

**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA CIVIL
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA CIVIL
TECNOLOGIA AMBIENTAL E RECURSOS
HÍDRICOS**

Previsão de Vazões Naturais Diárias ao Reservatório de Três
Marias Usando a Técnica de Redes Neurais

MESTRANDA: LUANA FERREIRA COSTA GOMES

ORIENTADORA: SUZANA MARIA GICO LIMA MONTENEGRO

CO-ORIENTADOR: MEUSER JORGE DA SILVA VALENÇA

RECIFE, JUNHO DE 2006

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Previsão de Vazões Naturais Diárias ao Reservatório de Três
Marias Usando a Técnica de Redes Neurais

LUANA FERREIRA COSTA GOMES

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO DA UNIVERSIDADE FEDERAL DE PERNAMBUCO COMO PARTE DOS REQUISITOS NECESSÁRIOS À OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA CIVIL.

G633p **Gomes, Luana Ferreira Costa**

Previsão de vazões naturais diárias ao reservatório de Três Marias usando a técnica de redes neurais / Luana Ferreira Costa Gomes. – Recife: O Autor, 2006.

xii, 122 f., il., figs., tabs.

Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Depto. de Engenharia Civil, 2006.

Inclui referências bibliográficas e anexos.

**1. Engenharia Civil. 2. Previsão de vazões – Três Marias.
3. Redes Neurais. I. Título.**

624 CDD (22.ed.)

BCTG/2006-64

BANCA EXAMINADORA:



Profª Suzana Maria Gico Lima Montenegro (Presidente), D. Sc.



Prof. Carlos de Oliveira Galvão, D. Sc.



Prof. José Almir Cirilo, D. Sc.

RECIFE, JUNHO DE 2006.

DEDICATÓRIA

Dedico este trabalho a meus pais, irmãos e a meu noivo Anderson, pela paciência, confiança e apoio incondicionais.

“O desafio é a força que movimenta os vencedores. É essa força que faz o homem construir barragens, para, entre outros, gerar energia elétrica. E essa energia é uma das fontes de desenvolvimento de toda a população.”

ÍNDICE GERAL

LISTA DE FIGURAS	i
LISTA DE TABELAS	vi
LISTA DE SIGLAS	viii
AGRADECIMENTOS	ix
RESUMO	xi
ABSTRACT	xii
1. Introdução	01
1.1 ASPECTOS GERAIS	01
1.1.1 Previsão de Vazões no Setor Elétrico	02
1.2 RELEVÂNCIA	04
1.3 OBJETIVOS DA PESQUISA	06
1.3.1 Gerais	06
1.3.2 Específicos	06
1.4 DESCRIÇÃO DA DISSERTAÇÃO	07
2. Revisão da Literatura	09
2.1 APLICAÇÃO DE GEOESTATÍSTICA	10
2.2 APLICAÇÃO DE REDES NEURAIS	14
3. Materiais e Métodos	22
3.1 DESCRIÇÃO DA ÁREA EM ESTUDO	22
3.2 MÉTODOS	25
3.2.1 Estatística	25
3.2.1.1 Dendograma	25
3.2.1.2 Análise de Componentes Principais – ACP	26
3.2.2 Geoestatística	26
3.2.3 Redes Neurais	31
3.2.3.1 Redes MLP	33

3.2.3.1.1 Momentum	35
3.2.3.1.2 Algoritmo de Levenberg-Marquardt (LM)	35
3.2.3.1.3 Método de Validação Cruzada	36
3.2.3.1.4 Método de Nguyen-Widrow	37
3.2.3.2 Redes NSRBN	37
3.2.4 Softwares Usados	43
4. Aplicações	44
4.1 ATUAIS ERROS DE PREVISÕES AO RESERVATÓRIO DE TRÊS MARIAS	44
4.2 BASE DE DADOS	45
4.2.1 Informações pluviométricas	46
4.2.1.1 Rede de postos pluviométricos pesquisados	46
4.2.1.2 Preenchimento de falhas e análise de consistência das informações	48
4.2.1.3 Análise geoestatística	49
4.2.2 Informações fluviométricas	49
4.2.2.1 Rede de postos fluviométricos pesquisada	49
4.2.2.2 Análise de correlação e preenchimento de falhas	50
4.2.3 Análise Estatística	53
4.2.3.1 Função de auto-correlação e correlação cruzada	53
4.2.3.2 Matriz de correlação, dendograma (agrupamento), e análise de componentes principais (ACP)	54
4.2.4 Redes Neurais	55
4.2.4.1 Rede MLP (Multi Layer Perceptron)	57
4.2.4.2 Rede NSRBN	60
5. Resultados e Discussões	62
5.1 ANÁLISE DOS DADOS	62
5.1.1 Informações pluviométricas	62
5.1.2 Informações fluviométricas	74

5.2 SIMULAÇÕES COM AS REDES NEURAIAS	94
6. Conclusões e Recomendações	111
7. Referências Bibliográficas	115
8. Anexos	121

LISTA DE FIGURAS

Figura 1.1 – Disposição dos aproveitamentos localizados na Bacia do São Francisco	05
Figura 1.2 – Fluxograma do trabalho	07
Figura 3.1 – Localização da sub-bacia 40 no estado de Minas Gérias	22
Figura 3.2 - Forma característica de um semi-variograma	29
Figura 3.3 - Modelos matemáticos aplicados a variogramas	30
Figura 3.4 – Estrutura da rede MLP	33
Figura 3.5 – Arquitetura combinatorial	38
Figura 3.6 – Arquitetura de um bloco polinomial	40
Figura 4.1 – Esquema com os dias previstos no PMO e suas revisões	44
Figura 4.2 – Postos pluviométricos da selecionados	48
Figura 4.3 – Disposição dos postos fluviométricos em operação da CEMIG localizados na área de drenagem de Três Marias	50
Figura 4.4 – Sub-bacias da área de drenagem de Três Marias	50
Figura 4.5 – Polígono de Thiessen dos postos pluviométricos selecionados da ANA e CEMIG	53
Figura 4.6 – Exemplos dos EPMA obtidos com 5 diferentes rodadas – período úmido	59
Figura 4.7 – Exemplos dos EPMA obtidos com 5 diferentes rodadas – período seco	60
Figura 5.1 – Precipitação média anual – validação cruzada, semi-variograma esférico	64
Figura 5.2 – Isolinhas da precipitação média anual e do erro obtida na interpolação por Kriging	65
Figura 5.3 – Precipitação média mensal – validação cruzada, semi-variograma exponencial	65
Figura 5.4 – Isolinhas da precipitação média mensal e do erro obtida na interpolação por Kriging	65

Figura 5.5 - Precipitação média do período úmido (nov a abr) – validação cruzada, semi-variograma esférico	66
Figura 5.6 – Isolinhas da precipitação média do período úmido e do erro obtida na interpolação por Kriging	66
Figura 5.7 – Precipitação média do período seco (mai a out) – validação cruzada, semi-variograma esférico	67
Figura 5.8 - Isolinhas da precipitação do período úmido e do erro obtida na interpolação por Kriging	67
Figura 5.9 – Disposição dos Postos Pluviométricos operados pela ANA/CPRM, localizados na área de drenagem da UHE Três Marias	68
Figura 5.10 – Disposição dos Postos Pluviométricos operados pela ANA/CPRM e os postos da CEMIG usados nas análises geoestatísticas, localizados na área de drenagem da UHE Três Marias	68
Figura 5.11 – Disposição dos Postos Pluviométricos operados pela ANA/CPRM selecionados e os postos da CEMIG usados nas análises geoestatísticas, localizados na área de drenagem da UHE Três Marias	69
Figura 5.12 – Precipitação média anual – validação cruzada, semi-variograma gaussiano	70
Figura 5.13 – Isolinhas da precipitação média anual e do erro obtida na interpolação por Kriging	71
Figura 5.14 – Precipitação média mensal – validação cruzada, semi-variograma esférico	71
Figura 5.15 – Isolinhas da precipitação média mensal e do erro obtida na interpolação por Kriging	72
Figura 5.16 – Precipitação média do período úmido (nov a abr) – validação cruzada, semi-variograma esférico	72
Figura 5.17 – Isolinhas da precipitação média do período úmido e do erro obtida na interpolação por Kriging	72
Figura 5.18 – Precipitação média do período seco (mai a out) – validação cruzada, semi-variograma esférico	73

Figura 5.19 – Isolinhas da precipitação média do período seco e do erro obtida na interpolação por Kriging	73
Figura 5.20 – Regressão potencial entre os postos fluviométricos de Porto Indaiá e Porto Pará	74
Figura 5.21 – Regressão linear simples entre os postos fluviométricos de Ponte dos Vilelas e Ponte Nova do Paraopeba	75
Figura 5.22 – Regressão potencial entre os postos fluviométricos de Porto das Andorinhas e Iguatama	75
Figura 5.23 – Regressão exponencial entre os postos fluviométricos de Porto Mesquita e Porto Pará	76
Figura 5.24 – Regressão linear simples entre os postos fluviométricos de Porto Mesquita e Ponte Nova do Paraopeba	76
Figura 5.25 – FAC dos postos de Porto Indaiá e Porto Mesquita	77
Figura 5.26 – FAC dos postos de Ponte Nova do Paraopeba e Ponte dos Vilelas Jusante	78
Figura 5.27 – FAC da vazão natural de Três Marias e do posto de Porto das Andorinhas	78
Figura 5.28 – FAC dos postos de Porto Pará e Iguatama	78
Figura 5.29 – FAC da precipitação média calculada por Thiessen	79
Figura 5.30 – Matriz de Correlação de dados semanais de postos da área a montante da UHE Três Marias	81
Figura 5.31 – Matriz de Correlação de dados semanais de postos da área a montante da UHE Três Marias	82
Figura 5.32 – Dendograma com informações diárias representativas da área a montante da UHE de Três Marias	83
Figura 5.33 – Dendograma com informações diárias representativas da área a montante da UHE Três Marias	83
Figura 5.34 – Análise de Componentes Principais com dados semanais representativo da área a montante da UHE Três Marias	85
Figura 5.35 – Análise de Componentes Principais com dados semanais representativo da área a montante da UHE Três Marias	86

Figura 5.36 – Correlação semanal entre a vazão natural de Três Marias com o posto de Porto Indaiá	87
Figura 5.37 – Correlação semanal entre a vazão natural de Três Marias com o posto de Iguatama	87
Figura 5.38 – Correlação semanal entre a vazão natural de Três Marias com o posto de Poto Mesquita	88
Figura 5.39 – Correlação semanal entre a vazão natural de Três Marias com o posto de Porto das Andorinhas	88
Figura 5.40 – Correlação semanal entre a vazão natural de Três Marias com o posto de Ponte dos Vilelas Jusante	89
Figura 5.41 – Correlação semanal entre a vazão natural de Três Marias com a chuva média	89
Figura 5.42 – Correlação semanal entre a vazão natural de Três Marias com a vazão de Ponte Nova do Paraopeba	90
Figura 5.43 – Correlação semanal entre a vazão natural de Três Marias com a vazão de Porto Pará	90
Figura 5.44 - Correlação de dados semanais da vazão natural de Três Marias com a chuva média calculada por Thiessen (com postos CEMIG)	91
Figura 5.45 - Correlação de dados semanais da vazão natural de Três Marias com a chuva média calculada por Kriging	92
Figura 5.46 – Dendograma com informações diárias representativas da área a montante da UHE Três Marias	93
Figura 5.47 – Análise de Componentes Principais com dados semanais representativo da área a montante da UHE Três Marias	93
Figura 5.48 – EPMA das melhores RN do período úmido – dados de validação	94
Figura 5.49 – EPMA das melhores RN do período úmido – dados de teste	95
Figura 5.50 – EPMA dos 12 dias de previsão	97
Figura 5.51 – EPMA dos 12 dias de previsão	97
Figura 5.52 – EPMA das melhores RN do período seco – dados de validação	99
Figura 5.53 – EPMA das melhores RN do período seco – dados de teste	99
Figura 5.54 – EPMA dos 12 dias de previsão – análise por “casos”	101

Figura 5.55 – EPMA dos 12 dias de previsão – análise por “dias”	101
Figura 5.56 – Vazão semanal prevista e observada com a rede MLP	103
Figura 5.57 – Vazão semanal prevista e observada com a rede NSRBN	110

LISTA DE TABELAS

Tabela 4.1 – Erros das previsões de vazões diárias ao reservatório de Três Marias no PMO	44
Tabela 4.2 – Erros das previsões de vazões semanais ao reservatório de Três Marias no PMO – 1ª semana	45
Tabela 4.3 – Postos pluviométricos da CEMIG	47
Tabela 4.4 – Postos Fluviométricos da CEMIG em operação na área a montante do reservatório de Três Marias	52
Tabela 4.5 – Significado das principais variáveis utilizadas no estudo	54
Tabela 4.6 – Média e desvio padrão dos dados semanais usados	55
Tabela 4.7 – Variáveis de entrada consideradas na rede neural MLP do período úmido	58
Tabela 4.8 – Variáveis de entrada consideradas na rede neural MLP do período seco	59
Tabela 4.9 – Variáveis de entrada consideradas na rede neural NSRBN do período úmido	61
Tabela 4.10 – Variáveis de entrada consideradas na rede neural NSRBN do período seco	61
Tabela 5.1 – Coeficiente de correlação entre os postos selecionados na área a montante de Três Marias	63
Tabela 5.2 – Resultados dos semi-variogramas ajustados e da aplicação da interpolação de Kriging (validação cruzada)	64
Tabela 5.3 – Postos Pluviométricos ANA/CPRM selecionados	69
Tabela 5.4 – Resultados dos semi-variogramas ajustados e da aplicação da interpolação de Kriging (validação cruzada)	70
Tabela 5.5 – Correlação entre as informações fluviométricas	74
Tabela 5.6 – Informações obtidas na regressão múltipla	77
Tabela 5.7 – Lags obtidos para FAC máxima de 0,6	79
Tabela 5.8 – Lags obtidos para correlação cruzada máxima de 0,6 entre a vazão natural de Três Marias e a variável indicada	80
Tabela 5.9 – Resultados de validação da rede neural do período úmido	96

Tabela 5.10 – Resultados do teste, com rede correspondente aos dados de validação	96
Tabela 5.11 – Resultados de validação da rede neural do período seco	100
Tabela 5.12 – Resultados do teste, com rede correspondente aos dados de validação	100
Tabela 5.13 – Erros da previsão de vazões diárias com os melhores modelos da Rede MLP	102
Tabela 5.14 – Erros da Previsão de vazões semanais com os melhores modelos da Rede MLP	102
Tabela 5.15 – Resultados das simulações do período úmido para os 7 dias de previsão	103
Tabela 5.16 – Resultados das simulações do período úmido para os 12 dias de previsão	104
Tabela 5.17 – Resultados das simulações do período seco para os 7 dias de previsão	107
Tabela 5.18 – Resultados das simulações do período seco para os 12 dias de previsão	107
Tabela 5.19 – Previsão de vazões diárias com os melhores modelos da Rede NSRBN	109
Tabela 5.20 – Previsão de vazões semanais com os melhores modelos da Rede NSRBN	109

LISTA DE SIGLAS

ANA – Agência Nacional de Águas

CEMIG – Companhia Energética de Minas Gerais

MLP – Multi Layer Perceptron

NSRBN – Non-linear Sigmoidal Regression Blocks Networks

ONS – Operador Nacional do Sistema Elétrico

SIN – Sistema Interligado Nacional

AGRADECIMENTOS

Á Deus, por estar sempre me iluminando e me dando forças para continuar em busca dos meus sonhos e ideais.

Aos meus pais, Nininha e Humberto, por darem sempre prioridade aos meus estudos e de meus irmãos, e por sempre me mostrarem a importância do mesmo.

Aos professores Suzana e Mêuser, por aceitarem este desafio e pelas orientações de extrema importância.

Aos professores deste programa de pós-graduação, pelos ensinamentos.

Aos amigos de mestrado, em especial Emanuelle Matos, Márcio Kuntze, Guilherme Peplau e Tereza Duarte, e aos de Graduação, em especial a Alessandra Maciel, Bianca Malta e Manoela Correia, pela amizade, companheirismo e pelos momentos de muita alegria que passamos juntos.

Ao meu noivo Anderson Paiva pelo companheirismo, paciência e incentivo incondicionais. Também o agradeço pela revisão de texto e por sempre me salvar quando a elaboração das macros do excel e o script do surfer me tiravam a paciência.

Aos amigos do laboratório de Hidráulica Thaisa Alcoforado, Osvalcílio Furtunato, Anderson Lima e João Salgueiro pelas ajudas com geoestatística.

Ao ONS, empresa em que trabalho, pelo apoio, em especial aos amigos do Núcleo Norte-Nordeste e Supervisão de Hidrologia, pelo incentivo e amizade. A Graça Camelo, gerente do Núcleo Norte-Nordeste pelo apoio, incentivo e pelos muitos ensinamentos de sempre.

A Giovanni Coimbra, amigo do ONS, por todo o apoio nesse meu início de vida profissional. Também pela correção de texto e pelas idéias sempre muito úteis aplicadas em parte desse trabalho.

Aos amigos do ONS, Luiz Guilherme Guilhon e Márcio Cataldi. O primeiro pelas conversas técnicas sobre o tema trabalhado e o segundo pela valiosa ajuda com o surfer. Também a amiga e meteorologista Christiane Osório pelo texto sobre o processo de formação da precipitação no Alto São Francisco.

Aos meus sogros Sra Fátima e Sr Paiva pela disponibilidade x
computacional e pelo incentivo.

Á CEMIG, ANA e ANEEL por disponibilizarem as informações
pluviométricas e fluviométricas usadas nesse trabalho.

RESUMO

No planejamento energético do Sistema Interligado Nacional – SIN as metas de geração de todas as usinas hidrelétricas, termelétricas e nucleares são definidas visando sempre à otimização energética, respeitando as restrições hidráulicas, elétricas e a inflexibilidade das usinas térmicas. Um dos principais insumos na definição da geração das usinas hidrelétricas é a previsão de vazões.

Na elaboração dessas previsões, diversos modelos podem ser utilizados. Pode-se citar como exemplo os modelos físicos, estatísticos e os baseados na técnica de redes neurais. O uso da técnica de redes neurais tem se intensificado cada vez mais, uma vez que, modelos baseados nessa técnica são de fácil aplicação e têm proporcionado resultados satisfatórios.

A análise prévia das informações que serão usadas na calibração e utilização da rede neural traz ganhos significativos no desempenho da mesma. Considerando informações pluviométricas e fluviométricas da bacia, existem diversas metodologias para analisar esses dados.

Esse trabalho teve o objetivo de calibrar um modelo de previsão de vazões naturais médias diárias, 12 dias à frente, utilizando a técnica de redes neurais, considerando as vazões naturais e a informação advinda de postos hidrométricos, a montante de Três Marias, no rio São Francisco, fazendo uso da geoestatística como ferramenta para seleção de postos pluviométricos e análise de precipitação média da área.

Na calibração das redes neurais foram utilizadas duas metodologias: rede MLP e rede NSRBN. Os resultados mostraram a importância da análise das informações de entrada da rede neural. Comparando os erros de previsão das duas metodologias, observou-se que a rede MLP produziu um erro de previsão um pouco menor para os dados diários e maior para os dados semanais. Ressalta-se que a técnica de redes neurais, principalmente para os dados diários, produz resultados bem superiores aos modelos estatísticos.

ABSTRACT

In the short term planning for energy supply goals are established encompassing the electric subsystems, the generation of all hydro plants, thermal plants and nuclear plants integrating the SIN (Integrated National System for Energy Supply), always aiming an optimized supply, respecting the hydraulic and electrical restrictions and the inflexibility of the thermal plants. One of the main issues to the hydro plants exploitations is the runoff forecast.

Different types of models can be used to elaborate this forecast,. It can be cited as examples the physical models, the statistical models and models based on the technique of neural network. The use of the technique of neural network has been intensified lately, because models based on this technique are easily applied and yields satisfactory results.

The previous analysis of the information that will be used in the calibration and use of the neural network can bring significant positive effects in tits performance. Many different methodologies are available for the analyses of flow and rainfall data,

This work had the objective of calibrating a model based on the neural network technique, for forecasting daily average natural outflows, 12 days ahead, considering natural outflows and flow and rainfall data, for the Três Marias reservoir, in São Francisco River, using Geostatistics as a tool for selecting rainfall stations and data and estimating the average precipitation of the area.

Two methods were used for calibration of the neural network: network MLP and NSRBN. The results have shown the importance of the input data for the neural network models. Comparing these two methods, observed that network MLP had a slightly better performance for the daily data and worst performance to the weekly data. It has also been observed that the technique of neural network, mainly for the daily data, yields better results when compared to statistical models.

Capítulo 1

Introdução

1.1 ASPECTOS GERAIS

Nas últimas décadas as diferentes tecnologias de modelagem vêm evoluindo rapidamente, propiciando ao homem a oportunidade de ampliar seu conhecimento sobre a natureza e compreender cada vez mais seus processos físicos e biológicos.

Na tentativa de representar o comportamento dos fenômenos naturais, os modelos hidrológicos utilizam como um dos insumos fundamentais para sua calibração e aferição, as vazões médias diárias, semanais ou mensais, dependendo do horizonte de estudo e da aplicação desejada.

Na hidrologia trabalha-se com diversas classificações para os modelos hidrológicos. Uma dessas principais classificações é feita na diferenciação dos modelos determinísticos e estocásticos. Segundo Chow (1964), os modelos estocásticos (ou estatísticos) são aqueles em que as variáveis associadas estão associadas a uma probabilidade de ocorrência, enquanto os modelos determinísticos são aqueles que seguem uma lei definida, que não é a da probabilidade.

Modelos hidrológicos denominados chuva-vazão são aplicados com o objetivo de, a partir de uma precipitação conhecida, simular a resposta da bacia em termos de vazão para uma determinada seção fluvial de interesse. Esses modelos são de extrema importância no planejamento e gestão de recursos hídricos, incluindo o dimensionamento e operação de estruturas, pois, na maioria dos casos, o tamanho das séries históricas de vazão é relativamente pequeno, enquanto as séries históricas de medição de precipitação são de tamanho consideravelmente maior. Estes modelos possuem diversas formulações matemáticas, todas bastante semelhantes no sentido de tentar modelar os diferentes componentes do ciclo hidrológico e, em alguns casos, a variabilidade espacial de seus possíveis inter-relacionamentos, considerando parâmetros que possuem forte significado físico. Os parâmetros desses modelos devem ser calibrados primeiramente com uma parte do histórico de chuva na bacia e de vazões passadas. Em seguida, de posse dos parâmetros já calibrados, procede-

se à verificação da previsão de vazões para outro período do histórico, mais recente, avaliando-se os erros obtidos ao se comparar os valores previstos com os ocorridos.

Contudo, a implementação e a calibração de tais modelos apresentam normalmente várias dificuldades, requerendo sofisticadas ferramentas matemáticas, quantidades significativas de parâmetros para calibração, além de vasta experiência com o modelo.

Por esse motivo, o uso da técnica de redes neurais tem se intensificado cada vez mais, uma vez que, modelos baseados nessa técnica são de fácil aplicação e têm proporcionado resultados satisfatórios.

As redes neurais artificiais (RNAs) são uma técnica integrante dos Sistemas Inteligentes, inspirados no comportamento humano ou da natureza. As RNAs vêm sendo sugeridas para análise de séries temporais devido a sua capacidade de tratar com relações de entrada-saída de dados não-lineares, destacando sua habilidade de aprendizado e capacidade de generalização, associação e busca paralela. Estas qualidades as tornam capazes de identificar e assimilar as características mais marcantes das séries, tais como sazonalidade, periodicidade, tendência, entre outras, muitas vezes camufladas por ruídos.

Contudo, a análise prévia das informações que serão usadas na calibração e utilização da rede neural pode trazer ganhos significativos no desempenho da mesma. Considerando informações pluviométricas e fluviométricas da bacia selecionada, existem diversas metodologias para analisar esses dados. Destaca-se nesse trabalho a utilização da geoestatística como uma ferramenta para seleção de postos pluviométricos e análise da precipitação da área.

1.1.1 Previsão de Vazões no Setor Elétrico

No âmbito do Setor Elétrico, a obtenção de previsões de vazões permite que os diferentes modelos de operação, otimização e simulação energética possam avaliar as condições operacionais das usinas hidrelétricas e termelétricas no futuro, propiciando subsídios para a tomada de decisões em relação à melhor forma de operar o Sistema Interligado Nacional - SIN, explorando toda sua capacidade hidrológica e provocando ganhos sinérgicos. Estas previsões de vazões podem permitir também uma tomada de decisão no sentido de minimizar

os efeitos de uma cheia excepcional, desde que feita com relativa precisão e antecedência.

As previsões de vazões interferem nas atividades de outros usuários da água como a irrigação, a pesca, o turismo, o lazer e a navegação, visto que as defluências dos principais reservatórios do Brasil são definidas com base nestas previsões.

Conforme o uso que se faça das previsões de vazões, trabalha-se com diferentes horizontes e intervalos de discretização. No setor elétrico, trabalha-se com três horizontes de previsões de vazões, a saber: previsões de longo prazo, previsões de médio prazo e previsões de curto prazo.

A previsão de vazões de longo prazo, abrangendo um horizonte de até um ano à frente, é feita a partir de um histórico de vazões naturais médias mensais, geralmente, utilizando o modelo estocástico PREVIVAZM (MACEIRA et al, 2003), que é um modelo estocástico univariado de previsão de vazões em base mensal, para um horizonte de até 12 (doze) meses à frente. O modelo, se baseia em 80 (oitenta) combinações de estrutura de correlação periódica ou estacionária, tipos de transformação.

Para as avaliações de longo prazo utiliza-se também o GEVAZP (CEPEL, 2001), que é um modelo estocástico multivariado de geração de séries sintéticas de vazões incrementais e totais afluentes aos aproveitamentos hidroelétricos, que se baseia em modelos estocásticos convencionais estacionários ou periódicos e utiliza informações de até 11 (onze) meses anteriores. Esse modelo gera cenários, para os quais se estabelecem probabilidades de ocorrência, incluindo, dessa forma, a incerteza dessas vazões. Esses resultados são utilizados, por exemplo, para subsidiar as tomadas de decisão relativas ao planejamento da operação para o horizonte anual, sinalizando o estado do armazenamento dos reservatórios do SIN e o risco de déficit de energia para cada subsistema.

No horizonte de previsão de médio prazo, a partir de um histórico de vazões naturais médias semanais e utilizando o modelo estocástico PREVIVAZ (CEPEL 1997), são realizadas estimativas de vazões naturais médias semanais. Estas previsões têm por objetivo subsidiar o Programa Mensal de Operação - PMO, realizado pelo Operador Nacional do Sistema Elétrico - ONS e suas revisões semanais, realizadas normalmente às quartas-feiras. O PREVIVAZ constitui-se

num modelo estocástico univariado de previsão de vazões em base semanal, para um horizonte de até 6 (seis) semanas à frente. O modelo, que se baseia em 94 (noventa e quatro) combinações de estrutura de correlação, tipos de transformação e métodos de estimação de parâmetros, utiliza informações de até 4 (quatro) semanas anteriores.

Quanto aos modelos de previsão de vazões para o curto prazo, isto é, para duas semanas à frente em base diária, há no Brasil inúmeras metodologias em uso. Algumas dessas metodologias já vêm sendo aplicadas pelas empresas do setor elétrico e outras estão ainda em fase de teste para algumas bacias hidrográficas. Primeiramente, podem-se citar os modelos físicos ou determinísticos baseados na relação entre chuva e vazão, conhecidos também como modelos chuva-deflúvio, citados anteriormente. O ONS calibra anualmente o modelo estocástico de previsão de vazões naturais médias diárias PREVIVAZH (Carvalho, 2001), que é um modelo de previsão de vazões em base diária, para um horizonte de até 14 (quatorze) dias à frente, baseado na desagregação de previsão semanal a partir da tendência inferida das últimas vazões passadas e de séries sintéticas diárias de vazões naturais de forma ponderada.

1.2 RELEVÂNCIA DA PESQUISA

Nos processos de planejamento e programação da operação do SIN, na etapa de planejamento de curto prazo, um dos principais produtos é o PMO. Este documento se constitui numa referência técnica mensal, com revisão semanal, para as metas de geração de todas as usinas hidrelétricas, termelétricas e nucleares integrantes do SIN, visando sempre à otimização energética, respeitando as restrições hidráulicas, elétricas e a inflexibilidade das usinas térmicas. Além disso, são estabelecidas também as metas de intercâmbio entre os subsistemas elétricos e outras diretrizes operativas em etapas semanais para o horizonte mensal.

Para a elaboração do PMO e suas revisões, um dos principais insumos é a previsão de vazões semanais para todos os locais de aproveitamentos hidrelétricos, da primeira à última semana operativa de cada mês. A cada revisão, as previsões de vazões para todas as semanas remanescentes são sempre corrigidas.

Para o subsistema Nordeste são usadas as previsões de vazões naturais dos seguintes aproveitamentos: Sobradinho, Itaparica, Complexo Paulo Afonso, Xingó, Boa Esperança, Pedra do Cavalo e Itapebí.

Nos processos de acompanhamento e previsão de vazões elaborados pelo ONS, para os aproveitamentos da bacia do São Francisco, indicados na Figura 1.1, a vazão natural ao reservatório de Sobradinho é formada pela vazão natural de Três Marias e Queimado (propagadas até o reservatório de Sobradinho) adicionada à vazão incremental do trecho Três Marias/Queimado – Sobradinho. As vazões naturais dos aproveitamentos de Itaparica, Complexo Paulo Afonso e Xingó são calculadas levando em consideração a vazão natural de Sobradinho, nestes reservatórios, adicionando-se as vazões incrementais em cada trecho.



Figura 1.1 – Disposição dos aproveitamentos localizados na Bacia do São Francisco.

A previsão de vazões afluentes e a operação do aproveitamento de Três Marias, localizado na cabeceira do rio São Francisco, são de extrema relevância no processo de operação dos demais reservatórios a jusante, uma vez que interfere totalmente nas definições energéticas para o Subsistema Nordeste. As previsões de vazões naturais a Sobradinho são afetadas pela qualidade das previsões realizadas para Três Marias e Queimado e pela previsão da vazão incremental do trecho Três Marias/Queimado – Sobradinho.

Como atualmente tem-se uma boa previsão para a vazão incremental do trecho Três Marias/Queimado a Sobradinho, com erros da ordem de 4% e 12% (ONS, 2004), para a primeira e segunda semanas, respectivamente, é importante que sejam envidados esforços, principalmente, na pesquisa de novos modelos de previsão para o reservatório de Três Marias, devido a sua maior representatividade, em relação ao reservatório de Queimado, contribuindo para a redução dos desvios das previsões de vazões no processo de elaboração do PMO, incluindo suas revisões semanais, para os aproveitamentos da bacia do rio São Francisco.

1.3 OBJETIVOS DA PESQUISA

1.3.1 Geral

Este trabalho apresenta a calibração de um modelo de previsão de vazões naturais médias diárias, 12 dias à frente, utilizando a técnica de redes neurais, considerando a informação advinda de diferentes postos pluviométricos e fluviométricos, a montante do aproveitamento hidrelétrico de Três Marias, bem como a informação de vazões naturais passadas desse reservatório.

1.3.2 Específicos

Este trabalho tem os seguintes objetivos específicos:

- Selecionar e analisar as informações dos postos hidrométricos (fluviométricos e pluviométricos) da área a montante do reservatório de Três Marias;
- Selecionar a melhor metodologia de redes neurais a ser usada na previsão de vazões;
- Definir qual a melhor topologia de rede neural para a previsão de vazões naturais afluentes ao reservatório de Três Marias.

A figura 1.2, abaixo, mostra o fluxograma do trabalho.

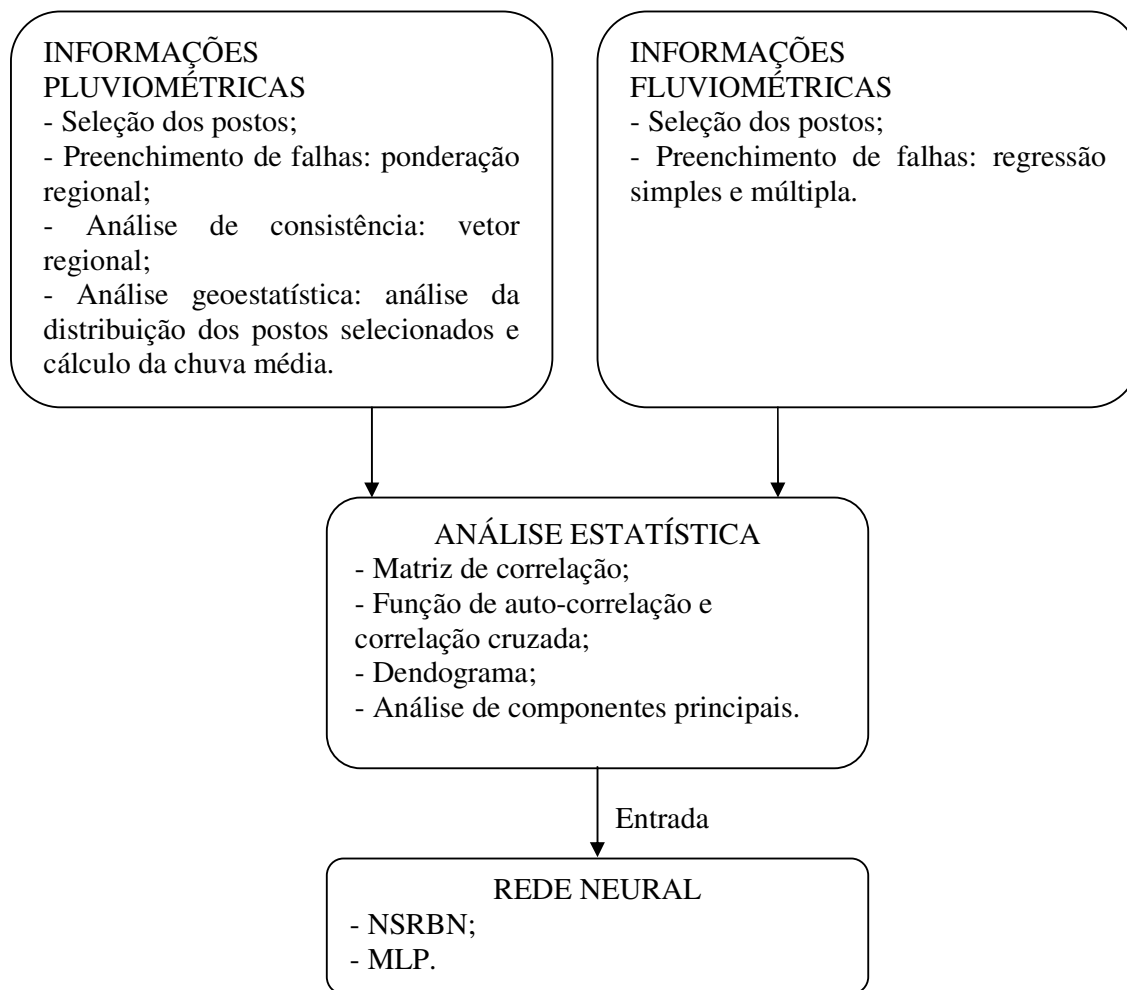


Figura 1.2 – Fluxograma do trabalho.

1.4 DESCRIÇÃO DA DISSERTAÇÃO

Os capítulos que se seguem estão organizados da seguinte forma:

No capítulo 2 é feita uma revisão bibliográfica sobre a aplicação de geoestatística e sobre a aplicação da técnica de redes neurais. São apresentados alguns estudos realizados para bacias hidrográficas brasileiras, no caso da aplicação de geoestatística; e de aplicações na área de recursos hídricos, no caso da técnica de redes neurais.

No capítulo 3 é mostrada a área em estudo, as metodologias utilizadas.

No capítulo 4 são mostradas as aplicações:

- Postos pluviométricos usados, período disponível dos seus históricos, preenchimento de falhas;
- Postos fluviométricos usados, período disponível dos seus históricos, preenchimento de falhas;
- Variáveis e parâmetros usados nas simulações com a técnica de redes neurais.

No capítulo 5 são apresentados os resultados das análises realizadas nos dados básicos e das simulações com a técnica de redes neurais:

- Postos pluviométricos: correlação entre os postos, resultados do preenchimento de falhas, resultados obtidos na análise de consistência e análise Geoestatística efetuadas;
- Postos fluviométricos: correlação entre os postos, resultados do preenchimento de falhas e os resultados obtidos;
- Análises estatísticas efetuadas: função de auto-correlação, função de correlação cruzada, matriz de correlação, correlação entre a vazão natural a Três Marias e as demais informações, análise de componentes principais (ACP) e dendograma.
- Resultados obtidos das simulações com a técnica de redes neurais e comparação com os atuais erros do PMO.

No capítulo 6 são apresentadas as conclusões e são feitas recomendações para futuros trabalhos.

No capítulo 7, são listadas as referências bibliográficas utilizadas nos 6 capítulos anteriores.

Capítulo 2

Revisão da Literatura

A importância da obtenção de informações precisas de precipitação para o desenvolvimento de pesquisas hidrológicas merece ser destacada. Muitos métodos são usados nas análises de precipitação. Em relação à avaliação espacial da precipitação podem-se citar alguns métodos: inverso da distância, Thiessen, Kriging.

O conhecimento do comportamento da variabilidade temporal e espacial da precipitação tem sido objeto de estudo de diversos pesquisadores. Nesse sentido, muitos trabalhos têm sido apresentados usando a técnica da geoestatística que também é usada em aplicações na área de geologia e recursos hídricos subterrâneos. A geoestatística é uma ferramenta para avaliar a distribuição temporal e espacial da precipitação, baseada na técnica de Kriging (Krige, 1951), a qual utiliza o modelo probabilístico associado à localização espacial dos pontos amostrados, possibilitando, desta forma, uma análise da variabilidade entre os valores reais observados. Muitos autores têm mostrado que a utilização da técnica da Krigagem para o processo de interpolação produz resultados mais reais que quase todos os outros métodos de interpolação. Vale salientar que, ao contrário das demais metodologias, nesse método o erro de interpolação pode ser aferido, e a interpolação leva em consideração a dependência espacial da variável em análise na área, ou seja, sua estrutura de correlação. Um dos mais poderosos instrumentos usados para inferir o comportamento espacial da variável a ser explorada é o variograma.

A utilização da ferramenta geoestatística na análise da distribuição espacial de precipitações foi proposta inicialmente por Delhome (1978) e desde então tem sido largamente aplicada na hidrologia espacial (Holawe e Dutter, 1999; Lou et al., 2003; Rodrigues e Silans, 2002).

O comportamento das vazões dos rios tem sido bastante estudado por vários autores. Uma das áreas de conhecimento que tem sido cada vez mais explorado na última década é a de inteligência artificial, onde a técnica de redes neurais está incluída.

A rede neural artificial (RNA) imita os neurônios humanos, recebendo e enviando sinais. Ela pode aceitar uma série de dados de entrada e produzir, a partir destes, uma ou mais saídas, através da busca de uma função que descreva o sistema. Esta analogia feita com os neurônios humanos é baseada na transmissão de sinais elétricos. De maneira simplificada, o neurônio humano pode ser entendido como um corpo celular com duas ramificações: dendritos e axônios. A transferência de sinais elétricos é feita através do axônio de um neurônio para os dendritos de um outro neurônio vizinho, em um processo chamado sinapse. A forma com que estas sinapses interconectam os neurônios define a topologia das redes neurais.

Assim, visando conhecer um pouco mais sobre a aplicação da geoestatística e da técnica de redes neurais, nessa ordem, a seguir são descritos alguns trabalhos.

2.1 APLICAÇÃO DA GEOESTATÍSTICA

Holawea e Dutterb (1999) analisaram as informações de chuva diária, através da Geoestatística, de mais de 400 estações, na Áustria, as quais tinham um período de mais de 20 anos de dados, focando período úmido e seco sem levar em consideração informações das estações do ano. Mostraram que o variograma é um dos mais poderosos instrumentos para medir o comportamento temporal e espacial da precipitação. Os autores analisaram como essas variáveis podem revelar interessantes características da estrutura espacial e temporal da precipitação na Áustria, o qual mostrou, por exemplo, características de agrupamentos climatológicos. Uma das mais importantes perspectivas obtidas é a possibilidade da estimação das variáveis do variograma para lugares onde não tem postos pluviométricos. Para todas as análises foi usado o software GEOSAN (Dutter, 1996). Nas duas direções analisadas foi encontrados, aproximadamente, os mesmos parâmetros do variograma.

Pinto e Naghettini (1999) trabalharam com a bacia do Alto São Francisco, especificamente sub-bacia 40, e apresentaram duas regiões homogêneas, a região Norte e a região Sul. Mostraram que a mesma distribuição com parâmetros de forma diferentes, pode estar indicando que a distribuição espacial do nível de

umidade e a intensificação dos processos associados às precipitações máximas diárias também diferem nas duas regiões. Assim, na região Sul seriam processos frontais e/ou convectivos intensificados por efeitos orográficos, e na região Norte uma menor intensificação pela ação do relevo.

Gomes e Silans (2000) calcularam a precipitação média da bacia do Gramame-PB através dos seguintes métodos: polígonos de Thiessen; interpolação pelo inverso da distância ao quadrado; Shepard, e Krigagem. Observou-se que o método geoestatístico forneceu sempre resultados quase idênticos ao método da interpolação pelo inverso da distância ao quadrado e difere relativamente pouco do método de Thiessen. Já o método de Shepard apresentou resultados diferentes dos demais. O trabalho apresentado mostrou que, para uma rede de pluviômetros bastante densa, o método de Thiessen para calcular a precipitação média numa bacia hidrográfica, gera valores excelentes mesmo quando existir uma variabilidade espacial importante.

Lima e Lima (2000) realizaram a estimativa da precipitação média no mês de abril para duas regiões, bacia do rio Pacoti no estado do Ceará, e bacia do rio dos Sinos, no estado do Rio Grande do Sul através da média aritmética, médias ponderadas pelo inverso da distância, kriging e superfícies de tendências. Os resultados obtidos mostraram que o método de kriging apresentou resultados menos precisos do que os métodos mais simples de média aritmética e de médias ponderadas com o inverso da distância, devido à imprecisão na representação da variabilidade espacial do processo de precipitação, em razão do número reduzido de pontos de observação e das características do processo de precipitação que apresenta baixa dependência espacial com pontos próximos se comparado, por exemplo, com dados de topografia. Os autores concluíram que de nada adianta aumentar a complexidade de processos matemáticos de interpolação se características físicas do processo em estudo não forem incluídas e se não houver um perfeito entendimento da variabilidade espacial do processo.

Ferreira et al. (2000) calcularam a pluviometria média anual, através da metodologia de kriging, em três regiões do Estado do Ceará: Serra da Ibiapaba, litoral e interior. Foram ajustados diferentes semi-variogramas para as regiões da

Serra da Ibiapaba e do interior. A metodologia de kriging, quando da existência de pequena quantidade de postos pluviométricos, mostrou-se mais eficiente no cálculo da precipitação média anual sobre uma região que o método de Thiessen.

Rodrigues e Silans (2002) fizeram uma análise geoestatística no período das Ondas de Leste (de Abril a Agosto) no estado da Paraíba. Foi possível representar a variabilidade espacial residual através de um semi-variograma climático.

Rodrigues e Silans (2003) apresentaram uma metodologia para otimização da rede de monitoramento pluviométrico, baseada na geoestatística. Utilizou-se o método da Krigagem para procurar minimizar a variância de estimação da precipitação média sobre uma área. O trabalho mostrou a existência de um semi-variograma climático para a região em estudo e ser indispensável separar os eventos chuvosos em função do movimento das massas de ar que lhes deram origem. Neste caso, após filtrar um componente determinístico que considera a distância ao mar e a altitude, os autores obtiveram, para os eventos provocados pelas ondas de leste, um semi-variograma climático e mostraram que é possível minimizar a variância de estimação da precipitação média sobre uma determinada área. Para isto, foi mostrada uma técnica de otimização baseada na simulação pelo processo de recozimento "Annealing Simulated", para otimizar a rede de monitoramento pluviométrico da região em estudo, reduzindo a variância de estimação através do aumento do número de postos pluviométricos e determinação de uma localização ótima para os mesmos.

Amaral et al. (2004) trabalharam na avaliação de precipitações máximas anuais (de 1, 3, 5 e 10 dias) nas sub-bacias 40, 41 do Alto São Francisco e 56 do rio Doce, para a obtenção de suas médias máximas espaciais através de métodos geoestatísticos. Na região proposta de estudo encontrou-se maior homogeneidade nos períodos de precipitações máximas para a região das bacias do rio Paraopeba, alto rio das Velhas e alto rio Doce, ou seja, a parte central da região de estudo. As mesmas características de comportamento foram encontradas para os eventos de todas as durações de um mesmo ano

hidrológico, sendo pequenas as diferenças entre os alcances. Contudo, quatro anos apresentaram comportamento diferente dos demais e se ajustaram melhor ao modelo linear. Os meses que se destacaram quanto a maior frequência de ocorrência de eventos máximos de precipitação foram dezembro e janeiro, como era esperado para o estado de Minas Gerais.

Almeida et al. (2004) apresentaram uma análise da variabilidade espacial da precipitação média anual e da precipitação média no trimestre úmido na bacia do Rio Ipanema, aplicando a metodologia da geoestatística, e também avaliaram o processo de formação das chuvas através da distribuição espacial de coeficientes de irregularidade. As chuvas apresentaram baixa altura pluviométrica na região central da bacia apresentando maior auto-correlação na direção nordeste sudoeste. A variável pluviométrica foi mais bem representada por uma distribuição normal, tanto para a amostra de valores mensais, quanto para totais anuais de precipitação. Após ter sido verificada uma dependência espacial nos cenários, foram utilizados os modelos de variograma definidos para a construção de mapas de isolinhas de precipitação utilizando a interpolação kriging. Percebeu-se que 43% das estações representam um comportamento de chuvas convectivas e outros 57% para chuvas frontais. Foi mostrado que o uso da aplicação da metodologia geoestatística tem se apresentado uma boa ferramenta nos estudos de variabilidade espacial de parâmetros hidrológicos e que o uso de coeficientes de irregularidade tem se apresentado coerente quando necessária a pesquisa da formação das precipitações e suas caracterizações.

Igreja et al. (2005) apresentaram uma metodologia para dimensionamento de redes pluviométricas considerando a precisão das estimativas e o custo de operação das estações. A geoestatística foi utilizada por apresentar a variância de estimação da krigagem como um parâmetro eficaz na indicação da precisão de redes pluviométricas. A rede estudada foi a da bacia do Rio São Mateus, localizada no norte do estado do Espírito Santo, com 18 postos de monitoramento. Foi mostrado que, a partir de determinado número de estações, com o acréscimo de postos de monitoramento, o ganho de precisão não é muito significativo. Uma função objetivo, que considera tanto a variância de estimação

(precisão) quanto o custo de operação, foi minimizada para que se encontrasse a seleção das melhores estações dentre as existentes. A rede estudada pôde ter o número de estações reduzido pela metade, fazendo com que os custos envolvidos fossem consideravelmente menores sem que houvesse perda significativa na precisão da estimativa da precipitação média anual sobre uma área, otimizando a relação custo x precisão.

Assim, cada vez mais, muitos trabalhos têm mostrado, com êxito, a aplicação da Geoestatística para aferir a variabilidade temporal e espacial da precipitação. Essas análises são feitas com diferentes aplicativos e períodos de dados. Destaca-se em especial, o potencial da geoestatística para estimativa da precipitação média espacial em uma área.

2.2 APLICAÇÃO DE REDES NEURAIS

Shamseldin (1997) trabalharam na modelagem dos processos chuva x vazão, com a técnica de redes neurais. O autor usou diferentes informações como entrada da rede neural: chuva, histórico sazonal e informações de postos próximos. Os resultados foram comparados com outros três modelos: modelo simples linear (SLM), modelo sazonal baseado na perturbação linear (LMP) e o modelo linear de perturbação (NNLPM) e mostraram que as redes neurais são promissoras no contexto da modelagem dos processos chuva x vazão, mas, como os demais modelos analisados, têm resultados variáveis.

Valença (1997) fez uma análise da aplicação de redes neurais perceptron de múltiplas camadas na área de recursos hídricos, em especial na simulação do processo de transformação de chuva em vazão e na previsão de vazões médias mensais. Os resultados obtidos no processo chuva vazão foram comparados com os modelos conceituais MOHTSAR e SMAP (Lopes, 1981), enquanto os de previsão de vazões foram comparados com aqueles já encontrados com a utilização de técnicas de regressão múltipla e metodologia de Box e Jenkins. Segundo o autor, os resultados encontrados para os dois estudos com o uso de redes neurais foram de ótima qualidade quando comparados com aqueles obtidos com as metodologias tradicionais. Mostrou-se que a grande vantagem da técnica

de redes neurais está na sua versatilidade de permitir que o processo chuva vazão incorpore a representatividade de cada posto pluviométrico em vez de se trabalhar com chuva média por Thiessen (por exemplo).

Sajikumar (1999) calibrou um modelo mensal chuva x vazão. O autor obteve bons e importantes resultados usando relativamente poucas informações, o que mostra a importância do estudo para lugares, por exemplo, como a Índia, que dispõe de poucos dados. Foram usados dados do rio Lee (na Inglaterra) e Thuthapuzha (na Índia). Os resultados mostraram a eficiência das redes neurais para modelagem do processo chuva x vazão para um período de informações inferiores a 6 anos.

Sperb et al. (1999) tiveram por objetivo demonstrar a aplicabilidade de redes neurais artificiais na previsão de enchentes para a Bacia do Rio Itajaí-Açu até a cidade de Blumenau, através da proposição de um modelo de transformação chuva em vazão para previsão de vazões em tempo real. O modelo utilizou dados de precipitação de estações pluviométricas distribuídas na bacia de contribuição para previsão da descarga no ponto de interesse. Devido à não linearidade entre as variáveis, o algoritmo de treinamento escolhido foi o de "Levenberg-Marquardt". Apesar de uma modelagem bastante simples, esta aplicação prática resultou em 93% de acerto nas previsões, o que indica serem as redes neurais uma alternativa promissora para tais previsões.

Coulibaly et al. (2000) mostraram um critério de parada (early stopped - STA) que foi introduzido para treinar as redes neurais feed-forward multi-layer (FNN) para a previsão de vazão em tempo real a um reservatório. O FNN foi treinado usando o critério de parada com o método de otimização Levenberg-Marquardt Backpropagation (LMBP). O método proposto leva vantagem nos métodos de LMBP e de validação cruzada por evitar sub e super-treinamento nas redes FNN e melhorar o desempenho da generalização. A metodologia é avaliada usando a série hidrológica de Chute-du-Diable, hidrelétrica no norte de Quebec (Canadá), e o desempenho do modelo é comparado com um modelo estatístico (ARMAX-KF) e um modelo conceitual em operação na hidrelétrica (PREVIS).

Analisando os resultados obtidos, o método proposto pareceu ser uma ferramenta eficaz para a previsão de vazão ao reservatório em tempo real. Na experiência considerada, o uso do STA reduziu em quatro vezes o tempo de treinamento da rede (mais rápida convergência) e os resultados de teste obtidos usando o STA-LMBP indicam que o método pode fornecer o desempenho melhor e de confiança da generalização do que o método LMBP sozinho.

Kyrillos (2000) mostrou o potencial da utilização dos sistemas integrados de análise hidrológica nos processos de planejamento e tomada de decisão no gerenciamento dos recursos hídricos, através de um Sistema de Suporte a Decisão – SSD. Dentre outras aplicações, foi usado um modelo de redes neurais (Qnet) para previsão de vazões em tempo real, para o posto de Morpará, na bacia do Rio São Francisco. Foi utilizado um algoritmo backpropagation, com aprendizado supervisionado e função de ativação sigmóide. Foram utilizados dados diários de vazão dos postos de São Francisco, Carinhanha, Bom Jesus da Lapa e Morpará, além de dados de chuva de 6 postos do rio das velhas, através da chuva média calculada por Thiessen. De todas as arquiteturas analisadas, levaram a melhores resultados aquelas que se destinavam a simular para um horizonte mais próximo e aquelas que agregavam mais neurônios de entrada. Segundo a autora, as arquiteturas que relacionavam apenas vazões apresentaram ajustes excelentes. A introdução da chuva como entrada da rede neural, trouxe mais ruído à simulação, mas mesmo assim os resultados foram razoáveis.

Valença e Ludermir (2001) avaliaram a viabilidade do uso de Redes Neurais na previsão de vazões semanais. Os autores ajustaram os 52 melhores modelos do tipo auto-regressivo média-móvel periódico – PARMA(p,q) e compararam os resultados com os obtidos para outros modelos de Redes Neurais – Redes *Perceptron* de Múltiplas Camadas, Redes Recorrentes de Elman e Jordan, Redes GMDH e as Redes NSRBN. A estrutura experimental para o treinamento e teste dos modelos envolveu a série de vazões médias semanais incrementais afluentes ao reservatório de Sobradinho, no período de 1929 a 1988. Nos vários testes realizados foram basicamente testados: a sensibilidade aos

pesos iniciais e os critérios de seleção. Efetuou-se também uma análise da sensibilidade quanto à mudança na taxa de aprendizagem e no termo momento, verificando-se a variação da taxa de momento entre 0,1 e 0,001 com o termo momento fixo 0,1 onde se concluiu que não foi obtido nenhum ganho de precisão, mas alterava-se significativamente o tempo de aprendizagem. Ou seja, um valor para a taxa de aprendizagem pequeno 0,001 duplicava o tempo de treinamento. Em termos de erro percentual médio absoluto, as redes com melhores desempenhos foram, nessa ordem: NSRBN, GMDH, MLP e Jordan, Elman, Box-Jenkins.

Olivo et al. (2002) mostraram os resultados obtidos pela aplicação de redes neurais artificiais ao problema de previsão de cheias fluviais. A estrutura de rede escolhida foi do tipo *multilayer perceptron* e o sistema foi construído na forma de uma “caixa-preta”, em que o nível do rio em uma seção de interesse é explicado através da observação do nível do mesmo rio em uma seção a montante. O estudo de caso foi realizado com 13 eventos de cheias observadas no rio Itajaí (SC). Os resultados mostraram a potencialidade do uso desta ferramenta para resolver problemas hidrológicos.

Lima e Ferreira (2003) analisaram a performance de diversos modelos de Redes Neurais Artificiais (RNA) aplicados ao processo chuva-deflúvio. A série histórica analisada conteve dados da precipitação mensal na bacia e da vazão mensal no exultório, compreendidos entre o período de Janeiro de 1963 e Dezembro de 2000. As diversas configurações de RNA's foram obtidas com a variação do número de nós na camada de entrada, na camada intermediária e pela variação do número de anos usados para o treinamento e para a validação da rede. Os testes dos modelos foram realizados com dados de precipitação e deflúvio no posto da Fazenda Cajazeiras, no Estado do Ceará. O deflúvio mensal foi previsto a partir de dados de precipitação dos meses anteriores e corrente e em alguns modelos utilizando também o deflúvio de meses anteriores. Os resultados mostraram a grande influência que tem a configuração da rede neural e o número de anos de treinamento e de validação, sendo que o período ideal de treinamento encontrado está basicamente na metade da série histórica, sendo a

outra parte da série destinada à validação. Observou-se que o número de nós na camada intermediária teve pouca influência nos resultados, sendo o número ideal a ser utilizado determinado somente pelos resultados das simulações numéricas, não existindo, a priori, um método que determine esse número ideal. Contudo, observou-se que a característica dos nós de entrada influencia significativamente os resultados. As melhores estatísticas na previsão do deflúvio no mês t foram obtidas com dados na entrada de precipitação, correspondentes aos períodos t , $t-1$ e $t-2$, em conjunto com o deflúvio no mês $t-1$. Entretanto, o volume acumulado no período foi mais bem previsto com os mesmos dados de precipitação, excluindo-se o deflúvio no mês $t-1$ e o volume acumulado também foi mais bem previsto quando se utiliza somente dados de precipitação no vetor de entrada.

Ballini et al. (2003) apresentaram um modelo de rede neural nebulosa para previsão de vazões sazonais. O modelo foi baseado em um método de aprendizado construtivo onde grupos de neurônios competem quando a rede recebe uma nova entrada e onde a rede aprende os parâmetros fundamentais para definir as regras nebulosas e funções de pertinência para cada variável de entrada. O modelo foi aplicado para o problema de previsão de vazões médias mensais de três usinas hidroelétricas situadas em diferentes regiões do Brasil. O desempenho do modelo foi comparado com métodos convencionais usados para previsão de vazões. Os resultados mostraram que a rede neural nebulosa forneceu um melhor desempenho para previsão um passo à frente, com erros significativamente menores que as outras abordagens.

Marinho et al. (2004) apresentaram uma comparação entre os modelos *adaptive network based fuzzy inference system*, conhecido como ANFIS, modelos periódicos auto-regressivos e uma rede neural multi-camadas para previsão de vazões sazonais. O trabalho também propôs o estudo da relação entre o modelo periódico auto-regressivo e o número de neurônios na camada de entrada dos modelos ANFIS, rede neural multi-camadas (MLP), baseado no critério BIC. Estes modelos foram ajustados aos dados de vazões das hidrelétricas de Tucuruí e Boa Esperança, para realizar previsão um passo à frente. Os resultados mostraram que o modelo ANFIS para fazer previsão de vazões afluentes médias

mensais um passo à frente apresentou erros menores comparados com as outras metodologias.

Méndez et al. (2004) trabalharam com previsão de vazões mensais com os modelos estatísticos Box-Jenkins e as diárias também com redes neurais. A aplicação foi para 3 pontos de medição no rio Xallas, localizado no Nordeste da Espanha, e os dados compreenderam o período de 1992 a 1999. Observou-se que os modelos Box-Jenkins são adequados para análises a nível mensal do comportamento chuva x vazão na bacia. Já para intervalos diários, os modelos Box-Jenkins não apresentaram bons resultados, mas o modelo de redes neurais apresentou-se adequado para essa análise.

Rajurkar et al. (2004) trabalharam com a modelagem de vazões diárias durante eventos de cheias, usando a técnica de redes neurais. O processo de transformação chuva x vazão foi analisado associando uma metodologia linear simples com um modelo de redes neurais. O estudo usou informações de duas grandes áreas na Índia e outras cinco áreas usado pela organização mundial da meteorologia (WMO) para comparação entre os modelos hidrológicos que estão sendo atualmente utilizados. O estudo demonstrou que o modelo analisado mostrou-se eficiente para locais de diferentes posições geográficas, o que prova a versatilidade do mesmo. Mais importante, a substituição de dias anteriores de vazão por um termo que representa a vazão estimada por um modelo linear melhorou muito o desempenho da rede neural.

Lima et al. (2005) analisaram a adequabilidade dos modelos de RNAs ao processo de transformação chuva-vazão e compararam seu desempenho com a performance do modelo hidrológico conceitual MODHAC. Várias tipologias de modelos de RNAs foram testadas de modo a verificar suas adequabilidades e a influência de parâmetros nos ajustes. Os testes dos modelos de RNAs foram realizados com dados dos postos fluviométricos Icó e Podimirim, na bacia do rio Jaguaribe (CE). Foram realizadas simulações, de modo a combinar e melhor avaliar os modelos de redes neurais para três diferentes configurações: i) número e parâmetro dos nós na camada de entrada; ii) período de calibração e

verificação; iii) número de nós na camada intermediária; e iv) função de transferência. Foi possível constatar que em todos os modelos, os melhores resultados foram obtidos para os períodos no meio da série. No entanto, deve-se ter cuidado com a existência de eventos extremos, dada sua influência no treinamento e na validação. Constatou-se a superioridade da modelagem por RNAs sobre a conceitual, em todos os modelos testados em ambos os postos estudados.

Silva e Silva (2005) aplicaram um modelo RNA na previsão mensal de consumo de água para os sistemas de abastecimento e distribuição de água do Litoral Norte do Estado de São Paulo. Concluiu-se que a tecnologia de Redes Neurais Artificiais atendeu a expectativa e alcançou bons resultados ao ser aplicada. A rede neural Universal – Região foi a que apresentou o melhor resultado, apresentando erros percentuais médios de 8,12% e 11,92% em suas previsões para dados desconhecidos. As RNAs de Ubatuba, Caraguatatuba e Ilhabela apresentaram taxas de erro percentual médio elevadas em suas novas previsões. A RNA Universal - Região apresentou o melhor resultado por ter sido treinada com um volume maior de dados, recebendo um número maior de exemplos e pôde aprender o comportamento das variações de consumo de água das cidades, podendo assim, realizar o cálculo das previsões com maior precisão. Outro fator que influenciou o bom desempenho da RNA Universal-Região foi à organização uniforme dos dados de entrada, que apresentaram proporcionalmente informações de cada cidade.

Valença e Valença (2005) mostraram a aplicação de redes neural artificiais (ANNs) para modelar os processos chuva vazão e vazão vazão. Os autores apresentaram os princípios básicos de uma rede neural, as arquiteturas mais comuns, o algoritmo de treinamento, a divisão dos dados, normalização, validação cruzada e métricas para avaliar o desempenho. Os autores utilizaram as informações de Guarapiranga no período de janeiro de 1978 a outubro de 1986 (chuva x vazão), onde obtiveram um erro percentual médio absoluto foi de 53,9% e do erro padrão foi de 4,1 m³/s. Através do programa *Neurohidro* (Valença 2005), também foi realizado a previsão de vazões naturais médias mensais afluentes ao reservatório de Boa Esperança. Esta série histórica compreendeu o período que

vai de 1931 a 2003 (vazão x vazão) e obteve-se um erro percentual médio absoluto de 22,1% e erro padrão de 111,5 m³/s.

Assim, ao longo da última década, cada vez mais, muitos trabalhos têm mostrado a prosperidade da aplicação da metodologia de redes neurais em diversas áreas e em particular na previsão de vazões, para intervalos temporais diários e mensais. Muitos dos quais compararam a aplicações dos tradicionais modelos estatísticos com os modelos que usam a técnica de redes neurais, com esses últimos sempre com resultados bem superiores.

Contudo, poucos trabalhos têm mostrado a preparação (seleção e análise) dos dados de entrada da rede neural. Essa preparação dos dados pode melhorar ainda mais os resultados, visto que muitos dos resultados da aplicação de diversas metodologias de redes neurais estão consolidados.

Capítulo 3

Materiais e Métodos

3.1 DESCRIÇÃO DA ÁREA EM ESTUDO

O Alto São Francisco está localizado entre os paralelos 18º e 21º de latitude sul e os meridianos 43º30' e 46º40' de longitude no estado de Minas Gerais (figura 01), em uma zona climática tropical. A área estende-se desde as cabeceiras, na Serra da Canastra, município de São Roque de Minas, até à cidade de Pirapora (MG), abrangendo as sub-bacias dos rios das Velhas, Pará e Indaiá, além das sub-bacias dos rios Abaeté e Jequitaí.

A Agência Nacional de Energia Elétrica - ANEEL divide a bacia do rio São Francisco em nove sub-bacias, sendo a região estudada a que corresponde à sub-bacia 40. Esta sub-bacia representa a área de drenagem do reservatório de Três Marias e parte da área de drenagem do Alto São Francisco (sub-bacias 40 e 41). Segundo a ANA/SPR, o Alto São Francisco possui 167 municípios, 93% de urbanização, 62,9hab/km² e uma disponibilidade hídrica de 6.003 m³/hab/ano.

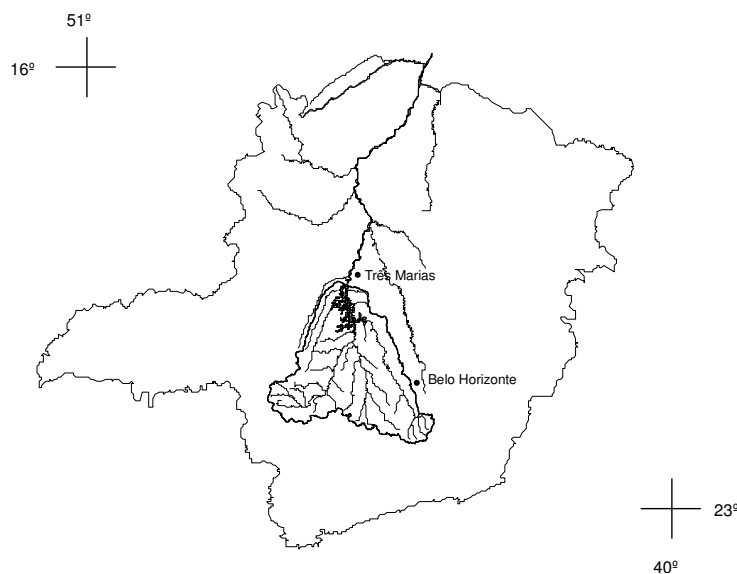


Figura 3.1 – Localização da sub-bacia 40 no estado de Minas Gerais (CPRM,1996).

A usina hidrelétrica de Três Marias é de propriedade da CEMIG, tem uma potência instalada de 396 MW, está localizada em -18⁰ 12' 50" e -45⁰ 15' 30", e

teve o início de suas obras em maio de 1957 sendo concluída em janeiro de 1961. Sua construção teve como principais objetivos: regularização do curso das águas do rio São Francisco nas cheias periódicas, melhoria da navegabilidade, utilização do potencial hidrelétrico, e fomento da indústria e irrigação. O reservatório tem uma área máxima de 1010km² e mínima de 316 km², correspondendo a volumes de 19.528 e 4.250hm³, respectivamente. O nível máximo operativo é 572,50m, e o nível mínimo operativo é 549,20m, a área de drenagem é de 50.732 km² e a vazão média de longo termo (MLT) de 689 m³/s.

As unidades geomorfológicas existentes na bacia do Alto São Francisco são o Planalto Dissecado do Centro-Sul e Leste de Minas, a Depressão São Franciscana, o Planalto do São Francisco, o Quadrilátero Ferrífero e a Serra do Espinhaço.

O Alto São Francisco apresenta topografia ligeiramente acidentada, com serras e terrenos ondulados e altitudes de 1.600 a 600m. O divisor leste é formado pelas montanhas da Serra do Espinhaço, estreitas e alongadas na direção N-S, e com altitudes de 1.300 a 1.000m. Do lado oeste, destacam-se as Serras Gerais de Goiás, cujas cotas oscilam entre 1.200 e 800m.

Segundo a ANA (2004), o clima predominante na região é o Tropical Úmido e temperado de altitude, insolação média anual de 2400h e evaporação média anual de 1000mm.

As características climáticas mais importantes na região em estudo são a predominância de temperaturas medianas a elevadas durante todo o ano e a existência de duas estações bem marcadas: uma estação chuvosa no verão e outra seca no inverno.

A bacia do Alto São Francisco apresenta temperatura média anual variando entre 19 e 23°C, sendo que as menores temperaturas são encontradas ao sul da bacia devido à influência orográfica, nas altitudes mais elevadas da serra do Espinhaço e nas cabeceiras da bacia ocorrendo um aumento gradativo à medida que se alcançam as latitudes mais baixas. Esses valores indicam a predominância de temperaturas medianas a elevadas na região durante quase todo o ano, principalmente na primavera e no verão.

Com relação às precipitações, os totais pluviométricos anuais variam, de forma geral, entre 1700mm nas cabeceiras da bacia a 1000mm na confluência do

rio São Francisco com rio das Velhas. O trimestre mais chuvoso contribui com cerca de 55 a 60% do total anual precipitado, correspondendo, para a quase totalidade da bacia, aos meses de novembro-dezembro-janeiro. A exceção ocorre na extremidade ocidental da bacia, onde os meses mais chuvosos são dezembro, janeiro e fevereiro. Já o trimestre mais seco, correspondente aos meses de junho, julho e agosto, contribui com menos de 5% da precipitação anual.

Por se situar na região Sudeste, o Alto São Francisco é afetado pela passagem de sistemas frontais, e a interação desses com a convecção tropical, denominada de Zona de Convergência do Atlântico Sul - ZCAS. Um percentual significativo da precipitação nesse trecho da bacia deve-se à atuação desses sistemas (Climanálise). Sistemas de escala subsinótica, tais como linhas de instabilidade, também ocasionam precipitação. Estas são de curta duração, em oposição às associadas as frentes e ZCAS que duram alguns dias e abrangem áreas maiores.

O Alto São Francisco apresenta valores de vazão média específica entre 14,3 e 21,7l/s.km². As enchentes ocorrem com maior freqüência nos afluentes, condicionadas pelo padrão de chuvas e relevo da região. São registradas inundações regulares nos rios Pará, Paraopeba, das Velhas e Jequitaiá.

A Floresta predominante na região úmida, apresentando-se, também, nas regiões subúmidas secas e úmidas, ao longo dos rios e riachos, onde ocorre maior umidade do solo, formando floresta de galerias ou mata ciliar. A floresta ocorre, ainda, nas regiões de clima subúmido seco e transicional para semi-árido, onde há presença de solos de alta fertilidade.

O Cerrado predomina nas regiões de clima úmido e subúmido e de solos de baixa fertilidade.

Ainda segundo a ANA (2004), a região do Alto São Francisco tem as seguintes características físicas: área de 100.076km², declividade do rio principal de 0,70 a 0,20m/km, contribuição de vazão média anual de 42%, vazão média anual máxima de 1303m³/s (em Pirapora, fevereiro), vazão média anual mínima de 637m³/s (em Pirapora, agosto).

3.2 MÉTODOS

3.2.1 Estatística

3.2.1.1 Dendograma

Na análise de agrupamentos (cluster analysis) a similaridade entre duas amostras pode ser expressa como uma função da distância entre os dois pontos representativos destas amostras no espaço n-dimensional.

Existem muitas maneiras de procurar agrupamentos no espaço n-dimensional. A maneira matematicamente mais simples consiste em agrupar os pares de pontos que estão mais próximos, usando a distância euclidiana (equação 3.1), e substituí-los por um novo ponto localizado na metade da distância entre eles.

$$X_{ab}^2 = \sum_{j=1}^n (d_{aj} - d_{bj})^2 \quad \text{eq 3.1}$$

Este procedimento, quando repetido até que todos os pontos sejam agrupados, leva à construção do dendograma; onde, no eixo horizontal são colocadas as amostras e, no eixo vertical, o índice de similaridade (s_{ij}) entre os pontos i e j , calculado segundo a seguinte expressão:

$$s_{ij} = 1 - \frac{d_{ij}}{d_{m\acute{a}x}} \quad \text{eq. 3.2}$$

onde: d_{ij} é a distância entre os pontos i e j e $d_{m\acute{a}x}$ é a distância máxima entre qualquer par de pontos.

Os dendogramas, portanto, consistem em diagramas que representam a similaridade entre pares de amostras (ou grupos de amostras) numa escala que vai de um (identidade) a zero (nenhuma similaridade).

A suposição básica de interpretação um dendograma é: quanto menor a distância entre os pontos, maior a semelhança entre as amostras. Os dendogramas são especialmente úteis na visualização de semelhanças entre

amostras ou objetos representados por pontos em espaço com dimensão maior do que três, onde a representação de gráficos convencionais não é possível.

3.2.1.2 Análise de Componentes Principais – ACP

A Análise de Componentes Principais (ACP) tem o objetivo de simplificar a estrutura de um conjunto de variáveis e explicar o máximo possível à variância total destas, através da determinação de valores calculados como combinações das variáveis originais. Os Componentes Principais (CP's) devem ser usados para reduzir a dimensionalidade dos dados e são formados pela transformação de um conjunto de "p" variáveis, em um conjunto de variáveis descorrelatadas. Estas novas variáveis são combinações lineares das variáveis originais e são obtidas em ordem decrescente de importância de tal forma que, por exemplo, o primeiro componente principal é responsável pelo máximo possível da variabilidade total nos dados originais. Os Componentes são extraídos na ordem do mais explicativo para o menos explicativo, ou seja, são obtidos em ordem decrescente de máxima variância. Teoricamente o número de Componentes é sempre igual ao número de variáveis. Entretanto, alguns poucos Componentes são responsáveis por grande parte da explicação total da variância da série.

3.2.2 Geoestatística

Algumas variáveis hidrológicas apresentam a propriedade de variar espacialmente e com uma certa dependência de valores no mesmo espaço, conduzindo à necessidade de um estudo geoestatístico. É justamente nos problemas em que a estatística clássica tem limitações que a geoestatística tem suas maiores aplicações.

Nas análises com a geoestatística, quando os postos de observações são dispersos na área de interesse, muitas vezes é necessário considerar tendência nas informações desses postos. A consideração da tendência terá um efeito significativo na interpolação para grandes áreas com poucas informações e quando se necessita de grandes extrapolações. Normalmente dispõe-se de três tipos de opções na análise da tendência: não considerar, tendência linear e quadrática. Quando se usa a Krigagem ordinária ou quando os dados são

dispersos uniformemente na área de interesse, recomenda-se não considerar a tendência.

O teste da normalidade das informações também se faz necessário nos estudos com a geoestatística. Caso os dados originais não sigam uma distribuição normal é necessário fazer uma transformação logarítmica nos mesmos.

O modelo de variabilidade espacial foi desenvolvido por Krige (1951) a partir de semi-variogramas definidos pela equação 3.3,

$$\gamma(h) = 1/2 \text{ var}[Z(x) - Z(x+h)] \quad \text{eq. 3.3}$$

onde estão correlacionadas as diferenças dos valores observados $Z(x)$ com a distância de separação entre amostras h da variável regionalizada - VR.

Os semivariogramas expressam a dependência espacial entre as amostras. Havendo dependência espacial, podem-se estimar valores da propriedade em estudo, para os locais não amostrados dentro do campo, sem tendências e com variância mínima, através da *krigagem*, seguindo a proposta de Krige.

Quando o gráfico do semivariograma é idêntico para qualquer direção de h , ele é chamado isotrópico e representa uma situação bem mais simples do que quando é anisotrópico, onde o semi-variograma precisa passar por transformação. Segundo Vieira (2000), para garantir a precisão da interpolação, em vez de se preocupar com a escolha do método de transformação da anisotropia, apenas deve-se limitar à faixa de distância no qual se utiliza o semivariograma.

Na krigagem, o procedimento é semelhante ao de interpolação por média móvel ponderada, exceto que aqui os pesos são determinados a partir de uma análise espacial, baseada no semivariograma experimental.

O termo *kriging* abrange um conjunto de métodos, sendo os mais usuais os seguintes:

- *Krigagem simples*: assume que as médias locais são relativamente constantes e de valor muito semelhante à média da amostragem que é conhecida. A média da população é utilizada para cada estimativa local, em conjunto com os pontos vizinhos estabelecidos como necessários para a estimativa.

A krigagem simples é baseada numa função aleatória $Z(x)$, estacionária, com média m e covariância $C(h)$ (eq 3.4).

$$Z_{KS}(x_0) = \sum_{\alpha=1}^n \lambda_{\alpha} Z(x_{\alpha}) + \left[1 - \sum_{\alpha=1}^n \lambda_{\alpha} \right] \cdot m \quad \text{eq.3.4}$$

Onde λ_{α} são pesos determinados para minimizar a variância da estimativa.

- *Krigagem Ordinária*: as médias locais não são necessariamente próximas da média da amostragem, usando-se apenas os pontos vizinhos para a estimação.

A krigagem ordinária não requer o prévio conhecimento da média m . O estimador pontual da krigagem ordinária é dado por:

$$Z_{KO}^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i), \quad \text{com } \sum_{i=1}^n \lambda_i = 1 \quad \text{eq.3.5}$$

- *Co-Krigagem*: é uma extensão do anterior a situações em que duas ou mais variáveis são espacialmente dependentes e a variável que se quer estimar não está amostrada com a intensidade com que estão as outras variáveis dependentes. Utilizando-se assim os valores destas e as suas dependências para estimar a variável requerida.

Seguindo a proposta da krigagem, para o estabelecimento do semivariograma, todos os dados são organizados em pares ordenados (x,y) , os quais são representados por uma variável locada no espaço, x_i , proveniente de um conjunto de valores observados de tamanho n , $\{z(x_i), i = 1, n\}$, num campo de área S .

Considerando que a correlação entre duas variáveis $Z(x)$ e $Z(x+h)$ seja inversamente proporcional a distancia h , que as separa, a função de covariância diminui o seu valor a partir da origem, $C(0)$, e o semi-variograma $\gamma(h) = C(0) - C(h)$ aumenta a partir da origem. É a partir destas relações que são definidos os parâmetros do variograma.

O primeiro parâmetro a ser verificado é o “efeito pepita”, C_0 , definido como uma constante na equação do semivariograma, revelando a descontinuidade do mesmo para distâncias menores que a mínima amostrada e representando a componente da variabilidade espacial que não pode ser relacionada a uma causa específica, normalmente é tido como uma conseqüência do número limitado de observação para pequenas distâncias ou de erros da medição da variável. Outro

parâmetro de ajuste é o “sill” ou patamar ($C_0 + C_s$), que corresponde à máxima variância da amostra. Outro parâmetro a ser determinado é o alcance (a), definido pela distância dentro da qual, as amostras apresentam-se correlacionadas espacialmente. À medida que aumenta a distância entre os pontos a serem comparados, maior é a semivariância, até que, a partir de determinada distância (range = alcance), a semivariância estabiliza num valor (sill = patamar) que é igual à variância dos dados da população considerada, significando que já não existe qualquer relação entre os pares de dados considerados a esta distância. A figura 3.2 mostra a forma característica de um semivariograma.

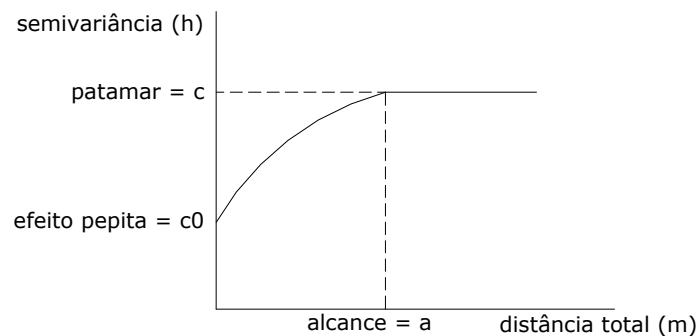


Figura 3.2 - Forma característica de um semi-variograma.

A parametrização do variograma é normalmente feita ajustando valores experimentais a modelos conhecidos, através de procedimentos específicos. É habitual encontrar dois tipos de variogramas elementares, variogramas com e sem patamares. Entre os variogramas com patamares o mais comum é o esquema esférico que apresenta um comportamento linear junto à origem, existindo ainda o variograma exponencial e o gaussiano. Quanto aos variogramas sem patamar é comum encontrar uma classe de funções que dependem de h^a (h =distância entre as variáveis e a =alcance), em que o variograma linear é um caso particular.

Os modelos Gaussiano, Exponencial e Esférico estão equacionados nas equações 3.6, 3.7 e 3.8, respectivamente, onde a_g é $a/\sqrt{3}$, para o modelo gaussiano, e a_s é $a/3$, para o modelo esférico. Para o modelo exponencial, o valor “ a ” é propriamente o alcance do variograma.

Modelo Gaussiano:

$$\gamma(h) = \begin{cases} C_0 + C_S \left[1 - \exp\left(-\frac{h^2}{a_g^2}\right) \right] & h \neq 0 \end{cases} \quad \text{eq.3.6}$$

Modelo Exponencial:

$$\gamma(h, C_0, C_S, a_e) = \begin{cases} C_0 + C_S \left[1 - \exp\left(-\frac{|h|}{a_e}\right) \right] & h \neq 0 \end{cases} \quad \text{eq.3.7}$$

Modelo Esférico:

$$\gamma(h, C_0, C_S, a_s) = \begin{cases} C_0 + C_S \left[\left(\frac{3}{2}\right)\left(\frac{|h|}{a_s}\right) - \left(\frac{1}{2}\right)\left(\frac{|h|}{a_s}\right)^3 \right] & 0 < |h| \leq a_s \\ C_0 + C_S & a_s < |h| \end{cases} \quad \text{eq.3.8}$$

Variograma em h^α : $\gamma(h) = a.h^\alpha$, $0 \leq \alpha \leq 2$, o variograma linear é um caso particular deste, quando $\gamma(h) = a.h$.

A figura 3.3 mostra a forma dos modelos que são aplicados a variogramas: exponencial, gaussiano, linear e esférico.

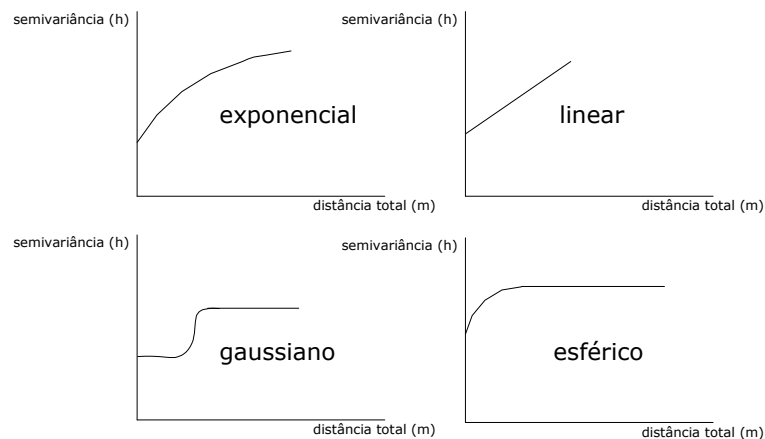


Figura 3.3 - Modelos matemáticos aplicados a variogramas (fonte: Surfer (2002)).

No ajuste de semivariogramas experimentais, sempre existe um certo grau de incerteza sobre as hipóteses de estacionaridade assumidas, modelos selecionados, parâmetros ajustados e qualidade da krigagem. Esta incerteza representa o erro da estimativa que pode ser avaliado pelo procedimento denominado de validação cruzada, que serve para validar o modelo de variograma utilizado para representar a estrutura de correlação da variável

estudada. Nessa técnica, cada ponto medido é excluído e seu valor estimado com os dados restantes, ou seja, ela estima os valores da variável em estudo sobre os mesmos pontos amostrados e compara os novos valores com os dados medidos, calculando o erro da interpolação. Os tipos erros normalmente utilizados nesse processo de validação são: médio, médio quadrático, médio padrão, médio normalizado e o desvio padrão normalizado. A interpolação ideal segue os seguintes padrões:

- Erro médio quadrático = erro médio padrão;
- Erro médio normalizado = 0;
- Desvio médio normalizado = 1.

Como a interpolação por krigagem está fortemente associada ao modelo de semivariograma escolhido, a validação cruzada pode ser usada para a decisão do modelo a ser ajustado. Determinado o modelo de melhor ajuste à amostra, a técnica de interpolação kriging é utilizada para representar a variabilidade espacial.

Segundo classificação proposta por Cambardella et al. (1994) semivariogramas podem ser considerados como caracterizando forte, moderada ou fraca dependência espacial. O critério proposto para classificação é a relação $Co / (Co + Cs) \times 100$. Valores inferiores a 25% caracterizam forte dependência espacial, entre 25% e 75% moderada e acima de 75%, fraca.

A Krigagem também pode ser usada no cálculo da precipitação média de uma área. Neste caso, utilizando o software Surfer, por exemplo, procede-se inicialmente a determinação da precipitação de toda a área, através da interpolação de Kriging, determina-se o volume dessa precipitação e o divide pela área total da bacia.

3.2.3 Redes neurais

As variantes de uma rede neural são muitas, e combinando-as, pode-se mudar a arquitetura conforme a necessidade da aplicação. Basicamente, os itens que compõem uma rede neural são os seguintes: i) conexões entre camadas; ii) número de camadas intermediárias; iii) quantidade de neurônios; iv) função de transferência e v) algoritmo de aprendizado/treinamento.

As camadas escondidas situam-se entre a camada de entrada e a camada de saída da rede neural. Não existe uma regra que defina o número de camadas escondidas, e o número de neurônios da camada escondida também não obedece a nenhuma regra específica. Contudo, a utilização de um grande número de camadas escondidas não é recomendada, porque cada vez que o erro médio durante o treinamento é utilizado para atualizar os pesos das sinapses da camada imediatamente anterior, ele se torna menos útil ou preciso. Testes empíricos com a rede neural *MLP backpropagation* não demonstraram vantagem significativa no uso de duas camadas escondidas ao invés de uma para problemas menores. Por isso, para a grande maioria dos problemas utiliza-se apenas uma camada escondida.

Com relação ao número de neurônios nas camadas escondidas, este é geralmente definido empiricamente, os melhores resultados são obtidos através de um processo de tentativa e erro.

A topologia define a maneira com que os vários neurônios constituintes de uma rede neural são interconectados.

As sinapses representam barreiras que, de um certo modo, modulam os sinais que passam através delas. Esta modulação representa a *força* da sinapse. Em um neurônio artificial, a força da sinapse é chamada de peso w . Como os neurônios possuem um grande número de dendritos, eles podem receber sinais de vários neurônios simultaneamente.

A entrada líquida é dada pela seguinte equação:

$$net_j = \sum_{i=1}^N w_{ij} x_i - \phi$$

Onde: w_{ij} é o peso;

x_i é o valor da entrada;

ϕ é o limiar.

Os neurônios de uma rede neural são geralmente dispostos em camadas (*layers*), que recebem as mesmas denominações descritas para os tipos de neurônio:

- **Camada de Entrada:** recebe os dados de entrada e não realiza nenhum processamento sobre os dados, servindo apenas para distribuição dos mesmos para os neurônios da camada subsequente;
- **Camadas Escondidas:** suas entradas e saídas permanecem dentro do sistema;
- **Camada de Saída:** gera os dados de saída;

A Função de Ativação é a responsável pela não linearidade das redes neurais. Existem diversas funções de ativação, entre elas:

- Função Sigmóide (logística):

$$\varphi(v) = \frac{1}{1 + e^{-v}}$$

- Função Tangente Sigmóide:

$$\varphi(v) = \frac{2}{1 + e^{-2v}} - 1$$

3.2.3.1 Redes Multi-Layer Perceptrons - MLP com BackPropagation

Em uma rede neural MLP, os neurônios (isto é, unidades de processamento), são arranjados em camadas. A estrutura deste modelo é composta por uma ou mais camadas de neurônios entre a camada de entrada e a camada de saída, como mostra a figura 3.4.

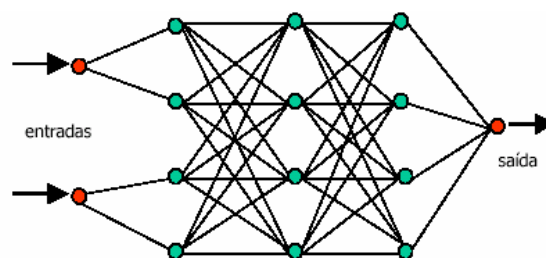


Figura 3.4 – Estrutura da rede MLP (Marinho, 2004).

As redes de múltiplas camadas só oferecem vantagens sobre as redes de uma única camada se a função de ativação da camada escondida for não-linear.

O fluxo das informações podem ser Forward ou Feed-Forward. As redes neurais *Feed-Forward* são redes de uma ou mais camadas de processadores, cujo *fluxo de dados* é sempre em *uma única direção*, isto é, não existe realimentação.

As Redes Neurais devem ser configuradas de modo que a apresentação de um conjunto de entradas produza o conjunto de saídas desejado.

A Convergência é a capacidade da Rede Neural de aprender todos os padrões do conjunto de treinamento. Se a rede neural for *pequena*, não será capaz de *armazenar* todos os padrões necessários. Isto é, a rede não deve ser *rígida* a ponto de não *modelar fielmente os dados*. Se a rede for *muito grande* (*muitos parâmetros = pesos*), não responderá corretamente a padrões nunca vistos. Isto é, a rede não deve ser *excessivamente flexível* a ponto de *modelar* também o *ruído*.

O processo que ocorre durante os ajustes dos pesos sinápticos é chamado de aprendizado ou treinamento, realizado alterando-se os pesos sinápticos segundo regras estabelecidas. Assim, a informação é armazenada em uma Rede Neural pelos pesos sinápticos.

Uma rede pode se especializar demasiadamente em relação aos exemplos contidos na base de aprendizado. Este tipo de comportamento leva a um problema de aprendizado conhecido como super-aprendizado (*over-training / over-fitting*). Normalmente o over-fitting pode ser detectado/evitado através do uso de um teste de generalização por validação cruzada (cross-validation) (Osório e Bittencourt, 2000).

O algoritmo backpropagation define uma maneira sistemática de atualização dos pesos das diversas camadas, a partir da seguinte idéia central: “Os erros dos neurônios da camada escondida são determinados pela propagação reversa dos erros dos neurônios da camada de saída”.

As equações de aprendizado são determinadas utilizando o método de gradiente descendente (Haykin, 1998), porém buscando-se minimizar o erro global da camada de saída. Ou seja, cada peso sináptico “i” do elemento processador “j” é atualizado proporcionalmente ao negativo da derivada parcial do erro deste processador com relação ao peso, como mostra a equação abaixo.

$$w \leftarrow w - \eta \cdot \nabla E(w) \quad \text{eq. 3.9}$$

Onde: w representa o peso a ser ajustado.

A velocidade de treinamento η tem a função é ponderar as correções dos pesos, de modo a evitar grandes correções; principalmente no final do processo,

quando boa parte dos pesos já foi corrigida e uma nova correção muito acentuada pode comprometer todas as correções já efetuadas.

3.2.3.1.1 Momentum

Uma das abordagens utilizadas para diminuir o risco da irregularidade real da função erro (composta de diversos máximos e mínimos locais) consiste na introdução de um termo adicional chamado de “momento”.

O termo “momento” introduzido tem por objetivo permitir o aumento da taxa de aprendizagem sem que ocorram oscilações nos pesos ajustados, bem como atuar no aumento da velocidade de convergência. A equação 3.10 mostra a equação de ajuste dos pesos, como o *momento* introduzido.

$$w \leftarrow w - \eta \cdot \nabla E(w) + \alpha \cdot \delta w_{old} \quad \text{eq. 3.10}$$

Onde, α é o *momentum*, uma constante que é multiplicada pela variação de pesos da iteração anterior.

A saída calculada pela rede será comparada com o conjunto de valores reais e, em função do erro calculado, todos os pesos serão novamente corrigidos. Este procedimento de aprendizado é repetido diversas vezes, até que, para todos os processadores da camada de saída e para todos os padrões de treinamento, o erro seja menor do que o especificado. O erro máximo estipulado e o número de iterações são parâmetros que são definidos pelo usuário antes de iniciar a execução do algoritmo de treinamento da rede.

O algoritmo de backpropagation apresenta ótimos resultados na maioria dos casos, havendo contudo alguns problemas, tais como o processo longo de aprendizado (resultado de má escolha de η e α), e fenômenos como *network paralysis* (paralisia do treinamento, que é contornada diminuindo-se o valor de η) e mínimos locais.

3.2.3.1.2 Algoritmo de Levenberg-Marquardt (LM)

Um dos métodos mais rápidos para o treinamento de redes neurais feedforward de tamanho moderado (até algumas centenas de pesos sinápticos) é o método de Levenberg-Marquardt, que é uma variação do Backpropagation de modo a obter maior velocidade de convergência.

Enquanto o back-propagation padrão utiliza a descida de gradiente como método de aproximação do mínimo da função erro, o algoritmo de Levenberg-Marquardt (LM) utiliza uma aproximação pelo método de Newton. Esta aproximação é obtida a partir da modificação do método de Gauss-Newton introduzindo-se o parâmetro μ , conforme a eq. 3.11:

$$\Delta w = (J^T J + \mu I)^{-1} J^T \varepsilon \quad \text{eq. 3.11}$$

onde:

- w representa a diferença entre os pesos inicial e final;
- μ é um escalar que controla a derivação dos erros, permitindo que o termo $(J^T J)$ possa ser invertido;
- J é o jacobiano dos erros da camada de saída. Cada elemento desta matriz J representa uma derivada parcial de um elemento da matriz de erros com o seu correspondente peso;
- I é a matriz identidade. Como ela é multiplicada pela constante α , gera-se uma nova matriz contendo como elementos apenas valores de α ;
- ε é o vetor de erros calculados.

3.2.3.1.3 Método da Validação Cruzada

A utilização de procedimentos de validação cruzada (VC) ou técnicas de regularização (Prechelt, 1998; Reed, 1993), diminuem os problemas em relação ao sobre-treinamento da rede e conseqüentemente amenizam os efeitos das redes sobre dimensionadas (Castro et al., 1999).

O treinamento usando validação cruzada deve ser interrompido quando o erro na validação começar a crescer de forma consistente.

Assim, para este método devem-se dividir os padrões em três conjuntos:

- treinamento: padrões usados para modificar os pesos;
- validação: padrões usados para verificar o problema de overfitting;
- teste: padrões para testar o desempenho do modelo final.

Além disso, para que a Rede Neural realize um bom aprendizado são necessários alguns procedimentos tais como a normalização e a seleção ou combinação de variáveis. Estes procedimentos conseguem diminuir o tamanho da rede, acelerar o aprendizado e melhorar a generalização da Rede Neural.

A normalização pode ser feita através do escalonamento linear para o caso de dados com comportamento linear e, para o caso de dados com comportamento não linear, pode se usar transformação logarítmica.

3.2.3.1.4 Método de Nguyen-Widrow

Este método é usado na inicialização dos pesos e bias que proporciona rapidez no treinamento em relação ao tradicional random.

Este algoritmo escolhe valores com o objetivo de distribuir a região ativa de cada neurônio no layer (camada) uniformemente através do espaço da entrada da camada.

O método é baseado em uma análise geométrica da resposta da camada escondida para uma simples entrada e a análise é estendida para o caso de diversas entradas, usando transformada de Fourier.

3.2.3.2 Redes NSRBN (Adaptado de Valença (2005b))

A arquitetura de uma rede neural, de acordo com a possibilidade de mudanças no seu tamanho durante o processo de treinamento, pode ser classificada em estática ou dinâmica. Uma rede é dita de arquitetura estática quando o número de camadas e unidades de processamento permanece constante desde a concepção da rede até a finalização do seu treinamento. Por outro lado, uma rede de arquitetura dinâmica tem como característica principal um processo de mudanças no tamanho durante o treinamento. A principal vantagem de um algoritmo com esse tipo de abordagem é que este realiza uma aproximação parcimoniosa do problema, resultando em uma melhor generalização.

Dentre estas redes construtivas uma que tem sido bastante aplicada na área de recursos hídricos são as redes neurais compostas com blocos de regressões sigmóides não-lineares (NSRBN – Non-linear Sigmoidal Regression Blocks Networks) proposta por Valença (1999). Essas redes têm uma poderosa capacidade de mapeamento no sentido de que qualquer função contínua definida em um conjunto compacto em RN pode ser uniformemente aproximada (Valença, 1999).

Dentre as características principais das Redes Compostas por Blocos de Regressões Sigmóides Não-Lineares NSRBN (Non Linear Sigmóidal Regression Blocks Networks) duas merecem destaque: uma estrutura compacta, por possuir apenas uma camada escondida e um algoritmo construtivo baseado em um método de aprendizado construtivo que é responsável pela definição de sua complexidade ótima.

As Redes NSRBN são redes com blocos de polinômios homogêneos, com uma função de ativação na camada intermediária (por exemplo: tangente hiperbólica) e, na camada de saída, uma ativação mais complexa, uma vez que realizam uma regressão logística não linear. A figura 3.5 mostra a arquitetura de uma rede NSRBN.

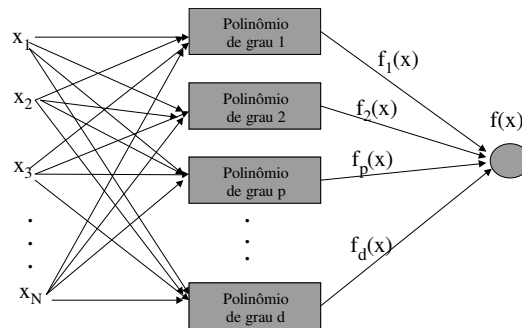


Figura 3.5 – Arquitetura combinatorial.

Nessa nova classe de redes proposta, um algoritmo construtivo é utilizado para incrementar a estrutura da rede, partindo-se de uma pequena estrutura, até que se atinja uma estrutura de tamanho ótimo, segundo critérios estatísticos de seleção adequados.

A proposta do modelo NSRBN é de se construir a rede por blocos de polinômios homogêneos, utilizando uma função de ativação adequada (tangente hiperbólica, por exemplo), para os neurônios escondidos. Quanto aos neurônios de saída, estes têm uma ativação mais complexa, uma vez que realizam uma regressão logística não-linear com relação à saída dos neurônios escondidos. O primeiro passo para a construção de um modelo combinatorial consiste em dividir a função $f(x)$ em blocos compostos de termos de igual grau, como apresentado na figura 3.5.

$$f(x) = \tilde{f}_1(x) + \tilde{f}_2(x) + \dots + \tilde{f}_d(x) \quad \text{eq. 3.12}$$

O bloco de grau p ($p=1,2,\dots,d$), de acordo com o polinômio de Kolmogorov-Gabor (teorema de Weierstrass), pode ser expresso como uma soma de produtos das entradas, isto é:

$$f_p(x) = \sum_{i_1=1}^N \sum_{i_2=i_1}^N \dots \sum_{i_N}^N a_p(i_1, i_2, i_3, \dots, i_N) \cdot x_1^{q_{p1}(i_1)} \cdot x_2^{q_{p2}(i_2)} \cdot \dots \cdot x_N^{q_{pN}(i_N)} \quad \text{eq.3.13}$$

em que, $q_{p1}(i_1) + q_{p2}(i_2) + q_{p3}(i_3) + \dots + q_{pN}(i_N) = p$ e $p \geq q_{pm}(i_m) \geq 0$ para $m=1,2,3,\dots,N$.

Tem-se como objetivo realizar o mapeamento de cada um dos polinômios homogêneos de grau p .

Uma forma simples seria realizar todos os produtos das variáveis de entrada procurando aproximar diretamente a hipersuperfície não-linear. Entretanto, isto não é prático, a menos que d e N sejam relativamente pequenos (modelo combinatorial). Logo, para se ter um modelo eficiente, faz-se necessário que ele possua funções de ativação que sejam capazes de realizar vários termos dos polinômios homogêneos de forma otimizada; isto é, sem que ocorra um crescimento explosivo no número de parâmetros necessários para ajustar a função desejada. Com esse objetivo, propõe-se uma nova classe de redes neurais que tenha por base o algoritmo combinatorial (composta pela soma de blocos) com uma estrutura de blocos similar a uma rede MLP, sendo que, diferentemente destas, os neurônios de saída realizam a técnica estatística de regressão logística não-linear com relação aos neurônios escondidos.

Uma rede NSRBN é uma rede combinatorial composta pela soma de p blocos ($p=1,\dots,d$) com estrutura similar a uma MLP, na qual as unidades escondidas desses blocos realizam uma regressão sigmóide linear das entradas (tal qual uma MLP), e as unidades de saída realizam uma regressão do tipo sigmóide não-linear das unidades escondidas (tal qual uma regressão logística não-linear). Essa nova estrutura formada é um aproximador universal com forte capacidade de mapeamento não-linear.

Utiliza-se aqui o termo sigmóide, significando qualquer função sigmoidal, como uma generalização da regressão logística clássica.

A estrutura proposta para cada bloco é composta de apenas uma camada escondida, tendo N_h unidades escondidas em cada bloco com função de ativação do tipo:

$$f(x) = (\sigma_{net(h)} + \theta_h) \quad \text{eq.3.14}$$

em que $net(h)$ é a entrada líquida para cada neurônio escondido, h é um número inteiro que representa as unidades escondidas em cada bloco (variando de 1 até a ordem p do bloco), θ_h é a tendência e $\sigma_{net(h)}$ é, por exemplo, a função tangente hiperbólica dadas por:

$$\sigma_{net(h)} = \frac{e^{net(h)} - e^{-net(h)}}{e^{net(h)} + e^{-net(h)}} \quad \text{eq.3.15}$$

Poder-se-ia utilizar outras funções, como é o caso da função recíproca, isto é,

$$f_r(x) = \frac{net(h)}{1 + |net(h)|} \quad \text{eq. 3.16}$$

O valor de p correspondente ao grau do bloco é quem vai definir o número de neurônios em cada bloco, isto é: o bloco de grau 1 ($p=1$) tem $h=1$, logo um neurônio escondido; o bloco de grau 2 ($p=2$) terá 2 neurônios escondidos ($h=1$ e $h=2$) e assim sucessivamente, de tal forma que o bloco de grau p terá $h=p$ e portanto p neurônios escondidos.

A estrutura proposta para cada bloco é similar à da figura 3.5, na qual apresenta-se o mapeamento para um bloco de grau p (a ordem dos neurônios escondidos h variará de 1 até p), para i variáveis de entrada (onde $i = 1, \dots, N$) e uma única unidade de saída.

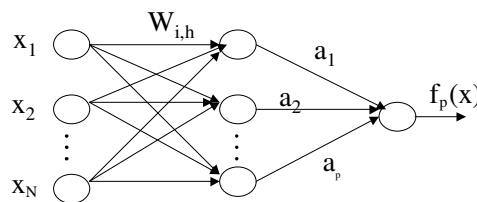


Figura 3.6 – Arquitetura de um bloco polinomial.

Esse bloco é uma rede *feedforward* completamente conectada com uma única camada escondida. A diferença básica desse bloco de grau p para uma rede MLP é que os neurônios de saída são expressos como uma regressão sigmóide não-linear das unidades escondidas.

Seja $\mathbf{x} = [1, x_1, x_2, x_3, \dots, x_N]^T$ o vetor coluna das variáveis de entrada, no qual o elemento x_i denota um elemento qualquer dele. As entradas são ponderadas pelos vetores de pesos $\mathbf{w} = [w_{0h}, w_{1h}, w_{2h}, \dots, w_{Nh}]^T$, $h=1, \dots, p$, nos quais p é a ordem do bloco. Portanto, utilizando-se a função de ativação estabelecida na equação 3.25, a saída de cada unidade escondida, $f(\text{net}(h))$, desse bloco de grau p pode ser escrito como:

$$f(\text{net}(h)) = (\sigma_{\text{net}(h)} + \theta_h) \quad \text{eq. 3.17}$$

em que: $\text{net}(h) = \mathbf{w}_h^T \mathbf{x}$, ou seja, $\text{net}(h) = \sum_{i=1}^N w_{ih} x_i + w_{0h}$, $h=1, 2, \dots, p$.

Logo, a unidade de saída (que realiza uma regressão logística não-linear) desse bloco pode ser representada pela seguinte equação:

$$f_p(x) = a_1(f(\text{net}(1)))^1 + a_2(f(\text{net}(2)))^2 + a_3(f(\text{net}(3)))^3 + \dots + a_p(f(\text{net}(p)))^p \quad \text{eq. 3.18}$$

sendo $\mathbf{a} = [a_1, a_2, a_3, \dots, a_p]^T$ o vetor dos pesos das unidades escondidas para a unidade de saída.

Fazendo-se uso da equação 3.17, a função $f_p(x)$ pode ser escrita da seguinte forma:

$$f_p(x) = a_1(\sigma_{\text{net}(1)} + \theta_1) + a_2(\sigma_{\text{net}(2)} + \theta_2)^2 + \dots + a_p(\sigma_{\text{net}(p)} + \theta_p)^p \quad \text{eq.3.19}$$

ou

$$f_p(x) = \sum_{h=1}^p a_h (\sigma_{\text{net}(h)} + \theta_h)^h \quad \text{eq.3.20}$$

com h representando a ordem dos neurônios na camada escondida, ($h=1, \dots, p$), do bloco de grau p .

De acordo com a figura 3.5, a arquitetura de uma rede NSRBN é composta pela soma de blocos com grau variando de 1 a d . Aqui, está sendo utilizado o bloco p (figura 6) para representar qualquer um desses blocos.

Portanto, usando-se esta arquitetura, uma função desconhecida f em R^N pode ser aproximada por uma rede NSRBN, através da seguinte expressão:

$$f(x) = \sigma_{\text{net}(o)} \left(\sum_{p=1}^d f_p(x) \right) \quad \text{eq. 3.21}$$

em que: $\sigma_{\text{net}(o)}$ uma função de ativação adequada para a unidade de saída.

Por outro lado, a construção das redes NSRBN é feita de forma incremental, de maneira que pode-se representá-la em uma forma similar a de um

bloco (equação 3.18), conforme apresenta-se a seguir. Considere o passo do algoritmo para o qual se adiciona-se à rede o bloco de ordem d , f_d , tendo-se anteriormente adicionado os blocos f_0, f_1, \dots, f_{d-1} . Fazendo-se $f_0 \equiv 0$, uma função f pode ser representada por:

$$f(x) = \sigma_{net(o)} \left(\sum_{p=0}^{d-1} f_p(x) + f_d(x) \right) \quad \text{eq 3.22}$$

na qual os pesos dos blocos f_0, f_1, \dots, f_{d-1} são congelados ao adicionar-se o bloco de ordem d , f_d .

Isso posto, pode-se escrever $f(x)$ da seguinte forma:

$$f(x) = \sigma_{net(o)} \{ c_0 + c_1(\sigma_{net(i)} + \theta_1) + c_2(\sigma_{net(2)} + \theta_2)^2 + \dots + c_d(\sigma_{net(d)} + \theta_d)^d \} \quad \text{eq. 3.23}$$

na qual o vetor $\mathbf{c} = [c_1, c_2, \dots, c_d]$ representa os pesos das unidades escondidas para as unidades de saída e a constante c_0 está representando todos os blocos congelados até a ordem $d-1$.

A equação 3.20 é, portanto, idêntica a uma regressão logística não-linear quando a função de ativação utilizada é a sigmóide logística tradicional.

A potencialidade dessas redes está fundamentada nas características impostas pela regressão sigmóide não-linear na unidade de saída.

O algoritmo construtivo para as NSRBN tem como base um método de aprendizagem construtiva (também chamado de auto-organização na teoria das redes GMDH). O princípio construtivo pode ser formulado como segue: quando um modelo incrementa gradualmente sua complexidade, certos critérios, que são chamados de critérios de seleção ou função objetivo e que têm a propriedade de complemento externo, passam através de um valor mínimo global. O encontro deste mínimo indicará a existência de um modelo de complexidade ótima.

Portanto, usando-se essa arquitetura proposta uma função qualquer $f(x)$ pode ser aproximada pelo uso direto de uma rede NSRBN por dividir $f(x)$ em blocos homogêneos de ordem 1 até o bloco de ordem d . Dessa forma, a função total $f(x)$ (polinômio completo) pode ser aproximada pela soma das funções parciais $f_p(x)$, $p=1, \dots, d$ (polinômios homogêneos), de maneira incremental podendo ser representada matematicamente por:

$$f(x) = g(\hat{f}_1(x) + \hat{f}_2(x) + \dots + \hat{f}_p(x) + \dots + f_d(x)) \quad \text{eq. 3.24}$$

O modelo proposto tem como princípio manter congelados os pesos de todos os blocos anteriores ao atualmente otimizado, de tal forma que o número de pesos a ser otimizado a cada grau de complexidade não aumente explosivamente, como no algoritmo combinatorial tradicional. O algoritmo de abordagem construtiva procede como se descreve a seguir.

Seja p um passo do algoritmo para o qual o bloco $f_p(x)$ é adicionado à rede, então uma função qualquer $f(x)$ é sucessivamente aproximada por:

$$f(x) = g(f_1(x)); \quad \text{eq. 3.25}$$

$$f(x) = g(\hat{f}_1(x) + f_2(x)); \quad \text{eq. 3.26}$$

$$f(x) = g(\hat{f}_1(x) + \hat{f}_2(x) + \dots + f_p(x)) \quad \text{eq. 3.27}$$

em que os pesos dos blocos $\hat{f}_{p-1}(x)$ são congelados quando o bloco de ordem p é adicionado à rede e g representa uma função de ativação adequada para a unidade de saída.

3.2.4 Softwares Usados

Abaixo são descritos os softwares especializados utilizados nessa dissertação.

- Matlab: foi usado o Toolbox Redes Neurais desse software nas simulações da rede MLP;
- NeuroHidro: foi usado nas simulações da rede NSRBN;
- ArcGis: foi usado nas análises da Geoestatística, para as informações médias anuais, e na determinação do polígono de Thiessen;
- Surfer: foi usado na interpolação de Kriging dos dados diários de precipitação;
- Autocad: foi usado na obtenção dos mapas e na delimitação e cálculo da área da chuva média por Thiessen;
- Sirius: foi usado nas determinações da matriz de correlação, dendograma e na análise de componentes principais.

Capítulo 4

Aplicações

4.1 ATUAIS ERROS DE PREVISÕES AO RESERVATÓRIO DE TRÊS MARIAS

O ONS mensalmente realiza o Programa Mensal de Operação – PMO, com revisões semanais. Esse PMO é elaborado nas terças-feiras e as revisões semanais às quartas-feiras e compreende sempre informações médias semanais operativas de sábado a sexta, conforme mostra a figura 4.1.

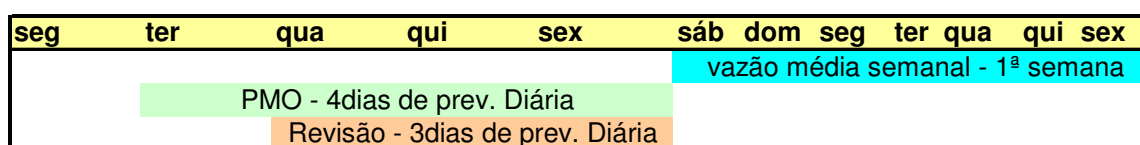


Figura 4.1 – Esquema com os dias previstos no PMO e suas revisões.

Como citado no capítulo 1, essas previsões semanais são obtidas pelo modelo Previvaz.

Para completar a semana operativa (terça ou quarta à sexta) necessita-se de vazões diárias que são solicitadas ao Agente de Geração responsável pelo aproveitamento, no caso do reservatório de Três Marias a CEMIG, ou obtida através do modelo PrevivazH. Nas tabelas 4.1 e 4.2 são apresentados os atuais erros de previsão de vazões naturais diárias e semanais do reservatório de Três Marias.

Tabela 4.1 – Erros das previsões de vazões diárias ao reservatório de Três Marias no PMO.

Ano	EPMA (%)	
	Cemig	PrevivazH
2001	18,1	46,9
2002	21,5	50,1
2003	16,7	27,1
Média:	18,8	41,4

Tabela 4.2 – Erros das previsões de vazões semanais ao reservatório de Três Marias no PMO – 1ª semana operativa (sábado a sexta).

Ano	EPMA (%)		
	úmido	seco	Média
1999	-	-	34,2
2000	-	-	36,8
2001	-	-	44,7
2002	44,0	40,2	42,1
2003	40,6	22,4	31,5
2004	35,2	10,8	23,0
Média:	40,0	24,5	35,4

4.2 BASE DE DADOS

Neste trabalho foram usadas informações de estações pluviométricas, fluviométricas e de vazão natural do reservatório de Três Marias. Todos os dados pluviométricos e fluviométricos utilizados foram da base de dados da CEMIG, com complementação de dados fluviométricos da ANEEL e de dados pluviométricos da ANA. A vazão natural do reservatório de Três Marias corresponde à existente na base de dados do ONS.

Na vazão natural procura-se retirar a interferência do homem na natureza. Assim, a vazão natural é considerada como as vazões ocorridas, obtidas através do balanço hídrico do reservatório, mais evaporação no reservatório, e mais vazões consuntivas da área de drenagem do reservatório.

A vazão natural de Três Marias do período 1987 a 2001 correspondeu à vazão obtida no projeto de revisão da série de vazão natural (ONS, 2003), o qual teve como metodologia o balanço hídrico do reservatório, a inclusão da vazão evaporada no reservatório e do uso consuntivo da área de montante, e desagregação das médias mensais afluentes, modulando com a soma dos postos de montante (Porto Pará, Porto Mesquita, Porto das Andorinhas e Porto Indaiá). O período de 2002 a 2004 correspondeu às informações obtidas na reconstituição da série de 2002/2003 e de 2004, tendo a afluência sido determinada pelo balanço hídrico do reservatório, e a vazão natural obtida através de modulação, com uma média móvel centrada de 5 dias até o limite de 1300 m³/s.

As aplicações realizadas neste trabalho utilizaram informações pluviométricas e fluviométricas de parte da bacia do Alto São Francisco, tendo

sido de extrema relevância a obtenção destas informações básicas junto CEMIG, ANA e ANNEL.

4.2.1 Informações pluviométricas

A seguir são mostrados os postos pluviométricos usados, período disponível dos seus históricos, processos utilizados no preenchimento de falhas e na análise Geoestatística efetuadas.

4.2.1.1 Rede de postos pluviométricos pesquisada

Foi realizado um inventário das informações pluviométricas objetivando proceder a seleção de postos confiáveis e que possibilitasse a execução de estudos. A CEMIG dispõe de uma rede telemétrica com 40 postos telemedidos, os quais compõem a maior parte da rede pluviométrica utilizada neste trabalho.

Foi fornecido pela superintendência de planejamento da operação de sistemas da geração e transmissão da CEMIG o histórico disponível de totais pluviométricos diários dos postos atualmente em operação na bacia do São Francisco. A tabela 4.3 mostra a relação desses postos, juntamente com o período do histórico disponível das informações.

A seleção dos postos pluviométricos e análise das informações recebidas focaram dois pontos: postos pluviométricos que estão atualmente em operação e com histórico sem muitas falhas. De início foram selecionados 19 postos pluviométricos, posteriormente, foram excluídos os postos Fazenda Macarujá e UHE Salto Paraopeba por apresentarem um histórico com curto período de observação. Os postos selecionados estão apresentados na figura 4.2(a).

Analisando a figura 4.2 (a), observa-se que a área próxima ao reservatório não está sendo coberta pelos postos selecionados da CEMIG, então foi feita complementação dessa rede adicionando-se os postos de Porto Pará e Porto Mesquita, os quais terão seu histórico composto com informações da CEMIG (até 1999) e da ANEEL (a partir de 2000). A figura 4.2 (b) mostra a disposição de todos os postos pluviométricos selecionados, destacando os postos de Porto Pará e Porto Mesquita .

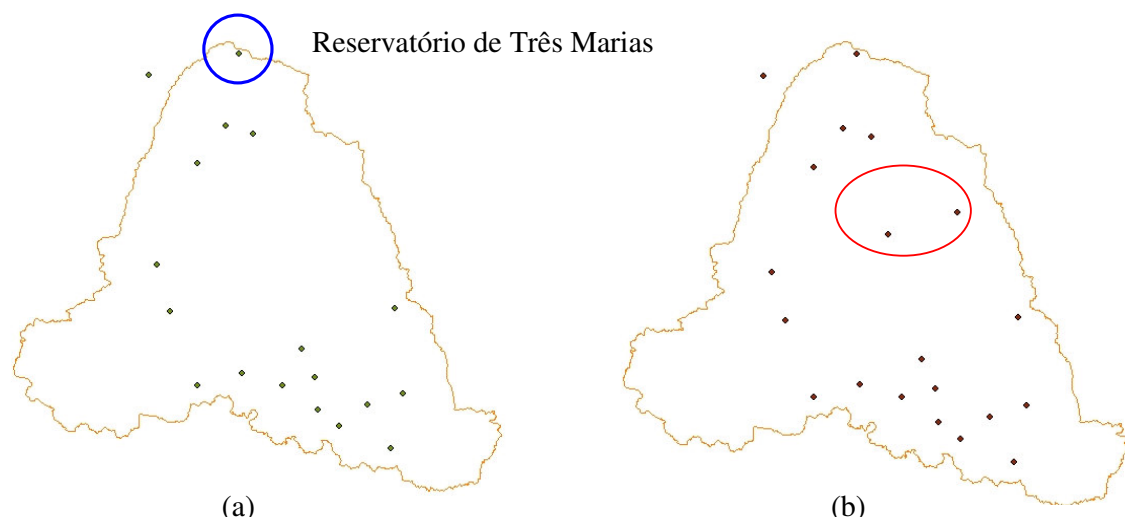


Figura 4.2 - Postos pluviométricos da selecionados. a) só os postos selecionados da CEMIG. b) postos selecionados da CEMIG, com a adição dos postos de Porto Pará e Porto Mesquita.

4.2.1.2 Preenchimento de falhas e análise de consistência das informações

Devido à maior quantidade de informações pluviométricas disponíveis, trabalhou-se com informações do período de janeiro/1987 a dezembro/2004, onde essas informações foram submetidas a um preenchimento de falhas, quando necessário, e uma análise de consistência.

O preenchimento de falhas foi feito para informações diárias e pelo método da ponderação regional, com os três postos de maior correlação. Vale salientar que em algumas poucas ocasiões não foi possível usar estes três postos na ponderação regional porque as falhas coincidiram, então, usaram-se os postos disponíveis para esse preenchimento.

A análise de consistência foi feita a nível mensal e através do método do vetor regional.

4.2.1.3 Análise geoestatística

Objetivando verificar a variabilidade espacial da precipitação nos postos selecionados foi feita uma análise geoestatística. Essa análise contemplou informações médias mensais, no período de 1987 a 2004, agrupadas das seguintes formas: período úmido (novembro a abril), período seco (maio a outubro), total anual e total mensal. Nessa análise excluiu-se o posto de São Gonçalo do Abaeté por estar localizado fora da área de drenagem da UHE Três Marias (conforme mostrado anteriormente nas figuras 4.2(a) e (b)). A base cartográfica utilizada foi resultado do Projeto ONS/THEMAG (ONS, 2003).

Foram ajustados os parâmetros do variograma teórico a partir de um variograma experimental, considerado neste estudo isotrópico, para os modelos Esférico, Exponencial e Gaussiano, e na interpolação usou-se a Krigagem Ordinária.

4.2.2 Informações fluviométricas

A seguir são mostrados os postos fluviométricos usados, período disponível dos seus históricos e metodologia usada no preenchimento de falhas.

4.2.2.1 Rede de postos fluviométricos pesquisada

Assim como o item 4.2.1, primeiramente foram usados os postos fluviométricos da CEMIG atualmente em operação na área de drenagem do reservatório de Três Marias. A tabela 4.4 apresenta esses postos, juntamente com o período do histórico disponível. A figura 4.3 ilustra a distribuição espacial desses postos e a figura 4.4 ilustra as áreas de drenagem existentes a montante do reservatório de Três Marias.



Figura 4.3 – Disposição dos postos fluviométricos em operação da CEMIG localizados na área de drenagem de Três Marias.

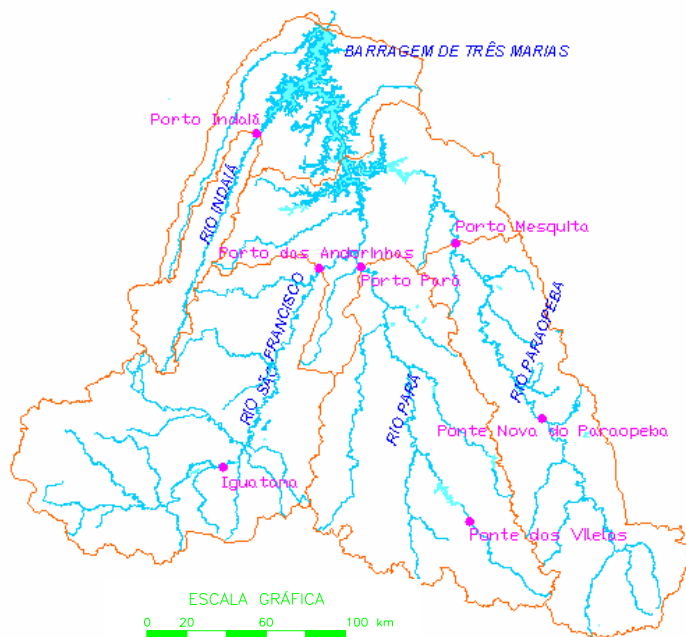


Figura 4.4 – Sub-bacias da área de drenagem de Três Marias.

4.2.2.2 Análise de correlação e preenchimento de falhas

Foi feita uma análise do coeficiente de correlação dos postos fluviométricos selecionados.

O período de 01/jan/1980 a 31/12/2004 corresponde a 6574 dias. Os postos de Porto das Andorinhas e Porto Mesquita não possuem nenhuma falha nesse período, o Porto Indaiá teve 7 dias com falhas (0,1%), Porto Pará teve 5 dias com falhas (0,1%), Iguatama teve 2660 dias com falhas (40,5%), Ponte Nova do Paraopeba teve 40 dias com falhas (0,6%) e Ponte dos Vilelas Jusante teve 2271 dias com falhas (34,5%). O preenchimento de falhas dessas séries foi feito a partir de duas metodologias:

1. A partir do posto mais correlacionado, usando uma regressão linear simples ou regressão potencial;

2. A partir de uma correlação múltipla, sendo que, para os postos de Porto Pará, Porto Mesquita, Porto Indaiá, Porto das Andorinhas e Ponte Nova do Paraopeba não foram consideradas as informações de Iguatama e Ponte dos Vilelas Jusante, devido à quantidade de falhas desses postos, o que limitaria o histórico analisado (já que nesta análise o dia em que há falhas em qualquer dos postos é desconsiderado) e poderia prejudicar o processo de preenchimento das falhas. Na análise dos postos de Iguatama e Ponte dos Vilelas Jusante foram consideradas as informações de todos os postos fluviométricos disponíveis. No preenchimento de falhas do posto de Ponte dos Vilelas Jusante foram consideradas as informações de Ponte Nova do Paraopeba já com suas falhas preenchidas, visto que, usando os dados falhos deste posto foram obtidos diversos valores de vazão negativa.

4.2.3 Análise Estatística

Na análise estatística, trabalhou-se com vazão média semanal (sábado a sexta) e precipitação média total semanal (sábado a sexta).

A precipitação média da área foi calculada através da metodologia do polígono de Thiessen e, conforme mostrado no item 4.2.1.3, pelo método de Kriging.

Depois de determinado o polígono de Thiessen, usando o software Arcgis, obteve-se um coeficiente proporcional à área de influência de cada posto pluviométrico, usando o software Autocad. Esses coeficientes serviram para determinar a precipitação média diária da área, que foi posteriormente usada para encontrar a precipitação total semanal. A figura 4.5 ilustra o polígono de Thiessen para os postos selecionados (CEMIG + ANA).

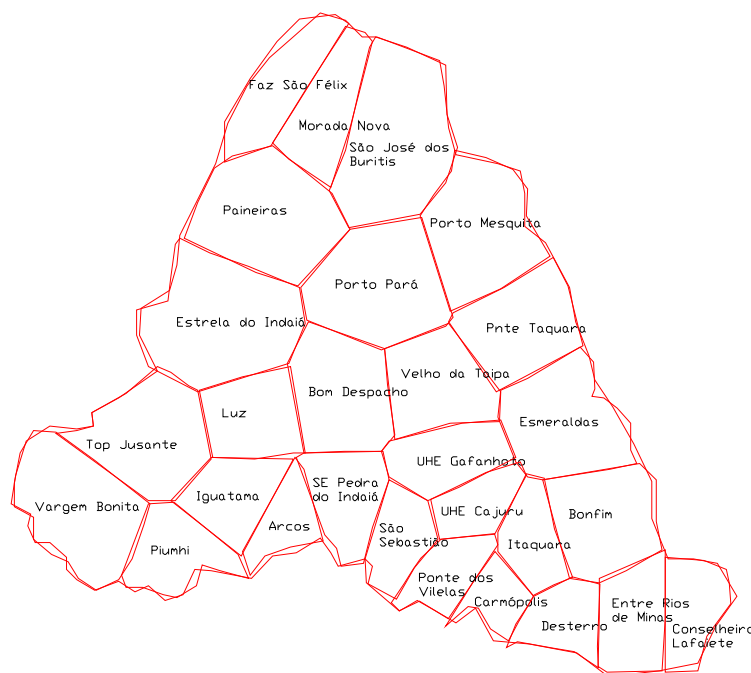


Figura 4.5 – Polígono de Thiessen dos postos pluviométricos selecionados da ANA e CEMIG.

4.2.3.1 Função de auto-correlação e correlação cruzada

Através de programas elaborados em planilhas de Excel, foi feita a determinação da defasagem temporal através do cálculo da função de auto-

correlação amostral (FAC) e correlação cruzada, e plotagem com o Lag (defasagem).

A função de auto-correlação e correlação cruzada é útil em algumas situações porque fornece uma visão de como a dependência da série cai com o 'lag' ou separação k entre pontos da série.

- Auto-correlação: A auto-correlação expressa a independência entre as variáveis. As informações diárias foram usadas na determinação da autocorrelação da vazão natural de Três Marias, dos postos fluviométricos e da precipitação média da área.

- Correlação Cruzada: As informações diárias foram usadas na determinação da correlação cruzada da vazão natural de Três Marias com os postos fluviométricos e com a precipitação média da área

4.2.3.2 Matriz de correlação, dendograma (agrupamento), e análise de componentes principais (ACP)

A determinação da matriz de correlação, dendograma (agrupamento), e análise de componentes principais (ACP) foi efetuada através da versão demonstrativa do software Sirius, desenvolvido pela empresa Pattern Recognition Systems (PRS).

Foi calculado a matriz de correlação, o dendograma e feita uma ACP para as informações fluviométricas, com o preenchimento de falhas através da regressão linear simples e múltipla, e para as informações pluviométricas, com a chuva média calculada por Thiessen e Kriging.

A tabela 4.5 mostra o significado das principais variáveis utilizadas neste estudo.

Tabela 4.5 – Significado das principais variáveis utilizadas no estudo.

Variável	Significado
Ind	Vazão do posto fluviométrico de Porto Indaiá
Par	Vazão do posto fluviométrico de Porto Pará
Mqt	Vazão do posto fluviométrico de Porto Mesquita
And	Vazão do posto fluviométrico de Porto das Andorinhas
Vil	Vazão do posto fluviométrico de Ponte dos Vilelas Jusante

Igu	Vazão do posto fluviométrico de Iguatama
Ppb	Vazão do posto fluviométrico Ponte Nova do Paraopeba
Nat	Vazão natural de Três Marias
Pthiessen2 PrecT2	Chuva média calculada por Thiessen com os postos da CEMIG e da ANA
PrecT1	Chuva média calculada por Thiessen com os postos da CEMIG
PrecKri	Chuva média calculada por Kriging

A tabela 4.6 mostra a média e o desvio padrão dos dados semanais utilizados (preenchimento de falhas com regressão múltipla e chuva média por Thiessen, com os postos da CEMIG e da ANA).

Tabela 4.6 – Média e desvio padrão dos dados semanais usados.

Variável	Média	Desvio Padrão
Ind	34	51
Par	129	115
Mqt	121	123
And	215	218
Vil	25	27
Igu	145	144
Ppb	77	74
Nat	608	610
Pthiessen2	22	30

4.2.4 Redes Neurais

A previsão de vazões médias diárias ao reservatório de Três Marias, 12 dias à frente, foi feita através de duas metodologias. Na primeira usou-se uma rede MLP através do Software Matlab, Toolbox Redes Neurais. Na segunda, testou-se uma rede NSRBN em uma versão modificada do software Neurot 58 disponível em Valença (2005).

Como dados de entrada para as redes neurais, foram consideradas as informações dos postos pluviométricos, fluviométricos e a vazão natural ocorrida

do reservatório de Três Marias, com diversas defasagens. Abaixo segue a relação dessas variáveis com a respectiva sigla adotada nesse trabalho:

- Postos pluviométricos: Porto das Andorinhas (And), Porto Indaiá (Ind), Porto Mesquita (Mqt), Porto Pará (Par), Iguatama (Igu), Ponte Nova do Paraopeba (Ppb), Ponte dos Vilelas Jusante (Vil);
- Postos pluviométricos: usados na análise geoestatística e citados no capítulo 4. Com os postos da Cemig e da ANA (Th2) calculou-se a chuva média da área, pela metodologia de Thiessen;
- Vazão natural do reservatório de Três Marias ocorrida (Nat).

No processo de treinamento foi usado o método da Validação Cruzada, o qual utiliza um conjunto de dados independente para determinar o ponto de parada ótima durante o treinamento de forma a minimizar especialmente os riscos de superajustamento ou subajustamento. Desta forma o conjunto de dados foi dividido em três sub-conjuntos:

- treinamento: padrões usados para modificar os pesos;
- validação: padrões usados para verificar principalmente o problema de superajustamento (overfitting);
- teste: padrões para testar o desempenho do modelo final.

Nas duas redes treinadas o período de 1987 a 1995 foi usado para treinamento da rede, 1996 a 1999 para validação e 2000 a 2004 para testes.

Os resultados obtidos consideraram o EPMA (Erro Percentual Médio Absoluto), EP (Erro Padrão de Predição) e EMA (Erro Médio Absoluto).

$$EPMA = \frac{1}{N} \left[\sum_{p=1}^N \frac{|Z_p - Z_o|}{Z_o} \right] * 100$$

$$EP = \frac{1}{N} \left[\sum_{p=1}^N (Z_p - Z_o)^2 \right]^{0,5}$$

$$EMA = \frac{1}{N} \left[\sum_{p=1}^N |Z_p - Z_o| \right]$$

Onde: Z_p representa a variável prevista, Z_o a variável observada e N o número total da amostra.

O erro padrão tende a valorizar os erros de grande intensidade.

A seguir são descritas as especificações de cada metodologia.

4.2.4.1 Rede MLP (Multi Layer Perceptron)

Foi usada uma rede MLP com uma camada de entrada, uma intermediária e uma camada de saída, treinadas com algoritmo de aprendizado supervisionado Levenberg-Marquardt (LM).

Como função de transferência foram usadas as funções *tansig* (tansigmoid), na camada intermediária e *purelin* (linear), na camada de saída.

Os dados de entrada e os de saída (vazão natural) foram normalizados, de forma a terem média zero e desvio padrão unitário e, através de uma ACP (Análise de Componentes Principais), foram desconsideradas as informações que contribuiriam com menos de 0,1% da variação total dos dados.

Foram usados os seguintes parâmetros no treinamento da rede:

- `net.trainParam.epochs = 100` → Número máximo de épocas no treinamento;
- `net.trainParam.goal = 0` → Performance meta;
- `net.trainParam.mu = 0.005` → Parâmetro ajustado Marquardt (μ);
- `net.trainParam.mu_dec = 0.1` → Fator de decrescimento para μ ;
- `net.trainParam.mu_inc = 10` → Fator de crescimento para μ ;
- `net.trainParam.mu_max = 1e-10` → Máximo valor para μ ;
- `net.trainParam.max_fail = 5` → Validação máxima de falhas;
- `net.trainParam.mem_reduc 1` → Fator para usar na memória/velocidade trade-off;
- `net.trainParam.min_grad = 1e-10` → Gradiente Mínimo;
- `net.trainParam.show = 25` → N^o de épocas entre o progresso mostrado;
- `net.trainParam.time = inf` → Tempo máximo para treinar, em segundos;

Levando em consideração os parâmetros mostrados anteriormente, o modelo foi configurado para interromper o treinamento quando uma 60 condições abaixo fosse satisfeita:

- O número máximo de épocas (repetições) fosse alcançado;
- O tempo máximo de simulação fosse excedido;

- A performance fosse minimizada até a meta;
- O gradiente performance ficasse abaixo do gradiente mínimo;
- μ excedesse μ_{max} ;
- A performance da validação tenha aumentado mais que max_{fail} (validação máxima de falhas) desde que o último tempo tenha decrescido (quando usando validação).

Foram ajustadas duas redes neurais: uma para o período úmido (novembro a abril) e outra para o período seco (maio a outubro).

A tabela 4.7 mostra as variáveis de entrada simuladas no período úmido, onde t corresponde à variável no dia da previsão, t-1 com um dia de defasagem, t-2 com dois dias de defasagem, assim por diante.

Tabela 4.7 – Variáveis de entrada consideradas na rede neural MLP do período úmido.

período	n° de entradas	caso	Variáveis de entrada																																				
per_úmido	35	a	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t-3	And_t-2	And_t-1	And_t	Ind_t-3	Ind_t-2	Ind_t-1	Ind_t	Par_t-1	Par_t	Igu_t-3	Igu_t-2	Igu_t-1	Igu_t	Ppb_t-3	Ppb_t-2	Ppb_t-1	Ppb_t	Vil_t-3	Vil_t-2	Vil_t-1	Vil_t	Mqt_t-3	Mqt_t-2	Mqt_t-1	Mqt_t	Par_t-3	Par_t-2	Nat-3	Nat-2	Nat-1		
	19	b	Th2_t-3	Th2_t	And_t-3	And_t	Ind_t-3	Ind_t	Mqt_t-3	Mqt_t	Par_t-3	Par_t	Igu_t-3	Igu_t	Ppb_t-3	Ppb_t	Vil_t-3	Vil_t	Nat-3	Nat-2	Nat-1																		
	11	c	Th2_t-3	And_t-3	Ind_t-3	Mqt_t-3	Par_t-3	Igu_t-3	Ppb_t-3	Vil_t-3	Nat-3	Nat-2	Nat-1																										
	14	d	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t-3	Ind_t-3	Mqt_t-3	Par_t-3	Igu_t-3	Ppb_t-3	Vil_t-3	Nat-3	Nat-2	Nat-1																							
	12	e	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t-3	Ind_t-3	Mqt_t-3	Par_t-3	Igu_t-3	Ppb_t-3	Vil_t-3	Nat-3	Nat-2	Nat-1																							
	9	f	Th2_t	And_t	Ind_t	Mqt_t	Par_t	Igu_t	Ppb_t	Vil_t	Nat-1																												
	12	g	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t	Ind_t	Mqt_t	Par_t	Igu_t	Ppb_t	Vil_t	Nat-1																									
	8	h	And_t	Ind_t	Mqt_t	Par_t	Igu_t	Ppb_t	Vil_t	Nat-1																													
	3	i	Nat-3	Nat-2	Nat-1																																		

A tabela 4.8 mostra as variáveis de entrada simuladas no período seco, onde t corresponde à variável no dia da previsão, t-1 com um dia de defasagem, t-2 com dois dias de defasagem, assim por diante.

Para cada caso, tanto no período úmido quanto no seco, foram feitas diversas simulações para definir o número de neurônios da camada intermediária. Essas simulações iniciavam com 2 neurônios e paravam quando o EPMA, após duas rodadas consecutivas, aumentavam.

Tabela 4.8 – Variáveis de entrada consideradas na rede neural MLP do período seco.

período	n° de entradas	caso	Variáveis de entrada																																		
per_seco	35	a	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t-3	And_t-2	And_t-1	And_t	Ind_t-3	Ind_t-2	Ind_t-1	Ind_t	Par_t-1	Par_t	Igu_t-3	Igu_t-2	Igu_t-1	Igu_t	Ppb_t-3	Ppb_t-2	Ppb_t-1	Ppb_t	Vil_t-3	Vil_t-2	Vil_t-1	Vil_t	Mqt_t-3	Mqt_t-2	Mqt_t-1	Mqt_t	Par_t-3	Par_t-2	Nat-3	Nat-2	Nat-1
	12	b	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t-3	Ind_t-3	Mqt_t-3	Par_t-3	Igu_t-3	Ppb_t-3	Vil_t-3	Nat-1																							
	21	c	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t-1	And_t	Ind_t-3	Ind_t-2	Ind_t-1	Ind_t	Mqt_t-1	Mqt_t	Par_t-1	Par_t	Igu_t-3	Igu_t-2	Ppb_t-3	Ppb_t-2	Vil_t-3	Vil_t-2	Nat-1														
	23	d	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t-1	And_t	Ind_t-3	Ind_t-2	Ind_t-1	Ind_t	Mqt_t-1	Mqt_t	Par_t-1	Par_t	Igu_t-3	Igu_t-2	Ppb_t-3	Ppb_t-2	Vil_t-3	Vil_t-2	Nat-3	Nat-2													
	17	e	And_t-1	And_t	Ind_t-3	Ind_t-2	Ind_t-1	Ind_t	Mqt_t-1	Mqt_t	Par_t-1	Par_t	Igu_t-3																								
	3	f	Nat-3	Nat-2	Nat-1																																

Na inicialização dos pesos e bias foi usado o método de Nguyen-Widrow. Com isso, não foi dada a semente para inicialização dos pesos e bias, conseqüentemente, cada vez que se executava o modelo, para os mesmos dados de entrada e objetivo, obteve-se uma diferente resposta da rede neural.

Assim, em cada simulação, definido pelo número de entradas e número de neurônios na camada intermediária, o modelo foi executado 5 vezes, sendo o modelo escolhido o que teve melhor EPMA para os dados de verificação, sempre observando os correspondentes erros dos dados de teste. As figuras 4.6 e 4.7 mostram exemplos desses resultados, para os períodos úmido e seco, respectivamente, onde se observa que é pequena a variação de uma simulação para a outra, tanto nos dados de validação quanto nos dados de teste.

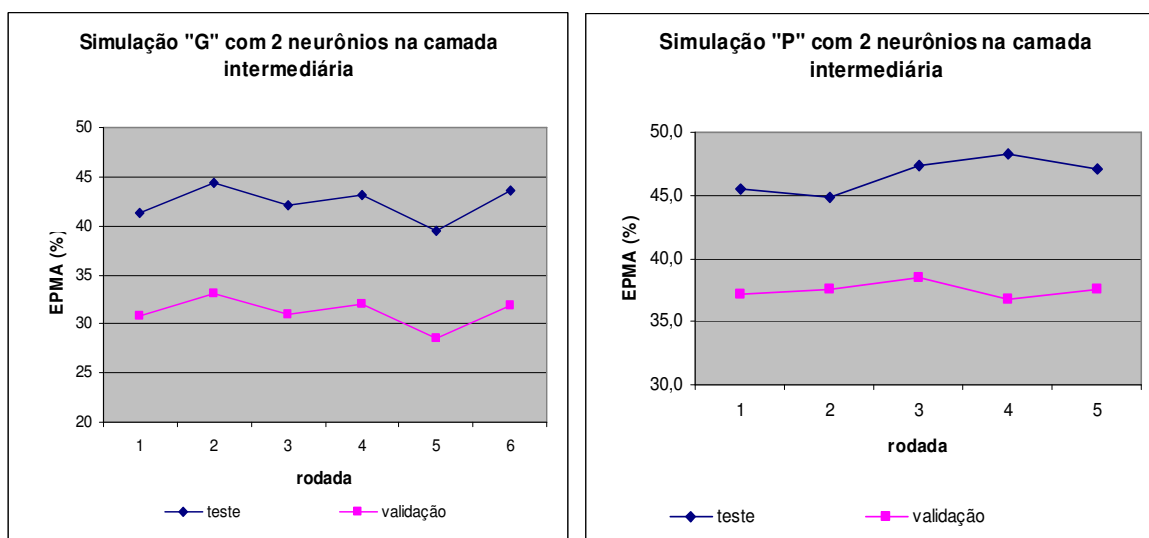


Figura 4.6 – Exemplos dos EPMA obtidos com 5 diferentes rodadas – período úmido.

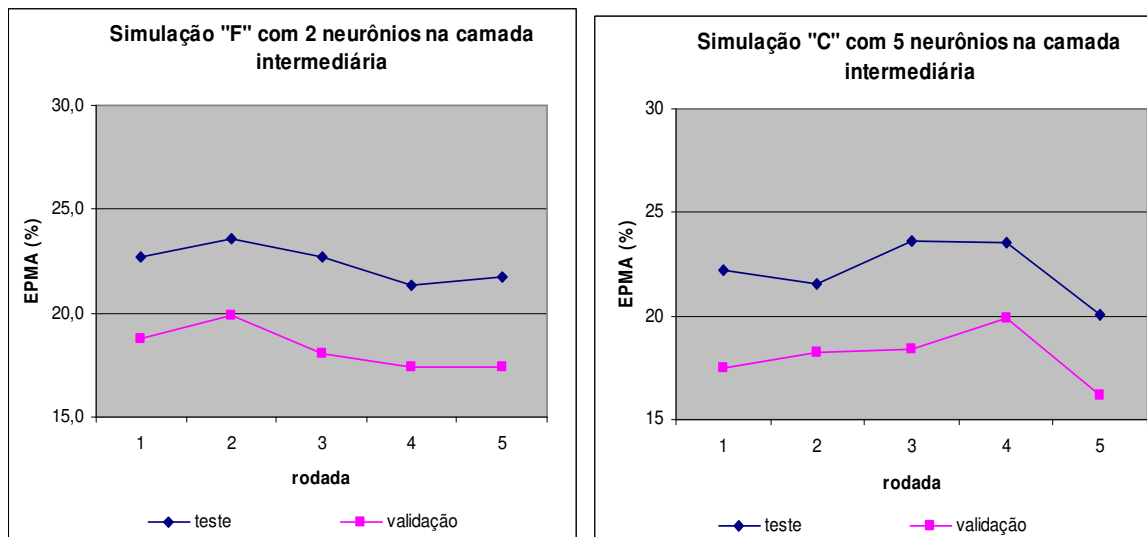


Figura 4.7 – Exemplos dos EPMA obtidos com 5 diferentes rodadas – período seco.

4.2.4.2 Rede NSRBN

Foram estudados diversos casos variando as variáveis de entrada, os parâmetros do algoritmo (taxa de aprendizagem, semente para geração e momentum) e os critérios de parada (época mínima e época máxima).

Seguindo recomendação do manual do NeuroHidro, para reduzir a variabilidade das informações e dar a mesma importância a todas, os dados de entrada da vazão natural foram logaritmadas.

Como no caso da rede MLP, foram ajustadas duas redes neurais: uma para o período úmido da área em estudo (novembro a abril) e outra para o período seco (maio a outubro).

No período úmido, foram ajustadas duas redes neurais: uma para os 7 primeiros dias e outra para os últimos 5 dias, totalizando os 12 dias de previsão. A rede ótima foi tomada como a composição das duas redes. A tabela 4.9 mostra as variáveis de entrada nessas simulações.

Tabela 4.9 – Variáveis de entrada consideradas nas redes neurais NSRBN do período úmido.

período	n° de entradas	caso	Variáveis de entrada										
per_úmido	35	a	Th2_t-3 Ind_t Vil_t-3 Nat-2	Th2_t-2 Par_t-1 Vil_t-2 Nat-1	Th2_t-1 Par_t Vil_t-1	Th2_t Igu_t-3 Vil_t	And_t-3 Igu_t-2 Mqt_t-3	And_t-2 Igu_t-1 Mqt_t-2	And_t-1 Igu_t Mqt_t-1	And_t Ppb_t-3 Mqt_t	Ind_t-3 Ppb_t-2 Par_t-3	Ind_t-2 Ppb_t-1 Par_t-2	Ind_t-1 Ppb_t Nat-3
	19	b	Th2_t-3 Igu_t	Th2_t Ppb_t-3	And_t-3 Ppb_t	And_t Vil_t-3	Ind_t-3 Vil_t	Ind_t Nat-3	Mqt_t-3 Nat-2	Mqt_t Nat-1	Par_t-3 Par_t	Par_t Igu_t-3	Igu_t-3
	11	c	Th2_t-3 Nat-3	And_t-3 Nat-2	Ind_t-3 Nat-1	Mqt_t-3 Th2_t	Par_t-3 And_t-3	Igu_t-3 Ind_t-3	Ppb_t-3 Mqt_t-3	Vil_t-3 Par_t-3	Nat-3 Igu_t-3	Nat-2 Ppb_t-3	Nat-1 Vil_t-3
	14	d	Th2_t-3 Nat-3	Th2_t-2 Nat-2	Th2_t-1 Nat-1	Th2_t	And_t-3	Ind_t-3	Mqt_t-3	Par_t-3	Igu_t-3	Ppb_t-3	Vil_t-3
	12	e	Th2_t-3 Nat-1	Th2_t-2	Th2_t-1	Th2_t	And_t-3	Ind_t-3	Mqt_t-3	Par_t-3	Igu_t-3	Ppb_t-3	Vil_t-3
	21	f	Th2_t-3 Par_t	Th2_t-2 Igu_t-3	Th2_t-1 Igu_t-2	Th2_t Ppb_t-3	And_t-1 Ppb_t-2	And_t Vil_t-3	Ind_t-3 Vil_t-2	Ind_t-2 Mqt_t-1	Ind_t-1 Mqt_t	Ind_t Nat-1	Par_t-1
	9	g	Th2_t	And_t	Ind_t	Mqt_t	Par_t	Igu_t	Ppb_t	Vil_t	Nat-1		
	12	h	Th2_t-3 Nat-1	Th2_t-2	Th2_t-1	Th2_t	And_t	Ind_t	Mqt_t	Par_t	Igu_t	Ppb_t	Vil_t
	21	i	Th2_t-3 Ind_t	Th2_t-2 Ppb_t-3	Th2_t-1 Ppb_t-2	Th2_t Ppb_t-1	And_t-3 Ppb_t	And_t-2 Vil_t-3	And_t-1 Vil_t-2	And_t Vil_t-1	Ind_t-3 Vil_t	Ind_t-2 Nat-1	Ind_t-1
	27	j	Th2_t-3 Ind_t Par_t-3	Th2_t-2 Par_t-1 Par_t-2	Th2_t-1 Par_t Nat-3	Th2_t Ppb_t-3 Nat-2	And_t-3 Ppb_t-2 Nat-1	And_t-2 Ppb_t-1	And_t-1 Ppb_t	And_t Mqt_t-3	Ind_t-3 Mqt_t-2	Ind_t-2 Mqt_t-1	Ind_t-1 Mqt_t
	14	l	Th2_t-3 Ppb_t-3	Th2_t-4 Vil_t-3	Th2_t-3 Nat-1	Th2_t-2	Th2_t-1	Th2_t	And_t-3	Ind_t-3	Mqt_t-3	Par_t-3	Igu_t-3
	18	m	Th2_t-3 Ppb_t-3	Th2_t-4 Vil_t-3	Th2_t-3 Nat-5	Th2_t-2 Nat-4	Th2_t-1 Nat-3	Th2_t Nat-2	And_t-3 Nat-1	Ind_t-3	Mqt_t-3	Par_t-3	Igu_t-3
	6	n	Th2_t	And_t	Ind_t	Ppb_t	Vil_t	Nat-1					
6	o	Th2_t	And_t	Ind_t	Mqt_t	Par_t	Nat-1						

Para o período seco, também foram ajustadas duas redes neurais: uma para os 7 primeiros dias e outra para os últimos 5 dias, totalizando os 12 dias de previsão. A rede ótima foi tomada como a composição das duas redes. A tabela 4.10 mostra as variáveis de entrada nessas simulações.

Tabela 4.10 – Variáveis de entrada consideradas nas redes neurais NSRBN do período seco.

período	n° de entradas	caso	Variáveis de entrada										
per_seco	35	a	Th2_t-3 Ind_t Vil_t-3 Nat-2	Th2_t-2 Par_t-1 Vil_t-2 Nat-1	Th2_t-1 Par_t Vil_t-1	Th2_t Igu_t-3 Vil_t	And_t-3 Igu_t-2 Mqt_t-3	And_t-2 Igu_t-1 Mqt_t-2	And_t-1 Igu_t Mqt_t-1	And_t Ppb_t-3 Mqt_t	Ind_t-3 Ppb_t-2 Par_t-3	Ind_t-2 Ppb_t-1 Par_t-2	Ind_t-1 Ppb_t Nat-3
	12	b	Th2_t-3 Nat-1	Th2_t-2	Th2_t-1	Th2_t	And_t-3	Ind_t-3	Mqt_t-3	Par_t-3	Igu_t-3	Ppb_t-3	Vil_t-3
	21	c	Th2_t-3 Mqt_t	Th2_t-2 Par_t-1	Th2_t-1 Par_t	Th2_t Igu_t-3	And_t-1 Igu_t-2	And_t Ppb_t-3	Ind_t-3 Ppb_t-2	Ind_t-2 Vil_t-3	Ind_t-1 Vil_t-2	Ind_t Nat-1	Mqt_t-1
	23	d	Th2_t-3 Mqt_t	Th2_t-2 Par_t-1	Th2_t-1 Par_t	Th2_t Igu_t-3	And_t-1 Igu_t-2	And_t Ppb_t-3	Ind_t-3 Ppb_t-2	Ind_t-2 Vil_t-3	Ind_t-1 Vil_t-2	Ind_t Nat-3	Mqt_t-1 Nat-2
	19	e	Th2_t-3 Par_t	Th2_t Igu_t-3	And_t-1 Igu_t-2	And_t Ppb_t-3	Ind_t-3 Ppb_t-2	Ind_t-2 Vil_t-3	Ind_t-1 Vil_t-2	Ind_t Nat-1	Mqt_t-1	Mqt_t	Par_t-1
	9	f	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t-3	Ind_t-3	Ppb_t-3	Vil_t-3	Nat-1		
	9	g	Th2_t-3	Th2_t-2	Th2_t-1	Th2_t	And_t-3	Ind_t-3	Mqt_t-3	Par_t-3	Nat-1		

Capítulo 5

Resultados e Discussões

Em todos os processos de modelagem, a familiaridade com os dados disponíveis é de máxima importância. Os modelos de redes neurais não são exceção, e o pré-processamento das informações pode ter um efeito significativo no desempenho do modelo. Assim, a seguir são mostrados os resultados das análises dos dados, que servirão de entrada para a rede neural, bem como os resultados das simulações.

5.1 ANÁLISE DOS DADOS

5.1.1 Informações pluviométricas

A tabela 5.1 mostra o coeficiente de correlação das precipitações diárias dos postos pluviométricos selecionados, destacando (em azul) os três postos mais bem correlacionados com o posto em análise. Observa-se que, com exceção do posto de Arcos, todos os postos têm correlação acima de 0,50, o que foi considerado suficiente neste trabalho.

O Anexo 1 mostra a análise de consistência dos dados mensais de precipitação, através do vetor regional e de gráficos dupla-massa, observando-se uma boa consistência dessas informações. Contudo, os postos de Ponte dos Vilelas e Arcos apresentaram uma mudança de declividade no gráfico dupla-massa, causada pelo preenchimento de falhas das informações. Vale realçar, como comentado no parágrafo anterior, que o posto de Arcos não teve uma boa correlação com os demais postos pluviométricos da bacia.

A análise geoestatística dos dados de precipitação média histórica dos postos pluviométricos abrangeu os seguintes períodos: precipitação média anual, precipitação média mensal, precipitação média do período úmido (novembro a abril) e precipitação média do período seco (maio a outubro).

A tabela 5.2 apresenta os resultados do ajuste do variograma teórico a partir de um variograma experimental. O melhor semi-variograma foi determinado analisando os resultados da validação cruzada. Esta tabela apresenta o melhor semi-variograma (mostrado em destaque) e os resultados da aplicação da Krigagem Ordinária, onde, conseqüentemente, não foi considerada a remoção da tendência.

Tabela 5.1 – Coeficiente de correlação entre os postos selecionados na área a montante de Três Marias.

Posto / Correlação	Ponte dos Vilelas	Arcos	Bonfin	Carmópolis de Minas	Desterro de Entre Rios	Esmeraldas	Estrela do Indaiá	Itaguara	Luz	Morada Nova	Paineiras	Porto Mesquita	Porto Pará	São Gonçalo do Abaeté	São José dos Buritis	São Sebastião do Oeste	SE Pedra do Indaiá	UHE Cajuru	UHE Gafanhoto
Ponte dos Vilelas	1,00	0,30	0,68	0,76	0,71	0,57	0,48	0,68	0,53	0,32	0,35	0,49	0,48	0,34	0,39	0,67	0,54	0,72	0,68
Arcos	0,30	1,00	0,29	0,32	0,29	0,30	0,28	0,28	0,29	0,24	0,24	0,3	0,26	0,25	0,28	0,24	0,37	0,26	0,25
Bonfin	0,68	0,29	1,00	0,60	0,63	0,57	0,45	0,63	0,44	0,33	0,36	0,41	0,42	0,33	0,37	0,52	0,49	0,59	0,59
Carmópolis de Minas	0,76	0,32	0,60	1,00	0,68	0,50	0,43	0,63	0,50	0,32	0,34	0,41	0,45	0,33	0,37	0,60	0,51	0,60	0,59
Desterro de Entre Rios	0,71	0,29	0,63	0,68	1,00	0,52	0,43	0,62	0,46	0,36	0,32	0,39	0,4	0,30	0,35	0,55	0,47	0,56	0,57
Esmeraldas	0,57	0,30	0,57	0,50	0,52	1,00	0,47	0,53	0,45	0,39	0,42	0,53	0,49	0,39	0,43	0,50	0,45	0,54	0,57
Estrela do Indaiá	0,48	0,28	0,45	0,43	0,43	0,47	1,00	0,44	0,63	0,41	0,48	0,41	0,52	0,43	0,46	0,49	0,44	0,46	0,50
Itaguara	0,68	0,28	0,63	0,63	0,62	0,53	0,44	1,00	0,43	0,30	0,34	0,36	0,39	0,31	0,36	0,55	0,51	0,59	0,57
Luz	0,53	0,29	0,44	0,50	0,46	0,45	0,63	0,43	1,00	0,37	0,44	0,38	0,47	0,36	0,40	0,49	0,50	0,50	0,54
Morada Nova	0,32	0,24	0,33	0,32	0,36	0,39	0,41	0,30	0,37	1,00	0,53	0,41	0,44	0,50	0,63	0,33	0,34	0,35	0,38
Paineiras	0,35	0,24	0,36	0,34	0,32	0,42	0,48	0,34	0,44	0,53	1,00	0,43	0,44	0,50	0,55	0,39	0,39	0,36	0,42
Porto Mesquita	0,49	0,30	0,41	0,41	0,39	0,53	0,41	0,36	0,38	0,41	0,43	1,00	0,54	0,38	0,47	0,43	0,39	0,40	0,44
Porto Pará	0,48	0,26	0,42	0,45	0,40	0,49	0,52	0,39	0,47	0,44	0,44	0,54	1,00	0,43	0,48	0,43	0,41	0,44	0,47
São Gonçalo do Abaeté	0,34	0,25	0,33	0,33	0,30	0,39	0,43	0,31	0,36	0,50	0,50	0,38	0,43	1,00	0,51	0,34	0,33	0,33	0,35
São José dos Buritis	0,39	0,28	0,37	0,37	0,35	0,43	0,46	0,36	0,40	0,63	0,55	0,47	0,48	0,51	1,00	0,37	0,36	0,39	0,41
São Sebastião do Oeste	0,67	0,24	0,52	0,60	0,55	0,50	0,49	0,55	0,49	0,33	0,39	0,43	0,43	0,34	0,37	1,00	0,57	0,63	0,66
SE Pedra do Indaiá	0,54	0,37	0,49	0,51	0,47	0,45	0,44	0,51	0,50	0,34	0,39	0,39	0,41	0,33	0,36	0,57	1,00	0,51	0,54
UHE Cajuru	0,72	0,26	0,59	0,60	0,56	0,54	0,46	0,59	0,50	0,35	0,36	0,4	0,44	0,33	0,39	0,63	0,51	1,00	0,71
UHE Gafanhoto	0,68	0,25	0,59	0,59	0,57	0,57	0,50	0,57	0,54	0,38	0,42	0,44	0,47	0,35	0,41	0,66	0,54	0,71	1,00

Tabela 5.2 – Resultados dos semi-variogramas ajustados e da aplicação da interpolação de Kriging (validação cruzada).

Casos/Análises		Caso 1: Precipitação Média Anual			Caso 2: Precipitação Média Mensal		
Modelo Ajustado		Esférico	Exponencial	Gaussiano	Esférico	Exponencial	Gaussiano
Transformação		sem transformação			sem transformação		
Ordem para remoção de tendenciosidade		nenhuma			nenhuma		
Parâmetros do modelo	Patamar (Co+Cs)	1,3E-02	1,4E-02	1,4E-02	3,0E-06	3,4E-06	3,2E-06
	Efeito Pepita (Co)	4,9E-03	2,8E-03	5,6E-03	9,8E-07	5,0E-07	1,2E-06
	Co/(Co+Cs)*100	38,5	19,7	39,9	32,6	14,7	37,9
	Alcance (Km)	197,6	197,6	181,5	197,6	197,6	175,3
Significância do modelo (validação cruzada)	Médio	1,382	0,8859	1,801	2,69	2,016	3,579
	Médio Quadrático	96,56	96,88	94,39	142,2	142,5	138,9
	Médio Padrão	98,03	100,01	91,1	143,8	147,2	134,5
	Médio Normalizado	0,004053	-0,002167	0,008883	0,006853	0,0006525	0,01352
	Desvio Padrão Normalizado	0,9783	0,9704	1,03	0,9818	0,9721	1,028
Casos/Análises		Caso 3: Precipitação Média do Período Úmido (nov a abr)			Caso 4: Precipitação Média do Período Seco (mai a out)		
Modelo Ajustado		Esférico	Exponencial	Gaussiano	Esférico	Exponencial	Gaussiano
Transformação		sem transformação			sem transformação		
Ordem para remoção de tendenciosidade		nenhuma			nenhuma		
Parâmetros do modelo	Patamar (Co+Cs)	6,8E-03	7,6E-03	7,0E-03	1,6E-01	1,5E-01	2,1E-01
	Efeito Pepita (Co)	3,7E-03	3,0E-03	4,3E-03	0	0	9,56E-03
	Co/(Co+Cs)*100	53,6	39,5	61,9	0	0	4,7
	Alcance (Km)	170,4	197,6	160,4	197,6	197,6	197,6
Significância do modelo (validação cruzada)	Médio	0,2323	-0,2666	0,5888	1,554	1,036	2,357
	Médio Quadrático	83,07	84,27	81,79	18,9	19,54	19,72
	Médio Padrão	82,11	84,5	78,95	20,59	25,61	14,14
	Médio Normalizado	-0,004108	-0,01011	0,0003811	0,03824	0,01528	0,1052
	Desvio Padrão Normalizado	1,005	0,9928	1,028	0,9065	0,7458	1,357

Na validação cruzada mostrada na tabela 5.2 observa-se que, para todos os casos destacados (em vermelho), o erro médio quadrático e padrão são semelhantes. Também se observa que o erro médio normalizado está próximo a zero e o desvio médio normalizado está próximo a um.

A figura 5.1 mostra os gráficos da validação cruzada, ou seja, a precipitação prevista com a Krigagem e a observada, bem como o teste de normalidade efetuado e a figura 5.2, os mapas com as isolinhas da precipitação total anual e do erro de interpolação, para o semi-variograma esférico, que foi o melhor ajustado para este caso.

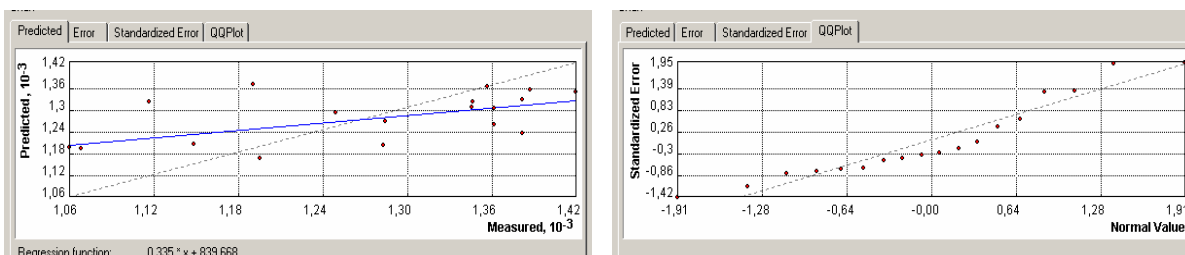


Figura 5.1 - Precipitação média anual – validação cruzada, semi-variograma esférico.

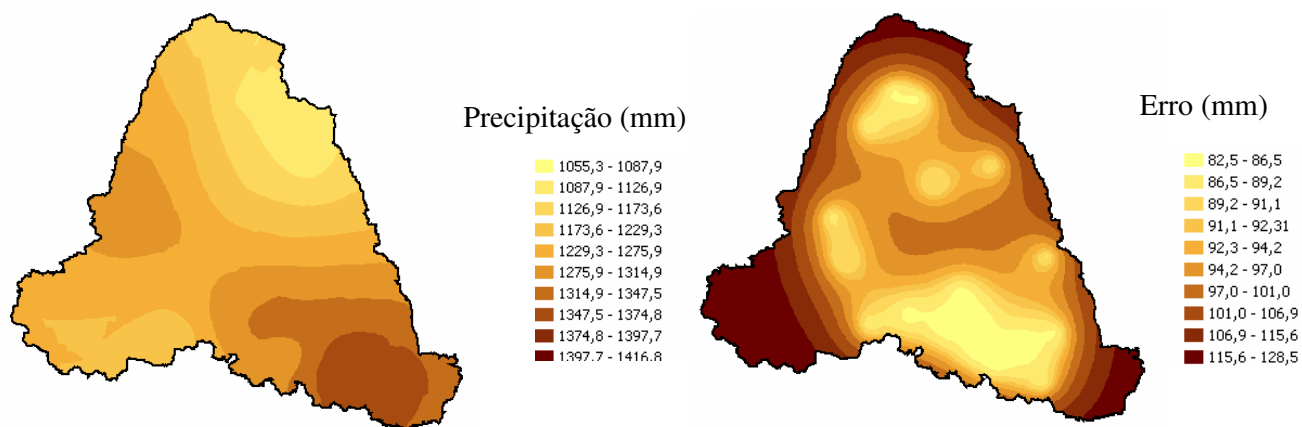


Figura 5.2 – Isolinhas da precipitação média anual e do erro obtida na interpolação por Kriging.

A figura 5.3 mostra o gráfico da validação cruzada, ou seja, precipitação prevista com a Krigagem e a observada, bem como o teste de normalidade efetuado, e a figura 5.4, os mapas com as isolinhas da precipitação total mensal e do erro de interpolação, para o semi-variograma exponencial, que foi o melhor ajustado para este caso.

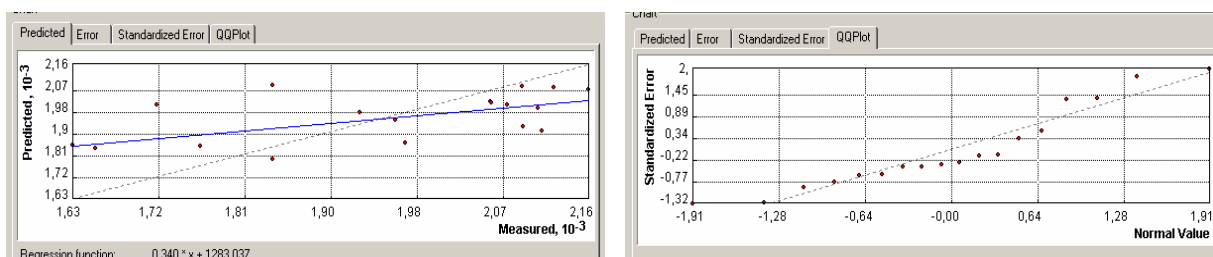


Figura 5.3 - Média da precipitação média mensal – validação cruzada, semi-variograma exponencial.

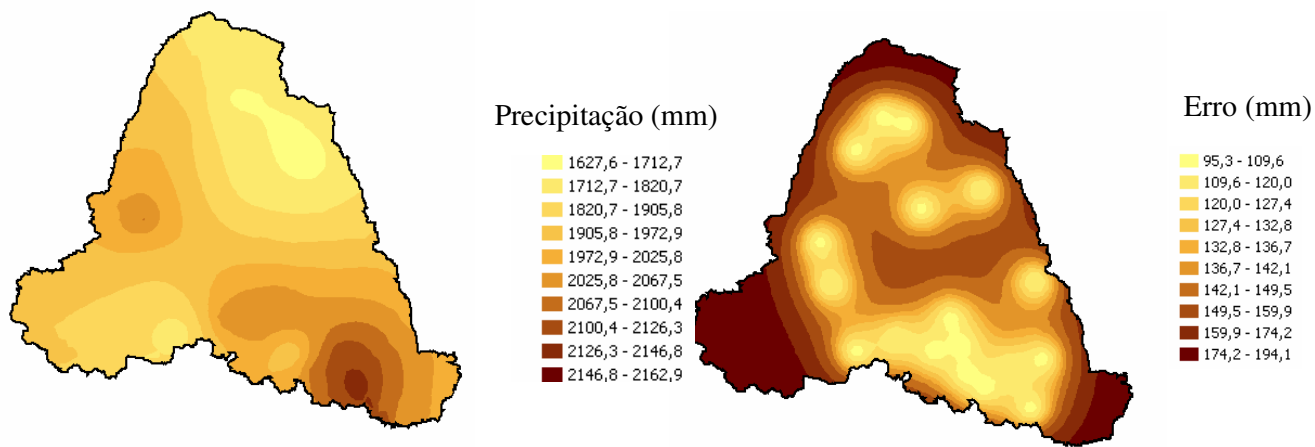


Figura 5.4 – Isolinhas da precipitação média mensal e do erro obtida na interpolação por Kriging.

A figura 5.5 mostra o gráfico da validação cruzada, ou seja, a precipitação prevista com a Krigagem e a observada, bem como o teste de normalidade efetuado e a figura 5.6, os mapas com as isolinhas da precipitação média histórica do período úmido e o erro de interpolação, para o semi-variograma esférico, que foi o melhor ajustado para este caso.

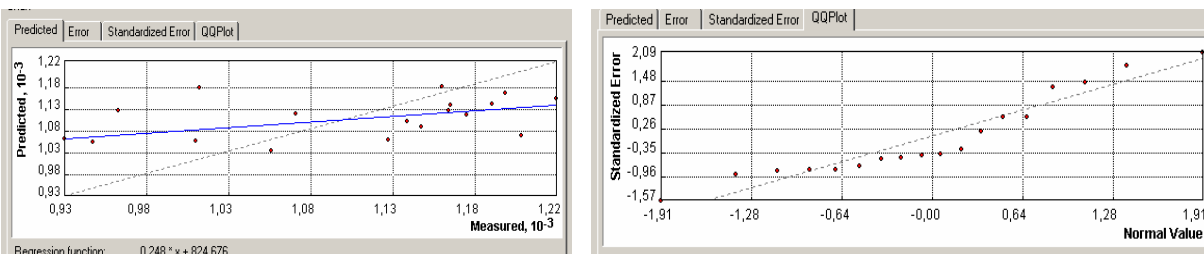


Figura 5.5 - Precipitação média do período úmido (nov a abr) – semi-variograma esférico.

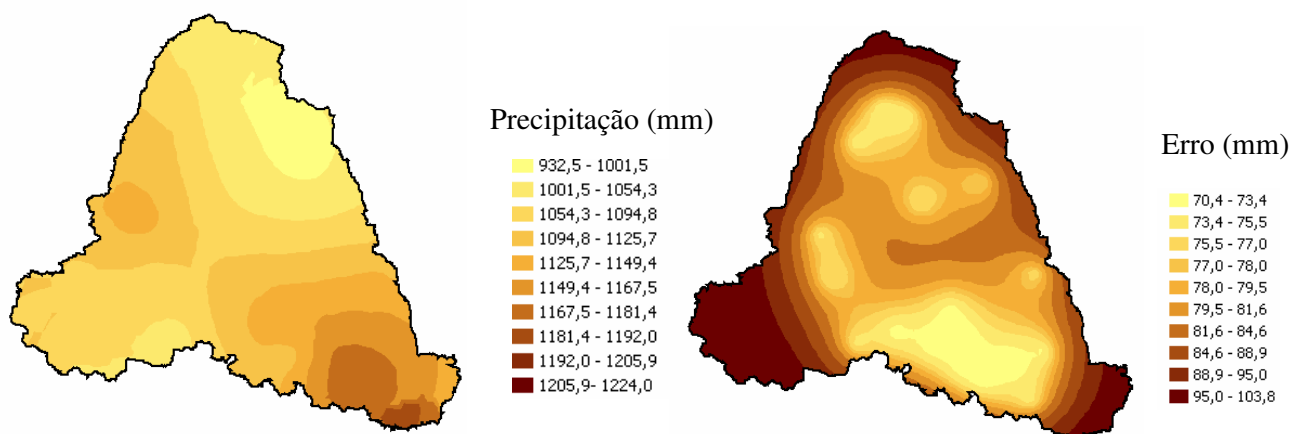


Figura 5.6 – Isolinhas da precipitação média do período úmido e do erro obtida na interpolação por Kriging.

A figura 5.7 mostra o gráfico da validação cruzada, ou seja, a precipitação prevista com a Krigagem e a observada, bem como o teste de normalidade efetuado e a figura 5.8, os mapas com as isolinhas da precipitação média do período seco e do erro de interpolação, para o semi-variograma esférico, que foi o melhor ajustado para este caso.

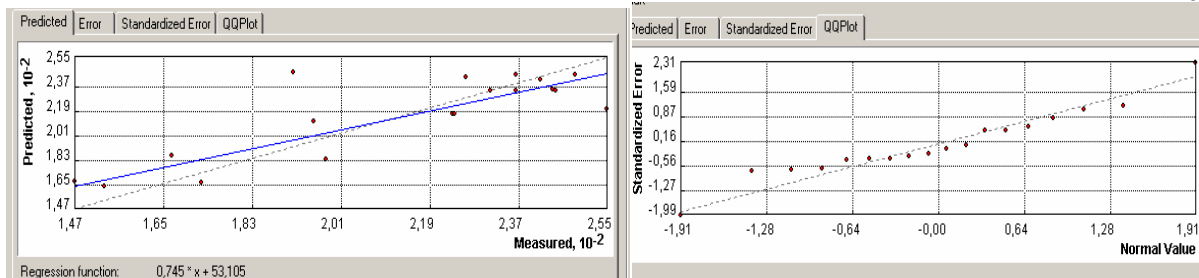


Figura 5.7 – Precipitação média do período seco (mai a out) – validação cruzada, semi-variograma esférico.

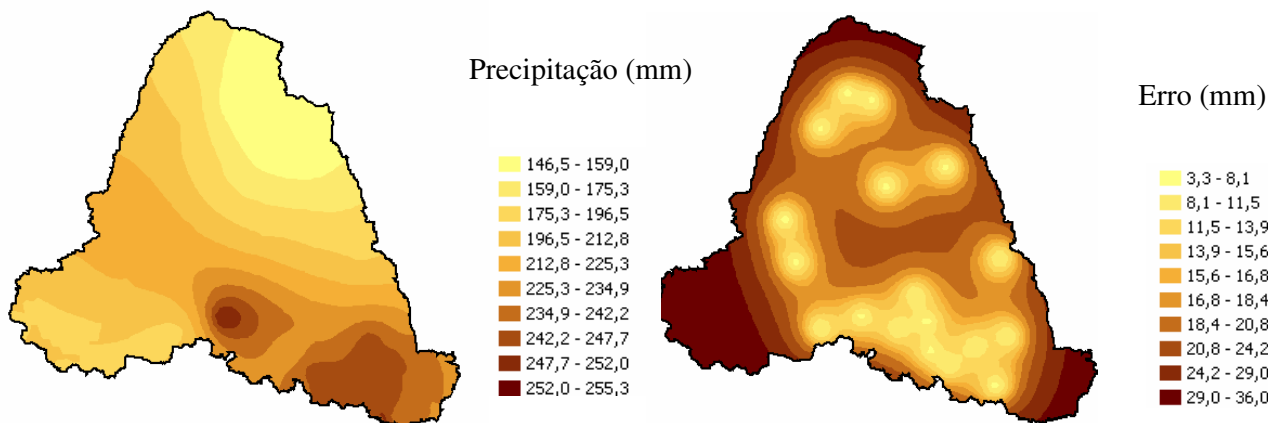


Figura 5.8 - Isolinhas da precipitação média do período seco e do erro obtida na interpolação por Kriging.

Observa-se que o valor do alcance é aproximadamente o mesmo para todas os variogramas ajustados e que a configuração dos mapas de isolinhas também se assemelha para todas as situações, a ordem de grandeza dos valores é que é alterada, ou seja, a estrutura de correlação é semelhante em todos os casos. Segundo classificação proposta por Cambardella et al. (1994), os semi-variogramas da precipitação total anual, mensal e do período úmido, podem ser classificados como de moderada dependência espacial. Apenas o da precipitação do período seco pode ser classificado como de forte dependência espacial.

Amaral et al (2004) mostraram que os valores de precipitação para a bacia do rio Doce foram mais baixos que as do São Francisco. Os alcances de grande parte dos eventos variaram entre 110 e 200 km, sendo descritos por modelos exponenciais ou esféricos de variogramas.

Nas figuras 5.1, 5.3, 5.5 e 5.7 observa-se que a validação cruzada, ou a plotagem das variáveis interpoladas por Kriging com as medidas, teve uma boa relação, destacando essa interpolação para o período seco. Nessas mesmas figuras também se observa que o

teste de normalidade das informações foi atendido, mais uma vez, destacando o desempenho das informações do período seco, não sendo necessário, em nenhum dos casos analisados, efetuar a transformação logarítmica dos dados.

Como era de se esperar, para todas as situações estudadas (média anual, média mensal, média do período úmido e média do período seco), a análise dos erros obtidos com a interpolação através da Krigagem mostra um incremento do erro nas áreas desprovidas de postos pluviométricos. Então, partiu-se para analisar as estações pluviométricas de responsabilidade da ANA/CPRM. A figura 5.9 mostra a disposição espacial dos postos da ANA/CPRM existentes na área em estudo e a figura 5.10 mostra a disposição desses postos (com estrela) e os postos da CEMIG selecionados para as análises geoestatísticas (círculo).

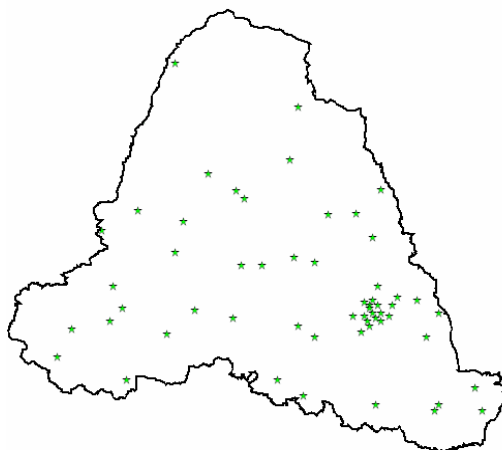


Figura 5.9 – Disposição dos Postos Pluviométricos operados pela ANA/CPRM, localizados na área de drenagem da UHE Três Marias.

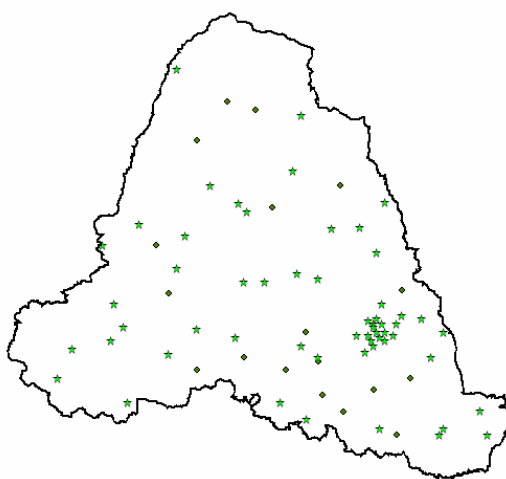


Figura 5.10 – Disposição dos Postos Pluviométricos operados pela ANA/CPRM e os postos da CEMIG usados nas análises geoestatísticas, localizados na área de drenagem da UHE Três Marias.

Com a análise da localização dos postos pluviométricos da ANA/CPRM e dos erros obtidos na Krigagem, foram selecionadas as áreas onde poderiam ser colocados postos pluviométricos adicionais, visando reduzir os erros de interpolação obtidos, com uma conseqüente melhora na determinação da chuva média da área. Analisando a disponibilidade de informações, foram selecionados os postos pluviométricos mostrados na tabela 5.3. A figura 5.11 ilustra a disposição desses postos juntamente com os da CEMIG. Ou seja, todos os postos que serão usados na nova análise geoestatística.

Tabela 5.3 – Postos Pluviométricos ANA/CPRM selecionados.

Código	Nome	Latitude	Longitude
1845002	Fazenda São Félix	-18,47	-45,65
1944031	Ponte da Taquara	-19,42	-44,55
2046013	Vargem Bonita	-20,33	-46,37
1944021	Velho da Taipa	-19,70	-44,93
1945008	Bom Despacho	-19,75	-45,25
1946000	Tapiraí-Jusante	-19,88	-46,03
2043005	Conselheiro Lafaiete	-20,67	-43,78
2044007	Entre Rios de Minas	-20,67	-44,07
2045002	Iguatama	-20,18	-45,70
2045012	Piumhi	-20,47	-45,95

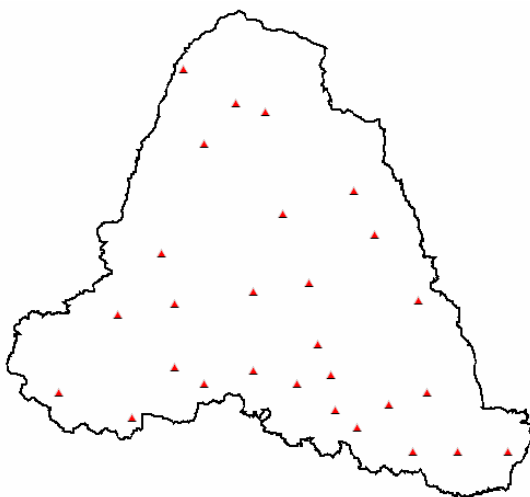


Figura 5.11 – Disposição dos Postos Pluviométricos operados pela ANA/CPRM selecionados e os postos da CEMIG usados nas análises geoestatísticas, localizados na área de drenagem da UHE Três Marias.

A tabela 5.4 mostra os resultados encontrados para a análise geoestatística, usando as mesmas características da análise anterior com as informações de todos os postos pluviométricos selecionados.

Tabela 5.4 – Resultados dos semi-variogramas ajustados e da aplicação da interpolação de Kriging (validação cruzada).

Casos/Análises		Caso 1: Precipitação Média Anual			Caso 2: Precipitação Média Mensal		
Modelo Ajustado		Esférico	Exponencial	Gaussiano	Esférico	Exponencial	Gaussiano
Transformação		sem transformação			sem transformação		
Ordem para remoção de tendenciosidade		nenhuma			nenhuma		
Parâmetros do modelo	Patamar (Co+Cs)	2,24E-06	2,57E-06	1,99E-06	5,68E-04	6,49E-04	5,07E-04
	Efeito Pepita (Co)	6,60E-07	3,45E-07	1,03E-06	1,50E-04	6,76E-05	2,46E-04
	Co/(Co+Cs)*100	29,5	13,4	51,8	26,5	10,4	48,4
	Alcance (Km)	225,8	255,8	208,1	230,1	255,8	214,6
Significância do modelo (validação cruzada)	Médio	-4,567	-7,021	-2,85	-7,301	-11,5	-4,378
	Médio Quadrático	114,4	115,1	115,9	177,7	178,4	180,4
	Médio Padrão	115,4	115,9	116,2	177,4	177,4	179,4
	Médio Normalizado	-0,02488	-0,03854	-0,01685	-0,02525	-0,04039	-0,01644
	Desvio Padrão Normalizado	1,004	1,006	1,01	1,012	1,016	1,017

Casos/Análises		Caso 3: Precipitação Média do Período Úmido (nov a abr)			Caso 4: Precipitação Média do Período Seco (mai a out)		
Modelo Ajustado		Esférico	Exponencial	Gaussiano	Esférico	Exponencial	Gaussiano
Transformação		sem transformação			sem transformação		
Ordem para remoção de tendenciosidade		nenhuma			nenhuma		
Parâmetros do modelo	Patamar (Co+Cs)	1,08E-04	1,38E-04	8,98E-04	2,82E-01	2,39E-01	3,18E-01
	Efeito Pepita (Co)	6,62E-05	4,46E-05	8,51E-04	0	0	3,22E-02
	Co/(Co+Cs)*100	61,5	32,3	94,8	0	0	10,1
	Alcance (Km)	211,9	255,8	186,5	255,8	255,8	255,8
Significância do modelo (validação cruzada)	Médio	-3,649	-4,967	-2,928	-0,4346	-1,116	0,5484
	Médio Quadrático	103,3	101,7	104,7	25,59	25,48	24,28
	Médio Padrão	103,2	102	103,9	23,27	28,87	21,76
	Médio Normalizado	-0,02514	-0,03252	-0,02155	-0,01365	-0,02845	0,01718
	Desvio Padrão Normalizado	1,002	0,9994	1,011	1,136	0,8992	1,115

Na validação cruzada mostrada na tabela 5.4 observa-se que, para todos os casos destacados (em vermelho), o erro médio quadrático e padrão são semelhantes. Também se observa que o erro médio normalizado está próximo a zero e o desvio médio normalizado está próximo a um.

A figura 5.12 mostra o gráfico da validação cruzada, ou seja, a precipitação prevista com a Krigagem e a observada, bem como o teste de normalidade efetuado e a figura 5.13, os mapas com as isolinhas da precipitação total anual e do erro de interpolação, para o semi-variograma gaussiano, que foi o melhor ajustado para este caso.

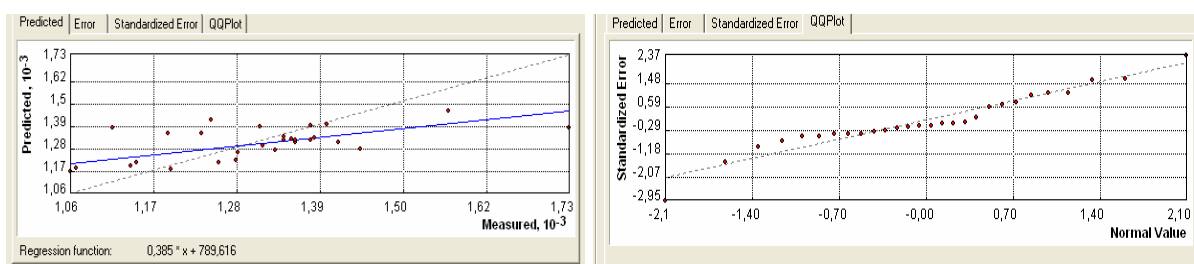


Figura 5.12 - Precipitação média anual – validação cruzada, semi-variograma gaussiano.

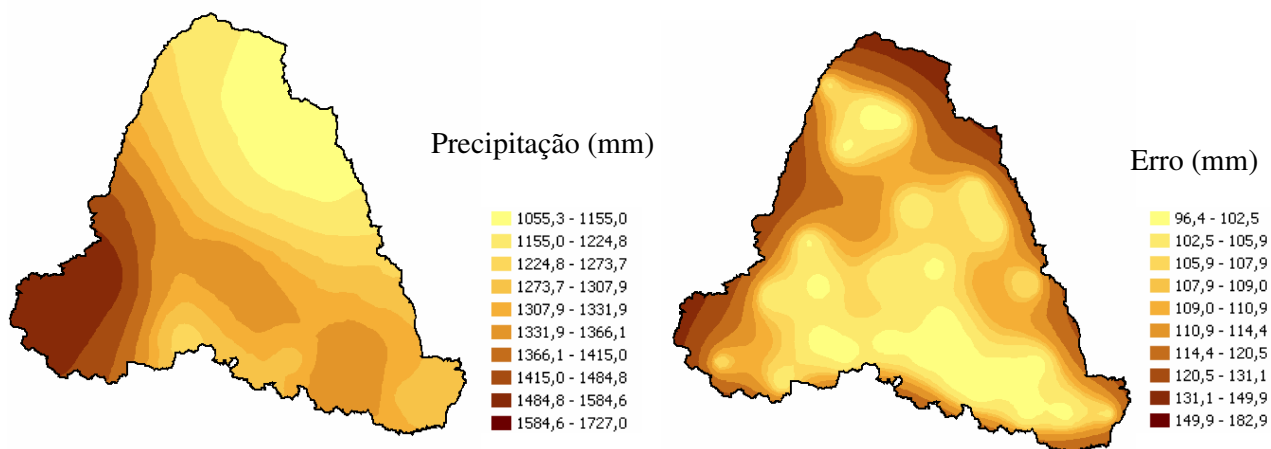


Figura 5.13 - Isolinhas da precipitação média anual e do erro obtida na interpolação por Kriging.

A figura 5.14 mostra o gráfico da validação cruzada, ou seja, a precipitação prevista com a Krigagem e a observada, bem como o teste de normalidade efetuado e a figura 5.15, os mapas com as isolinhas da precipitação total mensal e do erro de interpolação, para o semi-variograma esférico, que foi o melhor ajustado para este caso.

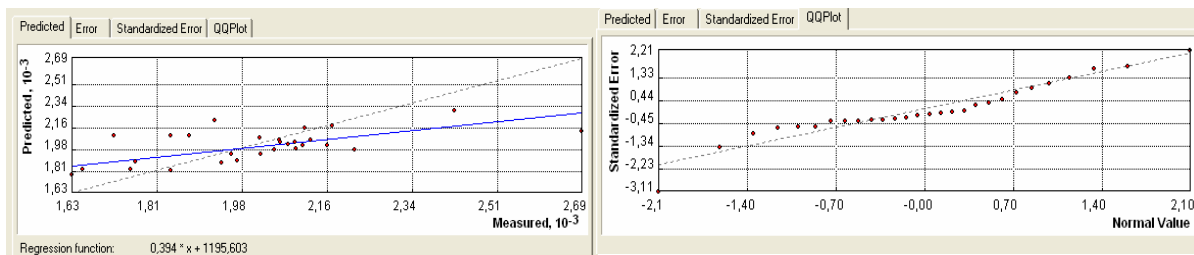


Figura 5.14 - Precipitação média mensal – validação cruzada, semi-variograma esférico.

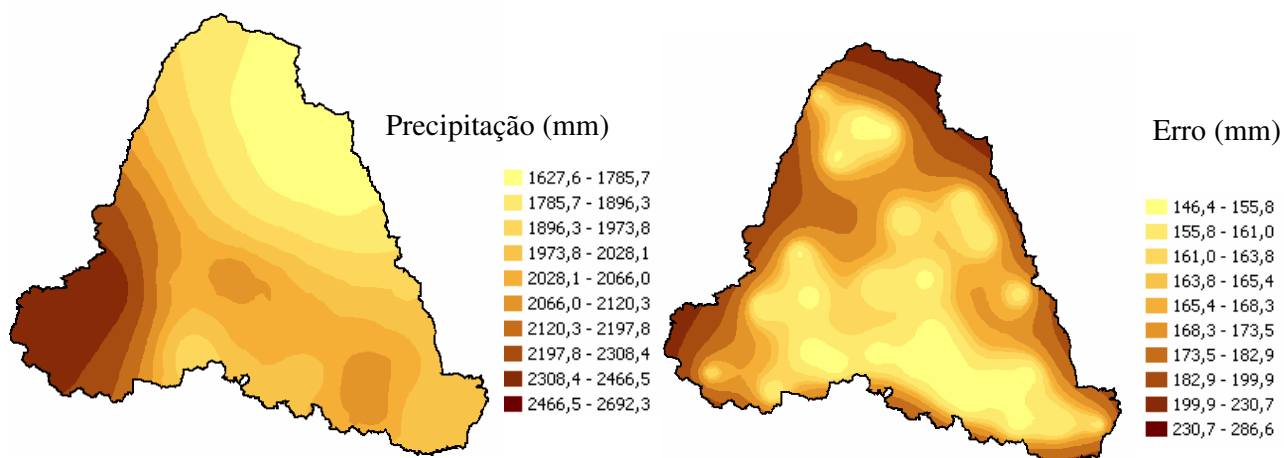


Figura 5.15 - Isolinhas da precipitação média mensal e do erro obtida na interpolação por Kriging.

A figura 5.16 mostra o gráfico da validação cruzada, ou seja, a precipitação prevista com a Krigagem e a observada, bem como o teste de normalidade efetuado e a figura 5.17, os mapas com as isolinhas da precipitação do período úmido e do erro de interpolação, para o semi-variograma esférico, que foi o melhor ajustado para este caso.

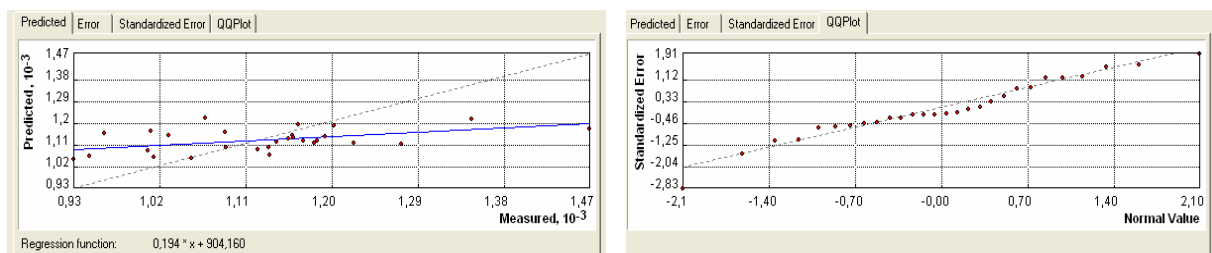


Figura 5.16 - Precipitação média do período úmido (nov a abr) – validação cruzada, semi-variograma esférico.

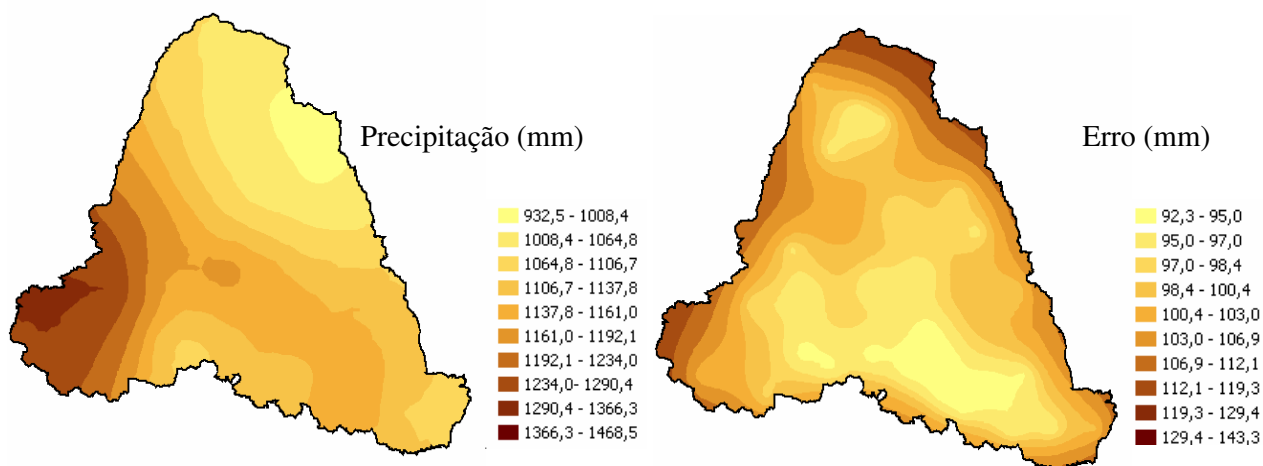


Figura 5.17 - Isolinhas da precipitação média do período úmido e do erro obtida na interpolação por Kriging.

A figura 5.18 mostra o gráfico da validação cruzada, ou seja, a precipitação prevista com a Krigagem e a observada, bem como o teste de normalidade efetuado e a figura 5.19, os mapas com as isolinhas da precipitação do período seco e do erro de interpolação, para o semi-variograma esférico, que foi o melhor ajustado para este caso.

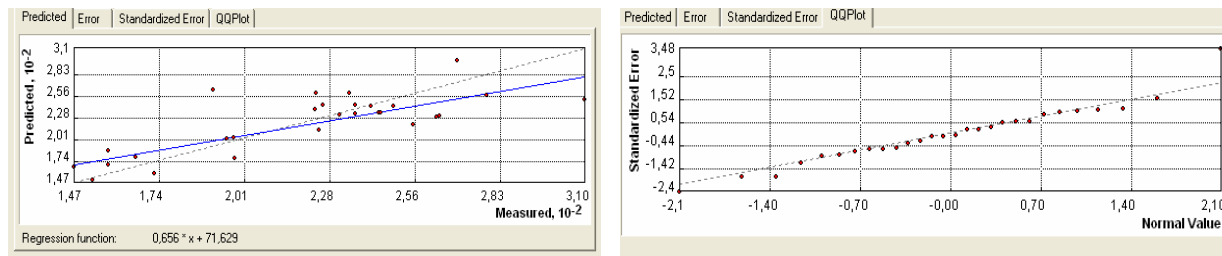


Figura 5.18 - Precipitação média do período seco (mai a out) – validação cruzada, semi-variograma esférico.

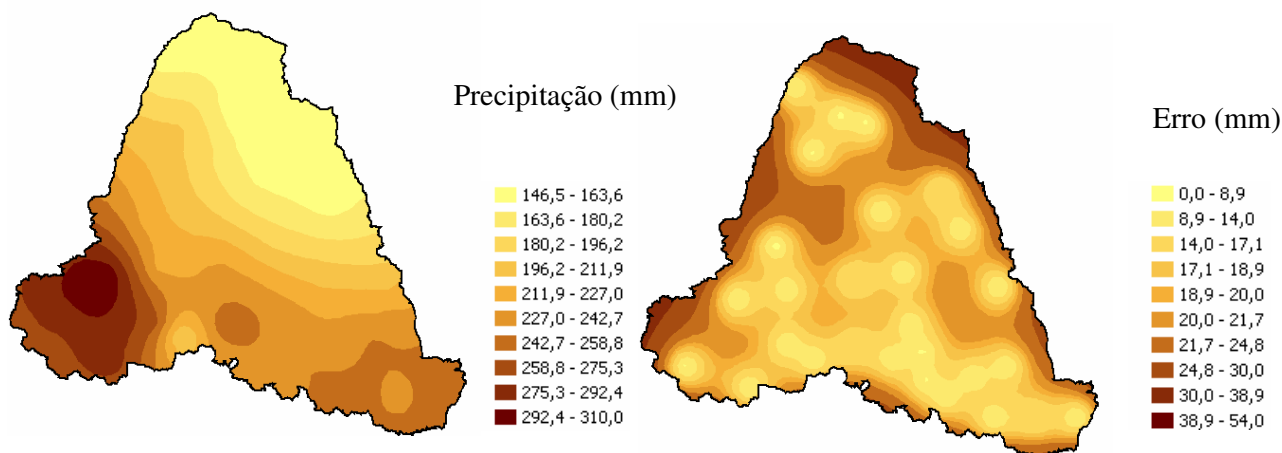


Figura 5.19 - Isolinhas da precipitação média do período seco e do erro obtida na interpolação por Kriging.

Mais uma vez, observa-se que o valor do alcance é aproximadamente o mesmo para todas as variogramas ajustados e que a configuração dos mapas de isolinhas também se assemelha para todas as situações. A ordem de grandeza dos valores é que é alterada, ou seja, a estrutura de correlação é semelhante em todos os casos. Também como no caso anterior, os semi-variogramas da precipitação total anual, mensal e do período úmido, podem ser classificados como de moderada dependência espacial e o da precipitação do período seco pode ser classificado como de forte dependência espacial.

Nas figuras 5.12, 5.14, 5.16 e 5.18 observa-se que a validação cruzada, ou a plotagem as variáveis interpoladas por Kriging com as medidas, teve uma boa relação. Nessas mesmas figuras observa-se o teste de normalidade das informações foi atendido e ficaram mais ajustados que a análise geoestatística anterior, considerando só os postos

da CEMIG, não sendo necessário, em nenhum dos casos analisados, efetuar a transformação logarítmica dos dados.

Como, com exceção da média da precipitação total anual, todos os semi-variogramas das informações mensais tiveram um melhor ajuste através do modelo esférico, também foi aplicada a Krigagem usando um semi-variograma esférico para os dados diários, através do software Surfer, para estimativa da precipitação média espacial. Vale salientar que não foi ajustado um novo semi-variograma para esses dados devido à dificuldade em assim fazer para os 6574 dias necessários. Na próxima seção parte esses resultados serão mais bem analisados.

5.1.2 Informações fluviométricas

Na tabela 5.5 estão contidos coeficientes de correlação dos postos fluviométricos selecionados. Observa-se que os dados dos postos fluviométricos têm uma correlação razoável, com coeficientes de correlação sempre superiores a 0,69.

Tabela 5.5 – Determinação entre as informações fluviométricas.

Postos / Correlação	IGUATAMA	PORTO DAS ANDORINHAS	PONTE NOVA DO PARAOPEBA	PORTO MESQUITA	PONTE DOS VILELAS JUSANTE	PORTO PARÁ	PORTO INDÁIA
IGUATAMA	1,00	0,88	0,74	0,78	0,68	0,84	0,69
PORTO DAS ANDORINHAS	0,88	1,00	0,71	0,77	0,60	0,83	0,66
PONTE NOVA DO PARAOPEBA	0,74	0,71	1,00	0,87	0,83	0,85	0,64
PORTO MESQUITA	0,78	0,77	0,87	1,00	0,68	0,93	0,65
PONTE DOS VILELAS JUSANTE	0,68	0,60	0,83	0,68	1,00	0,73	0,59
PORTO PARA	0,84	0,83	0,85	0,93	0,73	1,00	0,70
PORTO INDÁIA	0,69	0,66	0,64	0,65	0,59	0,70	1,00

As Figuras de 5.20 a 5.24 mostram os resultados do preenchimento de falhas através da regressão linear ou potencial.

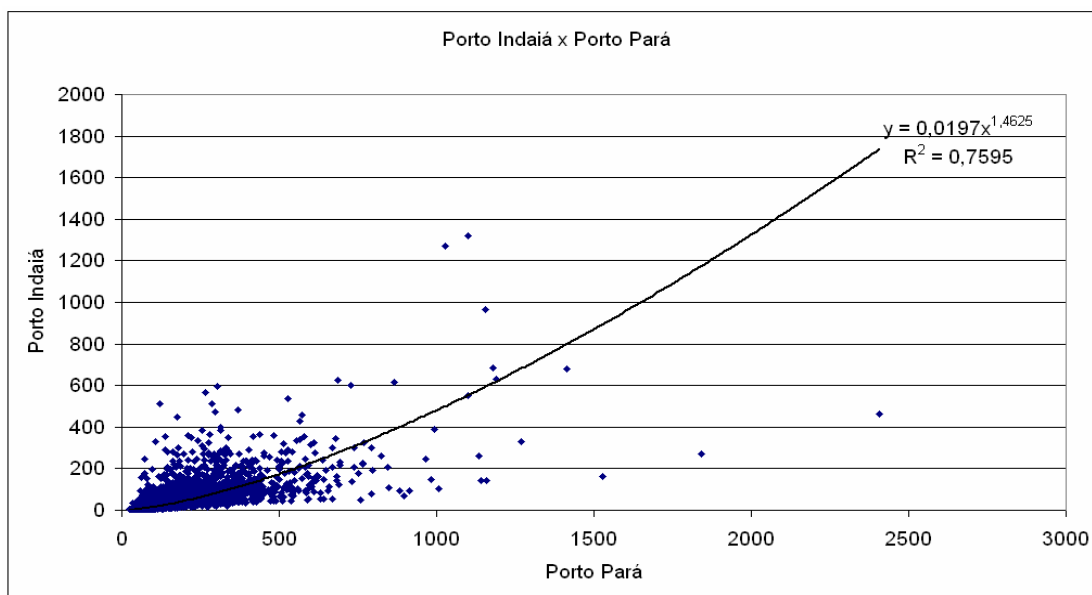


Figura 5.20 – Regressão potencial entre os postos fluviométricos de Porto Indaiá e Porto Pará.

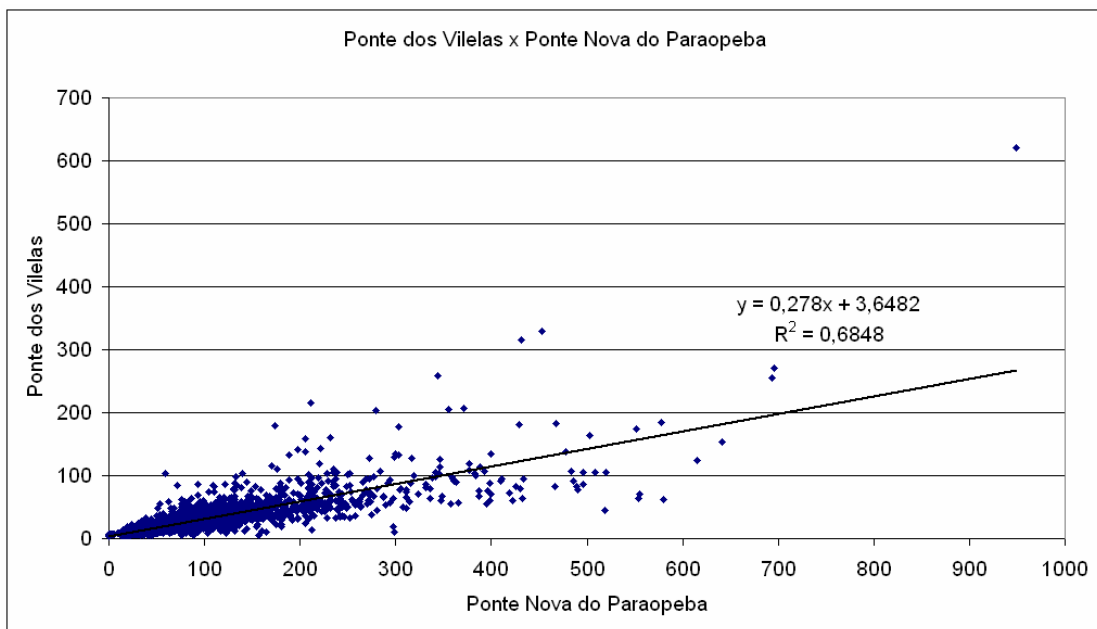


Figura 5.21 – Regressão linear simples entre os postos fluviométricos de Ponte dos Vilelas e Ponte Nova do Paraopeba.

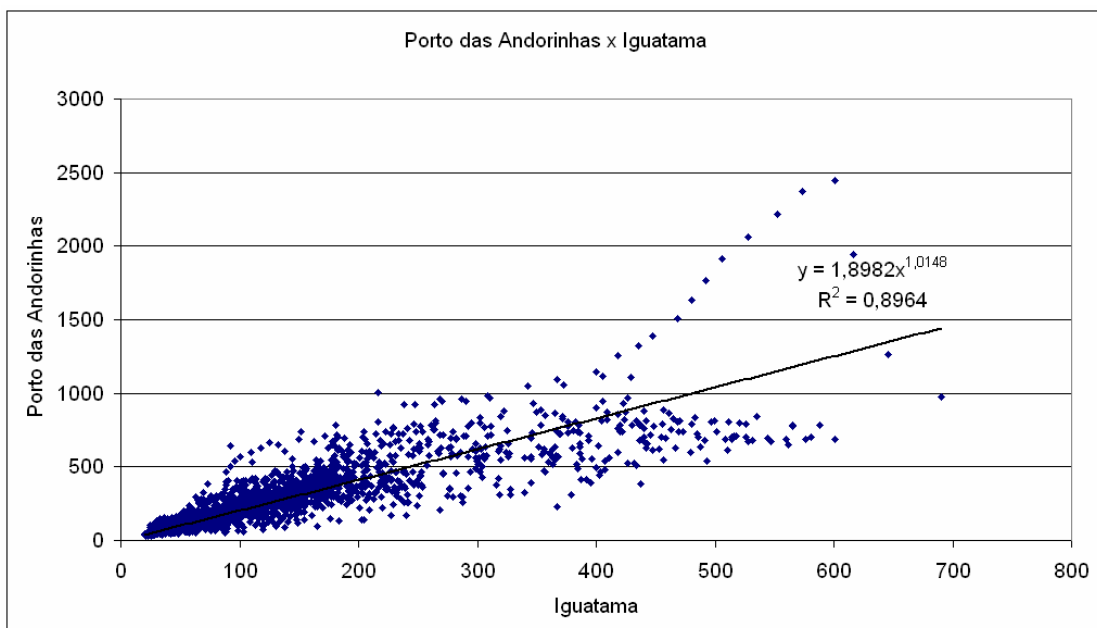


Figura 5.22 – Regressão potencial entre os postos fluviométricos de Porto das Andorinhas e Iguatama.

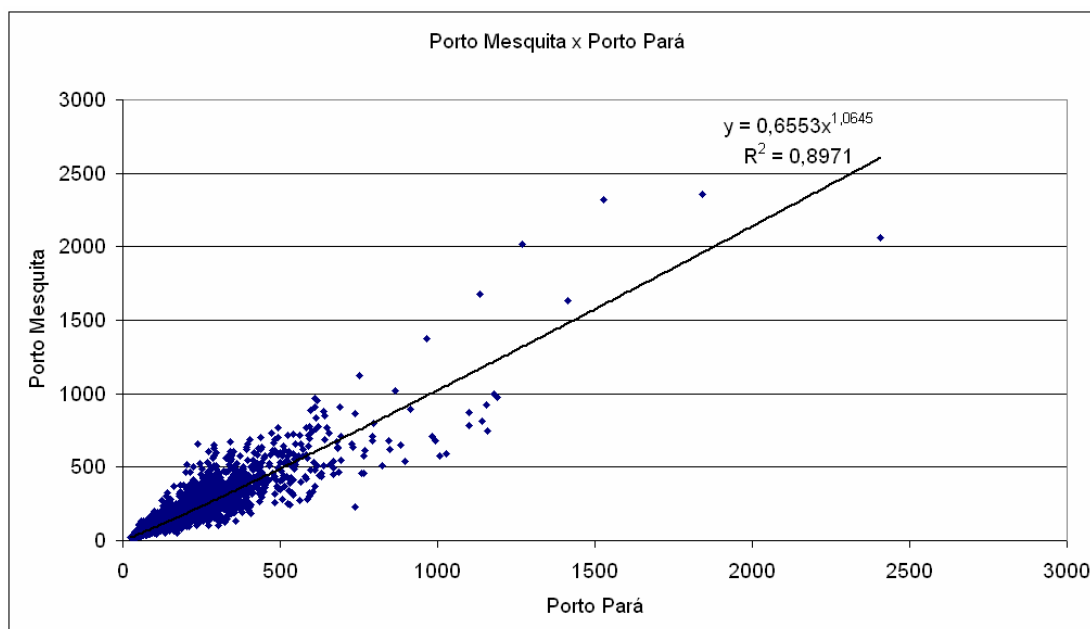


Figura 5.23 – Regressão potencial entre os postos fluviométricos de Porto Mesquita e Porto Pará.

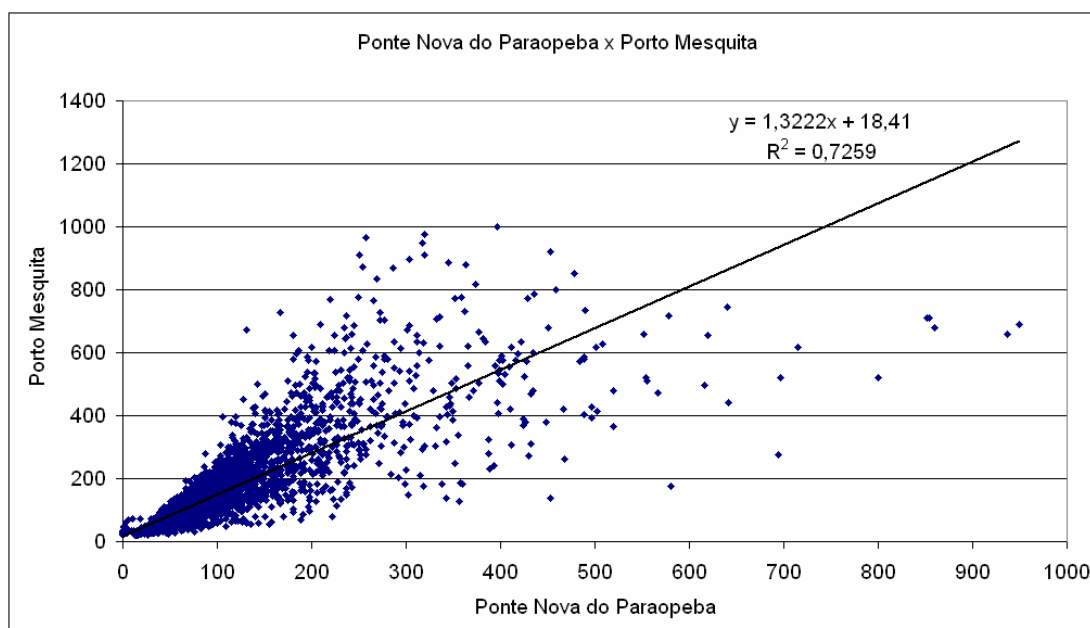


Figura 5.24 – Regressão linear simples entre os postos fluviométricos de Porto Mesquita e Ponte Nova do Paraopeba.

Observa-se que todos os coeficientes de correlação encontrados no preenchimento de falhas através da regressão linear simples ou potencial foram maiores que 0,68, o que leva a concluir que as informações fluviométricas usadas são bem correlacionadas.

A tabela 5.6 mostra os resultados obtidos no preenchimento de falhas através da regressão múltipla.

Tabela 5.6 – Informações obtidas na regressão múltipla.

Variáveis / Postos	a	b	c	d	e	f	g
<i>Estatística de regressão</i>							
R múltiplo	0,84	0,74	0,93	0,94	0,92	0,87	0,84
R-Quadrado	0,71	0,54	0,86	0,89	0,84	0,76	0,71
R-quadrado ajustado	0,71	0,54	0,86	0,89	0,84	0,76	0,71
Erro padrão	116,75	39,47	43,03	37,66	36,12	36,96	13,35
<i>Coefficientes</i>							
Interseção	23,18	-15,78	-4,61	14,08	-1,18	8,83	1,05
Variável X 1	f -0,56	a 0,08	b -0,03	c 0,48	a 0,31	d 0,25	e 0,05
Variável X 2	d 1,41	f 0,18	a 0,03	b 0,16	b 0,13	c 0,36	a -0,01
Variável X 3	c 0,25	d 0,17	f 0,49	a 0,15	c -0,07	b 0,16	b 0,00
Variável X 4	b 0,68	c -0,03	d 0,63	f 0,26	d 0,21	a -0,06	c -0,07
Variável X 5	- -	- -	- -	- -	f 0,07	- -	d 0,07
Variável X 6	- -	- -	- -	- -	g 0,38	- -	f 0,27

Porto das Andorinhas - a

Porto Indaiá - b

Porto Mesquita - c

Porto Pará - d

Iguatama - e

Ponte Nova do Paraopeba - f

Ponte dos Vilelas Jusante - g

Analisando-se o coeficiente R^2 (tabela 5.6) e os coeficientes da regressão simples ou potencial (figuras 5.20 ao 5.24), observa-se que há um ganho em se usar a regressão múltipla no preenchimento de falhas, e assim foi feito neste trabalho. Na próxima seção esses resultados serão mais bem explorados.

5.1.3 Análise Estatística

As figuras de 5.25 a 5.29 a seguir mostram os resultados da função de auto-correlação (FAC) para as informações de vazão e de precipitação média espacial da área, onde o lag representa a defasagem temporal da variável, nestas figuras, em dias.

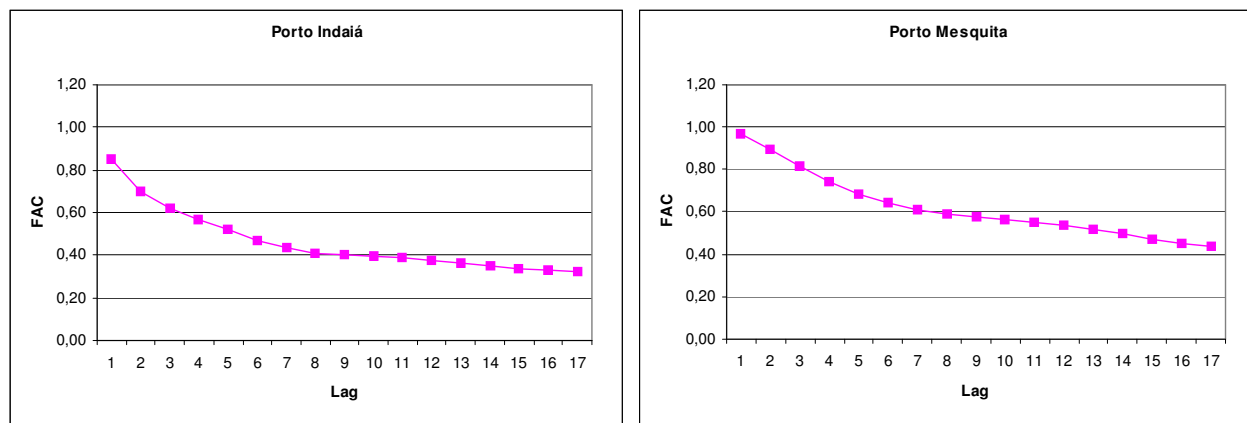


Figura 5.25 – FAC dos postos de Porto Indaiá e Porto Mesquita.

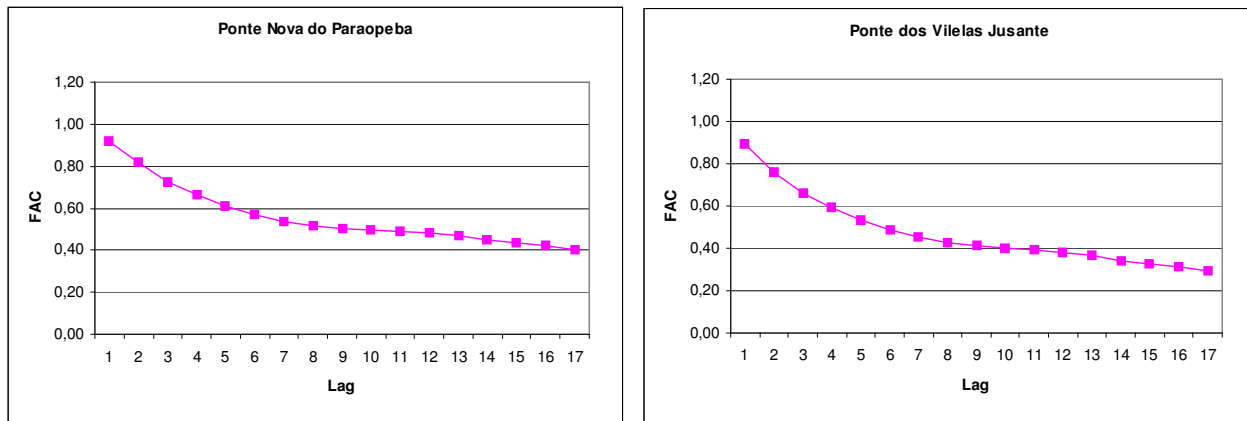


Figura 5.26 – FAC dos postos de Ponte Nova do Paraopeba e Ponte dos Vilelas Jusante.

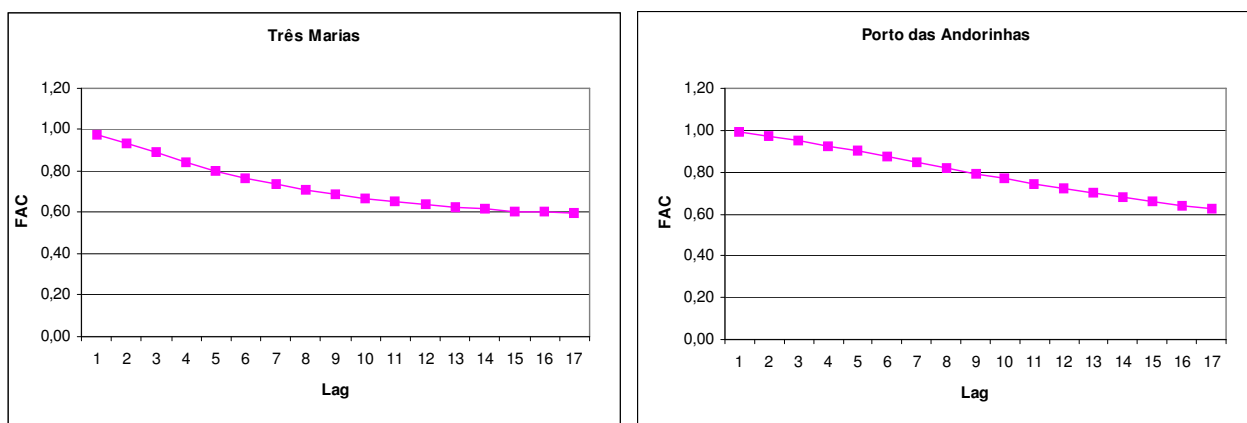


Figura 5.27 – FAC da vazão natural de Três Marias e do posto de Porto das Andorinhas.

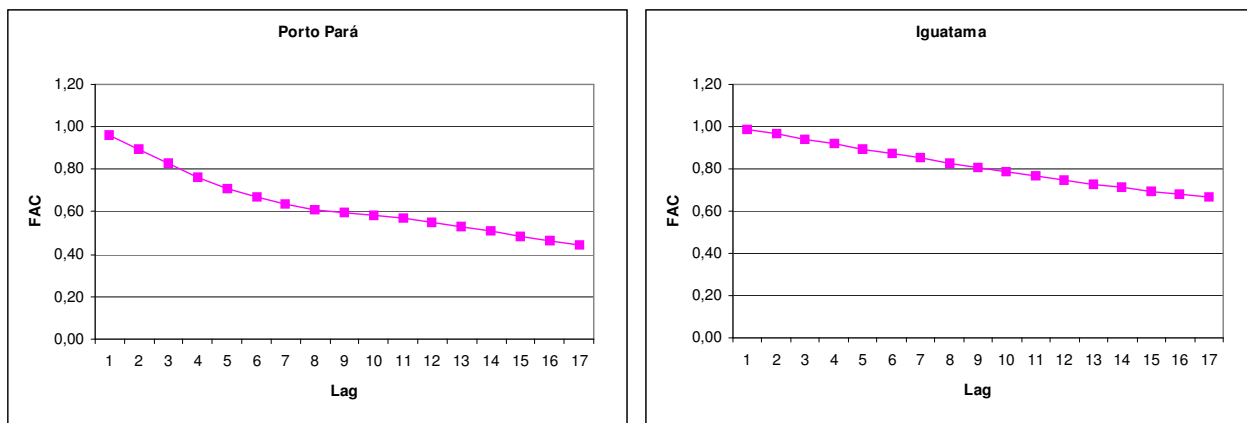


Figura 5.28 – FAC dos postos de Porto Pará e Iguatama.

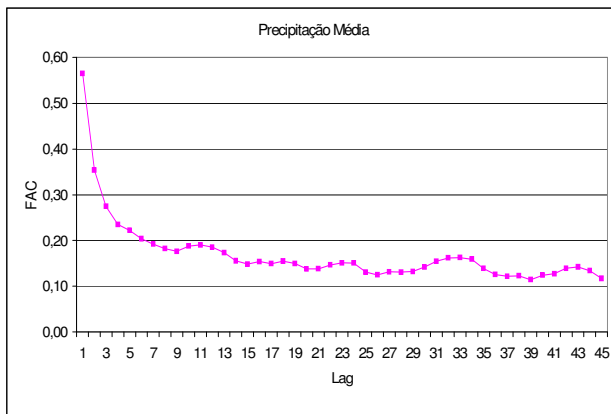


Figura 5.29 – FAC da precipitação média calculada por Thiessen.

Nas figuras de 5.25 a 5.28 observa-se que a partir de 0,6 a FAC começa a ficar constante com a variação do lag. Assim, considerando esse valor de 0,6 como limite da FAC, a tabela 5.7 mostra os lags obtidos.

Tabela 5.7 – Lags obtidos para FAC máxima de 0,6.

Variável		Lag
Vazão Natural	Três Marias	15
Postos Fluviométricos	Porto das Andorinhas	17
	Porto Indaiá	3
	Porto Mesquita	8
	Porto Pará	9
	Iguatama	20
	Ponte Nova do Paraopeba	5
	Ponte dos Vilelas Jusante	4

Nas simulações com redes neurais, a série deve ser analisada em termos de tendência e sazonalidade.

As figuras de 5.25 a 5.28 mostram que as séries de vazões naturais apresentam altos valores de correlação para defasagens temporais de até 17 dias, o que mostra que a série não pode ser considerada como uma seqüência aleatória. Os gráficos de auto-correlação servem também para se verificar a existência de sazonalidade na série. Neste caso o gráfico de auto-correlação aparece na forma de uma senóide (entre um lag e outro a FAC varia muito). No caso das séries de vazões, as figura de 5.25 a 5.28 mostram que não há sazonalidade marcante, para o lag analisado.

Os gráficos de auto-correlação podem ser analisados para verificar se a série pode ser adequadamente modelada por um modelo auto-regressivo linear. Ou seja, a previsão de um registro pode ser realizada como uma combinação linear dos registros anteriores e o gráfico de auto-correlação aparece como uma reta. Este não é o caso para as séries de vazão, indicando que a modelagem adequada para essas séries podem ser não lineares e ainda devem certamente contar com informações adicionais que, no caso, deverão vir das medidas de chuva na bacia.

Na figura 5.29 observa-se que a série de precipitação é menos auto-correlacionada que a série de vazão, mas não pode ser considerada como uma seqüência aleatória.

Também considerando um limite da correlação cruzada máxima de 0,6, a tabela 5.8 mostra os Lags obtidos para a correlação cruzada da vazão natural de Três Marias com os postos fluviométricos e com a chuva média da área.

Tabela 5.8 – Lags obtidos para correlação cruzada máxima de 0,6 entre a vazão natural de Três Marias e a variável indicada.

Variável		Lag
Postos Fluviométricos	Porto das Andorinhas	27
	Porto Indaiá	8
	Porto Mesquita	10
	Porto Pará	11
	Iguatama	26
	Ponte Nova do Paraopeba	10
	Ponte dos Vilelas Jusante	11
Postos Pluviométricos	Chuva Média por Thiessen	7

A plotagem da correlação cruzada de duas variáveis com o lag apresenta como varia a correlação entre duas séries de acordo com a defasagem de uma delas. O lag está relacionado com a representatividade e a área de drenagem do posto fluviométrico. Assim, analisando a tabela 5.8 observa-se que o posto de Porto Indaiá, que é o posto fluviométrico mais próximo do reservatório e tem menor área de drenagem que os outros postos fluviométricos, teve o menor lag com a vazão natural de Três Marias, que depende de uma vazão de até 8 dias deste posto.

Os postos de Porto Pará e Ponte dos Vilelas Jusante, Porto Mesquita e Ponte Nova do Paraopeba, estão, aos pares, na mesma sub-bacia (conforme mostrado na figura 4.4)

e tiveram o mesmo lag. Ou seja, mesmo estando mais próximo ao reservatório os postos de Porto Pará e Porto Mesquita não proporcionaram uma maior defasagem com a vazão natural que o posto de Ponte dos Vilelas Jusante e Ponte Nova do Paraopeba.

Observa-se que os lags obtidos para os postos de Porto das Andorinhas e de Iguatama, que estão na mesma sub-bacia, são bem próximos, ficando o primeiro com um lag maior, por estar mais próximo ao reservatório, conseqüentemente, com uma área de drenagem maior. Esses postos estão no curso principal do rio São Francisco e são os que têm maior representatividade e defasagem temporal com a vazão natural de Três Marias.

Esta tabela também mostra que a vazão natural de Três Marias tem uma relação temporal com até 7 dias de precipitação média da área.

Foi calculada para os postos da CEMIG a matriz de correlação usando-se o preenchimento de falhas dos postos fluviométricos através da regressão linear simples e a chuva média (figura 5.30). Nesta figura pode-se observar uma correlação positiva entre a vazão natural de Três Marias e a vazão do posto fluviométrico de Porto das Andorinhas. Também se observa uma grande correlação negativa entre a vazão natural e a vazão dos postos fluviométricos de Porto Indaiá, Ponte Nova do Paraopeba, Ponte dos Vilelas Jusante e entre a precipitação média da área.

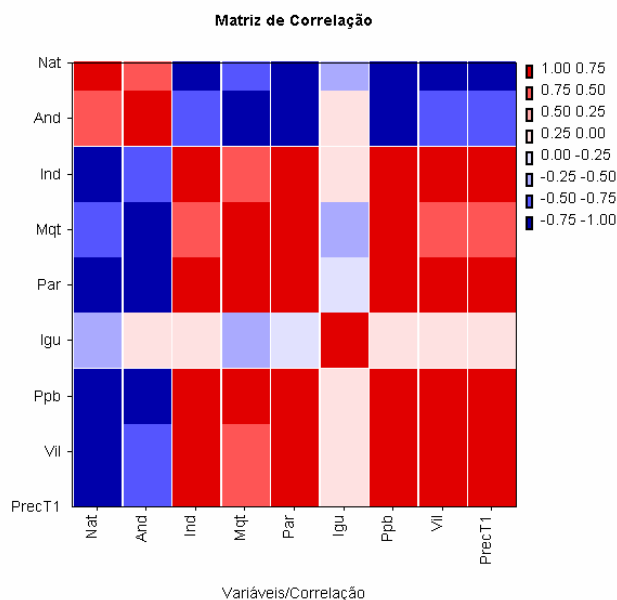


Figura 5.30 - Matriz de Correlação de dados semanais de postos da área a montante da UHE Três Marias.

Com os postos da CEMIG e da ANA, foi obtida a matriz de correlação usando-se o preenchimento de falhas através da regressão linear múltipla e a chuva média calculada por Thiessen (figura 5.31). Destaca-se a melhor correlação obtida entre a vazão natural e o posto de Iguatama, que tinham uma correlação, com a regressão simples, no intervalo de -0,25 a -0,50 e passaram a ter uma correlação no intervalo de -0,75 a -1,0.

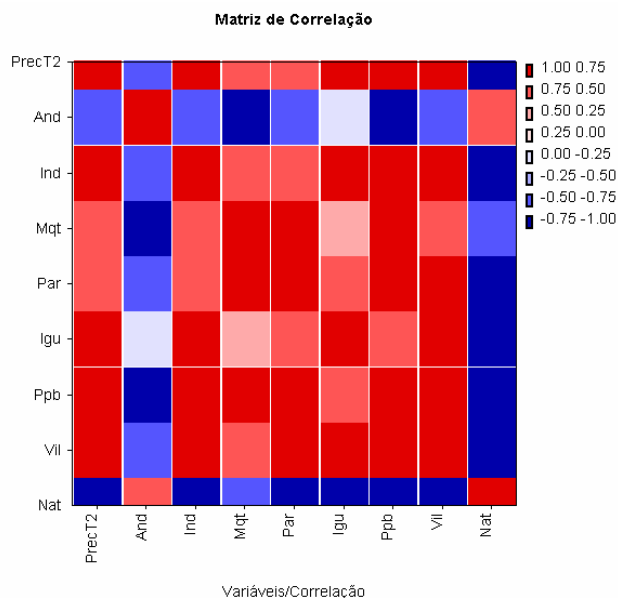


Figura 5.31 - Matriz de Correlação de dados semanais de postos da área a montante da UHE Três Marias.

Substituindo-se a chuva diária média espacial calculada por Thiessen pela calculada por Kriging não se observou diferença na matriz de correlação.

A forma gráfica mais usada para representar o resultado final dos diversos agrupamentos é o dendrograma. Nela estão dispostas linhas ligadas segundo os níveis de similaridade que agruparam pares de variáveis. Obteve-se o dendrograma (figura 5.32) usando-se o preenchimento de falhas através da regressão linear simples e a chuva média calculada apenas com os postos da CEMIG, onde se observa que a vazão natural não se agrupa diretamente com nenhuma outra informação analisada (postos fluviométricos e precipitação média da área), mas se agrupa com a junção dessas variáveis entre si.

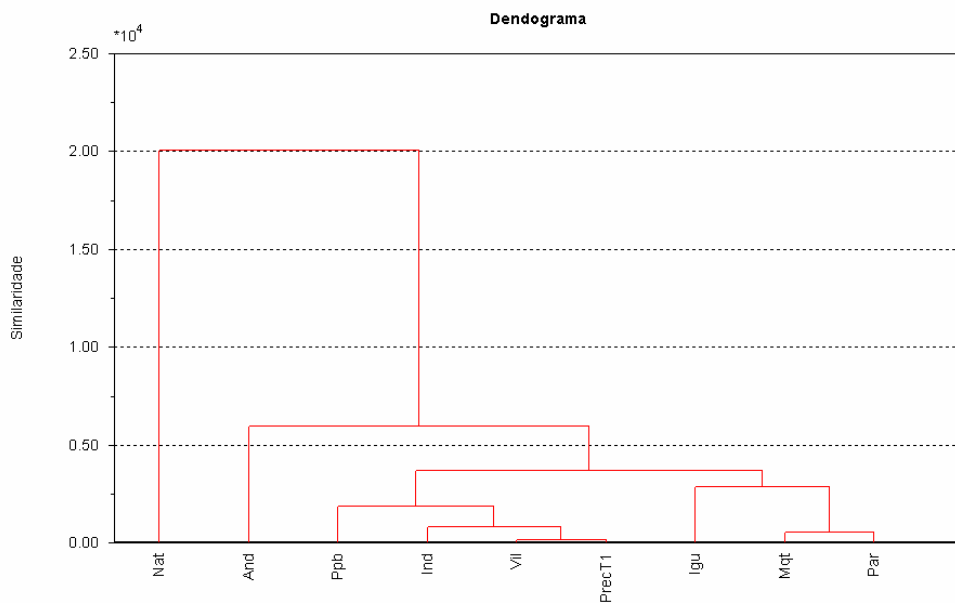


Figura 5.32 – Dendrograma com informações diárias representativas da área a montante da UHE de Três Marias.

Também foi obtido o dendrograma (figura 5.33) usando-se o preenchimento de falhas através da regressão linear múltipla e a chuva média calculada por Thiessen, com os postos da CEMIG e da ANA.

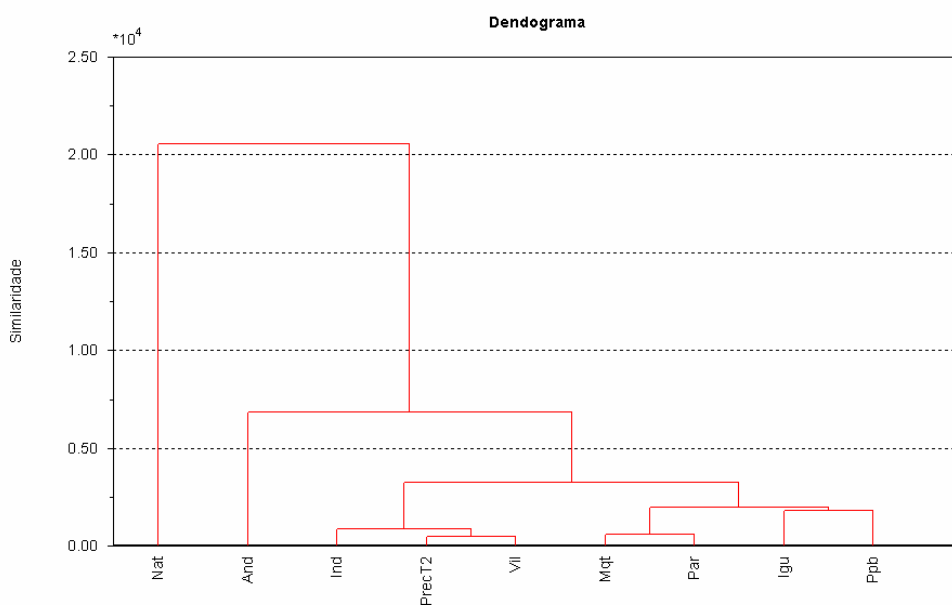


Figura 5.33 – Dendrograma com informações diárias representativas da área a montante da UHE Três Marias.

Das variáveis mostradas nas figuras 5.32 e 5.33 as de Ponte dos Vilelas Jusante, Iguatama e a chuva média foram as que sofreram as principais alterações: as duas primeiras devido ao preenchimento de falhas e a última, devida à metodologia de cálculo. Nestas figuras, observa-se que as informações de Ponte dos Vilelas Jusante e a chuva média continuaram com um agrupamento direto, contudo, aumentou a distância entre eles, o que representa uma redução na semelhança entre as séries. Nessas figuras, observa-se que os postos de Iguatama e Ponte Nova do Paraopeba reduziram a similaridade, o que representa um aumento na semelhança entre as séries. Observa-se também que aumentou a distância entre a vazão natural de Três Marias com o agrupamento das demais séries, com uma conseqüente redução da semelhança entre as séries.

Substituindo-se a chuva média calculada por Thiessen pela calculada por Kriging não se observou diferença significativa no dendograma obtido.

Efetou-se a Análise de Componentes Principais (ACP) (figura 5.34) usando-se o preenchimento de falhas através da regressão linear simples e a chuva média calculada com os postos da CEMIG. A componente CP 1 (96,84%), a CP2 (1,94%) e a CP3 (0,51%) juntas explicam 99,29% da variância total da série (e 0,71 representa o resíduo, ou seja, é explicada por nenhuma das três componentes). A CP1 é composta, principalmente, pela informação da vazão natural de Três Marias, da vazão no posto fluviométrico de Ponte dos Vilelas e Porto Indaiá, e da precipitação média da área. A CP2 é formada pela transformação linear das vazões dos postos fluviométricos de Porto das Andorinhas, Porto Indaiá, Porto Mesquita, Porto Pará, Iguatama e Ponte Paraopeba. Observa-se que a vazão natural é a variável mais importante para explicar os eventos de grande variação, sendo seguida, respectivamente, pela precipitação média da área e pelas informações dos postos fluviométricos de Ponte dos Vilelas, Porto Indaiá e Ponte Nova do Paraopeba.

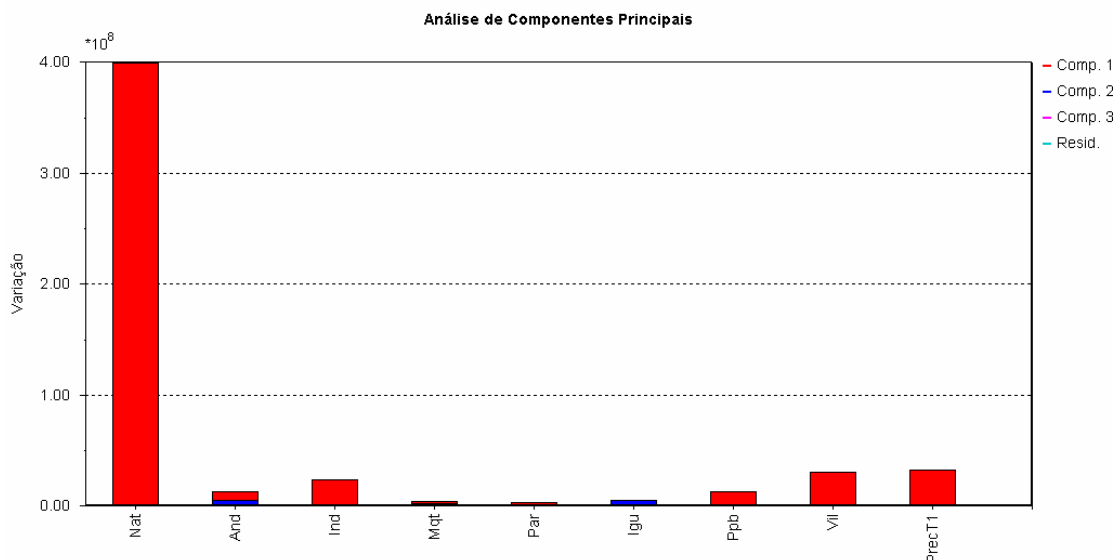


Figura 5.34 – Análise de Componentes Principais com dados semanais representativos da área a montante da UHE Três Marias.

A figura 5.35 mostra a Análise de Componentes Principais (ACP) obtida usando-se o preenchimento de falhas dos postos fluviométricos através da regressão linear múltipla e a chuva média calculada por Thiessen, com os postos da CEMIG e da ANA. A componente CP 1 (97,68%), a CP2 (1,51%) e a CP3 (0,43%) juntas explicam 99,62% da variância total da série (e 0,38 representa o resíduo, ou seja, é explicada por nenhuma das três componentes). Observa-se que a utilização da regressão linear múltipla no preenchimento de falhas e a o cálculo da chuva média também com os postos da ANA explicaram melhor a variância total da série (de 99,29 para 99,62%).

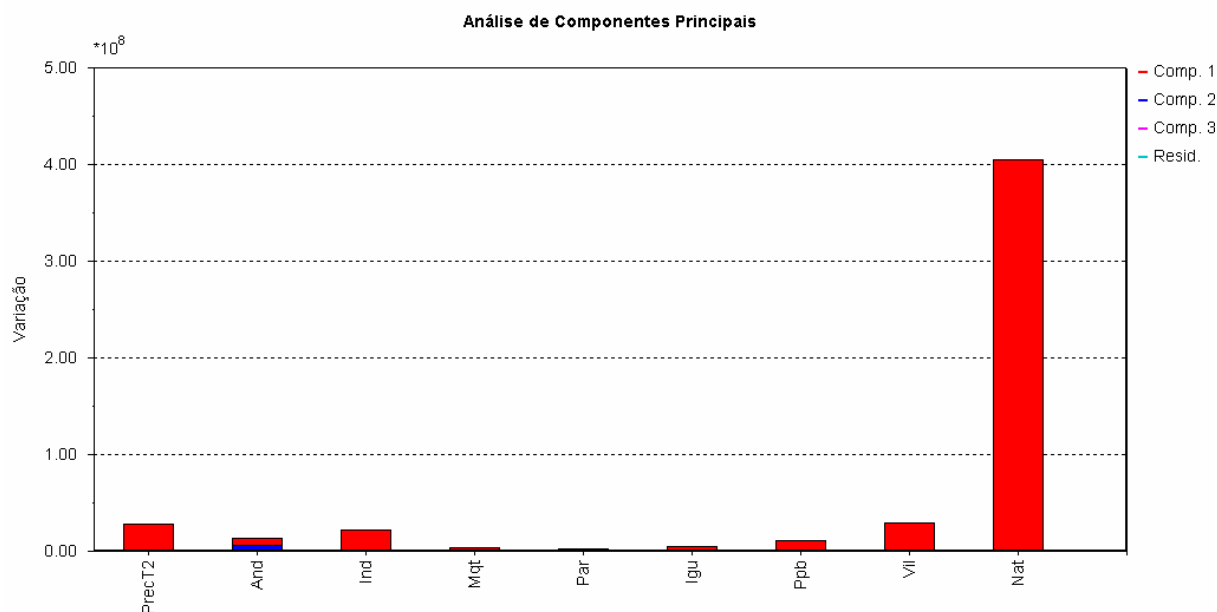


Figura 5.35 – Análise de Componentes Principais com dados semanais representativos da área a montante da UHE Três Marias.

Na Análise de Componentes Principais (ACP) efetuada usando-se o preenchimento de falhas dos postos fluviométricos através da regressão linear múltipla e a chuva média calculada por Kriging, com os postos da CEMIG e da ANA, obteve-se a componente CP 1 (97,66%), a CP2 (1,51%) e a CP3 (0,43%) que juntas explicam 99,60% da variância total da série. Observa-se que a utilização da chuva média por Kriging reduziu sensivelmente a explicabilidade da variância total da série (de 99,62 para 99,60%).

Comparando as figuras 5.17 e 5.18, observa-se que, na primeira, a vazão do posto fluviométrico de Iguatama precisava de uma nova componente (componente 2) para explicar a variância da sua série. Contudo, após efetuar-se o preenchimento de falhas através da regressão múltipla, essa variável se ajustou melhor à série, e foi explicada pela mesma componente das demais informações da bacia (componente 1).

Nas figuras de 5.36 a 5.43 são apresentadas as correlações semanal da vazão natural de Três Marias com os postos fluviométricos (com preenchimento de falhas através da regressão múltipla) e a precipitação média da área por Thiessen (usando os postos da CEMIG e da ANA).

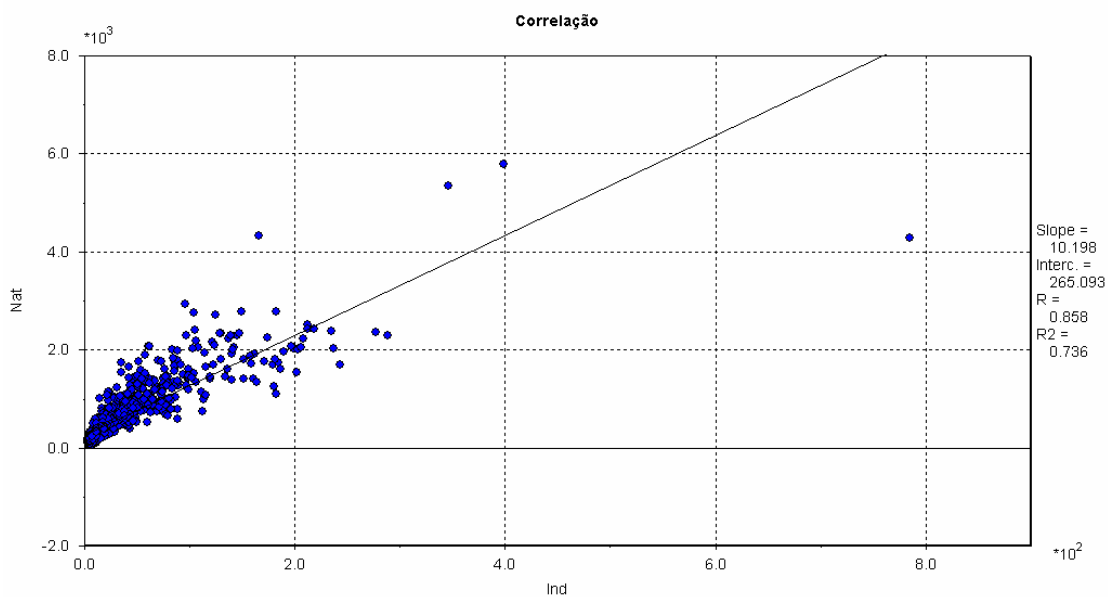


Figura 5.36 – Correlação semanal entre a vazão natural de Três Marias com o posto de Porto Indaíá.

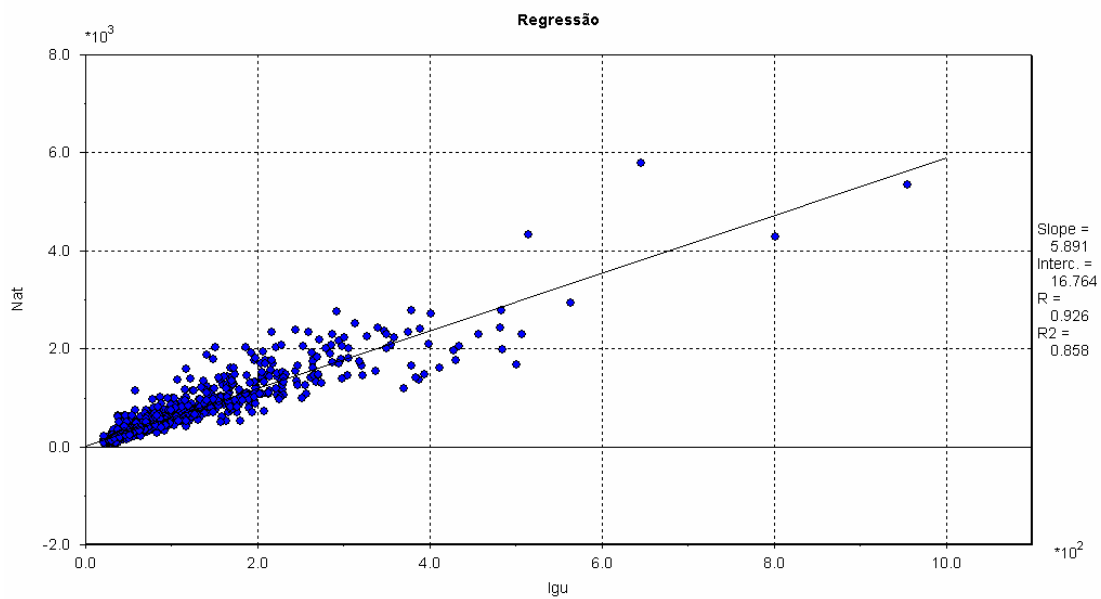


Figura 5.37 – Correlação semanal entre a vazão natural de Três Marias com o posto de Iguatama.

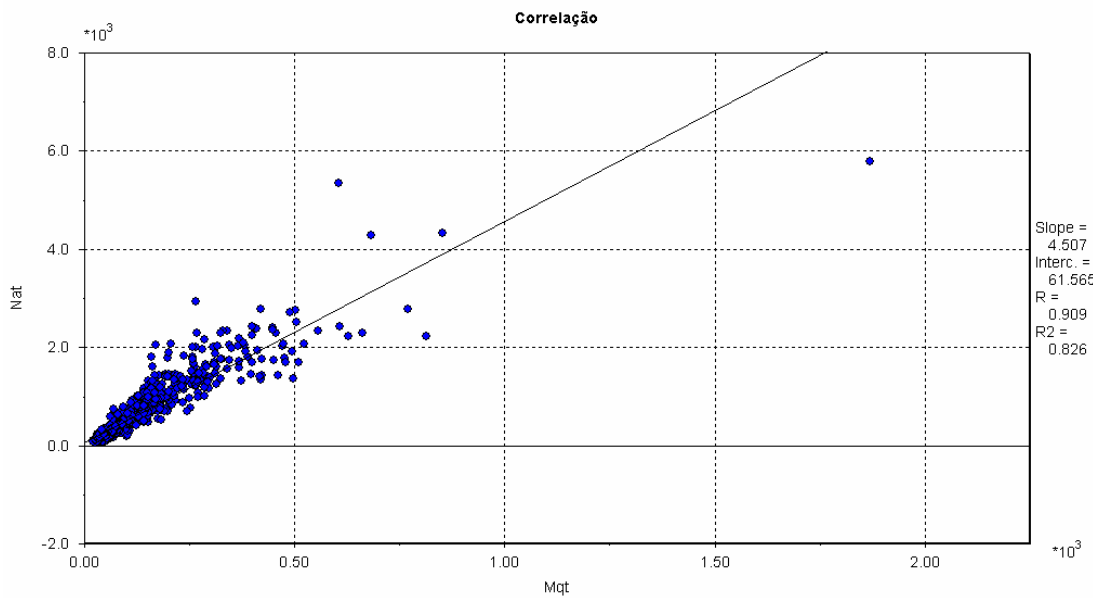


Figura 5.38 – Correlação semanal entre a vazão natural de Três Marias com o posto de Porto Mesquita.

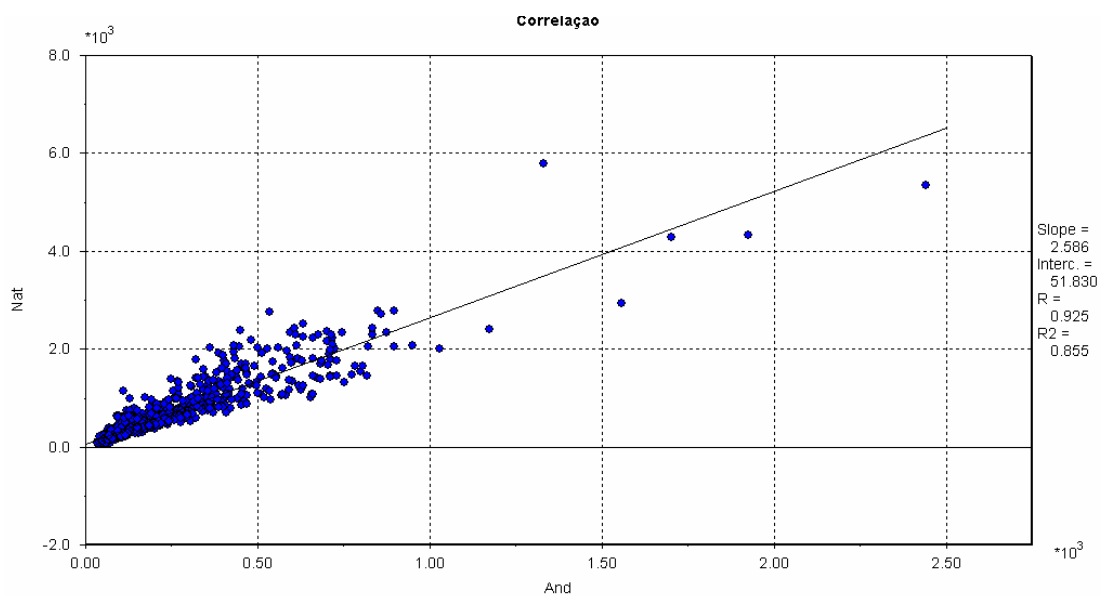


Figura 5.39 – Correlação semanal entre a vazão natural de Três Marias com o posto de Porto das Andorinhas.

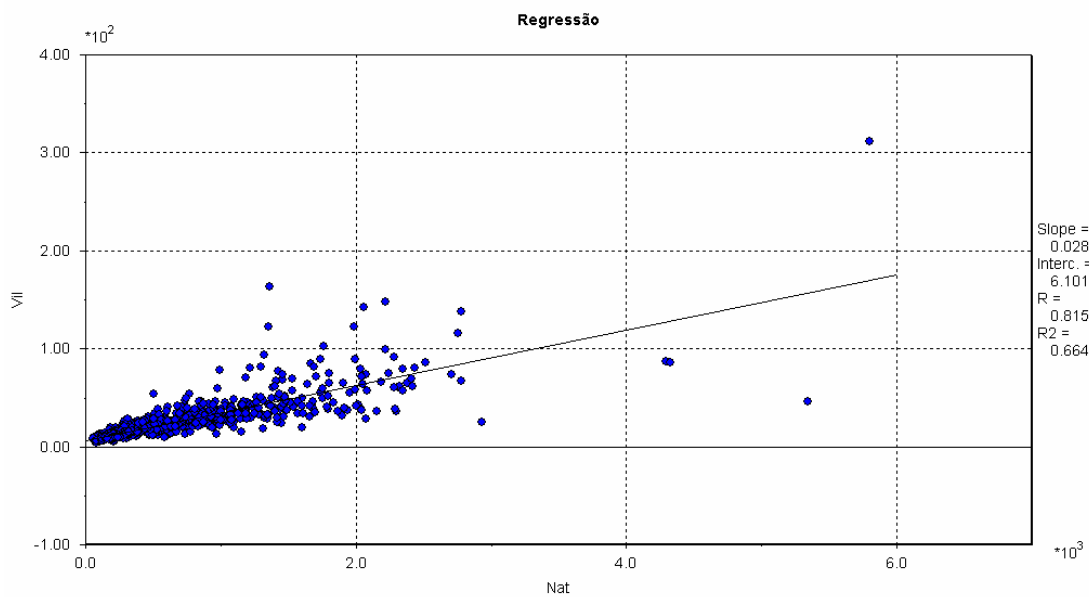


Figura 5.40 – Correlação semanal entre a vazão natural de Três Marias com o posto de Ponte dos Vilelas Jusante.

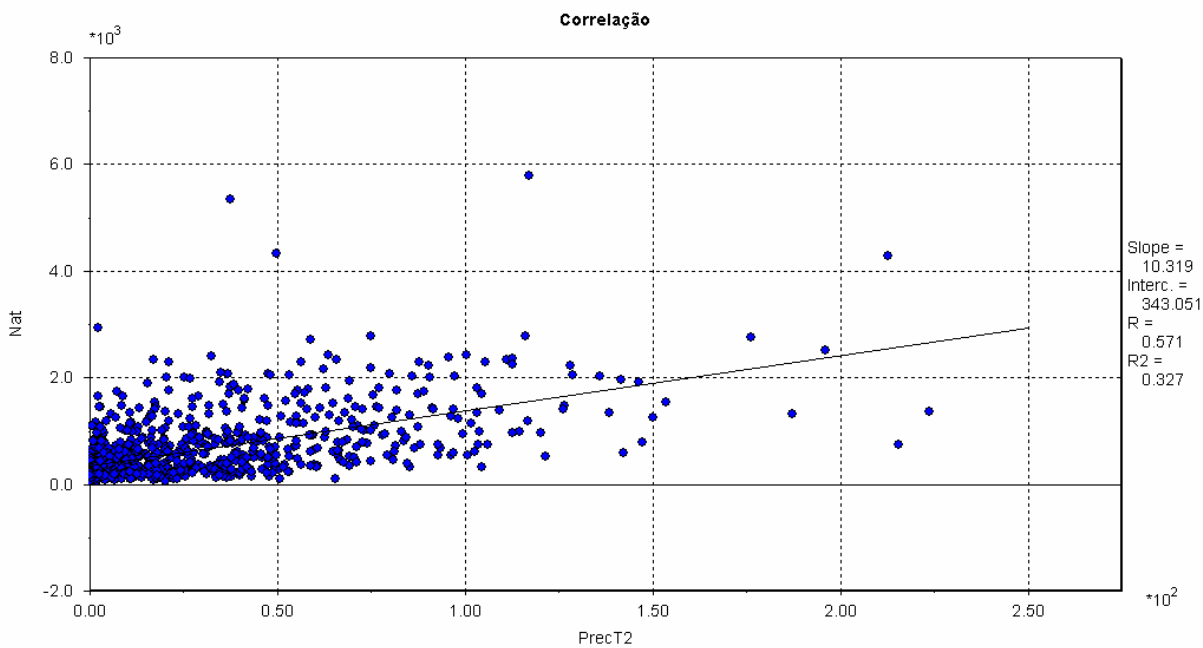


Figura 5.41 – Correlação semanal entre a vazão natural de Três Marias com a chuva média.

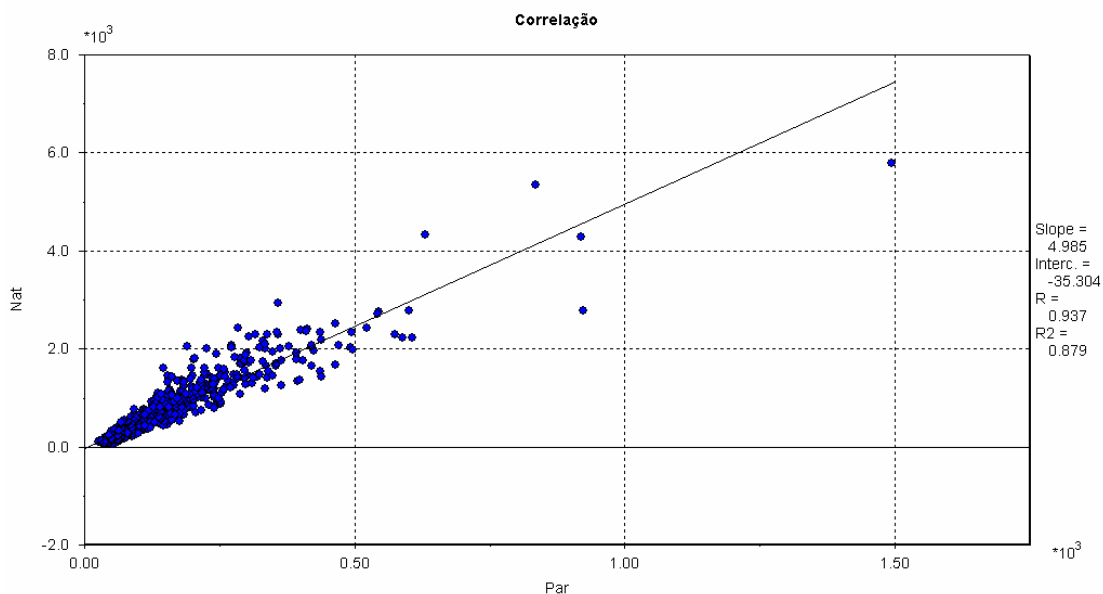


Figura 5.42 – Correlação semanal entre a vazão natural de Três Marias com a vazão de Ponte Nova do Paraopeba.

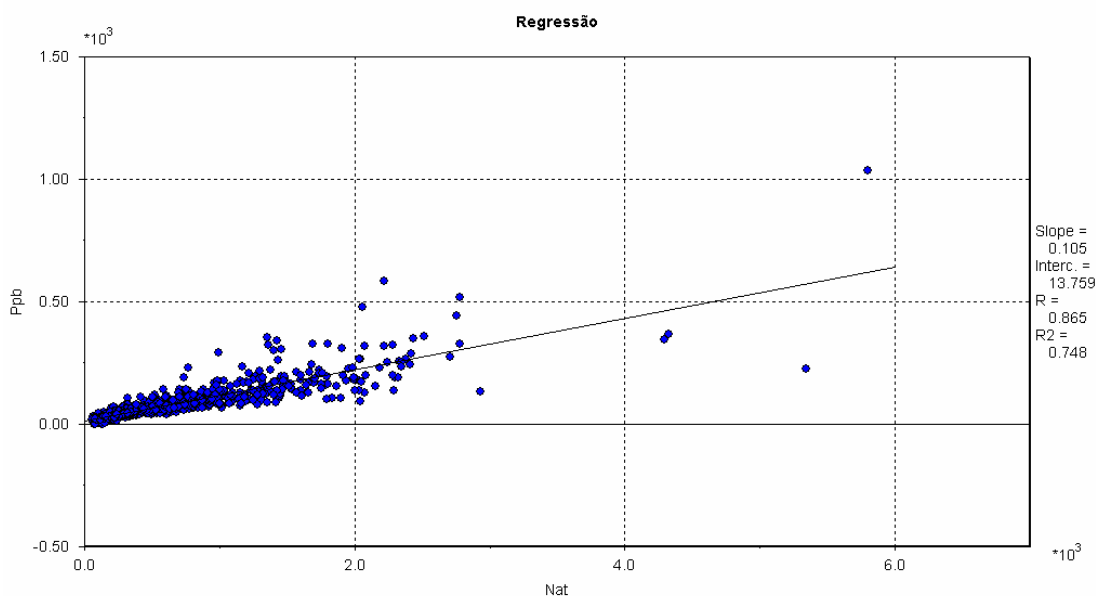


Figura 5.43 – Correlação semanal entre a vazão natural de Três Marias com a vazão de Porto Pará.

As figuras de 5.36 a 5.43 mostram que as séries dos postos fluviométricos usados como dados de entrada na rede neural têm uma boa correlação com a vazão natural de Três Marias, que é o dado de saída das simulações.

Foram elaboradas figuras da correlação semanal da vazão natural de Três Marias com a precipitação média da área por Thiessen (usando só os postos da CEMIG) e com a calculada por Kriging (figuras 5.44 e 5.45, respectivamente). Observa-se que o coeficiente de determinação do primeiro caso foi de 0,279 e do segundo 0,323 e, conforme a figura 5.41, com a chuva média por Thiessen (com os postos da ANA e CEMIG) o coeficiente de determinação foi de 0,327, ou seja, esta última opção apresentou uma melhor correlação com a vazão natural de Três Marias.

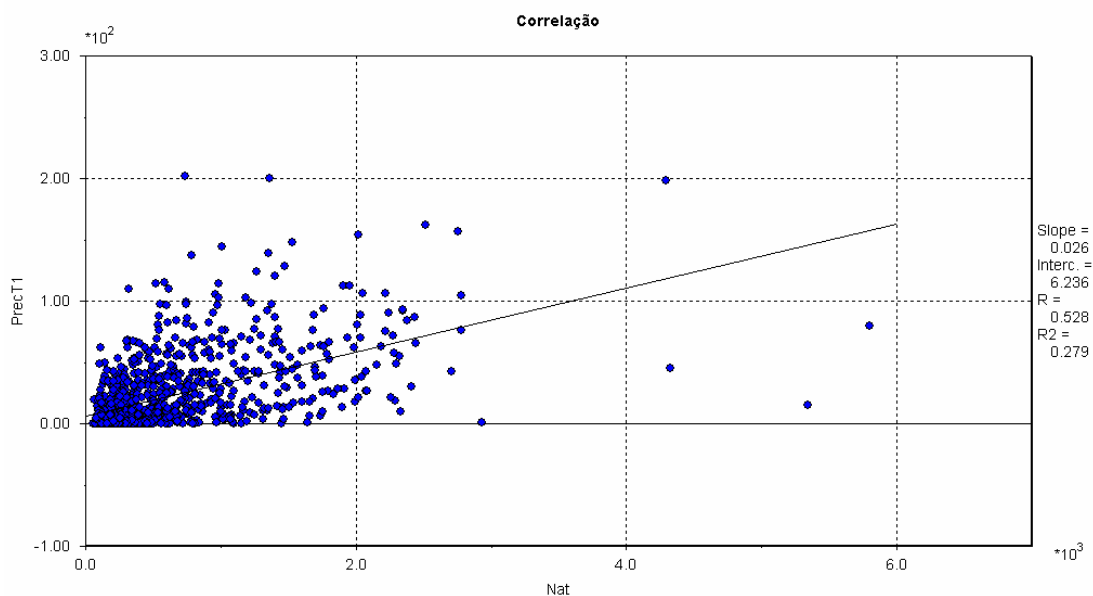


Figura 5.44 - Correlação de dados semanais da vazão natural de Três Marias com a chuva média calculada por Thiessen (com postos CEMIG).

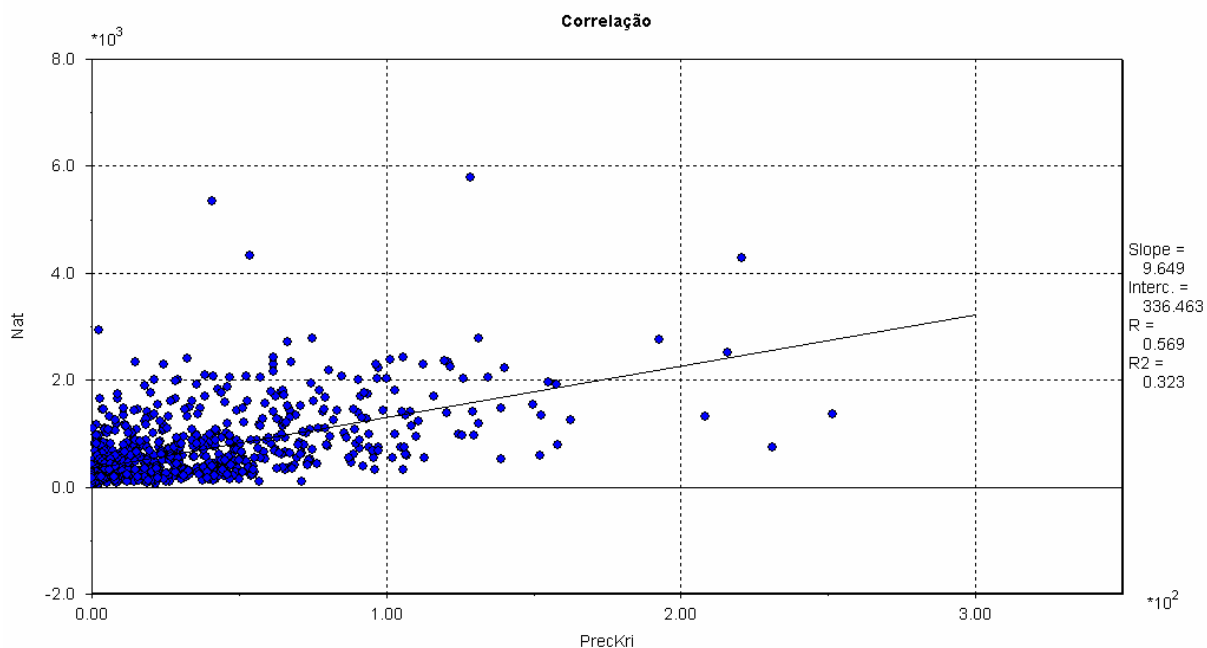


Figura 5.45 - Correlação de dados semanais da vazão natural de Três Marias com a chuva média calculada por Kriging.

As figuras 5.46 e 5.47 mostram, respectivamente, o dendograma e a ACP obtidos entre as informações dos postos fluviométricos, com o preenchimento de falhas através de regressão múltipla, e chuva média por Thiessen, com postos pluviométricos da ANA e CEMIG. Na ACP, a CP1 foi responsável por 88,20% da explicação da variabilidade da série, a CP2 por 8,31% e a CP3 por 1,17%. Juntas essas três componentes explicam 97,69% da variabilidade total da série, número sensivelmente inferior aos obtidos anteriormente, quando a série de vazão natural de Três Marias estava na análise.

Nessas figuras observa-se que, excluindo-se a vazão natural de Três Marias da análise, o posto de Porto das Andorinhas passa a ser o posto responsável pela explicação da variabilidade das séries.

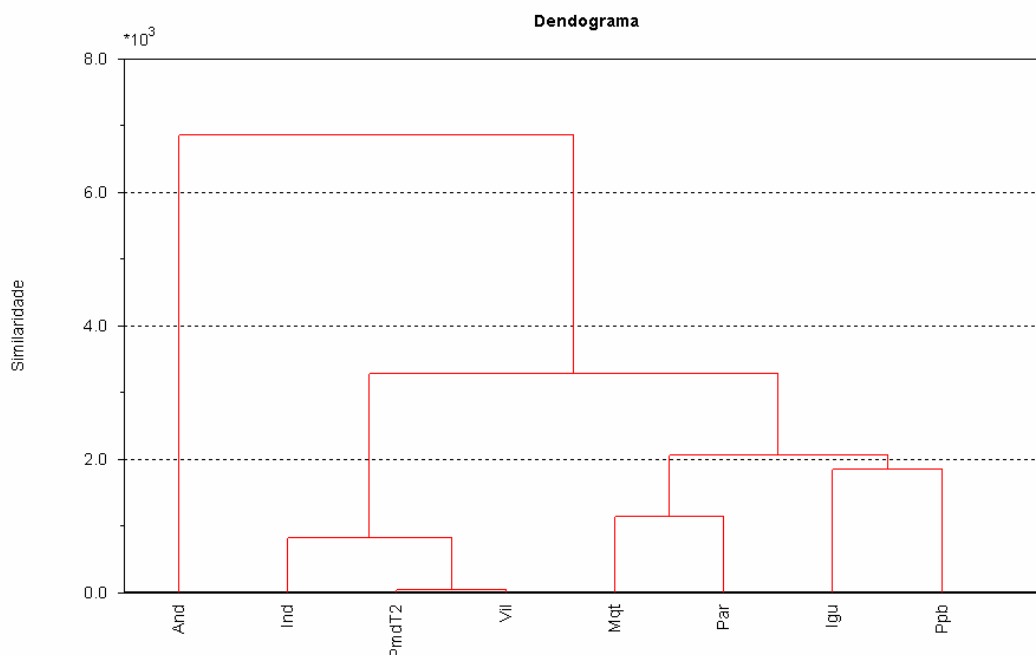


Figura 5.46 – Dendrograma com informações diárias representativas da área a montante da UHE Três Marias.

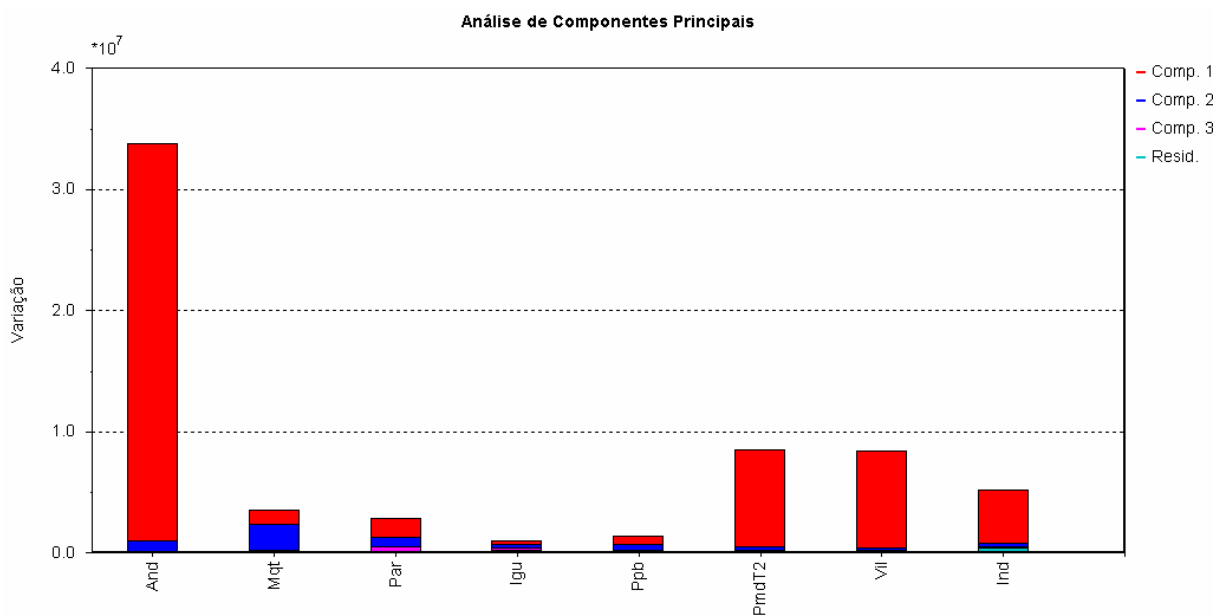


Figura 5.47 – Análise de Componentes Principais com dados semanais representativos da área a montante da UHE Três Marias.

Assim, baseado nas análises dos dados mostrados em todo o item 5.1, observa-se que:

- A chuva média calcula pelo método de Thiessen com os postos da CEMIG da e ANA foi a que melhor representou a bacia analisada;
- As melhores informações fluviométricas foram obtidas a partir do preenchimento de falhas com a regressão linear múltipla;
- Não foi observada sazonalidade nas informações dos postos fluviométricos, pluviométricos, nem na vazão natural ao reservatório de Três Marias;
- Os postos que melhor representam a variância total das séries são os de Porto das Andorinhas, Porto Indaiá, Ponte Nova do Paraopeba, Ponte dos Vilelas Jusantes, a precipitação média da área e a vazão natural de Três Marias;
- Os postos de Ponte Nova do Paraopeba, Ponte dos Vilelas Jusantes e Iguatama têm uma defasagem temporal, em relação à vazão natural de Três Marias, próximas às defasagens dos postos de Porto das Andorinhas, Porto Indaiá, Porto Mesquita e Porto Pará;

As melhores informações obtidas neste item 5.1, resumidas no parágrafo anterior, servirão de base para as informações de entrada das simulações com as redes neurais, mostradas no item 5.2, a seguir.

5.2 SIMULAÇÕES COM AS REDES NEURAIS

As figuras 5.48 e 5.49 mostram, respectivamente, os resultados da melhor rede neural MLP do período úmido, em termos de EPMA dos dados de validação, e os correspondentes erros de teste. As tabelas 5.9 e 5.10 mostram esses erros considerando o EPMA, EP e EMA em termos diários.

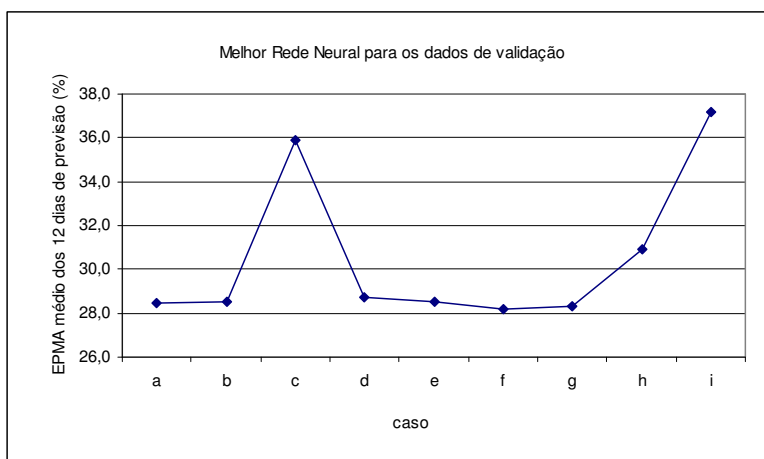


Figura 5.48 – EPMA das melhores RN do período úmido – dados de validação.

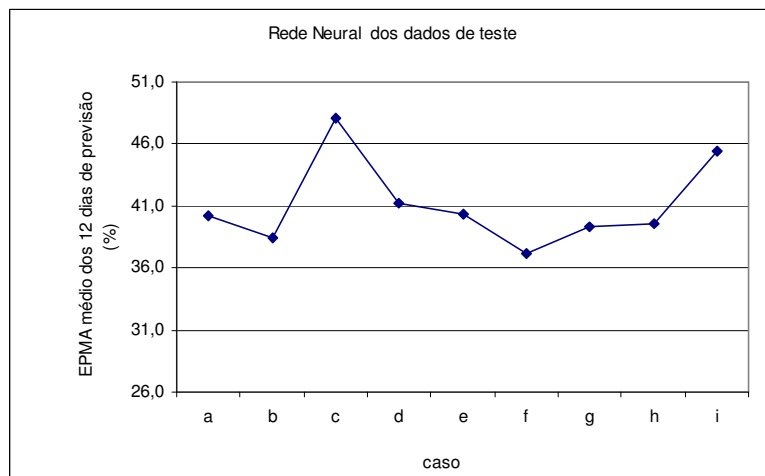


Figura 5.49 – EPMA das melhores RN do período úmido – dados de teste.

Tabela 5.9 – Resultados de validação da rede neural do período úmido.

caso	n° neurônios			EPMA médio 12 dias	erro																	
	entrada	tansig intermediária	purelin saida		Dia1			Dia2			Dia3			Dia4			Dia5			Dia6		
					EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA
a	35	2	12	28,5	9,6	220,8	105,9	11,2	243,9	121,4	14,3	285,7	147,9	18,2	344,2	186,9	22,3	414,6	229,9	27,4	485,3	278,0
b	19	2	12	28,6	12,8	223,8	126,8	13,0	216,4	133,8	15,1	251,5	148,1	18,3	314,6	185,0	21,5	395,6	225,0	26,5	476,2	272,1
c	11	2	12	35,9	12,0	182,9	114,8	16,7	289,2	170,3	23,0	397,7	232,8	29,2	488,6	286,3	34,1	555,5	327,4	38,0	609,1	358,9
d	14	2	12	28,7	11,1	208,7	111,1	12,0	227,9	127,1	14,8	262,6	150,3	18,5	317,2	189,0	22,5	393,4	229,8	27,6	473,2	277,3
e	12	5	12	28,5	11,5	301,1	124,3	12,6	302,7	136,6	15,7	333,8	160,4	18,6	359,0	193,2	22,2	419,2	231,1	27,0	487,2	275,0
f	9	3	12	28,2	9,3	192,8	99,2	14,6	238,2	134,8	15,1	290,3	152,2	18,7	350,8	189,6	21,7	408,1	221,0	27,0	485,5	269,6
g	12	3	12	28,3	10,7	186,3	112,3	13,3	231,6	139,8	15,7	274,3	155,7	18,4	335,1	187,9	20,8	396,8	220,1	25,7	474,7	264,4
h	8	3	12	30,9	16,4	214,5	153,6	14,4	261,0	150,6	15,6	354,4	177,5	20,2	425,9	225,2	26,0	501,7	273,1	32,0	566,9	316,4
i	3	2	12	37,2	10,0	175,2	104,5	16,8	310,0	177,5	23,7	413,9	237,3	30,2	500,2	290,8	35,4	571,5	331,0	39,9	631,2	366,3

caso	n° neurônios			EPMA médio 12 dias	erro																	
	entrada	tansig intermediária	purelin saida		Dia7			Dia8			Dia9			Dia10			Dia11			Dia12		
					EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA
a	35	2	12	28,5	32,3	541,6	317,4	36,2	586,0	349,1	39,3	625,1	374,3	42,0	656,0	395,6	43,8	680,8	408,4	45,0	698,4	418,1
b	19	2	12	28,6	31,7	540,0	313,9	35,8	588,3	346,7	39,1	628,5	374,0	41,6	660,7	395,4	43,1	685,8	408,3	44,2	702,5	418,4
c	11	2	12	35,9	41,2	651,5	385,5	43,9	682,0	406,7	45,9	703,6	421,0	47,5	719,9	433,2	49,0	734,1	444,6	50,1	746,0	456,4
d	14	2	12	28,7	32,6	535,8	318,7	36,6	583,5	351,5	39,6	623,3	375,8	41,8	654,5	393,8	43,3	678,8	405,9	44,3	696,5	414,7
e	12	5	12	28,5	32,0	549,3	317,4	36,2	594,5	348,9	39,0	630,6	372,0	41,2	661,2	394,0	42,7	680,9	403,9	43,8	697,8	411,1
f	9	3	12	28,2	33,7	551,9	316,1	35,2	607,6	341,7	38,2	646,9	369,8	40,0	679,0	391,6	41,6	704,2	408,8	43,1	720,9	422,5
g	12	3	12	28,3	31,6	566,1	310,6	35,5	600,5	342,2	39,2	653,1	370,8	41,6	670,7	392,2	43,0	689,9	404,2	44,1	706,7	415,8
h	8	3	12	30,9	36,0	616,6	348,5	38,4	655,4	371,3	40,7	685,8	391,0	42,0	711,9	407,9	43,7	732,1	424,9	45,2	747,8	439,9
i	3	2	12	37,2	43,3	677,9	394,8	46,2	710,6	418,5	48,2	732,5	434,6	49,7	747,3	448,0	51,1	758,2	459,6	51,9	765,4	469,2

Tabela 5.10 – Resultados do teste, com rede correspondente aos dados de validação.

caso	n° neurônios			EPMA médio 12 dias	erro																	
	entrada	tansig intermediária	purelin saida		Dia7			Dia8			Dia9			Dia10			Dia11			Dia12		
					EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA
a	35	2	12	40,3	45,6	460,3	318,1	51,2	492,1	347,7	55,4	511,1	368,7	59,3	527,3	386,4	62,5	540,2	400,8	65,0	550,8	413,7
b	19	2	12	38,5	42,1	456,7	308,1	48,2	488,4	339,7	52,5	507,9	361,7	56,0	523,3	377,7	59,1	536,0	392,8	61,4	547,0	405,1
c	11	2	12	48,1	52,8	526,9	374,8	56,1	539,4	389,2	59,5	552,0	404,5	63,0	558,6	417,8	66,5	563,9	429,9	69,2	572,4	439,7
d	14	2	12	41,2	45,6	457,1	313,2	51,3	490,8	344,6	55,8	512,6	368,5	59,5	530,8	387,9	62,5	547,8	404,7	64,7	559,8	416,5
e	12	5	12	40,4	43,1	448,1	304,8	50,6	485,5	340,3	54,8	508,9	363,4	55,7	525,2	379,0	60,7	541,4	398,8	64,9	555,3	414,7
f	9	3	12	37,2	45,1	453,1	315,8	46,6	487,3	336,4	50,2	508,4	356,9	52,0	524,2	370,0	54,2	537,9	384,3	56,8	548,5	397,6
g	12	3	12	39,3	43,4	455,9	313,2	49,5	486,9	342,2	53,2	504,2	361,7	57,0	517,8	378,7	60,2	532,8	394,3	63,2	548,3	411,2
h	8	3	12	39,6	47,0	486,4	337,5	50,1	509,0	358,6	53,3	526,4	377,0	55,2	542,4	391,4	57,6	551,5	404,7	60,3	556,4	415,2
i	3	2	12	45,5	52,6	525,9	377,5	56,1	538,6	394,3	59,2	549,3	408,8	61,8	556,3	419,1	64,0	559,0	426,3	65,6	567,6	435,3

caso	n° neurônios			EPMA médio 12 dias	erro																	
	entrada	tansig intermediária	purelin saida		Dia1			Dia2			Dia3			Dia4			Dia5			Dia6		
					EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA
a	35	2	12	40,3	12,1	181,6	111,0	14,5	205,7	125,4	19,5	268,5	161,7	25,9	323,9	204,7	32,8	376,2	246,6	39,2	420,9	282,2
b	19	2	12	38,5	18,3	209,9	136,7	16,7	224,0	140,8	19,4	274,2	166,7	23,6	321,9	199,2	28,9	371,4	234,5	35,2	414,8	270,5
c	11	2	12	48,1	21,3	287,7	160,6	25,7	351,8	201,1	31,5	395,7	245,0	38,4	445,8	290,2	44,6	490,5	331,5	48,9	514,2	353,4
d	14	2	12	41,2	18,3	214,5	132,4	18,3	236,3	138,8	21,4	272,3	164,1	26,2	320,8	198,4	32,2	371,4	237,4	38,8	416,0	275,2
e	12	5	12	40,4	20,6	213,4	136,5	17,7	231,6	139,2	25,2	271,8	173,4	24,6	308,0	191,5	29,0	358,7	229,0	38,2	404,3	268,5
f	9	3	12	37,2	15,3	177,3	108,4	21,1	213,5	141,1	18,1	256,6	155,4	23,1	309,1	194,7	28,6	356,1	230,8	34,9	404,7	267,9
g	12	3	12	39,3	14,8	188,0	111,6	16,9	207,7	128,7	20,4	261,8	160,7	25,6	319,2	202,5	30,8	368,2	238,9	36,9	414,0	273,7
h	8	3	12	39,6	17,6	207,7	142,3	16,9	225,3	141,9	19,0	290,3	176,3	24,4	354,3	221,7	32,7	407,2	266,2	41,5	453,0	308,5
i	3	2	12	45,5	14,4	270,6	137,1	18,7	342,9	189,8	27,0	397,0	241,5	35,3	441,6	285,1	42,8	484,6	329,1	48,1	511,5	355,4

Analisando a figura 5.48, observa-se que, com exceção do caso "c" os resultados obtidos não são muito diferentes.

As figuras 5.50 e 5.51 mostram os EPMA dos 12 dias de previsão para todos os casos simulados da rede MLP do período úmido.

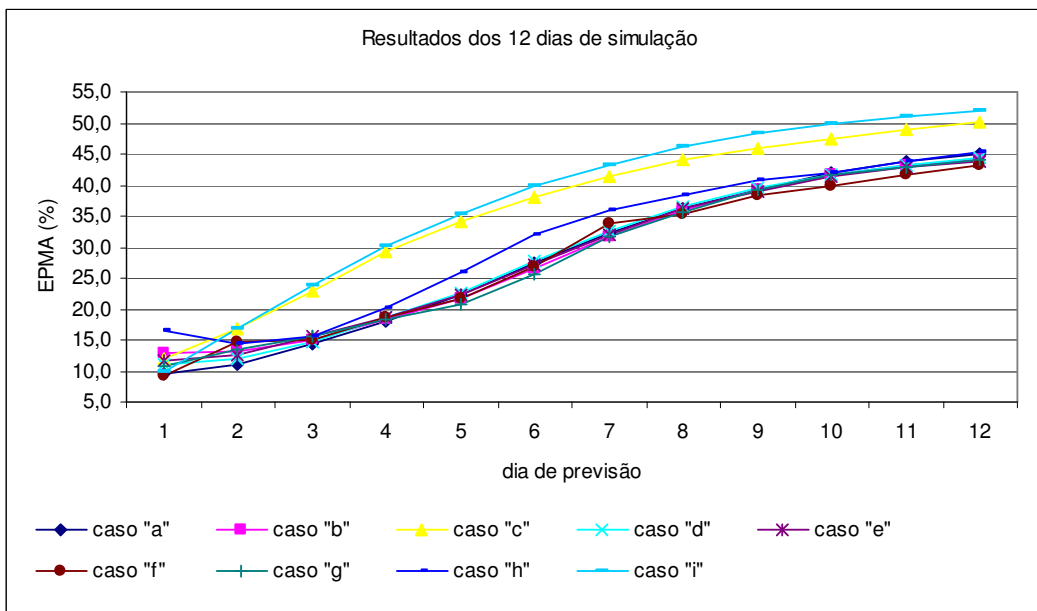


Figura 5.50 – EPMA dos 12 dias de previsão.

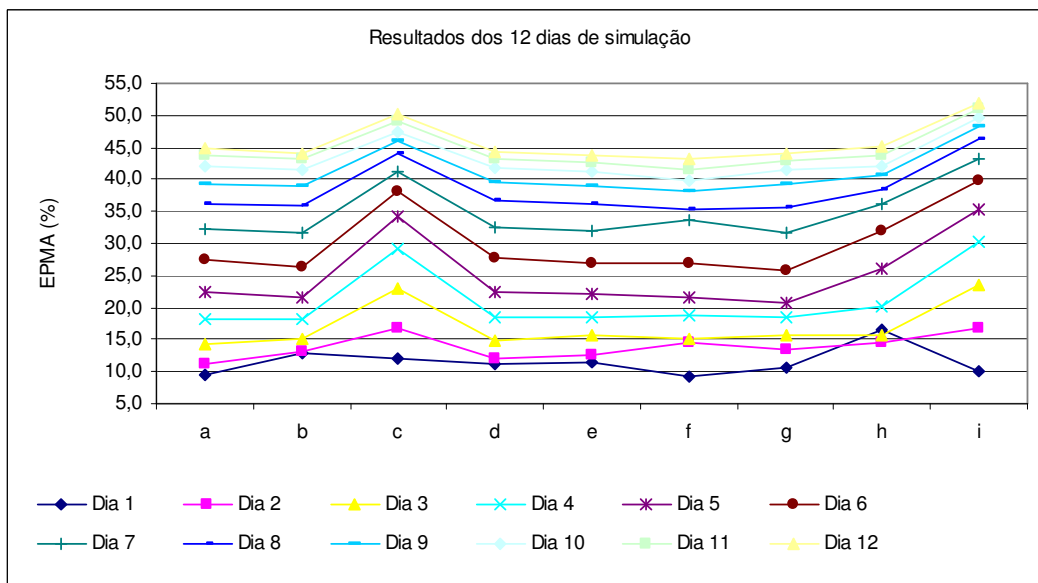


Figura 5.51 – EPMA dos 12 dias de previsão.

Analisando as figuras de 5.48 a 5.51 observa-se que o melhor modelo, em termos de EPMA médio dos 12 dias de previsão, é a simulação “f”, que considera apenas informações dos postos fluviométricos, pluviométricos e da vazão natural ocorrida, do dia em que se está fazendo a previsão. Comparando os desempenhos médios dos 12 dias de previsão, das simulações “f” e “g” observa-se:

- Em termos de EPMA: 28,2% - simulação “f” e 28,3% - simulação “g”;
- Em termos de EP: 496,6 m³/s - simulação “f” e 489,6 m³/s - simulação “g”;
- Em termos de EMA: 276,4 m³/s - simulação “f” e 276,3 m³/s - simulação “g”;

Assim, observa-se que o EPMA e EMA das duas simulações são bem próximos. Já em termos de EP, observa-se que a simulação “g” teve um desempenho um pouco melhor (7 m³/s). Contudo, a simulação “f” será a selecionada neste trabalho, visto que, como mostrado na figura 5.32, sua configuração leva a um erro de previsão para os dados de teste bem inferiores à simulação “g”.

Na simulação “d” foram colocados mais 3 dias de precipitação em relação à simulação “c”, onde se observa que os resultados foram bem melhores, mostrando a importância da precipitação para essas análises.

Na simulação “e” foram retirados 2 dias de vazão natural em relação à simulação “d”, onde se observa que os resultados pioraram um pouco para os primeiros dias de previsão, mas melhoraram um pouco para os últimos dias de previsão. Em termos médios dos 12 dias, o caso “e” teve um desempenho um pouco superior ao “d”.

A simulação “h” foi gerada a partir da “g” tirando as informações de precipitação. Observa-se que os resultados foram bem inferiores, mostrando, mais uma vez, a importância da precipitação nas simulações do período úmido.

Na simulação “i” foram consideradas apenas as informações de vazão natural do reservatório de Três Marias com um, dois e três dias de defasagem, onde se observa que os resultados da previsão foram de péssima qualidade.

As Figuras 5.52 e 5.53 mostram, respectivamente, os resultados da melhor rede neural MLP do período seco, em termos de EPMA dos dados de validação e os correspondentes erros de teste. As tabelas 5.11 e 5.12 mostram esses erros considerando o EPMA, EP e EMA em termos diários.

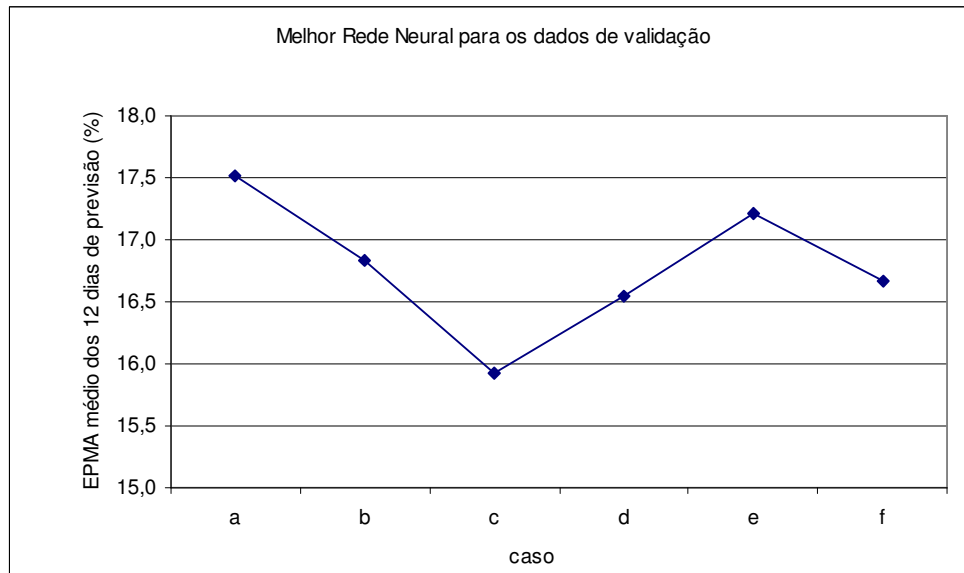


Figura 5.52 – EPMA das melhores RN do período seco – dados de validação.

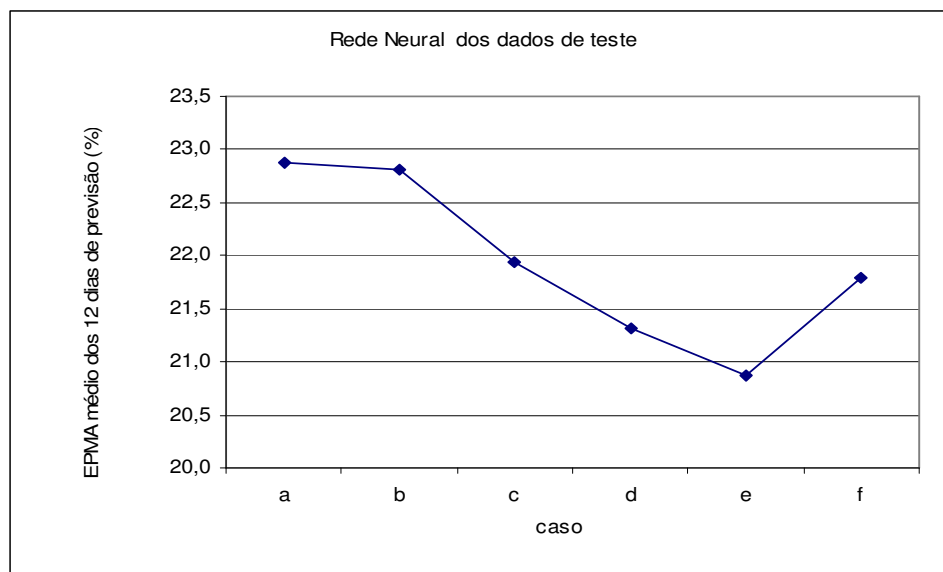


Figura 5.53 – EPMA das melhores RN do período seco – dados de teste.

Tabela 5.11 – Resultados de validação da rede neural do período seco.

caso	n° neurónios			EPMA médio 12 dias	erro																	
	entrada	tansig intermediária	purelin saída		Dia1			Dia2			Dia3			Dia4			Dia5			Dia6		
					EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA
a	35	5	12	17,5	8,5	32,2	21,1	7,2	30,1	17,9	8,0	37,8	19,6	11,2	59,4	28,3	14,4	76,4	38,1	16,9	97,3	46,7
b	12	4	12	16,8	7,9	30,2	18,9	7,5	30,3	18,9	9,7	41,8	24,6	9,7	57,0	27,1	14,1	70,7	35,2	14,8	90,6	41,1
c	21	4	12	15,9	7,2	29,4	17,5	7,6	27,0	17,0	7,4	37,5	19,7	9,5	54,9	25,1	11,4	69,4	31,3	14,8	88,6	40,1
d	23	2	12	16,6	12,7	46,8	30,6	11,7	42,6	28,3	12,9	48,1	30,0	12,3	61,7	32,4	12,0	75,2	35,7	13,9	95,2	42,1
e	17	2	12	17,2	5,9	26,9	15,0	7,1	31,1	17,9	9,2	45,0	23,6	11,7	64,8	30,4	14,4	80,9	37,5	16,8	100,2	44,7
f	3	3	12	16,7	6,5	24,5	15,1	5,6	36,0	17,1	9,0	52,6	25,5	11,4	71,4	32,5	13,9	85,8	39,0	16,4	104,4	46,2

caso	n° neurónios			EPMA médio 12 dias	erro																	
	entrada	tansig intermediária	purelin saída		Dia7			Dia8			Dia9			Dia10			Dia11			Dia12		
					EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA
a	35	5	12	17,5	19,2	115,3	53,8	21,1	125,0	59,6	23,0	132,5	64,7	25,3	138,6	70,7	26,6	143,3	74,3	28,6	148,5	79,8
b	12	4	12	16,8	17,6	109,0	49,0	21,7	121,1	58,3	20,6	131,4	62,0	24,1	135,9	67,4	24,7	140,6	70,3	29,8	147,5	77,4
c	21	4	12	15,9	16,5	109,2	46,8	19,3	120,0	54,7	21,4	128,9	60,6	22,8	135,6	64,3	25,6	141,6	69,8	27,5	147,4	74,0
d	23	2	12	16,6	17,0	114,8	50,3	18,1	126,7	55,7	20,1	135,1	61,0	21,1	139,9	64,0	22,5	144,0	67,5	24,3	148,7	71,3
e	17	2	12	17,2	19,1	118,2	51,7	21,3	129,5	58,2	23,2	137,7	63,8	24,7	142,8	67,6	26,0	146,7	71,0	27,3	150,9	74,8
f	3	3	12	16,7	19,2	121,6	53,8	20,9	132,4	59,4	22,2	139,7	64,1	23,6	144,1	67,7	24,9	147,4	71,2	26,3	151,1	74,9

Tabela 5.12 – Resultados do teste, com rede correspondente aos dados de validação.

caso	n° neurónios			EPMA médio 12 dias	erro																	
	entrada	tansig intermediária	purelin saída		Dia1			Dia2			Dia3			Dia4			Dia5			Dia6		
					EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA
a	35	5	12	22,9	12,0	30,5	21,3	12,0	32,7	21,7	14,2	37,8	25,3	17,1	43,4	30,0	19,0	46,2	32,5	20,4	48,9	34,4
b	12	4	12	22,8	12,3	32,7	21,6	13,1	35,7	23,5	17,8	45,2	30,8	15,7	42,2	28,5	23,0	51,0	37,3	20,8	51,2	35,4
c	21	4	12	21,9	12,8	33,7	23,9	14,6	37,1	26,3	13,6	38,6	25,8	15,2	42,6	28,1	18,0	46,6	31,6	21,6	52,0	36,8
d	23	2	12	21,3	14,7	36,7	25,4	15,5	38,1	26,9	16,5	40,5	28,7	16,9	42,3	29,9	18,5	44,4	31,9	19,5	46,4	32,9
e	17	2	12	20,9	10,2	29,6	18,9	11,9	33,4	21,6	14,1	37,5	25,3	16,8	41,9	29,2	19,4	46,2	32,7	20,7	48,7	34,2
f	3	3	12	21,8	14,8	45,7	28,0	14,3	48,0	28,1	15,9	50,6	31,1	17,5	50,7	33,2	20,4	54,1	37,0	21,4	55,8	38,3

caso	n° neurónios			EPMA médio 12 dias	erro																	
	entrada	tansig intermediária	purelin saída		Dia7			Dia8			Dia9			Dia10			Dia11			Dia12		
					EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA	EPMA	EP	EMA
a	35	5	12	22,9	22,8	52,9	37,9	25,7	58,4	42,5	28,6	64,1	47,3	33,2	73,5	54,5	33,3	74,3	55,2	36,3	81,1	59,9
b	12	4	12	22,8	23,7	55,4	39,6	28,6	62,3	45,8	24,0	60,4	42,1	29,0	65,4	47,4	29,2	67,2	48,6	36,6	75,4	56,4
c	21	4	12	21,9	21,8	52,1	36,0	25,3	57,8	41,2	27,7	62,0	44,6	28,6	64,2	45,4	31,3	68,4	48,5	32,6	70,8	49,9
d	23	2	12	21,3	22,5	49,9	36,4	23,1	53,5	37,7	25,1	57,3	40,3	26,3	60,7	42,5	27,7	63,5	44,2	29,6	66,0	46,3
e	17	2	12	20,9	22,2	51,1	36,1	24,0	54,5	38,5	26,0	59,3	41,7	27,3	60,4	43,6	28,3	62,7	45,2	29,5	64,9	47,0
f	3	3	12	21,8	23,5	57,7	41,1	24,2	59,5	42,3	25,6	62,2	44,6	26,6	64,2	46,1	28,1	67,0	48,7	29,2	69,0	49,9

As Figuras 5.54 e 5.55 mostram os EPMA dos 12 dias de previsão do período seco para todos os casos simulados.

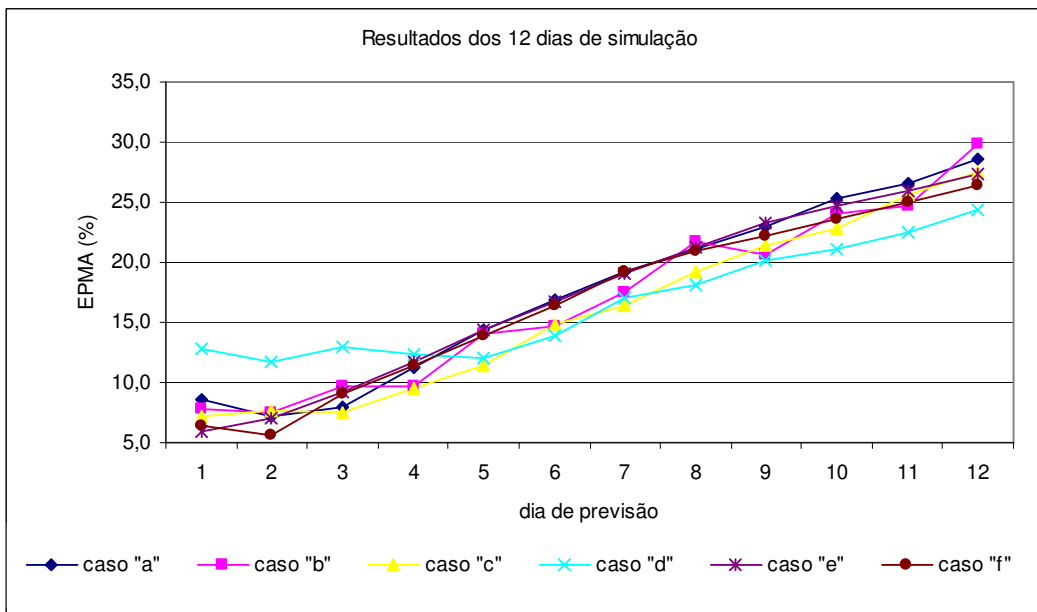


Figura 5.54 – EPMA dos 12 dias de previsão – análise por “casos”.

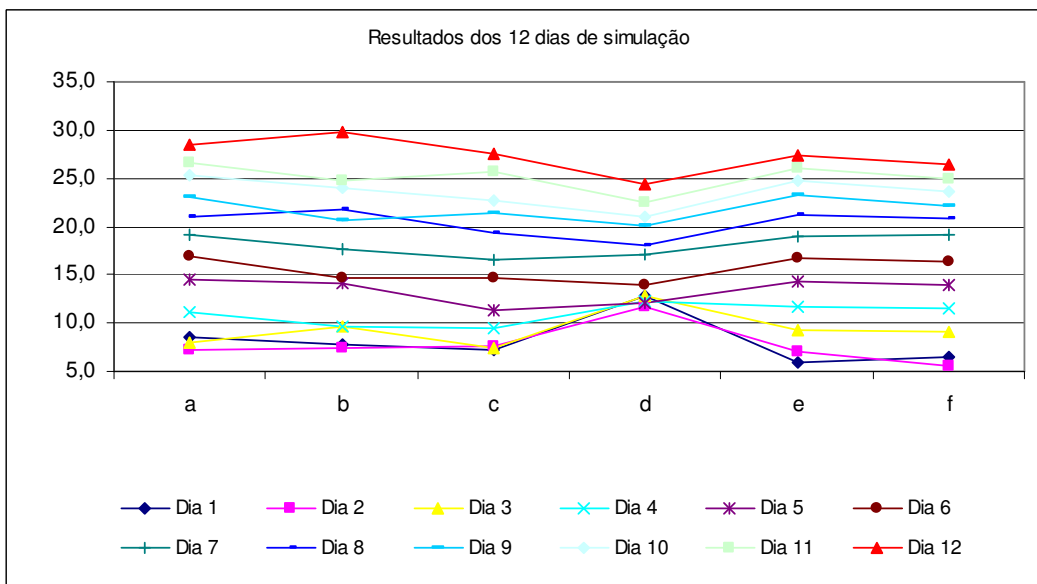


Figura 5.55 – EPMA dos 12 dias de previsão – análise por “dias”.

Analisando as figuras de 5.35 a 5.38 observa-se que o melhor modelo, em termos de EPMA médio dos 12 dias de previsão, é a simulação “c”, que considera informações dos postos fluviométricos, pluviométricos e da vazão natural ocorrida, para diversas defasagens.

A simulação “d” é formada pelas informações da simulação “c”, colocando-se dois dias de vazão natural. Observa-se que os resultados da previsão pioraram muito para os primeiros dias e melhoraram para os últimos dias. Para os 12 dias de previsão, os resultados pioraram.

A simulação “e” é formada pelas informações da simulação “c”, tirando-se apenas a precipitação. Observa-se que a previsão teve uma qualidade bem inferior, mostrando a importância da precipitação também para o período seco.

Na simulação “f” foram consideradas apenas as informações de vazão natural do reservatório de Três Marias com um, dois e três dias de defasagem, onde se observa que os resultados da previsão foram de qualidade bem inferior quando comparados com as previsões que consideram informações de postos pluviométricos e fluviométricos.

Selecionando a melhor rede neural MLP para os períodos úmido e seco e colocando na mesma base de dados do Programa Mensal de Operação - PMO, as tabelas 5.13 e 5.14 mostram o desempenho das previsões diárias e semanais, respectivamente; e a figura 5.56, a vazão semanal prevista e ocorrida. O Anexo 02 mostra essas vazões diárias e semanais previstas e ocorridas, através de gráficos anuais.

Tabela 5.13 – Erros da previsão de vazões diárias com os melhores modelos da Rede MLP.

Ano	EPMA (%)
2000	11,8
2001	17,9
2002	18,6
2003	14,1
2004	22,1
Média:	16,9

Tabela 5.14 – Erros da previsão de vazões semanais com os melhores modelos da Rede MLP.

Ano	EPMA (%)		
	úmido	seco	Média
2000	22,5	18,7	20,6
2001	45,3	27,0	36,1
2002	39,9	29,4	34,6
2003	34,2	19,5	26,9
2004	41,4	12,5	27,0
Média:	36,7	21,4	29,0

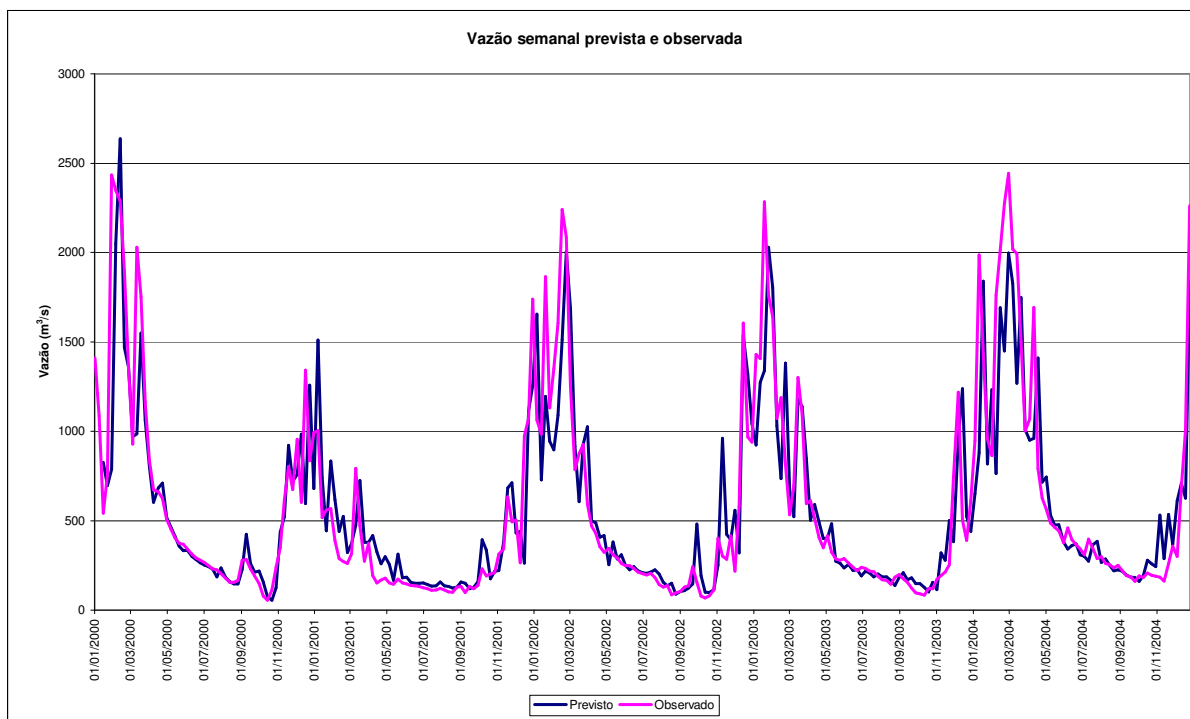


Figura 5.56 – Vazão semanal prevista e observada com a rede MLP.

As tabelas 5.15 e 5.16 mostram os parâmetros de entrada da rede NSRBN e os resultados encontrados para os 7 e 12 dias de previsão, respectivamente. Para as redes neurais com melhores desempenhos, essas tabelas também mostram os EPMA's dos primeiros 7 dias, dos últimos 5 dias e de todos os 12 dias de previsão.

Tabela 5.15 – Resultados das simulações do período úmido para os 7 dias de previsão.

Simul	Entradas							Obtidos			7 primeiros dias	
	%val	%teste	alfa	semente	beta	epoca min	epoca max	erro min	erro perc md abs	erro padrão	teste	validação
A	22	28	20	987654	40	100	500	1,53	18,75	242,02	23,5	20,4
	22	28	30	987654	50	200	400	1,51	18,96	260,46	23,2	19,0
	22	28	30	987654	30	200	400	1,60	18,68	250,33	23,3	20,6
B	22	28	50	987654	50	100	500	1,55	17,66	248,25	22,2	20,0
	22	28	30	987654	50	200	400	1,47	18,40	244,05	23,2	19,9
C	22	28	30	987654	50	200	400	2,75	27,15	331,76	30,1	25,6
D	22	28	30	987654	50	200	400	1,58	18,97	244,71		
	22	28	20	987654	60	200	400	1,57	19,2	248,31		
E	22	28	30	987654	50	200	400	1,78	20,01	257,84		
F	22	28	30	987654	50	200	400	1,67	23,74	283,72		
	22	28	50	987654	50	200	400	1,51	18,3	244,63	23,1	20,1
G	22	28	20	987654	50	200	400	1,47	17,51	232,19	22,1	19,7
	22	28	30	987654	50	200	400	1,64	20,57	245,22		
	22	28	50	987654	60	200	400	2,46	19,09	257,40		
	22	28	70	987654	50	200	400	1,45	19,7	246,71	24,6	20,2
H	22	28	20	987654	40	200	400	1,50	23,65	252,39	28,9	22,1
	22	28	30	987654	50	200	400	1,52	21,88	265,18	26,8	21,7
I	22	28	20	987654	50	200	400	1,51	18,91	239,65	23,7	20,3
	22	28	20	987654	50	300	500	1,51	19,4	242,75		
J	22	28	20	987654	50	200	400	1,77	19,43	252,37		
	22	28	30	987654	50	200	400	1,52	18,53	238,04	23,4	19,9
L	22	28	20	987654	50	200	500	1,45	19,26	247,47	24,1	20,4
	22	28	30	987654	50	200	400	1,45	19,61	247,59	24,4	20,8
M	22	28	20	987654	50	200	400	1,51	19,09	249,18		
	22	28	30	987654	50	200	400	1,67	18,95	261,12	22,8	19,1
N	22	28	20	987654	50	200	400	1,49	18,48	231,60	23,3	20,2
	22	28	30	987654	60	200	400	1,53	18,43	235,09		
O	22	28	20	987654	50	200	400	1,54	18,27	243,05		
	22	28	50	987654	60	200	400	1,86	18,27	252,60		

Tabela 5.16 – Resultados das simulações do período úmido para os 12 dias de previsão.

Simul	Entradas							Obtidos			12 dias		5 últimos dias		7 primeiros dias	
	%val	%teste	alfa	semente	beta	epoca min	epoca max	erro min	erro perc md abs	erro padrão	teste	validação	teste	validação	teste	validação
A	22	28	20	987654	50	100	500	4,92	28,72	320,70	36,7	29,8				
	22	28	30	987654	50	200	400	4,88	27,64	309,60	35,2	28,9				
	22	28	30	987654	30	200	400	5,04	27,06	313,18	34,5	29,4				
	22	28	30	987654	70	200	400	4,96	27,17	323,40	35,2	29,0				
B	22	28	20	987654	50	200	400	4,93	26,07	309,58						
	22	28	30	987654	50	200	400	4,99	24,79	304,13	31,6	27,4	43,6	37,2	23,0	20,4
C	22	28	30	987654	50	200	400	7,66	38,35	410,8						
D	22	28	30	987654	50	200	400	4,99	29,29	326,92	37,9	31,0	45,8	39,2	32,3	25,1
	22	28	20	987654	50	200	400	5,14	29,02	328,24						
E	22	28	20	987654	50	200	400	4,80	31,18	335,84	39,2	30,2	47,7	39,3	33,0	23,7
F	22	28	20	987654	50	200	400	4,72	29,24	315,73	37,3	29,1	46,6	38,2	30,6	22,5
	22	28	50	987654	50	200	400	5,24	27,81	343,38	34,5	30,0	45,0	38,6	27,0	23,8
G	22	28	20	987654	50	200	400	4,77	27,14	310,78						
	22	28	30	987654	50	200	400	4,70	28,24	322,43	36,1	29,1				
	22	28	50	987654	50	200	400	4,80	25,71	304,43	32,6	27,4				
	22	28	70	987654	50	200	400	4,97	27,93	334,74						
H	22	28	20	987654	50	200	400	4,65	29,65	317,22	37,9	29,4	46,9	38,4	31,5	23,0
	22	28	50	987654	50	200	400	4,46	29,52	319,59	37,6	28,3	47,9	37,5	30,2	21,8
I	22	28	20	987654	50	200	400	4,75	29,54	315,62	37,8	29,2	47,3	38,7	31,0	22,4
	22	28	20	987654	50	300	500	4,75	29,08	319,88						
J	22	28	20	987654	50	300	500	5,03	28,01	327,22	36,0	29,4	44,4	38,1	30,0	23,2
	22	28	20	987654	50	200	500	5,02	28,86	330,16	37,3	30,1	44,7	38,4	32,0	24,2
L	22	28	20	987654	50	200	500	4,45	28,39	319,98	35,9	27,8	45,9	36,5	28,8	21,5
	22	28	30	987654	50	200	500	4,56	29,80	338,85						
	22	28	30	987654	50	200	400	4,69	27,15	314,34	34,4	27,8	46,0	37,1	26,0	21,2
M	22	28	20	987654	50	200	400	4,72	27,85	321,26	35,6	29,0	44,7	37,3	29,2	23,0
	22	28	30	987654	50	200	400	4,74	26,13	306,42	33,4	28,1	45,4	37,3	24,8	21,5
N	22	28	20	987654	50	200	400	4,54	29,08	326,18						
	22	28	30	987654	50	100	500	5,90	31,17	337,23						
O	22	28	20	987654	50	200	400	4,43	30,70	326,68	39,7	29,8	47,6	38,0	34,0	23,9
	22	28	20	987654	20	200	400	4,48	28,84	323,21	37,8	29,4	45,8	37,9	32,1	23,3

Nas tabelas 5.15 e 5.16 observa-se que a alteração das variáveis básicas de entrada do programa NeuroHidro (alfa, semente, beta, épocas mínima e época máxima) os resultados não variam muito, o que facilita a simulação com esse modelo, visto que não são necessárias diversas simulações para ajuste desses parâmetros, como nas simulações com o software Matlab, para a rede MLP.

Ainda analisando as tabelas 5.15 e 5.16, observa-se que a calibração de uma rede para os 7 primeiros dias e depois outra para os 12 dias, os resultados de previsão dos 7 primeiros dias desse último são inferiores aos resultados da calibração direta, visto que a função objetivo da versão utilizada no NeuroHidro tem esse objetivo. Como exemplo observam-se os resultados da simulação L, em que, para os dados de validação da rede de 7 dias, tem um erro de 20,4 e 21,5, pegando os 7 primeiros dias da rede ajustada para os 12 dias.

A simulação “E” é formada pela simulação “D”, tirando-se dois dias de vazão natural. Analisando o erro mínimo, observa-se que as previsões pioraram para os primeiros 7 dias e melhoraram para os 12 dias de previsão.

A simulação “H” é formada pela simulação “G”, colocando-se três dias de precipitação. Analisando o erro mínimo, observa-se que as previsões pioraram para os primeiros 7 dias e melhoraram para os 12 dias de previsão.

A simulação “L” é formada pela simulação “G”, colocando-se cinco dias de precipitação. Analisando o erro mínimo, observa-se que as previsões tiveram o mesmo desempenho para os primeiros 7 dias e melhoraram para os 12 dias de previsão.

A simulação “M” é formada pela simulação “L”, colocando-se quatro dias de vazão natural. Analisando o erro mínimo, observa-se que as previsões pioraram para os primeiros 7 dias e para os 12 dias de previsão.

A simulação “N” é formada pela simulação “G” apenas com os postos mais representativos segundo a ACP, apresentados anteriormente na figura 5.18, ou seja, excluíram-se das simulações as informações dos postos fluviométricos de Porto Mesquita, Porto Pará e Iguatama. Analisando o erro mínimo, observa-se que as previsões pioraram para os primeiros 7 dias e melhoraram para os 12 dias de previsão.

A simulação “O” é formada pela simulação “G” apenas com os postos mais representativos da análise de correlação cruzada, ou seja, foram desconsiderados os postos que estavam na mesma bacia e tinham “lag” semelhante. Assim, trabalhou-se com os postos fluviométricos de Porto das Andorinhas, Porto Indaiá, Porto Mesquita, Porto Pará e com a chuva média da bacia. Analisando o erro mínimo, observa-se que as previsões pioraram para os primeiros 7 dias e melhoraram para os 12 dias de previsão.

Considerando como critério de seleção o erro mínimo do conjunto de validação, a melhor simulação para os 7 primeiros dias foi a “G”, que considera apenas informações fluviométricas, pluviométricas e de vazão natural do dia em que está fazendo a previsão, e a “L” para os últimos 5 dias de previsão, que, em relação à “G,” adiciona-se apenas mais 2 dias de precipitação.

Considerando os dados de validação e juntando as duas melhores redes NSRBN do período úmido, com um EPMA de 20,2% para os 7 primeiros dias e 36,5% para os 5 últimos dias de previsão, obteve-se um EPMA de **27,0%** para os 12 dias de previsão

As tabelas 5.17 e 5.18 mostram os resultados da rede NSRBN encontrados para os 7 e 12 dias de previsão do período seco, respectivamente.

Tabela 5.17 – Resultados das simulações do período seco para os 7 dias de previsão.

Simul	Entradas							Obtidos			7 primeiros dias	
	%val	%teste	alfa	semente	beta	epoca min	epoca max	erro min	erro perc md abs	erro padrão	teste	validação
A	22	28	20	987654	50	300	500	1,66	12,84	31,43	19,7	18,5
	22	28	30	987654	50	200	400	1,67	12,64	31,45	19,3	18,4
B	22	28	20	987654	60	200	400	1,65	12,74	33,11		
	22	28	30	987654	50	200	400	1,54	12,64	33,37	19,2	17,8
C	22	28	20	987654	50	200	400	1,78	12,90	34,09	19,1	18,5
	22	28	30	987654	50	200	400	1,75	12,83	34,21		
D	22	28	30	987654	50	200	400	1,62	12,43	31,49	15,5	13,2
	22	28	20	987654	50	200	400	1,76	12,85	31,72		
E	22	28	20	987654	50	200	400	1,62	12,96	32,25	19,9	18,1
	22	28	30	987654	50	200	400	1,61	13,10	32,39		
F	22	28	20	987654	50	200	600	1,62	12,13	32,25	18,6	18,0
	22	28	30	987654	50	100	500	1,63	12,29	31,48		
G	22	28	22	987654	50	200	400	1,30	13,19	33,6	16,5	13,0
	22	28	30	987654	50	200	400	1,40	12,16	32,43	18,7	17,4

Tabela 5.18 – Resultados das simulações do período seco para os 12 dias de previsão.

Simul	Entradas							Obtidos			12 dias		5 últimos dias		7 primeiros dias	
	%val	%teste	alfa	semente	beta	epoca min	epoca max	erro min	erro perc md abs	erro padrão	teste	validação	teste	validação	teste	validação
A	22	28	20	987654	50	100	500	4,68	18,15	49,23	25,0	23,9	28,0	24,7	22,8	23,2
	22	28	30	987654	50	200	400	4,65	18,15	40,35	24,9	23,8	28,0	24,6	22,8	23,2
B	22	28	20	987654	50	200	400	4,31	17,41	39,53	24,4	23,3	27,8	25,0	21,9	22,2
	22	28	30	987654	50	200	400	4,29	17,42	39,68						
C	22	28	30	987654	50	200	400	4,79	18,69	41,95	25,5	24,2	29,1	25,1	22,9	23,5
	22	28	30	987654	50	200	400	4,52	17,82	40,21						
D	22	28	20	987654	50	200	400	4,50	17,79	40,09	24,7	23,5	28,0	24,8	22,4	22,5
	22	28	30	987654	50	200	400									
E	22	28	20	987654	50	200	400	4,73	18,75	42,01						
	22	28	30	987654	50	200	400	4,23	17,45	39,27	24,6	23,6	28,2	25,4	22,1	22,2
F	22	28	20	987654	50	200	600	4,20	17,45	39,27	24,7	23,5	28,3	25,4	22,1	22,1
	22	28	30	987654	50	200	600									
G	22	28	30	987654	50	200	400	3,73	16,89	38,76	24,3	22,5	28,5	25,4	21,4	20,5
	22	28	20	987654	50	200	600	3,75	16,82	38,65	24,2	22,6	28,3	25,5	21,3	20,5

Nas tabelas 5.17 e 5.18 observa-se que a alteração das variáveis básicas de entrada do programa NeuroHidro (alfa, semente, beta, épocas mínima e época máxima) os resultados não variam muito, o que facilita a simulação com esse modelo, visto que não são necessárias diversas simulações para ajuste desses parâmetros, como nas simulações com o software Matlab, para a rede MLP.

Mais uma vez, analisando as tabelas 5.17 e 5.18 observa-se que a calibração de uma rede para os 7 primeiros dias e depois outra para os 12 dias, os resultados de previsão dos 7 primeiros dias desse último são inferiores aos resultados da calibração direta, visto que a função objetivo da versão utilizada no NeuroHidro tem esse objetivo. Como exemplo observam-se os resultados da simulação L, em que, para os dados de validação da rede de 7 dias, tem um erro de 18,0 e 20,5, pegando os 7 primeiros dias da rede ajustada para os 12 dias.

A simulação “D” é formada pela simulação “C”, colocando-se dois dias de vazão natural. Analisando o erro mínimo, observa-se que as previsões melhoraram para os primeiros 7 dias e para os 12 dias de previsão.

A simulação “F” é formada pela simulação “B” considerando os postos mais representativos segundo a ACP, apresentados anteriormente na figura 5.18. Analisando o erro mínimo, observa-se que as previsões pioraram para os primeiros 7 dias e melhoraram para os 12 dias de previsão.

A simulação “G” é formada pela simulação “B” considerando os postos mais representativos da análise de correlação cruzada, ou seja, foram desconsiderados os postos que estavam na mesma bacia e tinham “lag” semelhante. Analisando o erro mínimo, observa-se que as previsões melhoraram muito para os primeiros 7 dias e para os 12 dias de previsão.

Considerando como critério de seleção o erro mínimo do conjunto de validação, a melhor simulação para os 7 primeiros dias e para os 5 últimos dias foi a “G”, que considera informações fluviométricas com 3 dias de antecedência e do dia que se está fazendo a previsão dos postos fluviométricos de Porto das Andorinhas, Porto Indaiá, Porto Mesquita e Porto Pará, da chuva média e de vazão natural com 3, 2 e 1 dia de antecedência.

Considerando os dados de validação e juntando as duas melhores redes NSRBN do período seco, com um EPMA de 13,0% para os 7 primeiros dias e 25,4% para os 5 últimos dias de previsão, obteve-se um EPMA de **18,2%** para os 12 dias de previsão.

Selecionando o melhor modelo para os períodos úmido e seco e colocando na mesma base de dados do Programa Mensal de Operação - PMO, as tabelas 5.19 e 5.20 mostram o desempenho das previsões diárias e semanais, respectivamente; e a figura 5.57, a vazão semanal prevista e ocorrida. O Anexo 02 mostra essas vazões diárias e semanais previstas e ocorridas, através de gráficos anuais.

Tabela 5.19 – Previsão de vazões diárias com os melhores modelos da Rede NSRBN.

Ano	EPMA (%)
2000	12,9
2001	17,1
2002	20,8
2003	12,7
2004	21,8
Média:	17,1

Tabela 5.20 – Previsão de vazões semanais com os melhores modelos da Rede NSRBN.

Ano	EPMA (%)		
	úmido	seco	Média
2000	24,4	18,9	21,6
2001	46,4	18,6	32,5
2002	43,6	28,9	36,3
2003	32,7	15,8	24,2
2004	34,6	12,3	23,5
Média:	36,3	18,9	27,6

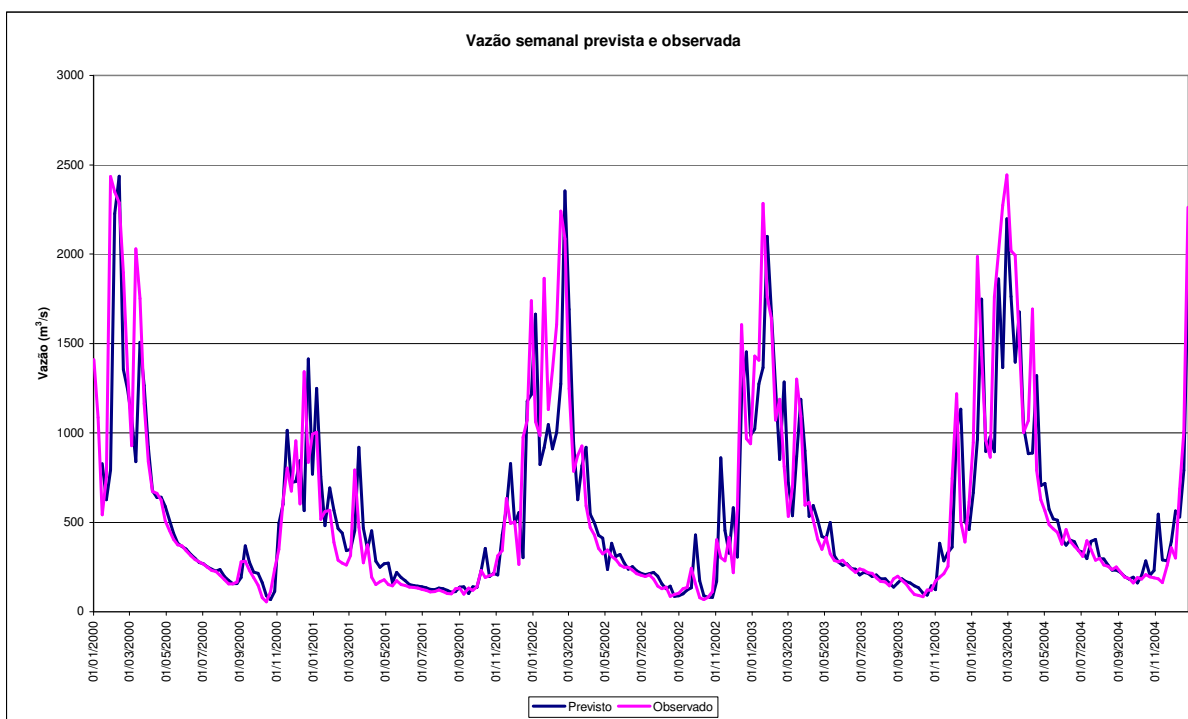


Figura 5.57 – Vazão semanal prevista e observada com a rede NSRBN.

Comparando as simulações com as redes MLP e NSRBN algumas observações são importantes:

- Neste trabalho, as simulações com a rede MLP, no software Matlab, foram bem mais trabalhosas que as efetuadas com a rede NSRBN, no software NeuroHidro. Um dos motivos é que, nas simulações com o Matlab, foi necessário fazer simulações para vários números de neurônios na camada intermediária;
- Considerando os dados de validação das redes MLP, obteve-se um EPMA de **28,2%** do período úmido e de **15,9%** do período seco, para os 12 dias de previsão. Também considerando os dados de validação das redes NSRBN, obteve-se um EPMA de **27,0%** do período úmido e de **18,2%** do período seco, para os 12 dias de previsão;
- Colocando as informações na base de dias do PMO, obteve-se um EPMA de 16,9% para os dados diários e 29,0% para os dados semanais nas simulações com a rede MLP e de 17,1% para os dados diários e 27,6% para os dados semanais nas simulações com a rede NSRBN.

Capítulo 6

Conclusões e Recomendações

Muitos trabalhos têm mostrado, com êxito, a aplicação da Geoestatística para aferir a variabilidade temporal e espacial da precipitação. Essas análises são feitas com diferentes aplicativos e períodos de dados.

Esforços se fazem necessários na obtenção de uma melhor previsão de vazões aos reservatórios do Sistema Interligado Nacional – SIN, visto que esses resultados interferem na otimização do despacho das gerações das usinas hidrelétricas.

Para isso, foi proposta neste trabalho a calibração de um modelo para previsão de vazões diárias, 12 dias à frente, baseado na técnica de redes neurais. Nesta técnica, muitas podem ser as variáveis de entrada do modelo. Neste trabalho foram escolhidas informações de postos fluviométricos, pluviométricos e da própria vazão natural do reservatório de Três Marias, dos dias anteriores à previsão.

Para ter essas informações de entrada com boa representatividade e qualidade, o que pode interferir bastante nos resultados das simulações, foi feito um preenchimento de falhas do histórico, uma análise geoestatística com as informações dos postos pluviométricos, e diversas análises estatísticas com todas as informações selecionadas.

A seguir são listadas as principais conclusões encontradas nesta fase da análise dos dados:

- A Função de Auto-Correlação mostrou que, para as defasagens temporais (lags) analisadas, não há efeito de sazonalidade nas séries de vazões. Também para as defasagens temporais analisadas, a informação da precipitação média espacial da bacia também não apresentou efeito de sazonalidade, mas se mostrou bem menos auto-correlacionada que as séries de vazões.
- A inclusão de postos pluviométricos nas áreas com grande erro de interpolação, apontadas pelo estudo geoestatístico, mostrou um deslocamento da área de precipitação máxima da bacia.

- A chuva média diária calculada com a metodologia de Kriging mostrou-se bem semelhante à calculada com o polígono de Thiessen, o que mostra que para uma rede de pluviômetros bastante densa, o método de Thiessen para calcular a precipitação média numa bacia hidrográfica, gera valores excelentes.

- Analisando os coeficientes de correlação e a ACP dos postos fluviométricos com a vazão natural de Três Marias, observou-se que o preenchimento de falhas das informações fluviométricas pelo método da regressão linear múltipla em substituição à regressão linear simples/potencial levou a informações mais coerentes com os demais dados da Bacia.

- A correlação cruzada mostrou que os postos de Ponte Nova do Paraopeba, Ponte dos Vilelas Jusantes e Iguatama têm uma defasagem temporal, em relação à vazão natural de Três Marias, próximas às defasagens dos postos de Porto das Andorinhas, Porto Indaiá, Porto Mesquita e Porto Pará, respectivamente.

Após a análise dos dados básicos, foram calibrados dois tipos de redes neurais: uma MLP e outra NSRBN. Tomando o período de 1987 a 1995 para treinamento da rede, 1996 a 1999 para validação e 2000 a 2004 para testes, obtiveram-se os seguintes resultados, em EPMA médio dos 12 dias de previsão:

- Dados de validação: rede MLP: período úmido = 28,2% e período seco = 15,9%; rede NSRBN: período úmido = 27,0% e período seco = 18,2%.

- Dados de teste: rede MLP: período úmido = 37,2% e período seco = 21,9%; rede NSRBN: período úmido = 33,5% e período seco = 22,8%;

Colocando na mesma base dos dias do PMO, obteve-se um EPMA da previsão diária média dos anos de 2001 a 2003 de 16,9% para a rede MLP, de 16,9% para a rede NSRBN, enquanto a atual previsão do PrevizazH é de 41,4% e da CEMIG 18,8%. Já para as vazões semanais, considerando a média dos anos de 2000 a 2004, obteve-se um erro de previsão de 29,0% para a rede MLP, de 27,6% para a rede NSRBN, enquanto a atual previsão do Previzaz é de 35,2%. Assim, pode-se concluir que a metodologia de redes neurais possibilita uma a

obtenção de uma previsão de vazões bem superiores às obtidas com os modelos estatísticos, tanto a nível diário quanto semanal. Cumpre salientar que esses resultados servem apenas como referência geral para a verificação do desempenho dos modelos analisados, visto que as simulações não seguiram rigorosamente as mesmas premissas.

Observou-se que os resultados das redes MLP e NSRBN foram bem próximos, com a rede NSRBN com um desempenho melhor para as previsões semanais e um pouco pior para as previsões diárias. Isto também se deve ao fato da rede NSRBN implementada no NeuroHidro ter função objetivo para os 7 dias de previsão, não visando o desempenho específico dos primeiros dias de previsão.

Contudo, neste trabalho, a fase de calibração da rede MLP, no software Matlab, foi bem mais trabalhosa que a da rede NSRBN, no software NeuroHidro. No primeiro caso, foi necessário definir o número de neurônios na camada intermediária, e para cada neurônio, fazer 5 simulações, visto que os resultados variavam um pouco de uma simulação para a outra, devido a forma de inicialização dos pesos e bias.

Depois de calibrada a rede neural, a implementação para aplicação da rede MLP pode ser facilmente feita em Excel, e a da rede NSRBN, através do NeuroHidro.

Vale destacar que a melhor simulação para a rede NSRBN do período seco foi formada pelos postos selecionados através da análise de correlação cruzada, reforçando a importância da análise dos dados na definição dos variáveis de entrada da rede neural.

Tanto nas simulações com a rede MLP quanto com a rede NSRBN, ficou clara que a adição de dias passados da vazão natural do reservatório de Três Marias pioram o desempenho das previsões para os 12 dias considerados. Também se observou a importância das informações de precipitação, tanto nas simulações do período úmido, quanto nas do período seco.

Na revisão bibliográfica, diversos autores compararam a aplicações dos tradicionais modelos estatísticos com os modelos que usam a técnica de redes

neurais, com esses últimos sempre com resultados bem superiores. Esse trabalho também comprova esta afirmativa.

Existe a necessidade de continuar as pesquisas para desenvolver modelos de previsão de vazões para o reservatório de Três Marias. Como sugestão para futuras pesquisas pode-se testar:

- Outras metodologias de redes neurais (ex: Fuzzy Recorrente);
- Modificações nas funções de transferência das camadas intermediária e de saída da rede MLP usada neste trabalho;
- Modificar o período da função objetivo da rede NSRBN.

Capítulo 7

Referências Bibliográficas

ALMEIDA, T. A. de; MONTENEGRO, S. M. G. L.; MONTENEGRO, A. A. de A.; OLIVEIRA, O. F. de. 2004. Análise da variabilidade espacial da precipitação da bacia do rio Ipanema, Pernambuco, Brasil. In: Anais do VII Simpósio de Recursos Hídricos do Nordeste, São Luís – MA.

ANA, 2004. Plano de Recursos Hídricos da Bacia do São Francisco. Resumo Executivo.

AMARAL, I. R.; GRIPP, A. H.; VERSIANI, B. R. 2004. Análise espacial de eventos máximos de precipitação pelo método geoestatístico: estudo das bacias do Alto São Francisco e Alto Rio Doce. In: Anais do VII Simpósio de Recursos Hídricos do Nordeste, São Luís– MA.

AUTODESK. Autocad LT-Brasil 2000. Copyright Patent No. 110-16003187.1999.

BALLINI, R.; SOARES, S.; ANDRADE, MARINHO G. 2003. Previsão de vazões médias mensais usando redes neurais nebulosas. In: Revista Controle & Automação, Vol.14 nº3.

BERTONI, J.C. e TUCCI, C. E. M. Precipitação. In: TUCCI, C. E. M. Hidrologia: Ciência e Aplicação. 2. ed. Porto Alegre: UFRGS/ABRH, 2001. p.177-252.

CASTRO, L.N., IYODA, E.M., PINHEIRO, E., VON ZUBEN F. 1999. Redes Neurais Construtivas: Uma Abordagem Comparativa. In: Anais do Proceedings of the IV Brazilian Conference on Neural Networks – IV Congresso Brasileiro de Redes Neurais pp. 102-107, São José dos Campos - SP – Brazil.

CAMBARDELLA, C. A., MOORMAN T. B., NOVAK J. M., PARKIN T. B., KARLEN, D. L., TURCO, R. F., KORNOPKA, A. E. 1994. Fieldscale variability of soils properties in Central Iowa Soils. In: Soil Science Society of America Journal, v. 58, p. 1501-1511.

CEPEL, 1997. Modelo de Previsão de Vazões Semanais Aplicado ao Sistema Hidroelétrico Brasileiro – PREVIVAZ, Relatório Técnico DPP/PEL 125/97, Rio de Janeiro, Brasil.

CEPEL, 2001. Modelo GEVAZP - Manual de Referência, 24 páginas.

CHOW, V.T. 1964. Handbook of applied hydrology. MacGraw-Hill Book Co. New York.

CPRM. 1996. Programa de Avaliação de Recursos Hídricos. Rede Hidrometeorológica e Caracterização Física da Bacia do Alto São Francisco - Sub-Bacia 40 Fase 1. Belo Horizonte.

CARVALHO, A. R. L. de. 2001. Previsão de Afluências Diárias para Horizontes de 1 a 7 dias Através de uma Abordagem Estocástica – Modelo PREVIVAZH. Universidade Federal do Rio de Janeiro – UFRJ, Departamento de Engenharia Civil. Dissertação de Mestrado.

COULIBALY, P.; ANCTIL, F. BOBÉE. B. 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. In: Journal of Hydrology, Volume 230, Pages 244-25.

DELHOME, J.P. 1978. Kriging in the Hydrosciences. Advances in Water Resources, v. 1. n. 5, p 251-266.

DUTTER, R.. Analysis of Spatial Data Using GEOSAN: Program System for Geostatistical Analysis. Handbook. Vienna University of Technology, Austria, 1996.

ESRI Inc. ArcMap 8.1. Copyright – US Patent No. 5710835. 2001.

FERREIRA, W. M. F.; NOGUEIRA, L. A. de A.; BEMFEITO, C. J. de S. 2000. Aplicação da Metodologia Kriging à Pluviometria do Estado do Ceará. In: Anais do V Simpósio de Recursos Hídricos do Nordeste, Natal – RN.

GOLDEN SOFTWARE Inc. Surfer 8.02. Copyright Patent No. WS-063644-1222.2002.

GOMES, R. S.; SILANS, A. M. B. P. de. 2000. Comparação de diversos métodos de cálculo da precipitação média sobre uma bacia hidrográfica. In: Anais do V Simpósio de Recursos Hídricos do Nordeste, Natal – RN.

GUILHON, L. G. F.; CARPIO, L. G. T. 2003. Modelo Heurístico para Previsão de Vazões Naturais Médias Semanais em Usinas Hidrelétricas Usando Chuva Equivalente. In: Anais do XV Simpósio Brasileiro de Recursos hídricos. Curitiba – PR.

HAYKIN, S. 1998. Neural Networks – A Comprehensive Foundation, Macmillan College Publishing Company, Inc.

HIEZ, G.; COCHONNEAU, G.; SÉCHET, P.; FERNANDES, U.M.. 1991. Aplicação do método do Vetor Regional à análise da pluviometria anual da bacia amazônica, 367-377. In: Anais do IX Simpósio Brasileiro de Recursos Hídricos, ABRH, Rio de Janeiro.

HOLAWEA, F; DUTTERB, R. 1999. Geostatistical study of precipitation series in Austria: time and space. In: Journal of Hydrology 219.

IGREJA, G. C.; COUTINHO, D. L.; REISEN, V. A.; TEIXEIRA, E. C. 2005. Aplicação do método geoestatístico no dimensionamento de redes pluviométricas considerando o custo de operação da estação de monitoramento. In: Anais do XVI Simpósio Brasileiro de Recursos hídricos, João Pessoa – PB.

KELMAN, J., DAMAZIO, J.M., COSTA, J.P., 1983, Geração de Séries Sintéticas Diárias - Modelo DIANA, Revista Brasileira de Engenharia, v.1, n.2, pp. 5-22.

KRIGE, D.G. 1951. A Statistical approach to some basic mine evaluation problems on the witwatersrand. J. Chem. Metall. Min. Soc. S. Afri., Johannesburg, v. 52, p 151-163.

KYRILLOS, D. De S. 2000. Sistema de suporte ao planejamento dos recursos hídricos. Estudo de caso: bacia do rio São Francisco. Dissertação de mestrado em Engenharia Civil. Departamento de Engenharia Civil – UFPE.

LIMA, H. V. C.; LIMA, L. C. T. M. 2000. Interpolação Espacial – ferramenta auxiliar na gestão dos recursos hídricos – análise comparativa de métodos. In: Anais do V Simpósio de Recursos Hídricos do Nordeste, Natal – RN.

LIMA, C. H. R.; FERREIRA, W. M. F. 2003. Análise de modelos de redes neurais aplicados ao processo chuva-deflúvio no Semi-Árido. In: Anais do XV Simpósio Brasileiro de Recursos hídricos, Curitiba – PR.

LIMA, H. V. C.; LIMA, L. C. T. M.; LIMA, C. H. R. 2005. Modelagens hidrológica conceitual e por redes neurais artificiais aplicadas ao processo chuva-vazão: uma análise comparativa. In: Anais do XVI Simpósio Brasileiro de Recursos hídricos, João Pessoa – PB.

LOU, A.P.F., ROTUNNO FILHO, O.C., ARAUJO, A.A.M., BARBOSA, M.C.D., AMARAL, V.M. 2003. Modelagem geoestatística aplicada a integração entre dados de postos pluviométricos e radar meteorológico. In: Anais do XV Simpósio Brasileiro de Recursos Hídricos, Curitiba. Anais em Cd-Rom.

LOPES, J.E.G., BRAGA JR., B.P.F., CONEJO, J.G.L. 1981. Simulação hidrológica: Aplicações de um modelo simplificado. In: Anais do III Simpósio Brasileiro de Recursos Hídricos, v.2, 42-62, Fortaleza.

MACEIRA, M. E. P.; COSTA, F. S.; DAMÁZIO, J. M.; DENÍCIO, M.; GUILHON, L. G. 2003. Modelo Estocástico de Previsão de Vazões Mensais - PREVIVAZM. In: Anais do XV Simpósio Brasileiro de Recursos Hídricos, Curitiba - PR.

MATLAB, 2002 Neural Network Toolbox, User Guide, MathWorks Inc.

MATHWORKS Inc. Matlab 7.1.0.246 (R14). Copyright – US Patent No.308160. 2005.

MARINHO, M. H. da N.; FILHO, M. G. de A.; BALLINI, R.; SANTOS, E. F. dos; FILHO, S. S. 2004. Previsão de vazões afluentes utilizando modelos periódicos autoregressivos, redes neurais e redes neurais nebulosas. In: Anais do VII Simpósio de Recursos Hídricos do Nordeste.

MÉNDEZ, M. C.; MANTEIGA, W. G.; BANDE, M. F.; SÁNCHEZ, J. M. P.; CALDERÓN, R. L. 2004. Modelling of the monthly and daily behaviour of the runoff of the Xallas river using Box–Jenkins and neural networks methods. In: Journal of Hydrology, Volume 296, Issues 1-4, 20, Pages 38-58.

OLIVO, A. A. de; SILVA, J. D. S. da; VIJAYKUMAR, N. L. 2002. Previsão de cheias fluviais usando redes neurais artificiais. In: Anais do VI Simpósio de Recursos Hídricos do Nordeste, Maceió – AL.

ONS, 2003. Revisão das séries de vazões naturais nas principais bacias do Sistema Interligado Nacional – SIN.

ONS, 2004. ACIOLI, G. C. L.; GOMES, L. F. C.; MAGALHÃES, J. K. M.. Previsão de vazões diárias ao reservatório de Sobradinho. In: Anais do VII Simpósio de Recursos Hídricos do Nordeste, São Luís – MA.

OSÓRIO, F., BITTENCOURT, J.R. 2000. Sistemas Inteligentes baseados em RNAs aplicados ao processamento de Imagens, I Workshop de Inteligência Artificial, UNISC.

PATTERN RECOGNITION SYSTEMS. Sirius 7.0 versão demonstrativa.

PINTO, E. J. DE A.; NAGHETTINI, M. 1999. Definição de regiões homogêneas e a regionalização de precipitações máximas diárias da bacia do

Alto São Francisco. In: Anais do XIII Simpósio Brasileiro de Recursos Hídricos, Belo Horizonte – MG.

PINTO, N.L. de S.; HOLTZ, A.C.T.; MARTINS, J.A.; GOMIDE, F.L.S. 2000. Hidrologia básica. São Paulo: Edgard Blücher. 78p.

PRECHELT, L. 1998. Automatic Early Stopping Using Cross Validation: Quantifying the Criteria, Neural Networks, vol. 11, n° 4, pp. 761-767.

RODRIGUES, A. A.; SILANS, A. M. B. P. de. 2002. Análise da variabilidade espacial da precipitação no litoral do Nordeste Brasileiro. In: Anais do VI Simpósio de Recursos Hídricos do Nordeste, Maceió – AL.

RODRIGUES, A. A.; SILANS, A. M. B. P. de. 2003. Metodologia para otimização de rede de monitoramento pluviométrico baseada na Geoestatística. In: Anais do XV Simpósio Brasileiro de Recursos Hídricos, Curitiba – PR.

REED, R. 1993. Pruning Algorithms – A Survey, IEEE Trans. on Neural Networks, vol. 4, n° 5, pp. 740-747.

RAJURKAR, M.P.; KOTHYARI, U.C.; CHAUBE, U.C. 2004. Modeling of the daily rainfall-runoff relationship with artificial neural network. In: Journal of Hydrology, Volume 285, Pages 96-113.

SAJIKUMAR, N.; THANDAVESWARA, B. S. 1999. A non-linear rainfall-runoff model using an artificial neural network. In: Journal of Hydrology, Volume 216, Pages 32-55.

SILVA, T. G. DE S.; SILVA, J. D. S. da. 2005. Projeto água viva: um estudo para aplicação de redes neurais artificiais na de consumo de água das cidades do litoral norte do estado de São Paulo. In: Anais do XVI Simpósio Brasileiro de Recursos Hídricos, João Pessoa – PB.

SHAMSELDIN, A. Y. 1997. Application of a neural network technique to rainfall-runoff modeling. In: Journal of Hydrology, Volume 199, Pages 272-294.

SPERB, R.M.; MODRO, N. R.; MARCONDES, P; FOES, P.A.; BARCIA, R. M.; PACHECO, R. C. S; TAVARES, G. 1999. Prevent: protótipo de um sistema de previsão de enchentes baseado em redes neurais. In: Anais do XIII Simpósio Brasileiro de Recursos Hídricos, Belo Horizonte – MG.

VALENÇA, M. J. S. 1997. Aplicação de redes neurais na área de recursos hídricos. In: Anais do XII Simpósio Brasileiro de Recursos Hídricos, Vitória – ES.

VALENÇA, M. J. S. 1999. Análise e projeto de redes neurais com abordagem construtiva para modelagem de sistemas complexos. Tese de Doutorado em ciência da computação. Departamento de Informática – UFPE.

VALENÇA, M.; LUDERMIR, T. 2001. Redes compostas por blocos de regressões sigmóides não-lineares: uma eficiente rede de alta ordem com aplicações na previsão de séries temporais. In: Anais do V Simpósio Brasileiro de Redes Neurais, Rio de Janeiro – RJ.

VALENÇA, M.; VALENÇA, A. 2005. Modelagem do processo chuva x vazão e vazão x vazão utilizando redes neurais. In: Anais do XVI Simpósio Brasileiro de Recursos Hídricos, João Pessoa – PB.

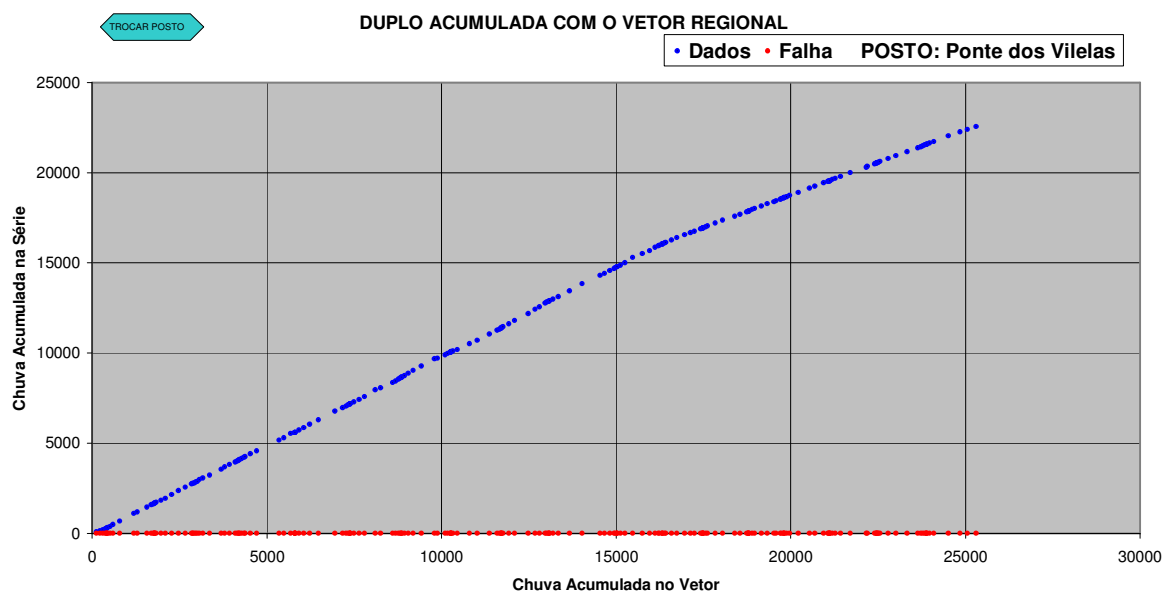
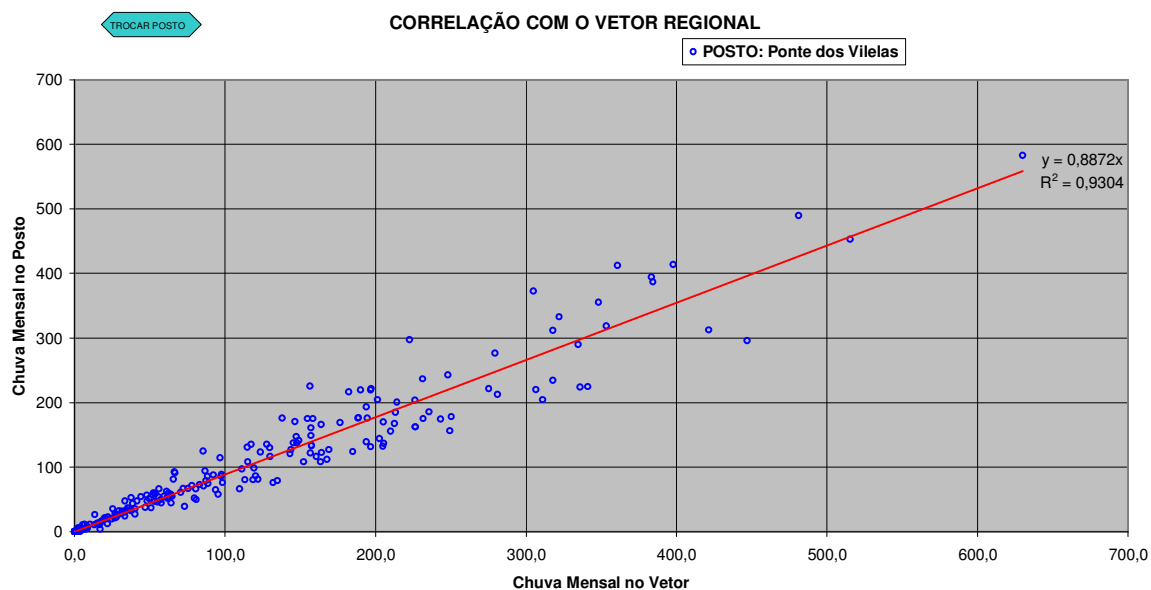
VALENÇA, M. J. S. 2005b. Aplicando redes neurais: um guia completo. Editora Livro rápido. Olinda, PE: Ed. Do Autor, 284p.

VIEIRA, F.R. 2000. Tópicos em ciência do solo. Campinas, São Paulo – SP. Vol.1. Págs 1 a 43.

8.0 Anexos

1. Resultados da aplicação da metodologia do Vetor Regional para a análise de consistência de informações de postos pluviométricos de postos localizados a montante da UHE de Três Marias;
2. Previsões de vazões diárias e semanais nos padrões do PMO.

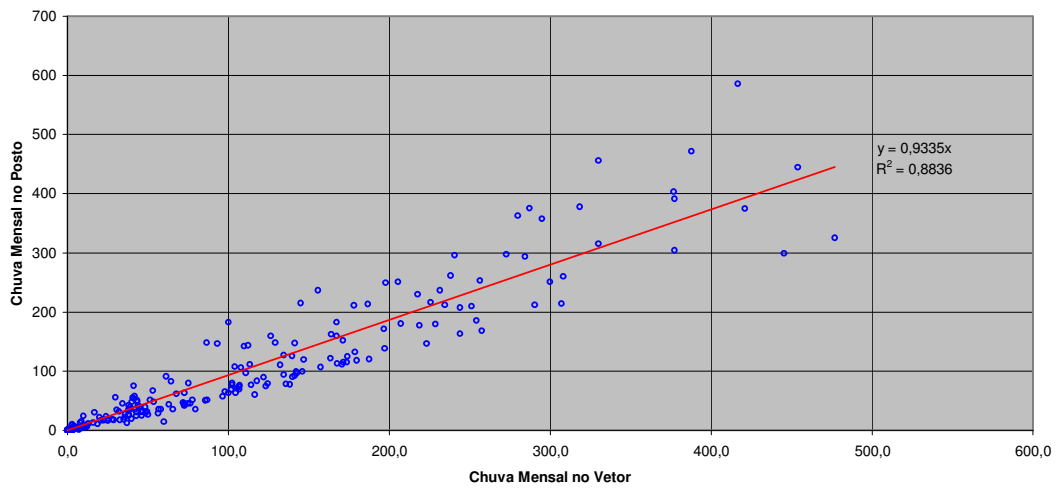
Anexo 01 – Resultados da aplicação da metodologia do Vetor Regional para a análise de consistência de informações de postos pluviométricos de postos localizados a montante da UHE de Três Marias



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

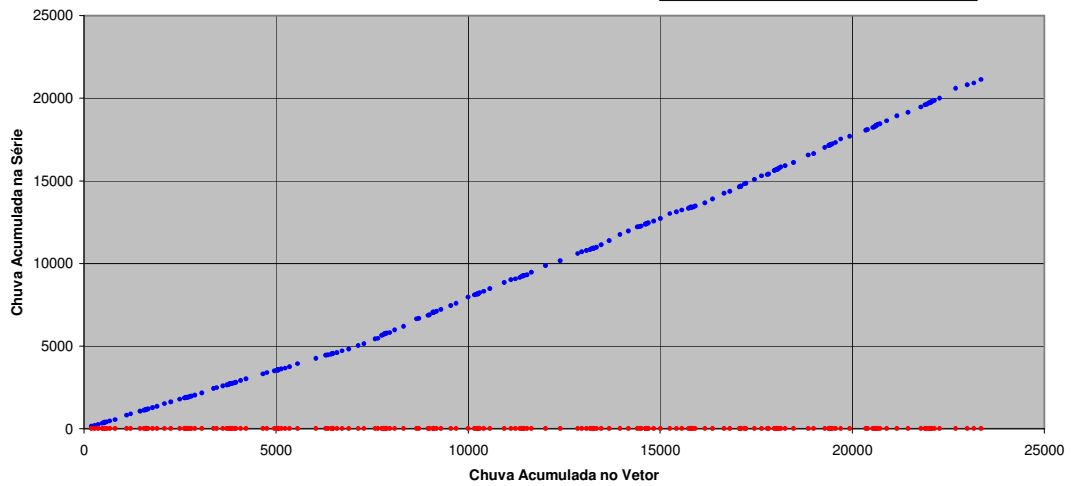
POSTO: Arcos



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

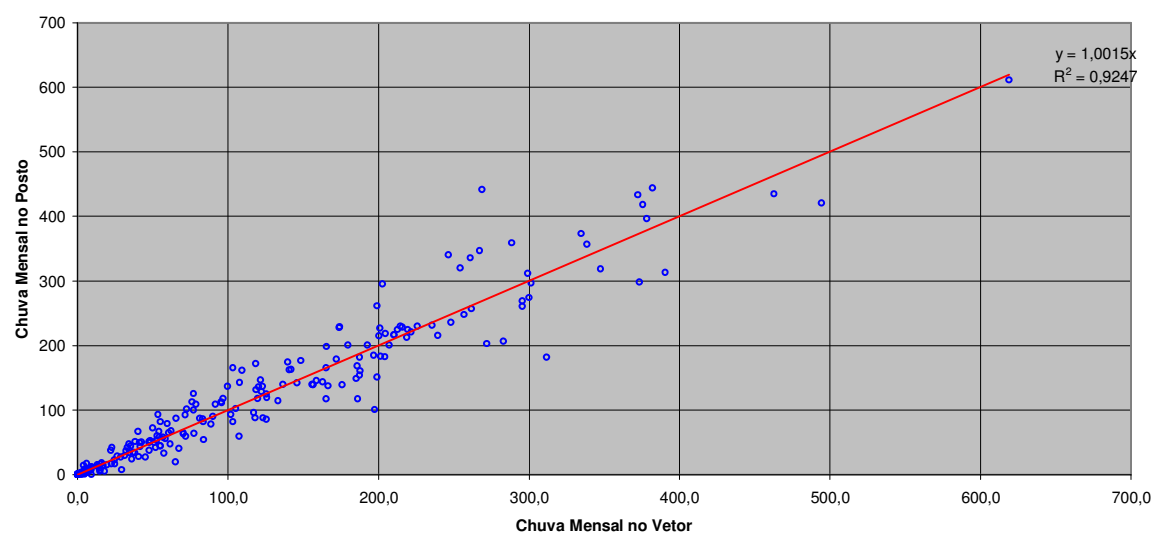
Dados Falha POSTO: Arcos



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

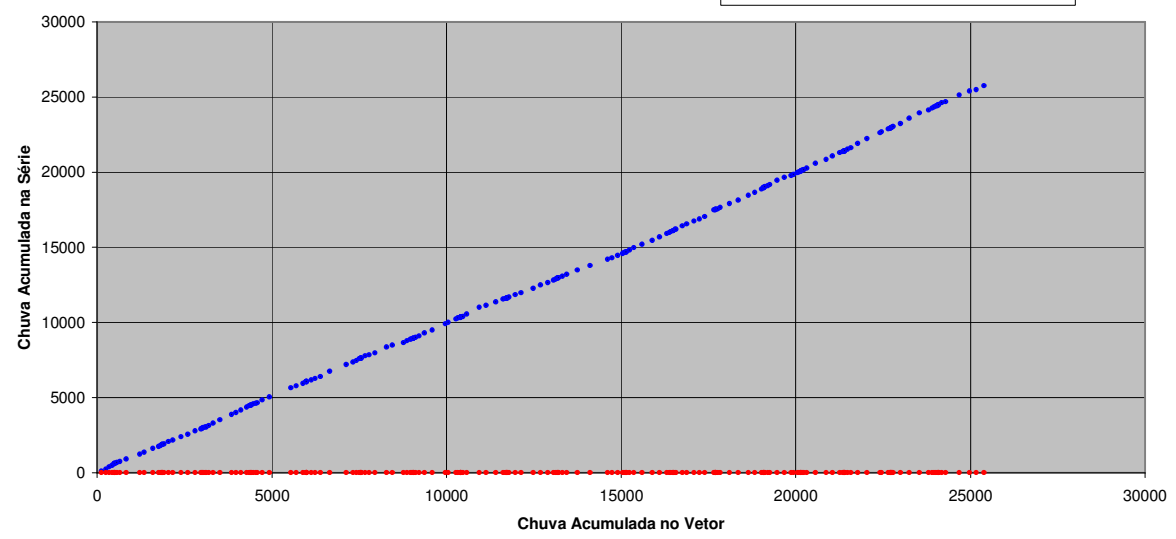
POSTO: Bonfin



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

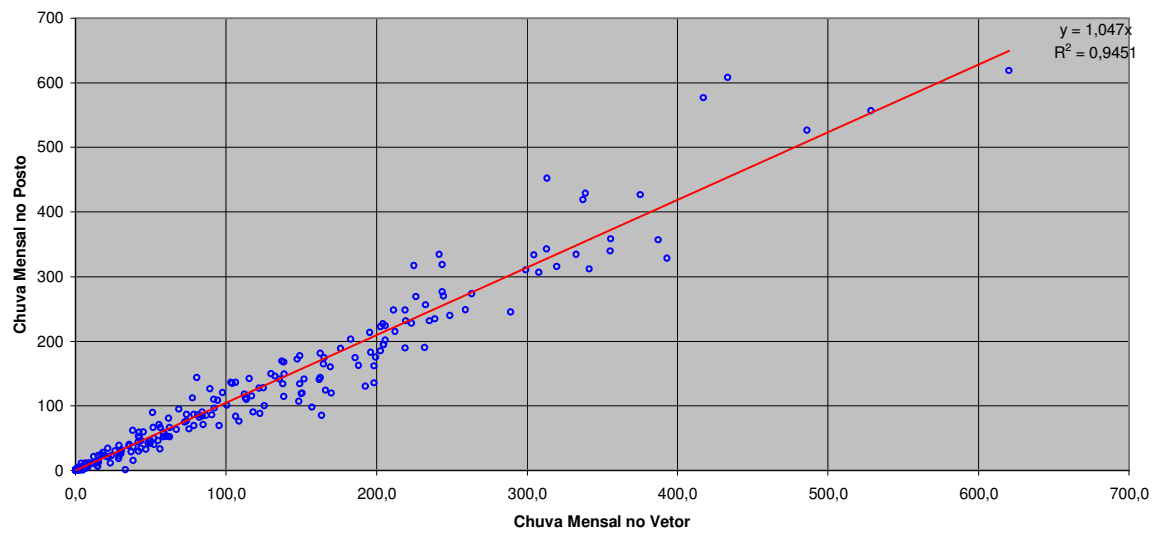
Dados Falha POSTO: Bonfin



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

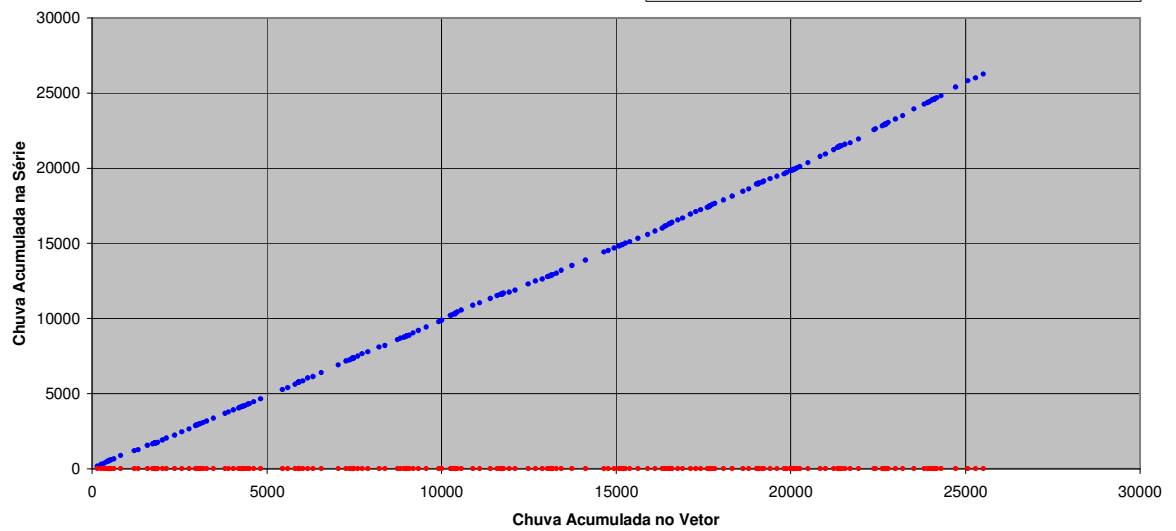
POSTO: Carmópolis de Minas



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

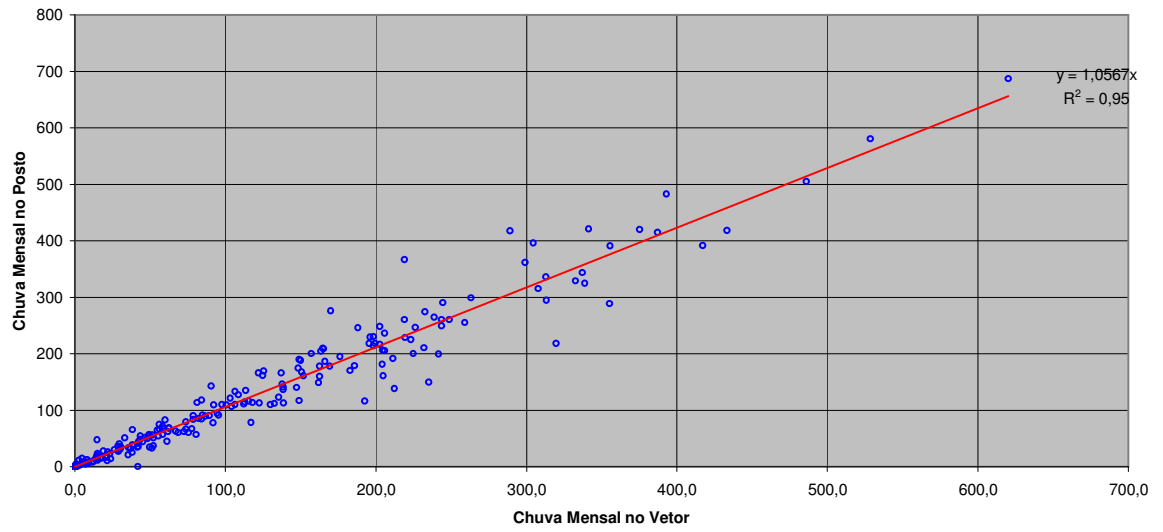
Dados Falha POSTO: Carmópolis de Minas



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

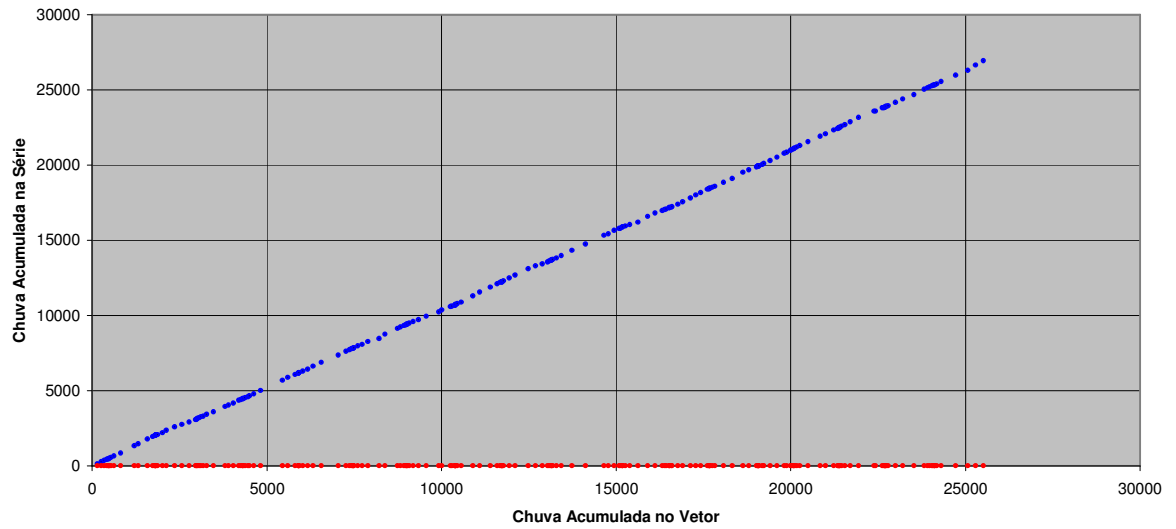
POSTO: Desterro de Entre Rios



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

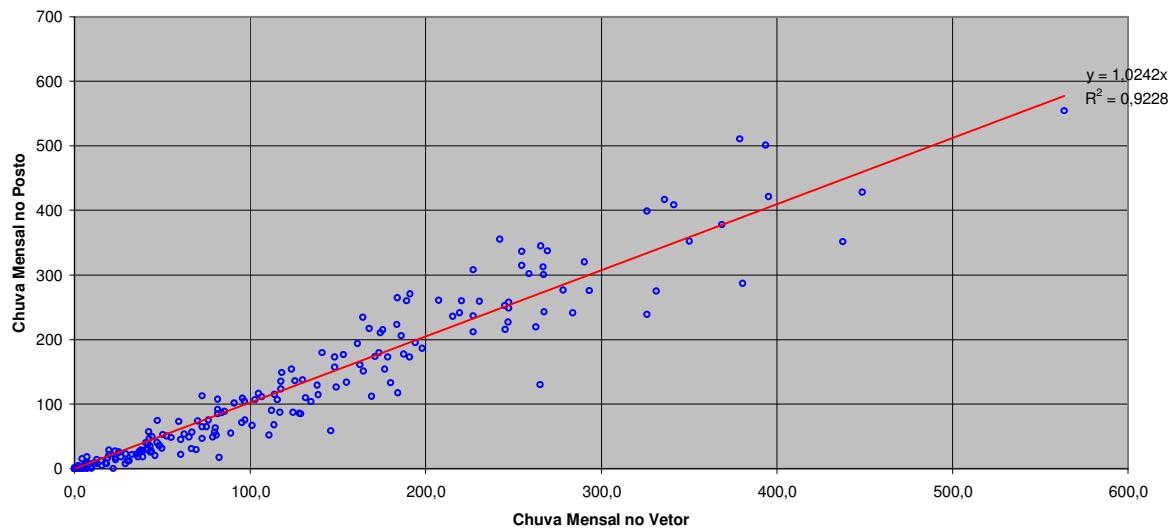
Dados Falha POSTO: Desterro de Entre Rios



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

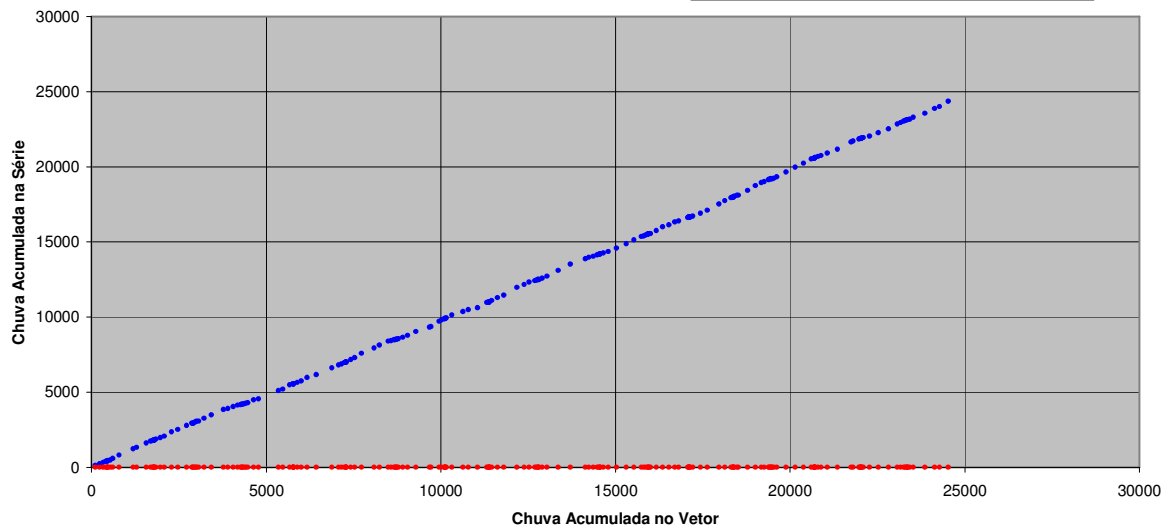
POSTO: Esmeraldas



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

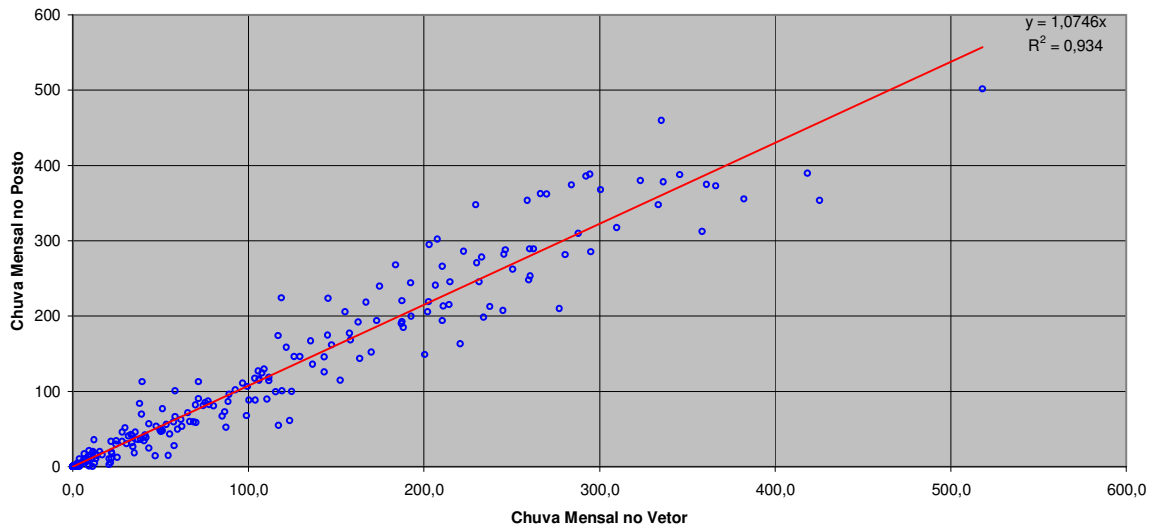
POSTO: Esmeraldas



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

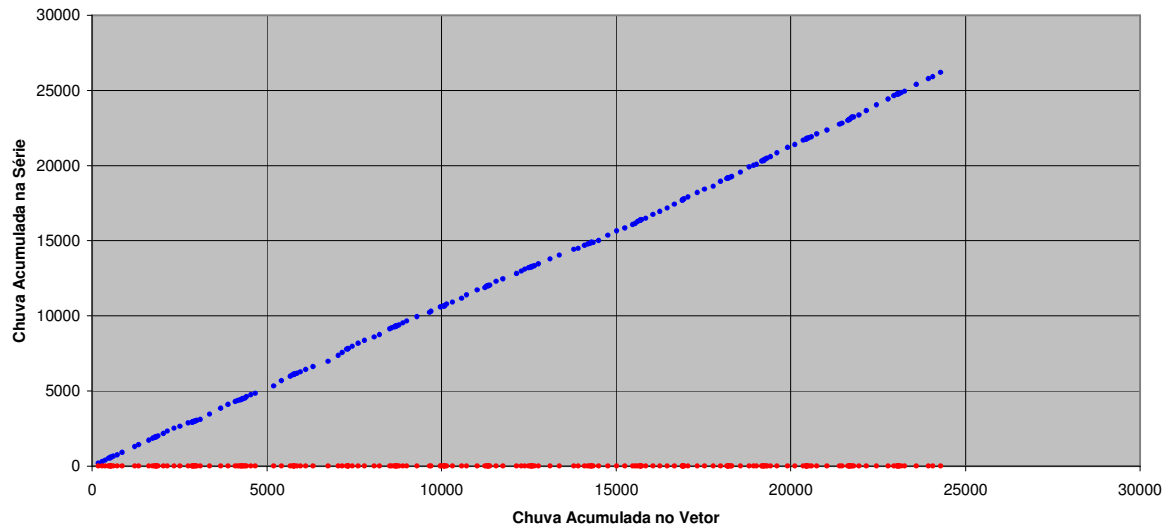
POSTO: Estrela do Indaiá



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

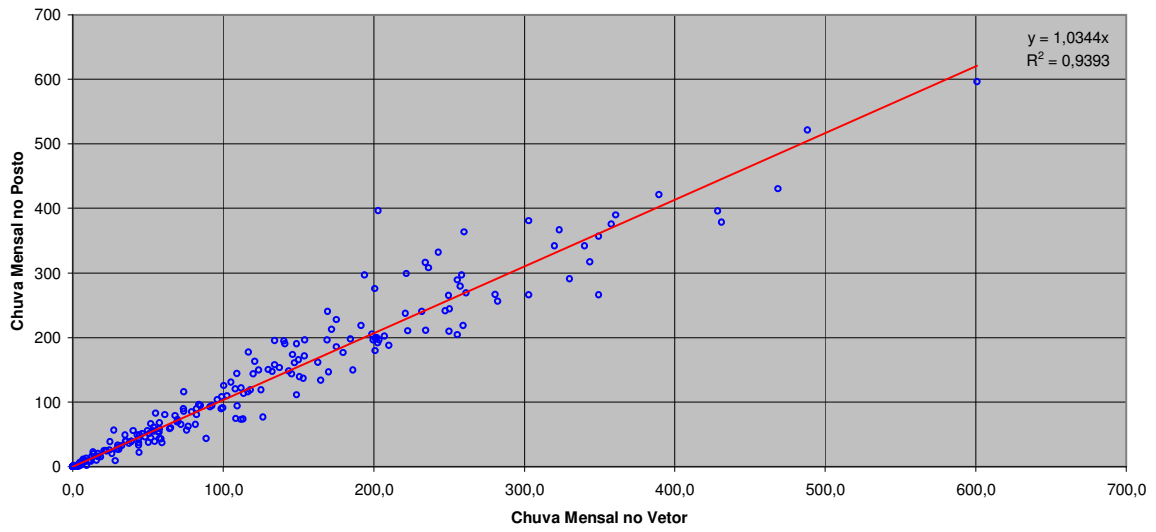
Dados Falha POSTO: Estrela do Indaiá



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

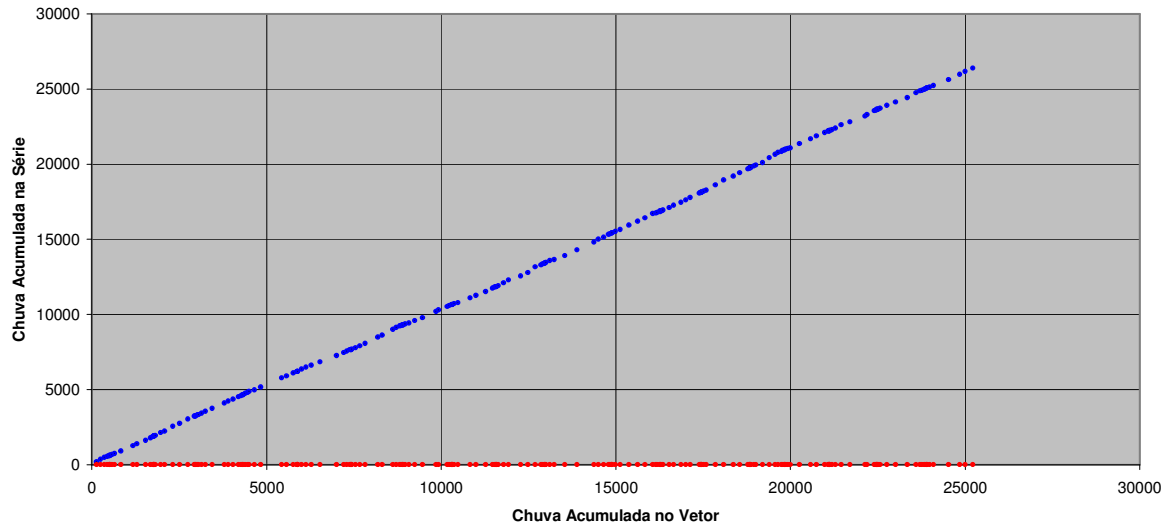
POSTO: Itaguara



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

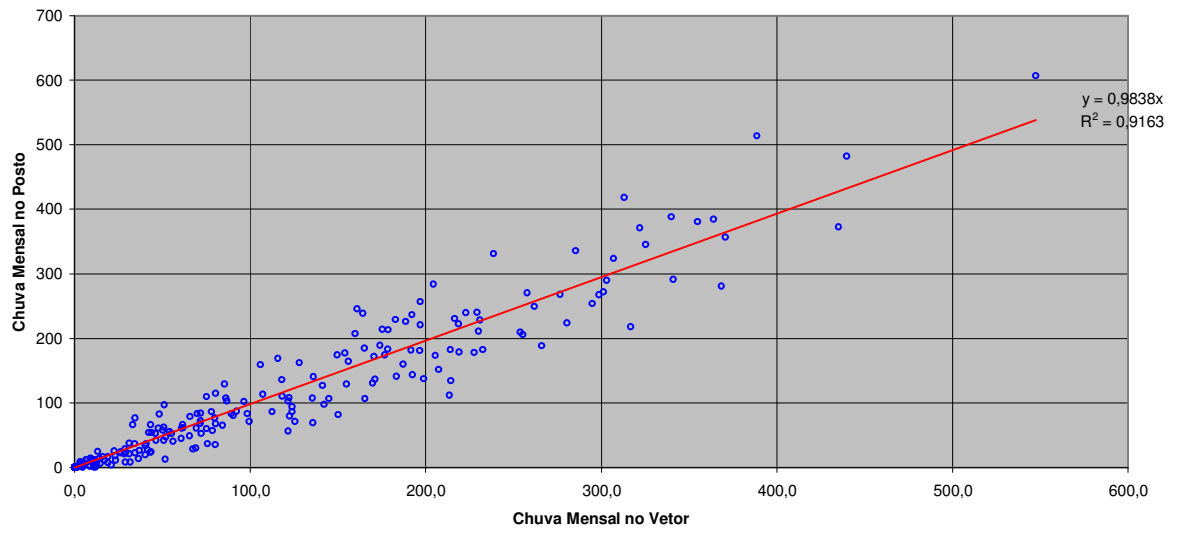
Dados Falha POSTO: Itaguara



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

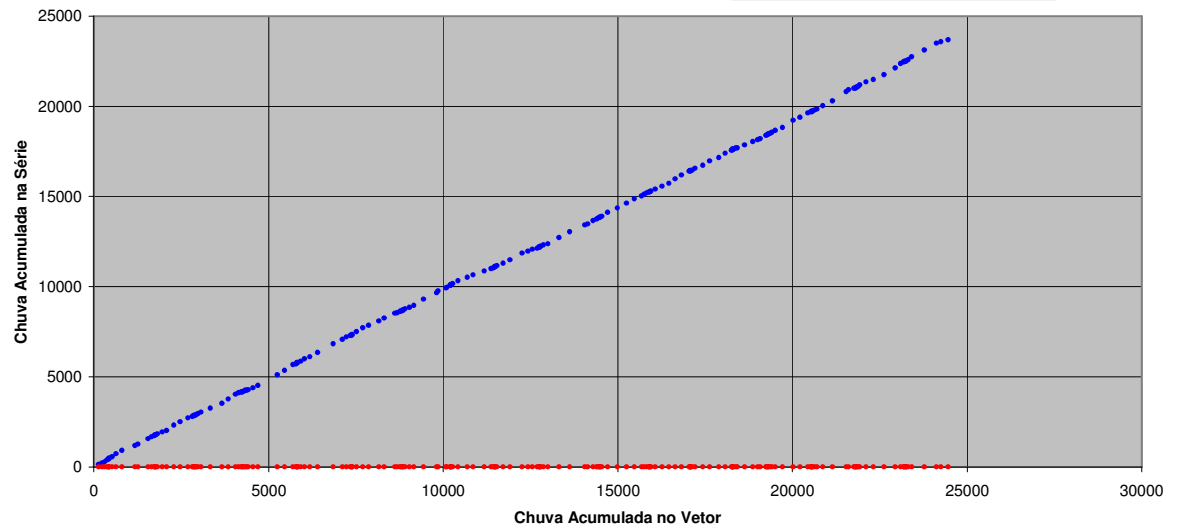
POSTO: Luz



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

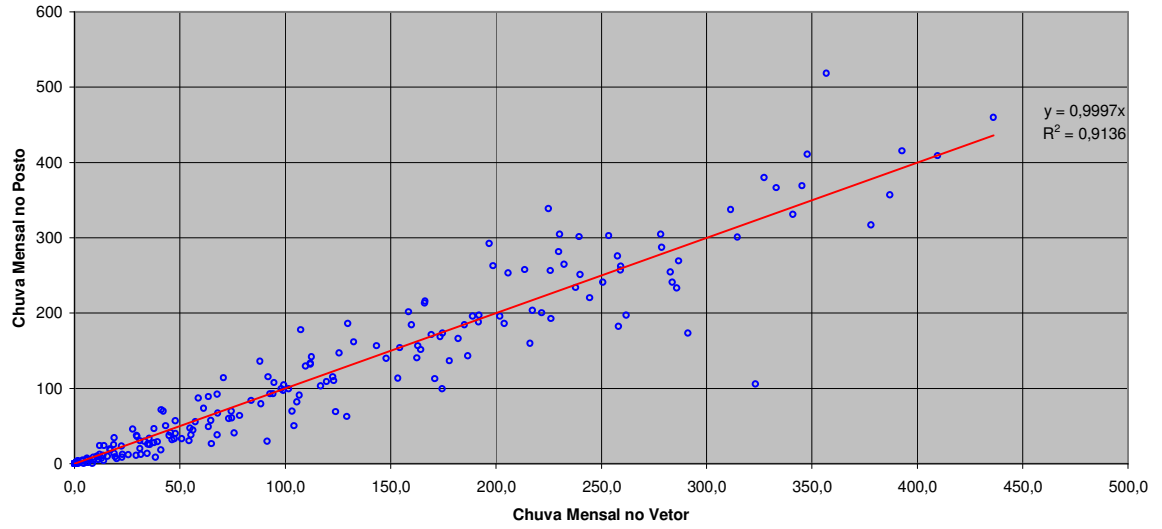
Dados Falha POSTO: Luz



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

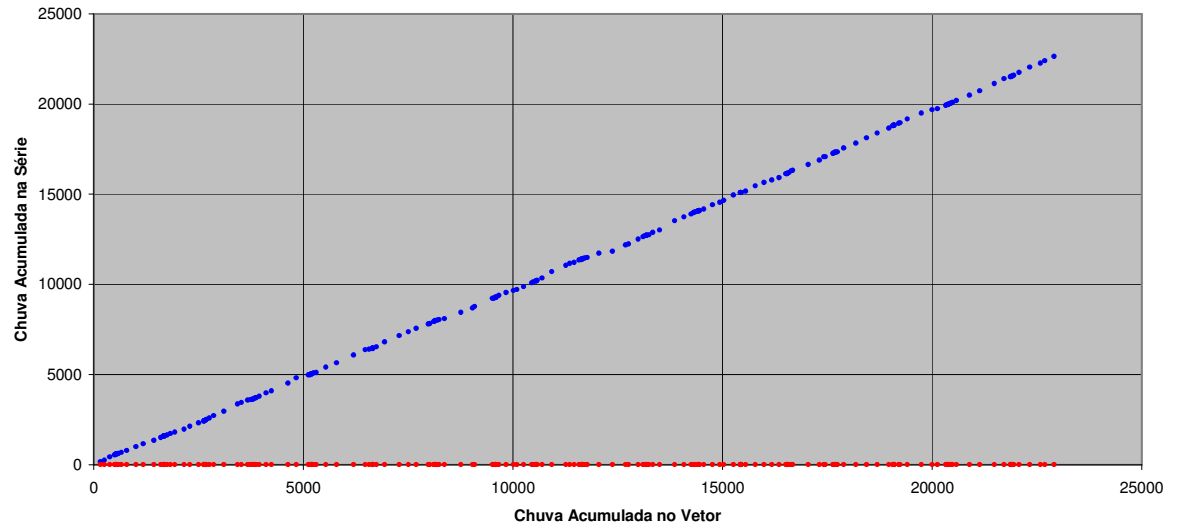
POSTO: Morada Nova



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

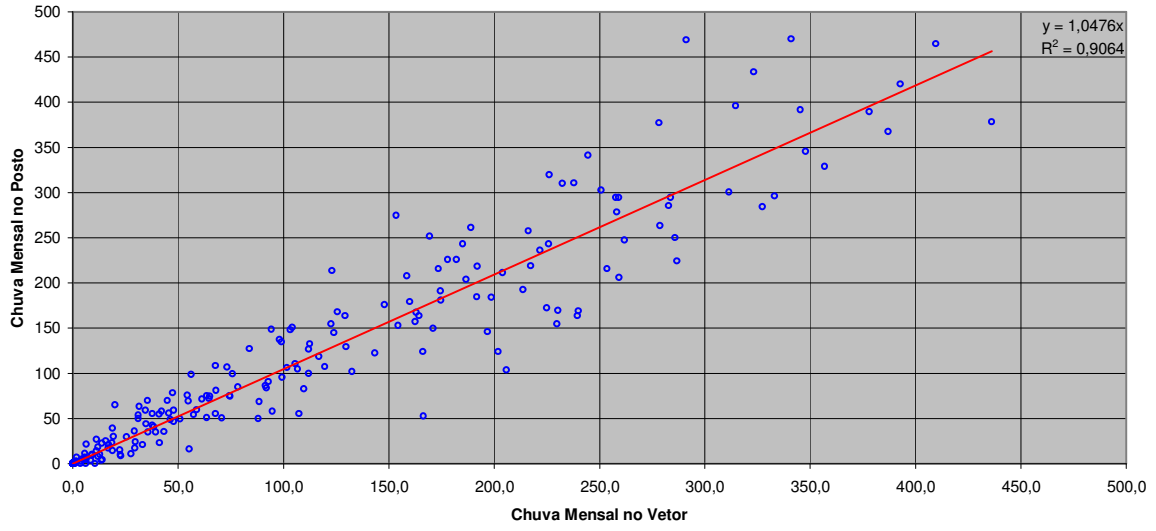
Dados Falha POSTO: Morada Nova



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

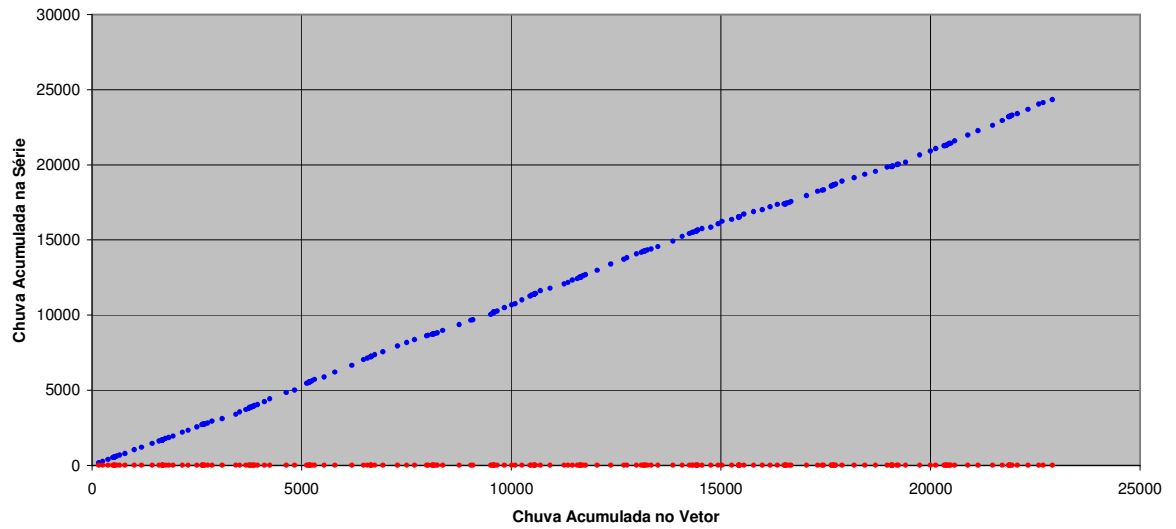
POSTO: Paineiras



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

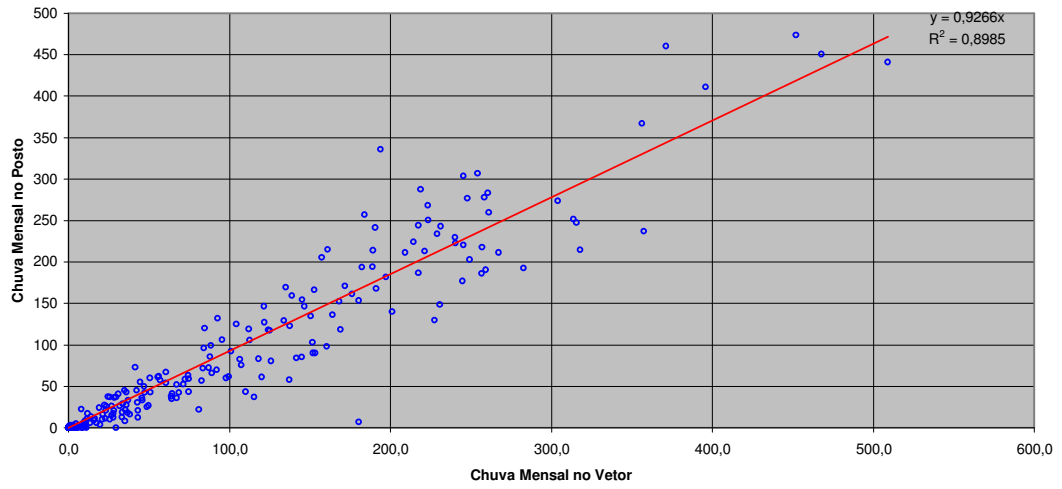
Dados Falha POSTO: Paineiras



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

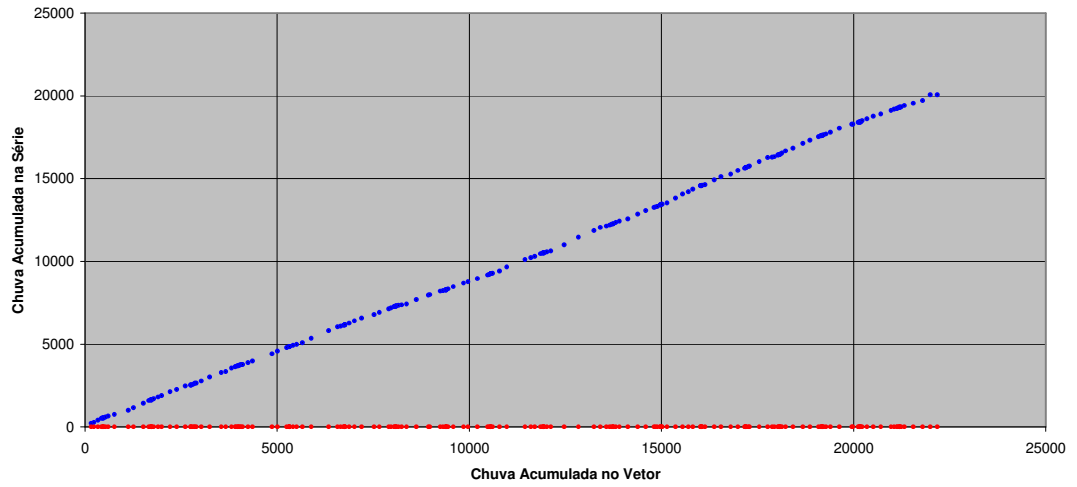
POSTO: Porto Mesquita



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

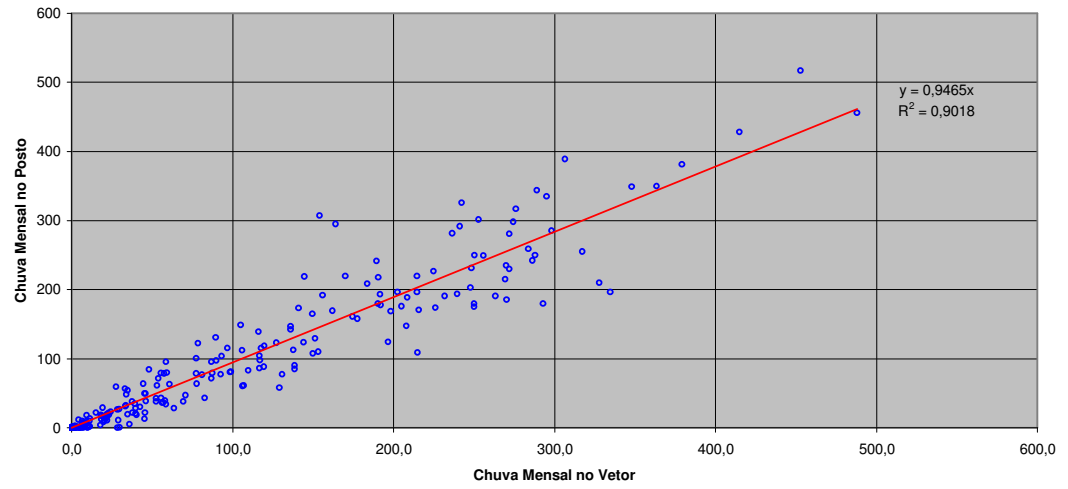
Dados Falha POSTO: Porto Mesquita



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

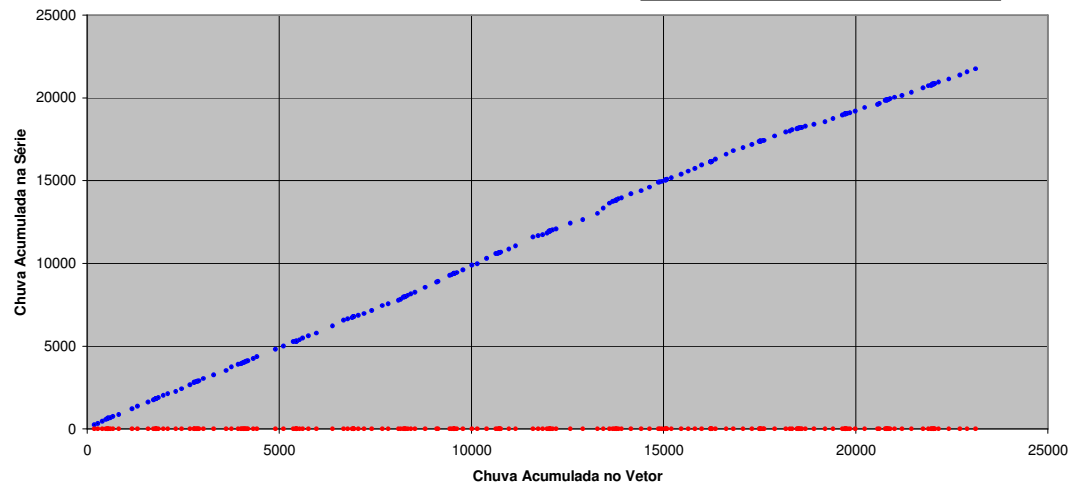
POSTO: Porto Pará



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

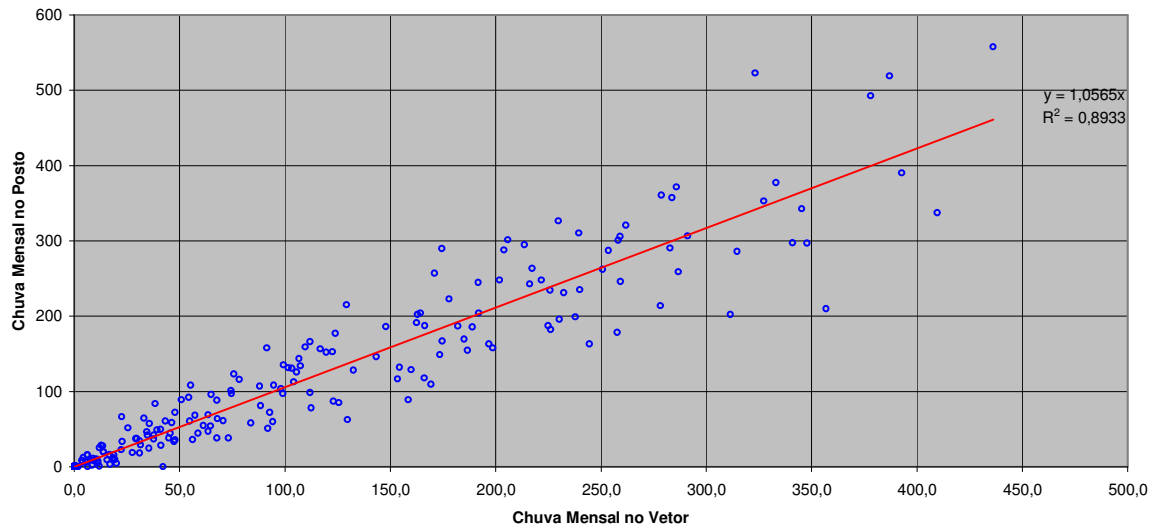
Dados Falha POSTO: Porto Pará



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

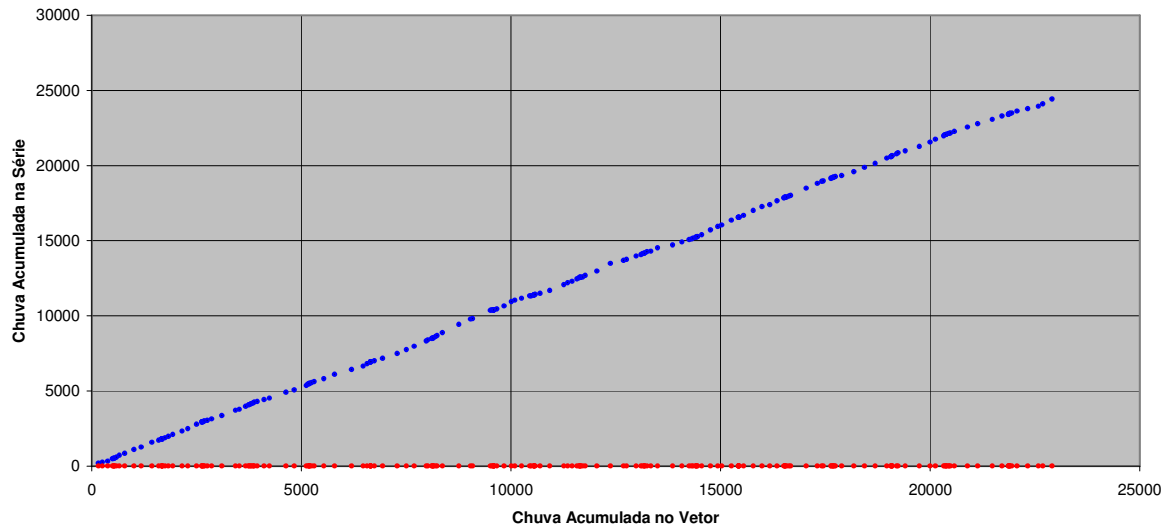
POSTO: São Gonçalo do Abaeté



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

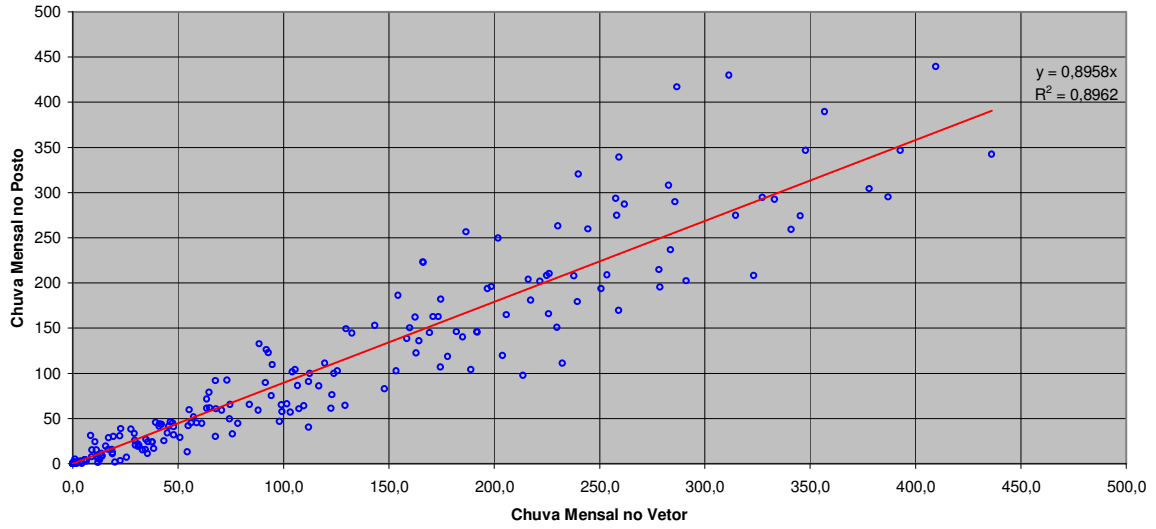
• Dados • Falha POSTO: São Gonçalo do Abaeté



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

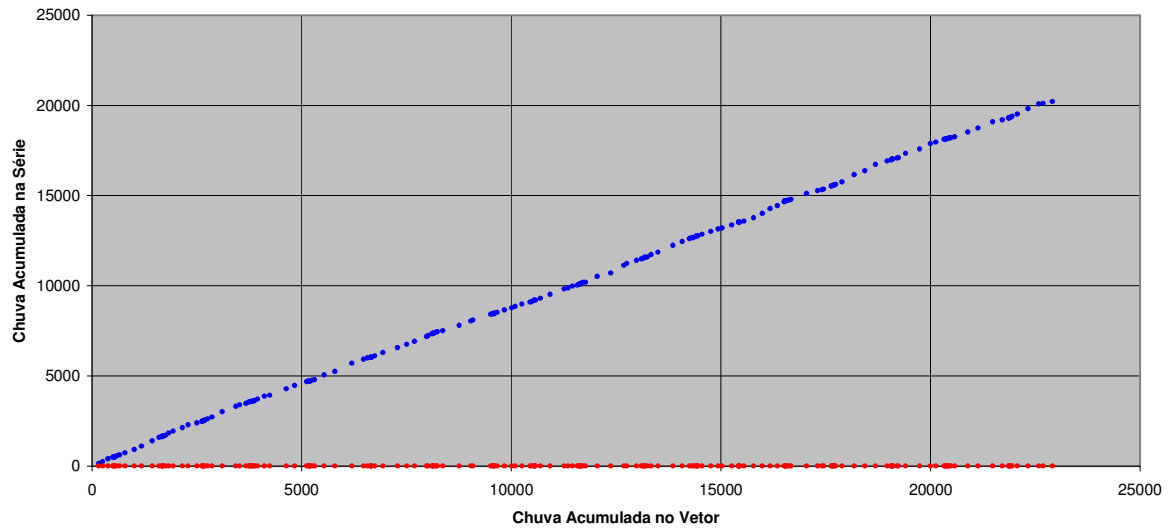
POSTO: São José dos Buritis



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

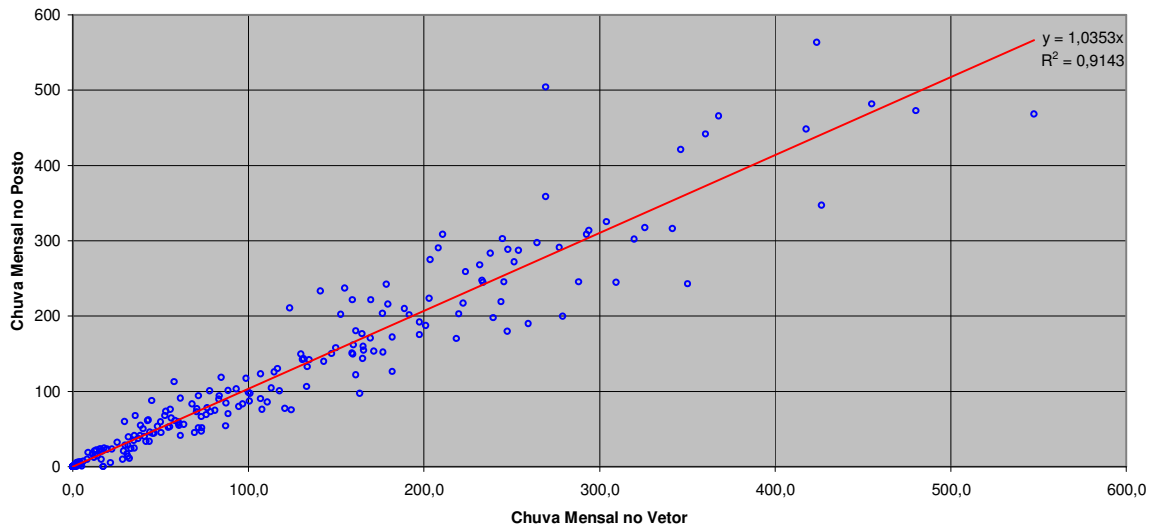
Dados Falha POSTO: São José dos Buritis



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

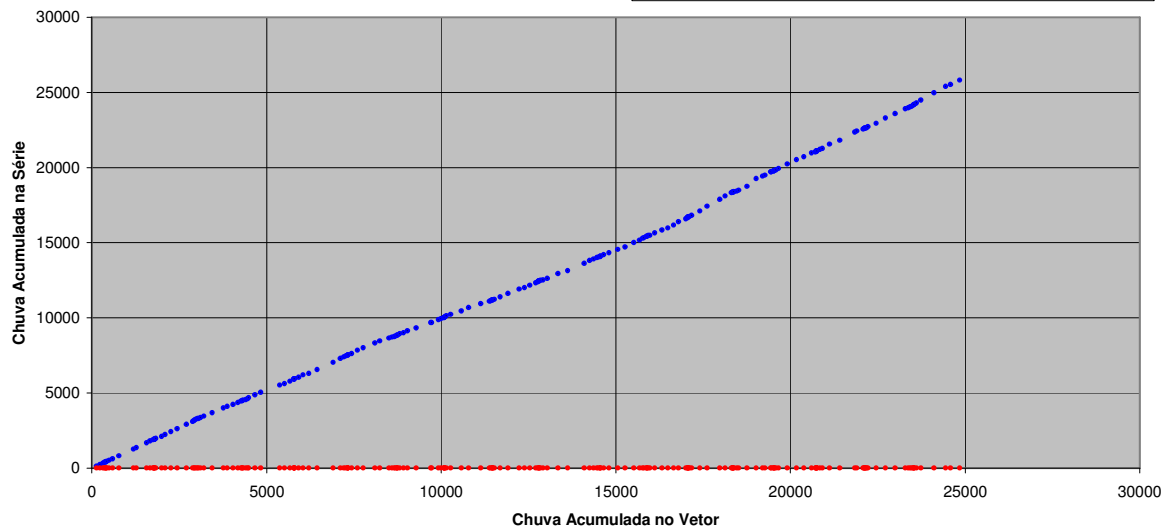
POSTO: São Sebastião do Oeste



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

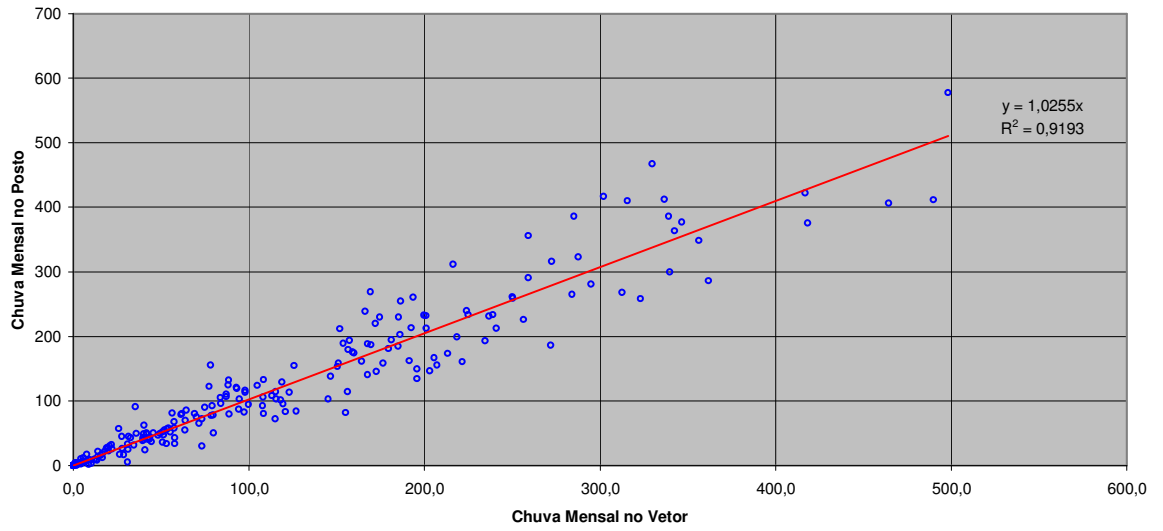
Dados Falha POSTO: São Sebastião do Oeste



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

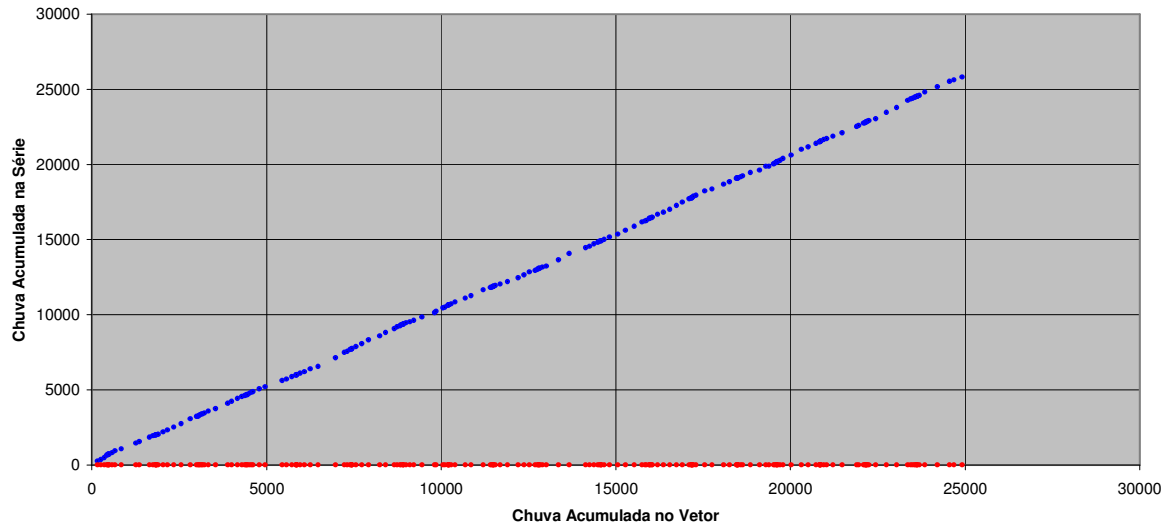
POSTO: SE Pedra do Indaiá



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

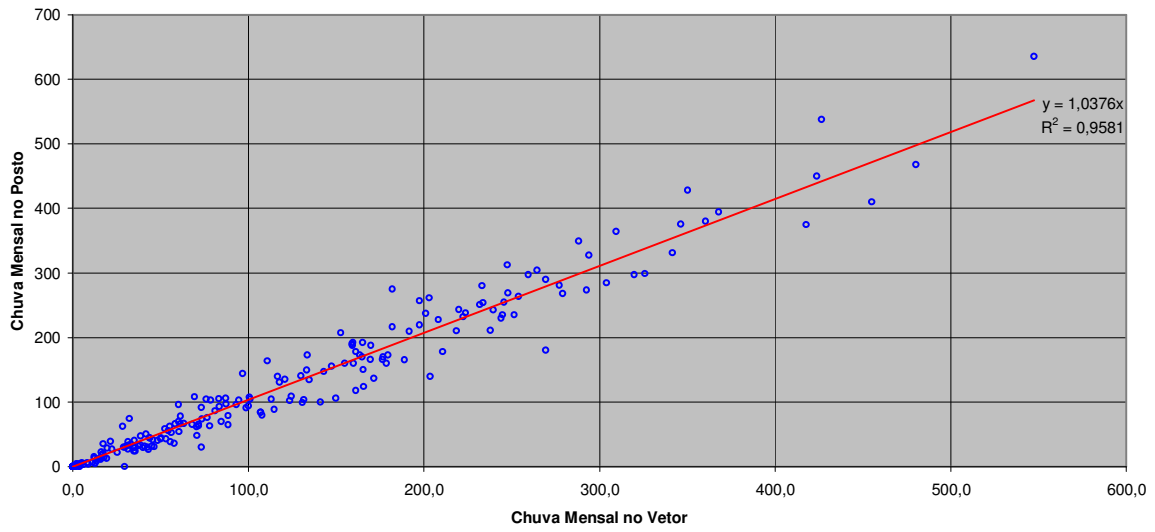
Dados Falha POSTO: SE Pedra do Indaiá



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

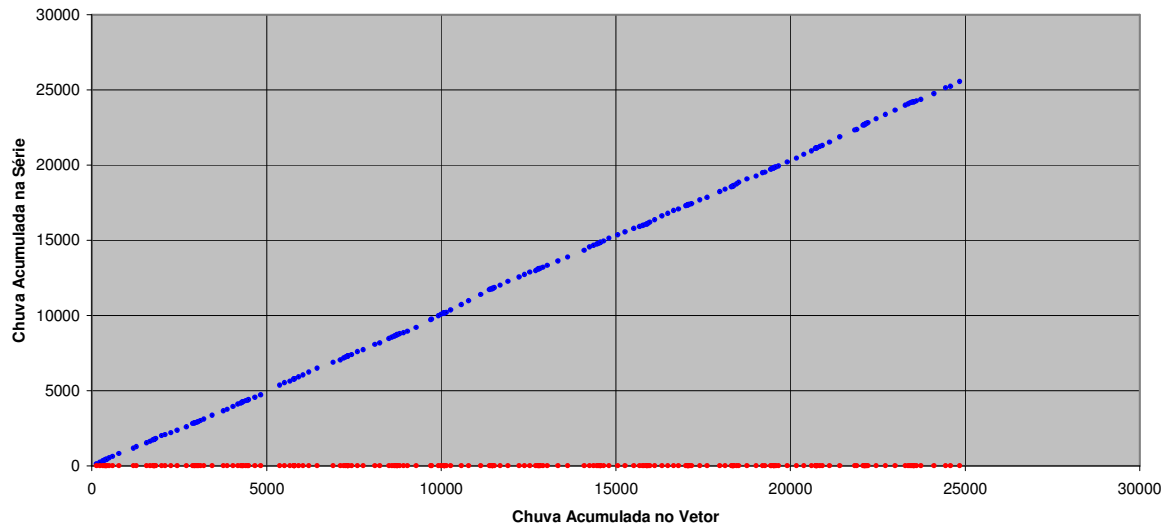
POSTO: UHE Cajuru



TROCAR POSTO

DUPLO ACUMULADA COM O VETOR REGIONAL

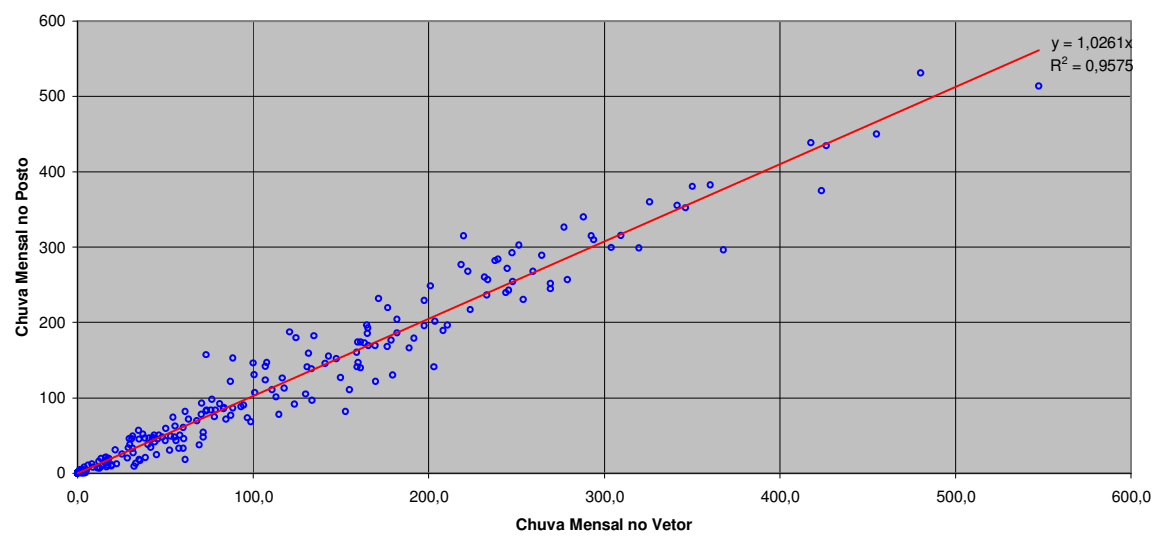
Dados Falha POSTO: UHE Cajuru



TROCAR POSTO

CORRELAÇÃO COM O VETOR REGIONAL

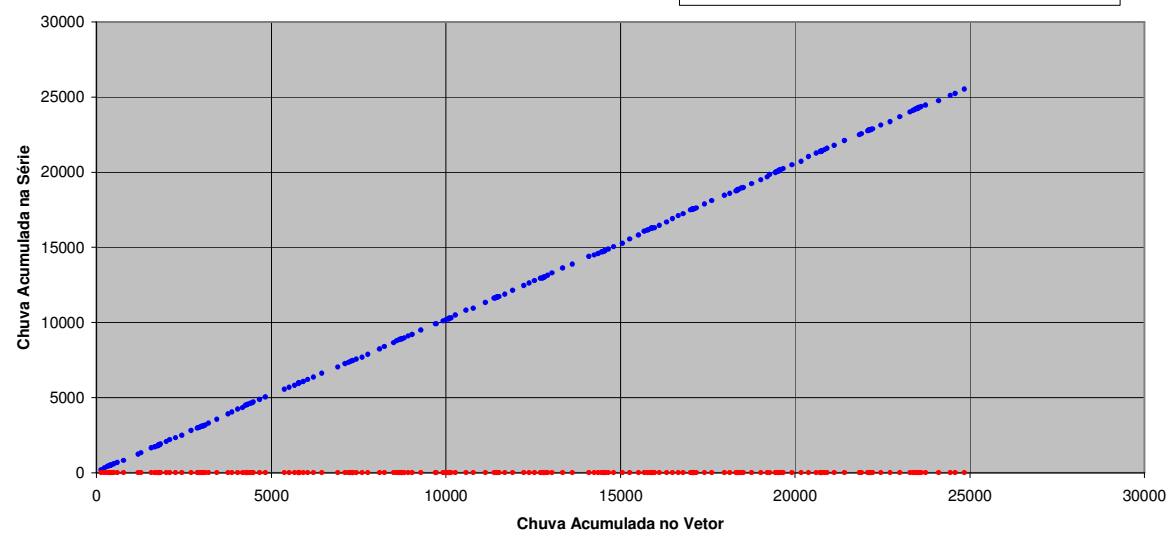
POSTO: UHE Gafanhoto



TROCAR POSTO

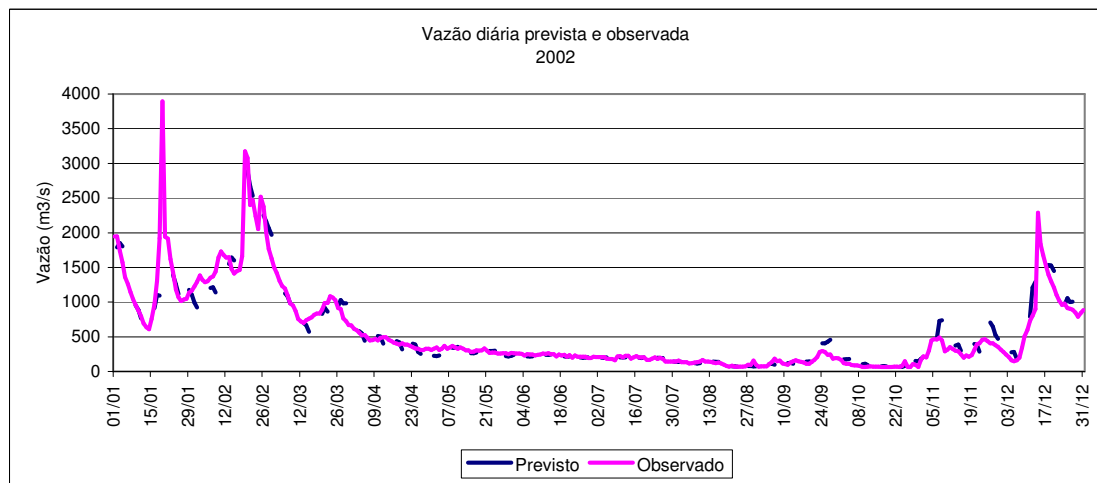
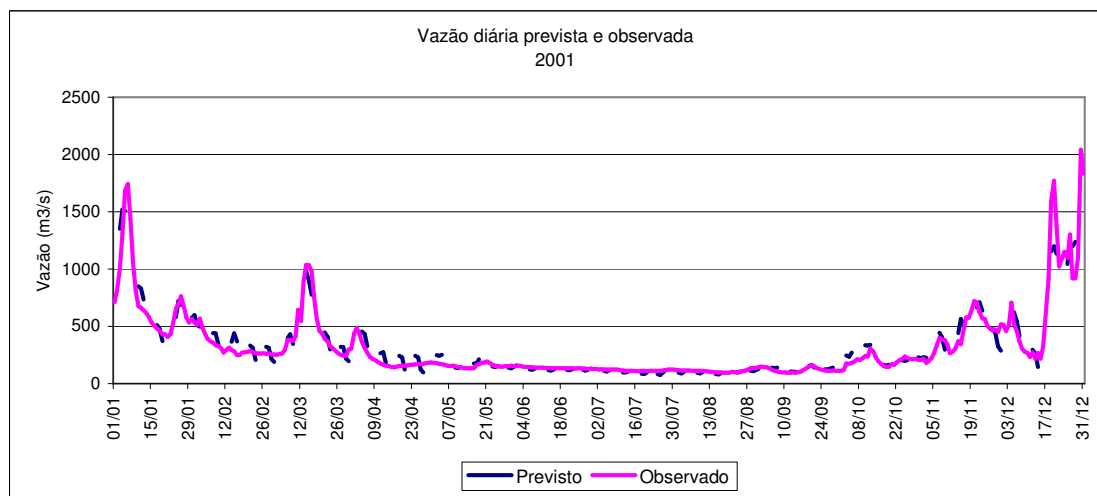
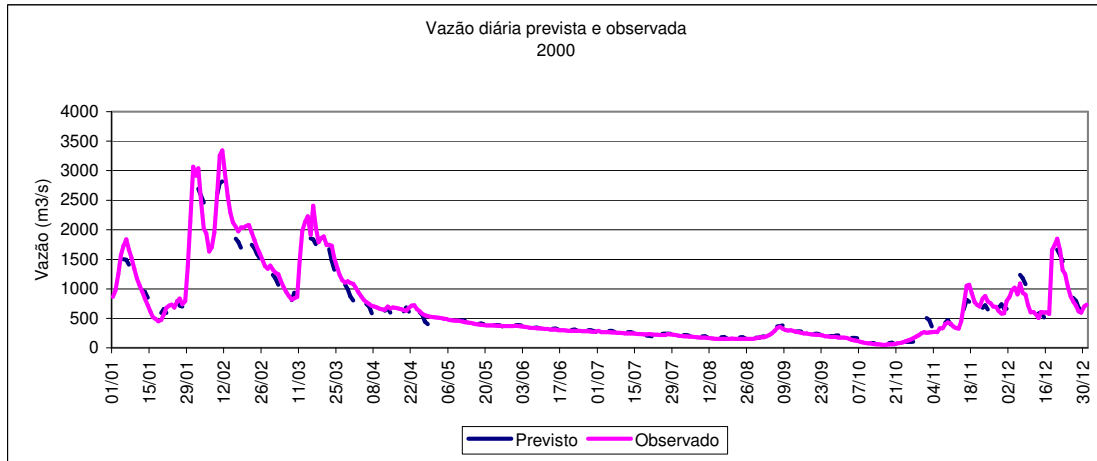
DUPLO ACUMULADA COM O VETOR REGIONAL

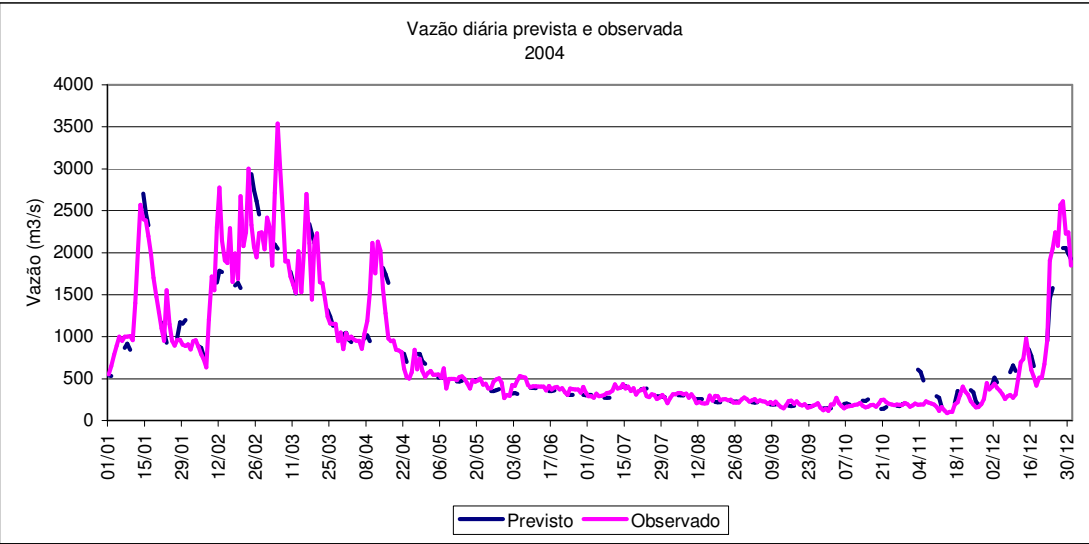
Dados Falha POSTO: UHE Gafanhoto



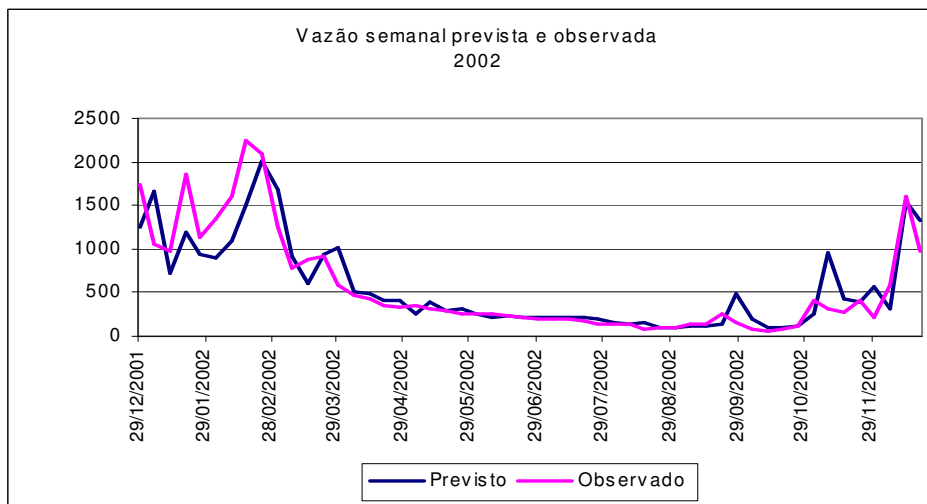
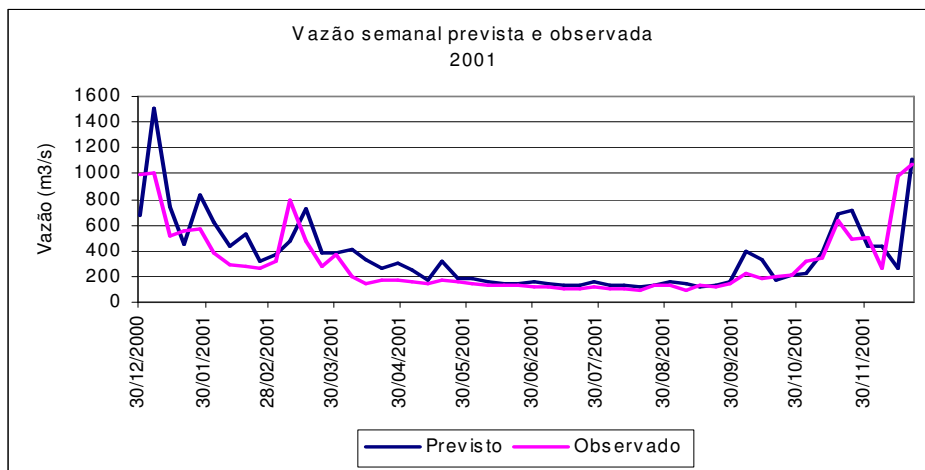
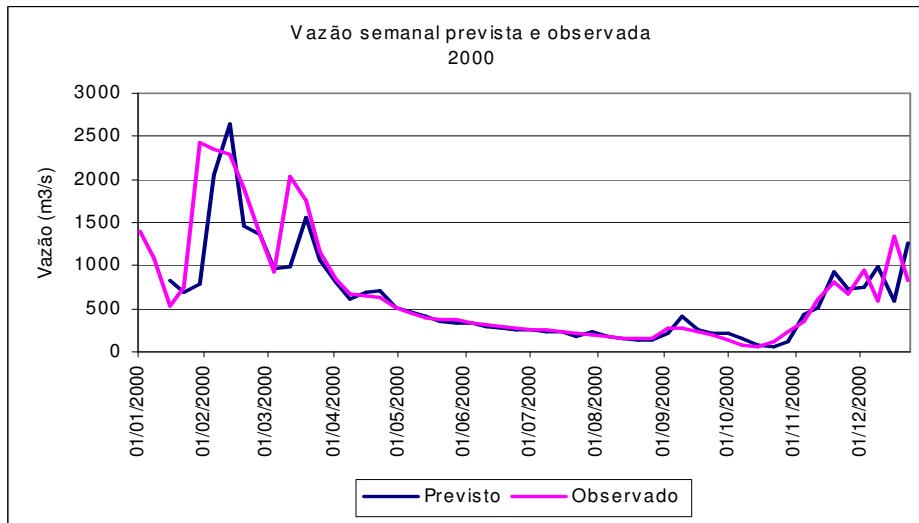
Anexo 02 – Previsões de vazões diárias e semanais nos padrões do PMO

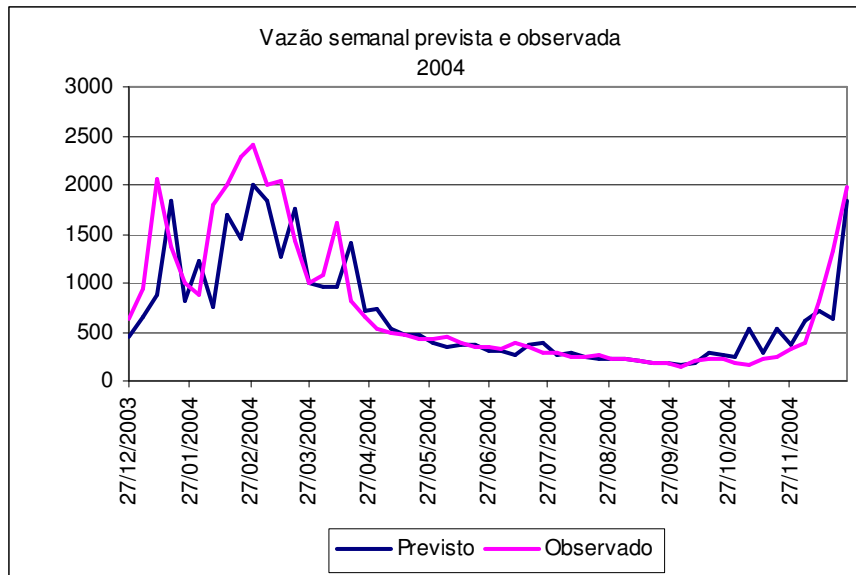
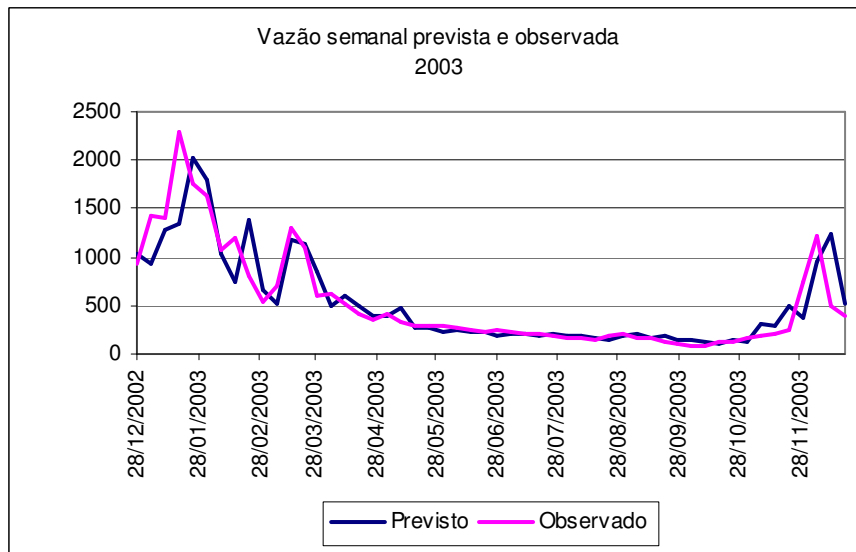
- Previsões de vazões diárias com a rede MLP



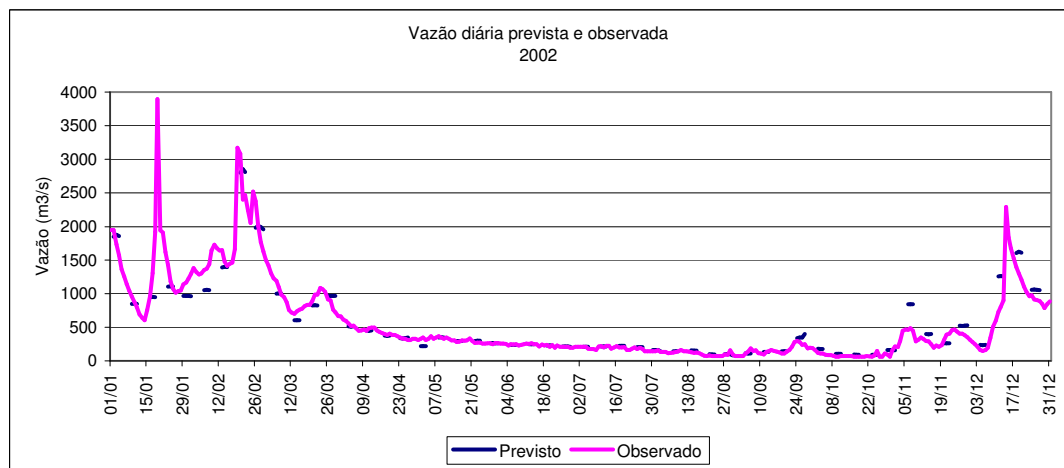
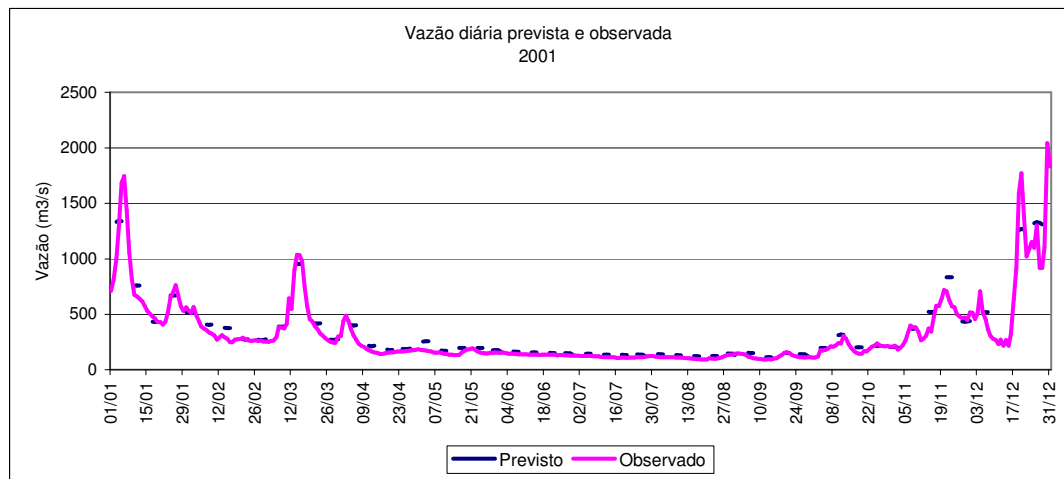
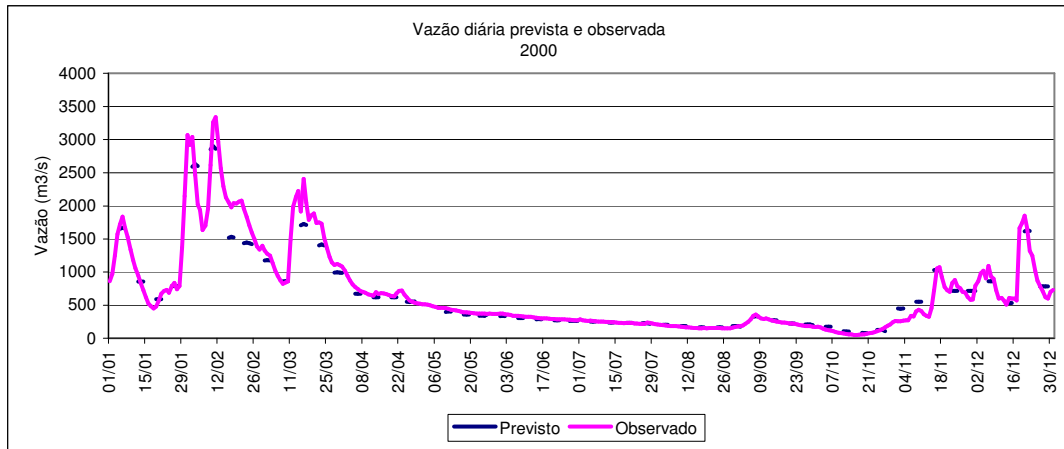


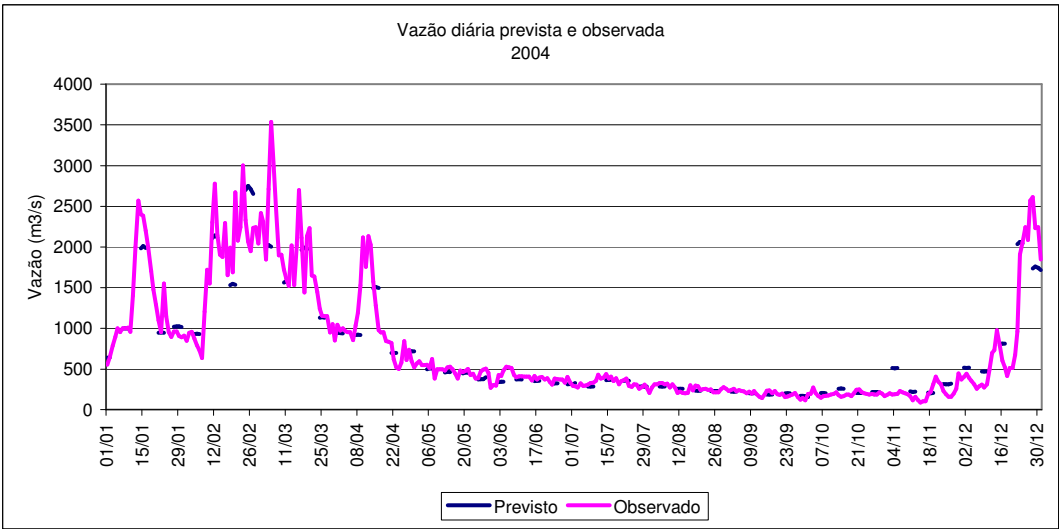
■ Previsões de vazões semanais com a rede MLP



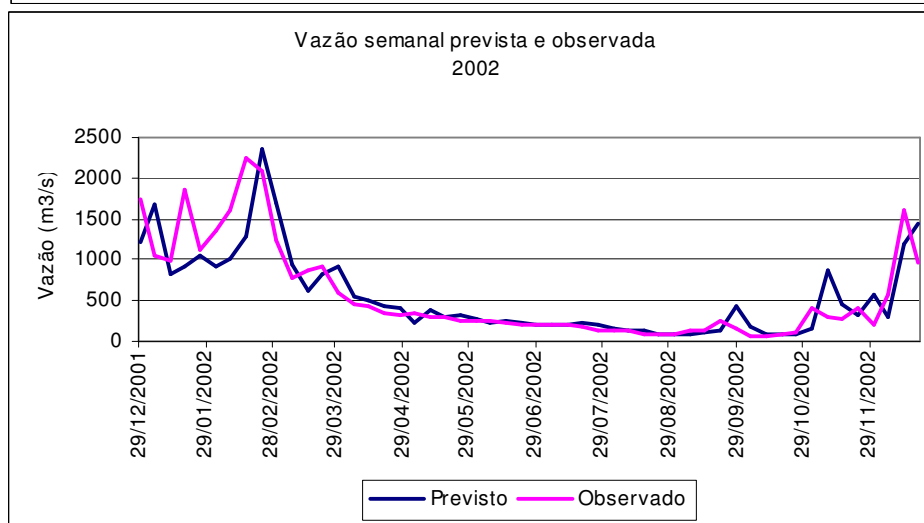
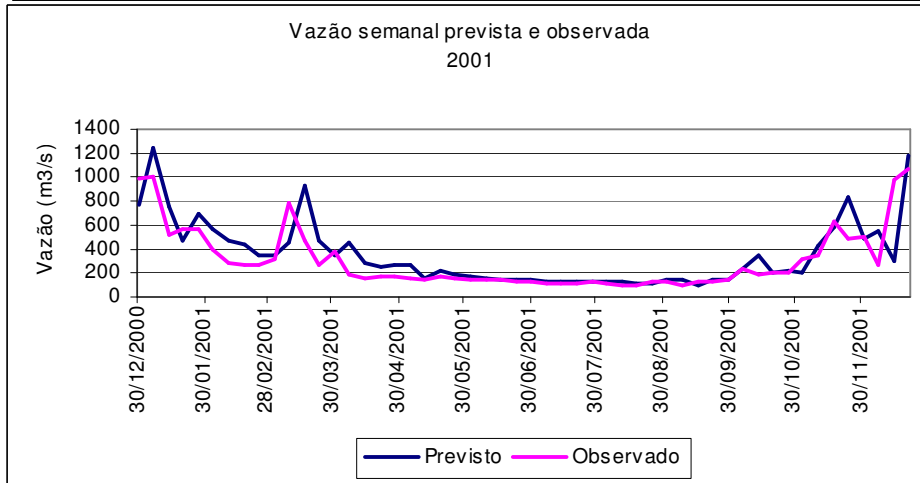
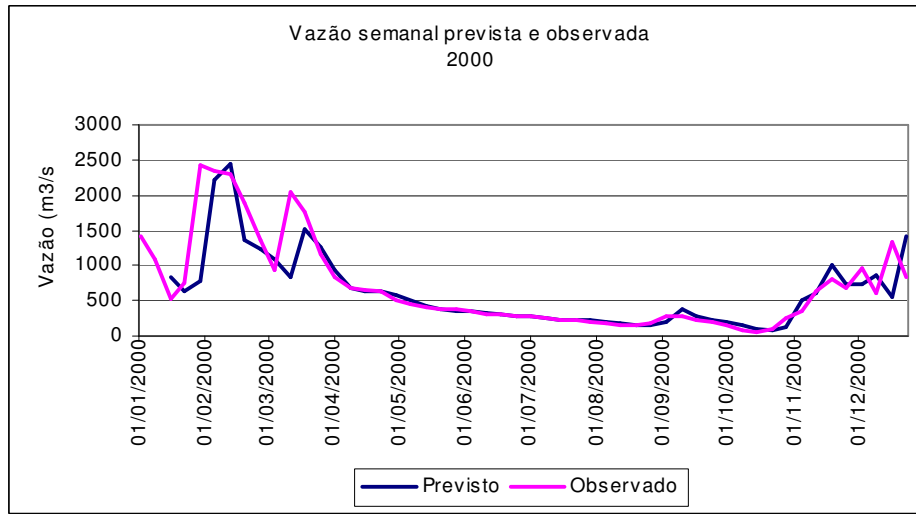


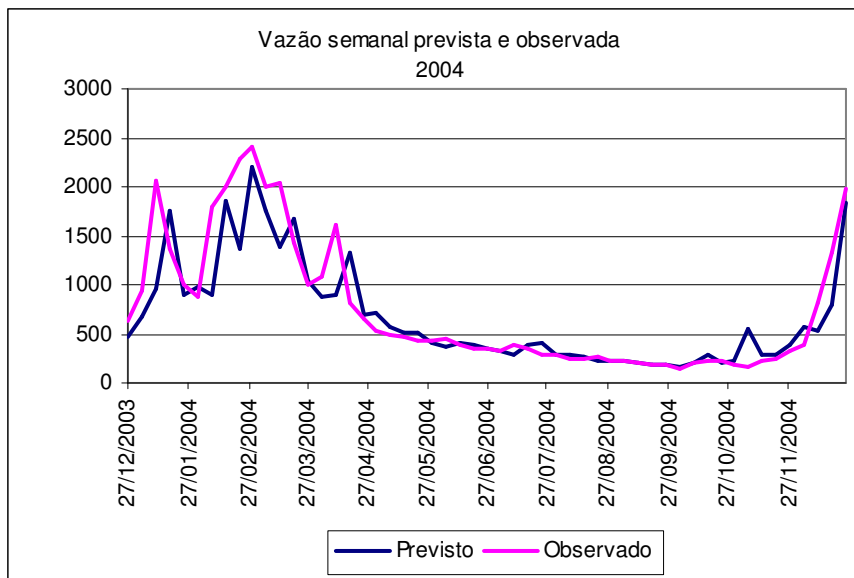
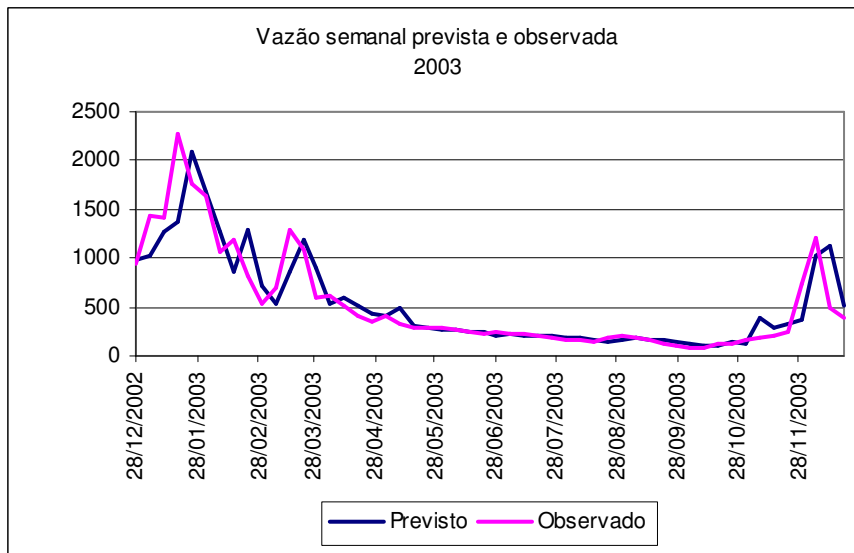
■ Previsões de vazões diárias com a rede NSRBN





- Previsões de vazões semanais com a rede NSRBN





Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)