



GIUSEPPE ALESSANDRO SIGNORIELLO

**Modelos Matemáticos para Previsão de  
Vazões Afluentes à Aproveitamentos  
Hidrelétricos**

68/2014

CAMPINAS

2014



**UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA MECÂNICA**

**GIUSEPPE ALESSANDRO SIGNORIELLO**

**Modelos Matemáticos para Previsão de  
Vazões Afluentes à Aproveitamentos  
Hidrelétricos**

Dissertação de Mestrado apresentada à Faculdade de Engenharia Mecânica da Universidade Estadual de Campinas como parte dos requisitos exigidos para obtenção do título de Mestre em Planejamento de Sistemas Energéticos.

Orientadora Prof<sup>ª</sup>. Dr<sup>ª</sup>. Ieda Geriberto Hidalgo

ESTE EXEMPLAR CORRESPONDE À  
VERSÃO FINAL DA DISSERTAÇÃO  
DEFENDIDA PELO ALUNO GIUSEPPE  
ALESSANDRO SIGNORIELLO E ORIENTADO  
PELA PROFA. DRA. IEDA GERIBERTO  
HIDALGO

A handwritten signature in black ink, appearing to read "Ieda Geriberto Hidalgo", is written over a horizontal dotted line.

**CAMPINAS**

**2014**

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca da Área de Engenharia e Arquitetura  
Rose Meire da Silva - CRB 8/5974

Si26m Signoriello, Giuseppe Alessandro, 1977-  
Modelos matemáticos para previsão de vazões afluentes à aproveitamentos hidrelétricos / Giuseppe Alessandro Signoriello. – Campinas, SP : [s.n.], 2014.

Orientador: Ieda Geriberto Hidalgo.  
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica.

1. Previsão de vazões. 2. Lógica fuzzy. 3. Usinas hidrelétricas. 4. Reservatórios - Projeto e construção. I. Hidalgo, Ieda Geriberto. II. Universidade Estadual de Campinas. Faculdade de Engenharia Mecânica. III. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Mathematical models to predict inflows to hydropower plants

**Palavras-chave em inglês:**

Inflow forecasting

Fuzzy logic

Hydroelectric plants

Reservoirs - Design and construction

**Área de concentração:** Planejamento de Sistemas Energéticos

**Titulação:** Mestre em Planejamento de Sistemas Energéticos

**Banca examinadora:**

Ieda Geriberto Hidalgo [Orientador]

Paulo de Barros Correia

Takaaki Ohishi

**Data de defesa:** 31-07-2014

**Programa de Pós-Graduação:** Planejamento de Sistemas Energéticos

UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA MECÂNICA  
COMISSÃO DE PÓS GRADUAÇÃO EM ENGENHARIA MECÂNICA  
PLANEJAMENTO DE SISTEMAS ENERGÉTICOS

DISSERTAÇÃO DE MESTRADO ACADÊMICO

# Modelos Matemáticos para Previsão de Vazões Afluentes à Aproveitamentos Hidrelétricos

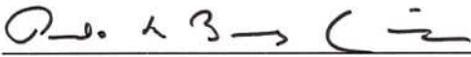
Autor: Giuseppe Alessandro Signoriello

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Ieda Geriberto Hidalgo

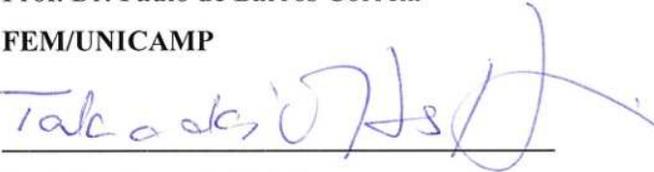
A Banca Examinadora composta pelos membros abaixo aprovou esta Dissertação:

  
\_\_\_\_\_

**Prof<sup>a</sup>. Dr<sup>a</sup>. Ieda Geriberto Hidalgo, Presidente**  
**FT/UNICAMP**

  
\_\_\_\_\_

**Prof. Dr. Paulo de Barros Correia**  
**FEM/UNICAMP**

  
\_\_\_\_\_

**Prof. Dr. Takaaki Ohishi**  
**FEEC/UNICAMP**

Campinas, 31 de julho de 2014

*À minha irmã Federica*

## **Agradecimentos**

Agradeço à professora Ieda Geriberto Hidalgo por me acolher sob sua valorosa orientação, pela confiança e cordialidade com que fui tratado por ela.

Agradeço à minha futura esposa Sayuri, companheira inigualável, pelo apoio, pela confiança, pela paciência e disponibilidade na revisão em português, pelo amor.

Meu eterno agradecimento aos meus pais e à minha irmã, que contribuíram e me apoiaram em todos os momentos da minha vida.

Aos amigos e colegas de convivência diária: Carolina, Daniele, Jéssica, Jhon, Jorge, Natalia, Simone e Viviana.

A todos os professores do 'Planejamento de Sistemas Energéticos'.

E a todos, que direta ou indiretamente, contribuíram para a realização e divulgação deste trabalho.

## Resumo

Este trabalho apresenta a comparação de dois modelos matemáticos desenvolvidos para prever vazões afluentes à usinas hidrelétricas. O objetivo é abordar os aspectos que determinam a qualidade do insumo fundamental para a programação da operação do sistema hidrotérmico brasileiro: a previsão de vazões.

A ferramenta de suporte à avaliação dos modelos matemáticos é o SISPREV, gerenciador/executor de estudos de previsão de vazões desenvolvido na UNICAMP. Esta ferramenta permite ao usuário prever vazões diárias e mensais com base em modelos de Regressão Linear (RL) e Sistema de Inferência Neuro-Fuzzy (SINF).

Avaliou-se a qualidade das previsões diárias e mensais dos modelos RL e SINF através da metodologia de mineração de dados Cross Industry Standard Process for Data Mining (CRISP-DM). A CRISP-DM é baseada em um modelo hierárquico de processos comumente usados na descoberta de conhecimento.

Os resultados mostram que o modelo RL apresenta um desempenho melhor para previsões diárias e o modelo SINF para as previsões mensais. Além disso, o modelo RL tem a tendência a ter bom desempenho de previsão nas situações típicas de chuva-vazão, enquanto os melhores índices de desempenho do modelo SINF caem nas condições atípicas, em particular com a contemporaneidade de altas vazões e baixas precipitações.

*Palavras Chave:* Previsão de vazões; Neuro-Fuzzy; Regressão Linear; Usinas hidrelétricas; Gerenciamento de reservatórios; Ferramentas computacionais.

## Abstract

This work presents a comparison between two different mathematical models developed to predict inflows to hydropower plants. The purpose is to explore the aspects that determine the quality of an important input variable for operation planning of the Brazilian hydrothermal system: the inflows forecasting.

The tool that supports the evaluation of the mathematical models is called SISPREV. It is a manager/runner of inflows forecasting studies developed at UNICAMP. This tool allows the user to predict daily and monthly inflows based on Linear Regression (RL) models and Neuro-Fuzzy Inference System (SINF).

In this thesis, was evaluated the quality of daily and monthly forecasts of RL and SINF models using the methodology Cross Industry Standard Process for Data Mining. CRISP-DM is used in the discovery of knowledge and based on a hierarchical process model.

The results show that the RL model performs better for daily predictions and the SINF model for monthly predictions. Furthermore, the RL model tends to have better performance in typical situations of rainfall-inflow, while the best performance indices of the SINF model fall in atypical conditions, in particular with the simultaneous high inflow rates and low precipitation.

*Key Words:* Inflow forecasting; Neuro-Fuzzy; Linear Regression; Hydroelectric plants; Reservoir management; Computational tool.

## Lista de Figuras

Figura 1.1	Oferta interna de energia elétrica por fonte - 2011 (EPE, 2013 – adaptada).	2
Figura 1.2	Capacidade de potência elétrica instalada – 1883 a 2011 (GW) (MME, 2012 – adaptada).	3
Figura 1.3	Estrutura da capacidade instalada – 1970 a 2011 (%) (MME, 2012).	3
Figura 1.4	Representação simplificada da integração entre os sistemas de produção e transmissão de energia elétrica (ONS, 2013).	4
Figura 3.1	Métodos de engenharia do conhecimento (BAUCHSPIESS, 2004).	22
Figura 3.2	Diagrama de dispersão.	24
Figura 3.3	Reta de RLS.	25
Figura 3.4	Significado da variância constante dos resíduos.	31
Figura 3.5	Natureza probabilística das variáveis do modelo de regressão (Rouhani, 2006).	33
Figura 3.6	Comunicação entre neurônios (BALLONE, 2013).	34
Figura 3.7	Modelo de neurônio artificial.	35
Figura 3.8	Ciclo de vida de uma RNA, baseado em Bialoskorski (2006).	36
Figura 3.9	Diagrama esquemático do sistema de inferência.	40
Figura 3.10	Estrutura de um SINF dividido em cinco camadas.	44
Figura 3.11	Estrutura geral do modelo (LUNA, 2011).	45
Figura 4.1	Estrutura geral da ferramenta de gestão de estudos de previsão de vazões.	51
Figura 4.2	Exemplo de resultado em forma gráfica do SISPREV.	52
Figura 4.3	Exemplo de resultado em forma de planilha do SISPREV.	53
Figura 5.1	Fases do CRISP-DM (CHAPMAN <i>et al</i> , 2000 - adaptada).	57
Figura 6.1	Localização das usinas hidrelétricas (AES, 2013).	66
Figura 6.2	Janela operativa do SQL para listar as vazões e precipitações da UHE analisada.	70
Figura 6.3	Água Vermelha – Gráfico de dispersão diária vazão vs. precipitação.	71
Figura 6.4	Bariri – Gráfico de dispersão diária vazão vs. precipitação.	71

Figura 6.5	Barra Bonita – Gráfico de dispersão diária vazão vs. precipitação.	71
Figura 6.6	Euclides da Cunha – Gráfico de dispersão diária vazão vs. precipitação.	71
Figura 6.7	Graminha – Gráfico de dispersão diária vazão vs. precipitação.	71
Figura 6.8	Ibitinga – Gráfico de dispersão diária vazão vs. precipitação.	71
Figura 6.9	Limoeiro – Gráfico de dispersão diária vazão vs. precipitação.	72
Figura 6.10	Nova Avanhandava – Gráfico de dispersão diária vazão vs. precipitação.	72
Figura 6.11	Promissão – Gráfico de dispersão diária vazão vs. precipitação.	72
Figura 6.12	Água Vermelha – Gráfico de dispersão mensal vazão vs. precipitação.	72
Figura 6.13	Bariri – Gráfico de dispersão mensal vazão vs. precipitação.	72
Figura 6.14	Barra Bonita – Gráfico de dispersão mensal vazão vs. precipitação.	73
Figura 6.15	Euclides da Cunha – Gráfico de dispersão mensal vazão vs. precipitação.	73
Figura 6.16	Graminha – Gráfico de dispersão mensal vazão vs. precipitação.	73
Figura 6.17	Ibitinga – Gráfico de dispersão mensal vazão vs. precipitação.	73
Figura 6.18	Limoeiro – Gráfico de dispersão mensal vazão vs. precipitação.	73
Figura 6.19	Nova Avanhandava – Gráfico de dispersão mensal vazão vs. precipitação.	73
Figura 6.20	Promissão – Gráfico de dispersão mensal vazão vs. precipitação.	74
Figura 6.21	Água Vermelha – Gráfico de dispersão diária vazão vs. precipitação média.	74
Figura 6.22	Bariri – Gráfico de dispersão diária vazão vs. precipitação média.	74
Figura 6.23	Barra Bonita – Gráfico de dispersão diária vazão vs. precipitação média.	75
Figura 6.24	Euclides da Cunha – Gráfico de dispersão diária vazão vs. precipitação média.	75
Figura 6.25	Graminha – Gráfico de dispersão diária vazão vs. precipitação média.	75
Figura 6.26	Ibitinga – Gráfico de dispersão diária vazão vs. precipitação média.	75
Figura 6.27	Limoeiro – Gráfico de dispersão diária vazão vs. precipitação.	75
Figura 6.28	Nova Avanhandava – Gráfico de dispersão diária vazão vs. precipitação média.	75

Figura 6.29	Promissão – Gráfico de dispersão diária vazão vs. precipitação média.	76
Figura 6.30	Vazões em correspondência da precipitação igual a zero encontradas pelas retas de RL.	76
Figura 6.31	Vazão diária observada mínima, média e máxima das usinas em ordem de vazão média.	77
Figura 6.32	Vazão mensal observada mínima, média e máxima das usinas em ordem de vazão média.	77
Figura 6.33	Precipitações diárias e mensais observadas.	78
Figura 6.34	Distribuição das observações mensais no plano vazão-precipitação.	79
Figura 6.35	Distribuição das observações mensais no plano vazão-precipitação.	80
Figura 6.36	Precipitação mensal de chuva medida na UHE Graminha no período 2003-2007.	81
Figura 6.37	Gráfico de dispersão vazão prevista vs. vazão observada das previsões diárias.	83
Figura 6.38	Distribuição da frequência do erro resíduo percentual nas previsões diárias.	84
Figura 6.39	Gráfico de dispersão MAPE – vazão observada- precipitação observada do modelo RL.	85
Figura 6.40	Gráfico de dispersão MAPE – vazão observada- precipitação observada do modelo SINF.	85
Figura 6.41	Repartição percentual do índice MAPE nos estudos diários.	86
Figura 6.42	Gráfico de dispersão precipitação observada vs. vazão observada, com indicação do MAPE.	86
Figura 6.43	Repartição percentual do índice MAPE na análise dos estudos diários com vazão média maior que 2.000 m <sup>3</sup> /s.	90
Figura 6.44	Confusion Index para as previsões diárias.	91
Figura 6.45	Gráfico de dispersão vazão prevista vs. vazão observada das previsões mensais.	91
Figura 6.46	Distribuição da frequência do erro resíduo percentual nas previsões diárias.	92
Figura 6.47	Gráfico de dispersão NS – vazão observada- precipitação observada do modelo RL.	93

Figura 6.48	Gráfico de dispersão NS – vazão observada- precipitação observada do modelo SINF.	93
Figura 6.49	Comparação dos modelos com o preditor de referência.	93
Figura 6.50	Repartição percentual do índice NS nos estudos mensais.	94
Figura 6.51	Gráfico de dispersão precipitação observada vs. vazão observada, com indicação do NS.	95
Figura 6.52	Confusion Index para as previsões mensais.	96
Figura 6.53	Precipitações médias observadas nas UHEs no período 2000-2011 e nos semestres 2000-2001, 2013-2014.	97
Figura 6.54	Graminha - Resultados das execuções dos modelos RL e SINF no semestre 10/2013 – 03/2014.	98
Figura 6.55	Euclides da Cunha - Resultados das execuções dos modelos RL e SINF no semestre 10/2013 – 03/2014.	98
Figura 6.56	Barra Bonita - Resultados das execuções dos modelos RL e SINF no semestre 10/2013 – 03/2014.	98
Figura 6.57	Bariri - Resultados das execuções dos modelos RL e SINF no semestre 10/2013 – 03/2014.	98
Figura 6.58	Ibitinga - Resultados das execuções dos modelos RL e SINF no semestre 10/2013 – 03/2014.	98
Figura 6.59	Promissão - Resultados das execuções dos modelos RL e SINF no semestre 10/2013 – 03/2014.	98
Figura 6.60	Água Vermelha - Resultados das execuções dos modelos RL e SINF no semestre 10/2013 – 03/2014.	99

## Lista de Tabelas

Tabela 3.1	Comparação entre as redes neurais artificiais e os SIFs (LUNA, 2003).	43
Tabela 5.1	Índices de desempenho usados nas previsões de vazões.	64
Tabela 6.1	UHEs codificadas no SISPREV.	67
Tabela 6.2	Postos de medição codificados no SISPREV.	67
Tabela 6.3	Variáveis de entrada para os modelos de previsões diária.	68
Tabela 6.4	Variáveis de entrada para os modelos de previsões mensal.	69
Tabela 6.5	Períodos escolhidos para previsões diárias e mensais.	81
Tabela 6.6	Distribuição percentual de MAPE considerado adequado.	87
Tabela 6.7	Avaliação do MAPE em períodos com vazões maiores de 2.000 m <sup>3</sup> /s.	89
Tabela 6.8	Distribuição percentual de NS considerado adequado.	96
Tabela 6.9	Índices NS relativos às previsões de vazões mensais no semestre outubro 2013 – março 2014.	99
Tabela A.1	Gráficos contendo os resultados das execuções dos modelos RL e SINF, apresentados com as vazões e precipitações observadas.	112
Tabela A.2	Índices de desempenho do modelo RL para as previsões diárias.	120
Tabela A.3	Índices de desempenho do modelo SINF para as previsões diárias.	122
Tabela B.1	Gráficos contendo os resultados das execuções dos modelos RL e SINF, apresentados com as vazões e precipitações observadas.	125
Tabela B.2	Índices de desempenho do modelo RL para as previsões mensais.	132
Tabela B.3	Índices de desempenho do modelo SINF para as previsões mensais.	134

## Lista de Abreviaturas e Siglas

AG	Algoritmos Genéticos
ANEEL	Agência Nacional de Energia Elétrica
CCEE	Câmara de Comercialização de Energia Elétrica
CI	Confusion Index
CMO	Custo Marginal de Operação
CMSE	Comitê de Monitoramento do Setor Elétrico
CNPE	Conselho Nacional de Política Energética
CRISP-DM	Cross Standard Process for Data Mining
CULMG	Centro Universitário do Leste de Minas Gerais
DM	Distância Multicritério
DPFT	Diferenças Primeiras da Função de Transferência
DW	Durbin-Watson
EM	Expectation Maximization
EPE	Empresa de Pesquisa Energética
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ME	Mean Error
MEL	Modelo Estocástico Linear
MIAS	Métodos de Inteligência Artificial Simbólicos
MLP	MultiLayer Perceptron
MLP-LM	MultiLayer Perceptron Levenberg-Marquardt
MME	Ministério das Minas e Energia
MMQ	Método dos Mínimos Quadrados
MPE	Mean Percentage Error
MSE	Mean Squared Error
NS	Nash-Sutcliffe
NS-LOG	Nash-Sutcliffe aplicado aos logaritmos
NSRBN	Non-linear Sigmoidal Regression Blocks Networks
ONS	Operador Nacional do Sistema Elétrico
PDE	Programação Dinâmica Estocástica

PELP	Plano de Expansão de Longo Prazo do Setor Elétrico
PMO	Programa Mensal da Operação
PUC-RJ	Pontifícia Universidade Católica do Rio de Janeiro
RL	Regressão Linear
RLM	Regressão Linear Múltipla
RLS	Regressão Linear Simples
RMSE	Root Mean Square Error
RNA	Rede Neural Artificial
RNB	Redes Neural Biológica
SD	Sistemas Difusos
SEB	Setor Elétrico Brasileiro
SH	Sistemas Híbridos
SIF	Sistema de Inferência Fuzzy
SIN	Sistema Interligado Nacional
SINF	Sistema de Inferência Neuro Fuzzy
SMAP	Soil Moisture Accounting Procedure
SWAT	Soil and Water Assessment Tool
TEP	Tonelada Equivalente de Petróleo
UB	Universidade de Brasília
UFL	Universidade Federal de Lavras
UFMG	Universidade Federal Minas Gerais
UFP	Universidade Federal do Paraná
UFRGS	Universidade Federal do Rio Grande do Sul
UFRJ	Universidade Federal do Rio de Janeiro
UFSC	Universidade Federal de Santa Catarina
UHE	Usina Hidrelétrica
UNICAMP	Universidade Estadual de Campinas
USP	Universidade de São Paulo

## Lista de Símbolos

$a$	número de componentes do vetor de entrada que contém informações de precipitações observadas e futuras.
$A$	conjunto Fuzzy .
$b$	número de componentes do vetor de entrada que contém informações de vazões observadas.
$b_k$	coeficiente técnico da reta de regressão linear.
$c_i$	centro que define cada sub-região associada à $i$ -ésima regra.
$C$	matriz que contém os parâmetros antecedentes do modelo SINF.
$d$	teste DW na regressão linear.
$D$	dias de precipitação considerados pelo modelo SINF.
$e_k$	diferença entre o valor observado da variável dependente e o valor predito pela equação de RL.
$f_A(x)$	pertinência dos elementos do conjunto $A$ no universo $X$ nos conjuntos Fuzzy.
$F$	Teste F na RL.
$g_i(x^h)$	grau de pertinência de cada padrão $x^h$ .
$h$	número de interações no modelo SINF.
$i$	índice de tempo da série temporal a ser predita.
$k$	número de variáveis independentes em modelos RL, RNA, Fuzzy, SINF.
$L$	número de entrada-saída usados para a otimização do modelo SINF.
$min$	menor valor no método MMQ.
$M$	número de sub-regiões que compõem o espaço de entrada representado, cada uma representada por uma regra de Fuzzy.
$n$	número de pontos no diagrama de dispersão.
$N$	número de previsões avaliadas.
$P(t)$	precipitação, variável de entrada, (função do período) para os modelos RL e SINF [mm].
$P_i^h$	probabilidade de escolher a regra $i$ .
$q$	previsão com tendência errada se $q=1$ , previsão com tendência certa se $q=0$ .
$Q(t)$	vazão, variável de entrada, (função do período) para os modelos RL e SINF [ $m^3/s$ ].
$r$	vetor de entrada que contém informações de precipitações observadas e futuras, no modelo SINF.
$R_i$	regras do modelo SINF.
$R^2$	coeficiente de Determinação da reta de RL.
$\bar{R}^2$	coeficiente de Determinação ajustado da reta de RL.
$s$	vetor de entrada que contém informações de vazões observadas, no modelo SINF.
$S_e$	erro padrão da estimação da reta de RL.

$t$	período de avaliação das variáveis de entrada em modelos RL e SINF medido em dias para previsões diárias e meses para previsões mensais.
$V_i$	matriz de covariância de dimensão $k \times k$ que define cada sub-região $M$ associada à $i$ -ésima regra, no modelo SINF.
$W_n$	ligações sinápticas no modelo neurônio artificial.
$x_k$	variável independente em modelos RL, RNA, Fuzzy e SINF.
$x'$	abscissas dos pontos do diagrama de dispersão na RL.
$x^h$	vetor de entrada no instante $h$ , no modelo SINF.
$y$	variável dependente em modelos RL, RNA, Fuzzy e SINF correspondente à vazão prevista [m <sup>3</sup> /s].
$y'$	ordenadas dos pontos do diagrama de dispersão na RL correspondente à vazão observada [m <sup>3</sup> /s].
$y^h$	saída de cada regra no modelo SINF [m <sup>3</sup> /s].
$\bar{y}$	média das vazões observadas no período considerado [m <sup>3</sup> /s].
$\bar{y}^{\bar{n}}$	saída do modelo SINF, para uma determinada entrada $x^h$ [m <sup>3</sup> /s].
$\mu_A(x)$	função de pertinência de $x$ em $A$ no conjunto Fuzzy.
$\theta_i$	vetor de coeficientes dos modelos locais lineares no modelo SINF.
$\mathfrak{S}(D, \Omega)$	função objetivo no modelo SINF.
$\varnothing^n$	vetor que contém os dados de entrada $x^h$ no modelo SINF.
$\Omega$	matriz que contém os parâmetros do modelo SINF.

# Sumário

<b>1. Introdução</b>	1
1.1 O Sistema Elétrico Brasileiro	2
1.1.1 Planejamento da Expansão do Setor Elétrico Brasileiro	6
1.1.2 Planejamento da Operação do Sistema Hidrotérmico	6
1.2 Características do Planejamento da Operação do Sistema Hidrotérmico	7
1.3 Definições em Modelagem Hidrológica	8
1.4 Importância da Previsão de Vazões	9
1.5 Previsão de Vazões no Setor Elétrico Brasileiro	9
1.6 Objetivo e Resultados Esperados	10
1.7 Estrutura do Trabalho	11
<b>2. Revisão da literatura</b>	12
2.1 Estado de Arte	13
2.1.1 Modelos Conceituais	15
2.1.2 Modelos Empíricos	15
2.1.3 Modelos Combinados	18
2.2 Revisão Bibliográfica	18
<b>3. Referencial teórico</b>	21
3.1 Métodos de Engenharia do Conhecimento (EC)	22
3.2 Regressão Linear (RL)	24
3.2.1 Regressão Linear Simples (RLS)	25
3.2.2 Regressão Linear Múltipla (RLM)	26
3.2.3 Medidas de Ajustamento	27
3.2.4 Testes de Significância Estatística	28
3.2.5 Os Pressupostos da Análise de Regressão	30
3.2.6 Aspectos Positivos e Negativos da RL	32
3.3 Rede Neural Artificial (RNA)	34
3.3.1 Rede MultiLayer Perceptron (MLP)	37
3.3.2 Aspectos Positivos e Negativos da RNA	37
3.4 Lógica Nebulosa Fuzzy	38
3.4.1 Sistema de Inferência Fuzzy (SIF)	40
3.4.2 Aspectos Positivos e Negativos do SIF	42
3.5 Sistema de Inferência Neuro Fuzzy (SINF)	43
3.5.1 Algoritmo Expectation Maximization (EM) na versão <i>offline</i>	44
3.5.2 Aspectos Positivos e Negativos do SINF	46
3.6 Comportamento Esperado de um Modelo de Previsão de Vazões Baseado em RL e de um Modelo Baseado em SINF	48

<b>4. Ferramentas Utilizadas</b>	50
4.1 SISPREV	51
4.1.1 Modelo de Regressão Linear Múltipla (RLM)	53
4.1.2 Modelo Sistema de Inferência Neuro Fuzzy (SINF)	53
4.2 Outras Ferramentas	55
<b>5. Metodologia</b>	56
5.1 Cross Industry Standard Process for Data Mining (CRISP-DM)	57
5.2 Avaliação dos Modelos	58
<b>6. Estudo de Casos</b>	65
6.1 Entendimento do Problema	66
6.2 Compreensão dos Dados	66
6.3 Preparação dos Dados	80
6.4 Modelagem e Avaliação	82
6.4.1 Dados Diários	82
6.4.2 Dados Mensais	91
6.5 Avaliação das previsões de vazões mensais de 10/2013 – 03/2014	97
<b>7. Conclusões e Trabalhos Futuros</b>	100
7.1 Conclusões	101
7.2 Trabalhos Futuros	103
<b>8. Referências Bibliográficas</b>	104
<b>Apêndices</b>	112
A Estudos de previsão de vazões diárias	112
B Estudos de previsão de vazões mensais	125

# Capítulo 1

## Introdução

Este capítulo está dividido em sete seções. Nas duas primeiras serão apresentadas as características do sistema elétrico brasileiro. Na terceira, quarta e quinta seção, destaca-se a importância da qualidade das previsões de vazões no planejamento e na programação da operação energética. Na sexta seção, define-se o objetivo da tese e os resultados esperados. Por fim, na sétima seção, encontra-se a organização deste trabalho.

## 1.1 O Sistema Elétrico Brasileiro

A matriz energética brasileira chegou a 257 milhões de Tonelada Equivalente de Petróleo (TEP) em 2012. A participação de fontes renováveis de energia, tais como: hidráulica, biomassa e eólica é expressiva. Diferente da maioria dos países do mundo, no Brasil, as fontes renováveis representam mais de 45% da matriz energética. O restante é devido ao petróleo e derivados, gás natural, carvão e urânio (EPE, 2013).

No Setor Elétrico Brasileiro (SEB), a energia hidráulica representa 81,9% da oferta interna da energia, conforme ilustrado na Figura 1.1 Para isso, o território nacional conta com mais de 140 Usinas Hidrelétricas (UHEs) (EPE, 2013).

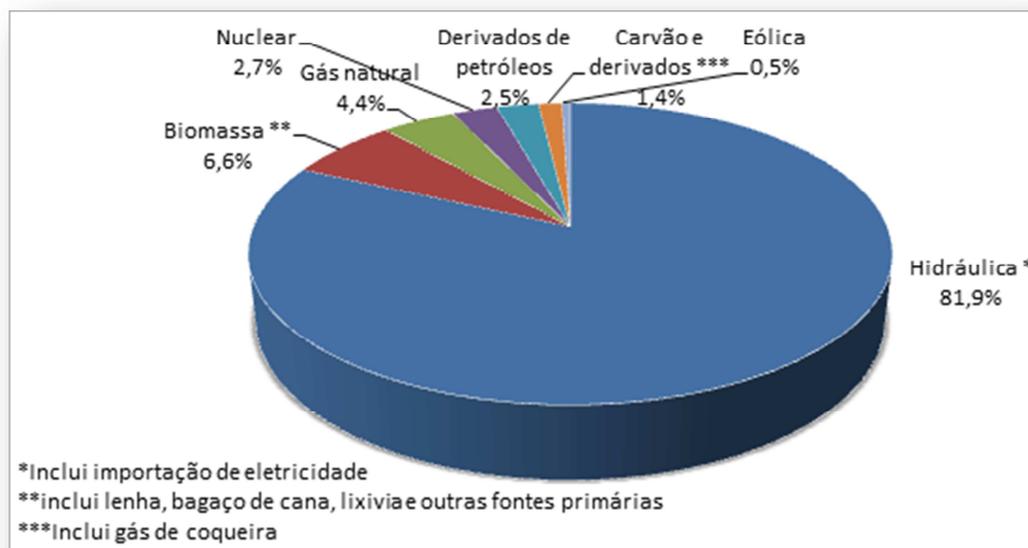


Figura 1.1 Oferta interna de energia elétrica por fonte - 2011 (EPE, 2013 – adaptada).

As atividades de planejamento da expansão do SEB procuram garantir, no médio e longo prazo, o equilíbrio entre a oferta e a demanda de energia com qualidade e confiabilidade. Regras econômico-financeiras permitem a remuneração atrativa dos investimentos, em harmonia com os princípios da modicidade tarifária e do desenvolvimento sustentável.

Sendo a energia hidráulica dependente de fatores climáticos, meteorológicos e ambientais, o planejamento da expansão do setor elétrico em um país como o Brasil, em desenvolvimento demográfico e econômico, torna-se um objetivo desafiador. Após o racionamento ocorrido em 2001

devido à diminuição dos níveis dos reservatórios (ANEEL, 2008), a capacidade do parque gerador brasileiro cresceu rapidamente. Em dez anos, a capacidade de 73.700 MW chegou a quase 120.000 MW em potência instalada em 2011 (Figura 1.2), com um considerável aumento percentual das usinas termoeletricas (Figura 1.3).

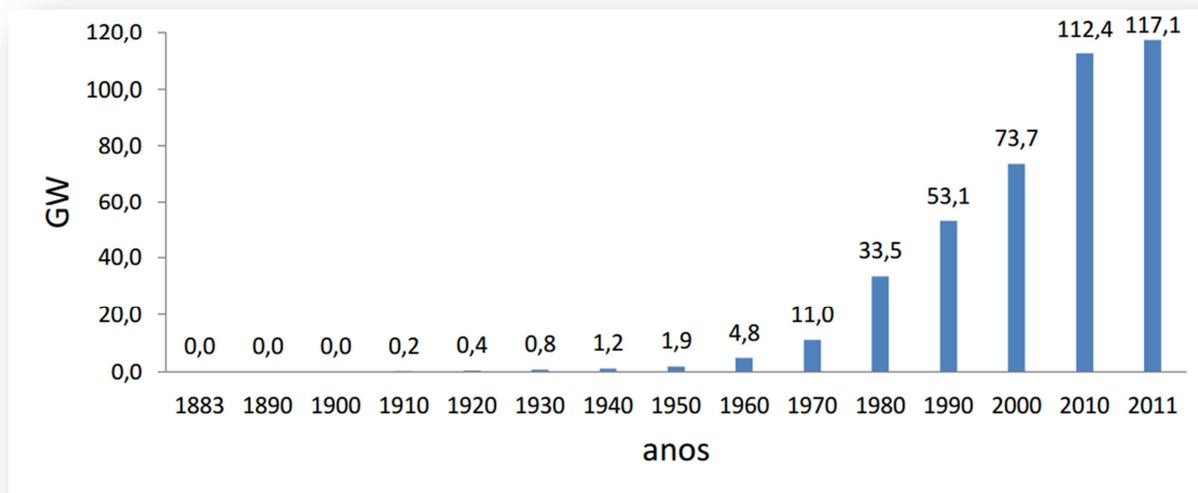


Figura 1.2 Capacidade de potência elétrica instalada – 1883 a 2011 (GW) (MME, 2012 – adaptada).

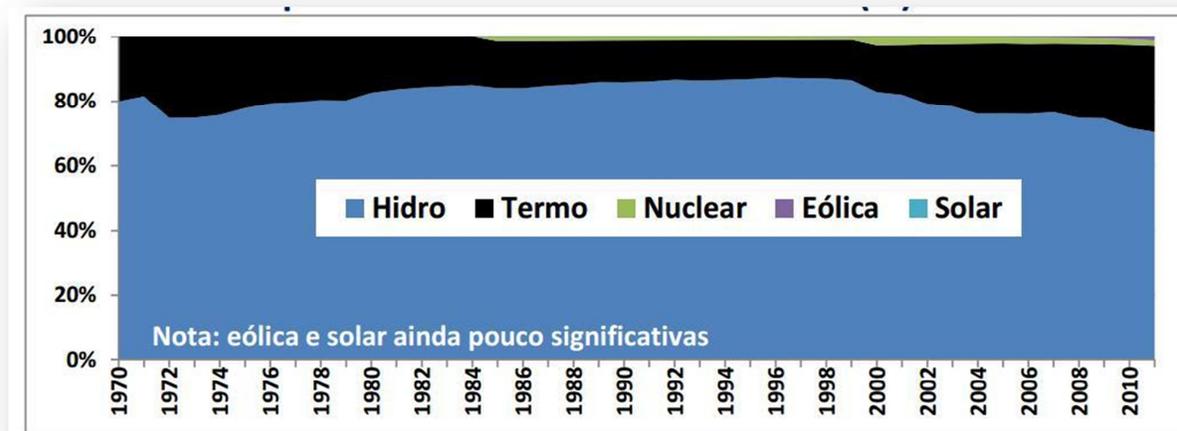
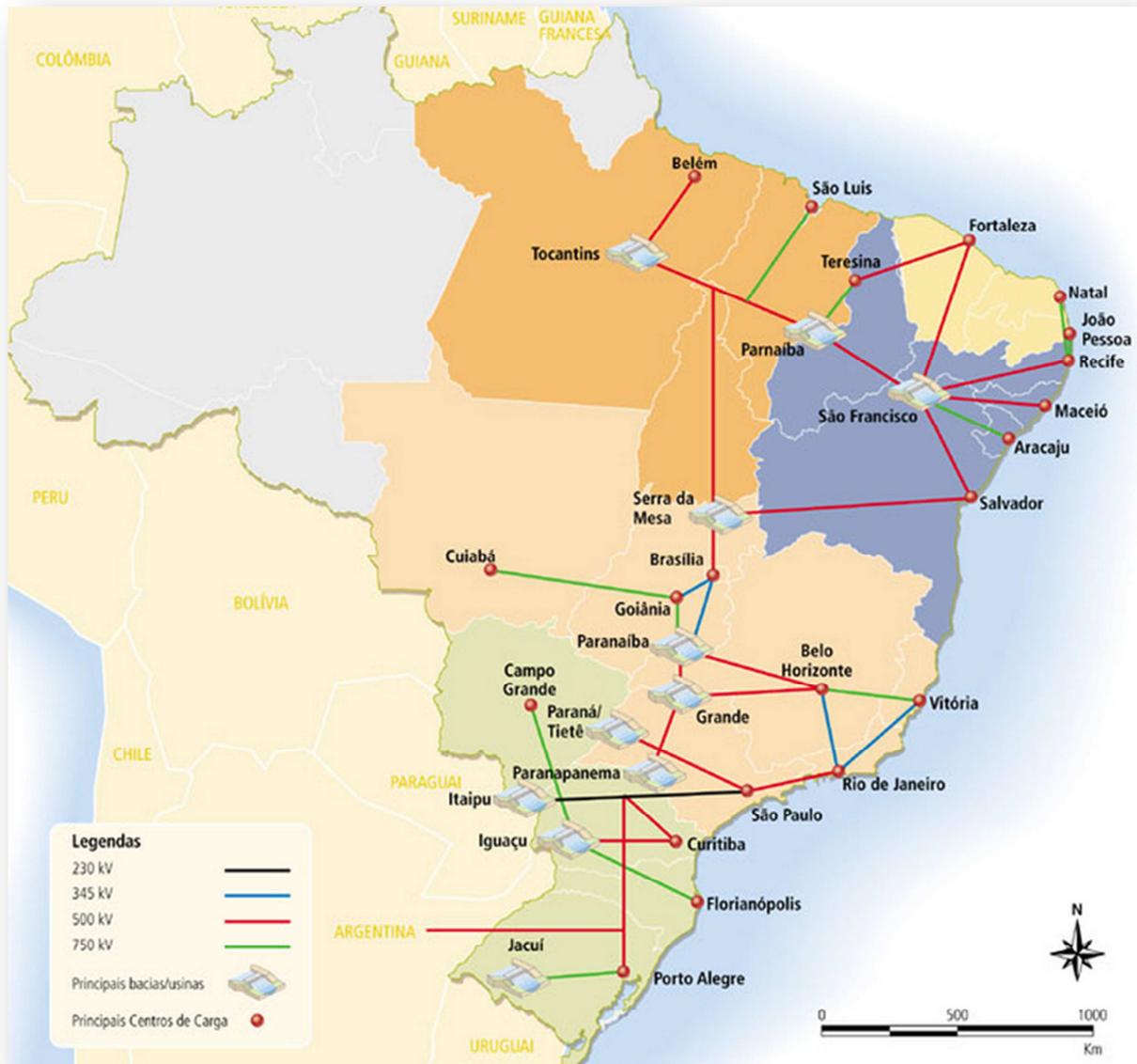


Figura 1.3 Estrutura da capacidade instalada – 1970 a 2011 (%) (MME, 2012).

Além do planejamento do parque gerador, torna-se igualmente desafiadora a programação das operações mensais, semanais e diárias. Tal dificuldade deriva do agrupamento das UHEs em bacias hidrográficas, sujeitas às condições climáticas e às restrições nas transmissões elétricas.

No território nacional, a maior parte das UHEs compõe o chamado Sistema Interligado Nacional (SIN). O SIN está dividido em quatro subsistemas: Sudeste/Centro-Oeste, Sul, Nordeste e Norte, que não cobre a totalidade do território brasileiro, devido à dificuldade do transporte físico da energia na região amazônica. Por isso, tem-se os Sistemas Isolados, que correspondem a 3% do território não atendidos pelo SIN (Figura 1.4).



*Figura 1.4 Representação simplificada da integração entre os sistemas de produção e transmissão de energia elétrica (ONS, 2013).*

Na década de 90, o modelo institucional do setor de energia elétrica passou por um processo de reformas cujo fim foi convergir para a redução da dimensão do Estado empresário. Em 2004, com as leis nº 10.847 e 10.848, estabeleceu-se a retomada da responsabilidade do planejamento pelo

Estado, de forma a reservar ao ente público os papéis de regulador, fiscalizador, mediador e formulador de políticas.

Os objetivos da criação do novo modelo, segundo o Ministério das Minas e Energia (MME), são: assegurar a eficiência na operação e prestação do serviço aos consumidores, garantir a modicidade tarifária e criar um ambiente regulatório estável que estimule a concorrência. O novo modelo mostra-se atrativo ao ingresso de novos investimentos privados no setor e mantém a orientação para as funções de planejamento setorial de longo, médio e curto prazo. A nova estrutura institucional do modelo elétrico brasileiro está orientada para as funções de planejamento setorial, segurança de suprimento, regulação e fiscalização, contando com as seguintes entidades:

- Conselho Nacional de Política Energética (CNPE) — órgão de assessoria da Presidência da República, multi-ministerial, presidido pelo MME, cujo objetivo é a formulação de políticas nacionais e diretrizes de energia, visando o aproveitamento racional dos recursos energéticos do país, a revisão periódica da matriz energética e a definição de diretrizes para programas setoriais específicos.
- MME — ministério cuja função é a formulação e implementação, no âmbito federal, da política energética nacional.
- Empresa de Pesquisa Energética (EPE) — empresa pública federal vinculada ao MME, que tem como missão atuar nos estudos voltados para o planejamento energético nacional, associados às projeções da composição da matriz energética nacional, do balanço energético nacional, do aproveitamento ótimo dos recursos hídricos, do licenciamento ambiental e, por fim, do planejamento da expansão da geração e transmissão de energia elétrica de curto, médio e longo prazos.
- Comitê de Monitoramento do Setor Elétrico (CMSE) — grupo instituído sob coordenação do MME, que tem como finalidade assegurar a continuidade e a segurança do suprimento de energia no País.
- Agência Nacional de Energia Elétrica (ANEEL) — vinculada ao MME, tem por objetivo a fiscalização e regulação das funções de geração, transmissão, comercialização e distribuição de energia elétrica em todo o território nacional.
- Operador Nacional do Sistema Elétrico (ONS) — pessoa jurídica de direito privado, sem fins lucrativos, que, sob regulação e fiscalização da ANEEL, tem como objetivo o controle das funções de geração e transmissão no SIN.
- Câmara de Comercialização de Energia Elétrica (CCEE) — pessoa jurídica de direito privado, sem fins lucrativos, que, sob regulação e fiscalização da ANEEL, tem como

objetivo a gestão dos processos de contratação de compra e venda de energia no novo modelo.

### **1.1.1 Planejamento da Expansão do Setor Elétrico Brasileiro**

O planejamento da expansão do SEB compreende duas etapas: planejamento de longo prazo e planejamento de médio prazo. O planejamento de longo prazo cobre um horizonte não inferior a vinte anos, observando um ciclo de atividades quadrienal, no qual se desenvolve o Plano de Expansão de Longo Prazo do Setor Elétrico (PELP). O PELP estabelece a estratégia de longo prazo para o setor no que concerne às novas fontes de geração, aos grandes troncos de transmissão e às expectativas de desenvolvimento tecnológico e industrial para o país. O desenvolvimento do PELP está sob a responsabilidade da EPE, que compatibiliza as necessidades de crescimento da oferta de energia no país com o desenvolvimento da matriz energética, com vistas à segurança de suprimento e o crescimento sustentado da matriz energética, segundo a natureza de suas fontes.

O planejamento de médio prazo, também sob a responsabilidade da EPE, cobre um horizonte não inferior a dez anos. No ciclo de atividades com periodicidade anual se definem o Plano Decenal de Expansão dos Sistemas Elétricos e o Programa Determinativo de Expansão (EPEa, 2013).

### **1.1.2 Planejamento da Operação do Sistema Hidrotérmico**

A otimização do despacho nacional visa minimizar o custo da operação e garantir o atendimento do mercado com segurança através de decisões sobre geração térmica e hidráulica, intercâmbios entre regiões e cortes de cargas (déficits). O planejamento da operação do sistema hidrotérmico, relacionado com a otimização do despacho, é uma atividade realizada pelo ONS, subdividida em planejamento a médio, curto e curtíssimo prazo.

No planejamento de médio e curto prazo são adotados métodos de otimização, os quais tratam as afluições aos aproveitamentos hidrelétricos através de cenários hidrológicos. O horizonte de estudo varia de um mês até cinco anos (GUILHON *et al.*, 2007).

O horizonte do planejamento a médio prazo, situado na faixa de dois a cinco anos, com discretização mensal, tem como objetivo definir as metas de geração de cada tipo de aproveitamento energético e minimizar o custo das operações. NEWAVE é o modelo computacional utilizado nessa etapa. Este modelo é do tipo periódico de ordem  $p$ ; isso significa que as afluições futuras dependem das afluições que ocorreram nos mesmos locais em até  $p$  meses

anteriores. Considera-se um histórico do regime hidrológico desde 1931. O modelo utiliza os métodos de Programação Dinâmica Estocástica (PDE) para construir as estratégias ótimas dos reservatórios.

No planejamento de curto prazo, cujo modelo computacional é o DECOMP, determina-se a política de geração de cada usina, sendo o período considerado anual, com uma discretização semanal, visando minimizar os custos operacionais do sistema. Partindo da previsão de vazões semanais para o mês em curso e dos cenários de afluições para o segundo mês, calcula-se o Custo Marginal de Operação (CMO), visando à minimização do custo imediato e futuro. Os resultados dessa modelagem determinam a política de geração para cada aproveitamento hidroelétrico.

No planejamento de curtíssimo prazo, o horizonte de operação é semanal, com discretização diária. Neste horizonte, o objetivo é determinar um cronograma de geração que atenda às metas estabelecidas pelo planejamento de curto prazo, em consonância com as restrições elétricas. O planejamento de curtíssimo prazo consegue, assim, suportar decisões de despacho de geração que buscam minimizar o custo de operação e garantir o atendimento do mercado com segurança. O modelo computacional utilizado para esta etapa é o DESSEM-PAT.

## **1.2 Características do Planejamento da Operação do Sistema Hidrotérmico**

O planejamento da operação do sistema hidrotérmico brasileiro é uma tarefa desafiadora devido às características do sistema. A predominância da hidroeletricidade no parque gerador de energia elétrica do Brasil dificulta esta tarefa. O elevado número de aproveitamentos com reservatórios de acumulação e as características de regularização plurianual do sistema brasileiro caracterizam o problema como de grande porte.

A interdependência operativa entre usinas pertencentes a uma mesma bacia limita a operação do sistema. Outras significativas restrições de usos múltiplos da água são: abastecimento urbano e industrial, navegação, irrigação, recreação, controle de cheias e pesca.

Como os recursos hídricos, representados pela água armazenada nos reservatórios e pelas afluições às usinas, são limitados, deve haver um compromisso entre o presente e o futuro. Essa característica torna o problema dinâmico, ou seja, o estado de armazenamento dos reservatórios no presente depende das decisões anteriores, e as decisões atuais influenciam a operação futura. Ademais, as conexões regionais de energia elétrica aumentam a complexidade do sistema hidrotérmico, pois as decisões de geração tornam-se sujeitas às restrições de intercâmbio.

O problema é também estocástico devido às características das vazões que dependem de chuvas e climas muito variados para sua formação. Isto leva a uma considerável incerteza na

informação hidrológica disponível. A não-linearidade das funções de produção hidrelétrica e de custo da complementação térmica aumentam a complexidade do problema.

O planejamento da operação energética do sistema hidrotérmico brasileiro é, portanto, um problema de otimização dinâmico, de grande porte, interconectado, com restrições de usos múltiplos da água, de intercâmbio, estocástico e não linear (HIDALGO, 2009). A previsão de vazão natural afluente aos aproveitamentos hidroelétricos apresenta-se como insumo fundamental para programação da operação do SIN.

### 1.3 Definições em Modelagem Hidrológica

Esta seção tem por objetivo definir os tipos de vazão afluente básicos para o desenvolvimento desta pesquisa.

- Vazão afluente – é a vazão que chega a um determinado ponto, em particular, um aproveitamento hidrelétrico. É determinante na avaliação da energia disponível. A vazão afluente depende das condições hidrológicas naturais da bacia hidrográfica e dos aproveitamentos existentes a montante.
- Vazão natural afluente – é a vazão que existia naturalmente antes de qualquer intervenção humana. Em rios com barragens, esta variável pode ser obtida por meio de cálculos.
- Vazão incremental – é a vazão proveniente da bacia intermediária entre dois aproveitamentos em cascata. Esta vazão pode ser calculada pela diferença entre a vazão afluente ao aproveitamento em referência e a vazão defluente do aproveitamento de montante, propagada ou estimada com base em informações de postos fluviométricos (DE OLIVEIRA *et al.*, 2012).

## 1.4 Importância da Previsão de Vazões

A geração de energia elétrica no Brasil e a programação da operação do sistema hidrotérmico têm grande dependência das vazões naturais e incrementais que afluem aos aproveitamentos hidrelétricos distribuídos por doze bacias. Considerando a importância estratégica da energia e a magnitude deste mercado, as condições de afluência influenciam diretamente os riscos de racionamento e o preço da energia. Prever eventos que podem trazer efeitos negativos para o país inteiro e tomar as medidas necessárias podem evitar situações extremas e danos para a economia e segurança nacional.

A qualidade da previsão de vazão para o ONS é significativa, pois é o órgão responsável por elaborar a previsão e a geração de cenários de vazões naturais médias diárias, semanais e mensais para todos os locais de aproveitamentos hidroelétricos. O uso de previsões hidrológicas de afluências naturais melhora o desempenho da operação do sistema, aumentando os benefícios e reduzindo o grau de incerteza devido ao regime aleatório das afluências hidrológicas futuras, o que mitiga riscos e custos.

Além disso, a previsão de vazões serve como insumo básico para o cálculo do CMO e para a definição dos montantes de geração de todos os aproveitamentos do SIN. Assim, o custo marginal, condicionado pela previsão de vazões, determina os preços da energia no curto prazo, influenciando diretamente no faturamento das empresas de geração (LANZETTI, 2008).

Enfim, uma boa previsão de vazões contribui para uma forma mais integrada de planejamento, programação, operação de tempo real e manutenção. Isso propicia elevados ganhos de qualidade, precisão e tempo (LAUDANNA *et al.*, 2007).

## 1.5 Previsão de Vazões no Setor Elétrico Brasileiro

As previsões diárias de vazão natural são feitas em sua grande parte pelo diferentes agentes de geração. O ONS consolida as previsões médias diárias enviadas pelo diferentes agentes de geração ou processadas pelo próprio órgão. A consolidação destas previsões visa o fechamento de vazão natural média da semana em curso, utilizada como insumo básico para a obtenção das previsões de vazões para as semanas seguintes. Uma vez consolidados os dados diários, o ONS dá início ao processamento dos modelos de previsão de vazões semanais para obter as previsões semanais até o fechamento do mês em curso para todos os postos base.

As previsões de vazões naturais no ONS são elaboradas de forma direta para os locais denominados de postos base, ou de forma indireta, para um posto não base. As equações de

regressão para o cálculo da vazão prevista dos postos base são definidas a partir do estudo da série histórica de vazões de cada aproveitamento e das relações estatísticas entre postos de uma determinada bacia. Para os postos não base, as previsões de vazões naturais totais são obtidas por regressões lineares, referentes a previsões realizadas para um posto base próximo.

Os modelos desenvolvidos para prever vazões podem utilizar dados de vazões (observadas) ou precipitações (observadas e previstas). No entanto, as previsões calculadas pelo ONS utilizam modelos da família PREVIVAZ, que não considera informações de previsão de precipitação (ONS, 2013). O modelo PREVIVAZ e outros modelos desenvolvidos atualmente no Brasil são descritos no próximo capítulo.

## **1.6 Objetivos e Resultados Esperados**

O objetivo principal desta pesquisa é abordar os aspectos que determinam a qualidade da previsão de vazões através da avaliação de dois diferentes modelos matemáticos desenvolvidos para prever vazões afluentes à UHEs brasileiras. Os modelos serão aplicados às UHEs localizadas nos rios Grande, Tietê, Pardo e Mogi Guaçu que pertencem à AES Tietê. Esta empresa é responsável por 20% da produção de energia no estado de São Paulo e 2% da produção nacional.

A avaliação será feita por meio de uma ferramenta, chamada SISPREV, que agrega um conjunto de modelos que constituem uma alternativa aos previsores existentes. Esta ferramenta realiza e gerencia estudos de previsão de vazões diárias e mensais para aproveitamentos hidrelétricos que pertencem ao SIN.

A ferramenta SISPREV é composta por modelos baseados em Regressão Linear (RL) e Sistema de Inferência Neuro Fuzzy (SINF). Embora o principal uso da ferramenta seja prever vazões afluentes, ela será empregada como instrumento de análise de desempenho dos modelos. Assim os resultados dos modelos serão comparados com valores de vazões reais calculadas.

Espera-se que a análise e a comparação entre os modelos matemáticos desenvolvidos para prever vazões contribua com a busca pela qualidade de uma variável importante para os modelos computacionais do SEB, a vazão afluente. Em particular, será objeto de estudo a avaliação das previsões em condições típicas e atípicas, como os períodos particularmente secos ou chuvosos. Outra avaliação que será realizada trata da análise do modelo com menor desvio das vazões previstas em relação às vazões observadas durante um determinado período.

## 1.7 Estruturação do Trabalho

Esta dissertação está dividida em oito capítulos. Neste primeiro capítulo, foi feita a introdução. No segundo capítulo, apresenta-se a revisão da literatura necessária para o desenvolvimento desta pesquisa. No terceiro capítulo, detalha-se o referencial teórico no qual baseia-se este trabalho, com destaque para as técnicas de RL e SINF. No quarto capítulo, é descrita a ferramenta utilizada na avaliação dos modelos matemáticos de previsão de vazões, o SISPREV. No quinto capítulo, mostra-se a metodologia Cross Standard Process for Data Mining (CRISP-DM), que representa um modelo de processo de mineração de dados. Para avaliar a metodologia proposta e analisar a qualidade das previsões, o sexto capítulo apresenta os estudos diário e mensal. No sétimo capítulo, são pontuadas as conclusões, observações e as sugestões de trabalhos futuros associados a esse tema. No oitavo capítulo, apresentam-se as referências bibliográficas e, por fim, os apêndices que detalham os resultados encontrados.

# Capítulo 2

## Revisão da literatura

Este capítulo está dividido em duas seções. Na primeira seção, apresenta-se o Estado da Arte dos modelos de previsão de vazões, subdividido em modelos conceituais, empíricos e combinados. Na segunda seção, mencionam-se alguns trabalhos que comparam diferentes modelos de previsão de vazões.

## 2.1 Estado da Arte

Com relação às metodologias de cálculo usadas, os modelos de previsão de vazões podem ser classificados em conceituais, empíricos ou combinados (TUCCI et al., 2007). Dentre as principais técnicas atualmente utilizadas, pode-se citar: modelos determinísticos, estocásticos, estatísticos e modelos de redes neurais.

### 2.1.1 Modelos Conceituais

Modelos conceituais ou físicos são aqueles cujas funções consideram os processos físicos do sistema hídrico para representar as variáveis. Eles geralmente são baseados na relação entre precipitação e vazão, conhecidos também como modelos chuva-vazão. A precipitação é transformada em vazão através da representação do balanço de água no solo e geração de escoamento; incluindo a interceptação, infiltração, evapotranspiração, escoamento subterrâneo e superficial.

Os modelos conceituais podem ser alimentados com chuva observada, transmitida em tempo real, ou com chuva prevista, para aumentar o horizonte de previsão. Se o modelo de previsão utiliza apenas chuva observada, a previsão de vazão fica limitada ao tempo de concentração na bacia. Tais modelos caracterizam-se por utilizarem como entradas diversas variáveis que possuem relação com a bacia hidrográfica e com o rio nos quais deseja-se realizar as previsões. Estes modelos tentam estabelecer uma relação direta entre as previsões de ocorrência das variáveis externas como chuva e neve (GUILHON, 2002).

Soil Moisture Accounting Procedure (SMAP) é um modelo determinístico conceitual de simulação hidrológica do tipo transformação chuva-vazão (LOPES et al., 1982). Ele é utilizado para os aproveitamentos da bacia do Rio Grande até Porto Colômbia e para todos os aproveitamentos da bacia do Rio Paranapanema. Foi originalmente desenvolvido para o intervalo de tempo diário e posteriormente foram apresentadas as versões horária e mensal, com algumas adaptações em sua estrutura. Os dados de entrada do modelo são os totais diários de chuva e o total diário de evaporação real. As principais vantagens do SMAP estão relacionadas à sua simplicidade e ao uso de pequeno número de parâmetros. Dos seis parâmetros do modelo, três são considerados fixos, com estimativas baseadas em processos físicos. Os três restantes são parâmetros calibráveis com base nas séries históricas de precipitação e vazão. Segundo os autores, para sua calibração, são necessários de 30 a 90 dias de dados de vazão.

A Duke Energy International Geração Paranapanema utilizou o modelo SMAP para a previsão de vazões afluentes às usinas na especificação e construção de um sistema

Previsão/Simulação para suporte à decisão da programação da operação e operação em tempo real do sistema gerador localizado no Rio Paranapanema. Segundo Laudanna *et al.* (2005), o sistema em uso desde janeiro de 2004 agregou um grande valor à cadeia de suporte à decisão da área de operação da Duke Energy. O sistema permitiu que o planejamento, a programação e a operação de tempo real trabalhassem de forma mais integrada, propiciando elevados ganhos de qualidade, precisão e tempo.

Para representar uma bacia, é necessário definir o espaço e o intervalo de tempo de cálculo. Neste sentido, os modelos conceituais podem ainda ser classificados como concentrados ou distribuídos. Os modelos concentrados são simples, pois consideram a bacia como um sistema único (por exemplo, o SMAP). No entanto, possuem limitações importantes quando a área em análise supera 500 km<sup>2</sup>, dependendo das características físicas dos processos e da bacia.

Os modelos distribuídos, por outro lado, subdividem a bacia em diversos elementos, como sub-bacias ou grades regulares. A diferença fundamental entre os modelos distribuídos por sub-bacias e por grades está na estrutura de relação entre as características da bacia e os parâmetros do modelo. Normalmente, os modelos discretizados em grades são utilizados com uma resolução maior e combinados com informações de uso do solo e de outras características físicas. Isso permite uma melhor relação entre parâmetro e característica física (DA SILVA, 2005). Os modelos distribuídos têm sido utilizados para representar bacias cujas variáveis, como a precipitação, características geológicas e o sistema de drenagem, são importantes no espaço e no tempo.

O MGB IPH é um modelo conceitual físico distribuído do tipo transformação chuva-vazão. Este modelo foi desenvolvido em 2001 por Walter Collischonn e Carlos E. M. Tucci, do Instituto de Pesquisas Hidráulicas da Universidade Federal do Rio Grande do Sul (UFRGS) (ONS, 2007). Para a previsão de vazões na bacia incremental à Itaipu, a Rhama Consultora Ambiental Ltda aplicou o modelo hidrológico MGB IPH no qual a bacia hidrográfica é subdividida em células quadradas com 10 km de lado. A topologia da bacia é representada pelo modelo através de uma estrutura que determina, em função da topografia da bacia, de onde vem e para onde vai a água que escoar por uma determinada célula.

Da Silva *et al.* (2005) integraram o modelo MGB IPH com o modelo regional de tempo Eta para a previsão de vazões de curto prazo na bacia do rio São Francisco (cerca de 639.000 km<sup>2</sup>). O modelo Eta, desenvolvido pela Universidade de Belgrado em conjunto com o Instituto de Hidrometeorologia da Iugoslávia, propõe previsões de fenômenos associados a precipitações, fornecidas duas vezes ao dia em uma grade regular com resolução de 40 km.

### 2.1.2 Modelos Empíricos

A formulação dos modelos empíricos não busca a representação dos processos físicos envolvidos. Os modelos empíricos são caracterizados pela rapidez na elaboração e na atualização dos parâmetros em tempo real. Os modelos estocásticos, estatísticos e de redes neurais são empíricos e baseiam-se nas informações provenientes das variáveis que descrevem o estado da bacia.

Entre os modelos empíricos, os modelos estocásticos tratam as afluições aos aproveitamentos hidroelétricos introduzindo o conceito de probabilidade de ocorrência das vazões (CHOW, 1964). Eles descrevem a estrutura probabilística de uma sequência de observações tomando como base as informações sazonais sobre o histórico de afluições para um determinado aproveitamento. Este histórico é dividido em duas partes. Com a primeira parte, estimam-se os parâmetros para os modelos utilizando diferentes métodos. Para a segunda parte do histórico, faz-se uma verificação dos erros associados a cada um dos métodos e seleciona-se aquele de menor erro médio quadrático.

Os modelos estocásticos podem ser subdivididos em modelos estacionários e modelos periódicos. No primeiro grupo pode-se citar o modelo ARMA (p,q). Já entre os modelos periódicos, cita-se o modelo PAR(p) (MAZER, 2003). Os modelos estocásticos também são subdivididos em modelos lineares e não lineares (ANDRIOLO, 2006).

Na maioria das bacias, as previsões de vazões naturais no ONS são obtidas utilizando modelos da família PREVIVAZ. Este é um modelo estocástico linear de previsão de vazões em base semanal, que utiliza uma combinação de 94 possíveis alternativas entre estruturas de correlação periódica ou estacionária, métodos de estimação de parâmetros e tipos de transformações. Este modelo não considera informações de previsão de precipitação, nem informações de chuva ou vazões observadas em estações existentes na bacia. O horizonte de previsão é de até seis semanas para ser utilizada no primeiro mês do planejamento de curto prazo. Este modelo é, portanto, executado ao final de cada mês para elaboração do Programa Mensal da Operação (PMO) do mês seguinte e, durante o mês em curso, é executado todas as semanas para a realização das revisões do PMO. Adota-se uma abordagem bastante utilizada para obtenção de previsões de vazões, que consiste no uso do valor esperado de modelos de séries temporais.

A previsão também considera o fenômeno denominado tendência hidrológica. Tendência hidrológica é o termo utilizado para se referir à tendência observada das séries históricas de vazões fluviais naturais. Por exemplo, se em uma determinada semana a vazão do rio foi superior ao valor médio para aquela semana, é provável que, na semana seguinte, a vazão do rio deverá continuar acima da média. Da mesma forma, se, na semana analisada, a vazão foi inferior à média da semana, na semana seguinte, a vazão do rio deverá continuar abaixo da média. Na terminologia de séries

temporais, a tendência hidrológica é conhecida como estrutura de dependência temporal, sendo quantificada pela função de autocorrelação estimada do registro de vazões (FAYAL, 2008).

O Modelo Estocástico Linear (MEL) trata da previsão diária de vazões empregando a conceituação dos chamados modelos função de transferência. Nesse tipo de modelo são utilizadas tanto informações de vazões do passado (parâmetros autoregressivos), quanto observações de outros fenômenos (por exemplo, vazões e precipitações acumuladas à montante do ponto de interesse). Em geral, a previsão é feita utilizando-se o conceito do valor esperado, podendo-se também empregar a estimativa feita por intervalo de confiança. O MEL desenvolvido para a previsão de vazões naturais na bacia incremental de Itaipu é do tipo misto, utilizando tanto dados de vazão observada, quanto de chuvas observadas e previstas para sua área a montante. Os chamados modelos mistos ou de função de transferência empregam técnicas de Análise de Séries Temporais (ONS, 2007).

A metodologia Diferenças Primeiras da Função de Transferência (DPFT) usa uma amostra de observações de eventos precipitação-vazão e um algoritmo iterativo para a identificação simultânea do hidrograma unitário (Função de Transferência) médio da bacia e de uma série de precipitações efetivas associadas a cada evento. Esta última particularidade permite a calibração e comparação de diferentes modelos não-lineares de transformação de precipitações totais em precipitações efetivas, ou Funções de Produção (ONS, 2007).

A RL é um método empírico estatístico para estimar o valor esperado da variável de vazão a partir de variáveis de entrada conhecidas previamente, tais como vazões e precipitações observadas e precipitações previstas. A aplicação da metodologia estatística é feita utilizando registros históricos, onde as variáveis independentes são utilizadas para um período de calibração. Em seguida, destina-se um período para validação dos parâmetros de regressão obtidos, os quais serão utilizados na fase de previsão. A eficácia deste método depende, em grande parte, da estabilidade das características principais do regime do curso d'água, ou seja, quando da utilização destes dados a bacia não deve ter sofrido nenhuma modificação hidrológica importante (desvio, construção de barragem, urbanização das margens, etc.).

Sousa Filho e Lall (2004) apresentaram resultados satisfatórios utilizando um modelo de RL para previsão de vazões naturais sazonais para um sistema de reservatórios no Estado do Ceará. Este método gerou valores observados de coeficiente de determinação  $R^2$  de 0,93.

Cervilha e Da Silva Filho (2010) chegaram a um erro médio menor que 10% utilizando um modelo de RL para a previsão de vazões no Rio Pardo, Rio Piracicaba, Rio Paraíba do Sul e Rio Jacaré. Consideraram este erro irrelevante por causa do alto grau de dispersão dos dados analisados.

Os modelos estatísticos de regressão dinâmica são uma categoria especial dos modelos de funções de transferência que combinam a dinâmica de séries temporais e o efeito de variáveis explicativas. O termo regressão dinâmica não indica que os parâmetros do modelo evoluem no

tempo, mas sim que a estrutura de dependência da série temporal também está incluída na regressão. Nos modelos de regressão dinâmica, a variável dependente (vazão prevista no n-ésimo dia, por exemplo) é explicada por seus valores defasados (vazões previstas nos dias n-1, n-2, etc.) e pelos valores atuais e passados de variáveis causais (precipitações na bacia e vazões a montante) ou exógenas (vazões em bacias vizinhas). O consórcio PUC-Rio desenvolveu o modelo estatístico de regressão dinâmica para a bacia incremental Porto Primavera – Rosana – Itaipu (ONS, 2007).

A Rede Neural Artificial (RNA) é uma técnica empírica inspirada no funcionamento do cérebro humano e, como tal, é composta por unidades básicas de processamento, os neurônios artificiais. As RNAs são construídas a partir de vários elementos simples de processamento que são geralmente implementados através de simulação em computadores digitais. Suas recentes aplicações na hidrologia têm-se mostrado particularmente efetiva na transformação chuva-vazão e na previsão de vazões.

No modelo MultiLayer Perceptron (MLP) várias unidades de processamento estão conectadas a uma única unidade de saída, através dos pesos sinápticos e da rede Perceptron. Os neurônios são conectados por arcos e a cada arco está associado um peso. A comunicação entre os neurônios é realizada através destas interconexões ponderadas (CRUZ *et al.*, 2010).

As redes neurais construtivas Non-linear Sigmoidal Regression Blocks Networks (NSRBN) utilizam blocos de sigmóides não-lineares como função de ativação. As redes neurais construtivas, ao contrário das redes de topologia fixa, constroem sua própria estrutura, incluindo unidades e/ou camadas intermediárias. Este fato é contrário às redes neurais de alta ordem que tomam como base polinômios multivariados e apresentam uma explosão no número de parâmetros livres necessários para aproximar uma dada função (GOMES *et al.*, 2010).

O Sistema de Inferência Fuzzy (SIF) é uma técnica de previsão de vazões diárias e mensais baseada em regras nebulosas de produção que fazem uso das variáveis linguísticas para executar um processo de tomada de decisão. A base de conhecimento destes sistemas consiste em um conjunto de regras de produção Fuzzy, que tem seu funcionamento baseado em três etapas: (1) *fuzzificação*, (2) procedimentos de inferência e (3) *defuzzificação* (RABELO *et al.*, 2011). Este sistema codifica na sua estrutura regras nebulosas adotando a forma "Se A, então B", onde A representa o antecedente, e B representa o conseqüente da regra nebulosa definido por um modelo linear das entradas do próprio modelo. Cada regra possui uma área de atuação determinada pelas funções de pertinência associadas. Finalmente, a previsão estimada é calculada pela ponderação entre as saídas locais (conseqüentes) e os respectivos graus de pertinência. Na literatura relacionada às previsões de vazões, encontram-se modelos híbridos Neuro Fuzzy: combinações de técnicas Fuzzy com RNAs.

O modelo de previsão Fuzzy Recorrente é um modelo híbrido baseado em RNAs. Foi aplicado por Evsukoff *et al.* (2007) para a bacia do Rio Iguaçu. Para a construção das regras do

modelo, foram utilizados cinco conjuntos Fuzzy na partição da variável de entrada (precipitação acumulada) e dois conjuntos Fuzzy para a variável de saída (vazão natural total diária). O modelo utiliza dados de previsão de precipitação de dez dias à frente fornecidos pelo modelo Eta. Os resultados numéricos mostram que o modelo é robusto e apresenta um bom desempenho para todos os aproveitamentos considerados.

Luna *et al.* (2007) apresentaram um sistema adaptativo baseado em regras nebulosas, sendo o algoritmo de aprendizado baseado no algoritmo de maximização da verossimilhança. O modelo foi aplicado à previsão de séries de vazões semanais do posto de Sobradinho, localizado na bacia do rio São Francisco, região nordeste do Brasil. Os resultados obtidos mostram que a proposta é uma alternativa eficiente para modelos de séries temporais, proporcionando uma estrutura simples e parcimoniosa.

### 2.1.3 Modelos Combinados

Os modelos combinados associam as características da metodologia de cálculo conceitual e empírica. Como exemplo, pode ser citado o modelo SMAP-MEL. Este modelo combina os resultados de dois modelos: SMAP (modelo determinístico de simulação hidrológica do tipo transformação chuva-vazão) e MEL (multivariado, que também considera informações de precipitação). Ele foi aplicado com sucesso na bacia piloto do rio Paraná para a previsão das vazões médias diárias e semanais afluentes à UHE Itaipu em um horizonte de curto prazo (doze dias à frente). Para isto, foram utilizados os dados hidrometeorológicos observados na bacia incremental de Itaipu e as previsões de precipitação acumuladas de 24 horas fornecidas pelo modelo meteorológico regional Eta (FADIGA Jr *et al.*, 2008).

## 2.2 Revisão Bibliográfica

O foco desta pesquisa é a avaliação dos resultados apresentados por modelos matemáticos desenvolvidos para prever vazões afluentes à UHEs. Portanto, considera-se adequado mencionar nesta seção alguns trabalhos que comparam diferentes modelos de previsão de vazões.

Figueiredo *et al.* (2007) apresentaram modelos de previsão de vazão natural para a Bacia do Rio Iguaçu baseados em previsões estatísticas e de RNAs. Os métodos estatísticos utilizados foram de Análise Multivariada com Incorporação de Informações Hidrometeorológicas, de Regressão Dinâmica e de Redes MLP-LM (Levenberg-Marquardt) MLP-Bayesiana, considerando as seguintes variáveis: vazões naturais e incrementais históricas nos pontos de interesse, vazões diárias observadas em postos fluviométricos, chuva média diária observada na bacia, médias das chuvas

médias diárias observadas na bacia, chuva prevista diária e média da chuva prevista. Os resultados obtidos comprovam a viabilidade da utilização de métodos quantitativos, RNAs e métodos estatísticos neste problema, que é altamente não-linear e muito complexo, permitindo obter previsões com boa precisão. O melhor desempenho foi para o modelo MLP-LM. Os modelos de RNAs (modelos não lineares) apresentaram melhores resultados que os modelos estatísticos em virtude do comportamento complexo da bacia do Iguaçu.

Gomes e Montenegro (2010) apresentaram a análise dos dados pluviométricos e fluviométricos da área à montante do reservatório de Três Marias, no rio São Francisco. Para isto, calibraram dois modelos baseados na técnica de RNAs para previsão de vazões naturais médias diárias, doze dias à frente: MLP e NSRBN. Os resultados obtidos mostraram que o modelo calibrado com a técnica de RNAs teve um desempenho superior ao modelo estocástico PREVIVAZ. Entre as RNAs analisadas, a NSRBN teve um desempenho um pouco superior a MLP. Também observou-se a importância das informações de precipitação, principalmente nas simulações do período úmido.

Cruz *et al.* (2010) apresentaram modelos de previsão de vazão natural para as Bacias no Estado de Minas Gerais. A metodologia DPFT é comparada com um modelo simples de RNAs, onde as não-linearidades inerentes ao processo de transformação precipitação-vazão são representadas por uma estrutura de rede MLP com quatro neurônios representativos das vazões e precipitações. Nos estudos de caso analisados, a metodologia DPFT apresentou um desempenho ligeiramente superior ao modelo de RNAs, mostrando, entretanto, menor regularidade nos resultados.

Luna *et al.* (2009) apresentaram uma comparação entre o SIF com abordagem bottom-up e o modelo SMAP, para previsão de vazões naturais médias diárias, aplicados para a bacia Porto Primavera, no Paraná. Estudos de simulação indicaram que o SIF apresenta um desempenho adequado, levando a uma promissora alternativa para a previsão de vazões diárias.

Guilhon *et al.* (2007) trataram do desenvolvimento de concorrência entre diferentes alternativas de modelos de previsão de vazões médias diárias, com horizonte de até doze dias à frente, utilizando informações de precipitação para aplicação em aproveitamentos hidroelétricos do SIN. Os aproveitamentos hidroelétricos considerados neste trabalho estão localizados nas bacias dos rios Iguaçu, Paraná e Paranaíba. Foram empregados diversos métodos de previsão, abrangendo modelagem física com modelos dos tipos conceitual, concentrado e distribuído, bem como metodologias híbridas. Combinaram-se tanto técnicas de inteligência artificial e modelos estocásticos diversos, quanto técnicas de inteligência artificial e mineração de dados. Foram escolhidos seis modelos:

- MLP — Modelo desenvolvido pela PUC-Rio que utiliza RNAs. Ele foi aplicado para a bacia do rio Iguaçu.
- Modelo de Regressão Dinâmica — Modelo que utiliza técnicas Estocásticas de Regressão Dinâmica, desenvolvido pela PUC-Rio, para as bacias dos rios Paranaíba e Paraná.
- Modelo Fuzzy Recorrente — Modelo híbrido utilizando Técnicas de Mineração de Dados, com emprego de Cubo OLAP, e Modelos de Inteligência Artificial, baseados em algoritmos Fuzzy Recorrentes, desenvolvido pela UFF em conjunto com a UFRJ.
- Modelo SMAP — Modelo físico concentrado, aplicado pelo LACTEC/PR.
- Modelo SMAP-MEL — Solução com modelo híbrido, que utiliza uma combinação do modelo físico concentrado SMAP e do MEL, desenvolvido pela Fundação Centro Tecnológico de Hidráulica da Universidade de São Paulo.
- Modelo MGB-IPH — Modelo Físico Distribuído complementado com Modelo Estocástico de Correção da Previsão, desenvolvido pela Rhama Consultoria Ambiental Ltda.

Os modelos selecionados tiveram um desempenho superior ao PREVIVAZ para os índices avaliados. Os resultados do modelo MGB-IPH e Fuzzy Recorrente foram superiores aos demais modelos em todos os índices. Recomendaram investir no desenvolvimento da qualidade da previsão de precipitação obtida com o modelo regional Eta, seja por modificação em sua parametrização, seja por alteração em sua escala, pois poderão trazer ainda mais benefícios aos resultados dos modelos chuva-vazão desenvolvidos.

# Capítulo 3

## Referencial Teórico

Neste capítulo serão apresentados os métodos da Engenharia do Conhecimento. Inicialmente, serão descritas as principais técnicas existentes, tais como: Sistemas Simbólicos, Sistemas Difusos, Algoritmos Genéticos, Métodos Estatísticos, Redes Neurais e Sistemas Híbridos. Em seguida, será analisado o referencial teórico que baseia este trabalho, com destaque para as técnicas de RL e SINF.

### 3.1 Métodos de Engenharia do Conhecimento (EC)

A EC é uma área de pesquisa que codifica o conhecimento de um especialista humano através do desenvolvimento de técnicas computacionais inspiradas em características tais como percepção, raciocínio e aprendizado. De acordo com a disponibilidade de dados e de teoria, diferentes métodos são indicados para cada situação particular. A Figura 3.1 apresenta os principais métodos: estatísticos, Algoritmos Genéticos (AG), sistemas baseados em RNAs, Sistemas Difusos (SD), Sistemas Híbridos (SH) e os Métodos de Inteligência Artificial Simbólicos (MIAS).

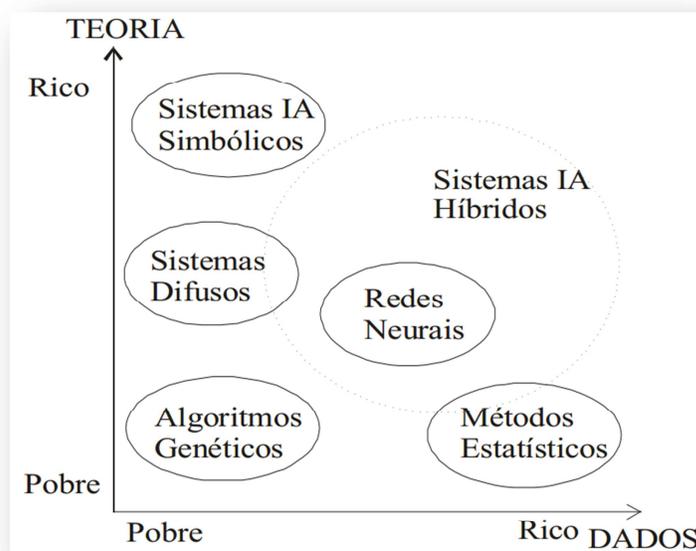


Figura 3.1 Métodos de Engenharia do Conhecimento (BAUCHSPIESS, 2004).

Quando existem apenas amostras representativas de um dado processo, sem regras que expliquem a sua geração, os métodos estatísticos permitem obter os melhores resultados. Este trabalho utiliza um modelo que se baseia no método estatístico de RL. Métodos de RL são geralmente utilizados para prever a demanda de um determinado item com base na previsão de outra variável que tenha relação com tal item. Segundo Figueirado *et al.* (2011), a análise de RL é o modelo estatístico mais usualmente empregado na ciência política contemporânea.

No extremo oposto do gráfico estão os MIAS. Neste caso, o que importa são as regras e o processo de inferência que permite resolver um determinado problema. Tal teoria lógica, baseada

em axiomas, prescinde de exemplos para a implementação do sistema especialista (LADEIRA, 1997).

Os AG não garantem que um conjunto de amostras rico em informações sobre o processo leve à solução ótima do problema. Estes métodos tiveram origem na esperança de que os processos de transmissão de material genético entre gerações de populações, levando eventualmente a indivíduos mais aptos, pudesse ser mimetizado em um programa de computador. Dado um conjunto inicial de possíveis soluções sub-ótimas, estas podem ser combinadas (cruzamento de material genético) sucessivamente (gerações) até obter-se a solução ótima do problema.

Os sistemas baseados em RNAs ocupam uma região intermediária em relação ao gráfico. De fato, uma das grandes vantagens desta abordagem é a possibilidade de treinamento das RNAs a partir dos dados. Não são necessárias regras ou uma teoria que descreva o processo, já que as RNAs simplesmente aprendem e adaptam-se a cada exemplo dado. O comportamento desejado é reforçado e o comportamento indesejado é reprimido até que o sistema realize a tarefa almejada. Aplicações de RNAs são inúmeras. As mais comuns são: controle de processos, robótica, sistema de suporte à decisão, classificação de dados, reconhecimento de padrões em linhas de montagem, filtros contra ruídos eletrônicos, análise de imagens, análise de voz, análise de aroma e odor.

Os SD, fundamentados pela lógica nebulosa ou Fuzzy, têm uma boa base teórica. Podem ser construídos a partir de regras formuladas por especialistas da aplicação. Amostras do processo não são necessárias. Estes sistemas utilizam uma lógica multi-valorada que permitem graus de pertinência e não verdades ou falsidades absolutas. Sistemas baseados na lógica nebulosa têm mostrado grande utilidade em uma variedade de operações de controle industrial e em tarefas de reconhecimento de padrões que se estendem desde o reconhecimento de texto manuscrito, até a avaliação de crédito financeiro. Existe também uma ampla variedade de outras áreas em que a lógica nebulosa pode desempenhar um papel diferente e inovador. Fornos de micro-ondas que medem a temperatura, umidade e forma dos alimentos para controlar o tempo de cozimento, aspiradores de pó que medem a quantidade de pó para ajustar a potência de sucção, sistemas administrativos e econômicos que usam centenas de regras para negociar ações, são algumas das aplicações mais recentes (CAMARGOS, 2002).

Este trabalho abordará uma das possibilidades mais interessantes em sistemas inteligentes, utilizando um modelo que, baseando-se nos SH, combina as características de RNAs com a Lógica Fuzzy, ou seja, sistemas Neuro-Fuzzy. Estes sistemas permitem criar um conjunto de regras que são treinadas com os dados da aplicação. Tem-se, assim, a componente de explicação dos sistemas especialistas combinada com a característica de aprendizagem dos sistemas neurais (BAUCHSPIESS, 2004). Segundo Teixeira *et al.* (2007), a união da rede Neural com a lógica

Fuzzy vem com o intuito de amenizar a deficiência de cada um destes sistemas fazendo com que tenhamos um sistema mais eficiente, robusto e de fácil entendimento.

### 3.2 Regressão Linear

A teoria da regressão consiste numa análise estatística com o objetivo de verificar a existência de uma relação funcional entre uma variável dependente e uma ou mais variáveis independentes ou explanatórias ou covariáveis ou regressoras. A variável  $x$  é a variável independente da equação, enquanto  $y = f(x)$  é a variável dependente das variações de  $x$ .

O modelo de regressão é chamado simples quando envolve uma relação causal entre duas variáveis. O modelo de regressão é múltiplo quando envolve uma relação causal com mais de duas variáveis, isto é, quando o comportamento de  $y$  é explicado por mais de uma variável independente  $x_1, x_2, \dots, x_k$ . O relacionamento entre variáveis pode ser do tipo linear (equação da reta ou do plano) ou não linear (equação exponencial, geométrica, etc.).

A teoria da regressão é utilizada para realizar previsões sobre o comportamento futuro de algum fenômeno da realidade descrito por meio dessas variáveis. Por exemplo, o volume do comércio internacional, onde a variável dependente é  $y$ , pode ser modelado como uma função linear do produto interno bruto dos países, cuja variável independente é  $x$ . Neste caso, são extrapoladas para o futuro as relações de causa-efeito já observadas no passado.

O conjunto finito de observações  $x$  e  $y$ , representável num diagrama de dispersão (Figura 3.2), corresponde a uma amostra do universo de informações ou população, a qual permitiria estabelecer a relação entre  $x$  e  $y$ .



Figura 3.2 Diagrama de dispersão.

A relação funcional é feita por meio de uma equação que explica a variação da variável dependente pela variação dos níveis das variáveis independentes. Verifica-se que os pontos do diagrama de dispersão não vão se ajustar perfeitamente à curva do modelo matemático proposto. Haverá, na maior parte dos pontos, uma distância entre os pontos do diagrama e a curva do modelo matemático. Isto acontece devido ao fato de o fenômeno que está em estudo não ser um fenômeno matemático e sim um fenômeno que está sujeito a influências estocásticas. Assim, o objetivo da regressão é obter um modelo matemático que melhor se ajuste aos valores observados de  $y$  em função da variação dos níveis da variável  $x$ .

### 3.2.1 Regressão Linear Simples (RLS)

O modelo de RLS consiste em uma variável dependente estar relacionada a uma variável independente por uma equação linear. A linha de regressão minimiza os desvios quadrados dos dados reais, ou seja, a soma dos quadrados dos erros de previsão é a mínima possível. Na Figura 3.3, observa-se a formação da reta no sistema cartesiano.

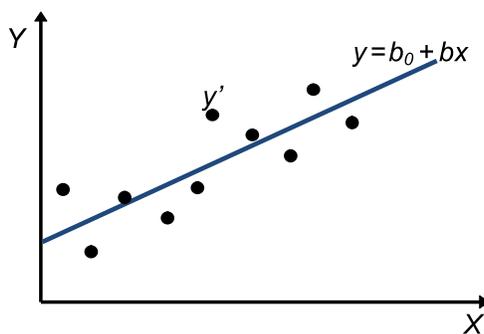


Figura 3.3 *Reta de RLS.*

A reta apresenta a seguinte equação:

$$y = b_0 + bx \quad (3.1)$$

A Equação 3.1 identifica o efeito da variável de previsão independente  $x$  sobre a demanda do produto em análise (variável dependente  $y$ ). O método utilizado para obter a relação funcional é denominado Método dos Mínimos Quadrados (MMQ). Esse método define uma reta que minimiza a soma de quadrados das distâncias entre os pontos  $y'$  do diagrama e os respectivos pontos  $y$  na reta.

A Equação 3.2 mostra assim a condição de trabalho deste método:

$$\min \sum (y - y')^2 \quad (3.2)$$

Para que a soma dos quadrados dos erros tenha um valor mínimo, devem ser aplicados os conceitos de cálculo diferencial com derivadas parciais. Como as incógnitas do problema são os coeficientes  $b_0$  e  $b$ , estrutura-se um sistema de duas equações.

Para encontrar os valores de  $b_0$  e  $b$ , segue a fórmula conforme descreve Gaither e Frazier (2006):

$$b = \frac{n(\sum x'y') - (\sum x')(\sum y')}{n(\sum x'^2) - (\sum x')^2} \quad (3.3)$$

$$b_0 = \frac{\sum y' - b(\sum x')}{n} \quad (3.4)$$

Os valores  $b_0$  e  $b$  acima correspondem aos parâmetros da Equação de regressão que minimiza as diferenças entre os valores de  $y$  (estimados pela regressão) e os de  $y'$  (levantados), sendo  $n$  o número de pontos no diagrama de dispersão e  $x'$  as abscissas dos pontos do diagrama de dispersão. Portanto, o problema de ajustar uma reta que melhor se adequa à nuvem de dados se reduz em calcular os parâmetros  $b_0$  e  $b$  da Equação de regressão.

### 3.2.3 Regressão Linear Múltipla (RLM)

A RLM envolve uma única variável dependente ( $y$ ) e duas ou mais variáveis independentes ( $x_k$ ,  $k = 1, 2, \dots$ ). A teoria é uma extensão da análise de RLS. De modo similar, a análise tem por objetivo estabelecer uma equação que possa ser usada para prever valores de  $y$  para valores dados das diversas variáveis independentes. A finalidade das variáveis independentes adicionais é melhorar a capacidade de predição em confronto com a RLS. A técnica de cálculo pode ser facilitada com o auxílio de álgebra de matrizes. A Equação 3.5 apresenta a forma geral de uma regressão múltipla.

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k \quad (3.5)$$

onde:

$b_k$  = coeficientes técnicos;

### 3.2.3 Medidas de Ajustamento

Para que se possa avaliar a qualidade da relação linear ajustada às informações amostrais das variáveis, duas medidas podem indicar o quão aproximada é a reta: Erro Padrão da Estimação  $S_e$  e Coeficiente de Determinação  $R^2$ .

#### Erro Padrão de Estimação

O Erro Padrão de Estimação mede o desvio médio entre os valores reais de  $y_n$  e os valores estimados  $y'_n$  para o  $n$ -ésimo caso. Ele informa de modo aproximado a extensão do erro entre os valores obtidos das estimativas e os valores de  $y$  fornecidos pela amostra. O que se busca é conseguir o menor valor possível de  $S_e$ . A fórmula de  $S_e$ , conforme Equação 3.6, é obtida da definição da variância da amostra  $S_e^2$ , com  $n-k$  graus de liberdade, sendo  $n$  o número de observações e  $k$  o número das variáveis explicativas, perde-se  $k+1$  graus de liberdade com as estimativas dos coeficientes técnicos da equação da reta mais a constante:

$$S_e = \sqrt{\frac{\sum_n (y_n - y'_n)^2}{[n - (k + 1)]}} \quad (3.6)$$

O Erro Padrão de Estimação avalia a precisão do cálculo da média amostral. Ao se obter uma amostra qualquer de tamanho  $n$ , calcula-se a média aritmética populacional. Provavelmente, se uma nova amostra aleatória for realizada, a média aritmética obtida será diferente daquela da primeira amostra. A variabilidade das médias é estimada pelo seu erro padrão. Observa-se que quanto melhor a precisão no cálculo da média amostral, menor será o erro padrão.

#### Coeficiente de Determinação

O Coeficiente de Determinação, é uma medida de ajustamento em relação aos valores observados. O  $R^2$  varia entre 0 e 1, indicando, em percentagem, o quanto o modelo consegue explicar os valores observados. Quanto maior o  $R^2$ , mais explicativo é o modelo e melhor ele se ajusta à amostra.

O Coeficiente de Determinação  $R^2$  é definido pela seguinte relação:

$$R^2 = 1 - \frac{\sum_n (y_n - y'_n)^2}{\sum_n (y_n - \bar{y})^2} \quad (3.7)$$

onde  $\bar{y}$  é a média da variável dependente.

$R^2$  é uma medida da efetividade geral da regressão linear. Quanto maior ele for, melhor a equação ajustada explica a variação nos dados. O coeficiente de determinação deve ser interpretado como a proporção de variação total da variável dependente  $y$  que é explicada pela variação da variável independente  $x$ . Por exemplo, se o  $R^2$  de um modelo é 0,8534, isto significa que 82,54% da variável dependente consegue ser explicada pelos regressores presentes no modelo.

O coeficiente de determinação é igual ao quadrado do coeficiente de correlação  $R$ . A partir do valor do coeficiente de determinação pode-se obter o valor do coeficiente de correlação. No exemplo acima, para um coeficiente de determinação  $R^2 = 0,8534$  obtém-se o coeficiente de correlação,  $R = 0,9237$ . O coeficiente de determinação é sempre positivo, enquanto que o coeficiente de correlação pode admitir valores negativos e positivos. Valores de  $R$  igual ou próximos de 1 ou  $-1$  indica que exige uma forte relação entre as variáveis. Com valores próximos de zero, o relacionamento entre as variáveis é pouco considerável.

A inclusão de inúmeras variáveis, mesmo que tenham pouco poder explicativo sobre a variável dependente, aumentarão o valor de  $R^2$ . Isto incentiva a inclusão indiscriminada de variáveis. Para combater esta tendência, pode-se usar uma medida alternativa do coeficiente de determinação, que penaliza a inclusão de regressores pouco explicativos. Trata-se do  $\bar{R}^2$  ajustado:

$$\bar{R}^2 = 1 - \frac{n-1}{n-(k+1)}(1 - R^2) \quad (3.8)$$

onde  $(k + 1)$  representa o número de variáveis explicativas mais a constante. A inclusão de mais variáveis com pouco poder explicativo prejudica o valor do  $\bar{R}^2$  ajustado, porque aumenta  $k$  em uma unidade, sem aumentar substancialmente o  $\bar{R}^2$ .

#### 3.2.4 Testes de Significância Estatística

A análise de regressão tem três objetivos gerais: modelar a relação entre a variável dependente  $y$  e uma ou mais variáveis independentes  $x$ , mensurar o erro ao usar a relação que prediz a variável dependente e medir o grau de associação entre a variável dependente e as independentes.

Para que esses objetivos sejam alcançados, uma série de testes estatísticos em relação à significância deve ser analisada. O propósito dos testes de significância estatística é determinar a confiança que pode ser depositada nos resultados da regressão e a sua aplicabilidade na população de valores possíveis. A significância estatística dos resultados obtidos na Análise de Regressão deve ser estabelecida antes do uso de tais resultados numa previsão.

### Teste de Significância dos Coeficientes de Regressão (Teste t)

Determina a significância (correlação) dos coeficientes da equação de regressão ( $b_0, b_1, b_2, \dots, b_k$ ) individualmente, ou seja, verifica se as variáveis  $x_k$  explicam a variabilidade de  $y$ . O questionamento essencial deste teste é se o valor atribuído a cada coeficiente é estatisticamente significativo e diferente de zero. Este teste consiste em calcular a variância de cada coeficiente da regressão e, através de sua raiz quadrada, estabelecer o erro padrão.

### Análise de variância (Teste F)

A análise de variância testa a significância geral da regressão, ou seja, confirma se há relação estatística significativa entre a variável dependente e uma ou mais variáveis explicativas. O Teste F indica se a equação de regressão é significativa – ou seja, se a relação funcional estabelecida entre a variável dependente e os efeitos combinados das variáveis independentes são relevantes. O valor do Teste F é determinado pelo quociente entre a variância explicada e a variância inexplorada. Esta relação pode ser expressa pela seguinte equação:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \quad (3.9)$$

onde  $R^2$  é o coeficiente de determinação,  $k$  é o número de variáveis explanatórias e  $n$  é o número de observações. Se a estatística F encontrada for de 35,29, isto significa que a equação de regressão tem um nível de significância de 35,29%, o que pode ser considerado baixo. O resultado do Teste F não deve ser considerado de forma isolada, uma vez que os coeficientes da regressão podem apresentar correlação significativa.

### Teste de Durbin-Watson (DW)

O teste DW compara se há ou não autocorrelação entre os resíduos, ou seja, se os resíduos do período  $t$  estão correlacionados com os do período  $t-1$ . O teste para detectar correlação serial é calculado da seguinte forma:

$$d = \frac{\sum_{k=2}^n (e_k - e_{k-1})^2}{\sum_{k=1}^n e_k^2} \quad (3.10)$$

onde os resíduos  $e_k$  são a diferença entre o valor observado da variável dependente e o valor predito:

$$e_k = y_k - y'_k \quad (3.11)$$

O intervalo da estatística de DW vai de 0 a 4, com um valor intermediário igual a 2. Assim, quando os erros são aleatórios, o valor DW é próximo de 2; quando há autocorrelação positiva, o valor é próximo de 0 e, quando é próximo de 4, indica autocorrelação negativa.

### 3.2.5 Os Pressupostos da Análise de Regressão

A análise de regressão de mínimos quadrados gera estimativas confiáveis e pode ser adequadamente utilizada se resultam satisfeitos seis pressupostos básicos distintos. Dito de outra forma, as estimações são eficientes desde que os pressupostos subjacentes à análise de regressão sejam devidamente respeitados.

De acordo com De Lurgio (1998) estes pressupostos são:

1. *Linearidade*. A relação projetada entre as variáveis está na forma correta – ou seja, parte-se do princípio de que a relação entre a variável dependente e as independentes é linear. A violação desse pressuposto impede que a estimação por mínimos quadrados produza o melhor estimador linear não-viesado.
2. *Homocedasticidade dos erros*. Se os resíduos não estão distribuídos ao longo da linha de regressão em torno de todo o intervalo de observações, o pressuposto da variância constante, ou homocedasticidade, é violado. A Figura 3.4, a seguir, ilustra o significado da variância constante dos resíduos:

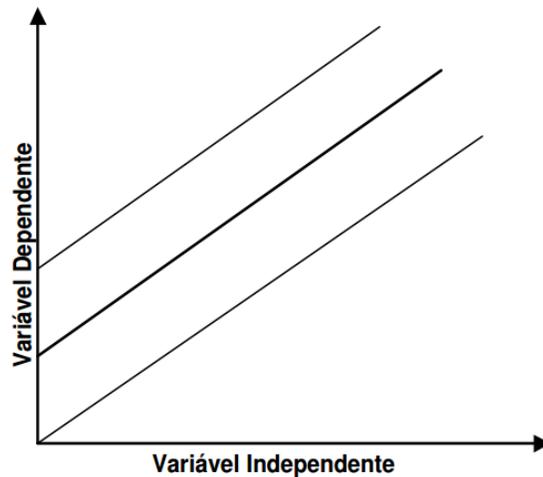


Figura 3.4 Significado da variância constante dos resíduos.

A ocorrência de variâncias não constantes nos resíduos é chamada de heterodasticidade. Sua ocorrência pode estar condicionada a especificações incorretas no modelo de regressão, e sua detecção é possível através do estudo residual dos erros. Se, à medida que o valor de  $y$  aumenta, os erros de predição também aumentam, tem-se heterodasticidade na variância.

Fundamentalmente, a violação desse pressuposto é preocupante na medida em que afeta a confiabilidade dos testes de significância. O teste Durbin-Watson pode indicar heterodasticidade e sua correção está vinculada à eliminação de algumas variáveis ou a transformação matemática do modelo, trazendo uniformidade dos erros percentuais ao longo da linha de regressão.

3. *Independência dos resíduos.* Isso significa que os resíduos do período  $t$  não são relacionados com os do período  $t-n$ , ou seja, são independentes dos dados passados. A falta de independência não afeta o valor dos parâmetros estimados, mas afeta diretamente as variâncias estimadas. Os fatores não incluídos no modelo (que compõem o termo de erro) não afetam sistematicamente o valor médio de  $y$  (os pontos positivos e negativos se anulam por serem equidistantes). A violação desse pressuposto compromete a consistência da estimativa do intercepto. Dessa forma, enquanto o coeficiente de regressão não é afetado, deve-se ter cuidado com a interpretação substantiva da constante.
4. *Normalidade dos resíduos.* Os erros são normalmente distribuídos em relação à linha de regressão; na média, sua influência pode ser desprezada, uma vez que o erro médio

apresenta um comportamento normalizado. Ao contrário, erros substancialmente assimétricos podem comprometer a eficiência do MMQ e podem levantar dúvidas quanto a razoabilidade de se estimar a média condicional de  $y$  a partir de  $x$ .

5. *A relação em análise inclui todas as variáveis importantes* – ou seja, se nem todas as variáveis independentes forem incluídas na relação, os coeficientes de regressão podem não ser precisos.
6. *Ausência de multicolinearidade.* A multicolinearidade é uma característica da série que se desenvolve quando duas ou mais variáveis independentes possuem forte correlação. A existência de multicolinearidade introduz erros grosseiros no resultado da regressão, produzindo erros na previsão da variável dependente. A ocorrência de multicolinearidade se dá quando um ou mais dos testes  $t$  assume resultados muito pequenos (não significativos) e os valores de  $R^2$  ou da estatística  $F$  são muito grandes.

A multicolinearidade é um problema frequentemente encontrado nos dados econômicos e de negócios tendo em vista a elevada correlação do tempo entre diferentes agregados como a população, a população economicamente ativa, o nível de renda disponível para consumo, vendas, estoques, custos, lucros, etc. A problemática da multicolinearidade em tais casos não deve ser desprezada, uma vez que a elevada correlação existente entre as mesmas pode prejudicar a sua utilização e, conseqüentemente, a modelagem.

Por exemplo, suponha que o modelo explicativo utiliza duas variáveis altamente correlacionadas,  $x_1$  e  $x_2$ , para explicar a variação de uma variável dependente qualquer,  $y$ . A variação total é formada pela variação associada a  $x_1$  mais a variação associada a  $x_2$  mais a variação comum entre  $x_1$  e  $x_2$ . O modelo de regressão de mínimos quadrados ordinários utiliza apenas a variação única de cada variável para estimar os coeficientes, ignorando a variância compartilhada. Quanto maior a correlação entre as variáveis independentes, menos informação estará disponível para estimar os coeficientes associados às variáveis explicativas.

### 3.2.6 Aspectos Positivos e Negativos da RL

O método de RL é um instrumento de previsão que precisa um número de variáveis inferior comparados aos modelos físicos e, por isso, permite chegar à solução de forma mais rápida. Apesar

da simplicidade e rapidez de análise, a RL pode apresentar desvantagens ao não respeitar devidamente os pressupostos básicos. Em particular, ressaltam-se os seguintes casos:

- Se as relações existentes entre as variáveis não são percebidas, efeitos antagônicos ou sinérgicos de efeito mútuo entre variáveis complicam a interpretação do fenômeno a partir das variáveis consideradas.
- O número das variáveis independentes é significativo e pode alterar o resultado da previsão. Assim, se nem todas as variáveis independentes forem consideradas na relação, os coeficientes de regressão podem ser imprecisos. Por exemplo, Guilhon (2002) observa no seu trabalho que, excluindo uma das duas variáveis independentes, o coeficiente de determinação passa da 0,71 a 0,2.

Além dos pressupostos básicos, para prever o valor de  $y$ , valores futuros das variáveis independentes precisam ser conhecidos. Isto pode representar uma dificuldade no recolhimento dos dados.

Por fim, é preciso reconhecer a natureza probabilística do modelo. A reta de regressão que se obtém através do MMQ indica a tendência, sendo tão somente uma aproximação da realidade. Como ilustrado na Figura 3.5, embora as  $x_k$  tenham seus valores fixados, são variáveis aleatórias com distribuição de probabilidades. Os valores de  $x$  e  $y$  são obtidos de uma amostra específica que contém distorções em relação à realidade.

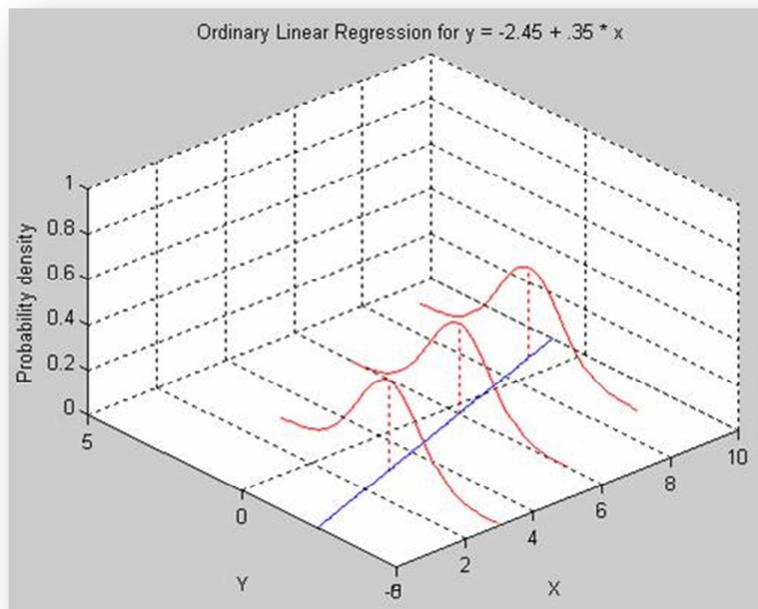


Figura 3.5 Natureza probabilística das variáveis do modelo de regressão (Rouhani, 2006).

### 3.3 Rede Neural Artificial (RNA)

RNAs são inspiradas no funcionamento do cérebro humano, ou seja, na rede neural biológica (Figura 3.6). São definidas como sendo uma forma de computação não-algorítmica, por não serem baseada em regras ou programas, sendo uma alternativa à computação algorítmica convencional (BRAGA, 2000). As RNAs são similares às Redes Neurais Biológicas (RNBs), pois a base de seu funcionamento é o neurônio. Segundo Bocanegra (2002), o neurônio artificial é a unidade fundamental de processamento de uma RNA, o qual recebe uma ou mais entradas, transformando-as em saídas.

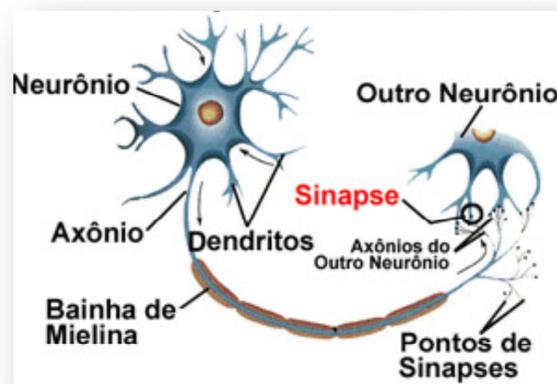


Figura 3.6 Comunicação entre neurônios (BALLONE, 2013).

De uma maneira geral, as atividades que compõe o desenvolvimento de uma RNA são:

- a) Definição do Problema: nesta etapa deve-se verificar se as características do problema a ser considerado são condizentes com o uso de uma RNA. Algumas características de uma boa aplicação para RNAs são: regras de resolução do problema desconhecidas ou difíceis de determinar e disponibilidade de um grande conjunto de exemplos. Alguns exemplos típicos de aplicação são: reconhecimento de formas, tratamento de sinal, visão, fala, previsão, modelagem, auxílio à decisão, e robótica.
- b) Escolha e adequação dos dados: para que se possa utilizar uma RNA é fundamental que se disponha de dados em quantidade e qualidade suficientes. Caso a quantidade de dados seja pequena, a rede não conseguirá criar um modelo suficientemente representativo para se ter um bom desempenho quando aplicado em situações reais após o seu desenvolvimento. Além disto, os dados

devem englobar todos os aspectos do problema em questão, a fim de que o modelo criado seja genérico. Em geral, tais dados precisam ser convertidos para um formato padrão para utilização pelas RNAs.

c) Treinamento da Rede: implica em aplicar o algoritmo de aprendizagem escolhido a cada um dos exemplos do conjunto de treinamento iterativamente, até que a rede atinja o comportamento desejado. Caso a rede não consiga atingir o desempenho mínimo, deve-se avaliar se os dados são adequados e realizar uma readequação dos mesmos, expandindo-os ou reduzindo-os.

d) Teste da Rede: após a rede ter sido treinada, deve-se testar a mesma para verificar se o modelo criado tem um índice de acerto aceitável em dados “novos”, ou seja, que ainda não foram aplicados à rede. Em caso positivo, passa-se ao uso em campo da rede neural. Se o desempenho não for adequado, deve-se retornar a uma etapa anterior e refazer aquela parte do processo, como, por exemplo, alterar o algoritmo de aprendizagem, a topologia da rede, ou eventualmente o próprio conjunto de dados.

e) Uso em Campo: se as etapas anteriores apresentaram sucesso, o modelo criado pode ser implementado para uso contínuo. Ainda assim, durante o uso podem surgir problemas, o que implica em retornar a uma etapa anterior e reiniciar o processo.

A Figura 3.7 ilustra o modelo de neurônio artificial proposto por McCulloch e Pitts em 1943 onde a  $x$  representa os estímulos de entrada, e  $w$  as ligações sinápticas.

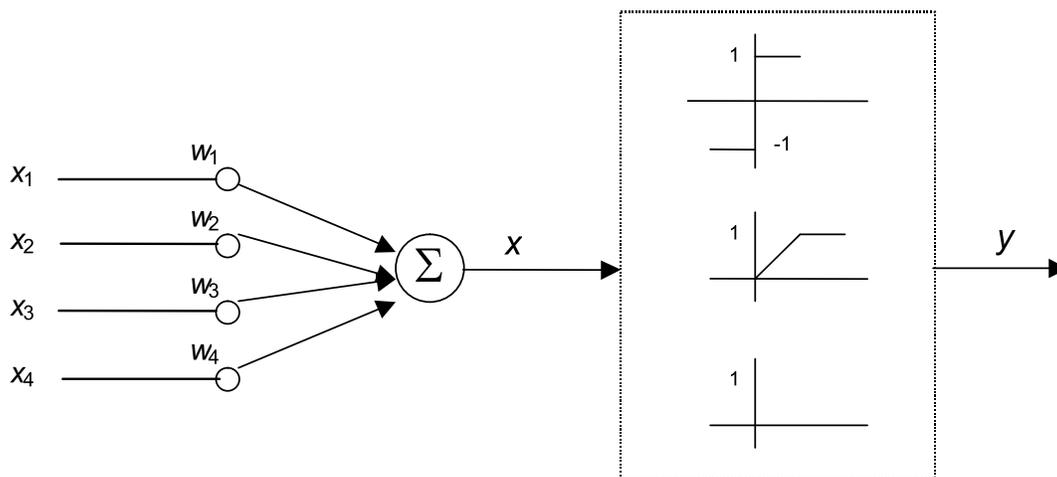


Figura 3.7 Modelo de neurônio artificial.

As ligações recebem um peso que representam seu potencial, chamado de pesos sinápticos, onde cada entrada vai ser multiplicada pelo peso de sua ligação, gerando um sinal tanto positivo ou negativo. Um sinal positivo é considerado excitatório e um sinal negativo, inibitório. De acordo

com Bialoskorski *et al.* (2006), as RNAs possuem quatro etapas de implementação que constituem seu ciclo de vida: definição da rede, treinamento, utilização da RNA e manutenção, como pode ser observado na Figura 3.8.

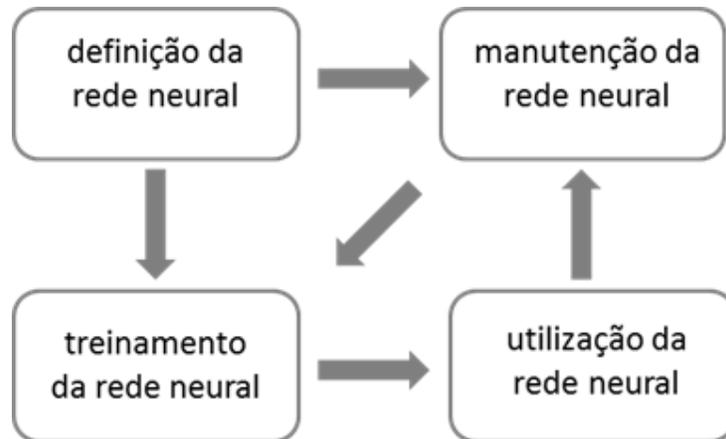


Figura 3.8 Ciclo de vida de uma RNA, baseado em Bialoskorski *et al.* (2006).

No ciclo de vida de uma RNA, a primeira etapa é a definição da rede neural. Nesta fase, são definidos o propósito da rede, o tipo de topologia que será utilizado, o tipo de treinamento e os dados de entrada. Em seguida, passa-se para a etapa de treinamento da rede, onde são inseridos os dados de entrada e feito o treinamento de acordo com as definições da etapa anterior. Posteriormente, segue a etapa de utilização da rede onde são extraídos os dados de saída, que são verificados e comparados; se a rede não atingir um valor próximo a seu objetivo, segue para a próxima fase. Na etapa de manutenção da rede, são feitos os ajustes necessários para que a rede seja treinada novamente e forneça dados mais confiáveis, caso exista a necessidade de alteração de seu modelo ou proposta, o que inicia outra vez a etapa de definição da rede, retomando o ciclo novamente.

Uns dos principais objetivos das redes neurais são modelos com boa capacidade de generalização, baseando no conjunto de dados utilizados na rede. Tendo determinado valor para a entrada e para uma saída desejável, a rede deve ser ajustada e treinada para que a diferença entre a saída da rede e a saída ideal, que é o erro, diminua a cada interação para se conseguir uma melhor generalização da rede.

### 3.3.1 Rede MultiLayer Perceptron (MLP)

A rede MLP é um dos paradigmas mais famosos e usados para implementar sistemas baseados em redes neurais. O processo de aprendizagem da rede MLP é dado da seguinte forma: um conjunto de padrões de entrada é apresentado à rede, que propaga estes sinais a cada camada. A utilização do algoritmo de aprendizagem *Backpropagation* consiste em comparar a saída desejada com a obtida pela rede, computando o erro associado com cada padrão de entrada. Este erro é propagado de trás para frente, ou seja, da camada de saída para a camada de entrada. Neste processo, os parâmetros da rede são ajustados. O método do gradiente descendente é usado para minimizar o erro global ao longo do processo de treinamento da rede.

### 3.3.2 Aspectos Positivos e Negativos da RNA

Entre as vantagens das RNAs pode-se indicar:

- A habilidade de lidar com informações incompletas e ruidosas, verificando-se a capacidade da rede em fornecer uma saída satisfatória;
- Não é necessária informação sobre o ambiente *a priori*, pois o aprendizado é feito através da apresentação de padrões à rede;
- Capacidade de aprender por meio de exemplos, e a habilidade de generalização a partir dos padrões de treinamento, fornecendo saídas satisfatórias a padrões novos que não foram vistos durante o treinamento.

Por outro lado, existe a dificuldade de justificar o comportamento das RNAs em determinadas situações:

- As RNAs são consideradas como “caixas pretas”, nas quais não se sabe por que a rede chega a um determinado resultado.
- Nas RNAs que utilizam o algoritmo de *backpropagation*, o tempo de treinamento tende a ser muito longo. Esta é uma limitação para arquiteturas muito grande ou grande quantidade de dados de treinamento.
- Encontra-se a dificuldade em determinar a arquitetura ideal das RNAs, de modo que ela seja grande o suficiente para conseguir resolver o problema e, ao mesmo tempo, pequena o suficiente para apresentar um treinamento rápido.
- Dependendo do tipo de aprendizagem, outra desvantagem nas RNAs é a necessidade de ter uma base de dados para o processo de treinamento.

Gomes e Montenegro (2010) reforçaram a importância da análise dos dados na definição das variáveis de entrada da rede neural. Observaram também a importância das informações de precipitação, principalmente nas simulações do período úmido. Concluíram que a metodologia de redes neurais possibilita a obtenção de uma previsão de vazões superiores às obtidas com os modelos estatísticos, para previsões diárias quanto semanais.

### 3.4 Lógica Nebulosa Fuzzy

Segundo Neto *et al.* (2006), a Lógica Nebulosa (do inglês Fuzzy Logic) é uma lógica multivalorada capaz de absorver informações vagas, normalmente descritas em uma linguagem natural e convertê-las para um formato numérico, de fácil manipulação computacional. A Lógica Nebulosa procura modelar o modo impreciso do raciocínio humano, auxiliando na habilidade humana de tomar decisões. Na Lógica Binária ou Lógica Booleana, uma declaração é falsa ou verdadeira, não havendo nada entre esses limites. Entretanto, entre a certeza de ser e a certeza de não ser (Lógica Binária), existem infinitos graus de incerteza (Lógica Fuzzy). Na Lógica Fuzzy, o raciocínio exato ou preciso (Lógica Binária) corresponde a um caso limite do raciocínio aproximado, sendo interpretado como um processo de composição Fuzzy. A Lógica Fuzzy foi desenvolvida na década de 1960 por Lofti A. Zadeh, da Universidade da Califórnia, em Berkeley (ZADEH, 1965). Ela combina lógica multivalorada, teoria probabilística e inteligência computacional para poder processar o conhecimento humano de representação simbólica (linguística). A Lógica Fuzzy tem por objetivo principal fazer com que as decisões tomadas pela máquina se aproximem cada vez mais das decisões humanas. Essa característica se acentua principalmente ao se trabalhar com uma grande variedade de informações vagas, imprecisas e incertas. A Lógica Fuzzy permite criar sistemas especialistas utilizando variáveis linguísticas para criar uma base de regras.

A Lógica Fuzzy é baseada na teoria dos Conjuntos Fuzzy. Esta é uma generalização da teoria dos Conjuntos Tradicionais para resolver os paradoxos gerados a partir da classificação “verdadeiro ou falso” da Lógica Clássica. Na Lógica Fuzzy, uma premissa varia em grau de pertinência de 0 a 1, o que leva a ser parcialmente verdadeira ou parcialmente falsa.

Em aplicações na engenharia, trabalha-se com números (temperatura, pressão, força etc) e as variáveis são contínuas. Para que um sistema Fuzzy possa ser utilizado em engenharia, faz-se necessário converter números (valores exatos) em variáveis linguísticas e vice-versa, pois a inferência Fuzzy utiliza variáveis linguísticas. Com a incorporação do conceito de “grau de pertinência”, a teoria dos Conjuntos Fuzzy estende a teoria dos Conjuntos Tradicionais. Os grupos

são rotulados qualitativamente (usando termos linguísticos, tais como: alto, baixo, quente, frio, longe, perto, etc) e os elementos deste conjunto são caracterizados variando o grau de pertinência (valor que indica o grau em que um elemento pertence a um conjunto). Por exemplo, um homem de 1,90 metro e um homem de 1,80 metro pertencem ao conjunto “alto”, embora o homem de 1,90 metro tenha um grau de pertinência maior neste conjunto. Geralmente os dados de entrada deste sistema são valores precisos, resultados de medições ou observações (conjuntos de dados, por exemplo).

### Conjuntos Fuzzy

A teoria dos Conjuntos Fuzzy é, em grande parte, uma extensão da teoria dos Conjuntos Tradicionais. Na teoria dos Conjuntos Tradicionais, os elementos de um conjunto  $A$  simplesmente pertencem ou não pertencem a um determinado universo  $X$ . A pertinência dos elementos do conjunto  $A$  no universo  $X$  pode ser definida como:

$$f_A(x) = \begin{cases} 1 & \text{para } x \in A \\ 0 & \text{para } x \notin A \end{cases} \quad (3.12)$$

Na teoria dos Conjuntos Fuzzy, a pertinência dos elementos do conjunto  $A$  no universo  $X$  pode assumir um número infinito de valores diferentes no intervalo  $[0,1]$ . Portanto, um conjunto nebuloso é um conjunto de pares ordenados:

$$A = \{\mu_A(x)/x\}; \quad x \in X \quad (3.13)$$

onde  $\mu_A(x)$  é a função de pertinência de  $x$  em  $A$  e é definida como o mapeamento de  $X$  no intervalo fechado  $[0,1]$ :

$$\mu_A(x): X \rightarrow [0,1] \quad (3.14)$$

Com base nessa concepção de conjuntos, Zadeh (1965) definiu também algumas propriedades e operações com conjuntos Fuzzy como a união e a interseção. É interessante ressaltar que propriedades dos conjuntos ordinários como distributividade, associatividade, comutatividade, involução e leis transitivas também valem para as operações com conjuntos Fuzzy. Tais resultados buscam estender os conceitos da teoria ordinária de conjuntos, incluindo-a como caso particular.

## Variáveis linguísticas

A lógica Fuzzy manipula variáveis linguísticas. Por variável linguística entende-se uma variável cujos valores são nomes ou sentenças ao invés de assumirem apenas valores específicos como ocorre com variáveis numéricas. Expressões linguísticas são típicas da natureza humana de tomar decisões. Por exemplo: “*Se estiver quente, vou ligar o ar condicionado no máximo*”. O valor de uma variável linguística (denominado valor linguístico ou valor Fuzzy) é uma sentença composta por termos primários (no exemplo da temperatura: alta, baixa, média), por conectivos lógicos (e, ou, não), e de modificadores (muito, pouco, extremamente). Cada valor linguístico, por sua vez, é representado por meio de um conjunto Fuzzy, com sua respectiva função de pertinência. Variáveis linguísticas servem tanto para caracterizar fenômenos complexos via uma abordagem vaga, quando são identificadas através de seus valores Fuzzy, quanto para uma caracterização quantitativa e precisa, quando se utiliza as funções de pertinência. Esta característica dual das variáveis linguísticas as torna convenientes para abordar problemas de forma quantitativa ou qualitativa.

O comportamento de tais sistemas é descrito por um conjunto de regras nebulosas do tipo: se  $x$  é  $A$  <premissa> então  $y$  é  $B$  <consequente>, que são uma forma de representar o conhecimento. Normalmente, denomina-se a proposição  $x$  é  $A$  como antecedente, enquanto  $y$  é  $B$  é chamado consequente (ou conclusão);  $x$  é chamada de variável premissa. Na concepção tradicional, ativa-se (dispara-se) uma regra Se-Então somente quando a variável premissa  $x$  é exatamente igual ao antecedente. Neste caso, a inferência produzida será:  $y$  é exatamente igual ao consequente.

### 3.4.1 Sistema de Inferência Fuzzy (SIF)

O SIF agrega os conceitos de conjuntos Fuzzy e variáveis linguísticas. A estrutura básica de um sistema de inferência é mostrada na Figura 3.9.

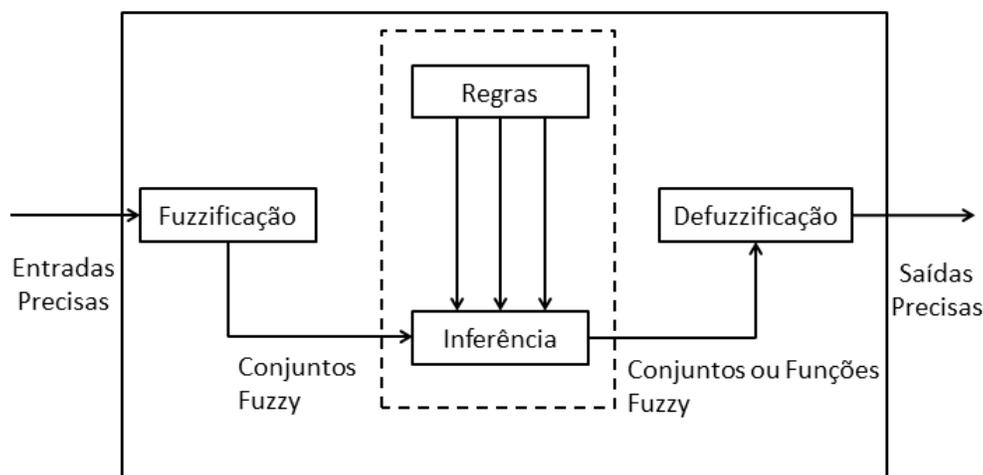


Figura 3.9 Diagrama esquemático do sistema de inferência.

No primeiro estágio, chamado Fuzzificação, efetua-se um mapeamento dos dados para os conjuntos de entrada. A Fuzzificação é o processo de conversão da entrada numérica em conjuntos difusos. É uma operação fundamental, pois em grande parte das aplicações de lógica difusa os dados são numéricos, sendo necessário então fuzzificar estes em conjuntos difusos. Um conjunto numérico  $x$  é convertido no conjunto difuso  $X$  através de um fuzzificador:  $X = \text{fuzzificador}(x)$ .

No estágio de inferência, ocorrem as operações com conjuntos Fuzzy ao longo de regras para processar, por meio de um mecanismo de inferência, as informações da entrada (Se) e produzir uma conclusão (Então). Dependendo do formato da saída do sistema, definem-se dois tipos de regras condicionais: regras linguísticas (Mandani) e regras de Takagi-Sugeno (TS). Nas regras linguísticas, tanto o antecedente (entrada) como o conseqüente (saída) são conjuntos difusos, do tipo:

Se ( $X$  é  $LX$ ) Então ( $Y$  é  $LY$ )

Nas regras de TS apenas o antecedente é um conjunto difuso. O conseqüente é função do antecedente, do tipo:

Se ( $X_1$  é  $LX_1$ ) e ( $X_2$  é  $LX_2$ ) e ... e ( $X_n$  é  $LX_n$ ), Então  $y = f(x_1, x_2, \dots, x_n)$

onde,  $x_1, x_2, \dots, x_n$  representam valores numéricos a cada uma das variáveis linguísticas do antecedente. O conseqüente corresponde a uma variável numérica, onde o resultado é obtido como uma função  $f$  dos valores numéricos dos antecedentes. Geralmente é uma função polinomial de ordem 0 ou 1, definindo-se o sistema difuso de (TS) de ordem 0 ou 1. Em se tratando de sistema de ordem 1, têm-se:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_mx_m = b_0 + \sum_{i=1}^m b_ix_i \quad b_i \in R, i = 1, 2, \dots, k \quad (3.15)$$

Em sistemas de ordem 0, a equação acima se reduz a  $y = b_0$ . O modelo Fuzzy TS é capaz de aproximar um sistema não-linear a uma combinação de vários sistemas lineares afins, através da decomposição de todo o espaço de entrada / saída com uma equação linear. Sistemas TS são capazes de representar, de forma aproximada ou exata, dinâmicas não lineares quaisquer, como a combinação de modelos lineares válidos localmente. Segundo Mozelli (2008) os sistemas TS são menos complexos do que outras classes de modelos Fuzzy, facilitando aspectos de identificação e modelagem.

Uma vez obtido o conjunto Fuzzy de saída através do processo de inferência, no estágio de defuzzificação é efetuada uma interpretação dessa informação. A defuzzificação é uma operação de conversão de um conjunto Fuzzy em um valor numérico. Esta defuzzificação pode ser feita através de vários métodos; os mais consagrados são a Média dos Máximos e o Centro de Gravidade. No primeiro, a saída determinística (não-nebulosa) é obtida tomando-se a média entre os dois

elementos extremos no universo que correspondem aos maiores valores das funções de pertinência. Com o Centro de Gravidade, a saída é o valor no universo que divide a área sob a curva em duas partes iguais (CORREA e BAESSA, 2007).

### 3.4.2 Aspectos Positivos e Negativos do SIF

Os aspectos positivos dos Sistemas Fuzzy são os seguintes:

- Os sistemas Fuzzy têm a capacidade de controlar sistemas com muitas variáveis de saída utilizando um só controlador Fuzzy, com um bom desempenho;
- Têm a habilidade de controlar processos com característica não-linear e de alta ordem, na qual a determinação do modelo matemático e o controle clássico do sistema são muito complexos;
- Implementam técnicas de controle baseadas na experiência de um especialista e em aspectos intuitivos, utilizando proposições linguísticas (regras) e entradas imprecisas.

Algumas das limitações que apresentam os Sistemas Fuzzy são as seguintes:

- Sofrem a influência da grande quantidade de parâmetros na configuração feita pelo usuário, algumas das quais são: número de funções de pertinência de cada variável, número de regras, seleção dos métodos de implicação e agregação, método de defuzzificação, assim como os parâmetros de cada função de pertinência.
- A precisão do sistema Fuzzy é limitada pela experiência do especialista na configuração dos parâmetros, a qual é determinada pelo conhecimento do processo pelo especialista.

A tabela 3.1 mostra a comparação entre as RNAs e os SIFs.

Tabela 3.1 Comparação entre as redes neurais artificiais e os SIFs (LUNA, 2003).

Redes Neurais Artificiais	Sistemas de Inferência Fuzzy
Conhecimento <i>a priori</i> não utilizado	Conhecimento <i>a priori</i> pode ser incorporado
Capacidade de aprendizado	Utiliza o conhecimento linguístico
Caixa preta	De fácil interpretação (regras SE – ENTÃO)
Algoritmo de aprendizado complexo	Fácil interpretação e implementação
Dificuldade para extração de conhecimento	Conhecimento disponível

### 3.5 Sistema de Inferência Neuro Fuzzy (SINF)

Os SH em geral são a sinergia obtida pela combinação de duas ou mais técnicas de modelagem. O SH Neuro Fuzzy é a combinação das técnicas de Fuzzy com RNAs. O foco destes sistemas está em obter um sistema mais poderoso e com menos deficiências. A ideia básica é a construção de um SIF numa estrutura paralela distribuída de tal forma que os algoritmos de aprendizado das redes neurais possam ser aproveitados nestes SHs para ajustar os parâmetros associados com as funções de pertinência do SIF (SANCHEZ, 2009). Os SINF combinam as vantagens destas duas abordagens, obtendo-se um sistema com capacidade de aprendizado, capaz de aproveitar informação linguística e numérica através da base de regras do sistema de inferência e utilizar conhecimento *a priori* para definir a estrutura do sistema.

Entretanto, a capacidade de aprendizado é, sem dúvida, a característica mais importante das redes neurais que os sistemas Fuzzy herdaram, gerando as redes neuro Fuzzy. Através da aprendizagem, estes dois componentes da inteligência computacional, quando combinados, transformam-se em um único sistema que supera as deficiências individuais (LUNA, 2003).

A Figura 3.10 apresenta a estrutura de um SINF dividido em cinco camadas.

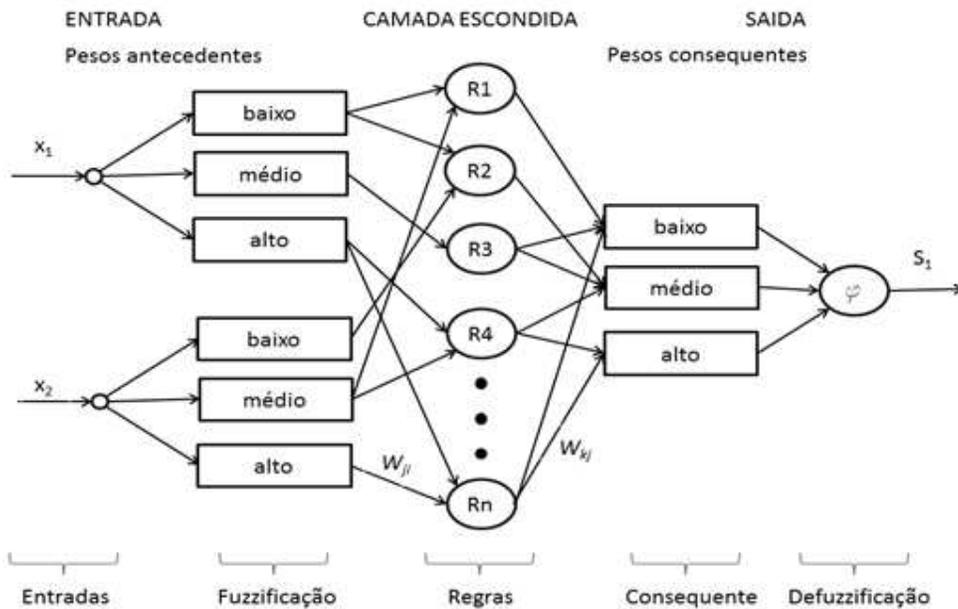


Figura 3.10 Estrutura de um SINF dividido em cinco camadas, (SANCHEZ, 2009 - adaptada).

A camada de entrada representa as variáveis de entrada, as quais são normalizadas e escalonadas dentro do intervalo numérico  $[0,1]$ . A segunda camada é a etapa de fuzzificação. Nesta etapa, os intervalos de cada variável de entrada são divididos em diversos níveis (Baixo, Médio e Alto), os quais indicam os pesos da rede para cada entrada. A terceira camada é definida pelas regras do SIF. A quarta camada é determinada pelos consequentes das regras, e a quinta camada, ou camada de saída, é a etapa de defuzzificação, onde se calcula o valor numérico de saída.

O tipo de aprendizado, em relação à forma de apresentação dos padrões, divide-se em aprendizado *offline* e aprendizado *online*. No aprendizado *online*, a atualização dos parâmetros da estrutura é feita para cada padrão apresentado. No aprendizado *off-line*, a atualização dos parâmetros é feita após a apresentação de todos os dados do conjunto de treinamento.

### 3.5.1 Algoritmo Expectation Maximization (EM) na versão *offline*

A estrutura do sistema é composta por um conjunto de  $M$  regras nebulosas de primeira ordem, do tipo TS. Seja  $x^h = [x_1^h, x_2^h \dots x_k^h] \in R^k$  o vetor de entrada no instante  $h, h \in Z_0^+$  e  $\widehat{y}^h \in R$  a saída do modelo, para uma determinada entrada  $x^h$ . O espaço de entrada representado  $x^h \in R^k$  é subdividido em  $M$  sub regiões, cada uma representada por uma regra Fuzzy (Figura 3.11).

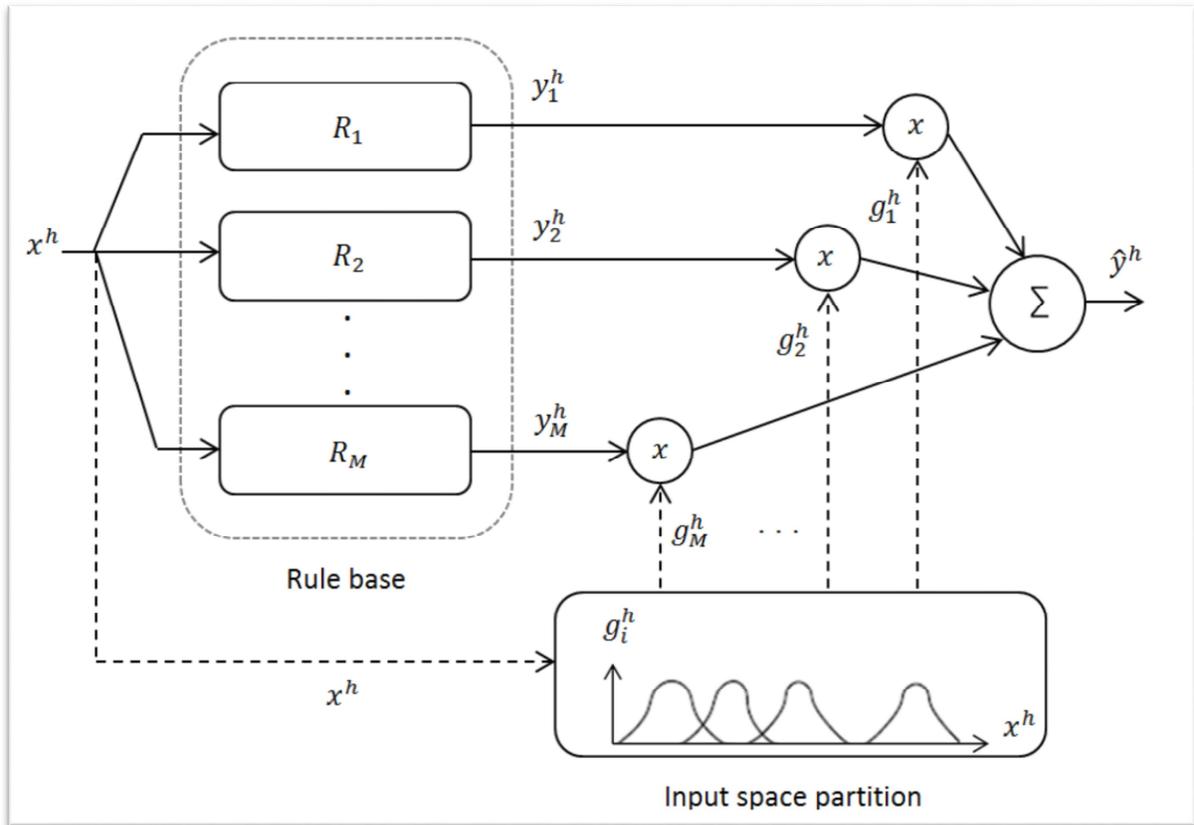


Figura 3.11 Estrutura geral do modelo (LUNA, 2011 - adaptada).

A representação do sistema baseado em regras nebulosas pode ser descrita por camadas. A partição do espaço de entrada é representada pelas primeiras duas camadas. Cada sub-região gerada é definida por um centro  $c_i \in R^k$  e por uma matriz de covariância  $V_i$  de dimensão  $k \times k$ . Assim, cada vetor de entrada terá um grau de pertinência associado a cada uma das sub-regiões geradas pela base de regras.

A primeira camada é a responsável pelo cálculo das funções de pertinência a cada interação  $h$ . Estas funções de pertinência são definidas pelo produto  $\alpha_i P_i^h$ , sendo  $\alpha_i P_i^h$  coeficientes positivos tal que  $\sum_{i=1}^M \alpha_i = 1$  e  $P_i^h = P[i|x^h]$  definida como:

$$P_i^h = P[i|x^h] = \left( \frac{1}{(2\pi)^{k/2} \det(V_i)^{1/2}} \right) \times \left( \exp \left\{ -\frac{1}{2} (x^h - c_i) V_i^{-1} (x^h - c_i)^T \right\} \right) \quad (3.16)$$

na qual  $V_i$  e  $c_i$  são a matriz de covariância e centro associados à  $i$ -ésima regra, respectivamente. A Equação 3.17 representa a probabilidade de escolher a regra  $i$ , dado que apenas a entrada  $x^h$  é conhecida, assim como os parâmetros  $V_i$  e  $c_i$ . Observa-se que  $P[i|x^h]$  é uma função

de densidade de probabilidade condicional, que pode ser interpretada como a probabilidade do vetor de entrada  $x^h$  pertencer à região ativa da  $i$ -ésima regra.

O grau de pertinência de cada padrão  $x^h$  é representado por  $g_i(x^h) \in [0,1]$ , dado por:

$$g_i(x^h) = g_i^h = \frac{\alpha_i \cdot P[i|x^h]}{\sum_{q=1}^M \alpha_q \cdot P[q|x^h]} \quad (3.17)$$

Estas funções de pertinência  $g_i^h$  são fornecidas pela segunda camada do modelo, tal que:

$$\sum_{q=1}^M g_i^h = 1 \quad (3.18)$$

Os coeficientes  $\alpha_i$ ,  $i = 1, \dots, M$  podem ser interpretados como índices indiretos de relevância para cada regra gerada durante o treinamento, de tal forma que, quanto mais alto seja o valor de  $\alpha_i$ , maior será a importância da  $i$ -ésima regra nebulosa.

A saída do sistema  $\hat{y}^h$  é calculada na última camada da estrutura do modelo, como uma combinação não linear dos consequentes  $y_i^h$  ponderados pelos seus respectivos graus de pertinência  $g_i^h$  ou seja:

$$\hat{y}^h = \sum_{i=1}^M g_i^h \times y_i^h \quad (3.19)$$

sendo, neste caso, a saída  $y_i^h$  de cada regra  $R_i$ , definida como:

$$y_i^h = \varphi^h \times \theta_i^T \quad (3.20)$$

na qual  $\varphi^h = [x_1^h, x_2^h \dots x_k^h]$  é um vetor que contém os dados de entrada  $x^h$  e o termo constante;

$\theta_i = [\theta_{i0}, \theta_{i1} \dots \theta_{ik}]$  é o vetor de coeficientes dos modelos locais lineares, de dimensão  $1 \times (k + 1)$  para  $i = 1, \dots, M$ , que representam os consequentes das regras nebulosas.

A dinâmica do modelo pode ser interpretada como um mecanismo de inferência nebulosa, sendo a base de regras composta por regras  $R_i$ ,  $i = 1, \dots, M$ , do tipo  $R_i^h$ : Se  $x^h$  pertence a  $i$ -ésima região com grau de pertinência  $g_i^h$  Então  $y_i^h = \varphi^h \times \theta_i^T$ .

### 3.5.2 Aspectos Positivos e Negativos do SINF

Os Aspectos positivos do SINF são os seguintes:

- Embora a análise e a síntese de SINF sejam mais complexas, estes modelos têm demonstrado desempenho superior em diversas aplicações, particularmente em identificação de sistemas não lineares e estimação de parâmetros, reconhecimento de padrões, controle de processos, previsão de vazões, aproximação de funções e previsão de séries financeiras, entre outras (LUNA, 2003);
- Os SINF combinam as vantagens dos sistemas Fuzzy e RNA, mostrando um enorme potencial para aplicações que combinem conhecimento qualitativo com robustez;
- A lógica Fuzzy provê uma interface de alto nível, de rápida computação e amigável para programar. Isso permite que o especialista se concentre nos objetivos funcionais ao invés dos detalhes matemáticos;
- O projetista não precisa ter conhecimento prévio do processo, levando a uma fácil adaptabilidade aos diferentes processos;
- Os SINF podem lidar de uma maneira melhor que as RNAs com o problema de ruído nos dados;
- Os SINF tem capacidade de auto-aprendizado, auto-organização e auto-direcionamento, imitando a capacidade humana do processo de tomada de decisão.

Algumas das limitações dos SINF são:

- Trabalham com um reduzido número de entradas, devido ao problema de explosão combinatória das regras. Portanto, um sistema com muitas entradas demanda um maior esforço computacional.
- Tem limitação na construção de sua própria estrutura devido à estrutura fixa; quando há a capacidade de alterar sua estrutura, são limitados pelo elevado número de regras.

Ballini e Soares (2003), apresentaram um modelo de rede neural nebulosa para previsão de vazões sazonais aplicado para o problema de previsão de vazões medias mensais de três usinas hidrelétricas situadas em diferentes regiões do Brasil. Os resultados mostraram que a rede neural nebulosa forneceu um melhor desempenho com erros significativamente menores que as outras abordagens, ao compará-la com os resultados obtidos pelos modelos de rede neural artificial e modelos de séries temporais.

Rabelo *et al.* (2011) encontraram uma vantagem na aplicação dos sistemas de inferência Fuzzy do tipo TS, a inclusão da experiência do grupo de pesquisa no estudo do comportamento otimizado dos reservatórios para geração de energia elétrica, na definição dos termos primários (conjuntos Fuzzy) da variável linguística e na construção da base de regras de produção. Por outro

lado algumas melhorias ainda podem ser investigadas, principalmente no ajuste das funções de pertinência, do número de regras e dos consequentes das regras, empregando-se técnicas como RNAs ou AG.

### **3.6 Comportamento Esperado de um Modelo de Previsão de Vazões Baseado em RL e de um Modelo Baseado em SINF**

O comportamento esperado dos modelos deve basear-se na compreensão das técnicas de RL e SINF. O modelo de RL pode alcançar o conhecimento através do MMQ e armazenar esse conhecimento nos coeficientes de regressão. No entanto, a RL envolve uma estrutura rígida e um conjunto rígido de premissas que representa um conjunto de pontos por uma função. Espera-se que um ponto localizado muito fora do conjunto de pontos não deva ser bem representado pelo modelo.

Ao contrário, o SINF não necessita determinar previamente as relações entre as variáveis dependentes e independentes, pois a forma das relações é estabelecida durante o processo de aprendizagem. Se uma relação linear entre as variáveis dependentes e independentes é adequada, os resultados do SINF devem aproximar-se às do modelo de regressão linear. Se uma relação não linear é mais adequada, as abordagens do SINF devem aproximar-se com maior precisão aos dados observados. No caso de estudo de previsões de vazões, dificilmente a relação entre variáveis dependentes e independentes serão lineares, e, por isso, em tese, os resultados fornecidos pelo SINF devem ser mais adequados.

Segundo Patuelli *et al.* (2006), as RNAs tem a habilidade de encontrar a solução com um grau de adequação ótimo quando as relações entre as variáveis não são inteiramente conhecidas, ou quando somente está disponível um conhecimento limitado do fenômeno. O compromisso desta flexibilidade é que os pesos sinápticos de uma rede neural não são facilmente interpretáveis. Portanto, se o objetivo é de explicar as relações entre as variáveis dependentes e independentes, seria melhor usar um modelo estatístico mais tradicional. No entanto, se a interpretabilidade do modelo não é importante, como no caso das previsões de vazões, é mais provável obter bons resultados usando um SINF. Espera-se, portanto, que esta habilidade vai favorecer os resultados do SINF em comparação aos da RL que falham quando a previsão é afetada por variáveis exógenas.

Segundo Wray *et al.* (1994), o desempenho das RNAs não é depreciado em virtude da ocorrência de multicolinearidade, da ausência de normalidade ou da existência de valores anômalos nos dados de entrada, ao contrário das técnicas tradicionais de previsão, como RL, que impõem para

seu uso hipóteses de normalidade nos dados e ausência de multicolinearidade e valores anômalos. Espera-se, portanto, que a previsão de precipitação usada pelo modelo de SINF possa trazer apenas efeitos positivos, embora este dado de entrada não esteja imune a erros.

Uma última consideração sobre o comportamento esperados dos modelos refere-se à diversidade dos dados de entrada. O modelo RL utiliza um número inferior de variáveis, portanto, espera-se um desempenho melhor onde as precipitações nos períodos de previsão de vazões não sejam tão diferentes dos períodos antes da previsão. Em tese, o desempenho do modelo SINF deveria ser indiferente em relação à variabilidade da precipitação, pois utiliza previsões de precipitação.

# Capítulo 4

## Ferramentas Utilizadas

Este capítulo apresenta ao leitor as ferramentas de apoio à aplicação da metodologia ilustrada nos próximos capítulos. A primeira ferramenta, o SISPREV, gerencia/executa os estudos de previsão de vazões e será usada na fase de modelagem. Além dele, Excel e Statistica serão utilizados para as fases de mineração, preparação, limpeza, transformação e avaliação dos dados.

## 4.1 SISPREV

O SISPREV é uma ferramenta de gerenciamento de estudos de previsão de vazões desenvolvida na UNICAMP durante um projeto de P&D com a empresa AES Tietê. Esta ferramenta permite ao usuário prever vazões diárias e mensais com base em modelos RL e SINF para UHEs e postos de medição. Cada um dos seus modelos realiza a previsão para intervalos diários e mensais de forma independente. As previsões diárias são executadas para quinze dias subsequentes, enquanto as previsões mensais são executadas para os próximos seis meses.

Os objetivos desta ferramenta, além de prever vazões diárias e mensais, são avaliar o desempenho dos modelos implementados, comparar modelos matemáticos para previsão de vazões e servir como ferramenta de pesquisa. Neste trabalho, a ferramenta em questão será empregada como instrumento de análise entre os modelos, a fim de que os resultados dos modelos sejam comparados com as vazões observadas.

Esse sistema foi implementado usando a linguagem técnica de programação C++ (HOLLINGWORTH *et al.*, 2001). O ambiente de desenvolvimento integrado utilizado chama-se C++ *Builder Enterprise Suite*, criado pela Borland. O mecanismo interno das consultas baseia-se na *Strutred Query Language* (SQL) (BOWMAN *et al.*, 1996). Para o desenvolvimento do banco de dados foi utilizado o Microsoft SQL SERVER 2008 EXPRESS *with Advanced Services* (MICROSOFT, 2013). O projeto de software fundamenta-se nos conceitos da programação orientada a objetos, tais como: encapsulamento, herança e polimorfismo.

A Figura 4.1 ilustra a estrutura física da ferramenta SISPREV.



Figura 4.1 Estrutura geral da ferramenta de gestão de estudos de previsão de vazões.

A base de dados contém os dados que serão utilizados como entrada para os modelos de previsão de vazões. Ela é composta por três tabelas: a) aproveitamentos; b) dados diários e; c) dados mensais. Os arquivos-texto contêm o registro completo das características dos estudos de previsão de vazões.

Este tipo de estrutura (base de dados e arquivo-texto) apresenta dois benefícios: além de permitir ao usuário alterar as características de um estudo para simular novas situações sem alterar as informações originais da base de dados, ela também facilita a troca de informações entre pesquisadores. De fato, em geral, arquivos-texto são menores que arquivos de base de dados.

A análise e a comparação entre os modelos matemáticos objetos desta pesquisa serão feitas observando os resultados apresentados pela interface do SISPREV. Ela permite a visualização dos resultados em forma gráfica e de planilhas. A Figura 4.2 ilustra a representação gráfica das trajetórias de vazão diária prevista e observada para a UHE Barra Bonita usando o modelo SINP. A Figura 4.3 mostra os mesmos resultados em formato de tabela. Se o estudo é aplicado para um período passado, a trajetória das vazões observadas é também apresentada. Para facilitar a avaliação dos resultados da previsão em função da chuva no período, um gráfico com os valores de precipitação é exibido.

A interface do SISPREV torna-se, assim, versátil para analisar o desempenho de um modelo, para ser utilizado como ferramenta de pesquisa e para comparação entre os modelos. Além disso, ele pode ser aplicado para qualquer conjunto de modelos matemáticos desenvolvidos para o mesmo propósito.

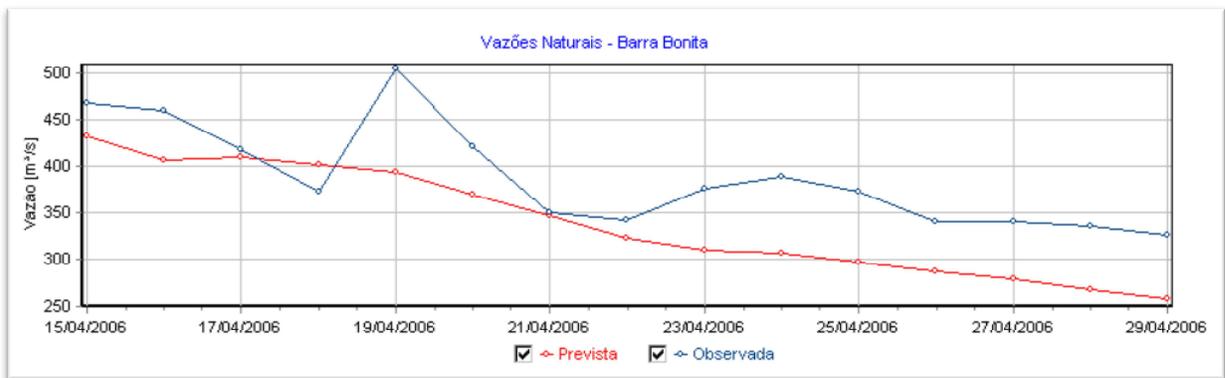


Figura 4.2 Exemplo de resultado em forma gráfica do SISPREV.

Barra Bonita		
Data	Prev (N)	Obs (N)
15/04/2006	431,9	468,0
16/04/2006	407,1	459,0
17/04/2006	409,5	418,0
18/04/2006	402,2	372,0
19/04/2006	394,3	505,0
20/04/2006	368,5	421,0
21/04/2006	346,0	350,0
22/04/2006	322,8	342,0
23/04/2006	309,4	376,0
24/04/2006	306,5	389,0
25/04/2006	296,9	372,0
26/04/2006	287,2	341,0
27/04/2006	278,7	341,0
28/04/2006	268,2	335,0
29/04/2006	257,8	325,0

Figura 4.3 Exemplo de resultado em forma de planilha do SISPREV.

#### 4.1.1 Modelo de Regressão Linear Múltipla (RLM)

O SISPREV utiliza o modelo de RLM para estimar o valor esperado da variável de interesse, ou seja, a vazão. Ele o faz a partir de duas variáveis explanatórias de entrada conhecidas previamente, que são: a) vazões observadas; e b) precipitações.

A aplicação da metodologia, tanto do modelo de previsão mensal quanto do modelo de previsão diário, é feita por modelos de ajustes estatísticos, utilizando-se de registros históricos, onde as variáveis independentes foram utilizadas para o período 2002-2008 de calibração. Em seguida, tem-se um período para validação dos parâmetros de regressão obtidos, os quais são utilizados na fase de previsão.

#### 4.1.2 Modelo Sistema de Inferência Neuro Fuzzy (SINF)

Os modelos de previsão de vazões diárias e mensais utilizando ferramentas da inteligência computacional são baseados em SINF – *offline*, tomando como dados de insumo a) vazões observadas; b) precipitações observadas; e c) precipitações previstas (fornecidas pela Somar Meteorologia). Estes sistemas codificam na sua estrutura regras nebulosas do tipo TS de primeira ordem, adotando a forma "Se A então B", onde A representa o antecedente, e B representa o consequente da regra nebulosa definido por um modelo linear das entradas do próprio modelo.

A inicialização do modelo utiliza um aprendizado construtivo baseado na versão *offline* do algoritmo EM que permite determinar valores faltosos mediante estimativas de máxima verossimilhança (JACOBS, 1998). Este algoritmo é composto por uma sequência iterativa de passos EM que maximiza a função de verossimilhança por meio do ajuste dos parâmetros do modelo. Finalmente, a vazão estimada é calculada pela ponderação entre as saídas locais (consequentes) e os respectivos graus de pertinência.

O período de 2005 até 2006 foi usado para a fase de aprendizagem do modelo SINF diário e 2000-2007 para aprendizagem do SINF mensal. Em primeiro lugar, a estrutura SINF é inicializada. Para definir o número das regras  $M$  na estrutura do modelo, o algoritmo é processado sobre dados de entrada-saída. Depois dessa inicialização, os parâmetros do modelo são reajustados através do algoritmo EM, com o objetivo de maximização da verossimilhança dos valores observados  $y^h$  para cada passo  $M$  do processo de aprendizado. A função objetivo é definida por:

$$\mathfrak{J}(D, \Omega) = \sum_{i=1}^L \ln \left( \sum_{i=1}^M g_i(x^h, C) \times P(y^h | x^h, \theta_i) \right) \quad (4.6)$$

onde  $L$  é o número de entrada-saída usados para a otimização do modelo  $D = \{x^h, y^h | h = 1, \dots, N\}$ ,  $\Omega$  contém os parâmetros do modelo e  $C$  contém apenas os parâmetros antecedentes. Para maximizar a verossimilhança é preciso saber a distribuição dos dados. O modelo é ajustado assumindo uma distribuição normal dos parâmetros observados, em virtude da falta do conhecimento da distribuição real de precipitação e vazões.

Um dado geral de entrada do modelo, que considera vazões, precipitações observadas e precipitações previstas é assim definido:

$$x^h = [x_1^h \quad x_2^h \quad \dots \quad x_k^h] = [r_1^h \quad \dots \quad x_a^h \quad s_1^h \quad \dots \quad s_b^h] \quad (4.7)$$

onde  $a \in Z^+$  indica o número de componentes do vetor de entrada que contém informações de precipitações observadas e futuras, representadas pelo  $r$ , e  $b \in Z^+$  indica o número de componentes do vetor de entrada que contém informações de vazões observadas  $s$ . Isto significa que, para prever a vazão no instante  $h$ , os dados de entrada de precipitação têm que considerar as defasagens dos últimos  $a-1$  dias  $D$  de precipitação acumulados, a precipitação acumulada pelo instante  $h$  dos dias  $D$  e as últimas vazões  $b$  gravadas até  $h-1$ , como variáveis de entrada. Para os passos sucessivos, são necessárias vazões e precipitações previstas.

Com o objetivo de melhorar o poder explanatório das séries temporais de chuva para cada usina hidrelétrica, o modelo SINF usa como variável de precipitação acumulada a precipitação dos

últimos dias  $D$ , onde  $D$  é selecionada através da maximização dos coeficientes de correlação entre vazão observada e precipitação acumulada durante esses dias  $D$ .

Depois que os dados de entrada e saída são construídos, um conjunto de dados é normalizado ao intervalo de unidade e, em seguida, dividido em dois subconjuntos, um conjunto de treinamento de dados e um conjunto de dados de teste. O conjunto de dados de treinamento é composto pelos primeiros dois anos de dados históricos, enquanto o conjunto de dados de teste foi constituído pelo último (LUNA, 2011).

## **4.2 Outras Ferramentas**

Para as fases de mineração, preparação, limpeza, transformação e avaliação dos dados serão utilizadas duas plataformas informáticas. Microsoft Excel 2010 será utilizado para a elaboração dos dados e a versão 8 do software Statistica, desenvolvido pela StatSoft, para as elaborações gráficas.

# Capítulo 5

## Metodologia

Neste capítulo, descreve-se a metodologia proposta para desenvolver a mineração de dados e analisar a qualidade das previsões. Enfim detalham-se os índices de desempenho mais usados na hidrologia para comparar e avaliar modelos de previsão de vazões.

## 5.1 Cross Industry Standard Process for Data Mining (CRISP-DM)

Para analisar o volume de dados disponível no SISPREV, escolheu-se a metodologia CRISP-DM, que pode ser traduzida como *processo padrão inter-industrial para mineração de dados*, por ser considerado o padrão de maior aceitação (LAROSE, 2005). O objetivo da Mineração de Dados é extrair ou minerar conhecimento de grandes volumes de dados através da procura de padrões consistentes, como regras de associação ou sequências temporais, a fim de detectar relacionamentos sistemáticos entre variáveis. A CRISP-DM foi projetada para possibilitar um modelo de processo genérico que possa ser especializado de acordo com as necessidades de uma empresa ou indústria em particular (CHAPMAN *et al.*, 2000).

Esta metodologia, baseada em um modelo hierárquico de processos, é constituída de seis fases organizadas de maneira cíclica, conforme mostrado na Figura 5.1. Apesar de ser composto por fases, o fluxo não é unidirecional, podendo ir e voltar entre as fases. Cada fase do modelo contém suas respectivas tarefas, e essas tarefas possuem relacionamentos entre si. Os relacionamentos entre as tarefas variam de acordo com os objetivos do processo de mineração, do interesse do usuário, do assunto tratado e das características dos dados. Além disso, cada atividade produz um ou mais artefatos que servirão ao propósito de documentar a execução do processo de mineração. Para uma melhor compreensão dessa metodologia, apresenta-se a seguir os detalhes de cada fase.

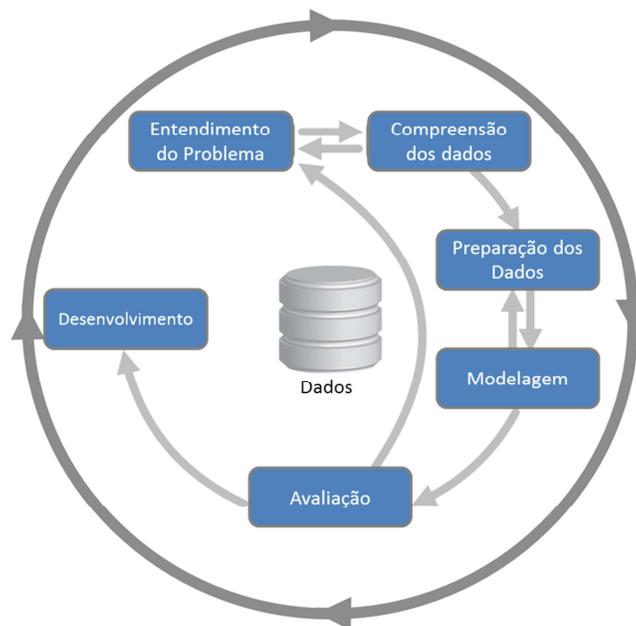


Figura 5.1 Fases do CRISP-DM (CHAPMAN *et al.*, 2000 - adaptada).

- **Entendimento do Problema:** esta fase busca identificar os objetivos que se deseja atingir com a mineração de dados. O principal objetivo deste trabalho será conhecer as características dos modelos de previsão de vazões diárias e mensais.
- **Compreensão dos Dados:** nesta etapa, identificam-se as informações relevantes e ocorre uma primeira familiarização com o seu conteúdo, descrição, qualidade e utilidade. Será analisado o banco de dados disponível.
- **Preparação dos Dados:** esta fase consiste em definir um conjunto de dados, a partir do qual serão validados os modelos de previsão de vazões RL e SINF. Nesta fase, serão escolhidos os períodos objetos de estudo e extraídos do banco de dados os parâmetros que fazem parte destes períodos. Estes parâmetros serão os dados de entrada dos modelos.
- **Modelagem:** são aplicadas as técnicas de modelagem mais apropriadas dependendo dos objetivos levantados no Entendimento do Problema. Neste trabalho, serão executados os modelos RL e SINF contidos no SISPREV sobre os períodos escolhidos, bem como produzidos os dados de saída, ou seja, as previsões de vazões diárias e mensais.
- **Avaliação:** consiste na avaliação do modelo, revendo os passos seguidos e verificando se os objetivos pretendidos foram alcançados. Neste estudo, serão avaliados os modelos RL e SINF.
- **Desenvolvimento:** nesta última etapa, define-se um conjunto de ações que conduz à organização e disponibilização do conhecimento adquirido. Este assunto será tratado nas conclusões.

## 5.2 Avaliação dos modelos

Em problemas de previsão de séries temporais, os índices de desempenho executam a importante tarefa de quantificar a qualidade da predição obtida. A medição da qualidade da previsão é determinante para calibrar, validar e comparar modelos de previsão. A calibração de um modelo hidrológico consiste na modificação de alguns parâmetros, para obtenção de um modelo que replique condições previamente conhecidas do processo natural modelado (VIESSMAN e LEWIS, 2003). Os modelos hidrológicos são calibrados com dados de um período de tempo e verificados com dados de outro período. Este último processo é conhecido como validação.

A comparação, foco deste trabalho, pode ser feita entre modelos de previsão diferentes ou entre um modelo de previsão a ser avaliado e um preditor de referência, utilizando índices de

desempenho cujas equações ponderam preditores de referência. Os índices de desempenho mais usados na literatura são os seguintes:

- Erro Médio - *Mean Error* (ME);
- Erro Absoluto Médio - *Mean Absolute Error* (MAE);
- Erro Quadrático Médio - *Mean Squared Error* (MSE);
- Raiz Quadrada do Erro Quadrático Médio - *Root Mean Square Error* (RMSE);
- Erro Percentual Médio - *Mean Percentage Error* (MPE);
- Erro Percentual Absoluto Médio - *Mean Absolute Percentage Error* (MAPE);
- U-THEIL;
- Coeficiente de eficiência de Nash-Sutcliffe (NS);
- Coeficiente de eficiência de Nash-Sutcliffe aplicado aos logaritmos (NS-LOG);
- Distância Multicritério (DM);
- *Confusion Index* (CI).

A seguir, serão detalhados os índices acima mencionados. A série temporal a ser predita é indicada por  $y'$ , e a previsão é indicada por  $y$ , ou seja,  $y'$  corresponde à vazão observada e  $y$  à vazão prevista.  $N$  é o número de previsões avaliadas e  $i$  seu índice de tempo.

O ME, Equação 5.1, é a soma das diferenças entre as vazões previstas e observadas, dividida pelo número de previsões avaliadas. O ME indica o quanto as previsões de vazões se desviam (em  $m^3/s$ ) em relação aos valores observados. O ME pode ser um número positivo ou negativo no intervalo  $(-\infty, +\infty)$  sendo que  $ME = 0$  pode indicar uma previsão perfeita. A vantagem deste índice é a de revelar se a previsão subestima (sinal -) ou sobrestima (sinal +) a vazão observada. No entanto, a sua desvantagem é a possibilidade de revelar um valor errôneo que seja igual ou próximo à zero, em caso de erros igualmente distribuídos em sobrestima e subestima. Além disso, considerando que a unidade de medida é em  $m^3/s$ , não podem ser comparadas bacias de características e tamanhos diversos.

$$ME = \frac{1}{N} \sum_{i=1}^N y_i - y'_i \quad (5.1)$$

O MAE, Equação 5.2, representa o ME em termos absolutos. Por ser um valor absoluto, perde-se a informação da sobrestima ou subestima do erro e, assim como o ME, não pode ser utilizado para avaliar diferentes bacias. A sua vantagem sobre o ME é que eventuais erros de sinal

oposto não se anulam. A tendência é que esse valor seja mais elevado para locais com maiores vazões e seja menor para locais com valores menores de vazões (GUILHERME *et al.*, 2007).

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - y'_i| \quad (5.2)$$

O MSE, Equação 5.3, é o quadrado da soma das diferenças entre as vazões previstas e observadas, dividido pelo número de previsões avaliadas. O MSE, em  $m^6/s^2$ , pode ser um valor no intervalo  $(0, +\infty)$ , sendo que  $ME = 0$  indica uma previsão perfeita. Por ser um valor quadrático, perde-se a informação da sobrestima ou subestima do erro e, assim como o MAE e o ME, resulta mais elevado para locais com maiores vazões. O MSE é utilizado em vários artigos que tratam especificamente da predição de longo-prazo no setor energético (CAMPOS, 2008).

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2 \quad (5.3)$$

O RMSE, Equação 5.4, é a raiz quadrada do MSE. Medido em  $m^3/s$ , é um valor no intervalo  $(0, +\infty)$ , sendo que  $ME = 0$  indica uma previsão perfeita. O RMSE é sempre maior ou igual ao MAE. Quanto maior a diferença entre eles, maior será a variância nos erros individuais na amostra. Se o RMSE for igual ao MAE, todos os erros da série são da mesma grandeza. No MAE, todas as diferenças individuais são ponderadas igualmente na média. A peculiaridade do RMSE é a sua sensibilidade a valores anômalos ou extremos (CARVALHO e ROMANEL, 2007). Isto significa que, na comparação entre modelos, o RMSE é mais útil quando grandes erros são particularmente indesejáveis, pois ele penalizará o modelo com este tipo de erro. Assim como o MAE, ME e MSE, perde-se a informação da sobrestima ou subestima do erro e não pode ser utilizado para avaliar diferentes bacias.

$$RMSE = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2\right)} \quad (5.4)$$

O MPE, Equação 5.5, representa a média da diferença entre os valores previstos e observados, expressa em porcentagem dos valores observados. Por isso, pode-se dizer que o MPE é o ME em forma de porcentagem. O MPE pode ser uma porcentagem positiva ou negativa no intervalo  $(-\infty, +\infty)$ , sendo que  $MPE = 0\%$  pode indicar uma previsão perfeita. Uma vantagem deste índice é a de revelar se a previsão mediamente subestima (sinal -) ou sobrestima (sinal +) a vazão observada. Além disso, diferentemente do ME, o MPE permite avaliar e comparar bacias diferentes, tendo em vista que representa um valor adimensional. A sua desvantagem é poder erroneamente revelar um valor igual ou muito próximo à zero, se houver erros igualmente distribuídos em sobrestima e subestima. O MPE é utilizado em vários artigos que tratam especificamente da previsão de longo-prazo no setor energético (CAMPOS, 2008).

$$MPE = \frac{1}{N} \sum_{i=1}^N \frac{y_i - y'_i}{y'_i} \times 100 \quad (5.5)$$

O MAPE, Equação 5.6, é o MAE em forma de porcentagem. Representa a média da diferença absoluta entre os valores previstos e observados, expressa em porcentagem dos valores observados. O MAPE é uma porcentagem positiva no intervalo  $(0, +\infty)$ , sendo que  $MAPE = 0$  indica uma previsão perfeita.

Por se tratar de um valor absoluto, perde-se a informação da sobrestima ou subestima do erro, mas, diferentemente do MAE, pode ser utilizado para avaliar diferentes bacias, por ser uma medida em porcentagem. Este valor resulta elevado para locais com vazões pequenas, pois qualquer desvio poderá ser relevante e resulta menor para locais com vazões maiores, pois, mesmo com grandes variações, estas influenciam em forma menor em relação ao total a ser previsto (GUILHERME *et al.*, 2007). Nos trabalhos na área de previsão de curtíssimo prazo, o índice MAPE é predominante (CAMPOS, 2008).

Segundo o ONS (2011), o MAPE, pode ser considerado como adequado para uma avaliação de curto prazo porque voltado para a avaliação de cada ponto isoladamente. Liew *et al.* (2007) apresentaram a seguinte classificação para simulações hidrológicas:  $MAPE < 10\%$ , muito bom;  $10\% \leq MAPE \leq 15\%$ , bom;  $15\% \leq MAPE \leq 25\%$ , satisfatório e  $MAPE > 25\%$ , o modelo é inadequado.

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - y'_i}{y'_i} \times 100 \right| \quad (5.6)$$

As equações dos próximos índices a serem tratados, U-THEIL, NS e NS-LOG, contêm a informação do preditor de referência. Nestes casos, o resultado da comparação contém a informação adicional do desempenho do modelo em relação ao preditor de referência.

U-Theil, Equação 5.7, é a razão entre a raiz quadrada do quadrado da soma das diferenças entre as vazões previstas e observadas, e a raiz quadrada do quadrado da soma das diferenças entre as vazões observadas no período  $i$  e no período  $i-1$ . O U-THEIL é uma porcentagem positiva no intervalo  $(0+\infty)$ . O U-Theil avalia o desempenho da previsão em relação à previsão de referência dita ingênua ou trivial. Define-se previsão ingênua ou trivial como a previsão cuja estimativa do valor futuro é igual ao valor atual.

O coeficiente U de Theil analisa a qualidade de uma previsão através dos seguintes valores:

- U -Theil  $> 1$ , significa que o erro do modelo é maior do que o da previsão ingênua;
- U -Theil  $< 1$ , significa que o erro do modelo é menor do que o da previsão ingênua.

Valores menores do que a unidade indicam um melhor desempenho em relação ao preditor padrão considerado (GUITIERREZ *et al.*, 2003).

$$U - THEIL = \frac{\sqrt{\sum_{i=1}^N (y_i - y'_i)^2}}{\sqrt{\sum_{i=1}^N (y'_i - y'_{i-1})^2}} \quad (5.7)$$

Na Equação 5.8 do indicador NS,  $\bar{y}$  representa a média das observações no período considerado. O segundo termo da Equação representa a razão entre o quadrado da soma das diferenças entre as vazões previstas e observadas e o quadrado da soma das diferenças entre as vazões observadas e a média das observações. Ressalta-se a analogia com o coeficiente de determinação  $R^2$  da regressão linear (Equação 3.7). O NS é adimensional e pode assumir valores no intervalo  $(-\infty, 1)$ . NS = 1 corresponde à previsão perfeita das vazões no período em foco. NS < 0 mostra que utilizar as previsões do modelo é pior do que utilizar o valor médio das observações.

Este índice é frequentemente usado na hidrologia para validar modelos hidrológicos (LUNA *et al.*, 2011). Segundo o ONS (2011), o NS, levando-se em consideração a média das observações, pode ser adequado para uma avaliação de longo prazo. O NS realiza previsões mais acertadas quando o aproveitamento encontra-se com vazões elevadas (ANDRADE, 2011). Segundo Tucci *et al.* (2003), valores acima de 0,70 mostram ajuste razoável e valores maiores que 0,80 são considerados bons. Santhi *et al.* (2001) apresentam a seguinte classificação para este coeficiente: NS  $> 0,65$  o modelo é considerado muito bom;  $0,54 \leq NS \leq 0,65$  o modelo é considerado bom e  $0,5 < NS < 0,54$ , satisfatório. Andrade *et al.* (2013) escolheram o coeficiente NS para calibrar e

validar o modelo Soil and Water Assessment Tool (SWAT) com valores de 0,66 e 0,87 para as fases de calibração e validação, respectivamente.

Na literatura científica este índice de desempenho encontra-se também sob a seguinte denominação: *Mass Curve Coefficient E* (WALLIS e TODINI, 1975). De fato, a Equação 5.9 do índice E é equivalente à Equação 5.8 do índice NS.

$$NS = 1 - \frac{\sum_{i=1}^N (y_i - y'_i)^2}{\sum_{i=1}^N (\bar{y} - y'_i)^2} \quad (5.8)$$

$$E = \frac{\sum_{i=1}^N (\bar{y} - y'_i)^2 - \sum_{i=1}^N (y_i - y'_i)^2}{\sum_{i=1}^N (\bar{y} - y'_i)^2} \quad (5.9)$$

O índice NS-LOG, Equação 5.10, utiliza a mesma Equação do índice NS, mas aplica o logaritmo a cada parâmetro. Assim como o NS, este indicador é adimensional e pode assumir valores no intervalo  $(-\infty, 1)$ . NS – LOG = 1 corresponde à previsão perfeita. NS < 0 mostra que utilizar as previsões do modelo é pior do que utilizar o valor médio das observações.

Este índice valoriza mais as vazões baixas do que o coeficiente NS, tornando-se menos adequado a vazões elevadas (GUILHON *et al.*, 2007). Além disso, assim como o NS, levando-se em consideração a média das observações, pode ser adequado para uma avaliação de longo prazo.

Segundo o ONS (2011), a utilização de uma grande quantidade de índices pode conduzir a dois problemas de difícil compreensão. O primeiro diz respeito à relação entre os valores de diversos indicadores. O segundo está relacionado à decisão que se deve tomar se encontrados indicadores com resultados conflitantes. Para solucionar estes problemas, o ONS passou a utilizar desde 2010 um indicador único. O índice proposto chama-se DM, cuja Equação é:

$$NS - LOG = 1 - \frac{\sum_{i=1}^N (\log(y_i) - \log(y'_i))^2}{\sum_{i=1}^N (\log(y_i) - \log(\bar{y}_i))^2} \quad (5.10)$$

O DM é um índice de desempenho híbrido, pois tem características pertencentes a duas categorias, NS e MAPE. Este índice é uma porcentagem positiva no intervalo  $(0, +\infty)$ , sendo que DM = 0% indica uma previsão perfeita. O DM calcula a distância euclidiana ao ponto ótimo entre

os indicadores (1-NS) e MAPE . Ao fazer uma composição de dois indicadores distintos, o indicador único DM abrange todas as características deles.

$$DM = \sqrt{(1 - NS)^2 + MAPE^2} \quad (5.11)$$

Alem dos índices quantitativos, o Confusion Index (CI) indica a tendência da previsão, ou seja, quantas vezes a predição, independentemente do erro cometido, acerta a direção da previsão. O CI calcula a proporção do tempo em que o modelo fornece tendências equivocadas. O cálculo é simplesmente a contagem desses eventos em uma mesma amostra (NOTINI, 2009). Na Equação 5.12,  $q = 1$  representa a previsão com tendência errada e  $q = 0$  a previsão com tendência correta. O CI é uma percentagem positiva no intervalo (0, 1), sendo que CI = 0 indica uma previsão cujos períodos não contêm tendências erradas.

$$CI = \frac{1}{N} \sum_{i=1}^N (q_i) \quad (5.12)$$

Resumindo as características dos índices de desempenhos analisados anteriormente, pode-se afirmar que, se a unidade de medida é adimensional, o índice resulta adequado para avaliar vazões elevadas. Ao contrário, se a unidade de medida é dimensional, por exemplo, m<sup>3</sup>/s, o índice resulta adequado para vazões baixas. Além disso, índices que incluem a média das observações são adequados para uma avaliação de longo prazo.

Segue a Tabela 5.1 que sintetiza as principais características dos índices apresentados.

*Tabela 5.1 Índices de desempenho usados para previsões de vazões.*

Índice	Unidade	Limites	Previsão Perfeita	Vazões elevadas	Vazões baixas	Curto prazo	Longo Prazo	Sub-sobrestima
ME	m <sup>3</sup> /s	$-\infty, +\infty$	0		X	X		X
MAE	m <sup>3</sup> /s	$-\infty, +\infty$	0		X	X		
MSE	m <sup>6</sup> /s <sup>2</sup>	0, +∞	0		X	X		
RMSE	m <sup>3</sup> /s	0, +∞	0		X	X		
MPE	adim.	$-\infty, +\infty$	0	X		X		X
MAPE	adim.	0, +∞	0	X		X		
U-THEIL	adim.	0, +∞	0			X		
NS	adim.	$-\infty, 1$	1	X			X	
NS-LOG	adim.	$-\infty, 1$	1		X		X	
DM	adim.	0, +∞	0	X		X	X	
CI	adim.	0, 1	0	X	X	X	X	

# Capítulo 6

## Estudo de Casos

Neste capítulo, apresenta-se a aplicação da metodologia proposta para avaliar modelos de previsão de vazões. Serão, assim, percorridas as etapas contidas na CRISP-DM. A aplicação da metodologia CRISP-DM prevê uma primeira fase de Entendimento do Problema onde são explicados os objetivos do estudo. Na fase sucessiva, de Compreensão dos Dados, será analisado o banco de dados disponível no SISPREV. Na fase de Preparação dos Dados, define-se o conjunto, que representa os dados de entrada dos modelos RL e SINF. Para as fases seguintes da metodologia, de Modelagem e Avaliação, apresenta-se um estudo de caso de previsões de vazões diárias e um estudo de caso de previsões de vazões mensais. Na fase de Modelagem, serão executados os modelos RL e SINF e produzidos os dados de saída, ou seja, as previsões de vazões diárias e mensais. Na fase de Avaliação, os resultados serão comparados com as vazões registradas e, através dos índices de desempenho, os modelos serão avaliados. O conteúdo da última etapa, Desenvolvimento, onde se disponibiliza o conhecimento adquirido, será tratado no Capítulo 7, Conclusões e Trabalhos Futuros.

## 6.1 Entendimento do Problema

O principal objetivo deste trabalho é conhecer as características dos modelos de previsão de vazões diárias e mensais implementados na ferramenta SISPREV. Pretende-se aferir o desempenho de cada um deles, bem como avaliar as previsões em condições típicas e particulares. Além disso, busca-se analisar as seguintes avaliações de pesquisa, em relação ao gerenciamento das UHEs:

- Estabelecer qual modelo (RL ou SINP) responde com maior precisão nas previsões de vazões afluentes mensais e diárias;
- Estabelecer se os modelos têm performance melhor em condições climáticas típicas ou particulares, como períodos chuvosos e de seca.

## 6.2 Compreensão dos Dados

A Figura 6.1 mostra a região estudada, situada no estado de São Paulo, e os rios onde as UHEs estão localizadas. As UHEs estão agrupadas em duas bacias hidrográficas. A primeira, formada pelo Rio Tietê, possui somente usinas com eclusas, enquanto a segunda, formada pelos Rios Grande, Pardo e Mogi Guaçu, possui usinas sem eclusas.

O banco de dados disponível contém as informações de vazões e precipitações relativas a doze anos, compreendendo o período de 2000 a 2011. Estas variáveis são divididas de acordo com as dez UHEs e os quinze postos de medição localizados nas mesmas bacias hidrográficas onde as UHEs estão localizadas.



Figura 6.1 Localização das usinas hidrelétricas (AES, 2013).

O SISPREV associa o nome das UHEs e dos postos de medição aos seus respectivos código ANEEL, conforme Tabela 6.1 e 6.2.

*Tabela 6.1 UHEs armazenadas no SISPREV.*

Código	Nome	Rio
61998080	Água Vermelha	Grande
62744080	Bariri	Tietê
62729080	Barra Bonita	Tietê
61818080	Euclides da Cunha	Pardo
61811080	Graminha	Pardo
62790080	Ibitinga	Tietê
61819080	Limoeiro	Pardo
61880000	Mogi Guaçu	Mogi Guaçu
62829580	Nova Avanhandava	Tietê
62820080	Promissão	Tietê

*Tabela 6.2 Postos de medição armazenados no SISPREV.*

Código	Nome
61812000	Abaixo Caconde
61807002	Abaixo Cascata das Antas
61811100	Bom Jesus
61802502	Cachoeira da Carmo
61860000	Cristália
62800000	Fazenda São Benedito
62776800	Gravião Peixoto
62752000	Invernada Recreio
62560000	Laranjal Paulista
62705000	Piracicaba
61991000	Porto Pontal
62376000	Rasgão
61945080	Rio Verde
61817004	São José do Rio Pardo
61874000	Sítio Marangone

Para cada dia/mês do período 2000-2011 e para cada UHE/posto de medição são associadas variáveis de precipitação e de vazão. A Tabela 6.3 mostra as variáveis de entrada para as previsões diárias e a Tabela 6.4 as variáveis de entrada para as previsões mensais. Assim, considerando o período de doze anos, o banco de dados do SISPREV disponibiliza aproximadamente 250.000 dados.

*Tabela 6.3 Variáveis de entrada para os modelos de previsão de vazões diárias.*

*P = precipitação [mm], Q= vazão [m<sup>3</sup>/s], t = período [dia].*

Variáveis de entrada	Previsões diárias	
	RL	SINF
P(t-5)	x	
P(t-4)	x	x
P(t-3)	x	x
P(t-2)	x	x
P(t-1)	x	x
P(t)	x	x
P(t+1)		x
P(t+2)		x
P(t+3)		x
P(t+4)		x
P(t+5)		x
P(t+6)		x
P(t+7)		x
P(t+8)		x
P(t+9)		x
P(t+10)		x
P(t+11)		x
P(t+12)		x
P(t+13)		x
P(t+14)		x
Q(t-3)		x
Q(t-2)	x	x
Q(t-1)	x	x

Tabela 6.4 Variáveis de entrada para os modelos de previsão de vazões mensais.

$P$  = precipitação [mm],  $Q$  = vazão [ $m^3/s$ ],  $t$  = período [mês].

Variáveis de entrada	Previsões mensais	
	RL	SINF
P(t-4)		X
P(t-3)		X
P(t-2)		X
P(t-1)	x	X
P(t)	x	X
P(t+1)		X
P(t+2)		X
P(t+3)		X
P(t+4)		X
P(t+5)		X
Q(t-3)		X
Q(t-2)	x	X
Q(t-1)	x	X

As Tabelas 6.3 e 6.4 mostram as variáveis precipitação  $P$  e vazão  $Q$  associadas aos intervalos de tempo  $t-n$  para os modelos RL e SINF, referindo-se a observações registradas. Os intervalos  $t+n$  referem-se a previsões de precipitações, requeridas só pelo modelo SINF. O banco disponibilizado pelo SISPREV associa às variáveis de previsão de precipitação nos intervalos  $t+n$  dados observados depois do intervalo em análise, pois o algoritmo que gerencia o arquivo vai substituir previsões de precipitação com as observações, uma vez que elas são disponíveis. Assim, neste estudo, o modelo SINF trabalhará sob hipótese de previsões de precipitações perfeitas. Lembra-se que o objetivo dos modelos de previsão é a determinação da vazão  $Q$  associada a intervalos futuros  $t+n$ , com  $n \in [1, 15]$  para previsões de vazões diárias e  $n \in [1, 6]$  para previsões mensais.

Na análise do banco de dados encontrou-se a incompletude dos parâmetros do modelo SINF em relação a todos os postos de medição e à UHE Mogi Guaçu. Assim, o estudo de comparação será feito sobre as usinas, conforme Tabela 6.1, com exceção da UHE Mogi Guaçu, completando nove UHEs no total.

Para analisar o conteúdo do banco de dados, decidiu-se visualizar os dados do banco em gráficos de dispersão apresentando vazão *versus* precipitação. Sendo as variáveis divididas de acordo com as nove UHEs, são gerados nove gráficos de dispersão diária, compreendendo vazões e precipitações observadas, e nove de dispersão mensal, compreendendo vazões e precipitações

observadas no período escolhido 2002-2010. O software Microsoft SQL SERVER 2008 EXPRESS foi usado com finalidade de extrair, numa sequência ordenada temporalmente e segundo o código da UHE, as precipitações e as vazões registradas no período em exame. Na Figura 6.2, apresenta-se a janela usada para essa análise.

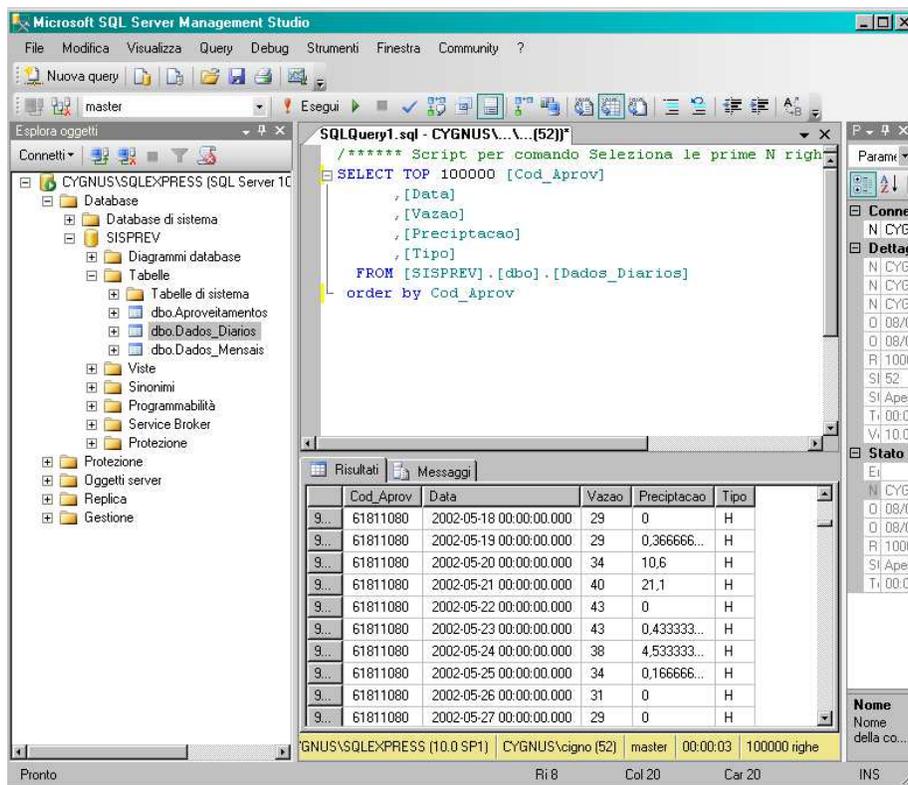


Figura 6.2 Janela operativa do SQL para listar as vazões e precipitações da UHE analisada.

Os dados obtidos foram elaborados com o software Statsoft Statistica para gerar os gráficos das Figuras 6.3 - 6.29. Com esta ferramenta foi possível também gerar a reta de RLS associada a cada diagrama de dispersão.

Figuras 6.3 – 6.11 Gráficos de dispersão diária vazão vs precipitação

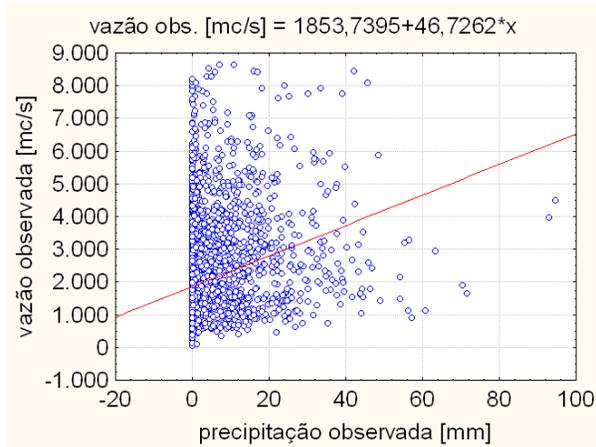


Figura 6.3 Água Vermelha.

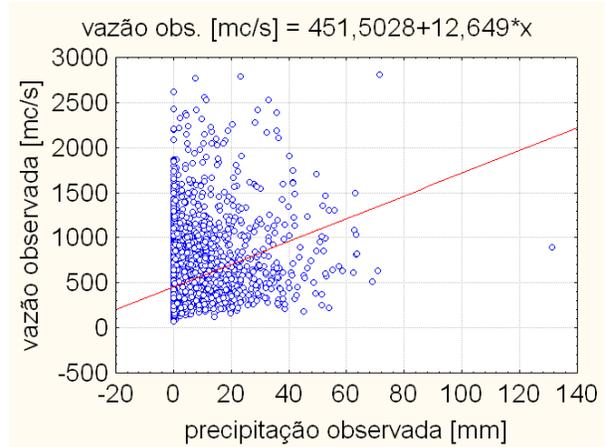


Figura 6.4 Bariri.

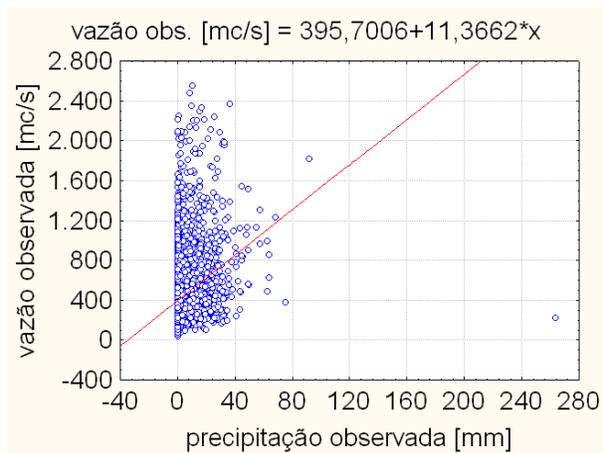


Figura 6.5 Barra Bonita.

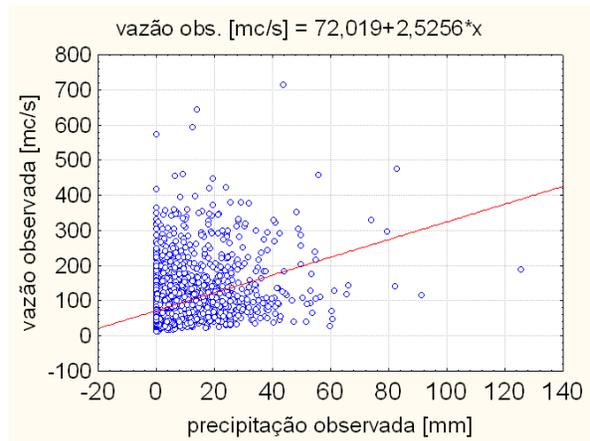


Figura 6.6 Euclides da Cunha.

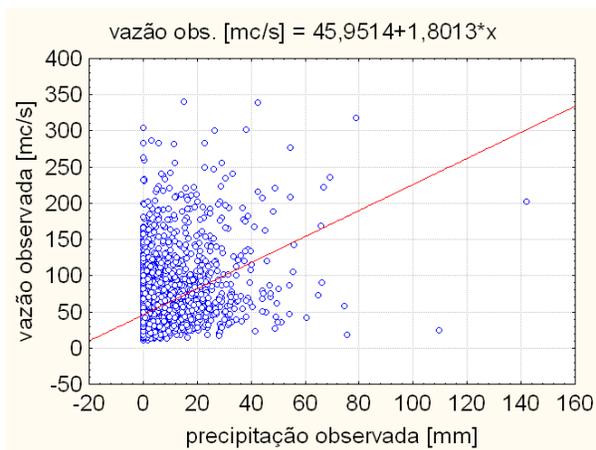


Figura 6.7 Graminha.

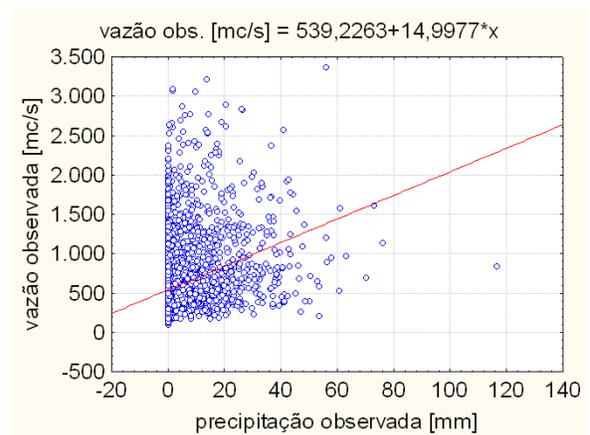


Figura 6.8 Ibitinga.

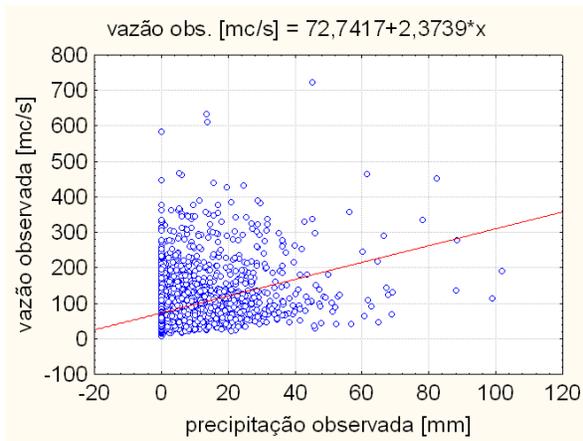


Figura 6.9 Limoeiro.

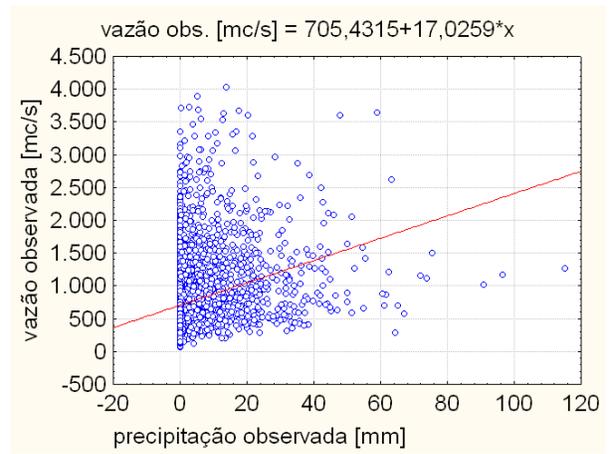


Figura 6.10 Nova Avanhandava.

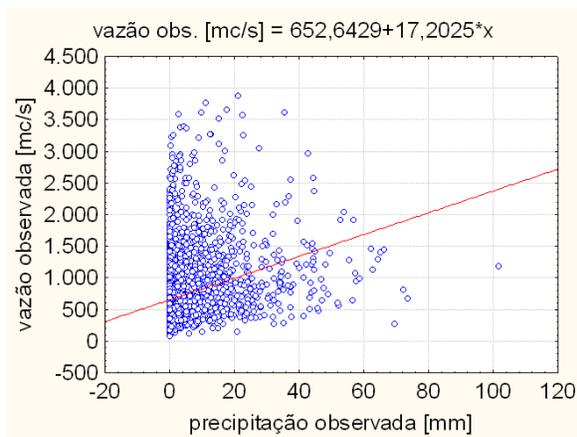


Figura 6.11 Promissão.

Figuras 6.12 – 6.20 Graficos de dispersão mensal vazão vs precipitação

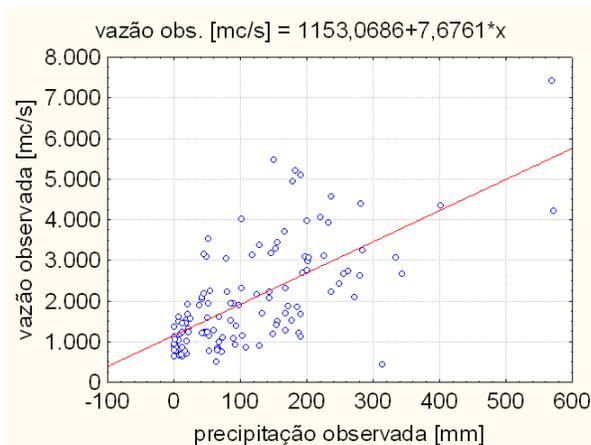


Figura 6.12 Água Vermelha.

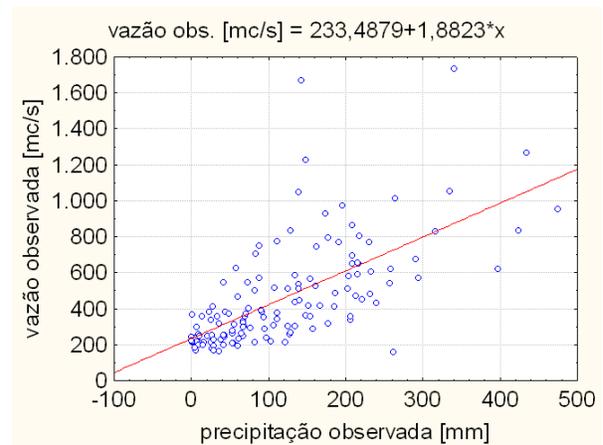


Figura 6.13 Bariri.

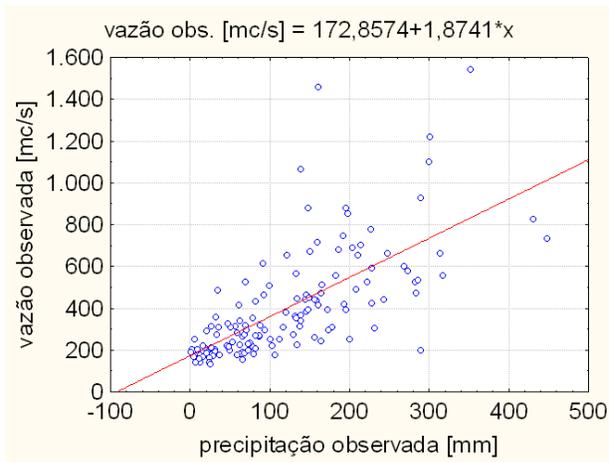


Figura 6.14 Barra Bonita.

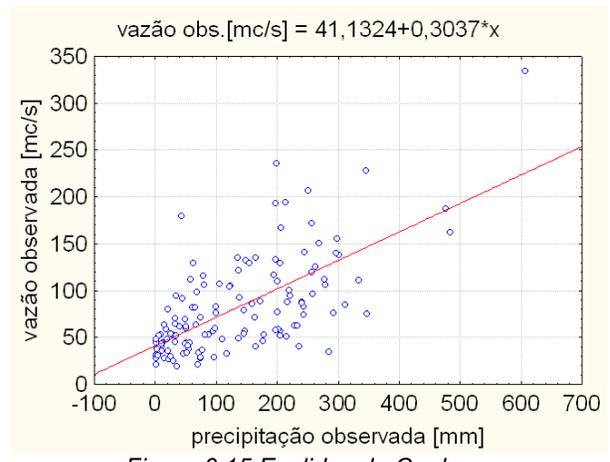


Figura 6.15 Euclides da Cunha.

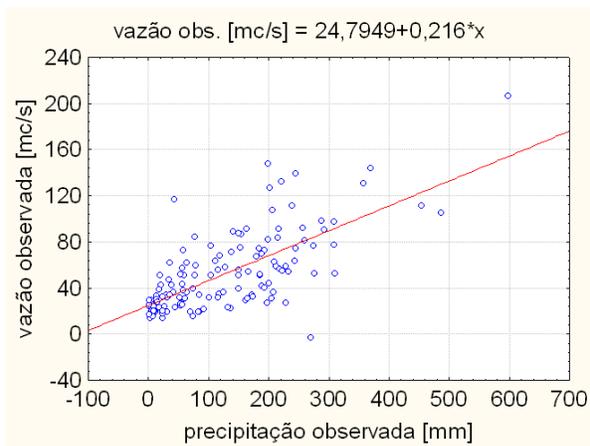


Figura 6.16 Graminha.

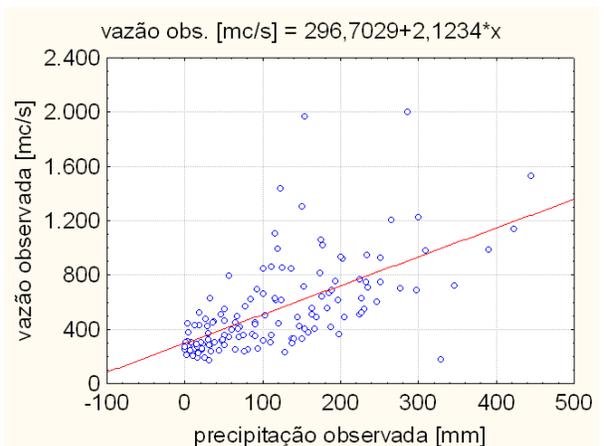


Figura 6.17 Ibitinga.

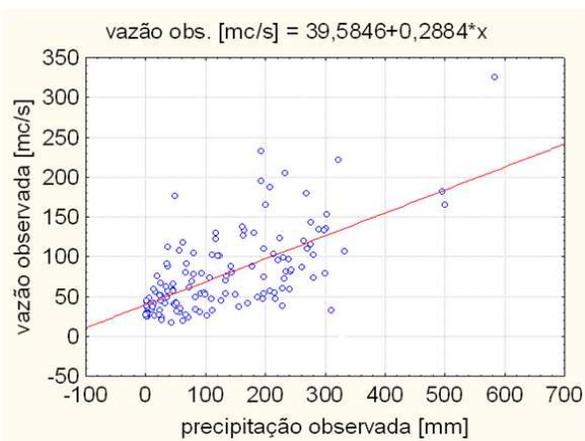


Figura 6.18 Limoeiro.

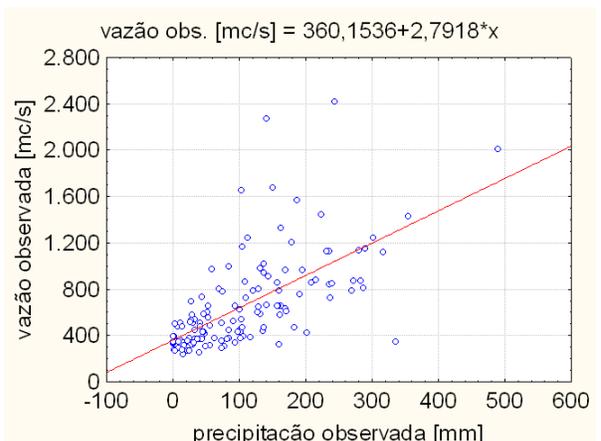


Figura 6.19 Nova Avanhandava.

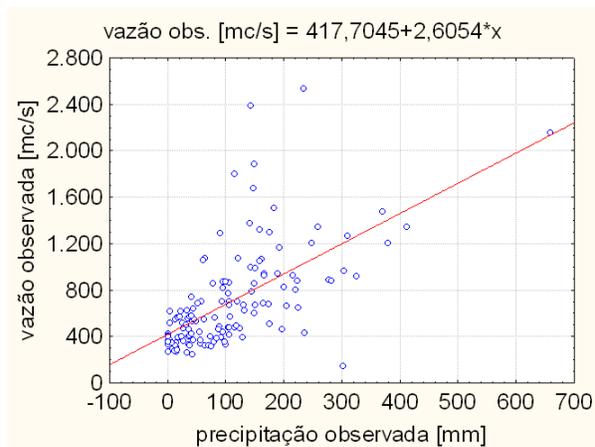


Figura 6.20 Promissão.

Pode-se observar que nos gráficos de dispersão diária, em ocorrência à precipitação igual a zero, a dispersão da vazão é maior do que os gráficos de dispersão mensal. Isto é devido ao fato de que a vazão diária é influenciada muito mais pelas precipitações dos períodos anteriores do que a vazão mensal. Assim, resulta que os valores da vazão mínima das retas de regressão simples são sempre menores nos gráficos de dispersão mensais.

Para incluir a influência das precipitações dos períodos anteriores na análise diária, decidiu-se reelaborar a dispersão em gráficos vazão *versus* precipitação média móvel dos intervalos diários  $t, t-1, t-2, t-3$ , conforme Figuras 6.21 – 6.29.

Figuras 6.21 – 6.29 Gráficos de dispersão diária vazão vs precipitação média móvel

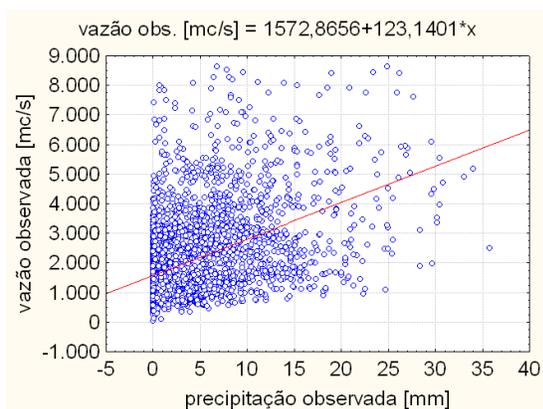


Figura 6.21 Água Vermelha.

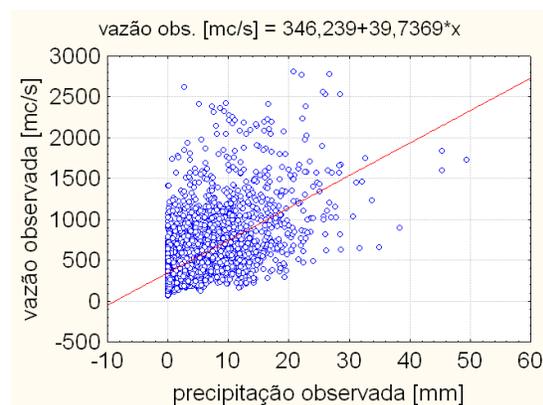


Figura 6.22 Bariri.

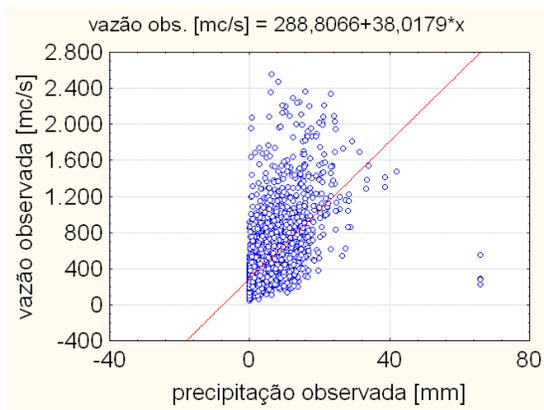


Figura 6.23 Barra Bonita.

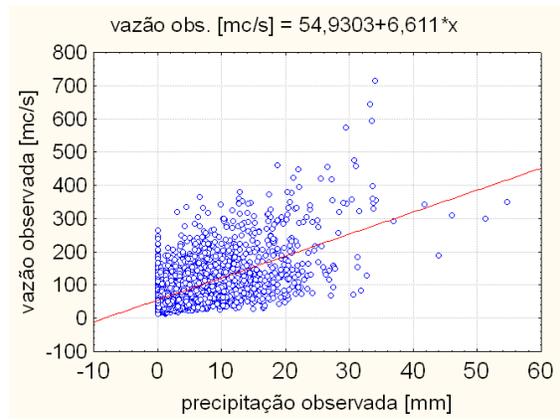


Figura 6.24 Euclides da Cunha.

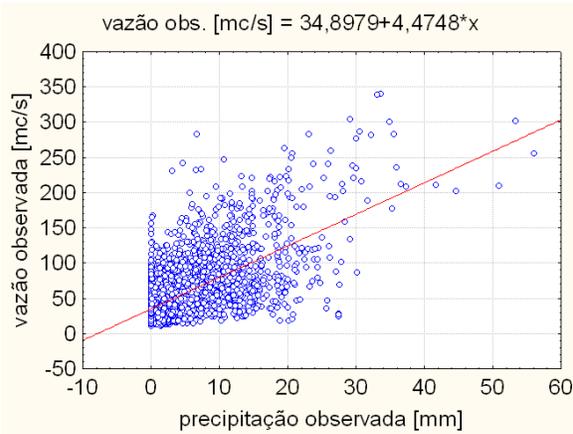


Figura 6.25 Graminha.

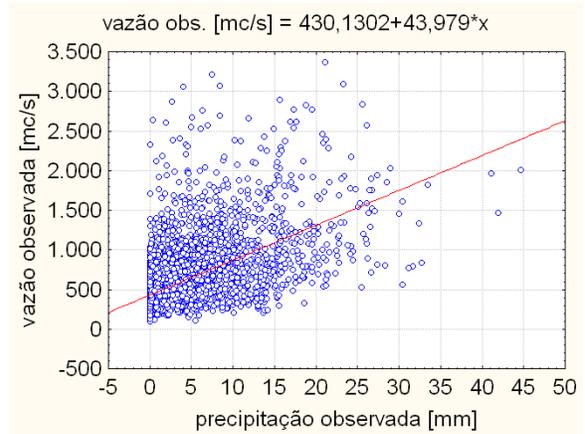


Figura 6.26 Ibitinga.

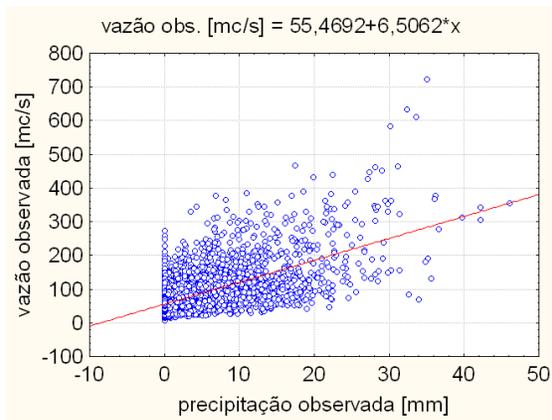


Figura 6.27 Limoeiro.

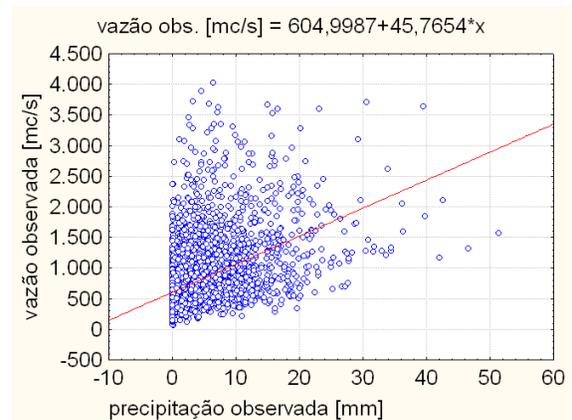


Figura 6.28 Nova Avanhandava.

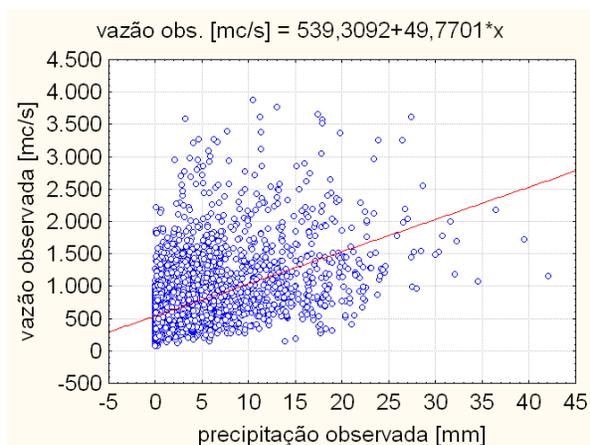


Figura 6.29 Promissão.

Nota-se que as retas de regressão associadas às dispersões diárias com média móvel (Figuras 6.21-6.29), têm valores de vazão mínima que mais se aproximam aos valores das retas de dispersão mensal (Figura 6.30).

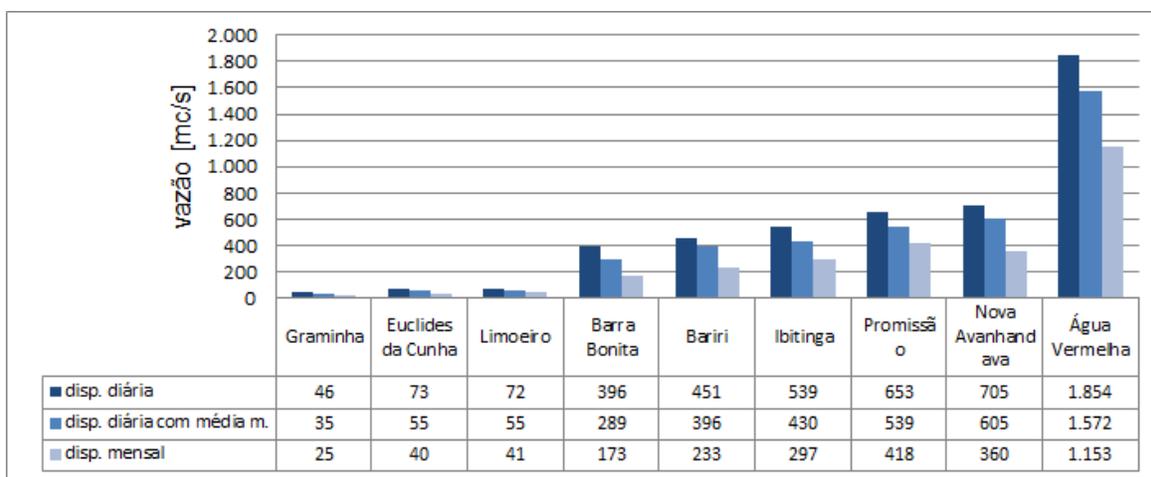


Figura 6.30 Vazões em ocorrência da precipitação igual a zero encontradas pelas retas de RL.

A elaboração dos dados permitiu ordenar as UHEs segundo as vazões registradas, da menor à maior, diárias e mensais, conforme as Figuras 6.31 e 6.32. Nota-se que as vazões são compreendidas em um amplo intervalo, onde a UHE maior, Água Vermelha, tem uma vazão média 50 vezes maior do que a UHE menor, Graminha.

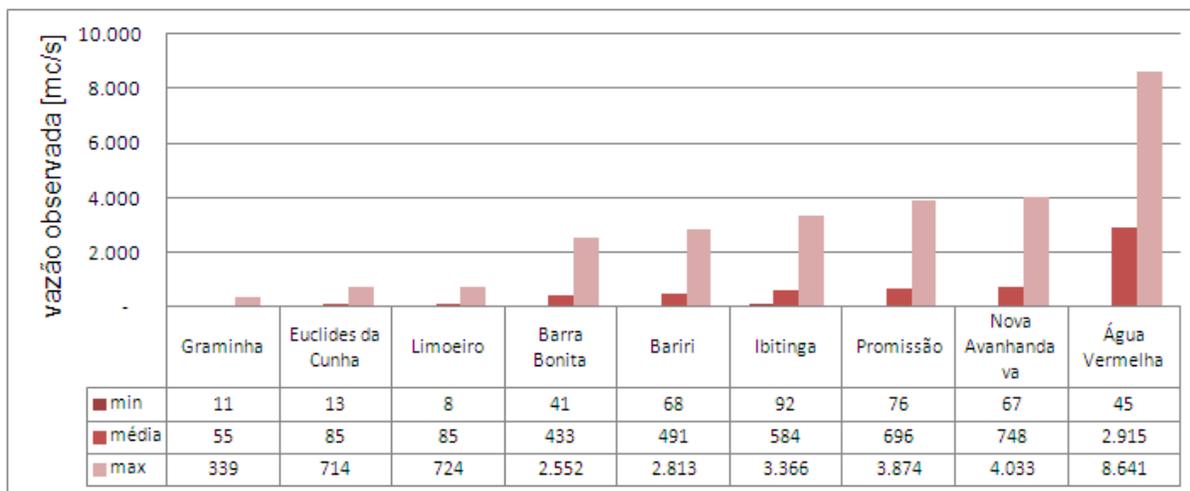


Figura 6.31 Vazão diária observada mínima, média e máxima das usinas em ordem crescente de vazão média.

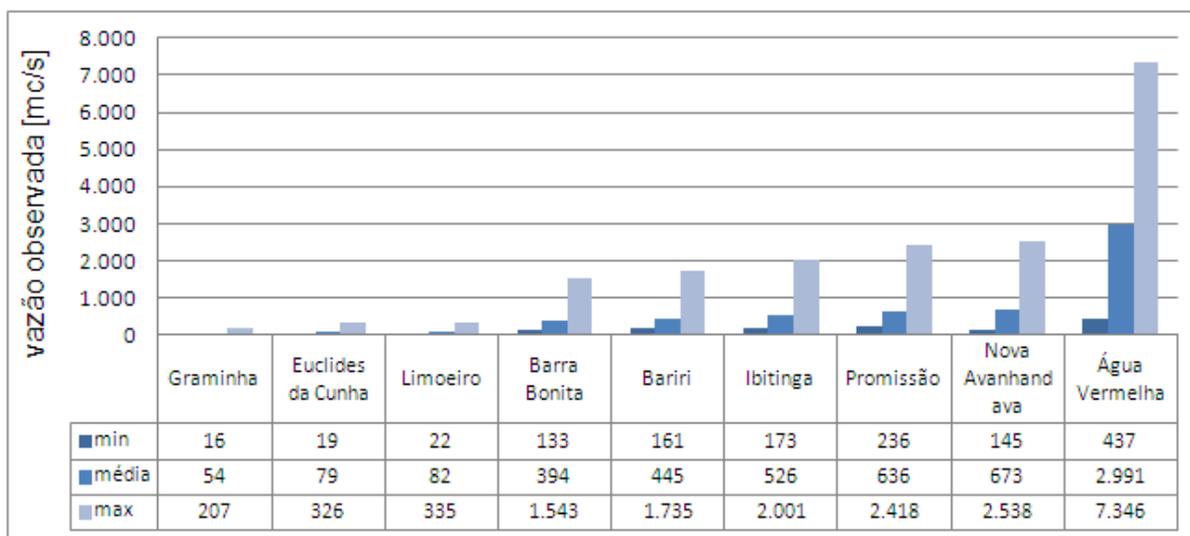


Figura 6.32 Vazão mensal observada mínima, média e máxima das usinas em ordem crescente de vazão média.

Desta análise do banco de dados também extraiu-se as precipitações diárias e mensais, mínima, média e máxima, registradas no período em análise (Figura 6.33). A delimitação dos valores mínimos, médios e máximos em relação às vazões e às precipitações será útil nas fases sucessivas, especialmente na análise de desempenho dos modelos.

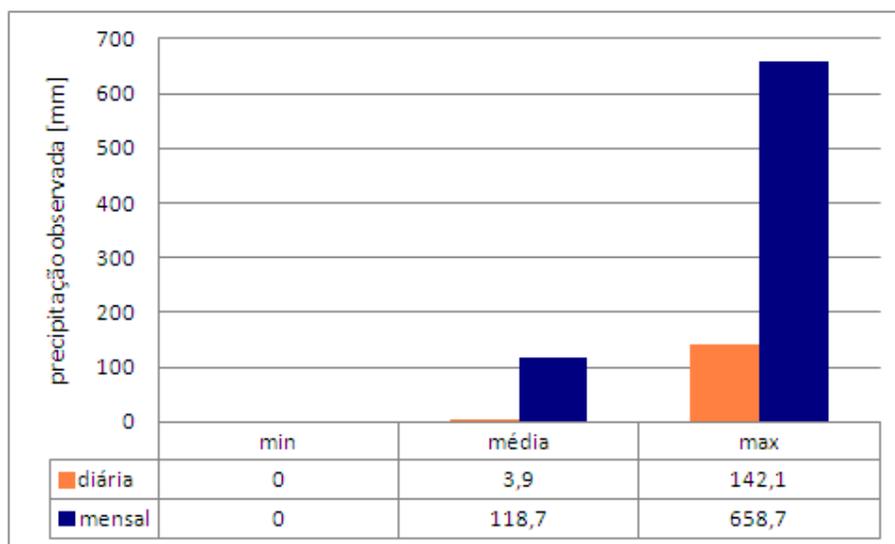


Figura 6.33 Precipitações diárias e mensais observadas.

A última análise desta etapa tem como objetivo representar todas as observações do banco de dados num plano cartesiano vazão-precipitação, a fim de estimar as distribuições percentuais delimitadas em vários quadrantes do plano. Na Figura 6.34, que mostra a distribuição diária das observações, nota-se que 72% das observações estão delimitadas no quadrante com vazão inferior a 1.000 m<sup>3</sup>/s e precipitação inferior a 7 mm/dia. Vazões maiores de 1.000 m<sup>3</sup>/s representam somente 16% das observações totais. Isto não significa que a previsão de vazões maiores de 2.000 m<sup>3</sup>/s tem menor importância, pois a potência de geração de energia elétrica das UHEs é diretamente proporcional à vazão registrada. Por isso, a qualidade de previsão de grandes vazões tem a mesma importância da qualidade de previsão de pequenas vazões, apesar de ter um peso percentual nas observações total muito inferior.

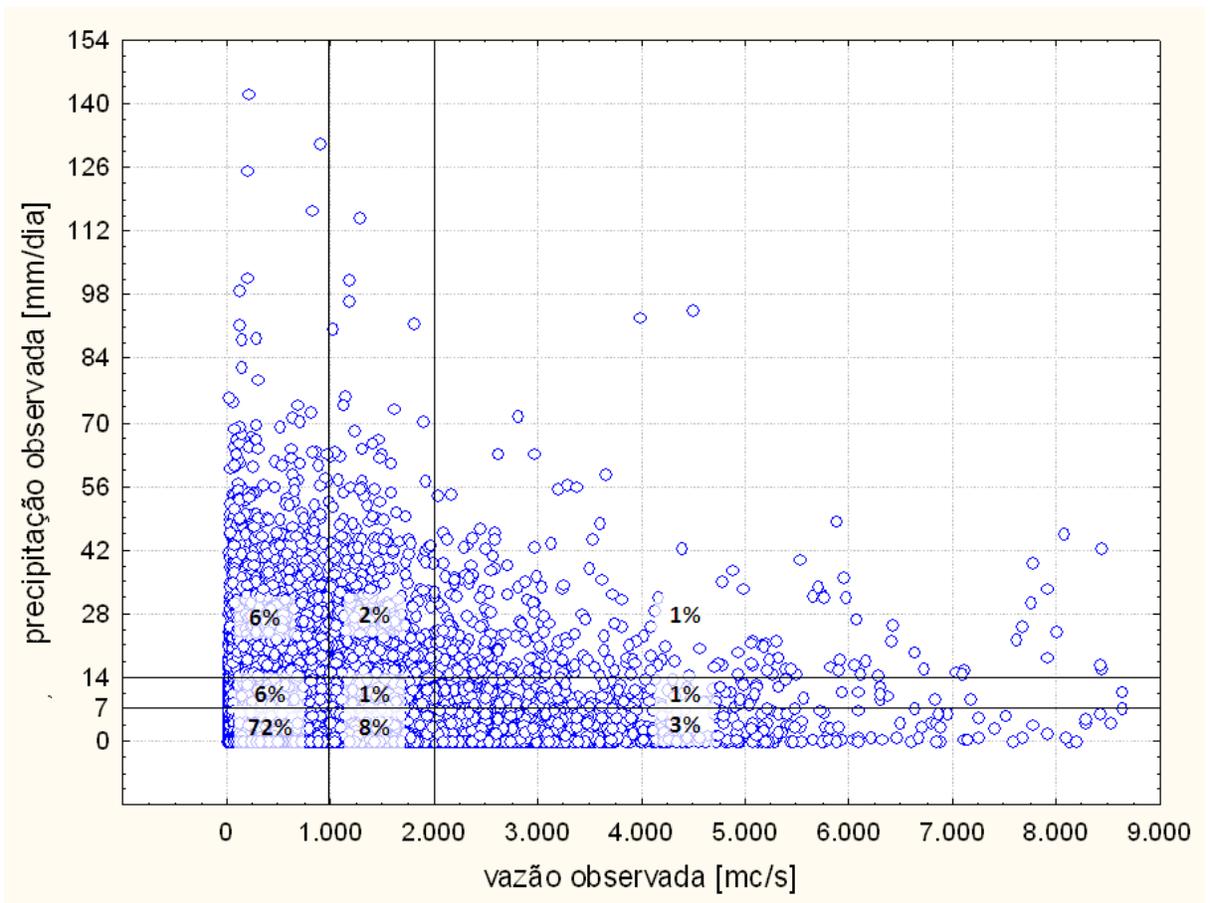


Figura 6.34 Distribuição das observações diárias no plano vazão-precipitação.

A Figura 6.35 mostra a distribuição das observações mensais no plano vazão-precipitação. Pode-se notar que a dispersão das vazões tem a mesma distribuição percentual das observações diárias. De fato, 16% das observações são maiores que 1.000 m<sup>3</sup>/s.

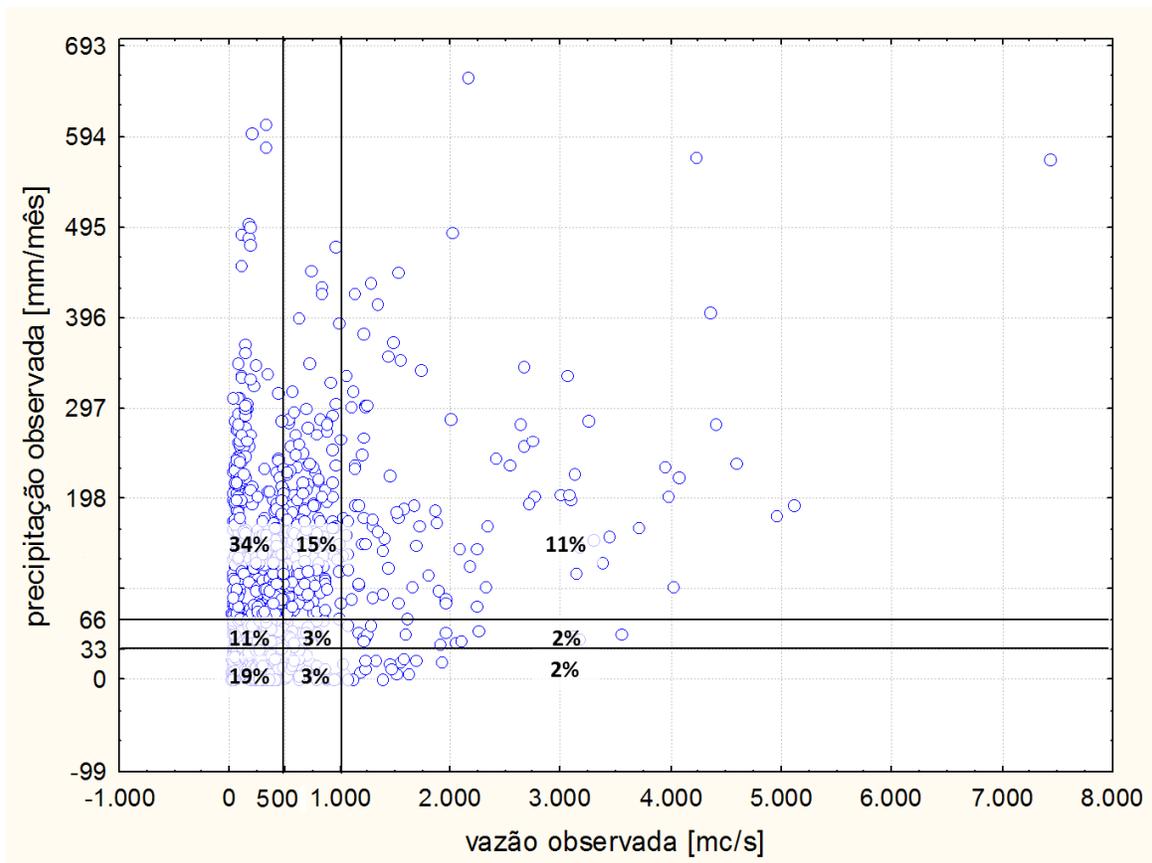


Figura 6.35 Distribuição das observações mensais no plano vazão-precipitação.

### 6.3 Preparação dos Dados

Nesta fase define-se um conjunto, extraído do banco de dados, a partir do qual serão validados os modelos. Este conjunto tem que ser uma amostra representativa das características da totalidade do banco de dados. Assim, o conjunto tem de incluir todas as UHEs e os períodos do ano com diferentes níveis de precipitação. A Figura 6.36 mostra a precipitação mensal observada na UHE Graminha no período 2003-2007. Nota-se que a estação mais chuvosa concentra-se nos meses de verão enquanto o período mais seco nos meses inverniais.

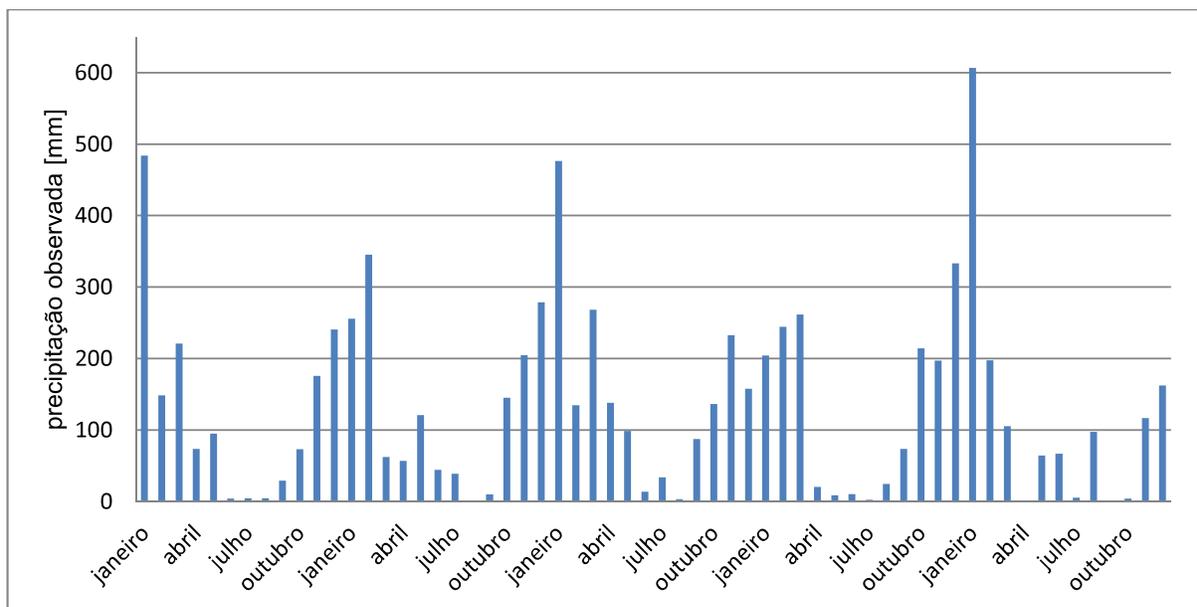


Figura 6.36 Precipitação mensal de chuva medida na UHE Graminha no período 2003-2007.

Considerando a sazonalidade das precipitações, escolhem-se onze períodos para previsão diária e onze períodos para previsão mensal, conforme Tabela 6.5. Os períodos escolhidos, conforme mostra as médias das precipitações observadas, cobrem tanto a estação seca quanto a chuvosa e a intermediária. Sendo nove as usinas em análise, os estudos resultantes serão 99 de previsão diária e 99 de previsão mensal. Sendo quinze os intervalos de tempo para a previsão diária e seis para a previsão mensal, por fim, cada modelo terá que fornecer 1.485 previsões de intervalos de tempo para a previsão diária e 594 para a previsão mensal.

Tabela 6.5 Períodos escolhidos para previsões diárias e mensais.

Previsões diárias		Previsões mensais	
prec. média observada diária nas bacias [mm/dia]		prec. média observada mensal nas bacias [mm/ mês]	
04/02/2002 - 18/02/2002	8,7	11/02 – 04/02	208,8
07/03/2003 - 21/03/2003	10,9	07/03 – 11/03	80,8
03/11/2003 - 17/11/2003	4,6	02/04 – 07/04	91,8
04/04/2004 - 18/04/2004	2,5	12/04 – 05/05	191,7
20/01/2005 - 03/02/2005	11,6	06/05 – 12/05	67,3
12/05/2006 - 26/05/2006	1,1	04/06 – 09/06	27,7
05/08/2006 - 19/08/2006	0,0	11/06 – 05/06	204,0
15/01/2007 - 29/01/2007	14,2	06/07 – 11/07	67,9
15/06/2007 - 29/06/2007	0,1	05/08 – 10/08	47,4
15/01/2008 - 29/01/2008	10,9	03/09 – 08/09	77,2
15/12/2009 - 29/12/2009	8,7	12/09 – 05/10	158,1
média	6,6	média	111,1

## 6.4 Modelagem e Avaliação

Foram executados os modelos de previsão RL e SINF para os períodos escolhidos e produzidos os resultados, ou seja, as previsões de vazões diárias e mensais de cada intervalo de tempo  $t$  que compõe cada período em análise. O modelo SINF forneceu resultados anômalos para previsões diárias de Água Vermelha no período 15–29/12/2009, de Ibitinga no período 15–29/01/2008, e de Promissão no período 20/01 – 03/02/2005. Valores anômalos do modelo SINF também foram encontrados em relação a todas as previsões mensais de Limoeiro e Nova Avanhandava, além de Água Vermelha nos períodos de 05-10/2008 e 03-08/2009. Como resultados anômalos, entendem-se vazões previstas negativas iguais, à zero ou positivas fora de escala (quatro ordens de grandeza maiores). Excluindo os estudos onde se verificaram valores anômalos, os períodos a serem avaliados diminuem a 96 diários e 75 mensais e os intervalos de tempo para previsão diária diminuem a 1.440 e para previsão mensal a 450.

### 6.4.1 Dados Diários

#### Modelagem

No Apêndice A, encontram-se os 96 gráficos dos resultados de todos os períodos escolhidos que descrevem os 1.440 intervalos de tempo observados e previstos pelos modelos, quinze para cada estudo de previsão. A figura 6.37 é uma síntese, pois reporta todos estes pontos num gráfico de dispersão vazão prevista *versus* vazão observada. Pode-se notar que os dois modelos tem uma previsão no complexo sub-estimada. De fato, as linhas de tendência das retas de regressão linear tem inclinação inferior à tendência que descreve uma previsão perfeita  $y = x$ .

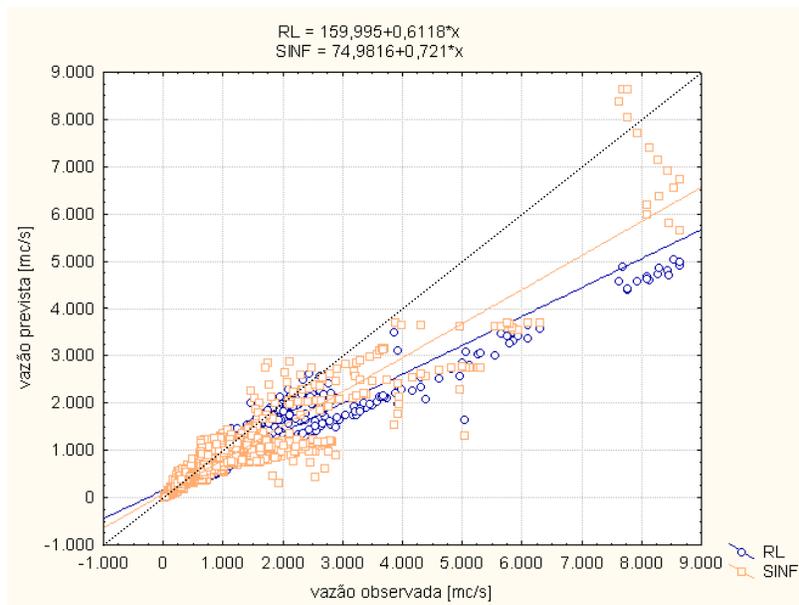


Figura 6.37 Gráfico de dispersão vazão prevista vs vazão observada das previsões diárias.

A Figura 6.38 mostra uma elaboração da Figura 6.37: a distribuição da frequência do erro resíduo percentual das previsões diárias dos modelos RL e SINF, utilizando como tentativa de modelagem a distribuição Gaussiana ou Normal. Em geral, confirma-se que as duas distribuições normalizadas têm a tendência a subestimar a previsão. De fato, a média  $\mu$  da distribuição SINF é melhor, com -5,51%, pois está mais próxima ao zero em comparação à distribuição RL com -7,12%. O desvio padrão  $\sigma$  da distribuição normalizada RL, com 16,02% é melhor, pois menor que o desvio padrão da distribuição normalizada SINF, com 29,63%.

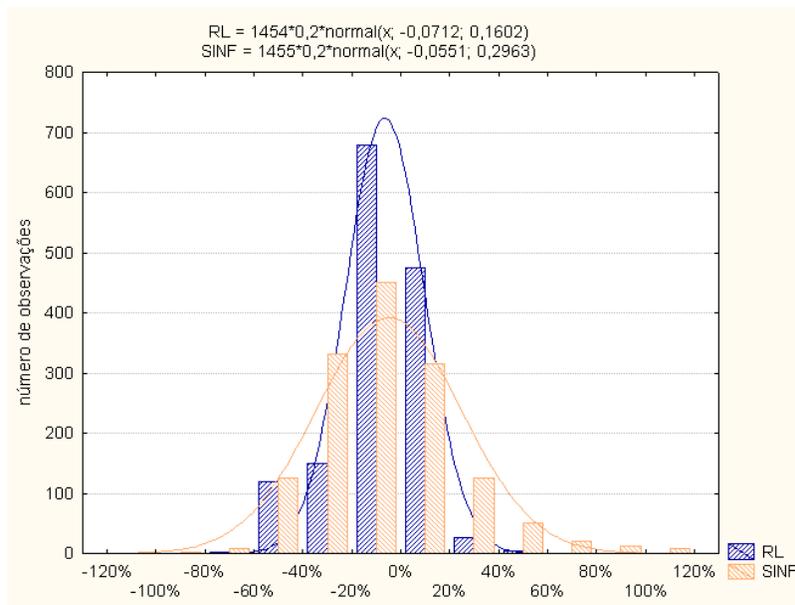


Figura 6.38 Distribuição da frequência do erro resíduo percentual nas previsões diárias.

### Avaliação

Esta fase prevê a análise dos 96 estudos diários com os índices de desempenho mostrados na sub-seção 5.2. O Apêndice B apresenta todos os onze índices resultantes para cada estudo. Nesta seção, escolheu-se aprofundar a análise com o índice MAPE para avaliar a qualidade, e com o índice CI para avaliar a tendência das previsões.

O índice MAPE é muito usado na literatura e, por se tratar de um valor absoluto, será útil na avaliação das diferentes UHEs. Este índice é adequado para avaliar previsões de curto prazo, como as diárias. Será usada a classificação de Liew *et al.* (2007), que define a qualidade da previsão para diferentes valores de MAPE (inadequata, satisfatória, boa, muito boa).

O índice CI será usado para avaliar a tendência da previsão, independentemente do erro cometido. Acredita-se que a análise da tendência seja uma apreciação de fundamental importância para avaliar uma previsão.

As figuras 6.39 e 6.40 mostram os gráficos tridimensionais das dispersões MAPE para os modelos RL e SINF, em ocorrência das vazões e precipitações médias registradas no período em exame. O gráfico do modelo RL, entre o MAPE de 0% a 25%, apresenta uma maior concentração de pontos com vazão inferior a 1000 m<sup>3</sup>/s e precipitação distribuída. O gráfico do modelo SINF, no que tange ao MAPE menor de 25%, tem uma dispersão dos pontos mais distribuída com presença de vazões maior de 1.000 m<sup>3</sup>/s.

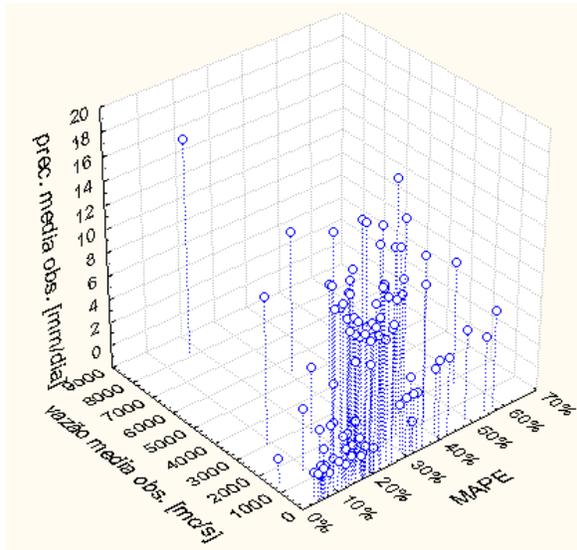


Figura 6.39 Grafico de dispersão MAPE – vazão observada- precipitação observada do modelo RL.

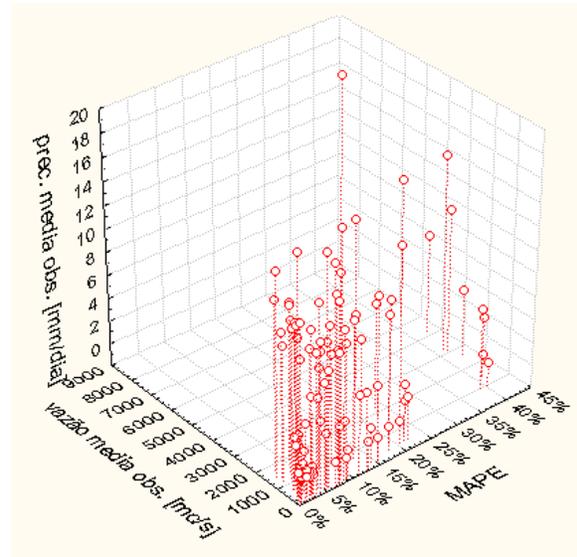


Figura 6.40 Grafico de dispersão MAPE – vazão observada- precipitação observada do modelo SINF.

A Figura 6.41 apresenta a repartição percentual do índice MAPE segundo a classificação de Liew *et al.* (2007) para cada UHE objeto deste estudo. As UHEs são ordenadas por vazão média. O resultado global penaliza as previsões SINF, pois indica que 10% das previsões RL contra 36% das SINF são inadequadas; 65% das RL contra 16% das SINF são muito boas.

Avaliando os resultados para cada UHE, o modelo de previsão SINF resulta melhor só para a UHE Água Vermelha, aquela com maior vazão. O modelo RL tem uma avaliação de previsão muito boa, maior de 50%, para seis UHEs, os resultados melhores do modelo SINF encontram-se na UHE Promissão com 20% de avaliação muito boa e na UHE Água Vermelha com 30% de avaliação muito boa.

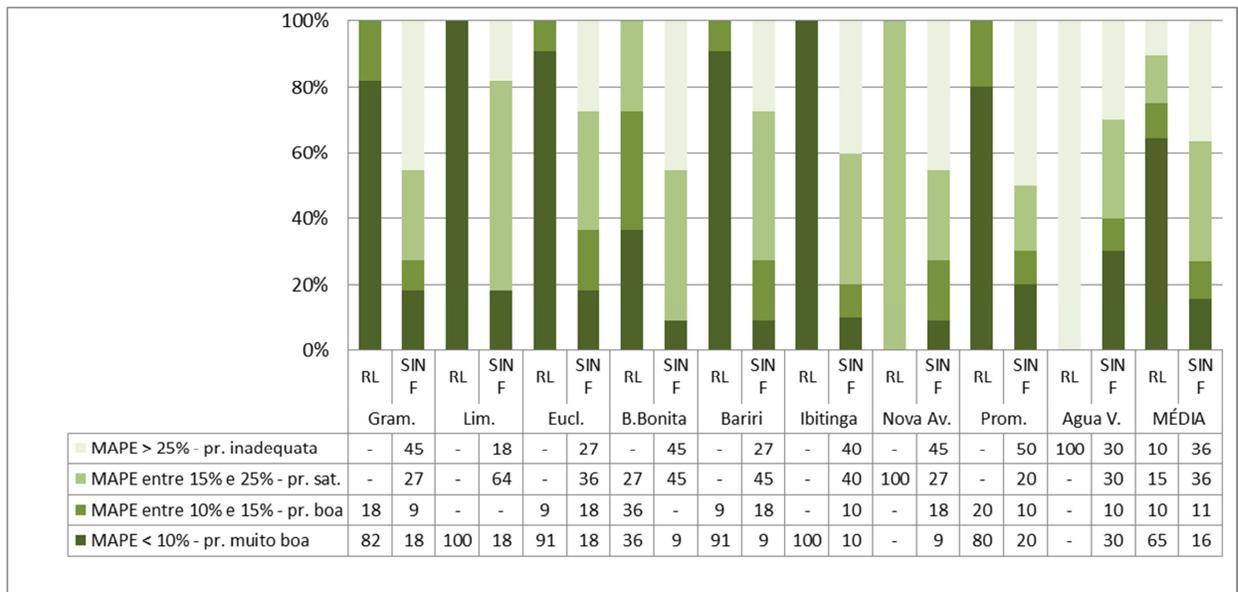


Figura 6.41 Repartição percentual do índice MAPE nos estudos diários.

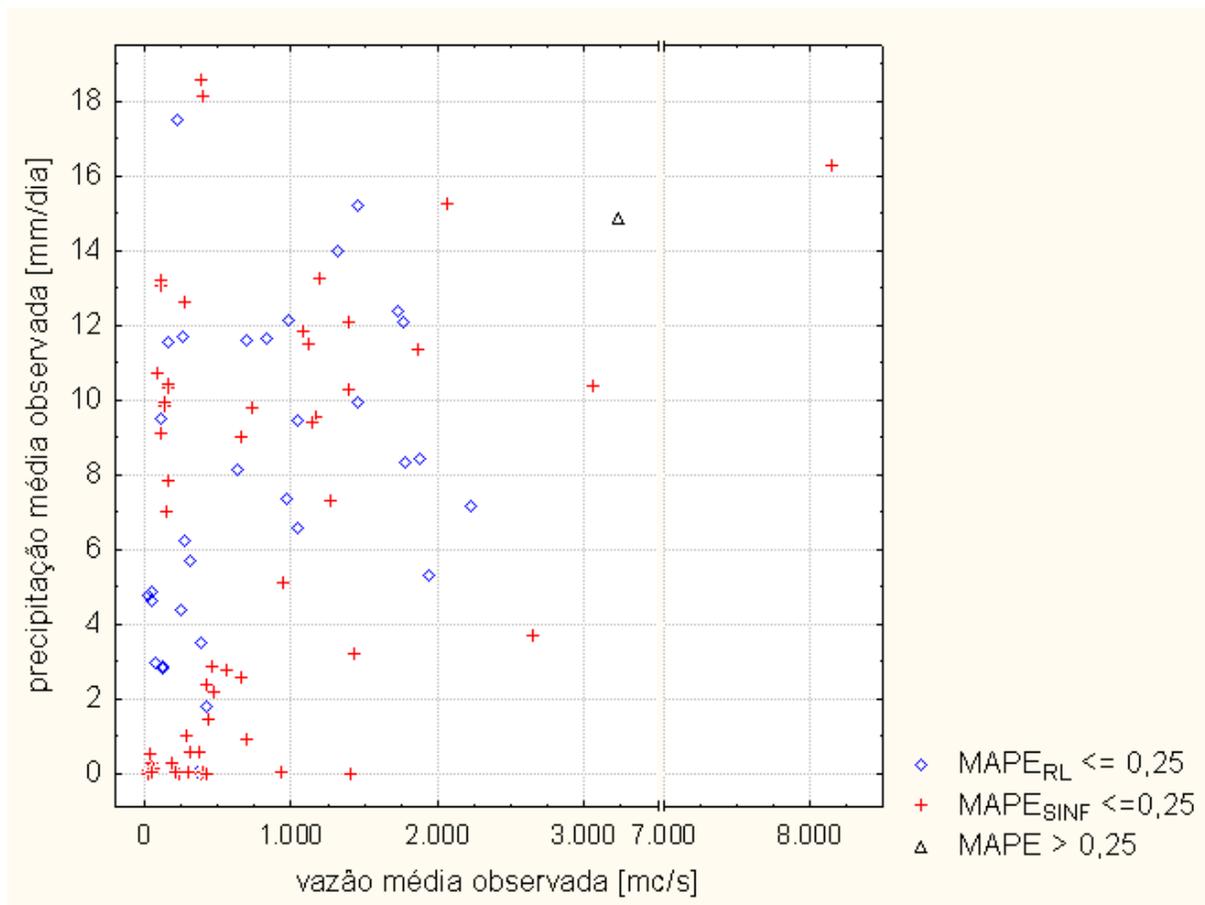


Figura 6.42 Gráfico de dispersão precipitação observada vs vazão observada, com indicação do MAPE.

A Figura 6.42 é um gráfico de dispersão-precipitação média observada *versus* vazão média observada nos 96 períodos objeto de estudo. A legenda indica, também, para cada ponto, se os modelos RL e/ou SINF têm um índice  $MAPE \leq 25\%$ , que corresponde a uma previsão adequada, ou se tem um índice  $MAPE > 25\%$ , que corresponde a uma previsão inadequada. Observa-se a complementaridade dos modelos, ou seja, só 3% dos pontos do gráfico não pertencem a uma previsão feita pelo RL ou pelo SINF com índice MAPE inadequado. Isto significa que os dois modelos têm desempenho de previsão que muda com as condições das variáveis chuva-vazão de um jeito complementar: o bom desempenho de um corresponde ao mau desempenho do outro e vice-versa.

A tabela 6.6 é uma elaboração da Figura 6.42. A ideia é a de dividir o plano cartesiano em nove quadrantes e fazer uma análise de cada porção do plano, a fim de individualizar o desempenho dos modelos com diferentes condições de chuva-vazão. Os percentuais referem-se à razão entre o número de previsões com  $MAPE \leq 25\%$  e o número de previsões totais no quadrante avaliado. Como analisado na Figura 6.41, o modelo RL tem um desempenho complexo melhor nas vazões menores de  $1.000 \text{ m}^3/\text{s}$ . Observa-se que, para vazão maior de  $2.000 \text{ m}^3/\text{s}$ , o modelo SINF tem um desempenho melhor do que o modelo RL. Em particular, ressalta-se que a maior diferença de desempenho está com precipitação média menor que  $7 \text{ mm}/\text{dia}$ , ou seja, com vazões elevadas e médio-baixo nível de precipitação.

Tabela 6.6 Distribuição percentual de MAPE considerado adequado.

	vazão < 1000 [m <sup>3</sup> /s]		vazão entre 1000 e 2000 [m <sup>3</sup> /s]		vazão > 2000 [m <sup>3</sup> /s]		vazão > 0 [m <sup>3</sup> /s]	
	RL	SINF	RL	SINF	RL	SINF	RL	SINF
precipitação > 14 [mm/dia]	100%	67%	100%	0%	33%	67%	71%	57%
prec. entre 7 e 14 [mm/dia]	100%	62%	100%	56%	33%	33%	95%	58%
precipitação < 7 [mm/dia]	95%	72%	50%	50%	0%	50%	88%	69%
precipitação > 0 [mm/dia]	97%	69%	90%	52%	25%	50%	90%	64%

Dos 96 estudos diários avaliados, somente oito se colocam na faixa com vazão maior de  $2.000 \text{ m}^3/\text{s}$ : dois no quadrante com precipitação menor de  $7 \text{ mm}/\text{dia}$ , três entre  $7$  e  $14 \text{ mm}/\text{dia}$  e três no quadrante com precipitação maior de  $14 \text{ mm}/\text{dia}$ . Assim, para conferir confiabilidade à hipótese de que o modelo SINF tem um desempenho melhor com vazões maiores de  $2.000 \text{ m}^3/\text{s}$ , decidiu-se avaliar trinta outros estudos para vazões maiores que  $2.000 \text{ m}^3/\text{s}$ : dez para cada quadrante desta coluna.

A seleção dos períodos foi possível através da análise do banco de dados com Excel, filtrando para vazões maiores de 2.000 m<sup>3</sup>/s períodos com precipitações médias entre os limites estabelecidos.

Tabela 6.7 Avaliação do MAPE em períodos com vazões maiores de 2.000 m<sup>3</sup>/s.

	UHE	data início	data fim	prec. méd. obs [mm/dia]	vaz. média obs. [m <sup>3</sup> /s]	MAPE RL	MAPE SINF
prec. média obs. > 14 [mm/dia]	Água Vermelha	03/01/2003	18/01/2003	15,8	2.926	42%	12%
	Água Vermelha	18/01/2003	02/02/2003	22,8	5.194	42%	9%
	Água Vermelha	16/01/2005	31/01/2005	14,6	4.885	42%	17%
	Água Vermelha	06/02/2006	21/02/2006	14,2	3.063	42%	12%
	Promissão	01/01/2007	16/01/2007	22,2	2.272	6%	35%
	Nova Avanhandava	01/01/2007	16/01/2007	28,9	2.438	25%	28%
	Nova Avanhandava	06/01/2007	21/01/2007	24,6	2.577	21%	14%
	Água Vermelha	13/01/2007	28/01/2007	15,4	8.021	42%	22%
	Água Vermelha	19/01/2008	03/02/2008	15,8	4.162	42%	53%
	Bariri	24/01/2010	08/02/2010	15,5	2.166	5%	30%
7 <= prec. média obs. <= 14 [mm/dia]	Água Vermelha	08/01/2002	23/01/2002	9,0	3.262	41%	10%
	Água Vermelha	27/02/2003	14/03/2003	8,1	2.591	41%	7%
	Água Vermelha	28/12/2004	12/01/2005	8,2	3.689	41%	13%
	Nova Avanhandava	20/01/2005	04/02/2005	7,1	2.223	23%	43%
	Água Vermelha	21/12/2006	05/01/2007	13,0	3.556	43%	12%
	Água Vermelha	25/01/2008	09/02/2008	12,5	5.066	42%	14%
	Água Vermelha	08/03/2008	23/03/2008	10,4	4.598	43%	47%
	Água Vermelha	21/01/2009	05/02/2009	9,4	5.137	44%	63%
	Ibitinga	20/01/2010	04/02/2010	8,3	2.653	4%	45%
	Nova Avanhandava	22/01/2010	06/02/2010	9,2	3.182	20%	44%
prec. média obs. < 7 [mm/dia]	Água Vermelha	19/02/2003	06/03/2003	5,4	3.222	41%	9%
	Água Vermelha	07/03/2002	22/03/2002	1,8	2.753	42%	5%
	Água Vermelha	11/01/2004	26/01/2004	4,3	3.035	42%	18%
	Água Vermelha	21/03/2004	05/04/2004	0,4	2.416	42%	5%
	Água Vermelha	21/05/2004	05/06/2004	1,9	2.174	42%	6%
	Água Vermelha	07/02/2005	22/02/2005	0,3	3.687	41%	14%
	Água Vermelha	05/04/2007	20/04/2007	1,6	2.243	42%	3%
	Água Vermelha	10/04/2008	25/04/2008	4,8	3.437	41%	63%
	Água Vermelha	14/03/2009	29/03/2009	5,5	3.358	42%	16%
	Água Vermelha	20/04/2009	05/05/2009	0,3	2.207	40%	85%

A Tabela 6.7 mostra os novos períodos avaliados, com indicação da UHE, data de início e de fim do período, precipitação média observada, vazão média observada e índices MAPE para cada modelo de previsão.

A figura 6.43 é uma síntese da Tabela 6.7, apresentando a repartição percentual do índice MAPE, segundo a classificação de Liew *et al.* (2007), para cada quadrante individualizado na faixa com vazões maiores de 2.000 m<sup>3</sup>/s. Os resultados confirmam a hipótese elaborada. Observa-se que o modelo SINF sempre tem índices de desempenhos melhores do que o modelo RL. Em particular, destaca-se que o resultado melhor do modelo SINF corresponde ao resultado pior do modelo RL e coloca-se no quadrante com precipitações médias observadas menores de 7 mm/dia. Neste caso, o modelo SINF tem uma percentual de inadequação igual a 20% e uma percentual com avaliação muito boa igual a 50%, contra 100% de inadequação do modelo RL.

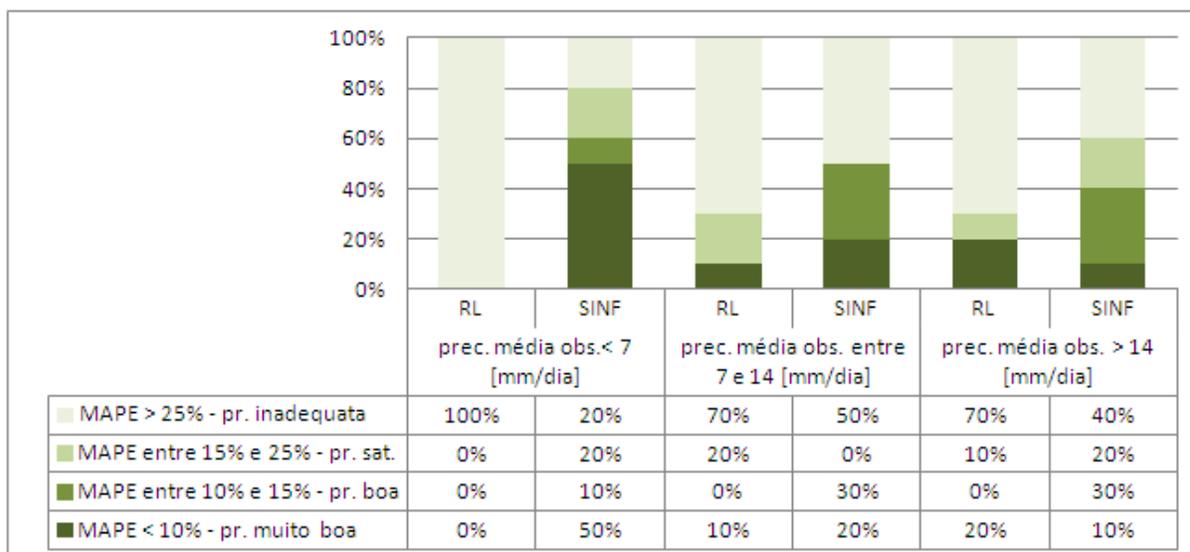


Figura 6.43 Repartição percentual do índice MAPE na análise dos estudos diários com vazão média maior que 2.000 m<sup>3</sup>/s.

A Figura 6.44 avalia o CI para cada UHE nos 96 estudos. A comparação é sempre favorável ao modelo RL, tendo para todas as UHEs valores inferiores de CI.

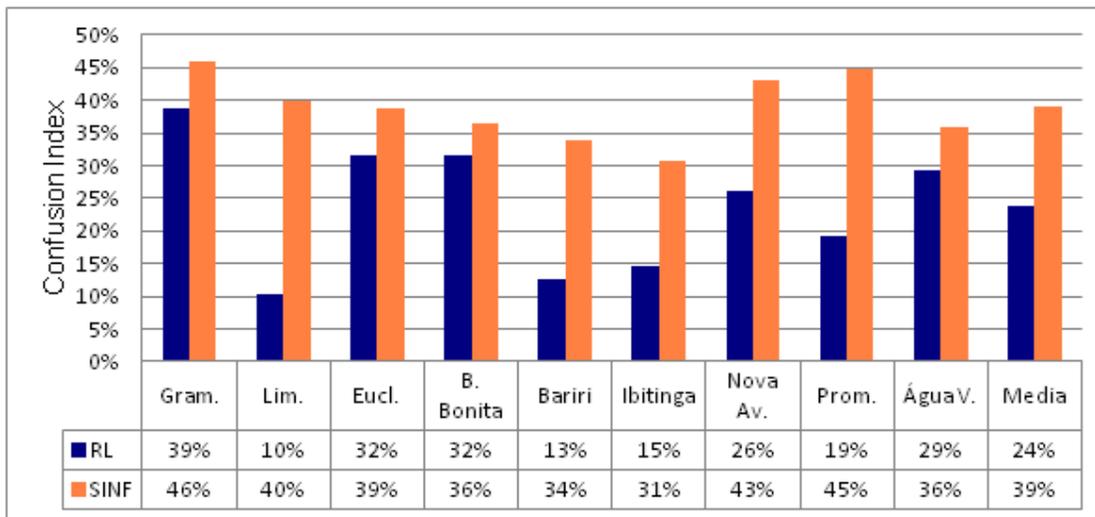


Figura 6.44 Confusion Index para as previsões diárias.

## 6.4.2 Dados mensais

### Modelagem

No Apêndice C encontram-se os 75 gráficos dos resultados de todos os períodos escolhidos que descrevem os 540 intervalos de tempo observados e previstos pelos modelos, seis para cada estudo de previsão. A Figura 6.45 reporta todos estes pontos num gráfico de dispersão-vazão prevista *versus* vazão observada. Pode-se notar que os dois modelos têm uma previsão em geral subestimada. De fato, as linhas de tendência das retas de regressão linear têm inclinação inferior à tendência que descreve uma previsão perfeita  $y = x$ .

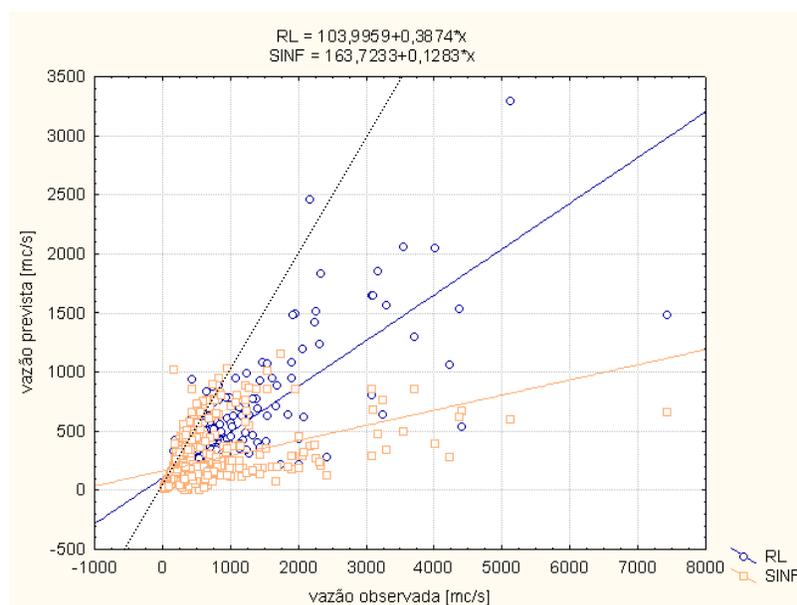


Figura 6.45 Gráfico de dispersão vazão prevista vs vazão observada das previsões mensais.

A Figura 6.46 mostra a distribuição da frequência do erro resíduo percentual das previsões diárias dos modelos RL e SINF. Neste gráfico foi utilizada, como tentativa de modelagem destas distribuições empíricas, a distribuição Gaussiana ou Normal. Em geral, confirma-se a tendência a subestimar as previsões para as duas distribuições normalizadas. De fato, a média  $\mu$  da distribuição RL é melhor, com -25,84%, pois mais próxima do zero do que a distribuição SINF, com -30,02%. O desvio padrão  $\sigma$  da distribuição normalizada RL, com 32,56%, é melhor, pois menor que o desvio padrão da distribuição normalizada SINF, com 51,88%.

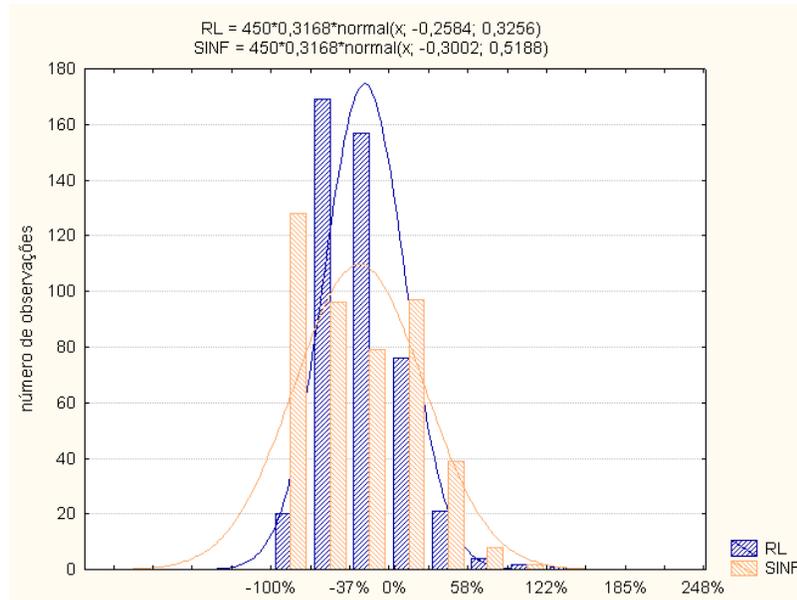


Figura 6.46 Distribuição da frequência do erro resíduo percentual nas previsões mensais.

### Avaliação

Esta fase prevê a análise dos 75 estudos mensais com os índices de desempenho mostrados na sub-seção 5.2. O Apêndice D apresenta todos os 11 índices resultantes para cada estudo. Nesta seção, escolheu-se aprofundar a análise com o índice NS para avaliar a qualidade, e com o índice CI, para avaliar a tendência das previsões.

O índice NS é muito usado na literatura e, por se tratar de um valor absoluto, será útil na avaliação das diferentes UHEs. Este índice é adequado para avaliar previsões de longo prazo como as mensais. Além disso, será interessante comparar os dois modelos também com este preditor de referência, ou seja, a média semestral das vazões. Por fim, na análise será usada a classificação de Santhi *et al.* (2001) que definem para diferentes valores de NS a qualidade da previsão.

As figuras 6.47 e 6.48 mostram os gráficos tridimensionais das dispersões NS, vazão média observada no período em exame e precipitação média observada para os modelos RL e SINF.

Nestas figuras, nota-se uma concentração de pontos mais acentuada no modelo RL. Os dois gráficos apresentam a maior concentração com NS entre 1 e -4, com vazão entre 2.000 m<sup>3</sup>/s e precipitação entre 30 mm/mês.

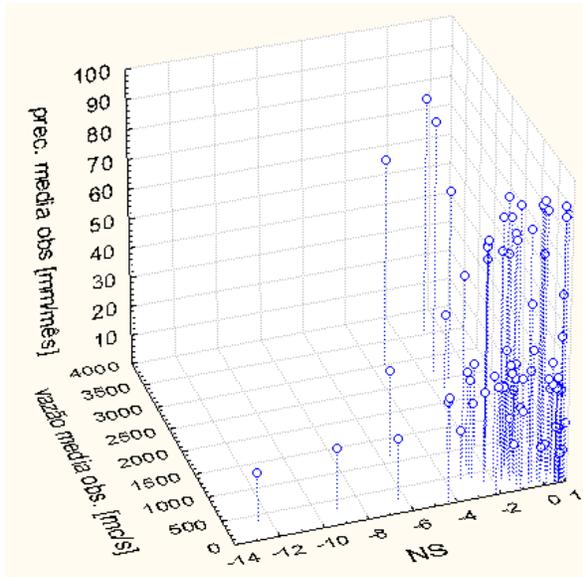


Figura 6.47 Gráfico de dispersão NS – vazão observada- precipitação observada do modelo RL.

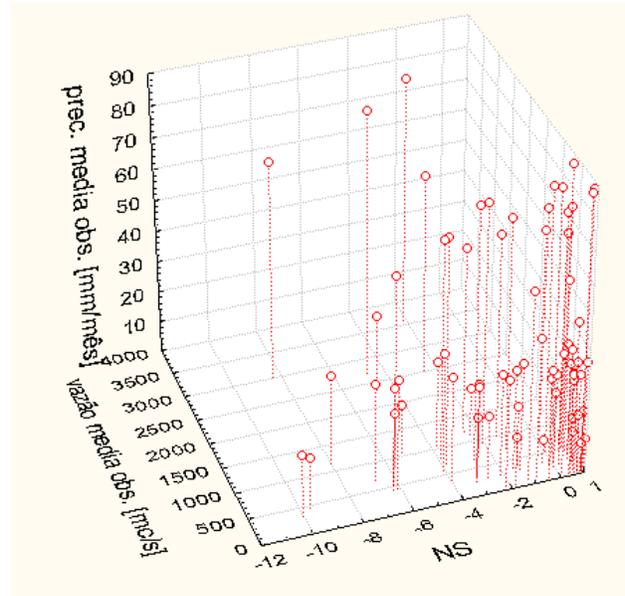


Figura 6.48 Gráfico de dispersão NS – vazão observada- precipitação observada do modelo SINF.

A figura 6.49 apresenta a comparação dos modelos com o preditor de referência, ou seja, a média das observações no semestre considerado: 65% dos valores do modelo RL e 60% dos valores do SINF tem NS<0. Isto significa que mais da metade das previsões dos dois modelos têm um desempenho pior em comparação ao preditor de referência.

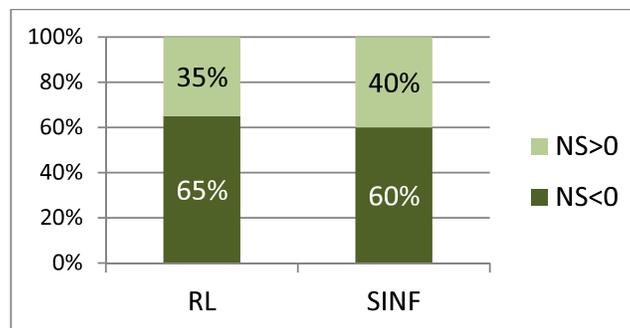


Figura 6.49 Comparação dos modelos com o preditor de referência.

A Figura 6.50 apresenta a repartição percentual do índice NS segundo a classificação de Santhi *et al.* (2001) para cada UHE objeto deste estudo. As UHEs são ordenadas por vazão média. O resultado global penaliza as previsões RL, pois indica que 77% das previsões RL contra 73% das SINF são inadequadas; 16% das RL contra 23% das SINF são muito boas.

Avaliando os resultados para cada UHE, o modelo de previsão RL resulta melhor para a UHE Graminha e Barra Bonita, enquanto o modelo SINF resulta melhor para a UHE Euclides da Cunha e Bariri. Nas previsões de vazões das UHEs Ibitinga, Promissão e Água Vermelha, as UHEs com vazões maiores, os dois modelos tem 100% de inadequação. Além disso, o RL tem uma avaliação de previsão muito boa, maior que 50% só para a UHE Graminha, enquanto os resultados mais performantes do modelo SINF encontram-se na UHE Graminha e Bariri, com 55% de avaliação muito boa.

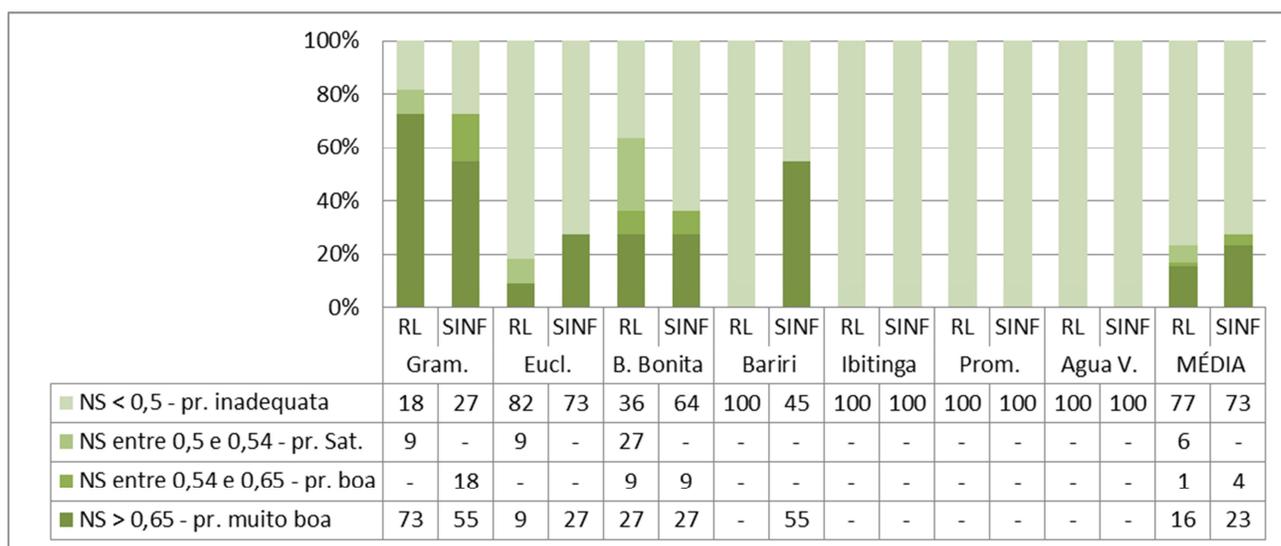


Figura 6.50 Repartição percentual do índice NS nos estudos mensais.

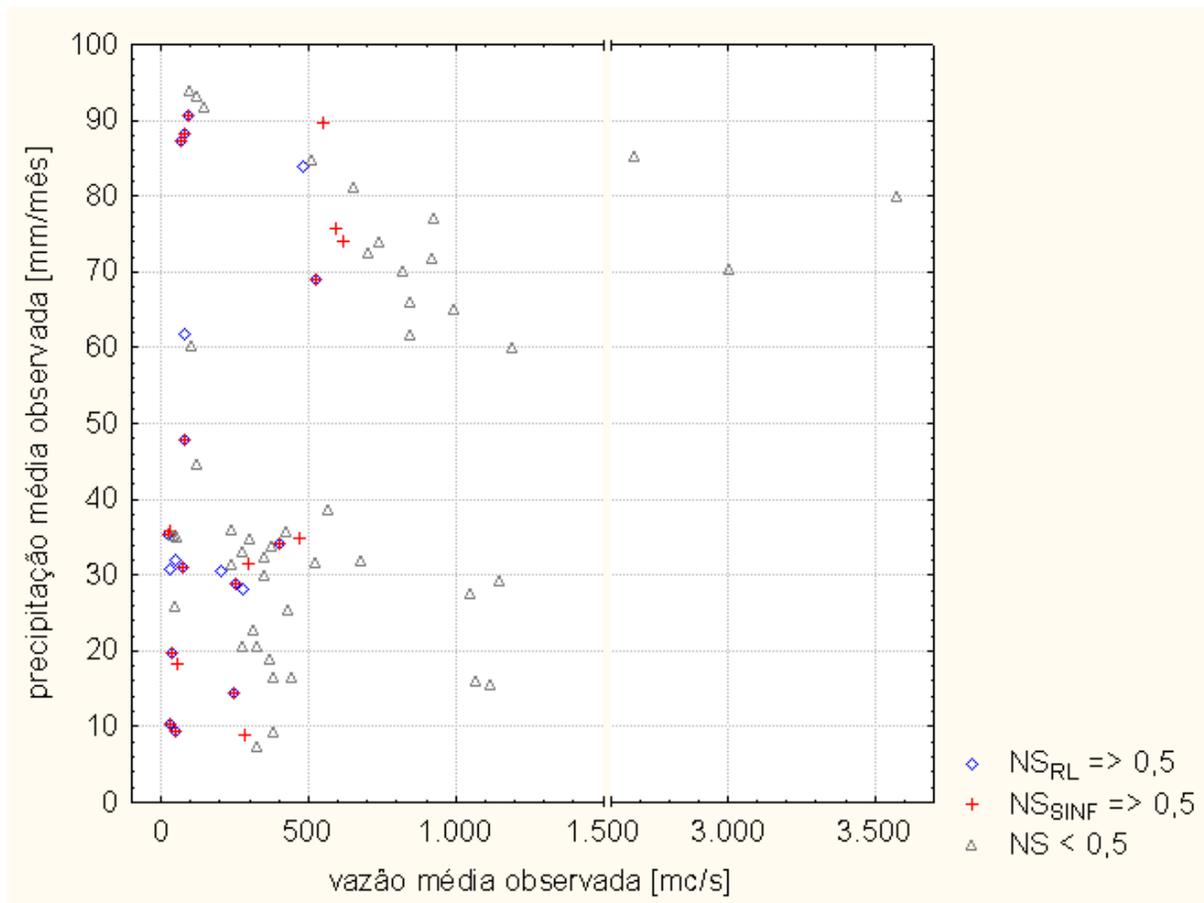


Figura 6.51 Gráfico de dispersão precipitação observada vs vazão observada, com indicação do NS.

A Figura 6.51 é um gráfico de dispersão precipitação média observada *versus* vazão média observada nos 75 períodos objeto de estudo. A legenda indica, para cada ponto, se os modelos RL e/ou SINF têm um índice  $NS \geq 0,5$ , que corresponde a uma previsão adequada, ou se têm um índice  $NS < 0,5$ , que corresponde a uma previsão inadequada. Observa-se que, para a maioria dos pontos, os dois modelos tem  $NS \geq 0,5$  e que a concentração de  $NS \geq 0,5$  encontra-se com vazão média  $< 500 \text{ m}^3/\text{s}$  para qualquer nível de precipitação.

Na tabela 6.8 elaborou-se a Figura 6.51 dividindo o plano cartesiano em nove quadrantes, a fim de avaliar cada porção do plano. As percentuais referem-se à razão entre o número de previsões com  $NS \geq 0,5$  e o número de previsões totais no quadrante avaliado.

No quadrante com vazão entre 500 e 1.000  $\text{m}^3/\text{s}$  e precipitação maior de 66 mm/mês, destaca-se a maior diferença de performance dos modelos a favor do modelo SINF.

Tabela 6.8 Distribuição percentual de NS considerado adequado.

	vazão < 500 [m <sup>3</sup> /s]		vazão entre 500 e 1000 [m <sup>3</sup> /s]		vazão > 1000 [m <sup>3</sup> /s]		vazão > 0 [m <sup>3</sup> /s]	
	RL	SINF	RL	SINF	RL	SINF	RL	SINF
precipitação > 66 [mm/mês]	57%	43%	9%	36%	0%	0%	24%	33%
prec. entre 33 e 66 [mm/mês]	29%	36%	0%	0%	0%	0%	20%	25%
precipitação < 33 [mm/mês]	37%	33%	0%	0%	0%	0%	29%	26%
precipitação > 0 [mm/mês]	38%	35%	6%	24%	0%	0%	25%	28%

A Figura 6.52 avalia o CI para cada UHE. A comparação é sempre favorável ao modelo SINF, com exceção para a UHE Água Vermelha.

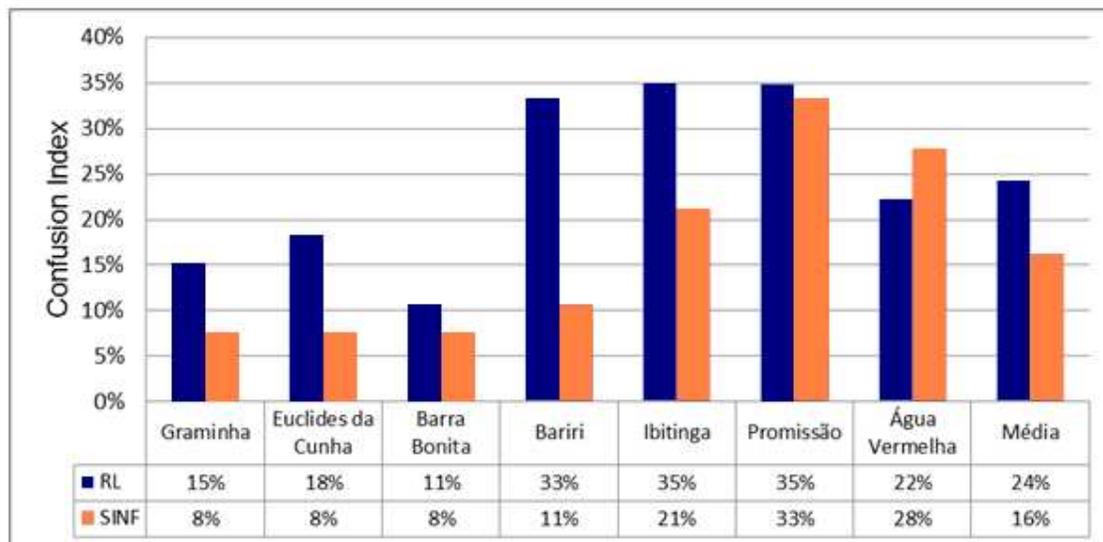


Figura 6.52 Confusion Index para as previsões mensais.

## 6.5 Avaliação das previsões de vazões mensais no semestre 10/2013 – 03/2014

Durante a redação deste trabalho, assistiu-se a uma estação de chuvas 2013-14 particularmente seca no Estado de São Paulo. Decidiu-se, assim, avaliar as previsões de vazões mensais no semestre 10/2013 – 03/2014 a fim de testar os modelos neste período atípico. A AES Tietê disponibilizou as observações diárias das precipitações e vazões observadas nos pontos de medição das UHEs durante o período 06/2013 – 03/2014. Estes dados foram elaborados a fim de fornecer aos modelos RL e SINF os dados de entrada, conforme Tabela 6.4, necessários para o lançamento do SISPREV.

A Figura 6.53 mostra as médias das precipitações mensais do semestre 10-03 observadas nas UHEs no período 2000-2011, no semestre mais seco deste período, ou seja, 10/2000 – 03/2001, e, enfim, no semestre 2013-2014. Lembra-se que logo depois do semestre 10/2000 – 03/2001, o Brasil sofreu o racionamento da energia, devido à diminuição de níveis dos reservatórios, ao qual seguiu um aumento considerável da capacidade do parque gerador de energia elétrica (veja Figura 1.2).

Nota-se que a precipitação total do semestre 2013-2014 é 27% inferior à precipitação ocorrida no semestre 2000-2001 e 39% inferior à precipitação do período 2000-2001. Em particular, evidencia-se que janeiro 2014 teve uma percentual de precipitação 70% inferior à precipitação média dos meses de janeiro (mês historicamente mais chuvoso) do período 2000-2001.

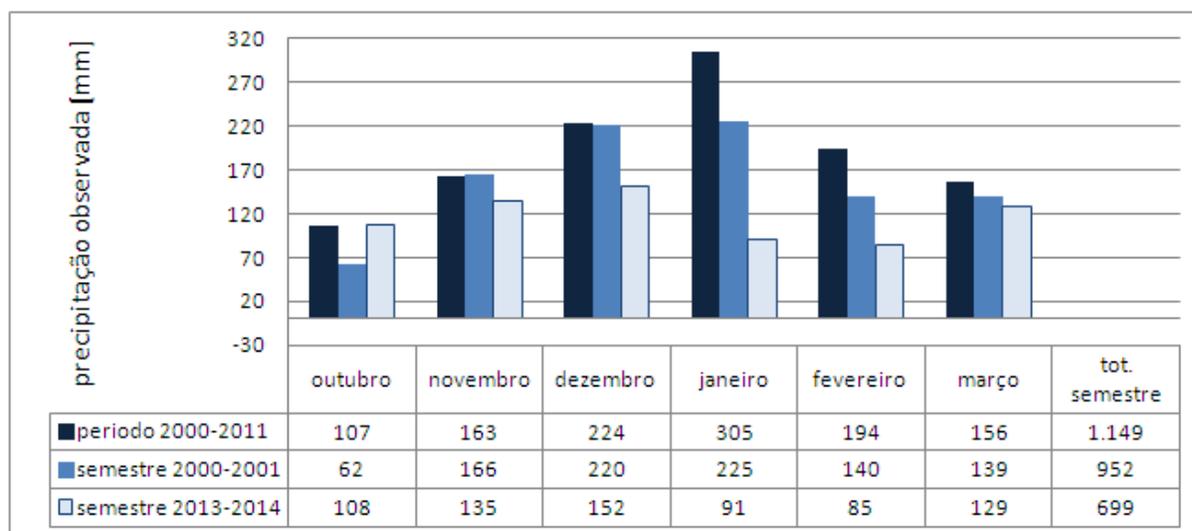


Figura 6.53 Precipitações médias observadas nas UHEs no período 2000-2011 e nos semestres 2000-2001, 2013-2014.

As Figuras 6.54 – 6.60 mostram os gráficos contendo os resultados das execuções dos modelos RL e SINF no semestre 10/2013 – 03/2014.

Figuras 6.54 – 6.60 Resultados das execuções dos modelos RL e SINF no semestre 10/2013-03/2014

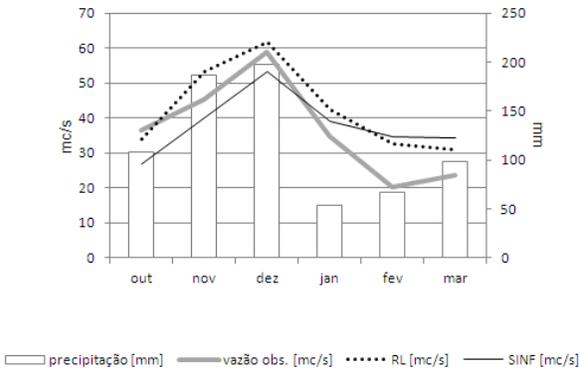


Figura 6.54 Graminha.

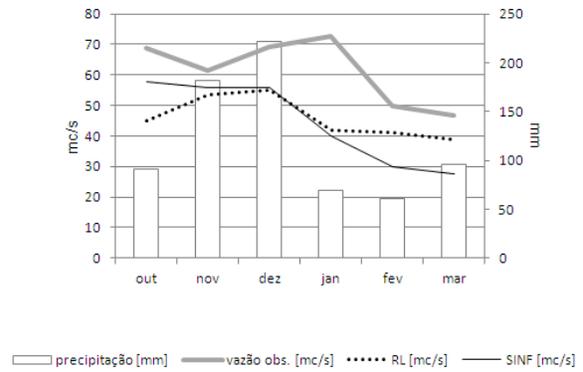


Figura 6.55 Euclides da Cunha.

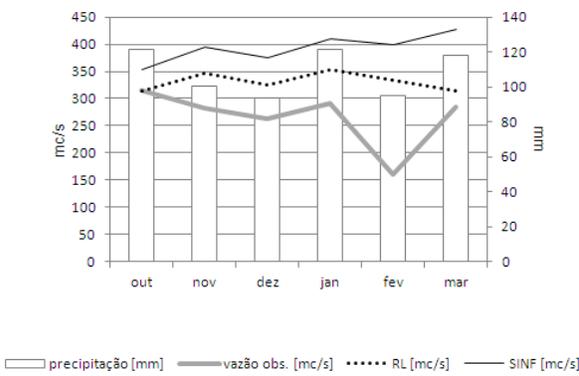


Figura 6.56 Barra Bonita.

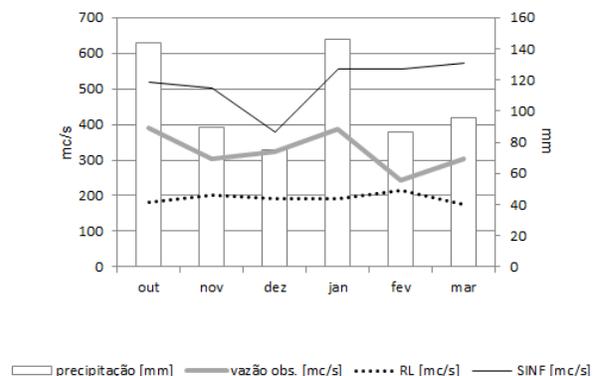


Figura 6.57 Bariri.

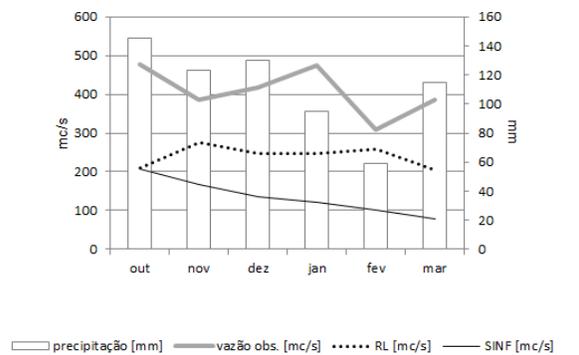


Figura 6.58 Ibitinga.

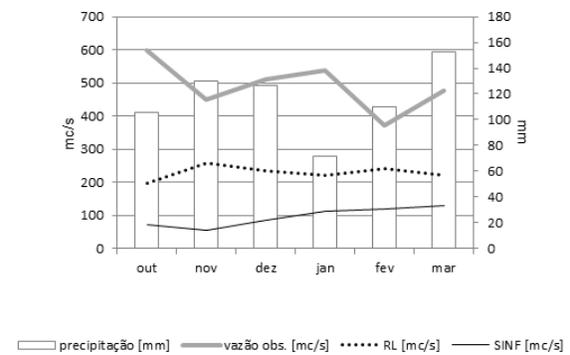


Figura 6.59 Promissão.

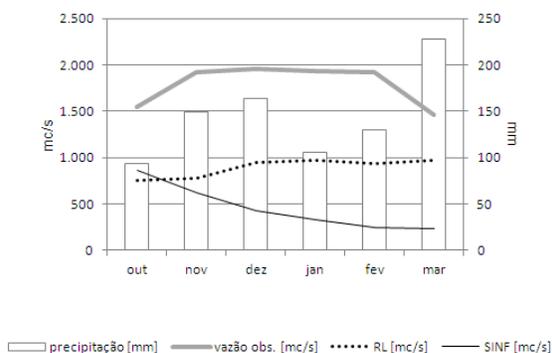


Figura 6.60 Água Vermelha.

Os dois modelos têm o melhor desempenho na previsão de vazão da UHE Graminha, aquela com vazão menor, conforme a Tabela 6.7, que mostra os índices NS para o semestre avaliado. Em particular, nota-se que, segundo a classificação de Santhi *et al.* (2001), o modelo RL tem desempenho considerado muito bom para a UHE Graminha e NS inadequado para as restantes UHEs. O modelo SINF tem desempenho bom para a UHE Graminha e inadequado para as restantes UHEs.

Tabela 6.9 Índices NS relativos às previsões de vazões mensais no semestre 10/2013 – 03/2014.

UHE	RL	SINF
Graminha	0,65	0,51
Euclides da Cunha	-2,19	-2,65
Barra Bonita	-1,88	-6,88
Bariri	-7,07	-15,46
Ibitinga	-8,84	-21,7
Promissão	-13,79	-30,73
Água Vermelha	-19,31	-44,5

# Capítulo 7

## Conclusões e Trabalhos Futuros

Neste capítulo, inicialmente é feito um breve resumo da dissertação com o objetivo de introduzir as conclusões finais. Na sequência, são listadas as conclusões e as observações relacionadas a este trabalho. Por fim, são apresentadas as perspectivas de desenvolvimento futuro que visam à melhoria da qualidade da previsão de vazões.

## 7.1 Conclusões

Conforme mencionado no Capítulo 1, o planejamento da operação do sistema hidrotérmico brasileiro tem grande dependência de variáveis estocásticas, como as vazões naturais e incrementais que afluem aos aproveitamentos hidrelétricos. Nesse contexto, a previsão de vazões apresenta-se como insumo fundamental para a programação da operação do sistema. A busca pela melhoria da qualidade das previsões de vazões é importante para a escolha de uma política operacional econômica, confiável e que contribui com a diminuição dos riscos de racionamento e econômicos para o sistema elétrico brasileiro. O objetivo deste trabalho foi abordar os aspectos que determinam a qualidade da previsão através da comparação entre os modelos de previsão de vazões RL e SINF.

A revisão da literatura apresentada no Capítulo 2 contribuiu para classificar os vários modelos de previsão segundo três categorias: conceituais, empíricos e combinados. No mesmo capítulo, foram mostrados os resultados de estudos realizados para comparar diferentes modelos de previsão de vazões.

Uma vez que o foco desta pesquisa é a avaliação de modelos matemáticos, no Capítulo 3 o referencial teórico dos modelos empíricos a serem avaliados, RL e SINF, foi minuciosamente analisado.

No Capítulo 4, apresentou-se o SISPREV, a ferramenta que gerencia os dois modelos de previsão. O SISPREV pode ser usado com duas finalidades: prever vazões diárias e mensais e servir como instrumento de análise dos modelos, por meio da comparação das vazões previstas com as vazões observadas. Neste trabalho, o SISPREV foi utilizado como instrumento de análise dos modelos RL e SINF.

No Capítulo 5, a metodologia CRISP-DM, usada para analisar o volume de dados disponível e avaliar os modelos implementados no SISPREV foi apresentada. Esta metodologia baseia-se em um modelo hierárquico de processos constituído de seis fases: Entendimento do Problema, Compreensão dos Dados, Preparação dos Dados, Modelagem, Avaliação e Desenvolvimento. No mesmo capítulo detalha-se o significado de onze índices de desempenho utilizados para avaliar a qualidade dos previsores.

No Capítulo 6, a fim de aferir a qualidade dos dois previsores de vazões e avaliar a metodologia proposta, foram apresentados dois estudos de previsões diárias e mensais. Estes estudos tinham os seguintes objetivos:

- Analisar as respostas dos modelos (RL e SINF) às previsões de vazões afluentes diárias e mensais;
- Avaliar os desempenhos dos modelos em condições climáticas típicas ou particulares.

Na fase de compreensão dos dados, analisou-se o banco de dados disponível com as informações (vazões e precipitações) relativas a doze anos, de 2000 até 2011, associadas a dez UHEs e quinze postos de medição. Na fase de preparação, definiu-se um conjunto de dados para cada estudo, extraído do banco de dados, a partir do qual foram validados os modelos. Considerando a sazonalidade das precipitações, foram escolhidos onze períodos para previsão diária e onze períodos para previsão mensal. Sendo nove as UHEs avaliadas, obteve-se um total de 99 estudos de previsão diária e 99 estudos de previsão mensal para cada modelo. Foram executados os modelos de previsão RL e SINF para os períodos escolhidos e produzidos os resultados, ou seja, as previsões de vazões diárias e mensais. Na fase de avaliação dos estudos diários, decidiu-se analisar os resultados com o índice MAPE para as previsões diárias, o NS para as previsões mensais e o CI para ambas. Além disso, na avaliação dos estudos diários, decidiu-se aumentar o número de períodos a serem avaliados, de 96 para 126, com o objetivo de melhorar a confiabilidade das análises de previsão de vazões com valores superiores a  $2.000 \text{ m}^3/\text{s}$ .

Dos resultados obtidos, podem-se listar as seguintes conclusões:

- Os dois modelos de previsão, RL e SINF, têm tendência à subestima, tanto para as previsões diárias quanto para as previsões mensais.
- De forma geral, o modelo RL tem um desempenho mais adequado para previsões diárias, enquanto o modelo SINF tem um desempenho mais adequado para as previsões mensais.
- Nas previsões diárias, o modelo RL tem 90% dos índices de previsões com avaliação considerada adequada, enquanto o modelo SINF tem 64%. Assim como já era previsto no Capítulo 3, os resultados dos modelos estão relacionados às características das técnicas. De fato, a RL, envolvendo uma estrutura rígida de premissas que associa um conjunto de pontos a uma função, apresenta um desempenho adequado para condições típicas de chuva-vazão, ou seja, nas condições em quem a vazão tende a crescer com o aumento da chuva. Em condições atípicas, por exemplo, para vazões médias maiores de  $2.000 \text{ m}^3/\text{s}$  e precipitações médias diárias menores que  $7 \text{ mm}/\text{dia}$ , o SINF tem 20% dos índices com avaliação satisfatória, 10% com avaliação boa e 50% com avaliação muito boa. De fato, a SINF não necessita determinar previamente as relações entre as variáveis dependentes e independentes, pois a forma das relações é estabelecida durante o processo de aprendizagem. Por isso, a SINF tem índices de desempenho considerados adequados mais distribuídos no plano chuva-vazão em comparação à RL.

Cabe ressaltar que, quanto maior a vazão registrada, maior o potencial de geração de energia. Este fato reforça a importância do estudo da qualidade das previsões nestas condições.

- Ressalta-se a complementaridade dos modelos: apenas 3% dos estudos diários avaliados não tiveram avaliação adequada em alguns dos modelos.
- Nas previsões mensais, os dois modelos tiveram um desempenho parecido, com melhor desempenho para as usinas com vazões médias menores.
- O índice CI mostrou-se coerente em relação aos índices MAPE e NS. Isto significa que bons resultados apresentados pelos índices MAPE e NS foram acompanhados de bons resultados apresentados pelo índice CI na comparação entre os modelos RL e SINF.

## 7.2 Trabalhos Futuros

Como sugestões de trabalhos futuros são listados três itens.

- **Recalibração dos modelos.** Nos estudos diários, a RL subestima 70% das previsões e a SINF, 67%. Nos estudos mensais a RL subestima 83% das previsões e a SINF, 73%. Além disso, o modelo SINF apresentou valores anômalos em 3% das previsões diárias e em 24% das previsões mensais. A recalibração dos modelos deve distribuir os erros de forma mais equilibrada e melhorar a qualidade das previsões.
- **Previsões de precipitação.** Atualmente, o banco de dados disponibilizado pelo SISPREV associa às previsões de precipitação dados observados e registrados conforme as informações tornam-se disponíveis. Assim, neste estudo, o modelo SINF trabalhou sob a hipótese de previsões de precipitação perfeitas. Propõe-se, para os trabalhos futuros, comparar as previsões de vazões do modelo SINF, trabalhando com previsões de precipitação perfeitas e com previsões de precipitação reais, a fim de avaliar o desempenho do modelo em relação aos erros nos dados de entrada.
- **Pontos de medição.** Outra sugestão para os trabalhos futuros é a de aprimorar os modelos, incluindo nas variáveis de entrada (veja Tabelas 6.3, 6.4), além do ponto correspondente à UHE, também pontos a montante. Por exemplo, pode ocorrer a não verificação de alguma precipitação perto da UHE e a vazão pode aumentar caso, no rio, distante alguns quilômetros, registrar-se uma precipitação que contribua com o aumento. Desta forma, aumenta-se a quantidade de informação que contribui com a qualidade da vazão prevista.

## Referências Bibliográficas

AES, 2013. *Usinas da AES Tietê*. Disponível em: [www.aestiete.com.br](http://www.aestiete.com.br) (Página inicial > Nossas usinas). Acesso em 14/05/2013.

ANDRADE, M.A.; DE MELLO, C.R.; BESKOW, S. *Simulação hidrológica em uma bacia hidrográfica representativa dos Latossolos na região Alto Rio Grande, MG*. Revista Brasileira de Engenharia Agrícola e Ambiental, vol.17, n.1, pp.69–76, Campina Grande, PB, UAEEA/UFCG, 2013.

ANDRADE, M.A. *Simulação hidrológica em uma bacia hidrográfica representativa dos Latossolos na região Alto Rio Grande*. Tese de mestrado, UFL, 2011.

ANDRIOLO, M.V. *Análise de Métodos não Lineares para Previsão de Vazões Médias Mensais*. Dissertação de mestrado, UFP, Curitiba, 2006.

ANEEL, 2008. *Atlas de Energia Elétrica no Brasil*. Agencia Nacional de Energia Elétrica. Disponível em: [www.aneel.gov.br](http://www.aneel.gov.br). Acesso em 06/07/2013.

BALLINI, R.; SOARES, S. *Previsão de vazões médias mensais usando redes neurais nebulosas*. Revista Controle & Automação, vol. 14, n. 3, Julho, Agosto e Setembro de 2003.

BALLONE, G. J. *Neurônios e Neurotransmissores*. Disponível em: [www.psiqweb.med.br](http://www.psiqweb.med.br). Acesso em: 16/09/2013.

BAUCHSPIESS, A. *Introdução aos Sistemas Inteligentes*. Universidade de Brasília, 2004. Disponível em: [www.ene.unb.br/adolfo/ISI](http://www.ene.unb.br/adolfo/ISI). Acesso em 04/02/2013.

BIALOSKORSKI, N.S.; NAGANO, M.S.; MORAES, M.B.C. *Utilização de redes neurais artificiais para avaliação socioeconômica: uma aplicação em cooperativas*. RAUSP, São Paulo-SP, vol. 41, n. 1, pp.59-68, janeiro/fevereiro/março, 2006.

BOCANEGRA, C.W.R. *Procedimentos para tornar mais efetivo o uso das redes neurais artificiais em planejamento de transportes*. Dissertação de mestrado, Escola de Engenharia de São Carlos - USP, 2002.

BOWMAN, J.S.; EMERSON, S.L.; DARNOVSKY, M. *The Practical SQL Handbook – Using Structured Query Language*. Third Edition. Addison Wesley, 1996.

BRAGA, A.; LUDEMIR, T.; CARVALHO, A. *Redes Neurais Artificiais: Teoria e Aplicações*. LTC Editora, Rio de Janeiro, 2000.

CAMARGOS, F.L. *Lógica Nebulosa: uma abordagem filosófica e aplicada*. UFSC, 2002. Disponível em: <http://www.inf.ufsc.br/~barreto/trabaluno/IAfernandoLC.pdf>. Acesso em 04/02/2013.

CAMPOS, R.J. *Previsão de séries temporais com aplicações a séries de consumo de energia elétrica*. Tese de mestrado. UFMG, 2008.

CARVALHO, J.V.; ROMANEL, C. *Redes neurais temporais aplicadas ao monitoramento de barragens*. RESI-Revista eletrônica de Sistemas de Informação Edição 10, n. 1, pp.1-9, 2007.

CERVILHA, G.B.; DA SILVA FILHO, A.C. *Metodos estocásticos e não estocásticos para a previsão da vazão em rios*. II Congresso de Iniciação Científica. Centro Universitário de Franca, 2010.

CHAPMAN, P.; CLINTON, J.; KERBER, R.; KHABAZA, T.; REINARTZ, T.; SHEARER, C.; WIRTH R. *CRISP-DM 1.0 Step-by-step Data Mining Guide*. SPSS, CRISP-DM consortium, 2000.

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

CORREA, M.V.; BAESSA, P.D.B. *Identificação de sistemas dinâmicos não-lineares utilizando a lógica Fuzzy*. Revista Principium, CULMG, vol. 1, n. 3, pp.68-86, dezembro de 2007.

CRUZ, M.F.M.; RODRIGUES, L.D.; VERSIANI, B.L. *Previsão de Vazões com a Metodologia DPFT e com Redes Neurais Artificiais*. Revista Brasileira de Recursos Hidricos, vol. 15, n. 1, 2010.

DA SILVA, B.C. *Previsão Hidroclimática de Vazão de Curto Prazo na Bacia do Rio São Francisco*. Tese do doutorado, UFRGS, 2005.

DE OLIVEIRA, G.L.C.P.; VIEIRA, R.; UMEZAWA, P.K. *Operação de controle de cheias na*

*bacia do Rio Uruguai*. Encontro para debates de assuntos de operação - XII edição, Brasília, 2012.

DE LURGIO, S.A. *Forecasting principles and applications*. 1.<sup>a</sup> edição: McGraw-Hill, p. 802, Singapore, 1998.

EPE, 2013. *Balanco Energético Nacional 2012*. Empresa de Pesquisa Energética - EPE. Disponível em: [www.ben.epe.gov.br](http://www.ben.epe.gov.br) . (Página inicial > Balanco Energético Nacional BEN > BEN 2012 Relatório final). Acesso em 13/04/2013.

EPEa, 2013. *Plano Decenal de Expansão de Energia PDE*. Empresa de Pesquisa Energética - EPE. Disponível em: [www.epe.gov.br](http://www.epe.gov.br) . (Página inicial > Planos e Balanços > Plano Decenal de Energia PDE). Acesso em 13/04/2012.

EVUSKOFF, A.G.; EBECKEN, N.F.F.; DE SOUZA, F.T.; ALEGRE, M.P; TERRA, G.S.; TAVARES, G.M.; DA HORA, A.F. *Uma Abordagem de Mineração de Dados para a Previsão de Vazões com Incorporação de Previsão de Precipitação da Bacia do Rio Iguaçu*. Workshop ONS Previsão de vazão, Rio de Janeiro, Brasil, 20-21 setembro 2007. Disponível em: [www.ons.org.br](http://www.ons.org.br) (Página inicial > Previsão de vazões> Artigos sessão). Acesso em 14/05/2013.

FADIGA, Jr, F.M.; LOPES, J.E.G.; MARTINS, J.R.S; DE BARROS, M.T.L.; DOS SANTOS, R.C.P. *Modelos de Previsão de Vazões para a Bacia Incremental à UHE Itaipu*. Revista Brasileira de Recursos Hídricos, vol. 13, n. 2, pp. 65-75, 2008.

FAYAL, M.A.A. *Séries temporais aplicadas ao Planejamento de Operação do Sistema Interligado Nacional - SIN*. Dissertação de mestrado, PUC/RIO, 2008.

FIGUEIREDO, F.F.; NUNES, F.; CARVALHO, R.E.; SANTOS, M.L.; BATISTA, M.; SILVA, J.JA. *O que Fazer e o que Não Fazer com a Regressão: pressupostos e aplicações do modelo linear de Mínimos Quadrados Ordinários (MQO)*. Revista Política Hoje, vol. 20, n. 1, 2011.

FIGUEIREDO, K.; BARBOSA, C.R.H.; DA CRUZ, A.V.A.; VELLASCO, M.; PACHECO, M.A.C.; CONTRERAS, R.J.; BARROS, M.; SOUZA, R.C.; MARQUES, V.S.; DUARTE, U.M. *Modelo de Previsão de Vazão com Informação de Precipitação Utilizando Redes Neurais*. Revista Brasileira de Recursos Hídricos, vol. 12, n. 3, pp. 69-82, 2007.

GAITHER, N.; FRAZIER, G. *Administração da produção e operações*. 8.<sup>a</sup> Edição Thomson

Learning, São Paulo, 2006.

GOMES, L.F.C.; MONTENEGRO, S.M.G.L. *Modelo Baseado na Técnica de Redes Neurais para Previsão de Vazões na Bacia do Rio São Francisco*. Revista Brasileira de Recursos Hídricos, vol. 15, n. 1, pp. 5-15, 2010.

GUILHERME, L.; GUILHON, F.; ROCHA, V.F. *Comparação de métodos de previsão de vazões naturais afluentes a aproveitamentos hidroelétricos*. Workshop ONS Previsão de vazão, Rio de Janeiro, Brasil, 20-21 de setembro de 2007. Disponível em: [www.ons.org.br](http://www.ons.org.br) (Página inicial > Previsão de vazões> Artigos sessão > Acesso em 14/05/2013).

GUILHON, L.G.F. *Modelo heurístico de previsão de vazões naturais médias semanais aplicado à usina de Foz do Areia*. Tese de mestrado. UFRG, 2002.

GUILHON, L.G.F.; ROCHA, V.F.; MOREIRA, J.C. *Comparação de Métodos de Previsão de Vazões Naturais Afluentes a Aproveitamentos Hidrelétricos*. Revista Brasileira de Recursos Hídricos, vol. 12, n. 3, pp. 13-20, 2007.

GUTIERREZ, J. L. C.; ROMANEL, C. *Aplicação de redes neurais na previsão de vazão através da fundação da ombreira esquerda da barragem Corumbà-I*. Comitê brasileiro de barragens XXV seminário nacional de grandes barragens, Salvador, 12-15 outubro de 2003.

HIDALGO, I.G. *Ferramentas e Metodologia para Consolidação de Dados de Usinas Hidrelétricas Brasileiras*. Tese de doutorado. UNICAMP, 2009.

HOLLINGWORTH, J.; SWART, B.; CASHMAN, M. *C++ Builder 5 Developer's Guide*. Sams Publishing, 2001.

JACOBS, R.A. *Increased rates of convergence through learning rate adaptation*. Neural Networks, vol. 1, pp. 295–307, 1988.

LADEIRA, M.H. *Representação de Conhecimento e Redes de Decisão*. Tese de doutorado. UFRGS, 1997.

LANZETTI, M.D.D. *Previsão de Vazão por Redes Neurais Artificiais e Transformada Wavelet*.

Dissertação de mestrado, PUC/RIO, 2008.

LAROSE, D.T. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley and Sons, Inc, 2005.

LAUDANNA, P.R.; COSTA, C.A.S.; RIBEIRO, T.S.; MORIBE, C.H.; LOPES, J.E.G.; CICOGNA, M.A. *Integração das áreas de planejamento, tempo real, manutenção e comercial, através de um sistema de suporte e decisão*. Encontro para debates de assunto de operação, IX ed., Goiás, Brasil, 25-29 de março de 2007.

LAUDANNA, P.R.; COSTA, C.A.S.; RIBEIRO, T.S.; MORIBE C.H.; LOPES, J.E.G.; CICOGNA M.A. *Sistema para a programação e operação em tempo real de sistemas hidroelétricos baseado em previsão de vazões e simulação hidráulica de reservatórios*. Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE), XVIII ed., Curitiba, Paraná, 16-21 de outubro de 2005.

LIEW, M.W.; VEITH, T.L.; BOSH, D.D.; ARNOLD, J.G. *Suitability of SWAT for the Conservation effects assessment project: A comparison on USDA-ARS watersheds*. Journal of Hydrological Research, v.12, pp.173-189, 2007.

LOPES, J.E.G.; BRAGA, B.P.F.; CONEJO, J.G.L. *SMAP - A Simplified Hydrologic Model: Applied Modeling in Catchment Hydrology*. Ed. V.P.Singh, Water Resources Publications, Littleton, Colorado, USA, pp. 167-176, 1982.

LUNA, I.R. *Redes Neurais Fuzzy Aplicadas em Identificação e Controle de Sistemas. Tese de mestrado*. UNICAMP, 2003.

LUNA, I.R.; SOARES, S.; BALLINI, R. *Modelo adaptivo baseado em regras nebulosas aplicado á previsão de series de vazões semanais*. VIII Congresso Brasileiro de Redes Neurais, Campinas (SP), 2007.

LUNA I.R.; SOARES, S.; LOPES, J.E.G.; BALLINI, R. *Verifying the use of evolving Fuzzy systems for multi-step ahead daily inflow forecasting*. 15th International Conference on Digital Object, 2009.

LUNA I.R.; HIDALGO I. G.; BARBOSA P.S.F; FRANCATO A.L. *Daily Reservoir Inflow Forecasting Using Fuzzy Inference Systems*. IEEE International Conference on Fuzzy Systems, Taipei, Taiwan, 2011.

MAZER, W. *Método não linear para previsões de vazões*. Dissertação de mestrado. UFP, Curitiba, 2003.

McCULLOCH, W.S; PITTS, W. *A logical calculus of the ideas immanent in nervous activity*. Bulletin of Mathematical Biology, vol. 52, n. 1/2, pp. 99-115, 1990.

MICROSOFT, 2013. *Microsoft® SQL Server® 2008 Express com Advanced Services*. Disponível em: [www.microsoft.com](http://www.microsoft.com). (Página inicial > download). Acesso em 26/08/2013.

MME, 2012. *Capacidade Instalada de Geração Elétrica*. Ministério de Minas e Energia - MME. Disponível em [www.mme.gov.br](http://www.mme.gov.br). (Página inicial > Boletins de Energia > 19 - Potência Instalada de Geração Elétrica - Brasil e Mundo). Acesso em 13/04/2013.

MOZELLI, L.A. *Controle Fuzzy para sistemas Takagi-Sugeno: condições aprimoradas e aplicações*. Tese de mestrado. UFMG, 2008.

NETO, L.B.; COELHO, P.H.G.; AMARAL, J.L.M.; MELLO, M.H.C.S. *Minicurso de sistema especialista nebuloso*. XXXVIII Simpósio Brasileiro Pesquisa Operacional, 2006.

NOTINI, H.H. *Ensaio sobre ciclos de negócios*. Tese de doutorado. Escola de Pós-Graduação em Economia da Fundação Getulio Vargas de Rio de Janeiro, 2009.

ONS, 2007. *Novo modelo de previsão de vazões com informação de precipitação para o trecho incremental de Itaipu*. Disponível em: [www.aneel.org.br](http://www.aneel.org.br) (Página inicial > Aplicações > Consulta pública > Documentos). Acesso em 14/05/2013.

ONS, 2011. *Relatório anual de avaliação das previsões de vazões - 2011*. Disponível em: [www.aneel.org.br](http://www.aneel.org.br) (Página inicial > Aplicações > Consulta pública > Documentos). Acesso em 14/05/2013.

ONS, 2013. *Previsão de Vazões*. Operador Nacional do Sistema Elétrico - ONS. Disponível em: [www.ons.org.br](http://www.ons.org.br) (Página inicial > Avaliação de Curto Prazo de Operação > Diretrizes para a Operação Eletroenergética > Hidrologia > Previsão de Vazões). Acesso em 13/04/2012.

PATUELLI, R.; REGGIANI, A.; NIJKAMP, P.; BLIEN, U. *New Neural Network Methods for*

*Forecasting Regional Employment*. Tinbergen Institute Discussion Paper, n. 2006/020/3, Feb. 2006.

RABELO, R.A.L; FERNANDES, R.A.S; CARNEIRO, A.A.F.M; BRAGA, R.T.V. *Uma Abordagem Baseada em Sistemas de Inferência Fuzzy Takagi-Sugeno Aplicada ao Planejamento da Operação de Sistemas Hidrotérmicos de Geração*. Revista Controle & Automação, vol. 22, n. 1, pp. 49-64, 2011.

ROUHANI, K.O. *Analysis, Theory and Design of Logistic Regression Classifiers Used For Very Large Scale Data Mining*. Tese de mestrado. University of Illinois at Chicago, 2006.

SANCHEZ E.C.M. *Controle por Aprendizado Acelerado e Neuro-Fuzzy de Sistemas Servo – Hidráulicos de Alta Frequência*. Tese de mestrado. PUC-RJ, 2009.

SANTHI, C.; ARNOLD, J. G.; WILLIAMS, J. R.; DUGAS, W. A.; SIRINIVASAN, R.; HAUCK, L. M. *Validation of the SWAT model on a large river basin with point and nonpoint sources*. Journal of the American Water Resources Association, vol. 37, pp.1169-1188, 2001.

SOUSA FILHO, F.A.; LALL, U. *Modelo de Previsão de Vazões Sazonais e Interanuais*. Revista Brasileira de Recursos Hídricos, vol. 9, n. 2, pp. 61-74, abril/junho de 2004.

TAKAGI, T.; SUGENO, M. *Fuzzy identification of systems and its applications to modeling and control*. IEEE Transactions on Systems, Man and Cybernetics, vol. 15, n. 1, pp. 116-132, 1985.

TEIXEIRA, R.A.; CORREA, M.V.; ASSIS, J.C.O. *Redes Neuro Fuzzy: algoritmo de treinamento e aplicações*. Revista Principium, CULMG, vol. 1, n. 3, pp.152-168, dezembro de 2007.

TUCCI, C. E. M.; CLARKE, R. T., COLLISCHONN, W., DIAS, P. L. S., OLIVEIRA, G. S. *Longterm flow forecasts based on climate and hydrologic modeling: Uruguay river basin*. Water Resources Research, vol. 39, n. 7, 1181, 11p., 2003.

TUCCI, C.E.M; COLLISCHONN, W. *Previsão de vazão em curto prazo com base na precipitação: modelos conceituais*. Workshop ONS Previsão de vazão, Rio de Janeiro, Brasil, 20-21 de setembro de 2007. Disponível em: [www.ons.org.br](http://www.ons.org.br) (Página inicial > Previsão de vazões> Artigos sessão > Acesso em 14/05/2013).

VISSMAN, W.; LEWIS, G.L. *Introduction to hydrology*. 5.ed. New Jersey: Upper Saddle River.

612p, 2003.

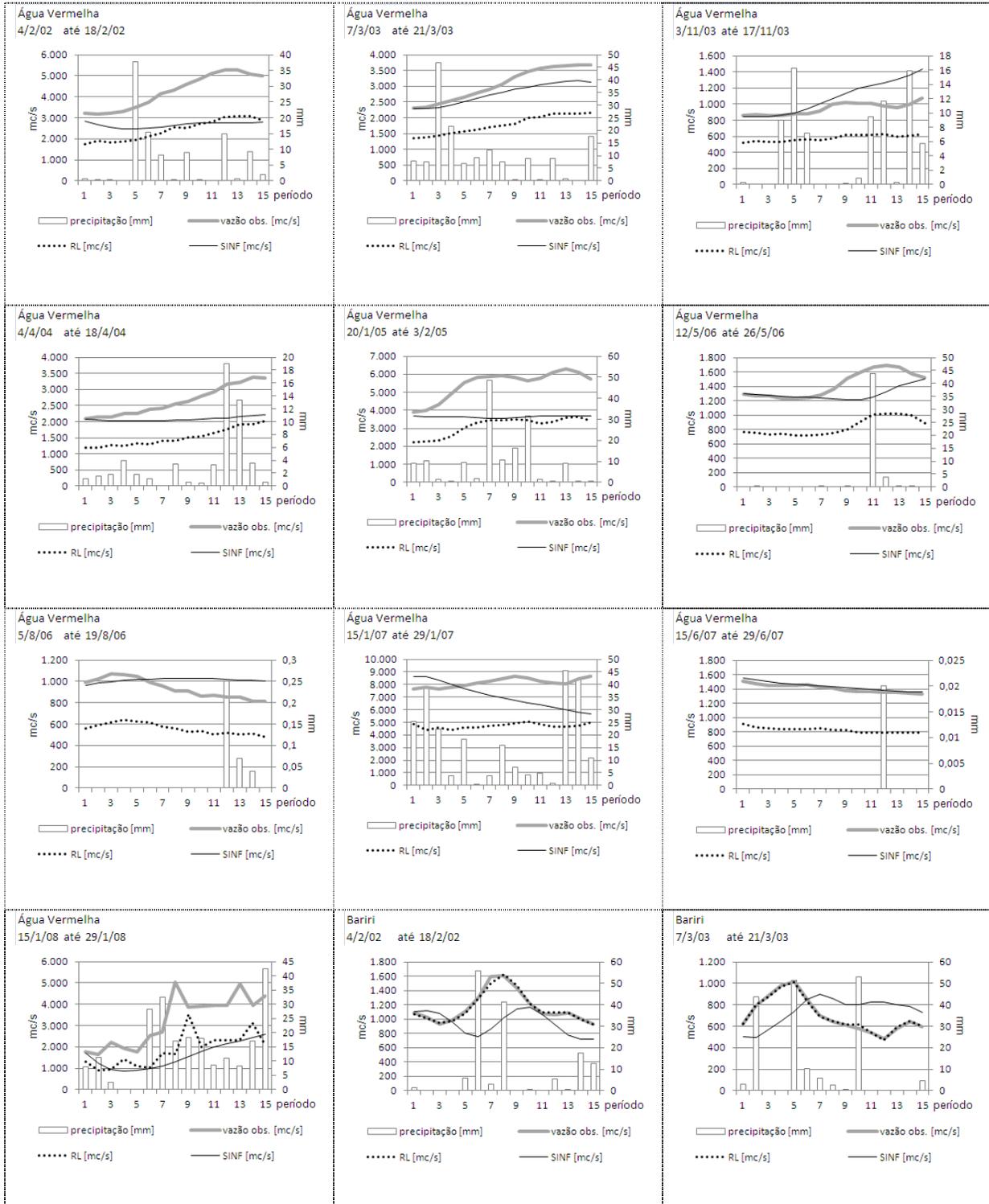
WALLIS, J.; TODINI, E.; *Comment upon the residual Mass Curve Coefficient*. Journal of Hidrology, vol. 24, pp 201-205, North-Holland Publishing Company, Amsterdam, 1975.

WRAY, B.; BEJOU, D.; PALMER, A. *Using neural network analysis to evaluate buyer-seller relationships*. European Journal of Marketing, Bradford, vol. 28, n.10, pp.32-48, 1994.

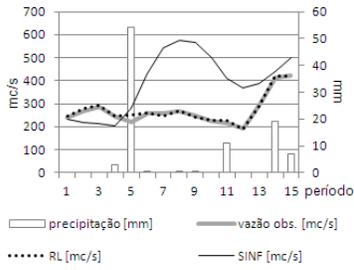
ZADEH, L.A. *Fuzzy sets. Information and Control*. Department of Electrical Engineering and Electronic research Laboratory, University of California, Berkeley, California, vol. 8, n. 3, pp. 338-353, 1965.

# APÊNDICE A – ESTUDOS DE PREVISÃO DE VAZÕES DIÁRIAS

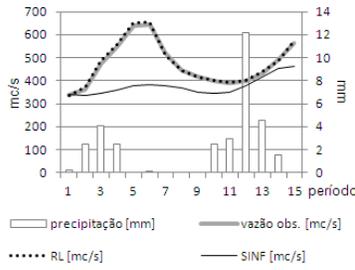
Tabela A.1 Gráficos contendo os resultados das execuções dos modelos RL e SINF, apresentados com as vazões e precipitações observadas.



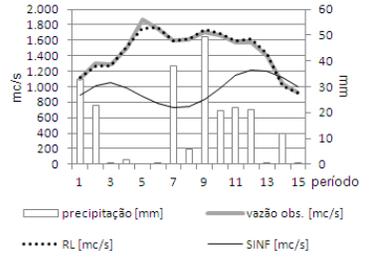
Bariri  
3/11/03 até 17/11/03



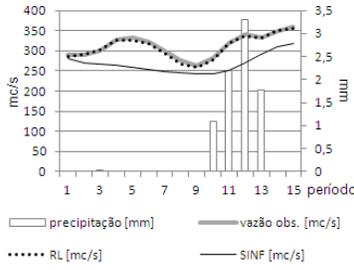
Bariri  
4/4/04 até 18/4/04



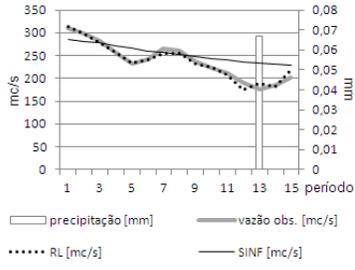
Bariri  
20/1/05 até 3/2/05



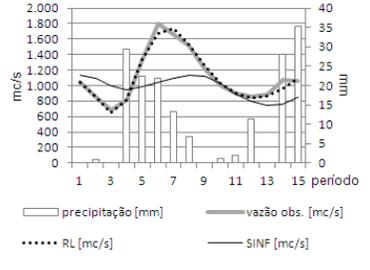
Bariri  
12/5/06 até 26/5/06



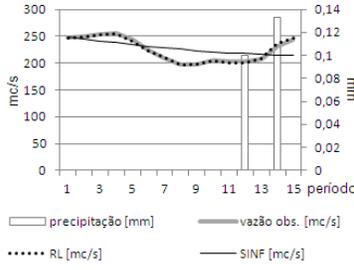
Bariri  
5/8/06 até 19/8/06



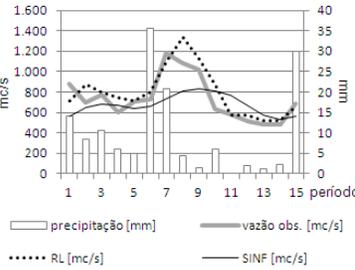
Bariri  
15/1/07 até 29/1/07



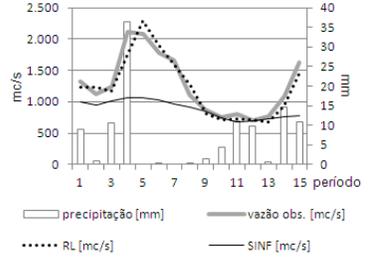
Bariri  
15/6/07 até 29/6/07



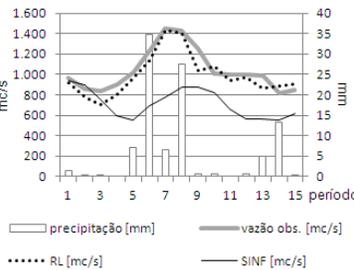
Bariri  
15/1/08 até 29/1/08



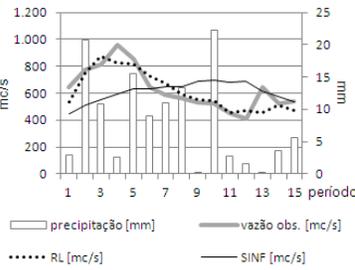
Bariri  
15/12/09 até 29/12/09



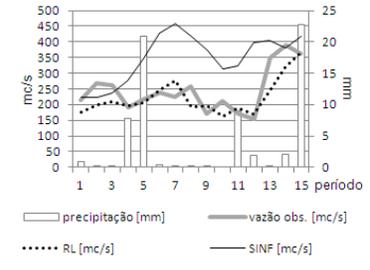
Barra Bonita  
4/2/02 até 18/2/02



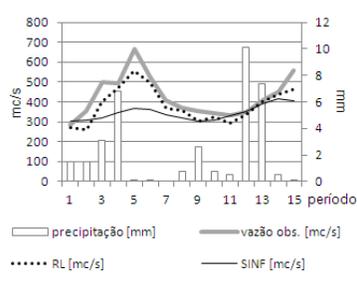
Barra Bonita  
7/3/03 até 21/3/03



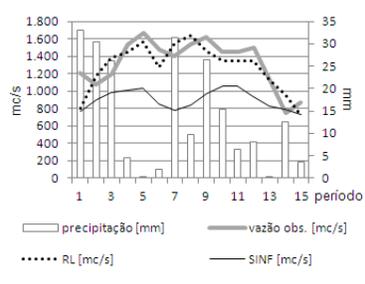
Barra Bonita  
3/11/03 até 17/11/03



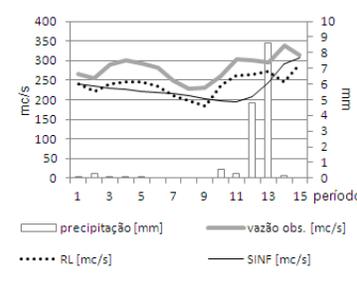
Barra Bonita  
4/4/04 até 18/4/04



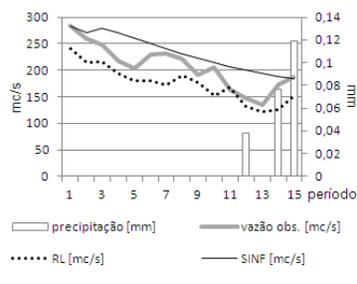
Barra Bonita  
20/1/05 até 3/2/05



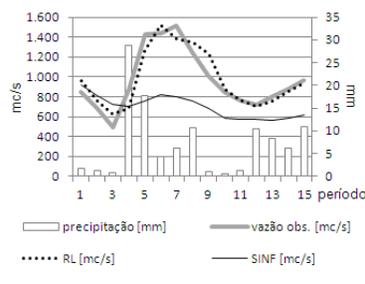
Barra Bonita  
12/5/06 até 26/5/06



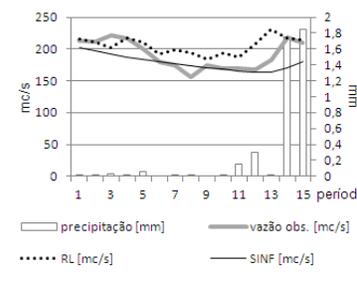
Barra Bonita  
5/8/06 até 19/8/06



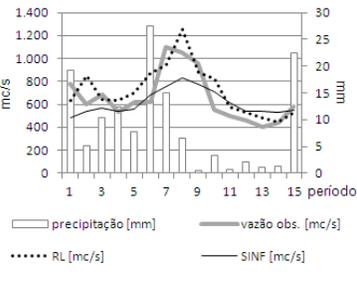
Barra Bonita  
15/1/07 até 29/1/07



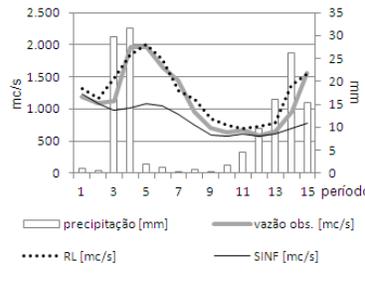
Barra Bonita  
15/6/07 até 29/6/07



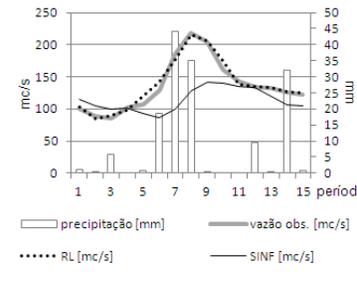
Barra Bonita  
15/1/08 até 29/1/08



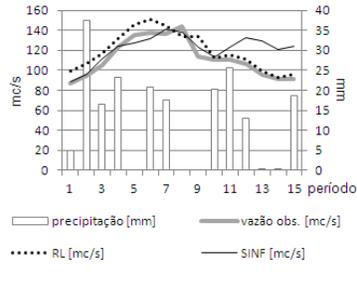
Barra Bonita  
15/12/09 até 29/12/09



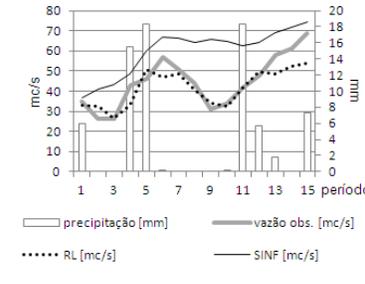
Euclides da Cunha  
4/2/02 até 18/2/02



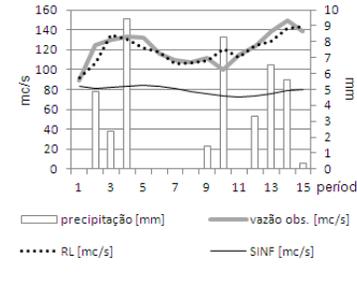
Euclides da Cunha  
7/3/03 até 21/3/03



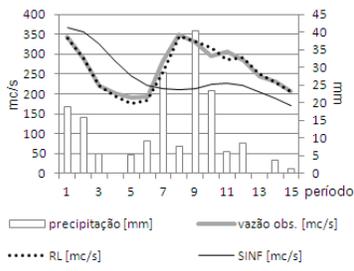
Euclides da Cunha  
3/11/03 até 17/11/03



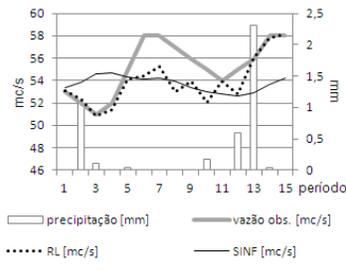
Euclides da Cunha  
4/4/04 até 18/4/04



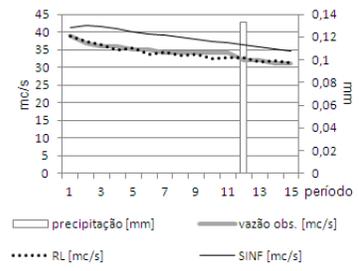
Euclides da Cunha  
20/1/05 até 3/2/05



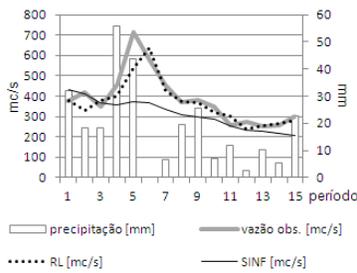
Euclides da Cunha  
12/5/06 até 26/5/06



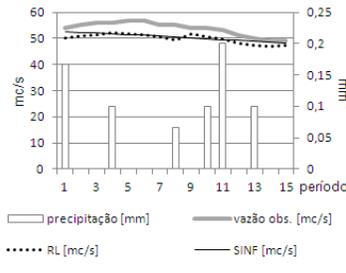
Euclides da Cunha  
5/8/06 até 19/8/06



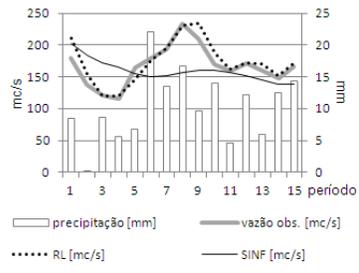
Euclides da Cunha  
15/1/07 até 29/1/07



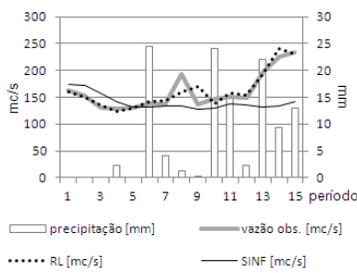
Euclides da Cunha  
15/6/07 até 29/6/07



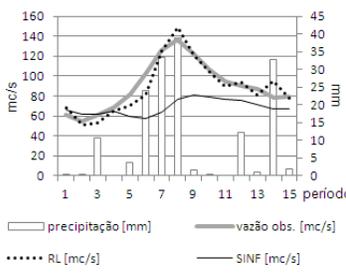
Euclides da Cunha  
15/1/08 até 29/1/08



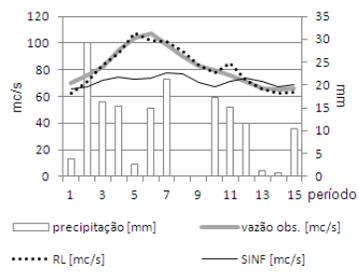
Euclides da Cunha  
15/12/09 até 29/12/09



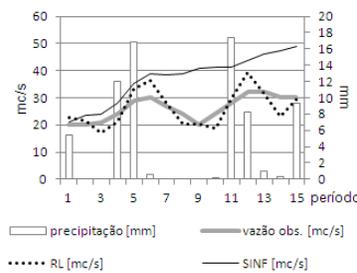
Graminha  
4/2/02 até 18/2/02



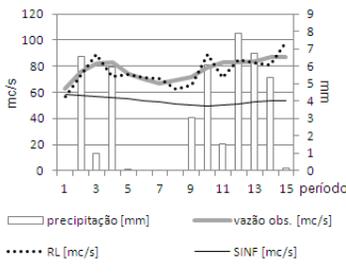
Graminha  
7/3/03 até 21/3/03



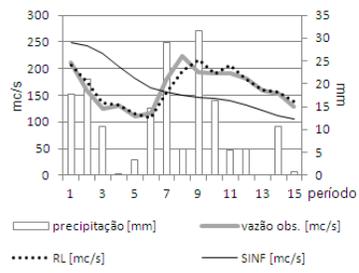
Graminha  
3/11/03 até 17/11/03

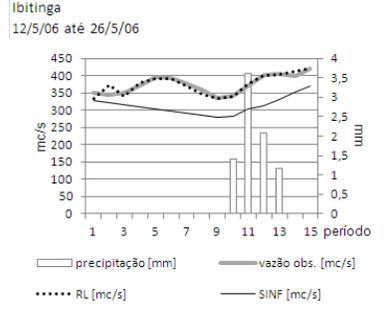
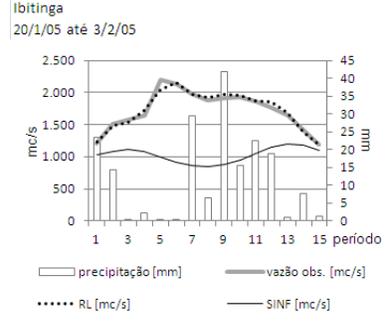
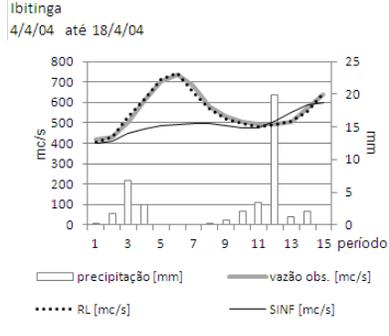
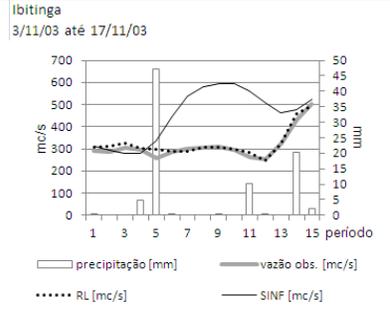
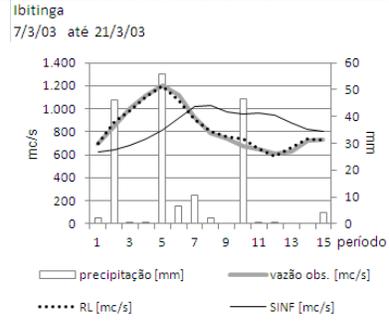
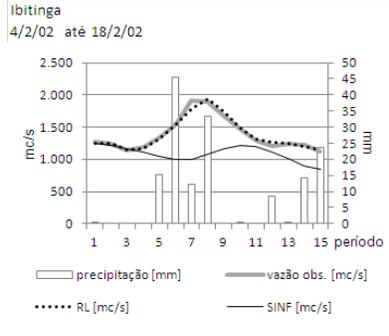
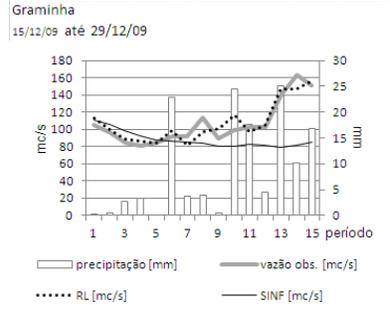
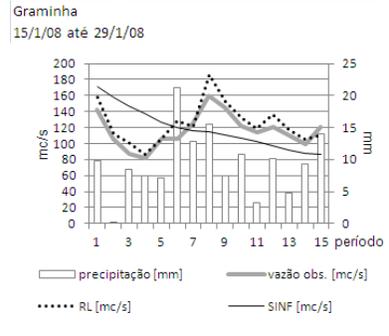
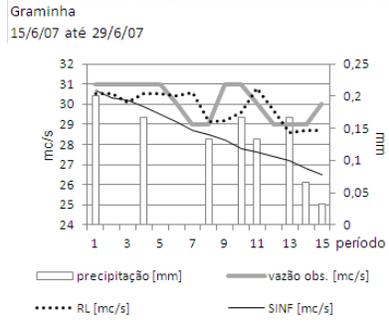
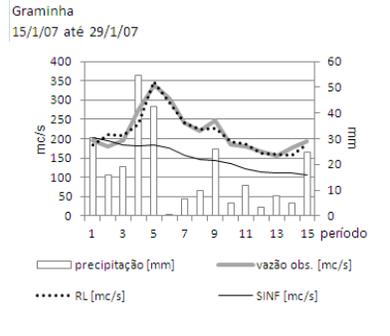
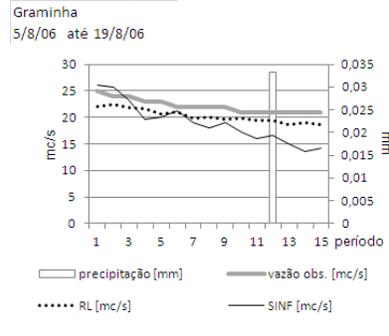
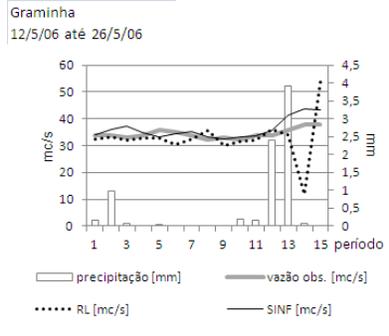


Graminha  
4/4/04 até 18/4/04

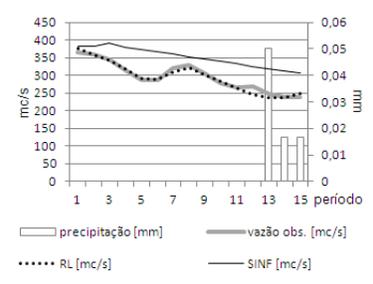


Graminha  
20/1/05 até 3/2/05

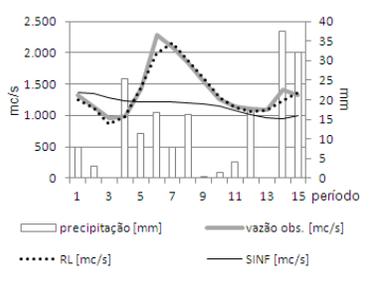




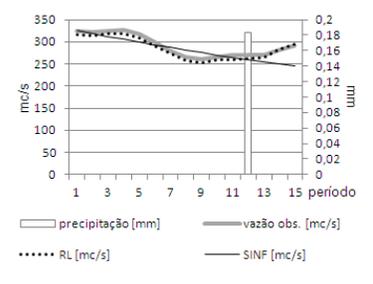
Ibitinga  
5/8/06 até 19/8/06



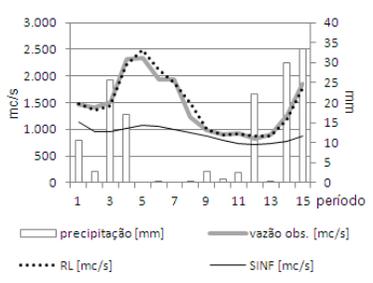
Ibitinga  
15/1/07 até 29/1/07



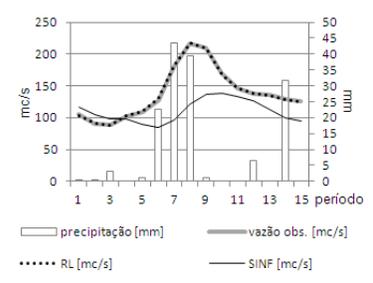
Ibitinga  
15/6/07 até 29/6/07



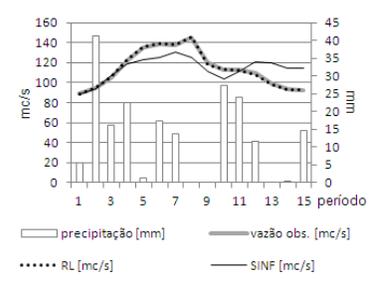
Ibitinga  
15/12/09 até 29/12/09



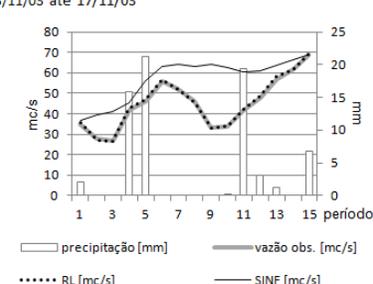
Limoeiro  
4/2/02 até 18/2/02



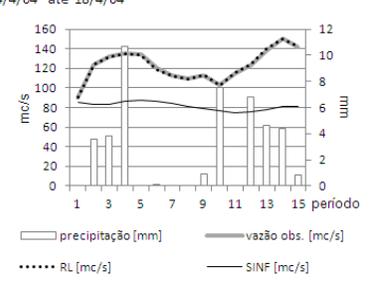
Limoeiro  
7/3/03 até 21/3/03



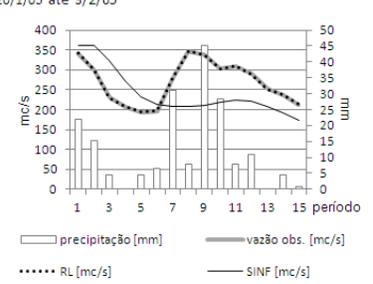
Limoeiro  
3/11/03 até 17/11/03



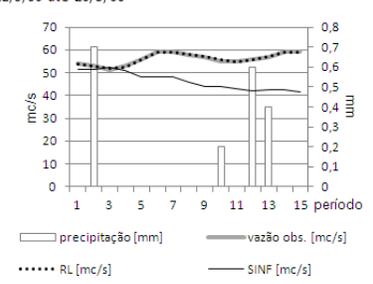
Limoeiro  
4/4/04 até 18/4/04



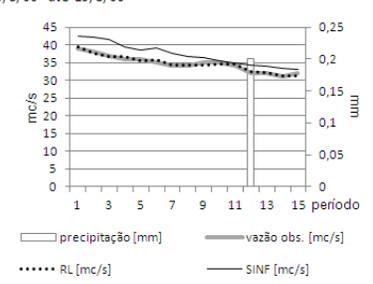
Limoeiro  
20/1/05 até 3/2/05



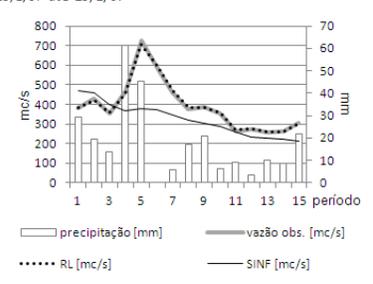
Limoeiro  
12/5/06 até 26/5/06



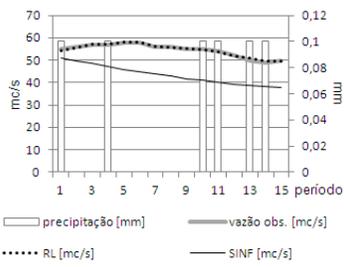
Limoeiro  
5/8/06 até 19/8/06



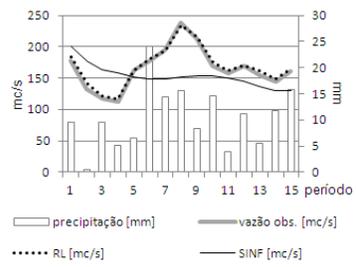
Limoeiro  
15/1/07 até 29/1/07



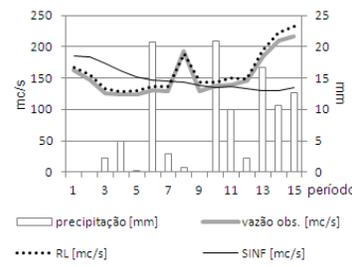
Limoeiro  
15/6/07 até 29/6/07



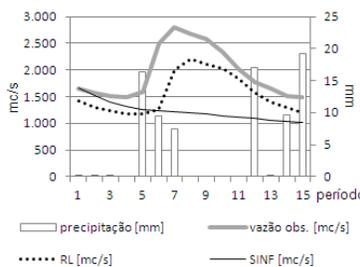
Limoeiro  
15/1/08 até 29/1/08



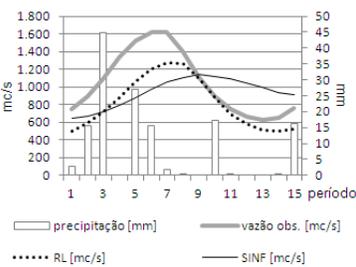
Limoeiro  
15/12/09 até 29/12/09



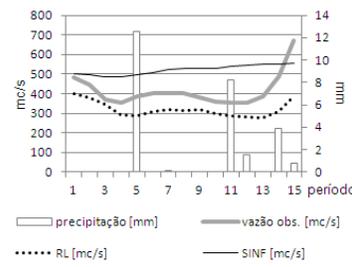
Nova Avanhandava  
4/2/02 até 18/2/02



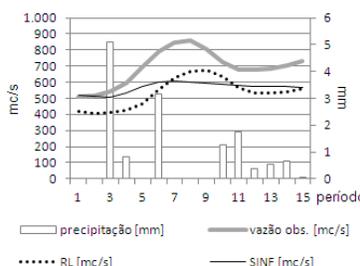
Nova Avanhandava  
7/3/03 até 21/3/03



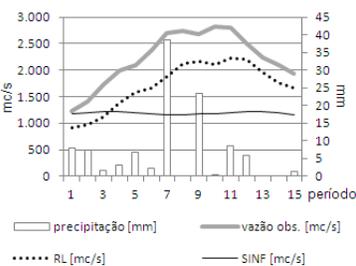
Nova Avanhandava  
3/11/03 até 17/11/03



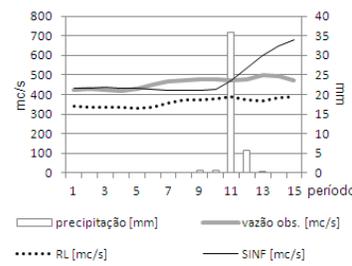
Nova Avanhandava  
4/4/04 até 18/4/04



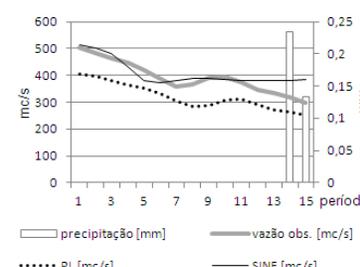
Nova Avanhandava  
20/1/05 até 3/2/05



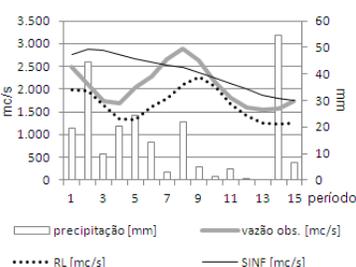
Nova Avanhandava  
12/5/06 até 26/5/06



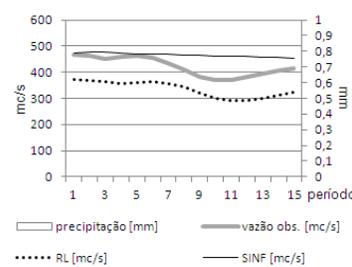
Nova Avanhandava  
5/8/06 até 19/8/06



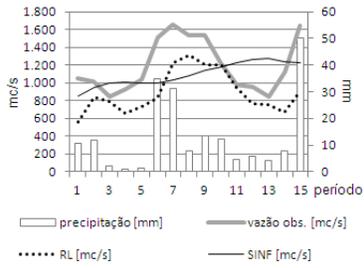
Nova Avanhandava  
15/1/07 até 29/1/07



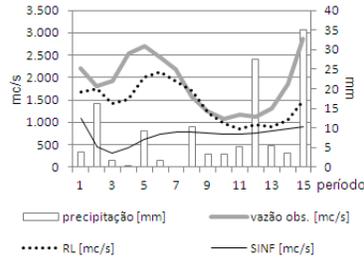
Nova Avanhandava  
15/6/07 até 29/6/07



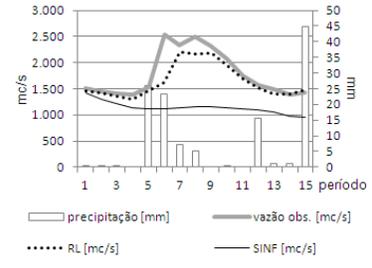
Nova Avanhandava  
15/1/08 até 29/1/08



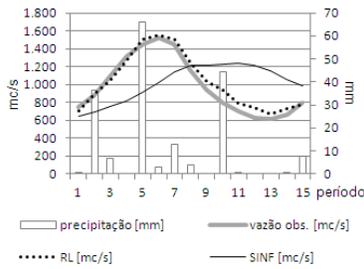
Nova Avanhandava  
15/12/09 até 29/12/09



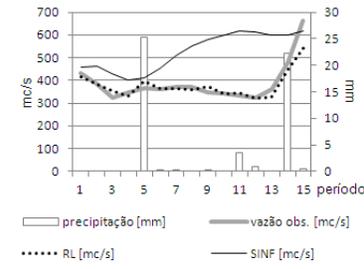
Promissão  
4/2/02 até 18/2/02



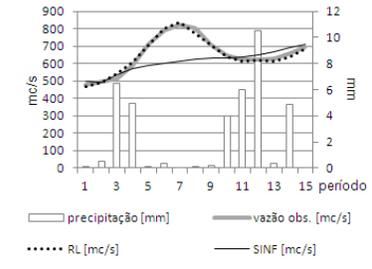
Promissão  
7/3/03 até 21/3/03



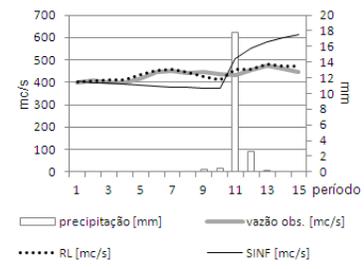
Promissão  
3/11/03 até 17/11/03



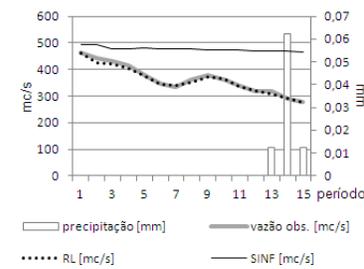
Promissão  
4/4/04 até 18/4/04



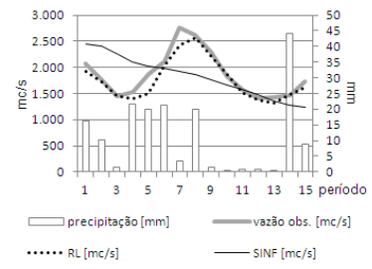
Promissão  
12/5/06 até 26/5/06



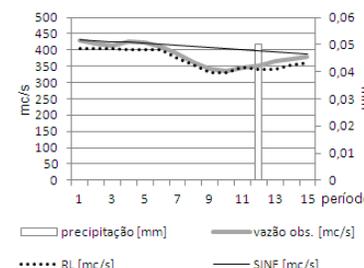
Promissão  
5/8/06 até 19/8/06



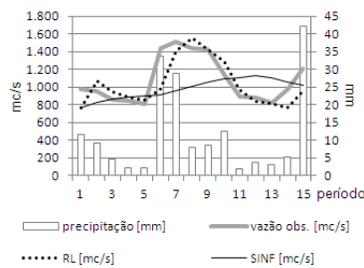
Promissão  
15/1/07 até 29/1/07



Promissão  
15/6/07 até 29/6/07



Promissão  
15/1/08 até 29/1/08



Promissão  
15/12/09 até 29/12/09

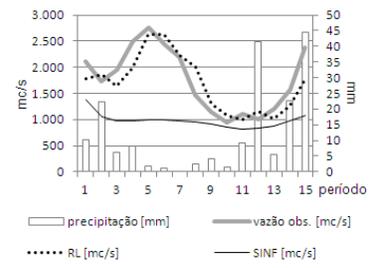


Tabela A.2 Índices de desempenho do modelo RL para as previsões diárias.

UHE	período	Q [m <sup>3</sup> /s]	P [mm]	ME [m <sup>3</sup> /s]	MAE [m <sup>3</sup> /s]	MSE [m <sup>6</sup> /m <sup>2</sup> ]	RMSE [m <sup>3</sup> /s]	MPE	MAPE	U-THEIL	NS	NS-LOG	DM	CI
Água Vermelha	4/2/02	4.250	6,54	-1.826	1.826	3.450.122	1.857	-43%	43,0%	8,91	-4,39	0,16	541%	27%
Água Vermelha	7/3/03	3.064	10,39	-1.283	1.283	1.697.641	1.303	-42%	41,8%	11,31	-5,50	0,13	651%	13%
Água Vermelha	3/11/03	944	5,10	-367	367	136.241	369	-39%	38,8%	10,08	-26,49	0,02	2750%	27%
Água Vermelha	4/4/04	2.647	3,69	-1.136	1.136	1.327.773	1.152	-43%	42,9%	10,13	-5,57	0,12	658%	27%
Água Vermelha	20/1/05	5.455	10,04	-2.327	2.327	5.514.953	2.348	-43%	42,8%	7,70	-8,46	0,11	947%	27%
Água Vermelha	12/5/06	1.428	3,21	-591	591	354.529	595	-41%	41,5%	9,27	-10,60	0,08	1161%	7%
Água Vermelha	5/8/06	935	0,02	-377	377	144.015	379	-40%	40,3%	11,42	-17,51	0,04	1852%	53%
Água Vermelha	15/1/07	8.153	16,28	-3.431	3.431	11.830.951	3.440	-42%	42,1%	18,57	103,80	0,01	10480%	27%
Água Vermelha	15/6/07	1.408	0,00	-582	582	338.847	582	-41%	41,3%	30,68	114,20	0,01	11520%	27%
Água Vermelha	15/1/08	3.233	14,88	-1.390	1.390	2.665.782	1.633	-42%	41,6%	1,97	-0,98	0,32	202%	60%
Bariri	4/2/02	1.163	9,55	-6	22	892	30	0%	1,8%	0,23	0,98	0,98	3%	13%
Bariri	7/3/03	701	11,61	3	11	207	14	1%	1,6%	0,16	0,99	0,99	2%	7%
Bariri	3/11/03	270	6,25	4	8	108	10	1%	3,1%	0,22	0,97	0,96	4%	7%
Bariri	4/4/04	471	2,19	4	6	55	7	1%	1,3%	0,12	0,99	0,99	1%	0%
Bariri	20/1/05	1.460	15,23	-2	26	1.384	37	0%	1,7%	0,21	0,98	0,99	3%	13%
Bariri	12/5/06	312	0,56	-3	4	17	4	-1%	1,2%	0,22	0,98	0,98	2%	13%
Bariri	5/8/06	238	0,00	-1	5	58	8	0%	2,6%	0,45	0,96	0,95	4%	13%
Bariri	15/1/07	1.115	11,49	-9	35	2.685	52	-1%	2,9%	0,22	0,97	0,98	4%	13%
Bariri	15/6/07	225	0,02	-1	2	4	2	0%	0,7%	0,18	0,99	0,99	1%	0%
Bariri	15/1/08	736	9,81	58	96	15.104	123	9%	12,9%	0,63	0,66	0,71	37%	33%
Bariri	15/12/09	1.268	7,32	-38	121	21.859	148	-3%	9,1%	0,44	0,90	0,93	14%	27%
Barra Bonita	4/2/02	1.041	6,57	-54	79	8.574	93	-5%	7,9%	0,71	0,79	0,78	23%	13%
Barra Bonita	7/3/03	632	8,11	-14	62	6.497	81	-1%	9,7%	0,69	0,71	0,66	30%	33%
Barra Bonita	3/11/03	246	4,39	-22	40	2.392	49	-7%	15,5%	0,74	0,50	0,41	52%	47%
Barra Bonita	4/4/04	427	2,38	-44	44	3.188	56	-10%	10,0%	0,66	0,68	0,73	33%	20%
Barra Bonita	20/1/05	1.324	13,99	-47	148	29.120	171	-3%	12,1%	0,71	0,58	0,58	44%	40%
Barra Bonita	12/5/06	280	0,99	-41	41	1.965	44	-14%	14,5%	1,85	-1,12	0,29	212%	40%
Barra Bonita	5/8/06	206	0,02	-32	33	1.335	37	-15%	15,4%	1,53	0,18	0,50	83%	40%
Barra Bonita	15/1/07	968	7,34	8	92	12.878	113	2%	9,9%	0,50	0,85	0,80	18%	7%
Barra Bonita	15/6/07	191	0,28	14	17	502	22	8%	9,5%	1,53	-0,06	-0,62	106%	53%
Barra Bonita	15/1/08	659	9,00	61	125	21.988	148	12%	19,5%	0,78	0,51	0,46	53%	33%
Barra Bonita	15/12/09	1.143	9,39	107	145	31.820	178	13%	14,9%	0,53	0,85	0,78	21%	20%
Euclides da Cunha	4/2/02	136	9,84	2	4	37	6	1%	3,6%	0,27	0,98	0,97	4%	7%
Euclides da Cunha	7/3/03	112	13,23	7	9	99	10	7%	7,7%	0,87	0,71	0,73	30%	27%
Euclides da Cunha	3/11/03	45	4,86	-3	5	46	7	-4%	11,2%	0,81	0,71	0,64	31%	33%
Euclides da Cunha	4/4/04	121	2,81	-2	6	75	9	-1%	5,2%	0,66	0,70	0,63	31%	33%
Euclides da Cunha	20/1/05	264	11,68	-3	8	133	12	-1%	3,2%	0,28	0,95	0,96	6%	13%
Euclides da Cunha	12/5/06	55	0,29	-1	1	4	2	-2%	2,4%	1,34	0,31	0,40	69%	33%
Euclides da Cunha	5/8/06	34	0,01	0	1	1	1	-1%	1,9%	0,84	0,88	0,86	13%	80%
Euclides da Cunha	15/1/07	386	18,56	-20	40	3.579	60	-4%	9,1%	0,63	0,77	0,79	25%	40%
Euclides da Cunha	15/6/07	54	0,05	-4	4	15	4	-7%	6,7%	3,71	-1,14	0,14	214%	47%
Euclides da Cunha	15/1/08	167	10,34	6	10	183	14	4%	5,6%	0,43	0,79	0,85	21%	7%
Euclides da Cunha	15/12/09	161	7,84	1	9	191	14	1%	5,9%	0,53	0,83	0,80	18%	27%
Graminha	4/2/02	90	10,74	-2	7	85	9	-2%	8,5%	0,77	0,85	0,87	17%	20%

Graminha	7/3/03	82	10,70	0	3	17	4	-1%	4,2%	0,57	0,91	0,92	10%	13%
Graminha	3/11/03	26	4,74	0	3	16	4	0%	12,5%	1,40	0,12	0,63	89%	40%
Graminha	4/4/04	77	2,96	-1	6	50	7	-2%	7,5%	1,32	0,10	0,62	90%	53%
Graminha	20/1/05	164	11,53	1	10	176	13	1%	6,0%	0,46	0,85	0,87	16%	27%
Graminha	12/5/06	34	0,53	-2	5	69	8	-5%	12,8%	6,28	-20,53	-0,09	2153%	60%
Graminha	5/8/06	22	0,00	-2	2	4	2	-9%	8,5%	3,54	-1,39	0,27	240%	80%
Graminha	15/1/07	217	17,50	-2	12	254	16	-1%	5,7%	0,41	0,91	0,89	11%	20%
Graminha	15/6/07	30	0,07	0	1	1	1	-1%	2,6%	1,18	-0,09	-0,23	109%	60%
Graminha	15/1/08	116	9,51	9	11	163	13	7%	9,0%	0,64	0,61	0,74	40%	13%
Graminha	15/12/09	106	9,09	1	8	95	10	2%	7,6%	0,61	0,84	0,80	18%	40%
Ibitinga	4/2/02	1.386	10,26	0	33	1.949	44	0%	2,2%	0,28	0,97	0,97	4%	13%
Ibitinga	7/3/03	832	11,67	2	20	677	26	1%	2,5%	0,26	0,98	0,97	3%	7%
Ibitinga	3/11/03	314	5,69	9	14	300	17	3%	4,5%	0,40	0,93	0,89	8%	27%
Ibitinga	4/4/04	560	2,75	-4	12	213	15	-1%	2,1%	0,24	0,98	0,98	3%	7%
Ibitinga	20/1/05	1.723	12,40	7	45	3.494	59	0%	2,5%	0,30	0,96	0,97	5%	13%
Ibitinga	12/5/06	375	0,55	0	8	118	11	0%	2,1%	0,57	0,83	0,84	18%	20%
Ibitinga	5/8/06	297	0,01	-2	7	85	9	-1%	2,4%	0,48	0,95	0,94	6%	20%
Ibitinga	15/1/07	1.396	12,09	-38	59	8.943	95	-3%	4,0%	0,31	0,94	0,95	7%	20%
Ibitinga	15/6/07	291	0,01	-6	7	56	7	-2%	2,5%	0,76	0,91	0,91	10%	20%
Ibitinga	15/12/09	1.451	9,95	11	78	11.074	105	1%	5,1%	0,30	0,96	0,97	7%	0%
Limoeiro	4/2/02	138	9,96	0	1	1	1	0%	0,6%	0,05	1,00	1,00	1%	0%
Limoeiro	7/3/03	114	13,06	0	1	1	1	0%	0,6%	0,08	1,00	1,00	1%	7%
Limoeiro	3/11/03	45	4,63	0	1	0	1	0%	1,3%	0,08	1,00	1,00	1%	0%
Limoeiro	4/4/04	123	2,88	0	0	0	1	0%	0,4%	0,05	1,00	1,00	0%	0%
Limoeiro	20/1/05	269	12,64	0	1	2	2	0%	0,4%	0,04	1,00	1,00	0%	0%
Limoeiro	12/5/06	56	0,13	0	0	0	0	0%	0,5%	0,24	0,97	0,97	3%	20%
Limoeiro	5/8/06	35	0,01	0	0	0	1	0%	1,4%	0,54	0,95	0,94	5%	47%
Limoeiro	15/1/07	391	18,13	1	6	55	7	0%	1,3%	0,08	1,00	1,00	1%	0%
Limoeiro	15/6/07	55	0,04	0	0	0	0	0%	0,4%	0,26	0,99	0,98	1%	27%
Limoeiro	15/1/08	166	10,42	3	4	20	4	2%	2,6%	0,16	0,98	0,97	3%	0%
Limoeiro	15/12/09	153	6,99	7	8	79	9	5%	5,2%	0,32	0,92	0,91	10%	13%
Nova Avanhandava	4/2/02	1.942	5,30	-396	396	231.416	481	-20%	19,6%	1,67	-0,02	0,39	104%	13%
Nova Avanhandava	7/3/03	1.046	9,44	-228	228	75.812	275	-21%	21,3%	1,68	0,40	0,61	64%	13%
Nova Avanhandava	3/11/03	417	1,79	-96	96	12.825	113	-22%	22,1%	1,86	-1,02	0,10	203%	40%
Nova Avanhandava	4/4/04	691	0,91	-156	156	26.497	163	-23%	22,6%	3,18	-1,31	0,30	232%	20%
Nova Avanhandava	20/1/05	2.223	7,15	-512	512	290.577	539	-23%	23,5%	2,62	-0,23	0,52	126%	20%
Nova Avanhandava	12/5/06	460	2,88	-99	99	10.004	100	-22%	21,5%	8,67	-13,53	0,05	1454%	53%
Nova Avanhandava	5/8/06	392	0,02	-72	72	5.461	74	-18%	18,2%	3,38	-0,58	0,37	159%	13%
Nova Avanhandava	15/1/07	2.064	15,27	-400	400	221.325	470	-19%	18,8%	1,65	-0,23	0,37	124%	20%
Nova Avanhandava	15/6/07	422	0,00	-86	86	7.590	87	-20%	20,4%	6,16	-5,03	0,14	603%	20%
Nova Avanhandava	15/1/08	1.192	13,25	-300	300	136.676	370	-24%	23,9%	1,49	-0,64	0,23	166%	33%
Nova Avanhandava	15/12/09	1.872	8,43	-437	457	345.913	588	-21%	22,5%	1,38	-0,01	0,42	104%	40%
Promissão	4/2/02	1.778	8,31	-135	144	69.018	263	-6%	6,8%	0,86	0,62	0,59	39%	27%
Promissão	7/3/03	984	12,13	42	65	5.209	72	5%	7,5%	0,49	0,95	0,91	9%	0%
Promissão	3/11/03	387	3,50	-10	23	1.354	37	-1%	5,3%	0,60	0,81	0,71	20%	27%
Promissão	4/4/04	653	2,55	-4	11	201	14	-1%	1,7%	0,26	0,98	0,98	3%	7%
Promissão	12/5/06	435	1,43	5	11	196	14	1%	2,5%	1,03	0,64	0,72	36%	27%
Promissão	5/8/06	363	0,01	-4	5	46	7	-1%	1,3%	0,32	0,98	0,98	2%	0%

Promissão	15/1/07	1.861	11,38	-100	109	22.324	149	-5%	5,7%	0,49	0,87	0,87	14%	20%
Promissão	15/6/07	384	0,00	-15	15	264	16	-4%	3,8%	1,17	0,75	0,78	25%	27%
Promissão	15/1/08	1.081	11,83	-45	127	29.205	171	-3%	11,3%	0,78	0,53	0,54	48%	20%
Promissão	15/12/09	1.768	12,09	-82	255	92.696	304	-2%	14,9%	0,76	0,73	0,73	31%	40%

Tabela A.3 – Índices de desempenho do modelo SINF para as previsões diárias.

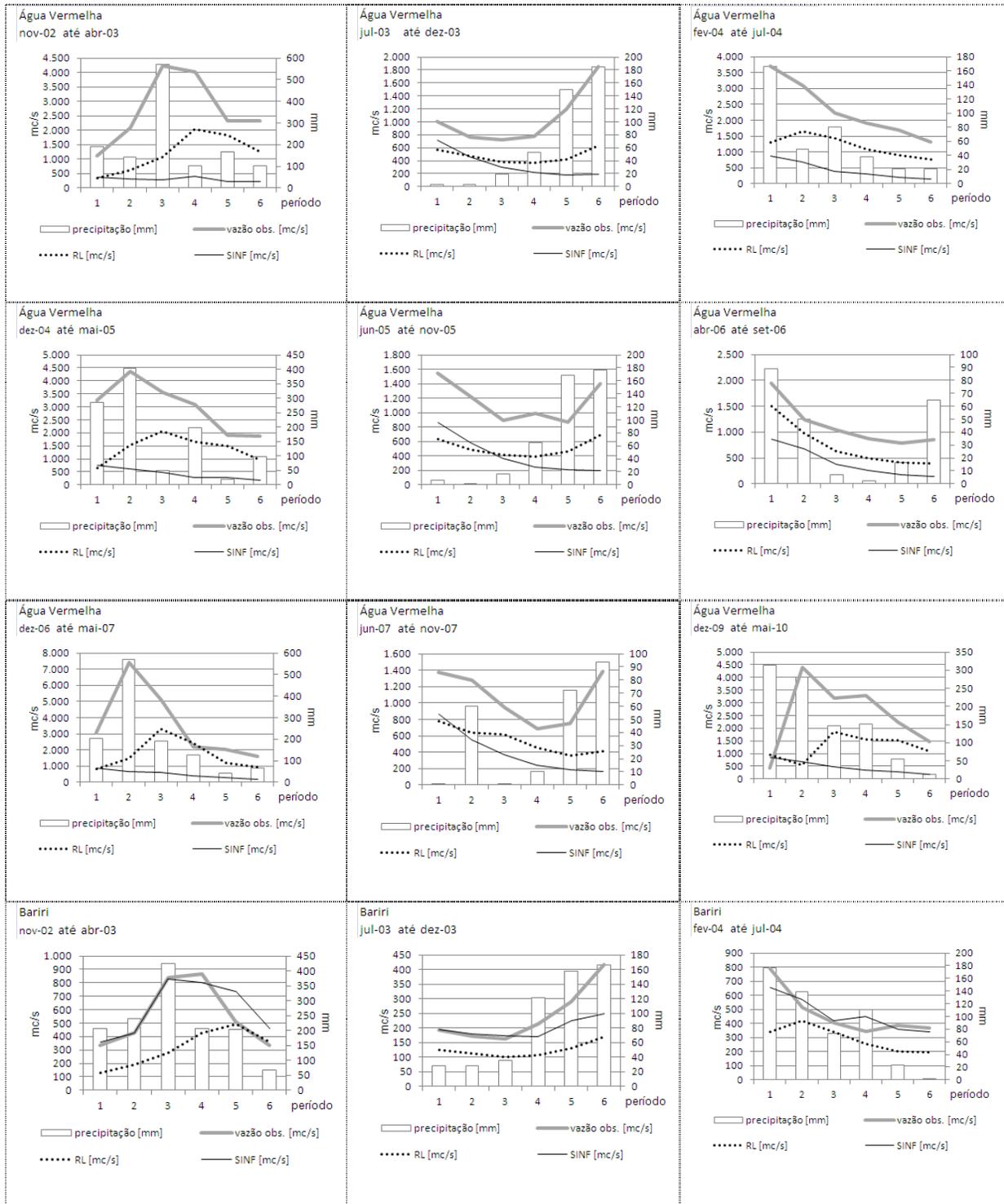
UHE	período	Q [m <sup>3</sup> /s]	P [mm]	ME [m <sup>3</sup> /s]	MAE [m <sup>3</sup> /s]	MSE [m <sup>6</sup> /m <sup>2</sup> ]	RMSE [m <sup>3</sup> /s]	MPE	MAPE	U- THEIL	NS	NS- LOG	DM	CI
Água Vermelha	4/2/02	4.250	6,54	-1.584	1.584	3.054.876	1.748	-35%	35,2%	8,56	-3,78	-0,05	479%	47%
Água Vermelha	7/3/03	3.064	10,39	-301	301	126.931	356	-9%	9,0%	3,15	0,51	0,55	49%	0%
Água Vermelha	3/11/03	944	5,10	137	146	38.753	197	14%	14,7%	5,53	-6,82	0,39	782%	27%
Água Vermelha	4/4/04	2.647	3,69	-566	566	477.205	691	-19%	19,5%	6,19	-1,36	-0,22	237%	27%
Água Vermelha	20/1/05	5.455	10,04	-1.810	1.810	3.874.016	1.968	-32%	31,6%	6,56	-5,65	-0,10	666%	53%
Água Vermelha	12/5/06	1.428	3,21	-130	142	43.659	209	-8%	9,0%	3,34	-0,43	-0,55	143%	53%
Água Vermelha	5/8/06	935	0,02	77	106	14.985	122	9%	12,0%	3,84	-0,93	-1,61	193%	53%
Água Vermelha	15/1/07	8.153	16,28	-1.069	1.455	2.755.201	1.660	-13%	17,6%	9,06	-23,40	-0,25	2441%	60%
Água Vermelha	15/6/07	1.408	0,00	29	29	1.066	33	2%	2,1%	1,70	0,64	0,74	36%	13%
Água Vermelha	15/1/08	3.233	14,88	-1.648	1.648	3.471.833	1.863	-49%	48,5%	2,25	-1,58	0,21	262%	27%
Bariri	4/2/02	1.163	9,55	-216	251	109.517	331	-16%	20,1%	2,53	-1,37	-0,18	238%	53%
Bariri	7/3/03	701	11,61	41	220	56.914	239	12%	32,8%	2,60	-1,26	-2,13	229%	47%
Bariri	3/11/03	270	6,25	122	143	32.089	179	50%	58,0%	3,95	-6,88	-0,13	790%	27%
Bariri	4/4/04	471	2,19	-95	96	16.079	127	-18%	18,1%	2,01	-0,87	-0,18	188%	13%
Bariri	20/1/05	1.460	15,23	-487	508	362.652	602	-30%	32,1%	3,45	-4,07	-0,24	508%	40%
Bariri	12/5/06	312	0,56	-44	44	2.370	49	-14%	14,0%	2,73	-2,06	0,11	306%	33%
Bariri	5/8/06	238	0,00	16	25	836	29	9%	11,5%	1,67	0,47	-0,95	54%	27%
Bariri	15/1/07	1.115	11,49	-141	246	102.674	320	-7%	20,6%	1,38	0,00	-0,77	102%	33%
Bariri	15/6/07	225	0,02	3	15	289	17	2%	6,9%	1,59	0,38	-1,49	63%	53%
Bariri	15/1/08	736	9,81	-58	156	36.641	191	-3%	20,3%	0,95	0,17	-1,19	86%	33%
Bariri	15/12/09	1.268	7,32	-386	386	282.072	531	-24%	24,3%	1,57	-0,32	0,04	134%	13%
Barra Bonita	4/2/02	1.041	6,57	-327	332	145.901	382	-30%	30,4%	2,96	-2,65	0,03	366%	33%
Barra Bonita	7/3/03	632	8,11	-23	159	37.333	193	3%	25,5%	1,71	-0,66	-4,38	168%	53%
Barra Bonita	3/11/03	246	4,39	102	113	18.958	138	51%	55,2%	2,14	-2,96	-0,50	400%	60%
Barra Bonita	4/4/04	427	2,38	-83	85	14.051	119	-16%	17,3%	1,39	-0,40	-0,20	141%	27%
Barra Bonita	20/1/05	1.324	13,99	-424	429	225.526	475	-30%	30,5%	2,41	-2,25	0,00	326%	33%
Barra Bonita	12/5/06	280	0,99	-51	51	3.432	59	-18%	17,8%	2,46	-2,69	0,09	370%	53%
Barra Bonita	5/8/06	206	0,02	27	27	1.165	34	15%	15,1%	1,51	0,28	0,08	73%	40%
Barra Bonita	15/1/07	968	7,34	-270	328	146.386	383	-22%	31,9%	1,76	-0,69	-0,22	172%	20%
Barra Bonita	15/6/07	191	0,28	-13	15	422	21	-6%	7,6%	1,39	0,11	-0,20	89%	33%
Barra Bonita	15/1/08	659	9,00	-43	131	25.460	160	-1%	19,2%	0,76	0,43	-0,61	60%	33%
Barra Bonita	15/12/09	1.143	9,39	-309	314	205.788	454	-22%	21,9%	1,38	0,02	0,32	101%	13%
Euclides da Cunha	4/2/02	136	9,84	-22	28	1.540	39	-12%	18,2%	1,73	-0,01	-0,28	103%	20%
Euclides da Cunha	7/3/03	112	13,23	9	12	279	17	10%	11,8%	1,54	0,17	-0,10	84%	33%
Euclides da Cunha	3/11/03	45	4,86	15	15	306	17	40%	39,6%	2,10	-0,96	-0,17	200%	33%
Euclides da Cunha	4/4/04	121	2,81	-42	42	2.032	45	-34%	33,7%	3,46	-7,19	-0,05	820%	33%
Euclides da	20/1/05	264	11,68	-19	68	5.604	75	-4%	25,7%	1,82	-0,95	-0,44	197%	33%

Cunha															
Euclides da Cunha	12/5/06	55	0,29	-1	3	8	3	-3%	4,7%	2,00	-0,54	-2,04	154%	47%	
Euclides da Cunha	5/8/06	34	0,01	4	4	18	4	12%	12,3%	4,74	-2,98	0,16	398%	60%	
Euclides da Cunha	15/1/07	386	18,56	-74	85	14.359	120	-16%	19,0%	1,24	0,06	0,33	96%	40%	
Euclides da Cunha	15/6/07	54	0,05	-3	3	13	4	-6%	5,8%	3,58	-0,88	-0,01	188%	53%	
Euclides da Cunha	15/1/08	167	10,34	-8	31	1.348	37	-1%	18,8%	1,44	-0,52	-2,98	153%	40%	
Euclides da Cunha	15/12/09	161	7,84	-20	30	1.809	43	-9%	15,9%	1,62	-0,62	-1,18	163%	33%	
Graminha	4/2/02	90	10,74	-21	23	896	30	-19%	22,3%	2,55	-0,57	-0,27	158%	20%	
Graminha	7/3/03	82	10,70	-11	12	246	16	-11%	13,6%	2,47	-0,33	-0,56	133%	27%	
Graminha	3/11/03	26	4,74	11	11	161	13	42%	42,3%	4,45	-7,65	0,16	866%	47%	
Graminha	4/4/04	77	2,96	-24	24	636	25	-30%	30,1%	4,88	-10,52	-0,06	1152%	53%	
Graminha	20/1/05	164	11,53	2	54	3.435	59	7%	35,7%	2,01	-1,89	-0,81	291%	27%	
Graminha	12/5/06	34	0,53	2	2	8	3	4%	6,1%	2,16	-1,55	0,46	255%	67%	
Graminha	5/8/06	22	0,00	-3	4	17	4	-15%	16,7%	8,00	-9,54	0,25	1054%	73%	
Graminha	15/1/07	217	17,50	-66	69	6.412	80	-29%	30,1%	2,08	-1,35	0,12	237%	33%	
Graminha	15/6/07	30	0,07	-2	2	4	2	-5%	5,2%	2,41	-3,48	0,17	448%	67%	
Graminha	15/1/08	116	9,51	2	29	1.099	33	5%	26,2%	1,70	-1,64	-0,94	266%	40%	
Graminha	15/12/09	106	9,09	-18	24	1.173	34	-13%	19,8%	2,20	-0,96	-1,01	197%	53%	
Ibitinga	4/2/02	1.386	10,26	-298	301	165.345	407	-19%	19,6%	2,62	-1,58	-0,26	259%	47%	
Ibitinga	7/3/03	832	11,67	23	231	64.393	254	9%	28,5%	2,56	-0,82	-2,51	184%	47%	
Ibitinga	3/11/03	314	5,69	139	144	34.391	185	48%	49,7%	4,47	-7,23	-0,13	824%	27%	
Ibitinga	4/4/04	560	2,75	-68	77	12.012	110	-11%	12,3%	1,85	-0,35	-0,22	136%	20%	
Ibitinga	20/1/05	1.723	12,40	-694	694	619.597	787	-38%	37,6%	3,96	-6,10	-0,14	711%	33%	
Ibitinga	12/5/06	375	0,55	-63	63	4.474	67	-17%	16,6%	3,80	-5,61	0,05	662%	27%	
Ibitinga	5/8/06	297	0,01	55	55	3.434	59	19%	19,5%	3,17	-1,14	-0,09	215%	33%	
Ibitinga	15/1/07	1.396	12,09	-235	344	206.036	454	-11%	22,1%	1,50	-0,40	-0,93	142%	27%	
Ibitinga	15/6/07	291	0,01	-8	14	327	18	-3%	4,9%	1,91	0,45	0,50	55%	40%	
Ibitinga	15/12/09	1.451	9,95	-544	544	461.284	679	-32%	32,5%	1,90	-0,84	0,07	187%	7%	
Limoeiro	4/2/02	138	9,96	-28	33	1.849	43	-16%	21,2%	1,95	-0,23	-0,16	125%	33%	
Limoeiro	7/3/03	114	13,06	0	10	169	13	1%	9,4%	1,30	0,49	-0,20	52%	33%	
Limoeiro	3/11/03	45	4,63	12	12	218	15	31%	31,1%	1,94	-0,50	-0,14	153%	40%	
Limoeiro	4/4/04	123	2,88	-41	41	1.951	44	-33%	32,5%	3,65	-7,02	-0,05	803%	33%	
Limoeiro	20/1/05	269	12,64	-27	66	5.536	74	-7%	24,1%	1,92	-1,03	-0,27	204%	33%	
Limoeiro	12/5/06	56	0,13	-10	10	126	11	-17%	17,3%	6,94	-19,71	-0,14	2071%	60%	
Limoeiro	5/8/06	35	0,01	3	3	8	3	7%	7,4%	2,94	-0,64	0,47	164%	47%	
Limoeiro	15/1/07	391	18,13	-67	90	15.004	122	-14%	20,2%	1,24	0,05	0,33	98%	40%	
Limoeiro	15/6/07	55	0,04	-11	11	133	12	-21%	20,6%	10,21	-15,40	0,09	1641%	53%	
Limoeiro	15/1/08	166	10,42	-12	34	1.544	39	-4%	20,5%	1,46	-0,49	-2,07	150%	33%	
Limoeiro	15/12/09	153	6,99	-5	33	1.705	41	1%	20,2%	1,49	-0,76	-3,38	177%	33%	
Nova Avanhandava	4/2/02	1.942	5,30	-713	714	796.079	892	-33%	32,9%	3,12	-2,51	-0,16	353%	20%	
Nova Avanhandava	7/3/03	1.046	9,44	-107	345	152.844	391	0%	33,6%	2,44	-0,22	-1,91	126%	33%	
Nova Avanhandava	3/11/03	417	1,79	104	120	16.619	129	28%	30,6%	2,16	-1,62	-0,58	264%	27%	
Nova Avanhandava	4/4/04	691	0,91	-125	126	21.662	147	-17%	16,9%	2,91	-0,89	-0,07	190%	40%	
Nova Avanhandava	20/1/05	2.223	7,15	-1.034	1.034	1.307.085	1.143	-43%	43,3%	5,62	-4,55	-0,07	557%	60%	
Nova Avanhandava	12/5/06	460	2,88	21	52	5.977	77	4%	10,9%	6,85	-7,68	0,26	868%	60%	
Nova Avanhandava	5/8/06	392	0,02	17	27	1.234	35	5%	7,7%	1,71	0,64	0,23	36%	27%	
Nova Avanhandava	15/1/07	2.064	15,27	314	420	279.233	528	18%	21,9%	1,92	-0,55	-0,28	156%	47%	

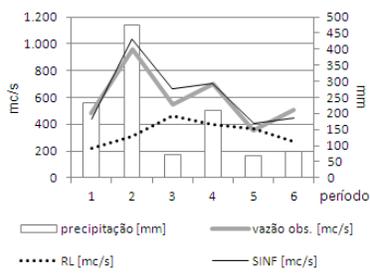
Nova Avanhandava	15/6/07	422	0,00	44	44	2.869	54	11%	11,2%	3,95	-1,28	-0,59	228%	53%
Nova Avanhandava	15/1/08	1.192	13,25	-94	270	106.734	327	-3%	21,7%	1,38	-0,28	-2,24	130%	73%
Nova Avanhandava	15/12/09	1.872	8,43	-1.153	1.153	1.720.432	1.312	-57%	57,0%	3,09	-4,03	-0,03	506%	33%
Promissão	4/2/02	1.778	8,31	-637	637	594.878	771	-33%	32,6%	2,53	-2,30	-0,12	332%	20%
Promissão	7/3/03	984	12,13	34	365	161.272	402	15%	41,5%	2,73	-0,66	-2,64	171%	47%
Promissão	3/11/03	387	3,50	140	145	29.168	171	40%	40,7%	2,79	-3,17	-0,26	419%	53%
Promissão	4/4/04	653	2,55	-44	65	9.205	96	-5%	8,9%	1,76	0,14	-0,15	87%	33%
Promissão	12/5/06	435	1,43	12	64	6.250	79	2%	14,3%	5,79	-10,51	0,24	1151%	67%
Promissão	5/8/06	363	0,01	113	113	15.105	123	34%	33,8%	5,90	-4,29	-0,32	530%	33%
Promissão	15/1/07	1.861	11,38	-38	384	228.342	478	1%	20,2%	1,60	-0,32	-0,25	133%	47%
Promissão	15/6/07	384	0,00	26	27	1.283	36	7%	7,6%	2,79	-0,22	-0,68	122%	40%
Promissão	15/1/08	1.081	11,83	-102	228	80.066	283	-5%	19,6%	1,34	-0,29	-1,56	130%	73%
Promissão	15/12/09	1.768	12,09	-793	793	910.442	954	-39%	39,4%	2,42	-1,61	0,00	264%	33%

# APÊNDICE B – ESTUDOS DE PREVISÃO DE VAZÕES MENSAIS

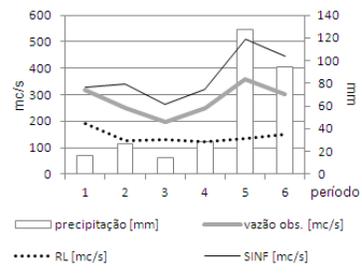
Tabela B.1 Gráficos contendo os resultados das execuções dos modelos RL e SINF, apresentados com as vazões e precipitações observadas.



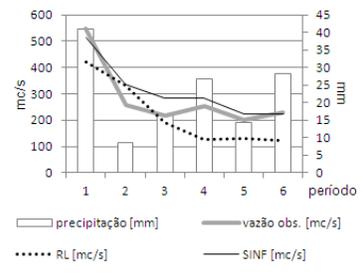
Bariri  
dez-04 até mai-05



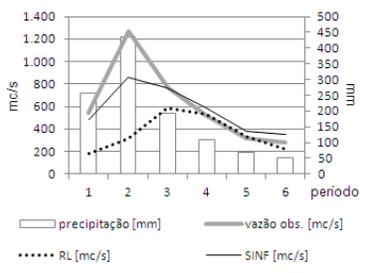
Bariri  
jun-05 até nov-05



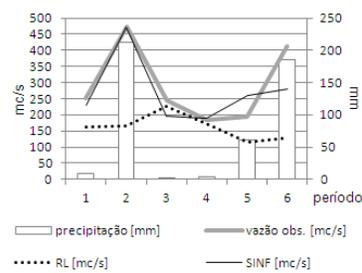
Bariri  
abr-06 até set-06



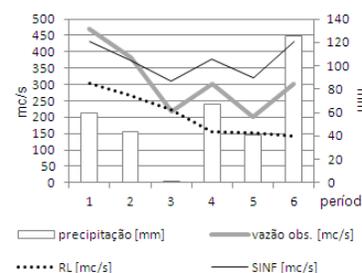
Bariri  
dez-06 até mai-07



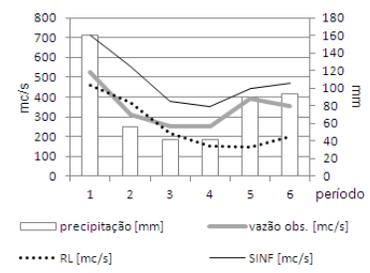
Bariri  
jun-07 até nov-07



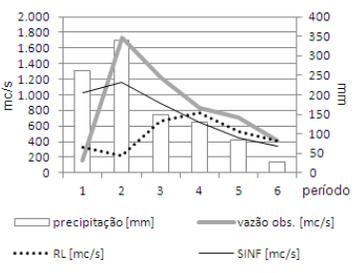
Bariri  
mai-08 até out-08



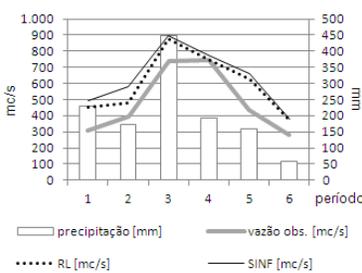
Bariri  
mar-09 até ago-09



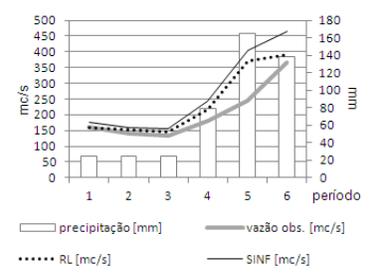
Bariri  
dez-09 até mai-10



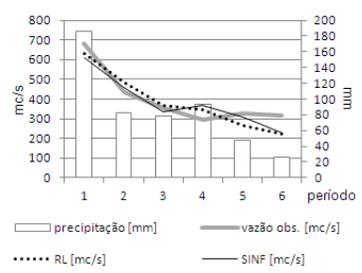
Barra Bonita  
nov-02 até abr-03



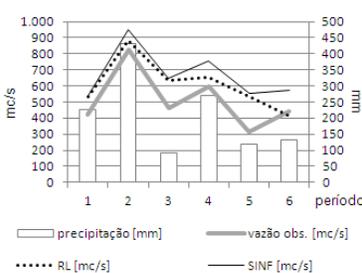
Barra Bonita  
jul-03 até dez-03



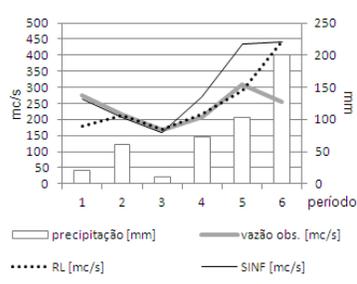
Barra Bonita  
fev-04 até jul-04



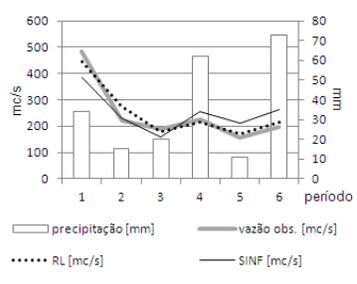
Barra Bonita  
dez-04 até mai-05



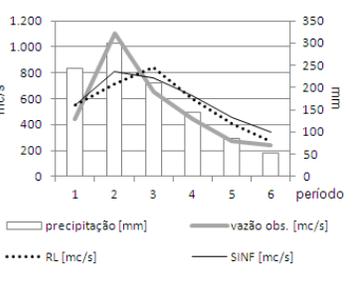
Barra Bonita  
jun-05 até nov-05



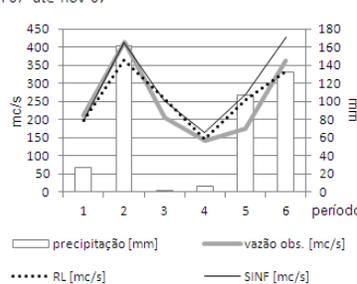
Barra Bonita  
abr-06 até set-06



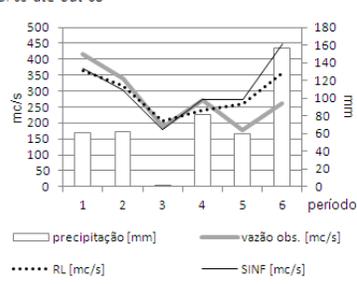
Barra Bonita  
dez-06 até mai-07



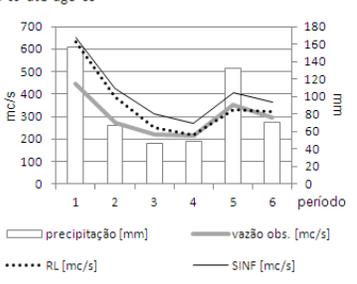
Barra Bonita  
jun-07 até nov-07



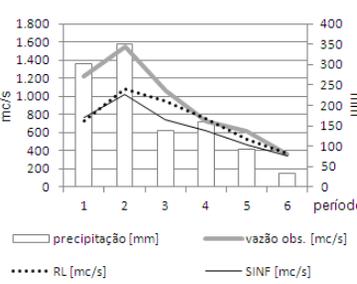
Barra Bonita  
mai-08 até out-08



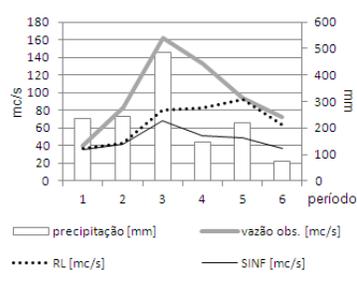
Barra Bonita  
mar-09 até ago-09



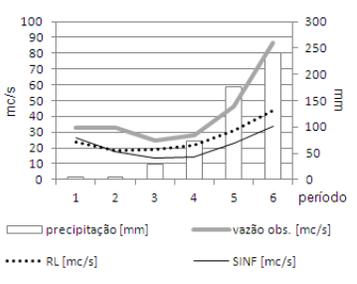
Barra Bonita  
dez-09 até mai-10



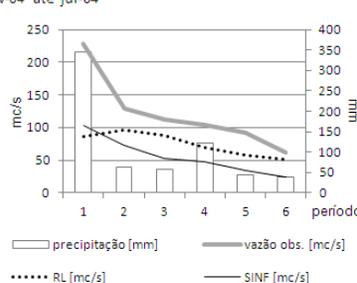
Euclides da Cunha  
nov-02 até abr-03



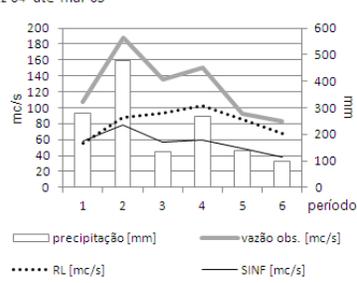
Euclides da Cunha  
jul-03 até dez-03



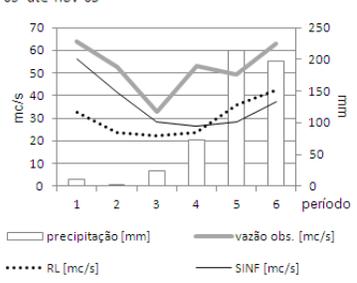
Euclides da Cunha  
fev-04 até jul-04



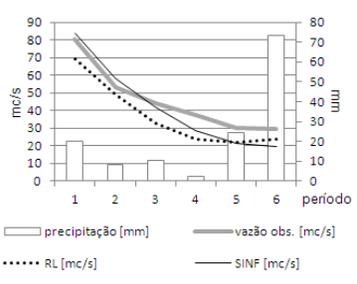
Euclides da Cunha  
dez-04 até mai-05



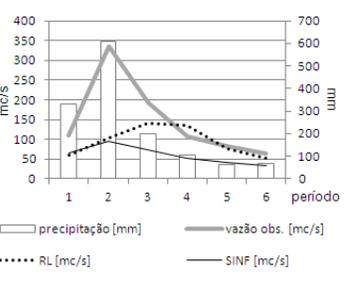
Euclides da Cunha  
jun-05 até nov-05



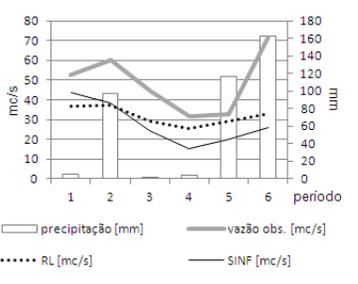
Euclides da Cunha  
abr-06 até set-06



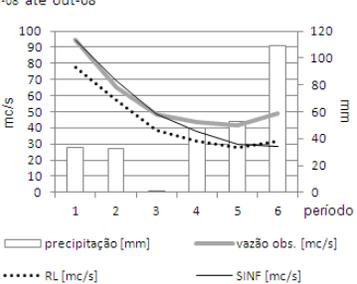
Euclides da Cunha  
dez-06 até mai-07



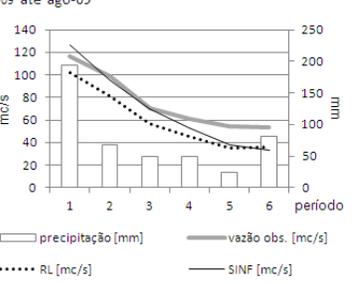
Euclides da Cunha  
jun-07 até nov-07



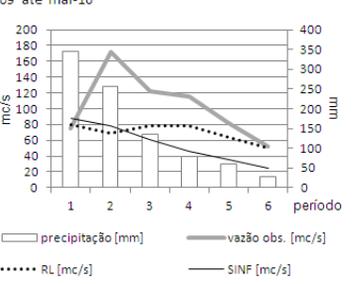
Euclides da Cunha  
mai-08 até out-08



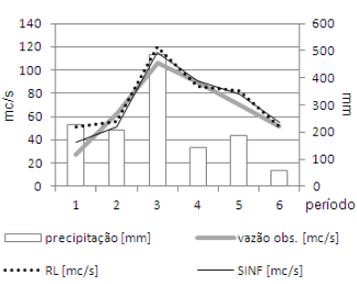
Euclides da Cunha  
mar-09 até ago-09



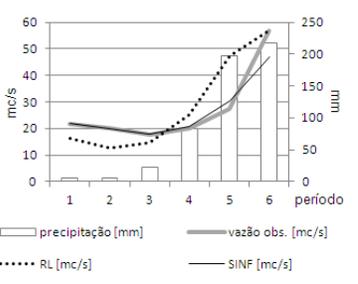
Euclides da Cunha  
dez-09 até mai-10



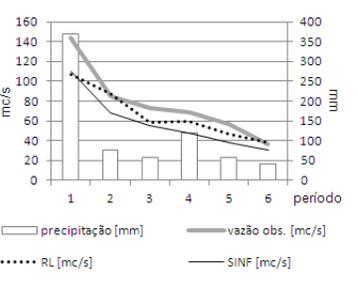
Graminha  
nov-02 até abr-03



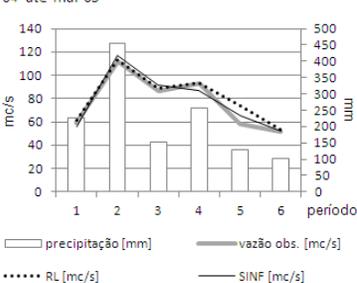
Graminha  
jul-03 até dez-03



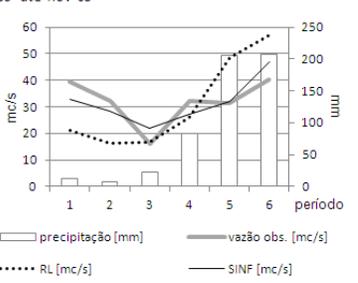
Graminha  
fev-04 até jul-04



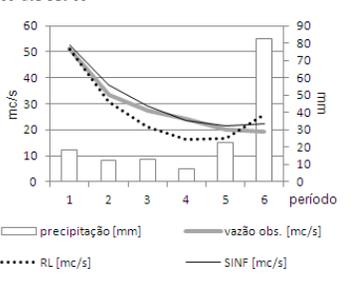
Graminha  
dez-04 até mai-05

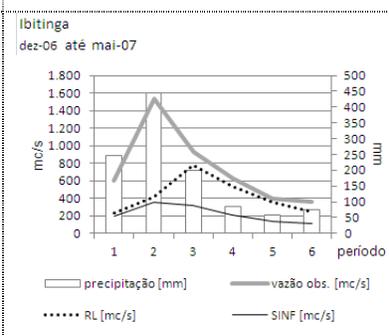
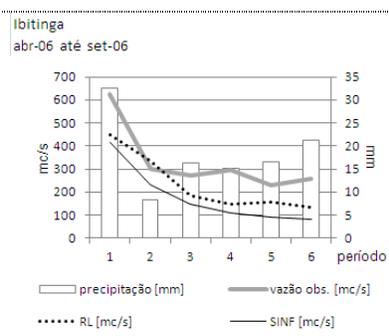
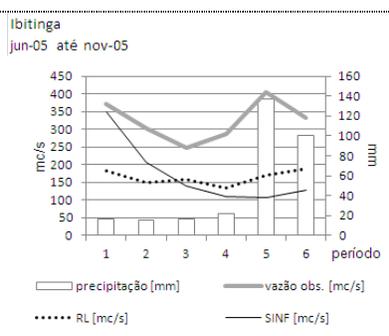
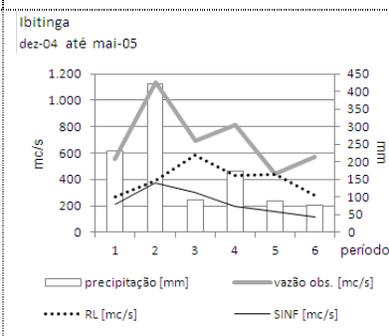
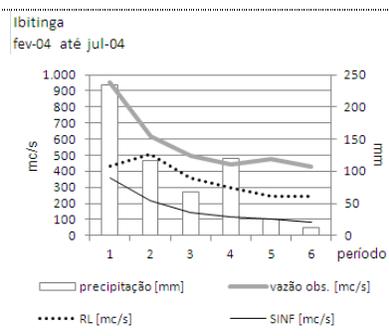
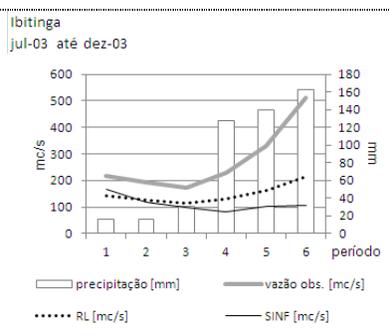
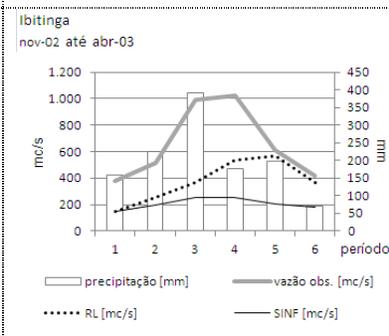
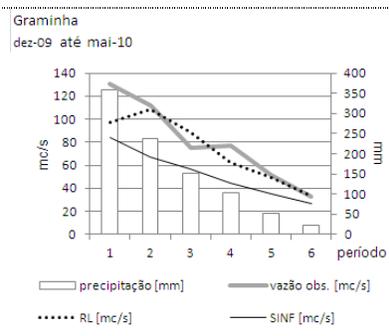
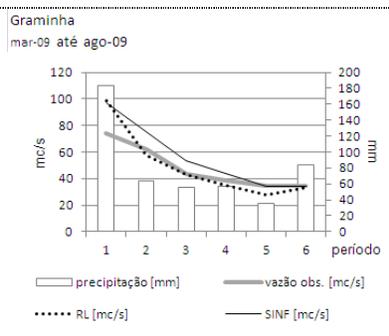
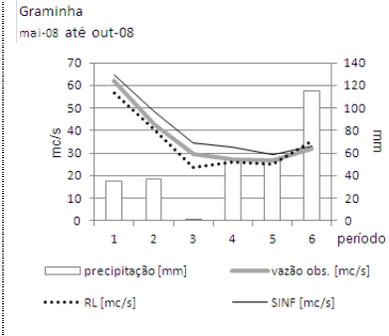
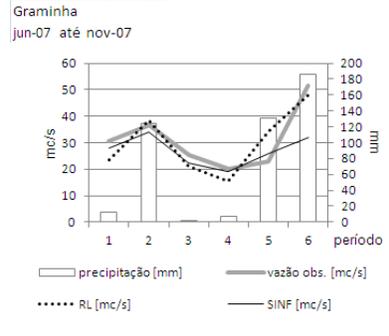
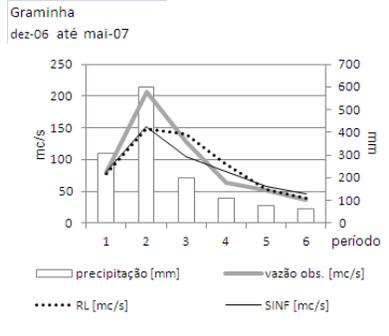


Graminha  
jun-05 até nov-05

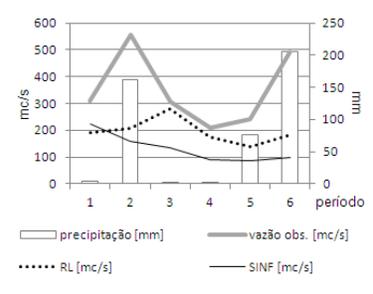


Graminha  
abr-06 até set-06

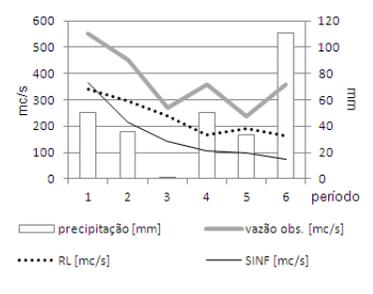




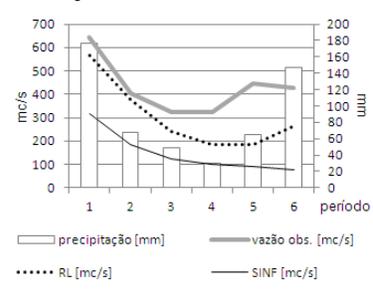
Ibitinga  
jun-07 até nov-07



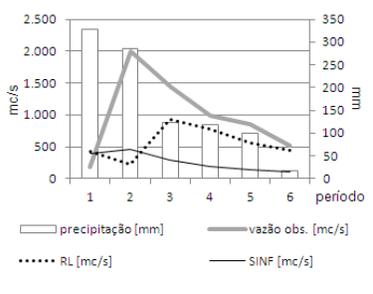
Ibitinga  
mai-08 até out-08



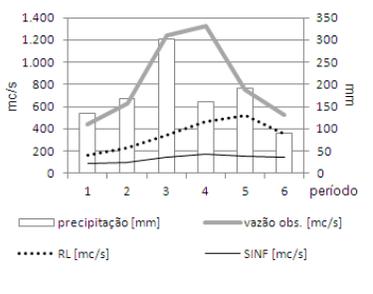
Ibitinga  
mar-09 até ago-09



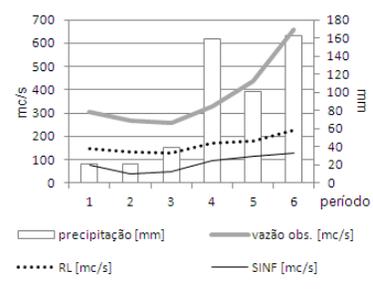
Ibitinga  
dez-09 até mai-10



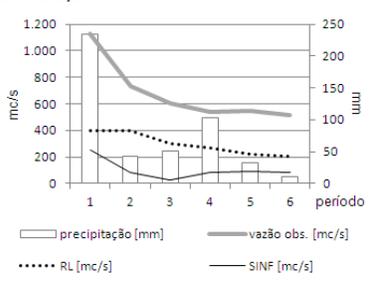
Promissão  
nov-02 até abr-03



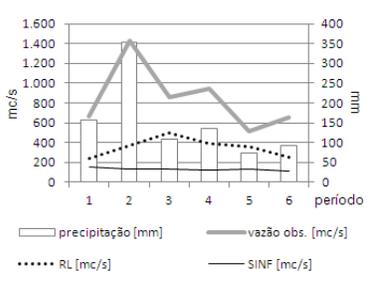
Promissão  
jul-03 até dez-03



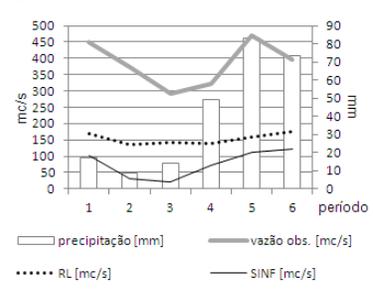
Promissão  
fev-04 até jul-04



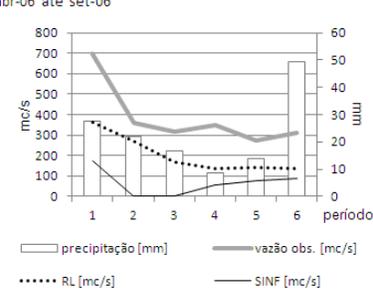
Promissão  
dez-04 até mai-05



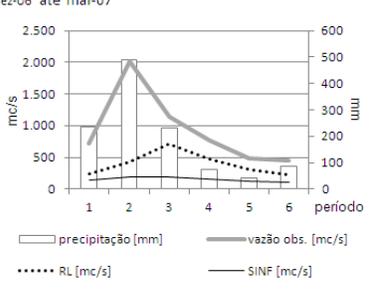
Promissão  
jun-05 até nov-05



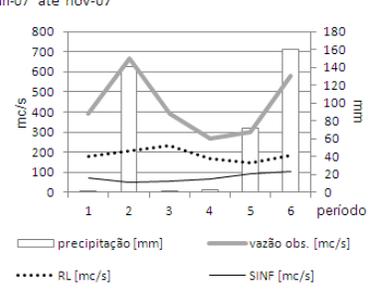
Promissão  
abr-06 até set-06



Promissão  
dez-06 até mai-07



Promissão  
jun-07 até nov-07



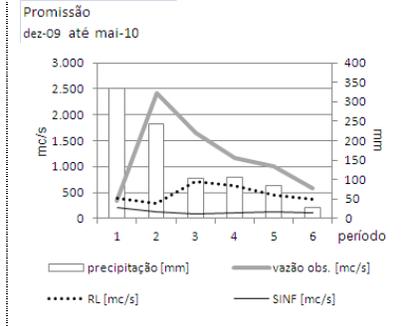
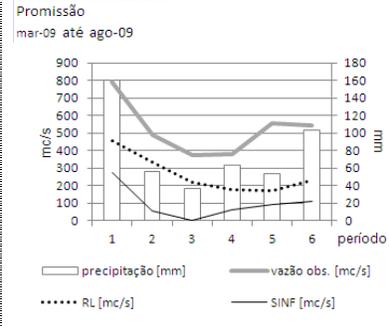
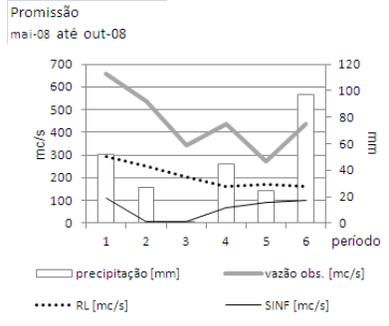


Tabela B.2 Índices de desempenho do modelo RL para as previsões mensais.

UHE	período	Q [m <sup>3</sup> /s]	P [mm]	ME [m <sup>3</sup> /s]	MAE [m <sup>3</sup> /s]	MSE [m <sup>6</sup> /m <sup>2</sup> ]	RMSE [m <sup>3</sup> /s]	MPE	MAPE	U- THEIL	NS	NS- LOG	DM	CI
Água Vermelha	nov-02	2.682	85,01	1.489	1.489	2.997.343	1.731	-55%	55,1%	1,44	-1,49	0,29	255%	17%
Água Vermelha	jul-03	1.052	27,47	-579	579	437.374	661	-52%	52,2%	1,91	-1,84	0,12	288%	17%
Água Vermelha	fev-04	2.325	25,04	1.141	1.141	1.697.404	1.303	-47%	46,7%	1,77	-1,48	0,21	252%	17%
Água Vermelha	dez-04	3.009	70,13	1.622	1.622	3.347.764	1.830	-51%	51,0%	1,96	-3,36	0,05	439%	17%
Água Vermelha	jun-05	1.152	29,05	-639	639	437.208	661	-55%	55,0%	1,84	-5,57	0,10	660%	33%
Água Vermelha	abr-06	1.120	15,31	-385	385	152.995	391	-38%	37,9%	1,11	0,05	0,47	103%	17%
Água Vermelha	dez-06	3.576	79,77	1.874	1.972	7.507.312	2.740	-43%	47,5%	1,09	-0,75	0,15	182%	17%
Água Vermelha	jun-07	1.072	15,77	-529	529	340.886	584	-47%	47,4%	1,67	-3,02	0,09	405%	17%
Água Vermelha	dez-09	2.506	63,89	1.257	1.423	3.435.388	1.853	-21%	59,2%	1,04	-1,05	-0,32	213%	50%
Bariri	nov-02	553	89,71	-240	248	101.362	318	-39%	40,9%	1,30	-1,10	0,17	214%	17%
Bariri	jul-03	241	35,83	-117	117	18.051	134	-45%	45,2%	2,04	-1,30	0,15	235%	0%
Bariri	fev-04	468	34,92	-175	175	47.984	219	-35%	35,1%	0,93	-0,99	0,18	202%	33%
Bariri	dez-04	591	75,86	-255	258	107.167	327	-38%	38,5%	1,02	-1,92	-0,07	295%	50%
Bariri	jun-05	279	20,49	-136	136	20.606	144	-48%	47,6%	2,08	-6,32	0,04	734%	50%
Bariri	abr-06	284	8,99	-63	88	9.080	95	-23%	32,4%	0,65	0,37	0,47	71%	50%
Bariri	dez-06	617	73,93	-256	263	180.484	425	-30%	32,1%	1,04	-0,63	0,04	166%	17%
Bariri	jun-07	294	31,47	-132	132	31.761	178	-37%	37,2%	1,09	-1,62	-0,15	264%	33%
Bariri	mai-08	313	22,49	-105	107	15.097	123	-31%	32,1%	1,00	-0,77	0,20	180%	33%
Bariri	mar-09	349	32,14	-91	111	17.221	131	-26%	32,2%	1,17	-0,93	0,21	196%	33%
Bariri	dez-09	846	65,90	-358	414	443.765	666	-10%	45,1%	0,94	-0,65	-0,48	171%	50%
Barra Bonita	nov-02	484	83,97	109	109	15.471	124	27%	27,4%	0,54	0,57	0,37	51%	17%
Barra Bonita	jul-03	205	30,41	35	36	3.040	55	16%	16,4%	0,92	0,53	0,78	49%	0%
Barra Bonita	fev-04	401	34,16	-15	54	3.390	58	-4%	14,7%	0,49	0,81	0,74	24%	17%
Barra Bonita	dez-04	512	84,63	96	107	16.037	127	23%	26,0%	0,45	0,40	0,00	65%	17%
Barra Bonita	jun-05	238	31,25	15	53	7.681	88	7%	20,3%	1,34	-2,49	0,23	349%	33%
Barra Bonita	abr-06	246	14,36	4	24	898	30	4%	9,8%	0,22	0,93	0,88	12%	0%
Barra Bonita	dez-06	525	69,06	38	167	39.977	200	19%	30,8%	0,57	0,53	0,35	56%	17%
Barra Bonita	jun-07	252	28,96	7	39	2.117	46	8%	17,0%	0,32	0,79	0,61	27%	0%
Barra Bonita	mai-08	276	28,09	15	50	3.444	59	10%	20,5%	0,59	0,51	-0,07	53%	17%
Barra Bonita	mar-09	300	34,75	56	63	8.304	91	17%	18,9%	0,54	-0,32	0,64	133%	0%
Barra Bonita	dez-09	920	71,61	-187	203	81.131	285	-15%	17,4%	0,68	0,49	0,65	54%	0%
Euclides da Cunha	nov-02	97	93,64	-31	31	1.844	43	-27%	26,7%	1,00	-0,15	0,37	118%	33%
Euclides da Cunha	jul-03	42	35,14	-16	16	400	20	-33%	33,3%	1,07	0,10	0,45	96%	17%
Euclides da Cunha	fev-04	121	44,54	-46	46	4.051	64	-33%	32,8%	0,61	-0,49	0,20	153%	17%
Euclides da Cunha	dez-04	126	92,96	-44	44	2.824	53	-32%	31,7%	1,05	-1,16	0,09	218%	17%
Euclides da Cunha	jun-05	52	34,78	-22	22	562	24	-42%	41,8%	1,46	-4,33	0,11	535%	17%
Euclides da Cunha	abr-06	46	9,29	-9	9	91	10	-21%	21,4%	0,68	0,70	0,72	37%	17%
Euclides da Cunha	dez-06	149	91,60	-54	63	9.940	100	-24%	31,9%	0,85	-0,16	0,22	120%	17%
Euclides da Cunha	jun-07	49	25,76	-17	17	420	20	-31%	31,0%	1,07	-1,05	0,05	208%	0%
Euclides da Cunha	mai-08	57	18,33	-13	13	180	13	-24%	24,3%	0,79	0,48	0,62	58%	0%
Euclides da Cunha	mar-09	76	31,10	-16	16	272	16	-24%	24,0%	1,06	0,52	0,66	54%	17%
Euclides da Cunha	dez-09	103	60,19	-33	35	2.400	49	-25%	27,2%	1,01	-0,59	-0,10	161%	50%
Graminha	nov-02	68	87,32	6	10	161	13	17%	22,3%	0,31	0,75	0,24	34%	0%

Graminha	jul-03	27	35,40	2	7	85	9	3%	29,5%	0,72	0,53	0,68	55%	17%
Graminha	fev-04	77	47,92	-11	13	306	17	-12%	13,9%	0,31	0,73	0,80	31%	17%
Graminha	dez-04	77	88,19	4	4	44	7	7%	6,8%	0,24	0,91	0,85	11%	0%
Graminha	jun-05	32	35,84	-1	13	202	14	-2%	36,3%	1,15	-2,21	0,23	323%	33%
Graminha	abr-06	29	10,43	-2	5	28	5	-8%	19,5%	0,67	0,77	0,67	30%	33%
Graminha	dez-06	94	90,54	-2	18	746	27	7%	16,1%	0,40	0,78	0,81	27%	0%
Graminha	jun-07	31	30,83	-1	5	39	6	-3%	20,6%	0,43	0,65	0,65	41%	0%
Graminha	mai-08	37	19,75	-2	3	14	4	-5%	9,5%	0,32	0,91	0,88	13%	17%
Graminha	mar-09	48	31,94	1	7	114	11	-1%	12,2%	0,40	0,50	0,87	51%	17%
Graminha	dez-09	80	61,82	-7	11	261	16	-5%	11,9%	0,39	0,76	0,86	26%	33%
Ibitinga	nov-02	655	81,10	-281	281	124.720	353	-40%	40,0%	1,24	-0,85	0,27	189%	17%
Ibitinga	jul-03	276	32,85	-127	127	22.975	152	-42%	42,4%	1,68	-0,71	0,23	176%	0%
Ibitinga	fev-04	569	38,48	-222	222	68.414	262	-37%	37,5%	1,03	-1,08	0,21	211%	50%
Ibitinga	dez-04	701	72,28	-304	304	147.346	384	-39%	38,7%	1,06	-1,83	0,00	286%	67%
Ibitinga	jun-05	324	20,45	-161	161	27.704	166	-49%	48,7%	2,15	-9,02	0,03	1003%	50%
Ibitinga	abr-06	330	7,34	-96	106	13.698	117	-30%	33,3%	0,67	0,25	0,46	82%	50%
Ibitinga	dez-06	741	73,91	-312	312	236.340	486	-34%	34,4%	0,97	-0,48	0,21	152%	17%
Ibitinga	jun-07	353	29,91	-157	157	40.549	201	-39%	38,5%	1,08	-1,50	-0,01	253%	33%
Ibitinga	mai-08	371	18,71	-138	138	24.514	157	-35%	35,1%	1,12	-1,17	0,16	219%	50%
Ibitinga	mar-09	428	35,67	-125	125	21.275	146	-31%	30,6%	1,23	-0,83	0,31	186%	17%
Ibitinga	dez-09	998	64,90	-437	517	602.234	776	-10%	54,5%	0,95	-0,71	-0,48	179%	33%
Promissão	nov-02	822	69,91	-478	478	315.934	562	-55%	55,0%	1,50	-1,68	0,19	273%	17%
Promissão	jul-03	374	33,67	-209	209	55.300	235	-54%	53,6%	2,15	-1,87	0,14	292%	0%
Promissão	fev-04	678	31,71	-378	378	168.082	410	-55%	55,0%	1,65	-2,68	0,14	372%	33%
Promissão	dez-04	845	61,57	-494	494	322.465	568	-55%	55,1%	1,25	-2,63	0,07	367%	50%
Promissão	jun-05	383	16,36	-230	230	55.837	236	-59%	59,3%	2,48	-12,51	0,02	1353%	50%
Promissão	abr-06	385	9,20	-182	182	39.326	198	-47%	47,5%	1,00	-0,94	0,26	200%	50%
Promissão	dez-06	928	77,03	-529	529	511.865	715	-51%	50,9%	1,04	-0,80	0,23	187%	17%
Promissão	jun-07	433	25,14	-247	247	78.822	281	-53%	53,5%	1,31	-2,79	0,05	383%	33%
Promissão	mai-08	447	16,34	-239	239	65.518	256	-52%	52,0%	1,52	-3,26	0,10	429%	50%
Promissão	mar-09	522	31,46	-257	257	74.231	272	-49%	49,2%	1,57	-2,81	0,15	384%	50%
Promissão	dez-09	1.196	59,74	-720	736	1.012.481	1.006	-45%	49,3%	1,07	-1,14	-0,04	219%	33%

Tabela B.3 – Índices de desempenho do modelo SINF para as previsões mensais.

UHE	período	Q [m <sup>3</sup> /s]	P [mm]	ME [m <sup>3</sup> /s]	MAE [m <sup>3</sup> /s]	MSE [m <sup>6</sup> /m <sup>2</sup> ]	RMSE [m <sup>3</sup> /s]	MPE	MAPE	U- THEIL	NS	NS- LOG	DM	CI
Água Vermelha	nov-02	2.682	85,01	-2.377	2.377	6.850.588	2.617	-86%	85,8%	2,20	-4,70	0,04	576%	67%
Água Vermelha	jul-03	1.052	27,47	-704	704	734.293	857	-62%	61,9%	2,55	-3,76	-0,05	480%	33%
Água Vermelha	fev-04	2.325	25,04	-1.895	1.895	3.918.194	1.979	-83%	83,0%	3,31	-4,73	0,14	578%	0%
Água Vermelha	dez-04	3.009	70,13	-2.569	2.569	7.147.402	2.673	-86%	85,8%	3,25	-8,31	0,08	935%	17%
Água Vermelha	jun-05	1.152	29,05	-741	741	597.412	773	-65%	65,3%	2,43	-7,98	0,08	901%	33%
Água Vermelha	abr-06	1.120	15,31	-706	706	529.706	728	-66%	66,0%	1,85	-2,30	0,21	337%	17%
Água Vermelha	dez-06	3.576	79,77	-3.077	3.077	13.231.711	3.638	-85%	84,6%	1,50	-2,09	0,16	320%	17%
Água Vermelha	jun-07	1.072	15,77	-676	676	523.691	724	-64%	64,0%	2,18	-5,18	0,10	622%	33%
Água Vermelha	dez-09	2.506	63,89	-2.036	2.177	5.937.773	2.437	-56%	88,5%	1,37	-2,54	-0,01	365%	33%
Bariri	nov-02	553	89,71	49	75	11.731	108	13%	16,3%	0,46	0,76	0,63	29%	17%
Bariri	jul-03	241	35,83	-43	50	5.823	76	-12%	16,1%	1,19	0,26	-0,08	76%	17%
Bariri	fev-04	468	34,92	-2	62	6.033	78	2%	12,8%	0,41	0,75	0,60	28%	33%
Bariri	dez-04	591	75,86	26	58	4.593	68	4%	10,8%	0,22	0,87	0,87	17%	0%
Bariri	jun-05	279	20,49	89	89	10.270	101	32%	31,9%	1,57	-2,65	0,23	366%	17%
Bariri	abr-06	284	8,99	27	40	2.247	47	13%	15,7%	0,37	0,84	0,65	22%	17%
Bariri	dez-06	617	73,93	-46	109	30.010	173	2%	16,1%	0,45	0,73	0,69	32%	0%
Bariri	jun-07	294	31,47	-23	46	4.036	64	-4%	16,1%	0,39	0,67	0,57	37%	0%
Bariri	mai-08	313	22,49	62	77	7.756	88	27%	29,9%	0,85	0,09	-0,90	96%	0%
Bariri	mar-09	349	32,14	135	135	22.317	149	41%	41,0%	1,19	-1,51	0,10	254%	17%
Bariri	dez-09	846	65,90	-97	384	216.731	466	66%	112,0%	0,43	0,19	-1,73	138%	0%
Barra Bonita	nov-02	484	83,97	146	146	25.844	161	36%	36,5%	0,70	0,28	0,15	81%	17%
Barra Bonita	jul-03	205	30,41	64	64	6.842	83	28%	28,0%	1,37	-0,05	0,66	108%	0%
Barra Bonita	fev-04	401	34,16	-17	49	3.190	56	-3%	13,1%	0,44	0,82	0,72	22%	33%
Barra Bonita	dez-04	512	84,63	157	157	26.603	163	36%	35,5%	0,60	0,01	-0,04	105%	0%
Barra Bonita	jun-05	238	31,25	58	66	9.091	95	22%	25,8%	1,62	-3,12	0,49	413%	17%
Barra Bonita	abr-06	246	14,36	5	47	3.183	56	8%	19,7%	0,33	0,73	0,37	33%	0%
Barra Bonita	dez-06	525	69,06	66	164	31.109	176	28%	36,9%	0,50	0,63	-0,23	52%	0%
Barra Bonita	jun-07	252	28,96	35	41	2.670	52	17%	19,8%	0,35	0,74	0,63	33%	0%
Barra Bonita	mai-08	276	28,09	32	63	7.985	89	17%	25,8%	0,94	-0,14	0,00	117%	17%
Barra Bonita	mar-09	300	34,75	106	106	14.463	120	35%	35,0%	0,90	-1,30	0,32	232%	0%
Barra Bonita	dez-09	920	71,61	-261	261	102.903	321	-24%	24,1%	0,89	0,35	0,64	69%	0%
Euclides da Cunha	nov-02	97	93,64	-51	51	3.441	59	-47%	46,6%	1,37	-1,15	0,19	220%	0%
Euclides da Cunha	jul-03	42	35,14	-20	20	649	25	-45%	45,0%	1,37	-0,46	0,31	153%	0%
Euclides da Cunha	fev-04	121	44,54	-66	66	5.047	71	-55%	54,8%	1,15	-0,86	0,34	194%	0%
Euclides da Cunha	dez-04	126	92,96	-69	69	5.448	74	-54%	53,8%	1,53	-3,17	0,13	420%	0%
Euclides da Cunha	jun-05	52	34,78	-16	16	335	18	-30%	30,0%	1,32	-2,18	0,17	319%	33%
Euclides da Cunha	abr-06	46	9,29	-3	6	50	7	-13%	17,6%	0,56	0,84	0,81	24%	0%
Euclides da Cunha	dez-06	149	91,60	-89	89	13.250	115	-54%	54,5%	0,99	-0,55	0,34	164%	0%
Euclides da Cunha	jun-07	49	25,76	-21	21	569	24	-42%	41,7%	1,30	-1,78	0,23	281%	17%
Euclides da Cunha	mai-08	57	18,33	-6	7	101	10	-13%	15,0%	0,70	0,71	0,69	33%	17%
Euclides da Cunha	mar-09	76	31,10	-6	10	142	12	-13%	15,5%	0,78	0,75	0,76	30%	0%
Euclides da Cunha	dez-09	103	60,19	-47	52	3.374	58	-43%	48,5%	1,20	-1,24	0,23	229%	17%
Graminha	nov-02	68	87,32	4	8	71	8	8%	14,5%	0,27	0,89	0,79	18%	0%
Graminha	jul-03	27	35,40	-1	2	17	4	0%	6,0%	0,33	0,91	0,93	11%	0%
Graminha	fev-04	77	47,92	-19	19	436	21	-25%	24,6%	0,57	0,61	0,71	46%	0%

Graminha	dez-04	77	88,19	2	4	23	5	2%	5,9%	0,16	0,95	0,95	8%	17%
Graminha	jun-05	32	35,84	0	5	27	5	2%	16,7%	0,43	0,57	0,38	46%	17%
Graminha	abr-06	29	10,43	2	2	5	2	7%	7,7%	0,28	0,96	0,93	9%	17%
Graminha	dez-06	94	90,54	-7	19	664	26	4%	18,5%	0,38	0,80	0,76	27%	0%
Graminha	jun-07	31	30,83	-4	5	68	8	-10%	14,1%	0,64	0,38	0,33	64%	0%
Graminha	mai-08	37	19,75	4	4	17	4	11%	11,3%	0,41	0,89	0,83	16%	0%
Graminha	mar-09	48	31,94	9	9	135	12	15%	15,2%	0,75	0,41	0,82	60%	17%
Graminha	dez-09	80	61,82	-28	28	981	31	-32%	32,3%	1,17	0,12	0,57	94%	17%
Ibitinga	nov-02	655	81,10	-449	449	251.234	501	-66%	65,7%	1,79	-2,72	0,12	378%	0%
Ibitinga	jul-03	276	32,85	-164	164	42.200	205	-53%	53,1%	2,31	-2,14	-0,03	319%	17%
Ibitinga	fev-04	569	38,48	-400	400	167.881	410	-72%	71,9%	2,22	-4,10	0,14	515%	17%
Ibitinga	dez-04	701	72,28	-474	474	252.018	502	-67%	67,4%	1,40	-3,84	0,12	489%	33%
Ibitinga	jun-05	324	20,45	-151	151	30.544	175	-46%	46,5%	2,54	-10,05	0,03	1106%	50%
Ibitinga	abr-06	330	7,34	-149	149	24.399	156	-48%	48,5%	0,95	-0,34	0,34	142%	33%
Ibitinga	dez-06	741	73,91	-519	519	370.039	608	-68%	68,4%	1,23	-1,32	0,26	242%	0%
Ibitinga	jun-07	353	29,91	-221	221	64.772	255	-60%	59,7%	1,40	-3,00	0,08	404%	33%
Ibitinga	mai-08	371	18,71	-204	204	44.973	212	-57%	56,9%	1,70	-2,97	0,18	401%	33%
Ibitinga	mar-09	428	35,67	-279	279	82.523	287	-66%	66,3%	2,20	-6,10	0,09	713%	17%
Ibitinga	dez-09	998	64,90	-731	802	838.048	915	-47%	85,9%	1,12	-1,38	0,05	253%	0%
Promissão	nov-02	822	69,91	-687	687	576.729	759	-82%	81,9%	2,04	-3,88	0,08	495%	0%
Promissão	jul-03	374	33,67	-290	290	96.549	311	-77%	77,4%	2,81	-4,00	0,12	506%	17%
Promissão	fev-04	678	31,71	-574	574	353.917	595	-85%	85,5%	2,78	-6,74	0,06	779%	17%
Promissão	dez-04	845	61,57	-715	715	598.838	774	-83%	82,9%	1,72	-5,74	0,03	679%	67%
Promissão	jun-05	383	16,36	-306	306	95.585	309	-81%	80,7%	3,30	-22,14	0,05	2315%	17%
Promissão	abr-06	385	9,20	-238	238	58.555	242	-76%	75,8%	4,75	-1,31	-2,48	243%	67%
Promissão	dez-06	928	77,03	-779	779	864.683	930	-81%	81,2%	1,36	-2,03	0,13	314%	17%
Promissão	jun-07	433	25,14	-360	360	151.213	389	-81%	81,0%	1,80	-6,28	0,01	732%	50%
Promissão	mai-08	447	16,34	-384	384	163.115	404	-85%	85,1%	2,45	-9,61	0,02	1064%	33%
Promissão	mar-09	522	31,46	-719	719	317.261	563	133%	133,1%	2,37	-6,16	-0,12	729%	33%
Promissão	dez-09	1.196	59,74	-1.073	1.073	1.647.547	1.284	-82%	82,3%	1,36	-2,48	0,04	357%	50%