

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE PESQUISAS HIDRÁULICAS

**REGIONALIZAÇÃO DE PARÂMETROS DE MODELO CHUVA-VAZÃO USANDO
REDES NEURAIS**

LAUDÍZIO DA SILVA DINIZ

Tese submetida ao Programa de Pós Graduação em Engenharia de Recursos Hídricos e Saneamento Ambiental da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Doutor em Engenharia.

Orientador: Prof. Dr. Robin Thomas Clarke

Co-Orientador: Prof. Dr. Alain Marie Passerat de Silans

Banca Examinadora

Prof. Dr. Carlos de Oliveira Galvão	UFCG – PB
Prof. Dr. João Batista Dias de Paiva	UFSM – RS
Prof. Dr. Olavo Correa Pedrollo	IPH/UFRGS
Prof. Dr. Walter Collischonn	IPH/UFRGS

Porto Alegre, agosto de 2008.

Este trabalho foi desenvolvido no Programa de Pós-Graduação em Engenharia de Recursos Hídricos e Saneamento Ambiental do Instituto de Pesquisas Hidráulicas da Universidade Federal do Rio Grande do Sul, sob a orientação do prof. Robin Thomas Clarke da Universidade Federal do Rio Grande do Sul e co-orientação do prof. Alain Marie Passerat de Silans da Universidade Federal da Paraíba.

Agradecimentos

- À Deus, que nunca me deixou só, pelas ricas bênçãos recebidas de suas divinas mãos.
- Aos meus pais que sempre acreditaram na educação e doaram, além de muito estímulo, quase tudo de seus escassos recursos para viabilizar nossos estudos, e aos meus cinco irmãos pelo apoio em forma de crédito que sempre depositaram em mim.
- A Roberta, Ananda, Matheus e Arthur, esposa e filhos, que souberam compreender o sentido de minha luta e sacrificaram muito de suas vidas pela adesão incondicional a este projeto.
- À Companhia de Água e Esgotos da Paraíba, na pessoa da Engenheira Aracilba Rocha, ex-Diretora Presidente, pela oportunidade para a realização do curso.
- Ao Conselho de Desenvolvimento Científico e Tecnológico, que ofereceu suporte financeiro durante nove meses do curso.
- Aos irmãos e amigos da Igreja Batista Regular no Bairro do Cristo Redentor, em João Pessoa-PB, que deram apoio irrestrito à minha família enquanto estive ausente.
- Ao meu orientador e co-orientador, os professores Robin Thomas Clarke e Alain Marie Passerrat de Silans, pelo esforço comum, paciência e dedicação.
- Ao amigo Carlos de Oliveira Galvão, pela cessão de uma de suas brilhantes idéias que culminou com a definição do tema deste trabalho.
- Ao professor Olavo Correa Pedrollo pelos comentários construtivos, os quais motivaram a introdução de outras reflexões até então não contempladas neste trabalho.
- Aos colegas Adriano Rolim, Joana D'Arc Medeiros, Luiz Gustavo, Márcio Nóbrega, Omar Barbosa e Walter Collischonn, com os quais tive o privilégio do mútuo aprendizado e da boa amizade.
- À Nadir Solari, ex-secretária do Programa de Pós-Graduação do IPH/UFRGS, pela paciente colaboração, cuja boa vontade e amizade nunca ofuscaram seu profissionalismo e eficiência.
- Aos amigos Marcos Célio e Adriano Lucena por me outorgarem afeto e encorajamento, neste e em outros trabalhos, igualmente difíceis, que me exigiram muito além do que eu sozinho poderia fazer.
- Aos colegas da Agência Executiva de Gestão das Águas do Estado da Paraíba – AESA, Ana Emília Duarte, Cecir Almeida, Ferdnando Cavalcanti, Guttemberg Silvino, Romildo Neto, Valéria Diniz e Washenayda Patrício, por me ajudarem em fases importantes deste trabalho, especialmente na etapa de obtenção das características físicas e climáticas das bacias.

À memória do meu pai Audízio Manoel Diniz
e do amigo Jorge Santiago Souto.

RESUMO

A aplicação dos modelos conceituais e semi-conceituais de transformação chuva-vazão depende de uma calibração prévia de seus parâmetros em locais com dados de vazões disponíveis e que guarde certa semelhança hidrológica com o local da aplicação. Para essa transferência de informações assumem-se algumas hipóteses que, em geral, não podem ser aferidas com dados reais, próprios de cada local, resultando em grandes incertezas sobre os resultados obtidos. Além do mais, a pouca quantidade de informações hidrológicas disponível impõe severas restrições à aplicabilidade dos modelos chuva-vazão, especialmente quando se necessita obter séries em escalas diárias ou mensais.

Este trabalho apresenta o desenvolvimento e validação de uma técnica de regionalização de parâmetros de modelos chuva-vazão baseado no uso de algoritmos genéticos e redes neurais, tendo o modelo SMAP – Soil Moisture Accounting Procedure (Lopes et al, 1981) mensal como objeto de avaliação. São empregados dados de 41 bacias hidrográficas localizadas no estado da Paraíba, no semi-árido do Nordeste brasileiro, entre as quais, 14 são áreas controladas por estações fluviométricas e 27 por reservatórios fluviais.

Para todas as bacias foram determinados 32 parâmetros físicos e climatológicos, sendo seis características das medidas lineares das bacias, três da forma, nove da rede de drenagem, sete do relevo, duas da capacidade de escoamento, três das classes de solos e duas da climatologia.

Uma análise dos dados disponíveis é feita em duas etapas: a primeira, consiste na avaliação da importância de cada bacia no conjunto dos dados, por meio da aplicação da técnica de aprendizado hierárquico não supervisionado; a segunda, consiste na seleção de atributos (características das bacias) usando técnicas de aprendizado supervisionado, a qual elimina dados redundantes e irrelevantes responsáveis pela diminuição da precisão e do entendimento sobre as hipóteses usadas quando as relações entre os parâmetros do modelo SMAP e as características são formuladas.

O modelo SMAP é calibrado de forma automática através do algoritmo SCE-UA (Shuffled Complex Evolution – University of Arizona), conforme Diniz (1994). Os parâmetros obtidos são relacionados às características físicas e climáticas das bacias hidrográficas através de redes neurais artificiais do tipo Perceptron Múltiplas Camadas – MLP.

Abstract

To use a conceptual or semi-conceptual hydrological rainfall–runoff model on a drainage basin without flow records, it must first have been calibrated at one or more sites where flow data are available and in a drainage basin having characteristics similar to those of the basin of interest. The transformation of rainfall into runoff needs some assumptions in both calibration and application sites that cannot be verified, leading to uncertainties. Moreover, the limited availability of hydrological data imposes some restriction on the application of rainfall-runoff models, particularly when the goal is to generate daily or monthly flow series.

This paper describes the development and validation of a parameter regionalization technique based on the use of genetic algorithms and artificial neural networks. The rainfall–runoff model used is the monthly version of the SMAP model. Data from 41 drainage basins in the Paraíba State of the semi-arid north-east of Brazil are used, 14 of which have a streamflow gauge station whilst 27 contain a dam at which reservoir levels are recorded. In each drainage basin, 32 physical and climatic attributes are calculated. Six of these describe linear dimensions of the basin, three are shape parameters, nine characterize the drainage network, seven describes the relief, two the run-off capacity, three the soil class and two the climatic conditions.

The analysis of the available data is made in two phases. The first consists of evaluating the importance of each basin in the whole ensemble, by means of a non-supervised hierarchical learning procedure, and the second consists of selecting attributes using a supervised learning procedure which eliminates redundancies and less important data that diminish the precision and understanding of the assumptions used when relationships between parameters and basin characteristics are formulated.

The SMAP model is calibrated in an automatic mode through SCE-UA algorithm, according to Diniz (1994). The acquired parameters are related do the physical and climate characteristics of the drainage basins through MLP – Multiple Layers Perceptrons artificial neural network.

Sumário

AGRADECIMENTOS -----	i
DEDICATÓRIA-----	ii
RESUMO-----	iii
ABSTRACT-----	iv
SUMÁRIO-----	v
LISTA DE TABELAS-----	viii
LISTA DE FIGURAS-----	x
LISTA DE SÍMBOLOS -----	xiii
CAPÍTULO I: INTRODUÇÃO -----	1
1.1 CARACTERIZAÇÃO DO PROBLEMA -----	1
1.2 JUSTIFICATIVA -----	2
1.3 OBJETIVOS -----	3
1.4 CONTRIBUIÇÕES -----	4
CAPITULO II: REGIONALIZAÇÃO HIDROLÓGICA-----	5
2.1 CONSIDERAÇÕES GERAIS-----	5
2.2 FATORES LIMITANTES DA REGIONALIZAÇÃO -----	6
2.3 ANÁLISE DOS DADOS USADOS NA REGIONALIZAÇÃO -----	7
2.4 MÉTODOS DE REGIONALIZAÇÃO -----	9
2.4.1 Equações empíricas -----	9
2.4.2 Regionalização por delimitação geográfica -----	9
2.4.3 Regionalização por mapeamento -----	9
2.4.4 Método da vazão índice -----	9
2.4.5 Métodos gráficos -----	11
2.4.6 Regionalização pelo método das células quadradas -----	12
2.4.7 Métodos estatísticos-----	12
2.4.8 Redes Neurais Artificiais-----	21
2.4.8.1 Base conceitual-----	22
2.5 VARIÁVEIS HIDROLÓGICAS REGIONALIZADAS -----	28
2.5.1 Considerações gerais -----	28
2.5.2 Vazão média -----	28
2.5.3 Vazão mínima-----	29
2.5.4 Vazão máxima -----	29

2.5.5	Regionalização da curva de regularização-----	30
2.5.6	Regionalização da curva de permanência -----	31
2.5.7	Regionalização dos parâmetros de uma distribuição de probabilidade ---	32
2.5.8	Regionalização do hidrograma unitário-----	33
2.5.9	Regionalização de parâmetros de modelos chuva-vazão -----	34
CAPÍTULO III: ÁREA DE ESTUDO -----		38
3.1	CARACTERIZAÇÃO FÍSICO-CLIMÁTICA DA ÁREA-----	38
3.1.1	Localização e extensão -----	38
3.1.2	Relevo e altimetria -----	40
3.1.3	Geologia-----	40
3.1.4	Solos-----	43
3.1.5	Vegetação e uso do solo -----	45
3.1.6	Clima -----	47
3.1.7	Hidrografia-----	47
3.1.8	Recursos hídricos -----	49
3.2	SELEÇÃO DAS BACIAS -----	49
CAPÍTULO IV: METODOLOGIA -----		52
4.1	ASPECTOS METODOLÓGICOS-----	52
4.2	SELEÇÃO DO MODELO CHUVA-VAZÃO-----	55
4.2.1	O modelo SMAP -----	55
4.3	DADOS PARA CALIBRAÇÃO DO MODELO SMAP -----	58
4.3.1	Fluviometria -----	58
4.3.2	Pluviometria -----	62
4.3.3	Evapotranspiração -----	63
4.4	CALIBRAÇÃO DO MODELO -----	68
4.4.1	Considerações gerais -----	68
4.4.2	O algoritmo de calibração SCE-UA -----	62
4.5	OBTENÇÃO DAS CARACTERÍSTICAS DAS BACIAS -----	75
4.5.1	Parâmetros morfológicos-----	75
4.5.2	Percentuais dos solos predominantes nas bacias -----	79
4.6	SELEÇÃO DE ATRIBUTOS -----	82
4.6.1	Considerações gerais-----	82
4.6.2	Técnicas de aprendizado de máquinas e avaliação dos dados-----	83
4.6.3	Avaliação de cada bacia no conjunto dos dados-----	89
4.6.4	Seleção de atributos-----	89

4.7 IDENTIFICAÇÃO DE RELAÇÕES REGIONAIS-----	94
4.7.1 Introdução -----	94
4.7.2 O método da bacia-meta -----	95
CAPÍTULO V: RESULTADOS E DISCUSSÕES-----	99
5.1 RESULTADOS DA CALIBRAÇÃO DO MODELO SMAP-----	99
5.2 CARACTERÍSTICAS DAS BACIAS-----	101
5.3 SELEÇÃO DE BACIAS E DE ATRIBUTOS-----	107
5.3.1 Avaliação de cada bacia no conjunto dos dados-----	107
5.3.2 Seleção de atributos-----	110
5.4 TREINAMENTO DA REDE NEURAL -----	119
5.4.1 Definição das arquiteturas da RNA -----	119
5.4.2 Treinamento da RNA sobre o parâmetro E2-----	120
5.4.3 Treinamento da RNA sobre o parâmetro STR-----	125
5.4.4 Treinamento da RNA sobre o parâmetro E1 -----	129
5.4.5 Treinamento da RNA sobre o parâmetro CINF -----	133
5.4.6 Identificação de Redes Neurais Regionais -----	136
5.4.7 Aplicação das Redes Neurais Regionais à bacias sem dados -----	143
CAPÍTULO VI: CONCLUSÕES E RECOMENDAÇÕES -----	145
6.1 CONCLUSÕES -----	145
6.1.1 Uso de bacias controladas por açudes -----	145
6.1.2 O modelo chuva-vazão -----	145
6.1.3 Seleção de atributos relevantes-----	146
6.1.4 Uso de Redes Neurais Artificiais -----	147
6.1.5 Conceito de bacia-meta -----	148
6.2 RECOMENDAÇÕES -----	149
6.2.1 Quanto ao uso de bacias controladas por açudes-----	149
6.2.2 Algoritmos para seleção de atributos -----	149
REFERÊNCIAS BIBLIOGRÁFICAS -----	150
ANEXO A: DETALHES DAS CARACTERÍSTICAS FÍSICAS	
DAS BACIAS HIDROGRÁFICAS E REFERÊNCIAS BIBLIOGRÁFICAS -----	164
ANEXO B: MATRIZ DE SIMILARIDADES-----	181
ANEXO C: CÓDIGO COMPUTACIONAL DO ALGORITMO SCE-UA	
E SUB-ROTINAS -----	190

Lista de Tabelas

Tabela 3.1: Distribuição dos usos do solo e cobertura vegetal -----	45
Tabela 3.2: Caracterização das bacias hidrográficas do Estado da Paraíba -----	49
Tabela 3.3: Distribuição das bacias de acordo com as áreas de drenagem -----	50
Tabela 4.1: Postos fluviométricos selecionados para os estudos -----	59
Tabela 4.2: Localização e período com dados disponíveis dos Postos fluviométricos -----	60
Tabela 4.3: Açudes selecionados para geração de vazões -----	61
Tabela 4.4: Postos pluviométricos selecionados para as bacias hidrográficas dos postos fluviométricos -----	63
Tabela 4.5: Postos pluviométricos selecionados para as bacias hidrográficas dos açudes -	65
Tabela 4.6: Localidades com evapotranspiração representativa -----	66
Tabela 4.7: Valores médios anuais de precipitação e evapotranspiração -----	67
Tabela 4.8: Características das medidas lineares da bacia -----	76
Tabela 4.9: Características da forma da bacia -----	76
Tabela 4.10: Características da rede de drenagem -----	77
Tabela 4.11: Características do relevo da bacia -----	78
Tabela 4.12: Características da capacidade de escoamento da bacia hidrográfica -----	78
Tabela 5.1: Parâmetros calibrados do modelo SMAP -----	100
Tabela 5.2: Parâmetros morfológicos das bacias hidrográficas -----	102
Tabela 5.3: Parâmetros morfológicos das bacias hidrográficas -----	103
Tabela 5.4: Parâmetros morfológicos das bacias hidrográficas -----	104
Tabela 5.5: Tipos de solos encontrados nas bacias hidrográficas -----	105
Tabela 5.6: Percentuais de ocorrência das classes de solos nas bacias -----	106
Tabela 5.7: Atributos selecionados através do algoritmo CFS -----	112
Tabela 5.8: Correlações entre o parâmetro do modelo SMAP e os atributos selecionados -	113
Tabela 5.9: Arquitetura da rede neural para cada parâmetro do SMAP -----	119
Tabela 5.10: Parâmetro E2 do modelo SMAP: fase de treinamento da RNA -----	122
Tabela 5.11: Parâmetro E2 do modelo SMAP: fase de validação -----	124
Tabela 5.12: Matriz de correlações entre os parâmetros do modelo SMAP -----	125
Tabela 5.13: Parâmetro STR do modelo SMAP: fase de treinamento da RNA -----	126
Tabela 5.14: Parâmetro STR do modelo SMAP: fase de validação -----	128
Tabela 5.15: Parâmetro E1 do modelo SMAP: fase de treinamento da RNA -----	130
Tabela 5.16: Parâmetro E1 do modelo SMAP: fase de validação -----	132
Tabela 5.17: Parâmetro CINF do modelo SMAP: fase de treinamento da RNA -----	134

Tabela 5.18: Parâmetro CINF do modelo SMAP: fase de validação-----	135
Tabela 5.19: Identificação das bacias-meta mais similares usadas para estimar os parâmetros do SMAP na bacia omitida -----	138
Tabela 5.20: Classificação das bacias segundo os resultados por parâmetro -----	139
Tabela 5.21: Contagem dos resultados segundo a classificação obtida no treinamento e na verificação da RNA-----	140
Tabela 5.22: Parâmetros obtidos na calibração e na regionalização-----	141

Lista de Figuras

Figura 1.1: Etapas da modelagem paramétrica-----	2
Figura 2.1: Estrutura do neurônio biológico-----	23
Figura 2.2: Modelo de um neurônio artificial -----	24
Figura 2.3: Arquitetura de uma RNA tipo MLP -----	25
Figura 2.4: Critério de parada do treinamento por validação cruzada-----	28
Figura 3.1: Situação e localização da área de estudo -----	39
Figura 3.2: Relevo e altimetria do estado da Paraíba-----	41
Figura 3.3: Mapa geológico do estado da Paraíba -----	42
Figura 3.4: Mapa de reconhecimento de solos -----	44
Figura 3.5: Cobertura vegetal -----	46
Figura 3.6: Classificação climática de Köppen -----	48
Figura 3.7: Localização das bacias selecionadas-----	51
Figura 4.1: Etapas do processo de regionalização de parâmetros do modelo chuva-vazão-----	54
Figura 4.2: Desenho esquemático do modelo SMAP -----	57
Figura 4.3: Precipitação média anual no estado da Paraíba-----	64
Figura 4.4: Fluxograma do algoritmo SCE-UA-----	73
Figura 4.5: Algoritmo de evolução competitiva -----	74
Figura 4.6: Distribuição espacial dos solos nas bacias do Sertão -----	80
Figura 4.7: Distribuição espacial dos solos nas bacias do Cariri e Curimataú-----	81
Figura 4.8: Fases do processo de KDD (Liu e Motoda, 2002) -----	82
Figura 4.9: Hierarquia do aprendizado indutivo -----	84
Figura 4.10: Topologia de uma rede de Kohonen -----	85
Figura 4.11: Single Link: menor distância entre dois clusters-----	87
Figura 4.12: Complete Link: maior distância entre dois clusters -----	88
Figura 4.13: Average Link: média das distâncias entre os dois clusters -----	88
Figura 4.14: Componentes do algoritmo CFS. Os dados de treinamento e teste são reduzidos para conter apenas os atributos selecionados pelo CFS. Com a redução da dimensionalidade os dados podem ser usados para as fases de indução e previsão. (adaptado de Hall, 1999)-----	94
Figura 4.15: Método da bacia-meta para construção de modelo hidrológico regional -----	98

Figura 5.1: Dendograma obtido por meio do algoritmo Single Link utilizando a medida euclidiana sobre os dados das 41 bacias -----	107
Figura 5.2: Dendograma obtido por meio do algoritmo Complete Link utilizando a medida euclidiana sobre os dados das 41 bacias -----	108
Figura 5.3: Dendograma obtido por meio do algoritmo Average Link utilizando a medida euclidiana sobre os dados das 41 bacias -----	109
Figura 5.4: Dendograma obtido por meio do algoritmo Ward utilizando a medida euclidiana sobre os dados das 41 bacias -----	110
Figura 5.5: Distribuição espacial dos grupos homogêneos de bacias hidrográficas -----	111
Figura 5.6: Dendograma obtido por meio do algoritmo Average Link utilizando a medida euclidiana sobre os dados das 36 bacias restantes -----	112
Figura 5.7: Relação entre o parâmetro STR do modelo SMAP e os atributos selecionados -----	114
Figura 5.8: Relação entre o parâmetro CINF do modelo SMAP e os atributos selecionados -----	115
Figura 5.9: Relação entre o parâmetro E1 do modelo SMAP e os atributos selecionados -----	116
Figura 5.10a: Relação entre o parâmetro E2 do modelo SMAP e os atributos selecionados -----	117
Figura 5.10b: Relação entre o parâmetro E2 do modelo SMAP e os atributos selecionados -----	118
Figura 5.11: Arquitetura da rede neural – parâmetro E2-----	121
Figura 5.12: Dendograma obtido para o parâmetro E2 por meio do algoritmo Average Link utilizando a medida euclidiana -----	123
Figura 5.13: Arquitetura da rede neural – parâmetro STR-----	125
Figura 5.14: Dendograma obtido para o parâmetro STR por meio do algoritmo Average Link utilizando a medida euclidiana-----	127
Figura 5.15: Arquitetura da rede neural – parâmetro E1-----	129
Figura 5.16: Dendograma obtido para o parâmetro E1 por meio do algoritmo Average Link utilizando a medida euclidiana -----	131
Figura 5.17: Arquitetura da rede neural – parâmetro CINF -----	133
Figura 5.18: Dendograma obtido para o parâmetro CINF por meio do algoritmo Average Link utilizando a medida euclidiana -----	136
Figura 5.19: Vazões observadas e simuladas - BF1 -----	140
Figura 5.20: Vazões observadas e simuladas - BF2 -----	140

Figura 5.21: Vazões observadas e simuladas – BF4-----	142
Figura 5.22: Vazões observadas e simuladas – BF6-----	142
Figura 5.23: Vazões observadas e simuladas – BF12 -----	142
Figura 5.24: Vazões observadas e simuladas – BA2 -----	143
Figura 5.25: Vazões observadas e simuladas – BA3 -----	143

Lista de Símbolos

- Q (m^3/s) vazão na secção de referência de uma bacia hidrográfica
- A (k^m^2) área de drenagem de uma bacia hidrográfica
- P (mm) precipitação medida em uma estação meteorológica
- P_m (mm) precipitação média sobre uma bacia hidrográfica
- c e n (adm) constantes da equação de regressão determinadas para a região
- T (anos) período de retorno
- Z matriz dos componentes principais
- X é a matriz dos dados
- A matriz dos coeficientes que definem uma transformação linear
- S (m^3) capacidade de armazenamento de um reservatório fluvial
- Dem ($m^3/mês$) demandas mensais para o abastecimento humano.
- $Evap$ ($m^3/mês$) volume mensal evaporado no espelho da água.
- Irr ($m^3/mês$) demandas totais para irrigação.
- V_{afiu} ($m^3/mês$) volume afluente ao reservatório fluvial
- Q_{reg} (m^3/s) vazão regularizável de um reservatório fluvial
- p garantia associada a vazão regularizável
- HU hidrograma unitário
- $E2$ (adm) parâmetro do modelo SMAP: controlador do escoamento superficial
- STR (mm) parâmetro do modelo SMAP: nível máximo do reservatório superficial, representa a capacidade de saturação do solo
- $E1$ (adm) parâmetro do modelo SMAP: responsável pelo controle das taxas de evapotranspiração na bacia
- $CINF$ (adm) parâmetro do modelo SMAP: representa a parcela da chuva que não alimenta o reservatório superficial
- BAn nomenclatura para bacia hidrográfica controlada por um açude
- BFn nomenclatura para bacia hidrográfica controlada por uma estação fluviométrica
- SU – Symmetrical Uncertainty: medida de correlação entre dois atributos ou entre um atributo e uma classe
- $GI(X|Y)$ Ganho de Informação que reflete a informação adicional acerca de X dado Y
- $H(X)$ entropia do atributo X
- $H(X,Y)$ entropia (ganho de informação) de X após observar a variável Y
- CFS – Correlation-based Feature Selection: algoritmo para seleção de atributos

- M_S medida de mérito da heurística de um subconjunto de atributos S contendo k atributos
- $\overline{r_{cf}}$ correlação média atributo-classe
- e $\overline{r_{ff}}$ inter-correlação média atributo-atributo
- WEKA – Waikato Environment for Knowledge Analysis: software que implementa algoritmos de aprendizagem de máquinas e mineração de dados
- Qnet Software que implementa uma rede neural MLP
- η taxa de aprendizagem de uma rede neural artificial
- s função de combinação de um neurônio de uma rede neural artificial
- $f(s)$ função de transferência entre neurônios de uma rede neural artificial
- ERR erro relativo entre duas variáveis (observada e calculada)
- PAE (adm) percentagem da área da bacia hidrográfica coberta por espelho de água
- Cmed (m/m) elevação média da bacia
- C_t (número de cursos de água/km²) coeficiente de torrencialidade
- DS (m/m) desnível específico
- I_p (adm) índice de declividade média da bacia
- K_f (adm) fator de forma da bacia
- L_d (km) comprimento da rede de drenagem
- L_t (km) comprimento do curso de água principal
- P_r (km) perímetro da bacia hidrográfica
- R_a (adm) índice das áreas
- R_b (adm) índice de bifurcação
- R_l (adm) índice dos comprimentos
- SIN sinuosidade do curso de água principal
- Solo 1 percentagem do solo pertencente à Classe I
- Solo 2 percentagem do solo pertencente à Classe II

CAPÍTULO I: INTRODUÇÃO

1.1 CARACTERIZAÇÃO DO PROBLEMA

Um dos problemas mais cruciais de todos os problemas sociais brasileiros, o drama das secas no Nordeste, especialmente na zona hidrológica semi-árida, tem sido objeto, desde o início do século, de importantes estudos e pesquisas sobre o comportamento da água e do solo nas bacias hidrográficas. Entretanto, um grande entrave no setor de recursos hídricos deve-se a descontinuidade no decurso dos projetos de pesquisas nas áreas mais importantes como é o caso da regionalização hidrológica, as quais concorreriam para a ampliação dos conhecimentos e metodologias sobre como desenvolver a região com base em suas limitações ambientais.

A escassez ou quase inexistência de dados fluviométricos com séries longas e confiáveis sobre os cursos d'água do Nordeste do Brasil, obriga os especialistas em recursos hídricos a adotarem procedimentos de cálculos de séries sintéticas de vazões a partir de dados de precipitação e evaporação disponíveis nesta região ou outras técnicas ainda mais simplificadas a exemplo da equação de Aguiar (Aguiar, 1940). Entretanto, a utilização de tais metodologias parte do pressuposto de que o modelo utilizado esteja devidamente calibrado e que os parâmetros determinados podem ser usados em outras bacias. Isto pode ser admitido, desde que, as bacias tenham alguma similaridade hidrológica, ainda assim com grandes restrições.

Devido a intermitência dos rios nordestinos, o principal meio de disponibilizar a água para os diversos usos é através da construção de açudes que barram os cursos d'água provocando o respectivo armazenamento. Como decorrência, vários postos fluviométricos instalados na rede de drenagem são inutilizados para fins de calibração de modelos hidrológicos, a menos que incorporem em sua estrutura o balanço hídrico e a operação dos açudes (Paiva et al, 1999).

Uma vez calibrado o modelo de geração de vazões em uma bacia hidrográfica em que se conheça a fluviometria, uma questão fundamental que se apresenta, como já mencionado, é sobre a aplicabilidade dos parâmetros deste modelo para outras regiões. Um procedimento bastante utilizado atualmente é o da similaridade hidrológica entre bacias hidrográficas, estabelecida a partir de critérios de classificação de zonas hidrológicamente homogêneas, para adoção dos parâmetros obtidos na calibração. Mesmo assim, a extrapolação do modelo para outras regiões ainda é feita sob severas limitações. A Figura 1.1 apresenta as etapas que constituem a modelagem paramétrica.

A regionalização de parâmetros de modelos hidrológicos pode ser realizada buscando-se relações úteis dos mesmos com as características físicas e climáticas das bacias hidrográficas, mas, de acordo com Tucci (1993), com algumas limitações devido a interdependência dos parâmetros, podendo criar tendenciosidade nas estimativas.

Aplicações recentes de Algoritmos Genéticos, que usam conceitos da teoria da evolução natural, na calibração de modelos chuva-vazão (Galvão & Valença, 1999; Duan et al, 1992; Duan et al, 1993; Sorooshian et al, 1998) indicam que os mesmos são bastante eficientes na localização dos valores ótimos globais dos parâmetros dos modelos e detêm capacidade para superar as dificuldades usualmente encontradas durante a calibração. Isso poderá se constituir numa importante ferramenta para minimizar as restrições próprias do processo de regionalização de parâmetros de modelos hidrológicos.

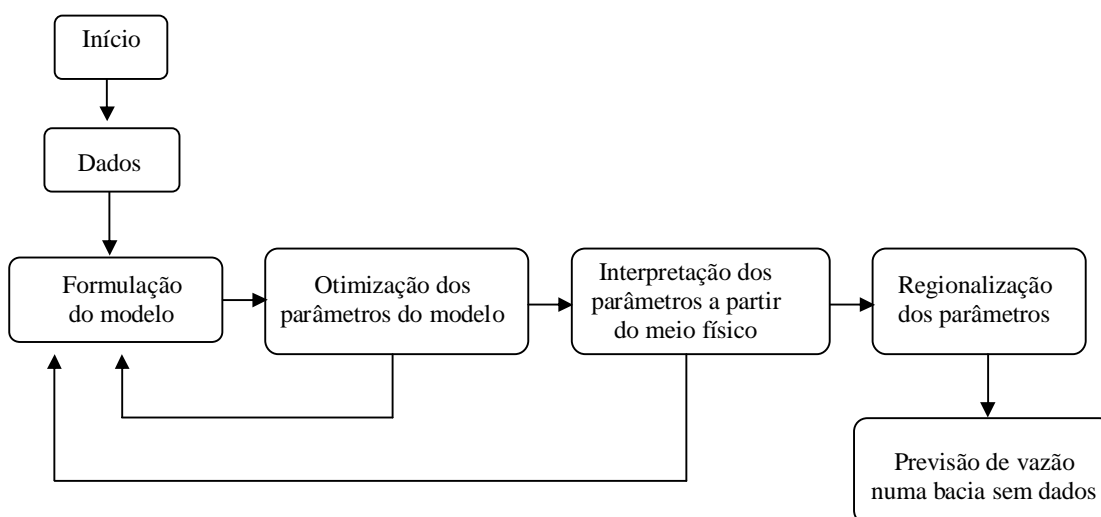


Figura 1.1 – Etapas da modelagem paramétrica

1.2 JUSTIFICATIVA

Cresce no Brasil a necessidade de se desenvolver métodos eficientes de regionalização hidrológica como consequência da implantação dos sistemas estaduais e do federal de gerenciamento de recursos hídricos e da escassez de dados obtidos através de medições em campo. Destacam-se os subsistemas de outorga e cobrança pelo uso da água bruta, o licenciamento ambiental e a elaboração dos planos diretores de bacias, os quais, igualmente, necessitam estimar curvas de permanência de vazões, na maioria das vezes não disponíveis.

A regionalização hidrológica pode ser realizada em duas categorias distintas, mas que, igualmente, utilizam as características físico-climáticas das bacias hidrográficas como variáveis explicativas: a) regionalização de características das séries de vazões (mínimas, médias, máximas, vazões da curva de permanência associada a uma dada garantia, vazões regularizáveis de reservatórios fluviais, quantis de vazões); b) regionalização de parâmetros de funções matemáticas e estatísticas e de parâmetros de modelos chuva-vazão, sendo esta última a utilizada neste trabalho.

A simplicidade é a principal qualidade da primeira abordagem. Sua maior deficiência está no fato de não usar séries de chuvas, disponíveis em maior quantidade e melhor qualidade em relação às séries de vazões.

A segunda categoria, além de permitir obter os resultados da primeira, pode fornecer resultados mais abrangentes com os benefícios decorrentes da geração de séries de vazões.

Várias pesquisas sobre regionalização já foram realizadas no Nordeste do Brasil, desde a antiga equação de Aguiar (Aguiar, 1940) para grandes e médias bacias, até o método de Cadier (Cadier, 1984; Cadier, 1996) para pequenas bacias. Entretanto, pouco se avançou na regionalização de vazão em intervalos diários e mensais, a partir do uso de modelos conceituais chuva-vazão, viabilizado com a associação dos respectivos parâmetros às características das bacias hidrográficas.

Embora as correlações entre parâmetros do modelo e características físicas da bacia hidrográfica tenham sido assumidas em muitos modelos, na realidade são pouco testadas e validadas para uma quantidade satisfatória de dados. Atualmente, os modelos chuva vazão são utilizados em bacias julgadas semelhantes àquela usada na calibração, só que, quase sempre, isso é feito subjetivamente.

A disponibilidade mais recente das técnicas de otimização baseadas em algoritmos genéticos e redes neurais aumenta as possibilidades de se desenvolver modelos hidrológicos com estruturas mais adequadas e a respectiva regionalização dos parâmetros.

1.3 OBJETIVOS

Os objetivos principais deste trabalho são:

a) Examinar a viabilidade de se usar características físicas (solo e morfologia) e climáticas (chuva, evaporação) para estimar os parâmetros de um modelo chuva-vazão através de relações regionais.

b) Estabelecer uma metodologia para a seleção de características físico-climáticas relevantes e não redundantes das bacias hidrográficas, visando diminuir a dimensionalidade dos dados e aumentar a eficiência das relações destes com os parâmetros do modelo SMAP.

c) Usar e avaliar as técnicas de redes neurais como opção para estimativa dos parâmetros de modelos conceituais chuva-vazão.

d) Propor critérios para uso e aplicação das técnicas definidas no item c.

1.4 CONTRIBUIÇÕES

As principais contribuições deste trabalho são:

- a) Introduzir a possibilidade do uso de dados de níveis de açudes para reconstituição de séries de vazões naturais, aumentando substancialmente a quantidade de estações fluviométricas no estado da Paraíba.
- b) Aplicar técnicas de mineração de dados na seleção de atributos irrelevantes e redundantes, permitindo reduzir a dimensionalidade dos dados e selecionar somente as características das bacias que melhor descrevem e discriminam o conjunto de dados, com conseqüente aumento da eficiência das redes neurais.
- c) Inserir o conceito de bacia-meta para permitir o eficiente uso de pequenas séries de dados no processo de regionalização hidrológica, particularmente na fase de verificação da rede neural.
- d) Aplicar técnicas de redes neurais na regionalização de parâmetros de modelos concentrados de transformação chuva-vazão com boas perspectivas de aplicação em qualquer categoria de modelos conceituais e semi-conceituais.

CAPÍTULO II: REGIONALIZAÇÃO HIDROLÓGICA

2.1 CONSIDERAÇÕES GERAIS

A idéia de regionalização em hidrologia além de ter uma antiga e longa história é de grande importância para o desenvolvimento e aplicação da engenharia de recursos hídricos, isto pelo fato das redes de observações hidroclimatológicas do Brasil, em particular, não poderem abstrair todas as informações contidas nos processos das bacias hidrográficas. Entre outras aplicações as técnicas de regionalização têm sido empregadas em vazões máximas e mínimas dos cursos de água, curvas de regularização de vazão e de permanência e precipitações máximas (Tucci, 2002).

A regionalização hidrológica é uma ferramenta muito popular usada para permitir aos hidrólogos utilizar dados medidos em um determinado lugar para aplicações em outros locais sem dados ou com dados de má qualidade. Considerada de uso complexo, nunca está totalmente concluída ou suficientemente confiável, não significando, entretanto, que não possa ser usada. Se corretamente empregada, torna-se muito importante nos estudos hidrológicos de grande escala. Usa princípios e lógicas desenvolvidas em uma grande variedade de disciplinas que, de alguma forma estão relacionados à hidrologia e a regionalização propriamente dita.

O processo de regionalização está presente no dia a dia das pessoas, as quais tendem a desenvolver mentalmente condutas de extrapolação de informações usando conceitos e experiências anteriores, obtidos durante suas próprias vidas, e as aplicam para decidirem sobre situações ainda não vividas. Suas conclusões serão tanto melhores quanto mais experiência e conhecimento elas tenham adquirido. O hidrólogo desenvolve ao longo de sua carreira profissional uma sensibilidade relativa sobre os fenômenos e processos hidrológicos, de tal forma que o permite fazer avaliações prévias dos problemas de recursos hídricos em geral, pela simples noção de ordem de grandeza de algumas variáveis hidrológicas.

Segundo Franchini & Suppo (1996) a regionalização deve ser feita com base nas características físicas e climáticas e, portanto, independentemente da continuidade geográfica. Duas regiões distintas podem ser consideradas homogêneas quanto a uma variável mas não o serem em relação a outras, o que, associado à escassez de dados observados, inviabiliza, muitas vezes, a análise direta sobre a variável em estudo, implicando na necessidade de se encontrar relações físicas entre a variável a regionalizar e outras mais facilmente determinadas. Definida a variável a regionalizar, uma das etapas mais importantes no processo é a seleção das variáveis explicativas (Andrade & Hawkins, 2000). A exclusão de

variáveis importantes ou a inclusão de variáveis insensíveis podem distorcer os resultados e inviabilizar a regionalização.

É evidente que os resultados decorrentes de uma regionalização nunca irão substituir a observação feita em campo, podendo-se distanciar dos valores reais, mas as decisões feitas com base nestes resultados serão sempre superiores às aquelas normalmente praticadas no meio técnico, que se baseiam em perigosas extrapolações apoiadas apenas em sentimentos individuais (Strobel, 1987).

2.2 FATORES LIMITANTES DA REGIONALIZAÇÃO

A confiabilidade de uma relação regional, ou mesmo sua validação em termos de uma região em particular, depende da qualidade dos dados, da homogeneidade da região, e da adequação das informações acerca das bacias hidrográficas. Algumas condições que limitam a validade de uma relação regional são as seguintes (Riggs, 1990):

- a) Série com pequena quantidade de dados ou pouco representativa. Uma série é representativa quando suas características não diferem, estatisticamente, de outras séries regionais de maior extensão.
- b) Número de estações pluvi-fluviométrica insuficiente para cobrir toda a extensão da bacia hidrográfica com sua heterogeneidade.
- c) Alterações nas estações de medições ou interrupções nas medições.
- d) Precipitação na bacia não bem definida, especialmente em regiões sensíveis aos efeitos orográficos.
- e) Rios que atravessam regiões com características físico-climáticas diferentes.
- f) Irregularidade na distribuição das chuvas, cujos efeitos são geralmente ponderados ao longo da bacia. Em regiões semi-áridas, por exemplo, onde esse efeito é mais pronunciado, uma parte da bacia que não contribui com a formação de uma dada vazão, dificilmente pode ser quantificada com segurança.
- g) Diferenças na geologia podem ocasionar grandes diferenças nas taxas de infiltração e no movimento da água subterrânea dentro de uma bacia ou entre bacias.
- h) A avaliação da similaridade entre bacias é fortemente dependente da escala, da unidade e da magnitude de cada variável. Uma alternativa para resolver este problema é a normalização dos dados.
- i) Dificuldade na seleção de variáveis explicativas.

Os dados geralmente requeridos para os estudos de regionalização de parâmetros de modelos chuva-vazão podem ser agrupados em quatro categorias:

- i. Observações básicas: consistindo de valores médios em um número estatisticamente representativo de bacias hidrográficas:
 - precipitação média diária ou mensal;
 - vazões médias diárias ou mensais;
 - evaporação potencial média de longo termo diária ou mensal.
- ii. Características físicas das bacias: incluindo, entre outros, os seguintes valores médios:
 - área de drenagem;
 - declividade média da bacia;
 - precipitação média anual;
 - evaporação potencial média anual;
 - coeficiente de escoamento médio anual;
 - distribuição de classes de texturas do solo;
 - distribuição dos tipos de vegetação;
 - porosidade média dos solos;
 - condutividade hidráulica saturada média.
- iii. Observações adicionais desejáveis: evaporação diária, dados meteorológicos.
- iv. Observações especiais para testes e avaliações detalhadas: características físicas obtidas de imagens de satélites ou de mapas em escalas detalhadas.

2.3 ANÁLISE DOS DADOS USADOS NA REGIONALIZAÇÃO

A qualidade dos dados hidrológicos é essencial para os estudos de regionalização, e uma análise criteriosa sobre suas condições deve sempre preceder a realização da regionalização (Tucci, 2002).

Para dados hidrológicos a serem usados em estudos de regionalização, dois tipos de erros são particularmente relevantes e presumíveis (Hosking & Wallis, 1997): primeiro, os valores dos dados são coletados ou transcritos incorretamente; segundo, as circunstâncias sobre as quais os dados foram coletados podem ter sido mudadas com o tempo.

Nestes casos, três tipos de diagnósticos podem revelar a presença de dados grosseiros na série considerada: a) análise individual dos valores da série de dados para verificar a existência de erros grosseiros; b) exame minucioso de cada estação de medição para identificação de possíveis valores atípicos ou repetidos, com provável origem na leitura ou na transcrição dos dados; c) comparação entre dados de diferentes estações de medições, para

tentar descobrir algum tipo de irregularidade nos dados. Se uma dada estação não harmonizar-se com outras aparentemente similares, pode-se desconfiar dos seus dados.

Numa análise inicial, o conjunto de dados disponível dentro da região é listado. As estações não adequadas são eliminadas, e um período comum das séries é selecionado. Geralmente, estações com menos de cinco anos de dados ou aquelas cujos rios são regulados ou controlados também são excluídas, a não ser que sejam considerados seus efeitos dentro do estudo.

Se as séries não são totalmente aleatórias ou apresentam tendências ou periodicidades o intervalo de tempo sobre o qual as medidas foram feitas se torna importante para uma comparação com outras séries disponíveis na região. Um gráfico de barras mostrando o período de dados de cada estação é útil na determinação do período base a ser usado. O período base deve incluir o máximo de informações possível, ou seja, o máximo número de estações e de anos.

Dados faltosos podem ser preenchidos por correlação inter-estações. Os dados obtidos por preenchimento não devem ser usados diretamente, mas apenas como apoio na determinação dos períodos de retorno representativos para os eventos registrados.

Testes estatísticos para examinar a existência de dados atípicos e tendências nas séries de dados são bem fundamentados na literatura (Barnett & Lewes, 1994). Para comparação de dados de diferentes estações, são bastante conhecidas as técnicas gráficas de dupla-massa e de quantis-quantis.

De acordo com Hosking & Wallis (1997), a técnica de momentos – L também pode ser usada para identificar dados incorretos, dados atípicos, tendências e oscilações na média da série de dados. A comparação das relações de momentos – L de diferentes estações de medições permite deduzir quais estações são grosseiramente discordantes do grupo como um todo.

No caso particular das variáveis explicativas, usadas para estimar os parâmetros ou as variáveis regionalizadas, estas devem ser de fácil obtenção e, quando obtidas de mapas, que sejam em escalas apropriadas (1:50000, 1:100000 ou 1:250000), conforme Tucci (2002).

Devido a forte correlação existente entre as variáveis explicativas, há a necessidade premente de se investigar e retirar da amostra as irrelevantes e redundantes, sob pena de insucesso da regionalização. O método de regressão stepwise (Draper & Smith, 1981) e as técnicas de aprendizado supervisionado (Hall, 2000) podem ser usadas para selecionar as variáveis mais importantes aos estudos de regionalização hidrológica.

2.4 MÉTODOS DE REGIONALIZAÇÃO

2.4.1 Equações empíricas

Um dos mais simples métodos de regionalização é o uso de equações empíricas relacionando a vazão de uma bacia hidrográfica com sua área de drenagem, dentro de uma região em particular, conforme a equação a seguir.

$$Q = c A^n \quad 2.1$$

onde: Q é a vazão em m³/s, A é a área em km², c e n são constantes determinadas para a região.

2.4.2 Regionalização por delimitação geográfica

A sua realização é feita por delimitação de áreas com características geográficas específicas, como o relevo, a topografia, o clima ou características das chuvas regionais (Matalas et al, 1975).

Em geral são delimitadas regiões formadas por áreas contíguas de estações fluviométricas vizinhas (Beable & McKerchar, 1982). São métodos arbitrários e subjetivos cujas regiões raramente dão a impressão de homogeneidade física (Hosking & Wallis, 1997).

2.4.3 Regionalização por mapeamento

Feito por meio de traçado de isolinhas expressando o comportamento da variável a ser regionalizada. Um exemplo típico deste método é a precipitação média anual de uma região expressa em termos de isoietas, as quais permitem conhecer a precipitação em um ponto qualquer da região onde não existem medições pluviométricas, por simples interpolação gráfica.

2.4.4 Método da vazão índice

A idéia básica do método da vazão-índice é aumentar a confiança das características mais freqüentes dentro de uma região. Se, dentro de uma área hidrologicamente homogênea, um número de estações hidrométricas tem sido operado e sido registrados os efeitos dos mesmos fatores meteorológicos, então uma combinação desses registros fornecerá, não um

registro longo, mas um mais representativo. Em linhas gerais, segundo Kite (1978), o método pode ser descrito como se segue.

Após a seleção e análise de consistência dos dados, o próximo passo é o cálculo dos períodos de retorno, T , para cada evento registrado em cada estação da região, usando-se a seguinte equação:

$$T = \frac{n+1}{m} \quad 2.2$$

onde: n é o tamanho da amostra e m é o número de ordem de um evento; $m = 1$ para o evento máximo e $m = n$ para o evento mínimo.

Para cada estação um gráfico de T versus o valor do evento é traçado. Nenhuma tentativa é feita para forçar um ajuste de qualquer distribuição matemática. O valor médio anual para a estação, tomado da distribuição extrema tipo I, é então selecionado da curva traçada no ponto $T = 2,33$. Há resultados experimentais (Benson, 1962) mostrando que o evento médio anual ocorre com um período de retorno de 2,33 anos. É preferido, no método da vazão-índice, obter-se o evento médio anual graficamente em vez de aritmeticamente.

A homogeneidade hidrológica regional é testada através do método de Dalrymple (1960) que consiste em plotar os limites de confiança definidos pela equação $2.25 \pm 6.33/\sqrt{n}$ para diferentes valores de n , válida para $T = 10$ anos, resultando no chamado gráfico teste.

Para cada estação da região estudada, calcula-se a relação entre os eventos de 10 anos de período de retorno e a média anual. Em seguida calcula-se a média entre os valores obtidos, representando a média regional. A relação média regional é multiplicada pelo evento médio anual de cada estação resultando num evento de 10 anos modificado para cada estação.

Os períodos de retorno correspondentes a esses eventos modificados são então encontrados para cada estação das suas respectivas curvas de frequências individuais, ou seja, T_E .

O período efetivo de registro de cada estação é determinado como o número de eventos anuais registrados mais a metade do número de eventos estimados para a estação através de correlação com outras estações, representado por N_E .

Depois, os pares de coordenadas (T_E, N_E) de cada estação são plotados no gráfico teste. Qualquer estação para a qual o ponto plotado estiver fora dos limites de confiança será excluída da região homogênea.

Para cada estação incluída na região homogênea, são calculadas relações entre eventos de diferentes períodos de retorno e a média anual, permitindo determinar os valores médios

dessas relações para a região. Um gráfico obtido pelo traçado entre essas relações médias e o período de retorno representa a curva de frequência regional e é, provavelmente, a relação mais representativa de toda região.

O próximo passo do método da vazão índice é o traçado de um gráfico entre a área de drenagem versus o evento médio anual para aquelas estações dentro da região homogênea. Uma alternativa a esse estágio é desenvolver uma equação por regressão múltipla entre o evento médio anual e características da bacia hidrográfica.

Para definir uma curva de frequência em qualquer local da região homogênea o evento médio anual é determinado a partir da equação obtida anteriormente e multiplicado pela relação média para os vários períodos de retorno, como determinado na curva regional.

2.4.5 Métodos gráficos

Representação gráfica das variáveis que definem a similaridade das bacias hidrográficas, a qual é verificada através do padrão de agrupamento demonstrado nos gráficos. Se forem usados mais de três parâmetros para definir a homogeneidade entre bacias, é quase impossível a expressão gráfica direta dos mesmos. Nesses casos, pode ser aplicada a técnica estatística multivariada denominada análise de componentes principais que consiste em encontrar outras variáveis (componentes principais) a partir de combinações lineares das variáveis originais. Essa ferramenta é mais útil quando os dois primeiros componentes explicam cerca de 90% da variação total, transformando o gráfico multidimensional em uma representação simples de duas variáveis.

Outro método simples de representação gráfica é aquele desenvolvido por Andrews (1972) e apresentado por Andrade e Hawkins (2000), no qual um ponto de várias dimensões pode ser representado por uma curva bidimensional da seguinte forma:

$$F(x) = 0,7071 u_1 + \text{sen}(x) u_2 + \text{cos}(x) u_3 + \text{sen}(2x) u_4 + \text{cos}(2x) u_5 \quad 2.3$$

onde:

u_1 = características físico-climáticas da bacia hidrográfica, sendo u_1 mais importante que u_2 e esta mais que u_3 e assim sucessivamente;

x = intervalo de variação da equação 2.3 ($-\pi < x < \pi$).

Esse método se aplica apenas para identificar bacias atípicas presentes em grupos formados a partir de outras técnicas, como a análise de grupos, por exemplo. A função de Andrews aplicada a cada bacia de um dado grupo é plotada em um único sistema de eixos coordenados, resultando numa nuvem de gráficos que possibilita, através da inspeção visual, verificar se existe ou não a presença de bacias diferentes do padrão do grupo.

2.4.6 Regionalização pelo método das células quadradas

A área em estudo é dividida em uma malha de quadrados uniformes identificados por coordenadas cartesianas. A cada quadrado é associado um conjunto de parâmetros que expressam as características físicas, informações climatológicas, meteorológicas e hidrológicas, de tal forma que os dados da bacia hidrográfica possam ser acumulados e avaliados automaticamente. A precisão dos resultados depende, também, das dimensões dos quadrados, os quais, por sua vez, dependem da magnitude da bacia, de seu relevo, entre outros.

2.4.7 Métodos estatísticos

No processo de regionalização algumas técnicas estatísticas têm se destacado como metodologia (Berger, 2001): análise de discriminante, análise de grupo, análise de correlação, análise de variância e análise de componentes principais.

Vários outros estudos têm usado técnicas estatísticas para identificar relações entre geomorfologia e hidrologia. Zecharias e Brutsaert (1988) identificaram oito variáveis morfológicas com um efeito sobre vazões subterrâneas, entre elas, a área e relevo da bacia e características hidromórficas como comprimento dos rios perenes. Isso foi feito usando análise de componentes principais com dados de 19 bacias localizadas ao longo do platô Appalachian. Os três primeiros componentes, representando tamanho, inclinação e forma, explicaram mais de 98% da variância. Sefton e Howarth (1998) usaram a mesma técnica para examinar relações entre resposta hidrológica dinâmica (parâmetros de perda e propagação) e descritores físicos da bacia como: morfologia, tipo de solo, cobertura vegetal e índices climáticos. Eles estimaram os componentes principais dos descritores físicos e obtiveram os componentes mais significativos em explicar as variáveis hidrológicas. Os quatro primeiros componentes explicaram 63% da variância, mas não pôde estabelecer relações entre os componentes e as seis variáveis hidrológicas.

Ao contrário de Zecharias e Brutsaert (1998), que apenas relacionaram fluxo de água subterrânea com índices morfológicos, Berger (2001) usou em sua análise variáveis hidrológicas explícitas ou modeladas. Ele demonstrou que as características fisiográficas e climáticas podem explicar uma grande porção das diferenças entre bacias no processo hidrológico modelado. Nem clima nem fisiografia sozinhos podem explicar a variabilidade observada entre as bacias. Os parâmetros topográficos usados foram a inclinação mediana, índice de relevo, densidade de drenagem, índice de umidade, capacidade de infiltração e um índice de eficiência da zona saturada. Para identificar quais combinações de variáveis melhor explicam as respostas da bacia foi utilizada uma regressão passo a passo.

a) Análise de Agrupamento

A análise de grupo é um método padrão de análise estatística multivariada cujo objetivo é dividir o conjunto de dados em sub-conjuntos com feições semelhantes e tem sido usado com sucesso na formação de regiões apropriadas para estudos de regionalização e de análise de frequência regional.

Mosley (1981) usou análise de grupos para formar grupos de bacias caracterizadas pela vazão específica média anual e o coeficiente de variação. Acreman (1985) usou análise de grupo hierárquico sobre uma matriz de características de 168 bacias hidrográficas para identificar grupos de bacias similares na Escócia, resultando em cinco regiões, quatro das quais julgou ser homogêneas.

A análise de grupos pode ser usada em consórcio com outras técnicas estatísticas, a exemplo da análise de componentes principais. Farhan (1984) usou análise de grupos para classificar afluentes do rio Jordão em regiões homogêneas com base em quatro componentes principais formados a partir de uma matriz de características das sub-bacias.

b) Análise de componentes principais

Em geral os objetivos da análise de componentes principais são a redução da dimensionalidade e a interpretação do problema.

A análise de componentes principais de um conjunto de dados X pode ser entendida de duas maneiras: a) algebricamente, consistindo em encontrar outras variáveis (componentes principais, Z) que são combinações lineares dos dados X ou da matriz de correlações R , cuja principal propriedade é de apresentar menor dimensionalidade e de serem não correlacionadas, o que não acontece com as variáveis originais; b) geometricamente, os

componentes principais representam a adoção de um novo sistema de eixos coordenados, obtidos pela rotação do sistema original.

As operações envolvidas na análise de componentes principais são feitas sobre a matriz de correlações ou a matriz de variâncias e covariâncias, oriundos do conjunto de dados iniciais, basicamente determinando-se seus autovalores e autovetores. Quando esses dados são apresentados em unidades diferentes deve-se trabalhar com a matriz de correlações, ou seja, com os dados padronizados. Em notação matricial, tem-se:

$$\mathbf{Z} = \mathbf{X} \cdot \mathbf{A} \quad 2.4$$

onde: \mathbf{Z} é a matriz dos componentes principais;

\mathbf{X} é a matriz dos dados;

\mathbf{A} é a matriz dos coeficientes que definem a transformação linear.

A interpretação dos resultados obtidos nesse tipo de análise é geralmente feita sobre o quanto poucos componentes principais podem explicar a variação total dos dados e, sobre a matriz de coeficientes, cuja magnitude mede a importância da variável correspondente na determinação do componente principal.

Uma aplicação desta ferramenta foi feita por Strobel (1987) para delimitar subregiões homogêneas quanto à precipitação, com base nos aspectos conjuntos da simultaneidade de ocorrência, altura precipitada e variabilidade da chuva. Ele usou a matriz de coeficientes para selecionar postos pluviométricos pertencentes à mesma subregião, adotando uma análise sobre cinco componentes principais para uma divisão da área em igual número de subregiões homogêneas.

Burn (1988) usou análise de componentes principais sobre séries de dados de cheias máximas anuais e classificou as estações fluviométricas de acordo com o conjunto rotacionado dos componentes principais, agrupando o mais próximo possível as estações semelhantes.

c) Análise de correlações canônicas

Dada uma matriz X ($n \times p$) contendo as características físico-climáticas (área da bacia, comprimento e declividade do rio principal, etc) e Y ($n \times q$) a matriz das variáveis a regionalizar (vazões máxima, média ou mínima anual, parâmetros de modelos chuva-vazão,

curva de regularização). O objetivo é obter-se duas novas variáveis (U, V), através das seguintes relações:

$$U = a_1x_1 + a_2x_2 + \dots + a_px_p \text{ ou } U = \mathbf{a}^T \mathbf{x} \quad 2.5$$

$$V = b_1y_1 + b_2y_2 + \dots + b_qy_q \text{ ou } V = \mathbf{b}^T \mathbf{y} \quad 2.6$$

As duas novas variáveis U e V maximizam a correlação entre os dois grupos de variáveis x e y. Ao plotar os valores dos pares (U, V) em duas dimensões, pode-se encontrar relações entre os dois grupos de variáveis.

O problema principal é encontrar os vetores **a** e **b** a partir das matrizes **X** e **Y**, os quais são obtidos pela maximização da correlação entre U e V, sujeito às restrições $\text{var}(U)=1$ e $\text{var}(V)=1$.

Se as características das bacias forem boas predictoras das variáveis a regionalizar, o padrão dos pontos nos dois espaços serão similares, permitindo examinar as características das bacias no espaço U_1 e U_2 , vizinhos à bacia sem dados, para estimar a variável regionalizada. Em outras palavras, se houver alta correlação entre as variáveis calculadas, os primeiros dois pares de variáveis canônicas (V_1, V_2) e (U_1, U_2) são suficientes para representar as relações entre as características das bacias e as variáveis regionalizadas.

Torranin (1972) e Rice (1972), e DeCoursey (1973) apresentaram as primeiras aplicações de correlações canônicas em hidrologia.

Segundo Ouarda (2001), apesar das vantagens potenciais, a análise de correlações canônicas tem sido pouco usada no campo da hidrologia. Em uma análise de frequência regional de cheias, por exemplo, as correlações canônicas podem ser usadas para investigar a estrutura de correlação entre dois conjuntos de variáveis, representados pelas características das bacias e os picos de cheias.

d) Análise discriminante

Consiste em encontrar uma transformação ou combinação de características da bacia que minimiza a relação da diferença entre as médias multivariadas do grupo de características da bacia e a variância multivariada das características da bacia entre os grupos. A forma geral da função discriminante é:

$$S = L_0 + L_1U_1 + L_2U_2 + \dots + L_A V_A \quad 2.7$$

onde:

L_i = coeficientes

L_0 = constante

U_i = valores das características da bacia

S = escores discriminantes

Em notação matricial, a equação 2.7 pode ser escrita da seguinte forma:

$$\mathbf{S} = \mathbf{L}^T \mathbf{U} + \mathbf{L}_0 \quad 2.8$$

O escore discriminante linear para o i -ésimo grupo pode ser escrito:

$$S_i = (\mathbf{B}_i^T \mathbf{C}^{-1})\mathbf{U} - \mathbf{B}_i^T \mathbf{C}^{-1} \mathbf{B}_i/2 + \log W_i \quad 2.9$$

onde:

B_i = vetor que contém as médias de cada característica da bacia calculadas para o grupo i .

C = matriz de covariâncias para as características das bacias (assumido igual para todos os grupos).

W_i = probabilidade prévia da bacia U encontrar-se no grupo i . $W_i = N_i/\sum N_i$, onde N_i é o número de regiões do grupo i .

Sousa (1998), aplicou a análise de discriminante em conjunto com o método da análise classificatória e da análise de correlação canônica para delinear regiões homogêneas em Portugal Continental.

A análise classificatória permitiu-lhe identificar regiões similares em relação a variáveis pré-selecionadas; através da análise de correlação canônica determinou a vizinhança hidrológica de uma bacia, considerando que cada estação hidrométrica constitui o centro da região e, por fim, por meio da análise discriminante avaliou a correspondência entre grupos formados com base na análise classificatória (utilizando apenas as características hidrológicas das bacias) e aquelas que seriam formados se somente estivessem disponíveis as características fisiográficas das bacias.

e) Análise de Variância

A análise de variância (ANOVA) se constitui numa ferramenta estatística que pode ser empregada para dividir a variância entre um conjunto de componentes de um ou mais processos de origens diferentes (Clarke, 1988). É mais uma alternativa a ser utilizada nos estudos de regionalização hidrológica.

Wiltshire (1986) usou análise de variância da estatística das vazões para local divisões ótimas de características espaciais da bacia.

f) Momentos – L

De acordo com Versiani & Carneiro (2001), os momentos-L são um sistema alternativo de descrição de formas de funções de distribuição de probabilidades. São somas ponderadas dos valores esperados de estatísticas de ordem (Clarke & Montenegro, 1989). Os momentos-L amostrais são definidos pelas seguintes expressões:

$$l_1 = b_0 \quad 2.10$$

$$l_2 = 2 b_1 - b_0 \quad 2.11$$

$$l_3 = 6 b_2 - 6 b_1 + b_0 \quad 2.12$$

$$l_4 = 20 b_3 - 30 b_2 + 12 b_1 - b_0 \quad 2.13$$

onde:

$$b_r = \frac{1}{n} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} x_{j:n} \quad 2.14$$

n = tamanho da amostra

r = posição de x na série disposta em ordem crescente

Os dois primeiros momentos-L, denominados l_1 (média da amostra) e l_2 (medida de dispersão), se constituem medidas de localização e de escala, respectivamente. As relações de momento-L mais importantes são as seguintes:

$$l_2 / l_1 = \text{L-Coeficiente de Variação}$$

$$l_3 / l_2 = \text{medida de assimetria da distribuição}$$

$$l_4 / l_2 = \text{medida de curtose da distribuição}$$

A extrapolação de l_1 , l_2 , l_2/l_1 , l_3/l_2 e l_4/l_2 a outras bacias sem dados da variável regionalizada, pode ser feita por regressão com características físico-climáticas da bacia.

g) Regressão linear múltipla

Modelo de regressão linear clássico

Especificamente, o modelo de regressão linear com uma única variável dependente tem a seguinte forma (Johnson e Wichern, 1998):

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_r x_r + \varepsilon \quad 2.15$$

[resposta] = [média (dependendo de x_1, x_2, \dots, x_r)] + [erro]

O termo linear refere-se ao fato de que a média é uma função linear dos parâmetros desconhecidos $\beta_0, \beta_1, \beta_2, \dots, \beta_r$. As variáveis independentes podem ou não entrar no modelo como termos de primeira ordem.

Para n observações independentes de Y e os respectivos valores associados de x_i , o modelo completo é dado por:

$$\begin{array}{rcccccc} Y_1 & = & \beta_0 & + & \beta_1 x_{11} & + & \beta_2 x_{12} & + \dots & + & \beta_r x_{1r} & + & \varepsilon_1 \\ Y_2 & = & \beta_0 & + & \beta_1 x_{21} & + & \beta_2 x_{22} & + \dots & + & \beta_r x_{2r} & + & \varepsilon_2 \\ \cdot & & \cdot & & \cdot & & \cdot & & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot & & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot & & & \cdot & & \cdot \\ Y_n & = & \beta_0 & + & \beta_1 x_{n1} & + & \beta_2 x_{n2} & + \dots & + & \beta_r x_{nr} & + & \varepsilon_n \end{array} \quad 2.16$$

onde: os termos referentes ao erro têm as seguintes propriedades:

a) $E(\varepsilon_j) = 0$;

b) $\text{Var}(\varepsilon_j) = \sigma^2$ (constante) 2.17

c) $\text{Cov}(\varepsilon_j, \varepsilon_k) = 0, j \neq k$.

Em notação matricial o sistema de equações é escrito da seguinte maneira:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1r} \\ 1 & X_{21} & X_{22} & \dots & X_{2r} \\ 1 & X_{31} & X_{32} & \dots & X_{3r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nr} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_r \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad 2.18$$

ou ainda:

$$\mathbf{Y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad 2.19$$

sendo: \mathbf{Y} = vetor da variável independente (Y_n)

\mathbf{X} = matriz das variáveis dependentes (X_{nr})

$\boldsymbol{\beta}$ = vetor dos coeficientes

$\boldsymbol{\varepsilon}$ = vetor dos erros

A estimativa de $\boldsymbol{\beta}$ e $\boldsymbol{\varepsilon}$ pelo método dos mínimos quadrados é dada por:

$$\boldsymbol{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \mathbf{Y} \quad 2.20$$

$$\boldsymbol{\varepsilon} = \mathbf{Y} - \hat{\mathbf{Y}} \quad 2.21$$

As equações de regressão múltipla podem ser usadas para prever, a partir das características fisiográficas das bacias, valores de varias estatísticas hidrológicas, tais como: média, desvio padrão, quantis e quantis normalizados de cheias regionais. Os modelos de regressão linear, de acordo com Maidment (1992) têm sido muito usados para previsão de quantis de cheias em bacias sem dados, usando a seguinte equação:

$$y_i = \alpha + \beta_1 \log(\text{area}) + \beta_2 \log(\text{declividade}) + \dots + \varepsilon$$

Onde: y é o quantil de cheia; area é a área da bacia hidrográfica; declividade é a declividade do rio principal; α , β_1 e β_2 são os coeficientes regionais a serem estimados com dados obtidos em bacias com informações; ε é o erro da estimativa.

Modelo de Regressão Stepwise

O erro cometido ao se adotar um modelo de regressão linear, pode estar relacionado, além de sua adequabilidade aos dados, ao efeito da omissão de variáveis predictoras na

equação, às vezes por equívocos durante o processo de seleção ou por nunca terem sido efetivamente medidas ou obtidas em um intervalo inferior ao limite especificado (Draper & Smith, 1981). Igualmente, variáveis redundantes ou irrelevantes incluídas na equação, desqualificam o modelo de regressão obtido.

O Modelo Stepwise é um procedimento de seleção da melhor equação de regressão pela inclusão e/ou substituição sucessivas de variáveis preditoras. A ordem de inserção é determinada pela medida de importância das variáveis ainda não incluídas na equação; este nível de importância é avaliado pelo coeficiente de correlação parcial.

De acordo com Draper & Smith (1981), o procedimento básico do Modelo Stepwise é o seguinte:

1. Selecionar entre as variáveis preditoras Z a mais correlacionada com a variável dependente Y (suponha z_1) e encontrar a equação de regressão linear de primeira ordem $\hat{Y} = f(z_1)$.
2. Checar se z_1 é significativa. Se não for, descartá-la e adotar um modelo $Y = Y_{\text{médio}}$ como o melhor; caso contrário, pesquisar a segunda variável preditora a entrar na equação. Isto é feito, examinando o coeficiente de correlação parcial de todas as variáveis que ainda não entraram na equação, isto é, z_j , $j \neq 1$, com Y . O z_j com o maior coeficiente de correlação parcial com Y é então selecionado (suponha z_2), e a segunda equação de regressão é ajustada.
3. A regressão global é checada pela significância, a melhoria no valor de R^2 é anotado, e o F-valor parcial para as variáveis z_1 e z_2 (não apenas z_2) são examinados. O menor dos dois F's parcial é então comparado com um ponto percentual F apropriado, e a correspondente variável preditora é retida na equação ou rejeitada caso o teste seja significativo ou não. Um preditor que pode ter sido o melhor candidato a entrar na equação em um estágio anterior pode, num estágio posterior, ser considerado supérfluo por causa das relações entre ele e outras variáveis contidas na equação de regressão.
4. Parar o procedimento quando não mais existirem variáveis a serem removidas da equação e quando a próxima melhor variável não seja aceita na equação. Como é testada a entrada de cada variável na equação, conseqüentemente obtém-se uma série de avaliações do efeito sobre R^2 (coeficiente de correlação múltipla).

Existem na bibliografia outros métodos de seleção de variáveis, tais como o que avalia todas as regressões possíveis, o que procura o melhor sub-conjunto de variáveis e o de eliminação regressiva, mas, o Stepwise é tido, entre os que avaliam a relevância das

características, como o de melhor desempenho. É evidente que não existe um método superior a todos os outros, mas, cada caso requer um método particular.

2.4.8 Redes Neurais Artificiais

As redes neurais artificiais têm se destacado nos últimos anos nas aplicações relacionadas a recursos hídricos por oferecer diversos atrativos em relação às demais metodologias estatísticas e matemáticas. Entre outras, a aprendizagem por exemplos, independência do problema, processamento distribuído, paralelo e local e implementação realística.

Destaca-se entre as aplicações: French et al (1992) para previsões espaciais de campos de chuvas intensas; Raman e Sunilkumar (1995) para geração de séries sintéticas de vazões na bacia do rio Bharathapuzha no Sudeste da Índia; Hjelmfelt e Wang (1993) para operar em substituição ao hidrograma unitário; Kang et al (1993) para previsões de vazões horárias e diárias a partir do uso de quatro arquiteturas de redes neurais; Zhu et al (1994) em previsões de escoamentos de 1-hora, 2-horas e 3-horas para a bacia do rio Butternut Creek em Nova York; Hsu et al (1995) para desenvolver um algoritmo mais eficiente baseado nos mínimos quadrados e treinar uma RNA com dados chuva-vazão; Campolo et al (1999) para previsão de níveis de rios em período de baixas vazões; Diniz e Clarke (2001) para regionalização dos parâmetros do modelo chuva-vazão SMAP versão mensal; Mazvimavi (2003) para a estimativa de características de vazões em bacias sem dados em Zimbabwe.

O método de regressão linear múltipla em regionalização foi severamente criticado no artigo de McIntyre et al (2005). Este método pode ser inapropriado para a regionalização de parâmetros de modelos chuva-vazão porque assume a linearidade das relações destes com as características das bacias e que estas variáveis têm distribuição que se aproxima da distribuição normal. As duas hipóteses não são sempre válidas e, desta forma, outros métodos teriam que ser usados para identificar a influência das características preditivas das bacias, e subsequentes técnicas preditivas.

Uma das técnicas possíveis é o uso da RNA, que são capazes de modelar as relações não-lineares, e não necessitam assumir previamente uma distribuição básica específica para os dados (Orr et al, 1999). As redes neurais têm sido amplamente usadas em hidrologia tanto para previsão de vazões (Dolling & Varas, 2003), quanto para regionalização de parâmetros de modelos chuva-vazão (Heuvelmans et al, 2006).

De acordo com Valença (1999), os principais benefícios das RNA são:

- Não linearidade: permite a modelagem de uma grande diversidade de variáveis;
- Adaptatividade: capacidade de se adequar a um novo ambiente através de retreinamento;
- Aprendizagem: conhecimento extraído de exemplos;
- Modelagem de sistemas complexos: permite trabalhar com uma grande quantidade e diversidade de dados;
- Processamento paralelo: testa todas as possibilidades ao mesmo tempo;
- Tolerância a falhas: quando sujeita a condições adversas, sofre perda gradual do desempenho e não uma perda catastrófica;
- Capacidade de generalização quando treinada adequadamente: fornecer resposta razoável quando estimulada com novas informações.

Segundo Osório (1999), as desvantagens mais importantes no uso de redes neurais são as seguintes:

- Dificuldade de configuração das redes em relação à sua estrutura inicial e no que se refere aos parâmetros dos algoritmos de aprendizado (taxa de aprendizado, coeficiente de inércia e número de ciclos);
- Dificuldade de explicitar os conhecimentos adquiridos pela rede através de uma linguagem compreensível para um ser humano;
- Possível dificuldade de convergência e instabilidade, inerentes aos algoritmos de otimização empregados;
- Lentidão do processo de aprendizado.

2.4.8.1 Base conceitual

O neurônio biológico

O cérebro humano possui bilhões de neurônios, interligados uns aos outros, e forma uma elaborada rede neural. Recebe continuamente sinais de muitas fontes e processa-os para criar uma resposta adequada. Essa rede executa milhões de funções para sustentar a vida.

No neurônio ou célula nervosa, são identificadas quatro porções com funções distintas: o corpo celular ou soma, a árvore dendrital, o axônio e terminais pré-sinápticos, conforme ilustrado na Figura 2.1.

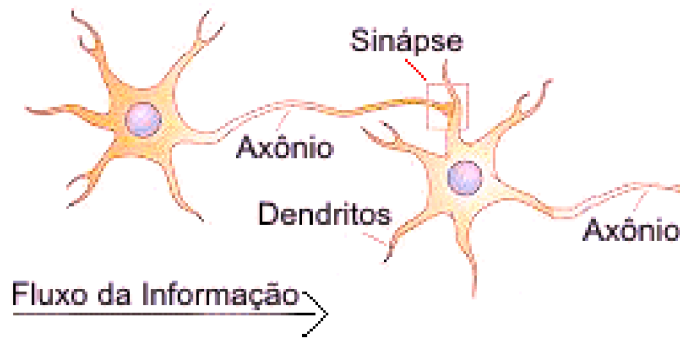


Figura 2.1: **Estrutura do neurônio biológico**

O corpo celular é a porção que contém o núcleo celular e o citoplasma. O axônio, geralmente único em cada neurônio, é responsável pela transmissão de informação na forma de pulsos elétricos regenerativos para as várias partes do sistema nervoso e do organismo. Os dendritos são a parte receptiva de informação do neurônio e se caracteriza por prolongamentos filamentosos que se arranjam em forma de árvores.

O contato entre o axônio de um neurônio e os dendritos de outro é feito através da sinapse que se caracteriza por propagar a informação em direção única. Existem cerca de 5000 a 95000 conexões sinápticas de um neurônio com outros.

O neurônio artificial

O neurônio artificial é uma estrutura lógica-matemática que procura simular a forma, o comportamento e as funções de um neurônio biológico. Assim sendo, os dendritos foram substituídos por *entradas*, cujas ligações com o corpo celular artificial são realizadas através de elementos chamados de *peso* (simulando as sinapses). Os estímulos captados pelas entradas são processados pela *função de soma*, e o limiar de disparo do neurônio biológico foi substituído pela *função de transferência*, conforme a Figura 2.2.

Redes neurais

Uma rede neural artificial – RNA consiste de um conjunto de elementos de processamento – EP, os quais são imitações dos neurônios biológicos, interligados entre si pelas chamadas conexões neurais (numa alusão à sinapse que é a conexão entre o terminal de um neurônio e o dendrite de outros neurônios), formando camadas. A RNA possui uma

camada de entrada de dados, uma ou mais camadas escondidas ou intermediárias e outra de saída.

Existem três características principais que descrevem uma rede neural, e que contribuem para sua habilidade funcional: arquitetura, dinâmica e aprendizado.

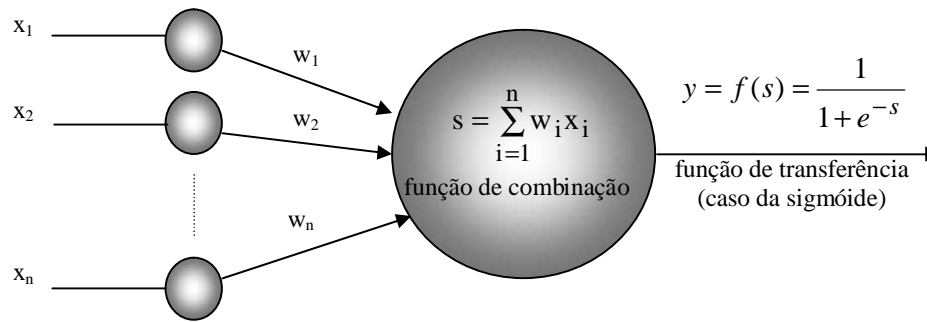


Figura 2.2: **Modelo de um neurônio artificial**
(adaptado de Loesch e Sari, 1996)

Arquitetura

Em função do processo de definição da arquitetura da rede elas podem ser: micro-arquitetura e meso-arquitetura.

A micro-arquitetura neural é definida por um neurônio. A operação de um neurônio se constitui de uma função de combinação ou propagação (s) e outra de transferência ou de ativação [$f(s)$], conforme Figura 2.2. O valor da função de transferência será um dos componentes de entrada do neurônio seguinte.

O nível meso-arquitetura de uma RNA é aquele em que a organização e arranjo físico dos Eps na rede são considerados, obedecendo ao seguinte roteiro:

- número de camadas da rede;
- número de neurônios por camada;
- tipo de conexão;
- grau de conectividade entre os Eps (um a um, plena, esparsa, aleatória, etc)

Fazem parte, portanto, da definição da arquitetura da rede, o número de camadas, o número de neurônios em cada camada e o tipo de conexão entre os neurônios e como as funções são interconectadas.

A Figura 2.3 apresenta a arquitetura de uma RNA do tipo backpropagation, feedforward ou perceptron de múltiplas camadas.

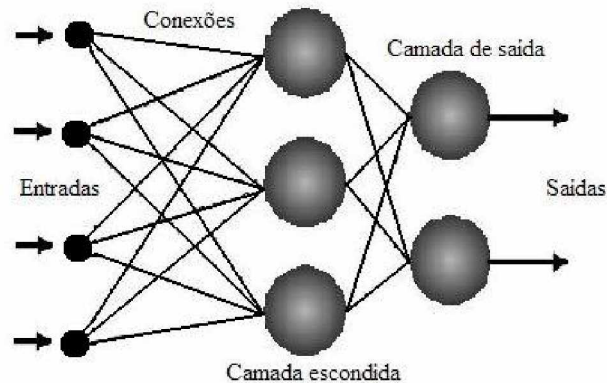


Figura 2.3: Arquitetura de uma RNA tipo MLP

A formulação matemática da rede neural da Figura 2.3 é feita da seguinte maneira:

$$\mathbf{O} = \mathbf{W}_{H,O} f(\mathbf{W}_{L,H} \cdot \mathbf{I} \cdot \beta)$$

Onde:

\mathbf{O} é o vetor das variáveis de saída ($\in \mathbb{R}^l$)

\mathbf{I} é o vetor das variáveis de entrada ($\in \mathbb{R}^m$)

β é o vetor de erros ($\in \mathbb{R}^n$)

$\mathbf{W}_{H,O}$ é a matriz de pesos entre as camadas escondida e de saída ($\in \mathbb{R}^{l \times n}$)

$\mathbf{W}_{L,H}$ é a matriz de pesos entre as camadas de entrada e escondida ($\in \mathbb{R}^{n \times m}$)

f é a função de transferência (sigmóide, gaussiana, hiperbólica secante, etc)

Dinâmica

Toda rede tem um modo específico no qual ela processa seus dados e está relacionado com sua estrutura. Por exemplo, em uma rede “feedforward” multicamadas todo o fluxo de informações é passado adiante, de uma camada para outra, conforme a figura anterior. Já em uma rede do tipo lateral, existe apenas uma camada e as informações transitam na mesma camada de um EP para outro e isso várias vezes.

No caso da rede “feedforward” cada EP recebe diversas entradas e produz somente uma saída. Os elementos da mesma camada atuam em paralelo, enquanto que de uma camada para

outra é sequencial. Quanto à sua conectividade pode ser parcialmente conectada ou completamente conectada.

Treinamento

Um tipo de classificação das redes neurais é quanto à forma de treinamento, que pode ser: a) não supervisionado: treinada para responder com grupos de padrões relativos à entrada; b) supervisionado: treinada para produzir saídas semelhantes a valores previamente conhecidos; c) híbrido: combina as duas abordagens anteriores.

No treinamento não supervisionado é classificado um conjunto de padrões de treinamentos em um número especificado de categorias sem ser mostrado antes como categorizar. A rede faz isto agrupando os padrões. Ela agrupa-os por sua proximidade em um espaço dimensional N , onde N é o número de padrões ou atributos. Redes Kohonen (Kohonen, 1990) são exemplos de redes não supervisionadas.

O treinamento supervisionado da RNA é a determinação dos pesos que faz o vetor de saída da rede se aproximar ao máximo possível do vetor desejado.

O treinamento da RNA, feito com o algoritmo de retropropagação, é a determinação dos pesos que faz o vetor de saída da rede se aproximar o máximo possível do vetor desejado. De forma resumida, o treinamento obedece aos seguintes passos (Loesch & Sari, 1996):

- a) Atribuir, aleatoriamente, valores aos pesos W_{ij}^k , preferivelmente próximos a zero.
- b) Sejam \mathbf{x} e \mathbf{d} os vetores de dados de entrada da rede e o desejado, respectivamente. No caso do presente estudo, \mathbf{x} é o vetor das características físico-climáticas das bacias hidrográficas e \mathbf{d} o vetor dos parâmetros do modelo SMAP. Aplicar o vetor \mathbf{x} na camada de entrada e propagar os sinais da rede até a camada de saída (operação feedforward). Seja \mathbf{y} o vetor de saída da rede. Calcular o erro quadrático $\epsilon^2 = \sum (d_i - y_i)^2$. Se for inferior ao valor de tolerância a rede é considerada treinada. Caso contrário, prosseguir no passo c.
- c) Fazer k igual à última camada.
- d) Para todos os elementos j da camada k calcular:

$$- \epsilon_j^k = d_j - x_j^k = d_j - y_j \text{ se } k \text{ for a última camada; ou}$$

$$- \epsilon_j^k = \sum_{i=1}^{N_{k+1}} \delta_i^{k+1} w_{j,i}^{k+1} \text{ se for uma camada oculta}$$

$$- \delta_j^k = \epsilon_j^k f'(s_j^k)$$

- e) $k \leftarrow k-1$. Se $k > 0$ ir para o passo d; caso contrário prosseguir no passo f.
- f) Recalcular todos os pesos de conexão da rede empregando a seguinte expressão:

$$W_{j}^k(n+1) = W_{j}^k(n) + 2 \eta \delta_j^k(n) X_j^k(n) \quad 2.22$$

onde: η é a taxa de aprendizagem.

- g) Tome outro par para treinamento (x, d) e retorne ao passo b.

Um dos aspectos mais relevantes na modelagem através das redes neurais é a definição da camada de entrada por meio das variáveis de entrada. É interessante pensar no problema que se deseja modelar e no modo como as RNA's funcionam, para se escolher quais informações devem ser fornecidas, de tal modo a obter-se um melhor desempenho.

No caso da regionalização de parâmetros de modelos chuva-vazão, o uso da técnica de redes neurais artificiais se deu numa etapa posterior à calibração do modelo adotado para as bacias hidrográficas integrantes do estudo de regionalização. As características físicas e climáticas das bacias compõem os dados de entrada chamados explicativos, enquanto os valores dos parâmetros do modelo obtidos na calibração serão os dados de entrada que orientarão o treinamento da rede, portanto um treinamento supervisionado.

Critério de parada do processo de treinamento da RNA

O critério de parada do algoritmo backpropagation não é bem definido, e geralmente é utilizado um número máximo de ciclos. Mas, devem ser considerados a taxa de erro médio por ciclo, e a capacidade de generalização da rede. Pode ocorrer que em um determinado instante do treinamento a capacidade de generalização comece a degenerar, causando o problema de “over-training”, ou seja, a rede se especializa no conjunto de dados do treinamento e perde a capacidade de generalização. Por outro lado, se o treinamento for parado prematuramente a rede ficará sub-ajustada e, igualmente, terá baixa capacidade de generalização.

De acordo com esse critério, quando o valor do erro, calculado para o conjunto de validação, alcança seu valor mínimo, o treinamento deve ser interrompido, conforme ilustrado na Figura 2.4.

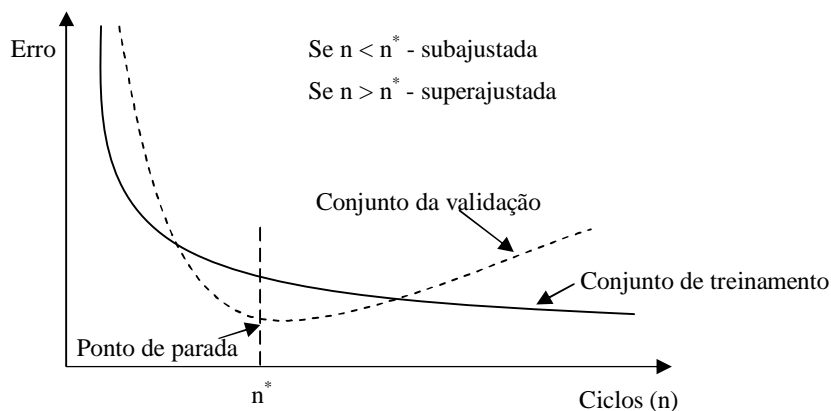


Figura 2.4: Critério de parada do treinamento por validação cruzada

Pode ocorrer que o ponto mínimo (ponto de parada) não seja o primeiro a acontecer como se presume na maioria das aplicações, mas em outro ponto da curva só revelado com a extensão do processo de treinamento. Sendo assim, é possível obter-se diversas combinações de pesos da RNA capazes de fornecer o valor da resposta, o que dificulta a seleção de um ponto de parada. Eventualmente, ocorrem casos com pontos ótimos globais e únicos.

2.5 VARIÁVEIS E FUNÇÕES HIDROLÓGICAS REGIONALIZADAS

2.5.1 Considerações gerais

Os estudos de regionalização envolvem desde a escolha dos métodos ou procedimentos a serem empregados até os dados da variável a regionalizar. Muitas vezes tais dados são obtidos em bacias com grandes áreas ou escalas muito pequenas, diferentes das situações para as quais se deseja o estudo, ocasionando erros e restrições na hora do uso da regionalização.

Como em geral os dados de vazão são mais raros e são muito suscetíveis a erros, os estudos se concentram mais nesta variável. Entretanto, outras variáveis e funções hidrológicas bastante estudadas são as precipitações pluviométricas, as curvas de regularização e de permanência e o hidrograma unitário.

Apresenta-se aqui as variáveis e funções mais comumente usadas em estudos de regionalização hidrológica.

2.5.2 Vazão média

Nas bacias onde não existem dados fluviométricos observados a obtenção da vazão média pode ser conseguida por regionalização, a qual pode ser feita de duas maneiras:

- a) Regionalizando a curva adimensional de probabilidades de vazões médias anuais. Neste caso, as vazões adimensionais de cada posto são plotadas em um mesmo gráfico e as regiões com mesma tendência são identificadas.
- b) Regionalizando os parâmetros da equação de regressão obtidos pela relação entre a vazão média e características físicas e climáticas da bacia. Em geral, a área da bacia (A) e a precipitação média (P) explicam a maior parcela da variabilidade da série de vazões médias.

$$Q = a A^b P^c \quad 2.23$$

Sendo, a, b e c os coeficientes regionais obtidos pelo método dos mínimos quadrados.

2.5.3 Vazão mínima

Em geral os estudos de regionalização de vazões mínimas são realizados de três formas (Tucci, 2002): 1) regionalização da curva de permanência ou parte dela; 2) da curva de probabilidades de vazões mínimas; 3) com base na adimensionalização da série de vazões, tal como utilizado para vazões médias.

A regionalização da vazão mínima com base no método de adimensionalização obedece a seguinte ordem:

- a) Séries de vazões para cada duração a ser utilizada;
- b) Curvas individuais de probabilidades de vazões mínimas;
- c) Equação de regressão entre as vazões mínimas e as características físico-climáticas das bacias hidrográficas.
- d) Definição das regiões homogêneas através do comportamento das curvas de probabilidades.

2.5.4 Vazão máxima

Os principais métodos de regionalização de vazões máximas utilizam a curva de probabilidade de cada estação fluviométrica e regressões simples ou multivariada.

Podem ser regionalizados valores de vazões máximas associadas a tempos de retorno específicos, obtidos das curvas individuais de probabilidades de cada estação, com base em características físicas e climáticas da bacia. De igual forma, podem ser regionalizados os parâmetros da distribuição estatística.

O método da curva adimensional consiste em determinar as curvas adimensionais e, pela tendência média destas, a curva média regional. A curva regional pode ser traçada manualmente ou ajustando-se uma distribuição teórica de probabilidades aos pontos médios obtidos a partir das curvas individuais.

2.5.5 Regionalização da curva de regularização

A curva de regularização é a função que relaciona a capacidade de armazenamento S de um açude com a vazão regularizável Q_{reg} , estando esta associada a uma dada garantia de atendimento p , representada pela seguinte equação:

$$S = f(Q_{reg}, p) \quad 2.24$$

Existem vários métodos para a determinação do volume regularizável, dentre os quais se destaca aquele baseado na equação da continuidade operacionalizada de forma interativa, segundo a seguinte equação:

$$VA - VE = \Delta S \quad 2.25$$

sendo:

VA = todos os volumes que afluem a seção (= Q);

VE = volumes efluentes (= Q_{reg} , incluindo a evaporação);

ΔS = variação do volume no intervalo de tempo considerado (= $S_{t+1} - S_t$)

A equação 2.23 adimensionalizada em relação à vazão média da bacia, pode ser escrita da seguinte forma:

$$S_a = f(\beta, p) \quad 2.26$$

onde: $\beta = Q_{reg}/Q_{med}$ e S_a é a capacidade máxima adimensional do açude, dado por:

$$S_a = S/Q_{med} \quad 2.27$$

Com base nos dados de cada reservatório se determina pares de pontos (S , Q_{reg}), com os quais se pode ajustar uma curva do tipo:

$$S_a = a \beta^b \quad 2.28$$

onde: a e b são constantes obtidas por análise de regressão.

A regionalização da equação 2.27 pode ser feita, segundo Medeiros et al (2002), de duas maneiras: a) estabelecer a regressão entre as constantes a e b da equação com base em características físicas das bacias estudadas; b) ajuste de uma curva média com postos de tendência semelhante.

Domokos (1990) e Tucci(1991) utilizaram o primeiro método sem sucesso, tentando regionalizar os parâmetros com os coeficientes de variação das bacias. O segundo método foi utilizado por Tucci (1991) e, depois, por Obregon (1999) usando séries estendidas através de um modelo chuva-vazão mensal.

A delimitação de regiões homogêneas em termos de regularização foi feita por agrupamento das curvas de regularização segundo as suas tendências, permitindo obter-se curvas regionais de regularização, válidas para o intervalo de vazão adimensional onde a curva pode ser usada. Sua aplicação em bacias que não tenham dados pode ser feita após regionalizar a vazão média Q_{med} em função das características físicas das bacias hidrográficas e combiná-las com as equações 2.26 e 2.27.

Medeiros *et al.* (2002) realizaram uma pesquisa, tendo como meta introduzir o conceito de diferentes garantias associadas ao atendimento da demanda na regionalização da curva de regularização e buscaram avaliar a metodologia simplificada apresentada por Tucci (1993) para considerar a evaporação na extrapolação da mesma. Obtiveram resultados satisfatórios, pelo menos para a região do estudo na bacia do rio Apodi, no Rio Grande do Norte.

2.5.6 Regionalização da curva de permanência

A curva de permanência indica a porcentagem de tempo que um determinado valor de vazão foi igualado ou ultrapassado dentro do período dos dados usados para construí-la. Esta curva é especialmente importante para estudos de vazões mínimas, por exemplo a vazão Q_{95} que é aquela com permanência de 95% do tempo.

De forma geral, sua regionalização é feita para vazões associadas a períodos de retornos específicos obtidas da curva empírica ou de uma distribuição de probabilidade ajustada. Outra opção é regionalizar os parâmetros da distribuição teórica com base nas características físicas e climáticas de bacias hidrográficas homogêneas do ponto de vista hidrológico.

Para uma dada vazão de permanência Q_p a equação matemática do tipo multiplicativo intrinsecamente linear é geralmente utilizada nos estudos de regionalização;

$$Q_p = \alpha \cdot A^a \cdot B^b \cdot C^c \dots Z^z \quad 2.29$$

onde:

A, B, C, ..., Z são variáveis que expressam as características das bacias;

α , a, b, c, ..., z são os parâmetros da equação determinados por análise de regressão;

Q_p = vazão de permanência p% para cada bacia (m^3/s).

Uma grande dificuldade é encontrar as características mais relevantes da bacia para explicar as vazões pontuais retiradas da curva de permanência. De Coursey (1973) usou análise de correlações canônicas para selecionar as características das bacias hidrográficas e depois determinou uma matriz de regressão múltipla para regionalizar vazões de cheias, de acordo com a seguinte expressão:

$$Q = a \cdot A + b \quad 2.30$$

onde:

Q é um vetor coluna das vazões de cheias associadas a vários períodos de retorno;

A é um vetor coluna das características da bacia hidrográfica;

a e b são, respectivamente, uma matriz e um vetor dos coeficientes de regressão.

Uma generalização da aplicação de De Coursey (1973) é a análise de correlação múltipla multivariada, que preserva a intercorrelação entre as variáveis dependentes.

2.5.7 Regionalização dos parâmetros de uma distribuição de probabilidade

Uma outra possibilidade de regionalização pode ser feita através da análise de regressão entre as características da bacia hidrográfica e os parâmetros de uma distribuição de probabilidades escolhida, em vez de trabalhar com os valores de vazões para um dado período de retorno. Os parâmetros são determinados para cada bacia da região usando-se o método dos momentos ou da máxima verossimilhança e, depois, equações de regressão podem ser desenvolvidas para estimar os parâmetros da distribuição a partir de características da bacia.

No caso da distribuição Log-Pearson III, por exemplo, a média, o desvio padrão e o coeficiente de assimetria são os parâmetros a regionalizar.

Aplicações nesse sentido foram feitas por Thomas e Benson (1969) usando cerca de setenta parâmetros das vazões do rio, incluindo vazões de cheias, volumes de cheias, médias mensais e anuais de vazão e distribuição estatística das vazões. Cada um destes parâmetros foi estudado em relação às características de mais de 30 bacias.

2.5.8 Regionalização do hidrograma unitário

O hidrograma unitário pode ser definido como o hidrograma resultante de um escoamento superficial devido a uma precipitação de volume unitário. Representa o escoamento superficial padrão de uma chuva uniformemente distribuída de 1 cm de altura.

A determinação de um hidrograma unitário (HU) pode ser feita pela resolução da seguinte equação (TVA, 1961), em notação matricial, equivalente ao método dos mínimos quadrados.

$$H = (P^T \cdot P)^{-1} \cdot (P^T \cdot Q) \quad 2.31$$

onde:

H = é o vetor coluna contendo as ordenadas do HU (n x 1);

P = é a matriz de precipitações efetivas (m x n), sendo m e n os números de ordenadas dos hidrogramas observado e HU, respectivamente;

Q = é o vetor coluna contendo as vazões observadas (m x 1).

Devido a dificuldade em se obter dados de chuvas e vazões, necessários para a construção do hidrograma unitário, vários pesquisadores estudaram uma alternativa para obtenção do mesmo a partir de características físicas da bacia e de informações inferidas do comportamento geral dos hidrogramas. Desenvolveram o denominado hidrograma unitário sintético (HUS), o qual se baseia em equações empíricas aplicáveis a determinadas regiões. Snyder (1938) foi um dos pioneiros a estabelecer um HUS, no seu caso, para bacias rurais da região das Montanhas Apalachianas, o qual, ainda hoje tem sido usado, particularmente para grandes bacias, embora não recomendados para tais casos.

Em áreas urbanas, Eagleson (1962) estudou um HUS com dados de cinco bacias em Louisville (Kentucky), particularizando os parâmetros do HUS de Snyder (1938) para aquela região.

Diaz e Tucci (1989) apresentaram um estudo de regionalização do HUS de Snyder usando dados de dezenove bacias urbanas brasileiras, relacionando os parâmetros que definem o hidrograma com as áreas das bacias (A), porcentagem de áreas impermeáveis (PAI), comprimento do talvegue (L), declividade do rio principal (I) e o comprimento do talvegue principal até o centro de gravidade da bacia (LCG).

Em linhas gerais a metodologia empregada consistiu em:

- Seleção dos eventos.
- Separação dos escoamentos.
- Determinação do HUS de cada evento.
- Obtenção do HUS médio para cada grupo de bacias.
- Regionalização dos HUS médios.

A regionalização foi feita por análise de regressão múltipla a partir da seguinte equação de potência:

$$Y = \alpha . A^a . B^b . C^c \dots Z^z \quad 2.32$$

onde:

A, B, C, ..., Z são variáveis que expressam as características das bacias (A, PAI, L, I, LCG);

α , a, b, c, ..., z são constantes determinadas por análise de regressão;

Y = variável dependente (vazão e tempo de pico, tempo de base e larguras nas quais a vazão corresponde a 25% e 50% da vazão de pico).

2.5.9 Regionalização de parâmetros de modelos chuva-vazão

A regionalização hidrológica tem como princípio a transferência de informações de bacias hidrográficas com dados para outras sem dados (Mosley, 1981). A regionalização de parâmetros de modelos chuva-vazão não é uma tarefa fácil pois se depara com diversas limitações, entre as quais os problemas inerentes ao processo de calibração dos modelos, interdependência dos parâmetros e dificuldades de se encontrar boas relações regionais, ou seja, alguns parâmetros dos modelos podem não se relacionar com os atributos das bacias, pelo menos os que usualmente se adotam (Abdulla & Lettenmaier, 1997).

Além destes, Heuvelmans et al (2006) destacam a seleção de descritores físicos das bacias como a primeira condição para se obter sucesso na regionalização de parâmetros de

modelos chuva-vazão. Diversos estudos iniciam de uma lista preliminar de descritores de bacias com base no conhecimento do modelo ou da área em apreço e, usando uma análise de componentes principais ou correlação selecionam as características mais apropriada (Sefton & Howarth, 1998; Mwakalila, 2003).

Conforme mencionado por Seibert (1999), o tipo e o número de descritores que são usados dependem do escopo geográfico do estudo. Em geral, grandes áreas apresentam enorme variabilidade nos parâmetros do modelo, requerendo uma quantidade maior de descritores físicos das bacias para explicar essa variabilidade. Heuvelmans et al (2006) citam como exemplo as variáveis climáticas que podem ser omitidas em pequenas áreas, mas pode ser um fator significativo em estudos de grandes áreas.

Apesar dos estudos de regionalização hidrológica já existirem à bastante tempo, a regionalização de parâmetros de modelos conceituais chuva-vazão só está sendo considerada mais recentemente.

As primeiras tentativas de regionalização de parâmetros de modelos chuva-vazão foram feitas à poucas décadas, entre outros, por James (1972) e Magette et al (1976).

Durante a última década grande número de disciplinas da área ambiental tem usado modelos hidrológicos no planejamento e gerenciamento dos recursos ambientais. As maiores dificuldades encontradas aparecem na aplicação dos modelos para grandes bacias quando foram calibrados para áreas pequenas e vice-versa, apontando para a necessidade da regionalização de equações ou parâmetros do modelo.

Devido ao aumento da demanda pelo uso dos modelos chuva-vazão, hoje muito mais viável face a evolução da tecnologia computacional, a sofisticação dos dados hidrometeorológicos e ao reduzido número de bacias com dados, cresce a necessidade por métodos regionais que permitam estimar os parâmetros destes modelos a partir de atributos das bacias que sejam facilmente obtidos. Kokkonen et al (2003), Hundecha & Bárdossy (2004) e Merz & Blöschl (2004) são exemplos do renovado interesse por esta matéria.

Riedel e Maniak (1999) desenvolveram estudos de regionalização dos parâmetros do modelo NAXOS que caracterizam o escoamento direto com dados da bacia hidrográfica do rio Aller do nordeste da Alemanha, sendo os demais parâmetros estimados na fase de calibração. O escoamento direto não teve relação com o tamanho da bacia e o comprimento do rio principal, mas variou com a declividade e o tipo de solo da bacia.

Um método para parametrização de modelos hidrológicos em mesoescala foi apresentado por Funke et al (1997), no qual os parâmetros do modelo são correlacionados com características da bacia hidrográfica, utilizando uma ferramenta GIS, através de equações empíricas obtidas por um processo de otimização numérica, análise determinística e

comparações. O método foi testado num conjunto de sete sub-bacias do rio Pruem com áreas entre 18 e 574 Km². Uma análise de regressão detectou relações entre parâmetros do modelo e características físicas conhecidas da bacia. Alguns parâmetros conceituais do modelo foram substituídos diretamente por parâmetros físicos da bacia e outros quantificados por funções de transferências, as quais servirão para parametrização e simulação em bacias sem estações de medições.

A aplicação de técnicas de regionalização aos parâmetros de um modelo concentrado com estrutura simples foi apresentada por Diniz e Clarke (2001), no qual o conjunto de parâmetros obtidos na calibração é associado às características das bacias através do uso de uma rede neural artificial. Cada uma das 14 bacias usadas no estudo foi omitida do processo de calibração e a rede neural foi usada para estimar os parâmetros do modelo para a bacia omitida. Comparações entre os parâmetros estimados para cada bacia omitida e os parâmetros obtidos calibrando o modelo aos dados daquela bacia, apresentaram “resultados bons para algumas bacias e piores para outras, provavelmente porque as bacias estão em regiões com diferenças hidrológicas”.

A bibliografia tem mostrado que existe forte correlação entre a geomorfometria de uma bacia hidrográfica e os processos hidrológicos observados em diferentes escalas. No trabalho de Hennrich et al (1999) há importância da geomorfometria por causa da forte correlação entre diferentes parâmetros morfométricos e índices hidrográficos, especialmente aqueles relacionados à morfometria do solo.

Embora muitos estudos objetivem a regionalização de um modelo hidrológico diferente, todos seguem uma mesma metodologia geral. Primeiro, um modelo chuva-vazão é calibrado para cada bacia com dados disponíveis. Isto é seguido pela aplicação de um método de regionalização que busca relacionar os parâmetros obtidos na calibração com certas características físico-climáticas das bacias.

O método estatístico de regressão múltipla multivariada é o mais empregado nos estudos de regionalização de parâmetros de modelos chuva-vazão. Alternativamente, o uso das redes neurais artificiais têm surgido mais recentemente com resultados satisfatórios (Mazvimavi, 2003; Heuvelmans et al, 2006). Não raro, estes dois métodos operam de forma consorciada com outro, a exemplo da análise de componentes principais, que reduz a dimensionalidade do problema e das técnicas de agrupamento por similaridade entre bacias, conforme mostrado por Burn & Boorman (1992).

A regressão linear tem sido de longe a técnica mais usada no contexto da regionalização, talvez pela facilidade de uso e de interpretação dos resultados. As redes neurais são modelos mais flexíveis e podem facilmente levar em conta a não linearidade e os

efeitos de inter relação entre parâmetros (Lek & Guégan, 1999; Maier & Dandy, 2001). Face à conhecida interdependência que caracteriza tanto os descritores físicos das bacias quanto os parâmetros dos modelos chuva-vazão, espera-se uma relação não linear entre um e outro. Desta maneira, um modelo de regionalização baseado em redes neurais artificiais se mostra bastante atrativo.

Fernandez et al (2000), apresentou uma nova metodologia em que a calibração do modelo e a relação dos seus parâmetros com atributos das bacias são perseguidos simultaneamente. Naturalmente, este método é computacionalmente exaustivo e, ao nosso ver, necessita-se estudar em maior profundidade os algoritmos de otimização que possam, ao mesmo tempo, minimizar os erros do modelo e maximizar a qualidade do ajuste das relações entre parâmetros do modelo e atributos das bacias.

CAPÍTULO III: ÁREA DE ESTUDO

3.1 CARACTERIZAÇÃO FÍSICO-CLIMÁTICA DA ÁREA

O estudo foi desenvolvido no âmbito do território paraibano, região semi-árida do Nordeste brasileiro. As 41 bacias hidrográficas selecionadas, usadas neste trabalho, estão distribuídas nas onze bacias hidrográficas que compõem a divisão oficial do estado da Paraíba, conforme mostrado na Figura 3.1.

3.1.1 Localização e extensão

O estado da Paraíba localiza-se no litoral oriental do Nordeste Brasileiro, no quadrilátero compreendido pelos paralelos de $06^{\circ} 02' 12''$ e $08^{\circ} 19' 18''$ de latitude sul e os meridianos de $34^{\circ} 45' 54''$ e $38^{\circ} 45' 45''$ de longitude oeste de Greenwich. Limita-se ao norte com o estado do Rio Grande do Norte, a leste com o Oceano Atlântico onde está situado o ponto mais oriental das Américas (Ponta de Seixas), a oeste com o estado do Ceará e ao Sul com o estado de Pernambuco.

A Paraíba conta com uma área oficial de $56.439,838 \text{ km}^2$ (Paraíba, 2003), o que representa 0,66% da superfície territorial brasileira e 3,63% da área da região Nordeste. Seu litoral se estende desde a foz do rio Guajú na divisa com o Rio Grande do Norte até a foz do rio Goiana, na divisa com o Estado de Pernambuco, perfazendo um total de 138 km.

Seu território está dividido em quatro mesorregiões geográficas (Mata Paraibana, Agreste Paraibano, Borborema e Sertão) e 23 microrregiões geográficas, cada uma com suas peculiaridades envolvendo o processo social, como determinante, o natural como condicionante e a rede de comunicação e de lugares como elementos de articulação espacial.

Possui 223 municípios e 278 distritos e uma população de 3.641.395 habitantes (IBGE, 2007), sendo 2.684.922 habitantes (73,7%) na área urbana, e 1.033.202 habitantes (26,3%) na área rural. A taxa de ocupação média é de $64,5 \text{ hab/km}^2$. De leste para oeste sua maior extensão linear é de 443 km e, na direção norte-sul, de 253 km.

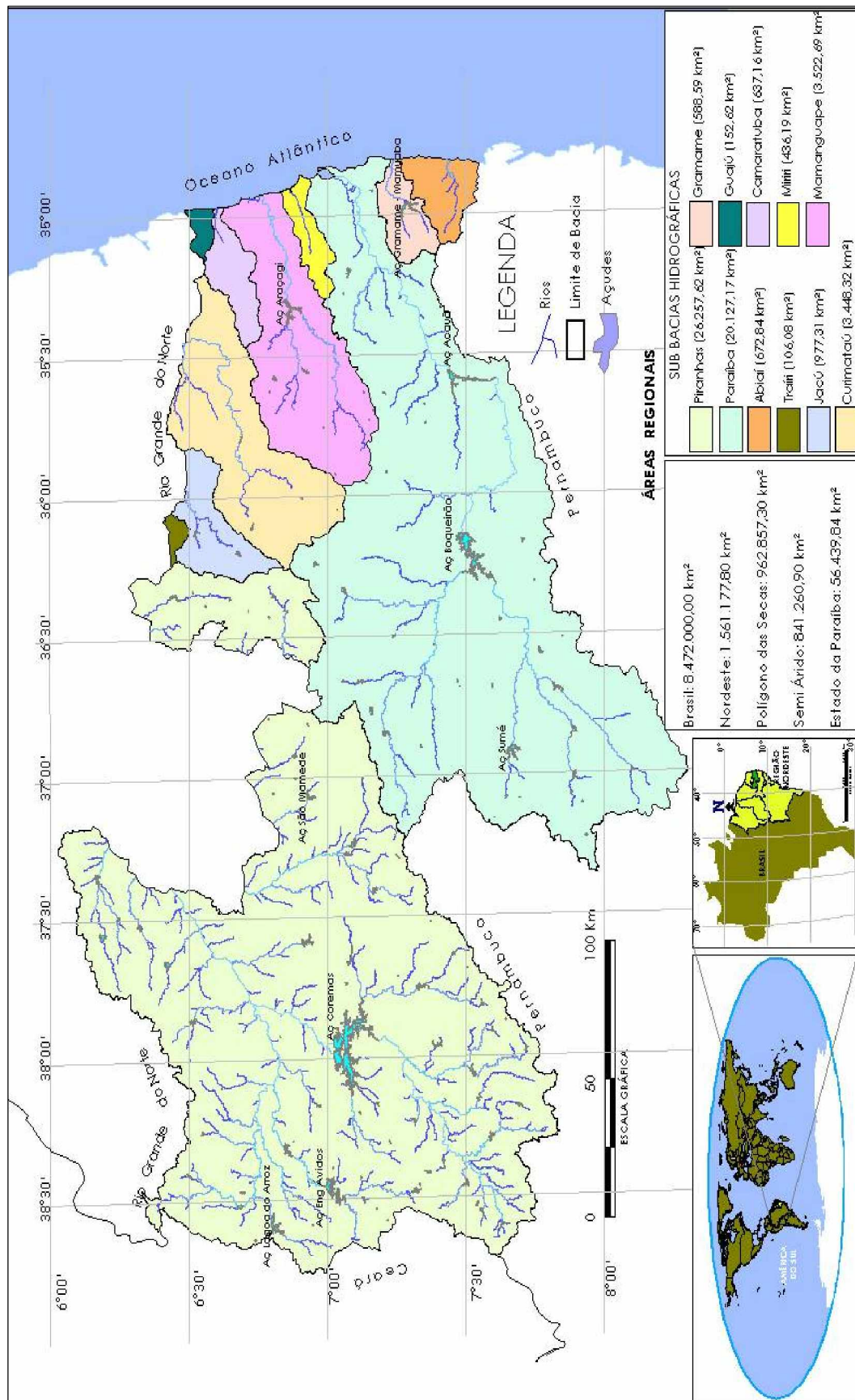


Figura 3.1: Situação e localização da área de estudo

3.1.2 Relevo e altimetria

Destacam-se cinco unidades morfológicas no relevo do Estado da Paraíba (SUDENE, 1972), conforme a Figura 3.2:

a) Baixada Litorânea: terrenos planos constituídos por sedimentos recentes sob as cotas mais baixas da orla marítima e adjacências, com altitudes entre 0 e 40 m;

b) Baixos Platôs Costeiros (Tabuleiros): superfície plana ou suavemente ondulada, acompanhando a direção da faixa costeira, com altitudes entre 40 e 100 m;

c) Depressão Sub-litorânea: – localizada entre os Tabuleiros e rebordo do Planalto da Borborema, caracterizada por áreas baixas com superfície colinosa de relevo suave ondulado, com altitudes até 300 m;

d) Planalto da Borborema: mais característico e elevado acidente da região Nordeste, separa a região litorânea baixa, drenada em direção ao leste, da região ocidental, drenada para o norte, com altitudes entre 500 e 800 m, encontrando-se nessa unidade o ponto culminante do Estado, Pico do Jabre com 1.197 m. O perfil topográfico no sentido leste-oeste, situado no paralelo 7°S, destaca uma verdadeira barreira que constitui a vertente oriental do Planalto da Borborema, sendo a mesma um obstáculo aos ventos dominantes;

e) Depressão Sertaneja: representada por toda a metade oeste do Estado da Paraíba, largamente aberta em direção ao norte pelo vale do Rio Piranhas, destacando-se a superfície de pediplanação com pequenos inselbergues e a bacia sedimentar cretácia, com altitude média entre 200 e 250 m, podendo chegar a 700m nas serras típicas do cristalino;

3.1.3 Geologia

O Estado da Paraíba tem cerca de 89% do seu território assente em rochas pré-cambrianas, e o restante em bacias sedimentares fanerozóicas, rochas vulcânicas cretácias, coberturas plataformais paleógenas/neógenas e formações superficiais quaternárias (Paraíba, 2004). De forma resumida, a geologia da Paraíba, conforme Figura 3.3, pode ser agrupada da seguinte maneira:

a) **Holoceno** – Restingas e Dunas, Aluviões e Mangues;

b) **Terciário** – Grupo Barreiras e Serra dos Martins.

c) **Cretácio** – Formação Itamaracá, Formação Gramame e Série Rio do Peixe.

d) **Pré-Cambriano (B)** – representado na Paraíba por micaxistos, filitos e quartzitos com muscovita e sericita da Série Ceará .

e) **Pré-Cambriano (CD)** – representado na Paraíba por Gnaisses e Migmatitos. São as mais disseminadas no território paraibano.

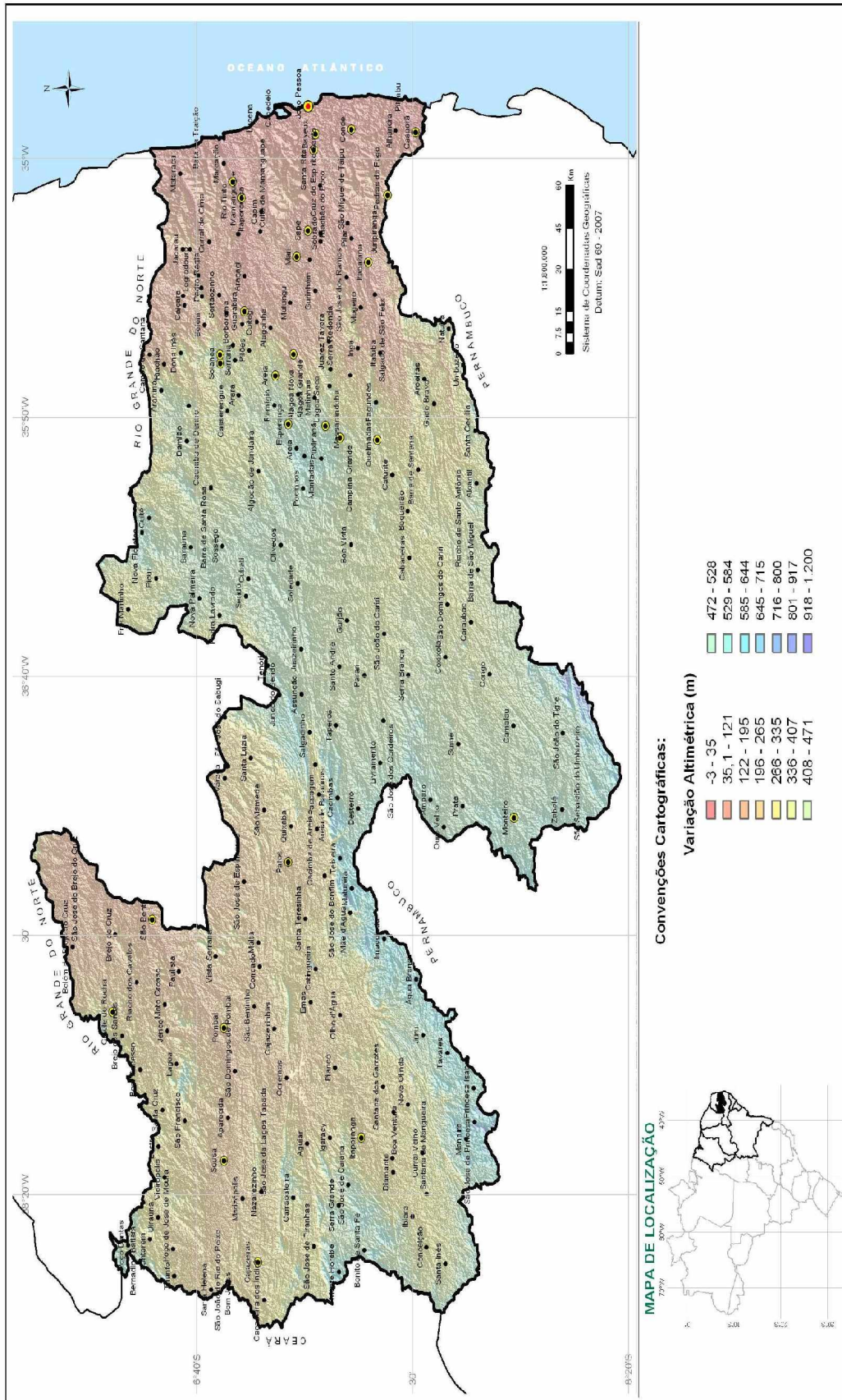


Figura 3.2: Relevo e altimetria do estado da Paraíba

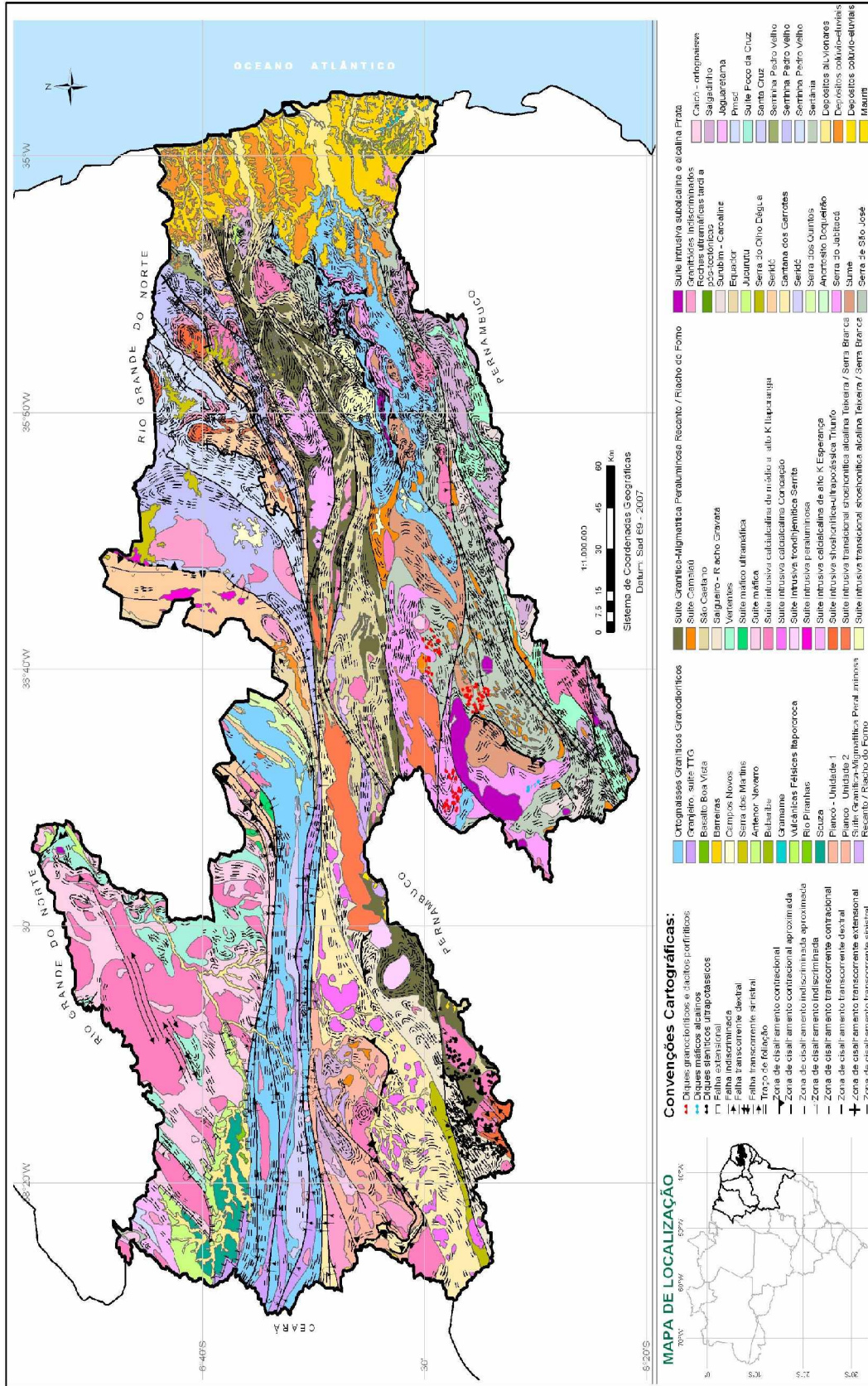


Figura 3.3: Mapa geológico do estado da Paraíba

3.1.4 Solos

Os solos do estado da Paraíba se apresentam conforme o Mapa de Reconhecimento de Solos (Figura 3.4) elaborado no âmbito do Plano Estadual de Recursos Hídricos (Paraíba, 2004) a partir do uso de imagens Landsat (escala 1:100.000), apoio de fotografias aéreas (1:70.000) e levantamentos de campo.

A diversidade de solos existente está diretamente relacionada ao clima regional e ao fato de 89% do seu território está assente em rochas ígneas e metamórficas, caracterizadas por uma grande variedade de minerais.

Aproximadamente 20% da área do Estado apresenta solos evoluídos, coincidindo com o clima quente e úmido do Litoral, de algumas serras, do Brejo e do corredor sub-úmido da Depressão Sub-Litorânea. Enquanto nas restingas e praias, predominam os solos arenosos, nas várzeas e nos mangues, solos aluviais e solos salinos e encharcados, respectivamente.

Na região do Brejo ocorrem solos Litólicos Eutróficos, Podzólicos Equivalentes, Terra Roxa estruturais e, em menor escala, Latossolos Vermelho-Amarelos em decorrência das rochas graníticas e os restos dos capeamentos sedimentares que ali se encontram.

Na Depressão Sub-Litorânea, a grande variedade de rochas cristalinas apresenta solos Podzólicos associados com Planossolos e solos pouco espessos do tipo Bruno-não-Cálcico.

Na porção semi-árida, os solos dominantes são do tipo Litossolos, Planossolos, Podzólicos e Regossolos associados a solos Bruno-não-Cálcicos.

Nas serras úmidas ocorrem os solos mais profundos tipo Regossolos, solos Podzólicos sempre associados a Litólicos e Bruno-não-Cálcicos.

No contexto geral, 80% do território do Estado é composto por solos Bruno-não-Cálcicos, solos Litólicos Eutróficos e solos Podzólicos.

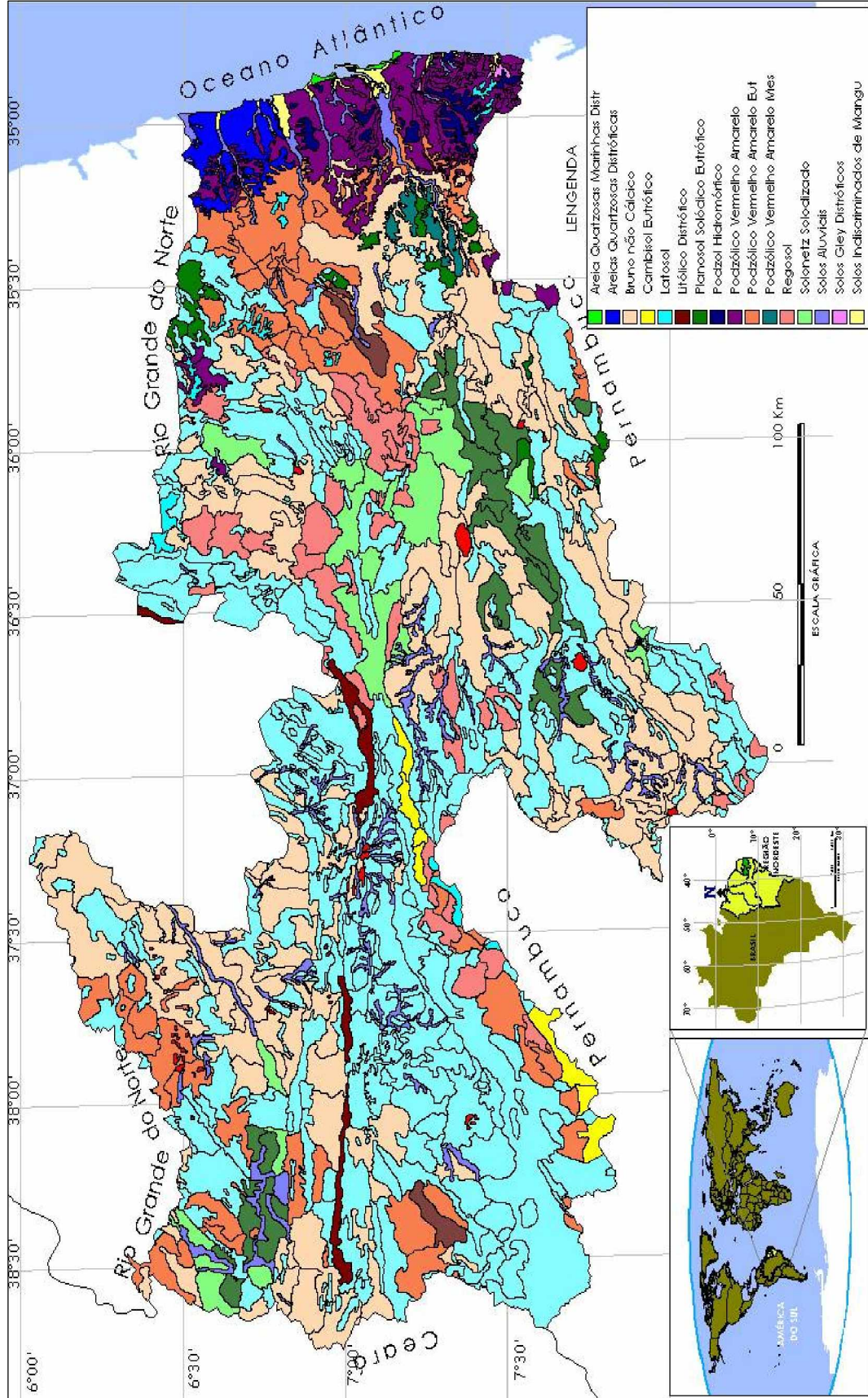


Figura 3.4: Mapa de reconhecimento de solos

3.1.5 Vegetação e uso do solo

Em decorrência da situação geográfica, formas de relevo, natureza das rochas e diversidade dos solos e clima, a vegetação do estado da Paraíba apresenta-se bastante variada, demonstrada pela multiplicidade das paisagens naturais e das formações vegetais distribuídas pelo seu território.

De acordo com o Mapa de Uso Atual e Cobertura Vegetal do Estado da Paraíba elaborado pela Atecel (2002), construído na Escala de 1:100.000, com base no levantamento feito por Lins & Medeiros (1994), domina na Paraíba as formações florestais definidas como Caatinga arbustiva arbórea aberta, Caatinga arbustiva arbórea fechada, Caatinga arbórea fechada, Tabuleiros costeiros, Mangues, Mata úmida, Mata semidecidual, Mata atlântica e Restinga, conforme ilustrado na Figura 3.5.

De acordo com a Tabela 3.1, o antropismo mostra-se bastante acentuado, atingindo aproximadamente 65% da área do Estado.

Tabela 3.1: Distribuição dos usos do solo e cobertura vegetal

Uso atual e cobertura vegetal	Ocupação no Estado	
	(km ²)	(%)
Caatinga arbustiva arbórea aberta	10.375,00	18,38
Caatinga arbustiva arbórea fechada	6.127,00	10,85
Caatinga arbórea fechada	1.312,000	2,32
Antropismo	35.420,000	62,76
Restinga	4,300	0,01
Tabuleiros costeiros	72,200	0,13
Mangues	98,500	0,17
Mata úmida	51,700	0,10
Mata atlântica	196,400	0,35
Área de preservação permanente	1.148,000	2,03
Mata semidecidual	139,000	0,25
Outros	1.495,738	2,65
TOTAL	56.439,838	100,00

FONTE: Paraíba, 2004.

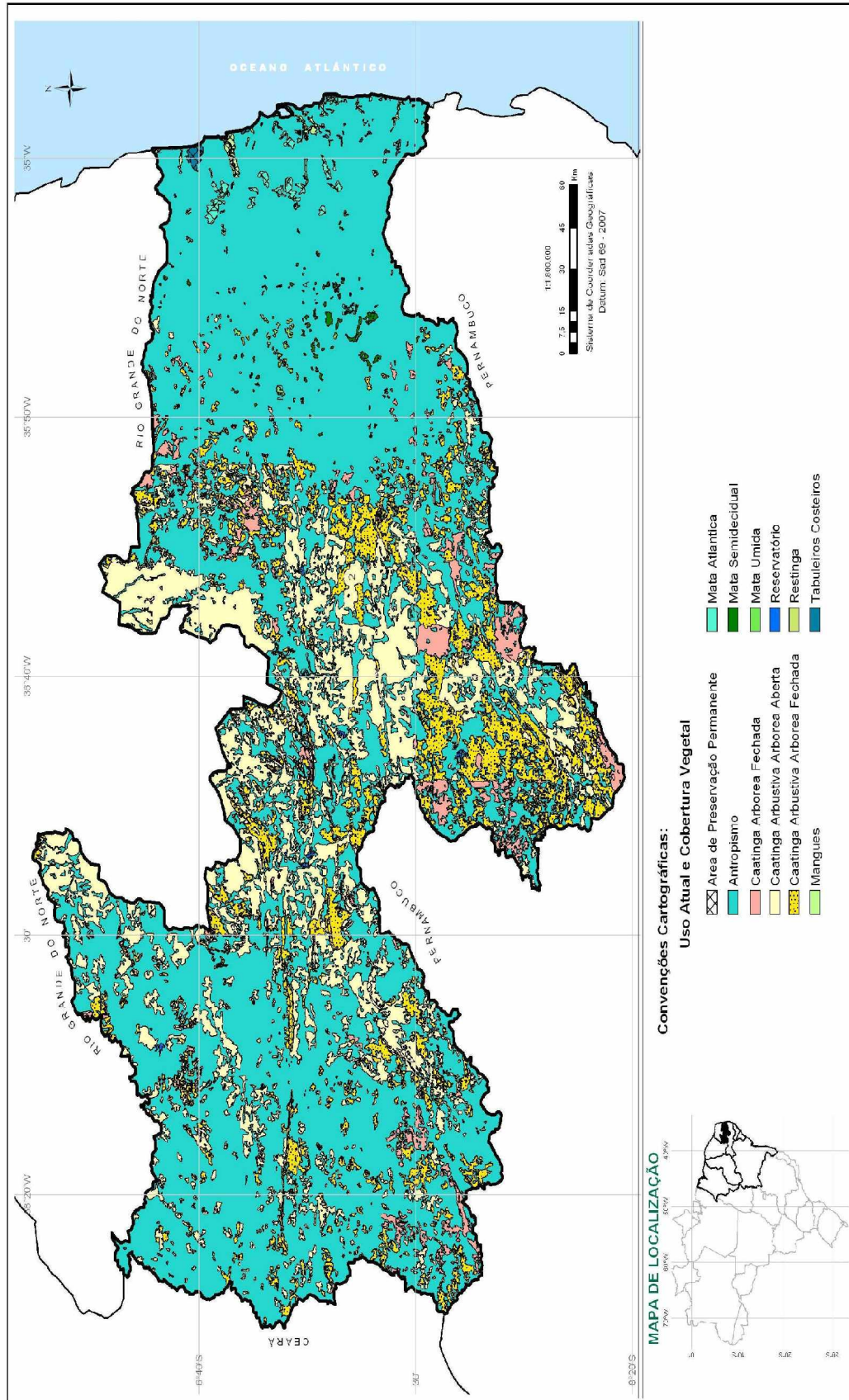


Figura 3.5: Cobertura vegetal

3.1.6 Clima

O estado da Paraíba, dada a sua localização dentro da faixa equatorial, é caracterizado por um clima quente com temperatura média anual de 26°C e pouca variação intra-anual e altos índices de evaporação, fazendo com que a distribuição espacial da temperatura mantenha uma dependência acentuada do relevo, da precipitação pluviométrica, umidade do ar, pressão atmosférica e insolação.

Pode ser classificado em três tipos climáticos segundo a metodologia de Köppen (1948): a) As' - quente e úmido com chuvas de outono-inverno; b) Bsh – semi-árido; c) Aw' – quente e úmido com chuvas de verão-outono, com variações mais localizadas, conforme ilustrado na Figura 3.6.

3.1.7 Hidrografia

A soma das áreas das bacias hidrográficas do Nordeste corresponde a 18% da área total das bacias brasileiras, colocando a Região em terceiro lugar, atrás apenas das Regiões Norte e Centro-Oeste. No entanto, apesar de sua extensão territorial, a hidrografia nordestina é considerada modesta devido ao caráter intermitente e irregular de boa parte de seus rios, cuja condicionante principal é o clima semi-árido, que domina grande parte da Região.

O Conselho Estadual de Recursos Hídricos – CERH dividiu o Estado da Paraíba, através da Resolução nº 02/2003, em onze bacias hidrográficas (Figura 3.1), definidas como: (1) Bacia Hidrográfica do Rio Piranhas, (2) Bacia Hidrográfica do Rio Paraíba, (3) Bacia Hidrográfica do Rio Abiaí, (4) Bacia Hidrográfica do Rio Gramame, (5) Bacia Hidrográfica do Rio Mirirí, (6) Bacia Hidrográfica do Rio Mamanguape, (7) Bacia Hidrográfica do Rio Camaratuba, (8) Bacia Hidrográfica do Rio Guaju, (9) Bacia Hidrográfica do Rio Curimataú, (10) Bacia Hidrográfica do Rio Jacu e (11) Bacia Hidrográfica do Rio Trairi.

As bacias dos rios Piranhas, Jacu, Curimataú, Trairi e Guaju, têm suas nascentes no estado da Paraíba e foz no estado do Rio Grande do Norte, sendo, portanto, de dominialidade federal. A caracterização individual de cada uma das bacias hidrográficas mencionadas é apresentada na Tabela 3.2.

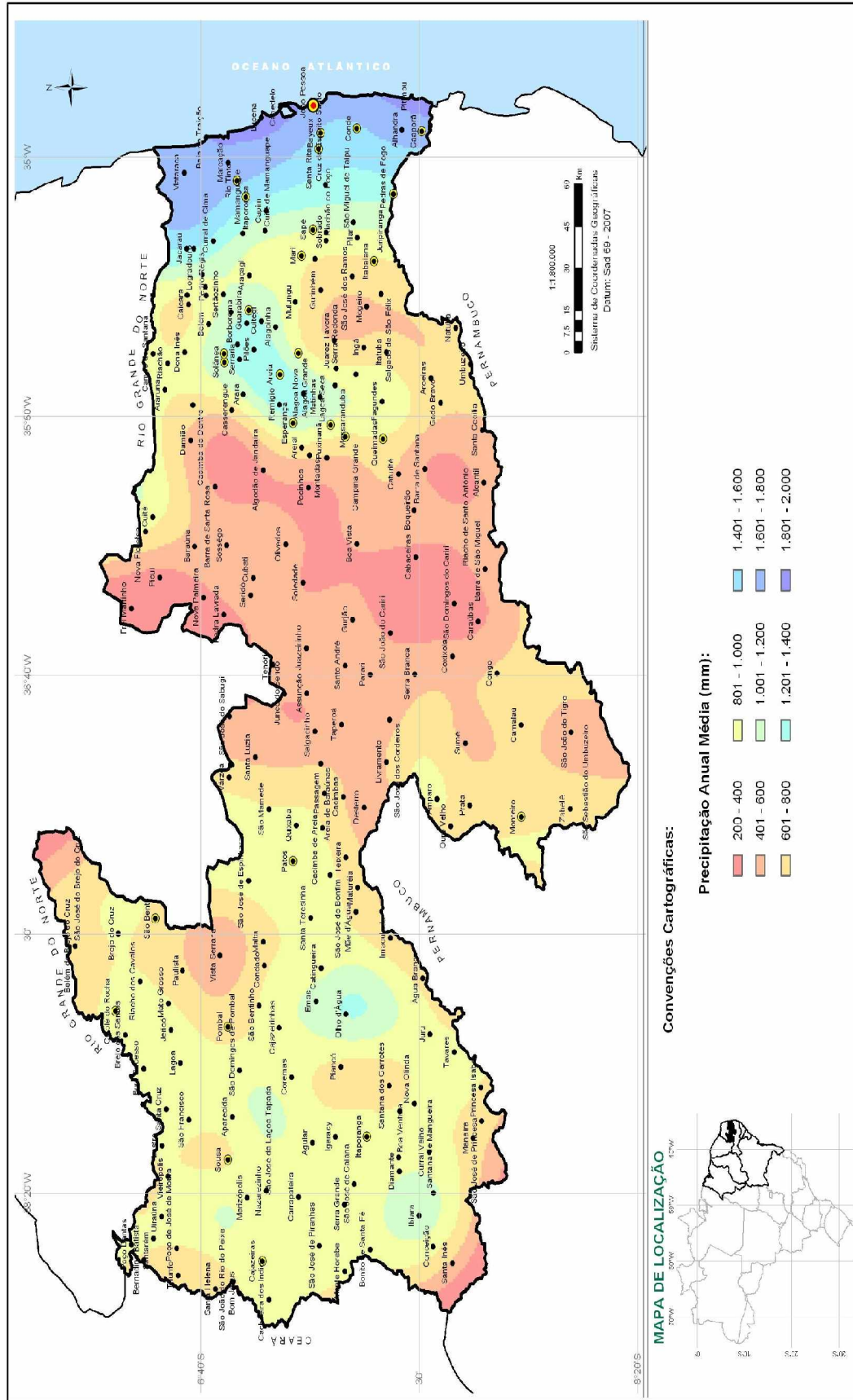


Figura 3.6: Classificação climática de Köppen

Tabela 3.2: Caracterização das bacias hidrográficas do estado da Paraíba

Bacia Hidrográfica	Localização	Área de drenagem (km ²)	Municípios na bacia			População (hab)	Comprimento do rio principal (km)
			sede	total	parcial		
Rio Piranhas	06°02' e 07°50' S 36°13' e 38°46' W	26.257,62	92	66	50	936.848	249,47
Rio Paraíba	06°51' e 08°18' S 34°48' e 37°21' W	20.127,17	71	43	48	1.836.046	358,67
Rio Abiaí	07°15' e 07°33' S 34°48' e 35°04' W	672,84	3	2	3	55.735	32,23
Rio Gramame	07°10' e 07°24' S 34°48' e 35° 11' W	588,59	2	0	6	43.483	50,31
Rio Miriri	06°50' e 07°02' S 34°51' e 35°19' W	436,19	2	0	7	22.450	47,04
Rio Mamanguape	06°38' e 07°11' S 34°55' e 35°58' W	3.522,69	29	16	24	448.678	160,37
Rio Camaratuba	06°33' e 06°46' S 34°58' e 35°28' W	637,16	6	1	10	37.123	69,87
Rio Guajú	06°29' e 06°35' S 34°58' e 35°10' W	152,62	0	0	1	616	21,85
Rio Curimataú	06°25' e 07°04' S 35°12' e 36°16' W	3.448,32	13	5	17	149.470	137,04
Rio Jacu	06°26' e 06°51' S 35°49' e 36°18' W	977,31	4	0	9	38.534	66,89
Rio Trairí	6°24' e 6°30' S 36°03' e 36°14' W	106,08	0	0	2	1.276	16,33

3.1.8 Recursos hídricos

O Estado da Paraíba dispõe de 9.985 açudes, dos quais apenas 192 possuem capacidade superior a 300.000 m³, podendo acumular 4.113.519.138 m³. Os outros 9.793 açudes, com capacidades inferiores a 300.000 m³, podem acumular 1.300.822.539 m³.

A capacidade de regularização desses açudes é bastante reduzida devido as altas taxas de evaporação verificadas no Estado. De acordo com o Plano Estadual de Recursos Hídricos (Paraíba, 2004) a vazão regularizável de todos os açudes da Paraíba com garantia de 100% é de 23,69 m³/s, ou seja, 747.087.840 m³/ano que representa 18,2% da capacidade total dos 192 açudes maiores.

Os rios perenes são de pequena vazão e se concentram de maneira geral na região litorânea do Estado, destacando-se os rios Abiaí, Gramame, Camaratuba e Miri, além de outros de menor porte. Juntos oferecem uma vazão à fio d'água de 3,44 m³/s.

A disponibilidade hídrica referente a água subterrânea é de 4,85 m³/s, proveniente de milhares de poços construídos em aluviões, no cristalino e em áreas sedimentares.

3.2 SELEÇÃO DAS BACIAS

A seleção das bacias que seriam incluídas nos estudos foi feita buscando-se satisfazer o máximo dos critérios seguintes:

- i. A estação fluviométrica não deve ser afetada por desvios ou regularizações a montante; caso contrário, deve essa característica ser considerada no processo.
- ii. O período de dados disponível tenha pelo menos cinco anos disponíveis.
- iii. É preferível que a estação esteja em operação.
- iv. Os dados disponíveis devem ser confiáveis e representativos.
- v. As áreas das bacias devem cobrir uma grande variação de magnitude e tipos diferentes de clima e características da vegetação e solo.

O critério principal para seleção das bacias foi a disponibilidade de dados de vazões diárias ou mensais de boa qualidade. Neste sentido, foram consideradas aptas, além das bacias dos postos fluviométricos, as bacias dos açudes que dispunham de dados de níveis de água e com controle relativo de demandas. Resultou num total de 41 bacias hidrográficas, sendo 14 de estações fluviométricas e 27 de açudes monitorados, conforme Figura 3.7.

As áreas das bacias selecionadas variam de 9,5 km² a 17.220 km² distribuídas conforme a Tabela 3.3, portanto, representativas dos tamanhos de bacias usados em planejamento de recursos hídricos.

Tabela 3.3: Distribuição das bacias de acordo com as áreas de drenagem

Área da bacia hidrográfica (km ²)	Número de bacias
< 100	16
101 – 500	9
501 – 1000	2
1001 – 2000	5
2001 - 5000	6
5001 – 10000	1
> 10000	2

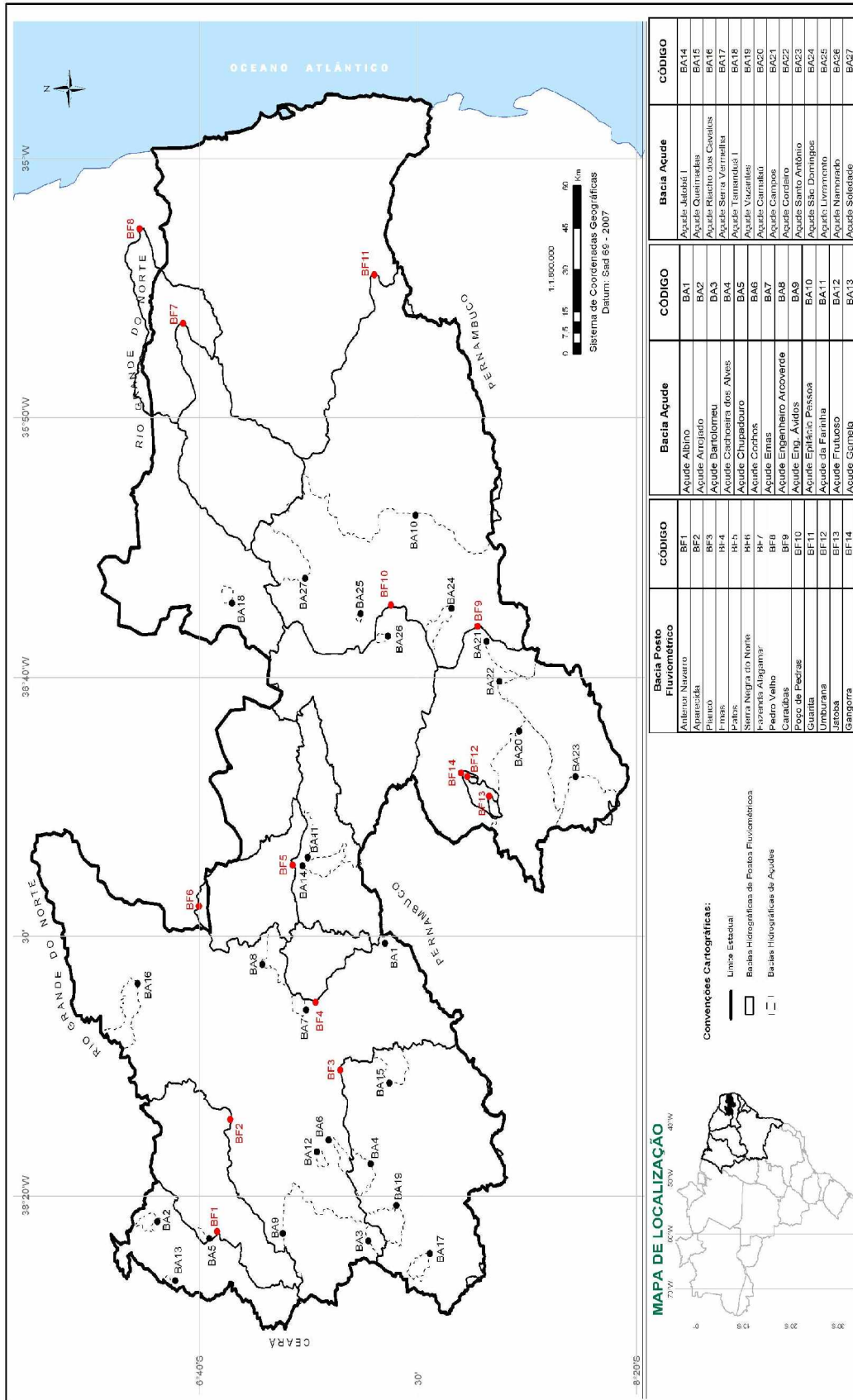


Figura 3.7: Localização das bacias hidrográficas selecionadas

CAPÍTULO IV: METODOLOGIA

4.1 ASPECTOS METODOLÓGICOS

Vários estudos de regionalização procuram estimar características ou estatísticas de vazões, tais como vazões médias mensais ou anuais, curvas de duração de vazões, a partir de características físico-climáticas de bacias hidrográficas, suficientes para certas aplicações em planejamento e gerenciamento de recursos hídricos. Entretanto, em outros problemas é necessário o conhecimento de séries temporais de vazões. Para tais situações o uso de modelos chuva-vazão com parâmetros regionalizados tem se mostrado uma alternativa interessante (Mazvimavi, 2003).

Qualquer estudo de regionalização hidrológica é motivado pela necessidade de obter-se informações inexistentes ou de difícil aquisição, como é o caso dos dados de vazões afluentes, para uma área ou região onde se pretende realizar uma aplicação.

Em geral, parte-se de um problema de engenharia de recursos hídricos (dimensionar uma barragem e seu sangradouro ou determinar sua vazão regularizável, definir a curva de permanência ou a vazão mínima de um rio para orientar um sistema de outorgas), cuja solução depende de informações hidrológicas locais, principalmente dados de vazões (mínimas, médias, máximas ou uma série completa), nem sempre disponíveis em quantidade e qualidade satisfatórias.

Após estabelecer a variável, parâmetro ou função a ser regionalizada, segue-se a etapa de aquisição dos respectivos dados regionais, incluindo as possíveis variáveis explicativas (características físicas e climáticas). Esta atividade se constitui da seleção de bacias hidrográficas com dados disponíveis e da separação das variáveis explicativas mais apropriadas para os estudos.

Posteriormente, procura-se relações regionais, por meio de técnicas matemáticas ou estatísticas, entre as variáveis dependentes e explicativas, de tal forma que estas últimas (de fácil determinação) permitam a obtenção das primeiras em áreas onde estas não existam. É evidente que as relações regionais são circunscritas às condições para as quais foram desenvolvidas.

No caso da regionalização de parâmetros de modelos chuva-vazão, objeto do presente trabalho, após a definição da área de estudo, neste caso as bacias hidrográficas localizadas no estado da Paraíba, cuja descrição encontra-se no capítulo anterior, a metodologia adotada, ilustrada na Figura 4.1, foi a seguinte:

1. Seleção do modelo chuva-vazão: modelo SMAP mensal;
2. Aquisição dos dados (chuva, vazão e evapotranspiração mensais) para calibração do modelo SMAP;
3. Calibração do modelo SMAP: algoritmo de calibração automática SCE-UA;
4. Obtenção das características físicas e climáticas para os estudos de regionalização;
5. Análise das bacias no conjunto dos dados e seleção das características mais importantes;
6. Identificação de relações regionais entre parâmetros do modelo SMAP e características físicas e climáticas das bacias usando redes neurais artificiais;
7. Definição das condições de aplicabilidade das relações regionais identificadas no item anterior.

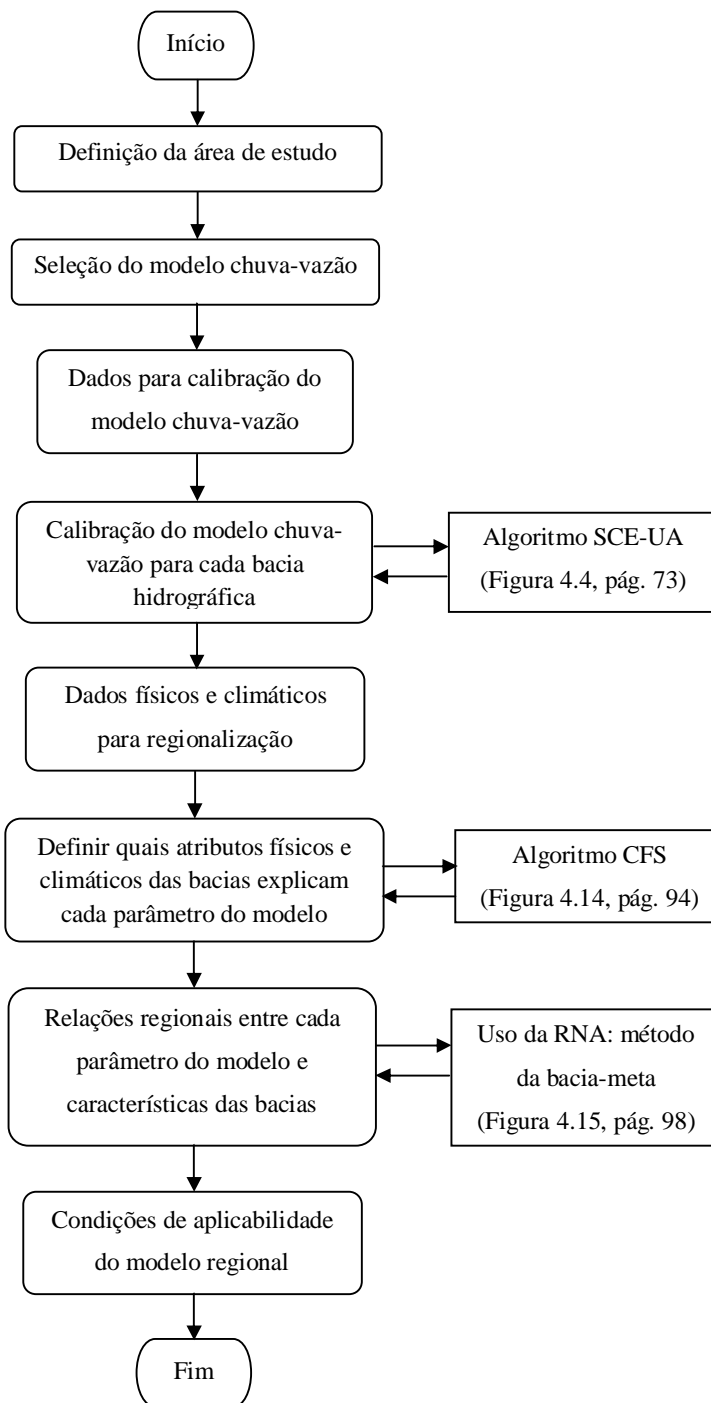


Figura 4.1: **Etapas do processo de regionalização de parâmetros do modelo chuva-vazão**

4.2 SELEÇÃO DO MODELO CHUVA-VAZÃO

Os modelos chuva-vazão podem ser, de forma geral, agrupados em duas categorias (Refsgaard & Knudsen, 1996): 1) modelos concentrados, cuja representação espacial dos processos hidrológicos se dá por meio de valores médios calculados na secção de referência da bacia hidrográfica (Wood, 1995); 2) modelos distribuídos, os quais utilizam equações contínuas que descrevem a variação temporal e espacial dos processos hidrológicos (Beven and O'Connell, 1982).

O sucesso na estimativa dos valores dos parâmetros de modelos distribuídos depende da disponibilidade de dados nas escalas temporal e espacial compatíveis com a estrutura do modelo adotado. Refsgaard and Knudsen (1996) compararam a habilidade de três modelos, um concentrado (NAM – *Nedbor-Afstromnings Model*), um semi-distribuído (WATBAL- *Water Balance*) e outro de base física (MIKE SHE – *Système Hydrologique Européen*) na modelagem de vazões em Zimbabwe, resultando num melhor desempenho do modelo NAM.

A grande maioria das aplicações de modelos chuva-vazão ocorre na escala mensal. O preenchimento de falhas em séries de vazões, os estudos de operação de reservatórios e de vazões regularizáveis com aplicações diretas na agricultura e em sistemas de abastecimento de água. Os modelos diários podem suprir tais necessidades, mas também requerem muito mais detalhes na sua estrutura, uma que os dados de calibração nem sempre são disponíveis.

Existem vários modelos concentrados e distribuídos em uso no Brasil e, para o presente trabalho, selecionou-se um modelo concentrado obedecendo a quatro critérios principais: a) aplicabilidade do modelo para a região de estudo; b) o menor número possível de parâmetros a calibrar para diminuir as incertezas dos seus valores; c) intervalo de cálculo compatível com os dados disponíveis na região; d) apropriado para comprovar as técnicas de regionalização propostas.

Entre os modelos estudados destacou-se o modelo concentrado SMAP - *Soil Moisture Accounting Procedure* em sua versão mensal, o qual, de acordo com Lopes et al (1981), possui uma estrutura relativamente simples e opera com dados de chuva e evapotranspiração potencial em intervalos mensais.

4.2.1 O Modelo SMAP

Sendo o modelo SMAP bastante conhecido pela comunidade técnica e já aplicado a diversas bacias hidrográficas da região Nordeste do Brasil, o mesmo foi adotado para desenvolvimento e comprovação das técnicas aqui propostas.

O modelo realiza o balanço da umidade do solo baseado em dois reservatórios lineares que representam o solo (camada superior) e o aquífero, conforme ilustrado na Figura 4.2.

A cada evento de precipitação P_m é feito um balanço de massa. Uma fração da precipitação é transformada em escoamento superficial ES em função da chuva, da taxa de umidade do solo, TU, e de um parâmetro controlador do escoamento superficial, E2.

A lâmina restante, P_m-ES , é adicionada a um reservatório que representa a camada superior do solo. Neste reservatório, a umidade é atualizada ao longo do tempo através de perdas por evapotranspiração real, função da evapotranspiração potencial, da taxa de umidade do solo e de um parâmetro calibrável do modelo (E1).

A outra saída desse reservatório é a recarga para o reservatório subterrâneo (aquífero). Nesta transferência de água é usado o conceito de capacidade de campo, isto é, se o nível do reservatório, RSOLO, for maior que a capacidade de saturação, STR, ocorrerá recarga, REC, para o reservatório subterrâneo, função da taxa de umidade do solo, do coeficiente de recarga, CREC, e do coeficiente de infiltração, CINF. Quando a umidade está próxima à capacidade de campo a recarga é nula.

No reservatório subterrâneo o nível de água existente, RSUB, é diminuído à taxa constante, KK, resultando o escoamento básico, EB. A soma dos escoamentos direto, ES, (superficial) e básico, EB, fornece a vazão no ponto de controle da bacia.

Em decorrência das características da área de estudo, predominantemente de solos rasos e assente sobre rochas impermeáveis, os resultados das calibrações do modelo SMAP revelaram a quase inexistência de recargas para o reservatório subterrâneo. O parâmetro CREC que governa estas transferências apresentou-se nulo em quase todas as bacias hidrográficas. Tal situação, identificada também por Alexandre et al (2005), comprometeu o significado do parâmetro de recessão KK, uma vez que sem recargas não haveria escoamentos subterrâneos.

Desta forma, os parâmetros CREC e KK foram eliminados do modelo e o processo de calibração se deu apenas sobre os quatro parâmetros que definem a dinâmica do reservatório superficial (STR, CINF, E1 e E2). Isto significa que a estrutura do modelo ficou reduzida a apenas um reservatório, o que é suficiente para manter a memória do balanço hídrico da bacia e é compatível não apenas com as características da região, mas também com o fato do intervalo de cálculo do modelo ser o mês e, nesse caso, as contribuições das bacias para o escoamento subterrâneo são de pouca relevância.

Com apenas um reservatório, o escoamento foi dividido em duas parcelas: a primeira, chamada de escoamento direto, ocorre por transbordamento do reservatório quando seu o nível de água (RSOL) superar a capacidade de armazenamento (STR); a segunda, formada a partir do

esvaziamento do reservatório, em forma de escoamento superficial e sub-superficial, determinada de maneira semelhante a ES descrita anteriormente para a estrutura original do modelo.

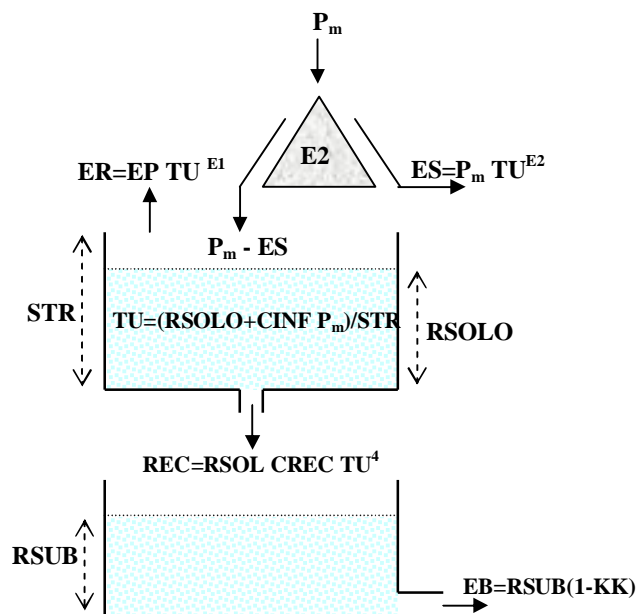


Figura 4.2: **Desenho esquemático do modelo SMAP**

Os dois processos hidrológicos mais importantes da região semi-árida do nordeste brasileiro são, indiscutivelmente, as precipitações e a evaporação, particularmente devido à grande variabilidade temporal e espacial do primeiro e as elevadas e constantes taxas do segundo. Assim, a modelação das equações que definem o escoamento superficial e as perdas por evaporação determinam os resultados finais do modelo.

De forma geral, os quatro parâmetros do modelo SMAP a serem calibrados podem ser associados a características ou comportamento da bacia hidrográfica. O parâmetro STR , nível máximo do reservatório superficial, representa a capacidade de saturação do solo e pode ser influenciado por fatores relacionados ao tipo e a profundidade do solo, capacidade de retenção, taxa de infiltração, espelhos de água na bacia e rede de drenagem. Tem forte correlação com o parâmetro $E2$, sendo este controlador do escoamento superficial e ligado a aspectos que tendem a gerar e facilitar o escoamento como a área, declividade, forma da bacia e sinuosidade dos cursos de água, à precipitação e evaporação.

O parâmetro $E1$ responsável pelo controle das taxas de evapotranspiração na bacia pode ser associado a informações como a área, a forma da bacia, a elevação da bacia e à própria rede de drenagem, além, é claro, das taxas potenciais de evapotranspiração. Mudanças em seus

valores podem ser compensadas por alterações no parâmetro CINF, o qual representa a parcela da chuva que não alimenta o reservatório superficial e pode ser afetado pelo índice de aridez e declividade da bacia, entre outros.

4.3 DADOS PARA CALIBRAÇÃO DO MODELO SMAP

4.3.1 Fluviometria

As bacias que compõem o estado da Paraíba estão inseridas nas sub-bacias 37 e 38 do Atlântico Sul, segundo código específico estabelecido pelo DNAEE – Departamento Nacional de Águas e Energia Elétrica, em consonância com a Portaria do Ministério das Minas e Energia, a partir do Decreto Nº 77.410, de 12 de abril de 1976.

Os dados fluviométricos usados neste trabalho são provenientes da rede fluviométrica do extinto DNAEE, atualmente sob a jurisdição da ANEEL – Agência Nacional de Energia Elétrica. Essa rede de coleta de dados fluviométricos foi, sucessivamente, operada pelos escritórios regionais do antigo DNAEE, por equipes do DNOCS – Departamento Nacional de Obras contra as Secas, em convênio com o DNAEE durante a década dos anos 70, e, posteriormente, pelos escritórios regionais da CPRM – Companhia de Pesquisa de Recursos Minerais.

A Agência Nacional de Águas – ANA disponibiliza pela Internet o Sistema HIDROWEB de Informações Hidrológicas, a partir do qual pôde-se identificar a existência de 72 postos fluviométricos no estado da Paraíba. Entretanto, dada uma grande variedade de problemas, selecionou-se apenas 14 postos com informações confiáveis (Tabelas 4.1 e 4.2), aptos a serem utilizados nos estudos hidrológicos aqui almejados, todos sob a responsabilidade da ANA, que os opera através da Companhia de Pesquisa de Recursos Minerais – CPRM por meio de convenio.

Os principais problemas verificados nos 72 postos fluviométricos do Estado da Paraíba foram identificados e sistematizados pela TC/BR - Riverside Technology Inc. (2001), resumidos da seguinte maneira:

- ausência total de dados limnimétricos e fluviométricos. Existem vários postos cadastrados onde o número de observações limnimétricas e fluviométricas é nulo ou insignificante;
- disponibilidade de séries de vazões diárias, sem sua respectiva série de alturas limnimétricas. Este problema, bastante usual, é consequência do processo de informatização deficiente dos dados, ocorrido na década de 70. Neste processo, séries completas de dados brutos foram extraviadas, após seu tratamento; isto levou a

dispor dos dados tratados sem a capacidade de verificar sua consistência devido à completa ausência de dados brutos;

- disponibilidade de curvas-chave sem os correspondentes dados de medição de descarga; este problema, que tem causas semelhantes ao anterior, impede também a análise de consistência dos dados tratados;
- ausência total, ou número insuficiente, de medições de descarga para a elaboração de uma curva-chave confiável;
- curvas-chave muito desatualizadas, isto é, mesmo dispondo de um número aceitável de medições de descarga, estas se concentram num período muito restrito de tempo, não sendo possível extrapolar seus resultados para todo o período em que se dispõe de alturas limnimétricas. Por exemplo, existem vários postos em que, mesmo dispondo de alturas limnimétricas a partir da década de 30, as medições de descarga se restringem ao período posterior a 1974. Nesse ano ocorreu a maior cheia já registrada e sabe-se que a grande maioria das seções sofreram mudanças consideráveis, não sendo válidas, fora deste período, as curvas-chave confeccionadas com medições de descarga anteriores a essa data.
- postos instalados a jusante de reservatórios de porte médio ou grande e funcionando após o início da operação daqueles, sendo portanto todos os registros influenciados pelo reservatório, como é o caso da bacia do Alto Piranhas onde tem-se praticamente todas as séries fluviométricas medidas a jusante dos reservatórios de São Gonçalo e Engenheiro Avidos.

Tabela 4.1: Postos fluviométricos selecionados para os estudos

Código	Sub-Bacia	Nome do posto	Altitude (m)	Área (km ²)	Latitude			Longitude		
					grau	min	seg	grau	min	seg
37260000	37	Antenor Navarro	240	1.580	-6	44	7	-38	26	53
37290000	37	Aparecida	170	3.720	-6	47	7	-38	5	13
37340000	37	Piancó	249	4.710	-7	12	16	-37	55	41
37360000	37	Emas	270	530	-7	6	38	-37	42	39
37430000	37	Patos	250	1.850	-7	1	22	-37	16	12
37440000	37	Serra Negra do Norte	160	3.330	-6	39	56	-37	24	5
38650000	38	Fazenda Alagamar	-	2.270	-6	36	19	-35	31	44
38680000	38	Pedro Velho	30	3.590	-6	26	42	-35	13	14
38830000	38	Caraúbas	493	5.120	-7	43	42	-36	30	9
38850000	38	Poço de Pedras	430	3.140	-7	23	53	-36	26	1
38880000	38	Guarita	50	17.220	-7	20	4	-35	22	22
*	38	Umburana/Sumé	590	10,40	-7	41	22	-36	59	9
*	38	Jatobá/Sumé	650	26,80	-7	46	25	-37	2	49
*	38	Gangorra/Sumé	590	137,40	-7	39	50	-36	58	20

*não fazem parte do banco de dados da ANA.

Tabela 4.2: **Localização e período com dados disponíveis dos Postos fluviométricos**

Código	Nome do posto	BACIA	Período com dados disponíveis	Chuva média (mm)	Vazão média (m ³ /s)
37260000	Antenor Navarro	Piranhas	nov/1963 a dez/1972; 1984 a dez/2001	888,2	2,330
37290000	Aparecida		jan/1985 a dez/2001	898,3	10,210
37340000	Piarcó		jan/1963 a dez/2001	877,0	11,710
37360000	Emas		nov/1963 a dez/1972; mai/1984 a dez/2001	1.100,0	1,945
37430000	Patos		jan/1985 a dez/2001	793,1	3,832
37440000	Serra Negra do Norte		mai/1921 a dez/1960; jan/1983 a dez/2001	804,5	5,688
38650000	Fazenda Alagamar	Curimataú	jan/1986 a dez/2002	602,4	1,129
38680000	Pedro Velho		jan/1975 a dez/2002	686,6	3,556
38830000	Caraúbas	Paraíba	mar/1973 a dez/2002	614,7	4,892
38850000	Poço de Pedras		jan/1970 a dez/2000	503,7	3,446
38880000	Guarita		jan/1970 a dez/2002	575,8	14,543
*	Umburana/Sumé		Out/1975 a set/1980	551,4	0,022
*	Jatobá/Sumé		Out/1974 a set/1980	681,8	0,055
*	Gangorra/Sumé		Out/1974 a set/1980	595,9	0,180

*não fazem parte do banco de dados da ANA.

Observa-se que os dados disponíveis no Estado são bastante reduzidos, consistindo-se em um fator restritivo para a realização de trabalhos de regionalização hidrológica. Diante dessa situação buscou-se aumentar o tamanho da amostra gerando series de vazões a partir dos níveis dos 133 açudes monitorados no Estado.

Entre esses açudes foram selecionados 27 em função da qualidade dos dados disponíveis distribuídos em todo o território paraibano. A Tabela 4.3 relaciona os açudes por bacia do Estado e algumas informações características.

Considerando-se a existência dos dados de níveis dos açudes no início de cada mês e as demais variáveis que compõem o balanço hídrico, foi obtida a calibração do modelo mensal pela aproximação dos volumes mensais acumulados nos açudes observados aos obtidos com o balanço hídrico. Os seguintes fatos simplificadores foram considerados:

- a) comportas de fundo dos açudes fechadas, e, portanto nula a vazão de regularização do rio;
- b) vazões retiradas do lago para irrigação, estimadas através da definição de um padrão de consumo anual, única variável desconhecida nos períodos de afluência nula, característica dos rios da região em cerca de sete a oito meses do ano;
- c) lâminas médias mensais de evaporação determinadas conforme Hargreaves & Sumani (1985);
- d) dados pluviométricos médios mensais referentes ao mesmo período dos dados de níveis;

e) demandas totais dos sistemas de abastecimento de água fornecidas pela Companhia de Água e Esgotos da Paraíba - CAGEPA.

Conhecida a variação do volume armazenado, a evaporação no espelho d'água e as demandas para o abastecimento humano, os volumes armazenados em cada açude foram determinados através da equação do balanço hídrico.

$$S_i = S_{i-1} - \text{Dem} - \text{Evap} - \text{Irr} + V_{\text{afllu}} \quad 4.1$$

Sendo: S_i = volume armazenado no açude no mês i .

S_{i-1} = volume armazenado no açude no mês anterior.

Dem = demandas mensais para o abastecimento humano.

Evap = volume mensal evaporado no espelho da água.

Irr = demandas totais para irrigação.

V_{afllu} = volume afluente ao açude (estimado pelo modelo SMAP mensal).

Tabela 4.3: Açudes selecionados para geração de vazões

Bacia Hidrográfica	Açude	Capacidade de acumulação (m ³)	Área de drenagem (km ²)	Período de observação*
Rio Piranhas	Albino	1.333.955,	9,5	2000 a 2003
	Arrojado	3.569.180,	30,1	1996 a 2002
	Bartolomeu I	17.570.556,	59,5	1996 a 2003
	Cachoeira dos Alves	10.611.198,	110,5	1996 a 2003
	Chupadouro	2.764.100,	17,8	2000 a 2003
	Cochos	4.200.000,	56,5	1996 a 2003
	Emas	2.013.750,	35,0	2000 a 2003
	Eng ^o Arcoverde	30.593.400,	126,9	1996 a 2003
	Eng ^o Avidos	255.000.000,	1.009,5	1995 a 2003
	Epitácio Pessoa	411.686.287,	10.659,0	1998 a 2004
	Farinha	25.738.500,	747,9	1995 a 2003
	Frutuoso II	3.517.220,	19,1	1995 a 2003
	Gamela	472.296,	12,8	1996 a 2003
	Jatobá I	17.516.000,	94,0	1995 a 2003
	Queimadas	15.625.338,	124,3	1995 a 2003
	Riacho dos Cavalos	17.699.000,	161,4	1997 a 2003
	Serra Vermelha	11.801.173,	55,7	1996 a 2003
Tamanduá I	1.261.125,	23,4	2000 a 2003	
Vazantes	9.091.200,	137,0	1995 a 2003	
Rio Paraíba	Camalau	46.437.520,	1.054,0	1995 a 2003
	Campos	6.594.392,	181,2	1995 a 2003
	Cordeiro	69.965.945,	1.665,8	1995 a 2003
	Santo Antonio	24.424.130,	340,6	1996 a 2003
	São Domingos	7.340.440,	65,5	2001 a 2003
	Livramento	2.432.420,	37,0	2000 a 2002
	Namorados	2.118.980,	14,2	2001 a 2003
Soledade	27.058.000,	313,1	1996 a 2003	

* Período comum entre os dados de chuva e níveis dos açudes

4.3.2 Pluviometria

A pluviometria do estado da Paraíba é caracterizada pela grande variabilidade temporal e espacial, destacando-se regiões onde as chuvas médias anuais são superiores a 1.500 mm, no litoral, e outras onde a média não passa de 250 mm, como é o caso do município de Cabaceiras no Cariri Paraibano, conforme mostrado na Figura 4.3.

A precipitação média é geralmente avaliada a partir de dados pontuais de postos existentes ao redor e no interior da própria bacia usando-se, entre outros, o método gráfico dos Polígonos de Thiessen, definido pela expressão:

$$P_m = \frac{\sum_{i=1}^n P_i a_i}{A} \quad 4.2$$

onde:

P_m - precipitação média;

a_i - área de influência do posto i ;

P_i - precipitação no posto i ;

A - área da bacia hidrográfica (km²);

n - número de postos que exerce influência na bacia.

A precipitação média diária foi calculada para cada bacia hidrográfica pelo Método de Thiessen. O ano hidrológico é usado neste estudo, começando em outubro e findando em setembro do ano seguinte.

Os períodos de determinação das chuvas médias nas bacias hidrográficas foram definidos em função da disponibilidade de dados de níveis de açudes e de dados de vazões fluviométricas, resultando para as bacias hidrográficas dos açudes nove anos hidrológicos (outubro/1994 a setembro/2003), e para bacias hidrográficas dos postos fluviométricos variou entre os anos de outubro/1963 a setembro/2002. Os dados foram obtidos da AESA – Agência Executiva de Gestão das Águas do Estado da Paraíba, da Empresa de Pesquisa Agropecuária do Rio Grande do Norte - EMPARN, da Fundação Cearense de Meteorologia e Recursos Hídricos - FUNCEME e do Instituto Tecnológico do Estado de Pernambuco - ITEP. As Tabelas 4.4 e 4.5 relacionam os postos pluviométricos utilizados para as bacias hidrográficas dos postos fluviométricos e dos açudes, respectivamente.

4.3.3 Evapotranspiração

Dado a inexistência de Estações evaporimétricas com abrangência espacial e temporal adequadas no estado da Paraíba, foram usados os dados de evapotranspiração potencial (ETP) determinados conforme Hargreaves & Samani (1985). Para cada bacia adotou-se as lâminas médias mensais referentes às médias calculadas entre os municípios localizados dentro e nas imediações da bacia, conforme resumido na Tabela 4.6.

A Tabela 4.7 apresenta os valores anuais de precipitação e evapotranspiração potencial médios para cada bacia hidrográfica.

Tabela 4.4: Postos pluviométricos selecionados para as bacias hidrográficas dos postos fluviométricos

Bacia Hidrográfica	Posto fluviométrico	Postos pluviométricos selecionados
Rio Piranhas	Antenor Navarro	Cajazeiras, Antenor Navarro, São Vicente, Uiraúna, Luiz Gomes, Icozinho.
	Aparecida	Aparecida, Cajazeiras, Jericó, Nazarezinho, Antenor Navarro, São Vicente, São Gonçalo, Uiraúna, Luiz Gomes, Alexandria, Icozinho.
	Piancó	Bonito de Santa Fé, Aguiar, Serra Grande, Ibiara, Coremas, Santana dos Garrotes, Conceição, Princesa Isabel, Jurú, Serra Talhada ¹ , Quixabinha ³ , Poço ³
	Emas	Mãe d'Água, Malta, Condado, Catingueira, Santa Terezinha, Imaculada.
	Patos	Imaculada, Santa Terezinha, Mãe d'Água, Salgadinho, São Mamede, Desterro, Teixeira, Passagem.
	Serra Negra do Norte	Imaculada, Santa Terezinha, Mãe d'Água, Malta, Serra Negra, Salgadinho, São Mamede, São João do Sabugi, Desterro, Teixeira, Passagem.
Rio Paraíba	Caraúbas	Coxixola, São José dos Cordeiros, Caraúbas, Congo, Prata, Santa Maria da Paraíba, São João do Tigre, São Sebastião do Umbuzeiro, Serra Branca, Bananeiras, Sertânia, Poção, Cimbres, Jabitacá.
	Poço de Pedras	Coxixola, Desterro, Gurjão, Juazeirinho, Passagem, Salgadinho, São João do Cariri, São José dos Cordeiro, Serra Branca, Bananeiras, Taperoá, Teixeira, Equador.
	Guarita	Juazeirinho, Salgadinho, Seridó, Soledade, Sossego, Taperoá, Desterro, Olivedos, Teixeira, Equador, Coxixola, Passagem, São João do Cariri, São José dos Cordeiros, Faz. Santa Tereza, Gurjão, Barra de São Miguel, Cabaceiras, Camalaú, Caraúbas, Congo, Prata, Santa Maria da Paraíba, São João do Tigre, São Sebastião do Umbuzeiro, Serra Branca, Bananeiras, Sertânia, Bodocongó, Poção, Cimbres, Jabitacá, Fagundes, Pocinhos, Mata Virgem, Vertentes, Taquaritinga, Alagoa Grande, Areia, Aroeiras, Ingá, Macaparana, Machados, Bom Jardim, Timbaúba.
	Umburana/Sumé	Postos 28 a 36 da BRS
	Jatobá/Sumé	Postos 32 a 48 da BRS
	Gangorra/Sumé	Postos 1 a 48 da BRS
Rio Curimataú	Fazenda Alagamar	Barra de Santa Rosa, Sossego, Olivedos, Algodão, Pocinhos, Araruna, Cacimba de Dentro, Salgado, Caiçara, Bananeira.
	Pedro Velho	Barra de Santa Rosa, Cuité, Sossego, Olivedos, Pocinhos, Araçagi, Areia.

¹ postos do Estado de Pernambuco ² postos do Estado do Rio Grande do Norte ³ postos do Estado do Ceará

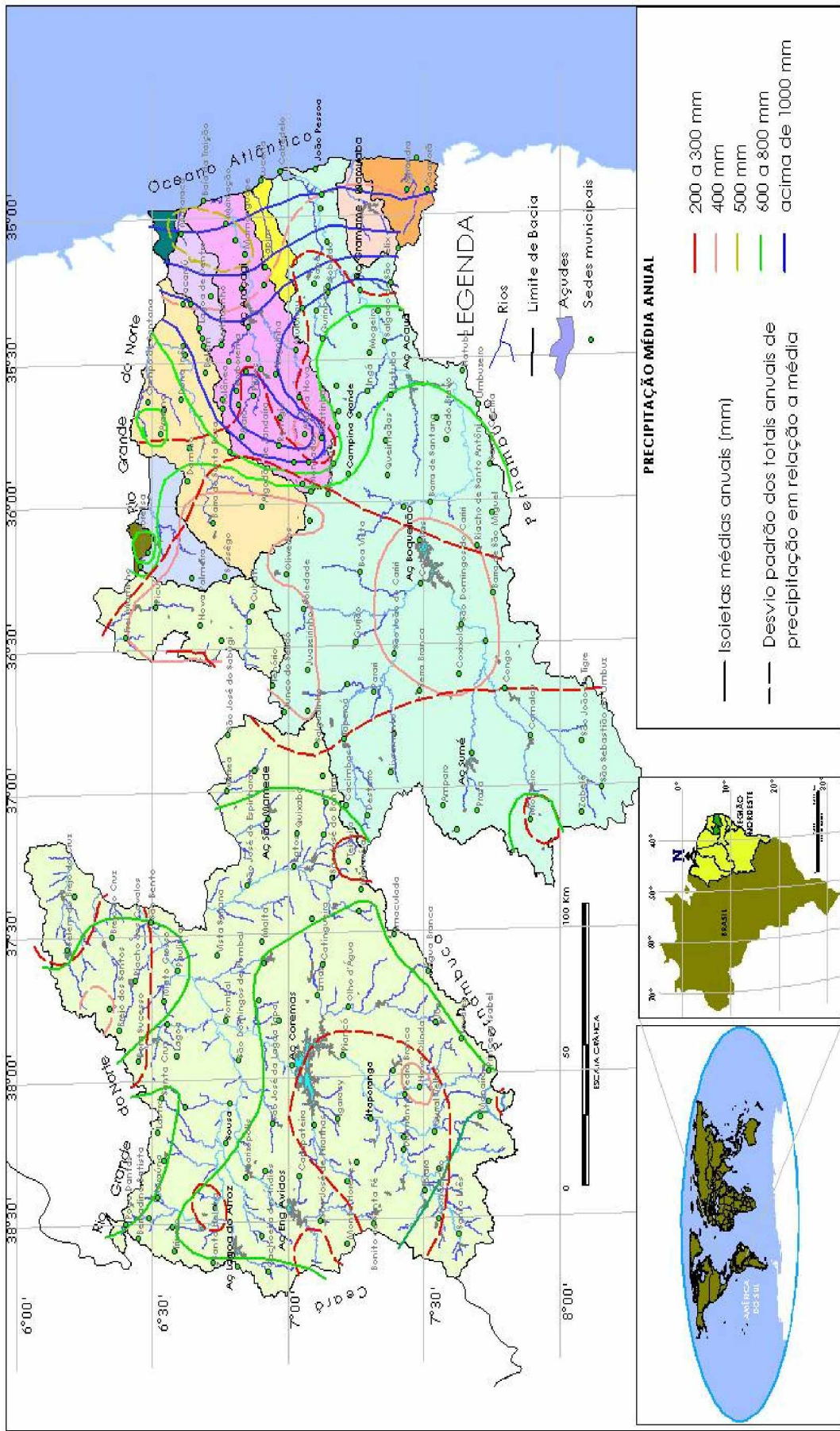


Figura 4.3: Precipitação média anual no estado da Paraíba

Tabela 4.5: Postos pluviométricos selecionados para as bacias hidrográficas dos açudes

Bacia Hidrográfica	Açude	Postos pluviométricos selecionados
Rio Piranhas	Albino	Santa Terezinha
	Arrojado	Uiraúna
	Bartolomeu I	Bonito de Sta Fé
	Cachoeira dos Alves	Itaporanga, São José de Caiana
	Chupadouro	Antenor Navarro
	Cochos	Igaracy, Itaporanga, São José de Caiana
	Emas	Emas
	Eng ^o Arcoverde	Condado, Malta
	Eng ^o Ávidos	Bonito de Santa Fé, Cajazeiras, Carrapateira, Monte Horebe, Sítio Arapuá-São José de Piranhas, São José de Piranhas, Serra Grande.
	Epitácio Pessoa	Boa Vista, Boqueirão, Coxixola, Juazeirinho, Pocinhos, Prata, Salgadinho, Soledade, Sossego, Taperoá, Desterro, Olivados, São João do Cariri, São José dos Cordeiros, Gurjão, Barra de São Miguel, Cabaceiras, Camalaú, Congo, Caraúbas, Santa Maria da Paraíba, São João do Tigre, São Sebastião do Umbuzeiro, Bananeiras, Sumé, Serra Branca.
	Farinha	Cacimba de Areia, Passagem, Salgadinho, Teixeira.
	Frutuoso II	Aguiar
	Gamela	Triunfo
	Jatobá I	Patos-EMBRAPA, São José do Bonfim, Teixeira
	Queimadas	Juru, Piancó, Santana dos Garrotes
	Riacho dos Cavalos	Brejo dos Santos, Catolé do Rocha, Riacho dos Cavalos.
	Serra Vermelha	Conceição
Tamanduá I	Cubati, Poço de Pedras, Pedra Lavrada.	
Vazantes	Bonito de Santa Fé, Ibiara.	
Rio Paraíba	Camalau	Sertânia II ¹ , Camalaú, Monteiro-Embrapa, Monteiro-Pilões, São João do Tigre, São Sebastião do Umbuzeiro.
	Campos	Jataúba ¹ , Caraúbas, Congo
	Cordeiro	Sertânia II ¹ , Camalaú, Monteiro-Pilões, São João do Tigre, São Sebastião do Umbuzeiro, Santa Maria da Paraíba-São João do Tigre, Poção, Jataúba ¹ , Congo.
	Santo Antonio	Sertânia II ¹ , São Sebastião do Umbuzeiro.
	São Domingos	Coxixola, São Domingos do Cariri.
	Livramento	Gurjão
	Namorados	São João do Cariri
Soledade	Olivados, Soledade, São Vicente do Seridó	

¹postos do Estado de Pernambuco ²postos do Estado do Rio Grande do Norte ³postos do Estado do Ceará

Tabela 4.6: Localidades com evapotranspiração representativa

Bacia Hidrográfica	Açude	Localidades*
Rio Piranhas	Albino	Imaculada
	Arrojado	Antenor Navarro
	Bartolomeu I	Bonito de Santa Fé
	Cachoeira dos Alves	Itaporanga
	Chupadouro	Antenor Navarro
	Cochos	Aguiar
	Emas	Catingueira
	Eng ^o Arcoverde	Condado
	Eng ^o Ávidos	São José de Piranhas, Eng ^o Ávidos, Bonito de Santa Fé
	Epitácio Pessoa	Boqueirão, Cabaceiras, Barra de São Miguel
	Farinha	Patos
	Frutuoso II	Aguiar
	Gamela	Antenor Navarro
	Jatobá I	Patos
	Queimadas	Nova Olinda
	Riacho dos Cavalos	Catolé do Rocha
	Serra Vermelha	Conceição
Tamanduá I	Pedra Lavrada	
Vazantes	Conceição	
Rio Paraíba	Camalau	Monteiro
	Campos	Sumé
	Cordeiro	Sumé
	Santo Antonio	São João do Tigre
	São Domingos	Cabaceiras
	Livramento	Desterro
	Namorados	São João do Cariri
Soledade	Soledade	
Bacia Hidrográfica	Posto fluviométrico	Localidades*
Rio Piranhas	Antenor Navarro	Antenor Navarro, Bom Jesus, Eng ^o Ávidos, São José de Piranhas.
	Aparecida	Antenor Navarro, Bom Jesus, Eng ^o Ávidos, São José de Piranhas, Nazarezinho, Sousa, Pombal.
	Piancó	Itaporanga, Piancó, Nova Olinda Princesa Izabel, Manaíra, Conceição, Bonito de Santa Fé.
	Emas	Catingueira, Mãe d'Água, Olho d'Água.
	Patos	Patos, Teixeira, Salgadinho, Imaculada, Mãe d'Água.
Serra Negra do Norte	Patos, Teixeira, Salgadinho, Imaculada, Mãe d'Água, Condado.	
Rio Paraíba	Caraúbas	Sumé, Monteiro, Cabaceiras.
	Poço de Pedras	São João do Cariri, Soledade, Desterro, Taperoá, Boa Vista, Cabaceiras.
	Guarita	São João do Cariri, Soledade, Desterro, Taperoá, Boa Vista, Cabaceiras, Bodocongó, Umbuzeiro, Ingá, Itabaiana.
	Umburana/Sumé	Sumé
	Jatobá/Sumé	Sumé
Gangorra/Sumé	Sumé	
Rio Curimataú	Fazenda Alagamar	Olivedos, Barra de Santa Rosa, Araruna, Bananeiras.
	Pedro Velho	Olivedos, Barra de Santa Rosa, Araruna, Bananeiras, Belém.

* nos casos com mais de uma localidade, foram usadas as médias aritméticas das lâminas destas localidades.

Tabela 4.7: Valores médios anuais de precipitação e evapotranspiração

Ordem	Bacia Hidrográfica	Precipitação	Evapotranspiração
		média anual (mm)	potencial média anual (mm)
1	Albino	623,9	1644,4
2	Arrojado	712,4	1931,5
3	Bartolomeu I	678,6	1752,7
4	Cachoeira dos Alves	911,5	1931,5
5	Chupadouro	946,4	1931,5
6	Cochos	969,5	1875,5
7	Emas	687,6	1925,5
8	Eng ^o Arcoverde	715,0	1903,8
9	Eng ^o Ávidos	852,6	1853,5
10	Epitácio Pessoa	450,0	1968,0
11	Farinha	545,9	1922,1
12	Frutuoso II	921,5	1875,5
13	Gamela	771,1	1931,5
14	Jatobá I	730,0	1922,1
15	Queimadas	672,5	1834,4
16	Riacho dos Cavalos	670,4	1897,8
17	Serra Vermelha	678,6	1931,5
18	Tamanduá I	449,4	1805,5
19	Vazantes	737,1	1931,5
20	Camalau	481,1	1715,5
21	Campos	357,8	1715,5
22	Cordeiro	445,3	1715,5
23	Santo Antonio	453,7	1667,9
24	São Domingos	375,4	1616,7
25	Livramento	507,8	1805,5
26	Namorados	466,6	1913,9
27	Soledade	343,8	1610,1
28	Antenor Navarro	561,1	1909,6
29	Aparecida	794,9	1912,4
30	Piancó	863,0	1830,8
31	Emas	1324,1	1932,2
32	Patos	969,5	1778,1
33	Serra Negra do Norte	843,3	1799,1
34	Fazenda Alagamar	545,2	1580,6
35	Pedro Velho	769,7	1645,3
36	Caraúbas	601,0	1682,6
37	Poço de Pedras	602,2	1726,2
38	Guarita	578,4	1652,1
39	Umburana/Sumé	551,4	1900,8
40	Jatobá/Sumé	681,8	1900,8
41	Gangorra/Sumé	595,9	1900,8

4.4 CALIBRAÇÃO DO MODELO

4.4.1 Considerações gerais

A fase de calibração contribui significativamente para a precisão dos modelos chuva-vazão. Essa precisão reflete-se diretamente no decurso da regionalização.

A calibração de um modelo chuva-vazão pode ser realizada através de três processos: tentativa-erro ou manual, automática e semi-automática.

A calibração manual é, em geral, extremamente trabalhosa. Ela exige um grande conhecimento do modelo e da região hidrológica em estudo, o que torna sua aplicação impraticável para modelos mais complexos, onde existe um grande número de parâmetros a serem estimados. A calibração automática, além de eliminar este problema, introduz conforto e menor subjetividade no processo. A atual realidade na área da informática e a necessidade de melhoria nos processos de calibração de modelos matemáticos em hidrologia têm propiciado um grande crescimento na aplicação dos métodos automáticos e semi-automáticos (Diniz et al, 1996).

Partindo-se do princípio de que a estrutura do modelo está definida e que os limites de variação dos parâmetros são conhecidos (definindo a incerteza inicial nos parâmetros), o processo clássico de calibração consiste em encontrar os melhores valores dos parâmetros para os quais a diferença entre as vazões observadas e simuladas seja mínima, ou melhor, o mais próximo de zero possível. Entretanto, não há uma maneira que não seja ambígua de se medir essa diferença (erros), a não ser pela comumente chamada função objetivo (Sorooshian et al, 1998). Quando se verifica na série de erros, padrões de estacionaridade e independência, a função objetivo mais popular é a soma dos quadrados dos erros médios.

Nesta matéria, um fato importante que tem se verificado ao longo dos últimos trinta anos de pesquisas é que sempre existe uma função objetivo particular mais adequada para calibrar um dado modelo do que outra [Chapman (1999), Diskin & Simon (1977), Servat & Dezetter (1991), Yapo et al (1996)], resultando em sérios problemas na fase de escolha da função objetivo, conforme discutido por Sorooshian & Dracup (1980).

Mesmo sendo a calibração dependente da função objetivo, da estrutura do modelo e dos dados da calibração, na maioria das pesquisas os erros decorrentes da estrutura do modelo ou são considerados desprezíveis ou absorvidos nos erros residuais (Sorooshian et al, 1998) e os erros contidos nos dados são geralmente tratados apenas através da análise estatística dos erros, esperando-se terem o mesmo comportamento estatístico dos dados de entrada.

Os métodos tradicionais de calibração de modelos chuva-vazão assumem que o objetivo primário é obter um melhor ajuste aos dados de vazões em cada estação fluviométrica,

obrigando a função objetivo a focar sobre os erros de cada estação individualmente. Nesta visão, cada estação é tratada independentemente, mesmo se a meta seja obter um modelo hidrológico regional.

Foi desenvolvido o método de calibração individual no qual os valores dos parâmetros são estimados em cada estação fluviométrica produzindo o melhor ajuste aos dados de vazões observados.

Os parâmetros do modelo SMAP foram calibrados utilizando o algoritmo de calibração automática SCE-UA (Shuffled Complex Evolution – University of Arizona), tendo como função objetivo o coeficiente de eficiência estabelecido por Nash & Sutcliffe (1970), conforme a Equação 4.3. No âmbito deste trabalho, outras funções objetivo, já consagradas na literatura, foram testadas e descartadas uma vez que levaram o modelo SMAP a resultados menos precisos (valores de correlação entre vazões observadas e simuladas abaixo daqueles obtidos com o uso da Equação 4.3), a exemplo do somatório das diferenças quadradas entre vazões observadas e simuladas.

$$R2 = 1 - \frac{\sum(Q_o(t) - Q_c(t))^2}{\sum(Q_o(t) - \bar{Q}_o)^2} \quad 4.3$$

sendo: $Q_o(t)$ a vazão mensal observada no mês t e $Q_c(t)$ a vazão calculada pelo SMAP no tempo t e Q_o a vazão média mensal observada da série disponível. Quanto mais próximo da unidade mais bem ajustadas estão as vazões observadas e calculadas. Segundo Collischonn (2001), o modelo pode ser considerado apropriadamente calibrado para valores de $R2$ acima de 0,75 e, aceitável, se o valor de $R2$ ficar entre 0,36 e 0,75.

Melhores resultados no processo de calibração foram obtidos substituindo-se os valores originais de vazões observadas e simuladas pela raiz quadrada de seus valores, talvez por permitir que os erros do modelo possam ser admitidos normalmente distribuídos com igual variância, ou seja, $Q_o^{1/2} = Q_s^{1/2} + \varepsilon$, onde $\varepsilon \approx N(0, \sigma^2)$.

4.4.2 O Algoritmo de calibração SCE-UA

O algoritmo SCE-UA realiza a pesquisa do ótimo global baseado nos mecanismos de seleção natural, combinando um processo de seleção artificial com operadores genéticos extraídos da natureza.

Neste algoritmo, uma população de "s" pontos é selecionada aleatoriamente no espaço de pesquisa dos parâmetros. Esta população é dividida em algumas comunidades (complexos ou

conjuntos) que evoluem independentemente uma das outras, ou seja, o espaço de pesquisa dos parâmetros é explorado em diferentes direções.

Após algumas gerações, as comunidades são envolvidas entre si, dando origem a novas comunidades através de um processo de mistura. Cada membro da comunidade (ponto selecionado) é um pai em potencial com habilidades para participar no processo de reprodução.

Um subcomplexo ou subconjunto extraído aleatoriamente de um complexo representa um casal reprodutor, sendo que o subcomplexo pode consistir em mais de dois membros. Para assegurar que o processo de evolução seja competitivo, é dada maior chance aos melhores pais (pontos de menor valor da função objetivo) para participar na geração de descendentes. A distribuição triangular de probabilidades garante essa competitividade. O método de busca direta de Nelder & Mead (1965) é aplicado a cada subcomplexo para produzir melhores descendentes. Esta estratégia utiliza as informações pertinentes aos subcomplexos para aperfeiçoamento da evolução.

Ocasionalmente, descendentes são substituídos por outros, gerados aleatoriamente no espaço de pesquisa dos parâmetros, para garantir que a evolução não se perca fora do espaço possível de busca. Essa mudança ocasional é uma analogia à mutação na teoria da genética. A mutação aumenta a quantidade de informações contidas na amostra. Os novos descendentes, assim formados, substituem os piores pontos do subcomplexo corrente. Isto garante que cada ponto tenha pelo menos uma chance de contribuir para o processo de reprodução antes que seja substituído ou descartado. Sendo assim, nenhuma informação contida na amostra é ignorada.

Para executar o algoritmo SCE-UA (Diniz, 1994), ilustrado na Figura 4.4, deve-se:

1. Definir o número de complexos $p \geq 1$ e $m \geq n+1$, onde n é o número de parâmetros. Calcular o tamanho da amostra $s=p \times m$. O valor de $m=2n+1$ foi utilizado nesse trabalho;
2. Selecionar s pontos (x_1, \dots, x_s) no espaço de pesquisa dos parâmetros $U \subset R^N$. Calcular o valor da função f_i em cada ponto x_i . Para essa seleção pode ser usada a distribuição uniforme;
3. Ordenar os s pontos em ordem crescente do valor da função objetivo f_i . Guardá-los em um conjunto $D=\{x_i, f_i, i=1, \dots, s\}$, de tal forma que $i=1$ represente o melhor ponto;
4. Dividir D em p complexos A^1, \dots, A^p , cada um contendo m pontos, de forma que $A^k=\{x_j^k, f_j^k / x_j^k=x_{k+p(j-1)}, f_j^k=f_{k+p(j-1)}, j=1, \dots, m\}$;
5. Evoluir cada complexo $A^k, k=1, \dots, p$, de acordo com o algoritmo de evolução competitiva esboçado separadamente;
6. Misturar os complexos da seguinte forma: Substitua A^1, \dots, A^p em D , de tal forma que $D=\{A^k, k=1, \dots, p\}$. Arrange D em ordem crescente do valor da função;

7. Verificar o critério de convergência. Se satisfeito, parar. Caso contrário, retornar para o passo 4.

No algoritmo de evolução competitiva requerido no passo 5, ilustrado na Figura 4.5, deve-se:

1. Selecionar q , α e β , de forma que $2 \leq q \leq m$, $\alpha \geq 1$ e $\beta \geq 1$. Os valores de $q=n+1$, $\alpha=7-P$ e $\beta=2n+1$ foram utilizados nesse trabalho;

2. Atribuir uma distribuição triangular de probabilidade acumulada para A^k , do tipo, $P_i = P_{i-1} + 2(m+1-i)/m(m+1)$, $i=1, \dots, m$, onde $P(0)=0$;

3. Selecionar os pais, escolhendo-se aleatoriamente q distintos pontos u_1, \dots, u_q de A^k de acordo com a distribuição especificada acima, definindo assim um "subcomplexo". Guardá-los em uma matriz $B = \{u_i, v_i, i=1, \dots, q\}$, onde v_i é o valor da função associado com o ponto u_i . Guardar em L as localizações de A^k que foram usadas para construir B ;

4. Gerar descendentes de acordo com o seguinte processo:

a) ordenar B e L de modo que os q pontos fiquem arranjados em ordem crescente do valor da função. Calcule o centróide g usando a expressão abaixo:

$$g = (1/q) \sum_{j=1}^{q-1} u_j$$

b) calcular um novo ponto $r = 2g - u_q$ (reflexão);

c) se r estiver dentro do espaço U , calcular o valor da função f_r e seguir para o passo d; caso contrário, calcular o menor hipercubo $H \subset \mathbb{R}^N$ que contem A^k , selecionar aleatoriamente um ponto z dentro de H , calcular f_z , fazer $r=z$ e $f_r=f_z$ (mutação);

d) se $f_r < f_1$, calcular $ex = 2r - g$ (expansão). Caso contrário, seguir para o passo g;

e) se ex estiver dentro do espaço U , seguir para o passo f. Caso contrário, selecionar aleatoriamente um ponto z dentro de H e fazer $ex=z$;

f) calcular a função f_{ex} . Se $f_{ex} < f_r$ fazer $u_q = ex$, $f_q = f_{ex}$ e seguir para o passo j; caso contrário, seguir para o passo h;

g) Se $f_r < f_q$, seguir para o passo h. Caso contrário, calcular $c = (g + u_q)/2$ e f_c . Seguir para o passo i;

h) Fazer $u_q = r$, $f_q = f_r$ e seguir para o passo j;

i) Se $f_c < f_q$, fazer $u_q = c$, $f_q = f_c$ e seguir para o passo j. Caso contrário, selecionar aleatoriamente um ponto z em H , calcular f_z , fazer $u_q = z$ e $f_q = f_z$ (mutação);

j) Repetir os passos "a" a "i" α vezes, sendo que $\alpha \geq 1$ é um parâmetro a ser especificado;

5. Substituir os pais pelos descendentes da seguinte forma: Substituir B em A^k usando as localizações originais guardadas em L. Organizar A^k em ordem crescente do valor da função;

6. Repetir os passos 2 a 5 β vezes, sendo que $\beta \geq 1$ é um parâmetro especificado que determina como cada complexo será evoluído.

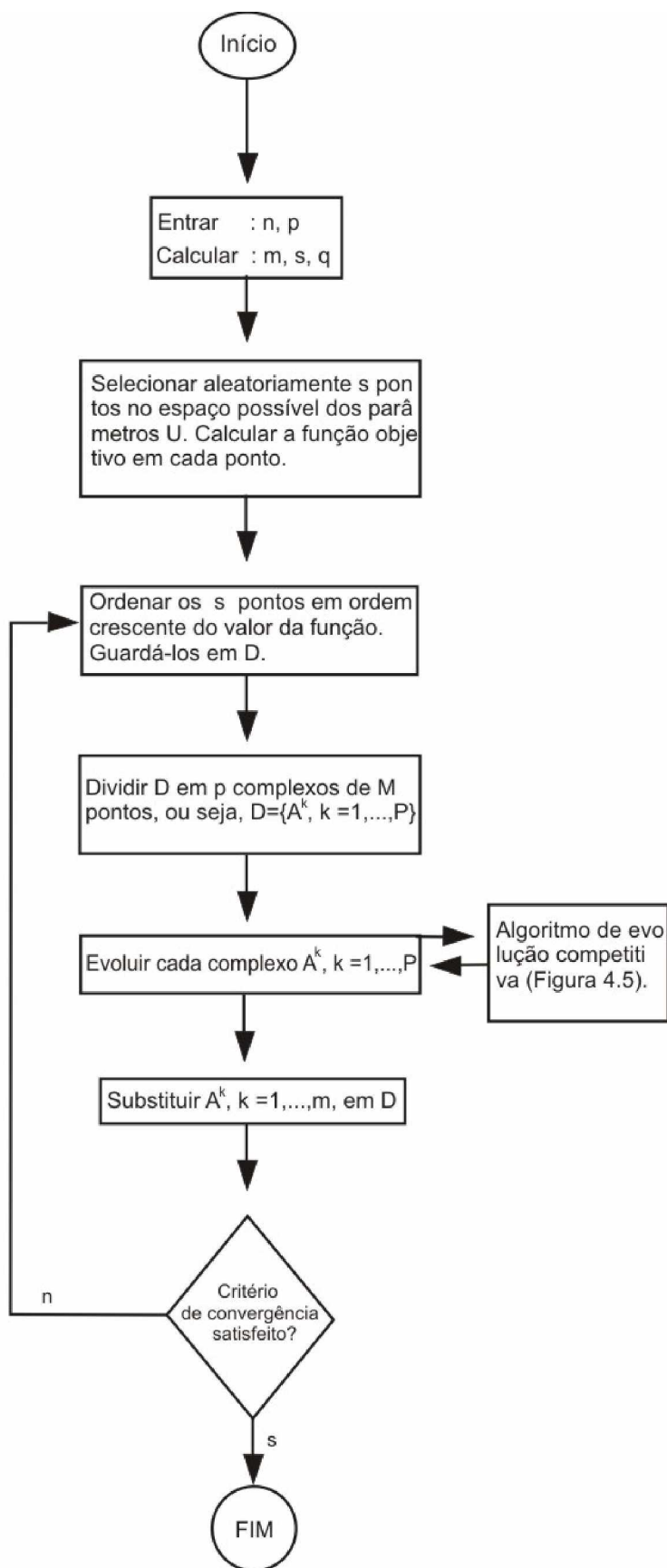


Figura 4.4: Fluxograma do Algoritmo SCE-UA (Duan et al, 1992)

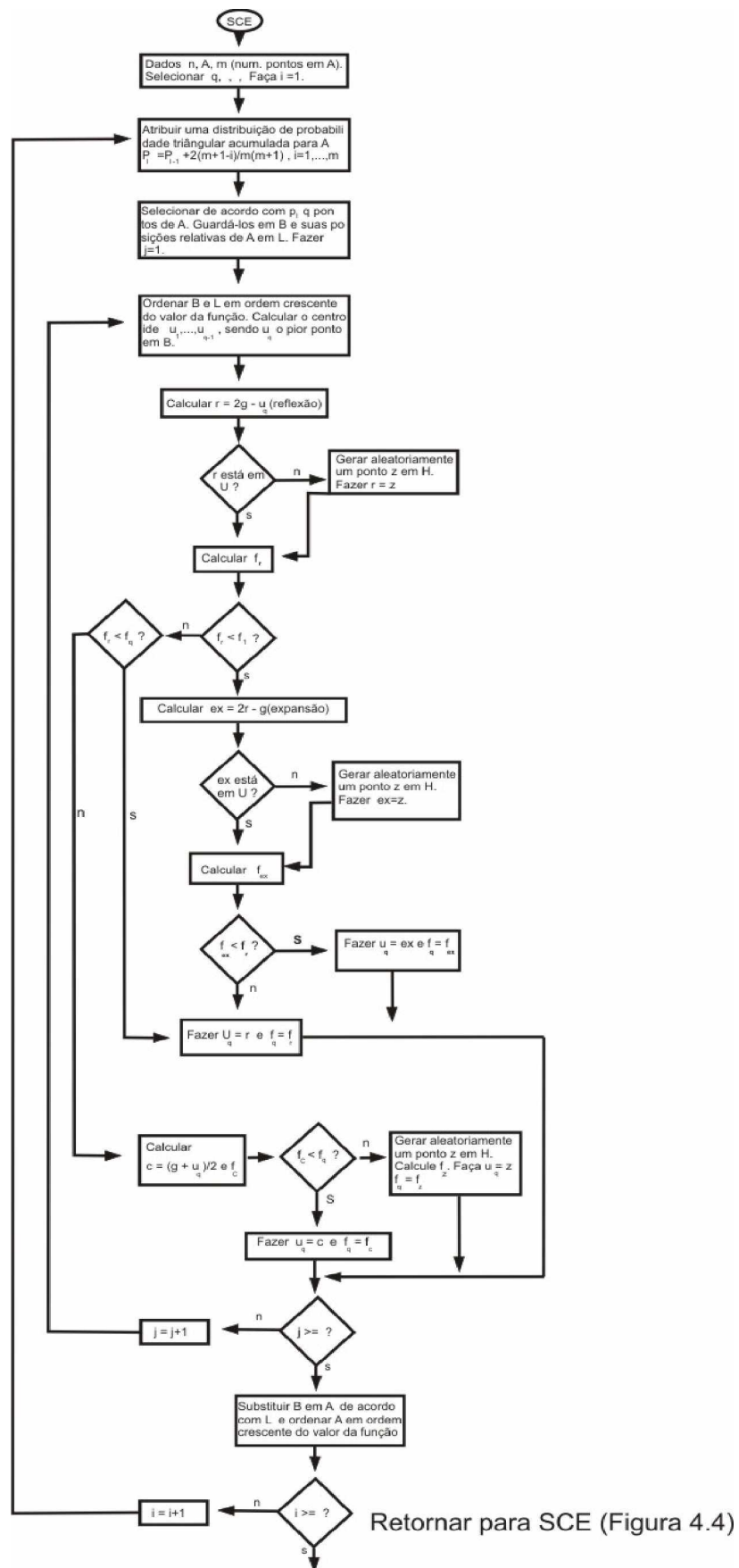


Figura 4.5: Algoritmo de evolução competitiva

4.5 OBTENÇÃO DAS CARACTERÍSTICAS DAS BACIAS

4.5.1 Parâmetros morfológicos

Para determinação dos parâmetros morfológicos foi necessário realizar duas etapas precedentes:

a. Delimitação das bacias – utilizando o mapa digitalizado e atualizado da hidrografia do Estado (Paraíba, 2004), foi possível delimitar as bacias dos 27 açudes. Nos casos das bacias dos açudes que continham outros açudes a montante, estas não foram delimitadas como bacias incrementais, exceto apenas para a bacia do açude Acauã que tem o açude Epitácio Pessoa (Boqueirão) a montante, o qual será considerado nos outros estudos do Açude Acauã como uma vazão afluente.

b. Geração de curvas de níveis - as curvas de níveis advindas das cartas da SUDENE digitalizadas não tinham o atributo “Z” (altitude) associado às mesmas, como também as “polylines” que as representavam não eram contínuas, apresentando quebras para linhas de mesma altitude. BAENA et al. (2003) destaca que os modelos digitais de elevação hidrologicamente consistentes (MDEHC’s), ou seja, representações digitais do relevo (no formato matricial ou TIN) são capazes de reproduzir, com exatidão, o caminho preferencial de escoamento da água superficial observado no mundo real, tendo ainda como grande vantagem deste processo de automação em relação aos procedimentos convencionais, a possibilidade de armazenamento e compartilhamento dos dados digitais.

Então, diante da necessidade de utilização destas curvas para obtenção de alguns parâmetros morfológicos, foi visto que o mais viável seria gerar um Modelo Digital de Elevação (MDE), uma vez que tinha-se acesso aos pontos cotados, georeferenciados e com os valores de altitudes associados, oriundos da SRTM (Shuttle Radar Topographic Mission), projeto internacional iniciado em 2000 e gerenciado pela Agência Nacional de Inteligência Geo-Espacial (NGA) e pela NASA (Rabus et al 2003), cobrindo todo o Estado com uma malha de 90 m x 90 m.

Desta forma, usando esta malha de pontos cotados conjuntamente com a extensão 3D ANALYST do software ArcView GIS do ESRI (Environmental Systems Research Institute), foi gerado o MDE para todo o estado da Paraíba a partir da interpolação dos pontos georeferenciados. Essa extensão do software ArcView utiliza uma técnica de interpolação baseada em diferenças finitas iterativas e é otimizado para ter a eficiência computacional dos interpoladores locais sem perder a continuidade da superfície proporcionada pelos interpoladores globais. Para dar mais credibilidade às curvas de níveis geradas, foi realizada uma validação comparando-as com as curvas de nível das Cartas da SUDENE.

Uma vez vencidas estas etapas e, levando-se em consideração os objetivos do estudo, foram selecionados e determinados os parâmetros morfológicos apresentados nas Tabelas 4.8 a 4.12, cuja descrição mais detalhada encontra-se no Anexo A.

Tabela 4.8: **Características das medidas lineares das bacias**

Parâmetro	Equação	Descrição	Unidade
Área (A)		Projeção horizontal inclusa entre os divisores topográficos.	km ²
Perímetro (P _f)		Comprimento da linha que define o divisor da bacia.	km
Linha de fundo (L)		Distância medida ao longo do curso de água principal desde a seção de referência até o divisor da bacia.	km
Comprimento do curso d'água (L _c)		Distância medida ao longo do curso de água principal desde a seção de referência até o final do canal.	km
Comprimento da rede de drenagem (L _d)		Somatório dos comprimentos de todos os talwegues da bacia.	km
Largura média (L _m)	A/L	Quanto mais L _m se aproxima de L, tanto mais a bacia é propícia a ocorrência de cheias instantâneas.	km

Tabela 4.9: **Características da forma das bacias**

Parâmetro	Equação	Descrição	Unidade
Índice de compacidade (K _c)	$K_c = \frac{0,28P}{\sqrt{A}}$	Relação entre o perímetro da bacia hidrográfica e o perímetro de um círculo de igual área.	Adm
Índice de circularidade (K _e)	$K_e = \frac{A}{A_0}$	Relação entre a área de um círculo que tem o perímetro igual ao da bacia e a área da bacia.	Adm
Fator de forma (K _f)	$K_f = \frac{L_m}{L}$	Relação entre a largura média e a linha de fundo da bacia.	Adm

Tabela 4.10: Características da rede de drenagem

Parâmetro	Equação	Descrição	Unidade
Ordem dos cursos de água (O_r)		Ordem do curso de água segundo Strahler (Anexo A).	Adm
Índice de bifurcação (R_b)	$R_b = \frac{N_u}{N_{u+1}}$	Média geométrica das relações entre números de talvegues de duas classes consecutivas.	Adm
Índice dos comprimentos (R_L)	L_u/L_{u+1}	Média geométrica das relações entre os comprimentos médios dos talvegues de duas classes consecutivas.	Adm
Índice das áreas (R_a)	$R_a = A_u/A_{u+1}$	Média geométrica das relações entre as áreas médias dos talvegues de duas classes consecutivas.	Adm
Densidade de drenagem (D_d)	$D_d = \frac{\sum L_i}{A}$	Quociente entre o comprimento total da rede de drenagem e a área da bacia hidrográfica.	km/km ²
Coefficiente de Torrencialidade (C_t)	$C_t = \frac{N}{A}$	Relação entre o número de cursos de água e a área da bacia.	n/km ²
Índice de rugosidade (IR)	$IR = D_d H$	Produto entre a densidade de drenagem D_d (km/km ²) e a amplitude altimétrica H (km)	Adm
Extensão média do escoamento superficial (L_e)	$L_e = \frac{A}{4L}$	Distância média em que a água da chuva teria que escoar sobre a superfície de uma bacia hidrográfica, caso o curso de água principal fosse uma linha reta.	km
Sinuosidade do curso d'água (SIN)	$SIN = \frac{L_r}{L_h}$	Relação entre o comprimento do rio principal e a sua projeção horizontal.	Adm

Tabela 4.11: Características do relevo das bacias

Parâmetro	Equação	Descrição	Unidade
Declividade máxima (I_{\max}) do rio	$I_{\max} = \frac{H_{\max} - H_{\min}}{L_r}$	Gradiente topográfico que afeta o movimento lateral da água subterrânea em áreas saturadas próximas da superfície se considerada a bacia como um todo.	m/m
Elevação média da bacia (C_{med})	$C_{\text{med}} = \frac{\sum c a}{A}$	c é a elevação média entre duas curvas de níveis consecutivas e a é a área entre as duas curvas de nível.	m/m
Retângulo equivalente	$l_r = \frac{K_c \sqrt{A}}{1,128} \left\{ 1 - \sqrt{1 - \left[\frac{1,128}{K_c} \right]^2} \right\}$ $L_r = \frac{K_c \sqrt{A}}{1,128} \left\{ 1 + \sqrt{1 - \left[\frac{1,128}{K_c} \right]^2} \right\}$	Retângulo com o mesmo perímetro, mesmo índice de compacidade e mesma distribuição hipsométrica que a bacia e é caracterizado pelo lado menor l_r e lado maior L_r .	km
Índice de declividade média da bacia (I_p)	$I_p = \frac{\sum \sqrt{a_i c_i}}{\sqrt{L_r}}$	Média ponderada das declividades do conjunto das áreas elementares para as quais pode-se considerar que a linha de maior declividade é constante.	Adm
Índice de declividade global (IG)	$IG = D/L_r$	D a diferença de altitude em metros entre as duas curvas de nível que abrangem 5% e 95% da área da bacia (localizada abaixo e acima delas).	m/m
Desnível específico (DS)	$DS = \frac{D\sqrt{A}}{L_r}$	É função da forma do contorno da bacia.	m/m

Tabela 4.12: Características da capacidade de escoamento das bacias hidrográficas

Parâmetro	Equação	Descrição	Unidade
L_{600}		Lâmina Média Anual escoada em uma bacia hidrográfica localizada na zona climática Sertão com precipitação anual média de 600 mm.	mm
PAE	AE/A	Porcentagem da área da bacia hidrográfica (A) coberta por espelho de água (AE).	Adm

4.5.2 Percentuais dos solos predominantes nas bacias

A base digital do mapa de solos do Estado, obtida junto à Agência Executiva de Gestão das Águas do Estado da Paraíba - AESA, juntamente com o software ArcView GIS, permitiram realizar o recorte das bacias hidrográficas no mapa de solos, podendo, a partir da visualização dos tipos de solo em cada bacia, associar informações de algumas delas, conforme ilustrado nas Figuras 4.6 e 4.7.

Foram encontrados doze tipos de solos nas 41 bacias hidrográficas estudadas. Visando reduzir o número de variáveis relativas ao solo a serem consideradas nos estudos de regionalização, foi introduzido o conceito de L_{600} (Molle & Cadier, 1992) para permitir agrupar solos com aptidões ao escoamento semelhantes, obtendo-se, desta forma, três classes distintas de solos.

O valor de L_{600} de uma bacia hidrográfica caracteriza sua capacidade de escoamento e corresponde à lâmina anual média para uma chuva anual de 600 mm numa bacia localizada na zona climática “Sertão”.

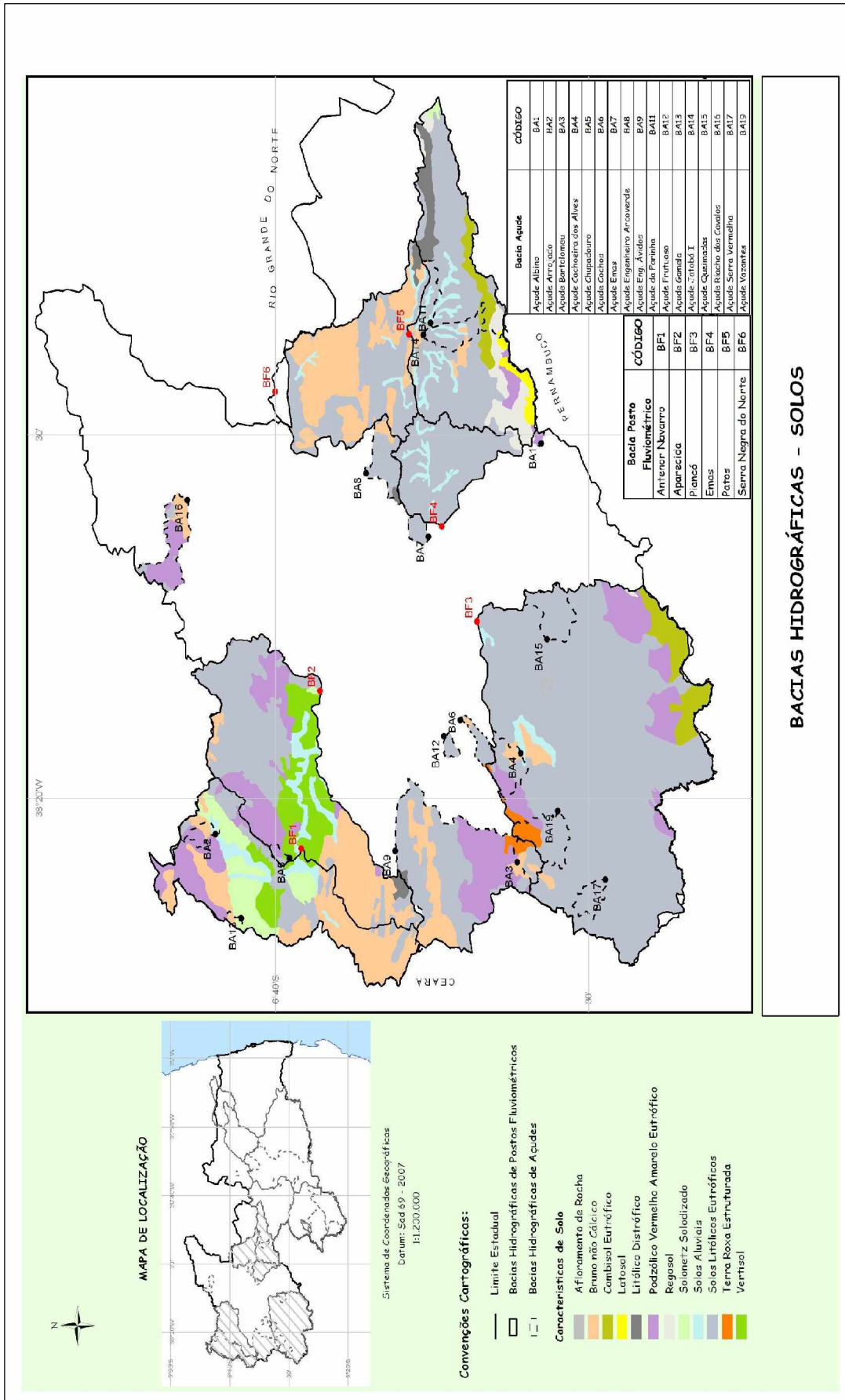


Figura 4.6: Distribuição espacial dos solos nas bacias do Sertão

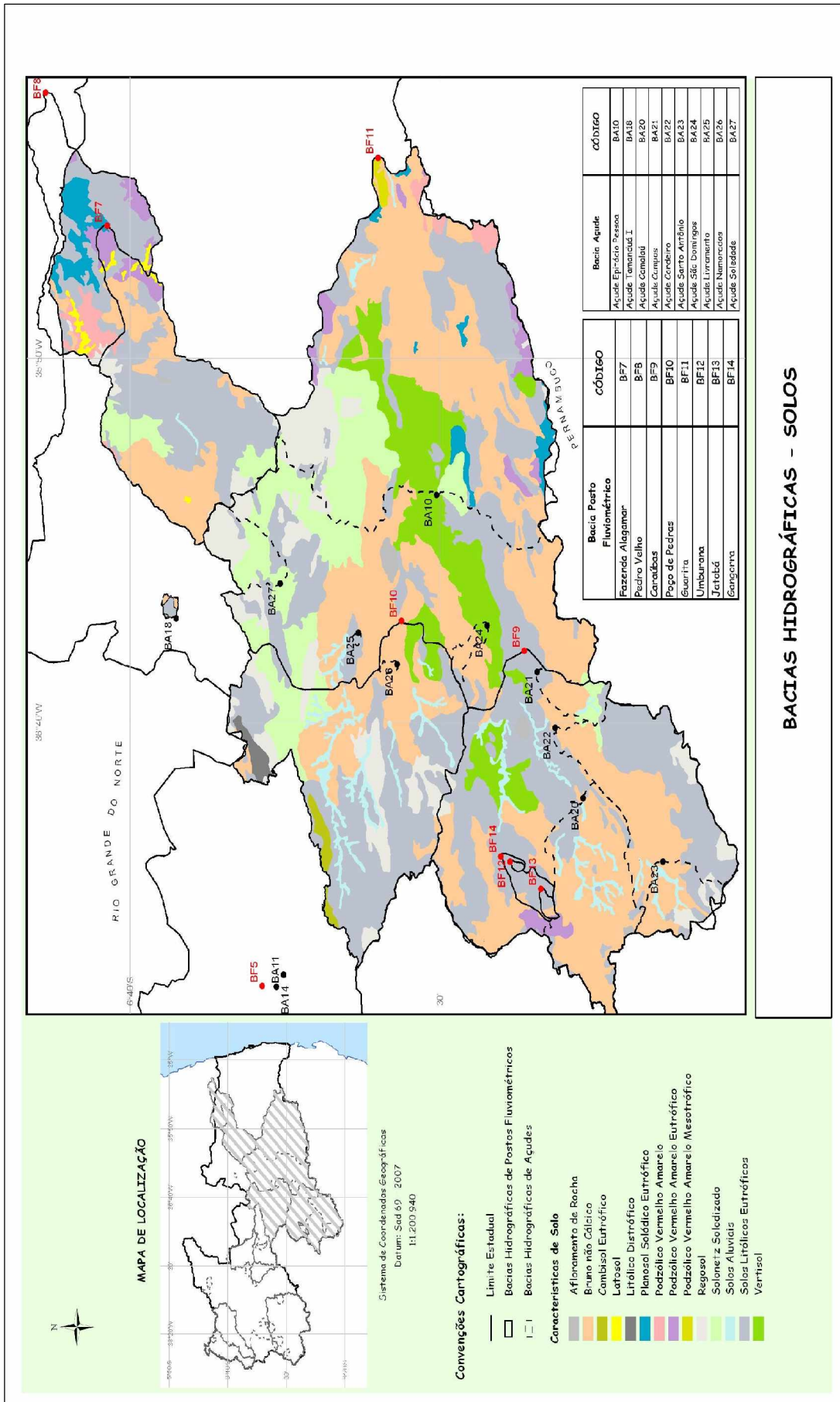


Figura 4.7: Distribuição espacial dos solos nas bacias do Cariri e Curimatá

4.6 SELEÇÃO DE ATRIBUTOS

4.6.1 Considerações gerais

Para que os dados das características das bacias hidrográficas apresentados anteriormente (item 4.5) possam tornar-se úteis ao trabalho de regionalização, é necessário que sejam preparados apropriadamente para a etapa posterior da busca por relações coerentes entre eles e os parâmetros do modelo SMAP.

Nos últimos anos tem se intensificado o uso do processo denominado descoberta de conhecimento em bases de dados (Knowledge Discovery on Databases – KDD), o qual combina técnicas de aprendizado de máquinas, reconhecimento de padrões, estatística, banco de dados e visualização para extrair conhecimento e/ou informação dos dados brutos disponíveis. Embora utilize técnicas consagradas há muito tempo nos ramos da estatística e da matemática, o fato de ter que lidar com grande quantidade de dados, por causa da atual facilidade de armazenamento eletrônico, e de sistematizar as ferramentas úteis ao processo, o acentua como alternativa para lidar com estes conjuntos complexos de dados.

De acordo com Fayyad et al. (1996), o KDD é compreendido pelas quatro fases apresentadas na Figura 4.8. Após a obtenção dos dados (características das bacias e parâmetros do modelo SMAP) na fase de *data warehousing*, seleciona-se na fase de pré-processamento aqueles considerados úteis para a tarefa seguinte, excluindo-se as bacias atípicas e, das bacias restantes, as características redundantes e irrelevantes, dando origem a uma base mais estabilizada dos dados. Dois atributos são redundantes entre si quando seus valores são correlacionados. Um atributo é irrelevante quando a sua inclusão ou retirada do conjunto de dados não melhora ou empobrece significativamente os resultados.

Na etapa denominada Mineração de Dados, escolhe-se uma ferramenta apropriada (redes neurais, algoritmos genéticos, modelos estatísticos) para procurar padrões interessantes e úteis nos dados. Neste estudo, procura-se padrões entre características físico-climáticas das bacias e os parâmetros do modelo SMAP obtidos por meio de calibração. A fase de pós-processamento compreende a validação e interpretação dos resultados, além da organização da informação encontrada, de maneira que ela possa ser verificada e reusada.

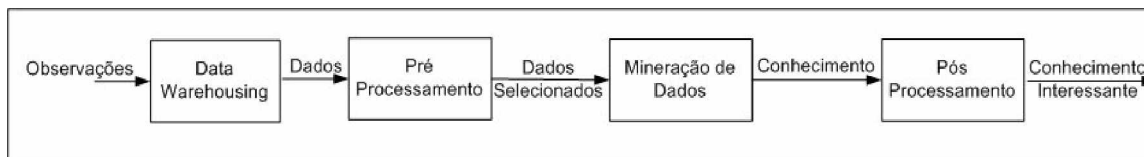


Figura 4.8: Fases do processo de KDD (Liu e Motoda, 2002)

A seleção de atributos tem sido nas últimas décadas um campo ativo e frutífero da pesquisa e do desenvolvimento no ramo da descoberta de padrões (Mitra et al., 2002), aprendizagem de máquinas (Liu & Motoda, 2002), mineração de dados (Dash et al., 2002) e estatística (Hastie et al., 2001). Tem como objetivo descobrir um subconjunto de características relevantes das bacias para a tarefa de regionalização, visando, entre outras coisas, tornar o processo de aprendizagem mais eficiente. Atributos redundantes prejudicam o desempenho da ferramenta de aprendizagem tanto na velocidade (devido à dimensionalidade dos dados) quanto no índice de acerto (devido à presença de informações redundantes) (Kira & Rendell 1992).

É uma tentativa de selecionar apenas as características que melhor descrevem e discriminam o conjunto de dados e suas estruturas latentes, reduzindo a dimensionalidade dos dados, aumentando a eficiência das ferramentas e a consequente qualidade dos resultados obtidos (Lee, 2005).

4.6.2 Técnicas de aprendizado de máquinas e avaliação dos dados

O aprendizado de máquinas pode ser definido como um conjunto de atividades que buscam adquirir conhecimento sobre um dado domínio através de experiências anteriores e observações do mundo real. Compreende o desenvolvimento de técnicas computacionais que permitem simular o processo de aprendizado e a construção de sistemas aptos a obter conhecimento automaticamente (Mitchell, 1997).

O processo de aprendizagem usa cinco estratégias identificadas de acordo com a complexidade da inferência: 1) hábito; 2) instrução; 3) dedução; 4) analogia; 5) indução. Este último, de maior interesse para as pesquisas relacionadas ao aprendizado de máquinas, caracteriza-se pelo raciocínio que parte do específico para o geral e pode ser dividido em supervisionado e não-supervisionado, conforme ilustrado na Figura 4.9.

O aprendizado supervisionado é aplicado quando existe um número expressivo de exemplos rotulados e o objetivo é induzir conceitos a partir de exemplos pré-classificados. Podem ser divididos em classificação (se os dados são discretos) e regressão (se os dados são contínuos).

Quando os exemplos são não rotulados, aplica-se o não supervisionado, portanto não possuem uma classe correspondente e, neste caso, o trabalho estará relacionado com o **agrupamento** dos exemplos (clustering), com uma descrição compacta de um subconjunto de dados denominado **sumarização**, ou com a caracterização, por meio de **regras de associação**, do quanto a presença de um conjunto de atributos implica na presença de algum outro

conjunto distinto de atributos nos mesmos exemplos (Martins, 2003). Aqui, cada bacia hidrográfica foi rotulada pelos parâmetros do modelo SMAP obtidos durante a calibração; suas características físico-climáticas são seus atributos.

Em função dos métodos adotados para definir os grupos, as técnicas de agrupamentos mais conhecidas são **particional**, **redes auto-organizáveis**, e **hierárquico** (Jain et al., 1999), destacando-se este último por reunir as características de maior interesse para o presente trabalho.

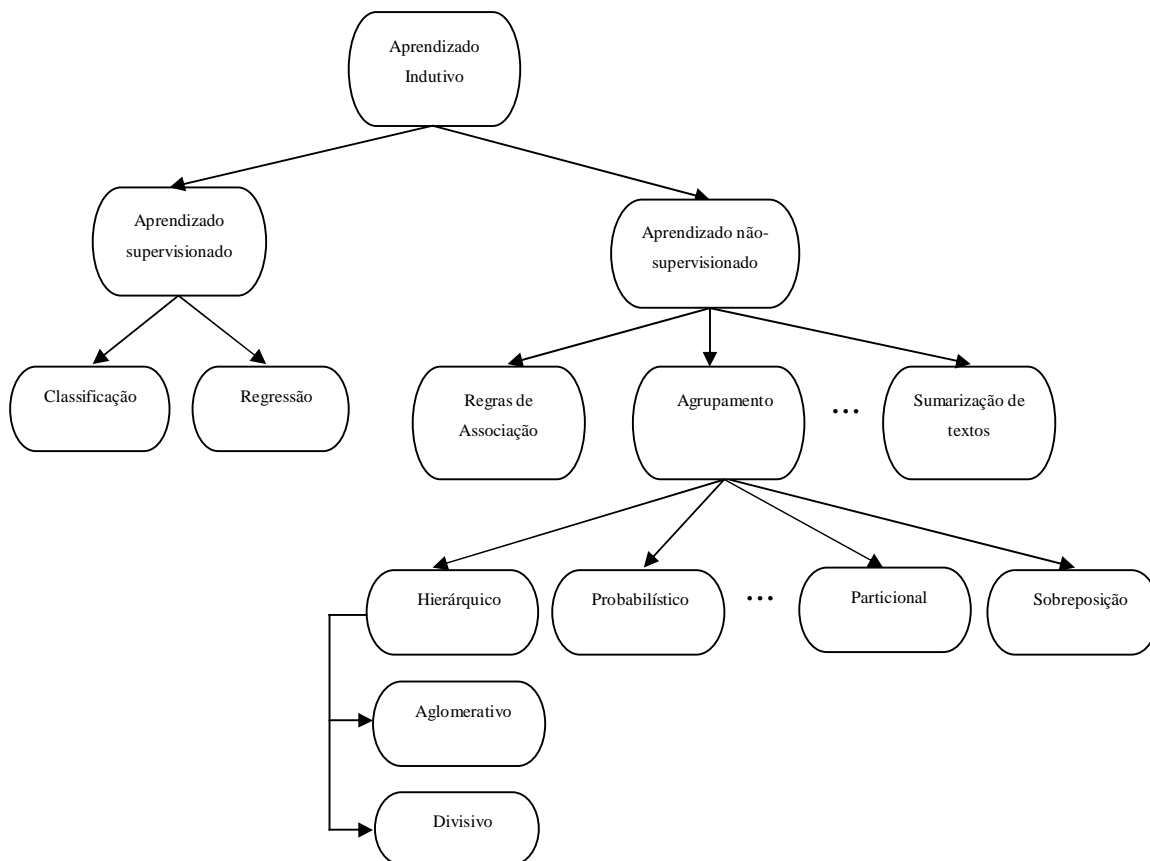


Figura 4.9: **Hierarquia do aprendizado indutivo**

Agrupamento particional: obtém uma partição simples dos dados em vez de uma estrutura de agrupamento, como o dendograma produzido através de técnicas de agrupamento hierárquico. Tem a vantagem sobre outros métodos em aplicações envolvendo grande quantidade de dados para os quais a construção de dendogramas é computacionalmente proibitiva. As duas principais desvantagens são a escolha antecipada do número de grupos, e serem propensos a caírem em mínimos locais. Conforme Han e Kamber (2001), os algoritmos de particionamento mais conhecidos são o *k-means* e o *K-medoids* e suas variações.

No algoritmo *k-means*, k objetos são selecionados aleatoriamente, dos quais cada um, inicialmente, representa a média do grupo. Para cada um dos objetos restantes, é feita a atribuição ao grupo ao qual o objeto é mais similar, baseado na distância entre o objeto e a média do cluster. Ele, então, calcula as novas médias para cada grupo. Este processo se repete até que o critério de parada seja satisfeito, em geral, o erro quadrático é usado.

O algoritmo *k-medoids*, em vez da média do grupo adotada no *k-means*, usa como representante o objeto que se encontra mais próximo ao centro de gravidade do grupo.

Agrupamento com base em redes auto-organizáveis: também denominados redes SOM - Self Organizing Map (Kohonen, 1990), o treinamento é não supervisionado e os dados de entrada são agrupados em conjuntos que agregam padrões semelhantes através do aprendizado competitivo. Realiza um mapeamento de um espaço contínuo de entrada para um espaço discreto de saída, onde as propriedades topológicas da entrada são preservadas. É uma rede linear de camada única totalmente conectada, cuja saída é organizada em uma ou duas dimensões.

No mapa de Kohonen os neurônios da camada de saída são organizados em uma estrutura uni ou bidimensional, conforme mostrado na Figura 4.10. A diferença entre as redes SOM e o agrupamento simples, que resulta de RNA's competitivas, é a organização ordenada dos grupos resultantes, que permite a visualização das distâncias relativas entre agrupamentos. Se estas relações (ditas topológicas) não forem importantes, uma rede competitiva de uma camada, simples, seria igualmente útil.

Os dois algoritmos mais conhecidos que implementam as redes SOM são o SOTA (Herrero et al., 2001) e o SomPack (Kohonen et al., 1996).



Figura 4.10: Topologia de uma rede de Kohonen

Clustering Hierárquico: A abordagem do agrupamento hierárquico constrói os agrupamentos de tal forma que exemplos pertencentes ao mesmo grupo possuem alta similaridade e unidades pertencentes a grupos diferentes possuem baixa similaridade. As técnicas hierárquicas criam árvores denominadas de dendogramas. As folhas da árvore representam exemplos individuais e os nós internos representam grupos de unidades (no presente caso grupos de bacias hidrográficas).

Dois estratégias podem ser utilizadas para implementar os algoritmos de agrupamento hierárquico (Metz, 2006): a) aglomerativo (bottom-up); b) divisiva (top-down). Na primeira, a árvore é construída de baixo para cima (das folhas para a raiz), ou seja, cada unidade (cada bacia hidrográfica) é considerada um grupo unitário, sendo iterativamente adicionados pares desses grupos de acordo com um índice de similaridade até que todos pertençam a um único grupo; na segunda, as árvores são construídas de cima para baixo e é iniciada com apenas um agrupamento com todos os exemplos, seguindo o caminho inverso da primeira abordagem, até alcançar o número de grupos desejados.

As vantagens do agrupamento hierárquico são sua simplicidade e o fato de produzir resultados mais fáceis de serem visualizados e interpretados do que os gerados pelos algoritmos de agrupamento particional. Porém, existem duas desvantagens: a primeira é que pequenos erros cometidos nos estágios iniciais dos algoritmos tendem a se propagar durante o processo, pois não tem a propriedade de voltar e reparar certos erros. A segunda desvantagem é que, à medida que os grupos crescem, o vetor que representa o grupo pode não representar mais nenhum dos elementos do grupo, tornando os padrões de expressão dos elementos menos relevantes.

Em geral, quatro etapas são realizadas no processo de agrupamento: 1) pré-processamento dos dados; 2) escolha de uma medida de similaridade; 3) escolha e aplicação de um algoritmo de agrupamento; 4) avaliação e interpretação dos resultados.

Pré-processamento dos dados: é a preparação e transformação dos dados, quando necessário, podendo os atributos assumirem valores na forma binária, discreta ou contínua. Neste último caso, o atributo pode ser representado por qualquer valor real dentro de um intervalo pré-definido. Além das três formas citadas, o agrupamento é influenciado pela escala dos valores dos atributos, a qual pode ser qualitativa (nominal ou ordinal) ou quantitativa (intervalo de valores ou proporção) (Metz, 2006). Pode ser exigido algum tipo de normalização dos valores, que considera a medida de similaridade a ser aplicada.

Medidas de similaridade: partindo do princípio de que os atributos são dimensões de um espaço multidimensional e a descrição de cada exemplo (cada bacia hidrográfica) corresponde a um ponto no espaço, a similaridade entre os exemplos pode ser medida por uma função de distância, de correlação ou de associação, sendo as duas primeiras medidas empregadas quando os dados são do tipo numérico (quantitativos). Quando os atributos são discretos utilizam-se as medidas de associação.

Neste trabalho adotou-se a distância euclidiana, definida pela Equação 4.4, talvez a mais utilizada em estudos de agrupamentos. Expressa a distância geométrica entre os exemplos num espaço multi-dimensional.

$$dist(E_1, E_2) = \sqrt{\sum_{i=1}^M (x_{1i} - x_{2i})^2} \quad 4.4$$

Algoritmos de agrupamentos: alguns algoritmos têm sido propostos, diferenciados entre si pelo método utilizado para identificar os pares de grupos mais semelhantes e por algumas heurísticas que promovem a melhoria dos resultados, destacando-se: Single Link, Complete Link, Average Link, Ward. Estes algoritmos operam sobre a matriz de similaridades e não dos valores originais para evitar a influência da magnitude de cada um.

Algoritmo Single Link: utiliza a técnica do vizinho mais próximo na qual a distância entre dois grupos é medida pela distância do par de exemplos mais próximo entre si, sendo cada exemplo pertencente a cada um desses grupos, respectivamente, conforme ilustrado na Figura 4.11 (Metz & Monard, 2006).

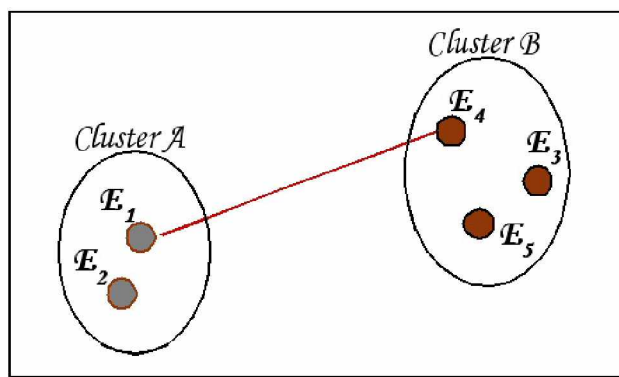


Figura 4.11: **Single Link**: menor distância entre dois clusters

Algoritmo Complete Link: utiliza a estratégia do vizinho mais distante. Ao contrário do método Single Link, essa estratégia determina a distância entre dois agrupamentos de acordo com a maior distância entre um par de exemplos, sendo cada exemplo pertencente a um cluster distinto, conforme ilustrado na Figura 4.12 (Metz & Monard, 2006). Desta forma, tem-se maior possibilidade de identificação de clusters menos alongados.

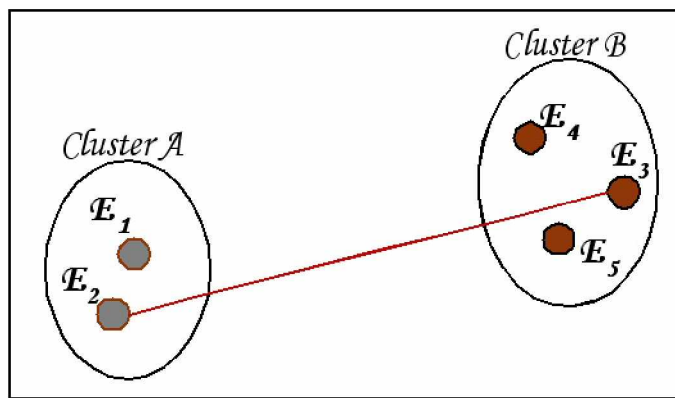


Figura 4.12: **Complete Link**: maior distância entre dois clusters

Algoritmo Average Link: a distância entre dois grupos é definida como a média da distância entre todos os pares de exemplos em cada grupo, cada par composto por um exemplo de cada grupo, conforme ilustrado na Figura 4.13 (Metz & Monard, 2006).

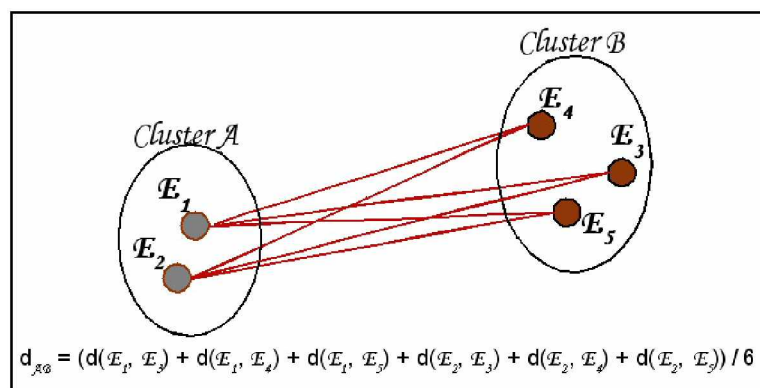


Figura 4.13: **Average Link**: média das distâncias entre os dois clusters

Algoritmo de Ward: O método de Ward é um método de agrupamento de dados que forma grupos de maneira a atingir sempre o menor erro interno entre os vetores que compõem cada grupo e o vetor médio do grupo. Isto equivale a buscar o menor desvio padrão entre os dados de cada grupo. No princípio, têm-se m grupos; ou seja, um grupo para cada vetor componente

da base de dados (um grupo para cada bacia hidrográfica). Neste estágio inicial o erro interno é nulo para todos os grupos, pois cada vetor que compõe cada grupo é o próprio vetor médio do grupo. Igualmente, o desvio padrão para cada grupo é nulo. Na etapa subsequente, cada possibilidade de aglutinação entre os grupos 2 a 2 é verificada, e é escolhido o agrupamento que causa o menor aumento no erro interno do grupo. São $m \times m$ verificações.

4.6.3 Avaliação de cada bacia no conjunto dos dados

A avaliação dos dados disponíveis neste trabalho foi feita em duas etapas distintas: 1) avaliação da importância de cada bacia no conjunto dos dados, com a aplicação da técnica de aprendizado hierárquico não supervisionado; 2) seleção de atributos, com o uso de técnicas de aprendizado supervisionado.

4.6.4 Seleção de atributos

Os métodos de Seleção de Atributos - SA em geral escolhem os atributos segundo duas abordagens: 1) pela avaliação individual; ou, 2) pela avaliação de subgrupos extraídos do grupo inicial de dados. O primeiro método propõe a remoção de atributos irrelevantes. O segundo método permite remover tanto atributos irrelevantes quanto redundantes. Qualquer método de seleção incluído nesta segunda classe pode ser caracterizado por seu interesse em relação às quatro questões seguintes, as quais determinam a natureza do processo de busca (Lee, 2005):

- A. o(s) ponto(s) de partida ou a direção em que a busca será realizada;
 - a. forward: inicia-se com um conjunto vazio de atributos, sendo passo-a-passo adicionado um novo atributo.
 - b. backward: inicia-se com um conjunto de todos os atributos, indo sucessivamente removendo-os.
 - c. bi-direcional (best first): é uma variação das duas abordagens anteriores, com duas buscas processadas simultaneamente.
 - d. randômica: não há uma direção específica na qual a busca será realizada, evitando cair em regiões de mínimos locais.
- B. a estratégia de busca;
 - a. busca completa: busca exaustiva para gerar um conjunto ótimo de atributos.
 - b. busca heurística: emprega algum tipo de heurística para conduzir a busca e evita que esta seja exaustiva.

- c. busca não-determinística: procura aleatoriamente pelos subconjuntos ótimos de atributos.
- C. o critério de busca;
 - a. embedded: embutida no próprio algoritmo de aprendizado.
 - b. filtro: utilizado para filtrar atributos sem considerar o algoritmo de aprendizado. Foi utilizado no presente trabalho.
 - c. wrapper: emprega o próprio algoritmo de aprendizado como uma “caixa preta”.
- D. o critério de parada (quando não é mais possível melhorar a performance).
 - a. quando remover ou adicionar atributos não melhora a precisão da estimativa para a classificação;
 - b. se ao revisar o subconjunto de atributos a precisão se degrada;
 - c. se o número de subconjuntos candidatos alcançar o outro extremo do espaço de busca;
 - d. quando o subconjunto de atributos selecionado separar perfeitamente todas as classes;
 - e. quando um parâmetro do sistema adotado como critério de parada for alcançado.

Medida de avaliação de atributos

Na abordagem de busca denominada filtro, os diferentes algoritmos de seleção de atributos podem ser categorizados em dois grupos distintos: 1) os que conferem pesos aos atributos para avaliá-los individualmente; 2) os que buscam subconjuntos ótimos de atributos.

No primeiro grupo, são imputados pesos aos atributos, os quais são ordenados com base nas suas capacidades preditivas, ou seja, o quanto cada um pode explicar o valor da classe (no nosso caso, quatro classes: E2, STR, E1, CINF). Um atributo é bom e será selecionado se o peso a ele associado for maior que um limite pré-estabelecido.

Os algoritmos do segundo grupo realizam a busca por subconjuntos de atributos guiados por uma medida de avaliação capaz de capturar a o nível de qualidade de cada subconjunto. O coeficiente de correlação linear pode ser adotado como esta medida com duas vantagens: primeiro, ajuda a remover atributos cuja correlação linear com a classe seja próxima de zero; segundo, ajuda a reduzir a redundância entre atributos selecionados. Porém, pode não ser apta para obter a correlação não linear frequentemente verificada em dados reais.

Para superar esta limitação e medir a correlação entre dois atributos ou entre um atributo e uma classe, adota-se em geral a medida de correlação denominada Symmetrical Uncertainty – SU (Press et al., 1992).

Para uma variável discreta X com N valores possíveis $\{x_i, i=1, \dots, N\}$, sua entropia é definida pela equação:

$$H(X) = - \sum_{i=1}^N P(x_i) \log_2(P(x_i))$$

A entropia de X após observar os valores de outra variável Y é definida como:

$$H(X, Y) = - \sum_{j=1}^N P(y_j) \sum_{i=1}^N P(x_i | y_j) \log_2(P(x_i | y_j))$$

onde: $P(x_i)$ é a probabilidade para todos os valores de X , e $P(x_i | y_j)$ é a probabilidade posterior de X dados os valores de Y . A quantidade pela qual a entropia de X diminui reflete o ganho de informação acerca de X dado Y . Dado outra variável discreta Y , o Ganho de Informação $GI(X|Y)$ ao incluí-la será:

$$GI(X|Y) = H(X) + H(Y) - H(X, Y)$$

onde: $H(X)$, $H(X, Y)$ e $GI(X|Y)$, representam respectivamente a entropia do atributo X , a entropia condicional de X dado Y , e o ganho de informação que é uma medida simétrica a qual reflete a informação adicional acerca de X dado Y .

De acordo com a equação acima, o atributo Y é considerado mais correlacionado ao atributo X do que a um atributo Z , se $GI(X|Y) > GI(Z|Y)$.

A simetria significa que a quantidade de informação ganha com um atributo Y após observar X é igual à quantidade de informação ganha com o atributo X após observar Y . A simetria é uma propriedade desejada na avaliação da inter-correlação atributo-atributo e na correlação atributo-classe, pois assegura que a ordem de dois atributos $[(X, Y)$ ou $(Y, X)]$ não afeta o valor de SU .

Quando aplicado em seleção de atributos, o ganho de informação é influenciado por múltiplos valores de atributos nominais e pelas divisões destes valores em intervalos. Daí, Hall et al. (2000) propôs para o cálculo simultâneo da contribuição do atributo e a correlação a seguinte expressão:

$$SU(X, Y) = 2 \left[\frac{GI(X|Y)}{H(X) + H(Y)} \right] \quad 4.5$$

Seu valor varia entre 0 e 1. Quando igual a 1 indica que o valor de um atributo prediz completamente o valor de outro. Por outro lado, o valor 0 indica que os dois atributos são independentes. Além disso, os atributos são tratados simetricamente (Yu & Liu, 2004).

Ela permite aproximar tanto a análise de relevância quanto a de redundância em duas fases: na primeira fase, a medida SU entre cada atributo (cada característica da bacia hidrográfica) na classe (parâmetro do modelo SMAP) é calculada e os atributos são ordenados de acordo com sua relevância na classe. Apenas os atributos com valores de SU superiores a um limiar mínimo adotado serão analisados na fase seguinte. Por não conhecer antecipadamente acerca dos dados deste trabalho, adotou-se um limiar igual a zero, por ser mais seguro, mesmo demandando mais tempo computacional.

Na segunda fase, os atributos são avaliados na ordem em que foram ordenados na fase anterior, de acordo com a redundância de uns em relação aos outros, produzindo um subconjunto final de atributos relevantes e não redundantes.

Algoritmo de seleção de atributos

Os quatro algoritmos mais frequentemente utilizados na abordagem filtro para seleção de atributos são: 1) ReliefF (Kira & Rendell, 1992); 2) FCBF – Fast Correlation-Based Filter (Yu & Liu, 2004); 3) CBF – Consistency-Based Filter (Liu & Setiono, 1996); 4) CFS – Correlation-based Feature Selection (Hall, 2000), sendo este último adotado no presente trabalho haja visto considerar a habilidade preditiva individual de cada atributo e o grau de redundância entre eles, preferindo subgrupos de atributos que têm alta correlação com a classe mas baixa inter-correlação. Portanto, trata tanto o problema da relevância quanto o da redundância de atributos.

O CFS é um algoritmo de filtro simples que ordena subconjuntos de atributos de acordo com uma correlação baseada numa função de avaliação heurística. A tendência da função de avaliação é encontrar subconjuntos que contenham atributos que sejam altamente correlacionados com a classe e não correlacionado com cada outro atributo. Devem ser ignoradas características irrelevantes porque têm baixa correlação com a classe. Características redundantes devem ser descartadas por apresentar alta correlação com uma ou mais das características restantes. A função de avaliação de subconjunto de atributos do CFS é apresentada a seguir:

$$M_S = \frac{k\bar{r}_{cf}}{\sqrt{k + k(k-1)\bar{r}_{ff}}} \quad 4.6$$

onde: M_S é uma medida de mérito da heurística de um subconjunto de atributos S contendo k atributos, $\overline{r_{cf}}$ é a correlação média atributo-classe, e $\overline{r_{ff}}$ é a inter-correlação média atributo-atributo. O numerador da Equação 4.6 pode dar a idéia do quanto um conjunto de atributos pode atuar como preditivo de uma classe; o denominador de quanta redundância há entre os atributos.

Segundo Hall (1999) a Equação 4.6 representa a parte mais importante do algoritmo CFS e impõe um ordenamento sobre os subconjuntos de atributos no espaço de busca de todos os subconjuntos de atributos possíveis. O subconjunto com alto mérito (medido pela Equação 4.6) encontrado durante a busca é usado para reduzir a dimensionalidade dos dados. A Figura 4.14 mostra os estágios do algoritmo CFS e como ele é usado na conjunção com um esquema de aprendizado de máquina (ML).

De acordo com Lee (2005) o algoritmo CFS classifica os subconjuntos de atributos de acordo com medidas de avaliação de separabilidade, como a Symmetrical Uncertainty - SU (Press et al., 1992). Para a seleção de atributos usando este algoritmo e a medida SU, é necessário que os atributos sejam discretizados. A implementação desse algoritmo permite realizar a busca de três formas diferentes: forward, backward e best first. Em relação à medida utilizada para determinar a importância dos atributos o algoritmo CFS usa as medidas de informação e de consistência.

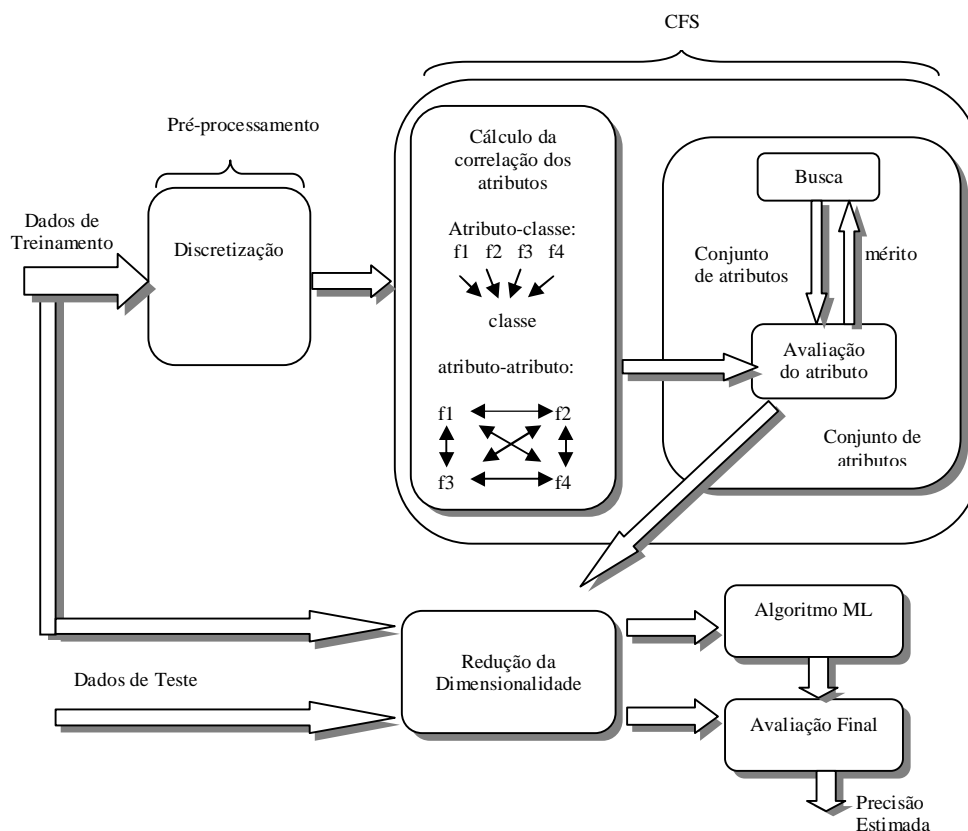


Figura 4.14 Componentes do algoritmo CFS. Os dados de treinamento e teste são reduzidos para conter apenas os atributos selecionados pelo CFS. Com a redução da dimensionalidade os dados podem ser usados para as fases de indução e previsão. (adaptado de Hall, 1999)

4.7 IDENTIFICAÇÃO DE RELAÇÕES REGIONAIS

4.7.1 Introdução

O objetivo deste capítulo é investigar a possibilidade de se estimar os valores dos parâmetros do modelo SMAP a partir de características de bacias hidrográficas, usando técnicas de redes neurais artificiais tipo Perceptron de Múltiplas Camadas - MLP.

Em geral os estudos de regionalização têm dado ênfase às estatísticas de certas características de vazões tais como as vazões anuais mínimas, médias ou máximas, curvas de duração de vazões e vazões médias mensais. Estas estatísticas têm sua relevância em diversos estudos de planejamento de recursos hídricos. Outros problemas, entretanto, requerem séries temporais de vazões, a exemplo dos estudos de vazões regularizáveis de reservatórios fluviais.

Um método comum para se obter estas séries em bacias sem estações fluviométricas é a extrapolação de registros de vazões de outra bacia que disponha de medições. As bacias

hidrográficas para as quais as séries de vazões tenham que ser estimadas podem não dispor de bacias comparáveis que possam ser usadas, tornando proibitiva qualquer extrapolação. Para tais casos, o uso de um modelo chuva-vazão com parâmetros regionalizados se apresenta como uma opção viável.

A avaliação dos valores dos parâmetros dos modelos de base física para bacias sem dados é teoricamente, possível uma vez que são supostamente mensuráveis. Entretanto, o sucesso no uso destes modelos em bacias sem dados tem sido limitado, devido à indisponibilidade dos dados requeridos nas escalas espacial e temporal, de tal forma que assegure uma modelagem verdadeiramente de base física. A bibliografia tem mostrado que, muitas vezes, quando o objetivo é derivar séries temporais de escoamentos, os modelos concentrados conseguem melhores resultados do que um modelo distribuído.

A regionalização de modelos concentrados é feita de duas maneiras: primeiro, por extrapolação dos parâmetros obtidos numa bacia com dados; ou, através de uma relação entre os parâmetros do modelo e características físicas e climáticas da bacia hidrográfica. Uma limitação desta última, em alguns casos, é que enquanto alguns dos parâmetros podem ser preditos, outros requerem calibração, o que não é possível em bacias sem dados (Manley, 1978). Há, todavia, evidências de que a estimativa dos valores de parâmetros para alguns modelos, a partir de características das bacias, é possível (Post & Jakeman, 1996; Fernandez et al., 2000; Yokoo et al., 2001).

Neste sentido, muitos estudos têm usado regressão múltipla para estimar os valores dos parâmetros. Porém, vários parâmetros se relacionam com as características das bacias de forma não linear e/ou têm relacionamentos complexos que não podem ser modeladas por regressão linear. Para tais situações as redes neurais podem se constituir uma importante alternativa.

4.7.2 O método da bacia-meta

Conceito

É a bacia retirada do conjunto de dados com o objetivo de orientar o treinamento da RNA. Desta forma, a RNA é treinada para reproduzir o valor do parâmetro do modelo SMAP para a bacia omitida a partir de suas características físico-climáticas. Sendo esta RNA capaz de reproduzir o parâmetro do modelo SMAP para a bacia-meta, o mesmo pode-se esperar para outra bacia que lhe seja similar.

Este conceito tem como inspiração o critério de parada do treinamento de uma rede neural tipo MLP: o critério da validação cruzada, o qual utiliza um conjunto de dados que

serve para informar quando o treinamento deve parar, evitando que a rede perca sua capacidade de generalização.

Embora o conjunto de dados de validação não participe do ajuste dos pesos, uma vez que é apresentado à rede apenas no final de cada ciclo, o fato de ser usado para cálculo do erro e, portanto, como indicador da qualidade do ajuste da rede, o institui naturalmente como meta da rede.

Neste trabalho, definiu-se como ponto de parada aquele que, simultaneamente, reproduza o conjunto de validação (bacia-meta) e tenha estabilizado ou se tornado insignificante a evolução do erro do treinamento.

Motivação

A pequena quantidade de dados disponíveis (36 bacias hidrográficas), limita a aplicação convencional da rede neural para a qual, em geral, se adota uma divisão dos dados em três amostras: treinamento (60% dos dados), validação (20% dos dados) e verificação (20% dos dados). Nesta divisão, as amostras resultantes seriam estatisticamente insuficientes, principalmente a referente ao treinamento que ficaria com apenas 22 bacias.

O conceito de avaliação de influência de valores individuais em uma série de dados, apresentado por Cook & Weisberg (1982), segundo o qual pequenas perturbações são introduzidas na formulação do problema para então se verificar seus efeitos sobre os resultados, fundamentaram o processo de reamostragem aqui adotado. O esquema através do qual os dados são modificados podem ser por adição ou deleção de casos, um de cada vez ou em sub-grupos. O diagnóstico por deleção tem sido o mais aceito e o mais usado.

Procedimento

O critério adotado, denominado de bacia-meta, consiste de uma amostra de validação, constituída de uma única bacia retirada do conjunto dos dados, que também passa a servir para caracterizar um modelo de RNA associado a ela. Isto é feito sucessivamente para cada bacia do conjunto num processo de amostragem com reposição. Posteriormente, para outra bacia, omitida inicialmente para fins de verificação, procura-se a mais similar entre aquelas e aplica-se o modelo associado.

Desta forma, o treinamento, validação e verificação da técnica de redes neurais, é caracterizada pela exploração máxima dos dados existentes, obedecendo aos seguintes procedimentos, a serem aplicados a cada um dos quatro parâmetros do modelo SMAP, conforme ilustrado na Figura 4.15.

- 1) Do conjunto das 36 bacias, omitir uma para a fase de verificação.
- 2) Das 35 bacias resultantes outra bacia (bacia de teste) é separada para orientar o treinamento da rede neural, ou seja, o processo buscará um conjunto de pesos que, ao mesmo tempo, ajuste o melhor possível as bacias do treinamento e reproduza a bacia de teste, aqui denominada de bacia-meta. Para cada bacia omitida, com vistas à verificação, resulta em 35 modelos de redes neurais e suas respectivas bacias-meta.
- 3) Do conjunto dos 35 modelos, um é selecionado para calcular o parâmetro da bacia separada para a verificação. Isto é feito com base na similaridade, medida pela distância euclidiana, que há entre esta bacia e a bacia-meta. No final, obteve-se 35 redes neurais treinadas, cada uma apta a determinar o parâmetro do modelo SMAP para sua respectiva bacia-meta.
- 4) Na fase de verificação, o parâmetro da bacia omitida é estimado por uma das 35 redes neurais cuja bacia-meta lhe seja mais similar.

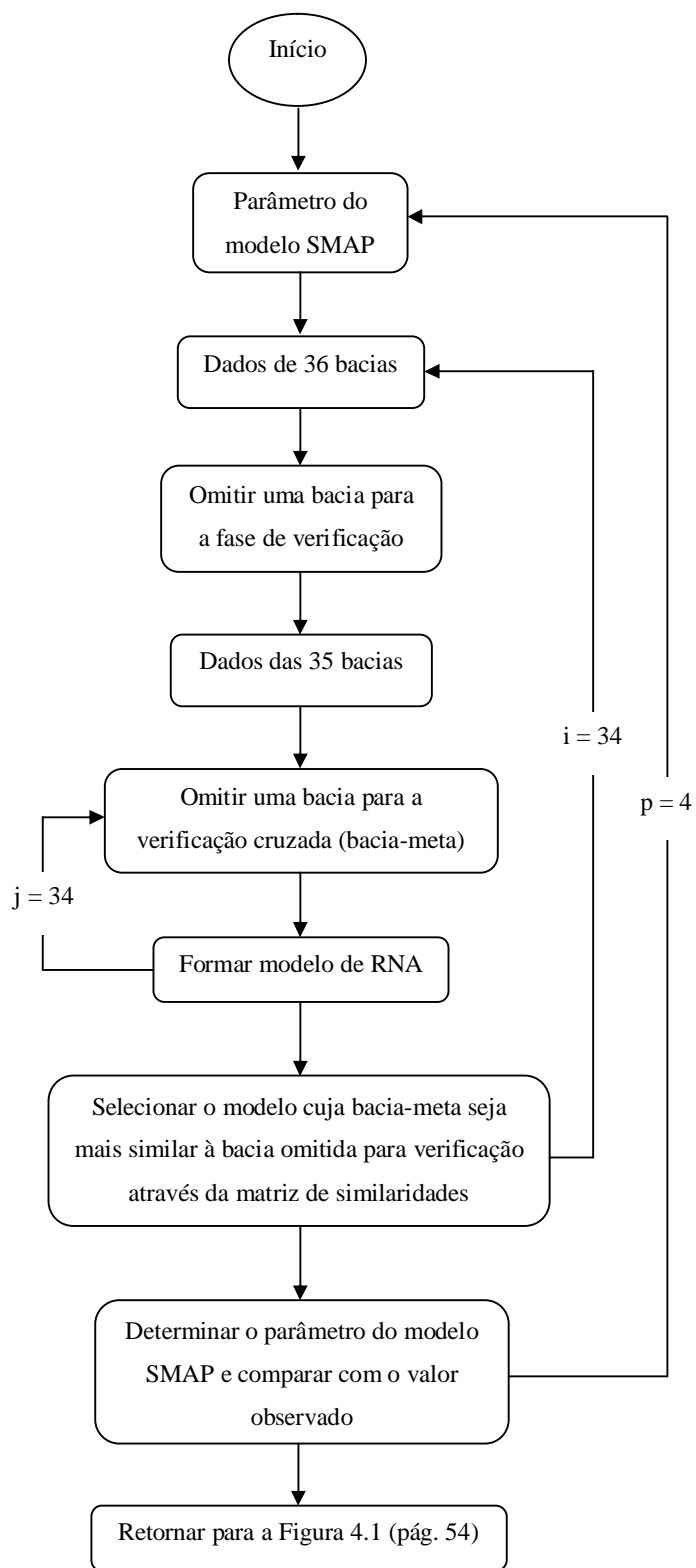


Figura 4.15: Método da bacia-meta para construção de modelo hidrológico regional

CAPÍTULO V: RESULTADOS E DISCUSSÕES

5.1 RESULTADOS DA CALIBRAÇÃO DO MODELO SMAP

O algoritmo SCE-UA foi implementado e executado neste trabalho em linguagem Fortran, em ambiente MS-DOS. O código computacional (Anexo C) é constituído pelo programa principal e três sub-rotinas, uma que executa o modelo chuva-vazão, a segunda para entrada de dados e a última para o cálculo da função objetivo (Diniz, 1994).

A rotina de entrada dos dados recebeu duas versões: uma para os dados das bacias hidrográficas das estações fluviométricas (chuva, vazão e evapotranspiração), e outra para os dados das bacias hidrográficas dos açudes. No segundo caso, além dos dados de chuva e evapotranspiração mensais, necessita-se das retiradas de água (demandas, regularização), curva cota-área-volume, cotas características da barragem e série dos níveis de água do açude. Isto, para permitir a obtenção das vazões naturais afluentes, por meio da equação do balanço hídrico.

A equação do balanço hídrico foi incorporada à rotina do modelo SMAP e a calibração se deu em relação aos níveis dos açudes em lugar dos valores das vazões afluentes.

O modelo foi considerado calibrado sempre que os valores dos parâmetros se repetissem em, pelo menos, nove vezes de dez tentativas.

Para cada bacia hidrográfica o processo de calibração foi iniciado a partir de pontos (valores dos parâmetros do modelo) iniciais gerados aleatoriamente. Os valores dos parâmetros do modelo SMAP resultantes da calibração para cada bacia estão apresentados na Tabela 5.1, com os respectivos coeficientes de correlação e de eficiência (Nash & Sutcliffe, 1970) entre as vazões observadas e calculadas, ou, no caso dos açudes, entre os volumes acumulados observados e calculados.

De acordo com os resultados mostrados na Tabela 5.1, das 41 bacias calibradas, vinte e sete foram apropriadamente calibradas, ou seja, com coeficiente de eficiência superior a 0,75 e, treze com calibração aceitável, ou seja, com coeficiente de eficiência entre 0,36 e 0,75. Apenas no caso da bacia hidrográfica do açude Riacho dos Cavalos (BA16), o coeficiente de eficiência ficou abaixo de 0,36.

Cabe destaque para as calibrações das bacias hidrográficas das estações fluviométricas, nas quais todas apresentaram coeficiente de eficiência acima de 0,75. As calibrações aceitáveis (13 casos) ficaram por conta das bacias dos açudes. Isto se deve, principalmente, à pouca quantidade de dados de níveis dos açudes e à indisponibilidade dos níveis acima das cotas das soleiras das barragens.

Tabela 5.1: Parâmetros calibrados do modelo SMAP

Ordem	Bacia Hidrográfica	Codigo	Parâmetros do SMAP				Calibração		
			E2	STR	E1	CINF	R	R2	Período
1	Albino	BA1	5,862	186,89	1,576	0,500	0,989	0,977	2000 a 2003
2	Arrojado	BA2	4,319	443,67	3,076	0,638	0,825	0,668	1996 a 2002
3	Bartolomeu I	BA3	9,908	441,39	1,021	0,190	0,896	0,787	1996 a 2003
4	Cachoeira dos Alves	BA4	3,752	121,07	6,659	0,889	0,813	0,659	1996 a 2003
5	Chupadouro	BA5	5,215	261,87	1,344	0,536	0,891	0,786	2000 a 2003
6	Cochos	BA6	0,890	2.141,65	2,004	0,900	0,813	0,651	1996 a 2003
7	Emas	BA7	4,341	61,95	4,545	0,893	0,982	0,963	2000 a 2003
8	Eng ^o Arcoverde	BA8	13,676	735,88	2,126	0,264	0,844	0,698	1996 a 2003
9	Eng ^o Ávidos	BA9	11,114	324,54	0,763	0,426	0,901	0,709	1995 a 2003
10	Epitácio Pessoa	BA10	2,299	282,62	1,008	0,494	0,914	0,792	1999 a 2003
11	Farinha	BA11	7,473	127,97	2,915	0,722	0,940	0,875	1995 a 2003
12	Frutuoso II	BA12	9,846	212,26	5,926	0,773	0,848	0,709	1995 a 2003
13	Gamela	BA13	2,342	87,95	4,311	0,900	0,838	0,670	1996 a 2003
14	Jatobá I	BA14	6,146	328,37	0,779	0,367	0,968	0,936	1995 a 2003
15	Queimadas	BA15	13,126	169,35	6,124	0,793	0,888	0,765	1995 a 2003
16	Riacho dos Cavalos	BA16	15,000	554,73	7,100	0,700	0,676	0,221	1997 a 2003
17	Serra Vermelha	BA17	1,012	6.941,06	0,069	0,000	0,958	0,913	1996 a 2003
18	Tamanduá I	BA18	1,752	5.000,00	11,08	0,635	0,828	0,673	2000 a 2003
19	Vazantes	BA19	9,564	221,00	6,054	0,807	0,777	0,550	1995 a 2003
20	Camalau	BA20	9,298	77,72	9,443	0,900	0,761	0,580	1995 a 2003
21	Campos	BA21	6,682	283,07	0,374	0,000	0,752	0,564	1995 a 2003
22	Cordeiro	BA22	6,839	389,38	0,617	0,000	0,919	0,681	1995 a 2003
23	Santo Antonio	BA23	15,000	148,74	0,463	0,369	0,935	0,849	1996 a 2003
24	São Domingos	BA24	14,915	247,95	4,281	0,681	0,948	0,880	2001 a 2003
25	Livramento	BA25	14,928	91,00	1,993	0,620	0,985	0,948	2000 a 2002
26	Namorados	BA26	15,000	265,97	0,821	0,029	0,916	0,674	2001 a 2003
27	Soledade	BA27	13,954	185,20	0,683	0,082	0,933	0,836	1996 a 2003
28	Antenor Navarro	BF1	7,587	285,40	5,825	0,534	0,996	0,992	1931 a 1933
29	Aparecida	BF2	1,640	1.012,12	0,389	0,000	0,945	0,893	1986 a 1990
30	Piancó	BF3	1,915	1.028,2	0,317	0,000	0,959	0,920	1974 a 1989
31	Emas	BF4	0,635	2163,14	0,477	0,450	0,946	0,890	1964 a 1968
32	Patos	BF5	1,389	1281,13	0,468	0,132	0,976	0,952	1985 a 1989
33	Serra Negra do Norte	BF6	2,120	957,1	0,755	0,181	0,960	0,921	1983 a 1989
34	Fazenda Alagamar	BF7	9,948	265,54	3,032	0,503	0,970	0,931	1987 a 1990
35	Pedro Velho	BF8	2,811	398,61	0,998	0,352	0,954	0,907	1985 a 1991
36	Caraúbas	BF9	3,401	267,66	2,783	0,632	0,925	0,854	1974 a 1991
37	Poço de Pedras	BF10	1,885	586,67	0,584	0,000	0,954	0,909	1971 a 1978
38	Guarita	BF11	2,522	718,86	0,939	0,190	0,960	0,920	1971 a 1991
39	Umburana/Sumé	BF12	6,053	590,07	0,135	0,000	0,996	0,992	1978 a 1980
40	Jatobá/Sumé	BF13	10,489	845,10	2,294	0,000	0,898	0,804	1975 a 1980
41	Gangorra/Sumé	BF14	5,616	803,42	1,293	0,000	0,917	0,839	1976 a 1980

5.2 CARACTERÍSTICAS DAS BACIAS

Na Tabela 5.2 estão anotadas as áreas das bacias hidrográficas, bem como os parâmetros morfológicos relativos às medidas lineares e à forma das bacias, sendo estes últimos obtidos por formulação a partir dos primeiros.

Quanto a estas características, destacam-se as bacias BA10 (açude Epitácio Pessoa) e BF11 (estação fluviométrica de Guarita) por apresentarem valores semelhantes entre si e por serem estes bem acima dos das demais bacias.

No que tange aos parâmetros que caracterizam a forma da bacia, em geral, não diferem uma das outras, exceção à bacia BA8 (açude Eng^o Arcoverde) cujos valores de Kc e Ke indicam, de forma simultânea, a vocação para ocorrência de cheias.

A Tabela 5.3 apresenta os parâmetros morfológicos que caracterizam a rede de drenagem. Quanto à estas características, as bacias BA18 (açude Tamanduá I) e BA25 (açude Livramento) são as que mais se distanciam do conjunto, particularmente influenciadas pelos valores do coeficiente de torrencialidade (Ct) e do índice dos comprimentos (RL), respectivamente.

Na Tabela 5.4 estão as sete características que tipificam o relevo das bacias hidrográficas. Devido à diversidade de localização das bacias e, conseqüentemente, variabilidade do relevo, cinco bacias mostraram-se atípicas ao conjunto como um todo.

As bacias BA10 (açude Epitácio Pessoa) e BF11 (estação fluviométrica de Guarita), por apresentarem baixo índice de declividade global (IG) e elevados valores dos comprimentos e das larguras do retângulo equivalente (lr e Lr). Estes últimos, naturalmente por causa dos tamanhos das bacias.

A bacia BA13 (açude Gamela) por ser bastante íngreme, apresenta valores de declividade máxima do rio (Imax), índice de declividade global (IG) e Desnível específico (DS) bem superiores aos valores do conjunto. Isto se deve ao fato de estar localizada em região serrana de divisor de águas.

As bacias BA23 (açude Santo Antônio) e BF5 (estação fluviométrica de Patos), pelos altos valores da declividade média (Ip). A elevação média da bacia BF5 (Cmed), também é bastante alta em relação às demais.

A Tabela 5.5 resume os doze tipos de solos encontrados no total das 41 bacias hidrográficas, assim como os valores de L_{600} (escoamento médio para uma chuva anual de 600 mm); os doze tipos de solos foram divididos em três classes conforme a maior, média e menor aptidão ao escoamento, respectivamente. A Classe I composta pelos solos 1, 6 e 9 da Tabela 5.5, a Classe II pelos solos 2, 5 e 11 e, a Classe III pelos solos 3, 7, 10 e 12. Assim, as

41 bacias hidrográficas apresentaram os percentuais de solos mostradas na Tabela 5.6. Nesta mesma tabela encontram-se os percentuais de espelhos de água sobre cada bacia hidrográficas (PAE).

Tabela 5.2: Parâmetros morfológicos das bacias hidrográficas

Ordem	Bacia Hidrográfica	A (km ²)	Medidas lineares da bacia (km)					Forma da bacia		
			Pr	L	Lt	Ld	Lm	Kc	Ke	Kf
1	Albino	9,5	13,8	3,9	3,5	3,5	2,4	1,25	0,64	0,61
2	Arrojado	30,1	35,8	12,6	12,0	56,5	2,9	1,65	0,37	0,23
3	Bartolomeu I	59,5	36,0	9,3	13,6	92,7	6,4	1,31	0,58	0,69
4	Cachoeira dos Alves	110,5	58,0	19,2	19,0	177,0	5,8	1,54	0,42	0,30
5	Chupadouro	17,8	19,1	8,5	7,7	28,3	2,1	1,27	0,62	0,24
6	Cochos	56,5	34,1	20,0	16,5	52,1	2,1	1,48	0,46	0,10
7	Emas	35,0	23,6	12,7	10,5	26,4	2,7	1,12	0,80	0,22
8	Eng ^o Arcoverde	126,9	69,2	10,5	8,5	162,0	12,0	1,72	0,34	1,14
9	Eng ^o Ávidos	1.009,5	178,0	51,0	48,3	1.264,2	19,8	1,57	0,41	0,39
10	Epitácio Pessoa	10.659,0	747,5	183,1	182,8	7.744,3	67,8	1,88	0,28	0,37
11	Farinha	747,9	196,9	68,5	62,8	803,2	10,9	2,02	0,25	0,16
12	Frutuoso II	19,1	19,1	7,7	7,4	23,4	2,5	1,22	0,67	0,32
13	Gamela	12,8	15,9	6,0	5,6	5,6	2,1	1,25	0,64	0,35
14	Jatobá I	94,0	53,8	25,5	25,2	134,5	3,7	1,55	0,41	0,14
15	Queimadas	124,3	56,6	25,0	24,7	101,5	5,0	1,42	0,50	0,20
16	Riacho dos Cavalos	161,4	87,1	24,7	23,8	208,0	6,5	1,92	0,27	0,26
17	Serra Vermelha	55,7	37,9	15,6	15,4	84,3	3,6	1,42	0,49	0,23
18	Tamanduá I	23,4	20,0	8,3	8,2	14,9	2,8	1,16	0,75	0,34
19	Vazantes	137,0	61,9	23,1	22,5	194,4	5,9	1,48	0,46	0,26
20	Camalau	1.054,0	186,5	80,7	71,9	620,8	13,1	1,61	0,39	0,16
21	Campos	181,2	63,8	32,5	32,2	86,6	5,6	1,33	0,57	0,17
22	Cordeiro	1.665,8	228,5	87,6	78,8	990,8	19,0	1,57	0,41	0,22
23	Santo Antonio	340,6	93,8	25,5	25,4	365,4	13,3	1,42	0,49	0,52
24	São Domingos	65,5	45,9	23,2	22,4	28,0	2,8	1,59	0,40	0,12
25	Livramento	37,0	24,0	4,0	3,7	5,7	1,51	1,06	0,90	0,37
26	Namorados	14,2	16,0	7,6	6,9	12,4	1,9	1,19	0,70	0,24
27	Soledade	313,1	81,2	35,9	35,6	208,3	8,7	1,29	0,61	0,24
28	Antenor Navarro	1.720,0	288,5	82,9	81,7	1.665,8	17,9	2,10	0,23	0,22
29	Aparecida	3.720,0	374,1	148,0	146,9	4.019,9	23,4	1,78	0,32	0,16
30	Piancó	4.710,0	381,6	145,9	145,4	7.056,9	31,6	1,57	0,40	0,22
31	Emas	530,0	135,3	46,9	45,3	511,1	12,8	1,55	0,42	0,27
32	Patos	1.850,0	270,8	80,3	78,1	1.843,5	20,8	1,85	0,29	0,26
33	Serra Negra do Norte	3.330,0	382,6	130,7	129,3	3.589,4	22,8	1,96	0,26	0,17
34	Fazenda Alagamar	2.270,0	244,4	125,8	125,5	1.362,9	16,9	1,48	0,45	0,13
35	Pedro Velho	3.590,0	345,2	193,8	193,8	2.466,5	17,8	1,65	0,37	0,09
36	Caraúbas	5.120,0	383,9	110,1	109,2	2.842,4	46,2	1,51	0,44	0,42
37	Poço de Pedras	3.140,0	347,7	130,3	129,1	2.371,5	24,7	1,72	0,34	0,19
38	Guarita	17.220,0	892,3	313,1	312,2	10.829,3	56,2	1,88	0,28	0,18
39	Umburana/Sumé	10,4	11,4	4,9	4,6	6,5	1,6	1,13	0,78	0,33
40	Jatobá/Sumé	26,8	22,1	9,8	9,7	14,8	2,6	1,23	0,66	0,26
41	Gangorra/Sumé	137,4	57,9	29,2	29,0	82,6	4,6	1,40	0,51	0,16

A – área da bacia; Pr – perímetro; L – linha de fundo; Lt – comprimento do curso de água principal; Ld – comprimento da rede de drenagem; Lm – largura média da bacia; Kc – índice de compacidade; Ke – índice de circularidade; Kf – fator de forma.

Tabela 5.3: Parâmetros morfológicos das bacias hidrográficas

Ordem	Bacia Hidrográfica	Características da rede de drenagem								
		Or	Rb	RL	Ra	Dd	Ct	IR	Le	SIN
1	Albino	1	1	0,50	0,40	0,37	0,11	52,3	0,85	1,28
2	Arrojado	4	3	0,34	0,65	1,54	0,11	522,7	1,56	2,04
3	Bartolomeu I	4	4	0,58	0,19	1,56	0,08	343,0	1,92	1,75
4	Cachoeira dos Alves	5	2	57,71	0,50	1,60	1,23	552,6	2,70	1,86
5	Chupadouro	3	4	0,24	0,41	1,60	0,28	127,7	0,73	1,26
6	Cochos	3	6	0,16	0,08	1,26	3,28	452,2	1,19	1,89
7	Emas	3	5	0,64	0,17	0,75	0,17	354,1	1,55	1,86
8	Eng ^o Arcoverde	5	2	0,82	0,93	1,28	0,02	344,6	3,57	0,96
9	Eng ^o Ávidos	6	3	0,58	0,92	1,25	0,00	688,8	5,98	1,14
10	Epitácio Pessoa	7	2	4,56	0,49	0,62	0,16	499,1	25,51	1,50
11	Farinha	5	5	0,46	1,80	1,07	0,01	746,4	3,78	1,27
12	Frutuoso II	3	3	1,99	0,32	1,22	1,15	483,7	0,95	1,47
13	Gamela	1	1	0,50	0,40	0,44	0,08	193,4	0,76	1,33
14	Jatobá I	5	2	0,61	0,86	1,43	0,03	944,4	1,32	1,41
15	Queimadas	4	2	0,49	0,34	0,82	0,27	269,6	2,43	1,93
16	Riacho dos Cavalos	4	6	0,26	0,72	1,29	0,04	644,3	1,82	1,07
17	Serra Vermelha	3	10	0,21	0,07	1,51	0,93	438,7	1,53	1,69
18	Tamanduá I	3	2	2,19	0,42	0,64	6,40	70,0	1,11	1,56
19	Vazantes	4	4	0,96	0,20	1,42	1,03	595,9	2,96	1,94
20	Camalau	5	2	0,49	0,31	0,59	0,14	300,4	3,99	1,09
21	Campos	3	2	0,84	0,32	0,48	0,09	224,6	2,31	1,64
22	Cordeiro	5	2	0,42	0,22	0,59	0,22	428,3	6,01	1,14
23	Santo Antonio	5	2	47,16	0,50	1,07	0,64	300,3	4,34	1,29
24	São Domingos	2	3	0,38	0,16	0,43	0,06	102,7	1,10	1,50
25	Livramento	2	2	129,00	1,00	0,94	0,49	18,8	0,58	1,40
26	Namorados	2	3	0,75	0,20	0,88	0,28	87,5	0,79	1,55
27	Soledade	4	2	0,11	0,17	0,67	0,18	46,6	3,21	1,46
28	Antenor Navarro	6	5	0,47	1,16	1,12	0,00	715,7	7,32	1,61
29	Aparecida	7	5	0,70	0,65	1,16	0,00	784,2	11,92	2,02
30	Piancó	7	5	1,42	0,15	1,53	0,00	1.409,3	11,80	1,49
31	Emas	5	3	5,78	0,69	0,85	0,01	636,0	4,46	1,35
32	Patos	6	4	9,34	0,17	1,10	0,00	1.072,2	6,40	1,20
33	Serra Negra do Norte	6	4	0,42	0,16	1,20	0,00	1.248,9	7,29	1,27
34	Fazenda Alagamar	5	1,5	0,57	0,41	0,64	0,00	401,2	6,72	1,59
35	Pedro Velho	5	7	0,15	0,11	0,71	0,00	512,3	7,32	1,64
36	Caraúbas	6	8	0,89	0,29	0,56	0,00	419,2	14,83	1,27
37	Poço de Pedras	6	6	0,83	0,11	0,74	0,00	439,8	9,26	1,49
38	Guarita	7	1	0,83	1,32	0,62	0,00	705,7	20,22	1,43
39	Umburana/Sumé	2	2	1,48	0,24	0,82	0,38	65,0	0,49	1,14
40	Jatobá/Sumé	2	3	0,99	0,24	0,59	0,16	125,0	0,74	1,13
41	Gangorra/Sumé	4	2	0,50	0,38	0,61	0,02	172,6	1,47	1,27

Or – ordem dos cursos de água; Rb – índice de bifurcação; RL – índice dos comprimentos; Ra – índice das áreas; Dd – densidade de drenagem; Ct – coeficiente de torrencialidade; IR – índice de rugosidade; Le – extensão média do escoamento superficial; SIN – sinuosidade do curso de água.

Tabela 5.4: Parâmetros morfológicos das bacias hidrográficas

Ordem	Bacia Hidrográfica	Características do relevo da bacia						
		Imax	Cmed	Lr	lr	Ip	IG	DS
1	Albino	0,41	40,72	4,90	1,94	13,82	18,35	56,6
2	Arrojado	6,79	40,86	15,37	2,39	24,48	18,84	114,3
3	Bartolomeu I	5,80	40,98	13,49	4,41	27,85	12,81	98,8
4	Cachoeira dos Alves	0,20	36,64	24,21	4,57	34,84	13,74	144,4
5	Chupadouro	9,84	41,65	6,94	2,56	14,30	10,38	43,7
6	Cochos	0,15	39,14	13,94	2,97	27,55	20,00	128,8
7	Emas	0,21	40,94	8,83	3,97	31,60	35,09	207,6
8	Eng ^o Arcoverde	0,14	38,68	30,15	4,21	25,63	5,38	60,6
9	Eng ^o Ávidos	0,05	40,96	74,89	13,48	72,98	5,57	176,8
10	Epitácio Pessoa	2,79	40,86	333,92	37,18	182,47	0,78	86,6
11	Farinha	4,35	40,21	89,39	8,37	64,56	5,84	159,8
12	Frutuoso II	0,33	38,73	6,56	2,92	28,90	48,65	212,9
13	Gamela	85,11	35,15	5,62	2,27	26,78	74,88	267,8
14	Jatobá I	28,07	40,85	22,52	4,17	33,15	14,77	143,2
15	Queimadas	0,14	40,26	22,58	5,50	31,87	8,18	91,1
16	Riacho dos Cavalos	0,08	38,49	39,13	4,13	33,36	7,61	96,7
17	Serra Vermelha	0,08	40,57	15,13	3,68	29,14	14,89	111,2
18	Tamanduá I	18,98	37,61	6,07	3,86	17,58	13,43	65,0
19	Vazantes	0,04	40,61	25,29	5,42	41,97	13,61	159,3
20	Camalau	5,30	40,86	79,27	13,30	51,36	2,13	69,1
21	Campos	15,31	37,89	24,16	7,50	51,15	6,57	88,4
22	Cordeiro	7,22	40,93	96,11	17,33	228,18	4,68	190,8
23	Santo Antonio	5,10	38,29	37,50	9,08	935,42	2,90	53,5
24	São Domingos	10,06	39,60	19,41	3,37	259,38	7,12	57,6
25	Livramento	3,79	40,86	9,30	3,98	7,26	9,60	23,7
26	Namorados	17,86	40,86	5,28	2,69	15,44	16,24	61,2
27	Soledade	2,87	40,86	29,81	10,50	34,57	4,15	73,4
28	Antenor Navarro	0,05	48,90	131,99	11,23	68,06	2,80	107,7
29	Aparecida	0,04	48,85	164,74	21,00	90,90	1,83	107,4
30	Piancó	0,03	48,93	160,75	28,69	135,24	2,92	198,6
31	Emas	0,07	48,99	56,56	10,63	67,21	6,89	169,0
32	Patos	0,03	283,22	120,56	13,88	221,54	4,25	173,7
33	Serra Negra do Norte	0,03	48,98	172,71	17,26	104,25	3,50	191,1
34	Fazenda Alagamar	0,05	48,96	100,08	21,25	95,66	2,91	134,1
35	Pedro Velho	0,11	48,96	148,05	23,31	117,69	3,48	204,5
36	Caraúbas	0,02	49,00	158,44	32,14	117,83	1,76	125,4
37	Poço de Pedras	0,02	48,98	151,40	21,23	84,56	1,49	84,3
38	Guarita	0,06	49,00	398,86	44,13	169,84	1,10	146,1
39	Umburana/Sumé	0,27	41,29	3,09	2,57	14,26	24,17	68,1
40	Jatobá/Sumé	0,14	48,28	7,68	3,28	22,72	16,99	85,4
41	Gangorra/Sumé	0,06	48,95	22,86	5,88	34,12	6,90	80,0

Imax – declividade máxima do rio; Cmed – elevação média da bacia; Lr – comprimento do retângulo equivalente; lr – largura do retângulo equivalente; Ip – índice de declividade média da bacia; IG – índice de declividade global; DS – desnível específico.

Tabela 5.5: Tipos de solos encontrados nas bacias hidrográficas

Ordem	Tipo de solo	L ₆₀₀ (mm)
1	Afloramento de Rocha	90
2	Bruno não Cálculo	37
3	Cambisol Eutrófico	30
4	Latosol	10
5	Litólicos	37
6	Planosol Solódico Eutrófico	70
7	Podzólico Vermelho Amarelo Eutrófico	25
8	Regosol	5
9	Solonetz Solodizado	125
10	Solos Aluviais	25
11	Terra Roxa Estruturada	37
12	Vertisol	25

L₆₀₀ - escoamento médio para uma chuva anual de 600 mm (Molle & Cadier, 1992).

De acordo com a Tabela 5.6, as bacias BA5 (açude Chupadouro) e BA8 (açude Eng^o Arcoverde) são as que possuem maior incidência de espelhos de água em suas bacias hidrográficas.

Quanto à ocorrência e distribuição dos solos, as três bacias BA13 (açude Gamela), BA27 (açude Soledade) e BF10 (estação fluviométrica de Poço de Pedras) podem ser consideradas atípicas dentro do conjunto por apresentarem solos da Classe I bem superior às demais, e portanto, maior vocação ao escoamento superficial.

Tabela 5.6: Percentuais de ocorrência das classes de solos nas bacias

Ordem	Bacia Hidrográfica	CLASSE I	CLASSE II	CLASSE III	L ₆₀₀ médio (mm)	PAE
1	Albino	0,00	0,00	0,86	22,20	0,026
2	Arrojado	0,01	0,30	0,69	29,60	0,032
3	Bartolomeu I	0,00	0,98	0,21	41,51	0,034
4	Cachoeira dos Alves	0,00	0,67	0,33	33,04	0,025
5	Chupadouro	0,00	0,38	0,61	29,31	0,063
6	Cochos	0,00	1,00	0,00	37,00	0,020
7	Emas	0,00	1,00	0,00	37,00	0,033
8	Eng ^o Arcoverde	0,03	0,98	0,00	38,96	0,052
9	Eng ^o Ávidos	0,00	0,73	0,28	34,01	0,026
10	Epitácio Pessoa	0,09	0,71	0,13	40,87	0,010
11	Farinha	0,01	0,75	0,11	31,95	0,008
12	Frutuoso II	0,02	1,00	0,05	41,10	0,036
13	Gamela	0,42	0,55	0,03	73,60	0,005
14	Jatobá I	0,00	0,75	0,25	34,40	0,027
15	Queimadas	0,00	1,00	0,00	37,00	0,024
16	Riacho dos Cavalos	0,00	0,35	0,65	29,20	0,038
17	Serra Vermelha	0,00	1,00	0,00	37,00	0,030
18	Tamanduá I	0,00	1,00	0,00	37,00	0,023
19	Vazantes	0,00	0,94	0,06	36,28	0,021
20	Camalau	0,01	0,88	0,12	36,46	0,008
21	Campos	0,23	0,72	0,05	56,64	0,010
22	Cordeiro	0,01	0,89	0,03	35,33	0,003
23	Santo Antonio	0,10	0,76	0,00	41,37	0,004
24	São Domingos	0,00	0,78	0,22	34,36	0,007
25	Livramento	0,00	1,00	0,00	37,00	0,003
26	Namorados	0,00	1,00	0,00	37,00	0,013
27	Soledade	0,43	0,15	0,00	61,45	0,012
28	Antenor Navarro	0,20	0,51	0,30	51,37	0,014
29	Aparecida	0,09	0,55	0,36	40,60	0,014
30	Piancó	0,01	0,85	0,15	36,40	0,011
31	Emas	0,00	0,89	0,08	35,08	0,030
32	Patos	0,03	0,65	0,17	32,65	0,009
33	Serra Negra do Norte	0,02	0,78	0,11	34,51	0,010
34	Fazenda Alagamar	0,13	0,79	0,06	46,38	0,001
35	Pedro Velho	0,15	0,70	0,11	43,55	0,002
36	Caraúbas	0,02	0,85	0,09	36,14	0,006
37	Poço de Pedras	0,30	0,44	0,13	50,08	0,005
38	Guarita	0,11	0,69	0,14	42,06	0,008
39	Umburana/Sumé	0,00	0,93	0,07	36,16	0,005
40	Jatobá/Sumé	0,06	0,19	0,75	31,18	0,002
41	Gangorra/Sumé	0,01	0,80	0,19	35,25	0,003

PAE – percentagem da área da bacia hidrográfica coberta por espelhos de água.

5.3 SELEÇÃO DE BACIAS E DE ATRIBUTOS

5.3.1 Avaliação de cada bacia no conjunto dos dados

Pela análise dos dados obtidos na seção anterior algumas bacias hidrográficas se diferem das demais do conjunto. Bacias íngremes, de áreas muito grandes, com solos rochosos ou com tendência à ocorrência de cheias são indicações da existência de bacias atípicas que podem influenciar os estudos de regionalização.

Pode ser preferível que tais bacias sejam retiradas da amostra para não influenciar negativamente os estudos de regionalização pretendidos.

Os dados das 41 bacias hidrográficas foram submetidos aos algoritmos de agrupamento hierárquico Single Link, Complet Link, e Average Link, através do Software Genstat for Windows Release 10.1, e ao algoritmo Ward, usando o Software Matlab Release 6.5, tendo como medida de similaridade a Distância Euclidiana. Resultaram os quatro dendogramas apresentados nas Figuras 5.1 a 5.4.

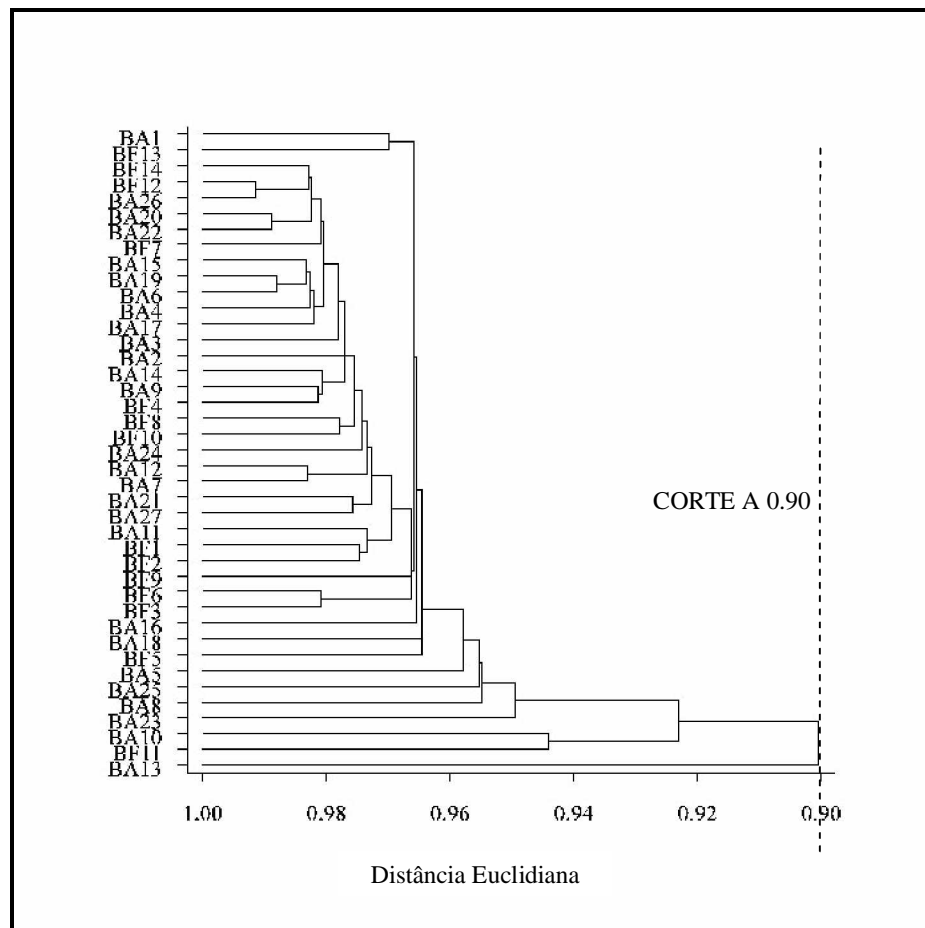


Figura 5.1: Dendrograma obtido por meio do algoritmo Single Link utilizando a medida euclidiana sobre os dados das 41 bacias.

Observando os quatro dendogramas correspondentes ao mesmo conjunto de 41 bacias e à mesma matriz de similaridade, pôde-se observar regularidade na representação da estrutura dos dados através dos algoritmos Complect Link e Average Link, evidenciando sempre a formação de três grupos com maior número de bacias e dois grupos menores com certo desbalanceamento no conjunto dos dados, para um corte na altura 0.90. O dendograma gerado a partir do algoritmo Single Link mostrou degeneração na formação dos grupos, dando a impressão de grupo único, impróprio, portanto para esta avaliação.

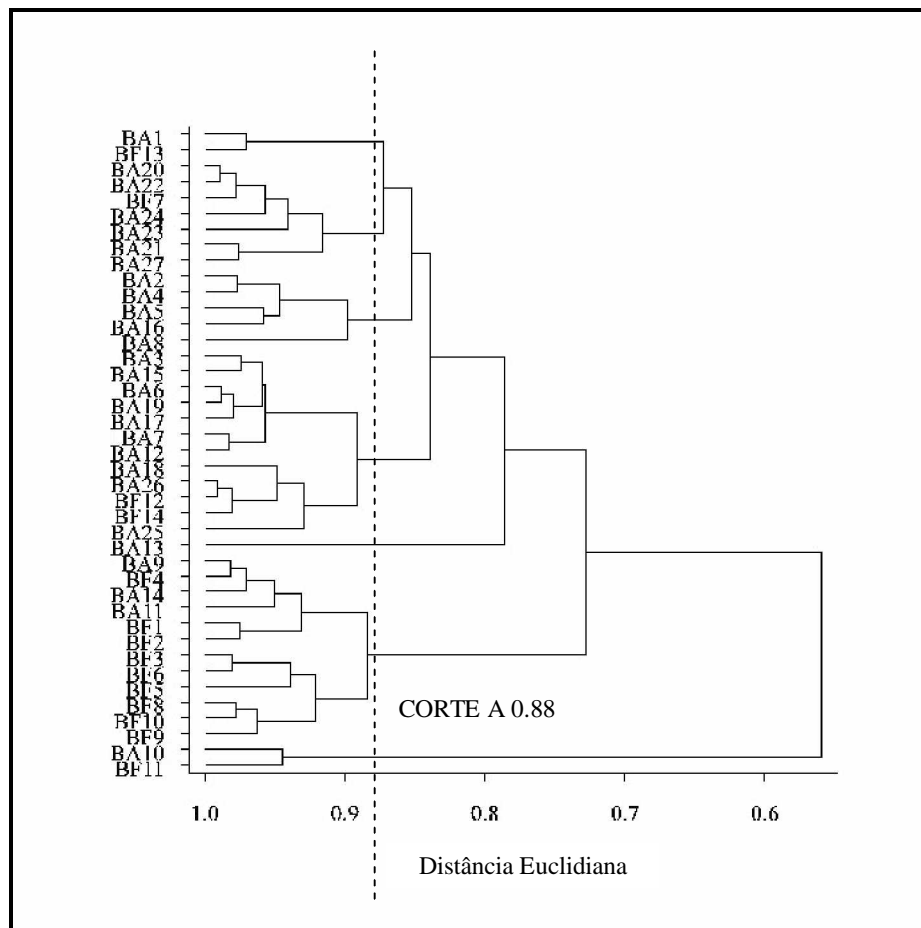


Figura 5.2: Dendograma obtido por meio do algoritmo Complete Link utilizando a medida euclidiana sobre os dados das 41 bacias.

Dessa forma, os dois grupos menores formados pelas bacias (BA10, BA13, BF11) e (BA1, BF13) em três dos quatro algoritmos, foram eliminados gerando um conjunto final de dados, agora com 36 bacias. Os estudos foram repetidos e os resultados revelaram a existência média de três grupos (Figura 5.5), sendo, porém, melhor representado pelo dendograma gerado pelo algoritmo Average Ling, conforme a Figura 5.6. Observa-se claramente maior

estabilidade no dendograma construído com as 36 bacias do que na situação anterior com 41 bacias.

Observa-se na Figura 5.5 que as bacias que integram o Grupo I são as de menor superfície (área média de 75,3 km²) e estão localizadas na sub-região do Sertão paraibano. O Grupo II integra as bacias de superfícies intermediárias (área média de 124,7 km²), localizadas nas sub-regiões do Cariri e Curimataú. O Grupo III, pelas bacias maiores (área média de 2.461,2 km²), distribuídas nas duas sub-regiões.

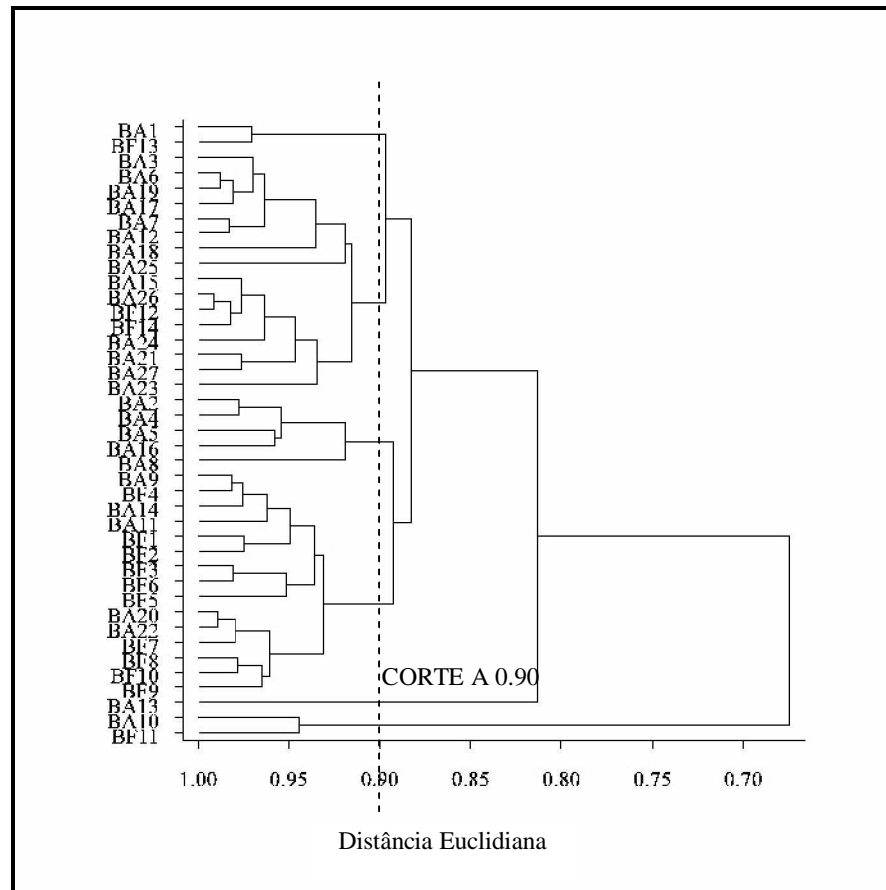


Figura 5.3: Dendograma obtido por meio do algoritmo Average Link utilizando a medida euclidiana sobre os dados das 41 bacias.

Observa-se que destoaram do conjunto de bacias as duas maiores (BA10 e BF11) e três das quatro menores (BA1, BA13 e BF13).

É compreensível que as características físicas e climáticas das bacias maiores sejam influenciadas pela dimensão espacial das mesmas e os valores obtidos não representam a variabilidade comum em bacias de grande porte.

As menores são fortemente comprometidas pela escala dos mapas usados para determiná-las. Além disso, podem estar localizadas em cabeceiras de rios e riachos, determinantes das declividades, da elevação e dos tipos de solos encontrados na bacia.

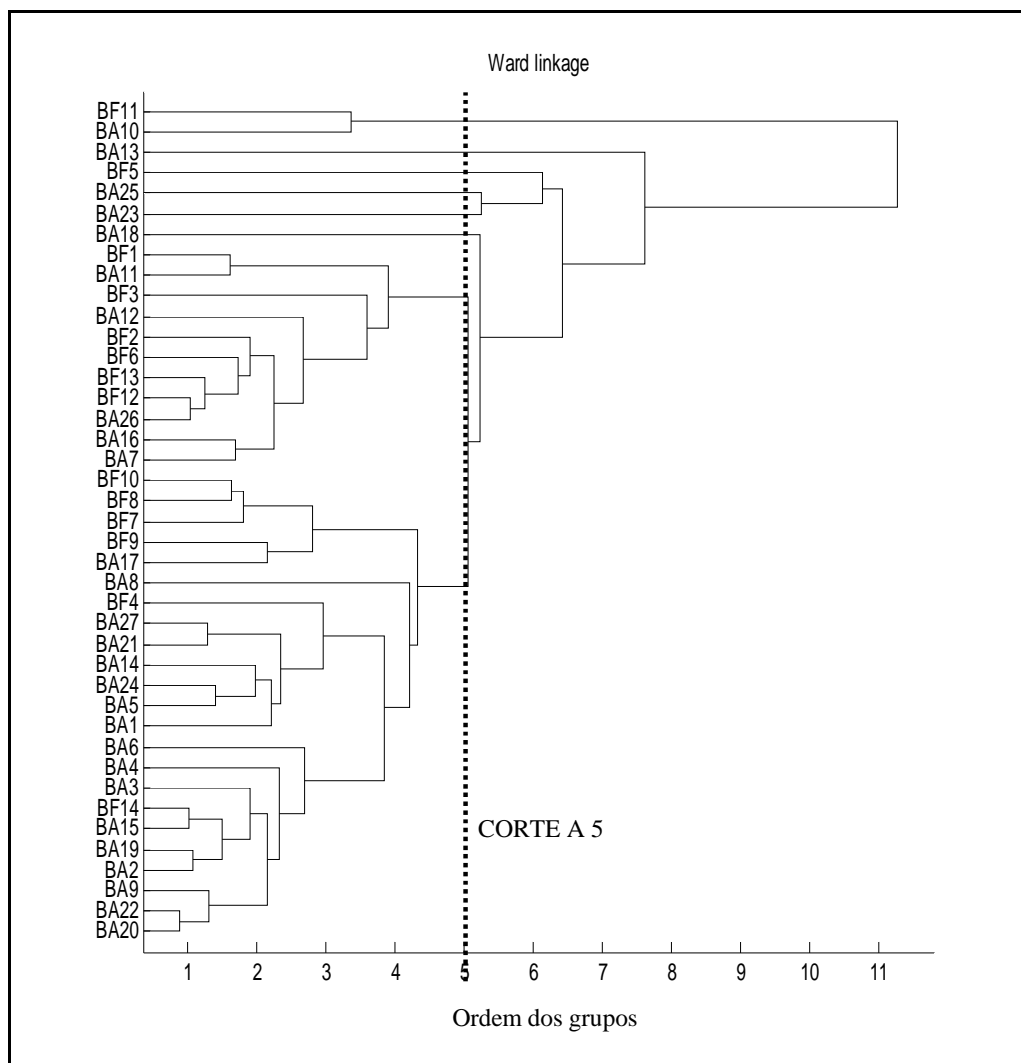


Figura 5.4: Dendrograma obtido por meio do algoritmo Ward utilizando a medida euclidiana sobre os dados das 41 bacias.

5.3.2 Seleção de atributos

A seleção de atributos utilizando o algoritmo CFS foi feita através da ferramenta WEKA - Waikato Environment for Knowledge Analysis (Witten & Frank, 2000), mantendo-se seus parâmetros configurados com os valores padrões e a busca forward. A Tabela 5.7 apresenta os atributos (características das bacias) selecionados pelo algoritmo CFS.

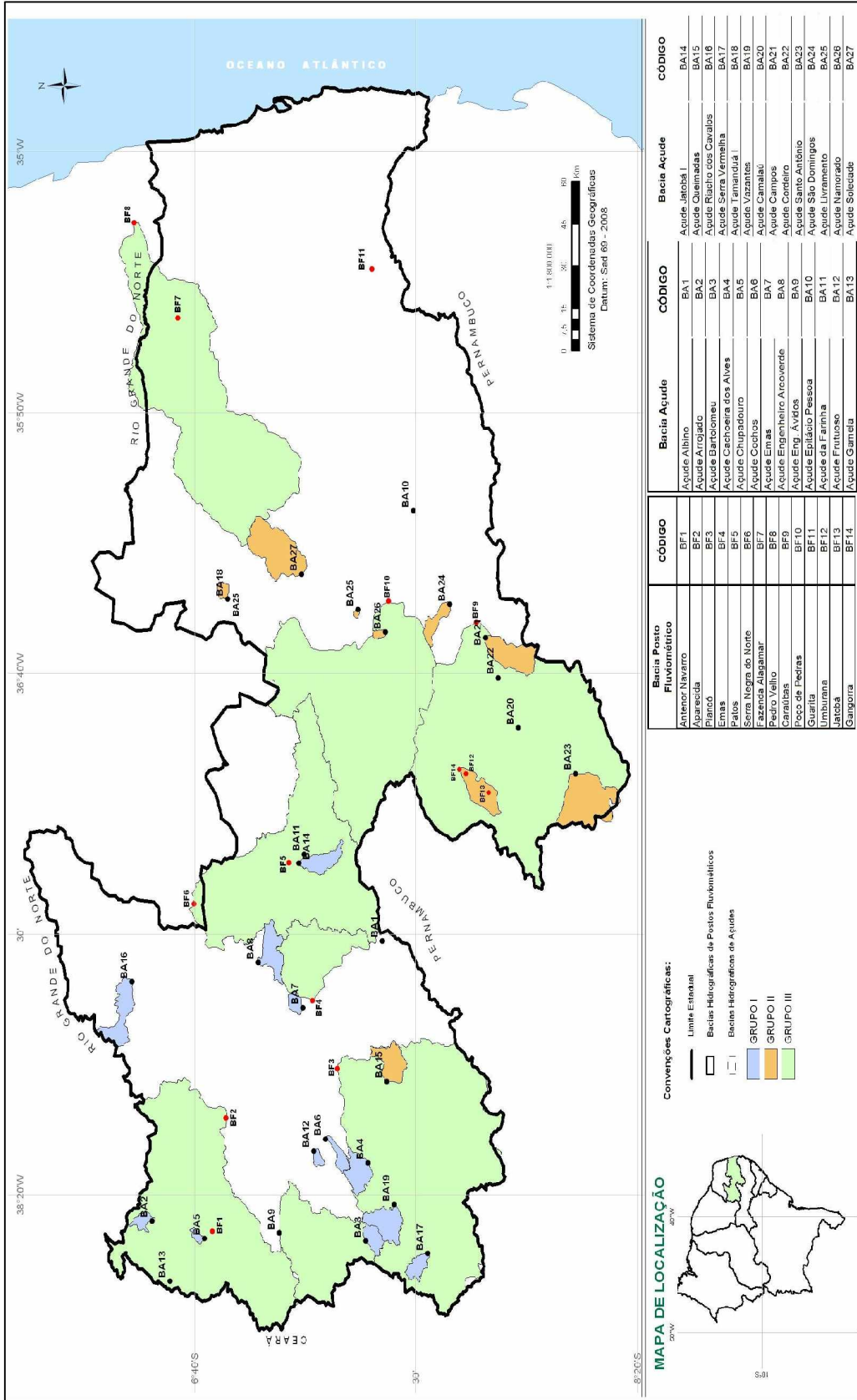


Figura 5.5: Distribuição espacial dos grupos homogêneos de bacias hidrográficas

Tabela 5.7: Atributos selecionados através do algoritmo CFS

Parâmetro do SMAP	Atributos selecionadas	Mérito (M _s)
E2	Pr, Lt, Kf, Rb, RL, Ra, Ct, SIN, Cmed, Ip, DS, P	0,629
STR	Rb, Ra, Ct, Solo 1, Solo 2, PAE	0,643
E1	Rb, Ct, Cmed, Solo 1, PAE	0,506
CINF	Ld, Ra, Ct, Cmed, Solo 1, PAE	0,373*

Pr – perímetro; Lt – comprimento do curso de água principal; Ld – comprimento da rede de drenagem; Kf – fator de forma; Rb – índice de bifurcação; RL – índice dos comprimentos; Ra – índice das áreas; Ct – coeficiente de torrencialidade; SIN – sinuosidade do curso de água; Cmed – elevação média da bacia; Ip – índice de declividade média da bacia; DS – desnível específico; PAE – percentagem da área da bacia hidrográfica coberta por espelhos de água.

*Após a inclusão de E1 entre os atributos de CINF o Mérito subiu para 0,764

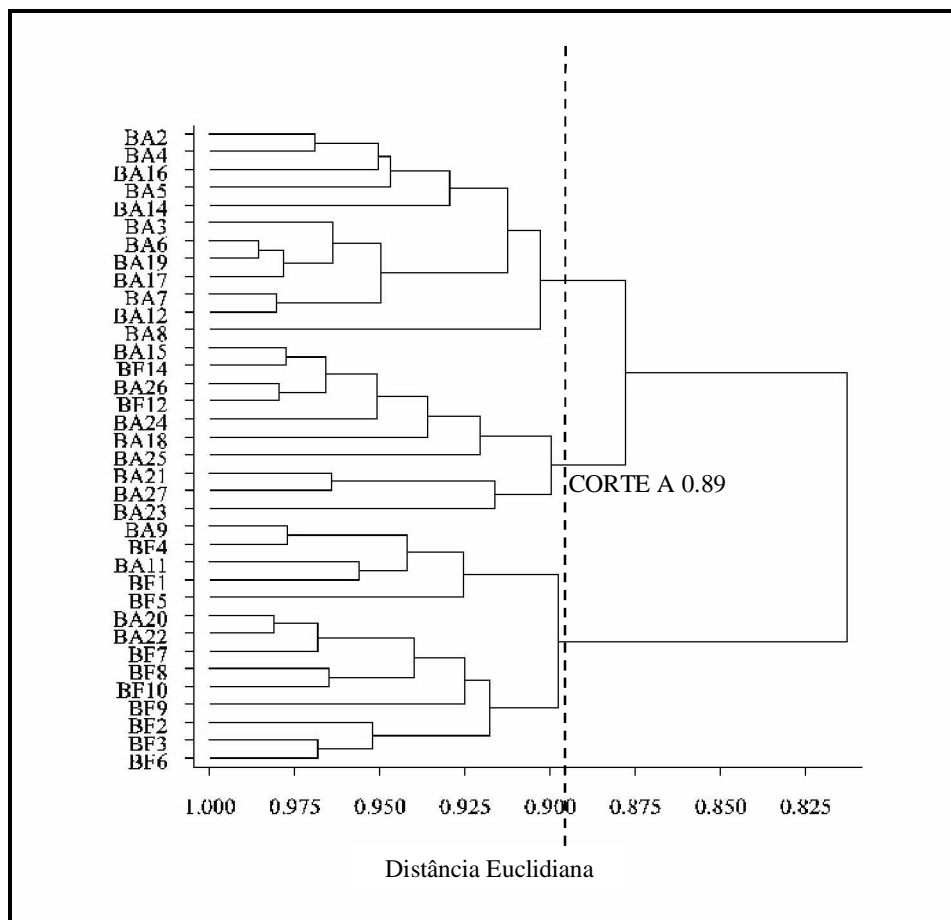


Figura 5.6: Dendrograma obtido por meio do algoritmo Average Link utilizando a medida euclidiana sobre os dados das 36 bacias restantes.

As Figuras 5.7 a 5.10 mostram as relações entre os parâmetros do modelo SMAP e os atributos selecionados para explicar cada um deles, projetados em duas dimensões. Observa-

se baixa relação linear entre eles, conforme representação numérica da Tabela 5.8. Essa situação aponta para o uso de uma metodologia que considere a estrutura não linear dos dados, como as técnicas de redes neurais.

Tabela 5.8: Correlações entre o parâmetro do modelo SMAP e os atributos selecionados

Atributo	Correlações (%)			
	STR	CINF	E1	E2
PAE	15,23	27,15	17,41	
Cmed		-18,65	-16,93	-24,72
Ct	57,01	29,10	49,53	-22,99
DS				-45,17
Ip				20,19
Kf				33,28
Ld		-34,89		
Lt				-40,36
P				-50,49
Pr				-39,89
Ra	-19,12	23,79		22,72
Rb	40,84		-15,19	-44,51
RI				25,84
SIN				-23,79
Solo 1	-16,47	-35,07	-21,50	
Solo 2	25,94			

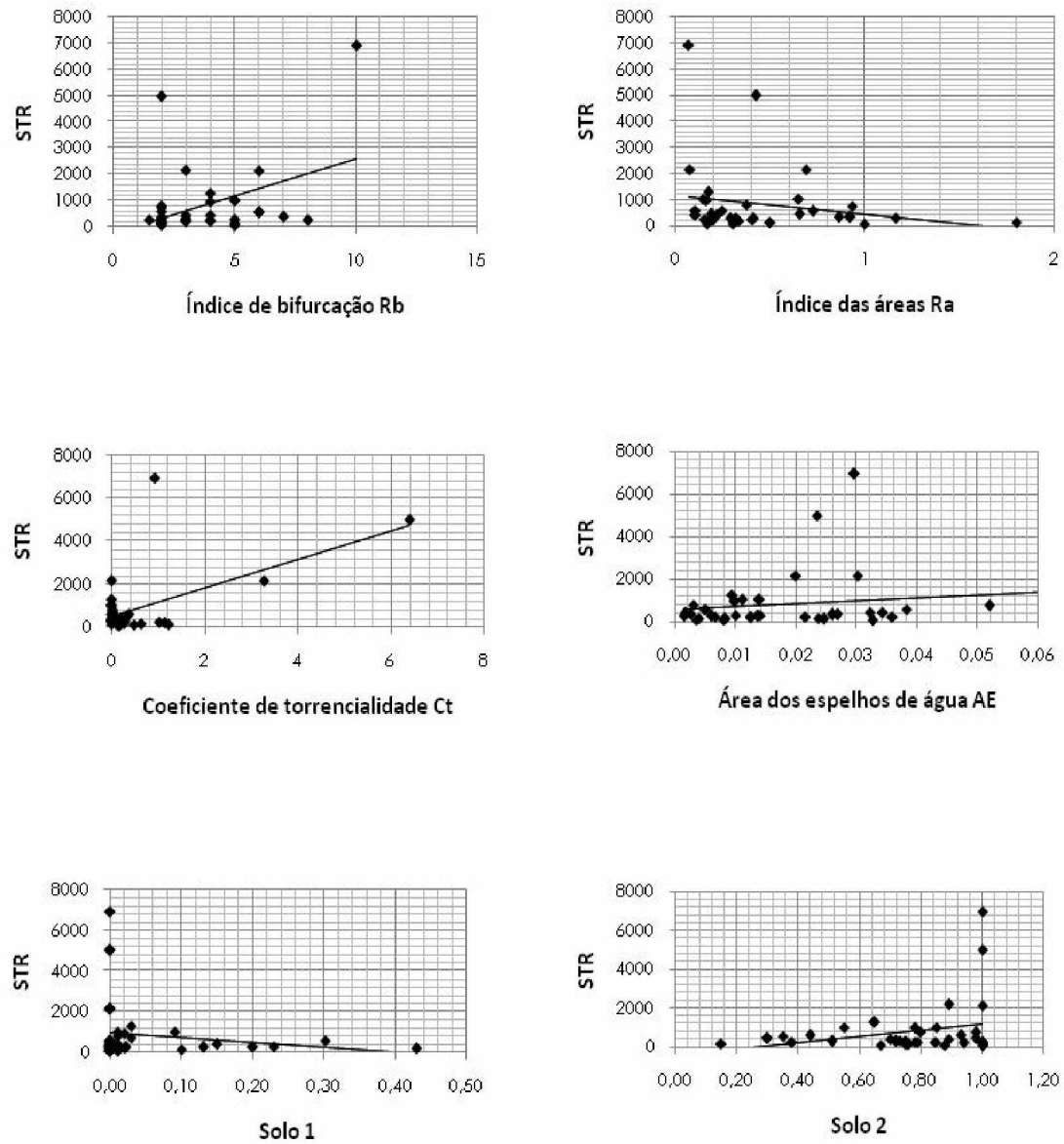


Figura 5.7: Relação entre o parâmetro STR do modelo SMAP e os atributos seleccionados

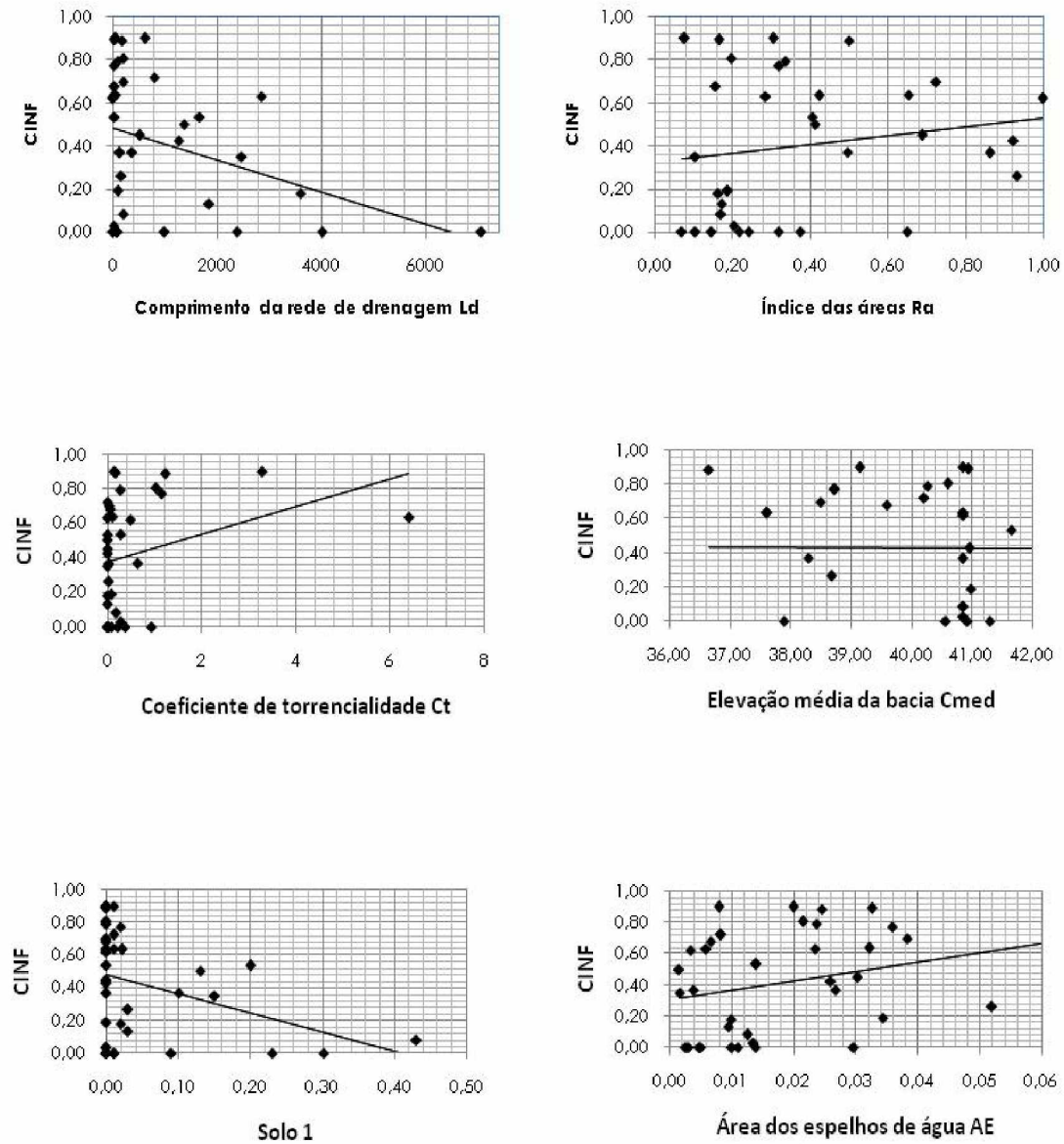


Figura 5.8: Relação entre o parâmetro CINF do modelo SMAP e os atributos selecionados

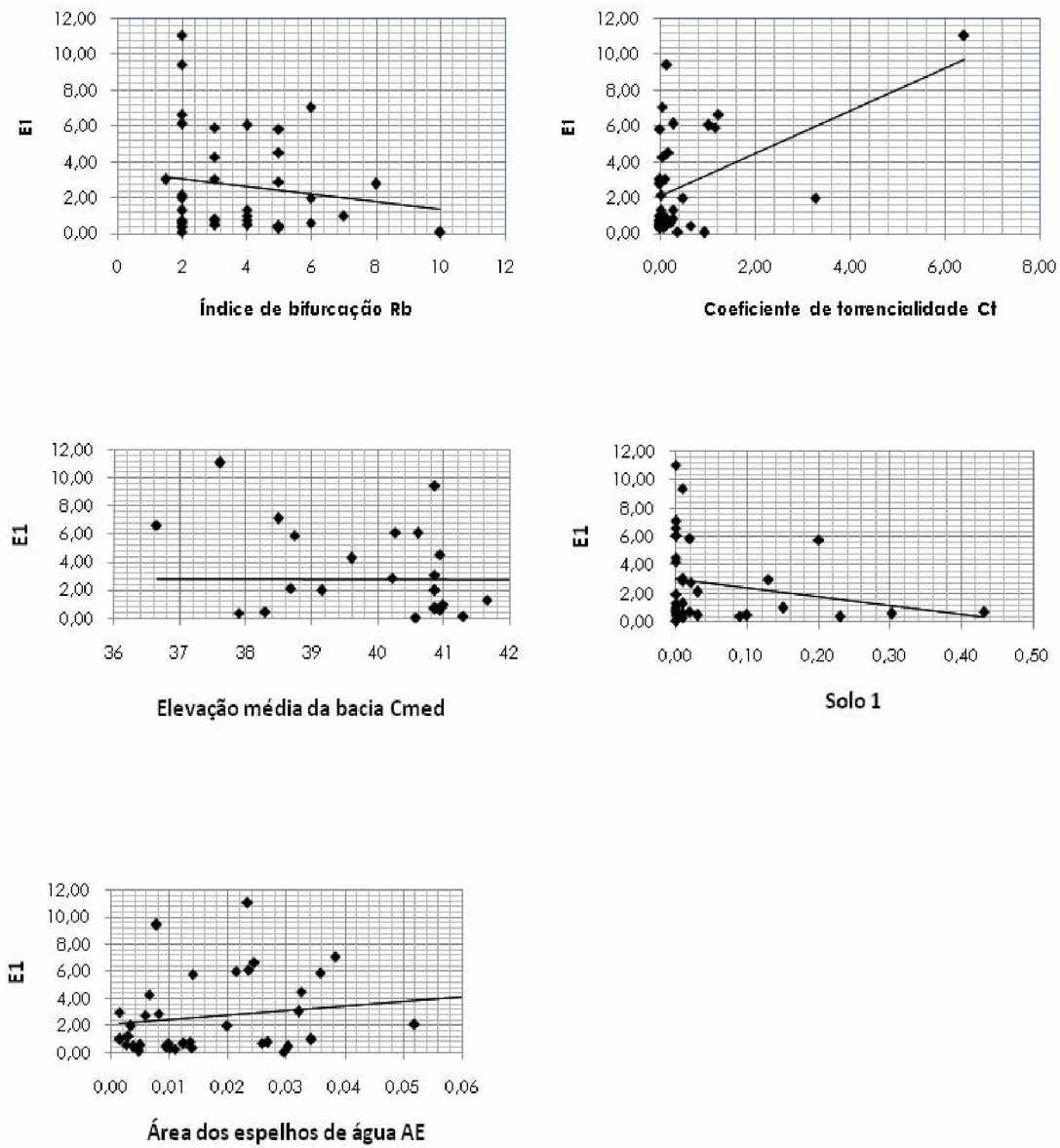


Figura 5.9: Relação entre o parâmetro E1 do modelo SMAP e os atributos seleccionados

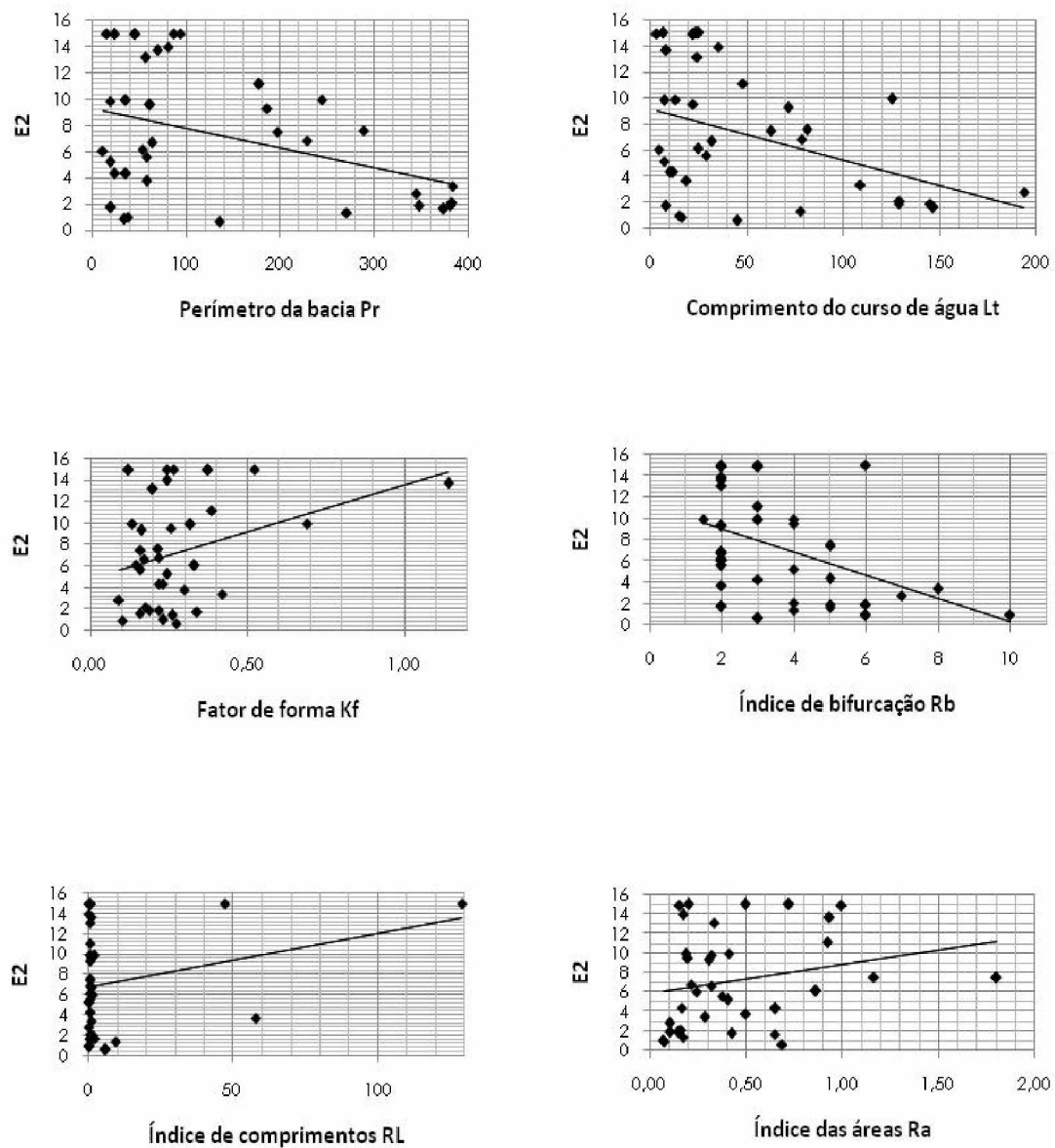


Figura 5.10a: Relação entre o parâmetro E2 do modelo SMAP e os atributos selecionados

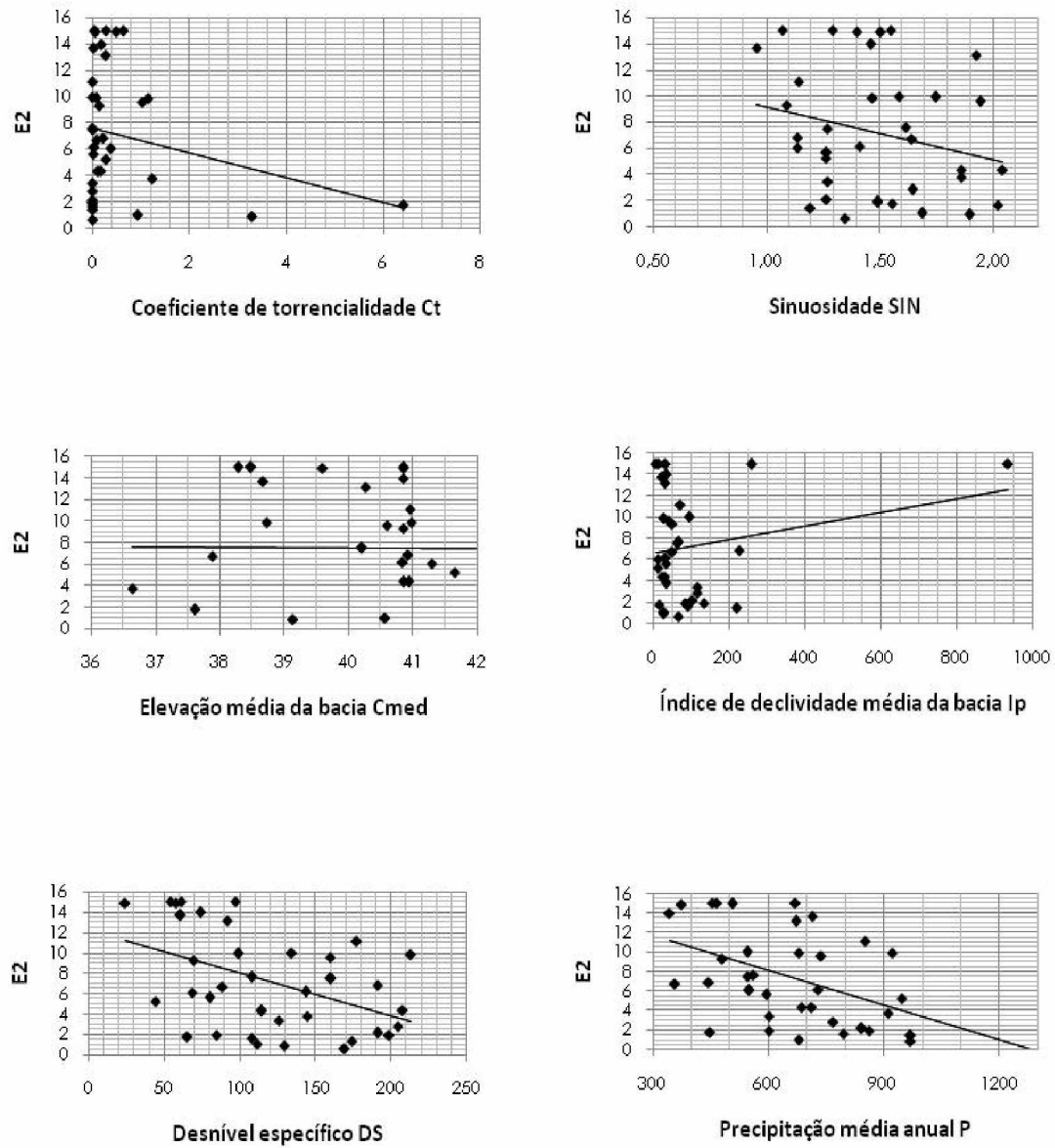


Figura 5.10b: Relação entre o parâmetro E2 do modelo SMAP e os atributos selecionados

5.4 TREINAMENTO DA REDE NEURAL

5.4.1 Definição das arquiteturas da RNA

Foi usado o sistema Qnet for Windows, Versão 2000 (<http://qnetv2k.com>). A arquitetura da rede pôde ser definida, uma para cada dos quatro parâmetros do SMAP, experimentando diversas combinações, variando o número de neurônios, o número de camadas e os tipos de funções de transferências.

A melhor arquitetura para os quatro casos variou conforme apresentado na Tabela 5.9. A topologia da rede neural é do tipo múltiplas camadas com conexões unidirecionais e plenamente conectada. Adotou-se as condições iniciais padrões do sistema Qnet, ou seja, $\eta = 0.01$ (taxa de aprendizado), $\alpha = 0.80$ (fator de momento ou termo de inércia que ajuda a estabilizar o treinamento) e o número de ciclos foi deixado sem limites.

O Qnet permite escolher entre usar os dados originais ou normalizados (valores entre 0 e 1). Após alguns testes usando uma ou outra opção, concluiu-se que a normalização dos dados permite obter melhores resultados finais. A expressão seguinte assegura que uma variável x_i , pertencente a uma série qualquer de dados, tenha seu valor normalizado, xn_i , entre 0 e 1. Nesta expressão, x_{\min} e x_{\max} são os valores mínimos e máximos da série, respectivamente.

$$xn_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad 5.1$$

A normalização dos dados diminui a influência da escala dos valores, assegurando que eles recebam igual tratamento durante o processo de treinamento, e melhora o desempenho do algoritmo de treinamento.

Tabela 5.9: Arquitetura da rede neural para cada parâmetro do SMAP

CAMADA	NÚMERO DE NEURÔNIOS				FUNÇÃO DE TRANSFERÊNCIA			
	E2	STR	E1	CINF	E2	STR	E1	CINF
Entrada	12	7	5	7	-	-	-	-
Escondida 1	5	3	3	4	gaussiana	gaussiana	hiperbólica secante	gaussiana
Escondida 2	-	2	2	-	-	gaussiana	hiperbólica secante	-
Saída	1	1	1	1	gaussiana	sigmóide	sigmóide	gaussiana

As funções de ativação ou de transferência nomeadas na Tabela 5.9 caracterizam-se pelas seguintes equações:

$$\text{Sigmóide: } y = f(s) = \frac{1}{1 + e^{-s}} \quad 5.2$$

$$\text{Gaussiana: } f(s) = e^{-s \cdot s} \quad 5.3$$

$$\text{Hiperbólica secante: } f(s) = \text{Sech}(s) \quad 5.4$$

$$\text{onde: } s \text{ é a função de combinação dada por: } s = \sum_{i=1}^n w_i x_i \quad 5.5$$

5.4.2 Treinamento da RNA sobre o parâmetro E2

Selecionada a arquitetura da rede neural que melhor representa a relação entre as características da bacia (Tabela 5.7) e os parâmetros do modelo SMAP (Tabela 5.1), ilustrada na Figura 5.11, definiu-se como ponto de parada aquele que, simultaneamente, reproduza o conjunto de validação cruzada (bacia-meta) e tenha estabilizado a taxa de decaimento do erro do treinamento, conforme item 4.7.3. Para a inspeção da qualidade do treinamento da rede e respectiva verificação, determinou-se o erro relativo entre os valores observados e calculados, a partir da Equação 5.6.

$$ERR = \frac{|E2_o - E2_c|}{E2_o} \quad 5.6$$

Para efeito de qualificação dos resultados obtidos, em função dos erros relativos, considerou-se quatro classes de erros, conforme se segue:

- a) ótimo : $ERR < 0,10$
- b) bom : $0,10 \leq ERR < 0,30$
- c) regular: $0,30 \leq ERR < 0,50$
- d) ruim : $ERR \geq 0,50$

A Tabela 5.10 apresenta os resultados do treinamento da rede neural cujos erros, na avaliação da bacia-meta são quase todos inferiores a 0,20, à excessão da bacia BA6, cujo resultado ($ERR=0,5225$) pode ser considerado no limite inferior da classificação ruim. É claro

que aí estão apenas os resultados referentes às bacias-meta mais similares às bacias separadas para verificação.

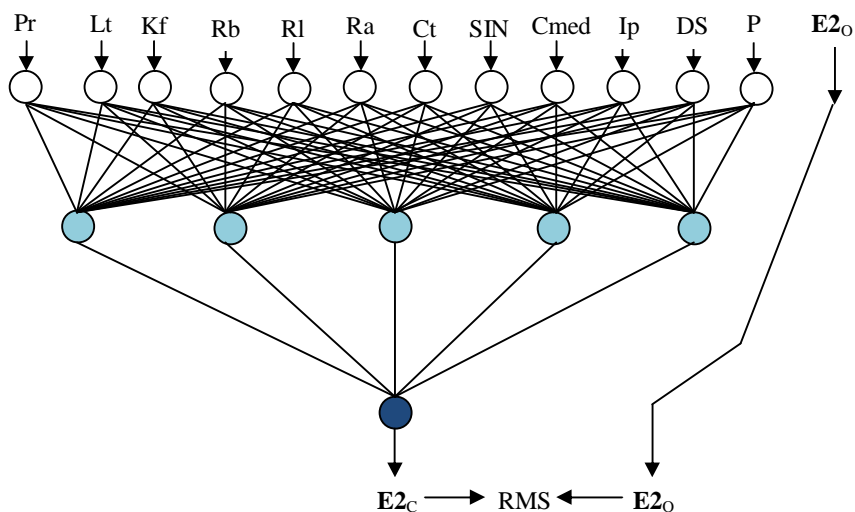


Figura 5.11: Arquitetura da rede neural – parâmetro E2

A fase de verificação da rede neural foi conduzida a partir da observação da matriz de similaridades e do respectivo dendograma (Figura 5.12), determinados com o uso dos dados das 36 bacias hidrográficas. É identificada, no conjunto das bacias-meta, a bacia mais similar àquela que foi separada para a validação, caracterizando qual modelo deve ser usado para reproduzir o parâmetro E2 da verificação. Isto foi feito para as 36 bacias, omitidas alternadamente, resultando nos valores apresentados na Tabela 5.11.

Tabela 5.10: Parâmetro E2 do modelo SMAP: fase de treinamento da RNA

Ordem	Bacia omitida para verificação	Bacia-meta mais similar	Valores de E2		Erro relativo ERR	R ² (%) Treinamento
			Obtido na calibração do SMAP	Obtido no treinamento da RNA		
1	BA2	BA15	13,126	13,130	0,0003	100,0
2	BA3	BA15	13,126	13,132	0,0005	100,0
3	BA4	BA19	9,564	9,560	0,0004	100,0
4	BA5	BA16	15,000	15,000	0,0000	100,0
5	BA6	BA17	1,012	1,010	0,0020	100,0
6	BA7	BA19	9,564	9,530	0,0036	100,0
7	BA8	BA23	15,000	15,059	0,0039	100,0
8	BA9	BA14	6,146	6,151	0,0008	100,0
9	BA11	BF1	7,587	7,670	0,0109	100,0
10	BA12	BA7	4,341	4,450	0,0251	100,0
11	BA14	BA9	11,114	11,091	0,0021	100,0
12	BA15	BA2	4,319	4,315	0,0009	100,0
13	BA16	BA5	5,215	5,219	0,0008	100,0
14	BA17	BA6	0,890	0,425	0,5225	100,0
15	BA18	BA26	15,000	14,912	0,0059	100,0
16	BA19	BA7	4,341	4,340	0,0002	100,0
17	BA20	BF14	5,616	5,630	0,0025	100,0
18	BA21	BA27	13,954	13,949	0,0004	100,0
19	BA22	BF7	9,948	9,950	0,0002	100,0
20	BA23	BA25	14,928	14,872	0,0038	100,0
21	BA24	BA26	15,000	15,000	0,0000	100,0
22	BA25	BA23	15,000	14,994	0,0004	99,9
23	BA26	BA27	13,954	13,947	0,0005	100,0
24	BA27	BA21	6,682	6,671	0,0016	100,0
25	BF1	BA11	7,473	7,453	0,0027	100,0
26	BF2	BF10	1,885	1,882	0,0016	100,0
27	BF3	BF6	2,120	2,120	0,0000	100,0
28	BF4	BA9	11,114	10,772	0,0308	100,0
29	BF5	BF6	2,120	2,040	0,0377	100,0
30	BF6	BF3	1,915	1,920	0,0026	100,0
31	BF7	BA22	6,839	6,697	0,0208	100,0
32	BF8	BF3	1,915	1,689	0,1180	100,0
33	BF9	BF10	1,885	1,890	0,0027	100,0
34	BF10	BF9	3,401	3,400	0,0003	100,0
35	BF12	BF14	5,616	5,620	0,0007	100,0
36	BF14	BF12	6,053	6,051	0,0003	100,0

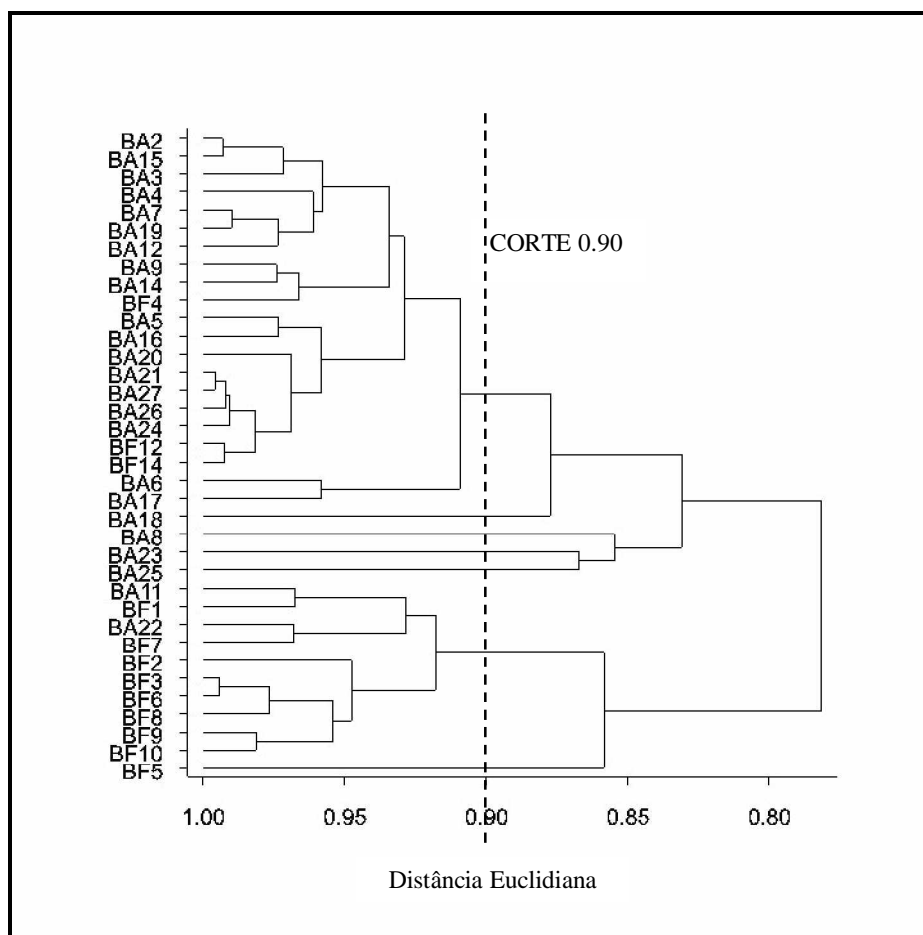


Figura 5.12: **Dendrograma obtido para o parâmetro E2 por meio do algoritmo Average Link utilizando a medida euclidiana.**

Considerando, a título de exemplo, a bacia BA2 como a primeira a ser separada para a fase de verificação, a bacia-meta passa a ser a BA15 e, vice-versa. Para a BA3, o dendrograma não permite a identificação visual, requerendo os valores da matriz de similaridade (Anexo B.1) para decidir qual, entre BA2 e BA15, seria a bacia-meta; no caso a bacia BA15, apresenta maior similaridade com BA3.

A Tabela 5.11 apresenta os resultados obtidos para as 36 bacias usadas no processo de verificação. Como se vê, das 36 bacias usadas nesta fase, apenas três apresentaram resultados ruins: BA18, BA25 e BF5 e outras duas, a BA8 e a BA17, resultados com e regular, respectivamente. Nota-se que, a bacia BA8 não

Para um corte no dendrograma da Figura 5.12 na altura 0.90, verifica-se que as bacias BA8, BA18, BA23, BA25 e BF5 destoam das demais e ficam isoladas, evidenciando certa impossibilidade de serem representadas por algum dos modelos construídos ou se constituir numa bacia-meta dentro do conjunto de bacias estudadas. Portanto, para efeito de determinação do parâmetro E2 de tais bacias, nenhuma das 36 redes neurais regionais

disponíveis poderia ser usada. Quanto a bacia BA17, mesmo sendo similar à bacia BA6, no conjunto as duas ficam relativamente isoladas, devendo a rede neural, para estes casos, ser usada com cautela.

Os erros cometidos ao adotar-se, para a bacia omitida, o valor do parâmetro E2 da bacia mais similar, são muito grandes e inviabilizam o uso direto dessa possibilidade.

Tabela 5.11: Parâmetro E2 do modelo SMAP: fase de verificação

Ordem	Bacia-meta mais similar	Bacia omitida para verificação	Valores de E2			Erro relativo ERR	
			Obtido na calibração do SMAP	bacia-meta	Obtido na verificação da RNA	ao adotar E2 da bacia-meta	ao adotar E2 obtido pela RNA
1	BA15	BA2	4,319	13,126	4,334	2,0391	0,0035
2	BA15	BA3	9,908	13,126	9,915	0,3248	0,0007
3	BA19	BA4	3,752	9,564	3,748	1,5490	0,0011
4	BA16	BA5	5,215	15,000	5,209	1,8763	0,0011
5	BA17	BA6	0,890	1,012	0,866	0,1371	0,0270
6	BA19	BA7	4,341	9,564	4,322	1,2032	0,0044
7	BA23	BA8	13,676	15,000	15,351	0,0968	0,1225
8	BA14	BA9	11,114	6,146	11,089	0,4470	0,0022
9	BF1	BA11	7,473	7,587	7,607	0,0153	0,0179
10	BA7	BA12	9,846	4,341	10,118	0,5591	0,0276
11	BA9	BA14	6,146	11,114	6,150	0,8083	0,0006
12	BA2	BA15	13,126	4,319	13,094	0,6710	0,0024
13	BA5	BA16	15,000	5,215	14,945	0,6523	0,0037
14	BA6	BA17	1,012	0,890	0,522	0,9991	0,4842
15	BA26	BA18	1,752	15,000	6,244	7,5616	2,5639
16	BA7	BA19	9,564	4,341	9,496	0,5461	0,0071
17	BF14	BA20	9,298	5,616	9,325	0,3960	0,0029
18	BA27	BA21	6,682	13,954	6,684	1,0883	0,0003
19	BF7	BA22	6,839	9,948	6,816	0,4546	0,0034
20	BA25	BA23	15,000	14,928	16,639	0,0048	0,1093
21	BA26	BA24	14,915	15,000	14,905	0,0057	0,0007
22	BA23	BA25	14,928	15,000	1,691	0,0048	0,8868
23	BA27	BA26	15,000	13,954	15,000	0,0697	0,0000
24	BA21	BA27	13,954	6,682	13,942	0,5211	0,0009
25	BA11	BF1	7,587	7,473	7,607	0,0150	0,0026
26	BF10	BF2	1,640	1,885	1,644	0,1494	0,0002
27	BF6	BF3	1,915	2,120	1,923	0,1070	0,0042
28	BA9	BF4	0,635	11,114	0,622	16,5026	0,0205
29	BF6	BF5	1,389	2,120	-2,402	0,5263	2,7293
30	BF3	BF6	2,120	1,915	2,116	0,0967	0,0019
31	BA22	BF7	9,948	6,839	10,108	0,3125	0,0161
32	BF3	BF8	2,811	1,915	2,488	0,3187	0,1149
33	BF10	BF9	3,401	1,885	3,410	0,4458	0,0026
34	BF9	BF10	1,885	3,401	1,891	0,8042	0,0032
35	BF14	BF12	6,053	5,616	6,050	0,0722	0,0005
36	BF12	BF14	5,616	6,053	5,620	0,0778	0,0007

5.4.3 Treinamento da RNA sobre o parâmetro STR

A interdependência entre os parâmetros é uma característica marcante nos modelos chuva-vazão. No caso do modelo SMAP, os parâmetros STR com E2 e, E1 com CINF, têm significativa correlação entre si, conforme mostrado na Tabela 5.12.

Tabela 5.12: Matriz de correlações entre os parâmetros do modelo SMAP

E2	1,000			
STR	-0,464	1,000		
E1	0,165	0,035	1,000	
CINF	0,161	-0,178	0,737	1,000
	E2	STR	E1	CINF

As respostas da rede neural para o parâmetro STR melhorou substancialmente com a introdução dos valores do parâmetro E2 entre as características explicativas, recurso também utilizado por Heuvelmans et al (2006). Desta maneira, a arquitetura final ficou definida conforme a Figura 5.13.

A Tabela 5.13 apresenta os resultados do treinamento da rede neural cujos erros, tanto no treinamento quanto na avaliação da bacia-meta, são bastante reduzidos. Igualmente ao caso do parâmetro E2, foram apresentados somente os resultados referentes às bacias-meta mais similares às bacias separadas para verificação.

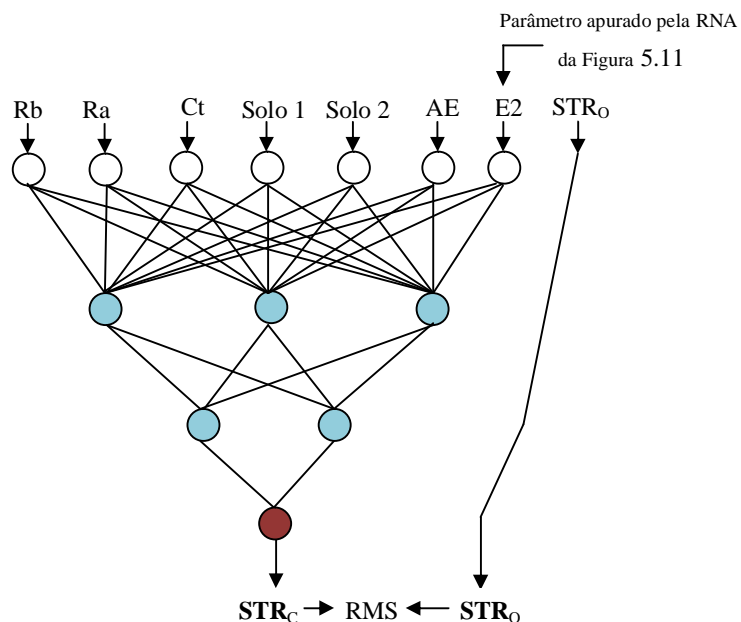


Figura 5.13: Arquitetura da rede neural – parâmetro STR

Para a fase de verificação, a identificação da bacia-meta mais similar se deu por meio da observação do dendograma mostrado na Figura 5.14, construído a partir da matriz de similaridades (Anexo B.2), já incluídos os dados do parâmetro E2 entre as características das bacias hidrográficas.

Tabela 5.13: Parâmetro STR do modelo SMAP: fase de treinamento da RNA

Ordem	Bacia omitida para verificação	Bacia-meta mais similar	Valores de STR		Erro relativo ERR	R ² (%) Treinamento
			Obtido na calibração do SMAP	Obtido no treinamento da RNA		
1	BA2	BA5	261,87	261,90	0,0001	99,91
2	BA3	BA12	212,26	212,18	0,0004	99,99
3	BA4	BA14	328,37	328,39	0,0001	99,99
4	BA5	BA2	443,67	443,42	0,0006	99,99
5	BA6	BA17	6.941,06	6.947,26	0,0009	99,97
6	BA7	BA12	212,26	212,30	0,0002	99,99
7	BA8	BA9	324,54	276,19	0,1490	99,99
8	BA9	BA8	735,88	735,08	0,0011	99,99
9	BA11	BF1	285,40	285,35	0,0002	99,99
10	BA12	BA3	441,39	449,16	0,0176	99,99
11	BA14	BA4	121,07	113,80	0,0600	99,99
12	BA15	BA26	265,97	261,57	0,0165	99,99
13	BA16	BA2	443,67	439,90	0,0085	99,99
14	BA17	BA6	2.141,65	2.141,92	0,0001	99,98
15	BA18	BA6	2.141,65	1.336,54	0,3759	99,99
16	BA19	BA3	441,39	449,40	0,0181	99,99
17	BA20	BA22	389,38	381,91	0,0192	99,99
18	BA21	BF7	265,54	112,15	0,5777	99,89
19	BA22	BF12	590,07	587,75	0,0039	99,99
20	BA23	BA24	247,95	251,56	0,0146	99,99
21	BA24	BA23	148,74	181,16	0,2180	99,99
22	BA25	BA23	148,74	148,66	0,0005	99,99
23	BA26	BA12	212,26	125,54	0,4086	99,99
24	BA27	BA21	283,07	330,81	0,1687	99,92
25	BF1	BA11	127,97	127,24	0,0057	99,92
26	BF2	BF5	1.281,13	1.252,92	0,0220	99,99
27	BF3	BF6	957,10	955,01	0,0022	99,99
28	BF4	BA4	121,07	118,88	0,0181	99,98
29	BF5	BF6	957,10	957,60	0,0005	99,73
30	BF6	BF3	1.028,2	820,76	0,2018	99,91
31	BF7	BA21	283,07	338,23	0,1949	99,92
32	BF8	BF9	267,66	262,73	0,0184	99,99
33	BF9	BF8	398,61	398,66	0,0001	99,99
34	BF10	BF8	398,61	408,58	0,0250	99,99
35	BF12	BA22	389,38	380,15	0,0237	99,99
36	BF14	BA22	389,38	389,35	0,0001	99,99

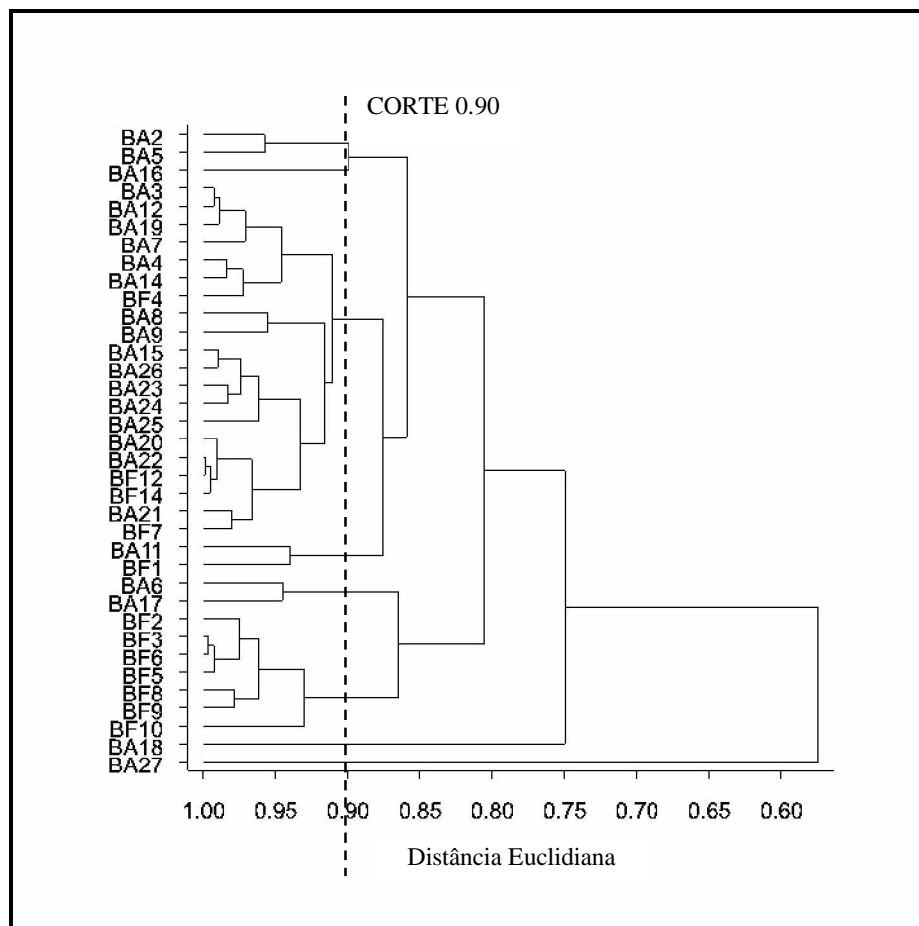


Figura 5.14: **Dendrograma obtido para o parâmetro STR por meio do algoritmo Average Link utilizando a medida euclidiana.**

Na fase de treinamento da rede neural uma (BA6) apresentou resultado ruim e duas (BA12, BF7), regular. As bacias (BA9, BA21, BA23 e BF3) se apresentam com resultados bons e as 29 restantes com resultados ótimos. Entre as ruins, apenas a BA6 ofereceu resultado ruim na fase de verificação. É facilmente identificável esta situação no dendrograma da Figura 5.14, no qual, ao aplicar um corte à altura 0.90, observa-se que esta bacia não se aninha com as demais.

Por outro lado, para algumas bacias (BA2, BA5 e BF1), ótimos resultados na fase de treinamento não redundaram em sucesso na verificação (Tabela 5.14), pelo mesmo motivo de desequilíbrio delas com as demais, conforme mostrado no dendrograma da Figura 5.14.

Destaca-se ainda o fato de outras bacias (BA9, BA21, BA23 e BF3), apresentarem resultados medianos no treinamento, mesmo estando, parte delas aninhadas ao conjunto. Pode ser atribuído, particularmente, aos resultados advindos da calibração do modelo SMAP, ou

estão sendo influenciados pela presença de bacias sem grau de semelhança razoável com o grupo.

Assim como ocorreu com o parâmetro E2, o elevado índice de erros acima de 0,50, em mais de 71% dos casos, inviabiliza a adoção direta dos valores de STR da bacia mais similar.

Tabela 5.14: Parâmetro STR do modelo SMAP: fase de verificação

Ordem	Bacia-meta mais similar	Bacia omitida para verificação	Valores de STR			Erro relativo ERR	
			Obtido na calibração do SMAP	bacia-meta	Obtido na verificação da RNA	ao adotar STR da bacia-meta	ao adotar STR obtido pela RNA
1	BA5	BA2	443,67	261,87	107,84	0.4098	0,7567
2	BA12	BA3	441,39	212,26	470,88	0.5191	0,0668
3	BA14	BA4	121,07	328,37	111,95	1.7122	0,0753
4	BA2	BA5	261,87	443,67	1.438,85	0.6942	4,4945
5	BA17	BA6	2.141,65	6.941,06	2.153,06	2.2410	0,0053
6	BA12	BA7	61,95	212,26	68,47	2.4263	0,1052
7	BA9	BA8	735,88	324,54	735,10	0.5590	0,0011
8	BA8	BA9	324,54	735,88	276,20	1.2675	0,1489
9	BF1	BA11	127,97	285,40	202,48	1.2302	0,5822
10	BA3	BA12	212,26	441,39	198,91	1.0795	0,0629
11	BA4	BA14	328,37	121,07	329,87	0.6313	0,0046
12	BA26	BA15	169,35	265,97	133,60	0.5705	0,2111
13	BA2	BA16	554,73	443,67	550,22	0.2002	0,0081
14	BA6	BA17	6.941,06	2.141,65	5.128,70	0.6915	0,2611
15	BA6	BA18	5.000,00	2.141,65	362,64	0.5717	0,9274
16	BA3	BA19	221,00	441,39	241,38	0.9972	0,0922
17	BA22	BA20	77,72	389,38	125,79	4.0100	0,6185
18	BF7	BA21	283,07	265,54	282,43	0.0619	0,0023
19	BF12	BA22	389,38	590,07	380,19	0.5154	0,0236
20	BA24	BA23	148,74	247,95	181,16	0.6670	0,2180
21	BA23	BA24	247,95	148,74	251,52	0.4001	0,0144
22	BA23	BA25	91,00	148,74	337,01	0.6345	2,7034
23	BA12	BA26	265,97	212,26	247,06	0.2019	0,0711
24	BA21	BA27	185,20	283,07	69,49	0.5285	0,6248
25	BA11	BF1	285,40	127,97	186,25	0.5516	0,3474
26	BF5	BF2	1.012,12	1.281,13	1.004,26	0.2658	0,0078
27	BF6	BF3	1.028,20	957,10	1.042,38	0.0691	0,0138
28	BA4	BF4	2.163,14	121,07	2.072,68	0.9440	0,0418
29	BF6	BF5	1.281,13	957,10	1.014,60	0.2529	0,2080
30	BF3	BF6	957,10	1.028,2	961,00	0.0743	0,0041
31	BA21	BF7	265,54	283,07	207,22	0.0660	0,2196
32	BF9	BF8	398,61	267,66	401,24	0.3285	0,0066
33	BF8	BF9	267,66	398,61	261,16	0.4892	0,0243
34	BF8	BF10	586,67	398,61	616,68	0.3206	0,0511
35	BA22	BF12	590,07	389,38	587,64	0.3401	0,0041
36	BA22	BF14	803,42	389,38	809,85	0.5153	0,0080

5.4.4 Treinamento da RNA sobre o parâmetro E1

A escolha da arquitetura da rede neural que melhor representa a relação entre as características da bacia e o parâmetro E1 do modelo SMAP, feita por exaustivas tentativas, resultou naquela ilustrada na Figura 5.15.

Face a interdependência existente entre E1 e CINF, a definição da rede neural só pôde ser realizada de forma conjunta, após identificar qual dos dois parâmetros é mais dependente um do outro. Pôde-se definir com relativa facilidade a estrutura de rede neural para o parâmetro E1, sem a inclusão de CINF entre suas características. Contrariamente, o parâmetro CINF, dependeu fundamentalmente da inclusão de E1 entre as características das bacias, para se ter uma estrutura de rede neural apropriada.

Pela análise do dendograma da Figura 5.16 e da respectiva matriz de similaridades (Anexo B.3) dá para presumir certa dificuldade no processo de treinamento da rede neural, haja visto o mesmo não demonstrar homogeneidade na formação de grupos de bacias. Um corte na altura 0.90 do dendograma revela o desbalanceamento do conjunto como um todo.

Uma possibilidade para melhorar a formação de grupos mais homogêneos seria a exclusão das bacias mais atípicas (BA18, BA27 e BF5) e refazer o dendograma bem como todo o procedimento de treinamento da RNA. Entretanto, estas bacias são importantes para explicar a variabilidade de outros parâmetros do SMAP.

A Tabela 5.15 apresenta os resultados do treinamento da rede neural. Em geral obteve-se bons resultados na fase de treinamento, porém, em alguns casos (BA21, BF4, BF6 e BF12) a reprodução do parâmetro da bacia-meta foi ruim e, regular, para a bacia BF7.

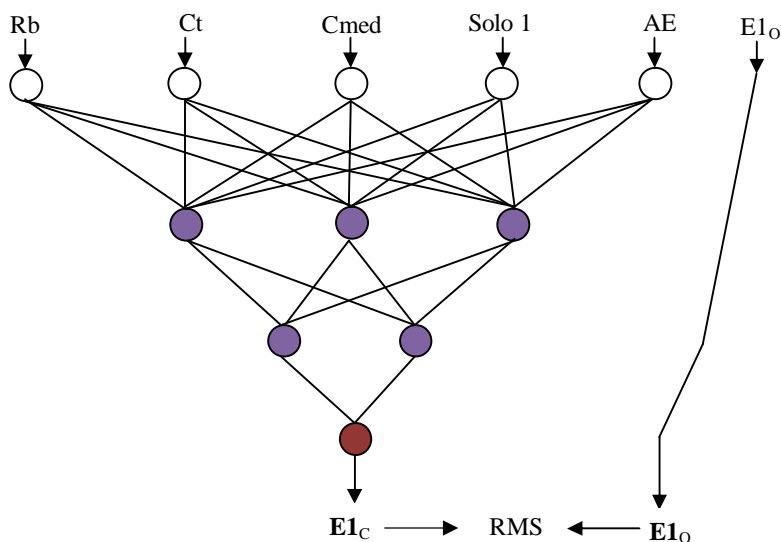


Figura 5.15: Arquitetura da rede neural – parâmetro E1

Tabela 5.15: Parâmetro E1 do modelo SMAP: fase de treinamento da RNA

Ordem	Bacia omitida para verificação	Bacia-meta mais similar	Valores de E1		Erro relativo ERR	R ² (%) Treinamento
			Obtido na calibração do SMAP	Obtido no treinamento da RNA		
1	BA2	BF4	0,477	0,479	0,0042	98,04
2	BA3	BA7	4,545	4,210	0,0737	98,31
3	BA4	BA15	6,124	6,070	0,0088	97,47
4	BA5	BA8	2,126	2,392	0,1251	99,59
5	BA6	BA17	0,069	0,070	0,0145	98,66
6	BA7	BA3	1,021	1,045	0,0235	95,13
7	BA8	BA5	1,344	1,337	0,0052	99,80
8	BA9	BF4	0,477	0,814	0,7065	99,23
9	BA11	BF3	0,317	0,309	0,0252	98,91
10	BA12	BA2	3,076	3,080	0,0013	88,36
11	BA14	BA15	6,124	6,228	0,0170	97,45
12	BA15	BA14	0,779	0,794	0,0193	99,46
13	BA16	BA7	4,545	4,566	0,0046	99,12
14	BA17	BA6	2,004	1,983	0,0105	98,67
15	BA18	BA6	2,004	1,667	0,1682	98,67
16	BA19	BA4	6,659	6,660	0,0002	99,14
17	BA20	BF12	0,135	0,224	0,6593	98,71
18	BA21	BF7	3,032	0,031	1,0000	99,13
19	BA22	BF14	1,293	1,476	0,1415	98,30
20	BA23	BF7	3,032	3,362	0,1088	99,03
21	BA24	BA26	0,821	0,803	0,0219	98,49
22	BA25	BF12	0,135	0,288	1,1333	98,87
23	BA26	BA24	4,281	4,282	0,0002	99,51
24	BA27	BA21	0,374	3,152	7,4278	96,32
25	BF1	BF10	0,584	0,454	0,2226	99,41
26	BF2	BF3	0,317	0,320	0,0095	99,24
27	BF3	BA11	2,915	2,920	0,0017	99,54
28	BF4	BA2	3,076	3,033	0,0140	98,30
29	BF5	BF6	0,755	-0,277	1,3669	96,98
30	BF6	BF3	0,317	0,318	0,0032	98,90
31	BF7	BA23	0,463	0,443	0,0432	99,06
32	BF8	BF9	2,783	2,805	0,0079	95,90
33	BF9	BF8	0,998	1,000	0,0020	96,55
34	BF10	BF1	5,825	5,112	0,1224	99,25
35	BF12	BA25	1,993	1,991	0,0010	97,07
36	BF14	BA22	0,617	0,621	0,0065	95,45

Entre as bacias ruins e regular, somente duas (BA21 e BF7) confirmaram na verificação os piores resultados obtidos no treinamento. Já as bacias (BA6, BA8, BA15, BA23, BA25, e BF3), listadas entre as que não lograram êxito no treinamento, se mostraram inadequadas na verificação, com erros acima de 0,5, conforme a Tabela 5.16.

Quase a totalidade dos erros cometidos, ao adotar-se os valores de E1 da bacia mais similar, são superiores a 0,50, corroborando com os casos dos parâmetros E2 e STR, o que inviabiliza a opção de uso direto da bacia-meta.

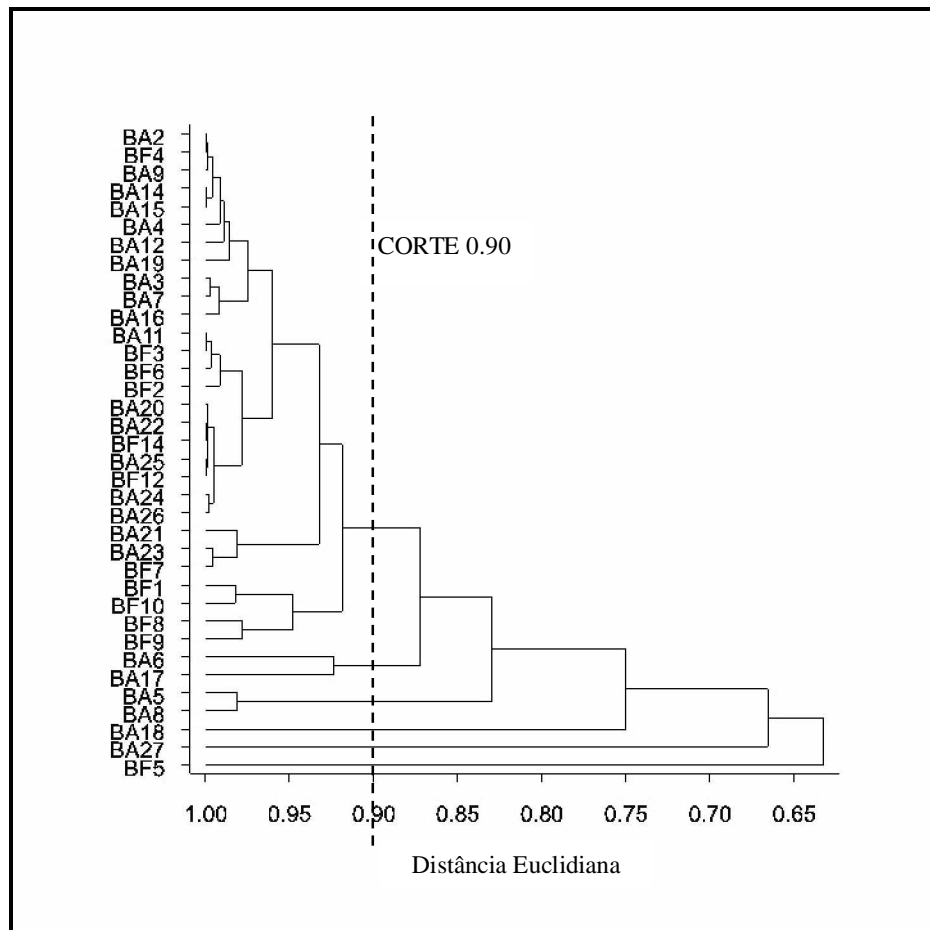


Figura 5.16: Dendrograma obtido para o parâmetro E1 por meio do algoritmo Average Link utilizando a medida euclidiana.

Tabela 5.16: **Parâmetro E1 do modelo SMAP: fase de verificação**

Ordem	Bacia-meta mais similar	Bacia omitida para verificação	Valores de E1			Erro relativo ERR	
			Obtido na calibração do SMAP	bacia-meta	Obtido na verificação da RNA	ao adotar E1 da bacia-meta	ao adotar E1 obtido pela RNA
1	BF4	BA2	3,076	0,477	3,005	0,9998	0,0231
2	BA7	BA3	1,021	4,545	1,021	3,4515	0,0000
3	BA15	BA4	6,659	6,124	7,132	0,0803	0,0710
4	BA8	BA5	1,344	2,126	8,126	0,5818	5,0461
5	BA17	BA6	2,004	0,069	1,951	1,0000	0,0264
6	BA3	BA7	4,545	1,021	4,400	0,7754	0,0319
7	BA5	BA8	2,126	1,344	2,174	0,3678	0,0226
8	BF4	BA9	0,763	0,477	0,832	0,3748	0,0904
9	BF3	BA11	2,915	0,317	2,870	0,9999	0,0154
10	BA2	BA12	5,926	3,076	6,695	0,4809	0,1298
11	BA15	BA14	0,779	6,124	1,379	6,8614	0,7702
12	BA14	BA15	6,124	0,779	6,195	0,9999	0,0116
13	BA7	BA16	7,100	4,545	7,036	0,3599	0,0090
14	BA6	BA17	0,069	2,004	0,537	28,0435	6,7826
15	BA6	BA18	11,083	2,004	1,612	0,8192	0,8545
16	BA4	BA19	6,054	6,659	6,519	0,0999	0,0768
17	BF12	BA20	9,443	0,135	8,297	1,0000	0,1214
18	BF7	BA21	0,374	3,032	3,109	7,1070	7,3128
19	BF14	BA22	0,617	1,293	0,648	1,0956	0,0502
20	BF7	BA23	0,463	3,032	0,445	5,5486	0,0389
21	BA26	BA24	4,281	0,821	3,720	0,9998	0,1310
22	BF12	BA25	1,993	0,135	1,639	0,9999	0,1776
23	BA24	BA26	0,821	4,281	0,874	4,2144	0,0645
24	BA21	BA27	0,683	0,374	9,017	0,4524	12,2020
25	BF10	BF1	5,825	0,584	3,905	0,9999	0,3296
26	BF3	BF2	0,389	0,317	1,108	0,1851	1,8483
27	BA11	BF3	0,317	2,915	0,164	8,1956	0,483
28	BA2	BF4	0,477	3,076	0,606	5,4486	0,2704
29	BF6	BF5	0,468	0,755	13,435	0,6132	27,7073
30	BF3	BF6	0,755	0,317	0,759	0,5801	0,0053
31	BA23	BF7	3,032	0,463	4,820	0,9998	0,5897
32	BF9	BF8	0,998	2,783	0,918	1,7886	0,0802
33	BF8	BF9	2,783	0,998	3,128	0,9996	0,1240
34	BF1	BF10	0,584	5,825	0,549	8,9743	0,0599
35	BA25	BF12	0,135	1,993	0,690	13,7630	4,1111
36	BA22	BF14	1,293	0,617	1,763	0,9995	0,3635

5.4.5 Treinamento da RNA sobre o parâmetro CINF

Merece destaque o fato da rede não reproduzir adequadamente o parâmetro CINF a partir das características apresentadas na Tabela 5.7, talvez pela forte interdependência que mantém com o parâmetro E1.

Neste caso, a introdução dos valores de E1 entre as variáveis explicativas foi a solução para permitir um bom treinamento (Heuvelmans et al, 2006), conforme ilustrado na Figura 5.17. O mérito do subconjunto de variáveis subiu de 0,373 para 0,764.

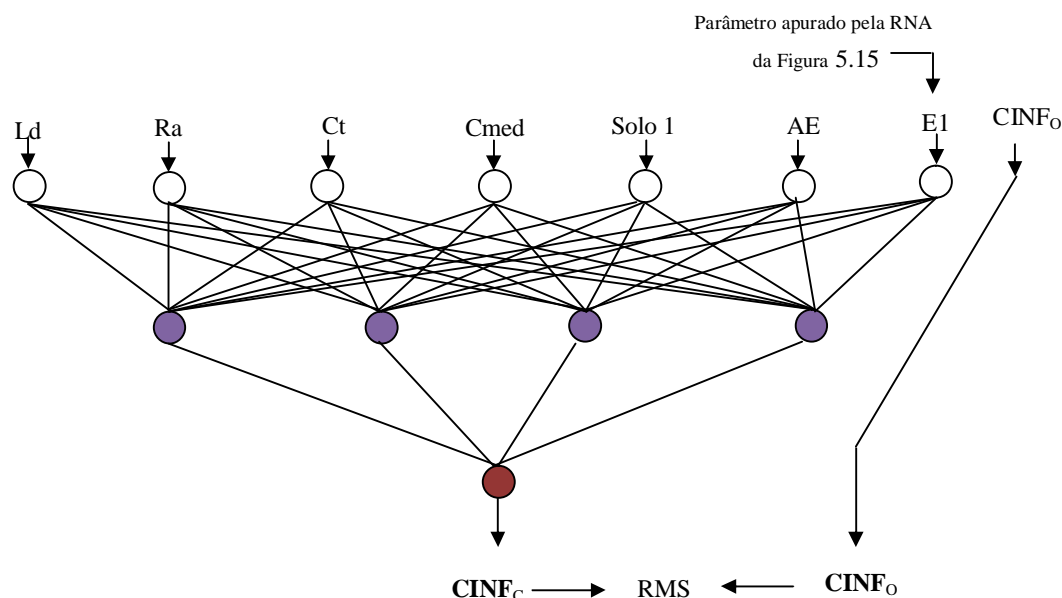


Figura 5.17: Arquitetura da rede neural – parâmetro CINF

A Tabela 5.17 apresenta os resultados do treinamento da rede neural. Em geral obteve-se ótimos resultados na fase de calibração, à exceção das bacias-meta (BA21 e BF6), particularmente na reprodução do parâmetro com valor nulo ou muito baixo.

Na Tabela 5.18 estão os resultados da fase de verificação. Apenas duas bacias (BA6 e BF6) apresentaram resultados ruins e outras duas regulares (BF4 e BF6). O dendograma da Figura 5.18 e a respectiva matriz de similaridades (Anexo B.4) ilustram o fato destas bacias não se aninharem às demais ser o motivo do insucesso nos resultados.

Igualmente aos outros três parâmetros do modelo SMAP, o CINF não poderia ser usado diretamente da bacia mais similar, por causa do elevado índice de erros que se cometeria, conforme mostrado na Tabela 5.18.

Tabela 5.17: Parâmetro CINF do modelo SMAP: fase de treinamento da RNA

Ordem	Bacia omitida para verificação	Bacia-meta mais similar	Valores de CINF		Erro relativo ERR	R ² (%) Treinamento
			Obtido na calibração do SMAP	Obtido no treinamento da RNA		
1	BA2	BA16	0,700	0,700	0,0000	99,83
2	BA3	BA17	0,000	0,020	0,0000	99,94
3	BA4	BA24	0,681	0,680	0,0015	98,71
4	BA5	BA2	0,638	0,640	0,0031	99,22
5	BA6	BA26	0,029	0,030	0,0345	99,86
6	BA7	BA12	0,773	0,760	0,0168	99,99
7	BA8	BF4	0,450	0,380	0,1556	99,83
8	BA9	BA14	0,367	0,367	0,0000	99,93
9	BA11	BF1	0,534	0,530	0,0075	99,88
10	BA12	BA7	0,893	0,773	0,1344	96,60
11	BA14	BA9	0,426	0,428	0,0047	99,94
12	BA15	BA19	0,807	0,836	0,0359	99,98
13	BA16	BA2	0,638	0,640	0,0031	99,88
14	BA17	BA6	0,900	1,004	0,1156	99,82
15	BA18	BA20	0,900	0,900	0,0000	99,70
16	BA19	BA15	0,793	0,829	0,0454	99,98
17	BA20	BA18	0,635	0,640	0,0079	99,67
18	BA21	BF10	0,000	0,000	0,0000	99,96
19	BA22	BF12	0,000	-0,111	0,0000	99,97
20	BA23	BF7	0,503	0,500	0,0060	99,96
21	BA24	BA4	0,889	0,890	0,0011	99,97
22	BA25	BA23	0,369	0,371	0,0054	99,96
23	BA26	BA6	0,900	0,900	0,0000	99,69
24	BA27	BA21	0,000	0,166	1,0000	99,07
25	BF1	BA11	0,722	0,720	0,0028	99,84
26	BF2	BF6	0,181	0,180	0,0055	99,43
27	BF3	BF6	0,181	0,247	0,3646	99,60
28	BF4	BA14	0,367	0,339	0,0763	99,95
29	BF5	BF6	0,181	0,289	0,5967	99,83
30	BF6	BF9	0,632	0,632	0,0000	99,93
31	BF7	BA23	0,369	0,365	0,0108	99,96
32	BF8	BF10	0,000	0,000	0,0000	99,97
33	BF9	BF6	0,181	0,180	0,0055	99,46
34	BF10	BF8	0,352	0,350	0,0057	99,95
35	BF12	BA22	0,000	0,000	0,0000	99,95
36	BF14	BA22	0,000	0,006	0,0000	99,93

Tabela 5.18: **Parâmetro CINF do modelo SMAP: fase de verificação**

Ordem	Bacia-meta mais similar	Bacia omitida para verificação	Valores de CINF			Erro relativo ERR	
			Obtido na calibração do SMAP	bacia-meta	Obtido na verificação da RNA	ao adotar CINF da bacia-meta	ao adotar CINF obtido pela RNA
1	BA16	BA2	0,638	0,700	0,636	0.0972	0,0031
2	BA17	BA3	0,190	0,000	0,194	1.0000	0,0210
3	BA24	BA4	0,889	0,681	0,755	0.2340	0,1507
4	BA2	BA5	0,536	0,638	0,538	0.1903	0,0037
5	BA26	BA6	0,900	0,029	0,891	0.9678	0,0100
6	BA12	BA7	0,893	0,773	0,760	0.1344	0,1489
7	BF4	BA8	0,264	0,450	0,345	0.7045	0,3068
8	BA14	BA9	0,426	0,367	0,438	0.1385	0,0282
9	BF1	BA11	0,722	0,534	0,719	0.2604	0,0041
10	BA7	BA12	0,773	0,893	0,811	0.1552	0,0492
11	BA9	BA14	0,367	0,426	0,365	0.1608	0,0054
12	BA19	BA15	0,793	0,807	0,836	0.0177	0,0542
13	BA2	BA16	0,700	0,638	0,741	0.0886	0,0586
14	BA6	BA17	0,000	0,900	0,043	-	-
15	BA20	BA18	0,635	0,900	0,724	0.4173	0,1402
16	BA15	BA19	0,807	0,793	0,829	0.0173	0,0273
17	BA18	BA20	0,900	0,635	0,850	0.2944	0,0555
18	BF10	BA21	0,000	0,000	0,072	-	-
19	BF12	BA22	0,000	0,000	0,052	-	-
20	BF7	BA23	0,369	0,503	0,368	0.3631	0,0027
21	BA4	BA24	0,681	0,889	0,816	0.3054	0,1982
22	BA23	BA25	0,620	0,369	0,633	0.4048	0,0210
23	BA6	BA26	0,029	0,900	0,003	30.0345	0,8965
24	BA21	BA27	0,082	0,000	0,100	1.0000	0,2195
25	BA11	BF1	0,534	0,722	0,533	0.3521	0,0019
26	BF6	BF2	0,000	0,181	0,008	-	-
27	BF6	BF3	0,000	0,181	1,073	-	-
28	BA14	BF4	0,450	0,367	0,378	0.1844	0,1600
29	BF6	BF5	0,132	0,181	0,900	0.3712	5,8182
30	BF9	BF6	0,181	0,632	0,180	2.4917	0,0055
31	BA23	BF7	0,503	0,369	0,502	0.2664	0,0020
32	BF10	BF8	0,352	0,000	0,271	1.0000	0,2301
33	BF6	BF9	0,632	0,181	0,634	0.7136	0,0032
34	BF8	BF10	0,000	0,352	0,048	-	-
35	BA22	BF12	0,000	0,000	0,085	-	-
36	BA22	BF14	0,000	0,000	0,023	-	-

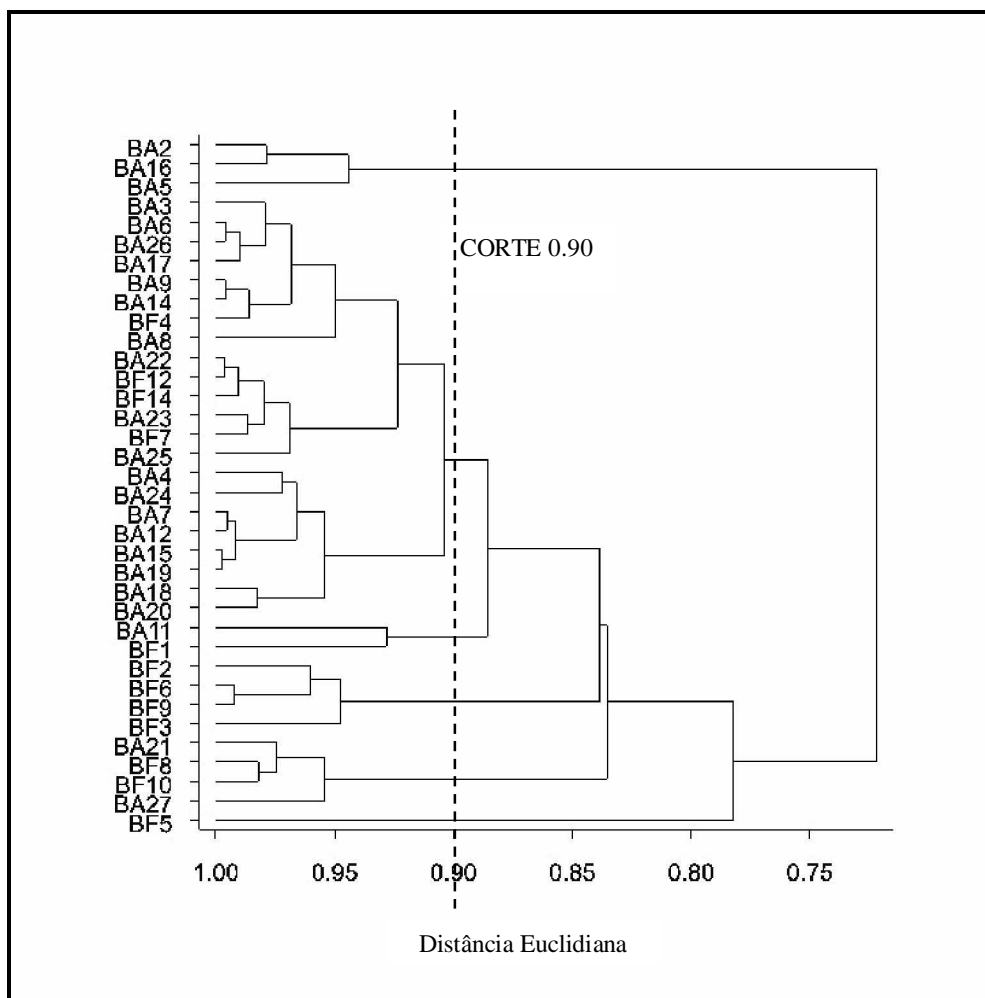


Figura 5.18: **Dendrograma obtido para o parâmetro CINF por meio do algoritmo Average Link utilizando a medida euclidiana.**

5.4.6 Identificação de Redes Neurais Regionais

A Tabela 5.20 apresenta, de forma resumida, para cada bacia omitida, as correspondentes bacias-meta usadas para se estimar os parâmetros do modelo SMAP. Apenas em três casos (BA17, BA27 e BF5), a mesma bacia-meta foi usada para a determinação dos quatro parâmetros do SMAP. Contrariamente, em outros doze casos, identificou-se uma bacia-meta para cada parâmetro.

Apesar de baixa interdependência entre os parâmetros STR e CINF em catorze casos os dois puderam ser explicados pelas mesmas bacias-meta. Já STR e E2, altamente correlacionados, figuram conjuntamente em nove bacias-meta.

As diferentes características físicas e climatológicas das bacias hidrográficas, suas variabilidades espacial e temporal, assim como a heterogeneidade pedológica e geológica nos

perfis verticais das bacias, são os principais responsáveis pela diversidade das bacias-meta selecionadas com vistas à estimativa dos parâmetros do modelo hidrológico; as semelhanças, pelos motivos inversos, ou seja, para os casos em que as características são proximamente iguais.

Confrontando-se a Tabela 5.20 com as Tabelas 5.11, 5.14, 5.16 e 5.18, pode-se observar que nenhuma bacia-meta apresentou, simultaneamente, resultados ruins para os quatro parâmetros, mostrando a inviabilidade de descarte em definitivo de qualquer bacia do conjunto. Por outro lado, há os casos em que a mesma bacia-meta figura como a mais similar para mais de uma bacia omitida.

A partir da observação dos erros de avaliação dos parâmetros contidos nas Tabelas 5.11, 5.14, 5.16 e 5.18, pôde-se classificar qualitativamente cada bacia-meta da Tabela 5.19 e construir as combinações apresentadas na Tabela 5.20.

A Tabela 5.21 apresenta uma contagem dos resultados, distribuídos segundo a classificação, nas fases de treinamento e verificação, para cada parâmetro do modelo SMAP.

Na Tabela 5.22 encontram-se os valores dos parâmetros obtidos na fase de calibração do modelo SMAP e determinados a partir dos modelos de redes neurais referentes a cada bacia-meta indicada na Tabela 5.19.

De posse destes valores dos parâmetros obteve-se as séries mensais de vazões para as 36 bacias hidrográficas, sendo apresentadas a seguir em forma de hidrogramas as séries correspondentes a cinco bacias hidrográficas de estações fluviométricas e duas bacias hidrográficas de açudes, escolhidas como cenários representativos das seguintes combinações entre bacias-meta.

BF1	BF2	BF4	BF6	BF12	BA2	BA3
E2 – ótimo	E2 – ótimo	E2 – ótimo	E2 – ótimo	E2 – ótimo	E2 – ótimo	E2 – ótimo
STR – regular	STR – regular	STR – regular	STR – ótimo	STR – ótimo	STR – ruim	STR – ótimo
E1 – regular	E1 – ruim	E1 – bom	E1 – ótimo	E1 – regular	E1 – ótimo	E1 – ótimo
CINF - ótimo	CINF - ótimo	CINF - bom	CINF - ótimo	CINF - regular	CINF - ótimo	CINF - ótimo

Na Figura 5.19, encontra-se o hidrograma da bacia BF1, onde pode-se observar para a regionalização um pico de vazão atípico, devido ao valor de E1 ter sido super-estimado e, isto, implicar em baixas perdas por evaporação. O efeito potencializador das vazões foi aumentado pela sub-estimação de STR.

De igual maneira, E1 foi super-estimado para a bacia BF2 (hidrograma da Figura 5.20), mas numa escala maior que na bacia anterior, o que aumentou ainda mais as vazões de picos.

Tabela 5.19: Identificação das bacias-meta mais similares usadas para estimar os parâmetros do SMAP na bacia omitida

Ordem	Bacia omitida para verificação	Bacia-meta usada para obter o parâmetro do modelo SMAP			
		E2	STR	E1	CINF
1	BA2	BA15	BA5	BF4	BA16
2	BA3	BA15	BA12	BA7	BA17
3	BA4	BA19	BA14	BA15	BA24
4	BA5	BA16	BA2	BA8	BA2
5	BA6	BA17	BA17	BA17	BA26
6	BA7	BA19	BA12	BA3	BA12
7	BA8	BA23	BA9	BA5	BF4
8	BA9	BA14	BA8	BF4	BA14
9	BA11	BF1	BF1	BF3	BF1
10	BA12	BA7	BA3	BA2	BA7
11	BA14	BA9	BA4	BA15	BA9
12	BA15	BA2	BA26	BA14	BA19
13	BA16	BA5	BA2	BA7	BA2
14	BA17	BA6	BA6	BA6	BA6
15	BA18	BA26	BA6	BA6	BA20
16	BA19	BA7	BA3	BA4	BA15
17	BA20	BF14	BA22	BF12	BA18
18	BA21	BA27	BF7	BF7	BF10
19	BA22	BF7	BF12	BF14	BF12
20	BA23	BA25	BA24	BF7	BF7
21	BA24	BA26	BA23	BA26	BA4
22	BA25	BA23	BA23	BF12	BA23
23	BA26	BA27	BA12	BA24	BA6
24	BA27	BA21	BA21	BA21	BA21
25	BF1	BA11	BA11	BF10	BA11
26	BF2	BF10	BF5	BF3	BF6
27	BF3	BF6	BF6	BA11	BF6
28	BF4	BA9	BA4	BA2	BA14
29	BF5	BF6	BF6	BF6	BF6
30	BF6	BF3	BF3	BF3	BF9
31	BF7	BA22	BA21	BA23	BA23
32	BF8	BF3	BF9	BF9	BF10
33	BF9	BF10	BF8	BF8	BF6
34	BF10	BF9	BF8	BF1	BF8
35	BF12	BF14	BA22	BA25	BA22
36	BF14	BF12	BA22	BA22	BA22

Num cenário mais favorável, com a presença de dois parâmetros ótimos e dois bons, as vazões da bacia BF4 foram muito bem reproduzidas, conforme hidrograma da Figura 5.21. Da mesma forma, como era de se esperar, com a presença dos quatro parâmetros ótimos, as vazões obtidas por regionalização para a bacia BF6 foram excelentes, como pode ser visto na Figura 5.22.

Tabela 5.20: Classificação das bacias segundo os resultados por parâmetro

Ordem	Bacia omitida para verificação	Bacia-meta usada para obter o parâmetro do modelo SMAP			
		E2	STR	E1	CINF
1	BA2	ótimo	ruim	ótimo	ótimo
2	BA3	ótimo	ótimo	ótimo	ótimo
3	BA4	ótimo	ótimo	ótimo	bom
4	BA5	ótimo	ruim	ruim	ótimo
5	BA6	ótimo	ótimo	ótimo	ótimo
6	BA7	ótimo	ótimo	ótimo	bom
7	BA8	bom	ótimo	ótimo	bom
8	BA9	ótimo	bom	ótimo	ótimo
9	BA11	ótimo	ruim	ótimo	ótimo
10	BA12	ótimo	ótimo	bom	ótimo
11	BA14	ótimo	ótimo	ruim	ótimo
12	BA15	ótimo	bom	ótimo	ótimo
13	BA16	ótimo	ótimo	ótimo	ótimo
14	BA17	regular	bom	ruim	bom
15	BA18	ruim	ruim	ruim	bom
16	BA19	ótimo	ótimo	ótimo	ótimo
17	BA20	ótimo	ruim	bom	ótimo
18	BA21	ótimo	ótimo	ruim	ótimo
19	BA22	ótimo	ótimo	ótimo	ótimo
20	BA23	ótimo	bom	ótimo	ótimo
21	BA24	ótimo	ótimo	bom	bom
22	BA25	ótimo	ruim	bom	ótimo
23	BA26	ótimo	ótimo	ótimo	ruim
24	BA27	ótimo	ruim	ruim	bom
25	BF1	ótimo	regular	regular	ótimo
26	BF2	ótimo	ótimo	ruim	ótimo
27	BF3	ótimo	ótimo	regular	ruim
28	BF4	ótimo	ótimo	bom	bom
29	BF5	ruim	bom	ruim	ruim
30	BF6	ótimo	ótimo	ótimo	ótimo
31	BF7	ótimo	bom	ruim	ótimo
32	BF8	bom	ótimo	ótimo	bom
33	BF9	ótimo	ótimo	bom	ótimo
34	BF10	ótimo	ótimo	ótimo	regular
35	BF12	ótimo	ótimo	ruim	regular
36	BF14	ótimo	ótimo	regular	regular

Para a bacia BF12, o valor super-estimado de E1 regionalizado foi, de certa forma, compensado pela super-estimativa de CINF, resultando em séries semelhantes de vazões (Figura 5.23).

Para a bacia BA2, a presença de um STR ruim, com valor muito baixo, ocasionou a geração de cheias além do real. Apesar de manter a forma, o hidrograma obtido por regionalização deslocou-se para cima em resposta à presença de picos de vazões (Figura 5.24).

Tabela 5.21: Contagem dos resultados segundo a classificação obtida no treinamento e na verificação da RNA

Classificação	Treinamento				Verificação			
	E2	STR	E1	CINF	E2	STR	E1	CINF
Ótimo	34	28	24	30	31	23	17	21
Bom	1	5	6	3	2	6	6	9
Regular	0	2	0	1	1	1	3	3
Ruim	1	1	6	2	2	6	10	3

Figura 5.19: Vazões observadas e simuladas - BF1

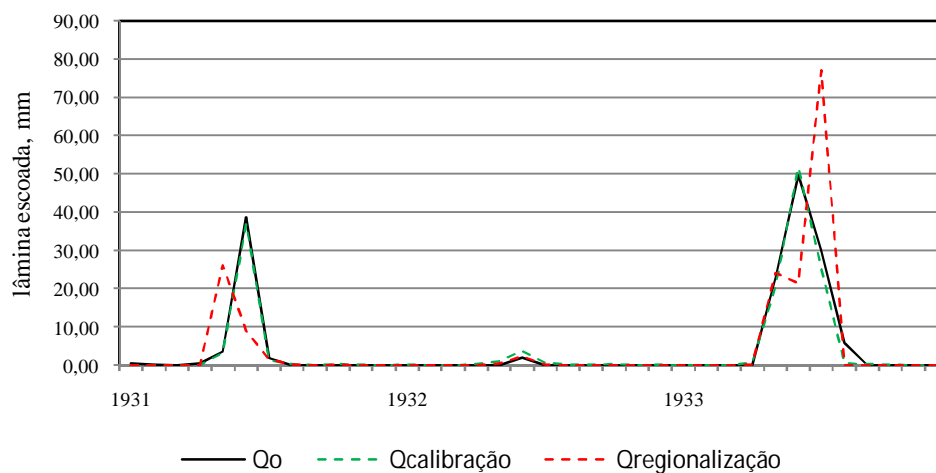


Figura 5.20: Vazões observadas e simuladas - BF2

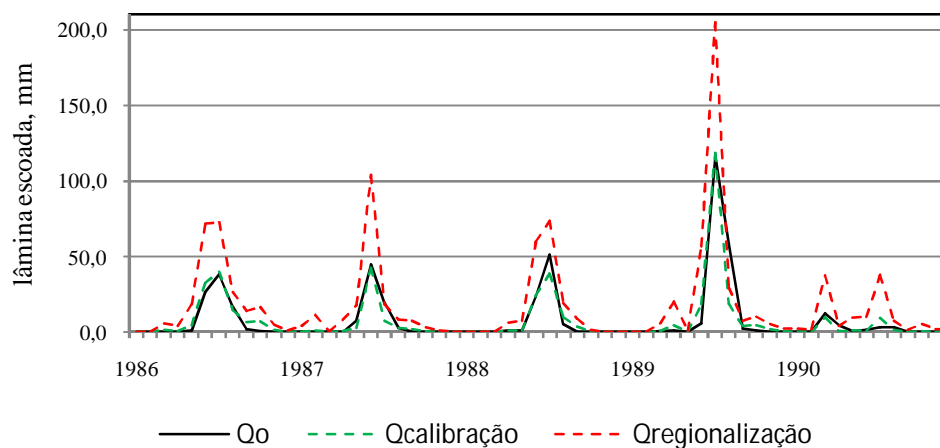


Tabela 5.22: Parâmetros obtidos na calibração e na regionalização

Ordem	Bacia omitida para verificação		Parâmetros obtido na calibração				Parâmetros obtidos na regionalização			
	denominação	Código	E2	STR	E1	CINF	E2	STR	E1	CINF
1	Arrojado	BA2	4,319	443,67	3,076	0,638	4,334	107,84	3,005	0,636
2	Bartolomeu I	BA3	9,908	441,39	1,021	0,190	9,915	470,88	1,021	0,194
3	Cachoeira dos Alves	BA4	3,752	121,07	6,659	0,889	3,748	111,95	7,132	0,755
4	Chupadouro	BA5	5,215	261,87	1,344	0,536	5,209	1.438,85	8,126	0,538
5	Cochos	BA6	0,890	2.141,65	2,004	0,900	0,866	2.153,06	1,951	0,891
6	Emas	BA7	4,341	61,95	4,545	0,893	4,322	68,47	4,400	0,760
7	Eng ^o Arcoverde	BA8	13,676	735,88	2,126	0,264	15,351	735,10	2,174	0,345
8	Eng ^o Ávidos	BA9	11,114	324,54	0,763	0,426	11,089	276,20	0,832	0,438
9	Farinha	BA11	7,473	127,97	2,915	0,722	7,607	202,48	2,870	0,719
10	Frutuoso II	BA12	9,846	212,26	5,926	0,773	10,118	198,91	6,695	0,811
11	Jatobá I	BA14	6,146	328,37	0,779	0,367	6,150	329,87	1,379	0,365
12	Queimadas	BA15	13,126	169,35	6,124	0,793	13,094	133,60	6,195	0,836
13	Riacho dos Cavalos	BA16	15,000	554,73	7,100	0,700	14,945	550,22	7,036	0,741
14	Serra Vermelha	BA17	1,012	6.941,06	0,069	0,000	0,522	5.128,70	0,537	0,043
15	Tamanduá I	BA18	1,752	5.000,00	11,083	0,635	6,244	362,64	1,612	0,724
16	Vazantes	BA19	9,564	221,00	6,054	0,807	9,496	241,38	6,519	0,829
17	Camalau	BA20	9,298	77,72	9,443	0,900	9,325	125,79	8,297	0,850
18	Campos	BA21	6,682	283,07	0,374	0,000	6,684	282,43	3,109	0,072
19	Cordeiro	BA22	6,839	389,38	0,617	0,000	6,816	380,19	0,648	0,052
20	Santo Antonio	BA23	15,000	148,74	0,463	0,369	16,639	181,16	0,445	0,368
21	São Domingos	BA24	14,915	247,95	4,281	0,681	14,905	251,52	3,720	0,816
22	Livramento	BA25	14,928	91,00	1,993	0,620	1,691	337,01	1,639	0,633
23	Namorados	BA26	15,000	265,97	0,821	0,029	15,000	247,06	0,874	0,003
24	Soledade	BA27	13,954	185,20	0,683	0,082	13,942	69,49	9,017	0,100
25	Antenor Navarro	BF1	7,587	285,40	5,825	0,534	7,607	186,25	3,905	0,533
26	Aparecida	BF2	1,640	1.012,12	0,389	0,000	1,644	1.004,26	1,108	0,008
27	Piancó	BF3	1,915	1.028,2	0,317	0,000	1,923	1.042,38	0,164	1,073
28	Emas	BF4	0,635	2163,14	0,477	0,450	0,622	2.072,68	0,606	0,378
29	Patos	BF5	1,389	1281,13	0,468	0,132	-2,402	1.014,60	13,435	0,900
30	Serra Negra do Norte	BF6	2,120	957,1	0,755	0,181	2,116	961,00	0,759	0,180
31	Fazenda Alagamar	BF7	9,948	265,54	3,032	0,503	10,108	207,22	4,820	0,502
32	Pedro Velho	BF8	2,811	398,61	0,998	0,352	2,488	401,24	0,918	0,271
33	Caraúbas	BF9	3,401	267,66	2,783	0,632	3,410	261,16	3,128	0,634
34	Poço de Pedras	BF10	1,885	586,67	0,584	0,000	1,891	616,68	0,549	0,048
35	Umburana/Sumé	BF12	6,053	590,07	0,135	0,000	6,050	587,64	0,690	0,085
36	Gangorra/Sumé	BF14	5,616	803,42	1,293	0,000	5,620	809,85	1,763	0,023

Na Figura 5.25, observa-se a influência negativa da presença de falhas na série histórica de níveis do açude (BA3), o que influenciou os resultados da calibração. Mesmo assim, a regionalização reproduziu com fidelidade o hidrograma obtido na calibração.

Figura 5.21: Vazões observadas e simuladas - BF4

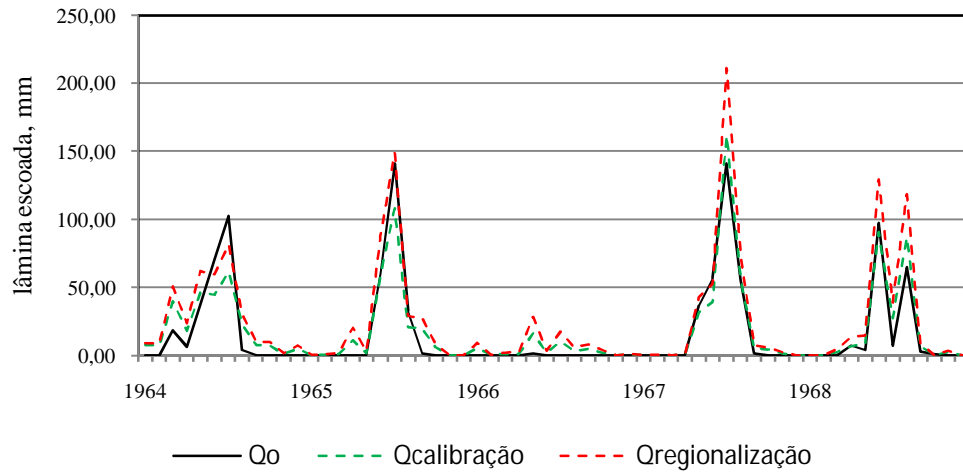


Figura 5.22: Vazões observadas e simuladas - BF6

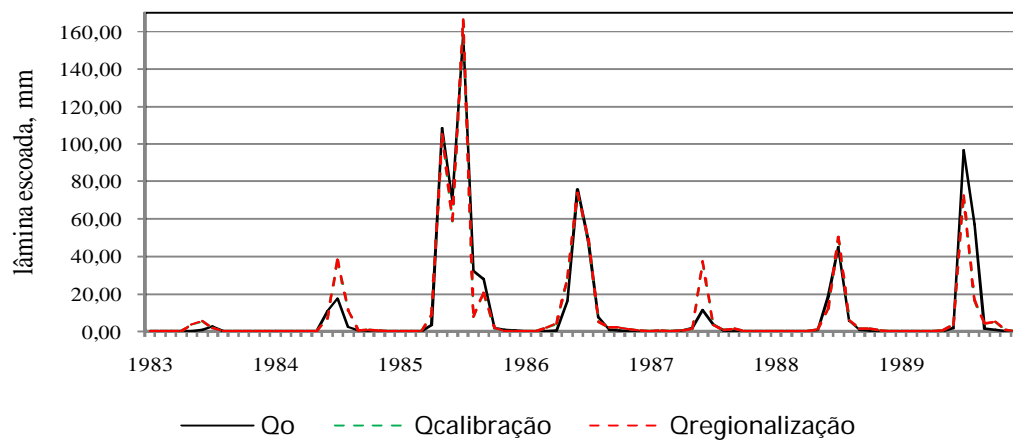


Figura 5.23: Vazões observadas e simuladas - BF12

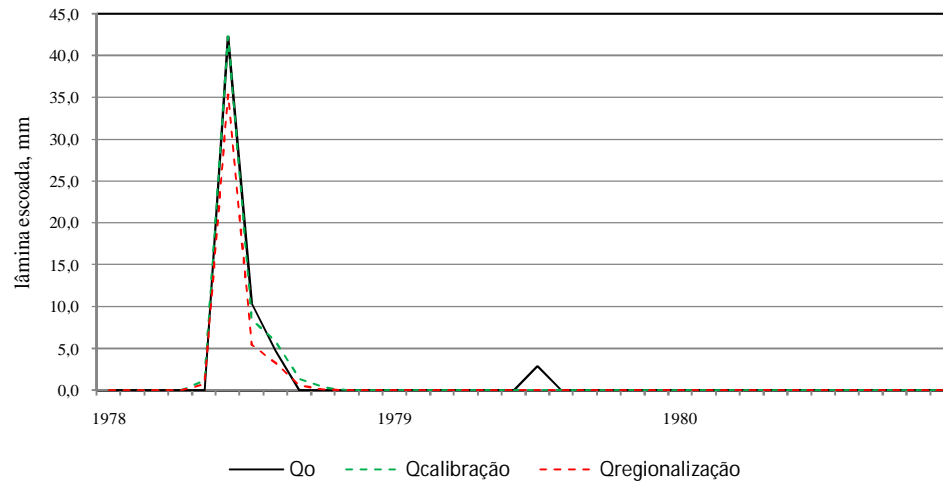


Figura 5.24: Vazões observadas e simuladas - BA2

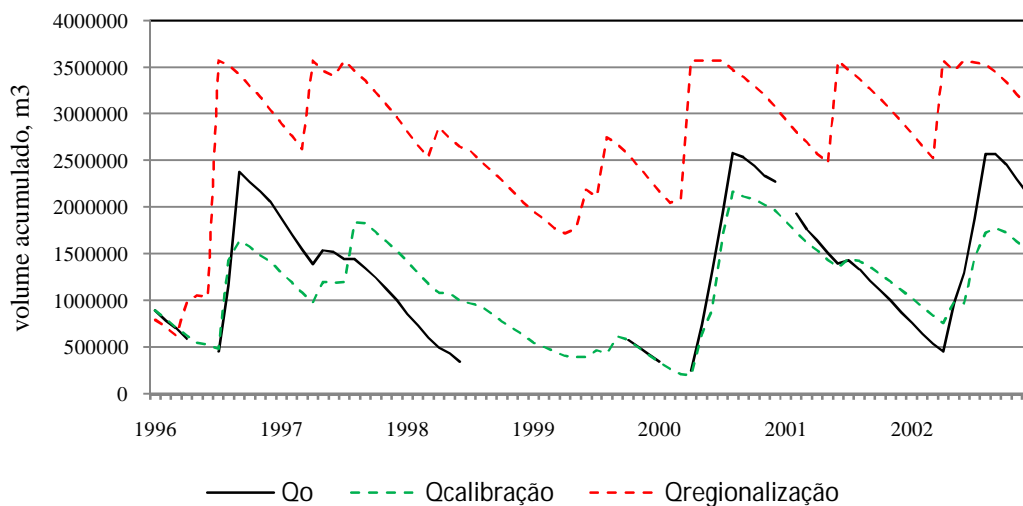
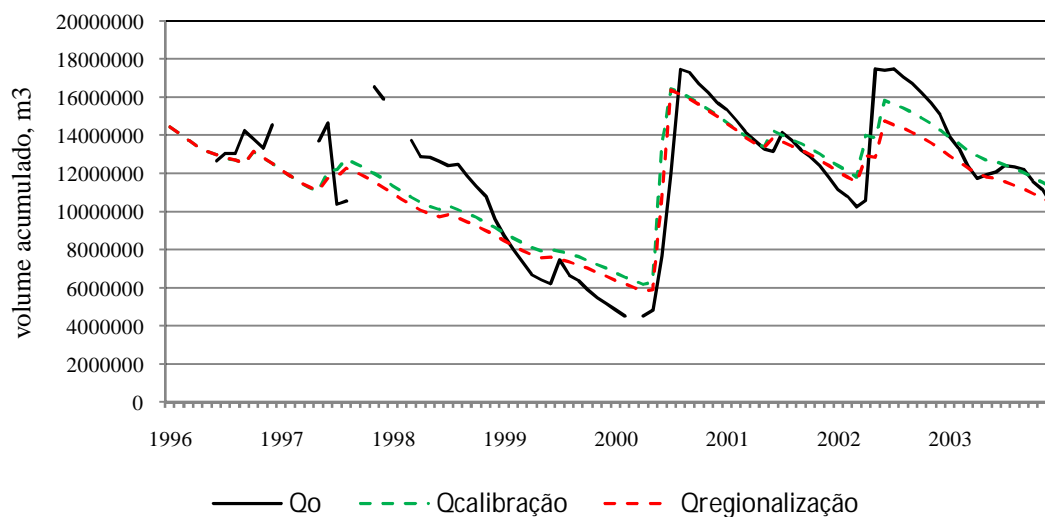


Figura 5.25: Vazões observadas e simuladas - BA3



5.4.7 Aplicação das Redes Neurais Regionais à bacias sem dados

A aplicação para bacias sem dados se dará em cinco fases:

- 1) Obter através de mapas as características físicas e climáticas da bacia sem dados (Ct, Cmed, DS, Ip, Kf, Ld, Lt, P, PAE, Pr, Ra, Rb, Ri, SIN, Solo 1, Solo2), necessárias à determinação dos quatro parâmetros do modelo SMAP (E2, STR, E1, CINF). Os valores de STR e CINF só podem ser calculados após obter-se os valores de E2 e

E1, respectivamente, uma vez que estes fazem parte das variáveis explicativas daqueles. É preferível que as escalas dos mapas sejam as mesmas usadas neste trabalho.

- 2) Determinar a matriz de similaridades e respectivo dendograma, incluindo no conjunto de bacias a bacia sem dados.
- 3) Identificar no conjunto das 35 redes treinadas, para cada parâmetro do modelo SMAP, qual bacia-meta é mais similar à bacia sem dados, selecionando assim o modelo de RNA; não devem ser usadas as bacias (BA26 e BF6) para estimar E2, nem as bacias (BA2, BA5, BA6, BA21, BA23 e BF1) para estimar STR, nem as bacias (BA6, BA8, BA15, BA21, BA23, BA25, BF3, BF6 e BF7) para estimar E1, nem (BA6 e BF6) para estimar CINF.
- 4) Determinar os valores dos parâmetros (E2, STR, E1 e CINF) a partir das respectivas RNA selecionadas.
- 5) A partir dos parâmetros estimados no item anterior, da série de chuvas mensais e das médias mensais de evapotranspiração, pode-se gerar a série de vazões mensais para a bacia sem dados.

CAPÍTULO VI: CONCLUSÕES E RECOMENDAÇÕES

6.1 CONCLUSÕES

As metas estabelecidas no presente estudo consistiram no desenvolvimento e comprovação de uma metodologia para regionalização de parâmetros de modelos chuva-vazão com base nas técnicas de redes neurais artificiais. No desenvolvimento dos trabalhos, algumas reflexões de natureza complementar permitiram agregar outras contribuições não menos relevantes aos estudos de regionalização hidrológica, particularmente aos aqui propostos, conforme apresentado a seguir.

6.1.1 Uso de bacias controladas por açudes

Um dos desafios básicos da análise hidrológica tem sido a falta de dados associados às regiões estudadas. A inexistência de estações fluviométricas em quantidade e qualidade satisfatórias constitui-se, no campo da regionalização, na principal limitação para se conseguir bons resultados. A disponibilidade de apenas 14 estações fluviométricas com dados confiáveis no estado da Paraíba seria inaceitável se não fosse o uso dos dados de níveis de 27 açudes monitorados pela Agência de Águas local.

A inclusão das bacias hidrográficas controladas por açudes juntamente com aquelas com medições fluviométricas permitiu formar uma série melhor de dados para uma boa cobertura geográfica com representação da diversidade físico-climática da região.

Dois problemas, entretanto, relacionados à qualidade das informações obtidas junto ao sistema estadual de monitoramento, se destacaram como relevantes obstáculos ao emprego desta alternativa, levando à rejeição dos dados de 106 açudes. Primeiro, a comum descontinuidade no registro histórico; segundo, a não anotação dos valores de níveis da água em cotas acima da soleira do sangradouro.

6.1.2 O modelo chuva-vazão

A modelagem hidrológica é uma importante ferramenta para se estimar elementos do ciclo hidrológico em áreas de interesse. Se a estrutura do modelo é baseada em princípios básicos da conservação de massa e energia, a estimativa dos parâmetros poderá ser uma difícil tarefa. A heterogeneidade das propriedades do solo e a extrema variabilidade espacial e temporal das variáveis meteorológicas (principalmente a chuva) limita a aplicabilidade dos

modelos de base física, principalmente para pequenas bacias hidrográficas. Para grandes bacias, os modelos conceituais ou semi-conceituais podem ser usados, mas preferencialmente os modelos distribuídos calibrados sobre eventos observados. A aplicação direta destes modelos sobre bacias sem dados é muito limitada face a elevada quantidade de dados requerida para sua calibração e aplicação.

O modelo SMAP, adotado como objeto de análise das técnicas aqui propostas, se destacou pelo fato de ser amplamente usado na região, ter uma estrutura muito simples e seus parâmetros poderem ser correlacionados à características físicas das bacias. Além disso, pode ser aplicado à análise de muitos problemas de engenharia de recursos hídricos.

Dada a simplicidade do modelo, a transferência hidrológica de seus parâmetros para bacias sem dados é de relativa facilidade com o uso das técnicas propostas no presente trabalho.

Ressalta-se que os resultados obtidos poderão orientar estudos semelhantes em modelos mais complexos, inclusive os distribuídos, bastando tratá-los como um somatório de modelos concentrados.

6.1.3 Seleção de atributos relevantes

A seleção de atributos (características físico-climáticas das bacias) se constituiu numa importante tarefa que antecedeu a construção do modelo de regionalização com uso de redes neurais artificiais. Como não era o objetivo central deste trabalho comparar metodologias diferentes, escolheu-se com auxílio da bibliografia, o algoritmo CFS (Hall, 2000), pois propõe tratar o problema de irrelevância e redundância de atributos.

Neste sentido, conta-se entre as contribuições deste trabalho a inclusão desta ferramenta, já consolidada em outras áreas do conhecimento, uma vez que, na maioria dos trabalhos de regionalização, empregam-se uma lista quase repetida de dados das bacias, por vezes sem qualquer critério técnico.

Dos 32 atributos adotados inicialmente, somente 16 foram selecionados como variáveis explicativas dos parâmetros do modelo SMAP, com a característica peculiar de não incluir aqueles normalmente usados em estudos de regionalização, como é o caso da área de drenagem. Cabe destacar que os 16 atributos nomeados foram divididos em sub-grupos para explicar individualmente cada parâmetro do modelo chuva-vazão, permitindo aumentar o número de graus de liberdade da amostra.

A questão da interdependência dos parâmetros do modelo SMAP foi considerada nos estudos por meio da inclusão de E2 e E1 entre as variáveis explicativas dos parâmetros STR e

CINF, respectivamente. Esta operação também permitiu a transferência de informações de um parâmetro para outro, redundando num aumento significativo da quantidade de variância contida nos parâmetros, explicada pelos respectivos atributos.

O desempenho da rede neural foi substancialmente melhorado com a redução da dimensionalidade do conjunto dos atributos. Isto ficou particularmente evidenciado após exaustivas tentativas de definição de uma estrutura de RNA que levasse em consideração todos os 32 atributos, a qual, invariavelmente, resultou em ótima calibração mas péssima verificação.

Sendo o algoritmo CFS um componente da ferramenta WEKA (Witten & Frank, 2000), e esta, importante linha de pesquisa da Universidade de Waikato, tem assegurado contínua atualização. Além do mais, sua aplicação a uma grande quantidade de problemas relacionados à seleção de atributos, lhe confere potencial de uso em regionalização hidrológica.

Destacaram-se como atributos mais importantes da amostra, por figurarem como variáveis explicativas da maioria dos parâmetros, o coeficiente de torrencialidade (C_t), a classe de solo tipo 1 (solo 1), a elevação média da bacia (C_{med}), a área da bacia hidrográfica coberta por espelho de água (AE) e o índice de bifurcação da bacia (R_b).

6.1.4 Uso de Redes Neurais Artificiais

Diversas variáveis hidrológicas requeridas em planejamento e gestão de recursos hídricos podem ser previstas ou regionalizadas a partir das características físico-climáticas das bacias hidrográficas. As redes neurais artificiais em geral propiciam melhores resultados do que os métodos estatísticos de regressão múltipla, para muitas destas variáveis, devido à não-linearidade característica de dados hidrológicos.

A principal vantagem do uso da técnica de RNA é o fato dela não requerer conhecimentos explícitos da bacia, dispensando-se uma análise teórica dos processos implícitos do ciclo hidrológico. No entanto, deve-se ressaltar que tal vantagem pode, eventualmente, reverter-se em uma desvantagem, uma vez que trata-se de um modelo do tipo caixa-preta, ou seja, não se pode conhecer os motivos que levaram tal modelo a um certo resultado. No caso de serem obtidos resultados inconsistentes, o bom senso do hidrólogo pode reparar as falhas de um modelo como o de RNA.

Visando superar os aspectos subjetivos contidos numa aplicação de RNA, buscou-se a definição de um procedimento capaz de extrair informações da amostra de dados e representar com a máxima eficiência as relações existentes entre os atributos das bacias hidrográficas e os parâmetros do modelo SMAP (obtidos no processo de calibração).

Ao mesmo tempo, uma boa calibração só é possível mediante um conjunto consistente de dados que guardem entre si fortes características de homogeneidade, garantindo a obtenção de pesos regionais que bem representam as bacias-meta selecionadas.

6.1.5 Conceito de bacia-meta

O conceito de bacia-meta, estabelecido neste trabalho, tem o objetivo de criar uma interface entre as bacias hidrográficas sem medições hidrométricas e o conjunto de bacias que formam a base regional do Estado. Foram construídos 36 modelos de redes neurais, cada um obtido pela determinação orientada dos pesos sinápticos que melhor ajustam os dados de 35 bacias. A calibração é orientada pela bacia-meta, que tem como objetivo final servir de guia para introduzir a bacia sem dados no modelo mais apropriado.

De posse das características físico-climáticas da bacia sem dados, realiza-se a consignação com os dados da base regional, através da matriz de similaridades e/ou do respectivo dendograma. Identificada a bacia-meta mais similar à bacia sem dados e, por conseguinte, o modelo de RNA associado, calcula-se o valor do parâmetro do modelo SMAP que será, no final, o resultado da interpolação das 35 bacias usadas para treinar a rede.

A simples adoção dos parâmetros referentes à bacia-meta mais similar à bacia sem dados, provocaria enormes erros, dado que o conjunto é que confere à RNA a possibilidade de captação da variabilidade dos atributos das bacias hidrográficas. Em outras palavras, como a RNA é capaz de reproduzir os parâmetros da bacia-meta, também será hábil em o fazer a outra bacia com características semelhantes.

Cabe mencionar dois aspectos de igual destaque: primeiro, a constatação de que a continuidade geográfica não implica em homogeneidade hidrológica; segundo, no caso dos parâmetros do modelo chuva-vazão SMAP, a homogeneidade verificada para um de seus parâmetros não significou o mesmo para outro parâmetro. Com decorrência destes fatos, nenhuma bacia, entre as 36 estudadas, poderia ser descartada do grupo, pois, ao se mostrar discrepante para um parâmetro, o mesmo não ocorreu para outro.

Apesar de terem sido retiradas da amostra as bacias hidrográficas muito grandes (BA10 e BF11) e as muito pequenas (BA1, BA13 e BF13), indicando certos limites para aplicação da metodologia proposta, somente a análise da matriz de similaridades decidirá se uma dada bacia pode ou não ter os seus parâmetros estimados pelo modelo de rede neural estabelecido.

6.2 RECOMENDAÇÕES

6.2.1 Quanto ao uso de bacias controladas por açudes

No que se refere ao uso dos açudes como alternativas às estações fluviométricas convencionais, no que tange à reconstituição das vazões naturais, o monitoramento a ser exercido deve contemplar as demandas hídricas, pelo menos em intervalos diários, e os níveis diários do espelho da água em régua que permitam uma precisão milimétrica, em lugar das régua atuais, graduadas a cada centímetro.

É muito importante o acompanhamento dos níveis, especialmente nos períodos de sangrias pelo vertedouro dos açudes. As descargas de fundo devem ser providas de medidores de vazão e o cadastro de usuários de montante deve ser mantido sempre atualizado.

Somente automatizando o sistema de coleta e envio da informação será possível resolver a maioria dos problemas das séries de níveis dos açudes, tendo-se em conta as dificuldades de gestão de pessoal e de acessibilidade aos locais em tempo de cheias.

6.2.2 Algoritmos para seleção de atributos

De acordo com Hall (1999), não existe um algoritmo de seleção de atributos que seja superior a todos os outros para todos as classes de problemas. Então, a identificação, através de pesquisas específicas, das potencialidades e limitações de diferentes algoritmos, poderá orientar a escolha, entre um ou outro, em contextos diferentes de dados e modelos hidrológicos.

Segundo Lee (2005), não existe ainda uma análise formal capaz de determinar previamente a superioridade de um algoritmo em relação a outro. Por isso, a avaliação do desempenho de algoritmos de seleção de atributos se constitui num tema essencial para o sucesso de boa parte dos estudos de regionalização hidrológica.

Para se obter uma maior compreensão das grandezas, nas relações entre a causa e o efeito que se estabelecem nos processos hidrológicos, notadamente representados através dos parâmetros de modelos chuva-vazão, a investigação de outros atributos, além dos 32 aqui apresentados, é uma relevante matéria para se buscar aqueles que mais expliquem, ou que sejam mais representativos da relação com os parâmetros do modelo chuva-vazão, e que sejam também o mais fácil possível de determinar ou mensurar.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abdulla, F. A.; Lettenmaier, D. P. (1997). Development of regional parameter estimation equations for a macroscale hydrologic model. Journal of Hydrology, v. 197, p. 230-257.
- Acreman, M. C. (1985). Predicting the mean annual flood from basin characteristics in Scotland. Hydrological Sciences Journal, v. 30, n. 1, p. 37-49.
- Aguiar, F. G., Estudo hidrométrico do Nordeste Brasileiro, IFOCS, Rio de Janeiro, jan/mar, 1940, v.13, n.1.
- Alexandre, A. M. B.; Martins, E. S.; Clarke, R. T.; Reis Jr, D. S. (2005). Regionalização de parâmetros de modelos hidrológicos. Anais do XVI Simpósio Brasileiro de Recursos Hídricos, João Pessoa-PB.
- Andrade, E. M., Hawkins, R. H. (2000). Aplicação da função de Andrews na avaliação da regionalização de bacias em regiões áridas e semi-áridas. Revista Brasileira de Recursos Hídricos, v. 5, n. 4, p. 17-24.
- Andrews, D. F. (1972). Plots of high-dimensional data. Biometrics, v. 28, p. 125-136.
- ATECEL - Associação Técnico Científico Enerto Luiz de Oliveira Júnior/UFPB/Instituto Nacional de Colonização e Reforma Agrária-PB (2002). Plano Diretor Sócio-Econômico e Ambiental do Estado da Paraíba. Campina Grande, PB.
- Baena, L. G. N.; Silva, D. D.; Pruski, F. F.; Calijuri, M. L. & Silva, E. O. (2003). Delimitação Automática e Determinação de Características Físicas de Bacias Hidrográficas Utilizando-se Modelos Digitais de Elevação Hidrologicamente Consistentes Gerados a Partir de Mapas com Diferentes Escalas. Anais do XV Simpósio Brasileiro de Recursos Hídricos, Curitiba, PR.
- Barnett, V.; Lewis, T. (1994). Outliers in statistical data. 3rd Edition, Wiley, Chichester, U. K.

- Beable, M. E.; McKerchar, A. I. (1982). Regional flood estimation in New Zealand. Water and Soil Technical Publication 20, Ministry of Works and Development, Wellington, N. Z.
- Benson, M. A. (1962). Evolution of Methods for Evaluating the occurrence of floods, USGS Water Supply Paper 1580-A.
- Berger, K. P., Entekhabi, D. (2001). Basin hydrologic response relations to distributed physiographic descriptors and climate. Journal of hydrology, v. 247, p. 169-182.
- Beven, K.; O'Connell, P. E. (1982). On the role of physically-based distributed modelling in hydrology. Report No. 81, Institute of Hydrology, Wallingford, UK.
- Burn, D. H. (1988). Delineation of groups for regional flood frequency. Journal of Hydrology, v. 7, p. 345-361.
- Burn, D. H.; Boorman, D. B. (1992). Estimation of hydrological parameters at catchments. Journal of Hydrology, v. 143, p. 429-454.
- Cadier, E. (1984). Método de avaliação dos escoamentos nas pequenas bacias do Semi-árido, SUDENE/ORSTON, Recife-PE, 75p.
- Cadier, E. (1996). Small watershed hydrology in semi-arid north-eastern Brazil: basin typology and transposition of annual runoff data. Journal of Hydrology, v. 182, n. 1, p. 117-141.
- Chow, Ven Te, 1964. Handbook of applied hydrology. New York, McGraw-Hill Book.
- Clarke, R. T. (1988). Análise de variância (ANOVA): uma ferramenta para regionalização hidrológica. Revista brasileira de engenharia, v. 6, n. 2, p. 93-105.
- Clarke, R. T.; Montenegro, L. E. (1989). O uso de momentos ponderados por probabilidade na regionalização de registros de vazão na bacia do rio Uruguai. Revista brasileira de engenharia, v. 7, n. 1, p. 103-113.

- Collischonn, W. (2001). Simulação hidrológica de grandes bacias. Tese de doutorado. Instituto de Pesquisas Hidráulicas da UFRGS, Porto Alegre-RS, 270 p.
- Cook, R. D.; Weisberg, S. (1982). Residuals and Influence in Regression. New York, Chapman and Hall, 230 p.
- Dash, M.; Choi, K.; Scheuermann, P.; Liu, H. (2002). Feature selection for clustering – a filter solution. In Proceedings of the Second International Conference on Data Mining, p. 115-122.
- DeCoursey, D. G. (1973). Objective regionalization of peak flow rates. Proceedings of Second International Symposium in Hydrology, p. 395-405, Water Resources Publication, Fort Collins, Colorado.
- Dalrymple, T. (1960). Flood frequency analyses. USGS Water Supply Paper, 1543-A.
- Diaz, O. B. P.; Tucci, C. E. M. (1989). Regionalização de hidrogramas unitários de bacias urbanas brasileiras. Revista brasileira de engenharia, v. 7. n. 2, p. 19-29.
- Diniz, L. S. (1994). Calibração automática de modelos chuva-vazão usando um algoritmo genético. Dissertação de mestrado. Universidade Federal da Paraíba, Campina Grande-PB, 130p.
- Diniz, L. S.; Góis, R. S. S.; Srinivasan, V. S. (1996). Application of a Genetic Algorithm for Calibration and Structural Modification of ‘Tank Model’. In: 6th International Conference: Penang, Malásia, 1996.
- Diniz, L. S.; Clarke, R. T. (2001). Regionalização de parâmetros de modelo chuva-vazão usando redes neurais. XIV Simpósio Brasileiro de Recursos Hídricos, Aracaju-SE.
- Diskin, M. H.; Simon, E. (1977). A procedure for the selection of objective functions for hydrologic simulation models, Journal of Hydrology, v. 34, p. 129-149.
- Draper, N. R.; Smith, H. (1981). Applied Regression Analysis. Second Edition. John Wiley & Sons, Inc.

- Duan, Q. Y.; Gupta, V. K.; Sorooshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization, Journal of Optimization Theory and Applications, v. 76, n. 3, p. 501-519.
- Duan, Q. Y.; Sorooshian, S.; Gupta, V. K. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resources Research, v. 28, n. 4, p. 1015-1031.
- Duan, Q.; Sorooshian, S.; Gupta, V. K. (1993). Shuffled complex evolution approach for effective and efficient global minimization, Journal of Optimization Theory and Applications, v.76, n.3, p.501-521.
- Eagleson, P. S. (1962). Unit hydrograph characteristics for sewered áreas. Journal of the ASCE Hydraulics Division, New York, v. 88, n. 2, p. 1-25.
- Fayyad, U. M.; Piatetsky-Shapiro, G.; Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. Menlo Park: AAAI Press.
- Fernandes, W.; Vogel, R. M.; Sankarasubramanian, A. (2000). Regional calibration of a watershed model. Hydrological Sciences Journal, v. 45, n. 5, p. 689-707.
- Franchini, M.; Suppo, M. (1996). Regional analysis of flow duration curves for a limestone region. Water Resource Management, v. 10, p. 199-218.
- French, M. N.; Krajewski, W. F.; Cuykendal, R. R. (1992). Rainfall forecasting in space and time using a neural network, Journal of Hydrology, v. 137, p. 1-37.
- Funke, R.; Schumann, A. H.; Schultz, G. A. (1997). An Objective method for regionalization of parameters of a physically based rainfall runoff model, International Conference on “Regionalization in Hydrology”, Braunschweig, Germany.
- Galvão, C. O.; Valença, M. J. S. (1999). Sistemas Inteligentes: Aplicações a Recursos Hídricos e Ciências Ambientais, Coleção ABRH de Recursos Hídricos; 7; Editora da UFRGS, 246p.

- Hall, M. (1999). Correlation-based feature subset selection for machine learning. PhD thesis, Department of computer science. <http://www.cs.waikato.ac.nz/~mhall/thesis.pdf>.
- Hall, M. (2000). Correlation-based feature selection for discrete and numeric class machine learning. In proc. Of the 17th Int. Conf. on Machine Learning, p. 359-366, San Francisco, CA. Morgan Kaufmann.
- Han, J. e Kamber, M. (2001). Cluster Analysis. In: Morgan Kaufmann Publishers (eds.), *Data Mining: Concepts and Techniques*, 1 ed., New York, USA, Academic Press.
- Hargreaves, G. H.; Samani, Z. A. (1985). Reference crop evapotranspiration from ambient air temperature. Chicago, Amer. Soc. Agric. Eng. Meeting.
- Hastie, T.; Tibshirani, R.; Friedman, J. (2001). The Elements of Statistical Learning. Springer.
- Helyane, B. B. (2006). Redução de dimensionalidade em bases de dados de expressão gênica. Dissertação de Mestrado, PUC, Curitiba, Paraná.
- Henrich, K.; Schmidt, J.; Dikau, R. (1999). Regionalization of geomorphometric parameters in hydrological modelling using GIS. IAHS Publication no. 254. In: IAHS Series of Proceedings and Reports, ISBN 1-901502-80-5, P.181-192.
- Heras, R., 1972. Manual de hidrologia: 1. princípios basicos en hidrologia. Instituto de hidrologia/Escuela de hidrologia, Madrid, 477 p.
- Herrero, J.; Valencia, A.; Dopazo, J. (2001). A hierarchical unsupervised growing neural network for clustering gene expression patterns. Bioinformatics, n.1, v. 17, p. 126-136.
- Heuvelmans, G.; Muys, B.; Feyen, J. (2006). Regionalization of the parameters of a hydrological model: comparison of linear regression models with artificial neural nets. Journal of Hydrology, n. 319, p. 245-265.
- Hjelmfelt, A. T.; Wang, M. (1993). Artificial neural network as unit hydrograph applications, paper presented at Engineering Hydrology Symposium, Hydraulics Division, Am. Soc. Of Civil Engineering, San Francisco, California.

- Horton, R. E., 1932. Drainage basin characteristics. Transactions, American Geophysical Union, v. 12, p. 350-361.
- Hosking, J. R. M; Wallis, J. R. (1997). Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press, CB2 2RU, UK.
- Hundecha, Y.; Bárdossy, A. (2004). Modeling the effect of land use changes on the runoff generation of a river basin through parameter regionalisation of a watershed model. Journal of hydrology, v. 292, p. 739-747.
- Hsu, K.; Gupta, H. V.; Sorooshian, S. (1995). Artificial neural network modeling of the rainfall-runoff process, Water Resources Research, n. 31, v. 10, p. 2517-2530.
- IBGE - Instituto Brasileiro de Geografia e Estatística (2007). Censo demográfico 2007. Editora IBGE.
- Jain, A. K.; Murty, M. N.; Flynn, P. J. (1999). Data Clustering: A Review. ACM Computing Surveys, v. 1, n. 3, p. 264-323.
- James, L. D. (1972). Hydrologic modeling, parameter estimation and watershed characteristics. Journal of Hydrology, v. 17, p. 283-307.
- Johnson, R. A.; Wichern, D. W. (1998). Applied multivariate statistical analysis, 4^a ed., 816 p., Prentice-Hall, Inc. Symon & Schuster / A Viacom Company, Upper Saddle River, New Jersey.
- Kang, K. W.; Jim, J. H.; Park, C. Y.; Ham, K. J. (1993). Evaluation of hydrological forecasting system based on neural network model, paper presented at 25th Congress of Intern. Assoc. for Hydraulic Research, Int. Assoc. for Hydr. Research, Tokyo, Japan.
- Kira, K.; Rendell, L. (1992). A practical approach to feature selection. In Proc. of the 9th Int. Conf. on Machine Learning, p. 249-256, Aberdeen, Scotland.
- Kite, G. W. (1978). Frequency and risk analyses in hydrology. Water Resources Publication, Fort Collins, Colorado, 224 p.

- Kohonen, T. (1990). The self-organizing map. Proceedings of the IEEE, v. 78, p. 1464-1480.
- Kohonen, T.; Hynninen, J.; Kangas, K.; Laaksonen, J. (1996). Sompak: the selforganizing map program package. Relatório técnico, Helsinki University of Technology, Laboratory of Computer and Information Science, Espoo.
- Kokkonen, T. S.; Jakeman, A. J.; Young, P. C.; Koivusalo, H. J. (2003). Predicting daily flows in ungauged catchments: model regionalisation from catchments descriptors at the Coweeta Hydrologic Laboratory. North Carolina, Hydrological Process, n. 17, v. 11, p. 2219-2238.
- Köppen, W. (1948). Climatologia - Con un estudio de los climas de la Tierra. Trad. P.R.H Perez - Fundo de Cultura Econômica - México.
- Lee, H. D. (2005). Seleção de atributos importantes para a extração de conhecimento de bases de dados. Tese de Doutorado, ICMC-USP, São Carlos - São Paulo.
- Lek, S.; Guégan, J. F. (1999). Artificial neural networks as a tool in ecological modelling, an introduction. Ecol. Model., n. 120, p. 65-73.
- Lins, J. R. T., Medeiros, A. N. (1994). Mapeamento da Cobertura Florestal Nativa Senhora do Estado da Paraíba. Projeto PNUD/FAO/IBAMA – BRA/87/0007/Governo do Estado da Paraíba.
- Liu, H.; Motoda, H. (2002). Feature selection with selective sampling. In 9th Int. Conf. on Machine Learning, p. 395-402, Sydney, Australia.
- Liu, H.; Setiono, R. (1996). A probabilistic approach to feature selection – a filter solution. In proc. Of the 13th Int. Conf. On Machine Learning, p. 319-327, Bari, Italy.
- Loesch, C.; Sari, S. T. (1996). Redes neurais artificiais: fundamentos e modelos. Editora da FURB, Blumenau, 166 p.

- Lopes, J. E. G.; Braga, B. P. F.; Conejo, J. G. L. (1981). Simulação hidrológica: aplicações de um modelo simplificado. IV Simpósio brasileiro de hidrologia e recursos hídricos, anais 1, p. 42-62.
- Magette, W. L.; Shanholtz, V. O.; Carr, J. C. (1976). Estimating selected parameters for the Kentucky watershed model from watershed characteristics. Water Resources Research, n. 12, v. 3, p. 472-476.
- Maidment, D. R. (1992). Handbook of hydrology. McGraw-Hill, Inc., NY.
- Maier, H. R.; Dandy, G. C. (2001). Neural network based modelling of environmente variables: a systematic approach. Math. Comput. Model, n. 33, p. 669-682.
- Manley, R. E. (1978). Simulation of flows in ungauged basins. Hydrological Sciences Journal, v. 23, n. 1, p. 85-101.
- Martins, C. A. (2003). Uma abordagem para pré-processamento de dados textuais em algoritmos de aprendizado. Dissertação de mestrado, USP, São Carlos, SP.
- Matalas, N. C.; Slack, J. R.; Wallis, J. R. (1975). Regional skew in search of a parent. Water Resources Research, v. 11, p. 815-826.
- Mazvimavi, D. (2003). Estimation of Flow Characteristics of Ungauged Catchments: Case Study in Zimbabwe. Ph.D. Thesis, Wageningen University, 188 p.
- Medeiros, J. D. F.; Diniz, L. S.; Tucci, C. E. (2002). Regionalização de curvas de regularização para diferentes garantias: rio Apodi - Mossoró/RN - ABRH. In: VI Simpósio de Recursos Hídricos do Nordeste. CD-ROM. Maceió - AL.
- Merz, R.; Blöschl, G. (2004). Regionalisation of catchment model parameters. Journal of Hydrology, n. 287, p. 95-123.
- Metz, J. (2006). Interpretação de clusters gerados por algoritmos de clustering hierárquico. Dissertação de mestrado, USP, São Carlos, SP.

- Metz, J.; Monard, M. C. (2006). Estudo e análise das diversas representações e estruturas de dados utilizadas nos algoritmos de clustering hierárquico. Relatório Técnico 269, ICMC – USP. http://www.imc.usp.br/~biblio/download/TR_269.pdf.
- McIntyre, N.; Lee, H.; Wheeler, H.; Young, A.; Wagener, T. (2005). Ensemble predictions of runoff in ungauged catchments. Water Resources Research, n. 41, v. 12, W12434, doi: 10.1029/2005WR004289.
- Mitchell, T. (1997). Machine Learning. MacGraw-Hill, Inc., New York, NY.
- Mitra, P.; Murthy, C. A.; Pal, S. K. (2002). Unsupervised feature selection using feature similarity. IEEE Transactions on Pattern Analysis and Machine Intelligence, n. 24, v. 3, p. 301-312.
- Molle, F.; Cadier, E. (1992). Manual do pequeno açude. Recife, SUDENE-DPG-PRN-DPP-APR.523 p., Convênio SUDENE/ORSTOM.
- Mosley, M. P. (1981). Delimitation of New Zealand hydrological regions. Journal of Hydrology, v. 49, p. 173-192.
- Mwakalila, S. (2003). Estimation of stream flows of ungauged catchments for river basin management. Phys. Chem. Earth, n. 28, p. 935-942.
- Nelder, J. A. ; Mead, R. (1965). A simplex method for functional minimization, Comput. J., n.7, v.4, p.308-313.
- Nouvelot, J. F., 1974. Planificação da implantação de bacias representativas. Recife, SUDENE-DRN, 91 p., Convênio SUDENE/ORSTOM.
- Obregon, E.; Tucci, C. E. M.; Goldenfum, J. A. (1999). Regionalização de vazões com base em séries extendidas: bacias afluentes à lagoa Mirim, RS. Revista Brasileira de Recursos Hídricos, v. 4, n. 1, p. 57-75.
- Orr, G.; Schraudolph, N.; Cummins, F. (1999). CS 449 Neural networks lectura notes. Willamette University, USA, <http://www.willamette.edu/~gorr/classess/cs449/>.

- Orsolini, H. E. et al, 2000. Hidrología: procesos y métodos. Editora UNR, Rosario, Argentina, 319 p.
- Osório, Fernando (1999). Redes neurais artificiais: do aprendizado natural ao aprendizado artificial. I Fórum de Inteligência Artificial. ULBRA Canoas, RS.
- Ouarda, T. B. M. J.; Girard, C.; Cavadias, G. S.; Bobée, B. (2001). Regional flood frequency estimation with canonical correlation analysis. Journal of Hydrology, v. 254, p. 157-173.
- Paiva, A. E. D.; Silans, A. M. B. P.; Albuquerque, D. J. S.; Almeida, C. N. (1999). Aplicação do modelo hidrológico distribuído Açumod à bacia hidrográfica do rio Taperoá – Estado da Paraíba, XIII Simpósio Brasileiro de Recursos Hídricos, CD-ROM, Belo Horizonte-MG.
- Paraíba (2003). Atlas do Estado da Paraíba – Informações para gestão do patrimônio natural. João Pessoa, Paraíba, Brasil, p.58.
- Paraíba (2004). Plano Estadual de Recursos Hídricos. João Pessoa, Paraíba.
- Post, D. A.; Jakeman, A. J. (1996). Relationships between catchment attributes and hydrological response characteristics in small Australian mountain ash catchments. Hydrological Processes, v. 10, p. 877-892.
- Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. (1992). Numerical Recipes en C: The Art of Scientific Computing. Cambridge University Press, New York.
- Rabus, B.; Eineder, M.; Roth, A.; Bamler, R. (2003). The shuttle radar topography mission - A new class of digital elevation models acquired by spaceborne radar. Journal of Photogrammetry & Remote Sensing, v. 57, p. 241-262.
- Raman, H.; Sunilkumar, N. (1995). Multi-variate modeling of water resources time series using artificial neural networks, Hydrology Sciences, n. 2, v. 40, p. 145-163.
- Raudkivi, A. J., 1979. Hydrology. Pergamon Press, 1ª Ed., 479 p.

- Refsgaard, J. C.; Knudsen, J. (1996). Operational validation and intercomparison of different types of hydrological models. Water Resources Research, v. 32, n. 7, p. 2189-2202.
- Rice, R. M. (1972). Using canonical correlation for hydrological predictions. Bull. Int. Assoc. Hydrolo. Sci. XVII v. 3, n. 10, p. 315-321.
- Riedel, G.; Maniak, U. (1999). Regionalization of parameters for direct runoff. IAHS Publication no. 254. In: the IAHS Series of Proceedings and Reports, ISBN 1-901502-80-5.
- Riggs, H. C. (1990). Estimating flow characteristics at ungauged. In: Regionalization in hydrology, IAHS publication n. 190, p. 159-169, IAHS Press, Institute of Hydrology, Wallingford, Oxfordshire, UK.
- Schumm, S. A., 1954. Evolution of drainage systems and slopes in Perth Amboy. New Jersey, Geological Society of America Bulletin, v. 67, p. 597-646.
- Sefton, C. E. M.; Howarth, S. M. (1998). Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. Journal of Hydrology, v. 211, p. 1-16.
- Seibert, J. (1999). Regionalisation of parameters for a conceptual rainfall-runoff model. Agricultural and Forest Meteorology, v. 89, p. 279-293.
- Servat, E.; Dezetter, A. (1991). Selection of calibration objective functions in the context rainfall-runoff modelling in a Sudanese Savana area. Journal des Sciences Hydrologiques, v. 36, n. 4, p. 307-330.
- Snyder, F. F. (1938). Synthetic Unit Graphs. Transactions of American Geophysical Union, v.19.
- Sorooshian, S.; Dracup, J. A. (1980). Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: correlated and heteroscedastic error cases. Water Resources Research, v. 16, n. 2, p. 430-442.

- Sorooshian, S.; Gupta, V.K.; Bastidas, L. A. (1998). Calibration of hydrologic models using multi-objectives and visualization techniques. Final Report EAR 9418147. Dep. Hydrology and Water Resources, University of Arizona, 78p.
- Sousa, P. C. (1998). Contribuição para a análise regional da distribuição de caudais de cheias em Portugal Continental. Dissertação de Mestrado em Engenharia de Rega e dos Recursos Agrícolas. Instituto Superior de Agronomia, Lisboa.
- Strobel, O. M. (1987). Identificação de regiões hidrologicamente homogêneas através de análise fatorial de componentes principais. Revista Brasileira de Engenharia, v. 5, n. 1, p. 31-52.
- SUDENE (1972). Levantamento exploratório. Reconhecimento de Solos do Estado da Paraíba. Boletim Técnico nº 05. Ministério do Interior, Rio de Janeiro.
- Thomas, D. M.; Benson, M. A. (1969). Generalization of streamflow characteristics from drainage basin characteristics. USGS Open-file Report, Washington, D. C., 45 p.
- TC/BR - Riverside Technology Inc. (2001). Plano Diretor de Recursos Hídricos da Bacia Hidrográfica do Rio Paraíba, Tomo I, Relatório Final de Diagnóstico, Paraíba, 142p.
- Torrain, P. (1972). Applicability of Canonical Correlation in Hydrology. Hydrology Paper 58. Colorado State University, Fort-Collins, CO 30 p.
- Tucci, C. E. M. (1991). Regionalização de curvas de regularização. Revista Brasileira de Engenharia, v. 9, n. 2, p. 53-69.
- Tucci, C. E. M. (1993). Hidrologia: ciência e aplicação, Editora da ABRH, Porto Alegre, 943p.,
- Tucci, C. E. M. (2002). Regionalização de Vazões. Editora da UFRGS-ABRH, 1ª Ed., Porto Alegre-RS, 256p.
- TVA (1961). Matrix operations in hydrograph computations. Research paper, n. 1, Tennessee Valley Authority, Knoxville, Tennessee, USA.

- Valença, M. J. S. (1999), Análise e projeto de redes neurais com abordagem construtiva para modelagem de sistemas complexos, Tese de Doutorado, UFPE/CCEN/Ciência da Computação, Recife-PE, 223p..
- Versiani, B. R.; Carneiro, R. M. F (2001). Identificações de regiões homogêneas na análise regional de precipitações e de vazões máximas na bacia do rio São Francisco. Revista brasileira de recursos hídricos, v. 6, n. 3, p. 67-80.
- Villela, S. M.; Mattos, A., 1975. Hidrologia aplicada. McGraw-Hill do Brasil, São Paulo, 245 p.
- Zecharias, Y. B.; Brutsaert, W. (1988). Recession characteristics of groundwater outflow and base flow from mountainous watersheds. Water Resources Research, v. 24, n. 10, p. 1651-1658.
- Zhu, M.; Fujita, M.; Hashimoto, N. (1994). Application of neural networks to runoff prediction, in Stochastic and Statistical Method in Hydrology and Environmental Engineering, v. 3, edited by K. W. Hipel et al., p. 205-216, Kluwer Acad., Norwell, Mass.
- Wiltshire, S. E. (1986). Regional flood frequency analysis II: multivariate classification of drainage basin in Britain. Hydrological Sciences Journal, v. 31, n. 3, p. 335-346.
- Wood, E. F. (1995). Heterogeneity and scaling land-atmospheric water and energy fluxes in climate systems. In: Feddes, R. A. (Ed), Space and time scale variability and interdependencies in hydrological processes. Cambridge University Press, p. 3-19.
- Witten, I. H.; Frank, E. (2000). Data Mining: Pratical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann, California.
- Yapo, P. O.; Gupta, H. V.; Sorooshian, S. (1996). Calibration of conceptual rainfall-runoff models: sensitivity to calibration data. Journal of Hydrology, v. 181, p. 23-48.

Yokoo, Y.; Kazama, S.; Sawamoto, M.; Nishimura, H. (2001). Regionalization of lumped water balance model parameters based on multiple regression, Journal of Hydrology, v. 246, p. 209-222.

Yu, L.; Liu, H. (2004). Efficiente feature selection via analysis of relevance and redundancy. Journal of Machine Learning Research, v. 5, p. 1205-1224.

**ANEXO A: DETALHES DAS CARACTERÍSTICAS FÍSICAS
DAS BACIAS HIDROGRÁFICAS**

A.1 BACIA HIDROGRÁFICA

Definição

Também denominada bacia fluvial ou bacia de contribuição, pode ser definida como a área que drena superficialmente a água proveniente da chuva para uma determinada seção de referência. Pode-se considerar que cada bacia está separada de outras que a rodeia por uma linha divisória das águas (Orsolini et al, 2000) traçada seguindo os pontos mais altos que lhes servem de borda (Figura A.1).

Os terrenos de uma bacia hidrográfica são delimitados por dois tipos de divisores de água (Villela e Mattos, 1975): o divisor topográfico e o hidrográfico, conforme indicado na Figura A.2. Observa-se que nem sempre há coincidência entre os dois divisores, embora na prática considere-se apenas o divisor topográfico. Em geral, essas diferenças têm pouca importância para as bacias grandes pois, em geral, são de pequena magnitude, salvo haja na bacia os fenômenos cársticos. Por outro lado, em bacias pequenas esta diferença absoluta pode ter grande importância relativa (Heras, 1972).

Traçado da linha divisória da bacia

Considerando-se coincidentes os divisores topográfico e hidrográfico seguem algumas regras práticas para o traçado do divisor topográfico da bacia hidrográfica (Molle, 1992).

- a) Traçar ou destacar no mapa a rede hidrográfica a montante da seção de referência, bem como a hidrografia das bacias adjacentes para orientar a definição dos divisores de água.
- b) A linha do divisor de água é sempre perpendicular às curvas de níveis.
- c) Quando a altitude do divisor vai aumentando, corta as curvas de níveis por seu lado convexo; ao contrário, quando a altitude vai diminuindo, corta pelo lado côncavo.
- d) Quando o relevo é suave e não há curvas de níveis próximas dos divisores, a linha divisória deve passar a meia distância das extremidades de dois riachos pertencentes a bacia estudada e a outra vizinha.
- e) O divisor de água sempre passa pelos pontos cotados no mapa mais próximos do traçado.

- f) Na ausência de informações claras da topografia as estradas podem ser utilizadas como divisores, uma vez que elas muitas vezes seguem a crista do relevo.
- g) Deve-se observar que a linha divisória nunca deve cortar um curso de água, exceto na seção de referência.

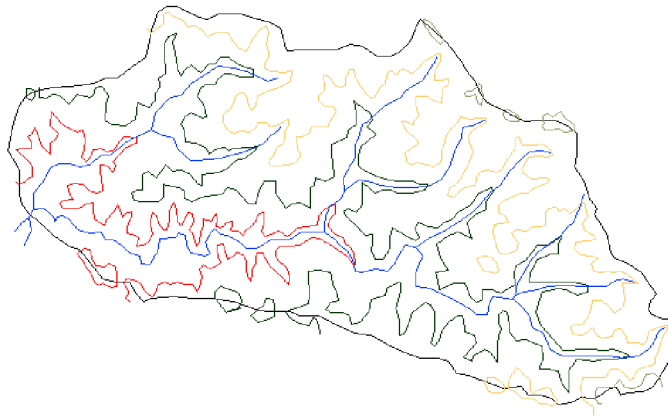


Figura A.1: **Desenho esquemático de uma Bacia hidrográfica**

A.2 CARACTERIZAÇÃO FÍSICA

As características físicas primárias de uma bacia hidrográfica são morfologia (área, forma, elevação, declividade, orientação e rede de drenagem), tipos de solos, capacidade de armazenar água e cobertura vegetal. O efeito de cada uma dessas características sobre o comportamento hidrológico da bacia é óbvio, porém sua quantificação é em geral de difícil avaliação, principalmente porque tais efeitos se sobrepõem. O efeito combinado de todos os fatores define o tamanho da bacia hidrográfica (Raudkivi, 1979). Duas bacias com áreas iguais, por exemplo, podem ter comportamentos muito diferentes.

As grandes bacias são sensíveis aos efeitos do armazenamento nos rios, lagos e depressões e insensíveis a variação e intensidade da chuva e ao uso do solo. Pequenas bacias são controladas pelo escoamento superficial, o uso do solo e a declividade e são também muito sensíveis à chuva. Na verdade, como uma grande bacia é formada por várias pequenas

bacias, a diferenciação das reações de cada bacia é, em parte, devido a descontinuidade e variabilidade dos fenômenos que ali ocorrem.

A grande dificuldade está em expressar a influência de cada característica física da bacia hidrográfica, mas para efeitos práticos pode-se estabelecer um número mínimo susceptível de uso para a classificação das bacias e facilitar os estudos por analogia.

Com base em mapas topográficos em escala apropriada, fotografias aéreas ou imagens de satélites, a bacia hidrográfica pode ser caracterizada por seus principais atributos morfométricos, além de vários outros derivados por meio de equações.

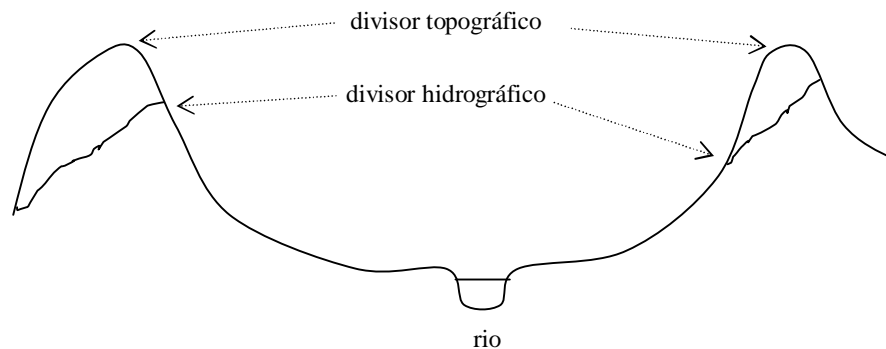


Figura A.2: Esquema dos divisores de uma bacia hidrográfica

Uma boa avaliação das características físicas de uma bacia hidrográfica depende, entre outros, da escala dos mapas topográficos usados. Segundo Orsolini (2000), podem ser consideradas as seguintes recomendações:

A (km ²)	Escala do mapa
menor que 100	25.000
entre 100 e 1.000	50.000
entre 1.000 e 10.000	200.000
maior que 10.000	500.000

No desenvolvimento de estudos para caracterização física das bacias é necessário o cálculo de alguns coeficientes que precisam dos valores das áreas entre curvas de níveis sucessivas, cuja escala em que se deve obtê-las pode ser estimada da seguinte maneira:

$$NU = \frac{H_M - H_E}{6}$$

sendo: H_M cota do terreno na seção de referência;

H_E cota máxima da bacia.

Se NU estiver compreendido entre 100 e 200 se planimetra as curvas de níveis de 100 em 100 metros, desprezando-se as cotas menores. Caso NU esteja entre 200 e 300, metros se planimetra as curvas de níveis de 200 em 200 metros, desprezando-se as cotas menores e, assim, sucessivamente.

A.2.1 Área e perímetro da bacia

A área e o perímetro são duas características da bacia hidrográfica sempre bem definidas. Duas bacias com áreas iguais podem ter comportamentos muito distintos. Com o objetivo de avaliar esse aspecto costuma-se usar duas figuras geométricas simples, o círculo e o quadrado, para comparar com a própria bacia. Quanto mais a forma da bacia se aproxima de uma das duas figuras, tanto menor é o tempo de concentração da bacia e, portanto, mais propícia à ocorrência de cheias.

Área (A)

A área também denominada de área de drenagem ou área de contribuição é a projeção horizontal da região contida entre seus divisores. A determinação da área é feita a partir de mapas, obtidos por fotografias aéreas, imagens de satélites ou por levantamento topográfico, utilizando planímetro ou pelo método das quadrículas usando papel milimetrado ou software gráfico que usa conceitos de trigonometria básica. A área se constitui a característica simples mais importantes da bacia hidrográfica para efeitos dos estudos hidrológicos.

O tamanho da área da bacia indica o quanto uma chuva será potencializada para gerar volume de água na seção de referência.

Perímetro (Pr)

O perímetro é o comprimento da linha divisória da bacia hidrográfica medida em sua planta baixa. Bacias com valores de perímetros iguais podem possuir áreas diferentes a depender de suas formas.

A.2.2 Características das medidas lineares da bacia

Além da área de drenagem e do perímetro existem outras medidas básicas importantes para caracterizar o tamanho da bacia.

Linha de fundo (L)

É a distância medida ao longo do curso de água principal desde a seção de referência até o divisor da bacia, conforme ilustrado na Figura A.3.

Comprimento de um curso de água (L_r)

É a distância medida ao longo do curso de água principal desde a seção de referência até o final do canal como indicado na Figura A.3. O curso de água principal é, geralmente o de maior extensão, corre mais ou menos na região central da bacia e é aquele que conduz os escoamento para fora da seção de referência.

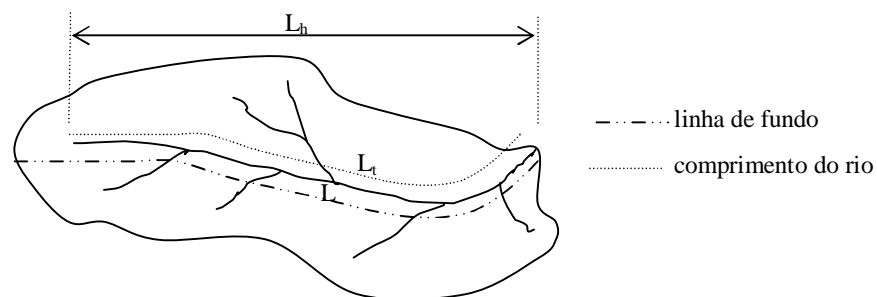


Figura A.3: **Comprimentos de uma bacia hidrográfica**

Comprimento da rede de drenagem (L_d)

É o somatório dos comprimentos de todos os talwegues da bacia. A sua magnitude indica se uma bacia é mal ou bem drenada.

Largura média (L_m)

Obtida de forma indireta através da relação entre a área da bacia e a linha de fundo. Quanto mais L_m se aproxima de L , tanto mais a bacia é propícia a ocorrência de cheias instantâneas.

$$L_m = \frac{A}{L}$$

A.2.3 Características da forma da bacia

Índice de compacidade (K_c)

Também conhecido como coeficiente de Gravelius é a relação entre o perímetro da bacia hidrográfica e o perímetro de um círculo de igual área. Para a mesma área, o círculo é a figura de menor perímetro e, portanto, em qualquer caso esse coeficiente será maior que a unidade e quanto mais a forma da bacia se aproxima de um círculo mais propícia será a ocorrência de cheias e mais próximo de um será K_c .

$$A = \pi r^2 \quad \therefore r = \sqrt{\frac{A}{\pi}}$$

$$K_c = \frac{\text{perímetro } P}{2\pi r} \quad \therefore$$

$$K_c = \frac{0.28P}{\sqrt{A}}$$

Índice de circularidade (K_e)

Com os mesmos princípios do índice de compacidade é definido pela relação entre a área de um círculo que tem o perímetro igual ao da bacia e a área da bacia, ou seja:

$$K_e = \frac{A}{A_0}$$

sendo: $A_0 (P^2/4\pi)$ a área do círculo que tem o mesmo perímetro da bacia

Os valores de K_e e K_c estão relacionados por: $K_e = K_c^{-2}$

Fator de forma (K_f)

É a relação entre a largura média (L_m) e a linha de fundo da bacia (L).

$$K_f = \frac{L_m}{L}$$

A exemplo do coeficiente de compacidade, K_f também é indicativo de tendência para enchentes na bacia. Uma bacia com fator de forma baixo ou muito alto é menos sujeita a enchentes que outra de mesmo tamanho porém com fator de forma próximo a unidade.

A.2.4 Características da rede de drenagem

Ordem dos cursos de água

Geralmente a análise das características de uma bacia hidrográfica é iniciada com a identificação da ordem dos cursos de água, a qual é feita com base em mapas que contêm todos os segmentos da rede de drenagem, sejam perenes, efêmeros ou intermitentes, pequenos ou grandes. De acordo com o critério de classificação introduzido por Horton e depois modificado por Strahler (Chow, 1964) aos pequenos cursos de água que não possuam tributários é atribuído ordem 1. A partir da confluência de dois cursos de água de ordem 1 se forma um curso de ordem 2; a partir da confluência de dois cursos de água de ordem 2 se forma um curso de ordem 3 e, assim, sucessivamente, conforme ilustrado na Figura A.4. O rio principal que passa na seção de referência é o segmento de maior ordem da bacia.

Sendo a ordem dos cursos de água um valor adimensional pode ser usada para comparar as geometrias de duas bacias completamente diferentes em escala linear.

Após a rede de drenagem ter sido numerada com a ordem de cada curso de água, os segmentos de cada ordem “u” são contados e atribuídos à variável N_u . O número N_u de qualquer ordem é maior que N_{u+1} e menor que N_{u-1} .

Índice de bifurcação (R_b)

É a média geométrica das relações entre números de talvegues de duas classes consecutivas (Schumm, 1954), ou seja, a relação entre o número de segmentos de cursos de água de uma dada ordem N_u e o de uma ordem imediatamente superior N_{u+1} .

$$R_b = \frac{N_u}{N_{u+1}}$$

Mesmo não sendo R_b o mesmo entre uma ordem e a sua próxima, tende a ser constante ao longo da série. Essa constatação é a base da seguinte Lei do número de cursos de água de Horton: *o número de segmentos de cursos de água de cada ordem forma uma seqüência geométrica inversa com a respectiva ordem*, ou seja:

$$N_u = R_b^{k-u}$$

Sendo k a ordem do rio principal na seção de referência da bacia hidrográfica.

O índice de bifurcação R_b caracteristicamente varia entre 3.0 e 5.0 para bacias cuja estrutura geológica não difere muito do padrão de drenagem (Chow, 1964), sem grandes diferenças entre regiões. É uma boa opção para comparar bacias com formas aproximadamente iguais, arredondadas ou alongadas, por exemplo.

Um importante resultado mostrado por Horton (1932) é que o número total de cursos de água de todas as ordens em um sistema de drenagem pode ser calculado se R_b e k são conhecidos pela seguinte expressão:

$$\sum_{i=1}^k N_u = \frac{R_b^k - 1}{R_b - 1}$$

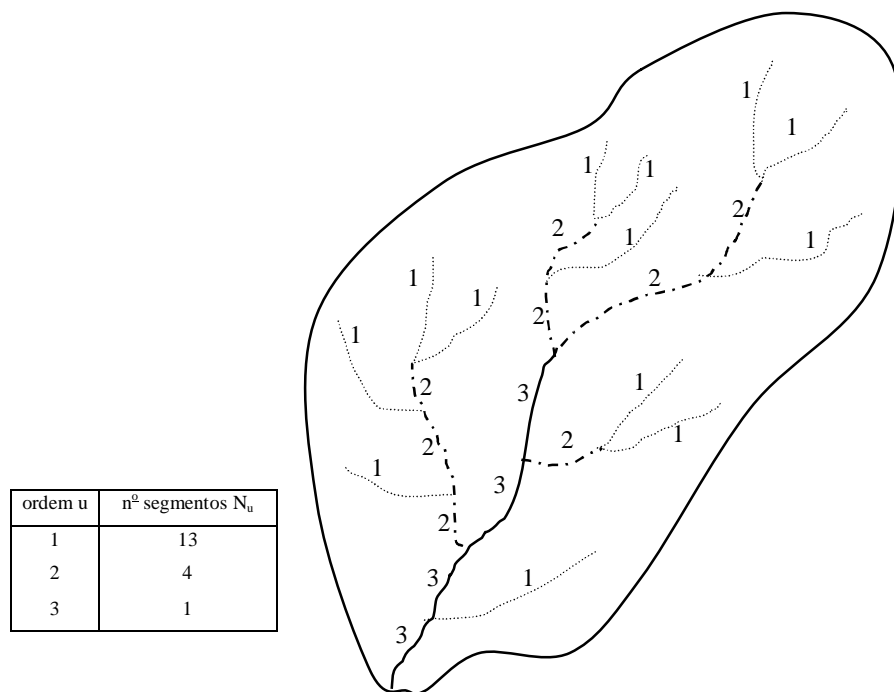


Figura A.4: Ordem dos cursos de água de uma bacia segundo Strahler (Chow, 1964)

Índice dos comprimentos (R_L)

É a média geométrica das relações entre os comprimentos médios dos talvegues de duas classes consecutivas (Schumm, 1954), ou seja, é a relação entre os comprimentos médios dos cursos de água L_u e L_{u+1} , respectivamente, referentes às ordens u e $u+1$.

Para se obter o comprimento médio dos cursos de água L_u de ordem “ u ”, o comprimento total é dividido pelo número de segmentos N_u daquela mesma ordem, então:

$$L_u = \frac{\sum_{i=1}^N L_i}{N_u}$$

O comprimento médio L_u é maior que L_{u-1} e menor que L_{u+1} . O índice de comprimentos R_L tende a ser constante através das sucessivas ordens de uma bacia (Horton, 1932). Além disso, o comprimento médio L_u para uma ordem qualquer pode ser relacionado com o comprimento médio dos cursos de água de primeira ordem L_1 , por meio da seguinte expressão, conhecida como a Lei dos comprimentos de Horton.

$$L_u = L_1 R_L^{u-1}$$

Índice das áreas (R_a)

A área média das sub-bacias dos cursos de água de ordem qualquer aumenta, como no caso dos comprimentos médios, com o aumento da ordem. Schumm (1954) expressou essa relação como a Lei das áreas dos cursos de água: *as áreas médias das sub-bacias dos cursos de água de cada ordem a uma função geométrica na qual o primeiro termo é a área média da sub-bacia de primeira ordem*, ou seja:

$$A_u = A_1 R_a^{u-1}$$

$$R_a = A_u/A_{u+1}$$

sendo: A_u a área média da bacia de ordem u ;

$$R_L = \frac{L_u}{L_{u+1}}$$

R_a é um índice de área análogo ao do comprimento R_L .

Densidade de drenagem (D_d)

É o quociente entre o comprimento total da rede de drenagem e a área da bacia hidrográfica, serve para auxiliar os índices de declividade e compacidade para entender os escoamentos de superfície. Muitas vezes é considerado um indicador chave da resposta hidrológica de uma bacia, dando a diferença em velocidade e tempo de permanência da água entre as cabeceiras e o rio principal. A densidade de drenagem também tem implicações sobre a extensão de áreas saturadas e a geração de deflúvios. O solo é provavelmente mais saturado dentro da rede de drenagem do que nas áreas localizadas nas cabeceiras da bacia. Um alto valor de densidade de drenagem indica uma densidade de rios relativamente alta e, então, uma resposta rápida de cheias.

$$D_d = \frac{\sum L_i}{A}$$

sendo: $\sum L_i$ o somatório dos comprimentos de todos os talwegues da bacia (km);

A é a área da bacia hidrográfica (km²).

Coefficiente de torrencialidade (C_t)

É definido pela relação entre o número de cursos de água e a área da bacia. É utilizado principalmente para estudos de cheias máximas, uma vez que fornece um índice relacionado com as características físicas e morfológica da bacia.

$$C_t = \frac{N}{A}$$

sendo: N o número de cursos de água e A a área da bacia hidrográfica (km²).

Índice de rugosidade (IR)

É um índice adimensional definido pelo produto entre a densidade de drenagem D_d (km/km²) e a amplitude altimétrica H (km). Elevados valores de IR indicam bacias sujeitas a cheias rápidas, face a grandes desníveis entre cabeceira e seção de referência associados a densa rede de drenagem.

$$IR = D_d H$$

Extensão média do escoamento superficial (L_e)

É a distância média em que a água da chuva teria que escoar sobre a superfície de uma bacia hidrográfica, caso o curso de água principal fosse uma linha reta, conforme a Figura A.5.

$$L_e = \frac{A}{4L}$$

Sendo: A a área da bacia hidrográfica (km²) e L a linha de fundo da bacia (km).

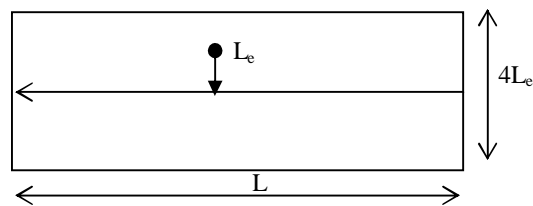


Figura A.5: Extensão média do escoamento superficial

Sinuosidade do curso d'água (SIN)

É a relação entre o comprimento do rio principal e a sua projeção horizontal, ou seja medido em linha reta. Um valor de SIN igual a unidade indica um rio sem sinuosidade.

$$SIN = \frac{L_t}{L_h}$$

A.2.5 Características do relevo da bacia

O relevo é a diferença de elevação entre dois pontos de referência. O relevo máximo de uma bacia é a diferença de elevação entre o ponto mais alto no seu divisor e o mais baixo na seção de referência.

Para um determinado tipo de terreno a divisão da água da chuva entre o escoamento superficial, a infiltração e a evapotranspiração é basicamente função do relevo, o qual pode ser caracterizado quantitativamente por alguns índices.

Declividades máxima (I_{\max}) e média (I_{med}) do rio

Indica, respectivamente, as declividades máxima e média do rio principal, que servem também para indicar a característica do relevo da bacia, sendo a declividade média mais representativa. A declividade máxima, aproximada pela expressão a seguir, serve para indicar uma estimativa aproximada do gradiente topográfico afetando o movimento lateral da água subterrânea em áreas saturadas próximas da superfície se considerada a bacia como um todo. Valores altos de I_{\max} se correlaciona com a eficiência de redistribuição lateral da umidade.

$$I_{\max} = \frac{H_{\max} - H_{\min}}{L_h}$$

sendo: L_h = distância horizontal entre os pontos mais baixo e mais alto do rio principal em metros;

H_{\max} e H_{\min} = cotas máxima e mínima do rio principal em metros.

Do perfil longitudinal do curso de água principal (elevação versus distância horizontal), pode-se determinar a declividade média como a inclinação da reta que divide as áreas acima e abaixo da curva em dois valores iguais, conforme a Figura A.6.

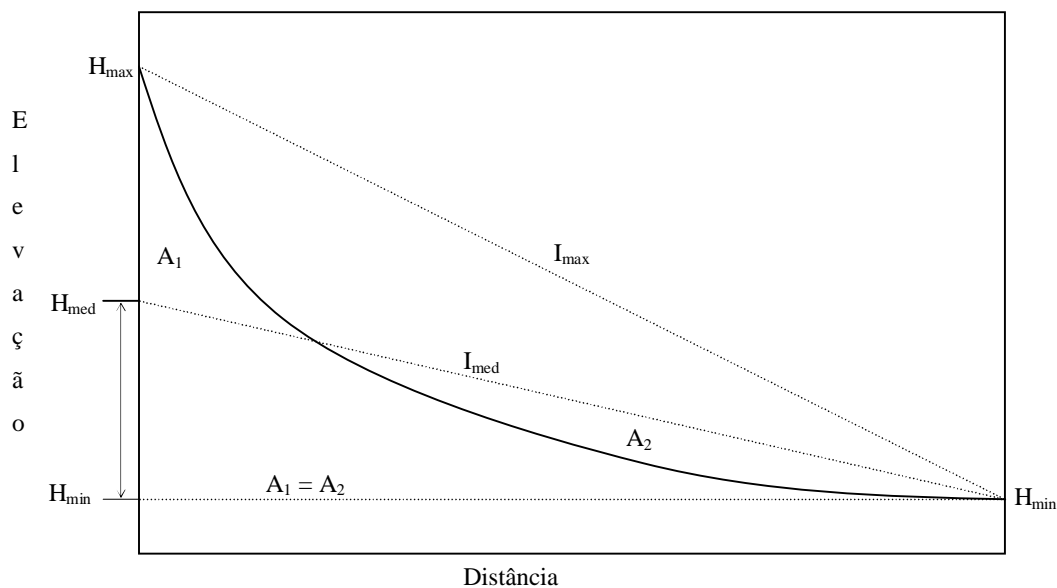


Figura A.6: Esquema das declividades de uma bacia hidrográfica

Curva hipsométrica

O relevo médio de uma bacia pode ser representado pela curva hipsométrica, que é o gráfico que mostra a variação com a elevação das percentagens de áreas compreendidas entre duas curvas de níveis. As percentagens 0% e 100% da área total correspondem às cotas máxima e mínima da bacia, respectivamente. A altitude mediana é aquela correspondente a 50% da área e a altitude média é determinada por meio de um retângulo de área equivalente à limitada pela curva hipsométrica e os eixos coordenados, sendo a altura do retângulo a elevação média, conforme ilustrado na Figura A.7. Segundo Berger (2001) a declividade mediana da superfície é mais representativo do que a declividade média, uma vez que evita a influência de pequenas áreas com declividades excessivas, ou muito inclinadas ou muito planas.

A curva hipsométrica é usada quando uma variável hidrológica tal como a chuva ou a cobertura vegetal, por exemplo, mostra uma tendência marcante para variar com a altitude. Sabe-se também que a altitude exerce influência sobre a temperatura média, a qual por sua vez, causa variações nas perdas por evapotranspiração (Villela, 1975) e, conseqüentemente, nos escoamentos.

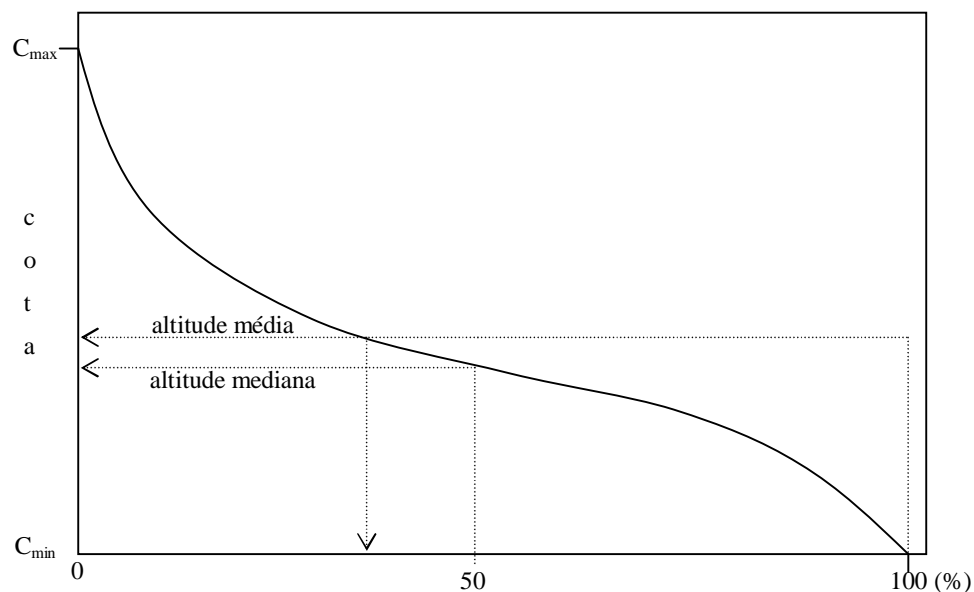


Figura A.7: Curva hipsométrica

Elevação média da bacia (C_{med})

Outro método, equivalente ao da curva hipsométrica, para determinação da altitude média da bacia é dado pela seguinte expressão:

$$C_{med} = \frac{\sum c a}{A}$$

sendo: C_{med} a altitude média da bacia

c a elevação média entre duas curvas de nível consecutivas

a é a área entre as duas curvas de nível

A é a área da bacia hidrográfica

Retângulo equivalente

O retângulo equivalente de uma bacia é o retângulo com o mesmo perímetro, mesmo índice de compacidade e mesma distribuição hipsométrica que a bacia e é caracterizado pelo lado menor l_r e lado maior L_r . Após traçar o retângulo, loca-se as curvas de nível paralelas a l_r obedecendo a distância proporcional natural entre elas.

Sejam: $A = L_r l_r$

$$P = 2(L_r + l_r) = \frac{K_c \sqrt{A}}{0,28}$$

Logo,

$$L_r = \frac{K_c \sqrt{A}}{1,128} \left\{ 1 + \sqrt{\left[1 - \left(\frac{1,128}{K_c} \right)^2 \right]} \right\}$$

$$l_r = \frac{K_c \sqrt{A}}{1,128} \left\{ 1 - \sqrt{\left[1 - \left(\frac{1,128}{K_c} \right)^2 \right]} \right\}$$

Índice de declividade média da bacia (I_p)

É definido pela média ponderada das declividades do conjunto das áreas elementares para as quais pode-se considerar que a linha de maior declividade é constante (Nouvelot, 1974).

$$I_p = \frac{\sum_1^n \sqrt{a_i c_i}}{\sqrt{L_r}}$$

sendo: a_i a fração de área da bacia entre as curvas de nível vizinhas C_i e C_{i-1}

c_i é o desnível entre estas duas curvas de nível

L_r é o comprimento do retângulo equivalente

Índice de declividade global (IG)

Menos preciso que I_p é expresso pela seguinte equação:

$$IG = D/L_r$$

Sendo: D a diferença de altitude em metros entre as duas curvas de nível que abrangem 5% e 95% da área da bacia (localizada abaixo e acima delas).

L_r o comprimento do retângulo equivalente (km)

A determinação de D pode ser feita a partir da curva hipsométrica da bacia ou apenas pela apreciação do mapa da bacia. Certamente IG não tem a mesma precisão que I_p mas a experiência mostra que as ligações entre estes dois índices são suficientemente estreitas para serem significativas (Nouvelot, 1974).

Desnível específico (DS)

É função da forma do contorno da bacia e dado pela seguinte expressão.

$$DS = \frac{D\sqrt{A}}{L_r}$$

A.3 TIPOS DE SOLO

Numa região onde o clima pode ser considerado homogêneo, o regime hidrológico vai depender, essencialmente, da aptidão do solo para infiltração. O critério principal geralmente usado para diferenciação do solo é a permeabilidade (Novelot, 1974).

Devido a existência de vários métodos para a estimativa da permeabilidade do solo (fórmula de Darcy, teste de bombeamento de poços, entre outros), P. Dubreuil, citado por Novelot (1974) propõe o seguinte critério como forma de diminuir a arbitrariedade na escolha de método, no qual a permeabilidade do solo é dividida em três classes, com base no coeficiente de permeabilidade K da fórmula de Darcy ($V = -K dh/ds$).

Solo impermeável: $K < 4 \times 10^{-6}$ m/s (15 mm/h). O escoamento superficial representa mais de 90% do total.

Solo de fraca permeabilidade: $2 \times 10^{-5} > K > 4 \times 10^{-6}$ m/s (15 a 75 mm/h). O escoamento superficial representa entre 60 e 90% do total.

Solo permeável: $K > 2 \times 10^{-5}$ m/s (75 mm/h). O escoamento superficial representa menos de 60% do total.

A.4 COBERTURA VEGETAL E USO DO SOLO

Durante uma pequena chuva as pessoas muitas vezes se abrigam sob uma árvore; elas sabem que as árvores podem abrigá-las por um breve intervalo de tempo, principalmente durante a fase de intercepção inicial da chuva. Isso mostra que uma bacia com cobertura vegetal produz menos escoamento superficial que uma outra sem a cobertura vegetal. Ao mesmo tempo, no caso do telhado de uma residência, a mesma chuva pode responder muito mais rapidamente com superficiais. Os dois exemplos ilustram que o tipo de cobertura do solo afeta significativamente as características do escoamento de uma bacia. Quando outras características de uma bacia hidrográfica, que não a cobertura do solo, são mantidas constantes, as características do escoamento (volume, propagação e picos de cheias) podem diferir significativamente.

Algumas variáveis descritoras da cobertura e uso do solo são usadas em estudos hidrológicos. Em muitos casos, uma descrição qualitativa é transformada em um índice quantitativo do escoamento potencial, por exemplo, o coeficiente de escoamento C da fórmula racional. Devido à falta de dados atualizados sobre o uso do solo e a cobertura vegetal, geralmente se considera apenas a vegetação natural predominante, como: floresta equatorial, tropical, galeria (às margens dos rios), caatinga tropical ou estepe desértica, por exemplo.

ANEXO B: MATRIZ DE SIMILARIDADE

**ANEXO C: CÓDIGO COMPUTACIONAL DO ALGORITMO SCE-UA E
AS SUB-ROTINAS DO MODELO SMAP, DA ENTRADA DE
DADOS EDO CÁLCULO DA FUNÇÃO OBJETIVO**

<pre> C PROGRAM OTSMAP C ***** PROGRAMA PRINCIPAL ***** C Algoritmo de Evolucao do Complexo Misturado C Laudizio da Silva Diniz - 30/03/93 C DIMENSION AX(25,15),AF(25),QQ(25),G(25),PT(25),XI(15), * XS(15),BX(15,15),BF(15),R(15),Z(15),C(15),HI(15),HS(15), * XX(60,15),F(60),XXD(60,15),FD(60),X(15),DX(15), * LL(15),EX(15) INTEGER Q,N,ALFA,BETA,L,M,I,T,J,LX,V,P,S,K,KP,U,INC,E, * PAR,NCH,IMPR REAL ACUM,XXX,FF,D,FR,FZ,FC,FO,EJ,VMXFO REAL*4 randval INTEGER*2 i100th OPEN(30, FILE='LPSMAP.DAT',STATUS='OLD') READ(30,22) VMXFO, EJ 22 FORMAT(2F12.4) C Num. Parametros (N), Num. Complexos (P) e Discretizacao (D) READ(30,25) N,P,D 25 FORMAT(2X,214,F6.1) WRITE(*,*) 'N, P, D ',N,P,D C Calculo de Valores Iniciais IMPR=0 M=2*N+1 S=P*M Q=N+1 ALFA=7-P BETA=2*N+1 NCH=0 c EJ=0.1 C Limite dos parametros DO 1 I=1,N READ(30,98) XI(I),XS(I) DX(I)=(XS(I)-XI(I))/(2.**D) 1 CONTINUE 98 FORMAT(2F12.4) C Amostragem Randomica de S Pontos CALL GETTIM(ihr,imin,isec,i100th) CALL SEED(i100th) CALL RANDOM(randval) DO 2 j=1,S DO 3 i=1,N CALL RANDOM(randval) xx(j,i)=xi(i)+int(randval*(2.**D))*dx(i) x(i)=xx(j,i) 3 CONTINUE CALL SMAP(X,FO,IMPR) NCH=NCH+1 F(J)=FO 2 CONTINUE C Ordena em Ordem Crescente da Funcao Objetivo 9 K=0 DO 4 j=1,s-1 IF(f(j) .GT. F(j+1)) THEN ff=f(j) f(j)=f(j+1) f(j+1)=ff go to 7 ENDIF k=k+1 go to 4 7 DO 5 i=1,N xxx=xx(j,i) xx(j,i)=xx(j+1,i) xx(j+1,i)=xxx 5 CONTINUE 4 CONTINUE IF(K .LT. S-1) GO TO 9 C Particao de Complexos 19 INC=1 DO 10 KP=1,P DO 11 J=1,M E=KP+P*(J-1) DO 12 I=1,N AX(J,I)=XX(E,I) </pre>	<pre> 12 CONTINUE AF(J)=F(E) 11 CONTINUE C Algoritmo de Evolucao Competitiva Complexa C L=1 65 ACUM=0 DO 31 I=1,M ACUM=ACUM+(2.*(M+1-I))/(M*(M+1)) PT(i)=ACUM 31 CONTINUE PT(0)=0. DO 32 t=1,M 32 QQ(t)=0 V=1 CALL GETTIM(ihr,imin,isec,i100th) CALL SEED(i100th) DO 34 i=1,1000 CALL RANDOM(randval) y=randval DO 35 J=0,(M-1) IF(y.GT.PT(j) .AND. y.LE.PT(j+1) .AND. QQ(j+1).EQ.0) THEN LL(V)=j+1 QQ(j+1)=1 V=V+1 ENDIF 35 CONTINUE IF(V .EQ. Q+1) GO TO 36 34 CONTINUE 36 DO 37 J=1,Q DO 38 I=1,N BX(J,I)=AX(LL(J),I) 38 CONTINUE BF(J)=AF(LL(J)) 37 CONTINUE JALFA=1 39 K=0 DO 40 J=1,Q-1 IF(BF(J) .GT. BF(J+1)) then FF=BF(J) BF(J)=BF(J+1) BF(J+1)=FF GO TO 41 ENDIF K=K+1 GO TO 40 41 DO 42 I=1,N XXX=BX(J,I) BX(J,I)=BX(J+1,I) BX(J+1,I)=XXX 42 CONTINUE LX=LL(J) LL(J)=LL(J+1) LL(J+1)=LX 40 CONTINUE IF(K .LT. Q-1) GO TO 39 C Calculo do Centroide DO 43 I=1,N SUM=0 DO 44 J=1,Q-1 SUM=SUM+BX(J,I) 44 CONTINUE G(I)=SUM/(Q-1) 43 CONTINUE C Operacao Reflexao DO 45 I=1,N R(I)=2.*G(I)-BX(Q,I) 45 CONTINUE LIM=0 DO 46 I=1,N IF((R(I) .GE. XI(I)) .AND. (R(I) .LE. XS(I))) THEN X(I)=R(I) LIM=LIM+1 ENDIF 46 CONTINUE </pre>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

```

IF(LIM .EQ. N) THEN
  CALL SMAP(X,FR,IMPR)
  NCH=NCH+1
  GO TO 47
ENDIF
C    Calculo do Menor Hipercubo e Sorteio do Ponto Z
DO 48 I=1,N
  HI(I)=9999999
  HS(I)=-9999999
DO 49 J=1,M
  IF(HI(I) .GT. AX(J,I)) HI(I)=AX(J,I)
  IF(HS(I) .LT. AX(J,I)) HS(I)=AX(J,I)
49 CONTINUE
CALL RANDOM(randval)
Z(I)=HI(I)+((HS(I)-HI(I))/(2.**D))*(INT(randval*(2.**D)))
X(I)=Z(I)
48 CONTINUE
CALL SMAP(X,FZ,IMPR)
NCH=NCH+1
DO 51 I=1,N
  R(I)=Z(I)
51 CONTINUE
FR=FZ
C    Operacao Expansao
47 IF(FR .LT. BF(1)) THEN
  LIM=0
DO 100 I=1,N
  EX(I)=2*R(I)-G(I)
  IF((EX(I) .GE. XI(I)) .AND. (EX(I) .LE. XS(I))) THEN
    X(I)=EX(I)
    LIM=LIM+1
  ENDIF
100 CONTINUE
IF(LIM .NE. N) THEN
DO 113 I=1,N
  CALL RANDOM(randval)
  Z(I)=HI(I)+((HS(I)-HI(I))/(2.**D))*(INT(randval*(2.**D)))
  X(I)=Z(I)
113 CONTINUE
ENDIF
CALL SMAP(X,FEX,IMPR)
NCH=NCH+1
IF(FEX .LT. FR) THEN
DO 101 I=1,N
101 BX(Q,I)=EX(I)
  BF(Q)=FEX
  GO TO 53
ENDIF
ENDIF
IF(FR .LT. BF(Q)) then
DO 52 I=1,N
  BX(Q,I)=R(I)
52 CONTINUE
  BF(Q)=FR
  GO TO 53
ENDIF
C    Operacao Contracao
DO 54 I=1,N
  C(I)=(G(I)+BX(Q,I))/2
  X(I)=C(I)
54 CONTINUE
CALL SMAP(X,FC,IMPR)
NCH=NCH+1
IF(FC .LT. BF(Q)) then
DO 55 I=1,N
  BX(Q,I)=C(I)
55 CONTINUE
  BF(Q)=FC
  GO TO 53
ENDIF
DO 56 I=1,N
  CALL RANDOM(randval)
  Z(I)=HI(I)+((HS(I)-HI(I))/(2.**D))*(INT(randval*(2.**D)))
  X(I)=Z(I)
56 CONTINUE
CALL SMAP(X,FZ,IMPR)
NCH=NCH+1
DO 57 I=1,N
  BX(Q,I)=Z(I)
57 CONTINUE
  BF(Q)=FZ
53 IF(JALFA .LT. ALFA) THEN
  JALFA=JALFA+1
  GO TO 39
ENDIF
DO 58 J=1,Q
DO 59 I=1,N
  AX(LL(J),I)=BX(J,I)
59 CONTINUE
  AF(LL(J))=BF(J)
58 CONTINUE
60 K=0
DO 61 J=1,(M-1)
  IF(AF(J) .GT. AF(J+1)) THEN
    ff=AF(J)
    AF(J)=AF(J+1)
    AF(J+1)=ff
  GO TO 63
ENDIF
  K=K+1
  GO TO 61
63 DO 62 I=1,N
  xxx=AX(J,I)
  AX(J,I)=AX(J+1,I)
  AX(J+1,I)=xxx
62 CONTINUE
61 CONTINUE
  IF(K .LT. (M-1)) GO TO 60
  IF(L .GE. BETA) GO TO 71
  L=L+1
  GO TO 65
71 DO 13 T=1,M
DO 14 U=1,N
  XXD(INC,U)=AX(T,U)
14 CONTINUE
  FD(INC)=AF(T)
  INC=INC+1
13 CONTINUE
10 CONTINUE
DO 80 T=1,S
DO 81 U=1,N
  XX(T,U)=XXD(T,U)
81 CONTINUE
  F(T)=FD(T)
80 CONTINUE
C    Ordena em Ordem Crescente da Funcao Objetivo
18 K=0
DO 15 J=1,S-1
  IF(F(J) .GT. F(j+1)) then
    ff=f(j)
    f(j)=f(j+1)
    f(j+1)=ff
  go to 17
ENDIF
  K=K+1
  GO TO 15
17 DO 16 I=1,N
  xxx=xx(j,i)
  xx(j,i)=xx(j+1,i)
  xx(j+1,i)=xxx
16 CONTINUE
15 CONTINUE
  IF(K .LT. S-1) GO TO 18
  WRITE(*,*) 'FO(1) = ',F(1)
  WRITE(*,*) 'VALORES DOS PARAMETROS'
  DO 50 par=1,N
  WRITE(*,*) xx(1,par)
50 CONTINUE
  AVG=0
  DPD=0
  DO 107 II=1,S
  107 AVG=AVG+F(II)
  AVG=AVG/S

```


<pre> DO 102 I=1,S 102 DPD=DPD+(F(I)-AVG)**2 DPD=SQRT(DPD/S) IF(DPD .LT. EJ .OR. F(1) .LE. VMXFO) GO TO 72 GO TO 19 72 CONTINUE WRITE(*,*) 'Valor Otimo de FO' WRITE(*,*) F(1) WRITE(*,*) 'Valores Otimos dos Parametros' WRITE(*,*) DO 21 I=1,N WRITE(*,*) XX(1,I) X(I)=XX(1,I) 21 CONTINUE IMPR=1 CALL SMAP(X,FO,IMPR) STOP END SUBROUTINE SMAP(X,FO,IMPR) CHARACTER DOUT*20 DIMENSION X(15),QO(27,12),QOA(27),QC(27,12),QCA(27), 1 PREC(27,12), EVTP(12),EVTPA(27),EVAPR(27,12), 2 EVAPRA(27),ANO(27),PRECA(27),NDM(31), 3 CEO(27,12),CEC(27,12) INTEGER ANOI,ANOF,ANO REAL KK,FO,PREC,PRECA,ED C Parametros Calibraveis do Modelo STR=X(1) CINF1=X(2) E1=X(3) E2=X(4) C Subrotina para Leitura de Dados LE=LE+1 WRITE(*,*) 'CHAMADA ',LE IF(LER .NE. 1) THEN CALL LEIA(ANOI,ANOF,PREC,QO,EVTP,ANO,SI,SUBIN) LER=1 ENDIF IF(IMPR .EQ. 1) THEN PRINT *, 'DIGITE O NOME EXATO DO ARQUIVO DE SAIDA.' READ(*,7) DOUT 7 FORMAT(A) OPEN(40, FILE=DOUT,STATUS='OLD') WRITE(40,25) 25 FORMAT(20X,20HPARAMETROS DO MODELO,/, ' STR 1 CINF1 E1 E2 KK CREC') WRITE(40,26) STR,CINF1,E1,E2,KK,CREC 26 FORMAT(F8.2,5F8.3,/) WRITE(40,27) 27 FORMAT(20X,'DINAMICA DOS RESERVATORIOS',/, ' ANO MES PREC EVTP *EVAPR QO QC ED ES EB RSOL SS REC * CEO CEC TU') ENDIF C CALCULO DA LAMINA ESCOADA NANO=ANOF-ANOI+1 RSOL=SI*STR C SS=(SUBIN/(1-KK)) DO 10 I=1,NANO DO 20 N=1,12 RSOL=RSOL+PREC(I,N)*(1-CINF1) ED=RSOL-STR IF(ED .LT. 0) ED=0 RSOL=RSOL-ED TU=RSOL/STR CES=RSOL*TU**E2 ES=(1-CINF1)*PREC(I,N)*TU**E2 IF(ES .GT. RSOL) ES=RSOL RSOL=RSOL-ES ER=EVTP(N)*TU**E1 IF(ER .GT. EVTP(N)) ER=EVTP(N) </pre>	<pre> IF(ER .GT. RSOL) ER=RSOL RSOL=RSOL-ER C REC=RSOL*CREC*TU**10 C RSOL=RSOL-REC C SS=SS+REC C EB=SS*(1-KK) C SS=SS-EB QC(I,N)=ES+ED EVAPR(I,N)=ER IF(PREC(I,N) .NE. 0) THEN CEO(I,N)=QO(I,N)*100/PREC(I,N) CEC(I,N)=QC(I,N)*100/PREC(I,N) ENDIF IF(IMPR .EQ. 1) THEN WRITE(40,11) ANO(I),N,PREC(I,N),EVTP(N),EVAPR(I,N),QO(I,N), 1 QC(I,N),ED,ES,EB,RSOL,SS,REC,CEO(I,N),CEC(I,N),TU 11 FORMAT(I5,I3,3F6.1,5F7.1,F8.2,5F6.2) ENDIF 20 CONTINUE 10 CONTINUE C*** BLOCO DE IMPRESSAO IF(IMPR .EQ. 1) THEN WRITE(40,33) 33 FORMAT(/,15X,'QUADRO DE VALORES ANUAIS ACUMULADOS',/, ' ANO 1PREC EVATP EVAPRA VAZ.OBS.VAZ.CAL. CEO CEC EREL. ') DO 30 I=1,NANO PRECA(I)=0.0 EVTPA(I)=0.0 EVAPRA(I)=0.0 QOA(I)=0.0 QCA(I)=0.0 30 CONTINUE DO 32 I=1,NANO DO 31 N=1,12 PRECA(I)=PRECA(I)+PREC(I,N) EVTPA(I)=EVTPA(I)+EVTP(N) EVAPRA(I)=EVAPRA(I)+EVAPR(I,N) QOA(I)=QOA(I)+QO(I,N) QCA(I)=QCA(I)+QC(I,N) 31 CONTINUE WRITE(40,34) ANO(I),PRECA(I),EVTPA(I),EVAPRA(I),QOA(I),QCA(I), 1 (QOA(I)/PRECA(I)),(QCA(I)/PRECA(I)),(QOA(I)- QCA(I))/QOA(I) 34 FORMAT(I4,5F8.1,3F8.2) 32 CONTINUE ENDIF CALL FUNOB(QC,QO,NANO,FO,R) IF(IMPR .EQ. 1) THEN WRITE(40,28) FO,R 28 FORMAT(30X,'FUNCAO OBJETIVO : FO =',2X,F8.3,/,30X,'COEF. *DE CORRELACAO (%) : R =',2X,F8.2,/) ENDIF IF(IMPR .EQ. 1) THEN OPEN(50, FILE='GRAFM.DAT',STATUS='OLD') WRITE(50,8) ANOI, ANOF 8 FORMAT(2I8) DO 22 I=1,NANO WRITE(50,19) ANO(I),(PREC(I,N),N=1,12) 22 CONTINUE DO 23 I=1,NANO WRITE(50,19) ANO(I),(QO(I,N),N=1,12) 23 CONTINUE DO 24 I=1,NANO WRITE(50,19) ANO(I),(QC(I,N),N=1,12) 24 CONTINUE ENDIF 19 FORMAT(I8,12F6.1) RETURN END </pre>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

```

SUBROUTINE FUNOB(QC,QO,NANO,FO,R)
DIMENSION QC(27,12),QO(27,12),QCC(27,12),QOO(27,12)
INTEGER NANO,I,N
REAL FO,SX,SY,SX2,SY2,SXY,R
FO=0
SX=0
SY=0
SXY=0
SX2=0
SY2=0
DO 50 I=1,NANO
DO 40 N=1,12
FO = FO + (SQRT(QO(I,N))-SQRT(QC(I,N)))**2.
SXY=SXY+SQRT(QO(I,N))*SQRT(QC(I,N))
SX=SX+SQRT(QO(I,N))
SY=SY+SQRT(QC(I,N))
SX2=SX2+SQRT(QO(I,N))**2.
SY2=SY2+SQRT(QC(I,N))**2.
40 CONTINUE
50 CONTINUE
NP=NANO*12
ME=0.
QMED=SX/NP
DO 70 I=1,NANO
DO 60 N=1,12
ME=ME+(SQRT(QO(I,N))-QMED)**2
60 CONTINUE
70 CONTINUE
FO=(FO/ME)
C   FO=SQRT(FO/NP)
R=(NP*SXY-SX*SY)/(SQRT((NP*SX2-SX**2.)*(NP*SY2-SY**2.)))
R=R*100
WRITE(*,*) 'FO   =',FO
WRITE(*,*) 'R (%) =',R
RETURN
END

```

```

SUBROUTINE LEIA(ANOI,ANOF,PREC,QO,EVTP,ANO,SI,SUBIN)
DIMENSION PREC(27,12),QO(27,12),EVTP(12)
INTEGER I,NANO,ANOI,ANOF,N,ANO(27),NDM(12)
REAL MED,DP
CHARACTER DIN*20
PRINT *, 'DIGITE O NOME EXATO DO ARQUIVO DE DADOS
HIDROLOGICOS.'
READ(*,10) DIN
OPEN(20,FILE=DIN, STATUS='OLD')
NDM(1)=31
NDM(2)=30
NDM(3)=31
NDM(4)=31
NDM(5)=28
NDM(6)=31
NDM(7)=30
NDM(8)=31
NDM(9)=30
NDM(10)=31
NDM(11)=31
NDM(12)=30
10  FORMAT(A)
READ(20,2) TITULG
C   WRITE(10,2) GERAL
2   FORMAT(20A4)
READ(20,3) ANOI,ANOF,AREA,SI,SUBIN,CT
3   FORMAT(2X,2I6,F7.1,3F7.2)
NANO=ANOF-ANOI+1
C   ARQUIVO DE PRECIPITACAO
READ(20,2) TITULP
DO 4 I=1,NANO
READ(20,5) ANO(I),(PREC(I,N),N=1,12)
4   CONTINUE
5   FORMAT(18,12F6.1)
C   PARAMETROS ESTATISTICOS DA CHUVA
MED=0
DP=0

```

```

DO 33 I=1,NANO
DO 34 N=1,12
MED=MED+PREC(I,N)
34  CONTINUE
33  CONTINUE
MED=MED/(NANO*12)
DO 35 I=1,NANO
DO 37 N=1,12
DP=DP+(PREC(I,N)-MED)**2
37  CONTINUE
35  CONTINUE
DP=SQRT(DP/(NANO*12-1))
WRITE(*,*) 'MEDIA   =',MED
WRITE(*,*) 'DESVIO PADRAO =',DP
C   ARQUIVO DE VAZoes
READ(20,2) TITULV
DO 6 I=1,NANO
READ(20,9) ANO(I),(QO(I,N),N=1,12)
6   CONTINUE
9   FORMAT(18,12F6.2)
C   ARQUIVO DE EVAPOTRANSPIRACAO
POTENCIAL
READ(20,2) TITULE
READ(20,7) (EVTP(N),N=1,12)
7   FORMAT(8X,12F6.1)
DO 8 N=1,12
8   EVTP(N)=CT*EVTP(N)
CLOSE(20)
RETURN
END

```

O PROGRAMA PRINCIPAL É O MESMO: A SEGUIR
AS SUBROTINAS QUE SOFRERAM ALTERAÇÕES
PARA INCORPORAR O BALANÇO HÍDRICO DOS
ACUDES

```

SUBROUTINE SMAP(X,FO,IMPR)
CHARACTER DOUT*20
DIMENSION
X(15),QO(15,12),QOA(15),QC(15,12),QCA(15),PREC(15,12),
1  EVTP(12),EVTPA(15),EVAPR(15,12),EVAPRA(15),ANO(15),
2  PRECA(15),NDM(31),CEC(15,12),COT(50),DEM(15,12),
3  ARE(50),VOL(50),VOLUME(15,12),ABH(15,12)
INTEGER ANOI,ANOF,ANO
REAL KK,FO,PREC,PRECA,CT,VMAX
C  Parametros Calibraveis do Modelo
STR=X(1)
CINF1=X(2)
E1=X(3)
E2=X(4)
C  Subrotina para Leitura de Dados
LE=LE+1
WRITE(*,*) 'CHAMADA ',LE
IF(LER .NE. 1) THEN
CALL LEIA(ANOI,ANOF,PREC,QO,EVTP,ANO,SI,SUBIN,AREA,
*COT,ARE,VOL,DEM,CT,VMAX)
LER=1
ENDIF
IF(IMPR .EQ. 1) THEN
c  PRINT *, 'DIGITE O NOME EXATO DO ARQUIVO DE SAIDA.'
c  READ(*,7) DOUT
c 7  FORMAT(A)
OPEN(40, FILE='patosI.out',STATUS='OLD')
WRITE(40,25)
25  FORMAT(20X,20HPARAMETROS DO MODELO,/, ' STR CINF1
E1
1  E2  KK  CREC')
WRITE(40,26) STR,CINF1,E1,E2,KK,CREC
26  FORMAT(F8.2,5F8.3,/)
WRITE(40,27)
27  FORMAT(20X,'DINAMICA DOS RESERVATORIOS',/, ' ANO MES
PREC EVTP
1  EVAPR  QC  ED  ES  EB  RSOL  SS  REC
2  CEC  TU')
ENDIF
C*** CALCULO DA LAMINA ESCOADA
NANO=ANOF-ANOI+1
RSOL=SI*STR
C  SS=(SUBIN/(1-KK))
DO 10 I=1,NANO
DO 20 N=1,12
RSOL=RSOL+PREC(I,N)*(1-CINF1)
ED=RSOL-STR
IF(ED .LT. 0) ED=0
RSOL=RSOL-ED
TU=RSOL/STR
CES=RSOL*TU**E2
ES=(1-CINF1)*PREC(I,N)*TU**E2
IF(ES .GT. RSOL) ES=RSOL
RSOL=RSOL-ES
ER=EVTP(N)*TU**E1
C  ER=RSOL*(1-EXP(-EVTP(N)/E1))
C  ER=RSOL*(1-E1**EVTP(N))
IF(ER .GT. EVTP(N)) ER=EVTP(N)
IF(ER .GT. RSOL) ER=RSOL
CWRITE(*,*) 'EVTP(N),E1,ER',EVTP(N),E1,ER,RSOL,ES
RSOL=RSOL-ER
C  REC=RSOL*CREC*TU**10
C  RSOL=RSOL-REC
C  SS=SS+REC
C  EB=SS*(1-KK)
C  SS=SS-EB

```

```

QC(I,N)=ES+ED
EVAPR(I,N)=ER
IF(PREC(I,N) .NE. 0) THEN
CEC(I,N)=QC(I,N)*100/PREC(I,N)
ENDIF
IF(IMPR .EQ. 1) THEN
WRITE(40,11)
ANO(I),N,PREC(I,N),EVTP(N),EVAPR(I,N),
1  QC(I,N),ED,ES,EB,RSOL,SS,REC,CEC(I,N),TU
11  FORMAT(15,I3,3F6.1,7F8.2,F6.1,F6.2)
ENDIF
20  CONTINUE
10  CONTINUE
C*** BLOCO DE IMPRESSAO
IF(IMPR .EQ. 1) THEN
WRITE(40,33)
33  FORMAT(/,28X,'QUADRO DE BALANCO HIDRICO',/, '
ANO VOLUME PRE
1C  EVATP AREA ESPELHO EVAPORACAO
DEMANDA VOL.AFLUENTE SALDO'
2)
DO 32 I=1,NANO
DO 31 N=1,12
WRITE(40,34)
ANO(I),QO(I,N),PREC(I,N),EVTP(N),ABH(I,N),ABH(I,N)*(E
1VTP(N)-
0.2*PREC(I,N))/1000,DEM(I,N),QC(I,N)*AREA*1000,VOLUME(I,
N)
34  FORMAT(14,F10.0,2F7.1,1X,F10.0,3X,F10.0,5X,F7.0,2X,2F10.0
)
31  CONTINUE
32  CONTINUE
ENDIF
CALL
FUNOB(QC,QO,NANO,FO,R,EVTP,CT,AREA,PREC,COT,AR
E,VOL,DEM,
1ABH,VOLUME,VMAX)
IF(IMPR .EQ. 1) THEN
WRITE(40,28) FO,R
28  FORMAT(30X,'FUNCAO OBJETIVO : FO
=',2X,F8.3,/,30X,'COEF.
*DE CORRELACAO (%) : R =',2X,F8.2,/)
ENDIF
CLOSE(40)
RETURN
END

```

```

SUBROUTINE
FUNOB(QC,QO,NANO,FO,R,EVTP,CT,AREA,PREC,COT,AR
E,VOL,DEM
1,ABH,VOLUME,VMAX)
DIMENSION
QC(15,12),QO(15,12),EVTP(12),PREC(15,12),ABH(15,12)
DIMENSION DEM(15,12),Z(15,12)
DIMENSION
COT(50),ARE(50),VOL(50),VOLUME(15,12),VOLO(200),VOLC(2
00)
INTEGER NANO,I,N,NP,T
REAL FO,SX,SY,SY2,SXY,R,CT,AREA,ME,VMAX
VI=QO(1,1)
VOLUME(1,1)=VI
DIF=0
T=0
N=0
CALL INTERP(COT,ARE,VOL,VI,ARR)
ABH(1,1)=ARR
DO 7 I=1,NANO
L=1
IF(I .EQ. 1 .AND. N .EQ. 0) L=2
DO 8 N=L,12
T=T+1
IF(N .EQ. 1) ABH(I,N)=ABH(I-1,12)
IF(N .NE. 1) ABH(I,N)=ABH(I,N-1)
Z(I,N)=ABH(I,N)

```

```

26 IF(N.EQ. 1) VOLU(I,N)=VOLU(I-1,12)+QC(I-1,12)*AREA*1000-
*DEM(I-1,12)-((Z(I,N)+ABH(I,N))/2)*(EVTP(12)-0.2*PREC(I-
1,12))/1000
IF(N.NE. 1) VOLU(I,N)=VOLU(I,N-1)+QC(I,N-1)*AREA*1000-
DEM(I,N-1)-
*((Z(I,N)+ABH(I,N))/2)*(EVTP(N-1)-0.2*PREC(I,N-1))/1000
VI=VOLU(I,N)
IF(VI.LT. 0) VI=0
IF(VI.GT. VMAX) VI=VMAX
CALL INTERP(COT,ARE,VOL,VI,ARR)
DIF=DIF+1
ABH(I,N)=ARR
IF(DIF.LE. 5) GOTO 26
C Transformação de variável
IF(QO(I,N).GE. 0) THEN
IF(VOLU(I,N).GE. 0) THEN
VOLO(T)=SQRT(QO(I,N))
VOLC(T)=SQRT(VOLU(I,N))
ELSE
VOLO(T)=SQRT(QO(I,N))
VOLC(T)=VOLU(I,N)
ENDIF
ELSE
IF(VOLU(I,N).GE. 0) THEN
VOLO(T)=QO(I,N)
VOLC(T)=SQRT(VOLU(I,N))
ELSE
VOLO(T)=QO(I,N)
VOLC(T)=VOLU(I,N)
ENDIF
ENDIF
DIF=0
8 CONTINUE
7 CONTINUE
FO=0
SX=0
SY=0
SXY=0
SX2=0
SY2=0
NP=0
DO 50 I=1,T
IF (VOLO(I).GT. 0 .AND. VOLO(I).LT. SQRT(VMAX)) THEN
NP=NP+1
FO = FO + (VOLO(I)-VOLC(I))**2.
SXY=SXY+VOLO(I)*VOLC(I)
SX=SX+VOLO(I)
SY=SY+VOLC(I)
SX2=SX2+VOLO(I)**2.
SY2=SY2+VOLC(I)**2.
ENDIF
50 CONTINUE
ME=0.
QMED=SX/NP
DO 75 I=1,T
IF (VOLO(I).GT. 0 .AND. VOLO(I).LT. SQRT(VMAX)) THEN
ME=ME+(VOLO(I)-QMED)**2.
ENDIF
75 CONTINUE
FO=(FO/ME)
C FO=ABS((SY-SX)/SX)
R=(NP*SXY-SX*SY)/(SQRT((NP*SX2-SX**2.)*(NP*SY2-SY**2.)))
R=R*100
WRITE(*,*) 'FO =',FO
WRITE(*,*) 'R (%) =',R
RETURN
END
SUBROUTINE INTERP(COT,ARE,VOL,VI,ARR)
DIMENSION COT(50),ARE(50),VOL(50)
INTEGER I,NT,K
REAL KV
c ===== SUBROTINA PARA INTERPOLACAO DA COTA E AREA
=====
K=1
IF (VI.EQ. 0) THEN
COTA=0

```

```

ARR=0
RETURN
ENDIF
15 IF (VI.EQ.VOL(K)) THEN
COTA=COT(K)
ARR=ARE(K)
RETURN
ENDIF
IF (VI.GT. VOL(K).AND. VI.LT. VOL(K+1)) THEN
KV=(VOL(K)-VI)/(VOL(K)-VOL(K+1))
ARR=ARE(K)-(ARE(K)-ARE(K+1))*KV
COTA=COT(K)-(COT(K)-COT(K+1))*KV
RETURN
ENDIF
K=K+1
GOTO 15
END

SUBROUTINE
LEIA(ANOI,ANOF,PREC,QO,EVTP,ANO,SI,SUBIN,AREA,
*COT,ARE,VOL,DEM,CT,VMAX)
DIMENSION PREC(15,12),QO(15,12),EVTP(12)
DIMENSION COT(50),ARE(50),VOL(50),DEM(15,12)
INTEGER I,NANO,ANOI,ANOF,N,ANO(15),NDM(12)
REAL VMAX,MED
CHARACTER DIN*20
OPEN(20, FILE='ARQG.IN', STATUS='OLD')
PRINT *, 'DIGITE O NOME EXATO DO ARQUIVO DE
DADOS HIDROLOGICOS.'
READ(*,10) DIN
OPEN(20,FILE=DIN, STATUS='OLD')
NDM(1)=31
NDM(2)=30
NDM(3)=31
NDM(4)=31
NDM(5)=28
NDM(6)=31
NDM(7)=30
NDM(8)=31
NDM(9)=30
NDM(10)=31
NDM(11)=31
NDM(12)=30
10 FORMAT(A)
READ(20,2) TITULG
2 WRITE(10,2) GERAL
2 FORMAT(20A4)
READ(20,3) ANOI,ANOF,AREA,SI,SUBIN,CT
3 FORMAT(2X,2I6,F7.1,3F7.2)
NANO=ANOF-ANOI+1
C ARQUIVO DE PRECIPITACAO
READ(20,2) TITULP
DO 4 I=1,NANO
READ(20,5) ANO(I),(PREC(I,N),N=1,12)
4 CONTINUE
5 FORMAT(I8,12F6.1)
C PARAMETROS ESTATISTICOS DA CHUVA
MED=0
DP=0
DO 33 I=1,NANO
DO 34 N=1,12
MED=MED+PREC(I,N)
34 CONTINUE
33 CONTINUE
MED=MED/(NANO*12)
DO 35 I=1,NANO
DO 37 N=1,12
DP=DP+(PREC(I,N)-MED)**2
37 CONTINUE
35 CONTINUE
DP=SQRT(DP/(NANO*12-1))
WRITE(*,*) 'MEDIA =',MED
WRITE(*,*) 'DESVIO PADRAO =',DP
READ(*,*)
C ARQUIVO DE VAZUES

```

<pre>READ(20,2) TITULV DO 6 I=1,NANO READ(20,9) ANO(I),(QO(I,N),N=1,12) 6 CONTINUE 9 FORMAT(18,12F10.0) C ARQUIVO DE EVAPOTRANSPIRACAO POTENCIAL READ(20,2) TITULE READ(20,7) (EVTP(N),N=1,12) 7 FORMAT(8X,12F6.1) DO 8 N=1,12 8 EVTP(N)=CT*EVTP(N) READ(20,12) TITULD DO 36 I=1,NANO READ(20,9) ANO(I),(DEM(I,N),N=1,12) 36 CONTINUE READ(20,12) TITULV READ(20,13) NT,VMAX DO 16 I=1,NT READ(20,19) COT(I),ARE(I),VOL(I) 16 CONTINUE 19 FORMAT(3F10.0) 12 FORMAT(20A4) 13 FORMAT(17,F10.0) CLOSE(20) RETURN END</pre>	
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--