

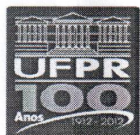
João Pedro Bazzo Vieira

Regionalização de Vazões com o Emprego de Redes Neurais Artificiais RBF

Trabalho apresentado como requisito parcial à obtenção do grau de bacharel em Engenharia Ambiental, Setor de Tecnologia da Universidade Federal do Paraná.

Orientador: Dr. Ricardo Carvalho de Almeida

Curitiba
2016



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE TECNOLOGIA
CURSO DE GRADUAÇÃO EM ENGENHARIA AMBIENTAL


TERMO DE APROVAÇÃO DE PROJETO FINAL

JOÃO PEDRO BAZZO VIEIRA

REGIONALIZAÇÃO DE VAZÕES COM O EMPREGO DE REDES NEURAIS ARTIFICIAIS RBF

Projeto Final de Curso, aprovado como requisito parcial para a obtenção do Diploma de Bacharel em Engenharia Ambiental no Curso de Graduação em Engenharia Ambiental do Setor de Tecnologia da Universidade Federal do Paraná, com nota **100**, pela seguinte banca examinadora:

Orientador: 
Ricardo Carvalho de Almeida, DSc., UFPR/DEA
Orientador

Membro 1: 
Akemi Kan, DSc., Solver Engenharia

Membro 2: 
Michael Mannich, DSc., UFPR/DEA

Curitiba, 04 de julho de 2016

Agradecimentos

Ao longo desses seis anos e meio de graduação, foram várias experiências, pessoas incríveis, muitos bons exemplos e ideias que pude levar. Sou grato a todos que colaboraram, de alguma forma, para a finalização desta graduação. Não poderia, no entanto, deixar de agradecer a algumas pessoas que foram cruciais nessa jornada.

Agradeço à minha família, meus pais Vieira e Iara e minha irmã Ana, por todo suporte, por acreditarem em mim e me estimularem sempre a sonhar grande.

À minha namorada e melhor amiga, Julia, por estar sempre ao meu lado. Obrigado pela parceria e por vibrar com todas as minhas conquistas.

Ao Ricardo Almeida, por ser meu professor e orientador, pelo exemplo e acompanhamento durante todo o Projeto Final. Este processo foi um aprendizado e tanto!

A banca avaliadora desse Projeto Final, Michael e Akemi, pelas críticas construtivas e contribuições ao longo do T.C.C. I e II.

A Empresa Júnior de Engenharia Ambiental & Bioprocessos e Biotecnologia, Ciclus Consultoria, por ser a maior imersão em empreendedorismo que um estudante de engenharia ambiental pode ter. Obrigado por inspirar valores.

Aos colegas do Centro de Apoio às Promotorias de Meio Ambiente do MP-PR, Luciane, Ellery, Paula, Mayara, Izabel e Wanderléia, obrigado pela amizade, pelo aprendizado proporcionado, e por me mostrarem o grande desafio de se trabalhar com questões ambientais dentro do setor público.

Aos meus amigos e professores da Universidade of Waterloo e a turma do *Science*, em especial ao Daniel e Jefferson, obrigado pela amizade e por terem dividido experiências acadêmicas e culturais que levarei para sempre.

A toda equipe da EnvEx Engenharia, em especial a Débora, Paulo, Roberta, André, Helder e Sandro; por me mostrarem aplicações das disciplinas do curso na

realidade das indústrias e na gestão ambiental de municípios. Foi bastante gratificante trabalhar com vocês.

Aos meus amigos, com quem dividi todas essas experiências e por serem motivo de tornar uma extensa e difícil graduação, muito mais divertida. Obrigado ao Rodrigo, Patrícia, Gabriele, Rafael, Érika, Melina, Leandro, Natália, Rena, Gabriel, Mateus e Einara.

Ao Leonardo, pelo interessante estudo sobre regionalização de vazões, e por fornecer os dados que foram utilizados neste trabalho.

A todos os professores da Universidade com quem interagi ao longo da graduação. Obrigado por me inspirar e me provocar a sempre aprender. As lições desta época estarão sempre vivas.

Por fim, a UFPR, como instituição, e ao CNPq, pelo suporte financeiro durante o Programa Ciências Sem Fronteiras. Graças a elas, pude ter as mais intensas experiências nesse período universitário, com aprendizados pessoais, profissionais e culturais que nunca pude imaginar. Tudo isso com financiamento público, em um País de grande desigualdade social e com soluções ambientais tímidas e insuficientes, que espero um dia poder retribuir.

Sumário

Resumo	xiii
1 Introdução	1
2 Objetivos	4
2.1 Objetivo Geral	4
2.2 Objetivos Específicos	4
3 Revisão Bibliográfica	5
3.1 Regionalização Hidrológica	5
3.2 Modelagem chuva-vazão	8
3.3 Redes Neurais Artificiais	9
3.3.1 Cérebro Humano e Neurônio Artificial	9
3.3.2 Redes Neurais tipo <i>Perceptron</i> Multicamadas	11
3.3.3 Redes Neurais tipo Função de Base Radial	13
3.4 Técnicas de Aprendizado	15
3.4.1 Mapas Auto-Organizados de Kohonen	16
3.5 Aplicações de RNAs em Recursos Hídricos	18
4 Metodologia	20
4.1 Metodologia Proposta	20
4.2 Área de Estudo	22
4.3 Dados Utilizados no Estudo	24

4.4	Pré-tratamento e Consistência dos Dados	25
4.5	Determinação das características das bacias	27
4.6	Seleção das variáveis relevantes para a Análise de Agrupamento	29
4.7	Seleção dos grupos semelhantes	30
4.8	Modelagem Chuva-Vazão	31
4.9	Regionalização de Vazões	32
4.9.1	Métricas de Validação	34
5	Resultados	36
5.1	Seleção dos grupos semelhantes	36
5.2	Modelagem Chuva-Vazão	40
5.3	Regionalização de Vazões	41
5.3.1	Modelo Chuva-Vazão I	41
5.3.2	Modelo Chuva-Vazão II	50
6	Discussão	56
6.1	Seleção dos grupos semelhantes	56
6.2	Modelagem Chuva-Vazão	58
6.3	Regionalização de Vazões	58
7	Conclusão	62

Lista de Tabelas

3.1	Diferenças entre as RNAs tipo <i>Radial Basis Function</i> e RNAs <i>Multi-layer Perceptron</i>	13
4.1	Estações fluviométricas e pluviométricas utilizadas no estudo. Fonte: Minucci (2014).	26
4.2	Características físicas e fisiográficas das bacias hidrográficas a ser utilizadas no estudo. Fonte: Minucci (2014).	28
4.3	Resultado da análise de componentes principais dos parâmetros levantados para as estações fluviométricas estudadas (Minucci, 2014).	29
5.1	Agrupamento A — formado a partir das quatro variáveis físicas/fisiográficas das bacias hidrográficas.	36
5.2	Agrupamento B — formado a partir das 11 variáveis físicas/fisiográficas das bacias hidrográficas.	37
5.3	Parâmetros de configuração de treinamento e operação da rede RBF, para o Modelo Chuva-Vazão I e II.	41
5.4	Estatísticas médias da calibração e operação do Modelo Chuva-Vazão I (Modelo de 4 entradas, 1 saída) de cada estação auxiliar.	42
5.5	Estatísticas médias da calibração e operação do Modelo Chuva-Vazão II (Modelo de 13 entradas, 1 saída) de cada estação auxiliar.	43

5.6	Estatísticas da regionalização para a estação teste Madeira Gavazzoni realizadas através do Modelo Chuva-Vazão I e do agrupamento A.	43
5.7	Estatísticas da regionalização para a estação teste Madeira Gavazzoni realizadas através do Modelo Chuva-Vazão I e do agrupamento B.	45
5.8	Estatísticas da regionalização para a estação teste Guampará realizadas através do Modelo Chuva-Vazão I e do agrupamento A.	47
5.9	Estatísticas da regionalização para a estação teste Guampará realizadas através do Modelo Chuva-Vazão II e do agrupamento B.	48
5.10	Estatísticas da regionalização para a estação teste Madeira Gavazzoni realizadas através do Modelo Chuva-Vazão II e do agrupamento A.	50
5.11	Estatísticas da regionalização para a estação teste Madeira Gavazzoni realizadas através do Modelo Chuva-Vazão II e do agrupamento B.	50
5.12	Estatísticas da regionalização para a estação teste Guampará realizadas através do Modelo Chuva-Vazão II e do agrupamento A.	53
5.13	Estatísticas da série regionalizada através do Modelo Chuva-Vazão II e do agrupamento B - a estação Guampará.	54
6.1	<i>Clusters</i> formados a partir de diferentes entradas do algoritmo SOM e o método <i>k-means</i> - proposto por Minucci (2014).	57
6.2	Resumo dos valores R^2 e NASH obtidos para os diferentes modelos e agrupamento utilizando, assim como as estatísticas encontradas por Minucci (2014).	59
6.3	Comparação entre os resultados obtidos por Minucci (2014) e pelo Autor (2016), com os valores observados - estação Madeira Gavazzoni.	60

6.4	Resultados obtidos por Minucci (2014) e pelo Autor (2016), com os valores observados - estação Guampará.	61
-----	--	----

Lista de Figuras

3.1	Representação do diagrama de blocos do sistema nervoso.	10
3.2	Neurônio Artificial.	11
3.3	Arquitetura de uma RNA do tipo <i>multilayer perceptron</i>	12
3.4	Rede Neural do tipo <i>Radial Basis Function</i>	14
3.5	Topologia de um mapa de Kohonen. Fonte: Adaptado de Haykin (1999) e Fausett (1994).	17
4.1	Abordagem adotada no estudo para regionalizar registros de vazões.	21
4.2	Estações hidrometeorológicas utilizadas no estudo. Fonte: Minucci (2014).	23
4.3	Abordagem adotada para aplicação dos agrupamentos.	31
4.4	Esquema adotado para a modelagem chuva-vazão.	32
4.5	Representação esquemática da abordagem adotada para o cálculo da vazão modelada média a partir de i estações.	33
4.6	Exemplo de valores individuais modelados ($Q_{Modeladas}$ 1, 2, 3 e 4) e valores modelados médios.	34
5.1	Agrupamento A — formado a partir de quatro variáveis físicas/fisiográficas das bacias hidrográficas. Fonte: Adaptado de Minucci (2014).	38
5.2	Agrupamento B — formado a partir de 11 variáveis físicas/fisiográficas das bacias hidrográficas. Fonte: Adaptado de Minucci (2014).	39

5.3	Relação Vazão específica média modelada x Vazão específica observada para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão I e do agrupamento A.	44
5.4	Vazão específica média modelada e observada ao longo do tempo para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão I e do agrupamento A.	44
5.5	Relação Vazão específica média modelada x Vazão específica observada para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão I e do agrupamento B.	45
5.6	Vazão específica média modelada e observada ao longo do tempo para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão I e do agrupamento B.	46
5.7	Vazão específica média modelada e observada ao longo do tempo para a estação teste Guampará através do Modelo Chuva-Vazão I e do agrupamento A.	47
5.8	Relação Vazão específica média modelada x Vazão específica observada para a estação teste Guampará através do Modelo Chuva-Vazão I e do agrupamento A.	48
5.9	Relação Vazão específica média modelada x Vazão específica observada para a estação teste Guampará através do Modelo Chuva-Vazão I e do agrupamento B.	49
5.10	Vazão específica média modelada e observada ao longo do tempo para a estação teste Guampará através do Modelo Chuva-Vazão I e do agrupamento B.	49
5.11	Relação Vazão específica média modelada x Vazão específica observada para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão II e do agrupamento A.	51

5.12	Vazão específica média modelada e observada ao longo do tempo para a estação teste Madeira Gavazzoni através do Modelo Chuva-Vazão II e do agrupamento A.	51
5.13	Relação Vazão específica média modelada x Vazão específica observada para a estação teste Madeira Gavazzoni através do Modelo Chuva-Vazão II e do agrupamento B.	52
5.14	Vazão específica média modelada e observada ao longo do tempo para a estação teste Madeira Gavazzoni através do Modelo Chuva-Vazão II e do agrupamento B.	52
5.15	Relação Vazão específica média modelada x Vazão específica observada para a estação teste Guampará através do Modelo Chuva-Vazão II e do agrupamento A.	53
5.16	Vazão específica média modelada e observada ao longo do tempo para a estação teste Guampará através do Modelo Chuva-Vazão II e do agrupamento A.	54
5.17	Relação Vazão específica média modelada x Vazão específica observada para a estação teste Guampará através do Modelo Chuva-Vazão II e do agrupamento B.	55
5.18	Vazão específica média modelada e observada ao longo do tempo para a estação teste Guampará através do Modelo Chuva-Vazão II e do agrupamento B.	55

Resumo

Com a ausência de monitoramento fluviométrico em parte dos rios de países em desenvolvimento, assim como a falta de operação e manutenção adequada das estações existentes, grande importância tem sido dada a procedimentos de transferência de informações de bacias hidrográficas com monitoramento fluviométrico para bacias carentes de registros. Tais procedimentos têm sido descritos como regionalização de vazões. O objetivo desta pesquisa é avaliar um método de regionalização de vazões que será aplicado em estações fluviométricas no interior do Estado do Paraná já estudadas anteriormente, com a proposição de diferentes técnicas para tal. O método consiste em regionalizar parâmetros de um modelo chuva-vazão por meio do uso de Redes Neurais Artificiais (RNAs) para bacias hidrográficas sem monitoramento fluviométrico, mas que apresentam medições históricas de temperatura e pluviometria. As técnicas propostas dentro desta abordagem foram o uso dos Mapas Auto-Organizados de Kohonen e RNAs do tipo Função de Base Radial. Na pesquisa foram estudadas 15 estações fluviométricas, das quais duas foram utilizadas como estações teste para validação do método. Em parte das simulações realizadas, foram verificados resultados satisfatórios para regionalização, com coeficientes de determinação e Nash-Sutcliffe elevados, quando confrontadas com as vazões observadas. Os resultados indicam que, assim como o método de regionalização proposto inicialmente, o uso das diferentes técnicas propostas neste trabalho também é viável para a regionalização de vazões, com o diferencial de que apresenta condições de aplicação para um maior número de bacias hidrográficas.

Palavras-chave: modelagem chuva-vazão, rede neural artificial, regionalização.

1. Introdução

Dadas as condições morfológicas favoráveis das bacias hidrográficas do estado do Paraná e a disponibilidade hídrica farta ao longo do ano, a região é considerada uma das mais propícias à instalação de aproveitamentos energéticos no Brasil (Minucci, 2014). Segundo SEMA (2010), a participação da hidroeletricidade deve continuar ainda muito significativa nos planos de expansão do setor. O Plano Decenal de Expansão de Energia 2023 prevê a inclusão de 14.679 MW de potência instalada de fonte hidrelétrica até o ano de 2023, sendo 525 MW (total de cinco aproveitamentos) no estado do Paraná. Deste total, compreende-se uma parcela de 2.000 MW relativa ao aumento da capacidade instalada de pequenas centrais hidrelétricas (PCHs) (MME, 2014).

Apesar do cenário de investimentos em aproveitamentos hidrelétricos, a distribuição de redes pluviométricas e fluviométricas em grande parte do território nacional tem se mostrado inadequada e, em alguns casos, insuficientes para atender demandas de projetos para este setor. Cita-se, por exemplo, a Portaria nº 463/2009 do Ministério de Minas e Energia (MME), que dispõe em seu Artigo 3º da obrigatoriedade de histórico de vazões médias mensais não inferiores a 30 anos, e gerado de maneira que seja o mais extenso e atualizado possível (MME, 2009).

Dessa forma, cresce no Brasil a necessidade de se desenvolver métodos eficientes de regionalização hidrológica como consequência da escassez de dados obtidos (Diniz, 2008). Segundo Saraiva (2010), há atualmente uma considerável demanda para aplicação desses métodos para bacias hidrográficas para as quais não existe

uma rede de monitoramento adequada.

Tucci (2007) define a regionalização como um conjunto de ferramentas que exploram ao máximo as informações existentes, visando a estimativa das variáveis hidrológicas em locais sem dados ou insuficientes. Por serem mais distribuídas e com séries históricas com duração normalmente mais extensa, os dados das redes de monitoramento climatológica e pluviométrica são muitas vezes utilizadas como entrada para a previsão de vazão em bacias carentes de postos fluviométricos, através de modelos matemáticos denominados chuva-vazão (Tucci, 1998).

Dentre as técnicas de modelagem chuva-vazão, o uso de Redes Neurais Artificiais — RNA tem sido empregado de maneira crescente para previsões de variáveis em recursos hídricos e aplicados para superar dificuldades encontradas em abordagens estatísticas tradicionais (Maier e Dandy, 1999). Seus principais benefícios, de acordo com Haykin (1999), são: não-linearidade (importante para representar fenômenos físicos intrinsecamente não-lineares); adaptatividade (capacidade de se adaptar devido à respostas do ambiente); mapeamento entrada-saída (A rede aprende com os exemplos que lhe são apresentados por construir um mapeamento entrada-saída do problema em questão).

Além disso, por conta das diferentes características físicas e climáticas das bacias hidrográficas, bem como a ausência destes parâmetros em certos locais, hidrologos têm regionalizado as características das bacias de interesse com base na experiência obtida em outras bacias. Tal medida, quando aplicada em conjunto com técnicas de modelagem chuva-vazão, constitui uma abordagem importante para solucionar o problema da falta de dados fluviométricos. O objetivo do trabalho consiste em investigar a aplicação da regionalização de parâmetros de um modelo chuva-vazão, calibrado através de RNAs tipo Função de Base Radial, para bacias que não contam com registros de fluviometria, apenas pluviometria e temperatura.

Esta pesquisa está dividida em sete capítulos, sendo essa introdução o Capítulo 1. No Capítulo 2 apresenta-se os objetivos geral e específicos. No Capítulo

3 apresenta-se a Revisão Bibliográfica, a qual expõe as principais características a cerca das técnicas de regionalização hidrológica, modelagem chuva-vazão e redes neurais artificiais. No Capítulo 4 apresenta-se a Metodologia, com o detalhamento dos dados utilizados no estudo e descrição dos procedimentos adotados para a seleção de agrupamentos, modelagem chuva-vazão e regionalização. No Capítulo 5 são apresentados os resultados da pesquisa, e no Capítulo 6 a discussão desses. Por fim, no Capítulo 7 são descritas as conclusões do estudo e recomendações para futuros trabalhos.

2. Objetivos

2.1 Objetivo Geral

O objetivo deste trabalho consiste em investigar a aplicabilidade da regionalização de vazões para bacias hidrográficas as quais não contam com dados de fluviometria disponíveis, apenas de pluviometria e climatologia, com o emprego de Redes Neurais Artificiais do tipo Função de Base Radial.

2.2 Objetivos Específicos

Os objetivos específicos foram:

- Investigar a aplicabilidade do uso de algoritmos de mapas auto-organizáveis para a definição das características semelhantes das bacias hidrográficas;
- Calibrar e validar modelo chuva-vazão através de RNAs do tipo Função de Base Radial a partir da variação de parâmetros de entrada de RNAs propostas para cada bacia;
- Explorar as possíveis semelhanças e diferenças dos resultados deste projeto com os resultados obtidos por trabalhos semelhantes, em especial o estudo desenvolvido por Minucci (2014).

3. Revisão Bibliográfica

3.1 Regionalização Hidrológica

As técnicas de regionalização podem ser empregadas para diversos fins. Segundo Tucci (2007), os objetivos principais são: melhor explorar as amostras pontuais, e, em consequência, melhorar as estimativas das variáveis hidrológicas; verificar a consistência das séries hidrológicas; e identificar a falta de postos de observação.

A regionalização pode ser elaborada de acordo com três grandes abordagens:

1. Funções estatísticas de variáveis hidrológicas

A determinação, por exemplo, de curvas de probabilidade de vazões máximas, médias ou mínimas ou ainda curva de probabilidade de precipitações máximas.

2. Funções específicas que relacionam variáveis

Destaca-se, entre as outras, as curvas de regularização ¹, infiltração e permanência².

Tucci (2007) cita dois procedimentos normalmente utilizados: a) ajuste de uma função matemática aos dados de cada posto de observação e regionalização dos

¹Curva resultante da combinação de vários volumes e suas respectivas vazões, para o caso de reservatórios.

²Relaciona a frequência acumulada da vazão ou nível de um rio e a probabilidade de ocorrerem vazões maiores ou iguais ao valor da ordenada

parâmetros da referida função; e b) adimensionalização da função, obtenção de uma curva média com base nas curvas adimensionais dos diferentes postos e a regressão entre a variável de adimensionalização e as características físicas e climáticas.

3. Parâmetros de modelos hidrológicos

Os critérios normalmente adotados para a regionalização dos parâmetros para outra bacia são: a) Determinação de equações de regressão entre o parâmetro a ser regionalizado e características físicas e climáticas das bacias; ou b) Através da definição do intervalo de variação possível dos parâmetros com base em informações características das bacias.

Contudo, esses procedimentos apresentam algumas limitações devido à interdependência dos parâmetros, o que pode criar tendenciosidade (Tucci, 2007).

Com relação aos dados básicos utilizados para a regionalização, os parâmetros hidrológicos, cartográficos e descritivos da região em estudo devem ser reunidos. Tucci (2007) ressalta que a regionalização deve ser orientada para que a escolha e seleção das informações sejam examinadas de acordo com a variável regionalizada.

Contudo, quanto à representatividade dos dados de vazão, DNAEE (1983) define como critérios mínimos a seleção de postos fluviométricos com uma série de pelo menos cinco anos de dados — requisito no qual permite que outras técnicas como a modelagem chuva-vazão para fins de regionalização seja explorada.

Nesse sentido, Minucci (2014) propôs uma nova abordagem para a regionalização de parâmetros de um modelo chuva-vazão, de uma maneira diferente dos métodos clássicos de regionalização. Visando tornar o método mais objetivo, as etapas adotadas foram as seguintes:

1. A partir de 15 estações fluviométricas do Estado do Paraná e as características fisiográficas e climatológicas de suas respectivas áreas de drenagem, definiram-

se, através de critérios estatísticos ³, quais eram as bacias com características semelhantes entre si;

2. Definiram-se duas estações fluviométricas “teste”, a partir das 15 totais, visando confrontar os dados observados com os calculados;
3. Foi desenvolvido um modelo chuva-vazão e selecionada a melhor RNA - do tipo Perceptron Multicamadas - para cada bacia;
4. Quanto às características fisiográficas e climatológicas das estações, foi realizada uma análise de componentes principais da amostra a fim de eliminar dados que são simplesmente combinações lineares de outros dados, e que, estatisticamente, não agregavam valor à amostra;
5. Foram definidos os grupos semelhantes (*clusters*) para as 15 estações fluviométricas a partir de técnica de análise de agrupamento (ou clusterização), seguida de análise discriminante;
6. A partir das métricas resultantes da análise discriminante, obteve-se a série regionalizada para as estações “teste”.

Tanto na aplicação do modelo chuva-vazão, quanto no método adotado para regionalização, os coeficientes de Nash-Sutcliffe e de determinação foram elevados (superiores a 0,80) e o erro médio quadrático foi baixo (inferiores a 10 m³/s), quando comparadas as vazões observadas e calculadas (Minucci, 2014).

Para maiores detalhes sobre as etapas de consistência de registros pluviométricos, fluviométricos e climatológicos, assim como o emprego da RNA tipo MLP, análise discriminante, análise dos componentes principais e o método da silhueta, consultar a pesquisa desenvolvida por Minucci (2014).

³Método das silhuetas - indica, a partir da definição inicial do número de grupos, o valor da silhueta S_i para as n estações. Quanto maior S_i , mais próximo as estações estão de seus próprios grupos e mais distantes estão dos outros grupos (Minucci, 2014).

3.2 Modelagem chuva-vazão

No contexto do ciclo hidrológico, os modelos “chuva-vazão” representam a relação entre a precipitação e a vazão, procurando contemplar os fenômenos da precipitação e sua distribuição espacial, evaporação, infiltração, percolação, escoamento superficial, sub-superficial e nos corpos d’água (Tucci, 1998).

Uma vez que os dados de precipitação geralmente se apresentam com séries históricas mais longas que as de vazão, é possível completar vazões desconhecidas para as bacias a partir dos dados de precipitação com o emprego desses modelos.

A estrutura dos modelos chuva-vazão é baseada nos seguintes elementos:

- **Discretização da bacia hidrográfica:** são utilizados critérios de subdivisão espacial da bacia;
- **Variáveis temporais de entrada:** as variáveis normalmente utilizadas são precipitação, vazão, e evapotranspiração potencial;
- **Estrutura básica de integração dos processos:** estrutura que se divide em bacia e canal, os quais podem ser expressos por equações matemáticas;
- **Aquisição de dados físicos da bacia:** Informações geralmente obtidas através de técnicas de Sensoriamento Remoto e Sistema de Informação Geográfica (SIG), como comprimento do rio, declividade, altitude, área de drenagem, cobertura do solo, entre outros.

Além disso, sua classificação é definida pela literatura hidrológica da seguinte maneira:

- **Modelos empíricos** - processo cujo a expressão matemática não possui relação com o comportamento físico dos processos;
- **Modelos conceituais** - utilização das equações de conservação, associadas a

uma ou mais equações empíricas que relacionam variáveis e parâmetros dos processos;

- **Físico-distribuídos** - são os modelos que utilizam as equações de continuidade e dinâmica dos processos envolvidos, a fim de integrar todos os processos físicos que ocorrem na bacia (Tucci, 1998).

Dentre os modelos empíricos, podem ser utilizadas, por exemplo, técnicas de regressão linear e não-linear, séries de Fourier, e o emprego de Redes Neurais Artificiais para relacionar a precipitação com a vazão.

3.3 Redes Neurais Artificiais

Os estudos na área de Redes Neurais Artificiais foram motivados a partir da observação do cérebro humano e sua capacidade de processamento de informações de uma forma completamente diferente do modo convencional do computador digital. Segundo Jain et al. (1999), a rede neural é “um processador distribuído paralelamente criado por simples unidades de processamento que tem propensão natural de armazenar conhecimento experimental e torná-lo disponível para uso”. A rede se assemelha ao cérebro em dois aspectos:

- O conhecimento é adquirido pela rede através de um processo de aprendizado;
- e
- Forças de conexões intersinápticas, denominadas pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

3.3.1 Cérebro Humano e Neurônio Artificial

O cérebro humano pode ser visto como um sistema de três estágios, de acordo com a Figura 3.1. O sistema central é o cérebro, representado pela Rede Neural, na qual continuamente recebe informação, a processa, e a torna apropriada para tomada

de decisão. As setas apontando da esquerda para direita indicam a transmissão progressiva de sinais de informação através do sistema. As setas direcionadas no sentido oposto demonstram a presença de realimentação no sistema. Os Receptores convertem estímulos do corpo humano (ou ambiente externo) em impulsos elétricos que transmitem informação para a Rede Neural. Os Atuadores transmitem impulsos elétricos gerados pela Rede Neural como saída, ou Resposta.

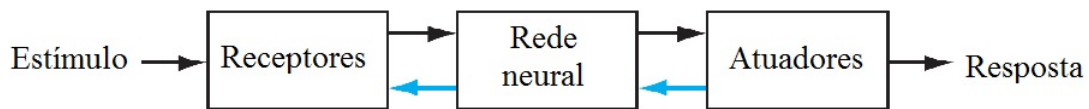


Figura 3.1: Representação do diagrama de blocos do sistema nervoso.

Fonte: Adaptado de Haykin (1999).

O neurônio é uma unidade de processamento de informação fundamental para a operação de uma rede neural. A Figura 3.2 mostra o modelo de um neurônio, que representa a base para o projeto de outros modelos de redes neurais mostrados neste trabalho.

Cada elemento do conjunto das sinapses é caracterizado por um peso ou força própria. Uma função de combinação é empregada para somar os sinais de entrada, os quais são ponderados pelos respectivos pesos sinápticos do neurônio. A função de ativação é o elemento das RNAs responsável por limitar a amplitude de saída de um neurônio a um valor finito (Haykin, 1999).

Na primeira etapa, são apresentadas à rede os dados de entrada, os quais se ligam ao neurônio através de sinapses. É realizado, então, um produto entre os dados de entrada $[x_1, x_2, \dots, x_m]$ de dimensão m e os respectivos pesos sinápticos $[w_{k1}, w_{k2}, \dots, w_{km}]$, seguido de um somatório de todas os produtos (representado na Figura 3.2 como Função de Combinação) e a adição de um viés b_k . A partir daí, aplicando função de ativação $\varphi(v_k)$ obtém-se a saída y_k .

A equação governante, portanto, é dada por:

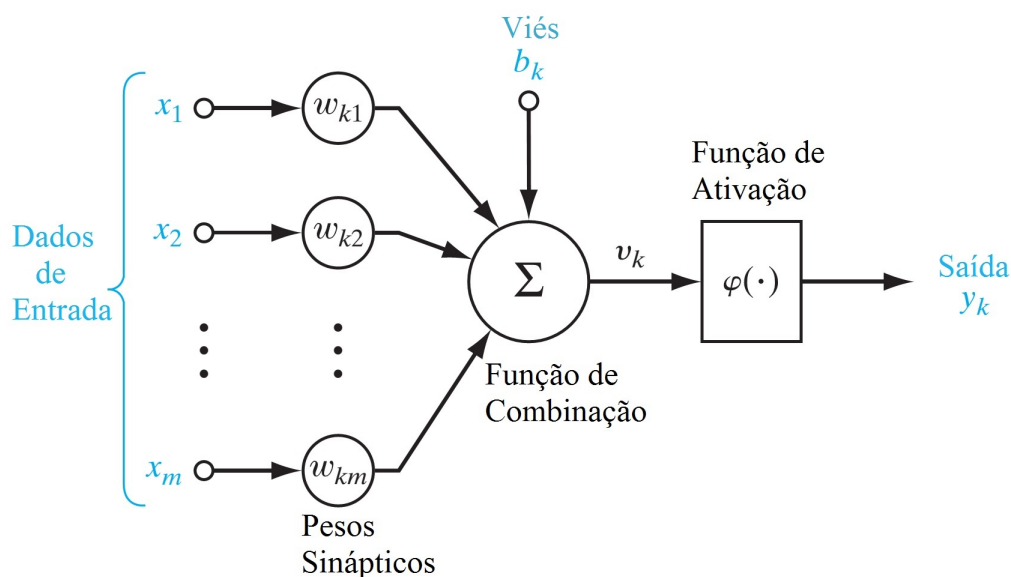


Figura 3.2: Neurônio Artificial.
 Fonte: Adaptado de Haykin (1999).

$$y_k = \varphi(v_k), \quad (3.1)$$

sendo

$$v_k = \sum_{j=1}^m w_{kj}x_j + b_k, \quad (3.2)$$

onde j é o contador m do número de dados de entrada.

O modo no qual os neurônios de uma rede neural são estruturados é intimamente relacionado com seu algoritmo de aprendizado para treinamento da rede.

3.3.2 Redes Neurais tipo *Perceptron* Multicamadas

Perceptron Multicamadas (MLP) é a forma mais difundida de aplicação do *perceptron* de Rosenblatt's (definido originalmente como uma única camada de rede neural) (Haykin, 1999). Suas principais características são:

- O modelo de cada neurônio da rede conta com uma função de ativação não-linear que é diferenciável;

- A rede contém essencialmente três camadas, sendo uma de entrada, uma de saída e uma (ou mais) intermediárias ou ocultas;
- A rede exibe um alto padrão de conectividade - o qual é determinado por pesos sinápticos da rede (Haykin, 1999).

A Figura 3.3 ilustra a estrutura de uma RNA do tipo MLP formada por duas camadas escondidas.

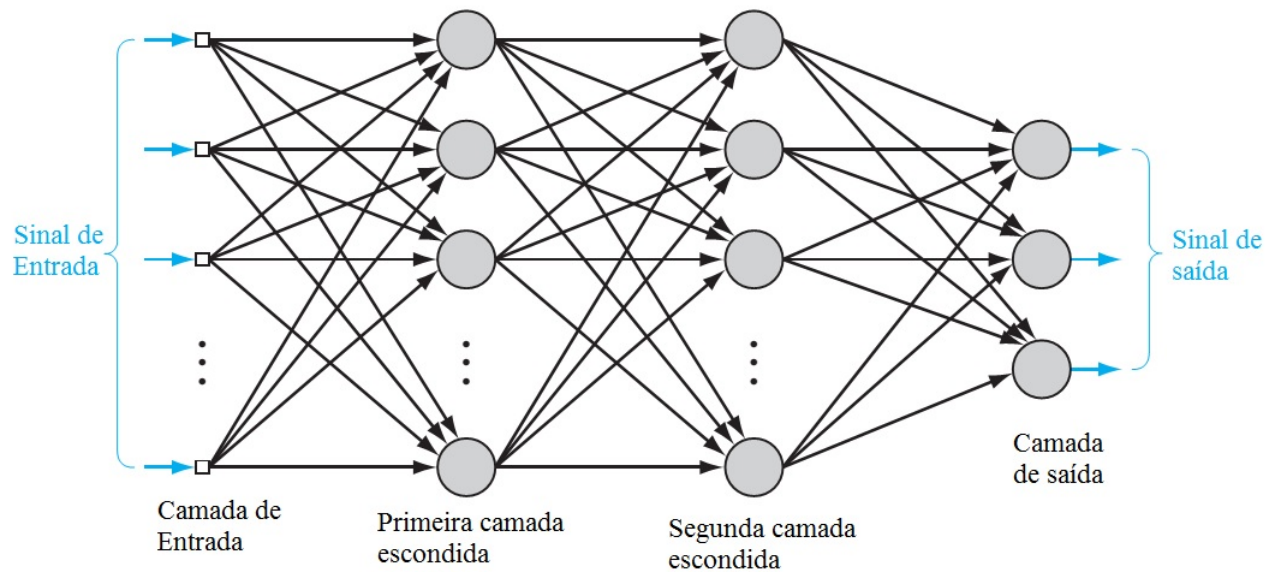


Figura 3.3: Arquitetura de uma RNA do tipo *multilayer perceptron*.
Fonte: Adaptado de Haykin (1999).

A propagação da informação nas redes MLPs se dá sempre no sentido progressivo (camada a camada), razão pela qual é classificada como *feedforward*. Os neurônios das camadas intermediárias recebem uma combinação linear dos dados de entrada, que gera um estímulo na função de transferência, a qual emite uma resposta. Nesse caso, a resposta é tomada com entrada na camada seguinte, que pode tanto ser uma camada intermediária quanto uma camada de saída.

3.3.3 Redes Neurais tipo Função de Base Radial

As Redes Neurais de Função de Base Radial — do inglês *Radial Basis Function* —, comumente denominadas Redes Neurais RBF, são descritas como Redes Neurais que utilizam funções de base radial como funções de ativação (Haykin, 1999).

De acordo com Castro e Castro (2000b), as RNAs do tipo RBF são redes supervisionadas, consideradas aproximadoras universais. Elas apresentam arquitetura semelhante às RNAs tipo MLP treinadas pelo algoritmo *Backpropagation*. As principais diferenças entre as Redes Neurais tipo RBF e MLP são apresentadas na Tabela 3.1.

Tabela 3.1: Diferenças entre as RNAs tipo *Radial Basis Function* e RNAs *Multilayer Perceptron*.

	RNA <i>Radial Basis Function</i>	RNA <i>Multilayer Perceptron</i>
Função de Ativação	Funções monotônicas	Função não-linear da distância entre o vetor de entrada e um vetor de referência
Determinação do valor de saída	Aproximações Locais — dado um vetor de entrada, tipicamente apenas algumas unidades escondidas apresentarão ativações significativas	Aproximações Globais — dado um vetor de entrada, muitas unidades escondidas contribuirão para a determinação do valor de saída
Pesos e Conectividade	Arquitetura simples, consistindo de duas camadas de pesos (1 ^a contém os parâmetros das funções de base radial, e 2 ^a forma combinações lineares das ativações das funções de base radial para gerar a saída)	Grande número de camadas pesos e um alto padrão de conectividade
Parâmetros	Treinadas em dois estágios (1 ^o - funções de base radial sendo determinadas primeiramente por meio de técnicas não supervisionadas, usando para tal os dados de entrada e a 2 ^o camada (de pesos) sendo por métodos lineares supervisionados, de rápida convergência)	Todos os parâmetros são usualmente determinados aos mesmos tempo, que incorre em um alto custo computacional
Comparações para mesmas aplicações		Erro Final (RBF) < Erro Final (MLP) Velocidade Convergência (RBF) > Velocidade Convergência (MLP) Capacidade Generalização (RBF) < Capacidade Generalização (MLP)

Fonte: Adaptado de Castro e Castro (2000b)

As redes RBF possuem em sua arquitetura uma camada escondida definida

por um conjunto de funções de base radial, na qual é centrada em uma particular coordenada do espaço multi-dimensional dos pontos que compõem o espaço de dados de entrada. Tais coordenadas caracterizam-se por definir um centro de uma região de aglomeração de pontos (*cluster*).

Na Figura 3.4, é apresentado um esquema genérico de uma RNA de função de base radial. A camada \mathbf{u} é a camada de entrada, seguida da camada intermediária de neurônios não-lineares φ_k e camada de saída y . Considera-se que a função de ativação da camada intermediária uma função Gaussiana, com o centro \mathbf{t}_k e variância σ_k^2 de um *cluster*. A camada y é a soma das saídas de cada Gaussiana, ponderadas pelo respectivo peso sináptico w_k , de modo que a combinação linear é dada por

$$y = \sum_{i=1}^{K-1} w_k \varphi_K(\mathbf{u}, \mathbf{t}_k, \sigma_k^2).$$

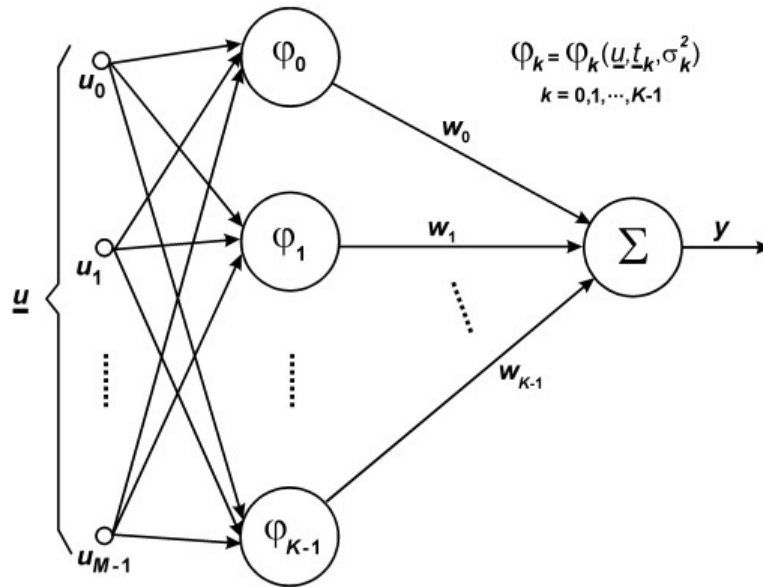


Figura 3.4: Rede Neural do tipo *Radial Basis Function*.
Fonte: Castro e Castro (2000a).

O raio de influência de uma coordenada em torno de seu centro é determinado pelo parâmetro σ através da função Gaussiana $\phi(\zeta)$ (mais comumente utilizada), de acordo com a Equação 3.3.

$$\phi(\zeta) = \exp\left(\frac{-\zeta^2}{2\sigma^2}\right) \quad (3.3)$$

O mapeamento não linear é expresso por funções de ativação gaussianas, da forma,

$$\varphi_k(n) = \varphi_k(\mathbf{u}(n), \mathbf{t}_k(n), \sigma_k^2(n)) = \exp\left(\frac{-1}{\sigma_k^2(n)} \|\mathbf{u}(n) - \mathbf{t}_k(n)\|^2\right), \quad (3.4)$$

onde $\mathbf{u}(n)$ representa o vetor de entrada \mathbf{u} no instante n , $\mathbf{t}_k(n)$ representa o vetor centro da k -ésima função de base radial, k é o número de funções de base radial e $\sigma_k^2(n)$ é a variância associada a cada uma das funções no instante n .

O processo de aprendizado é equivalente a ajustar uma superfície não-linear ao conjunto de dados, em um espaço multi-dimensional, de acordo com algum critério estatístico, e a generalização equivale a usar esta superfície multi-dimensional para interpolar outros pontos que não pertençam ao conjunto de treinamento, mas estejam em sua vizinhança.

Os dados representados através de redes neurais RBF são, portanto, expandidos com referência a um conjunto finito de funções de ativação neurais, chamadas funções de base radial. Assim, cada uma dessas funções define o centro de uma — entre várias possíveis — região de aglomeração de pontos do espaço de dados de entrada.

3.4 Técnicas de Aprendizado

Segundo Diniz (2008), o processo de aprendizado pode ser identificado de acordo com as diferentes complexidades de inferência: a) hábito; b) dedução; c) analogia; d) instrução, e e) indução. O maior interesse para pesquisas ligadas ao aprendizado de máquinas é a indução, cujo conceito se caracteriza pelo raciocínio que parte do específico para o geral, sendo dividido em supervisionado e não supervisionado.

O aprendizado supervisionado, como o nome sugere, ocorre a partir da apre-

sentação de pares de entrada e saída, de forma a “ensinar” a função que mapeia a entrada para a saída. O aprendizado não supervisionado, por sua vez, reconhece padrões nos dados de entrada, mesmo sem nenhum *feedback* de saída. Neste caso, o trabalho está relacionado com o agrupamento dos exemplos (*clustering*), que se dá através da organização do conjunto de padrões, e resultam em grupos baseados na similaridade (Jain et al., 1999).

Uma das técnicas de aprendizado não supervisionado amplamente empregada se dá através do uso de Redes Neurais Artificiais para fins de agrupamento. De acordo com Jain et al. (1999), algumas das importantes características de RNA para padrões de clusterização são:

- Processam vetores numéricos e, por isso, requerem que os padrões a serem representados contenham apenas características quantitativas;
- São intrinsicamente paralelas e com arquiteturas de processamento distribuídas;
- Podem ajustar seus “pesos” de interconexão adaptativamente.

Um dos exemplos mais conhecidos de aplicação de RNAs para clusterização é o Mapa Auto-Organizado (da literatura inglesa *Self-Organizing Maps* - SOM), desenvolvido por Kohonen (Castro e Castro, 2000a).

3.4.1 Mapas Auto-Organizados de Kohonen

Os mapas auto-organizados são redes competitivas que possuem a habilidade de formar mapeamentos que preservam a topologia entre o espaços de entrada e saída. Os neurônios de entrada da RNA competem entre si para serem ativados, com o resultado de que apenas um neurônio de saída (ou um neurônio por grupo) está “ativo”. O neurônio de saída, que “vence” a competição, é normalmente descrito como “neurônio vencedor” (Fausett, 1994). Os neurônios em uma rede auto-organizada se tornam seletivamente “sintonizados” a partir de estímulos (padrões de entrada)

ao longo do processo competitivo. A Figura 3.5 ilustra o esquema de um mapa auto-organizado.

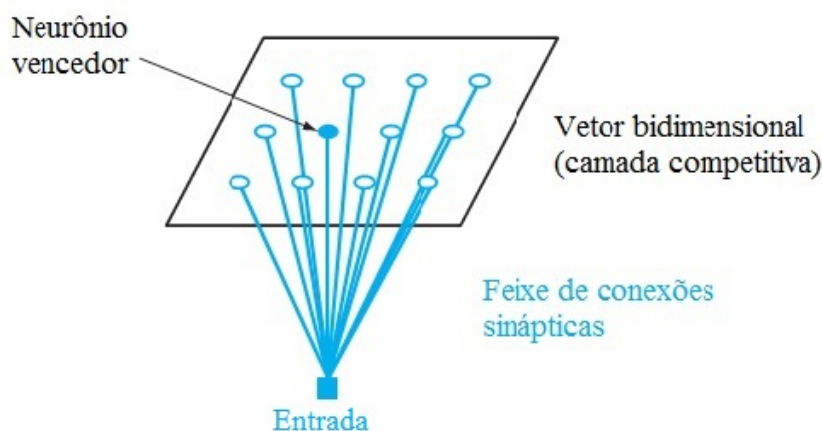


Figura 3.5: Topologia de um mapa de Kohonen. Fonte: Adaptado de Haykin (1999) e Fausett (1994).

Os algoritmos mais conhecidos que implementam as redes de Kohonen são o SOTA (Herrero et al., 2001) e o SOM (Kohonen et al., 1996). Segundo Haykin (1999), o algoritmo inicia selecionando valores aleatórios para os dados iniciais dos pesos sinápticos, e, à medida que o treinamento inicia, os pesos serão ajustados ao valor adequado em conformidade com as características da rede. Após esta inicialização, são realizados três grandes processos para a definição do mapa auto-organizado, cujas principais características estão resumidas a seguir:

1. **Competição:** Para cada vetor de entrada, os neurônios da rede irão computar seus respectivos valores através de uma função discriminante. O neurônio particular com o maior valor da função é declarado vencedor da competição;
2. **Cooperação:** O neurônio vencedor determina a localização espacial da topografia vizinha dos neurônios excitados, fornecendo, assim, uma base para a cooperação entre os neurônios vizinhos;
3. **Adaptação sináptica:** Este mecanismo permite aos neurônios excitados aumentar em seus valores individuais da função discriminante com relação ao seu

padrão de entrada através de ajustes aplicados a seus pesos sinápticos. A resposta de cada neurônio “vencedor” à sua subsequente aplicação de padrão de entrada é aumentada.

3.5 Aplicações de RNAs em Recursos Hídricos

As aplicações de RNAs para diferentes áreas em recursos hídricos tem aumentado de maneira expressiva (Maier e Dandy, 1999). Machado (2005) modelou a relação chuva-vazão para a bacia do Rio Jangada, localizada entre o estado do Paraná e Santa Catarina, através de RNAs tipo MLP e confrontou os resultados com o modelo conceitual IPHMEN⁴. Os resultados obtidos foram satisfatórios, com coeficientes de Nash-Sutcliffe e de determinação superiores aos obtidos pelo modelo IPHMEN, e valores inferiores de Erro Médio Quadrático quando comparado ao modelo conceitual.

Riad et al. (2003) desenvolveram um modelo chuva-vazão através de RNAs tipo MLP para o exutório do Rio Ourika (Marrocos), e obtiveram coeficientes de determinação superiores a 0,90 e erros médios quadráticos inferiores a 1,6%. Vos (2003) também obteve resultados satisfatórios, empregando RNAs tipo MLPs para a predição de vazão no exutório da bacia *Alzette-Pfaffenthal*, em Luxemburgo.

Dogan et al. (2014) verificaram o desempenho do método autoregressivo e de três métodos de RNAs para prever a vazão diária na estação *Çifteler* no rio *Sakarya* (Turquia), utilizando dois anos de dados para treinamentos e aproximadamente um ano e seis meses para teste. Os coeficientes de determinação do método autoregressivo e das RNAs tipo *back-propagation*, RBF e Redes Neurais Recorrentes foram satisfatórios, sendo 0,7547, 0,9495, 0,9479, e 0,9991, respectivamente.

Kagoda et al. (2010) utilizaram RNAs RBF para predição da vazão diária de 1-dia para o rio *Luvuvhu* (África do Sul) e verificaram valores do coeficiente

⁴Modelo de público domínio elaborado pelo Instituto de Pesquisas Hidráulicas da Universidade Federal do Rio Grande do Sul

de Nash-Sutcliffe de 0,988 e 0,98 durante a calibração e operação, com desvios de vazão (*percent of bias*) de cerca de -20% para ambos os casos. Kiliç et al. (2000) compararam a qualidade das predições de vazão do rio *Cankdere* (Istambul) obtidas com o emprego de RNAs tipo RBF e MLP, verificando resultados similares, porém com baixos coeficientes de determinação para ambos os métodos.

Quanto às técnicas de clusterização, Lin e Chen (2006) utilizaram 154 estações pluviométricas em Taiwan para análise regional de frequências. Após a seleção das variáveis de entrada (elevação, longitude, latitude e precipitação média mensal para cada mês), foram definidos oito *clusters*, empregando o algoritmo SOM, e duas outras técnicas tradicionais de clusterização (*k-means* e método de *Ward*). Verificou-se que o SOM pode identificar com mais precisão os *clusters* que os outros dois métodos. Contudo, quando se refere à regionalização de vazões através de RNAs tipo RBF, a literatura disponível para esta abordagem é escassa.

4. Metodologia

4.1 Metodologia Proposta

Este trabalho baseia-se na regionalização de parâmetros de um modelo chuva-vazão baseado em RNAs para bacias que não contam com dados de pluviometria, mas que apresentam séries históricas de precipitação e temperatura.

As etapas do método foram baseadas no estudo realizado por Minucci (2014), com algumas diferenças metodológicas, a serem apresentadas neste capítulo.

As principais etapas são descritas a seguir, e ilustradas pela Figura 4.1.

1. A partir de 15 estações fluviométricas do Estado do Paraná e das características fisiográficas e climatológicas de suas respectivas áreas de drenagem, foi definido, através do emprego de mapas auto-organizados de Kohonen, quais serão as bacias com características semelhantes entre si — definição dos *clusters*;
2. Foram adotadas as duas estações fluviométricas “teste” definidas por Minucci (2014), a partir das 15 totais, visando confrontar os dados observados com os calculados;
3. Foi desenvolvido um modelo chuva-vazão e selecionado a melhor RNA — do tipo Função de Base Radial — para cada bacia;
4. Em seguida, foram obtidos modelos chuva-vazão “médios” dos grupos semelhantes;

5. Foram obtidas as séries de vazões regionalizada para as estações “teste”;
6. As vazões modeladas foram confrontadas com os registros observados, para a validação do método.

As séries regionalizadas são geradas a partir da operação da RNA calibrada da estação auxiliar que utilizou como entrada os registros de precipitação e temperatura da Estação Teste (a ser regionalizada).

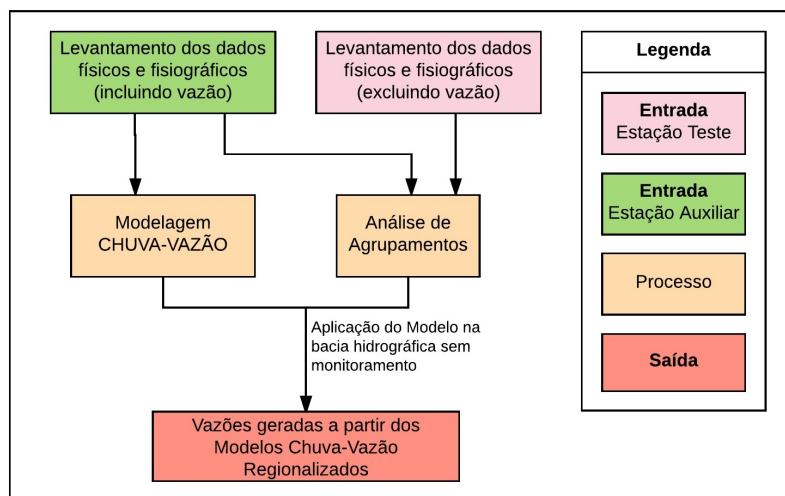


Figura 4.1: Abordagem adotada no estudo para regionalizar registros de vazões.

A regionalização abordada por Minucci (2014): Utilizou o método *k-means* para a etapa de clusterização, enquanto nesse trabalho foi utilizado os mapas auto-organizados de Kohonen; empregou RNAs tipo *perceptron* multicamadas para o modelo chuva-vazão, sendo RNAs tipo RBF as redes neurais utilizadas neste projeto. Tanto as variáveis físicas/fisiográficas das bacias hidrográficas, como seus registros de pluviometria, fluviometria e climatologia foram obtidos a partir de Minucci (2014), sendo as principais etapas para regionalização descritas nesse capítulo.

Nas Seções 4.8, 4.7 e 4.9, são detalhadas as etapas de formulação do Modelo Chuva-Vazão, Análise de Agrupamentos e Regionalização, respectivamente.

4.2 Área de Estudo

As bacias hidrográficas objeto de estudo são as seguintes: Balsa do Cantú, ETA Guarapuava, Fazenda Maracanã, Guampará, Madeireira Gavazzoni, Ponte do Goio Bang, Ponto Leôncio Primo, Porto Carriel, Porto Guarani, Porto Santo Antônio, Rio dos Patos, Salto Sapucaí, São Sebastião, Tereza Cristina e Tormenta Fazenda Trento. As bacias são ilustradas na Figura 4.2, e os métodos utilizados para a consistência e tratamento dos dados estão descritos na Seção 4.3.

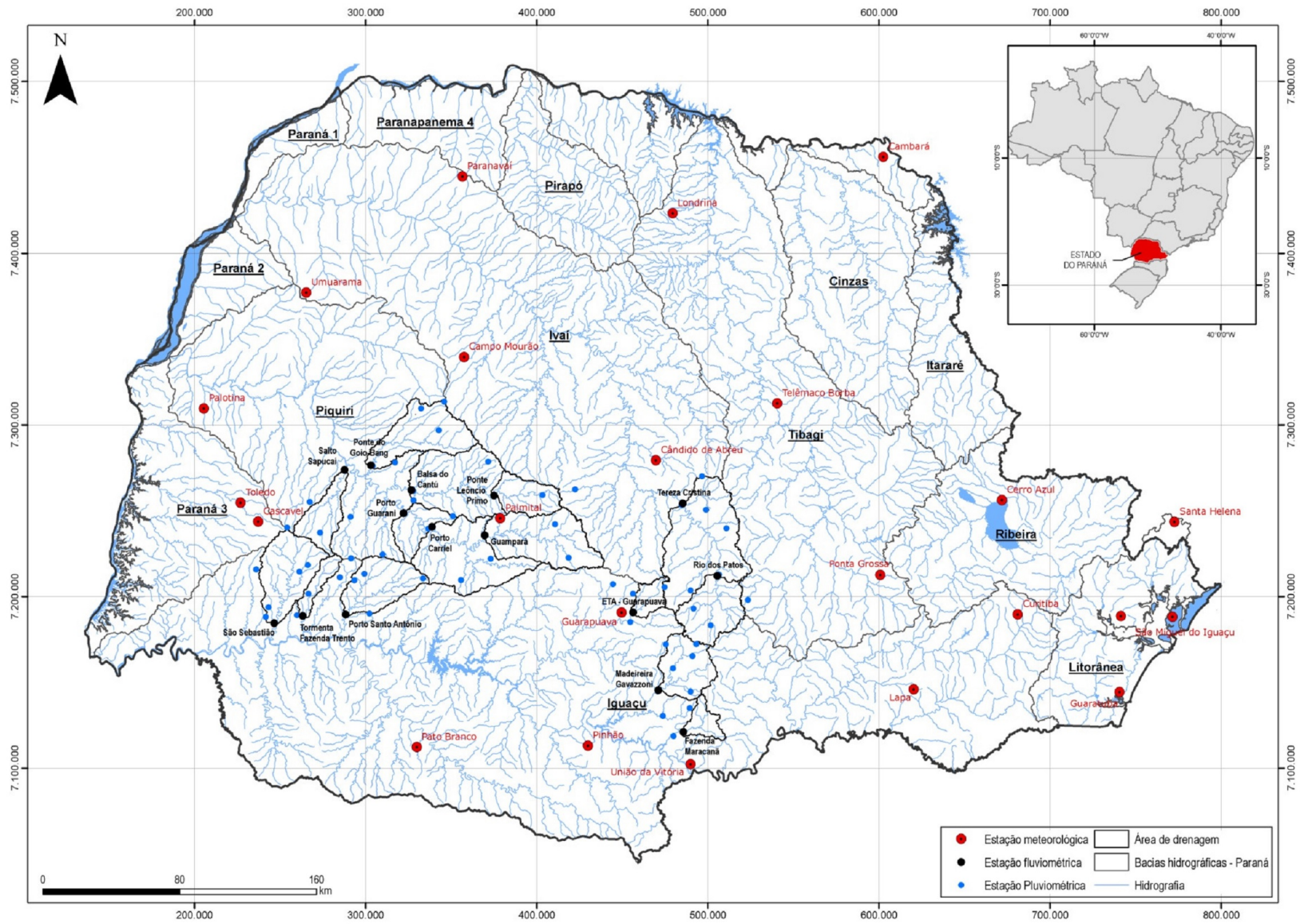


Figura 4.2: Estoes hidrometeorolgicas utilizadas no estudo. Fonte: Minucci (2014).

Os critérios utilizados para a definição da área estão listados a seguir:

1. Seleção de 15 estações fluviométricas com áreas de drenagem inferiores a 5.000 km² e com dados confiáveis e representativos, de vazão e precipitação, entre janeiro de 1988 a dezembro de 2011 (totalizando 16 anos), apresentando período de falhas não superiores a 5 meses sequenciais;
2. Bacias hidrográficas com baixa expansão urbana entre as datas compreendidas das medições;
3. Corpo hídrico não afetado por desvios ou regularização de vazão à montante da régua linimétrica.

4.3 Dados Utilizados no Estudo

As séries históricas consistidas foram obtidas a partir do estudo realizado por Minucci (2014). No referido trabalho, foi realizada a consistência e tratamento dos dados pluviométricos e fluviométricos das estações hidrometeorológicas, bem como o levantamento das características físicas e climatológicas das bacias.

Os dados hidrometeorológicos utilizados são provenientes do portal HIDROWEB da Agência Nacional de Águas — ANA, bem como de dados disponibilizados pelo Instituto Ambiental do Paraná. Cada parâmetro avaliado apresenta uma série histórica mensal de 168 registros. Os dados brutos foram os seguintes:

- Histórico de medições de cotas das estações fluviométricas;
- Resumo de descarga líquida;
- Levantamento dos perfis transversais das seções dos rios;
- Série histórica de precipitação das estações de observação.

A partir de uma série de medições da descarga líquida em uma estação fluviométrica, é possível determinar uma relação entre o nível da água (cota) e a

vazão. Esta relação entre a cota e a descarga líquida é denominada de curva-chave. A determinação da curva-chave possibilita o cálculo da série de descarga líquida (vazão) correspondente à série de medições da cota linimétrica (nível) realizadas na estação fluviométrica.

Os dados de perfil transversal são necessários para aplicação de técnicas de extrapolação da curva-chave, como, por exemplo, o método de Stevens, Manning ou logarítmico (Kennedy, 1984).

As estações fluviométricas e pluviométricas utilizadas no estudo estão relacionadas na Tabela 4.1.

4.4 Pré-tratamento e Consistência dos Dados

Pluviometria

Os dados de precipitação que apresentaram falhas foram estimados através do método da razão das normais (Equação 4.1), que permite estimar os valores faltantes de uma estação pluviométrica baseadas em outras estações auxiliares (Serrano, 1997).

$$\frac{P_1}{\bar{P}_1} = \frac{1}{(N-1)} \left(\frac{P_2}{\bar{P}_2} + \frac{P_3}{\bar{P}_3} + \frac{P_4}{\bar{P}_4} + \dots + \frac{P_N}{\bar{P}_N} \right), \quad (4.1)$$

onde, P_1 é o valor faltante de precipitação na estação 1; P_2, \dots, P_N é a precipitação nas estações auxiliares para o mesmo período; \bar{P}_1 é precipitação média anual da estação 1; $\bar{P}_2, \dots, \bar{P}_N$ é a média anual das estações auxiliares e N é o número de estações auxiliares (pelo menos três).

Fluviometria

Os períodos mensais faltantes da série histórica de vazão foram preenchidos através de equações de regressão. A estação utilizada para preenchimento foi determinada com base em sua semelhança hidrológica de escoamento superficial médio

Tabela 4.1: Estações fluviométricas e pluviométricas utilizadas no estudo. Fonte: Minucci (2014).

Código ANA	Estação Fluviométrica	A.D. (km ²)	Corpo hídrico	Estações pluviométricas associadas
64775000	Balsa do Cantú	2520,1	Rio Cantú	Altamira do Paraná, Balsa do Cantú, Barra Grande, Laranjal, Palmital e Roncador
65809000	ETA Guarapuava	312,8	Rio das Pedras	Guarapuava, Campo de Dentro, Invernadinha e Palmeirinhas
65415000	Fazenda Maracanã	323,6	Rio Palmital	Fazenda Maracanã, Pátio Velho, Salto do Vau e Santana
64764000	Guampará	1006,0	Rio Piquiri	Campina do Simão, Marquinhos, Palmital e Santa Maria do Oeste
65764000	Madeira Gavazzoni	935,7	Rio da Areia	Campina Bonita, Colônio Alegrave, Fazenda Zaniolo, Inácio Martins, Leonópolis e Madeira Gavazzoni
64785000	Ponte do Goio Bang	1341,8	Rio Tricolor	Campina da Lagoa, Guarani, Mamborê, Pensamento e Ponte do Goio-Bang
64773000	Ponte Leôncio Primo	696,6	Rio Cantú	Barra Grande, Roncador, Palmital e Pitanga
64767000	Porto Carriel	3569,3	Rio Piquiri	Campina do Simão, Campo Verde, Laranjal, Marquinhos, Palmital, Porto Carriel e Santa Maria do Oeste
64771500	Porto Guarani	4165,0	Rio Piquiri	Campina do Simão, Campo Verde, Laranjal, Marquinhos, Palmital, Porto Carriel e Santa Maria do Oeste
65970001	Porto Santo Antônio	1084,1	Rio Guarani	Goioxim, Guaraniaçu, Linha Mirim, Porto Santo Antônio, Terra Alta e Vila Guarani
64620000	Rio dos Patos	1097,7	Rio dos Patos	Apiaba, Itapara, Relógio, Rio dos Patos, Tijuco Preto e Inácio Martins
64790000	Salto Sapucaí	693,3	Rio Sapucaí	Corbélia, Ponte Tourinho – Braganey, Salto Sapucaí e São João do Oeste
65979000	São Sebastião	1311,6	Rio Andrada	Boi Preto, Cascavel – Ocepar, Flor da Serra, Linha Santa Catarina, Nelson Taborda, Rio do Salto, Santa Lúcia e São João do Oeste
64625000	Tereza Cristina	3567,9	Rio Ivaí	Apiaba, Bairro da Imbuia, Campo de Dentro, Inácio Martins, Itapara, Laranjeiras, Relógio, Rio dos Patos, Saltinho, Tereza Cristina e Tijuco Preto
65971050	Tormenta Fazenda Trento	319,6	Rio Tormenta	Catanduvas, Flor da Serra, Ibema, Ibira-cema, Nelson Taborda e Rio do Salto

A.D.: Área de Drenagem.

com a estação a ser preenchida. Minucci (2014) considerou como base um coeficiente de correlação (R^2) superior a 0,80.

Climatologia

As séries de dados das estações climatológicas apresentam poucos dias de falhas entre o período de 1998 a 2011. Desse modo, as falhas diárias foram preenchidas por dados brutos da estação mais próxima. O cálculo dos valores mensais foi realizado para as variáveis temperatura (incluindo seus valores mínimos, médios e máximos), umidade relativa e radiação solar.

4.5 Determinação das características das bacias

Após a consistência e tratamento dos dados, foram determinadas as características físicas e climatológicas das bacias, resumidas na Tabela 4.2, visando determinar o grau de similaridade dentre elas.

As características das estações hidrometeorológicas foram levantadas através de técnicas de Sistemas de Informações Geográficas (SIG) e parâmetros estatísticos das séries históricas, resultando nos seguintes parâmetros de interesse: temperatura média, coeficiente de variação da temperatura média, precipitação anual média, além dos parâmetros fisiográficos como altitude da estação pluviométrica, área de drenagem, densidade de drenagem, declividade média, altitude, latitude e longitude do contróide da bacia, totalizando 11 atributos.

Na Tabela 4.2 são apresentadas as características das bacias hidrográficas utilizadas no estudo. Destaca-se que as duas bacias hidrográficas “testes”, ou seja, as bacias que serão objeto da regionalização, estão destacadas em negrito.

Tabela 4.2: Características físicas e fisiográficas das bacias hidrográficas a ser utilizadas no estudo. Fonte: Minucci (2014).

Nome	Altitude (m)	Área de Drenagem (km ²)	Temperatura Média (C)	Densidade de drenagem (km/km ²)	Declividade Média (%)	Coord. UTM_x (m)	Coord. UTM_y (m)	Altitude do Centróide (m)	Precip. anual Thiessen (mm)	C.V. Precip.	C.V Temp.
Balsa do Cantú	363,0	2520,1	20,2	0,331	0,408	365 158	7 264 100	463	1746	51,39	13,51
ETA Guarapuava	963,0	312,8	18,1	0,282	0,635	466 041	7 199 190	1067	1814	58,17	16,41
Fazenda Maracanã	843,3	323,6	18,2	0,253	0,818	497 502	7 126 860	912	1617	60,28	18,88
Guampará	549,4	1006,0	19,9	0,317	0,415	403 426	7 233 170	818	1893	58,73	13,99
Madeira Gavazzoni	808,2	935,7	18,5	0,273	0,540	487 214	7 158 460	1133	1884	56,52	16,93
Ponte do Goio Bang	319,4	1341,8	20,6	0,310	0,396	333 307	7 294 420	568	1818	57,52	14,19
Ponto Leôncio Primo	494,5	696,6	20,1	0,326	0,738	393 992	7 258 480	643	1997	58,79	13,48
Porto Carriel	418,3	3569,3	20,1	0,335	0,303	379 712	7 227 640	721	1946	58,27	13,43
Porto Guarani	370,3	4165,0	20,1	0,334	0,269	372 591	7 229 420	709	1947	58,46	13,38
Porto Santo Antônio	350,4	1084,1	20,2	0,310	0,534	308 372	7 206 870	664	1995	57,30	14,87
Rio dos Patos	733,2	1097,7	18,8	0,295	0,211	502 337	7 191 890	746	1780	58,92	16,13
Salto Sapucaí	334,8	693,3	20,6	0,311	0,523	275 728	7 246 730	646	1906	60,28	14,83
São Sebastião	258,2	1311,6	20,6	0,291	0,436	254 495	7 213 920	539	1963	59,69	14,92
Tereza Cristina	488,1	3567,9	19,1	0,284	0,258	498 207	7 219 170	737	1729	57,46	15,75
Tormenta Fazenda Trento	350,9	319,6	20,5	0,269	0,715	272 350	7 207 560	621	1975	57,81	14,93

C.V.: Coeficiente de variação temporal.

4.6 Seleção das variáveis relevantes para a Análise de Agrupamento

Para que sejam selecionados os atributos que são mais representativos para descrição de uma bacia hidrográfica, é necessário que as características sejam reduzidas a uma dimensão que facilite o trabalho de processamento e que mantenha a amostra com grau de significância razoável. Diniz (2008) argumenta que “os 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)”.

No estudo conduzido por Minucci (2014), foi aplicada uma técnica de Estatística Multivariada — Análise dos Componentes Principais (ACP)— com o objetivo de reduzir a dimensão dos dados descritivos. De acordo com os resultados, resumidos na Tabela 4.3, foi verificado que as características mais relevantes das bacias, ou seja, os atributos que explicam maior porcentagem da variância, foram temperatura média, área de drenagem, precipitação média anual e longitude do centróide da bacia.

Tabela 4.3: Resultado da análise de componentes principais dos parâmetros levantados para as estações fluviométricas estudadas (Minucci, 2014).

Variável	Parâmetro Físico/Fisiográfico	% da variância
1	Temperatura Média	57%
2	Área de drenagem	76%
3	Precipitação média anual	86%
4	Longitude do centróide da área de drenagem	91%
5	Coefficiente de variação da precipitação	95%
6	Declividade média da área de drenagem	97%
7	Altitude do centróide da área de drenagem	99%
8	Latitude do centróide da área de drenagem	99%
9	Altitude da estação fluviométrica	100%
10	Coefficiente de variação da temperatura	100%
11	Densidade de drenagem	100%

No entanto, visando verificar a influência dos parâmetros físicos para a seleção dos agrupamentos de bacias, a análise de agrupamentos será realizada levando-se em

conta as quatro primeiras variáveis listadas na Tabela 4.3, assim como todas as 11 variáveis.

4.7 Seleção dos grupos semelhantes

Uma vez levantadas as variáveis físicas e fisiográficas das bacias hidrográficas, bem como a redução destas variáveis a um conjunto significativo o suficiente, foi realizada a clusterização, através do emprego do algoritmo SOM (*Self Organized Maps*) de Kohonem.

As principais configurações para o aprendizado da rede são:

- Dimensão do vetor de entrada
- Número de linhas e colunas da estrutura topológica (camada de saída)
- Número de épocas de treinamento
- Limite máximo do raio de vizinhança

Três agrupamentos finais foram adotados, visando comparar com os resultados verificados por Minucci (2014) - que utilizou o método *k-means* para a clusterização em três grupos. Forão analisados os resultados da clusterização para o seguinte conjunto de variáveis:

- **Agrupamento A** (Quatro variáveis): temperatura média, área de drenagem, precipitação média anual, longitude do centróide da área de drenagem;
- **Agrupamento B** (11 variáveis): temperatura média, área de drenagem, precipitação média anual, longitude do centróide da área de drenagem, coeficiente de variação da precipitação, declividade média da área de drenagem, altitude do centróide da área de drenagem, altitude da estação fluviométrica, coeficiente de variação da temperatura e densidade de drenagem.

A Figura 4.3 ilustra o esquema para a aplicação da metodologia de agrupamento entre as Bacias Hidrográficas do trabalho.

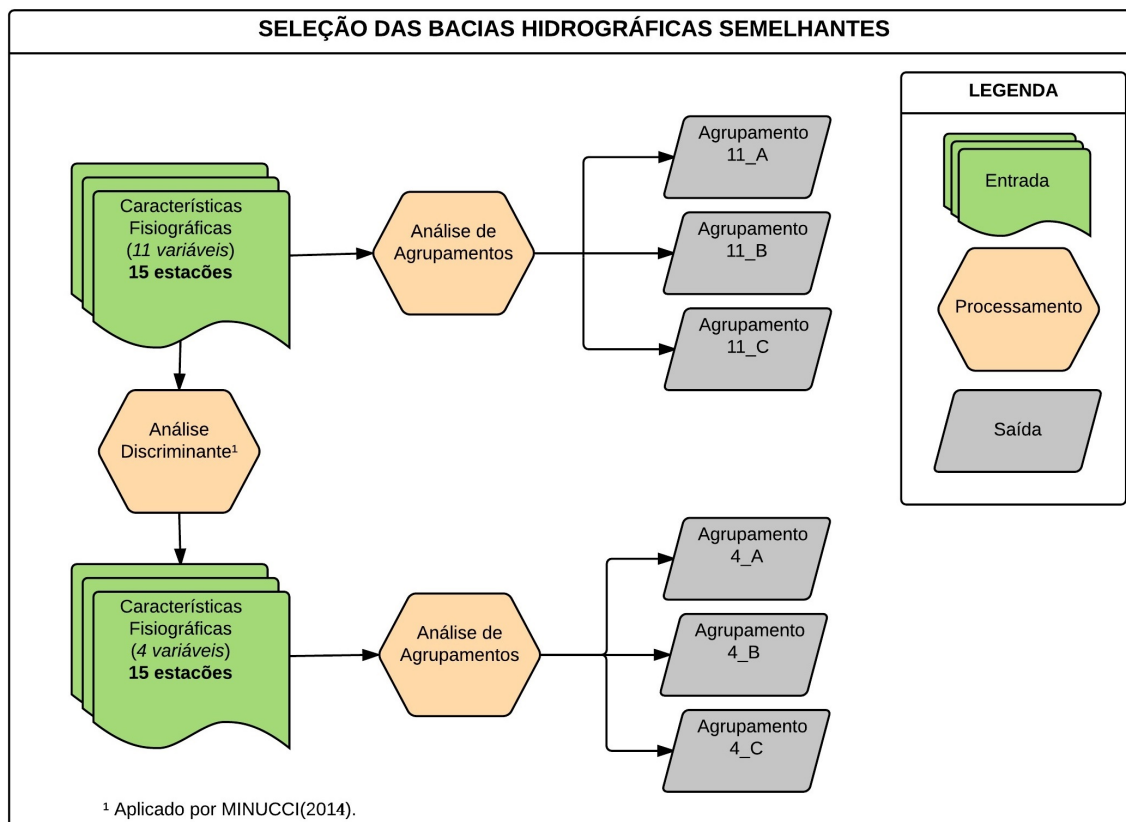


Figura 4.3: Abordagem adotada para aplicação dos agrupamentos.

4.8 Modelagem Chuva-Vazão

A modelagem chuva-vazão realizada no presente trabalho visam verificar a melhor configuração de treinamento da rede neural, a partir de dois diferentes conjuntos de dados de pluviometria e temperatura utilizados como entrada:

- **Modelo Chuva-Vazão I:** $P_m(t), P_m(t - 1), P_m(t - 2), T_m(t)$;
- **Modelo Chuva-Vazão II:** $P_1(t), P_1(t - 1), P_1(t - 2), P_2(t), P_2(t - 1), P_2(t - 2), P_3(t), P_3(t - 1), P_3(t - 2), P_4(t), P_4(t - 1), P_4(t - 2), T_m(t)$;

onde $P_1()$ — $P_4()$, referem-se a precipitação média mensal das estações pluviométricas 1—4; $P_m()$ a precipitação de Thiessen média da bacia; $T_m()$ a temperatura média da

bacia (calculadas através dos Polígonos de Thiessen); e as expressões t , $t - 1$, $t - 2$ ao mês atual, um e dois meses anteriores, respectivamente.

A Figura 4.4 apresenta o fluxograma de geração das RNAs de chuva-vazão.

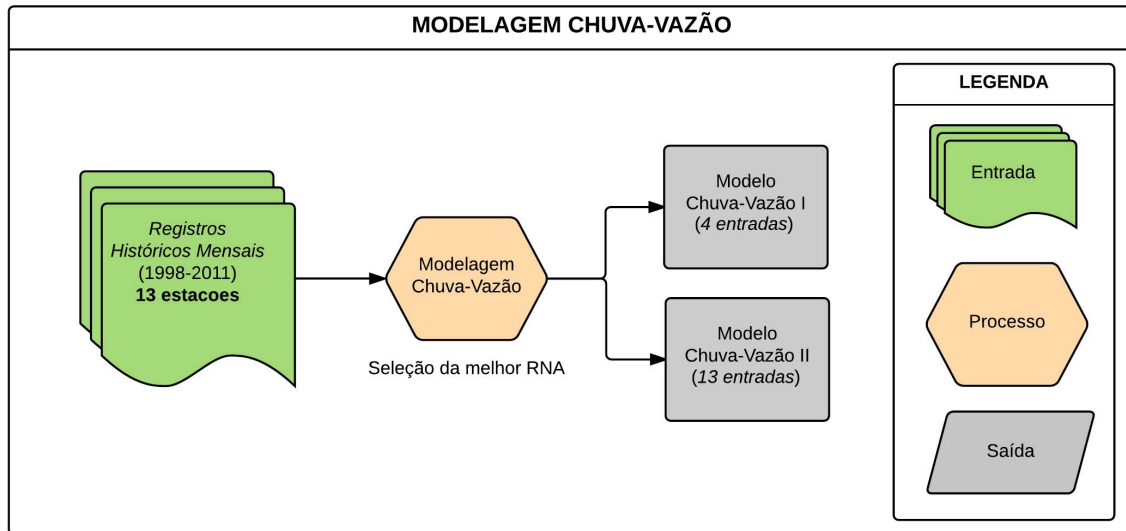


Figura 4.4: Esquema adotado para a modelagem chuva-vazão.

4.9 Regionalização de Vazões

A partir da definição da melhor configuração de treinamento da rede neural, assim como a seleção dos grupos semelhantes, foi realizada a regionalização dos modelos chuva-vazão a partir da operação das RNAs utilizando os dados das estações teste como entrada, e a vazão das estações auxiliares como saída.

Uma vez regionalizada as vazões das estações teste para cada estação auxiliar semelhante, foi calculada a média das vazões modeladas e confrontadas com a vazão observada nas estações teste. A Figura 4.5 exemplifica a abordagem adotada para regionalização das vazões para a estação teste.

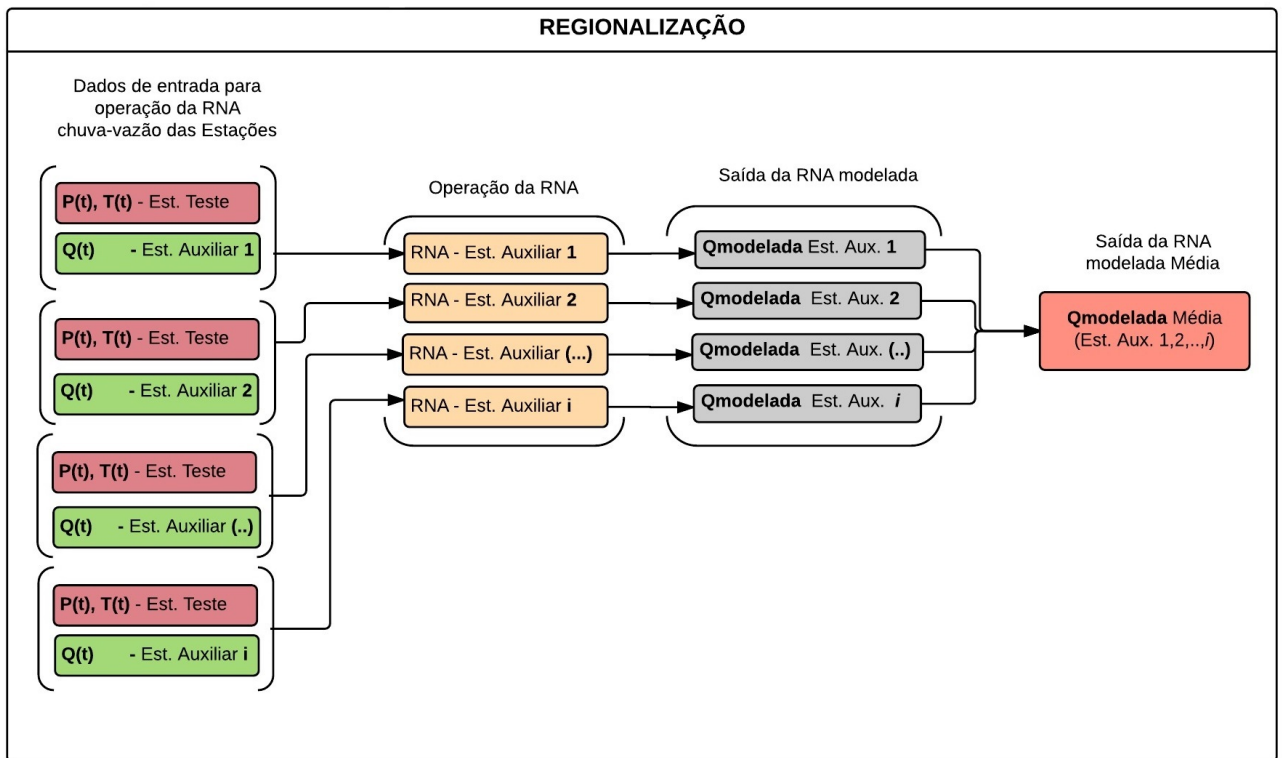


Figura 4.5: Representação esquemática da abordagem adotada para o cálculo da vazão modelada média a partir de i estações.

Tendo os modelos chuva-vazão calibrados e validados, os dados de entrada da RNA da estação auxiliar foram substituídos pelos registros de precipitação e temperatura da estação teste (regionalizada), sendo mantidos os registros de vazão da estação auxiliar.

Visando atenuar a variabilidade intrínseca de cada vazão gerada, é calculada a média das vazões modeladas, para também comparar com a vazão observada da estação teste. Wilks (2006) destaca que o comportamento temporal da média de conjunto (*ensemble mean*) é diferente da solução individual que utiliza a média dos valores iniciais de entrada¹, e conclui que, em geral, a melhor predição não é a que utiliza a melhor estimativa das condições iniciais, mas sim a que adota a média

¹Quando a função $f(x)$ é não-linear,

$$\frac{1}{n} \sum_i^n f(x_i) \neq f\left(\frac{1}{n} \sum_i^n x_i\right) \quad (4.2)$$

de conjunto. A Figura 4.6 mostra um exemplo de valores de vazão regionalizados através de RNAs calibradas das estações auxiliares 1, 2, 3 e 4; com destaque para a vazão regionalizada média e observada. Observa-se que a média dos valores modelados ($Q_Modelada-Média$), em geral, atenua os altos e baixos valores de vazões modeladas pelas estações auxiliares, se ajustando melhor a vazão observada ($Q_Observada$).

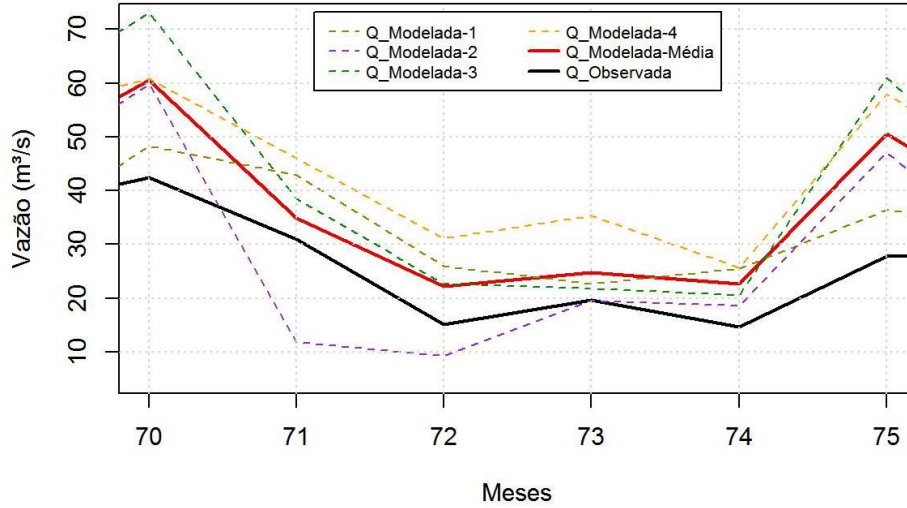


Figura 4.6: Exemplo de valores individuais modelados ($Q_Modeladas$ 1, 2, 3 e 4) e valores modelados médios.

4.9.1 Métricas de Validação

As métricas que foram utilizadas para avaliação da qualidade das previsões são as seguintes: Desvios de Vazões DV (Eq. 4.3), Média dos Valores Absolutos dos Desvios MAD (Eq. 4.4), Raiz do Erro Quadrático Médio $REQM$ (Eq. 4.5), Coeficiente de Nash-Sutcliffe $NASH$ (Eq. 4.6) e Coeficiente de Determinação R^2 (Eq. 4.7).

$$DV = \left(\frac{\sum_{t=1}^{t=T} Q_m^t - \sum_{t=1}^{t=T} Q_o^t}{\sum_{t=1}^{t=T} Q_o^t} \right) \times 100, \quad (4.3)$$

$$MAD = \frac{\sum_{t=1}^{t=T} |Q_m^t - Q_o^t|}{T}, \quad (4.4)$$

$$REQM = \sqrt{\frac{\sum_{t=1}^{t=T} (Q_o^t - \bar{Q}_m)^2}{T}}, \quad (4.5)$$

$$NASH = 1 - \frac{\sum_{t=1}^{t=T} (Q_m^t - Q_o^t)^2}{\sum_{t=1}^{t=T} (Q_o^t - \bar{Q}_o)^2}, \quad (4.6)$$

$$R^2 = \frac{\left(\sum_{t=1}^{t=T} (Q_o^t - \bar{Q}_o)(Q_m^t - \bar{Q}_m) \right)^2}{\sum_{t=1}^{t=T} (Q_o^t - \bar{Q}_o)^2 \sum_{t=1}^{t=T} (Q_m^t - \bar{Q}_m)^2}, \quad (4.7)$$

onde T é o número total de registros, Q_o é o valor de vazão observado e Q_m a vazão modelada.

5. Resultados

5.1 Seleção dos grupos semelhantes

Os resultados obtidos através do emprego do algoritmo SOM, desenvolvido em linguagem FORTRAN 95, são apresentados a seguir. Foram selecionadas as configurações da rede que permitiram verificar a estabilização das variações dentro de determinado *cluster*, a partir da variação de número de épocas, dimensionalidade (número de linhas e colunas) da estrutura topológica, e raio de influência na vizinhança. Os grupos formados e suas respectivas configurações são listadas nas Tabelas 5.1 e 5.2.

Tabela 5.1: Agrupamento A — formado a partir das quatro variáveis físicas/fisiográficas das bacias hidrográficas.

Guampará*	Balsa do Cantú	ETA - Guarapuava
Madeira Gavazzoni*	Porto Carriel	Fazenda Maracanã
Ponte Goio Bang	Porto Guarani	Ponto Leôncio Primo
Porto Santo Antônio	Tereza Cristina	Salto Sapucaí
Rio dos Patos		Tormenta Fazenda Trento
São Sebastião		

Configurações utilizadas: Épocas (150), Dimensionalidade (1 linha x 3 colunas), raio de influência (1).

*Estações teste adotadas.

Conforme verificado nas Tabelas 5.1 e 5.2, os agrupamentos formados se mantiveram diferentes a medida que se incluem variáveis físicas adicionais no processo de clusterização. Tanto a estação Guampará, como a Madeira Gavazzoni se deslocaram para diferentes grupos quando analisado o agrupamento com 11 variáveis.

Tabela 5.2: Agrupamento B — formado a partir das 11 variáveis físicas/fisiográficas das bacias hidrográficas.

Ponte Leôncio Primo	Balsa do Cantú	ETA - Guarapuava
Porto Santo Antônio	Guampará*	Fazenda Maracanã
Rio dos Patos	Ponto Goio Bang	Madeira Gavazzoni*
Salto Sapucaí	Porto Carriel	Tormenta Fazenda Trento
São Sebastião	Porto Guarani	
	Tereza Cristina	

Configurações utilizadas: Épocas (200), Dimensionalidade (1 linha x 3 colunas), raio de influência (1);

*Estações teste adotadas.

Apesar do deslocamento de grupos das estações Guampará e Madeira Gavazzoni, verifica-se que, no geral, as estações foram pouco deslocadas para outros grupos. Para as duas diferentes análises verificou-se no mínimo três estações em comum. As Figuras 5.1 e 5.2 destacam no mapa os agrupamentos A e B, respectivamente.

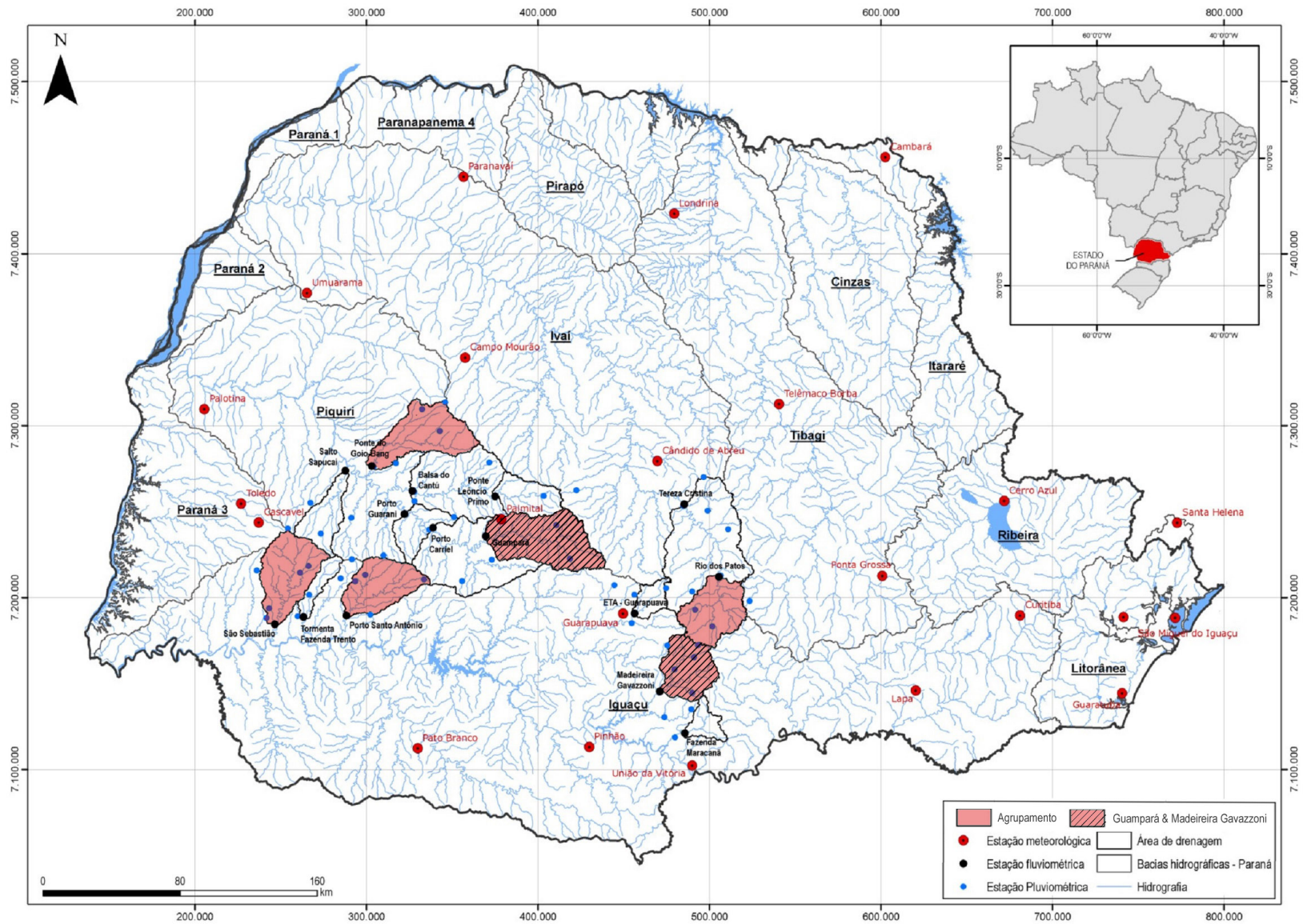


Figura 5.1: Agrupamento A — formado a partir de quatro variáveis físicas/fisiográficas das bacias hidrográficas. Fonte: Adaptado de Minucci (2014).

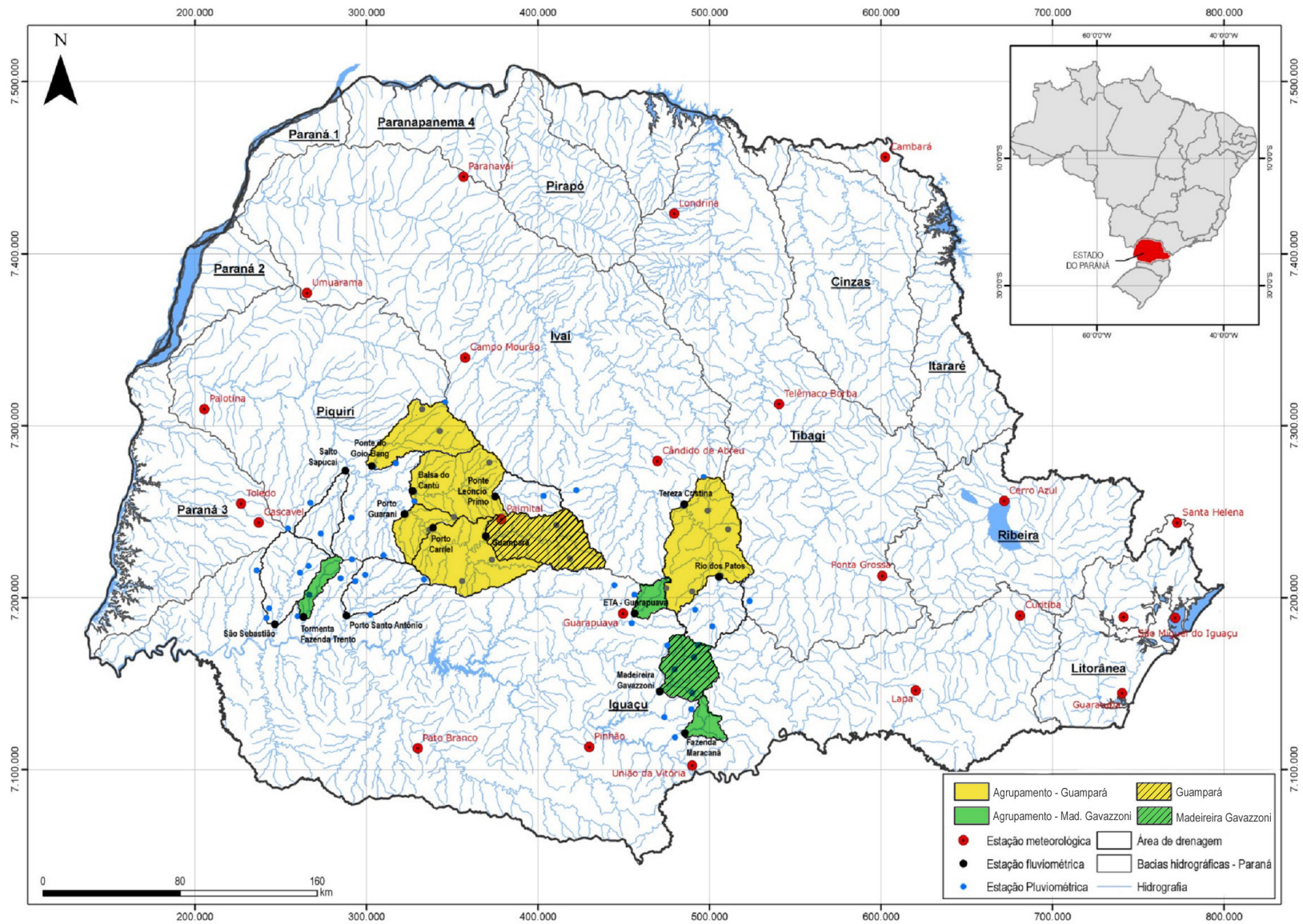


Figura 5.2: Agrupamento B — formado a partir de 11 variáveis físicas/fisiográficas das bacias hidrográficas. Fonte: Adaptado de Minucci (2014).

Conforme Figura 5.1, verificou-se que as estações definidas através do Agrupamento A se mantiveram distantes entre si, com estações localizadas tanto no extremo norte, como oeste e leste do conjunto de estações adotado para clusterização. Contudo, na Figura 5.2, referente ao Agrupamento B, se verificou que as estações semelhantes a estação Guampará se mantiveram mais próximas. Apenas a estação Tormenta Fazenda Trento (semelhante a Madeireira Gavazzoni), no caso do Agrupamento B, se manteve distante do restante das estações do conjunto.

5.2 Modelagem Chuva-Vazão

Diversos parâmetros de configuração de RNA foram testados visando obter melhor resultado de calibração, sendo as configurações que proporcionaram melhores resultados e atingiram o *overtraining*¹, listadas na Tabela 5.3.

O tempo médio computacional para o treinamento foi considerado rápido com relação ao tempo de processamento da rede MLP. Para um número neurônios na camada oculta de 15, 20, 25 e 30, o tempo aproximado foi de 5, 7, 9 e 12 minutos, respectivamente². As Tabelas 5.4 e 5.5, resumem as estatísticas médias relevantes dos modelos chuva-vazão I e II, respectivamente.

¹Quando a rede memoriza todos os registros de treinamento ao ter aprendido muitos exemplos de padrões de “entrada-saída” dos dados (Haykin, 1999).

²Processador Inter(R) Core(TM) 180 GHz e Memória RAM 8.0 GHz.

Tabela 5.3: Parâmetros de configuração de treinamento e operação da rede RBF, para o Modelo Chuva-Vazão I e II.

Variável	Valor
Número de iterações	10
Método de normalização dos dados	Mínimo/Máximo [-1,1]
Percentuais de dados para treinamento, teste e validação (respectivamente)	80%, 10% e 10%
Épocas de treinamento da 1ª camada da rede RBF	2 000
Taxa de aprendizado da 1ª camada da rede RBF	0,025
Erro máximo para treinamento da 1ª camada da rede RBF	10^{-7}
Épocas de treinamento da 2ª camada da rede RBF	10 000
Taxa de aprendizado da Variância σ_k^2 , Centro t_k e Pesos Sinápticos w_k da RBF	0,005
Função de ativação na saída	Linear
Número de neurônios da camada oculta da rede*	15, 20, 25 ou 30

*Variação de 15, 20, 25 ou 30 neurônios na camada oculta conforme melhores estatísticas de calibração de cada RNA.

5.3 Regionalização de Vazões

A seguir, serão apresentados os resultados da regionalização utilizando os modelos chuva-vazão I e II, para os agrupamento A e B. Em cada subseção são mostradas as vazões médias modeladas a partir dos dois diferentes agrupamentos.

5.3.1 Modelo Chuva-Vazão I

Os resultados da regionalização foram divididos por estação teste - Madeira Gavazzoni e Guampará, apresentados na sequência.

Tabela 5.4: Estatísticas médias da calibração e operação do Modelo Chuva-Vazão I (Modelo de 4 entradas, 1 saída) de cada estação auxiliar.

Estação	Calibração					Operação			
	REQM	NASH	EMA	ER	EM	REQM	NASH	EMA	DV
Balsa do Cantú	39,47	0,38	30,01	-28,86	0,0352	25,61	0,81	19,55	-22,43
ETA Guarapava	3,46	0,65	2,34	-12,66	0,0164	1,77	0,94	1,34	-7,31
Fazenda Maracanã	3,83	0,70	2,76	-14,13	0,0203	2,78	0,87	2,03	-13,50
Ponte do Goio Bang	10,21	0,40	7,65	-11,82	0,0284	6,92	0,79	5,40	-9,33
Porto Carriel	38,82	0,70	29,62	-19,21	0,0120	23,59	0,94	17,70	-12,81
Porto Guarani	45,65	0,72	35,01	-23,07	0,0137	30,42	0,92	23,12	-14,59
Porto Santo Antônio	17,13	0,24	11,67	-35,62	0,0238	9,29	0,89	6,63	-21,29
Rio dos Patos	13,23	0,69	9,18	-14,57	0,0135	8,71	0,90	6,31	-14,15
São Sebastião	18,50	0,43	13,42	-25,67	0,0238	11,88	0,87	8,41	-17,63
Tereza Cristina	35,53	0,66	26,05	-16,21	0,0155	21,45	0,92	16,02	-11,29
Tormenta Fazenda Trento	5,59	0,37	3,80	-43,25	0,0299	3,63	0,83	2,61	-24,79

REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; EMA: Erro Médio Absoluto; DV: Desvio de Vazões (%); ER: Erro Relativo; EM: Erro Médio.

Madeira Gavazzoni - Agrupamento A e B

As Tabelas 5.6 e 5.7 resumem as estatísticas médias e individuais das RNAs regionalizadas para os Agrupamentos A e B, respectivamente. As relações vazão específica média modelada e vazão observada são mostradas nas Figuras 5.3 e 5.5, com a reta parametrizada e a função identidade, visando comparar a qualidade das predições com relação ao ideal (1 : 1). As Figuras 5.4 e 5.6 apresentam o histórico de vazão específica ao longo do tempo para os Agrupamentos A e B, respectivamente.

Tabela 5.5: Estatísticas médias da calibração e operação do Modelo Chuva-Vazão II (Modelo de 13 entradas, 1 saída) de cada estação auxiliar.

Estação	Calibração					Operação			
	REQM	NASH	EMA	ER	EM	REQM	NASH	EMA	DV
Balsa do Cantú	38,14	0,13	29,96	-49,05	0,0331	21,34	0,87	17,43	-26,23
ETA - Guarapua	4,10	0,45	3,03	-22,00	0,0214	1,86	0,93	1,50	-9,02
Fazenda Maracanã	3,73	0,73	2,75	-18,72	0,0197	2,31	0,91	1,77	-12,85
Ponte do Goio Bang	9,72	0,44	7,52	-11,83	0,0257	6,42	0,82	4,93	-7,45
Porto Carriel	58,39	0,38	42,38	-40,86	0,0260	27,69	0,91	20,67	-18,85
Porto Guarani	59,09	0,57	43,31	-30,80	0,0236	31,54	0,91	23,74	-17,21
Porto Santo Antônio	19,81	0,06	14,14	-44,76	0,0297	9,38	0,89	7,06	-23,00
Rio dos Patos	17,09	0,46	12,89	-46,35	0,0217	10,50	0,86	8,09	-31,09
São Sebastião	22,23	0,09	15,62	-26,33	0,0337	12,60	0,85	9,19	-18,47
Tereza Cristina	39,65	0,60	30,15	-28,34	0,0196	20,51	0,92	15,90	-15,74
Tormenta Fazenda Trento	6,01	0,41	4,32	-50,00	0,0352	3,58	0,83	2,68	-23,33

REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; EMA: Erro Médio Absoluto; DV: Desvio de Vazões (%); ER: Erro Relativo; EM: Erro Médio.

Tabela 5.6: Estatísticas da regionalização para a estação teste Madeireira Gavazzoni realizadas através do Modelo Chuva-Vazão I e do agrupamento A.

Vazões	Cor.	R ²	NASH	REQM	MAD	DV(%)
Ponto do Goio Bang	0,798	0,637	0,609	13,626	2,345	8,26
Porto Santo Antônio	0,913	0,834	0,783	10,150	0,793	2,79
Rio dos Patos	0,928	0,862	0,668	12,566	7,097	25,00
São Sebastião	0,904	0,818	0,368	17,325	12,19	42,93
M.V.M.	0,932	0,869	0,793	9,926	5,606	19,74

Cor.: Correlação Linear; REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; MAD: Média Absoluta dos Desvios; DV: Desvio de Vazões (%).

M.V.M.: Média das Vazões Modeladas.

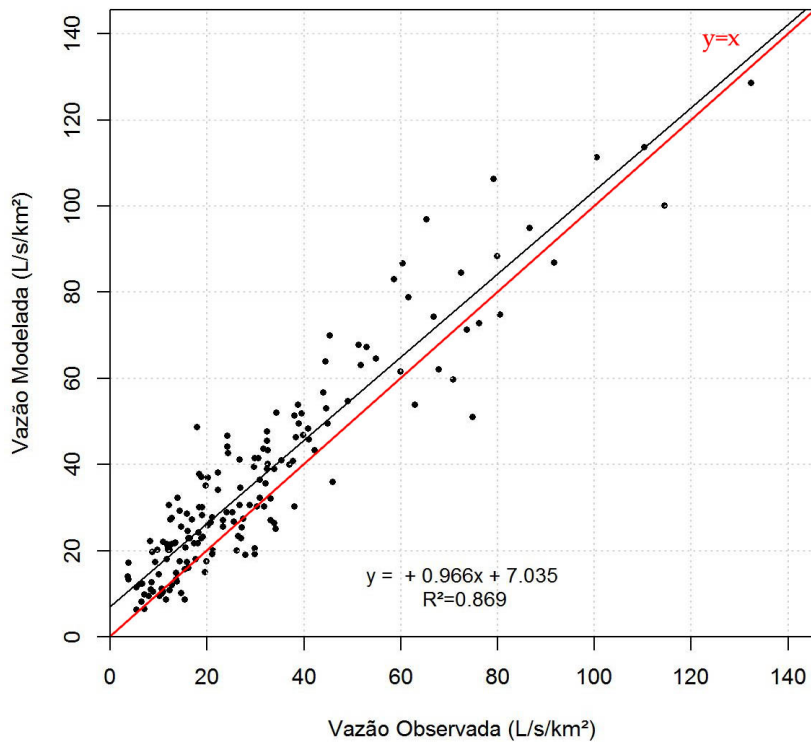


Figura 5.3: Relação Vazão específica média modelada x Vazão específica observada para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão I e do agrupamento A.

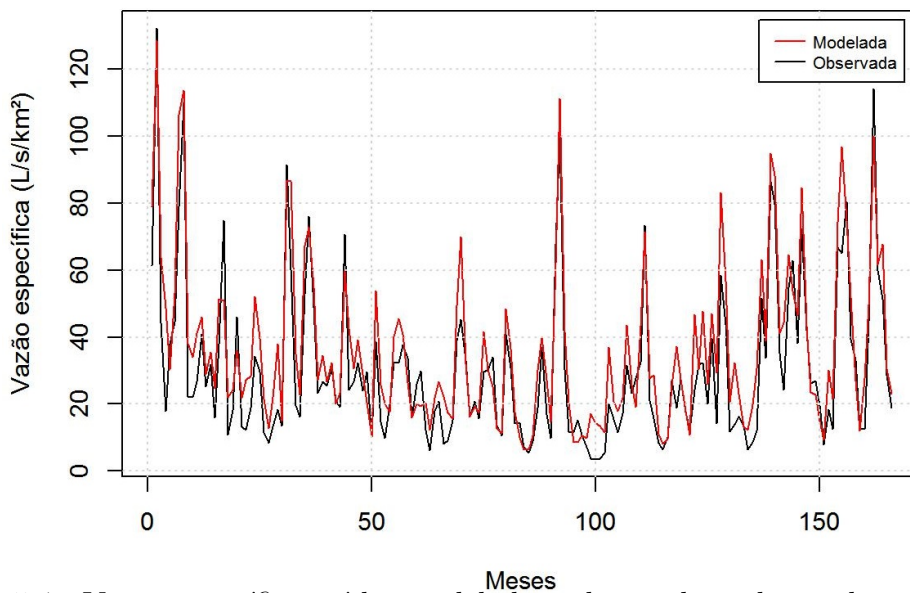


Figura 5.4: Vazão específica média modelada e observada ao longo do tempo para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão I e do agrupamento A.

Tabela 5.7: Estatísticas da regionalização para a estação teste Madeireira Gavazzoni realizadas através do Modelo Chuva-Vazão I e do agrupamento B.

Vazões	Cor.	R ²	NASH	REQM	MAD	DV(%)
ETA Guarapuava	0,943	0,890	-0,28	24,660	-18,969	-66,81
Fazenda Maracanã	0,906	0,821	-0,07	22,498	-16,517	-58,17
Tormenta Fazenda Trento	0,879	0,772	-0,22	24,074	-18,127	-63,84
M.V.M.	0,937	0,877	-0,18	23,673	-17,871	-62,94

Cor.: Correlação Linear; REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; MAD: Média Absoluta dos Desvios; DV: Desvio de Vazões (%).

M.V.M.: Média das Vazões Modeladas.

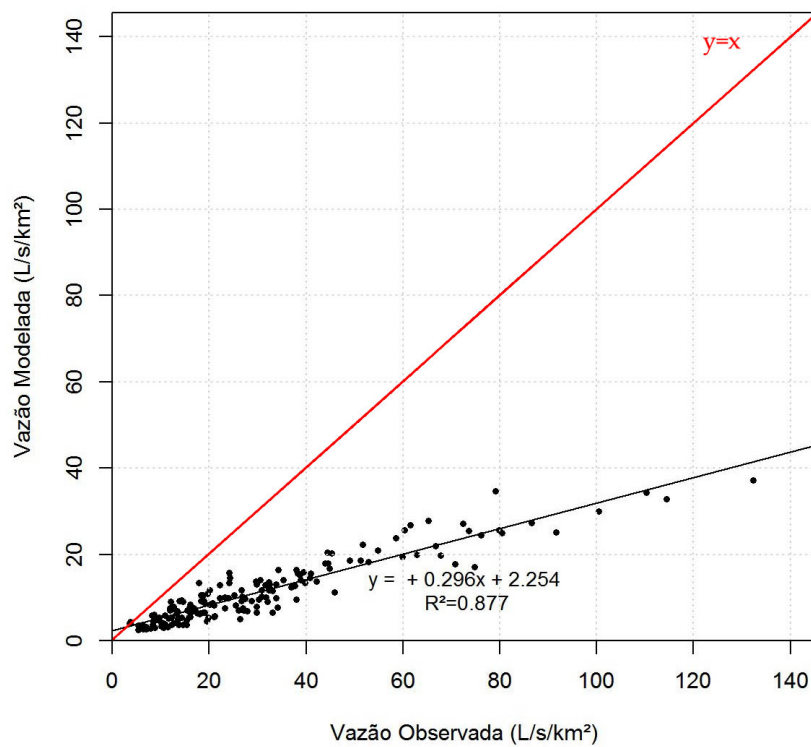


Figura 5.5: Relação Vazão específica média modelada x Vazão específica observada para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão I e do agrupamento B.

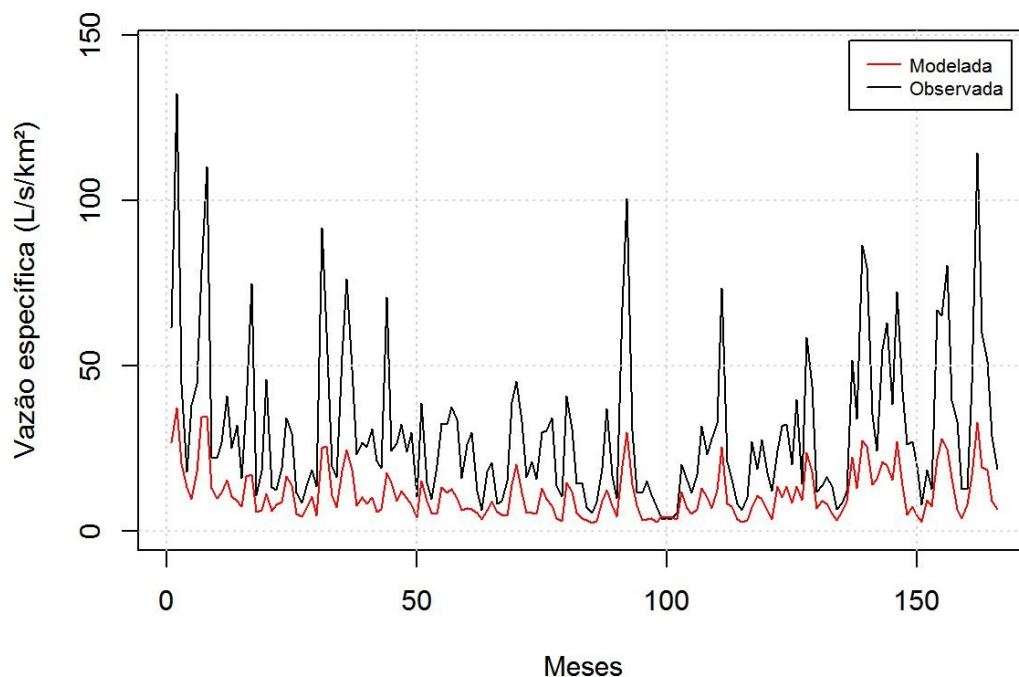


Figura 5.6: Vazão específica média modelada e observada ao longo do tempo para a estação teste Madeira Gavazzoni através do Modelo Chuva-Vazão I e do agrupamento B.

Guamará - Agrupamentos A e B

As Tabelas 5.8 e 5.9 resumem as estatísticas médias e individuais das RNAs regionalizadas para os Agrupamentos A e B, respectivamente. As relações vazão específica média modelada e vazão observada são mostradas nas Figuras 5.8 e 5.9, com a reta parametrizada e a função identidade, visando comparar a qualidade das predições com relação ao ideal (1 : 1). As Figuras 5.7 e 5.10 apresentam o histórico de vazão específica ao longo do tempo para os Agrupamentos A e B, respectivamente.

Tabela 5.8: Estatísticas da regionalização para a estação teste Guampará realizadas através do Modelo Chuva-Vazão I e do agrupamento A.

Vazões	Cor.	R ²	NASH	REQM	MAD	DV(%)
Ponto do Goio Bang	0,752	0,566	0,535	16,625	0,619	2,051
Porto Santo Antônio	0,908	0,824	0,795	11,047	-3,497	-11,579
Rio dos Patos	0,891	0,793	0,716	13,002	5,115	16,936
São Sebastião	0,881	0,777	0,594	15,536	8,149	26,985
M.V.M.	0,906	0,82	0,809	10,655	2,597	8,598

Cor.: Correlação Linear; REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; MAD: Média Absoluta dos Desvios; DV: Desvio de Vazões (%).

M.V.M.: Média das Vazões Modeladas.

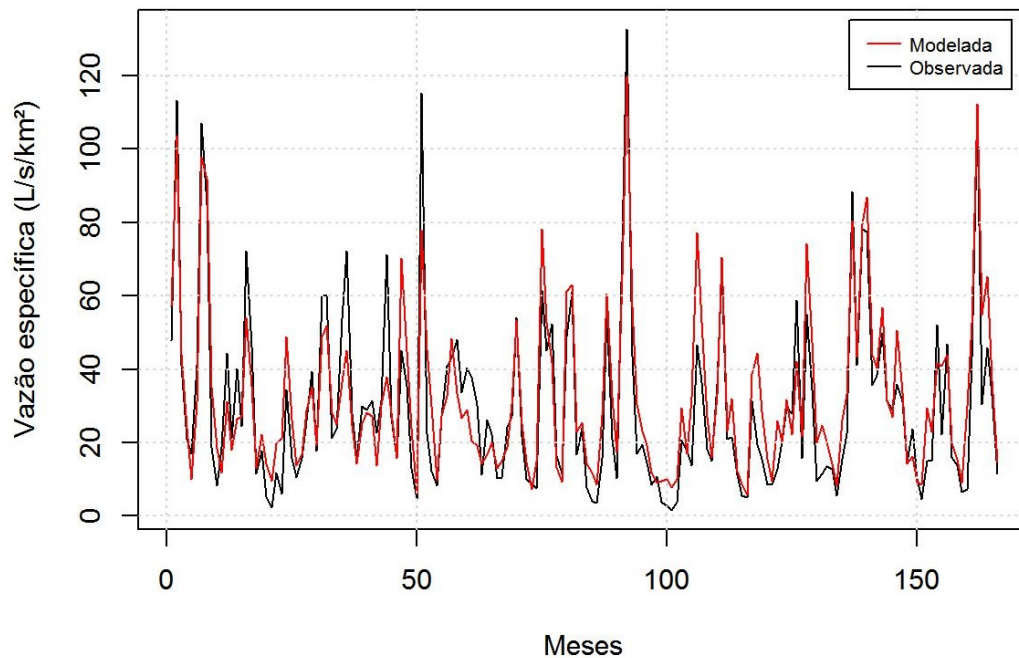


Figura 5.7: Vazão específica média modelada e observada ao longo do tempo para a estação teste Guampará através do Modelo Chuva-Vazão I e do agrupamento A.

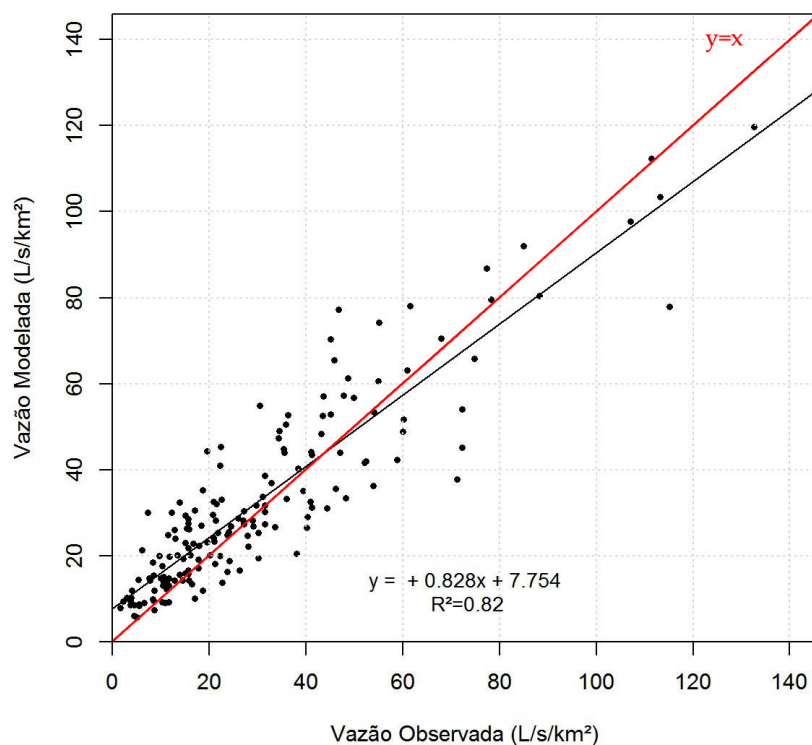


Figura 5.8: Relação Vazão específica média modelada x Vazão específica observada para a estação teste Guampará através do Modelo Chuva-Vazão I e do agrupamento A.

Tabela 5.9: Estatísticas da regionalização para a estação teste Guampará realizadas através do Modelo Chuva-Vazão II e do agrupamento B.

Vazões	Cor.	R ²	NASH	REQM	MAD	DV(%)
Balsa do Cantú	0,821	0,674	-6,676	67,541	56,255	186,27
Ponte do Goio Bang	0,752	0,566	0,535	16,625	0,619	2,05
Porto Carriel	0,930	0,865	-12,869	90,784	67,056	222,04
Porto Guarani	0,922	0,85	-19,380	110,05	82,711	273,87
Tereza Cristina	0,894	0,799	-10,705	83,401	63,956	211,77
M.V.M.	0,918	0,843	-6,775	67,973	54,119	179,20

Cor.: Correlação Linear; REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; MAD: Média Absoluta dos Desvios; DV: Desvio de Vazões (%).

M.V.M.: Média das Vazões Modeladas.

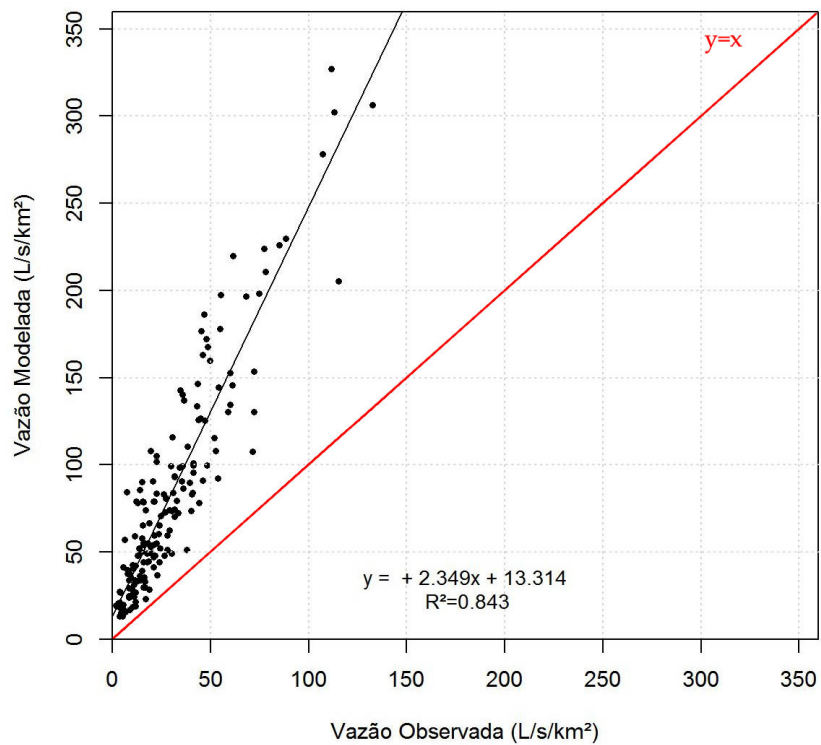


Figura 5.9: Relação Vazão específica média modelada x Vazão específica observada para a estação teste Guampará através do Modelo Chuva-Vazão I e do agrupamento B.

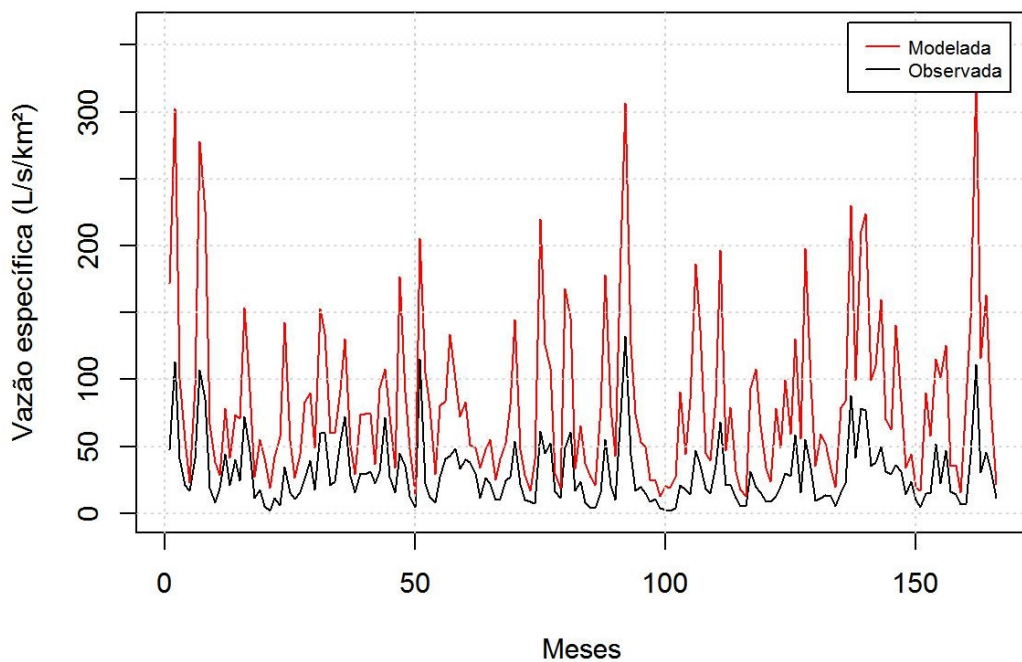


Figura 5.10: Vazão específica média modelada e observada ao longo do tempo para a estação teste Guampará através do Modelo Chuva-Vazão I e do agrupamento B.

5.3.2 Modelo Chuva-Vazão II

Madeira Gavazzoni - Agrupamento A e B

As Tabelas 5.10 e 5.11 resumem as estatísticas médias e individuais das RNAs regionalizadas para os Agrupamentos A e B, respectivamente. As relações vazão específica média modelada e vazão observada são mostradas nas Figuras 5.11 e 5.13, com a reta parametrizada e a função identidade, visando comparar a qualidade das predições com relação ao ideal (1 : 1). As vazões ao longo do tempo são mostradas nas Figuras 5.12 e 5.14, para os Agrupamentos A e B, respectivamente.

Tabela 5.10: Estatísticas da regionalização para a estação teste Madeira Gavazzoni realizadas através do Modelo Chuva-Vazão II e do agrupamento A.

Vazões	Cor.	R ²	NASH	REQM	MAD	DV(%)
Ponte do Goio Bang	0,770	0,593	0,544	14,72	3,180	11,199
Porto Santo Antônio	0,875	0,765	0,745	11,00	0,407	1,434
Rio dos Patos	0,878	0,77	0,558	14,49	7,361	25,925
São Sebastião	0,865	0,747	0,331	17,83	12,422	43,752
M.V.M.	0,895	0,801	0,726	11,40	5,842	20,578

Cor.: Correlação Linear; REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; MAD: Média Absoluta dos Desvios; DV: Desvio de Vazões (%).

M.V.M.: Média das Vazões Modeladas.

Tabela 5.11: Estatísticas da regionalização para a estação teste Madeira Gavazzoni realizadas através do Modelo Chuva-Vazão II e do agrupamento B.

Vazões	Cor.	R ²	NASH	REQM	MAD	DV(%)
ETA Guarapuava	0,885	0,783	-0,332	25,151	-19,034	-67,040
Fazenda Maracanã	0,903	0,816	-0,238	24,244	-18,530	-65,264
Tormenta Fazenda Trento	0,711	0,506	-0,350	25,319	-18,396	-64,792
M.V.M.	0,896	0,803	-0,292	24,769	-18,653	-65,699

Cor.: Correlação Linear; REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; MAD: Média Absoluta dos Desvios; DV: Desvio de Vazões (%).

M.V.M.: Média das Vazões Modeladas.

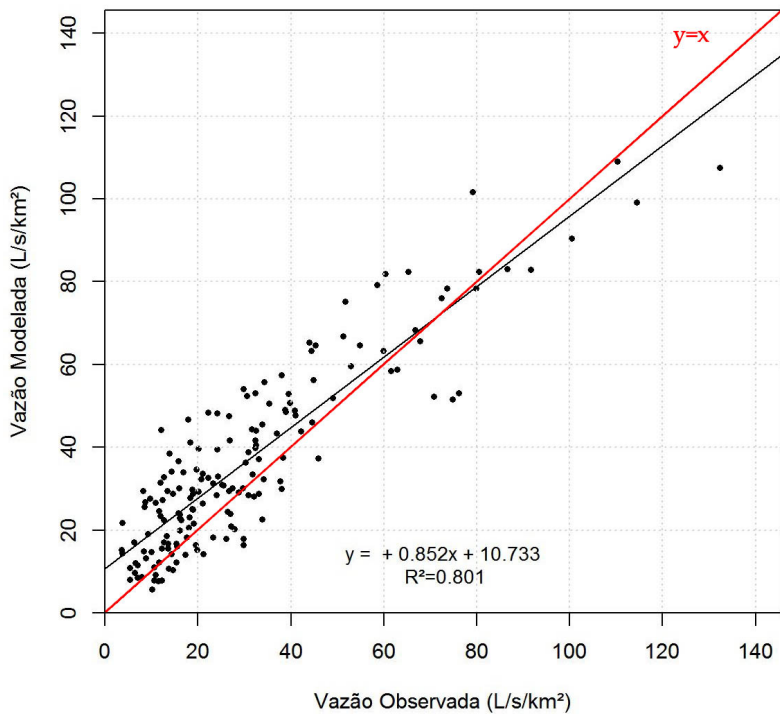


Figura 5.11: Relação Vazão específica média modelada x Vazão específica observada para a estação teste Madeira Gavazzoni através do Modelo Chuva-Vazão II e do agrupamento A.

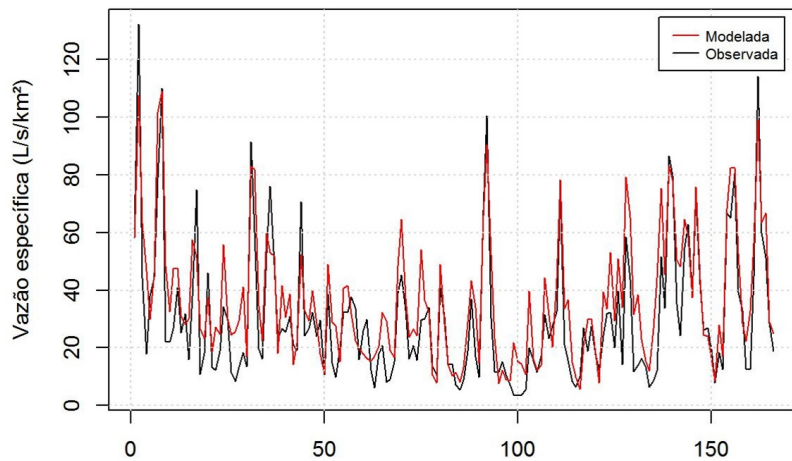


Figura 5.12: Vazão específica média modelada e observada ao longo do tempo para a estação teste Madeira Gavazzoni através do Modelo Chuva-Vazão II e do agrupamento A.

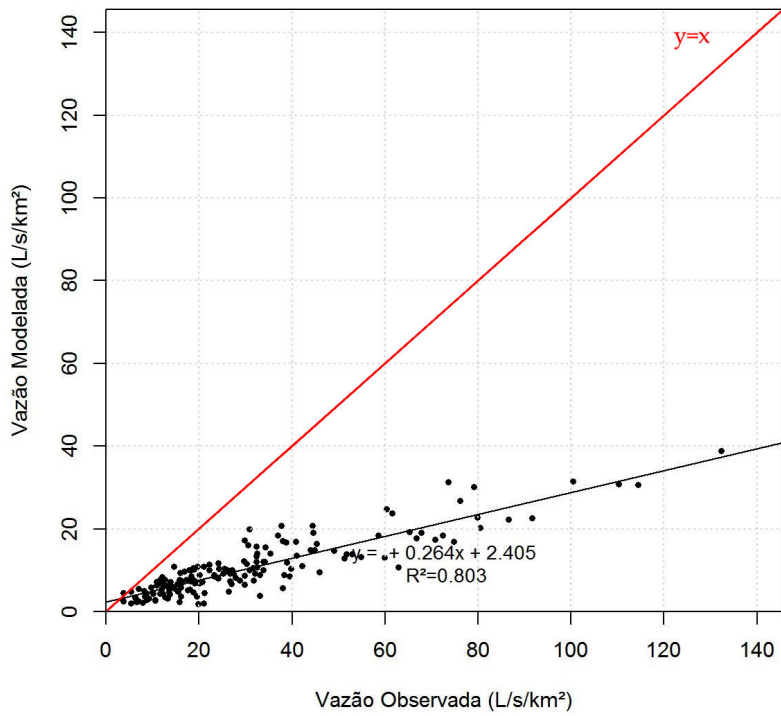


Figura 5.13: Relação Vazão específica média modelada x Vazão específica observada para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão II e do agrupamento B.

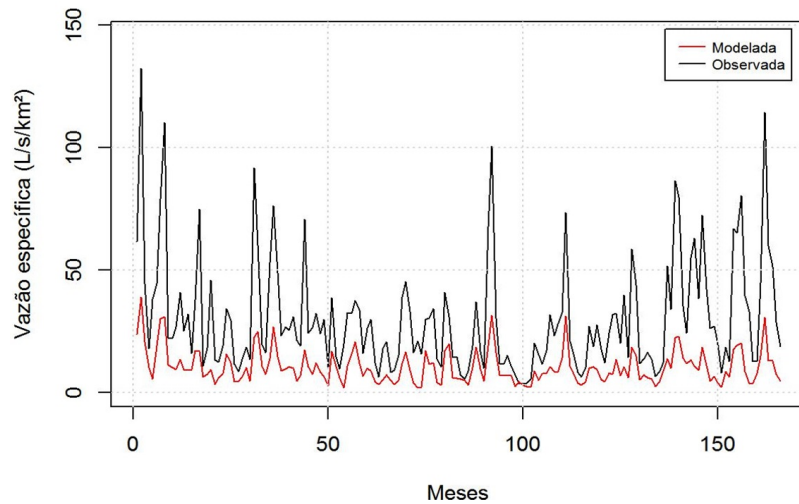


Figura 5.14: Vazão específica média modelada e observada ao longo do tempo para a estação teste Madeireira Gavazzoni através do Modelo Chuva-Vazão II e do agrupamento B.

Guamará - Agrupamento A e B

As Tabelas 5.12 e 5.13 resumem as estatísticas médias e individuais das RNAs regionalizadas para os Agrupamentos A e B, respectivamente. As relações

vazão específica média modelada e vazão observada são mostradas nas Figuras 5.15 e Figura 5.17, com a reta parametrizada e a função identidade, visando comparar a qualidade das predições com relação ao ideal (1 : 1). As vazões ao longo do tempo são mostradas nas Figuras 5.16 e 5.18, para os Agrupamentos A e B, respectivamente.

Tabela 5.12: Estatísticas da regionalização para a estação teste Guampará realizadas através do Modelo Chuva-Vazão II e do agrupamento A.

Vazões	Cor.	R ²	NASH	REQM	MAD	DV(%)
Ponte do Goio Bang	0,708	0,501	0,467	17,79	2,300	7,616
Porto Santo Antônio	0,866	0,750	0,739	12,458	-1,454	-4,815
Rio dos Patos	0,829	0,687	0,549	16,365	6,204	20,543
São Sebastião	0,830	0,690	0,446	18,153	9,686	32,074
M.V.M.	0,859	0,738	0,707	13,189	4,184	13,854

Cor.: Correlação Linear; REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; MAD: Média Absoluta dos Desvios; DV: Desvio de Vazões (%).

M.V.M.: Média das Vazões Modeladas.

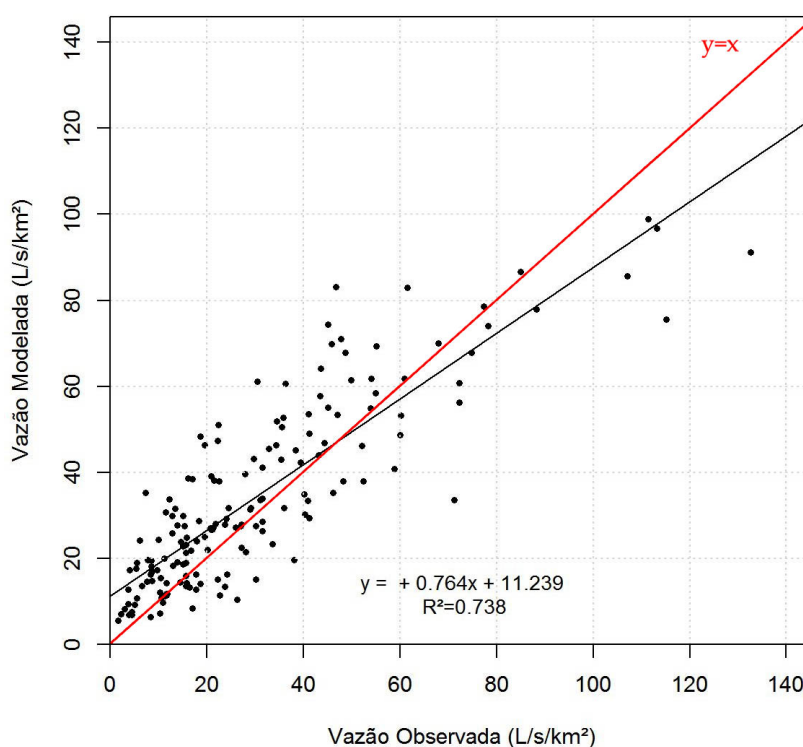


Figura 5.15: Relação Vazão específica média modelada x Vazão específica observada para a estação teste Guampará através do Modelo Chuva-Vazão II e do agrupamento A.

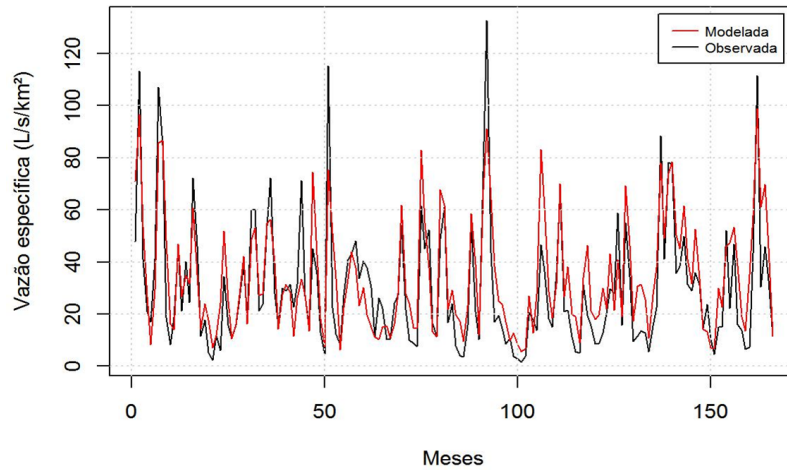


Figura 5.16: Vazão específica média modelada e observada ao longo do tempo para a estação teste Guampará através do Modelo Chuva-Vazão II e do agrupamento A.

Tabela 5.13: Estatísticas da série regionalizada através do Modelo Chuva-Vazão II e do agrupamento B - a estação Guampará.

Vazões	Cor.	R ²	NASH	REQM	MAD	DV(%)
Balsa do Cantú	0,815	0,665	-5,763	63,397	53,800	178,143
Ponto do Goio Bang	0,708	0,501	0,467	17,790	2,300	7,616
Porto Carriel	0,911	0,829	-14,481	95,915	73,925	244,782
Porto Guarani	0,905	0,818	-21,088	114,57	89,118	295,090
Tereza Cristina	0,842	0,709	-10,536	82,798	61,711	204,337
M.V.M.	0,895	0,8	-7,081	69,298	56,171	185,994

Cor.: Correlação Linear; REQM: Raiz do Erro Quadrático Médio; NASH: Coeficiente de Nash-Sutcliffe; MAD: Média Absoluta dos Desvios; DV: Desvio de Vazões (%).

M.V.M.: Média das Vazões Modeladas.

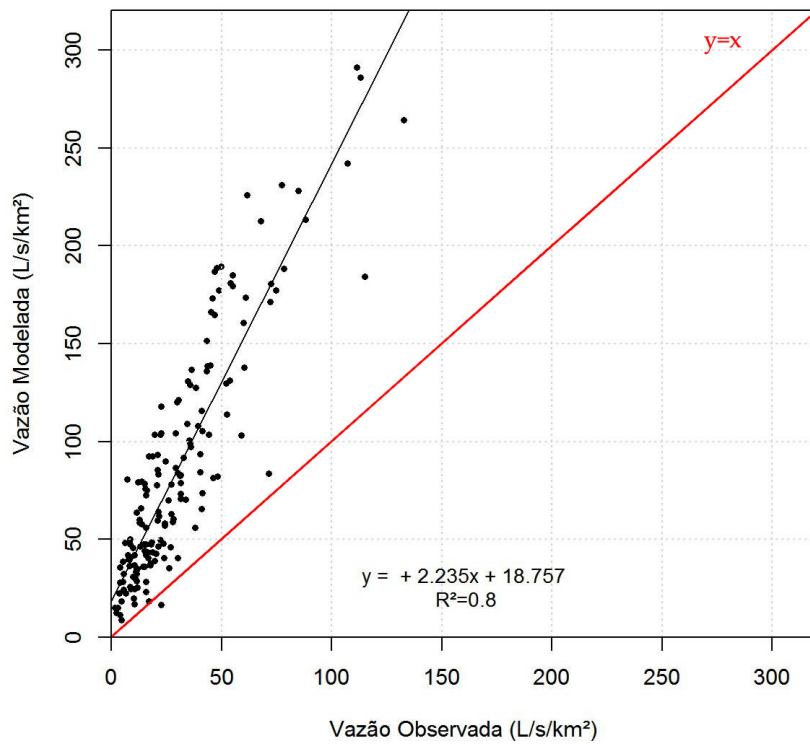


Figura 5.17: Relação Vazão específica média modelada x Vazão específica observada para a estação teste Guampará através do Modelo Chuva-Vazão II e do agrupamento B.

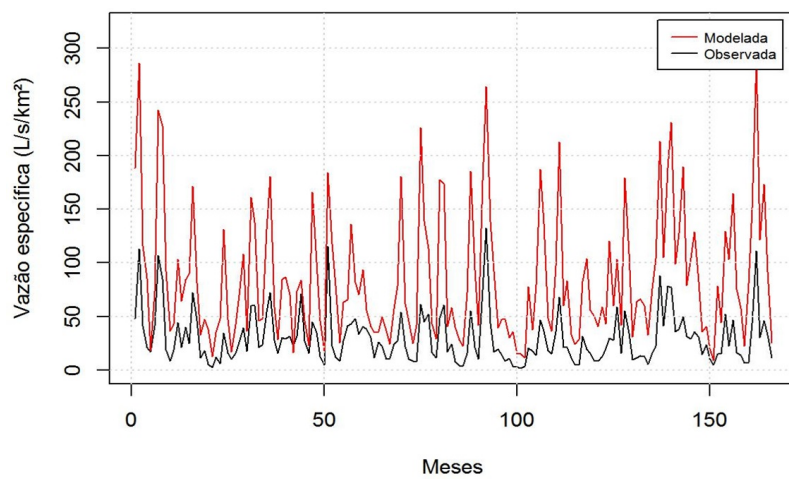


Figura 5.18: Vazão específica média modelada e observada ao longo do tempo para a estação teste Guampará através do Modelo Chuva-Vazão II e do agrupamento B.

6. Discussão

A discussão dos resultados está dividida conforme os itens a seguir:

- Seleção dos grupos semelhantes
- Modelagem chuva–vazão
- Regionalização de vazões

6.1 Seleção dos grupos semelhantes

A Tabela 6.1 resume os agrupamentos formados a partir da aplicação do algoritmo SOM — *Self Organized Maps* — (tanto para quatro variáveis físicas/fisiográficas, como para onze), assim como os agrupamentos gerados a partir da aplicação do algoritmo *k-means*.

Conforme a Tabela 6.1, verificou-se que os algoritmos de agrupamentos mantiveram-se com, no mínimo, duas estações similares, sendo:

- Porto Santo Antônio e São Sebastião;
- Balsa do Cantú, Porto Carriel, Porto Guarani, Tereza Cristina;
- Fazenda Maracanã e ETA Guarapuava.

Quando comparado os grupos formados a partir do agrupamento A e do definido por Minucci (2014) (os quais são formados a partir das mesmas quatro variáveis físicas), verifica-se que apenas um grupo obteve maior número de estações em

Tabela 6.1: *Clusters* formados a partir de diferentes entradas do algoritmo SOM e o método *k-means* - proposto por Minucci (2014).

Agrupamento A ¹	- <u>Guampará</u> - <u>Madeiraira Gavazzoni</u> - Ponte Goio Bang - Porto Santo Antônio - Rio dos Patos - São Sebastião	- Balsa do Cantú - Porto Carriel - Porto Guarani - Tereza Cristina	- ETA Guarapuava - Fazenda Maracanã - Ponto Leôncio Primo - Salto Sapucaí - Tormenta Fazenda Trento
Agrupamento B ²	- Ponto Leôncio Primo - Porto Santo Antônio - Rio dos Patos - Salto Sapucaí - São Sebastião	- Balsa do Cantú - <u>Guampará</u> - <u>Ponto Goio Bang</u> - Porto Carriel - Porto Guarani - Tereza Cristina	- ETA Guarapuava - Fazenda Maracanã - <u>Madeiraira Gavazzoni</u> - Tormenta Fazenda Trento
Minucci (2014) ³	- Ponto Goio Bang - Ponte Leôncio Primo - Porto Santo Antônio - Salto Sapucaí - São Sebastião - Tormenta Fazenda Trento	- Balsa do Cantú - Porto Carriel - Porto Guarani - Tereza Cristina - <u>Guampará</u>	- ETA Guarapuava - Fazenda Maracanã - Rio dos Patos - <u>Madeiraira Gavazzoni</u>

¹Agrupamento formado a partir da clusterização de quatro variáveis físicas/fisiográficas

²Agrupamento formado a partir da clusterização de 11 variáveis física/fisiográficas

³Agrupamento formado a partir do emprego do algoritmo *k-means* - utilizando 4 variáveis físicas/fisiográficas, por Minucci (2014).

comum - sendo quatro no total. Contudo, nesta situação, as estações teste Guampará e Madeiraira Gavazzoni se encontraram em diferentes grupos.

As estações teste também se deslocaram de grupo para as simulações deste estudo (Agrupamento A e B), apesar de serem verificadas 10 estações sem alteração de *cluster*. Contudo, uma vez que as estações teste do presente estudo se deslocaram de grupo, espera-se alterações mais significativas quando aplicada a regionalização de vazões.

Kalteh et al. (2007) destacam a importância de se introduzir dados de entrada “eficientes”, uma vez que a simples apresentação de todas as potenciais variáveis para a rede neural e a crença de que a rede identificará as mais cruciais pode criar problemas. As principais desvantagens dessa abordagem, além da necessidade de maior memória computacional, são: dificuldade de aprendizagem; a não convergência

e baixo desempenho do modelo; dificuldade de compreensão do modelo; e o aumento de ruído devido à introdução de variáveis irrelevantes (Kalteh et al., 2007; Bowden et al., 2005).

6.2 Modelagem Chuva-Vazão

Durante a calibração das RNAs Chuva-Vazão, verificou-se, de modo geral, baixos coeficientes de NASH para as 11 estações auxiliares modeladas. Para o Modelo Chuva-Vazão I (4 entradas, 1 saída), as estações Porto Carriel e Rio dos Patos apresentaram valores de NASH 0,66 e 0,7, respectivamente, com baixas REQM (Raiz do Erro Quadrático Médio); enquanto as estações Porto Carriel e Porto Guarani apresentaram valores de NASH de 0,7 e 0,72, respectivamente, com valores REQM mais elevadas (superiores a $38 \text{ m}^3/\text{s}$). Para a operação, verificou-se altos valores de NASH (superiores a 0,79) e baixas REQM para todas as estações.

O Modelo Chuva-Vazão II (13 entradas, 1 saída) apresentou estatísticas médias de calibração e operação inferiores aos verificados no Modelo I. Tal resultado é diferente ao verificado por Minucci (2014), que não obteve resultados satisfatórios quando utilizada apenas a precipitação média mensal (obtida pelo método de Thiessen). Para o referido trabalho, ao empregar o Modelo Chuva-Vazão II, “A variância inserida na amostra trazia mais informações às redes, fazendo com que elas “aprendessem” melhor as características no entorno da região e levassem à melhores resultados no momento da validação”. A razão pela qual verificou-se baixos valores de Nash-Sutcliffe durante a calibração se dá pelo fato de se utilizar menor percentual dos registros para a validação - 10% (17 registros de um total de 165).

6.3 Regionalização de Vazões

Foram definidos quatro diferentes resultados da regionalização de vazões para a estação teste Madeireira Gavazzoni e Guampará, totalizando oito simulações. A

Tabela 6.2 resume as simulações que obtiveram melhores resultados - sendo utilizado como critério de avaliação os coeficientes de determinação R^2 e NASH.

Tabela 6.2: Resumo dos valores R^2 e NASH obtidos para os diferentes modelos e agrupamento utilizando, assim como as estatísticas encontradas por Minucci (2014).

			Guampará	Madeira Gavazzoni
Modelo Chuva-Vazão I	Agrupamento A	R^2	0,82	0,87
		NASH	0,81	0,79
	Agrupamento B	R^2	0,843	0,877
		NASH	-6,775	-0,18
Modelo Chuva-Vazão II	Agrupamento A	R^2	0,74	0,88
		NASH	0,71	0,73
	Agrupamento B	R^2	0,82	0,803
		NASH	-7,08	-0,292
Minucci (2014)	Agrupamento	R^2	0,82	0,88
		NASH	0,82	0,88

Ressalta-se que os resultados obtidos por Minucci (2014) não são diretamente comparáveis com os modelos chuva-vazão acima apresentados, uma vez que os agrupamentos das estações semelhantes foram distintos. As Tabelas 6.3 e 6.4 resumem os valores modelados para as estações teste e seus respectivos registros observados.

Destaca-se que o Modelo Chuva-Vazão II utilizou 13 registros de dados como entrada para a rede, sendo quatro registros mensais médios de $p(t)$, $p(t-1)$ e $p(t-2)$ das quatro estações pluviométricas mais representativas da bacia, totalizando 12; e a temperatura média da bacia. Assim, as bacias hidrográficas que não contam com ao menos quatro estações pluviométricas não podem ter suas vazões regionalizadas a partir do modelo proposto. De modo análogo, as bacias que tem mais de 4 estações pluviométricas, e portanto maior quantidade de registro de precipitação, terão de selecionar apenas as quatro estações pluviométricas mais representativas e descartar os registros restantes. Tal abordagem pode tanto inviabilizar a aplicação da regionalização - para o caso de bacias com menos de quatro estações - assim como demandar a realização de análise discriminante e o conseqüente descarte de registro de dados de precipitação - para o caso de bacias com mais de quatro estações pluviométricas.

Os resultados verificados entre o Modelo Chuva-Vazão I e II demonstram que a utilização da precipitação média mensal de Thiessen como entrada pode melhorar os resultados de predição. Tal medida é vantajosa em relação à abordagem proposta por Minucci (2014), uma vez que nenhum registro histórico de estações pluviométricas das bacias hidrográficas necessitará ser descartado.

Conforme verificado, o desempenho da regionalização através das estações do Agrupamento B foi baixa para ambas as estações teste. A estação Madeireira Gavazzoni teve sua vazão média subestimada em 63%, enquanto a vazão média da estação Guampará foi superestimada em cerca de 184% com relação ao valor observado. Vale destacar que a mudança de *clusters* das estações teste foi significativa, quando verificadas as estações dos Agrupamentos A e B. Nesse sentido, destaca-se que a apresentação de todas as variáveis das bacias hidrográficas para a clusterização via algoritmo SOM foi crucial para mudança dos agrupamentos e o consequente baixo desempenho da regionalização.

As melhores estatísticas médias verificadas no presente estudo, Modelo Chuva-Vazão I - Agrupamento A, foram comparadas com os valores observados e os modelados por Minucci (2014). Conforme se observa nas Tabelas 6.3 e 6.4, as vazões específicas regionalizadas superestimaram a vazão observada, porém com desvios padrão similares.

Tabela 6.3: Comparação entre os resultados obtidos por Minucci (2014) e pelo Autor (2016), com os valores observados - estação Madeireira Gavazzoni.

	Regionalização (Minucci, 2014)	Regionalização (Autor, 2016)	Vazão Observada
Vazão Esp. Média (L/s/km ²)	30,49	36,34	30,34
Desvio Padrão (L/s/km ²)	21,34	24,19	23,36
Nash-Sutcliffe	0,88	0,79	-
R ²	0,88	0,87	-

Tabela 6.4: Resultados obtidos por Minucci (2014) e pelo Autor (2016), com os valores observados - estação Guampará.

	Regionalização (Minucci, 2014)	Regionalização (Autor, 2016)	Vazão Observada
Vazão Esp. Média (L/s/km ²)	28,09	32,60	30,2
Desvio Padrão (L/s/km ²)	22,25	22,21	24,45
Nash-Sutcliffe	0,82	0,81	-
R ²	0,82	0,82	-

7. Conclusão

No geral, com relação à análise de *clusters* realizada neste trabalho, verificou-se que o algoritmo SOM identificou 9 estações fluviométricas, de um total de 16, nos mesmos grupos dos formados por Minucci (2014) — considerando os mesmos dados de entrada; o que representa uma diferença razoável com relação aos resultados da aplicação do método tradicional *k-means*. Contudo, conforme exposto, tendo em vista que as estações teste foram encontradas entre as oito estações que mudaram de *cluster*, se verificaram maiores diferenças no momento da regionalização.

A modelagem chuva-vazão apresentou índice de determinação e Nash-Sutcliffe baixos durante sua calibração, e satisfatórios para a operação, sendo que os dados de precipitação média mensal de Thiessen foram os *inputs* de precipitação que apresentaram melhores desempenhos; o que pode representar uma vantagem na aplicabilidade da abordagem da regionalização reproduzida.

A série de vazões regionalizadas para as duas estações teste apresentaram resultados satisfatórios — em se tratando do Modelo Chuva-Vazão I, Agrupamento A — com índices de determinação, NASH e REQM bons, quando confrontados os registros observados e modelados. Contudo, comparando com as mesmas entradas propostas por Minucci (2014), as simulações resultaram em R^2 e NASH menores.

Para futuros estudos, sugere-se avaliar o desempenho da aplicação de RNAs RBF para a regionalização de cheias, uma vez que os desvios observados no presente estudo para vazões mais altas foram pouco significativos. Além disso, para a análise de agrupamentos, sugere-se analisar maior quantidade de bacias hidrográficas —

incluindo, por exemplo, bacias hidrográficas mais distantes, de modo a possibilitar o emprego da abordagem reproduzida em um número maior de casos.

Referências Bibliográficas

- Bowden, G. J., Dandy, G. C., e Maier, H. R. (2005). Input determination for neural network models in water resources applications: Part 1 - background and methodology. *Journal of Hydrology*, 301:75–92.
- Castro, F. C. C. e Castro, M. C. F. (2000a). Mapas Auto-Organizados de Kohonen - SOM. Relatório técnico, PUCRS, FENG, DEE - Mestrado em Engenharia Elétrica.
- Castro, F. C. C. e Castro, M. C. F. (2000b). Radial Basis Function. Relatório técnico, PUCRS, FENG, DEE - Mestrado em Engenharia Elétrica.
- Diniz, L. S. (2008). *Regionalização de parâmetros de modelo chuva-vazão usando redes neurais*. Doutorado em engenharia, Programa de Pós Graduação em Engenharia de Recursos Hídricos e Saneamento Ambiental, Universidade Federal do Rio Grande do Sul, Porto Alegre.
- DNAEE (1983). Inventário das Estações Fluviométricas. Relatório técnico.
- Dogan, E., Isik, S., Toluk, T., e Sandalci, M. (2014). Daily streamflow forecasting using artificial neural networks. páginas 1–13.
- Fausett, L. (1994). *Fundamentals of Neural Networks - Architectures, Algorithms and Applications*. Prentice Hall, Englewood, NJ, 3 edio.
- Haykin, S. (1999). *Neural Networks and Learning Machines*. PEARSON Prentice Hall, Canadá, 3 edio. Bibliografia: p. 847–938.

- Herrero, J., Valencia, A., e Dopazo, J. (2001). A hierarchical unsupervised growing neural network for clustering gene expression patterns. *Bioinformatics*, 17:126–136.
- Jain, A. K., Murty, M. N., e Flynn, P. J. (1999). Data Clustering: A Review. *ACM Computing Surveys*, 31(3):264–323.
- Kagoda, P. A., Ndiritu, J., Ntuli, C., e Mwaka, B. (2010). Application of radial basis function neural networks to short-term streamflow forecasting. *Physics and Chemistry of the Earth*, (35):571–581.
- Kalteh, A. M., PHjorth, e RBerndtsson (2007). Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environmental Modelling & Software*, 23:835 – 845.
- Kennedy, E. J. (1984). capítulo A10, página 59. U.S. Geological Survey, Washington, Estados Unidos.
- Kiliç, I., Cigizoglu, H. K., e Zuran, A. (2000). A comparison of three methods for the prediction of future stream flow data. *State Hydraulic Works (DSI) Turkey*, páginas 1 – 8.
- Kohonen, T. et al. (1996). SOM_PAK: The Self-Organizing Map Program Package. Relatório técnico, Helsinki University of Technology, Laboratory of Computer and Information Science.
- Lin, G. e Chen, L. (2006). Identification of homogenous regions for regional frequency analysis using the self-organizing map. *Journal of Hydrology*, 32(4):1–9.
- Machado, F. W. (2005). Modelagem chuva-vazão mensal utilizando redes neurais artificiais. Mestrado em engenharia de recursos hídricos e ambiental, Universidade Federal do Paraná, Curitiba.

- Maier, H. R. e Dandy, G. C. (1999). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software*, 15:101–124.
- Minucci, L. R. (2014). Regionalização de série de vazões mensais utilizando redes neurais artificiais. Mestrado em engenharia de recursos hídricos e ambiental, Programa de Pós Graduação em Engenharia de Recursos Hídricos e Ambiental - Universidade Federal do Paraná, Curitiba.
- MME (2009). Portaria nº 463, de 03 de dezembro de 2009, alterada pela Portaria MME nº 376, de 05 de agosto de 2015. Relatório técnico, Brasil. Suplemento.
- MME (2014). Plano Decenal de Expansão de Energia 2023. Relatório técnico, Brasília.
- Riad, S., Mania, J., Bouchau, L., e Najjar, Y. (2003). Rainfall-Runoff Model Using an Artificial Neural Network Approach. *Mathematical and Computer Modelling*, 40:839–846.
- Saraiva, I. S. (2010). Regionalização de parâmetros de modelos de simulação hidrológica para emprego em bacias hidrográficas sem monitoramento fluviométrico. Mestrado em meio ambiente e recursos hídricos, Universidade Federal de Minas Gerais, Belo Horizonte.
- SEMA (2010). Sumário Executivo: Plano Estadual de Recursos Hídricos do Paraná. Relatório técnico, Paraná.
- Serrano, S. E. (1997). *Hydrology for Engineers, Geologists and Environmental Professionals*. HydroScience Inc., Kentucky, 1 edio.
- Tucci, C. E. M. (1998). *Modelos Hidrológicos*. Cambridge, Porto Alegre, 2 edio.

Tucci, C. E. M. (2007). *Regionalização de Vazões*, capítulo 15, páginas 573–611. UFGRS - Universidade Federal do Rio Grande do Sul, ABRH - Associação Brasileira de Recursos Hídricos, Porto Alegre, 4 edio.

Vos, N. J. (2003). *Rainfall-Runoff Modelling Using Artificial Neural Networks*. Tese de Doutorado, Subfaculty of Civil Engineering, Delft University of Technology, Delft, Holanda.

Wilks, D. S. (2006). *Statistical Methods in the Atmospheric Sciences*. Elsevier, Department of Earth and Atmospheric Sciences, Cornell University, 2 edio.