoto orbital, a produção de mapas temáticos referentes a grandes extensões terrestres se tornou possível devido à possibilidade de aplicar métodos de reconhecimento de padrões que, a partir da identificação de tipologias de cobertura da paisagem em amostras, extraem parâmetros que alimentam modelos utilizados para rotular toda a área imageada. Esse procedimento é comumente conhecido como classificação de imagens (MATHER, 2004).

O processo de classificação de imagens tem passado por transformações em busca de novos procedimentos para melhorar seu desempenho, como o emprego de séries temporais de índices espectrais (i.e. NDVI, EVI) para classificação de culturas agrícolas relativas ao uso e cobertura do solo (WARDLOW, EGBERT e KASTENS, 2007), ou a integração de sistemas especialistas com classificadores convencionais para melhorar a acurácia da classificação (MASOCHA e SKIDMORE, 2011). Contudo, esses resultados têm alcançado valores de acurácia da ordem de 90%. Talvez uma das mais relevantes reflexões que se deva fazer em torno desta questão do patamar de exatidão temática seja em relação à escala e ao propósito daquela determinada aplicação do produto cartográfico obtido da classificação (MATHER, 1999).

A imagem digital orbital é o registro quase instantâneo da superfície da Terra, que se encontra em constante alteração, seja pelos processos humanos de interação com o meio, seja pela dinâmica natural dos processos de transformação da paisagem. Logo, o caráter temporal das imagens deve ser considerado ao definir que objetos ou classes informacionais se pretende distinguir durante o processo de classificação.

Caso contrário, confusões entre as classes informacionais podem ocorrer como parte do processo. Muitas vezes isto acontece devido à discrepância entre o número de classes informacionais que a paisagem apresenta e são registradas na imagem, e aquelas que se pretende identificar (ASSIS, 2008). Além disso, no processo de classificação a generalização de uma classe informacional através da rotulação da mesma por toda imagem é realizada a partir de parâmetros extraídos de amostras dessa classe. Tal procedimento pode ocasionar confusão ao realizar a generalização quanto maior for a área de interesse.

O emprego do teorema Bayesiano por diferentes abordagens aplicadas ao processo de classificação de imagens tem apresentado resultados superiores aos métodos tradicionais utilizados como medida de comparação (PINZ, 1996; ABKAR, MOHAMMED e MULDER, 2000; ZHU e TATEISHI, 2006; MALPICA, ALONSO e SANZ, 2007). Assis, Vieira e Silva (2010) conseguiram aumentar a exatidão temática de um mapa produzido pela classificação por método que faz uso de informações contextuais através de inferência Bayesiana. Diante do exposto, torna-se relevante fornecer subsídios ao processo de classificação, tais como informações externas, de modo a reduzir o efeito de 'ruído' provocado pelas classes informacionais (categorias temáticas) não contempladas na etapa de definição, mas que são inerentes à constituição da paisagem.

1.1 Objetivo

Avaliar duas metodologias empregadas como mecanismos de descrição da paisagem, obtidos pela incorporação de informações contextuais, para uso em processo de classificação de imagens do sensoriamento remoto com intuito de produzir um mapa temático de uso e ocupação do solo.

1.2 Questão de pesquisa

A informação contextual no âmbito desta pesquisa é representada através do conhecimento especialista, expresso por sua expectativa acerca da ocorrência de condições que definem a paisagem em determinado local ou região. Por tal princípio é razoável refletir sobre maneiras de se viabilizar formalmente a expressão desse conhecimento em âmbito espacial por meio de cenários.

Uma vez que a utilização de informações contextuais para representação da paisagem no processo de classificação de imagens pode melhorar seu resultado (ASSIS, VIEIRA e SILVA, 2010), espera avaliar o desempenho de duas metodologias particulares e distintas para este propósito, trata-se da Regressão Logística Binomial (RLB) e da Avaliação Multicritérios (AMC), a primeira é estatística enquanto a última não. A questão de pesquisa que se faz é: qual das metodologias, RLB ou AMC é mais adequada para representação da paisagem e, portanto, oferece melhores subsídios para classificação de imagens?

2 Revisão de literatura

2.1 Regressão Logística

A regressão logística é geralmente adequada para descrever e testar hipóteses acerca das relações entre uma variável resposta do tipo categórica e uma ou mais variáveis explicativas do tipo contínuas, porém não necessariamente (PENG, LEE e INGERSOLL, 2002), no qual o caso mais simples refere-se ao binomial, que admite apenas dois resultados possíveis. O objetivo da regressão logística é prever a categoria resultante corretamente para casos particulares através do modelo mais parcimonioso (SAHA, 2011). Seu conceito matemático central é o *logit*, logaritmo natural da razão de chances (*odds ratio*) (PENG, LEE e INGERSOLL, 2002). As aplicações da regressão logística são observadas em diversas áreas do conhecimento, desde biometria (LARSEN et al., 2000), ciências sociais (PENG, LEE e

INGERSOLL, 2002), educação (SAHA, 2011), proposição de cenários para simulação hidrológica (OÑATE-VALDIVIESO e SENDRA, 2010) e até a classificação de imagens do sensoriamento remoto (ASSIS, 2008).

Aplicado ao Sensoriamento Remoto, a Regressão Logística Binomial (RLB) pode ser entendida como um método estatístico capaz de descrever a relação entre variáveis independentes de valores contínuos (i.e., variáveis de ambiente que constituem as informações contextuais) e uma variável dependente binária (i.e., uma categoria das amostras de treinamento). Essa relação pode ser expressa matematicamente pela Equação 1.

$$P(\theta = 1|X) = \frac{\exp(\sum_{k=0}^k B_k X_k)}{1 + \exp(\sum_{k=0}^k B_k X_k)} \quad (1)$$

onde P é a probabilidade da variável dependente θ ser 1; X é o vetor de variáveis independentes tal que $X = (x_0, x_1, x_2, \dots, x_k)$; e B é vetor dos coeficientes estimados tal que $B = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)$ e k é o número de variáveis independentes. Para linearizar a Equação 1 e eliminar os valores de probabilidade nos limites próximos de 0 e 1, emprega-se a transformação logística, obtendo a Equação 2.

$$P(\theta)' = \log[P(\theta = 1|X)] \quad (2)$$

Após a transformação logística, $P(\theta)'$ pode assumir qualquer valor no intervalo de “-” a “+” infinito.

Como $\text{logit}(P) = \log[P/(1 - P)] = \exp \theta$, aplica-se a transformação *logit* obtendo o modelo de regressão linear padrão conforme a Equação 3.

$$\text{logit}[P/(1 - P)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + e \quad (3)$$

em que o termo e representa o erro do modelo. Uma vez determinado o vetor dos coeficientes β , aplica-se o modelo apresentado na Equação 1 para obter os valores de probabilidade.

Para estimar os coeficientes do modelo (β), é necessário empregar algum método de estimação como o da Máxima Verossimilhança (MV), pois usualmente apresenta melhores resultados que o dos Mínimos Quadrados (MMQ).

Como a Função Densidade de Probabilidade (FDP) assumida é a Distribuição Binomial ($\theta \sim \text{Bin}(n, y)$), sua MV é dada pela Equação 4.

$$L(\theta) \propto \prod_{i=1}^n \theta_i^{y_i} (1 - \theta)^{(1-y_i)} \quad (4)$$

onde $L(\theta)$ é a verossimilhança de θ ; θ_i é o valor predito da variável dependente para a amostra i ; y_i é o valor observado da variável dependente para a amostra i . Note que apesar da Distribuição Binomial¹ ser definida em função de n e y , na Equação 4 foi apresentada apenas em função de y , isto porque n é fixo e unitário. Definição que se justifica por características do Domínio da análise (i.e., o pixel), onde apenas uma categoria temática pode ocorrer, logo, o espaço de realização de eventos é unitário. O espaço amostral, por sua vez, é a própria classe informacional. Desse modo, nota-se que os resultados da RLB são expressos em termos de probabilidade.

Modelos obtidos por Regressão Logística podem ser avaliados por vários métodos, tais como LogLikelihood (Equação 5), Pseudo R^2 (Equação 6) e a estatística ROC².

$$\text{LogLikelihood} = -2 \ln[L(\hat{\theta})] \quad (5)$$

no qual $[L(\hat{\theta})]$ é o valor da função de verossimilhança para o modelo ajustado completo.

$$\text{Pseudo}R^2 = 1 - \frac{\ln[L(\hat{\theta})]}{\ln(L0)} \quad (6)$$

onde $L0$ é o valor da função de verossimilhança assumindo todos os seus coeficientes como 0, exceto o intercepto.

Para comparação entre os modelos, pode-se utilizar a medida AIC (*Akaike Information Criterion*), que emprega o princípio da parcimônia, penalizando o modelo com maior número de parâmetros. Sua medida pode ser determinada através da Equação 7.

¹ A forma proporcional da Distribuição Binomial é dada por $p(y|\theta, n) \propto \theta^y (1 - \theta)^{(n-y)}$

² *Relative Operating Characteristic* é assunto abordado no tópico de avaliação dos métodos RLB e MCE na seção 2.3.

$$AIC = -2 \ln[L(\hat{\theta})] + 2m \quad (7)$$

onde m é número de parâmetros. Quanto menor for o valor de AIC, melhor o modelo.

2.2 Tomada de Decisão por Agregação Multicritérios

Inicialmente devem ser definidos os termos decisão e critério para o entendimento apropriado do texto. O termo decisão é definido como uma escolha entre alternativas disponíveis, e o termo critério refere-se a uma evidência que pode ser medida e avaliada, responsável por fundamentar uma decisão (EASTMAN et al., 1995). Em um processo de tomada de decisão com base na agregação de critérios múltiplos, a estrutura de combinação desses critérios é central na determinação da relação entre eles, onde, por um lado espera-se obter uma solução em que todos os critérios sejam satisfeitos e pelo outro, a solução de que apenas um deles já satisfaz (YAGER, 1988). Essa condição de oposição diz respeito respectivamente aos operadores *and* e *or*. Como alternativa, o operador de agregação conhecido por Média Ponderada Ordenada (*Ordered Weighted Averaging – OWA*) foi apresentado (YAGER, 1988). Às vezes mencionada como uma opção de análise por Avaliação Multi-Critérios (AMC) ou *MCE (MultiCriteria Evaluation)* distinta, a Combinação Linear Ponderada, ou *Weighted Linear Combination – WLC* é, na verdade, um caso particular de *OWA* (EASTMAN, 2009).

O operador *OWA* não se trata de uma ferramenta estatística, é a técnica fundamental que governa o método de Avaliação Multicritérios, ou seja, a regra de decisão. Devido à sua extensa possibilidade de aplicações, a técnica de operadores *OWA* tem sido discutida em muitos artigos (FULLÉR, 1996).

A definição (CHICLANA et al., 2007) de um operador *OWA* de dimensão n é expressa por uma função $F: R^n \rightarrow R$ em que $R = [0,1]$, com um vetor de ponderação $W = (w_1, \dots, w_n)$ associado tal que satisfaça as condições $w_i \in [0,1]$ e $\sum_{i=1}^n w_i = 1$, e deve agregar um conjunto de valores $\{p_1, \dots, p_n\}$ conforme a Equação 8.

$$F(p_1, \dots, p_n) = \sum_{i=1}^n w_i \cdot p_{\sigma(i)} \quad (8)$$

onde $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ é uma permutação tal que $p_{\sigma(i)} \geq p_{\sigma(i+1)}, \forall i = 1, \dots, n-1$, ou seja, $p_{\sigma(i)}$ é o i ésimo maior valor do conjunto $\{p_1, \dots, p_n\}$, que representa o critério ordenado.

Usualmente, os valores dos pesos (W) dos critérios são determinados através do Processo Hierárquico Analítico (*Analytical Hierarchy Process - AHP*), pois seu emprego para análises espaciais é consolidado em ambiente de Sistemas de Informação Geográfica (SIG) (EASTMAN et al., 1995; VIEIRA, ASSIS e BECKER, 2010). O *AHP* é uma teoria geral de medição utilizada para produzir uma escala de comparação pareada, discreta ou contínua (SAATY, 1987). Baseia-se em uma matriz quadrada recíproca (matriz A) onde os n critérios são dispostos nas n linhas e colunas. Os valores de A referem-se à medida subjetiva de comparação entre critérios em análise através da importância relativa entre eles. Desse modo, um valor a_{ij} qualquer representa a importância relativa do critério na linha i em relação a outro da coluna j . Conseqüentemente tem-se que $a_{ij} = 1/a_{ji}$ ou $a_{ij} = 1$ quando $i = j$. Os pesos são obtidos pelo cálculo do principal *autovetor* da matriz A . Entretanto, como os valores de A são medidas subjetivas comparativas, há a necessidade de verificar se os julgamentos de importância relativa entre os critérios são coerentes. Uma forma para se verificar isso é através da Razão de Consistência (*Consistency Ratio - CR*), uma medida que indica a probabilidade de que as avaliações da matriz A foram geradas aleatoriamente (EASTMAN et al., 1995). Saaty (1987) sugere que se deve considerar 10% de inconsistência na matriz A porque a prioridade de consistência para se obter um explicação coerente de um conjunto de fatos deve divergir, por uma ordem de magnitude, da prioridade de inconsistência, que é um erro na medida de consistência. Portanto, nos casos onde o CR exceda 10%, deve-se reavaliar os julgamentos. Saaty (1987) argumenta ainda que o valor CR não deve ser muito pequeno, como por exemplo 1%, porque a inconsistência por si só é importante uma vez que sem ela, novos conhecimentos que alteram a ordem das preferências, simplesmente não seriam admitidos.

No ambiente dos SIG, *MCE* é uma das técnicas disponíveis aplicadas para análise espacial de problemas que envolvem temas múltiplos, relacionados a um determinado assunto (VIEIRA, ASSIS e BECKER, 2010). O método de Avaliação Multicritérios é estabelecido por uma sequência de processos, com possibilidade de revisão em várias etapas. Inicialmente são selecionados os temas-critério que participam da análise e podem se tratar de grandezas distintas (i.e., mapas de distâncias, declividades, temperatura, uso do solo, entre outros). Referem-se, no âmbito desta proposta, às variáveis de ambiente que definem as informações contextuais. Como o objetivo é viabilizar a comparação entre critérios que se encontram em diferentes escalas de valores, deve-se padronizá-los em uma mesma escala. Existem várias abordagens para esse propósito, de modo que, o que influencia a escolha por uma delas seja a característica inerente ao próprio dado que se pretende padronizar. O uso de funções de transformação oferece flexibilidade para a maior parte das situações. No ambiente do SIG Idrisi são funções muito utilizadas para este propósito, as *fuzzy*: linear ou sigmoide, dos tipos monotônica crescente ou decrescente, ou a do tipo complexa³. Funções *fuzzy* assinalam um grau de pertinência variável entre zero e um para cada objeto da classe de objetos de um conjunto *fuzzy*, caracterizando-o (ZADEH, 1965).

A função sigmoide adota o modelo de transformação de valores dado pela Equação 9 e a função linear é apresentada pela Equação 10 caso seja monotônica crescente.

$$f(u) = \cos^2 \alpha \quad (9)$$

$$f(u) = \frac{(x - x_a)}{(x_b - x_a)} \quad (10)$$

no qual $\alpha = \frac{\pi}{2} [(x - x_c)/(x_d - x_c)]$ no caso de pertinência ser monotonicamente decrescente, ou $\alpha = \frac{\pi}{2} [1 - (x - x_a)/(x_b - x_a)]$ caso seja monotonicamente crescente e $u = 1$ quando $x < x_c$ ou $x > x_b$, respectivamente; u é o valor mapeado x na nova escala $[0, 1]$, x é o valor que se quer escalonar; x_a , x_b , x_c e x_d definem os limites do intervalo. A função linear monotônica decrescente é dada pela Equação 11.

³ Pontos definidos pelo usuário, linearmente interpolados.

$$f(u) = -1 \left[\frac{(x - x_d)}{(x_d - x_c)} \right] \quad (11)$$

Essa operação é realizada pixel a pixel por toda a imagem.

Diante do exposto tem-se que os resultados da AMC são expressos em termos de graus de adequabilidade, no intervalo de [0,1].

2.3 Avaliação pela estatística ROC

ROC é uma técnica de análise originária da teoria de detecção de sinais (SWETS, 1988) amplamente utilizada em diversas áreas, aplicável a qualquer modelo que faz previsões de uma categoria homogênea em uma célula da imagem (PONTIUS JR. e SCHNEIDER, 2001). Na sua forma gráfica, a técnica também conhecida pelo termo “*Receiver Operating Characteristics*” é utilizada para visualizar, organizar e selecionar classificadores com base no seu desempenho (FAWCETT, 2006).

A estatística *ROC* é determinada pela métrica de comparação de falsos e verdadeiros positivos dos percentis de um mapa de adequabilidade. Na prática, verdadeiro positivo significa que houve coincidência da classe de interesse na imagem de referência e na imagem avaliada. Falso positivo significa que a classe de interesse identificada na imagem avaliada não corresponde à classe de interesse na imagem de referência. Mais detalhes sobre a análise *ROC* podem ser consultados em Fawcett (2006) e sua aplicação em Pontius Jr. e Schneider (2001).

A estatística *ROC* é obtida determinando a área sob a curva conforme a Equação 12.

$$ROC = \sum_{i=1}^n (x_{i+1} - x_i)(y_i + y_{i+1} - y_i/2) \quad (12)$$

onde x_i é taxa de falsos positivos para o cenário i , y_i é a taxa de verdadeiros positivos para o cenário i , e n o número de grupos de adequabilidade (PONTIUS JR. e SCHNEIDER, 2001). Os valores de *ROC* variam entre 1 e 0.5 e representam respectivamente, coincidência total ou simplesmente ao acaso.

A estatística *ROC* reúne características necessárias para avaliar os resultados dos dois métodos propostos. Ambos os métodos de análise, da RLB e da AMC, permitem utilizar dados contínuos que pertençam ao Domínio dos números Reais e seus resultados são expressos no intervalo de [0,1], em termos de probabilidade e adequabilidade, respectivamente.

O método que apresentar melhor desempenho na avaliação deve ser utilizado como gerador do cenário para classificação.

3 Métodos

3.1 Base de dados acessória

A incorporação de dados auxiliares externos para definir a contextura de informações sobre um determinado fenômeno – ou classe informacional, além das bandas espectrais de uma imagem de satélite, é fortemente dependente da disponibilidade de uma base de dados espaciais acessória. Essa é constituída de mapas de variáveis físicas que devem auxiliar a caracterização das condições regionais propícias nas quais se verifica a predominância de ocorrência de algum tipo específico de uso ou cobertura do solo. Modelos digitais como os de elevação, declividades, morfologia do terreno são exemplos comuns desse tipo de base, bem como as imagens dos índices espectrais de vegetação como a NDVI (*Normalized Difference Vegetation Index*), entre outros. As principais fontes geradoras desses dados são o Modelo Digital de Elevação (MDE) e a imagem de satélite.

3.1.1 Modelo Digital de Elevação

Para gerar o MDE da área de estudo foram necessários os planos de informação referentes a hidrografia e altimetria, esta última discriminada em pontos cotados e

curvas de nível, obtidas através das cartas topográficas vetoriais do mapeamento sistemático e disponibilizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE). A escala dos mapas digitais é de 1:50.000 e sua edição data de 1972, cujas curvas de nível da altimetria apresentam equidistância vertical de 20m. Foram necessárias 18 cartas para abranger toda a área da bacia. Em seguida realizou-se o pré-processamento dos dados digitais obtidos, com a verificação da consistência de toda a base. As curvas de nível, quando apresentaram inconsistências, foram editadas manualmente em conformidade com a hidrografia mapeada. Esta última, por sua vez, foi orientada no sentido da direção de escoamento. Após a etapa de edição pôde-se produzir o MDE com resolução de 15 x 15m, através do módulo TOPOGRID disponível no software Arc/INFO. Posteriormente esse MDE foi reamostrado pelo algoritmo do *vizinho mais próximo* para a dimensão do pixel coincidente com a imagem de satélite utilizada.

Sabendo-se que o MDE é uma superfície produzida por interpolação a partir de informações de altitude disponíveis em determinados locais estrategicamente levantados para melhor caracterizar a topografia local; e ainda que seus valores são determinados por um estimador, conseqüentemente a incerteza dos mesmos é inerente ao processo, assim sendo o produto gerado pode apresentar tanto super quanto subestimativas. Ao se tratar da dinâmica dos processos relacionados ao escoamento superficial, as subestimativas são as mais problemáticas. Elas são observadas nos MDE como depressões espúrias, as quais são células cuja circunvizinhança apresenta valores de elevação superiores ao seu, o que impede o fluxo do escoamento superficial ao longo do seu percurso. Diante do exposto, procedeu-se a eliminação destas, permitindo a obtenção de um MDE isento de depressões espúrias. O MDE é gerador de diversas variáveis de ambiente topográfico, tais como: Modelo Digital de Declividades (MDD), Categorias Morfológicas do Terreno (CMT), modelo de direções da face do terreno (MDF).

3.1.2 Imagem de satélite

Para abranger toda área da BHRP utilizou-se duas cenas da plataforma orbital Landsat 5, sensor TM (*Thematic Mapper*), com resolução espacial de 30 metros,

ambas datadas de 26/08/2010, obtidas gratuitamente no repositório de imagens do Instituto Nacional de Pesquisas Espaciais (INPE). Como as cenas são do mesmo dia, são referentes a uma mesma órbita, a 217, variando apenas o ponto, 74 para região de jusante da bacia (maior parte) e 75 para a região de cabeceira.

As cenas foram mosaicadas e devido à necessidade de sobreposição da base, a imagem foi ortorretificada⁴, em seguida foi feito um recorte na imagem gerada em relação aos limites da bacia hidrográfica.

Para o processo de ortorretificação foram utilizadas como base de referência para amostragem de pontos de controle duas outras cenas correspondentes às de 2010, ou seja, sensor TM, órbita 217, pontos 74 e 75 (mesmo sensor, órbita e ponto), previamente ortorretificadas pela NASA, porém datadas de 2006 e 2007, e também disponibilizadas gratuitamente no repositório do INPE. Adotou-se a estratégia *Automatic Point Matching (APM)* como procedimento de amostragem, disponível no software ERDAS 9.1.

A relação de métodos e parâmetros utilizados no processo de ortorretificação, bem como seus respectivos valores de dados encontram-se relacionados no Quadro (III) 1.

Método (M)/Parâmetro(P)	Descrição/Valor
Reamostragem (M)	Vizinho mais próximo
Seleção de pontos de controle (M)	Automatic Point Matching (APM)
Número de pontos amostrados (P)	96
Modelo de transformação geométrica (M)	Projetiva 3D
RMS (P)	0,34 metros
MDE – altitude mínima (P)	322,6
MDE – altitude máxima (P)	1.457,4

Quadro (III) 1 – Sumário de métodos e parâmetros utilizados no processo de ortorretificação.

Nota: O parâmetro RMS é expresso em unidades de referência da base utilizada.

O valor referente ao parâmetro RMS (*Root Mean Square* ou Erro médio quadrático) não corresponde ao procedimento de checagem com pontos de controle levantados em campo e sim aos pontos de controle amostrados automaticamente nas imagens

⁴ Apesar do bom senso cartográfico sugerir que a ordem mais adequada de procedimentos seja a de ortorretificação em seguida mosaico, portanto, o inverso do que foi feito, vale esclarecer que a área coberta pelo MDE na imagem da cena da órbita 217 ponto 75 é muito pequena. Ou seja, apenas a porção da região de cabeceira teria referência de altitude, o que poderia ser prejudicial na qualidade do processo de ortorretificação devido à dificuldade de encontrar pontos de controle. Como o MDE não foi gerado em um formato retangular, mas sim no formato irregular da bacia hidrográfica, haveria cerca de 90% da cena sem cobertura pela elevação na referida cena, uma vez que a maior parte da bacia, e portanto o MDE, estão na cena da órbita 217 ponto 74.

digitais já ortorretificadas e utilizadas como referência cartográfica. Por isso seu valor tão baixo. Após a imagem ter sido ortorretificada e o MDE ter sido corrigido de depressões espúrias, os produtos derivados de ambos apresentaram grande coincidência de superfícies.

3.2 Definição das classes informacionais

A definição das classes informacionais representa a etapa que efetivamente subsidia a elaboração de IFVP. Após diversas visitas a campo constatou-se que a ocupação do solo na bacia do rio Piranga é predominantemente caracterizada pelas classes mata secundária (A), café (B), eucalipto(C), pastagem (D), solo exposto (E) e área agrícola (F), ilustrada na Figura (III) 1.

Como se trata de uma região de relevo forte ondulado a montanhoso, há ocorrência de sombras provocadas pela superfície escarpada, além de áreas com alguma porção coberta por pequenas barragens, represas e rios mais largos. Algumas áreas de queimada também são observadas, o que impossibilita inferir qual cobertura havia no local. Embora não representem a tipologia de uma cobertura comum, a análise de inspeção visual na imagem colorida possibilitou identificar mais essas três classes: superfície sombreada, superfície aquosa e áreas de queimada, que devem ser distinguidas.

As áreas urbanizadas não foram abordadas como uma classe informacional objetivo da classificação, por dois motivos: 1) por se tratar de um ambiente misto, onde várias categorias são observadas em conjunto, o que provoca confusão ao classificador; 2) sua ocorrência na região aparenta pouca representatividade espacial, apenas as cidades de Viçosa e Ponte Nova ocupam área maior.

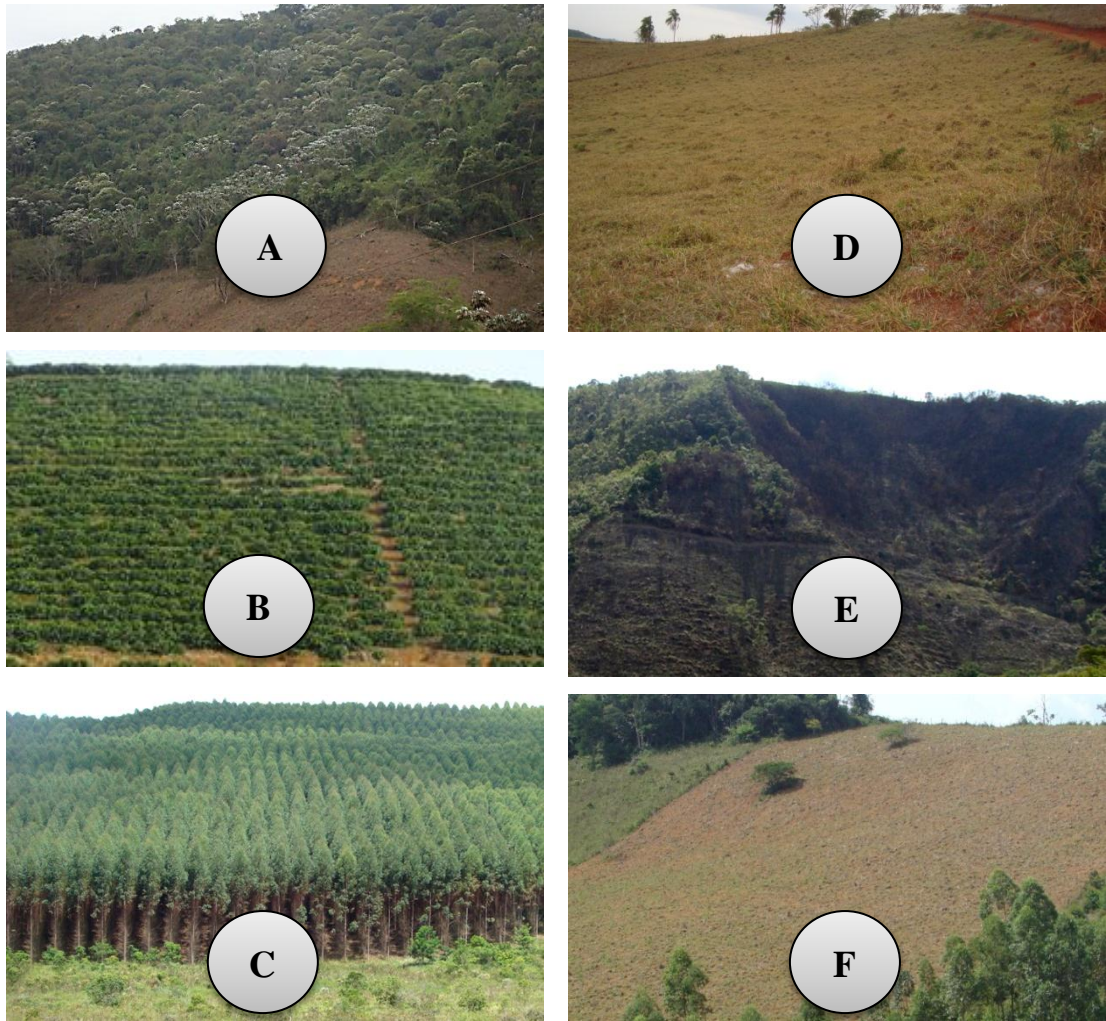


Figura (III) 1 – Fotografias dos tipos de uso e ocupação do solo mais comuns encontrados na BHRP, os quais sejam: mata secundária (A); café (B); eucalipto (C); pastagem (D); solo exposto (E); e área agrícola (F).

3.3 Amostragem das classes informacionais

Estabeleceu-se o conjunto amostral de referência para treinamento e avaliação dos métodos a partir de visitas a campo aliadas a inspeção visual em imagens do *Google Earth*. Visitas a campo apoiaram o levantamento das amostras das classes de uso e cobertura do solo, enquanto a inspeção visual apoiou a obtenção das amostras para as classes sombra, água e queimadas. As quantidades de pixels relativas a cada classe espectral para as etapas de treinamento e avaliação dos métodos são apresentadas na Tabela (III) 1.

Tabela (III) 1 – Número de pixels utilizados nas etapas de treinamento e teste dos métodos para cada classe informacional.

Id.	Classe Informacional	Treino	Avaliação	Total
1*	Eucalipto jovem	1081	139	1220
2	Pastagem	1269	186	1455
3	Café	667	91	758
4	Área agrícola	670	114	784
5	Mata	1181	137	1318
6	Área queimada	709	82	791
7	Superfície sombreada	475	65	540
8	Superfície aquosa	437	67	504
9	Solo exposto	489	52	541
10*	Eucalipto maduro	678	50	728

Fonte: o autor.

Nota: As classes eucalipto jovem e eucalipto maduro identificadas pelo símbolo “*” foram agrupadas na classe “eucalipto” após a classificação.

3.4 Imagens de representação da paisagem

Ao valorar a informação de expectativa de ocorrência de uma classe informacional para determinada região, como uma forma de representação da paisagem, a consequência é a variação desse valor pela imagem. Ou seja, ao tratar os valores de expectativa como Probabilidades se determinados pela RLB ou como Adequabilidade se obtidos pela AMC, *a priori*, tem-se como efeito ser espacialmente pouco ou muito informativo conforme o critério formulado pelo especialista ao formalizar o contexto da análise. Esta é uma característica muito útil para se explorar pela regra de decisão, pois onde houver maior probabilidade/adequabilidade *a priori* de ocorrência de uma determinada classe, maior o papel que a informação especialista desempenhará.

Como as imagens de representação da paisagem são Imagens Fundamentadas em Valores de Probabilidade (IFVP) para o caso da RLB ou de Adequabilidade (IFVA) para o caso da AMC, optou-se por padronizar o termo IFVP para ambos (probabilidade/adequabilidade), indistintamente. Isto para facilitar referências futuras no texto. Uma análise comparativa das IFVP encontra-se resumida no Quadro (III) 2.

Vantagens	Desvantagens
1 – Elaborada apenas uma vez para cada classe (i.e., apenas uma alteração significativa no modelo de uso e ocupação do solo justificaria a necessidade de atualização)	1 – Demanda mais tempo para especificar o contexto (i.e., produzir/acessar os modelos digitais da base de dados acessória, bem como garantir a compatibilidade com a base existente)
2 – Pode ser utilizada novamente em procedimentos futuros (i.e., para fins de monitoramento)	2 – Demanda mais espaço para armazenamento (i.e., para cada classe informacional, uma IFVP é gerada)
3 – Ser independente do número de classes informacionais (i.e., o aparecimento de uma nova classe não implica na invalidade das imagens já produzidas para as demais classes)	3 – Preconiza expertise do analista (i.e., há necessidade de um especialista para dizer quais variáveis de ambiente estão associadas à ocorrência de uma determinada classe)
4 – Flexibilidade quanto ao método de obtenção (i.e., podem ser produzidas por maneiras diferentes)	4 – Necessitam de informações auxiliares de base de dados adicionais (i.e., além das bandas espectrais)
5 – Fácil atualização pelo paradigma Bayesiano (i.e., a IFVP de uma classe pode ser parte da base de informação contextual para geração de uma outra IFVP)	5 – São dependentes do espaço geográfico (i.e., o conjunto de informações contextuais de uma mesma classe pode não ser adequado para diferentes regiões)
6 – Caráter informativo espaço-seletivo (i.e., a informação a <i>priori</i> varia espacialmente conforme o grau de susceptibilidade de ocorrência da classe).	

Quadro (III) 2 – Vantagens e desvantagens das Imagens Fundamentadas em Valores de Probabilidade (IFVP).

Em um conjunto de n imagens IFVP, o valor de um determinado pixel p na IFVP k , localizado coluna i e na linha j é definido por $p_{(i,j,k)} = [0, 1]$. Para o conjunto de n IFVP, tem-se: $\sum_{k=1}^n p_{i,j,k} = [0, \leq n]$.

Para gerar as IFVP optou-se por dois métodos distintos em relação à supervisão de um analista, a Regressão Logística Binomial e a Avaliação Multicritérios. A opção pelos métodos de geração das IFVP via RLB e AMC, em relação à supervisão de um analista, justifica-se por características implícitas de cada um em relação ao tratamento da informação subjetiva na análise. No método RLB há necessidade de supervisão do expert para coleção de amostras de treinamento nos locais onde se observa ocorrência de determinada categoria temática como forma de especificar a informação contextual, ou seja, trata-se de um procedimento supervisionado. Por outro lado, na AMC a informação contextual é construída pela especificação de cenários sem necessidade da supervisão do analista, isto é, apenas com base em critérios de âmbito geral, o que caracteriza um processo não supervisionado.

3.5 Classificação híbrida

Durante um processo de classificação convencional assume-se que os valores amostrais dos vetores espectrais de uma determinada classe informacional seguem uma distribuição Multivariada Normal. Adota-se então o método de Estimação da Máxima Verossimilhança para determinar os estimadores dos parâmetros de uma função de máxima *verossimilhança* Gaussiana Multivariada, a partir dos quais serão feitas as generalizações.

No processo de classificação proposto, adicionalmente considera-se que a IFVP é utilizada como representação da informação *a priori* no processo de classificação. Desse modo, em uma operação pixel a pixel na imagem, o uso da informação *priori* via IFVP combinado pela função de máxima *verossimilhança* é a probabilidade *a posteriori*, de aquele pixel pertencer a uma determinada classe de interesse, o que configura a aplicação de uma abordagem de classificação híbrida.

4 Resultados e discussões

4.1 Definição das informações contextuais

Classes informacionais que representam superfícies sombreadas e aquosas são correntemente confundidas quando submetidas ao processo de classificação de imagens realizadas por análise de variância com amostras extraídas de bandas espectrais. No entanto, superfícies aquosas são geralmente observadas em condições topográficas distintas daquelas que favorecem a ocorrência de áreas sombreadas (i.e. declividades e posição angular, morfologia, entre outros aspectos). Embora esses alvos sejam muito semelhantes no domínio espectral, o ambiente que contextualiza sua ocorrência é diferente.

As variáveis de ambiente utilizadas para representar o contexto de ocorrência das classes informacionais foram as seguintes: Modelo Digital de Elevação (MDE), Modelo Digital de Declividades (MDD), Modelo Digital de Direções do Terreno,

para o Norte/Sul (MDN – *Northness*) cuja superfície topográfica é voltada para o norte ou sul e Modelo Digital de Direções no qual elas são voltadas para Leste/Oeste (MDL – *Eastness*). Os modelos de direções N/S e L/O são obtidos respectivamente por operações trigonométricas de cosseno e seno, aplicadas no Modelo Digital *Aspect* (DOMAÇ e SÜZEN, 2006). Além das variáveis topográficas, obteve-se ainda a variável espectral NDVI⁵.

4.2 IFVP obtidas por RLB

Os resultados detalhados do processo de RLB estão apresentados no Apêndice (III) A, onde pode-se observar os valores dos coeficientes do vetor paramétrico β , bem como as medidas de avaliação da qualidade mais usuais, tais como Pseudo R^2 (Equação 6), AIC (Equação 7) e ROC (Equação 12). Com interpretação semelhante à adotada ao coeficiente de qualidade de ajuste de modelos lineares R^2 , os maiores valores de Pseudo R^2 indicam melhor ajuste. Isto também acontece com os valores da estatística ROC, ou seja, maiores valores indicam melhores ajustes. Entretanto, a interpretação para o parâmetro AIC refere-se aos menores valores como indicadores dos melhores ajustes.

Devido à semelhança de contexto de algumas classes informacionais, optou-se por confeccionar modelos representativos para um consórcio de classes, que compartilham características contextuais. Esse foi o caso das classes mata, café e eucalipto, que ocorrem em condições muito semelhantes, portanto, a probabilidade e a adequabilidade de um deles é explicativa para os outros dois. O que as diferenciam é apenas o porte arbóreo. As três foram agrupadas no cenário “classes de porte arbóreo alto”. Seguindo o mesmo princípio, formulou-se um cenário para associar as classes: pastagem, área agrícola e área de queimadas, identificadas pelo cenário “classes de porte arbóreo baixo”. Para cada cenário foi ajustado modelo de RLB. As classes superfície sombreada, superfície aquosa e solo exposto tiveram seus modelos especificados individualmente, pois suas características contextuais são notadamente

⁵ Obtido através de operações aritméticas das bandas espectrais na região de intensidade de radiação eletromagnética do InfraVermelho próximo (IV) e Vermelho (V) por $NDVI = (IV - V)/(IV + V)$.

diferenciadas. Uma versão resumida da tabela disponível no Apêndice (III) A, com os melhores modelos eleitos, encontra-se na Tabela (III) 2.

Tabela (III) 2 – Parâmetros de avaliação de qualidade para seleção de modelos.

Classe	Modelo	ROC	Pseudo R ²	AIC
Porte arbóreo alto	Mod.1	0,88	0,17	5040,5
	Mod.2	0,88	0,18	5230,3
	Mod.3*	0,88	0,17	4947,3
Porte arbóreo baixo	Mod.1	0,85	0,12	4262,9
	Mod.2	0,86	0,13	2260,6
	Mod.3*	0,86	0,12	2215,5
Superfície sombreada	Mod.1	0,868	0,12	1049,2
	Mod.2	0,864	0,09	859,1
	Mod.3*	0,863	0,10	710,5
Superfície aquosa	Mod.1*	0,987	0,38	601,8
	Mod.2	0,989	0,32	708,8
	Mod.3	0,887	0,04	696,5
Solo exposto	Mod.1	0,906	0,11	922,5
	Mod.2	0,939	0,12	1054,3
	Mod.3*	0,942	0,12	780,3

Fonte: o autor.

Nota: o símbolo “*” denota o modelo escolhido

Os valores em negrito (Tabela (III) 2) indicam os melhores parâmetros de determina classe a que se referem, portanto, o melhor modelo segundo aquele critério. Percebe-se que dificilmente todos os parâmetros indicaram concordância quanto ao melhor ajuste, o que dificulta a seleção. Contudo, esses parâmetros obedecem a certa hierarquia de importância, na qual Pseudo R² tem a menor e AIC a maior. Mesmo assim sua medida auxilia a avaliar o desempenho dos modelos quando analisada em conjunto.

Respectivamente as classes superfície aquosa e solo exposto tiveram ótimo desempenho medido por ROC. Frequentemente os valores de ROC foram aproximados entre modelos da mesma classe, impondo a decisão pela escolha do melhor à medida AIC.

As diferenças entre os modelos adotados para cada classe se caracterizam pelas informações contextuais, ou seja, quais variáveis de ambiente topográficas e espectrais foram utilizadas como variáveis explicativas. A partir dos modelos

selecionados pelo critério de melhor desempenho segundo os critérios de avaliação, determinou-se as respectivas IFVP aplicando-se a Equação 1.

4.3 IFVP determinadas por MCE

Talvez o termo probabilidade do acrônimo IFVP deva ser repensado para a definição dos valores de pertinência determinados pelo método de Agregação Multicritérios. Esse argumento é devido aos valores resultantes do processo *OWA* representarem variações de intensidade de adequabilidade a um determinado objetivo, escalonados no intervalo de zero a um, e não uma probabilidade propriamente dita. Apesar disso, a semântica do termo IFVP será mantida no decorrer do texto.

Foram gerados cenários de adequabilidade para caracterizar a ocorrência das classes informacionais conforme as informações contextuais disponíveis em forma de variáveis de ambiente topográfico e espectral. A abordagem *OWA* é mais flexível que a *RLB*, pois possibilita a inserção de variáveis categóricas na análise. Desse modo, além das variáveis *NDVI*, *MDE* e *MDD*, fez-se uso da variável *Categorias Morfológicas do Terreno (CMT)*. Para isso, utilizou-se o algoritmo para classificação automática da topografia do terreno a partir do *MDE* chamada *TOPOSHAPE*, implementada no software *Idrisi*. Utilizou-se ainda o modelo *Aspect*, caracterizador da direção de declividade do terreno – numa abordagem semelhante aos *MDN* e *MDL* adotados no procedimento de *RLB*, porém com valores angulares isentos de transformação.

Em abordagem similar àquela adotada no método *RLB*, formulou-se os mesmos dois cenários para representar consórcios de classes informacionais de *Porte Arbóreo Alto* e *Baixo*. Exatamente as mesmas classes foram empregadas na formulação respectiva de cada cenário. As classes *superfície sombreada*, *superfície aquosa* e *solo exposto* foram tipificadas separadamente.

A partir da Tabela (III) 3 até a Tabela (III) 7 são apresentados os resultados dos parâmetros utilizados nas análises *OWA*, bem como seus parâmetros de medida de qualidade. Todas as tabelas compartilham as mesmas características de parâmetros. Há nelas uma divisão que representa as duas etapas que dividem a análise de *Agregação Multicritérios*. No campo *fatores* há a relação de variáveis de ambiente

utilizadas para constituição dos cenários. Nessa parte há ainda a relação de valores dos pesos obtidos pela técnica AHP para cada fator, bem como o grau de consistência (*CR*) derivado da matriz de comparação pareada da importância relativa entre os fatores. Os pontos de controle utilizados no procedimento de padronização *fuzzy* de cada fator são exibidos. Note-se que nas ocasiões em que apenas os parâmetros x_a e x_b foram preenchidos, trata-se da função monotônica crescente. Por outro lado, quando apenas os parâmetros x_c e x_d são preenchidos, a função é monotonicamente decrescente. Um caso particular é apresentado quando todos são preenchidos, o que indica duas fases de comportamento da função, uma de elevação outra de decaimento. Isto a caracteriza como uma função simétrica, crescente de a até b e decrescente de c até d .

Na parte *cenários* há a especificação dos valores assumidos aos pesos ponderados na configuração de cada um deles. Pelo menos três cenários foram simulados para cada classe informacional (ou consórcio de classes). Nas tabelas pode-se verificar também os valores de *Risco*, dado por $(n - 1)^{-1} \sum_{i=1}^n (n - i) O_i$, *Trade-off* dado por $1 - [n \sum_{i=1}^n (O_i - n^{-1})^2 (n - 1)^{-1}]^{\frac{1}{2}}$ e *ROC* para cada cenário. Onde i é a ordem do peso, O_i é o peso ordenado e n a quantidade de fatores. A partir dos pesos ordenados, a medida *Risk* tem por objetivo identificar o risco do cenário configurado através da sua localização na área estratégica de decisão em relação aos operadores de risco mínimo (*AND*) e de risco máximo (*OR*), no eixo x . Complementarmente, a medida *trade-off* localiza no eixo y , a compensação entre os fatores. Valores menores de *Risk* indicam maior atuação do operador *AND* e, portanto, uma postura mais conservadora na análise. Valores maiores de *Risk*, próximos à unidade, indicam o contrário, ou seja, que o operador *OR* governa a análise. A compensação apresentada por *Trade-off* fornece a medida acerca do nível de privilégio atribuído para os fatores, a partir dos valores de seus pesos. Segundo (EASTMAN, 2009), conforme a maioria do peso ordenado é assinalada para a esquerda, direita ou centro dos pesos ordenados, é determinada a posição na dimensão do risco no espaço estratégico de decisão (Figura (III) 2).

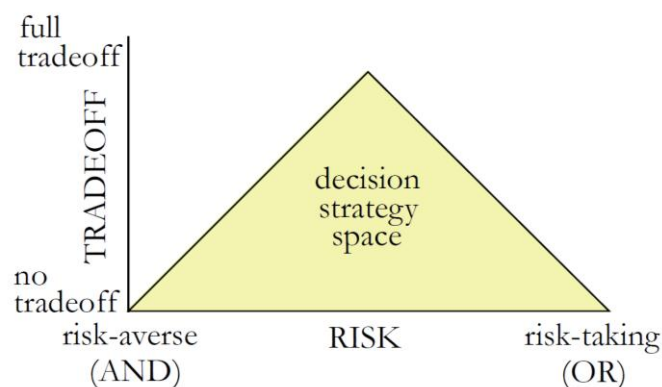


Figura (III) 2 – Espaço estratégico de decisão.

Fonte: Eastman (2009), p.141.

Um cenário com razoabilidade preconiza uma boa compensação entre fatores, o que indica a contribuição efetiva dos mesmos para a análise.

Para o consórcio de classes “vegetação de porte arbóreo alto” (Tabela (III) 3), o cenário que apresentou melhor desempenho do valor da estatística *ROC* foi justamente o que apresentou maior *Risk* entre os três. Percebe-se, contudo, que seu *trade-off* foi alto.

Tabela (III) 3 – Resultados OWA para o consórcio de classes “Vegetação de Porte Alto”.

Classes de Vegetação de Porte Arbóreo Alto									
Fatores					Cenários				
Pontos de	NDVI	MDE	MDD	CMT	Pesos Ordenados	(i)	(ii)	(iii)	
Controle/Pesos	0,5254	0,1393	0,2788	0,0565	P.ord.1	0,25	0,4	0,16	
x_a	0,4	300	0	-	P.ord.2	0,25	0,3	0,22	
x_b	0,8	600	10	-	P.ord.3	0,25	0,2	0,28	
x_c	-	900	30	-	P.ord.4	0,25	0,1	0,34	
x_d	-	1200	60	-	Risk	0,5	0,67	0,4	
CR		0,05			Trade-off	1	0,74	0,85	
					ROC	0,851	0,865	0,793	

Fonte: o autor.

Entretanto, para o consórcio de classes “vegetação de porte baixo” (Tabela (III) 4), o melhor desempenho foi obtido pelo cenário com mais baixo *Risk* e com *trade-off* médio.

Tabela (III) 4 – Resultados OWA para o consórcio de classes “Vegetação de Porte Baixo”.

Classes de Vegetação de Porte Arbóreo Baixo							
Fatores				Cenários			
Pontos de	NDVI	CMT	MDD	Pesos Ordenados	(i)	(ii)	(iii)
Controle/Pesos	0,637	0,2583	0,1047	P.ord.1	0,333	0,5	0,15
x_a	-0,4	-	-	P.ord.2	0,333	0,3	0,3
x_b	-0,2	-	-	P.ord.3	0,333	0,2	0,55
x_c	0,2	-	30	Risk	0,5	0,65	0,3
x_d	0,4	-	60	Trade-off	1	0,7	0,65
CR		0,03		ROC	0,81	0,794	0,82

Fonte: o autor.

Para as classes individuais superfície sombreada (Tabela (III) 5), superfície aquosa (

Tabela (III) 6) e solo exposto (Tabela (III) 7), os melhores desempenhos foram observados para os cenários com *Risk* médio e *trade-off* máximo. Trata-se de um caso particular de *OWA*, no qual o operador se posiciona exatamente no meio entre *AND* e *OR*, e a compensação entre fatores é total. Tal caso é conhecido por *Weighted Linear Combination (WLC)*, onde uma interpretação é raciocinar em termos de média, isto é, a contribuição de cada fator é a média dos fatores. Caso os fatores tenham sido elaborados com cautela para especificar significado próximo da interpretação real do fenômeno, essa é uma opção com resultados quase sempre satisfatórios ao custo de menor esforço.

Tabela (III) 5 – Resultados OWA para a classe superfície sombreada.

Classe superfície sombreada							
Fatores				Cenários			
Pontos de	Aspect	MDD	NDVI	Pesos Ordenados	(i)	(ii)	(iii)
Controle/Pesos	0,4286	0,4286	0,1429	P.ord.1	0,333	0,52	0,12
x_a	170	30	0,1	P.ord.2	0,333	0,29	0,3
x_b	220	50	0,5	P.ord.3	0,333	0,19	0,58
x_c	230	-	-	Risk	0,5	0,67	0,27
x_d	280	-	-	Trade-off	1	0,71	0,6
CR		0		ROC	0,95	0,95	0,93

Fonte: o autor.

Tabela (III) 6 – Resultados OWA para a classe superfície aquosa.

Classe superfície aquosa							
Fatores				Cenários			
Pontos de	MDD	NDVI	CTM	Pesos Ordenados	(i)	(ii)	(iii)
Controle/Pesos	0,2583	0,637	0,1047	P.ord.1	0,333	0,15	0,55
x_a	-	-	-	P.ord.2	0,333	0,55	0,18
x_b	-	-	-	P.ord.3	0,333	0,3	0,27
x_c	1	-0,1	-	Risk	0,5	0,43	0,64
x_d	10	0,1	-	Trade-off	1	0,65	0,67
CR		0,03		ROC	0,95	0,94	0,95

Fonte: o autor.

Tabela (III) 7 – Resultados OWA para a classe Solo Exposto.

Classe Solo Exposto							
Fatores				Cenários			
Pontos de	NDVI	CMT	MDD	Pesos Ordenados	(i)	(ii)	(iii)
Controle/Pesos	0,637	0,2583	0,1047	P.ord.1	0,333	0,53	0,12
x_a	-0,1	-	30	P.ord.2	0,333	0,29	0,3
x_b	-0,01	-	60	P.ord.3	0,333	0,18	0,58
x_c	0,01	-	-	Risk	0,5	0,68	0,27
x_d	0,1	-	-	Trade-off	1	0,69	0,6
CR		0,03		ROC	0,94	0,93	0,84

Fonte: o autor.

4.4 Classificação híbrida de imagem via IFVP

Apenas os melhores resultados da avaliação pela estatística *ROC* em relação aos métodos propostos, da RLB e da AMC, para obtenção de IFVP foram utilizados para compor as probabilidades *a priori* utilizadas no processo de classificação híbrida.

Os valores de probabilidade para o método de RLB foram obtidos pelo modelo específico escolhido (Apêndice (III) A). As imagens IFVP resultantes foram empregadas como informação *a priori* no método de classificação da máxima verossimilhança Gaussiana. Utilizou-se as referidas imagens para os consórcios de

classe conforme o porte arbóreo, além das classes individuais superfície aquosa, superfície sombreada e solo exposto.

Conforme exposto, os processos de classificação, híbrida e convencional para efeito de avaliação, foram conduzidos no SIG Idrisi. A imagem classificada resultante foi avaliada segundo índice de concordância Kappa⁶. Se comparada com a metodologia de classificação convencional pela máxima verossimilhança Gaussiana, os resultados pelo método da RLB foram inferiores. O Kappa obtido pela classificação convencional foi de 0,828, enquanto que o fornecido pelo método *RLB* foi de 0,827. Aparentemente o método introduziu informações discrepantes ao processo, contribuindo para um desempenho pior.

Em relação às IFVP produzidas pelo método de Agregação Multicritérios, o desempenho também não foi significativamente melhor, fornecendo o valor de Kappa igual a 0,833. Através do método *OWA* foi possível produzir imagens de adequabilidade como alternativa para incorporar conhecimento no processo de classificação.

Entre os métodos *RLB* e *OWA*, o primeiro é expressivamente mais caro sob vários aspectos como tempo necessário para coleccionar amostras de treinamento e para realizar o processamento. Além disso, seus resultados foram inferiores. Por outro lado, o método *OWA* se mostrou efetivo e promissor para gerar IFVP. A partir de seus melhores resultados obtidos pela estatística ROC foi gerada a carta-imagem de uso e cobertura do solo de 2010 (Figura (III) 3).

⁶ É uma medida de avaliação da qualidade temática de produtos obtidos por classificação de imagens do sensoriamento remoto amplamente documentado pela literatura especializada. É calculado a partir de uma matriz de contingência entre a imagem classificada e outra de referência. Mais detalhes podem ser consultados em (Congalton, 1991).

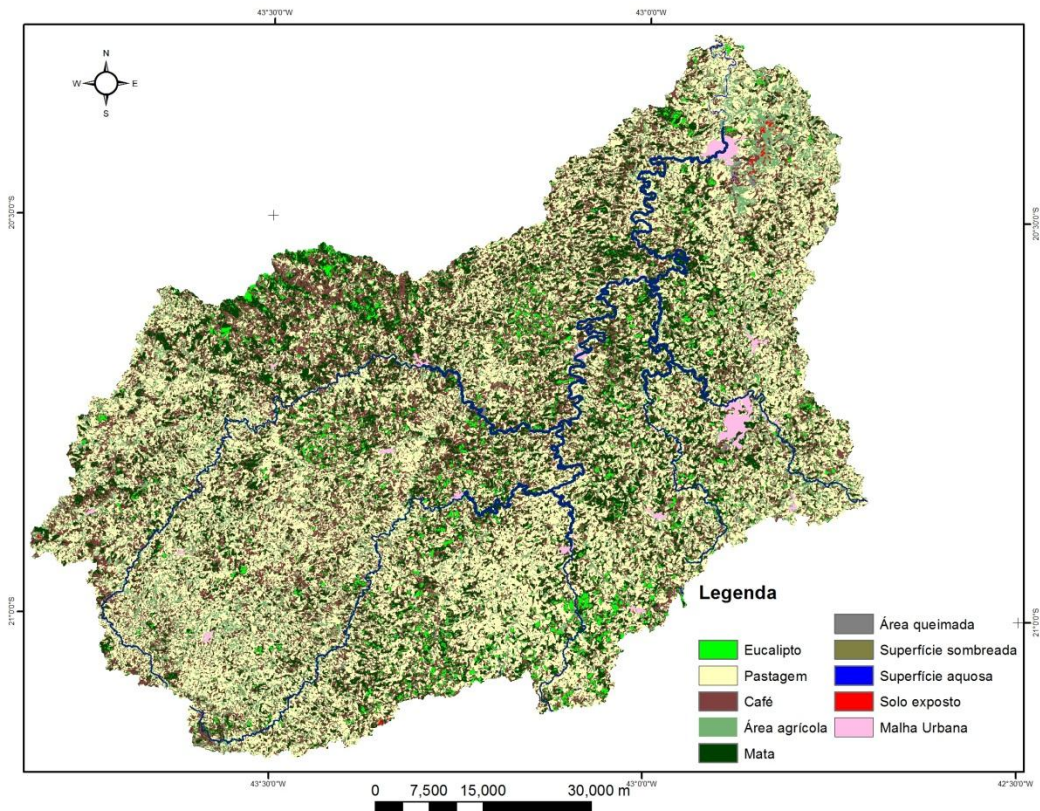


Figura (III) 3 – Imagem de uso e cobertura do solo de 2010 da bacia hidrográfica do rio Piranga.

Cabe mencionar que a técnica de incorporar informações contextuais (DANIELS, 2006) através de variáveis de ambiente (DOMAÇ e SÜZEN, 2006) representadas por valores de probabilidade (ASSIS, VIEIRA e SILVA, 2010) tem sido documentada com sucesso para o propósito de classificação de imagens de satélite. Contudo, a abordagem utilizada nesse estudo, apesar de não ter apresentado contribuição significativa, tem caráter de ineditismo.

As classes de uso e cobertura obtidas na carta-imagem classificada (Figura (III) 3) apresentaram as condições mais comuns de representação da paisagem na BHRP especificamente para o segundo semestre de 2010. Isto é, a classe área queimada muito provavelmente não seria objeto de análise caso a imagem fosse referente ao período chuvoso. Essa mesma classe não trata de uma ocupação propriamente dita, porém não era recomendável ignorá-la, pois seria certamente uma região de confusão no espaço característico, portanto, fonte de erro.

Conforme pode ser observado nos valores apresentados na Tabela (III) 8, em termos quantitativos, os usos e coberturas predominantes ocorrem por pastagens com cerca de 40%, seguida pela cultura de café (26%), mata secundária (15%) e agricultura incluindo outras culturas (12%). As plantações de eucalipto (jovem e maduro)

representam cerca de 3% da área. Somadas representam mais de 90% da paisagem verificada na região naquele período. Chama-se atenção para a ocorrência de solo exposto em apenas 0,35% de área ocupada. A expressiva área identificada como sombra é forte indicadora da superfície rugosa que se tem na região.

Tabela (III) 8 – Quantificação do uso e ocupação do solo na BHRP.

Classe Informacional	Área (ha)	%
Eucalipto	21.375,99	3,24
Pastagem	264.743,46	40,09
Café	172.494,36	26,12
Área agrícola	82.645,74	12,52
Mata secundária	99.739,62	15,10
Queimada	1.741,86	0,26
Superfície sombreada	9.217,98	1,40
Superfície aquosa	1.264,59	0,19
Solo exposto	2.316,06	0,35
Malha urbana	4.827,15	0,73

Fonte: o autor.

O resultado referente à classe café foi certamente superestimado. Houve certa confusão entre café e eucalipto jovem e até mata secundária, conforme pode ser constatado pela matriz de contingência apresentada no Apêndice (III) B. Assim sendo, parte expressiva rotulada como café pode ser na realidade, outra tipologia de cobertura do solo. Uma explicação plausível é devido a semelhança do porte arbóreo do eucalipto ainda jovem com o café. Adicionalmente, parte da confusão com mata se justifica devido a semelhança espectral entre o café e a mata secundária nessa época do ano, o mês de agosto. Por se tratar de um período de final de estiagem e boa parte da vegetação de mata secundária ser do tipo caducifólio, seu vigor vegetativo é reduzido. Apesar disso, o método da RLB já foi aplicado ao processo de classificação de imagens com sucesso (ASSIS, 2008), porém este ocorreu através de inferência Bayesiana e não isoladamente como neste estudo. Tal fato leva a crer que a abordagem aqui proposta poderia passar por uma releitura e, portanto, este trabalho tem caráter exploratório.

5 Conclusões

Nas condições em que o estudo foi realizado pode-se concluir que os objetivos propostos foram alcançados satisfatoriamente, pois foi possível avaliar as metodologias de RLB e de AMC aplicadas à descrição da paisagem para posterior uso no processo de classificação de imagens e então produzir um mapa de uso e cobertura do solo.

Em relação à avaliação das metodologias, concluiu-se que:

- a metodologia da RLB não foi eficiente para representar a paisagem através dos valores de probabilidade *a priori* no processo de classificação de imagens e sua efetiva implementação além de laboriosa, demandou considerável disponibilidade de recursos de processamento computacional;
- o método de agregação multicritérios OWA se mostrou promissor, pois apresentou facilidade de elaborar cenários e foi consideravelmente mais permissivo em relação à incorporação de fatores – possibilitando manipular dados de representação categóricos e contínuos;
- a maior parte dos cenários utilizados para representar a paisagem foi constituída de apenas 3 variáveis de informações contextuais – a menor quantidade permitida no método OWA, o que representa uma desvantagem em relação ao método da RLB, que permite apenas uma;
- as IFVP geradas pela metodologia OWA foram mais efetivas que aquelas geradas pela RLB, contudo, a suspeita de que essas imagens elevariam o grau de exatidão temática do processo de classificação pelo aumento da quantidade de variáveis discriminantes, não se confirmou.

Adicionalmente, foi possível responder à questão de pesquisa, havendo confirmado a metodologia da AMC como a mais adequada e, portanto, recomendada para representação da paisagem através da confecção de cenários, oferecendo melhores subsídios ao processo de classificação de imagens.

6 Referências

ABKAR, A. A.; MOHAMMED, A. S.; MULDER, N. J. Likelihood-based image segmentation and classification: a framework for the integration of expert knowledge in image classification procedures. **International Journal of Applied Earth-Observation and Geoinformation**, 2, 2000. 104-119.

ASSIS, C.; VIEIRA, C. A. O.; SILVA, F. F. **Improvement of the accuracy on image classification process through incorporation of contextual information**. Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Leicester-UK: University of Leicester. 2010. p. 341-344.

ASSIS, L. C. **Uso de Informações Contextuais no Processo de Classificação de Imagens do Sensoriamento Remoto**. Universidade Federal de Viçosa. Viçosa-MG, p. 121. 2008. Mestrado em Informações Espaciais.

CHICLANA, F. et al. Some induced ordered weighted averaging operators and their use for solving group decision-making problems based on fuzzy preference relations. **European Journal of Operational Research**, 182, 2007. 383-399.

CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, 37, 1991. 35-46.

DANIELS, A. E. Incorporating domain knowledge and spatial relationships into land cover classifications: a rule based approach. **International Journal of Remote Sensing**, 27, n. 14, 2006. 2949-2975.

DOMAÇ, A.; SÜZEN, M. L. Integration of environmental variables with satellite images in regional scale vegetation classification. **International Journal of Remote Sensing**, 27, n. 7, 2006. 1329-1350.

EASTMAN, J. R. **IDRISI Taiga: Guide to GIS and Image Processing**. Clark University. Worcester, MA, p. 325. 2009.

EASTMAN, J. R. et al. Raster procedures for multi-criteria/multi-objective decisions. **Photogrammetric Engineering & Remote Sensing**, 61, n. 5, May 1995. 539-547.

FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, 27, 2006. 861-874.

FULLÉR, R. OWA operators in decision making. In: CARLSSON, C. **Exploring the Limits of Support Systems**. [S.l.]: [s.n.], 1996. Cap. 3, p. 85-104.

LARSEN, K. et al. Interpreting Parameters in the Logistic Regression Model with Random Effects. **Biometrics**, 56, n. 3, Sep. 2000. 909-914.

MALPICA, J. A.; ALONSO, M. C.; SANZ, M. A. Dempster-Shafer theory in geographic information systems: a survey. **Expert Systems with Applications**, 32, 2007. 47-55.

MASOCHA, M.; SKIDMORE, A. K. Integrating conventional classifiers with a GIS expert system to increase the accuracy of invasive species mapping. **International Journal of Applied Earth Observation and Geoinformation**, 13, 2011. 487-494.

MATHER, P. M. Land cover classification revisited. In: ATKINSON, P. M.; TATE, N. J. **Advances in Remote Sensing and GIS Analysis**. Chichester: Wiley & Sons, 1999. p. 7-16.

MATHER, P. M. **Computer processing of remotely-sensed images: an introduction**. 3. ed. Chichester: John Wiley & Sons, 2004. ISBN ISBN 0-470-84919-3.

OÑATE-VALDIVIESO, F.; SENDRA, J. B. Application of GIS and remote sensing techniques in generation of land use scenarios for hidrologica modeling. **Journal of Hydrology**, 395, 2010. 256-263.

PENG, C.-Y. J.; LEE, K. L.; INGERSOLL, G. M. An Introduction to Logistic Regression Analysis and Reporting. **The Journal of Educational Research**, 96, n. 1, 2002. 3-14.

PINZ, A. Active fusion - a new method applied to remote sensing image interpretation. **Pattern Recognition Letters**, 17, 1996. 1349-1359.

PONTIUS JR., R. G.; SCHNEIDER, L. C. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. **Agriculture, Ecosystems and Environment**, 85, 2001. 239-248.

SAATY, R. W. The Analytic Hierarchy Process - What it is and how it is used. **Mathematical Modeling**, 9, n. 3, 1987. 161-176.

SAHA, G. Applying Logistic Regression Model to the examination results data. **Journal of Reliability and Statistical Studies**, 4, n. 2, 2011. 105-117.

SWETS, J. A. Measuring the Accuracy of Diagnostic Systems. **Science**, 240, 3 June 1988. 1285-1293.

VIEIRA, C. C. A.; ASSIS, L. C.; BECKER, M. **Identification of Priority Zones for Conservation Considering Water Resources in the Alto Paraguai Watershed Region**. Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Leicester: [s.n.]. 2010. p. 433-436.

WARDLOW, B. D.; EGBERT, S. L.; KASTENS, J. H. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. **Remote Sensing of Environment**, v. 108, p. 290-310, 2007.

YAGER, R. R. On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decisionmaking. **IEEE Transactions on Systems, Man, and Cybernetics**, 8, n. 1, 1988. 183-190.

ZADEH, L. A. Fuzzy Sets. **Information and Control**, 8, 1965. 338-353.

ZHU, L.; TATEISHI, R. Fusion of multisensor multitemporal satellite data for land cover mapping. **International Journal of Remote Sensing**, 27, 2006. 903-918.

IV. Simulação de processos hidrológicos de evento extremo na bacia hidrográfica do rio Piranga.

1 Introdução

Um grave problema ambiental observado na bacia do rio Doce, incluindo a região do rio Piranga na sua nascente, é a ocorrência de inundações recorrentes em áreas urbanas e ribeirinhas, isto porque algumas cidades ocuparam a planície de inundação dos rios. O desmatamento indiscriminado, o manejo inadequado do solo associado a despejos irregulares da mineração e de resíduos industriais e domésticos, criaram condições favoráveis à formação de processos erosivos responsáveis pelo assoreamento dos rios que, de tempos em tempos provocam o alagamento de parte dessas planícies causando graves problemas à população (CONSÓRCIO ECOPLANLUME, 2010). Há estudos em outros países que indicam aumento dos picos e volumes de inundação em decorrência da deterioração do uso e ocupação dos solos, entretanto a cobertura de vegetação não é tão efetiva para inundações de elevados períodos de retorno nesses lugares (SAGHAFIAN et al., 2008).

Registros históricos permitem verificar que os efeitos catastróficos das inundações provocam tragédias em zonas rurais ribeirinhas onde há o transbordamento da calha do rio, entretanto, esses são devastadores quando as ondas de cheia atingem as áreas urbanas de uma cidade.

A apresentação de qualquer proposta séria para enfrentamento do problema não procede sem antes passar por uma ampla análise dos fatores envolvidos. Dessa forma, se faz necessário conhecer os papéis desses fatores bem como suas características através de análise para que, soluções apropriadas possam ser cogitadas. Nesse contexto, o suporte dos Sistemas de Informações Geográficas (SIG) para o levantamento de informações hidrológicas associadas às características do ambiente, possibilita análise integrada da dinâmica do processo hidrológico em detalhe nas sub-bacias.

Alguns estudos hidrológicos apoiados por SIG exploraram as vantagens da análise hidrológica especializada como ferramenta de gerenciamento dos recursos hídricos em grandes bacias hidrográficas (AL-ABED, ABDULLA e KHYARAH, 2005), ou

como fundamentação para determinação de um modelo hidrológico local (MENDAS, 2010). Independente do propósito, já foi verificado que a distribuição espacial dos processos hidrológicos é mais sensível à estimação de eventos extremos de escoamento superficial que a abordagem convencional (KUMAR, SAMANIEGO e ATTINGER, 2010).

Variáveis topográficas que caracterizam as formas do terreno constituem elementos de alta relevância na determinação da vazão em corpos d'água, uma vez que suas características influenciam a velocidade do volume escoado e na conservação do solo, na identificação de locais adequados para manobras de intervenção entre outras. Modelos que tratam da distribuição espacial da água na bacia hidrográfica requerem dados com base em características topográficas tais como limites da bacia hidrográfica e sub-bacias, declividade, comprimento da encosta, forma do declive, aspecto e canais de drenagem (MOORE et al., 1993), entre outros.

2 Objetivo

O objetivo deste estudo é realizar simulação de processos hidrológicos na bacia hidrográfica do rio Piranga com vistas a avaliar um cenário crítico para a identificação de problemas relacionados ao escoamento superficial na região com suporte de análise espacial em ambiente SIG para subsidiar a obtenção de parâmetros de modelos hidrológicos adotados.

3 Revisão de literatura

3.1 Sistema de Simulação Hidrológica

Não é interessante analisar cada componente necessário à análise hidrológica de maneira individual e independente. Determinado software utilizado para simulação

de processos hidrológicos disponibiliza um conjunto limitado de métodos e modelos, que por sua vez precisam ser parametrizados com algum grau diferenciado de acurácia, influenciada pela qualidade da base de dados, e cujos valores podem ser determinados de forma mais ou menos dificultosa a depender do sistema utilizado. Todos os componentes são interdependentes e devem ser analisados de forma integrada.

Dentre o rol de ferramentas de simulação hidrológica disponíveis, talvez um dos mais adotados e documentados na literatura especializada seja o Sistema de Simulação Hidrológica – *Hydrologic Modeling System* HEC-HMS, disponibilizado pelo Centro de Engenharia Hidrológica – *Hydrologic Engineering Center* (HEC), vinculada ao Instituto de Recursos Hídricos – *Institute of Water Resources* (IWR), um dos centros de expertise do Corpo de Engenheiros do Exército dos Estados Unidos – *US Army Corps of Engineers* (USACE). Tal afirmação é suportada por aplicações desse sistema para objetivos diversos, como obter a hidrógrafa de uma onda de cheia em uma pequena bacia (YUSOP, CHAN e KATIMON, 2007), avaliar a intensificação de inundações em decorrência da alteração no uso do solo (SAGHAFIAN et al., 2008) ou, ainda, os impactos hidrológicos de mudanças climáticas (MEENU, REHANA e MUJUMDAR, 2012).

Em várias abordagens o emprego do HEC-HMS aparece fortemente vinculado ao SIG e às tecnologias do Sensoriamento Remoto Orbital (SRO). São, portanto, dados comuns à simulação de eventos de inundação com HEC-HMS aqueles referentes aos modelos de dados espaciais que descrevem a topografia do terreno (i.e., como o Modelo Digital de Elevação – MDE), o mapa de uso e cobertura do solo – quase sempre obtido por SRO, e os mapas de tipos de solos (OLANG e FÜRST, 2011). Há ainda os dados hidrometeorológicos das séries históricas de estações fluviométricas, pluviométricas e dos radares meteorológicos (MENDAS, 2010; VERMA, JHA e MAHANA, 2010).

3.2 Modelo Digital de Elevação Condicionado à Hidrografia

A composição de modelos digitais de terreno que representem descrição mais próxima com a paisagem observada e, portanto, contemplem as condições

topográficas e de ocupação do solo é tarefa fundamental para a simulação de processos hidrológicos (simulação hidrológica), pois é neste cenário que ocorre o escoamento superficial, frequentemente o maior responsável pela ocorrência de inundações. Aparentemente óbvio então que para a representação mais adequada de uma inundação através da propagação por ondas de cheia nos rios⁷ é necessário haver forte correspondência entre o modelo que representa a topografia do terreno, especialmente a conformidade dos vales, com a hidrografia natural (i.e., real) do lugar. Entretanto, tal coincidência é praticamente impossível em um MDE sem que este seja submetido a algum algoritmo de pré-processamento *a priori*. Desse modo, aspectos de atenção especial a serem considerados prudentemente no âmbito de uma análise hidrológica são relativos: ao MDE, pois dele deriva-se a rede de drenagem e demais parâmetros fisiográficos da bacia; ao Mapa de Tipos de Solos; e à Carta-Imagem de Uso Cobertura do Solo. Juntos fundamentam a simulação hidrológica através da parametrização de modelos e, como dizem Elfert e Bormann (2010), esta informação é um fator chave que controla o comportamento hidrológico das áreas de captação.

Alguns estudos tem destinado atenção especial no tratamento da qualidade da informação derivada do MDE (HENGL, HEUVELINK e VAN LOON, 2010; THOMMERET, BAILLY e PUECH, 2010). Também se comprovou que o aumento na qualidade da informação topográfica horizontal e da exatidão vertical do MDE (obtidos por exemplo através de batimetria do canal) pode provocar a redução da parcela de área inundada em estudos de simulação (COOK e MERWADE, 2009). Contudo, é necessário também estar atento à qualidade do próprio MDE e, a depender da aplicação, determinados procedimentos de ajuste são aceitos. No caso de extração de valores de parâmetros hidrológicos é recomendável que se utilize de procedimentos que garantam a integridade das informações do modelo, como o condicionamento do modelo de elevação à rede de drenagem mapeada. Esse outro modelo, conhecido como Modelo Digital Hidrograficamente Condicionado (MDEHC), pode ser derivado a partir do MDE e da hidrografia em representação matricial com parâmetros espaciais equivalentes aos seus. Apesar de interferir ligeiramente na forma dos canais naturais, esse modelo é adequado à simulação

⁷ O termo canais é empregado como sinônimo no decorrer do texto

hidrológica, pois se aproxima mais do modelo gravitacional, essencial para representar o escoamento superficial.

4 Métodos

Segundo definição de Houaiss, (2009), o termo simulação se refere ao teste, experiência ou ensaio em que se reproduz artificialmente uma situação, ou as condições reais de um meio, fenômeno, entre outros, frequentemente realizado com emprego de modelos. Devido se tratar de uma proposta para conhecer parte dos problemas associados às inundações através de simulações, estabeleceu-se um cenário conhecido desse evento para calibrar os parâmetros dos modelos hidrológicos adotados.

Para obter a configuração dos cenários, empregou-se componentes metodológicos conforme aqueles apresentados resumidamente na ilustração da Figura (IV) 1⁸.

A qualidade dos dados de entrada, a escolha de ferramentas apropriadas para manipulação e análise dos dados, bem como a seleção dos métodos e modelos adequados à proposta de simulação foram, em conjunto com a calibração de parâmetros e validação do modelo adotado, componentes fundamentais para se alcançar credibilidade dos resultados.

⁸ A pintura que ilustra o cenário de representação de uma onda é recorte do quadro *The Ninth Wave* de Ivan Konstantinovich Aivazovsky.

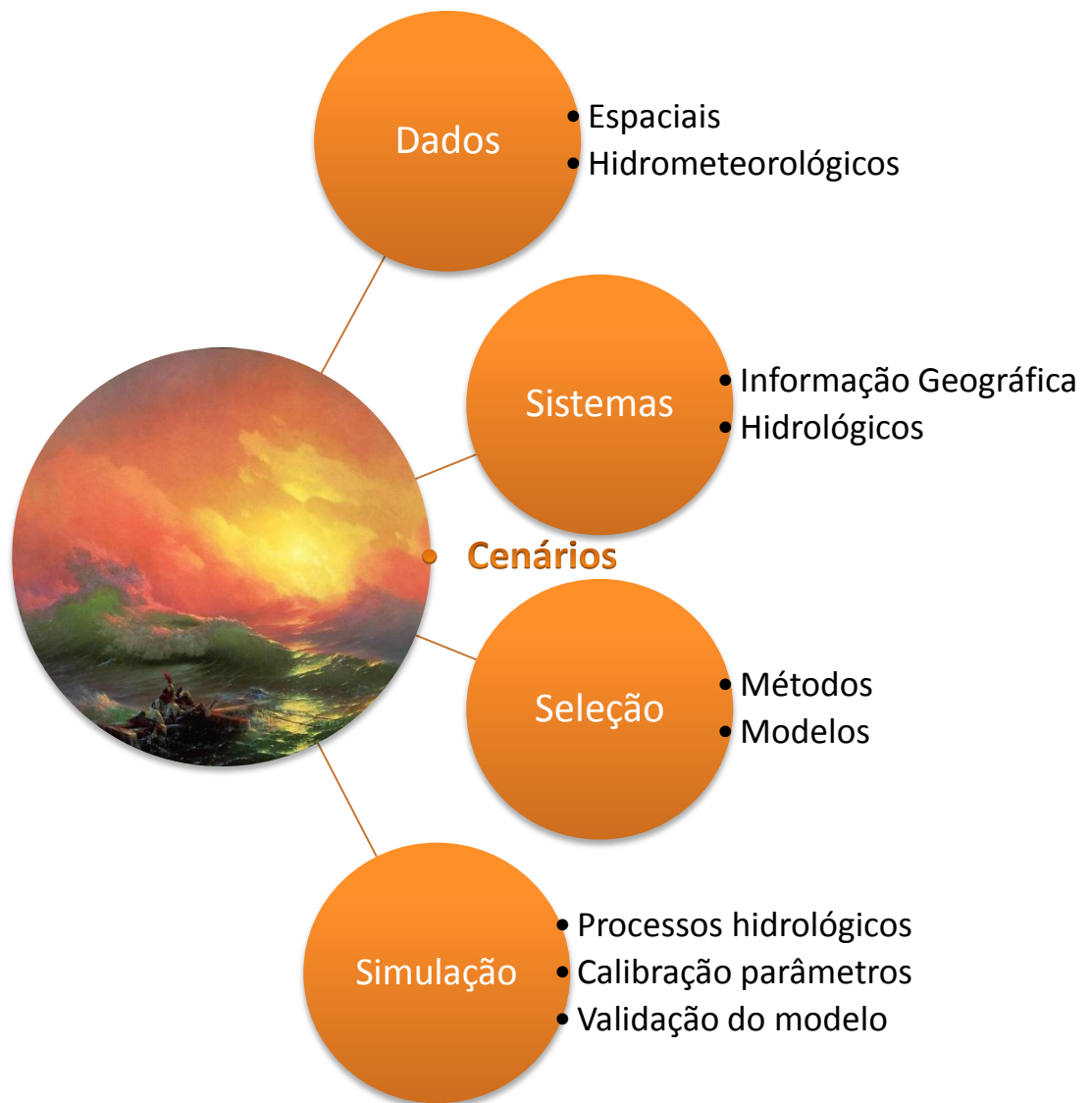


Figura (IV) 1 – Principais processos metodológicos utilizados nesta pesquisa.

4.1 Dados de entrada: hidrometeorológicos e espaciais

4.1.1 Dados hidrometeorológicos

Os dados hidrometeorológicos consistiram de séries históricas de altura máxima precipitada diária registrada por estações pluviométricas e vazões médias diárias obtidas pelo método da curva chave a partir da leitura de cota em estações fluviométricas localizadas na bacia hidrográfica do rio Piranga. A localização das

estações fluviométricas delimitou a área de estudo como sendo as respectivas áreas de drenagem de cada uma (Figura (IV) 2). A região completa de estudo foi definida pela área de captação à montante da estação fluviométrica localizada à jusante da cidade de Ponte Nova, no rio Piranga, cuja área delimitada é pouco inferior à da bacia desse rio, e que o exutório localiza-se na sua junção com o rio do Carmo, onde a partir desse ponto passa a se denominar rio Doce.

Conforme a disponibilidade de registros ininterruptos de valores nas séries históricas e a disposição espacial do arranjo de estações pluviométricas e fluviométricas, foram definidos os pares *Pluvio-Fluvio (PF)*, para representar a precipitação incidente e vazão correspondente para cada área de captação. Para cada par de registros hidrometeorológicos *PF* definiu-se um período comum da série histórica para análise.

4.1.2 Modelo Digital de Elevação Hidrograficamente Condicionado

O condicionamento do MDE à hidrografia mapeada foi realizado pelo algoritmo proposto por (RIBEIRO, MEITNER e VEIGA, 2006), e pode ser dividido em três grandes etapas, a saber: 1) rasterização da hidrografia mapeada, isto é, converter a hidrografia vetorial para sua representação matricial equivalente; 2) correção do perfil da hidrografia e; 3) refinamento de taludes ao longo da hidrografia. Foi implementada uma versão doméstica deste algoritmo no SIG ArcGIS. O MDEHC pode ser observado na ilustração da Figura (IV) 2.

Para verificar se o condicionamento do MDE à hidrografia mapeada estava adequado, procedeu-se com o seguinte teste empírico:

- gerou-se uma rede de drenagem numérica a partir do MDEHC⁹;
- converteu-se a rede numérica gerada para o modelo de dados vetorial;
- extraiu-se os pontos de confluências da rede numérica vetorial;
- extraiu-se também os pontos de confluências da hidrografia vetorial mapeada;

⁹ Mais especificamente, a rede de drenagem foi obtida a partir do Modelo Digital de Fluxo Acumulado, que por sua vez foi gerado do Modelo Digital de Direções do Escoamento e este foi produzido pelo MDEHC.

- comparou-se as posições dos pontos de confluências da rede de drenagem gerada e da hidrografia mapeada, através de operador de distância tipo *buffer*.

Foi possível constatar um nível de coincidência dos pontos de confluências superior a 10%. Ou seja, esses pontos se localizavam num raio de menos de 10 metros de distância entre si. Como a resolução espacial do MDEHC era de 15 metros, equivale dizer que esses pontos se encontravam na mesma posição, portanto, o MDE estava satisfatoriamente condicionado à hidrografia mapeada.

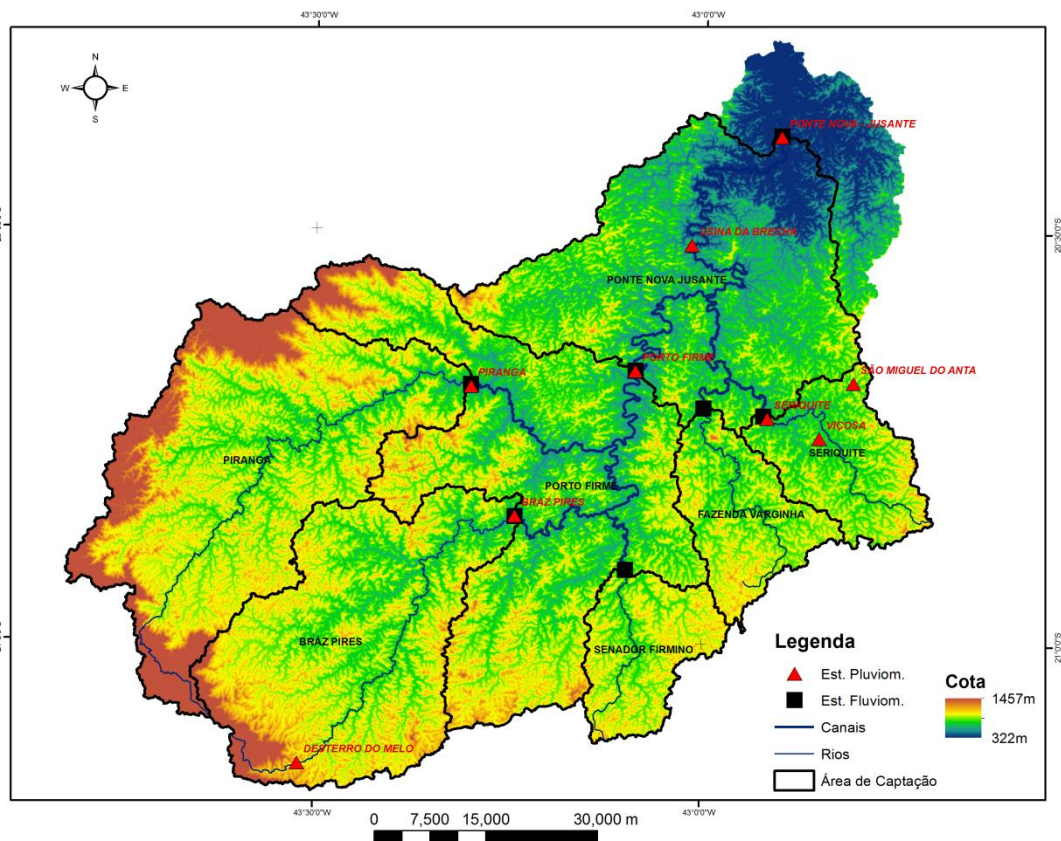


Figura (IV) 2 – Modelo digital de elevação condicionado à hidrografia, delimitação das áreas de drenagem de cada estação fluviométrica bem como sua identificação, estações pluviométricas, principais rios e seus respectivos canais evidenciados.

4.1.3 Imagem de Uso e Cobertura do Solo

Obtida a partir de classificação de imagem do sensor TM a bordo do satélite Landsat 5 de Agosto de 2010, a imagem resultante foi mosaificada, ortorretificada e classificada a partir de duas cenas da mesma órbita, porém, diferentes pontos. Como

o instante de tomada foi praticamente o mesmo, não houve necessidade de submetê-la ao processamento para atenuação dos efeitos atmosféricos como etapa de pré-processamento. A imagem classificada foi avaliada quanto à sua qualidade temática através do coeficiente de concordância Kappa, obtendo resultado de 0,83, numa escala de 0 a 1, considerado satisfatório. As categorias de uso e cobertura do solo mais comuns na bacia hidrográfica do rio Piranga são pastagem, café, mata secundária e outras culturas, conforme pode ser observado na ilustração da Figura (IV) 3.

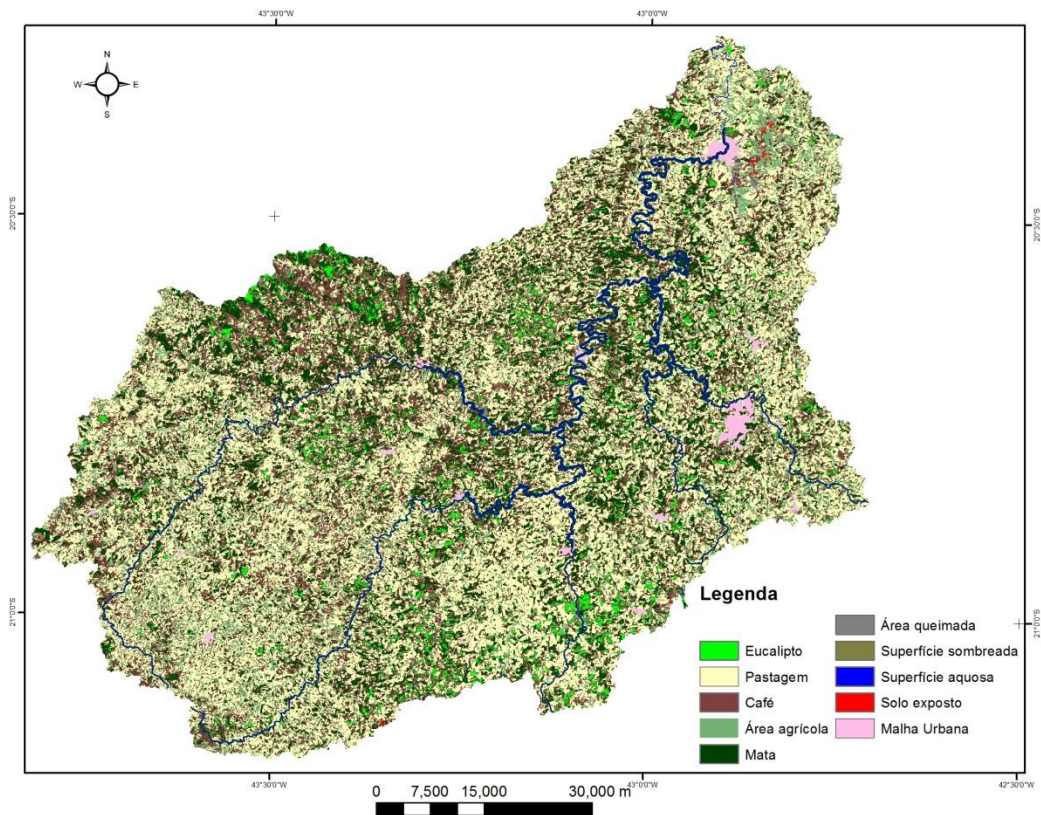


Figura (IV) 3 – Imagem de uso e cobertura do solo elaborada por processo de classificação de imagem do satélite Landsat 5 de agosto de 2010 da bacia hidrográfica do rio Piranga.

4.1.4 Mapa de Tipos de Solos

O mapa de solos (Figura (IV) 4) utilizado foi um recorte referente à área do rio Piranga do original “Levantamento de solos e aptidão agrícola da porção mineira da bacia do rio Doce” produzido e publicado por Fernandes Filho et al. (2010). Todas as

tipologias de solo que ocorrem na região apresentam horizonte A moderado, fase floresta tropical subperenifólia e suas especificidades quanto ao grupo predominante são:

- Latossolo Vermelho-Amarelo distrófico (LVAd) típico de textura argilosa com relevo forte ondulado (de 35% a 50%) no eixo Nordeste-Sudoeste (NE-SO) e ondulado/forte ondulado (de 50 a 60%) na região da nascente (SO) do alto piranga, e de textura argilosa/muito argilosa com relevo forte ondulado/montanhoso na porção Sudeste (SE);
- Latosso Vermelho distrófico (LVd) típico de textura muito argilosa com relevo forte ondulado e forte ondulado/montanhoso em cerca de 30% e 35% respectivamente, ambos ocorrem a Noroeste (NO) na região da nascente;
- Argissolo Vermelho-Amarelo eutrófico (PVAe) típico, de textura argilosa e relevo ondulado representando 40% e ocorrência bem definida na região do baixo Piranga próximo à sua foz, à Norte/Nordeste (N-NE);
- Argissolo Vermelho distrófico (PVD) típico de textura argilosa e relevo forte ondulado em cerca de 50%, aparece ao Norte (N) da região do alto rio Piranga;
- Cambissolo Háptico Tb¹⁰ distrófico (CXbd) típico, de textura média e relevo montanhoso (50%), ocorre a extremo Sudoeste na região da nascente.

¹⁰ Refere-se à capacidade de troca catiônica correspondente à fração argila, atividade baixa.

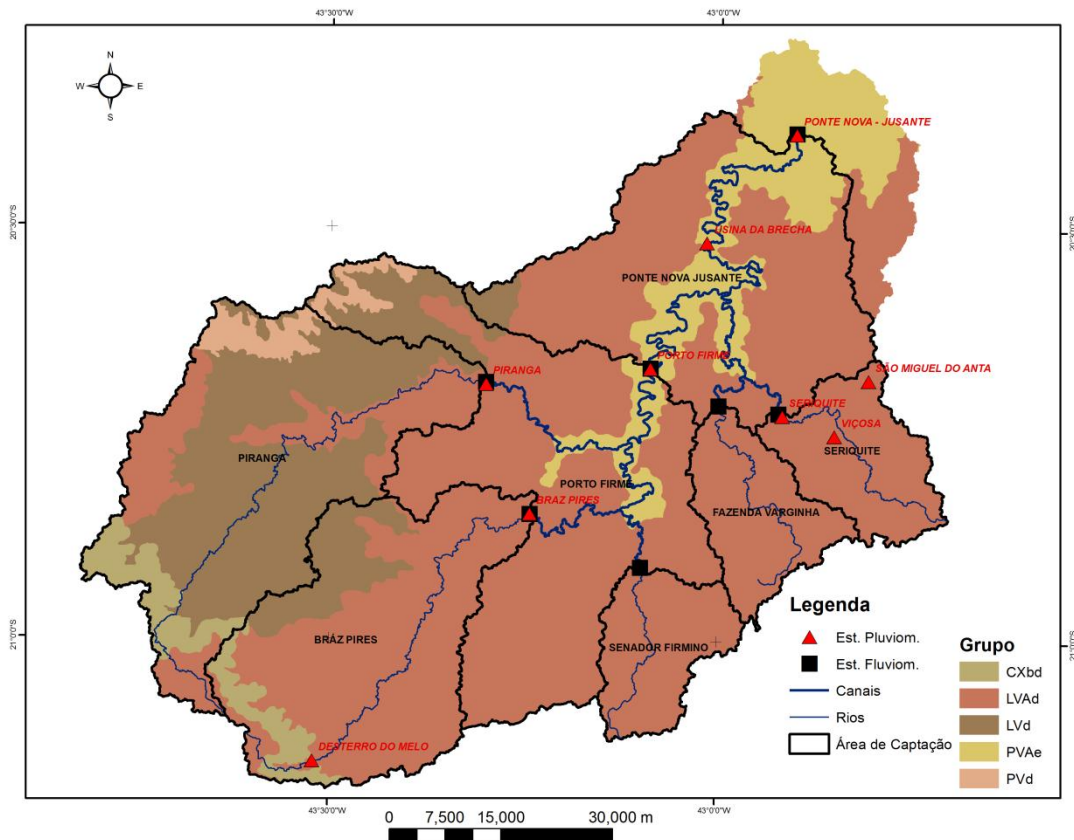


Figura (IV) 4 – Principais grupos tipológicos de solos ocorrentes na bacia hidrográfica do rio Piranga, delimitação das áreas de drenagem de cada estação fluviométrica em análise bem como sua identificação, estações pluviométricas, principais rios e seus respectivos canais evidenciados.

4.2 Sistemas de Informação Geográfica e de Simulação Hidrológica

Os sistemas informatizados empregados para manipulação e análise das informações de âmbito hidrológico nas dimensões espacial e temporal são ferramentas de grande utilidade e elevado potencial de aplicação. Recentes programas aplicativos têm convergido para integração¹¹ entre Sistemas de Informações Geográficas (SIG) e Sistemas de Simulação Hidrológica (SSH). O último precisa de grande quantidade de parâmetros para alimentar seus modelos e o primeiro é adequado para obter tais informações com agilidade.

Nesta pesquisa utilizou-se os sistemas para tratamento de informações espaciais ArcGIS 10 e para simulação de processos hidrológicos o HEC-HMS 3.5.

¹¹ Inicialmente ArcHydro Tools, posteriormente incorporado ao GeoHMS, GeoRAS (*River Analysis System*), entre outros. Todos são exemplos de *plug-ins* para o ArcGIS.

4.3 Seleção de métodos e modelos hidrológicos

Há princípios fundamentais que delineiam a seleção de métodos e modelos utilizados na simulação hidrológica que devem ser considerados sob a perspectiva de uma avaliação criteriosa acerca de vários aspectos. Através de questionamentos, esses princípios são:

- *Qual cenário se pretende simular?* O consórcio de métodos adotados inevitavelmente irá privilegiar alguma(s) característica em detrimento de outra(s);
- *Quais processos hidrológicos são necessários para configurar o cenário?* Pode haver processos essenciais para representar determinado cenário que, ou os dados, ou as ferramentas, ou a complexidade inerente inviabilizem a análise;
- *Os métodos disponíveis são adequados para os processos?* Todo método é fundamentado em premissas e hipóteses simplificativas que limitam sua aplicação (i.e., área de estudo, característica da bacia, tempo de duração do evento simulado, entre outras);
- *Quais parâmetros essenciais ao método são acessíveis?* Determinados parâmetros não podem ser calibrados, ao invés, devem ser determinados por levantamento (i.e., como medidas de área ou comprimento, entre outros), experimentação (i.e., em condições de campo) ou consulta a tabelas de referência¹²;
- *As escalas dos dados espaciais e temporais são adequadas aos métodos?* Vários métodos em hidrologia foram idealizados para características espaciais (i.e., dimensões da bacia, tipo de uso e cobertura, entre outros) e temporais (i.e., eventos discretos ou contínuos) diferenciadas, de modo que utilizá-las inadvertidamente pode ser prejudicial à análise.

¹² Deve-se ter cautela ao consultar essas fontes de referência uma vez que as condições ambientais ou fenomenológicas observadas na determinação dos valores referenciais podem ser significativamente distintas daquelas encontradas no estudo de caso aplicado.

O programa HEC-HMS oferece métodos capazes de simular várias etapas do ciclo hidrológico, dentre as quais optou-se por aquelas mais representativas dos eventos de inundação: (i) precipitação pluvial, (ii) escoamento superficial, (iii) transformação chuva-deflúvio e (iv) propagação em canais.

4.3.1 Precipitação pluvial

A partir dos dados das séries históricas das estações pluviométricas localizadas na área de estudo produziu-se mapas de precipitação pluvial diária pelo método do Inverso da Potência das Distâncias (IPD), um mapa para cada dia do período de duração dos eventos analisados na simulação de processos hidrológicos. Desse modo, foram elaborados conjuntos de mapas para representar eventos, um deles utilizado para calibração de parâmetros dos modelos, referente a uma chuva que ocorreu entre 17 de janeiro e 10 de fevereiro de 1980, e o outro, para efeito de avaliação dos modelos, relativo a uma chuva de 30 de Dezembro de 1997 a 09 de janeiro de 1998. A partir dos valores precipitados dos mapas de representação espacial da chuva, determinou-se pela estatística da média, o valor diário para cada área de drenagem e, a partir dessas estimativas, produziu-se os ietogramas dos eventos de calibração e avaliação dos modelos.

4.3.2 Método de determinação do escoamento superficial

Para quantificar o escoamento superficial utilizou-se o método do *National Resource Conservation Service-Curve Number* (NRCS-CN), conhecido como Número da Curva. Esse método fornece relações entre as abstrações iniciais, I_a , e os números da curva CN , baseado em experimentos conduzidos em pequenas bacias hidrográficas (KUMAR e BHATTACHARJYA, 2011). O método permite estimar a lâmina de escoamento superficial a partir de dados de precipitação e outros parâmetros da bacia (PRUSKI, BRANDÃO e SILVA, 2004). Fundamenta-se em uma tabela que associa

um determinado valor (de 25 a 100) às condições de tipos de solos e de uso e cobertura. Através dos experimentos foi possível evidenciar que a relação entre a infiltração acumulada (I) e a infiltração potencial (S) é equivalente à relação entre o escoamento superficial total (ES) e a diferença entre a precipitação total (PT) e as abstrações iniciais (I_a), ou seja, a precipitação efetiva (P_e), conforme a Equação 1.

$$\frac{I}{S} = \frac{ES}{PT - I_a} \quad (1)$$

em que $PT - I_a = P_e$. Os experimentos conduzidos permitiram verificar ainda que I_a corresponde a cerca de 20% de S (PRUSKI, BRANDÃO e SILVA, 2004).

Após as abstrações iniciais, o processo de escoamento superficial começa, podendo ser expresso pela reorganização da Equação 1, mostrada na Equação 2.

$$ES = \frac{(PT - 0,2S)^2}{(PT + 0,8S)} \quad (2)$$

A análise de vários hidrógrafas de diferentes bacias possibilitou obter S por meio da Equação 3.

$$S = \frac{25.400}{CN} - 254 \quad (3)$$

em que CN é o número da curva (*Curve Number*).

A partir das características descritivas do mapa de tipologias de solos (Figura (IV) 4) estabeleceu-se sua categorização conforme os grupos A, B, C ou D, definidos pelo método *NRCS-CN*. Segundo a classificação, de A até D, os grupos apresentam taxas decrescentes de infiltração da água no solo. Ou seja, os solos arenosos encontram-se no grupo A enquanto os mais argilosos no grupo D. Após essa etapa, o mapa de uso e cobertura foi empregado para definir o número de curva para cada grupo. O valor do CN composto adotado para cada área de captação foi efetivamente determinado pela porcentagem média observada.

4.3.3 Método de propagação de cheias em canais

É importante ressaltar que os modelos disponíveis no HEC-HMS são exclusivos para simular a propagação de cheias em canais abertos, não sendo indicados para outras situações. Dentre as opções de modelos disponíveis para simular a propagação de cheias em canais, adotou-se o modelo de *Muskingum-Cunge* para este estudo. É o mais indicado para a maior parte das situações conforme consta na tabela de orientação para seleção de modelos de propagação em canais do Manual Técnico de Referência do HEC-HMS (USACE, 2000). Sua única e exclusiva limitação em relação a outros métodos está na capacidade de representar o efeito de remanso provocado por reservatórios, quando este for significativo. Isto porque o método assume condição de fluxo uniforme. Na região de estudo encontra-se a Usina da Brecha, mas como é a fio d'água, seu reservatório não configura fator impeditivo ao emprego do método de *Muskingum-Cunge* inclusive nesse trecho do rio.

O único método disponível para simular efeitos de remanso no HEC-HMS é o método de *Puls* Modificado, quando se conhece a relação de volume-vazão (USACE, 2000). Entretanto, mesmo esse método apresenta limitações. A estrutura do HMS utiliza um modo de processamento do tipo montante-jusante, sendo limitado para contemplar efeitos de remanso à jusante.

O método de *Muskingum-Cunge* é parametrizado pelas dimensões da seção do canal, seu comprimento e declividade média, além do coeficiente de rugosidade de Manning n . Os valores desses parâmetros, apresentados na Tabela (IV) 1, foram inicialmente obtidos através de análises espaciais, visitas a campo e consulta a imagens de satélites.

Tabela (IV) 1 – Parâmetros do Método de Propagação de cheias em canais *Muskingum-Cunge*

Canal	Comprimento aproximado (m)	Declividade (m/m)	Manning's n	Forma da seção	Largura média (m)
Canal_P	35.466,12	0,095	0,535	Retangular	30
Canal_BP	28.614,84	0,129	0,190	Retangular	15
Canal_SF	16.949,65	0,113	0,170	Retangular	5
Canal_FV	12.943,25	0,123	0,447	Retangular	5
Canal_S	14.475,30	0,117	0,474	Retangular	6
Canal_BPSF	24.193,95	0,124	0,494	Retangular	30

Canal	Comprimento aproximado (m)	Declividade (m/m)	Manning's n	Forma da seção	Largura média (m)
Canal_PF	24.359,58	0,105	0,613	Retangular	45
Canal_PFPN	42.349,70	0,114	0,732	Retangular	50
Canal_FVS	32.078,76	0,109	0,160	Retangular	10
Canal_PNJ	79.654,35	0,089	0,732	Retangular	60

Fonte: o autor.

4.3.4 Método de transformação chuva-deflúvio

O método de transformação chuva-deflúvio é responsável por simular o processo de escoamento superficial oriundo do excesso de precipitação em uma bacia hidrográfica (USACE, 2000). O software HEC-HMS oferece dois métodos para este propósito, através de modelos empíricos como o Hidrograma Unitário (HU) e um modelo conceitual, da Onda Cinemática. O conceito fundamental do HU assume que o processo de escoamento é linear, de modo que um valor de escoamento superior ou inferior a uma unidade é simplesmente um múltiplo da unidade de escoamento do hidrograma (USACE, 2000). O HU é determinado por uma representação discreta do escoamento potencial (precipitação excedente), calculado para cada intervalo de tempo, conforme a Equação 4 (USACE, 2000).

$$Q_n = \sum_{m=1}^{n \leq M} P_m U_{n-m+1} \quad (4)$$

em que Q_n é a ordenada da hidrógrafa no instante $n\Delta t$; P_m é a altura precipitada excedente no intervalo de $m\Delta t$ até $(m+1)\Delta t$; M é total de intervalos discretos de precipitação; e U_{n-m+1} é a ordenada no instante $(n-m+1)\Delta t$.

As pesquisas do NRCS sugerem que a relação entre o pico do HU e tempo de ascensão é expressa pela Equação 5.

$$U_p = C \frac{A}{T_p} \quad (5)$$

em que A é a área de drenagem, C é uma constante de conversão, igual a 2,08 no SI¹³; e T_P é o tempo de ascensão. O tempo de ascensão é relacionado à unidade de precipitação excedente através da Equação 6.

$$T_P = \frac{\Delta t}{2} + t_{lag} \quad (6)$$

em que Δt é a duração da precipitação no período da simulação e t_{lag} é a diferença entre o centro de massa de precipitação excedente e o pico de vazão do HU (USACE, 2000).

Neste estudo empregou-se o método do Hidrograma Unitário do NRCS (*NRCS-Unit Hydrograph*). O intervalo da precipitação correspondeu ao período dos eventos avaliados na etapa de simulação e avaliação dos modelos. A opção pelo método do NRCS-HU para simulação do processo de transformação chuva-deflúvio se deveu por ser facilmente parametrizável e ainda, adicionalmente, sua estimativa inicial poder ser calibrada por registros históricos de estações fluviométricas.

4.3.5 Calibração dos modelos

A partir de registros fluviométricos é possível calibrar parâmetros dos modelos adotados nas simulações por aferição entre valores simulados e observados (USACE, 2000). Obviamente que os valores iniciais dos parâmetros dos modelos devem ser escolhidos com cautela para aperfeiçoar o processo de calibração, os quais devem pelo menos apresentar referência em literatura. Os registros fluviométricos são, portanto, essenciais para calibração dos modelos.

A calibração de parâmetros no software HEC-HMS é feita em duas etapas, na primeira duas hidrógrafas são comparadas conforme critérios de uma função objetivo (i.e., *Peak-weighted RMS error*, *Sum of absolute errors*, *Sum of squared residuals*, *Percent error in peak*). Uma hidrógrafa refere-se aos valores de vazão observados enquanto a outra representa os simulados a partir das estimativas iniciais dos parâmetros. Na segunda etapa, um algoritmo de busca (i.e., *Univariate-Gradient*

¹³ Sistema Internacional

Algorithm ou *Nelder and Mead Algorithm*) procura matematicamente por valores alternativos dos parâmetros que minimizem os valores da função objetivo (USACE, 2000).

Utilizou-se a função objetivo *Peak-weighted RMS error* e o algoritmo de busca *Nelder and Mead* para calibração dos parâmetros dos modelos. Essa função objetivo foi escolhida por ser uma medida implícita de comparação da magnitude dos picos, volumes e tempos de ascensão de duas hidrógrafas (USACE, 2000). O algoritmo de busca de *Nelder and Mead* foi escolhido por avaliar todos os parâmetros simultaneamente e determinar qual parâmetro ajustar, diferente do *Univariate-Gradient*, que calcula e ajusta um único parâmetro por vez (KUMAR e BHATTACHARJYA, 2011). Foi utilizado um evento para calibrar modelo referente a uma tempestade que ocorreu de janeiro a fevereiro de 1979.

4.3.6 Avaliação e validação do desempenho do modelo

Para avaliar o desempenho do modelo, foi empregado o método estatístico conhecido como Coeficiente de Eficiência de Nash e Sutcliffe (NSE), frequentemente utilizado para esse propósito (ROCHA et al., 2012). O coeficiente NSE é apresentado na Equação 7. Seus valores oscilam entre $-\infty$ a 1, indicando melhor ajuste quanto maior for seu valor, ou seja, NSE igual a 1 indica um ajuste perfeito.

$$NSE = 1 - \frac{\sum_{i=1}^n (E_{obs} - E_s)^2}{\sum_{i=1}^n (E_{obs} - \bar{E})^2} \quad (7)$$

em que *NSE* é o valor do coeficiente de Nash e Sutcliffe, E_{obs} são os valores do evento observado, E_s são os valores do evento simulado e \bar{E} é a média dos valores do evento observado. O coeficiente de *NSE* mede a adequação do ajuste gráfico dos valores simulados em comparação com os observados (MEENU, REHANA e MUJUMDAR, 2012)

Adicionalmente utilizou-se o coeficiente de determinação R^2 , calculado pela Equação 8, que indica a correlação entre os valores observados e simulados (MEENU, REHANA e MUJUMDAR, 2012).

$$R^2 = \frac{\sum(Q_{obs} - \overline{Q_{obs}}) \cdot (Q_{sim} - \overline{Q_{sim}})}{\sqrt{[\sum(Q_{obs} - \overline{Q_{obs}})^2 \cdot \sum(Q_{sim} - \overline{Q_{sim}})^2]}} \quad (8)$$

onde Q_{sim} é o valor simulado, Q_{obs} é o valor observado, $\overline{Q_{sim}}$ é a média do valor simulado, e $\overline{Q_{obs}}$ é a média do valor observado. O coeficiente de determinação R^2 varia de 0.0 a 1.0, respectivamente o pior e o melhor ajuste.

Para validação do modelo foi utilizado um evento de precipitação de longa duração referente a uma tempestade de dezembro de 1996 a janeiro de 1997.

5 Resultados e discussões

Os resultados da simulação dos processos hidrológicos são apresentados em relação aos procedimentos de calibração dos parâmetros do modelo e também de avaliação do seu desempenho. Para isso foram utilizados registros pluviométricos e fluviométricos diários de eventos de inundação distintos e conhecidos.

A configuração topológica empregada para representação física da bacia do rio Piranga no processo de simulação hidrológica é apresentada na ilustração da Figura (IV) 5. Os elementos de representação topológica adotados foram: sub-bacias, canais e nós. Para cada sub-bacia foi produzida uma hidrógrafa como resultado do escoamento em sua foz, que por sua vez foi conectada a outro elemento (nó ou canal), em relação ordinal unária com o elemento à jusante. Desse modo, a estrutura conceitual resultante foi semelhante à representação tipo árvore na qual o último elemento é o exutório da bacia.

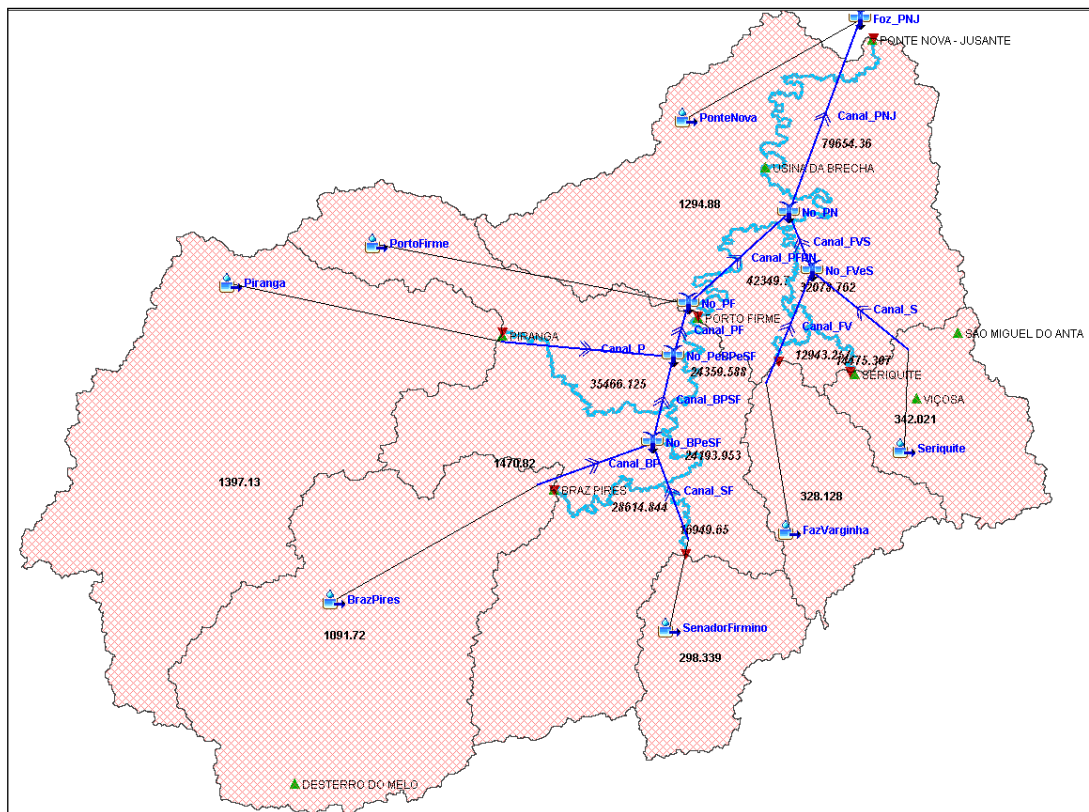


Figura (IV) 5 – Modelo de representação topológica da bacia do rio Piranga no HEC-HMS

O modelo meteorológico utilizado para descrever a precipitação foi o do ietograma especificado pelo usuário, cujos valores diários para cada evento simulado foram obtidos pelos mapas de precipitação, para cada área de drenagem, conforme descrito no item 4.3.1.

5.1 Calibração de modelos

Os parâmetros calibrados foram CN , I_a e o $NRCS-Lag$, respectivamente o número da curva, as abstrações iniciais e o $Lag-time$, esse último derivado do tempo de concentração¹⁴. Tais parâmetros referem-se aos métodos que representam dois processos distintos, o escoamento superficial e a transformação chuva-deflúvio. À exceção de CN , os outros são os parâmetros comumente calibrados quando se utiliza

¹⁴ Aproximadamente 60% do tempo de concentração. Vários estudos indicam que T_{lag} geralmente varia entre 0,5 e 0,75 * T_c . (USACE, 2000)

os métodos adotados (VERMA, JHA e MAHANA, 2010). Apesar de não ser comum otimizar valores de CN, esse foi o único modo de se alcançar alguma concordância entre as hidrógrafas observada e simulada, conferindo razoabilidade quanto a qualidade da etapa de calibração do modelo.

Na Tabela (IV) 2 estão apresentados os códigos das estações fluviométricas cujos nomes identificaram cada área de drenagem e os valores da série histórica de vazões utilizados nas etapas de calibração e validação da simulação. Encontram-se também os valores iniciais dos parâmetros dos métodos para quantificação do escoamento superficial e transformação chuva-deflúvio, respectivamente o Número da Curva e Hidrograma Unitário para cada área de drenagem.

Tabela (IV) 2 – Identificação das áreas de drenagem que compõem a bacia do rio Piranga e valores iniciais e otimizados dos parâmetros.

Área de drenagem			Parâmetros					
Código	Nome	km ²	CN	Iniciais		Otimizados		
				Ia (mm)	Lag (min)	CN	Ia (mm)	Lag (min)
56028000	Piranga	1397	79.0	10.0	363.6	79.1	16.1	1626.1
56055000	Braz Pires	1091	77.2	10.0	333.1	77.1	14.1	1590.1
56065000	Senador Firmino	298	81.8	10.0	199.7	82.1	18.1	798.1
56090000	Fazenda Varginha	328	82.2	100.0	205.6	52.1	243.1	631.1
56085000	Seriquite	342	83.1	100.0	207.2	63.1	368.1	624.1
56075000	Porto Firme	1470	77.8	100.0	374.3	48.1	204.4	980.1
56110005	Ponte Nova Jusante	1294	78.3	100.0	364.9	48.1	233.1	830.1

Fonte: o autor.

A ilustração da Figura (IV) 6 apresenta duas hidrógrafas, uma de valores observados e a outra de valores simulados no processo de calibração. São exibidos ainda os valores dos coeficientes, de NSE e de determinação R^2 , respectivamente 0,53 e 0,75.

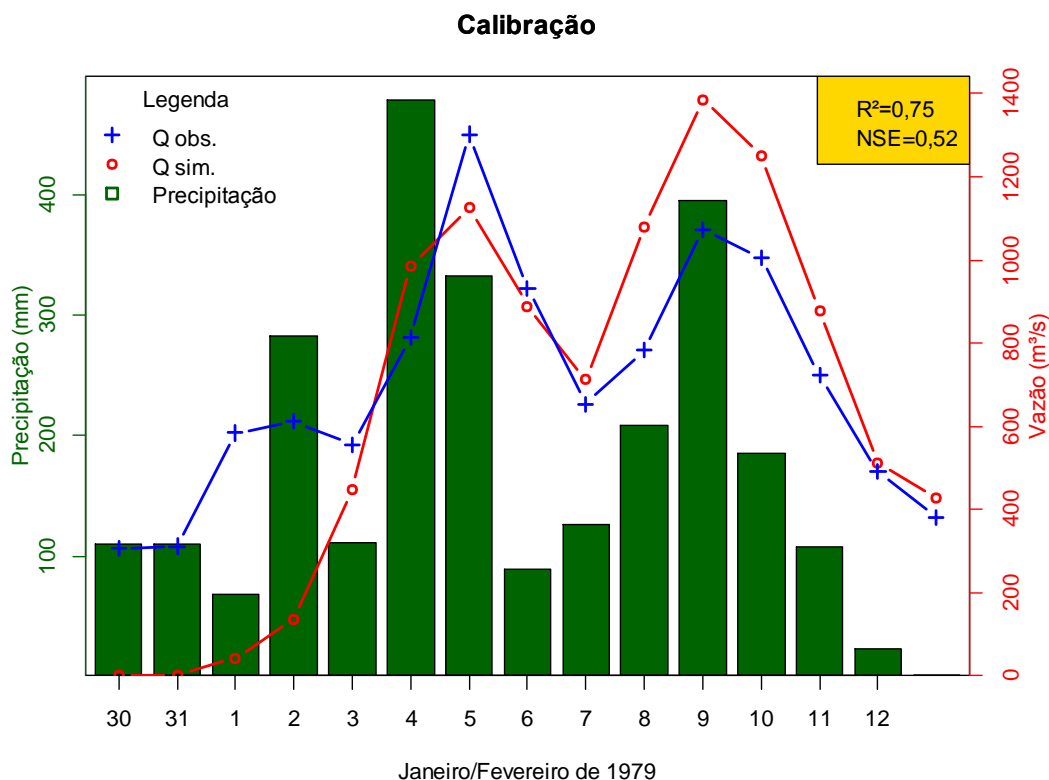


Figura (IV) 6 – Resultado de simulação hidrológica da etapa de calibração de parâmetros dos modelos. A linha azul refere-se os valores de vazão observados enquanto a linha vermelha é referente aos simulados, barras em cor verde representam alturas precipitadas diárias na bacia durante o período do evento.

Considerou-se para efeito da avaliação quantitativa da simulação, somente valores a partir do 5 dia (inclusive) por caracterizar o comportamento de ascensão do pico mais significativo e de maior magnitude do evento. Pelo gráfico apresentado na Figura (IV) 6 percebe-se que a ascensão da hidrógrafa simulada é suave nos quatro primeiros dias, não conseguindo acompanhar o comportamento dos valores de vazão observados nesse período. Após, no entanto, sua trajetória segue com valores próximos aos da hidrógrafa de valores observados, notadamente no primeiro pico, apresentando coincidência com os tempos de ascensão, porém, subestimando a magnitude alcançada, em relação ao observado. Diferente do primeiro pico, o segundo exibiu mais discrepância, com evidente superestimação do valor simulado em relação ao observado.

Segundo os critérios de avaliação da qualidade da simulação, sejam qualitativos pela inspeção visual comparativa entre as hidrógrafas observada e simulada, sejam quantitativos pelos coeficientes de NSE e de determinação R^2 , considerou-se os resultados obtidos para calibração adequados para representar o comportamento da

bacia do rio Piranga frente a um evento de precipitação de expressiva magnitude. Portanto, os parâmetros calibrados foram utilizados para validação do modelo em um evento distinto.

5.2 Validação do modelo

Para validação do modelo foi identificado e utilizado evento de inundação associado a uma chuva de longa duração, iniciada em 30 de dezembro 1997 e finalizada em 9 de janeiro de 1998. Os resultados da validação estão apresentados resumidamente no gráfico da Figura (IV) 7. Pode-se verificar que a vazão simulada apresentou ascensão muito semelhante à vazão observada, entretanto, a magnitude do seu pico foi da ordem 1.954,5 m³/s e ocorreu em 6 de janeiro de 1998, um dia após o pico de vazão observado de 1.647 m³/s, isto é, o superestimou em aproximadamente 15%. Convém ressaltar que o valor de vazão simulado em 5 de janeiro de 1998, dia do pico de vazão observado, foi apenas 7% superior, atingindo 1.776,1 m³/s. Apesar do decaimento da hidrógrafa simulada ter divergido da hidrógrafa observada em volume, os ângulos de inclinação e decaimento de ambas foram muito semelhantes, o que significa dizer que a propagação da onda de cheia foi bem representada pelo modelo e parte dessa concordância deve-se ao condicionamento do MDE à hidrografia mapeada.

Os resultados em relação ao cômputo dos volumes também foram semelhantes, com valores simulados aproximadamente 10% superiores aos observados, respectivamente 941.921,5 (x1000) e 842.270,4 (x1000) m³. Conforme pode ser verificado na ilustração da Figura (IV) 7, as áreas das hidrógrafas foram semelhantes. Apesar disso, todos os valores simulados em relação à magnitude do pico e volume escoado foram superestimados em relação aos observados, mas essa diferença foi percentualmente pequena. Tais resultados foram promissores uma vez que não se pode negligenciar o fato dos registros hidrometeorológicos serem médias diárias, o que impõe uma incerteza intrínseca aos dados e, conseqüentemente ao processo de simulação e também ao modelo para representação de eventos extremos.

A avaliação da qualidade do desempenho do modelo em relação aos coeficientes de NSE e de determinação R² forneceu bons resultados, com valores de 0,64 e 0,88,

respectivamente. Por comparação pode-se dizer que os resultados obtidos com a validação apresentaram qualidade superior àqueles obtidos com a etapa de calibração do modelo.

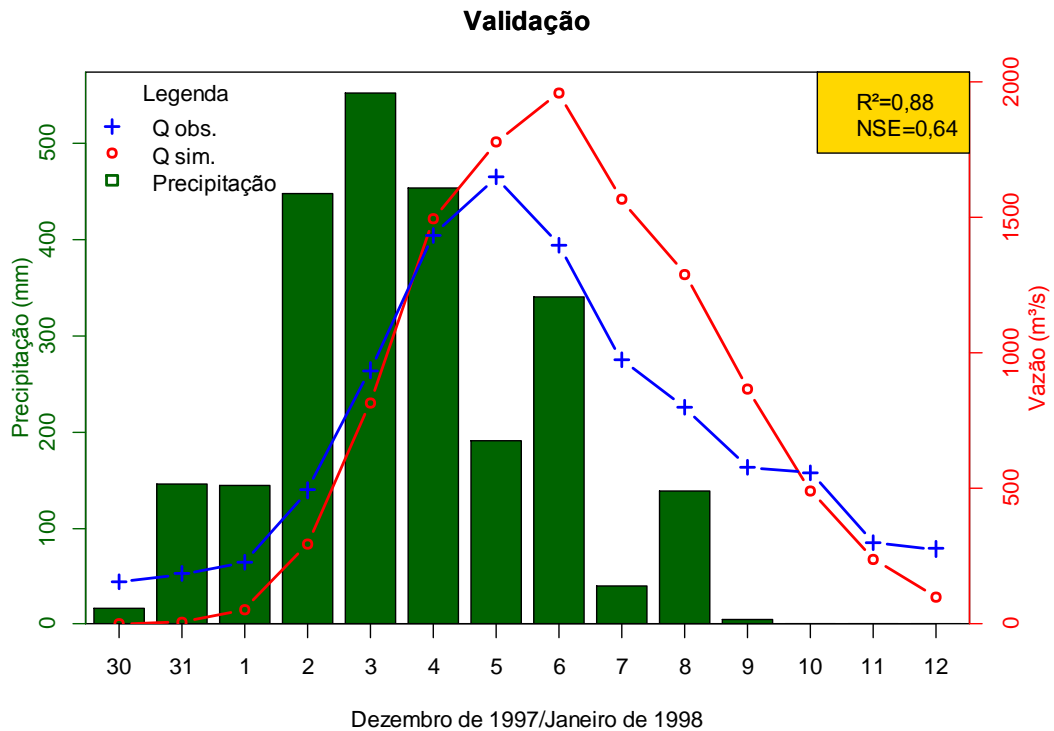


Figura (IV) 7 – Resultado de simulação hidrológica da etapa de validação dos modelos. A linha azul refere-se os valores de vazão observados, a linha vermelha é relativa aos valores de vazão simulados, e as barras em cor verde representam alturas precipitadas diárias na bacia.

5.3 Limitações do estudo

Verificou-se que um fator limitador para caracterizar eventos de cheia mais expressivos diz respeito à disponibilidade de registros históricos de vazão relativos aos eventos de inundação, devido principalmente às circunstâncias relacionadas à operação de medição. Em alguns casos não é possível realizar a observação, de modo que os registros associados à inundação não são preenchidos, portanto não estão disponíveis na série histórica daquela estação fluviométrica. Além disso, como os registros hidrometeorológicos são valores médios diários, apresentam algum grau de incerteza inerente e esta é consequentemente propagada para as análises dos

processos hidrológicos. Por exemplo, parâmetros como o tempo de concentração são dificilmente estimados ou calibrados com confiabilidade e grau de precisão desejável.

Uma vez que houve diferença em torno de treze anos entre a data de aquisição da imagem de uso e cobertura do solo e a data do evento simulado para validação dos parâmetros e, o método de representação do escoamento superficial utilizado é sensível às condições de cobertura do solo, alguma imprecisão relativa a essa informação foi incorporada nos resultados, ainda que em pequena proporção. Isto porque as características topográficas e socioeconômicas da região não tem sido propícias para promover expressivas alterações no modelo regional de exploração do solo nesse intervalo.

Outra fonte de imprecisão diz respeito às diferentes escalas dos mapas da base de dados espaciais. Suspeita-se que influenciaram preponderantemente nos resultados, especialmente para as áreas de drenagem menores, por serem mais sensíveis às escalas pequenas em relação à representatividade da informação extraída para parametrização dos modelos.

6 Conclusões e recomendações

Os resultados alcançados permitem atestar que os objetivos propostos foram satisfeitos com êxito, pois foi possível simular dois eventos críticos de inundação, um para calibração de modelos e o outro para avaliação do seu desempenho. Para validação dos modelos reproduziu-se uma vazão de inundação histórica com boa acurácia, o que permitiu avaliar as situações ambientais relacionadas ao evento. Além disso, o uso do ambiente dos Sistemas de Informações Geográficas como ferramenta de suporte para as análises espaciais e hidrológicas viabilizou a obtenção de parâmetros para simulação de processos hidrológicos com precisão e agilidade. Ao se considerar a base de dados, metodologia e condições de condução do estudo foi possível concluir que:

- a obtenção de parâmetros para os modelos permitiu verificar que as condições naturais da bacia relacionadas a vários fatores, tais como a interconexão de

canais sinuosos aliada à topografia acidentada de declividade acentuada, as características dos tipos de solos com baixa taxa de infiltração e os elevados volumes precipitados favoreceram conjuntamente as condições de inundação na região, independente da capacidade de alteração do ambiente provocada pelo homem;

- mesmo ao se considerar as incertezas inerentes ao processo de medição hidrometeorológica, as diferenças de escalas da base cartográfica utilizada na extração de parâmetros para os modelos hidrológicos e a ordem de grandeza dos eventos simulados, pode-se afirmar que foi possível avaliar satisfatoriamente o comportamento da bacia do rio Piranga frente a um evento de inundação, pois as discrepâncias percentuais entre valores simulados e observados foram pequenas;
- o modelo foi otimizado para estimar vazões de inundação provenientes de chuvas de longa duração, a partir de registros históricos de precipitações que causaram graves consequências na bacia, portanto, espera-se que seja capaz de prever volumes de escoamento superficial com maior precisão para eventos que apresentem características semelhantes às utilizadas na sua calibração;

Recomenda-se que os modelos do Numero da Curva e do Hidrograma Unitário do NRCS sejam utilizados para simulação de processos hidrológicos de escoamento superficial e transformação chuva-deflúvio relativos a evento extremo na bacia hidrográfica do rio Piranga. Trabalhos futuros devem se direcionar a técnicas para elevar a acurácia dos modelos, através, por exemplo, de meios de obtenção das medidas hidrométricas em intervalos discretos, menores que um dia. Outros métodos e modelos para simulação dos processos hidrológicos na bacia do rio Piranga devem ser testados e avaliados.

7 Referências

AL-ABED, N.; ABDULLA, F.; KHYARAH, A. A. GIS-hydrological models for managing water resources in the Zarqa River basin. **Environmental Geology**, 47, 2005. 405-411.

CONSÓRCIO ECOPLAN-LUME. **Plano Integrado de Recursos Hídricos da Bacia Hidrográfica do Rio Doce e Planos de Ações para as Unidades de Planejamento e Gestão de Recursos Hídricos no Âmbito da Bacia do Rio Doce.** [S.l.], p. 478. 2010.

COOK, A.; MERWADE, V. Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping. **Journal of Hydrology**, 377, 2009. 131-142.

ELFERT, S.; BORMANN, H. Simulated impact of past and possible future land use changes on the hydrological response of the Northern German lowland 'Hunte' catchment. **Journal of Hydrology**, v. 383, p. 245-255, 2010.

FERNANDES FILHO, E. I.; et al. **Levantamento de solos e aptidão agrícola da porção mineira da bacia do Rio Doce.** Belo Horizonte: Fundação Estadual de Meio Ambiente (FEAM), 2010.

HENGL, T.; HEUVELINK, G. B. M.; VAN LOON, E. E. On the uncertainty of stream networks derived from elevation data: the error propagation approach. **Hydrology and Earth Systems Science**, v. 14, p. 1153-1165, 2010.

HOUAISS, A.; VILLAR, M. S.; FRACO, F. M. M. **Dicionário Houaiss da língua portuguesa.** Rio de Janeiro: Objetiva, 2009.

KUMAR, D.; BHATTACHARJYA, R. K. Distributed Rainfall Runoff Modeling. **International Journal of Earth Sciences and Engineering**, 4, n. 6, October 2011. 270-275.

KUMAR, R.; SAMANIEGO, L.; ATTINGER, S. The effects of spatial discretization and model parametrization on the prediction of extreme runoff characteristics. **Journal of Hydrology**, 2010. ISSN doi: 10.1016/j.jhydrol.2010.07.047.

MEENU, R.; REHANA, S.; MUJUMDAR, P. P. Assessment of hydrologic impacts of climate change in Tunga-Bhadra river basin, India with HEC-HMS and SDSM. **Hydrological Process**, 2012.

MENDAS, A. The contribution of the digital elevation models and geographic information systems in a watershed hydrologic research. **Applied Geomatics**, 2, 2010. 33-42.

MOORE, I. D. et al. GIS and Land Surface Subsurface Process. In: GOODCHILD, M. F.; PARKS, B. O.; STEYAERT, L. T. **Environmental Modeling with GIS.** [S.l.]: [s.n.], 1993. p. 196-230.

OLANG, L. O.; FÜRST, J. Effects of land cover change on flood peak discharges and runoff volumes: model estimates for the Nyando River Basin, Kenya. **Hydrological Process**, 25, 2011. 80-89.

PRUSKI, F. F.; BRANDÃO, V. S.; SILVA, D. D. **Escoamento Superficial**. 2. ed. [S.l.]: UFV, 2004. 87 p. ISBN ISBN: 85-7269-154-5.

RIBEIRO, C. A. A. S.; MEITNER, M. J.; VEIGA, M. M. **Environmental protection in Brazil: where the truth lies**. Proceedings of 26th ESRI International User Conference. San Diego, CA: [s.n.]. 2006. Acesso em: 5 Mai. 2011. Disponível em.

ROCHA, E. O. et al. The contribution of conservation practices in reducing runoff, soil loss, and transport of nutrients at the watershed level. **Water Resources Management**, 26, 2012. 3831-3852.

SAGHAFIAN, B. et al. Flood intensification due to changes in land use. **Water Resources Management**, 22, 2008. 1051-1067.

THOMMERET, N.; BAILLY, J. S.; PUECH, C. Extraction of thalweg networks from DTMs: application to badlands. **Hydrology and Earth Systems Science**, v. 14, p. 1527-1536, 2010.

USACE. **Hydrologic Modeling System HEC-HMS: Technical Reference Manual**. U.S. Army Corps of engineers. [S.l.], p. 145. 2000.

VERMA, A. K.; JHA, M. K.; MAHANA, R. K. Evaluation of HEC-HMS and WEPP for simulating watershed runoff using remote sensing and geographical information system. **Paddy and Water Environment**, 8, 2010. 131-144.

YUSOP, Z.; CHAN, C. H.; KATIMON, A. Runoff characteristics and application of HEC-HMS for modeling stormflow hydrograph in an oil palm catchment. **Water Science & Technology**, 8, 2007. 41-48.

V. CONCLUSÕES GERAIS

Esta pesquisa apresenta sucintamente, na forma de estudos individuais, os principais métodos aplicados aos estudos hidrológicos conduzidos na bacia hidrográfica do rio Piranga.

Os resultados permitiram concluir que:

- o método de análise das séries históricas de precipitações diárias máximas através de modelos hierárquicos por inferência Bayesiana viabilizou a obtenção de dois modelos de representação regional, um adequado à região do alto/médio e o outro à região do baixo rio Piranga. Esses modelos apresentaram a vantagem de fornecer estimativas de altura máxima diária precipitada associada aos maiores quantis da distribuição, ou seja, a elevados períodos de retorno, com variância pequena. Essas estimativas são úteis para suportar projetos de obras de engenharia na região;
- a abordagem metodológica de introdução de informações contextuais para aperfeiçoar o processo de classificação de imagens de satélite através da especificação de modelos digitais de representação da paisagem por dois métodos distintos, a avaliação multicritérios e a análise de regressão logística, não apresentou desempenho significativamente superior à abordagem convencional (i.e., por análise de variância das amostras de treinamento das categorias a serem classificadas). Entretanto, a avaliação multicritérios apresentou melhor desempenho e facilidade de aplicação que o método por análise de regressão logística. O resultado desse estudo foi a produção de uma imagem de uso e cobertura do solo referente ao mês de agosto de 2010 para a região da bacia hidrográfica do rio Piranga;
- o estudo sobre simulação de processos hidrológicos de inundação com uso de SIG para manipulação da base de dados e obtenção de parâmetros para posterior utilização no HEC-HMS, possibilitou obter um modelo de chuva-vazão para representação do comportamento da bacia durante uma chuva de longa duração causadora de enchente (segundo registros fluviométricos históricos). Uma vez que o modelo obtido permitiu reproduzir um evento de inundação conhecido com boa acurácia, poderá subsidiar órgãos e gestores de recursos hídricos da bacia do rio Piranga na análise e avaliação de medidas estruturais e não estruturais para atenuação dos efeitos de uma inundação;

- a abordagem proposta para análise hidrológica de bacias hidrográficas através das frentes: i) caracterização da precipitação regional para verificar a distribuição espacial das chuvas; ii) identificação do uso e cobertura do solo para distinguir a condição predominante; e iii) simulação de chuva-vazão com enfoque em eventos extremos, oferece mecanismos para a compreensão dos diferentes atores associados às inundações sob uma perspectiva integrada;
- exceto pelos SIGs proprietários ArcGIS, Erdas e Idrisi, os demais sistemas utilizados em situações particulares (i.e., R, OpenBUGS, HEC-HMS) foram softwares livres. Com alguma adaptação pode-se aplicar a metodologia de tratamento dos dados espaciais em SIGs livres.

APÊNDICE II

Apêndice (II) A – Série anual de eventos máximos diários de precipitação para cada estação pluviométrica em análise

Ano*	e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	e11	e12	e13	e14	e15
41/42	62.0	102.4	56.2	123.0	114.0	63.0	-	-	-	-	-	-	-	-	-
42/43	60.2	117.8	61.6	101.0	91.0	97.6	-	-	-	-	-	-	-	-	-
43/44	75.7	104.8	76.0	54.2	138.0	75.0	-	-	-	-	-	-	-	-	-
44/45	84.6	79.4	66.0	53.4	92.0	60.0	-	-	-	-	-	-	-	-	-
45/46	54.4	84.3	96.0	70.2	75.0	95.0	-	-	-	-	-	-	-	-	-
46/47	60.0	29.4	63.8	105.2	98.0	100.2	-	-	-	-	-	-	-	-	-
47/48	114.5	38.3	91.0	90.3	100.0	54.0	-	-	-	-	-	-	-	-	-
48/49	105.8	39.4	102.8	66.2	77.0	93.0	-	-	-	-	-	-	-	-	-
49/50	82.0	78.1	56.2	68.3	58.2	63.0	-	-	-	-	-	-	-	-	-
50/51	125.0	92.0	89.8	87.4	75.2	82.6	-	-	-	-	-	-	-	-	-
1/52	107.0	106.2	97.6	50.5	87.2	74.2	-	-	-	-	-	-	-	-	-
52/53	110.8	87.2	84.2	58.2	145.0	63.2	-	-	-	-	-	-	-	-	-
53/54	94.4	53.3	64.7	60.8	132.0	86.4	63.3	-	-	-	-	-	-	-	-
54/55	96.6	52.0	70.4	70.5	44.2	78.2	62.0	-	-	-	-	-	-	-	-
55/56	70.6	54.0	70.3	50.0	125.0	73.2	52.9	-	-	-	-	-	-	-	-
56/57	91.0	99.3	99.4	80.0	75.2	52.4	80.2	-	-	-	-	-	-	-	-
57/58	95.0	48.3	95.4	85.0	72.4	58.0	40.5	-	-	-	-	-	-	-	-
58/59	81.2	88.4	94.2	60.0	71.4	53.2	40.2	-	-	-	-	-	-	-	-
59/60	98.4	94.0	82.1	160.0	67.2	83.5	70.2	52.3	-	-	-	-	-	-	-
60/61	96.8	93.0	80.0	240.0	105.4	105.5	29.4	92.2	-	-	-	-	-	-	-
61/62	81.8	74.0	62.0	105.2	79.4	90.4	32.3	94.0	-	-	-	-	-	-	-

Ano*	e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	e11	e12	e13	e14	e15
62/63	24.0	72.0	50.5	41.0	67.2	31.0	28.0	81.2	-	-	-	-	-	-	-
63/64	61.6	58.0	91.8	80.0	83.0	50.2	25.3	61.0	-	-	-	-	-	-	-
64/65	122.2	76.0	106.2	101.3	90.2	43.2	79.0	63.5	-	-	-	-	-	-	-
65/66	122.0	100.3	122.8	110.3	94.2	49.8	132.0	56.5	84.3	-	-	-	-	-	-
66/67	86.6	38.3	78.6	102.1	70.4	47.6	70.0	103.2	87.6	34.0	42.8	-	27.4	-	-
67/68	57.7	82.4	55.8	82.4	121.0	42.4	100.0	100.2	58.8	83.2	43.3	24.0	37.2	90.0	-
68/69	72.4	81.0	61.2	120.0	115.8	81.0	76.1	82.4	75.4	42.8	77.6	25.0	38.2	59.2	-
69/70	105.0	59.4	137.2	60.4	76.2	75.4	71.5	70.0	84.0	29.6	50.2	81.8	76.6	56.8	-
70/71	46.3	80.0	63.3	98.0	55.0	49.8	59.5	71.6	60.0	29.8	12.2	81.7	12.6	75.7	-
71/72	99.4	123.3	60.0	85.0	40.0	90.2	64.2	60.4	78.3	69.6	69.0	83.7	62.2	75.8	-
72/73	82.8	129.0	86.0	74.2	110.2	116.4	72.3	87.4	75.0	89.8	58.0	68.7	65.6	70.0	-
73/74	81.2	81.5	78.0	89.3	42.5	74.8	47.2	61.4	47.0	56.2	52.0	85.6	51.6	45.6	-
74/75	66.8	54.0	70.0	75.0	60.1	66.0	56.4	112.4	50.2	89.6	39.0	70.8	69.9	106.0	77.3
75/76	49.6	82.2	64.0	104.2	63.0	111.5	82.1	94.0	40.8	68.4	66.0	67.8	77.8	93.0	70.0
76/77	86.2	98.8	53.2	54.6	69.0	90.0	73.3	89.0	50.2	62.2	76.0	73.6	56.8	84.4	100.2
77/78	54.8	112.5	67.0	81.7	58.4	64.2	89.0	99.2	69.2	46.2	77.0	55.8	62.8	73.6	58.6
78/79	83.4	121.0	67.5	88.2	100.4	93.2	121.4	79.4	63.6	90.0	88.1	120.0	75.4	86.6	86.6
79/80	71.2	61.2	79.1	100.0	55.1	83.2	73.0	81.2	42.6	208.4	112.0	70.0	76.8	163.9	110.0
80/81	57.2	66.3	61.8	69.4	63.0	66.2	64.0	73.1	49.3	68.8	69.0	60.6	86.4	83.6	88.0
81/82	74.2	93.0	142.2	71.0	62.4	121.0	133.2	91.2	72.5	72.0	59.1	154.0	88.4	112.6	102.0
82/83	94.0	80.2	79.0	75.6	62.8	77.0	72.4	149.2	74.1	70.6	63.0	68.0	52.2	50.0	54.0
83/84	78.2	64.2	77.9	60.0	107.4	68.5	52.0	77.3	35.6	64.6	65.3	53.0	59.8	45.8	98.1
84/85	89.6	82.0	70.8	90.0	100.5	89.0	97.3	98.7	91.2	92.2	72.0	73.0	74.3	100.9	64.0
85/86	105.0	67.2	68.4	55.0	66.4	39.1	92.2	70.3	42.3	134.0	79.0	44.0	118.6	184.8	90.5
86/87	75.2	54.0	57.8	70.2	40.0	57.0	88.0	61.2	84.5	61.8	57.0	13.8	42.0	60.9	65.6
87/88	84.6	128.4	57.4	118.4	85.0	72.6	60.2	87.6	49.3	109.2	57.2	69.4	96.4	110.8	79.4

Ano*	e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	e11	e12	e13	e14	e15
88/89	61.6	91.0	69.8	53.0	79.4	50.5	83.2	74.0	35.4	52.2	62.0	45.0	49.0	56.7	73.0
89/90	53.2	88.0	59.0	118.8	55.2	88.2	120.0	86.6	84.6	74.2	97.0	51.3	117.2	67.3	62.6
90/91	84.2	85.3	90.3	75.1	99.8	100.0	128.0	93.8	56.3	97.6	107.0	53.5	52.4	74.8	71.0
91/92	64.8	141.0	127.2	100.0	53.7	100.0	55.5	74.0	77.6	106.8	58.0	78.2	87.4	49.4	85.3
92/93	100.3	83.0	77.3	101.0	111.5	70.6	83.2	100.3	70.3	57.2	80.6	78.8	62.2	82.0	90.2
93/94	122.2	83.0	96.9	89.1	51.0	95.0	69.0	125.0	56.5	77.2	83.3	72.1	80.6	109.3	89.0
94/95	72.0	70.0	94.1	83.3	55.0	74.0	58.0	38.6	66.0	65.6	64.4	72.1	64.6	66.0	62.8
95/96	82.4	97.5	83.4	81.5	90.0	46.0	80.0	77.0	52.5	77.8	78.0	85.9	57.8	73.0	66.0
96/97	106.4	98.1	103.0	86.0	101.6	75.0	98.3	140.0	88.5	85.4	86.3	95.5	101.2	92.2	117.5
97/98	101.6	135.7	48.7	66.1	69.1	79.0	72.2	53.2	82.1	44.8	88.2	59.5	37.2	133.4	110.2
98/99	79.6	65.3	55.2	49.3	61.0	75.0	117.0	56.0	75.5	102.4	65.2	57.6	58.6	73.3	60.0
99/00	115.5	145.7	98.6	101.0	82.0	120.4	74.5	75.0	77.7	70.0	78.3	86.9	112.4	87.4	118.0
00/01	48.2	102.3	61.5	64.0	117.0	185.0	70.7	58.0	96.4	132.4	80.1	87.7	60.5	108.6	72.3
01/02	68.9	82.6	72.3	62.0	101.3	52.3	63.0	91.4	90.5	52.2	50.1	88.9	78.1	69.1	72.8
02/03	62.2	62.9	76.4	67.8	70.2	70.9	70.6	77.8	77.2	59.6	64.4	73.1	93.3	71.6	158.3
03/04	161.2	120.5	86.8	103.0	113.0	52.2	79.6	96.2	100.0	102.2	80.0	74.5	98.5	133.4	83.3
04/05	83.8	121.3	78.2	68.2	100.0	80.5	64.5	81.0	76.0	93.9	84.0	58.5	75.9	80.3	96.1
05/06	57.7	57.3	54.0	79.0	66.0	95.5	77.1	58.8	63.7	76.7	87.4	61.5	51.5	90.5	63.3
06/07	62.9	94.3	62.1	63.5	74.6	65.7	69.2	61.0	94.5	35.7	69.0	51.9	67.9	92.7	73.5
07/08	65.2	113.7	62.5	74.0	107.3	57.3	72.0	98.7	72.5	63.3	69.2	47.4	86.7	80.6	72.7
08/09	100.4	117.0	93.7	84.4	114.0	96.3	152.8	91.0	148.7	113.8	102.0	67.2	108.5	94.4	79.9
09/10	63.6	83.6	64.3	135.0	66.0	85.3	75.0	83.0	49.3	66.2	56.0	62.6	98.5	80.2	45.5
10/11	114.5	115.0	88.6	94.5	87.0	60.0	72.9	98.5	62.5	76.2	71.8	66.1	99.8	76.4	113.0

Fonte: o autor.

Nota: o ano se refere ao período chuvoso do ano hidrológico, ou seja, de outubro do anterior a março do seguinte.

Apêndice (II) B – Teste diagnóstico de convergência de BGR aplicado ao par de cadeias paralelas dos modelos independentes.

Parâm.	CSRF**			Parâm.	CSRF			Parâm.	CSRF		
	Estimate	0,975	PSRF		Estimate	0,975	PSRF		Estimate	0,975	PSRF
$\eta(e1)$	1,000	1,001	1,000	$\mu(e1)$	1,000	1,000	1,000	$\sigma(e1)$	1,000	1,000	1,000
$\eta(e2)$	1,000	1,000	1,000	$\mu(e2)$	1,000	1,000	1,000	$\sigma(e2)$	1,000	1,001	1,000
$\eta(e3)$	1,000	1,000	1,000	$\mu(e3)$	1,000	1,000	1,000	$\sigma(e3)$	1,000	1,000	1,000
$\eta(e4)$	1,000	1,000	1,000	$\mu(e4)$	1,000	1,000	1,000	$\sigma(e4)$	1,000	1,001	1,000
$\eta(e5)$	1,000	1,000	1,000	$\mu(e5)$	1,000	1,001	1,000	$\sigma(e5)$	1,000	1,000	1,000
$\eta(e6)$	1,000	1,000	1,000	$\mu(e6)$	1,000	1,000	1,000	$\sigma(e6)$	1,000	1,000	1,000
$\eta(e7)$	1,000	1,000	1,000	$\mu(e7)$	1,000	1,000	1,000	$\sigma(e7)$	1,000	1,000	1,000
$\eta(e8)$	1,000	1,001	1,000	$\mu(e8)$	1,000	1,000	1,000	$\sigma(e8)$	1,000	1,000	1,000
$\eta(e9)$	1,000	1,002	1,000	$\mu(e9)$	1,000	1,000	1,000	$\sigma(e9)$	1,000	1,000	1,000
$\eta(e10)$	1,000	1,000	1,000	$\mu(e10)$	1,000	1,000	1,000	$\sigma(e10)$	1,000	1,000	1,000
$\eta(e11)$	1,000	1,000	1,000	$\mu(e11)$	1,000	1,001	1,000	$\sigma(e11)$	1,000	1,000	1,000
$\eta(e12)$	1,000	1,000	1,000	$\mu(e12)$	1,000	1,000	1,000	$\sigma(e12)$	1,000	1,000	1,000
$\eta(e13)$	1,000	1,000	1,000	$\mu(e13)$	1,000	1,000	1,000	$\sigma(e13)$	1,000	1,001	1,000
$\eta(e14)$	1,000	1,001	1,000	$\mu(e14)$	1,000	1,000	1,000	$\sigma(e14)$	1,000	1,001	1,000
$\eta(e15)$	1,000	1,000	1,000	$\mu(e15)$	1,000	1,000	1,000	$\sigma(e15)$	1,000	1,000	1,000

Fonte: o autor.

Notas: Brooks, Gelman, and Rubin convergence diagnostics:

Iterations used = 130001:239991;

Multivariate Potential Scale Reduction Factor = 1.002279+0i

Apêndice (II) C – Teste diagnóstico de convergência de RL aplicado à cadeia unificada dos modelos independentes.

Param.	Thin**	Burn-in	Total	LB	DF.	Param.	Thin	Burn-in	Total	LB	DF.	Param.	Thin	Burn-in	Total	LB	DF.
$\eta(e1)$	1	2	3775	3746	1,008	$\mu(e1)$	1	2	3709	3746	0,990	$\sigma(e1)$	1	2	3801	3746	1,015
$\eta(e2)$	1	2	3767	3746	1,006	$\mu(e2)$	1	2	3717	3746	0,992	$\sigma(e2)$	1	2	3789	3746	1,011
$\eta(e3)$	1	2	3699	3746	0,987	$\mu(e3)$	1	2	3841	3746	1,025	$\sigma(e3)$	1	2	3883	3746	1,037
$\eta(e4)$	1	2	3703	3746	0,989	$\mu(e4)$	1	2	3817	3746	1,019	$\sigma(e4)$	1	2	3823	3746	1,021
$\eta(e5)$	1	2	3773	3746	1,007	$\mu(e5)$	1	2	3777	3746	1,008	$\sigma(e5)$	1	2	3745	3746	1,000
$\eta(e6)$	1	2	3787	3746	1,011	$\mu(e6)$	1	2	3737	3746	0,998	$\sigma(e6)$	1	1	3776	3746	1,008
$\eta(e7)$	1	2	3791	3746	1,012	$\mu(e7)$	1	2	3790	3746	1,012	$\sigma(e7)$	1	2	3821	3746	1,020
$\eta(e8)$	1	2	3721	3746	0,993	$\mu(e8)$	1	2	3669	3746	0,979	$\sigma(e8)$	1	2	3810	3746	1,017
$\eta(e9)$	1	2	3724	3746	0,994	$\mu(e9)$	1	2	3738	3746	0,998	$\sigma(e9)$	1	1	3761	3746	1,004
$\eta(e10)$	1	2	3764	3746	1,005	$\mu(e10)$	1	2	3813	3746	1,018	$\sigma(e10)$	1	2	3720	3746	0,993
$\eta(e11)$	1	2	3693	3746	0,986	$\mu(e11)$	1	2	3714	3746	0,991	$\sigma(e11)$	1	2	3706	3746	0,989
$\eta(e12)$	1	2	3782	3746	1,010	$\mu(e12)$	1	2	3758	3746	1,003	$\sigma(e12)$	1	2	3831	3746	1,023
$\eta(e13)$	1	2	3677	3746	0,982	$\mu(e13)$	1	2	3775	3746	1,008	$\sigma(e13)$	1	1	3759	3746	1,003
$\eta(e14)$	1	2	3703	3746	0,989	$\mu(e14)$	1	2	3726	3746	0,995	$\sigma(e14)$	1	2	3793	3746	1,013
$\eta(e15)$	1	2	3727	3746	0,995	$\mu(e15)$	1	1	3765	3746	1,005	$\sigma(e15)$	1	2	3818	3746	1,019

Fonte: o autor.

Notas: Raftery and Lewis convergence diagnostic:

Quantile = 0.025;

Accuracy = +/- 0.005;

Probability = 0.95;

Chain: coda1::coda2

Apêndice (II) D – Altura da precipitação máxima diária e seu respectivo período de retorno calculado por Análise de Frequência dos registros da série histórica das estações pluviométricas em análise.

T (anos)	Altura precipitada (mm)														
	e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	e11	e12	e13	e14	e15
5	104,3	111,2	94,2	101,2	106,9	94,6	89,6	98,6	84,6	96,8	83,8	95,7	83,7	106,0	100,9
10	114,5	121,3	102,4	117,5	115,6	100,2	120,1	102,2	92,0	110,8	91,3	103,7	88,2	121,8	113,7
20	122,2	131,5	124,5	127,6	127,7	117,9	132,0	129,0	98,4	133,4	105,0	115,3	111,8	153,7	120,1
30	123,7	138,5	132,6	148,4	135,2	120,7	133,5	141,2	113,5	156,6	108,5	117,6	131,3	170,9	141,3
40	129,6	141,6	137,8	170,1	138,9	129,1	140,2	144,7	134,2	189,0	110,7	118,2	146,4	180,2	-
50	139,8	142,9	139,2	192,7	140,9	147,1	146,8	148,2	-	-	-	-	-	-	-
70	160,2	145,6	142,1	237,7	144,8	183,2	-	-	-	-	-	-	-	-	-

Fonte: o autor.

APÊNDICE III

Apêndice (III) A – Parâmetros de avaliação da qualidade e seleção de modelos.

CLASSE	MODELO	ROC	M	-2Ln(L(θ))	AIC	Pseudo R ²
Porte arbóreo alto	$\text{logit}(\theta_1) = -14,2628 + 15,243780*\text{NDVI} + 0,001807*\text{MDD} - 0,001739*\text{MDE}$	0,88	4	5032,5	5040,5	0,17
	$\text{logit}(\theta_2) = -15,5418 + 15,284721*\text{NDVI} + 0,091280*\text{MDN} + 0,067280*\text{MDL} - 0,000086*\text{MDD}$	0,88	5	5220,26	5230,3	0,18
	$\text{logit}(\theta_3) = -15,3436 + 14,848571*\text{NDVI} + 0,148972*\text{MDN}$	0,88	3	4941,3	4947,3	0,17
Porte arbóreo baixo	$\text{logit}(\theta_1) = -6,2305 - 8,142071*\text{NDVI} + 0,042789*\text{MDN} - 0,041278*\text{MDL} + 0,015009*\text{MDD}$	0,85	5	4252,9	4262,9	0,12
	$\text{logit}(\theta_2) = -6,3201 - 7,994183*\text{NDVI} + 0,012219*\text{MDD}$	0,86	3	3799,2	2260,6	0,13
	$\text{logit}(\theta_3) = -6,1080 - 7,548122*\text{NDVI} - 0,131140*\text{MDN}$	0,86	3	3837,6	2215,5	0,12
Eucalipto	$\text{logit}(\theta_1) = -20,0936 + 21,034573*\text{NDVI} - 0,000104*\text{MDD}$	0,95	3	2254,6	2432,3	0,23
	$\text{logit}(\theta_2) = -19,7359 + 20,434242*\text{NDVI} - 0,000084*\text{MDD} + 0,073148*\text{MDL} - 0,140193*\text{MDN}$	0,93	5	2205,5	2233,4	0,22
	$\text{logit}(\theta_3) = -20,8970 + 22,424984*\text{NDVI} + 0,021203*\text{MDL} + 0,081508*\text{MDN}$	0,93	4	2424,3	2224,5	0,25
Pastagem	$\text{logit}(\theta_1) = -7,0560 - 5,931608*\text{NDVI} + 0,014543*\text{MDN}$	0,84	3	2227,4	2381,6	0,05
	$\text{logit}(\theta_2) = -6,9798 - 6,334780*\text{NDVI} - 0,042102*\text{MDL} + 0,162455*\text{MDN} + 0,000260*\text{MDD}$	0,83	5	2214,5	1133,5	0,06
	$\text{logit}(\theta_3) = -6,9703 - 5,981271*\text{NDVI} - 0,007530*\text{MDL} - 0,107056*\text{MDN}$	0,83	4	2373,6	1300,5	0,05
Café	$\text{logit}(\theta_1) = -16,3039 + 15,224074*\text{NDVI} - 0,001286*\text{MDE} + 0,001322*\text{MDD}$	0,75	4	1125,5	1235,2	0,15
	$\text{logit}(\theta_2) = -15,7348 + 13,214145*\text{NDVI} - 0,000209*\text{MDE}$	0,77	3	1294,5	1145,2	0,13
	$\text{logit}(\theta_3) = -16,0442 + 13,349194*\text{NDVI} + 0,122722*\text{MDN} + 0,120144*\text{MDL}$	0,77	4	1227,2	1197,6	0,13

CLASSE	MODELO	ROC	M	-2Ln(L(θ))	AIC	Pseudo R ²
Área agrícola	$\text{logit}(\theta_1) = -7,5658 + 0,000113*\text{MDD} - 7,131872*\text{NDVI}$	0,80	3	1139,2	1156,1	0,10
	$\text{logit}(\theta_2) = -7,4588 - 7,500822*\text{NDVI} + 0,051674*\text{MDN} + 0,249821*\text{MDL} + 0,000128*\text{MDD}$	0,82	5	1187,6	1898,38	0,11
	$\text{logit}(\theta_3) = -7,5015 - 7,432452*\text{NDVI} - 0,067881*\text{MDL}$	0,83	3	1150,1	2190,13	0,12
Mata	$\text{logit}(\theta_1) = -13,3586 + 0,008561*\text{MDD} + 9,140624*\text{NDVI}$	0,86	3	1892,38	1835,363	0,09
	$\text{logit}(\theta_2) = -13,1360 + 8,872734*\text{NDVI} - 0,127769*\text{MDL} - 0,034180*\text{MDN} + 0,011248*\text{MDD}$	0,86	5	2180,13	941,6	0,09
	$\text{logit}(\theta_3) = -13,2598 + 9,303224*\text{NDVI} - 0,055097*\text{MDL}$	0,86	3	1829,363	965,8	0,09
Área queimada	$\text{logit}(\theta_1) = -9,2567 - 9,813520*\text{NDVI} - 0,000168*\text{MDE} + 0,069104*\text{MDD} - 0,375205*\text{MDN}$	0,93	5	931,6	1183,8	0,24
	$\text{logit}(\theta_2) = -9,2940 - 9,758029*\text{NDVI} + 0,068024*\text{MDD} + - 0,198550*\text{MDN}$	0,93	4	957,8	1051,22	0,23
	$\text{logit}(\theta_3) = -7,2414 - 8,837569*\text{NDVI} - 0,053458*\text{MDL}$	0,97	3	1177,8	861,1	0,21
Superfície sombreada	$\text{logit}(\theta_1) = -12,7549 + 0,091338*\text{MDD} - 0,200031*\text{MDL} - 0,291410*\text{MDN}$	0,87	4	1043,22	712,5	0,12
	$\text{logit}(\theta_2) = -11,7173 + 0,083376*\text{MDD} + 0,226448*\text{MDL} - 0,324143*\text{MDN} - 2,637874*\text{NDVI}$	0,86	5	851,1	603,8	0,09
	$\text{logit}(\theta_3) = -11,7831 + 0,082376*\text{MDN} - 3,367385*\text{NDVI} + 0,085957*\text{MDD}$	0,86	4	704,5	710,8	0,10
Superfície aquosa	$\text{logit}(\theta_1) = -7,8931 - 10,071192*\text{NDVI} + 0,000074*\text{MDD}$	0,99	3	597,8	698,5	0,38
	$\text{logit}(\theta_2) = -7,4563 - 9,266354*\text{NDVI} + 0,000594*\text{MDD} - 0,000485*\text{MDE}$	0,99	4	702,8	924,5	0,32
	$\text{logit}(\theta_3) = -4,8262 + 0,008144*\text{MDD} - 0,007756*\text{MDE}$	0,89	3	692,5	1056,3	0,04
Solo exposto	$\text{logit}(\theta_1) = -7,7661 - 7,404507*\text{NDVI} + 0,000185*\text{MDD}$	0,91	3	918,5	782,309	0,11
	$\text{logit}(\theta_2) = -7,5437 - 7,751743*\text{NDVI} - 0,031359*\text{MDL} + 0,069208*\text{MDN}$	0,94	4	1048,3	2260,6	0,12
	$\text{logit}(\theta_3) = -7,9420 - 7,591665*\text{NDVI} + 0,291331*\text{MDN} + 0,000298*\text{MDD}$	0,94	4	774,309	2215,5	0,12

Fonte: o autor.

Apêndice (III) B – Matriz de contingência resultante da avaliação da classificação híbrida via IFVP obtida por MCE

	Classes de Referência										Total	Erro de Comissão
	Eucalipto jovem	Pastagem	Café	Área agrícola	Mata secundária	Área queimada	Superfície sombreada	Superfície aquosa	Solo exposto	Eucalipto maduro		
Eucalipto jovem	111	0	19	0	2	0	0	0	0	6	138	0,1957
Pastagem	0	170	12	16	0	1	0	0	0	0	199	0,1457
Café	16	0	56	0	4	0	0	0	0	1	77	0,2727
Área agrícola	0	15	0	98	0	0	0	0	2	0	115	0,1478
Mata secundária	2	0	2	0	113	0	1	0	0	10	128	0,1172
Área queimada	0	0	0	0	0	79	0	1	0	0	80	0,0125
Superfície sombreada	1	0	0	0	3	2	64	2	0	0	72	0,1111
Superfície aquosa	0	0	0	0	0	0	0	64	0	0	64	0
Solo exposto	0	1	0	0	0	0	0	0	50	0	51	0,0196
Eucalipto maduro	9	0	2	0	15	0	0	0	0	33	59	0,4407
Total	139	186	91	114	137	82	65	67	52	50	983	
Erro de Omissão	0,2014	0,086	0,3846	0,1404	0,1752	0,0366	0,0154	0,0448	0,0385	0,34		

Fonte: o autor.