

Maiana Santos Lopes

SISTEMA DE SUPORTE PARA PREVISÃO E GERAÇÃO DE SÉRIES
SINTÉTICAS DE VAZÕES

Campinas
2014



Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Maiana Santos Lopes

SISTEMA DE SUPORTE PARA PREVISÃO E GERAÇÃO DE SÉRIES SINTÉTICAS DE VAZÕES

Dissertação de mestrado apresentada a Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos exigidos para a obtenção do título de Mestra em Engenharia Elétrica, área de concentração: Energia Elétrica.

Orientador: Secundino Soares Filho
Co-orientadora: Ivette Luna Huamani

Este exemplar corresponde versão final da dissertação defendida pela aluna, e orientada pelo Prof. Dr. Secundino Soares Filho

Campinas
2014

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

L881s Lopes, Maiana Santos, 1985-
Sistema de suporte para previsão e geração de séries sintéticas de vazões /
Maiana Santos Lopes. – Campinas, SP : [s.n.], 2014.

Orientador: Secundino Soares Filho.
Coorientador: Ivette Luna Huamani.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de
Engenharia Elétrica e de Computação.

1. Previsão de vazões. 2. Sistema de suporte a decisão. 3. Usinas
hidrelétricas. 4. Análise de séries temporais. I. Soares Filho, Secundino, 1949-. II.
Huamani, Ivette Luna. III. Universidade Estadual de Campinas. Faculdade de
Engenharia Elétrica e de Computação. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Support system for prediction and generation of synthetic series
streamflow

Palavras-chave em inglês:

Streamflow forecasting
Decision support system
Hydroelectric plants
Time series analysis

Área de concentração: Energia Elétrica

Titulação: Mestra em Engenharia Elétrica

Banca examinadora:

Secundino Soares Filho [Orientador]
Cássio Dener Noronha Vinhal
Takaaki Ohishi

Data de defesa: 07-03-2014

Programa de Pós-Graduação: Engenharia Elétrica

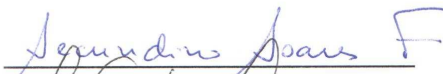
COMISSÃO JULGADORA - TESE DE MESTRADO

Candidata: Maiana Santos Lopes

Data da Defesa: 7 de março de 2014

Título da Tese: "Sistema de Suporte para Previsão e Geração de Séries Sintéticas de Vazões"

Prof. Dr. Secundino Soares Filho (Presidente):



Prof. Dr. Cássio Dener Noronha Vinhal:



Prof. Dr. Takaaki Ohishi:



Resumo

A previsão de vazões médias mensais é um insumo fundamental para o planejamento da operação das usinas hidrelétricas do Sistema Interligado Nacional (SIN). Durante os últimos anos, diferentes modelos baseados em inteligência computacional têm sido sugeridos para esse problema. A principal contribuição desta dissertação é o desenvolvimento de um sistema de suporte para a previsão e geração de séries sintéticas de vazões mensais, necessárias para o planejamento da operação das usinas do SIN. Este sistema permite analisar o desempenho de modelos clássicos de geração de séries sintéticas e de previsão de vazões, permitindo comparações entre um conjunto específico de modelos clássicos de séries temporais e de inteligência computacional para todas as usinas hidrelétricas do SIN.

Palavras-chave: Previsão de vazões, geração de séries sintéticas, vazões mensais, sistema de suporte.

Abstract

The prediction of monthly average inflows is a fundamental input for the operation planning of the hydroelectric plants of the National Interconnected System (SIN). During the last years, different models based on computational intelligence have been suggested for this problem. The main contribution of this dissertation is the development of a support system for the prediction and generation of synthetic series of monthly average inflows, necessary for planning the operation of the plants of SIN. This system permits to analyze the performance of different models for forecasting and generation of synthetic series of inflows, allowing the comparison between a set of models based on classical time series and computational intelligence for all the hydropower plants of the SIN.

Key-words: Streamflow forecasting, generation of synthetic series, monthly inflow, support system.

Sumário

1	Introdução	1
1.1	Descrição do Problema	1
1.2	Objetivo	2
2	Previsão de Vazões	3
2.1	Introdução	3
2.2	Modelos de previsão de vazões	4
2.2.1	Modelos autorregressivos	4
2.2.2	Redes neurais multicamadas	9
2.2.3	Bases de regras nebulosas	17
2.2.4	Rede neurofuzzy	24
2.2.5	Previsões agregadas	34
3	Geração de Séries Sintéticas	37
3.1	Introdução	37
3.2	Modelos lineares de geração de séries sintéticas	38
3.2.1	Modelos log-normal com três parâmetros	38
3.2.2	Modelo periódico log-normal com três parâmetros $BCPAR(p_m)$	42
3.3	Modelos não-lineares de geração de séries sintéticas	43
4	HydroPrev	47
4.1	Introdução	47
4.2	HydroPrev	47
5	HydroSeries	63
5.1	Introdução	63
5.2	HydroSeries	63
6	Estudos de Caso	69
6.1	Previsão de Vazões	69
6.2	Geração de Séries Sintéticas	77
6.3	Previsão de vazões sobre séries sintéticas	80
7	Conclusões e Perspectivas	83

AOS MEUS PAIS LUIZ CARLOS E
MARIA LÚCIA

Agradecimentos

Agradeço,

primeiramente a Deus, por me proteger e guiar meus passos.

Ao Professor Secundino Soares Filho que tive o imensurável prazer de tê-lo como meu orientador. Meus sinceros agradecimentos pela orientação e pela experiência de vida que o senhor me proporcionou.

À minha co-orientadora Ivette Luna pela orientação, paciência, pelas conversas agradáveis e pela amizade.

À professora Rosângela Ballini(Unicamp), ao professor Marinho G. de Andrade Filho (USP) e ao professor Ricardo Menezes (Unifal) pelas contribuições deste trabalho

aos amigos do COSE André Toscano e Mônica Zambelli que deram o “ponta pé” inicial e o incentivo no início deste trabalho, e ao Makoto Kadowaki que nunca mediu esforços em me auxiliar.

À “família” coseana Elma Pereira, Leonardo Martins, Marcos Leone, Anibal Azevedo, João Borsoi, Liz Alvarez, Luis Germán, Miguel Paredes, Mateus Neves, pelos adoráveis almoços e cafés.

Aos amigos de outros departamentos Silvio Segura, Floriano Torres, Laura Assis, Fernanda Trindade, Jorge Fernando, Hugo Valadares.

Aos amigos do Grupo de Oração Universitário.

Aos meus eternos “irmãos”Arlindo, Jordana e Patrícia. Patrícia minha irmã sempre presente, obrigada por ter sido minha “família” neste tempo em Campinas, por todo apoio e carinho. Arlindo e Jordana a distância nunca será uma dificuldade para nós. Obrigada pela presença constante em minha vida.

A Ilezi de Carvalho, minha amiga-mãe, por ter tornado pouco dolorosa a distância de casa com todo seu carinho e cuidado.

Ao meu amado namorado Cássio Fujisawa, pelo companheirismo e amor .

A Lígia Fujisawa, Kátia Fujisawa e Márcia Fujisawa.

Aos meus avós, tios e primos pelo incentivo e pela lição mais linda, família é a coisa mais importante que temos. Agradeço especialmente ao meu vovô João dos Santos Teodoro e minha querida bisavó Maria que não puderam ver a conclusão deste trabalho.

Aos meus pais, por serem meu porto seguro e pelo infinito amor e dedicação.

A educação tem raízes amargas, mas os seus frutos são doces.

Aristóteles

Lista de Figuras

2.1	Rede MLP de uma camada intermediária com bias.	10
2.2	Generalização dos neurônios de uma rede MLP para L=3	14
2.3	Estrutura geral do FIS.	19
2.4	Função de pertinência discretizada.	25
2.5	Modelo geral do neurônio nebuloso.	27
2.6	Rede neurofuzzy.	27
2.7	Função de decodificação.	28
2.8	Vazões anuais	35
3.1	A função densidade de probabilidade normal generalizada.	41
3.2	Geração de séries sintéticas	45
4.1	O programa HYDROLAB e seu módulo de previsões de vazões HYDROPREV.	48
4.2	Criando um novo estudo no HYDROPREV	49
4.3	Escolha do Posto Hidrológico	50
4.4	Cenário ajustado para o treinamento das redes	51
4.5	Médias mensais das vazões	52
4.6	Histórico das vazões	53
4.7	Execução do treinamento	54
4.8	Resultado das previsões para treinamento e teste	55
4.9	Trajectoria dos erros de previsões	56
4.10	Resultados vistos por rede	57
4.11	Erros percentuais absolutos médio de cada passo	58
4.12	Desvio padrão do resultado previsto	58
4.13	Desvio padrão - Passos à frente	59
4.14	Visualização dos resultados na forma de arquivo	59
4.15	Análise dos resultados com o auxílio das escalas horizontais	60
4.16	Recursos avançados - Propriedades aba Geral	60
4.17	Recursos avançados - Propriedades aba Rede	61
4.18	Módulo de Previsão Mensal	61
4.19	Módulo de Previsão Mensal	62
4.20	Vazões prevista pelo módulo Previsão Mensal	62

5.1	Módulo HYDROSERIES	64
5.2	Módulo HYDROSERIES	65
5.3	Dados de um estudo do módulo HYDROSERIES	66
5.4	Aba para visualização do histórico de vazões	66
5.5	Escolha do modelo RNF	67
5.6	Série Gerada	67
5.7	Aba para comparações estatísticas	68
5.8	Visualização dos resultados na forma de arquivos	68
6.1	Resultado do treinamento do modelo RNF para Furnas	70
6.2	REQM do mês de junho para o posto de Furnas	71
6.3	REQM do mês de dezembro para o posto de Furnas	71
6.4	Resultado preditivo da abordagem agregada	72
6.5	REQM das previsões da abordagem agregada para o mês de junho	72
6.6	REQM das previsões da abordagem agregada para o mês de dezembro	73
6.7	Comparação entre previsão mensal e abordagem agregada para o mês de junho .	73
6.8	Comparação entre previsão mensal e abordagem agregada para o mês de dezembro	74
6.9	Previsão do modelo PAR	74
6.10	REQM das previsões do modelo PAR para o mês de junho	75
6.11	REQM das previsões do modelo PAR para o mês de dezembro	75
6.12	REQM das previsões do modelo PAR para o mês de junho	75
6.13	REQM das previsões do modelo PAR para o mês de dezembro	76
6.14	Médias mensais das séries sintéticas geradas	77
6.15	Coefficientes de assimetria das séries sintéticas geradas	78
6.16	Medidas de curtose das séries sintéticas geradas	79
6.17	Desvio padrão das séries sintéticas geradas	79
6.18	REQM das previsões do modelo PAR sobre séries sintéticas do mês de junho . .	80
6.19	REQM das previsões do modelo PAR sobre séries sintéticas do mês de dezembro	81
6.20	REQM das previsões do modelo PAR sobre séries sintéticas do mês de junho . .	81
6.21	REQM das previsões do modelo PAR sobre séries sintéticas do mês de dezembro	82

Introdução

1.1 Descrição do Problema

O conhecimento da disponibilidade da água é indispensável, visto que seu uso é importante em diversas aplicações como sistemas de irrigação, navegação, transporte, pesca, geração de energia, dentre outros. No Brasil, a geração de energia elétrica é predominantemente hidráulica, sendo que a maior parte da capacidade instalada no país é composta por usinas hidrelétricas distribuídas em doze bacias hidrográficas localizadas em diferentes regiões. A fim de construir usinas, bem como operá-las é necessário o estudo do comportamento da água. Devido a complexidade dos fenômenos hídricos, existe a necessidade de estabelecer modelos capazes de representar com nível de precisão satisfatório o ciclo hídrico. No caso de usinas hidrelétricas, a água proveniente da chuva direciona-se aos rios que, por sua vez, alimentam os reservatórios dessas usinas que recebem diferentes volumes fluviométricos, o que reflete diretamente na operação do reservatório de uma usina.

A produção energética de um sistema hidroelétrico depende da série de vazões afluentes às diversas usinas. Por exemplo, quando se considera um aproveitamento hídrico, a série de vazões no local é usualmente utilizada como dado de entrada para um modelo de simulação e/ou otimização (Ballini 2000).

A especificação de um modelo de série temporal é um processo que começa na observação e análise dos dados e finaliza com a definição do modelo mais adequado para a série em estudo. Para tal, diversas etapas devem ser consideradas, tais como pré-processamento da série, seleção de entradas, identificação do modelo e, finalmente, o ajuste e otimização dos parâmetros do modelo, também chamada de etapa de estimação, a qual, geralmente, considera alguma métrica de desempenho de interesse na aplicação, para fins de validação.

Entretanto, de um modo geral, análises estatísticas relacionadas às vazões encontram limitações relacionadas ao tamanho das séries históricas disponíveis, que em sua maioria são insuficientes. Uma boa alternativa para superar essas limitações é a geração sintética de série de vazões. Esta solução é utilizada por vários pesquisadores ao longo dos últimos anos, que desenvolveram modelos baseados em propriedades determinísticas e estocásticas, aplicados a varias escalas de tempo como anual, mensal e diária.

Nas últimas décadas, modelos baseados na inteligência computacional têm sido amplamente

sugeridos nas mais diversas aplicações, destacando-se as redes neurais, os sistemas nebulosos e os modelos híbridos neurais nebulosos. Particularmente, a utilização destes modelos para a previsão de séries temporais vem sendo proposta por diversos autores, tendo como exemplos os trabalhos descritos por (Weigend, Ruberman & Rumelhart 1992), (Maier & Dandy 2000), (Crone 2005), (Marinho, Ballini & Soares 2005), (Akbari, van Overloop & Afshar 2010), (Evsukoff, de Lima & Ebecken 2011).

Em geral, os resultados mostraram evidências sobre a capacidade dos modelos mencionados de representar adequadamente as incertezas e não linearidades embutidas nos problemas respectivamente abordados. Por outro lado, modelos para geração de séries sintéticas de vazões são ferramentas necessárias ao planejamento e alocação de recursos hídricos. Atualmente, modelos de geração de séries sintéticas são amplamente utilizados na simulação de cenários visando a otimização do planejamento eletroenergético. Nestas áreas, apesar da disponibilidade de históricos de vazões, um grande número de sequências, que repliquem as características hidrológicas das séries é requisito necessário para a obtenção de melhores resultados.

Tratando-se agora de modelos de geração de séries sintéticas de vazões, esses tornaram-se uma ferramenta muito útil para a otimização dos recursos hídricos. Atualmente, são amplamente utilizados em simulações e técnicas de otimização, tentando-se produzir resultados comparáveis para as várias alternativas de planejamento, projeto ou operação de sistemas de recursos hídricos. Nestas áreas, a disponibilidade de longos históricos de vazões com um grande número de sequências, que se assemelhe aos dados em termos de vários parâmetros especificados, é necessário para que possa conseguir resultados confiáveis.

1.2 Objetivo

O objetivo deste trabalho é desenvolver um sistema de suporte para a previsão e geração de séries sintéticas de vazões mensais, necessárias para o planejamento e a operação das usinas do SIN. Este sistema permite analisar o desempenho de modelos de geração de séries sintéticas e de previsão de vazões, além de permitir comparações entre um conjunto específico de modelos de séries temporais e da inteligência computacional através de estudos de caso considerando postos hidrológicos diversos.

Previsão de Vazões

2.1 Introdução

Os modelos de previsão de vazões podem ser classificados como estocásticos, ao introduzir o conceito de probabilidade de ocorrência das vazões (Chow 1964), ou determinísticos caso o modelo não siga a lei das probabilidades (Tucci 1998).

Modelos estocásticos de previsão de vazões tomam como base o histórico de aflúências para um determinado aproveitamento o qual é dividido em duas partes. A primeira parte do histórico é usada para construir o conjunto de dados dentro da amostra a ser usada no ajuste ou estimação dos parâmetros do modelo previamente especificado e de acordo com a técnica de otimização adequada ao modelo. A segunda parte do histórico constitui o conjunto de dados da amostra, que utiliza-se para fins de validação e teste do modelo ajustado.

Os modelos determinísticos por sua vez caracterizam-se por utilizarem como entrada diversas variáveis que possuem forte relação com a bacia hidrográfica e com o rio cujas vazões serão modeladas e previstas, tentando estabelecer um relacionamento mais direto entre as vazões e a ocorrência das variáveis externas como chuva e neve, os diferentes reservatórios por onde passa a água e suas interdependências, as condições de uso do solo na bacia e a formação das aflúências propriamente ditas.

Nos últimos anos, redes neurais artificiais tornaram-se extremamente conhecidas para a previsão em várias áreas, incluindo finanças, carga elétrica e recursos hídricos. A utilização de redes neurais para previsão de séries temporais vem sendo proposta por diversos autores, entre eles (Lapedes & Farber 1987), (Weigend, Rumelhart & Huberman 1990), (Lachtermacher & Fuller 1995).

As redes neurais e técnicas de séries temporais têm sido comparadas em vários estudos. Comparações de redes neurais com outras abordagens ainda não são conclusivas. Alguns estudos mostram que redes neurais artificiais são melhores que os métodos tradicionais (Weigend, Rumelhart & Huberman 1991a) enquanto outras pesquisas concluem o contrário (Tang, Almeida & Fishwick 1991). As explicações para tais contradições podem ser descritas por diferentes fatores tais como: a estrutura da rede neural, o tipo de série (estacionária ou não-estacionária) usada nos estudos e a relação de tamanho da rede e o número de entradas do modelo das séries temporais.

Nas seções a seguir, apresenta-se os modelos de previsão de vazões implementados no Sistema HydroLab (Cicogna 2003), cuja programação foi iniciada em 2003, utilizados para fins de comparação de desempenho dos mesmos. Neste trabalho, embarcou-se os novos modelos de previsão além do acréscimo de novas ferramentas. O programa HydroLab é um sistema de suporte à decisão para o planejamento e programação da operação de sistemas hidrotérmicos. O termo suporte à decisão é devido à característica que o sistema tem em gerenciar um conjunto de modelos de otimização, previsão e simulação.

2.2 Modelos de previsão de vazões

Os modelos implementados nesse trabalho para previsão de vazões são modelos univariados de séries temporais. Tais modelos consideram como variáveis explicativas apenas os atrasos da própria série, neste caso, de vazões naturais afluentes em reservatórios do sistema hidroelétrico brasileiro.

A primeira abordagem a ser tratada são modelos lineares denominados modelos autorregressivos periódicos $PAR(p_m)$. Como abordagem alternativa, tem-se modelos não lineares baseados em inteligência computacional. Os modelos utilizados para a construção de modelos não lineares de séries temporais são a rede neural multicamadas (MLP), a base de regras nebulosas (FIS) e a rede neurofuzzy (RNF).

A seguir, apresenta-se a descrição detalhada dos modelos mencionados, assim como seus métodos de otimização de parâmetros.

2.2.1 Modelos autorregressivos

Muitas séries temporais revelam um comportamento periódico ou sazonal que pode ser modelado por processos periódicos. Os processos periódicos foram propostos originalmente em (Gladyshev 1961) e várias aplicações desses modelos na área de hidrologia são apresentadas em (Vecchia 1985a), (Vecchia 1985b) e (MecLeod 1994). Séries com autocovariância periódicas não devem ser modeladas como processos sazonais Autoregressivos de Medias Móveis (ARMA) proposto em (Box, Jenkins & G. C. Reinsel 1994)), por serem estes modelos na realidade estacionários.

Uma classe de processos periódicos amplamente utilizada é a dos processos periódicos autorregressivos (PAR), originalmente propostos por Thomas & Fiering (1962a) e amplamente analisados em (Pagano 1978), que apresenta as propriedades assintóticas destes modelos assumindo normalidade do processo, e em (Troutman 1979) que mostra as condições necessárias e suficientes para estacionariedade periódica.

Esta seção apresenta os modelos autorregressivos (AR), os modelos autorregressivos periódicos (PAR) e os métodos de identificação da ordem e de estimação dos parâmetros destes modelos.

Estrutura geral

Modelos autorregressivos Seja $x_t, t = 1, \dots, n$, uma série temporal estacionária, por exemplo a de vazões médias anuais. Vamos assumir que esta série segue um modelo AR(p), dado por:

$$x_t = c + \sum_{j=1}^p \phi_j x_{t-j} + a_t, \quad (2.1)$$

em que a_t é o ruído com $E(a_t) = 0$, $Var(a_t) = \sigma_a^2$ e $E(a_t a_{t+k}) = 0$, para $k \neq 0$. Os parâmetros do modelo são os coeficientes ϕ_1, \dots, ϕ_p ; o termo constante c e a variância σ_a^2 , tais que $c > 0$ e as raízes do polinômio $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ estejam todas fora do círculo unitário. Portanto (2.1) descreve um modelo estacionário no sentido fraco (Box, Jenkins & Reinsel 2008).

Considerando o modelo (2.1), tem-se que:

$$E(x_t) = \frac{c}{1 - \sum_{j=1}^p \phi_j}, \quad (2.2)$$

$$Var(x_t) = \frac{\sigma_a^2}{1 - \sum_{j=1}^p \phi_j \rho_j}, \quad (2.3)$$

em que $\rho_j, j \geq 1$ define a função de autocorrelação (FAC) da série x_t , dado por:

$$\rho_j = \frac{E(x_t - \mu)(x_{t-j} - \mu)}{\sigma^2}, \quad (2.4)$$

em que denota-se por $\mu = E(x_t)$ e $\sigma^2 = Var(x_t)$.

A função de autocorrelação $\rho_j, j \geq 1$ satisfaz a equação de Yule-Walker (Box et al. 2008), dada por:

$$\rho_j = \phi_j \rho_{j-1} + \dots + \phi_p \rho_{p-1}, \quad j \geq 1, \quad (2.5)$$

em que $\rho_0 = 1$.

A função de autocorrelação parcial de k -ésima ordem é dada pela solução do sistema de equações definido pelas equações de Yule-Walker para um processo AR(k), tal que $j = 1, \dots, k$

$$P_k \phi_k = \rho_k$$

com

$$P_k = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{pmatrix}$$

em que $\phi_k = (\phi_{k1}, \dots, \phi_{kk})^T$ e $\rho_k = (\rho_1, \dots, \rho_k)^T$. O último coeficiente ϕ_{kk} é a autocorrelação de ordem k .

Modelos autorregressivos periódicos $PAR(p_m)$ O modelo $PAR(p_m)$ é descrito matematicamente por:

$$\left(\frac{x_{t(r,m)} - \mu_m}{\sigma_m} \right) = \sum_{j=1}^{p_m} \phi_{j,m} \left(\frac{x_{t(r,m)-j} - \mu_{m-j}}{\sigma_{m-j}} \right) + a_{t(r,m)}, \quad (2.6)$$

em que $x_{t(r,m)}$ é a vazão média mensal para o ano r e o mês m ; μ_m é a média do mês m ; σ_m é o desvio padrão do mês m ; $\phi_{1,m}, \dots, \phi_{p_m,m}$ são os coeficientes do modelo autorregressivo para cada mês m ; $a_{t(r,m)}$ é um ruído não correlacionado para o ano r e o mês m ; com $E(a_{t(r,m)}) = 0$, $Var(a_{t(r,m)}) = \sigma_m^{(a)2}$ dada por:

$$\sigma_m^{(a)2} = \sigma_m^2 \left(1 + \sum_{j=1}^{p_m} \rho_j^m \right) \quad (2.7)$$

em que ρ_j^m é a autocorrelação entre $x_{t(r,m)}$ e $x_{t(r,m)-j}$, dada por:

$$\rho_j^m = E \left\{ \left(\frac{x_{t(r,m)} - \mu_m}{\sigma_m} \right) \left(\frac{x_{t(r,m)-j} - \mu_{m-j}}{\sigma_{m-j}} \right) \right\} \quad (2.8)$$

Método de ajuste dos parâmetros

Modelos autorregressivos O método de ajuste dos parâmetros dos modelos AR consiste na identificação da ordem p do modelo, e da estimação do termo constante c e dos coeficientes ϕ_1, \dots, ϕ_p , representados vetorialmente por $(c, \phi_1, \dots, \phi_p)^T$. Para estimar estes parâmetros vamos considerar uma série x_1, x_2, \dots, x_n e as estimativas para a função de autocorrelação dadas por:

$$r_k = \frac{1/(n-k) \sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{s^2}, \quad (2.9)$$

calculadas para $k = 1, 2, \dots, k < n/4$, em que

$$\bar{x} = \frac{\sum_{t=1}^n x_t}{n}, \quad (2.10)$$

$$s^2 = \frac{\sum_{t=1}^n (x_t - \bar{x})^2}{n-1}. \quad (2.11)$$

Podemos estimar os parâmetros $\phi_j, j = 1, \dots, p$ do modelo AR(p) resolvendo-se a equação de Yule-Walker dada por:

$$\begin{bmatrix} 1 & r_1 & \cdots & r_{p-1} \\ r_1 & 1 & \cdots & r_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{bmatrix} \quad (2.12)$$

Os estimadores da FAC, dadas por (2.9), têm variância assintótica dada por $Var(r_k) \approx 1/n$ (Box et al. 1994). Estas variâncias podem ser usadas para testar as hipóteses $H_0 : \rho_k = 0$ contra $H_1 : \rho_k \neq 0$; para $k = 1, \dots, p$. A hipótese H_0 não pode ser rejeitada se o intervalo $r_k \pm 2/\sqrt{n}$ contém o número 0. Neste caso, se diz que a correlação de ordem k não é significativa.

A função de autocorrelação parcial (FACP) também pode ser estimada escrevendo as equações de Yule-Walker para um modelo AR(k), como:

$$\begin{bmatrix} 1 & r_1 & \cdots & r_{k-1} \\ r_1 & 1 & \cdots & r_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k-1} & r_{k-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \hat{\phi}_{1k} \\ \hat{\phi}_{2k} \\ \vdots \\ \hat{\phi}_{kk} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix} \quad (2.13)$$

O conjunto de valores ϕ_{kk} , $k = 1, 2, \dots$ representam a FACP do modelo autorregressivo. Esta função é outra forma de representar a estrutura de correlação do processo autorregressivo. Em um processo AR(p), a função de autocorrelação parcial ϕ_{kk} , $k = 1, 2, \dots$ será diferente de zero para $k \leq p$ e zero para $k > p$. Um teste de hipótese também pode ser formulado para decidir se um determinado valor ϕ_{kk} é nulo ou não. Para testar a hipótese $H_0 : \phi_{kk} = 0$ considere a aproximação feita para a variância do estimador $\hat{\phi}_{kk}$, $Var[\hat{\phi}_{kk}] \approx 1/n$ e a considera rejeição de H_0 com nível de significância $\alpha < 5\%$ se o zero não está contido no intervalo $\hat{\phi}_{kk} \pm 2/\sqrt{n}$. A FACP pode ser usada para determinar a ordem do modelo AR, usando-se a propriedade de que $\phi_{kk} = 0$ para $k > p$. (Stedinger 2001) sugere um critério para escolha da ordem p que consiste em procurar o maior valor de k , tal que, todas as estimativas $\hat{\phi}_{kk}$ são significativas, ou seja, rejeita-se H_0 . Portanto, segundo este critério, tem-se que p é o maior valor de k tal que a hipótese nula $H_0 : \phi_{kk} = 0$ é rejeitada. Por razões práticas escolhe-se $k \leq p^*$ (neste trabalho $p^* = 12$).

O ajuste do modelo AR(p) é feito, após a determinação da ordem p do modelo, estimando-se os coeficientes ϕ_j , $j = 1, \dots, p$ com a equações de Yule-Walker.

Modelos autorregressivos periódicos Segundo (Stedinger & Taylor 1982), a identificação de um modelo $PAR(p_m)$ consiste na determinação da ordem p_m do modelo de cada mês, que é realizado por meio das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) da série. A autocorrelação ρ_k^m entre $x_t(r, m)$ e $x_{t(r, m) - k}$ é dada por $\rho_k^m = E [x_{t(r, m)} x_{t(r, m) - k}]$

Considerando o modelo $PAR(p_m)$, tem-se que:

$$\rho_k^m = \phi_{1,m} E [x_{t(r, m) - 1} x_{t(r, m) - k}] + \dots + \phi_{p_m, m} E [x_{t(r, m) - p_m} x_{t(r, m) - k}], \quad k \geq 1, \quad (2.14)$$

em que $\rho_0^m = 1$.

Fixando-se o mês m e variando k de 1 até p_m em (2.14) encontra-se para cada mês m , um conjunto de equações de Yule-Walker. Assim, para um mês m qualquer, as equações de Yule-Walker são dadas por:

$$\begin{bmatrix} 1 & \rho_1^{m-1} & \cdots & \rho_{p_m-1}^{m-1} \\ \rho_1^{m-1} & 1 & \cdots & \rho_{p_m-2}^{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p_m-1}^{m-1} & \rho_{p_m-2}^{m-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \hat{\phi}_1^m \\ \hat{\phi}_2^m \\ \vdots \\ \hat{\phi}_{p_m}^m \end{bmatrix} = \begin{bmatrix} \rho_1^m \\ \rho_2^m \\ \vdots \\ \rho_{p_m}^m \end{bmatrix} \quad (2.15)$$

Denotando por ϕ_{kj} , o j -ésimo parâmetro autorregressivo de um processo PAR de ordem k em que ϕ_{kk} é o último parâmetro deste processo. As equações de Yule-Walker para cada período m

podem ser reescritas da seguinte forma:

$$\begin{bmatrix} 1 & \rho_1^{m-1} & \cdots & \rho_{p_m-1}^{m-1} \\ \rho_1^{m-1} & 1 & \cdots & \rho_{p_m-2}^{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p_m-1}^{m-1} & \rho_{p_m-2}^{m-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \widehat{\phi}_{k1} \\ \widehat{\phi}_{k2} \\ \vdots \\ \widehat{\phi}_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1^m \\ \rho_2^m \\ \vdots \\ \rho_{p_m}^m \end{bmatrix} \quad (2.16)$$

Os valores dos parâmetros $\{\phi_{kk}\}$, $k = 1, 2, \dots$ correspondem a função de autocorrelação parcial do processo PAR, no mês m . Esta função representa também a estrutura de dependência do processo estocástico ao longo do tempo. Em um processo periódico autorregressivo de ordem p_m , a função de autocorrelação parcial é valor do último parâmetro ϕ_{kk} dado em 2.16 e será diferente de zero para k menor ou igual a p_m e zero para k maior que p_m .

Portanto, a identificação de um modelo $PAR(p_m)$ consiste em determinar as ordens, p_m , mais apropriadas dos operadores autorregressivos de cada mês, $m = 1, 2, \dots, 12$. Isto pode ser feito obtendo-se estimativas de $\{\phi_{kk}\}$, $k = 1, 2, \dots$ que são calculadas, substituindo em (2.16) as autocorrelações ρ_i^{m-j} , para $i = 1, \dots, p_m$ e $j = 0, \dots, p_m - 1$ pelos respectivos valores amostrais. Para testar a hipótese nula $H_0 : \phi_{kk} = 0$ para todos os valores de $k = 1, 2, \dots$, utiliza-se o resultado assintótico o qual afirma que, se a ordem do operador autorregressivo de um mês m qualquer, é p_m , então o estimador $\widehat{\phi}_{kk}^m$ de ϕ_{kk}^m , para $k > p_m$ tem distribuição aproximadamente Normal de média zero e variância $1/n$. Portanto não se rejeita H_0 com nível de significância $\alpha < 5\%$ se o zero está contido no intervalo $\widehat{\phi}_{kk}^m \pm 2/\sqrt{n}$. Neste caso se diz que a estimativa $\widehat{\phi}_{kk}^m$ é não significativa. Um procedimento de identificação da ordem p_m , consiste em procurar, para cada mês m , a maior ordem p_m tal que todas as estimativas $\widehat{\phi}_{kk}^m$ para $k > p_m$ não sejam mais significativas.

Este procedimento clássico de identificação da ordem dos modelos p_m pode levar a aceitação de processos como sendo ruídos brancos, quando, na verdade, trata-se de modelos que possuem baixos valores de ρ_k^m e ϕ_{kk}^m . Valores intermediários baixos de ϕ_{jj}^m para $j < p_m$ também pode ser considerados não significativos e isso leva a identificação de modelos com parâmetros intermediários nulos $\phi_{kj} = 0$, para alguns $j < p_m$.

Stedinger (2001) afirma que uma modelagem que considera individualmente cada mês não pode produzir o conjunto de modelos mais adequado. Sem maiores detalhes matemáticos e somente considerando a natureza dos fenômenos naturais envolvidos, o autor argumenta que não faria sentido considerar que a vazão de determinado mês dependa de vários meses anteriores, ou seja, o modelo $PAR(p_m)$ tem ordem p_m elevada $p_m = 6$ e para o mês seguinte, $m + 1$ a ordem seja pequena $p_{m+1} = 1$. Assim como também não é razoável supor que, devido a parâmetros intermediários nulos, a vazão de um mês m dependa da vazão do mês $m - 2$ e não do mês $m - 1$. Estes resultados corroboram a idéia de que a forma clássica de identificação das ordens p_m pode não ser a mais adequada. (Stedinger 2001) propõe como critério para a seleção das defasagens j significantes, que para cada período m , encontre-se a maior ordem j tal que todas as estimativas $\widehat{\phi}_{kk}^m$ para $k < j$ sejam significativas. Este critério não admite valores intermediários de $\widehat{\phi}_{kk}^m$ não significativos, o que ocorre na forma clássica.

2.2.2 Redes neurais multicamadas

Redes neurais artificiais tornaram-se extremamente conhecidas para previsão em várias áreas, incluindo finanças, carga elétrica e recursos hídricos. A utilização de redes neurais para previsão de séries temporais vem sendo proposta por diversos autores, entre eles (Lapedes & Farber 1987), (Weigend et al. 1990) e (Lachtermacher & Fuller 1995), mostrando a viabilidade de utilizar estes modelos.

As comparações de redes neurais com outras abordagens ainda são contraditórias. Alguns estudos concluem que redes neurais artificiais são melhores que os métodos tradicionais (Weigend et al. 1991a), enquanto outros concluem o contrário (Tang et al. 1991). As explicações para tais contradições podem ser descritas por diferentes fatores, tais como: a estrutura da rede neural, o tipo de série (estacionária, não-estacionária) usada nos estudos e a relação do tamanho da rede e o número de entradas das séries temporais.

Lapedes & Farber (1987) aplicaram redes neurais multi-camadas (*Multi-Layer Perceptron - MLP*) para previsão de duas séries temporais caóticas geradas por um processo determinístico não-linear. Os resultados indicaram que os modelos MLP podem “aprender”, simulando sistemas não-lineares, como excelentes propriedades de previsão, comparadas a métodos tradicionais.

Tang et al. (1991) compararam modelos MLP e Box & Jenkins, usando séries de passageiros de vôo internacional, vendas de carros domésticos e vendas de carros importados nos EUA. Os resultados mostraram que o desempenho do modelo Box & Jenkins foi superior ao modelo de redes neurais para previsão a curto prazo. Já para previsão de longo prazo, os modelos de redes neurais apresentaram um melhor desempenho.

Lachtermacher & Fuller (1995) descreveram uma metodologia híbrida para a aplicação de redes neurais em análise de séries temporais. A metodologia de Box & Jenkins foi utilizada como um procedimento para explorar importantes relações nos dados da série e estas informações foram usadas para definir uma estrutura inicial para a rede neural reduzindo, assim, os parâmetros a serem estimados e os dados exigidos. Dois tipos de séries foram estudadas: séries estacionárias e não estacionárias. Observou-se que o modelo de redes neurais, para a maioria das séries testadas, obteve um desempenho igual ou superior quando comparados aos modelos de Box & Jenkins.

Maier & Dandy (2000) fazem uma revisão dos trabalhos propostos para aplicação de previsão de variáveis de recursos hídricos, mostrando as várias técnicas de escolha do critério de desempenho do modelo, o pré-processamento dos dados, a determinação das entradas e a arquitetura da rede, técnicas de treinamento e validação do modelo.

Ballini (1996) propôs o uso de redes neurais multi-camadas com algoritmo *back-propagation* para previsão de chuva-vazão diária. Francelin, Ballini & Andrade (1996) aplicaram redes neurais MLP para previsão de vazões a longo prazo e compararam com os modelos de séries temporais Box & Jenkins. Os resultados mostraram desempenho promissor para a rede MLP, apresentando erros percentuais máximos menores que os métodos de Box & Jenkins.

O desenvolvimento, implementação, comparação e validação de algoritmos de previsão de séries temporais de vazões com o uso de técnicas de inteligência computacional, com destaque para redes neurais em simbiose com algoritmos de agrupamento e reconhecimento de padrões foi analisado e os resultados mostraram a viabilidade da utilização dessas técnicas para o problema

de previsão (Magalhães, Ballini, Molck & Gomide 2004), (Magalhães, Ballini, Gonçalves & Gomide 2004).

A arquitetura (ou topologia) da rede MLP consiste de uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída. A camada de entrada, cujos neurônios pertencentes a esta camada são denominados unidades de entrada, difundem os sinais de entrada para as camadas seguintes sem nenhuma modificação. As camadas intermediárias transmitem informações através das conexões entre as camadas de entrada e saída. Os neurônios pertencentes a esta camada são denominados unidades intermediárias. Finalmente, a camada de saída, cujos neurônios são chamados unidades de saída, transmite a resposta da rede neural à entrada aplicada na camada de entrada.

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. Entretanto, não há comunicação entre as unidades de uma mesma camada.

Cada neurônio da rede tem uma função não-decrescente e diferenciável, conhecida como função de ativação, a qual pode ser uma função linear, logística, ou tangente hiperbólica. A estrutura geral de uma rede neural MLP com uma camada oculta é descrita a seguir.

Estrutura geral

A Figura 2.1 mostra a arquitetura de uma rede MLP com uma camada intermediária e entradas de polarização ou bias, onde $\mathbf{x} = [x_0, x_1, x_2, \dots, x_n]$; $\mathbf{y} = [y_1, y_2, \dots, y_p]$ e w_{ij} e v_{jk} são os pesos da camada intermediária e de saída respectivamente para $i = 0, \dots, n$; $j = 0, \dots, M$; $k = 1, \dots, p$; sendo M o número de neurônios da camada intermediária e x_0 a entrada de polarização.

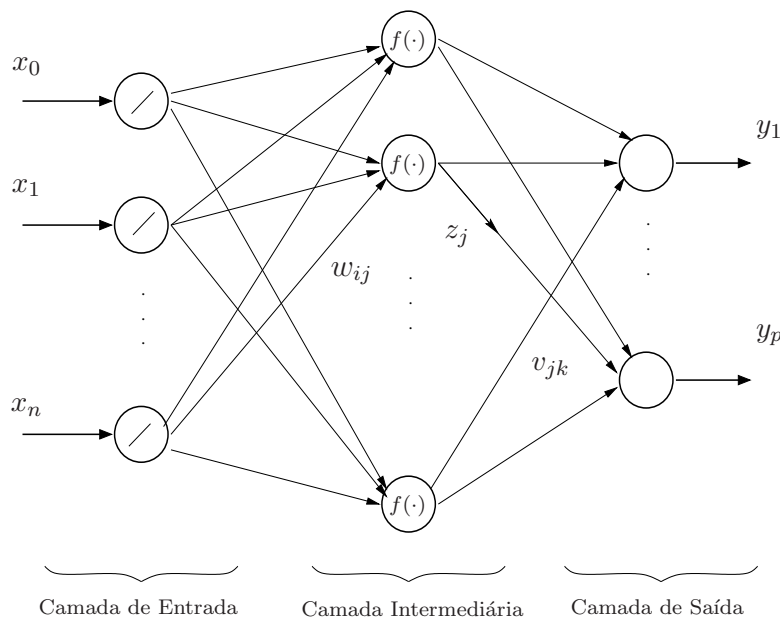


Figura 2.1: Rede MLP de uma camada intermediária com bias.

As conexões entre os neurônios das distintas camadas da rede, do mesmo jeito que entre os neurônios biológicos, são denominadas sinapses, cada sinapse está ponderada por um peso que

representa a intensidade da conexão, e a fluência da informação está sempre direcionada em um único sentido entrada-saída, ou seja, não existe realimentação (*feedforward*).

Considerando bias na rede MLP e funções de ativação lineares na camada de entrada, define-se a saída z_j do j -ésimo neurônio da camada intermediária como:

$$z_j = f \left(\sum_{i=0}^{i=n} w_{ij} \cdot x_i \right) \quad (2.17)$$

onde $x_0 = +1$ e $f(\cdot)$ é a função de ativação na camada intermediária. A k -ésima saída da rede y_k é calculada como:

$$y_k = \sum_{j=0}^M v_{jk} \cdot z_j \quad (2.18)$$

sendo M o número de neurônios da camada intermediária, $z_0 = +1$ e considera-se que os neurônios da camada de saída possuem a função identidade como função de ativação.

Método de ajuste dos parâmetros

Um método de treinamento supervisionado altamente popular é o método do gradiente do erro, comumente denominado algoritmo de retropropagação, proposto pela primeira vez por Werbos (1974) e redescoberta e popularizada por Rumelhart & McClelland (1986). O nome de retropropagação deve-se o fato do sinal de erro ser *retropropagado* através da rede, camada por camada.

O algoritmo de retropropagação consiste basicamente de duas etapas (Ballini 2000): durante a primeira etapa (*forward*), as entradas são apresentadas e propagadas através da rede, camada por camada e mantendo os pesos fixos; logo, a saída é comparada com a saída desejada, calculando assim o sinal de erro. Na segunda etapa, o erro é propagado da frente pra traz (*backward*), sendo os pesos atualizados via o Método do Gradiente Descendente (Haykin 1994). Seja:

- z_j^l a saída do j -ésimo neurônio na camada l ;
- w_{ij}^l o peso entre o neurônio i da camada $l - 1$ e o neurônio j da camada l ;
- \mathbf{x}_p o p -ésimo padrão de entrada;
- z_i^0 o i -ésimo componente do vetor de entrada;
- y_j a j -ésima saída desejada;
- M_l o número de neurônios da l -ésima camada;
- N o número de padrões de treinamento;
- L o número de camadas da rede neural;
- $f_l(\cdot)$ a função de ativação dos neurônios na camada l .

sendo $z_j^0 = x_j$, $z_0^l = 1$ (bias) e w_{j0}^l os pesos das bias. A saída do j -ésimo neurônio da camada l é dada por:

$$z_j^l = f_l \left[\sum_{i=0}^{M_{l-1}} w_{ij}^l \cdot z_i^{l-1} \right] \quad (2.19)$$

O objetivo é minimizar o Erro Quadrático Médio (EQM), tal que:

$$J(\mathbf{w}) = \frac{1}{N} \sum_{p=1}^N J_p(\mathbf{w}) \quad (2.20)$$

onde $J_p(\mathbf{w})$ é o erro quadrático total para o p -ésimo padrão de entrada-saída, definido da maneira seguinte:

$$J_p(\mathbf{w}) = \frac{1}{2} \sum_{q=1}^{M_L} \epsilon_q^2 \quad (2.21)$$

com:

$$\epsilon_q = (z_q^L - y_q) \quad (2.22)$$

sendo y_q a q -ésima saída desejada, com $q = 1, \dots, M_L$

Para minimizar a função da equação (2.20) via o método do gradiente descendente, é necessário calcular a derivada parcial de J_p em relação a cada peso da rede. Assim, as regras para efetuar o ajuste destes pode ser escrito da maneira seguinte:

$$w_{ij}^l(k+1) = w_{ij}^l(k) - \alpha \left. \frac{\partial J(\mathbf{w})}{\partial w_{ij}^l} \right|_{\mathbf{w}(k)} \quad (2.23)$$

ou seja:

$$w_{ij}^l(k+1) = w_{ij}^l(k) - \alpha \sum_{p=1}^{M_L} \left. \frac{\partial J_p(\mathbf{w})}{\partial w_{ij}^l} \right|_{\mathbf{w}(k)} \quad (2.24)$$

sendo α a taxa de aprendizado. Utilizando a regra da cadeia têm-se:

$$\frac{\partial J_p(\mathbf{w})}{\partial w_{ij}^l} = \frac{\partial J_p(\mathbf{w})}{\partial z_j^l} \cdot \frac{\partial z_j^l}{\partial w_{ij}^l} \quad (2.25)$$

onde:

$$\frac{\partial z_j^l}{\partial w_{ij}^l} = \frac{\partial}{\partial w_{ij}^l} \left[f_l \left(\sum_{m=0}^{M_{l-1}} w_{mj}^l \cdot z_m^{l-1} \right) \right] \quad (2.26)$$

seguindo a regra da cadeia:

$$\frac{\partial z_j^l}{\partial w_{ij}^l} = f_l' \left(\sum_{m=0}^{M_{l-1}} w_{ij}^l \cdot z_m^{l-1} \right) \cdot \frac{\partial}{\partial w_{ij}^l} \left[\sum_{m=0}^{M_{l-1}} w_{mj}^l \cdot z_m^{l-1} \right] \quad (2.27)$$

para logo obter:

$$\frac{\partial z_j^l}{\partial w_{ij}^l} = f_l' \left(\sum_{m=0}^{M_{l-1}} w_{mj}^l \cdot z_m^{l-1} \right) \cdot z_i^{l-1} \quad (2.28)$$

assim:

$$\frac{\partial J_p(\mathbf{w})}{\partial w_{ij}^l} = \frac{\partial J_p(\mathbf{w})}{\partial z_j^l} \cdot f_l' \left(\sum_{m=0}^{M_{l-1}} w_{mj}^l \cdot z_m^{l-1} \right) \cdot z_i^{l-1} \quad (2.29)$$

onde o termo $\partial J_p(\mathbf{w})/\partial z_j^l$ representa a sensibilidade de $J_p(\mathbf{w})$ com respeito da saída z_j^l (Hush & Horne 1993). A saída z_j^l manifesta a sua influência através das seguintes camadas. Seja:

$$\delta_j^l = \frac{\partial J_p(\mathbf{w})}{\partial z_j^l} \cdot f_l' \left(\sum_{m=0}^{M_{l-1}} w_{mj}^l \cdot z_m^{l-1} \right) \quad (2.30)$$

a equação (2.29) é reescrita como:

$$\frac{\partial J_p(\mathbf{w})}{\partial w_{ij}^l} = \delta_j^l \cdot z_i^{l-1} \quad (2.31)$$

Assumindo uma rede MLP com uma camada intermediária ($L = 3$) a modo de exemplo, o erro ϵ_q é calculado como:

$$\epsilon_q(k) = z_q^3(k) - y_q(k) \quad (2.32)$$

para $q = 1, \dots, M_3$, considerando primeiro, o caso em que o neurônio pertence a camada de saída (ver figura 2.2.2), utilizando a equação (2.21), tem-se:

$$\frac{\partial J_p(\mathbf{w})}{\partial z_q^3} = \epsilon_q(k) \quad (2.33)$$

deste modo:

$$\delta_q^3(k) = \epsilon_q(k) \cdot f_3' \left(\sum_{m=0}^{M_2} w_{mq}^3 \cdot z_m^2 \right) \quad (2.34)$$

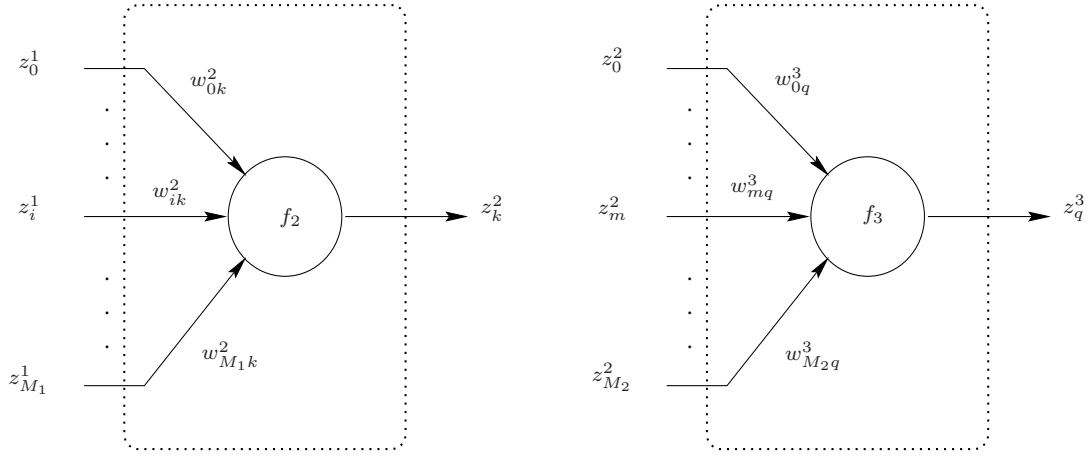


Figura 2.2: Generalização dos neurônios de uma rede MLP para $L=3$

para $q = 1, \dots, M_3$, sendo M_3 o número de saídas da rede. Deste modo, o gradiente para os pesos da camada de saída pode ser escrito como:

$$\Delta w_{iq}^3(k) = \left. \frac{\partial J_p(\mathbf{w})}{\partial w_{iq}^3(k)} \right|_{\mathbf{w}(k)} = \epsilon_q(k) \cdot f_3' \left(\sum_{m=0}^{M_2} w_{mq}^3 \cdot z_m^2 \right) \cdot z_i^2 \quad (2.35)$$

para $i = 0, \dots, M_2$ e $q = 1, \dots, M_3$. Considerando agora os neurônios da camada intermediária, δ_j^2 é calculado como:

$$\delta_j^2(k) = \frac{\partial J_p(\mathbf{w})}{\partial z_q^3} \cdot \frac{\partial z_q^3}{\partial z_j^2} \cdot f_2' \left(\sum_{m=0}^{M_1} w_{mj}^2 \cdot z_m^1 \right) \quad (2.36)$$

calculando as derivadas:

$$\frac{\partial J_p(\mathbf{w})}{\partial z_q^3} = \frac{1}{2} \sum_{q=1}^{M_3} \frac{\partial \epsilon_q^2(k)}{\partial z_q^3} = \sum_{q=1}^{M_3} \epsilon_q(k) \quad (2.37)$$

$$\frac{\partial z_q^3}{\partial z_j^2} = \frac{\partial}{\partial z_j^2} \left[f_3 \left(\sum_{j=0}^{M_2} w_{jq}^3 \cdot z_j^2 \right) \right] = f_3' \left(\sum_{j=0}^{M_2} w_{jq}^3 \cdot z_j^2 \right) \cdot w_{jq}^3 \quad (2.38)$$

substituindo as equações (2.37) e (2.38) na equação (2.36), tem-se:

$$\delta_j^2(k) = \sum_{q=1}^{M_3} \epsilon_q(k) \cdot f_3' \left(\sum_{j=0}^{M_2} w_{jq}^3 \cdot z_j^2 \right) \cdot w_{jq}^3 \cdot f_2' \left(\sum_{m=0}^{M_1} w_{mj}^2 \cdot z_m^1 \right) \quad (2.39)$$

finalmente, substituindo a equação (2.34) na equação (2.39), obtém-se:

$$\delta_j^2(k) = \left[\sum_{q=1}^{M_3} \delta_q^3(k) \cdot w_{jq}^3 \right] \cdot f_2' \left(\sum_{m=0}^{M_1} w_{mj}^2 \cdot z_m^1 \right) \quad (2.40)$$

Desta forma, o gradiente para atualizar os pesos da camada intermediária utilizando as equações (2.40) e (2.29) é calculada como:

$$\Delta w_{ij}^2 = \frac{\partial J_p(\mathbf{w})}{\partial w_{ij}^2} \Big|_{\mathbf{w}(k)} = \left[\sum_{q=1}^{M_3} \delta_q^3(k) \cdot w_{jq}^3 \right] \cdot f_2' \left(\sum_{m=0}^{M_1} w_{mj}^2 \cdot z_m^1 \right) \cdot z_i^1 \quad (2.41)$$

para $i = 0, \dots, M_1$ e $j = 0, \dots, M_2$, sendo neste caso $z_i^1 = z_i^0 = x_i$.

As equações (2.41) e (2.35) definem os gradientes dos pesos da camada de entrada e saída respectivamente a utilizar durante a atualização dos pesos utilizando o método de retropropagação, isto, para a rede MLP de uma camada intermediária. A equação (2.19) é uma generalização da saída para qualquer neurônio de uma rede multicamada, enquanto a equação (2.31) resume o método do gradiente para qualquer neurônio de uma rede MLP.

Função de ativação Para o cálculo dos gradientes locais de cada neurônio da rede MLP é preciso conhecer a derivada da função de ativação. A diferenciabilidade da função de ativação é, portanto, uma condição necessária. Uma das funções de ativação mais utilizadas em aplicações é a função logística, definida como segue:

$$y_j(n) = f_j(\nu_j(n)) = \frac{1}{1 + \exp^{-\nu_j(n)}} \quad (2.42)$$

onde $\nu_j(n)$ é o nível de ativação do neurônio j no instante n . Esta função é limitada, assumindo valores no intervalo $(0, 1)$. Assim, diferenciando a equação (2.42) em relação a $\nu_j(n)$, tem-se:

$$f_j'(\nu_j(n)) = \frac{\exp(-\nu_j(n))}{[1 + \exp(-\nu_j(n))]^2} = y_j(n)[1 - y_j(n)] \quad (2.43)$$

permitindo concluir que o cálculo da derivada requer poucos recursos computacionais.

Generalização Generalização é uma das mais importantes características de redes neurais, podendo ser influenciada por três fatores:

1. O tamanho e a eficiência do conjunto de treinamento, que deve ser muito bem escolhido, pois deve expressar todas as características relevantes acerca do problema no universo considerado;
2. A arquitetura da rede neural: números de camadas e neurônios intermediários;
3. A complexidade do problema a ser tratado.

A rede neural é dita ter uma boa generalização se, fornecidos os conjuntos de dados de treinamento e teste, a relação entre os dados de entrada/saída calculada pela rede já treinada está correta ou tão próxima quanto desejável para os padrões pertencentes ao conjunto de teste, podendo ser obtida através da validação cruzada.

Taxa de aprendizado O algoritmo de retropropagação fornece uma aproximação da trajetória no espaço dos pesos em direção ao mínimo da função custo (equação (2.20)), e é computada via o método do gradiente descendente.

A taxa de aprendizado α determina o passo na variação dos pesos sinápticos, assim, quanto menor seja α , menor será a variação dos pesos de uma iteração para a próxima e no caso do algoritmo de retropropagação, a trajetória no espaço dos pesos será mais suave. Por outro lado, se α toma valores muito grandes, embora a velocidade de convergência aumente, a variação dos pesos entre iterações será também maior podendo virar a rede instável (Haykin 1994).

Utilizando o treinamento por reforço, como se verá nos próximos capítulos, dependendo das regras para atualizar os pesos, é necessário utilizar taxas de aprendizado que regulem a mudança dos pesos por iteração, e as considerações mencionadas acima para o caso do algoritmo de retropropagação valem também para este caso.

Crítérios de Parada Não existe critério de parada para o algoritmo de retropropagação que garanta a convergência da rede. Contudo, existem alguns critérios que podem ser utilizados para dar término ao processo de treinamento. Segundo (Haykin 1994), dois critérios de parada são comumente utilizados:

1. Considerando as propriedades dos mínimos locais ou globais da superfície de erro. Sendo $\mathbf{w} = \mathbf{w}^*$ um mínimo local ou global, o vetor gradiente calculado para este mínimo será nulo. Assim, considera-se que a convergência têm sido alcançada quando a norma euclidiana do vetor gradiente alcança um limiar determinado.
2. Outro critério de parada é quando o erro quadrático médio ou função objetivo é estacionária e suficientemente pequena;

Um outro critério de parada a considerar é a capacidade de generalização da rede. Uma rede neural com capacidade de generalização deve ser capaz de uma vez fornecidos padrões de treinamento e teste diferentes, o mapeamento entre os dados entrada-saída estimados pela rede já treinada é tão eficiente quanto desejável, isto para os padrões de teste. Uma forma de garantir esta propriedade, é aplicando uma ferramenta estatística denominada validação cruzada.

O mecanismo de validação cruzada é aplicado da maneira seguinte: gerar um conjunto de padrões de treinamento e um conjunto de padrões de teste, este ultimo para avaliar a eficiência da rede neural. Durante o processo de treinamento, após cada ajuste de pesos, é necessário avaliar a rede tanto com o conjunto de treinamento como com o conjunto de teste. Quando o erro de teste comece a ter uma tendência crescente, o processo de treinamento se da por concluído.

A validação cruzada é também utilizada para avaliar topologias, como por exemplo para determinar o número de neurônios na camada intermediária, a topologia ótima será aquela que apresente um melhor desempenho.

Inicialização dos Pesos A qualidade e eficiência de aprendizado supervisionado em redes multicamadas, depende não só da estrutura da rede, funções de ativação e regras de aprendizagem, também depende dos valores iniciais dos parâmetros ajustáveis. Em geral, não é possível

ter um conhecimento prévio de valores ótimos para estes parâmetros, pois isto depende tanto do conjunto de treinamento como da natureza da solução. O conjunto inicial de pesos a ser utilizado influi diretamente na velocidade de aprendizado e na qualidade da solução final. Uma escolha inicial pouco adequada pode gerar problemas de mínimos locais pobres ou de saturação prematura.

2.2.3 Bases de regras nebulosas

Um sistema nebuloso é um modelo baseado em regras linguísticas do tipo se-então. Estas regras são as responsáveis por representar o conhecimento adquirido pelo especialista, ou extraído de dados, os quais fornecem informações através de variáveis de entrada e saída do sistema. Dado que não existe uma maneira formal e efetiva de extração do conhecimento, a análise dos dados disponíveis para a extração de informação pelo especialista, com o objetivo de determinar o número de regras adequadas e assim, definir a estrutura do sistema nebuloso, torna-se uma tarefa difícil de ser resolvida (Er & Wu 2002).

Geralmente, sistemas baseados em regras nebulosas do tipo se-então são compostos por um mecanismo de inferência nebulosa, uma base de regras e um *defuzzificador*, o qual é o responsável pelo processamento das saídas locais de cada uma das regras nebulosas (Luna 2003). Já o mecanismo de inferência é o encarregado de mapear a transformação nebulosa do espaço de entrada para o espaço de saída, utilizando por sua vez, a base de regras que compõe o sistema. Os sistemas nebulosos ou modelos baseados em regras nebulosas do tipo Takagi-Sugeno (Takagi & Sugeno 1985) têm se convertido em uma poderosa ferramenta na área de engenharia para modelagem e controle de sistemas complexos (Angelov & Filev 2004), pois estes utilizam uma combinação não linear de modelos simples, geralmente lineares, localizados em sub-regiões do espaço de entrada, de tal forma que o problema global é dividido em problemas “locais”, sendo a saída final do sistema uma combinação não-linear das soluções locais estimadas. Estes sistemas são capazes de processar o conhecimento de forma compreensível e de manipular incertezas e imprecisões visando solucionar problemas reais complexos (Zimmermann 1987).

Existem duas abordagens gerais para a otimização de modelos baseados em regras nebulosas. A primeira abordagem engloba os algoritmos de aprendizado realizados de forma *offline* ou em batelada. Esta abordagem consiste em definir uma estrutura do modelo que será mantida fixa durante o processo de aprendizado, no qual os parâmetros do modelo definido previamente serão ajustados. Para tal, o conjunto total de dados entrada-saída disponíveis para o ajuste será utilizado a cada época do processo.

Uma abordagem diferente é proposta em (Luna, Soares & Ballini 2007a), onde, embora o algoritmo de aprendizado seja baseado em uma abordagem em batelada, este utiliza em paralelo ao ajuste do modelo, critérios de adição de regras nebulosas, assim como critérios de poda, modificando a estrutura inicial durante o aprendizado. Após a etapa de ajuste construtivo, a estrutura é mantida fixa e a previsão múltiplos passos à frente é realizada para um conjunto de séries financeiras e econômicas.

A vantagem do aprendizado em batelada é que, devido à utilização de todos os dados disponíveis para o ajuste do modelo a cada época do processo de aprendizado, em geral, obtém-se ótimos locais adequados, porém, com um maior custo computacional, já que a cada época do

aprendizado, processa-se o total de dados disponíveis para o treinamento. Além disso, a estrutura e parâmetros do modelo se mantêm constante durante e após o aprendizado, independente de existir ou não variações consideráveis na dinâmica do problema em estudo.

No entanto, devido à própria complexidade dos problemas reais, os quais são, em geral, de natureza não linear, torna-se impossível saber com antecedência o número adequado de regras que irão compor a estrutura do modelo. Este problema é normalmente tratado via tentativa e erro. Estruturas de diferentes tamanhos são construídas, ajustadas e testadas. O modelo com melhor desempenho é escolhido como o mais adequado.

Como alternativa a esta abordagem, surge um segundo grupo de algoritmos de aprendizado, o qual engloba os algoritmos de aprendizado de natureza dinâmica ou incremental, também conhecidos na literatura como modelos com aprendizado *online*. Nesta abordagem, uma estrutura inicial é modificada a cada iteração, a medida que novos dados entrada-saída são apresentados ao modelo. A estrutura é modificada acrescentando ou eliminando regras nebulosas no decorrer do aprendizado, de acordo com critérios de adição de regras e de punição ou poda previamente estabelecidos. Ou seja, não existe uma estrutura fixa, já que tanto a estrutura do modelo como os parâmetros são variáveis no tempo.

O trabalho proposto em (Hong & White 2009) sugere a utilização de um modelo dinâmico baseado em regras nebulosas, denominado de DNFLMS, para a constituição de um modelo hidrológico para a previsão um passo à frente da vazão diária do rio Pupu Springs, localizado no vale Takaka na Nova Zelândia.

Por outro lado, o trabalho apresentado em Luna, Soares, Lopes & Ballini (2009), mostra a capacidade de predição dos modelos dinâmicos baseados em regras nebulosas, quando utilizados na previsão vários passos à frente das vazões incrementais médias diárias para a UHE de Porto Primavera, Brasil. Além disso, o modelo FIS é proposto em (Zambelli, Luna & Soares 2009), mostrando a utilização de uma base de regras com aprendizado *offline*, para a previsão um passo à frente de séries de vazões anuais.

Este modelo surgiu da simplificação do modelo construtivo proposto em (Luna, Soares & Ballini 2007b). O sistema de inferência proposto (FIS, do termo em inglês *Fuzzy Inference System*), é baseado em regras nebulosas e em modelos de primeira ordem, compondo um sistema nebuloso do tipo Takagi-Sugeno (TS) (Takagi & Sugeno 1985). Assim, o FIS codifica na sua estrutura uma base de regras do tipo **Se-Então**, adotando a seguinte representação:

$$\mathbf{Se} < \text{antecedente} > \mathbf{Ent\~{a}o} < \text{consequente} >$$

A parte do antecedente representa uma sub-região do espaço de entrada. Esta sub-região define a região ativa associada à regra nebulosa. A parte do consequente é definida por um modelo local que contribuirá com uma parcela na estimativa da saída final. A estrutura geral do FIS é descrita a seguir.

Estrutura geral

A estrutura do FIS é composta por um conjunto de M regras nebulosas de primeira ordem, do tipo Takagi-Sugeno (TS) (Takagi & Sugeno 1985), e é ilustrada na Figura 2.3.

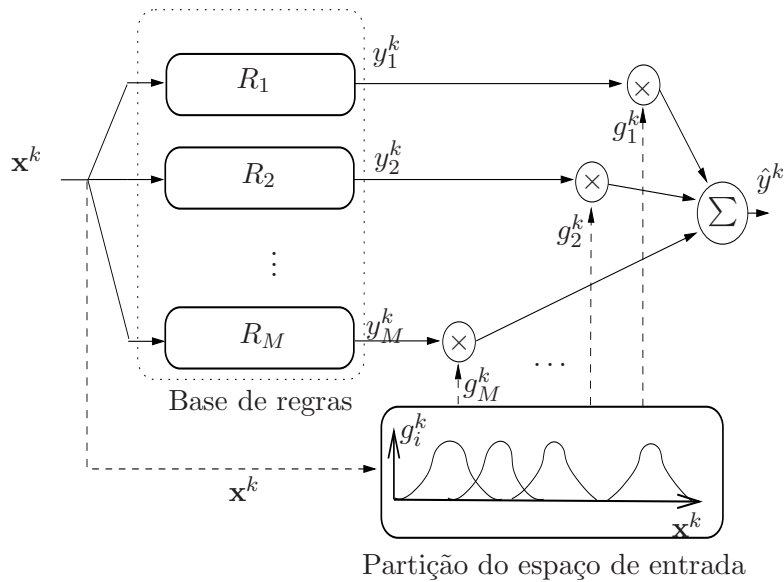


Figura 2.3: Estrutura geral do FIS.

Seja $\mathbf{x}^k = [x_1^k, x_2^k, \dots, x_p^k] \in \mathbb{R}^p$ o vetor de entrada no instante k , $k \in \mathbb{Z}_0^+$; $\hat{y}^k \in \mathbb{R}$ é a saída do modelo que representa em termos de previsão, a estimativa um passo à frente da série em estudo, dada a entrada \mathbf{x}^k .

O espaço de entrada, representado por $\mathbf{x}^k \in \mathbb{R}^p$ é particionado em M sub-regiões, sendo cada sub-região representada por uma regra nebulosa. Os antecedentes de cada regra nebulosa R_i (do tipo **Se-Então**), são definidos pelos respectivos centros $\mathbf{c}_i \in \mathbb{R}^p$ e matrizes de covariância $\mathbf{V}_i |_{p \times p}$. Os consequentes são representados por modelos locais lineares com saídas denotadas por y_i , $i = 1, \dots, M$ e definidas de acordo com:

$$y_i^k = \phi^k \times \theta_i^T, \quad (2.44)$$

onde $\phi^k = [1 \ x_1^k \ x_2^k \ \dots \ x_p^k]$; $\theta_i = [\theta_{i0} \ \theta_{i1} \ \dots \ \theta_{ip}]$ é o vetor de coeficientes do modelo local para a i -th regra nebulosa.

A cada padrão de entrada associa-se um grau de pertinência referente às sub-regiões do espaço de entrada. Os graus de pertinência são calculados utilizando funções de pertinência gaussianas $g_i(\mathbf{x}^k)$, dadas pela seguinte equação:

$$g_i(\mathbf{x}^k) = g_i^k = \frac{\alpha_i \cdot P[i | \mathbf{x}^k]}{\sum_{q=1}^M \alpha_q \cdot P[q | \mathbf{x}^k]}, \quad (2.45)$$

onde α_i são coeficientes positivos que satisfazem a condição de unicidade

$$\sum_{i=1}^M \alpha_i = 1$$

e $P[i | \mathbf{x}^k]$ é definida de acordo com a Equação (2.46) a seguir

$$P[i | \mathbf{x}^k] = \frac{1}{(2\pi)^{p/2} \det(\mathbf{V}_i)^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x}^k - \mathbf{c}_i) \mathbf{V}_i^{-1} (\mathbf{x}^k - \mathbf{c}_i)^T \right\}, \quad (2.46)$$

onde $\det(\cdot)$ representa o operador determinante. Assim, a saída do FIS $y(k) = \hat{y}^k$, é calculada como uma combinação não linear dos consequentes y_i^k ponderados pelos seus respectivos graus de pertinência g_i^k , ou seja:

$$\hat{y}(\mathbf{x}^k) = \hat{y}^k = \sum_{i=1}^M g_i^k y_i^k. \quad (2.47)$$

A Equação (2.46) também pode ser interpretada como a probabilidade do vetor de entrada \mathbf{x}^k pertencer à região ativa da i -ésima regra, já que, a medida que o dado de entrada encontra-se mais próximo do i -ésimo centro, maior será o valor do i -ésimo grau de pertinência. Em outras palavras, cada sub-região nebulosa é representada pelo respectivo centro \mathbf{c}_i , assim como os parâmetros de dispersão codificados na diagonal principal da matriz de covariância associada \mathbf{V}_i .

Assim, 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

$$R_i^k : \quad \text{Se } \mathbf{x}^k \text{ pertence a } i\text{-ésima região com grau} \\ \text{de pertinência } g_i^k \quad \text{Então } y_i^k = \phi^k \times \theta_i^T.$$

Método de ajuste dos parâmetros

A obtenção do modelo final é realizada em duas etapas: a inicialização do modelo e o ajuste dos seus parâmetros.

Inicialização Uma das formas comumente utilizadas para a inicialização de modelos baseados em regras nebulosas consiste na utilização de algoritmos de agrupamento, os quais como será visto, influenciam diretamente na estrutura e otimização dos modelos.

Neste trabalho, a inicialização do modelo é realizada utilizando o algoritmo de agrupamento conhecido como algoritmo SC ou *Subtractive Clustering*, o qual foi originalmente proposto em (Chiu 1994). O algoritmo SC agrupa o conjunto de dados entrada-saída -utilizados para fins de ajuste do FIS- em um total de M conjuntos, sendo o processo realizado de forma não supervisionada.

O algoritmo SC proposto em (Chiu 1994), vem sendo um algoritmo bastante empregado para a geração de modelos baseados em regras, devido a sua simplicidade e capacidade para determinar de forma automática o número de centros necessários para representar de forma adequada o conjunto de dados, o qual é uma grande vantagem quando comparado a outros algoritmos de agrupamento, como por exemplo, o algoritmo tradicional *Fuzzy C-Means* (FCM) (Bezdek 1981).

A idéia base do algoritmo SC é a seguinte. O algoritmo define um índice de potencial P_k para cada dado entrada-saída, com $k = 1, \dots, N^0$. O potencial de cada dado depende diretamente da distância deste para todos os outros dados do conjunto apresentado. Assim, quanto menor for a distância do dado aos outros, seu potencial será maior. Desta forma, o potencial pode ser definido como:

$$P_k = \sum_{j=1}^N \exp\{-\alpha \|\psi_k - \psi_j\|^2\} \quad (2.48)$$

com $\psi_k = [\mathbf{x}^k \ y^k]$, $\alpha = 4/r_a^2$, onde \mathbf{x}^k é o vetor de entrada ao modelo composto por p componentes; y^k é a saída unidimensional e r_a é um termo constante e positivo que representa o parâmetro de dispersão associado a cada grupo.

Deve-se observar que, todo dado localizado em uma região com uma densidade alta de dados, terá um potencial maior, tendo assim, uma possibilidade mais alta de se tornar um novo centro. Além disso, o parâmetro r_a definirá o raio de influência na vizinhança de cada um dos centros, ou seja, a região na qual os centros exercerão a sua influência. Assim, r_a é o primeiro hiperparâmetro do modelo proposto.

Seja ψ_1 o dado entrada-saída com o maior potencial P_1^* . Assim, o potencial para todos os outros dados serão atualizados como:

$$P_k = P_k \times (1 - \exp\{-\beta \|\psi_k - \psi_1\|^2\}) \quad (2.49)$$

com $\beta = 4r_a^2/r_b^2$ e $r_b/r_a = r_{ba}$. Desta maneira, quanto mais próximo esteja um dado entrada-saída de um centro já definido, menor será o seu potencial, sendo menor a sua possibilidade de se tornar um novo centro. Neste caso, a relação de proximidade está diretamente vinculada ao valor de r_{ba} , o qual torna-se assim no segundo hiperparâmetro do FIS, e em conjunto com r_a influencia diretamente na estrutura do modelo FIS a otimizar na segunda etapa de ajuste de parâmetros.

O Algoritmo 2.2.3 resume os passos que devem ser seguidos para a definição dos centros.

Assim, após a identificação do número de regras, os parâmetros do FIS são inicializados seguindo os seguintes critérios:

1. $\mathbf{c}_i^0 = \psi_i^0|_{1..p}$, onde $\psi_i^0|_{1..p}$ é composto pelos primeiros p elementos do i -ésimo centro encontrado pelo algoritmo SC;
2. $\sigma_i^0 = 1.0$;
3. $\theta_i^0 = [\psi_i^0|_{p+1} \ 0 \ \dots \ 0]_{1 \times p+1}$, onde $\psi_i^0|_{p+1}$ é a $p+1$ -ésima coordenada do i -ésimo centro encontrado pelo algoritmo SC;
4. $\mathbf{V}_i^0 = 10^{-4}\mathbf{I}$, onde \mathbf{I} é a matriz identidade de ordem p ;
5. $\alpha_i^0 = 1/M$, o que garante pesos iniciais homogêneos para todas as M regras nebulosas.

Algorithm 1 Algoritmo de agrupamento Subtrativo - SC

Sejam N dados entrada-saída $[x_1^k, \dots, x_p^k, y^k]$, $k = 1, \dots, N$ disponíveis para o ajuste do modelo de previsão. Definir o valor de r_a e r_{ba} . Calcular o potencial de todos os N dados utilizando a Equação (2.48) e selecionar o primeiro centro como sendo igual ao padrão com maior potencial (P_1^*).

Passo 1. Atualizar os potenciais P_k de todos os dados utilizando a Equação (2.49).

Passo 2. Verificar qual é o dado com potencial mais alto, sendo o potencial maior denotado por P_k^* ;

Passo 3. Se $P_k^* > \bar{\varepsilon}P_1^*$, com $\bar{\varepsilon} = 0,5$, então ψ_k^* é definido como um novo centro, ir ao Passo 5. Senão, ir ao Passo 4.

Passo 4. Se $P_k^* < \underline{\varepsilon}P_1^*$, com $\underline{\varepsilon} = 0,15$, então o possível centro é rejeitado e o processo termina. Senão, ir ao Passo 6.

Passo 5. Atualizar os potenciais utilizando a Equação (2.49) e voltar ao Passo 2.

Passo 6. Seja d_{min} a distância mínima entre ψ_k^* e todos os outros centros já selecionados. Se

$$\frac{d_{min}}{r_a} + \frac{P_k^*}{P_1^*} \geq 1$$

então ψ_k^* é definido como um novo centro e retorne ao Passo 5. Senão, ψ_k^* é rejeitado e $P_k^* = 0$. Selecionar o próximo dado com maior potencial e defini-lo como um possível novo centro. Verificar novamente a condição do Passo 6.

Ajuste Após a inicialização do FIS, os seus parâmetros são re-ajustados utilizando o algoritmo da Maximização da Verossimilhança (EM), na sua versão *offline*, através de uma sequência de passos E-M (*Expectation-Maximization*), originalmente proposto para o ajuste de modelos de mistura de especialistas (Jacobs, Jordan, Nowlan & Hinton 1991).

O objetivo do algoritmo EM é encontrar um conjunto de valores para os parâmetros do FIS, tal que a função de verossimilhança seja maximizada. Em termos práticos, trabalha-se com o logaritmo da função de verossimilhança dos valores observados y^k . Esta função objetivo é denotada por \mathcal{L} , e definida como segue

$$\mathcal{L}(D, \mathbf{\Omega}) = \sum_{k=1}^N \ln \left(\sum_{i=1}^M g_i(\mathbf{x}^k, \mathbf{C}) \times P(y^k | \mathbf{x}^k, \theta_i) \right) \quad (2.50)$$

onde $D = \{\mathbf{x}^k, y^k | k = 1, \dots, N\}$, $\mathbf{\Omega}$ representa o conjunto de parâmetros a otimizar e \mathbf{C} contém apenas os parâmetros dos antecedentes (centros e matrizes de covariância). No entanto, para maximizar $\mathcal{L}(D, \mathbf{\Omega})$, é necessário estimar os parâmetros incompletos ou desconhecidos h_i^k durante o passo E. Este dado faltante, é definido como a probabilidade a posteriori de \mathbf{x}^k pertencer à região ativa da i -ésima regra nebulosa.

Quando o algoritmo EM é adaptado para o ajuste de sistemas nebulosos, h_i^k pode ser interpretado como a estimativa a posteriori do grau de pertinência g_i^k , dada pela Eq. (2.45). Assim,

h_i^k é calculada de acordo com

$$h_i^k = \frac{\alpha_i P(i | \mathbf{x}^k) P(y^k | \mathbf{x}^k, \theta_i)}{\sum_{q=1}^M \alpha_q P(q | \mathbf{x}^k) P(y^k | \mathbf{x}^k, \theta_q)} \quad (2.51)$$

para $i = 1, \dots, M$. Estas estimativas são chamadas a posteriori, pois são calculadas assumindo que y^k , $k = 1, \dots, N$ já é conhecida. Além disso, a probabilidade condicional $P(y^k | \mathbf{x}^k, \theta_i)$ é dada por:

$$P(y^k | \mathbf{x}^k, \theta_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{[y^k - y_i^k]^2}{2\sigma_i^2}\right) \quad (2.52)$$

com σ_i^2 estimada como segue:

$$\sigma_i^2 = \left(\sum_{k=1}^N h_i^k [y^k - y_i^k]^2 \right) / \sum_{k=1}^N h_i^k \quad (2.53)$$

Portanto, o algoritmo EM para a otimização dos parâmetros do FIS pode ser resumido nos seguintes passos:

1. Passo E: estimar h_i^k utilizando a Eq. (2.51);
2. Passo M: maximizar a função objetivo da Eq. (2.50) e atualizar os parâmetros dos modelos, com valores ótimos dados por:

$$\alpha_i = \frac{1}{N} \sum_{k=1}^N h_i^k \quad (2.54)$$

$$\mathbf{c}_i = \left(\sum_{k=1}^N h_i^k \mathbf{x}^k \right) / \sum_{k=1}^N h_i^k \quad (2.55)$$

$$\mathbf{V}_i = \left(\sum_{k=1}^N h_i^k (\mathbf{x}^k - \mathbf{c}_i)' (\mathbf{x}^k - \mathbf{c}_i) \right) / \sum_{k=1}^N h_i^k \quad (2.56)$$

para $i = 1, \dots, M$, N é o total de dados entrada-saída disponíveis para o ajuste do FIS. Em todas estas equações, \mathbf{V}_i foi considerada como uma matriz diagonal positiva, como uma alternativa para simplificar o problema e evitar soluções infactíveis. Uma solução ótima para θ_i é obtida a partir da solução da equação seguinte:

$$\sum_{k=1}^N \frac{h_i^k}{\sigma_i^2} (y^k - \phi^k \times \theta_i) \cdot \phi^k = 0 \quad (2.57)$$

onde σ_i é o desvio padrão referente a cada saída local y_i , $i = 1, \dots, M$, com σ_i^2 definido pela Eq.(2.53). Finalizado o passo M, calcular o novo valor de $\mathcal{L}(D, \mathbf{\Omega})$.

3. Se a convergência é atingida, logo, o processo é finalizado, senão, voltar ao passo 1.

2.2.4 Rede neurofuzzy

A metodologia desenvolvida em (Ballini 2000) para previsão de vazões, é baseada nos modelos de redes neurais e sistemas nebulosos.

A utilização de redes neurais para previsão de séries temporais vem sendo proposta por diversos autores, entre eles (Lapedes & Farber 1987), (Weigend et al. 1992), (Cottrell, Girard, Girard, Mangeas & Muller 1995), (Lachtermacher & Fuller 1995), mostrando a viabilidade de utilizar estes modelos.

Entretanto, a capacidade de mapeamentos complexos das redes neurais cresce com o número de camadas e neurônios, acarretando maior tempo de processamento, bem como considerável soma de dados.

Frequentemente, dados do mundo real apresentam ruídos, podendo conter contradições e imperfeições. Tolerância a imprecisão e incertezas é também exigido para considerar tratabilidade e robustez. Sob essas hipóteses, modelos de análise de dados baseados em conjuntos nebulosos vêm sendo empregados atualmente em diversas áreas (Kacprzyk & Fedrizzi 1992).

À medida que as teorias relacionadas às redes neurais e aos sistemas nebulosos se desenvolviam, observou-se que estas abordagens poderiam ser unificadas, resultando em sistemas de características mais completas no que se refere às suas capacidades de processamento inteligente. Além destas abordagens exibirem características comuns, isto é, tratamento de incertezas e raciocínio aproximado, cada uma apresenta propriedades exclusivas que contribuem para compensar deficiências individuais da outra. De um lado, as redes neurais falham em representar o conhecimento de forma explícita, mas exibem sua inclinação para o aprendizado. Por outro lado, os sistemas nebulosos são fracos sob o ponto de vista da aquisição automática de conhecimento, porém valem-se amplamente de suas características com relação à manipulação de termos linguísticos. O resultado desta simbiose se concretizou nas redes neurofuzzy ou nos chamados sistemas *neurofuzzy* (Pedrycz & Gomide 1998).

Esta abordagem vem sendo sucessivamente aplicada para modelagem de sistemas não-lineares e estimação de parâmetros (Figueiredo & Gomide 1998), (Wang 1994), reconhecimento de padrões (Caminhas, Tavares & Gomide 1996), controle de processo (Lee & Lin 1991), previsão de séries econômicas (Kuo & Xue 1999), previsão de carga (Bakirtzis, Theocharis, Kiartzis & Satsios 1995) e previsão de séries de vazões (Ballini 2000).

Em (Ballini, Soares & Andrade 1998) e (Ballini, Soares & Andrade 1999) foi utilizada uma estrutura de rede neurofuzzy para previsão de vazões, em que os parâmetros que definem as funções de pertinência do sistema nebuloso são obtidos através de um método de aprendizagem construtivo. Este modelo está sendo empregado para a previsão de vazões médias mensais e sua estrutura será descrita a seguir.

Estrutura da Rede Neurofuzzy

A estrutura da rede neurofuzzy (RNF) foi desenvolvida com base em duas características essenciais: o mapeamento das regras nebulosas na estrutura é direto e o processamento neural deve equivaler totalmente a um mecanismo de inferência nebuloso. Essas propriedades são condições necessárias para garantir que o sistema proposto tenha uma natureza dual, ou seja, o

sistema pode ser visto ora como um sistema baseado em regras nebulosas, ora como uma rede neurofuzzy (Pedrycz & Gomide 1998).

A rede é baseada no mecanismo de inferência nebuloso, codificando uma base de regras na forma de “Se um conjunto de condições é satisfeito Então um conjunto de consequentes é inferido”, o qual pode ser definido da seguinte forma:

Entradas:	$x_1 \text{ é } A_1 \text{ e } \dots \text{ e } x_M \text{ é } A_M$
Se $x_1 \text{ é } A_1^1 \text{ e } \dots \text{ e } x_M \text{ é } A_M^1$ Então $y \text{ é } w_1$	
Se $x_1 \text{ é } A_1^i \text{ e } \dots \text{ e } x_M \text{ é } A_M^i$ Então $y \text{ é } w_i$	
Regras:	\dots
Se $x_1 \text{ é } A_1^N \text{ e } \dots \text{ e } x_M \text{ é } A_M^N$ Então $y \text{ é } w_N$	
Saída: $y = f(\mathbf{x})$	

sendo $x_j, j = 1, \dots, M$ uma variável nebulosa e M o número de entradas; A_j e $A_j^i, i = 1, \dots, N$, são conjuntos nebulosos representados por funções de pertinência sendo N o número de regras nebulosas; y é um número real definido no espaço de saída. Os valores de $w_i(\mathbf{x}) = w_i$, são números reais e representam os pesos sinápticos correspondentes aos consequentes das regras nebulosas, sendo que, para cada vetor de entrada $\mathbf{x} = (x_1, \dots, x_j, \dots, x_M)'$, está associado um único peso sináptico w_i .

Para efeito de cálculo e representação de conhecimento na RNF, todas as funções de pertinência são discretizadas, ou seja, para cada conjunto nebuloso Z define-se uma função de pertinência $Z(\cdot) : U \rightarrow [0, 1]$, sendo U o universo de discurso, tal que $Z(x_j) = z_k$, se $x_j \in I_{j_k} = (x_{j_{k_I}}, x_{j_{k_F}}]$, sendo I_{j_k} o k -ésimo intervalo de discretização em relação à entrada $x_j, k = 1, \dots, Q$ (Figura 2.4).

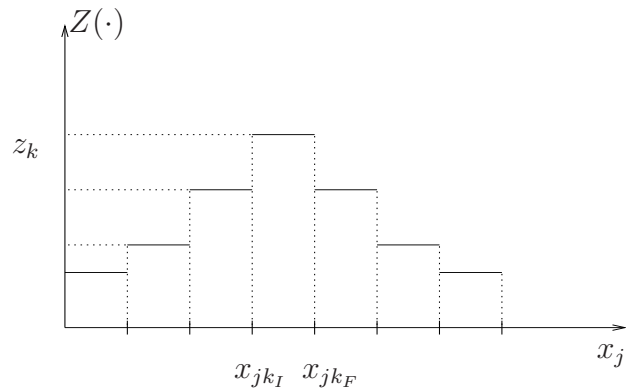


Figura 2.4: Função de pertinência discretizada.

Desta forma, define-se a_{j_k} como sendo o valor da função de pertinência do conjunto nebuloso A_j , para todo $x_j \in I_{j_k}$, ou seja, $A_j(x_j) = a_{j_k}$. O valor de $a_{j_k}^i$ é também definido como sendo o valor da função de pertinência do conjunto nebuloso A_j^i , isto é, $A_j^i(x_j) = a_{j_k}^i$, se $x_j \in I_{j_k}$.

Para a implementação do mecanismo de inferência nebuloso foi utilizado o método proposto

por (Yager & Filev 1994), sendo a saída y determinada a partir de três estágios de processamento, como segue:

1. *Comparação*: Para cada regra i , $i = 1, \dots, N$, e para cada antecedente j , $j = 1, \dots, M$, calcula-se a medida de possibilidade P_j^i entre os conjuntos nebulosos A_j e A_j^i , da seguinte forma:

$$P_j^i(\mathbf{x}) = S_k \{T(A_j(\mathbf{x}), A_j^i(\mathbf{x}))\} \quad (2.58)$$

sendo S e T operadores lógicos, correspondentes a uma s -norma e uma t -norma, respectivamente, e $\mathbf{x} = (x_1, \dots, x_j, \dots, x_M)'$ o vetor de entrada.

A notação $S_k(\cdot, \cdot)$ é a aplicação da norma $S(\cdot, \cdot)$ para todo intervalo de discretização $k = 1, \dots, Q$ relativos à variável do vetor de entrada x_j .

2. *Agregação dos Antecedentes*: Para cada regra i , $i = 1, \dots, N$, agregam-se os valores de cada um dos antecedentes através de uma t -norma. O resultado desta agregação é chamado de nível de ativação H^i , definido por:

$$H^i(\mathbf{x}) = T_j \{P_j^i(\mathbf{x})\} \quad (2.59)$$

sendo a notação $T_j(\cdot, \cdot)$ a aplicação da norma $T(\cdot, \cdot)$ para todo $j = 1, \dots, M$ relativo ao vetor de entrada \mathbf{x} .

3. *Agregação das Regras*: A saída y do sistema de inferência, ou seja, a etapa de defuzzificação dos dados, é realizada através de uma média ponderada entre os consequentes e os níveis de ativação calculados em (2.59) para cada regra, dada por:

$$y(\mathbf{x}) = \frac{\sum_{i=1}^N H^i(\mathbf{x}) w_i}{\sum_{i=1}^N H^i(\mathbf{x})} \quad (2.60)$$

A RNF é constituída por unidades de processamento, chamadas de neurônios nebulosos. Um modelo geral do neurônio nebuloso é mostrado na Figura 2.5. Matematicamente, este neurônio pode ser representado da seguinte forma:

$$y(\mathbf{x}) = \varphi(\phi(\Psi(\mathbf{x}))) = \varphi(\phi(\psi_1(x_1), \dots, \psi_M(x_M)))$$

o qual $\Psi : \mathfrak{R}^M \rightarrow \mathfrak{R}^M$ é o operador sináptico, $\phi : \mathfrak{R}^M \rightarrow \mathfrak{R}$ é o operador de agregação e $\varphi : \mathfrak{R} \rightarrow \mathfrak{R}$ é a função de ativação.

O modelo RNF apresenta uma arquitetura não recorrente com 5 camadas, como mostra a Figura 2.6. Pode-se verificar que o mecanismo de inferência descrito através dos procedimentos de comparação, agregação dos antecedentes e agregação das regras é isomórfico à estrutura da rede.

A primeira camada é dividida em M grupos de neurônios, cada um associado a uma variável nebulosa de entrada. Cada neurônio nesta camada representa um intervalo de discretização I_{jk} , correspondente ao espaço de entrada. Ou seja, há tantos neurônios nesta camada quantos forem

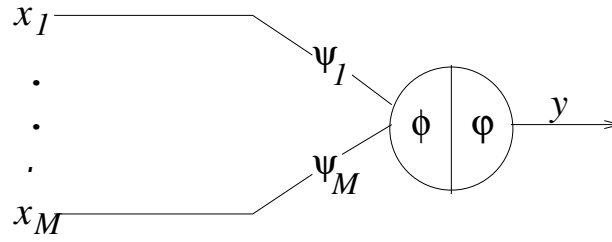


Figura 2.5: Modelo geral do neurônio nebuloso.

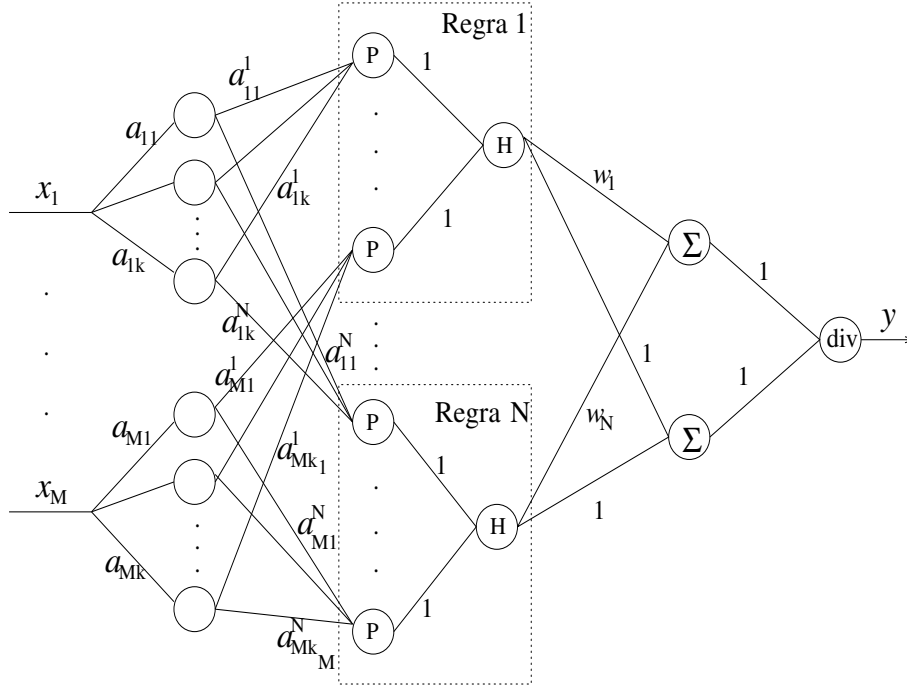


Figura 2.6: Rede neurofuzzy.

os intervalos de discretização. Cada grupo de neurônios é responsável por transformar as entradas não nebulosas para que as camadas seguintes possam tratar a informação adequadamente. Estes grupos geram os valores a_{jk} referentes à função de pertinência do conjunto nebuloso A_j . Assim, um neurônio de entrada recebe um sinal simples, decodifica e o transmite para a segunda camada. O sinal a_{jk} é transmitido pelo k -ésimo neurônio localizado no j -ésimo grupo.

Definindo-se ψ e ϕ como funções identidade, a saída $a_{jk} \in \{0, 1\}$ é dada por $a_{jk} = \varphi(x_j)$. Para um intervalo $I_{jk} = (x_{jk_I}, x_{jk_F}]$, a função decodificada $\varphi(\cdot)$ para este neurônio, através do k -ésimo intervalo, mostrada na Figura 2.7, é dada por:

$$a_{jk} = \varphi(x_j) = \begin{cases} 1, & \text{se } x_j \in (x_{jk_I}, x_{jk_F}] \\ 0, & \text{caso contrário} \end{cases}$$

A segunda camada contém N grupos (número de regras pré-definido), cada qual contendo M

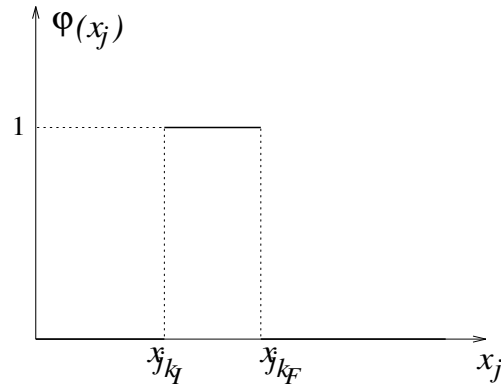


Figura 2.7: Função de decodificação.

neurônios (número de antecedentes de cada regra). Esta camada representa o primeiro estágio de inferência chamado comparação. O j -ésimo neurônio do i -ésimo grupo representa, através dos pesos sinápticos $a_{j_k}^i$, a função de pertinência do conjunto nebuloso A_j^i . Portanto, o i -ésimo grupo é composto por neurônios que representam as funções de pertinência de todos os conjuntos nebulosos correspondentes aos antecedentes da i -ésima regra. Além disso, cada neurônio da i -ésima regra calcula a comparação P_j^i entre os conjuntos nebulosos A_j e A_j^i . Assim, o k -ésimo neurônio do j -ésimo grupo da primeira camada, cuja saída é dada por a_{j_k} , se conecta com o i -ésimo grupo da segunda camada através de pesos sinápticos $a_{j_k}^i$. Considerando-se que $\psi = t$ -norma, $\phi = s$ -norma e φ é a função identidade, a saída do j -ésimo neurônio é a medida de possibilidade P_j^i dada pela equação (2.58).

A terceira camada contém N neurônios, cada um com M entradas. Para cada grupo i da segunda camada, um neurônio na terceira camada realiza agregação dos antecedentes (segundo estágio de inferência). Todas as conexões possuem peso sináptico unitário.

A saída desta camada é dada pela equação (2.59), assumindo que $\psi =$ função identidade, $\phi = t$ -norma e $\varphi =$ função identidade.

A quarta camada possui dois neurônios, ambos com N entradas. Ou seja, cada neurônio i da terceira camada está conectado com os dois neurônios da quarta camada. O operador de agregação associado com estes neurônios é a soma algébrica.

Um dos neurônios conecta-se com o i -ésimo neurônio da camada anterior através das sinapses, com pesos w_i . Os operadores ψ e ϕ são, respectivamente, o operador algébrico e a soma algébrica, e φ é a função identidade. As saídas são constituídas pelo numerador da equação (2.60).

O outro neurônio também é conectado com todos os neurônios da camada anterior com pesos sinápticos unitários. Considerando $\psi =$ função identidade, $\phi =$ soma algébrica e $\varphi =$ função identidade, a saída equivale ao denominador da inferência nebulosa definida em (2.60).

Finalmente, a última camada consiste de um único neurônio para calcular o quociente dos sinais de entrada, ou seja, para calcular $y(\mathbf{x})$ (equação (2.60)). Portanto, as conexões possuem pesos unitários e os operadores ψ , ϕ e φ são função identidade, operador de divisão e a função identidade, respectivamente.

Cada uma das camadas da RNF possui uma associação com os diversos estágios da inferên-

cia nebulosa, descritos de (2.58) a (2.60). Além disso, os pesos sinápticos representam tanto o conhecimento da rede neural como também representam as funções de pertinência dos antecedentes e consequentes das regras nebulosas. Assim, a RNF codifica um conjunto de N regras em sua estrutura, processa as informações através de neurocomputação, executando raciocínio nebuloso.

Método de Ajuste dos Parâmetros

O método de ajuste dos parâmetros da rede neuro-fuzzy, é chamado de método de aprendizagem ou algoritmo de aprendizagem

A estratégia de aprendizagem da RNF se divide em duas fases. A primeira fase corresponde à aprendizagem não supervisionada, ou seja, a aprendizagem se processa sem que as saídas sejam fornecidas, envolvendo somente uma parte da arquitetura, mais especificamente a segunda camada. Nesta fase, a rede aprende as funções de pertinência dos antecedentes das regras nebulosas enquanto ajusta os pesos sinápticos associados a esta camada. O método utilizado durante esta etapa é do tipo competitivo, baseado nas redes do tipo Kohonen. A segunda fase adota um método supervisionado, baseado no método do gradiente para aprendizagem dos consequentes das regras, e envolve apenas o ajuste dos pesos da quarta camada. Assim, a RNF deve aprender os pesos sinápticos a_{jk}^i relacionados aos neurônios da segunda camada, bem como os pesos sinápticos w_i da quarta camada.

O algoritmo de aprendizagem consiste na apresentação dos pares de entrada/saída à rede, o qual o t -ésimo par é dado por $(\mathbf{x}(t), y_d(t))$, sendo $\mathbf{x}(t) = (x_1(t), \dots, x_M(t))'$ o vetor de entrada com M coordenadas e $y_d(t)$ a saída desejada.

Considera-se que os conjuntos nebulosos, $A_j^i(\cdot)$, são representados por funções de pertinência do mesmo tipo, simétricas e completamente definidas pelos valores modais c_j^i e de dispersão ρ^i . Por exemplo, funções do tipo Gaussianas ou funções triangulares. Determinando-se o valor modal e de dispersão de cada função de pertinência $F_j^i(\cdot)$ e definidos os intervalos de discretização I_{jk} , é possível determinar os pesos sinápticos a_{jk}^i na forma:

$$a_{jk}^i = F_j^i(x_j), \quad \text{tal que} \quad x_j \in I_{jk} = \left(x_{jk_I}, x_{jk_F} \right]$$

Os valores modal e de dispersão de cada função $F_j^i(\cdot)$ devem ser aprendidos durante o treinamento, segundo o método de aprendizado competitivo, para determinar os pesos sinápticos a_{jk}^i da segunda camada.

A função erro que determina o índice de desempenho do modelo, necessária à aplicação do método baseado no gradiente descendente, é definida por:

$$E(\mathbf{x}(t)) = \frac{1}{2}(y_d(t) - y(t))^2 \quad (2.61)$$

sendo $y(t)$ a saída da rede em relação à entrada $\mathbf{x}(t)$, dada pela equação (2.60).

O objetivo do processo de treinamento é ajustar os parâmetros (valor modal, dispersão e pesos) da rede, minimizando $E(\mathbf{x}(t))$. Desta forma, o ajuste dos pesos sinápticos w_i é proporcional

à direção oposta ao gradiente $\partial E(\mathbf{x}(t))/\partial w_i(t)$ e é dado da seguinte forma:

$$w_i(t+1) = w_i(t) - \eta \frac{\partial E(\mathbf{x}(t))}{\partial w_i(t)} \quad (2.62)$$

sendo η a taxa de aprendizagem. Derivando a equação (2.61) em relação a w_i , tem-se:

$$\frac{\partial E(\mathbf{x}(t))}{\partial w_i} = \frac{(y_d(t) - y(t)) H^i(\mathbf{x}(t))}{\sum_{i=1}^n H^i(\mathbf{x}(t))} \quad (2.63)$$

Substituindo a equação (2.63) na equação (2.62), o ajuste dos pesos sinápticos $w_i(t)$ é rescrito da seguinte forma:

$$w_i(t+1) = w_i(t) + \eta \left(\frac{(y_d(t) - y(t)) H^i(\mathbf{x}(t))}{\sum_{h=1}^n H^h(\mathbf{x}(t))} \right) \quad (2.64)$$

Método de Aprendizagem O método de aprendizagem ou algoritmo de aprendizagem capacita a RNF a adquirir novos conhecimentos codificando novas regras nebulosas em sua estrutura, alterando a arquitetura da rede. Portanto, o número de regras nebulosas codificadas pela rede, correspondente ao número de grupos de neurônios na segunda camada, não é constante, ou seja, $N = N(t)$.

A estrutura da rede é alterada basicamente devido a dois motivos: o desempenho desejado não é satisfeito pela rede ou todos os grupos neurais da segunda camada são tais que o nível de ativação é igual a zero. Nestes casos, a rede acrescenta novos elementos à sua estrutura de maneira a codificar uma nova regra nebulosa. Portanto, os grupos de neurônios da segunda camada competem entre si sempre que a rede recebe uma nova entrada $\mathbf{x}(t)$. O grupo do neurônio vencedor guia os procedimentos de ajuste dos pesos sinápticos.

O algoritmo utiliza como fatores de aprendizagem para ajuste dos valores modais $c_j^i(t)$ das funções de pertinências $F_j^i(\cdot)$ e para ajuste dos consequentes das regras nebulosas $w_i(t)$, utilizam-se os fatores de aprendizagem $\beta^i(t)$ e $\eta \in (0, 1)$, respectivamente. O fator $\beta^i(t)$ é definido como $\beta^i(t) = \epsilon/(\alpha^i(t) + 1)$, sendo $\alpha^i(t)$ o número de vezes que o valor modal é ajustado até a t -ésima entrada e $\epsilon \in (0, 1]$. Finalmente, para o ajuste da dispersão das funções de pertinência é utilizado um fator de redução $\gamma \in (0, 1)$.

O algoritmo de aprendizagem pode ser definido através dos seguintes passos:

1. Inicialização:

- 1.1 Inicialize o número de regras. Neste caso, o número de regras inicial é $N(0) = 0$ para $t = 0$.
- 1.2 Defina os perfis das funções de pertinência $F_j^i(\cdot)$ que formam as regras nebulosas.
- 1.3 Defina a função $\beta^i(t)$.

- 1.4** Estabeleça os valores para ϵ , η , e γ .
- 1.5** Defina o valor inicial da dispersão das funções de pertinência, $\Theta \in \mathfrak{R}$.
- 1.6** Apresente o par entrada/saída inicial $(\mathbf{x}(t), y_d(t))$ à RNF. A rede codifica a primeira regra nebulosa após receber o primeiro par de entrada/saída. A seguir, o valor de N é atualizado e os valores modais são definidos como sendo cada componente do vetor de entrada e o conseqüente como a saída desejada. Inicialize o fator de aprendizagem $\alpha^i(t)$, ou seja:

$$N(t+1) = N(t) + 1 \quad (2.65)$$

Para $i = 1$ faça:

$$c_i^j(t+1) = x_j(t), \quad j = 1, \dots, M \quad (2.66)$$

$$\rho^i(t+1) = \Theta \quad (2.67)$$

$$w_i(t+1) = y_d(t) \quad (2.68)$$

$$\alpha^i(t+1) = 0 \quad (2.69)$$

$$t = t + 1 \quad (2.70)$$

2. Estímulo:

Apresente à rede um novo par $(\mathbf{x}(t), y_d(t))$ proveniente do sistema a ser modelado.

3. Competição:

A entrada $\mathbf{x}(t)$ é propagada através da rede até que seja calculada a saída da terceira camada, determinando-se, assim, o resultado da agregação dos antecedentes de cada uma das $N(t)$ regras. Dois casos são possíveis:

- 3.1** A entrada $\mathbf{x}(t)$ é caracterizada por pelo menos uma regra, ou seja, $\exists i \mid H^i(\mathbf{x}(t)) \neq 0$. A regra $\xi(t)$ vence a competição se $H^{\xi(t)}(\mathbf{x}(t)) \geq H^i(\mathbf{x}(t)), \forall i \in \{1, \dots, N(t)\}$.

Prossiga o processo de aprendizagem em **4.1**.

- 3.2** A entrada $\mathbf{x}(t)$ não é caracterizada por nenhuma regra, isto é, $H^i(\mathbf{x}(t)) = 0, \forall i \in \{1, \dots, N(t)\}$. Neste caso, a distância $D(\mathbf{c}^i(t), \mathbf{x}(t))$ é determinada entre o vetor de entrada e o vetor de valores modais $\mathbf{c}^i(t) = (c_1^i(t) \dots c_M^i(t))'$ e o grupo vencedor $\xi(t)$ é tal que $D(\mathbf{c}^{\xi(t)}(t), \mathbf{x}(t)) \leq D(\mathbf{c}^i(t), \mathbf{x}(t)), \forall i \in \{1, \dots, N(t)\}$.

Prossiga o processo de aprendizagem em **4.2**.

4. Ajuste e Atualização da Rede:

- 4.1** Se a entrada $\mathbf{x}(t)$ foi classificada por alguma regra (caso **3.1**), a saída da terceira camada, determinada para realizar a competição, é propagada até a última camada, calculando-se, assim, a saída da rede. Então os seguintes procedimentos são realizados:

- (a) Se a saída da rede satisfizer o desempenho desejado, ou seja, $\|y_d(t) - y(t)\| \leq \delta$, então os parâmetros são ajustados de acordo com as seguintes equações:

$$c_j^{\xi(t)}(t+1) = c_j^{\xi(t)}(t) + \beta^i(t)(x_j(t) - c_j^{\xi(t)}(t)), \quad j = 1, \dots, M \quad (2.71)$$

$$c_j^i(t+1) = c_j^i(t), \quad i = 1, \dots, N(t), i \neq \xi(t), \quad j = 1, \dots, M \quad (2.72)$$

$$\rho^i(t+1) = \rho^i(t), \quad i = 1, \dots, N(t) \quad (2.73)$$

$$w_i(t+1) = w_i(t) + \eta \left(\frac{H^i(\mathbf{x}(t))(y_d(t) - y(t))}{\sum_{h=1}^{N(t)} H^h(\mathbf{x}(t))} \right), \quad i = 1, \dots, N(t) \quad (2.74)$$

$$\alpha^{\xi(t)}(t+1) = \alpha^{\xi(t)}(t) + 1 \quad (2.75)$$

$$\alpha^i(t+1) = \alpha^i(t), \quad i = 1, \dots, N(t), i \neq \xi(t) \quad (2.76)$$

$$N(t+1) = N(t) \quad (2.77)$$

Prossiga o processo de aprendizagem em 5.

- (b) Pode ser que a saída da rede não satisfizesse o desempenho esperado, $\|y_d(t) - y(t)\| > \delta$ e exista um vetor de valores modais de uma determinada regra q igual ao padrão de entrada ($\mathbf{x}(t) = \mathbf{c}^q(t) = (c_1^q(t) \dots c_M^q(t))'$), ou seja, a rede possui uma regra com antecedentes iguais ao padrão de entrada porém o consequente não está sendo calculado corretamente devido à interferência indesejável de outras regras. Para solucionar este problema, os valores modais atuais são mantidos e as dispersões são reduzidas para que as regras nebulosas diferentes de q não influenciem no resultado da entrada $\mathbf{x}(t)$. Assim, os seguintes ajustes são realizados:

$$c_j^i(t+1) = c_j^i(t), \quad i = 1, \dots, N(t), \quad j = 1, \dots, M \quad (2.78)$$

Para $i = 1$ até $N(t)$ e $i \neq q$ faça:

Enquanto $H^i(\mathbf{x}(t)) \neq 0$ faça:

$$\rho^i(t) = \gamma \rho^i \quad (2.79)$$

$$\rho^i(t+1) = \rho^i(t) \quad (2.80)$$

Seja $k \neq q$ tal que $D(\mathbf{c}^k(t), \mathbf{x}(t)) \leq D(\mathbf{c}^i(t), \mathbf{x}(t)), \forall i \in \{1, \dots, N(t)\}$, então:

$$\rho^q(t+1) = D(\mathbf{c}^k(t), \mathbf{x}(t)) \quad (2.81)$$

Os demais ajustes são definidos pelas seguintes equações:

$$w_q(t+1) = y_d(t) \quad (2.82)$$

$$w_i(t+1) = w_i(t), \quad i = 1, \dots, N(t), \quad i \neq q \quad (2.83)$$

$$\alpha^q(t+1) = \alpha^q(t) + 1 \quad (2.84)$$

$$\alpha^i(t+1) = \alpha^i(t), \quad i = 1, \dots, N(t), \quad i \neq q \quad (2.85)$$

$$N(t+1) = N(t) \quad (2.86)$$

Prossiga o processo de aprendizagem em 5.

- (c) Se a saída da rede não satisfizer o desempenho desejado, $\| y_d(t) - y(t) \| > \delta$ e não existir um vetor de valores modais igual ao vetor de entrada, ou seja, $\mathbf{x}(t) \neq \mathbf{c}^i(t) = (c_1^i(t) \dots c_M^i(t))'$, $\forall i \in \{1, \dots, N(t)\}$, isto significa que uma nova regra deve ser adicionada à estrutura da rede. As equações de ajuste dos parâmetros são:

$$c_j^i(t+1) = c_j^i(t), \quad i = 1, \dots, N(t), \quad j = 1, \dots, M \quad (2.87)$$

Reduza a dispersão de cada função $F_j^i(\cdot)$ para que as regras existentes deixem de influenciar no resultado da entrada $\mathbf{x}(t)$.

Para $i = 1$ até $N(t)$ faça:

Se $H^i(\mathbf{x}(t)) \neq 0$ então:

$$\rho^i(t) = \gamma \rho^i \quad (2.88)$$

$$\rho^i(t+1) = \rho^i(t) \quad (2.89)$$

$$w_i(t+1) = w_i(t), \quad i = 1, \dots, N(t) \quad (2.90)$$

$$\alpha^{\xi(t)}(t+1) = \alpha^{\xi(t)}(t) + 1 \quad (2.91)$$

$$\alpha^i(t+1) = \alpha^i(t), \quad i = 1, \dots, N(t), \quad i \neq \xi(t) \quad (2.92)$$

Realiza-se a inserção de uma nova regra à estrutura da rede. Os novos valores modais são iguais a cada componente do vetor de entrada, a dispersão é igual à distância $D(\mathbf{c}^{\xi(t)}(t), \mathbf{x}(t))$ e o consequente é igual à saída desejada $y_d(t)$:

$$N(t+1) = N(t) + 1 \quad (2.93)$$

Para $i = N(t+1)$ faça:

$$c_j^i(t+1) = x_j(t), \quad j = 1, \dots, M \quad (2.94)$$

$$\rho^i(t+1) = D(\mathbf{c}^{\xi(t)}(t), \mathbf{x}(t)) \quad (2.95)$$

$$w_i(t+1) = y_d(t) \quad (2.96)$$

$$\alpha^i(t+1) = 0 \quad (2.97)$$

Prossiga o processo de aprendizagem em 5.

- 4.2** Se o vetor de entrada $\mathbf{x}(t)$ não foi classificado por nenhuma regra, ou seja, o grupo vencedor $\xi(t)$ foi escolhido de acordo com o caso 3(b), então é necessário alterar a estrutura da rede, acrescentando uma nova regra capaz de codificar o padrão $(\mathbf{x}(t), y_d(t))$. Os ajustes são realizados da seguinte forma:

$$c_j^i(t+1) = c_j^i(t), \quad i = 1, \dots, N(t), \quad j = 1, \dots, M \quad (2.98)$$

$$\rho^i(t+1) = \rho^i(t), \quad i = 1, \dots, N(t) \quad (2.99)$$

$$w_i(t+1) = w_i(t), \quad i = 1, \dots, N(t) \quad (2.100)$$

$$\alpha^i(t+1) = \alpha^i(t), \quad i = 1, \dots, N(t) \quad (2.101)$$

$$N(t+1) = N(t) + 1 \quad (2.102)$$

Os valores modais para a nova regra são iguais a cada componente do vetor de entrada, a dispersão é igual a $D(\mathbf{c}^{\xi(t)}(t), \mathbf{x}(t))$ e o conseqüente é igual a saída desejada:

Para $i = N(t + 1)$ faça:

$$c_j^i(t + 1) = x_j^i(t), \quad j = 1, \dots, M \quad (2.103)$$

$$\rho^i(t + 1) = D(\mathbf{c}^{\xi(t)}(t), \mathbf{x}(t)) \quad (2.104)$$

$$w_i(t + 1) = y_d(t) \quad (2.105)$$

$$\alpha^i(t + 1) = 0 \quad (2.106)$$

Prossiga o processo de aprendizagem em 5.

5. Teste de Parada:

Se ainda há padrões a serem treinados, volte ao passo 2. Caso contrário, fim do treinamento.

O método de aprendizagem capacita a rede a adquirir novos conhecimentos quando necessário, fornecendo uma forma automática de aprender os parâmetros para um modelo nebuloso sem a interferência ou participação de um especialista.

2.2.5 Previsões agregadas

Uma proposta desse trabalho é o uso de previsões de vazões baseada em um histórico de vazões anuais. Em Zambelli et al. (2009), as previsões de vazões mensais baseadas em previsões de vazões anuais desagregadas mensalmente apresentaram melhor desempenho na otimização do despacho hidrotérmico do que previsões baseadas diretamente nas vazões mensais. Os resultados justificam-se pela variabilidade das séries anuais, a qual é menor se comparada com a variabilidade apresentada nas séries de vazões mensais.

Além disso, enquanto os modelos anuais estimam a vazão agregada doze meses à frente em um único passo, os modelos mensais aumentam o horizonte de previsão a cada mês, realimentando erros de previsão dos passos anteriores ao modelo a cada nova previsão, deteriorando o desempenho com o aumento do horizonte de previsão (Lopes, Luna, Ballini & Soares 2012).

Com base nesses estudos preliminares, procedeu-se ao ajuste dos modelos agregados. Estes modelos são obtidos a partir do ajuste de modelos de vazões agregadas resultantes a partir da soma de doze meses consecutivos à frente para cada instante (mês) do histórico de vazões mensais. A partir desse histórico agregado realiza-se o ajuste de um modelo agregado. Os valores estimados um passo à frente pelo modelo agregado são desagregados em doze valores mensais considerando os fatores de participação médios mensais do histórico (Média de longo termo ou MLT) no valor histórico médio anual (Figura 2.8). Assim, estes doze valores corresponderão às doze estimativas das vazões dos próximos meses de um até doze passos à frente.

O modelo ajustado sobre as vazões anuais de cada posto considerado é a rede RNF, sendo o procedimento de normalização, identificação e ajuste dos parâmetros dos modelos igual ao descrito para o caso dos modelos mensais.

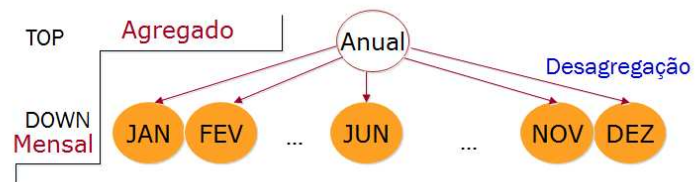


Figura 2.8: Vazões anuais

Geração de Séries Sintéticas

3.1 Introdução

O conhecimento da disponibilidade de água é de suma importância, como foi dito anteriormente, não só para a vida em geral como também economicamente. Principalmente para o Brasil, que tem a maior parte de sua energia produzida por usinas hidrelétricas (Martinez 2001). A geração de energia hidrelétrica consolida-se como uma atividade de extrema importância para a economia em geral. A fim de construir usinas bem como operar reservatórios de forma eficiente, faz-se necessário estudar o comportamento da água. Devido a complexidade dos fenômenos hídricos, existe a necessidade de estabelecer modelos capazes de representar com nível de precisão satisfatório o comportamento da água. No caso de usinas hidrelétricas por exemplo, a água proveniente da chuva direciona-se aos rios que, por sua vez, alimentam os reservatórios dessas usinas, o que reflete diretamente na operação do reservatório de uma usina.

De um modo geral, análises estatísticas relacionadas às vazões encontram limitações relacionadas ao tamanho das séries históricas disponíveis, que em sua maioria são insuficientes. Uma boa alternativa para superar essas limitações é a geração de séries sintéticas de vazões. Esta alternativa é utilizada por vários pesquisadores ao longo dos últimos anos, que desenvolveram modelos baseados em propriedades determinísticas e estocásticas, aplicados a várias escalas de tempo (anual, mensal, diária), de acordo com o mérito de cada estudo.

Segundo Wilks (1999), uma série sintética é o resultado da associação de séries de dados reais com números aleatórios produzidos por algoritmos computacionais a fim de gerar sequências de variáveis aleatórias que se assemelham aos dados reais.

Modelos de geração de séries sintéticas de vazões tornaram-se uma ferramenta muito útil para o planejamento do uso de recursos hídricos. Atualmente, são amplamente utilizados em simulações e técnicas de otimização, visando produzir resultados comparáveis para as várias alternativas de planejamento, projeto ou operação de recursos hídricos.

Os primeiros modelos para geração de vazões surgiram no início do século XX com o trabalho pioneiro de (Hazen 1914), que obteve uma sequência prolongada de vazões, combinando as vazões anuais de quatorze rios individuais. Vários trabalhos surgiram depois deste. Mas um trabalho que deu origem aos modelos atuais foi o de Thomas & Fiering (1962b). Uma revisão destes modelos é apresentada em (Lawrence & Kottagoda 1977) e (Phien & Ruksasilp 1981).

Os modelos para geração de séries sintéticas de vazões podem ser classificados preliminarmente em modelos para geração de séries de vazões diárias e semanais, modelos para séries de vazões mensais e modelos para séries vazões anuais. Os modelos para séries diárias e semanais geralmente utilizam dados climatológicos. O mais comum é o que usam dados de precipitação. No entanto, modelos para séries mensais e anuais também podem fazer uso destas variáveis. Os modelos para séries anuais, em geral, são mais simples porque estas séries apresentam, na maioria das vezes, um padrão estacionário e baixa correlação plurianual. Contudo é importante que estes modelos reproduzam a dependência de longa duração ou fenômeno de Hurst (Lin & Lye 1994).

Em seguida descreve-se os modelos implementados no módulo HydroSeries que foi acrescentado no sistema HydroLab. Este módulo será descrito no Capítulo 5.

3.2 Modelos lineares de geração de séries sintéticas

Os modelos lineares considerados neste trabalho para a geração de séries sintéticas de vazões diferem dos utilizados para fins de previsão, pois estes consideram uma distribuição log-normal generalizada das vazões mensais. Estes modelos são denominados modelos periódicos autorregressivo log-normal de três parâmetros $BCPAR(p_m)$, ou simplesmente $BCPAR$, que usam como referencial teórico o modelo anual lognormal de três parâmetros descrito a seguir, para o caso de vazões anuais.

3.2.1 Modelos log-normal com três parâmetros

A seguir o modelo autorregressivo log-normal com três parâmetros considerando vazões anuais é descrito. Este modelo será o referencial teórico para a descrição do modelo periódico autorregressivo log-normal de três parâmetros $BCPAR(p_m)$, a ser usado na geração de series sintéticas de vazões afluentes mensais.

Estrutura geral e método de ajuste dos parâmetros

Considerando o histórico de vazões anuais $x_t, t = 1, \dots, n$ e o modelo AR(p) ajustado para esta série temporal dado por:

$$x_t = c + \sum_{j=1}^p \phi_j x_{t-j} + a_t, \quad (3.1)$$

em que a_t é o ruído com $E(a_t) = 0$, $Var(a_t) = \sigma_a^2$ e $E(a_t a_{t+k}) = 0$, para $k \neq 0$.

A elaboração de um modelo para geração de vazões anuais, tem como ponto principal a identificação da distribuição de probabilidade do ruído a_t , o mais próximo possível da distribuição real. Entre as diversas características observadas na prática, da distribuição de a_t , a mais difícil de ser modelada é a assimétrica. Muitos autores tem adotado a suposição de que a_t tem distribuição log-normal com três parâmetros (Stedinger & Taylor 1982), (Pereira, Oliveira, Costa & Kelman 1984), (Jardim, Maceira & ao 2001). Esta suposição é considerada assumindo-se que o processo transformado por $\xi_t = \ln(a_t + \tau)$ tem distribuição normal $N(\mu_\xi, \sigma_\xi^2)$. Neste contexto, o

parâmetro τ deve ser estimado de forma que $a_t + \tau > 0$, para todo $t > 0$. Um método simples de estimação de τ é o método dos quantis, proposto por (Boswell, Ord & Patil 1979), que considera a série ordenada $a_{(1)} < a_{(2)} < \dots < a_{(n)}$ e denotamos por m_a a mediana da série a_t . Então o estimador de τ é dado por:

$$\hat{\tau} = \frac{m_a^2 - a_{(1)}a_{(n)}}{a_{(1)} + a_{(n)} - 2m_a}, \quad (3.2)$$

Os parâmetros da distribuição normal $N(\mu_{\xi_t}, \sigma_{\xi_t}^2)$ são estimados de forma a preservar a média e a variância do ruído a_t , usando-se as seguintes relações:

$$\mu_{\xi} = \frac{1}{2} \ln\left(\frac{\sigma_a^2}{\theta_2 - \theta}\right), \quad (3.3)$$

$$\sigma_{\xi}^2 = \ln(\theta), \quad (3.4)$$

$$\theta = 1 + \left(\frac{\sigma_a}{\mu_a + \tau}\right)^2. \quad (3.5)$$

Os parâmetros μ_{ξ} e σ_{ξ}^2 são estimados usando as equações (3.3)-(3.5), considerando $\hat{\tau}$ dado em (3.2), $E(a_t) = \mu_a = 0$ e σ_a^2 estimado dado por:

$$\hat{\sigma}_a^2 = \frac{1}{n} \sum_{t=1}^n \left(x_t - \hat{c} - \sum_{j=1}^p \hat{\phi}_j x_{t-j} \right)^2. \quad (3.6)$$

A limitação do modelo definido com as equações (3.3)-(3.5) é a suposição de que o ruído a_t segue uma distribuição log-normal com três parâmetros ou de forma equivalente que $\xi_t = \ln(a_t + \tau)$ segue uma distribuição normal $N(\mu_a, \sigma_a^2)$.

Para contornar as dificuldades encontradas na assimetria das distribuições de vazões podemos fazer uso de transformações. Estamos propondo uma generalização deste modelo considerando uma família de transformações, proposta em (Box & Cox 1964), na qual podemos transformar o ruído a_t em um ruído $\xi_t = \Lambda(a_t, \lambda)$ de forma a termos uma distribuição para ξ_t aproximadamente simétrica. Para isso vamos considerar um valor do parâmetro λ que define uma transformação específica dada por:

$$\xi_t = \begin{cases} \frac{(a_t + \tau)^\lambda - 1}{\lambda} & \text{se } \lambda \neq 0 \\ \ln(a_t + \tau) & \text{se } \lambda = 0 \end{cases} \quad (3.7)$$

em que ξ_t tem distribuição simétrica denotada aqui por $F(\xi_t; \mu_{\xi}, \sigma_{\xi}, s)$. O parâmetro μ_{ξ} é denominado parâmetro de locação e σ_{ξ} é o parâmetro de dispersão. A vantagem da abordagem proposta é que a família das distribuições simétricas é muito vasta e amplamente estudada. No entanto vamos considerar uma família de distribuição dentro da classe das distribuições simétricas, denominadas de distribuição normal generalizada, (Nadarajah 2005). Assim, um procedimento sistemático pode ser adotado para a escolha da distribuição que melhor represente a verdadeira distribuição do ruído transformado ξ_t . Para usar a equação (3.7), o parâmetro τ é

estimado usando a equação (3.2) e o valor específico do parâmetro λ da transformação, que torna a distribuição de ξ_t simétrica e é calculado considerando-se uma sequência de valores de λ em um determinado intervalo $[\lambda_{min}, \lambda_{max}]$ e escolhendo-se λ^* tal que o coeficiente de assimetria do processo ξ_t , que denotamos aqui por $\widehat{\gamma}_{\xi_t}$, seja aproximadamente zero. O procedimento adotado para encontrar λ^* consiste em calcular as estimativas $\widehat{\gamma}_{\xi_t}$ dada por:

$$\widehat{\gamma}_{\xi_t} = \frac{\widehat{\mu}_{\xi_t}^{(3)}}{[\widehat{\mu}_{\xi_t}^{(2)}]^{3/2}}, \quad (3.8)$$

em que $\widehat{\mu}_{\xi_t}^{(k)}$ é o estimador de k -ésimo momento do ruído ξ_t , para um certo valor de λ .

Portanto o parâmetro λ^* é dado por:

$$\lambda^* = \min_{\lambda \in (\lambda_{min}, \lambda_{max})} (\widehat{\gamma}_{\xi_t}^2) \quad (3.9)$$

O valor λ^* pode ser calculado numericamente ou simplesmente graficamente observando no gráfico de $\widehat{\gamma}_{\xi_t}^2$ por λ , o valor λ^* para o qual $\widehat{\gamma}_{\xi_t}^2(\lambda^*) \approx 0$.

O objetivo da transformação dada em (3.7), com λ^* satisfazendo a equação (3.9) é obter uma função de distribuição aproximadamente simétrica para o ruído ξ_t . Várias distribuições simétricas podem ser testadas e usadas para diferentes situações práticas. Neste trabalho, vamos considerar a distribuição normal generalizada, proposta por (Nadarajah 2005). Esta distribuição tem como caso particular a distribuição normal e isto faz com que o modelo proposto tenha o modelo log-normal com três parâmetros como caso particular, quando $\lambda^* = 0$. Além desta, a distribuição normal generalizada pode acomodar diferentes formas de distribuições simétricas, quando são considerados diferentes valores dos parâmetros, como pode ser visto na Figura 3.1.

Considerando o processo ξ_t com distribuição normal generalizada (Nadarajah 2005), então a função densidade de probabilidade de ξ_t é dada por:

$$f(\xi_t) = K \exp - \left| \frac{\xi_t - \mu_\xi}{\sigma_\xi} \right|^s, \quad (3.10)$$

para $-\infty < \xi_t < +\infty$, $-\infty < \mu_\xi < +\infty$, $\sigma_\xi > 0$ e $s > 0$. A constante normalizadora é dada por:

$$K = \frac{s}{2\sigma \sqrt{1/s}}. \quad (3.11)$$

No caso particular, em que $s = 2$ a equação (3.10) define a distribuição normal $N(\mu_\xi, \sigma_\xi)$. A equação (3.10) tem também como caso particular a distribuição de Laplace para $s = 1$.

A Figura 3.1 ilustra possíveis formas da função de densidade definida em (3.10), para $\mu_\xi = 0$, $\sigma_\xi = 1$ e $s = 1, 2, 5, 10$.

A estimação dos parâmetros μ_ξ, σ_ξ e s desta distribuição pode ser feita pelo método dos momentos ou pelo método da máxima verossimilhança (Nadarajah 2005).

O algoritmo para gerar ξ_t da distribuição normal generalizada:

Passo 1: Gerar $U \sim Gama(1 + 1/s, 1)$

Passo 2: Gerar $V \sim U(0, 1)$

Passo 3: Usando o método da função inversa, temos a função de distribuição acumulada

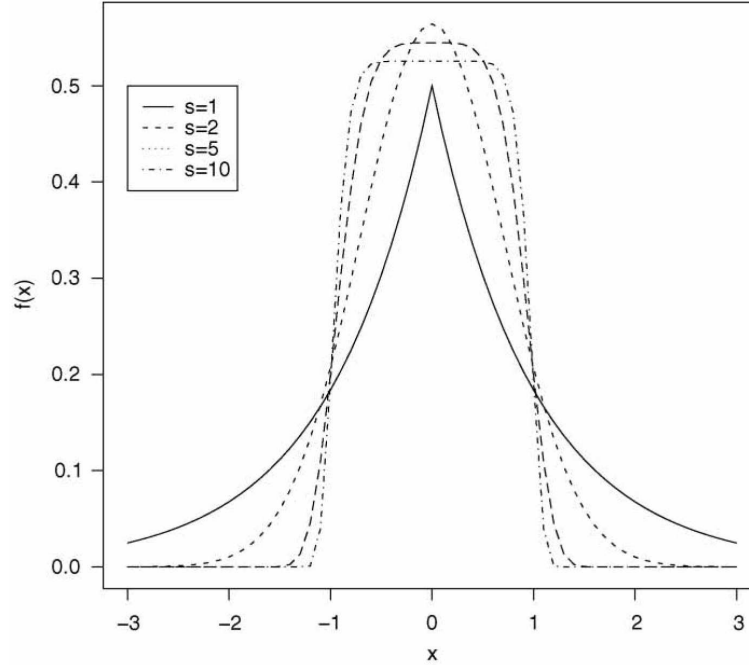


Figura 3.1: A função densidade de probabilidade normal generalizada.

$$F(\xi_t|u) = \frac{\xi_t - \mu_\xi}{2\sigma_\xi u^{1/s}} + \frac{1}{2}$$

Se $F(\xi_t|u) = V$, então

$$\frac{\xi_t - \mu_\xi}{2\sigma_\xi u^{1/s}} + \frac{1}{2} = V$$

$$\frac{\xi_t - \mu_\xi}{\sigma_\xi u^{1/s}} = 2V - 1$$

Se $V \sim U(0, 1)$ e denotando por $Y = 2V - 1$, então $Y \sim U(-1, 1)$, assim temos:

$$\xi_t = \sigma_\xi u^{1/s} Y + \mu_\xi,$$

Passo 4: Gerando-se valores de ξ_t podemos calcular o ruído a_t por:

$$a_t = \begin{cases} (\lambda \xi_t + 1)^{1/\lambda} - \tau, & \lambda \neq 0, \\ \exp(\xi_t) - \tau, & \lambda = 0, \end{cases} \quad (3.12)$$

Passo 5: A vazão a ser gerada é calculada por:

$$x_t = c + \sum_{j=1}^p \phi_j x_{t-j} + a_t \quad (3.13)$$

3.2.2 Modelo periódico log-normal com três parâmetros $BCPAR(p_m)$

O histórico de vazões mensais disponível para ajuste dos modelos consiste em uma das possíveis realizações do processo estocástico gerador da série. O procedimento para geração de séries sintéticas é realizado a partir de um modelo ajustado à série histórica e, após, sinteticamente gerar tantas séries temporais quanto se necessite. Portanto a elaboração do modelo estocástico para a série de vazões é um passo fundamental para a geração de séries sintéticas de vazões.

Neste trabalho vamos considerar, para modelar as séries sintéticas de vazões mensais, os modelos $BCPAR(p_m)$ dado por:

$$\left(\frac{x_{t(r,m)} - \mu_m}{\sigma_m} \right) = \sum_{j=1}^{p_m} \phi_{j,m} \left(\frac{x_{t(r,m)-j} - \mu_{m-j}}{\sigma_{m-j}} \right) + a_{t(r,m)} \quad (3.14)$$

Um restrição ao uso do modelos (3.14), para geração de séries de vazões mensais é que deve-se ter sempre $x_{t(r,m)} > 0$. Para atender esta restrição, os ruídos $a_{t(r,m)}$ devem satisfazer a condição:

$$a_{t(r,m)} > - \left(\frac{\mu_m}{\sigma_m} \right) - \sum_{j=1}^{p_m} \phi_{m,j} \left(\frac{x_{t(r,m)-j} - \mu_{m-j}}{\sigma_{m-j}} \right) \quad (3.15)$$

O modelo de geração de séries sintéticas deve ser aplicado diretamente à série temporal original e deve ser capaz de lidar com ruídos que apresentam um forte coeficiente de assimetria. Além disso, é necessário que a distribuição dos ruídos $a_{t(r,m)}$ seja adequadamente definida para que a restrição imposta pela equação (3.15) seja satisfeita. A solução adotada, segundo (Stedinger & Taylor 1982) e (Pereira et al. 1984) foi ajustar uma distribuição log-normal com três parâmetros para os ruídos $a_{t(r,m)}$. Portanto, vamos supor que $a_{t(r,m)}$ tem uma distribuição log-Normal com três parâmetros, tal que podemos definir um processo aleatório $\xi_{t(r,m)}$, com distribuição normal $N(\mu_m^{(\xi)}, \sigma_m^{(\xi)2})$, que está relacionado ao processo $a_{t(r,m)}$ pela equação:

$$\xi_{t(r,m)} = \ln(a_{t(r,m)} + \tau_m) \quad (3.16)$$

A equação (3.16) impõe uma nova restrição para $a_{t(r,m)}$, ou seja, $a_{t(r,m)} > -\tau_m$. Portanto, comparando-se as equações (3.15) e (3.16), tem-se uma possível definição para o parâmetro τ_m dada por:

$$\tau_m = \left(\frac{\mu_m}{\sigma_m} \right) + \sum_{j=1}^{p_m} \phi_{m,j} \left(\frac{x_{t(r,m)-j} - \mu_{m-j}}{\sigma_{m-j}} \right) \quad (3.17)$$

Os parâmetros $\mu_m^{(\xi)}$ e $\sigma_m^{(\xi)2}$ devem ser calculados de forma a preservar os dois primeiros momentos do ruído $a_{t(r,m)}$, denotados por $E(a_{t(r,m)}) = \mu_m^{(a)}$ e $Var(a_{t(r,m)}) = \sigma_m^{(a)2}$. Isto é feito usando-se as relações entre os parâmetros da distribuição normal $N(\mu_m^{(\xi)}, \sigma_m^{(\xi)2})$ e os da distribuição log-normal com três parâmetros $LN(\mu_m^{(a)}, \sigma_m^{(a)2}, \tau_m)$ dadas por:

$$\mu_m^{(\xi)} = \frac{1}{2} \ln \left(\frac{\sigma_m^{(a)2}}{\theta_m^2 - \theta_m} \right) \quad (3.18)$$

$$\sigma_m^{(\xi)2} = \ln(\theta_m) \quad (3.19)$$

em que θ_m é definido por

$$\theta_m = 1 + \frac{\sigma_m^{(a)2}}{(\mu_m^{(a)} - \tau_m)^2} \quad (3.20)$$

Portanto após a identificação da ordem (p_m) e o ajuste dos modelos BCPAR(p_m), estima-se $\sigma_m^{(a)2}$ e τ_m com as equações (2.7) e (3.17). Os parâmetros $\mu_m^{(\xi)}$ e $\sigma_m^{(\xi)2}$ são estimados com as equações (3.18) - (3.20).

Um modelo para geração de uma série sintética de vazões mensais pode ser encontrado substituindo-se na equação (3.14), o valor do ruído $a_{t(r,m)}$ retirado da equação (3.16), ou seja, $a_{t(r,m)} = \exp\{\xi_{t(r,m)}\} - \tau_m$ e a expressão de τ_m dada em (3.17). Após algumas manipulações matemáticas obtém-se uma equação final para a geração de uma série sintética no ano r para o mês m . Esta equação final é dada por:

$$x_{t(r,m)} = \sigma_m \exp(\xi_{t(r,m)}) \quad (3.21)$$

Portanto gerando-se o processo $\xi_{t(r,m)} \sim N(\mu_m^{(\xi)}, \sigma_m^{(\xi)2})$ e substituído-se na equação (3.21), pode ser gerada uma vazão $x_{t(r,m)}$. Epletindo-se este procedimento recursivamente, tem-se como resultado a geração de uma série sintética de vazões mensais. Observa-se que dependência entre as vazões $x_{t(r,m)}$ e $x_{t(r,m)-j}$ são consideradas no modelo de geração devido a presença do parâmetro τ_m introduzido no cálculo dos parâmetros $\mu_m^{(\xi)}$ e $\sigma_m^{(\xi)2}$.

3.3 Modelos não-lineares de geração de séries sintéticas

Os modelos não-lineares utilizados para previsão de vazões são também utilizados para fins de geração. No entanto, estes seguem um procedimento distinto, uma vez que estes não consideram hipótese alguma sobre a distribuição das séries de vazões. Deve-se notar que, no caso do modelo FIS, o algoritmo da maximização da verossimilhança é utilizado apenas para o ajuste de modelos de previsão, pois este considera uma distribuição normal conjunta das vazões. Para eliminar esta hipótese, o modelo FIS quando utilizado para geração de séries sintéticas é utilizado sem reajuste dos seus parâmetros, considerando apenas os centros e desvios predefinidos na etapa de inicialização, tendo como consequentes modelos do tipo *singletons* ou constantes que também são ajustados pelos mínimos quadrados. O desempenho do FIS em termos de geração considerando e omitindo esta hipótese é discutido em (Luna, Ballini, Soares & Silva Filho 2011).

Os detalhes sobre o procedimento adotado pelos modelos não lineares na geração de séries sintéticas é descrito a seguir.

Pré-processamento dos dados A assimetria na distribuição dos dados de vazões é uma característica das séries em estudo que não pode ser desprezada. Existem picos de vazão e períodos de seca que fogem do comportamento médio das vazões ao longo do histórico. Embora os modelos descritos tenham se mostrado adequados em termos de previsão um passo à frente considerando uma distribuição simétrica nos dados (como a normal, por exemplo), este desempenho não se mantém quando os mesmos modelos são utilizados para fins de geração de séries sintéticas, pois uma distribuição simétrica nos impediria de replicar os valores extremos

da série. Assim, surge a necessidade de utilizar modelos livres de qualquer pressuposto sobre a distribuição das séries para fins de geração, como é o caso do modelo FIS e da rede neural MLP.

Para tornar a série de vazão estacionária, é necessário remover o componente sazonal. Para isso realiza-se a seguinte transformação nos dados:

$$\tilde{z}_{m,t} = \frac{y_{m,t} - \hat{\mu}_m}{\hat{\sigma}_m} \quad (3.22)$$

na qual, $\hat{\mu}_m$ e $\hat{\sigma}_m$ são a média e desvio padrão estimados para cada mês m , respectivamente. Após, toma-se os valores máximos e mínimos mensais da variável $\tilde{z}_{m,t}$, obtendo-se assim valores pertencentes ao intervalo $[0, 1]$.

Identificação dos modelos Para o modelo FIS o número de atrasos é escolhido pelo critério Bayesian Information Criterium (BIC). No caso da rede MLP, o número de atrasos é determinado pela minimização do erro médio quadrático (MSE) (Morettin & Toloí 2004).

Dessa forma, para ajustar os modelos para um determinado mês i calcula-se os critérios BIC ou MSE variando o número de atrasos $p = 1, \dots, 6$. O valor de p que apresentar menor índice é escolhido para representar as informações dos padrões de vazões. O critério MSE também é empregado para a escolha do número de neurônios na camada intermediária, taxa de aprendizado e número de épocas no caso do modelo MLP.

Geração da Série Sintética De forma geral, uma série temporal $y_{m,t}$ com não estacionariedade sazonal pode ser representada por dois componentes, um sazonal $s_{m,t}$ e outro estacionário $\tilde{z}_{m,t}$:

$$y_{m,t} = s_{m,t} + \tilde{z}_{m,t} + a_{m,t} \quad (3.23)$$

sendo $a_{m,t}$ um componente aleatório com média zero e variância constante σ_a^2 , t representa o ano e m o mês.

Para a geração das vazões sintéticas mensais, são determinados modelos mensais, ou seja, doze modelos independentemente. Os modelos são ajustados com o objetivo de se estimar o componente estacionário.

O componente aleatório é selecionado a partir do conjunto de resíduos estimados ($e_{m,t}$) dado pela diferença entre a vazão observada $y_{m,t}$ do mês m e ano t e o valor estimado $\hat{y}_{m,t}$ pelo modelo mensal, sobre o conjunto de dados utilizados durante o ajuste. Formalmente,

$$\begin{bmatrix} e_{1,t} \\ e_{2,t} \\ \vdots \\ e_{12,t} \end{bmatrix} = \begin{bmatrix} \tilde{z}_{1,t} \\ \tilde{z}_{2,t} \\ \vdots \\ \tilde{z}_{12,t} \end{bmatrix} - \begin{bmatrix} \hat{y}_{1,t} \\ \hat{y}_{2,t} \\ \vdots \\ \hat{y}_{12,t} \end{bmatrix}$$

sendo que $e_{m,t}$ é a variável aleatória independente entre si, de média zero e variância constante para $m = 1 \dots, 12$. O valor de $\hat{y}_{m,t}$ pode ser estimado por um modelo não linear. Nesta dissertação, os modelos usados foram os modelos MLP e FIS (Figura 3.2). Como o componente aleatório é independente, $e_{m,t}$ é sorteado da amostragem aleatória simples sem reposição, (bootstrapping),

a partir do conjunto de valores compostos pelos erros de estimação de $y_{m,t}$ na etapa de ajuste dos modelos. Um trabalho similar na geração de séries sintéticas é proposto em (Ochoa-Rivera, García-Bartual & Andreu 2002). A seleção do componente estocástico é realizada utilizando um gerador de números pseudo-aleatórios com distribuição uniforme.

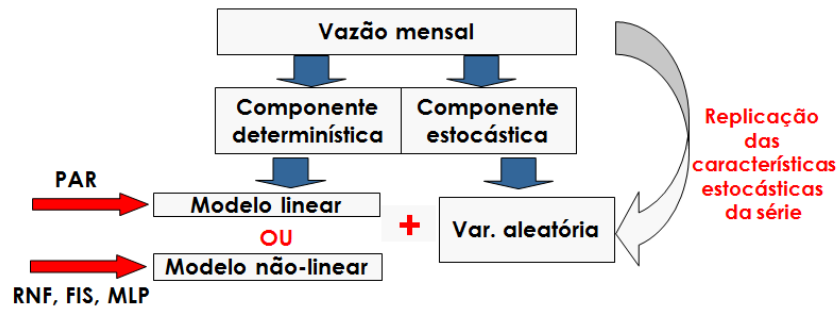


Figura 3.2: Geração de séries sintéticas

HydroPrev

4.1 Introdução

O programa HYDROLAB é um sistema de suporte à decisão para o planejamento e programação da operação de sistemas hidrotérmicos. Ou seja, um programa que possibilita ao usuário gerenciar um conjunto de modelos de otimização, previsão e simulação. Esses modelos auxiliam na solução dos problemas de planejamento das decisões de operação do sistema e da programação da operação, e estão implementados em diferentes módulos dentre os quais, os módulos HydroPrev e HydroSeries são temas desta pesquisa.

4.2 HydroPrev

O HydroPrev é um sistema de suporte a decisão para modelos de previsão de vazões. Esses modelos são ajustados com base nos históricos de vazões naturais, incrementais, agregadas e sintéticas importados da base de dados HYDRODATA. Atualmente, este módulo contempla os modelos de previsão descritos anteriormente.

- Redes neurofuzzy (RNF) (Ballini 2000)
- Modelo auto-regressivo periódico $PAR(p_m)$
- Sistema de inferência nebulosa (FIS, do termo em inglês, *Fuzzy Inference System* (Luna et al. 2007b))
- Modelo de redes neurais multicamadas (MLP) (Haykin 2001)

Vale lembrar que, a partir do banco de dados HYDRODATA, a aplicação dos modelos pode ser estendida, facilmente, para todas as usinas despachadas pelo Operador Nacional do Sistema Elétrico (ONS).

Em se tratando de estudos de previsão de vazões, existem dois procedimentos possíveis no módulo HYDROPREV. O primeiro procedimento inclui a seleção e o ajuste dos modelos de series temporais. O segundo procedimento utiliza os modelos ajustados na previsão de vazões.

Por seleção, entende-se escolha da usina a ser considerada e seleção das entradas relevantes para o modelo de previsão. Já a fase de ajuste dos modelos considera o ajuste dos parâmetros dos modelos utilizando um subconjunto de dados do histórico de vazões. O modelo ajustado deve ser validado em um segundo subconjunto de dados, diferente do primeiro, para aferir assim o desempenho do modelo. A fase de previsão permite utilizar os modelos ajustados e validados no primeiro procedimento, para realizar a estimação de vazões futuras.

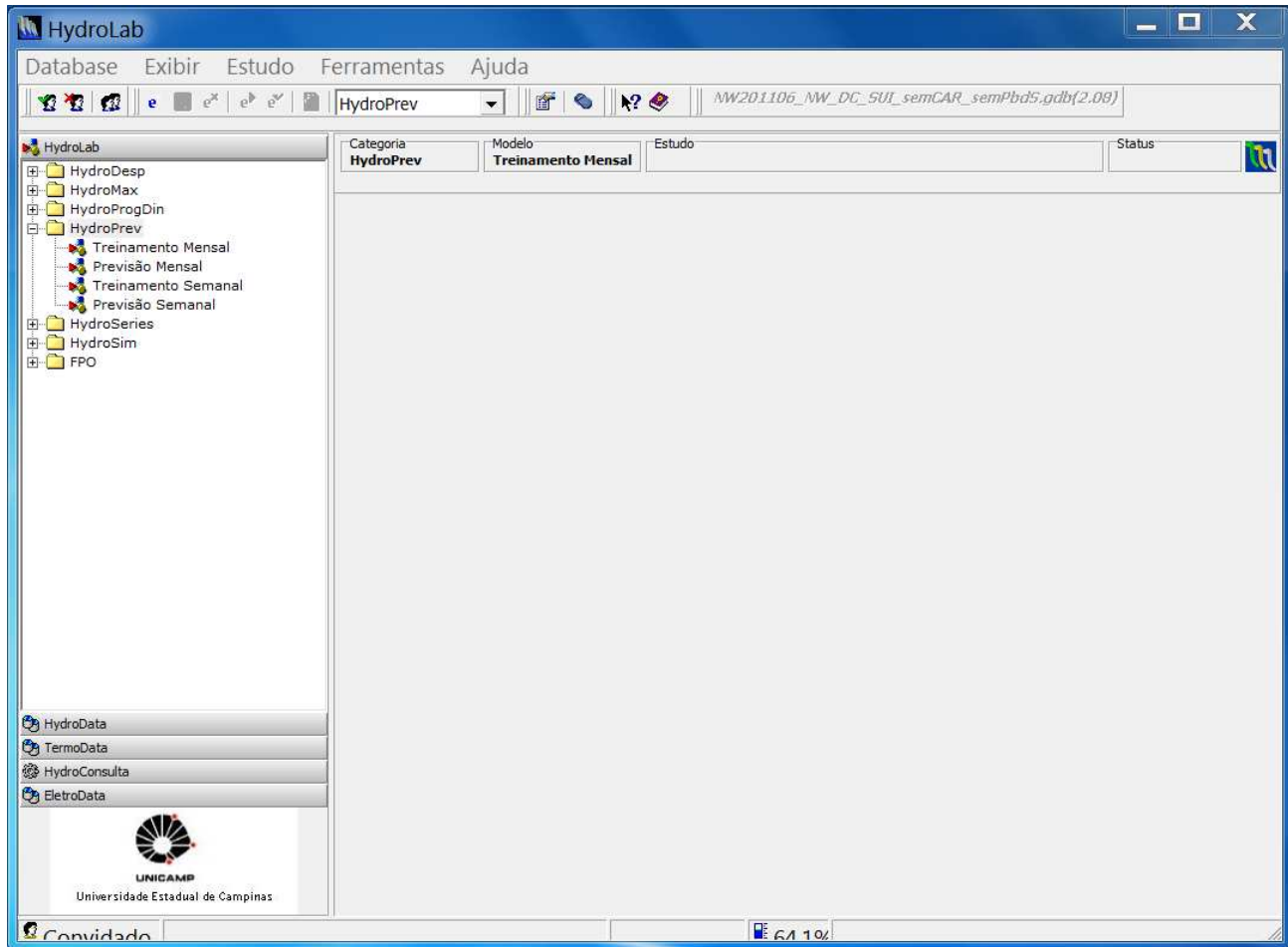


Figura 4.1: O programa HYDROLAB e seu módulo de previsões de vazões HYDROPREV.

Como pode ser visto na Figura 4.1, o módulo HYDROPREV constitui uma categoria de modelos de previsão de vazões dentro do programa HYDROLAB, subdividida em:

Treinamento Mensal: módulo para o ajuste de modelos de previsão de vazões mensais. Neste módulo o usuário pode configurar um estudo composto por doze modelos, ou seja, um modelo para cada mês do histórico de vazões.

Previsão Mensal: módulo para a realização de previsões utilizando os modelos ajustados no módulo anterior. Com esse módulo pode-se realizar previsões sobre o histórico de vazões para verificar a acurácia do modelo ajustado, ou estimar vazões futuras.

Treinamento Semanal: módulo para o ajuste de redes neurofuzzy sobre dados de vazões semanais. Neste caso, considera-se somente um modelo por série temporal.

Previsão Semanal: assim como na previsão mensal, utiliza as redes ajustadas no módulo de treinamento semanal para a realização de estimativas de vazões semanais.

Módulo Treinamento Mensal

Neste item, descreve-se a fase de treinamento ou ajuste dos modelos de vazões mensais no módulo HYDROPREV. Este módulo visa a inserção de todos os modelos anteriormente citados. A descrição a seguir utiliza como exemplo a rede neurofuzzy para fins de ilustração do sistema.

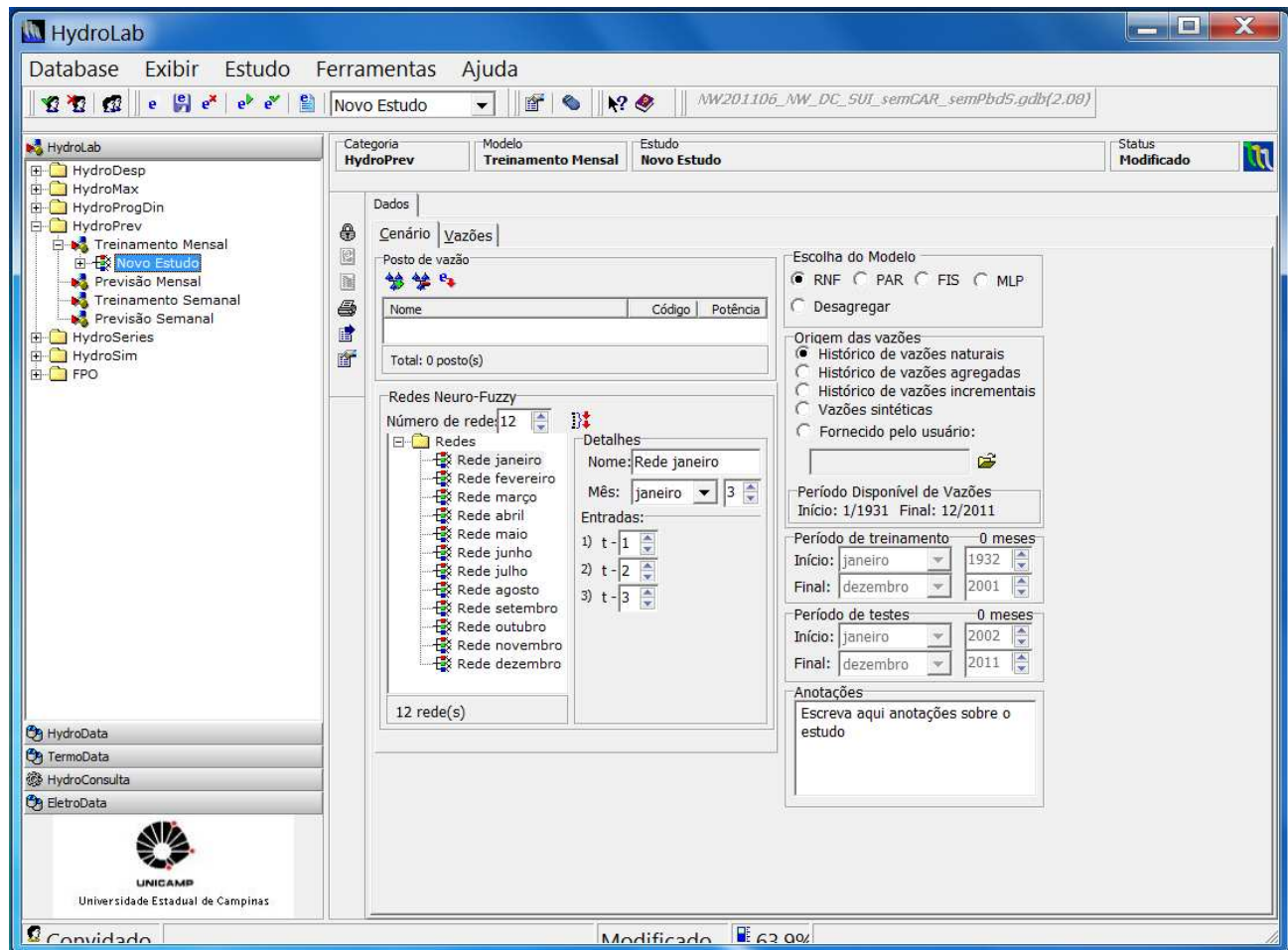


Figura 4.2: Criando um novo estudo no HYDROPREV

No quadro de estudos à esquerda da Figura 4.1 vê-se os módulos que constituem o HydroPrev. Selecionado o módulo **Treinamento Mensal**, o primeiro passo é criar um novo estudo. Este estudo é criado na função **Estudo** ou no botão **e*** que se encontra na barra de ferramentas presente na parte superior do sistema, como mostra a Figura 4.2.

Aba de Dados

No quadro de conteúdo à direita da figura, vê-se a aba **Dados** a qual contém as informações sobre o estudo de caso criado. A aba **Dados** é dividida em duas abas: **Cenário** e **Vazões**. Nesta primeira aba, as informações incluem nome, código do posto de vazão e a potência instalada

(em MW) da usina em estudo (campo **Posto de vazão**). Estas informações, assim como o histórico de vazões são acessadas via o banco de dados HYDRODATA.

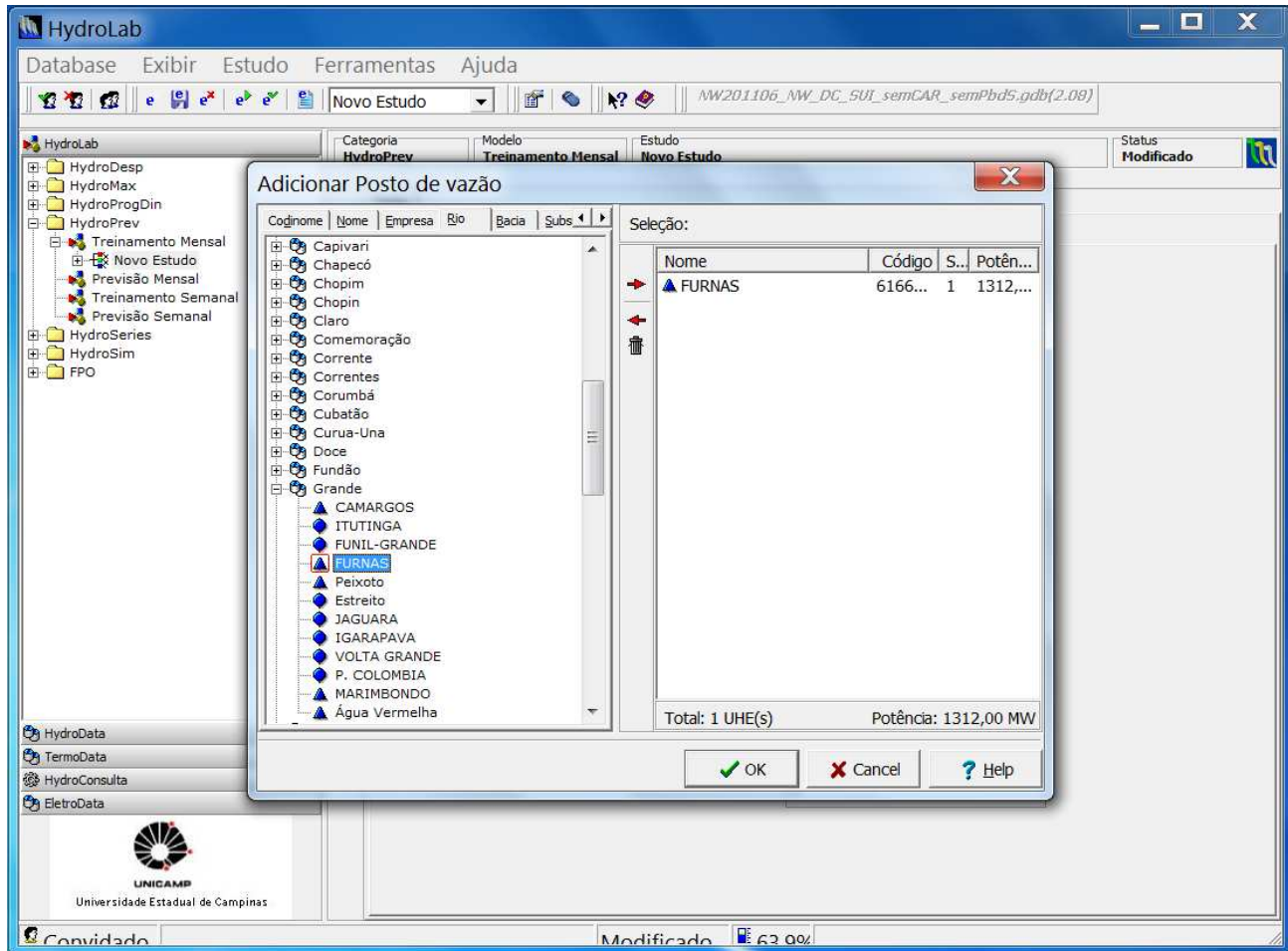


Figura 4.3: Escolha do Posto Hidrológico

Ainda na aba **Cenário**, vê-se o campo chamado Redes NeuroFuzzy. Neste campo, o usuário pode escolher o **Número de redes** a ser ajustado e também escolher e visualizar as características de cada um dos modelos ajustados para cada mês. Neste campo, tem-se a opção de escolha do número de entradas para cada uma das redes que serão ajustadas. Também tem-se a opção de replicar o número de entradas de um determinado mês para todas as demais meses.

Mais a direita da aba **Dados**, no campo **Escolha do Modelo**, o usuário seleciona o modelo que deseja usar para o treinamento das redes. Neste caso, os modelos implementados são os modelos apresentados no Capítulo 6.1. No campo abaixo, o usuário pode escolher a origem das vazões (campo **Origem das vazões**), o qual definirá o conjunto de dados para o qual o modelo será ajustado, podendo ser visualizado na aba **Vazões**. No campo **Período Disponível de Vazões**, o usuário visualizará o horizonte disponível das vazões, ou seja, em qual período se inicia e finaliza os dados disponíveis para o posto escolhido. Logo, os subconjuntos de treinamento (campo **Período de treinamento**) e de teste (campo **Período de teste**) podem ser especificados. Para cada um dos subconjuntos o usuário pode definir o início e o fim (no caso o mês e o ano).

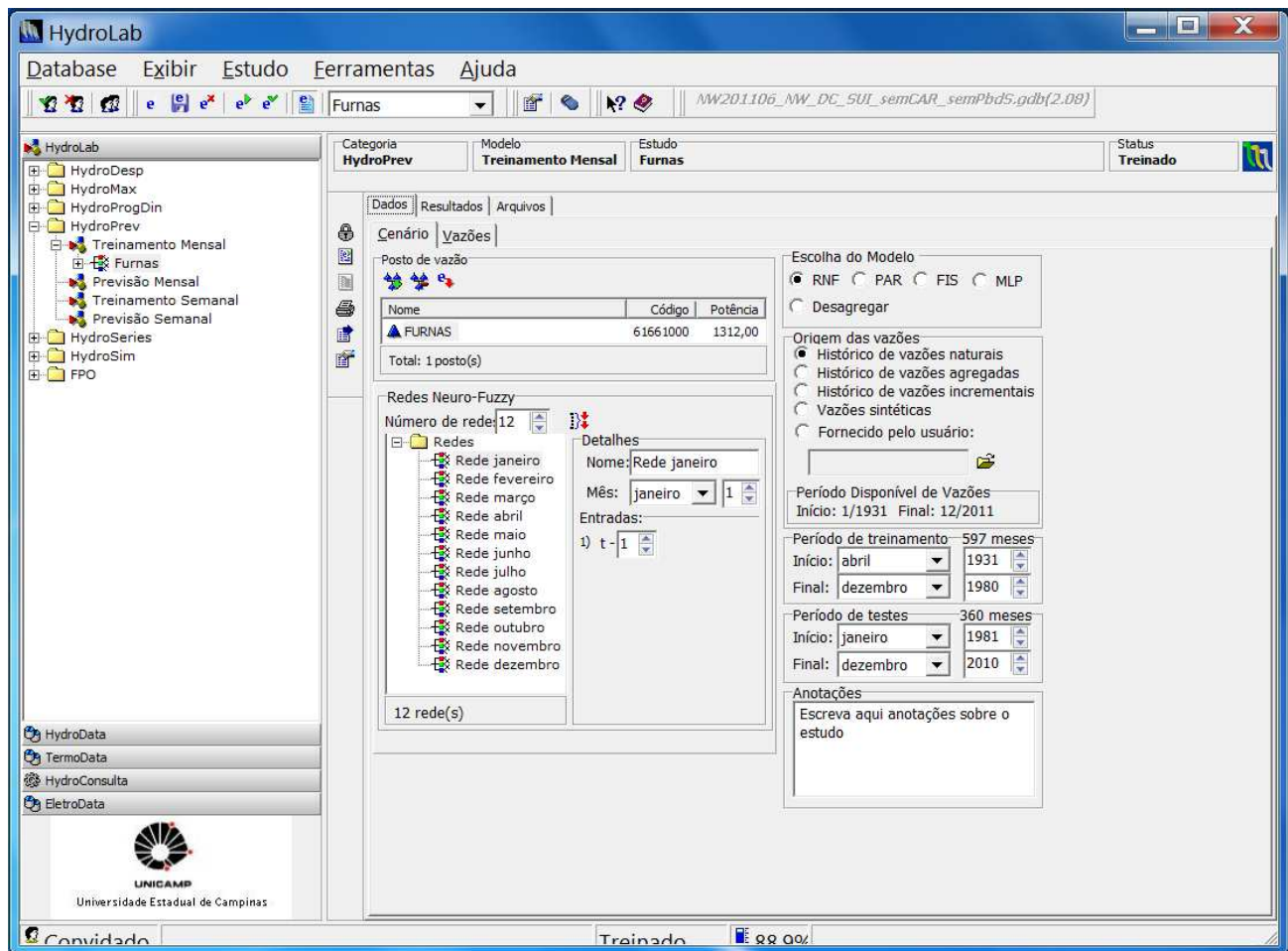


Figura 4.4: Cenário ajustado para o treinamento das redes

O próximo passo é selecionar o posto hidrológico considerado para o estudo. Como exemplo de utilização do módulo de previsão de vazões do sistema HYDROLAB, escolheu-se o posto de vazão de Furnas (Figura 4.3), com período de treinamento de abril de 1931 a dezembro de 1980 e período de teste de janeiro de 1981 a dezembro de 2010. Para o posto selecionado, deve-se escolher o número de entradas para cada rede no campo **Redes Neurofuzzy** na aba **Cenário** (Figura 4.4). Definida a configuração de cada rede, deve-se salvar o estudo na função **Estudo**.

A aba **Vazões** que se encontra na subdivisão da aba **Dados**, permite que o usuário visualize informações à respeito da vazão do posto escolhido (Figura 4.5). O primeiro campo que pode-se observar nesta aba é **Origem das Vazões**, em seguida tem-se duas abas: a aba **Médias Mensais** apresenta a média histórica mensal, enquanto a aba **Histórico** (Figura 4.6) ilustra os históricos de vazão do posto.

A seguir, executa-se o estudo (Figura 4.7). Durante a execução, vê-se a aba **Progresso do Treinamento** a qual contém duas subdivisões: **Progresso** e **Relatório**. Na aba **Progresso**, tem-se basicamente as informações sobre o tempo de execução para o ajuste das redes. Na aba **Relatório**, tem-se um resumo sobre o processo de ajustes dos modelos, com as seguintes informações: número de entrada e regras ajustadas para cada rede.

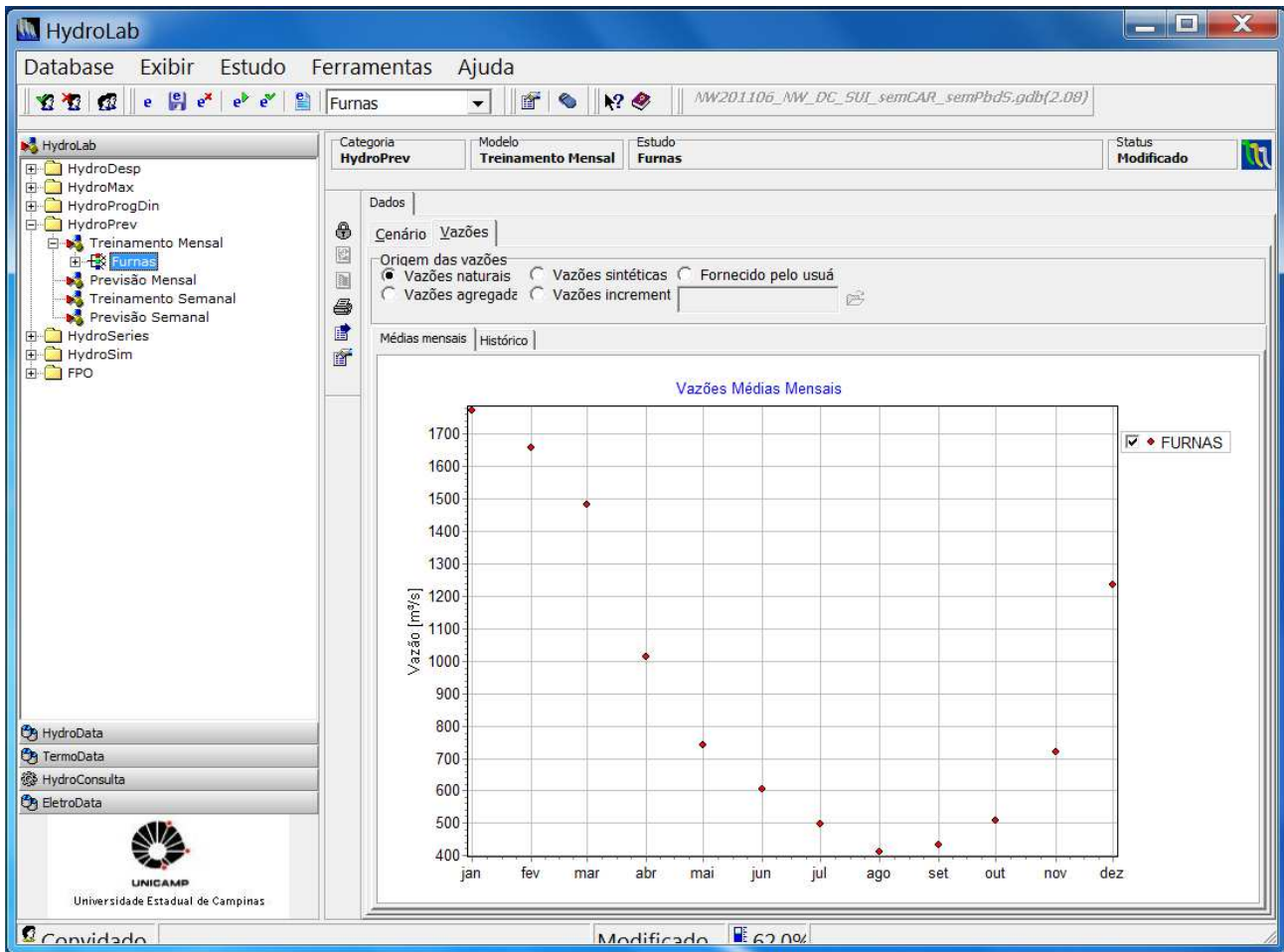


Figura 4.5: Médias mensais das vazões

Aba de Resultados

Após o ajuste das redes, uma nova aba é criada. Esta aba, chamada de **Resultados** subdivide-se em quatro abas: **Geral**, **Rede**, **Passo** e **Desvio Padrão** (Figura 4.7). A aba **Geral** contém os resultados médios mensais de todo o período de treinamento e de todo período de teste, com os respectivos erros globais para os dois períodos.

Na Figura 4.8, apresenta-se uma visão geral do arranjo gráfico dos resultados do estudo para o posto de vazões de Furnas (aba **Geral**). Nota-se que as trajetórias de vazão possuem estatísticas associadas aos erros de treinamento e de testes das redes ajustadas.

Além disso, a aba **Geral** tem uma subdivisão, chamada de **Trajectoria**, na qual pode-se visualizar a trajetória de erros de previsão, tanto no conjunto de dados de ajuste ou treinamento, como no conjunto de teste (Figura 4.8).

A apresentação dos resultados é auxiliada por cores, uma vez que a quantidade de informações é relativamente grande para uma adequada visualização. Definiu-se a cor azul para representar os resultados associados ao período de teste. A diferença de cores também é utilizada na análise dos erros de treinamento e testes das redes. Para os resultados de treinamento e teste, apresentam-se os valores observados e os estimados.

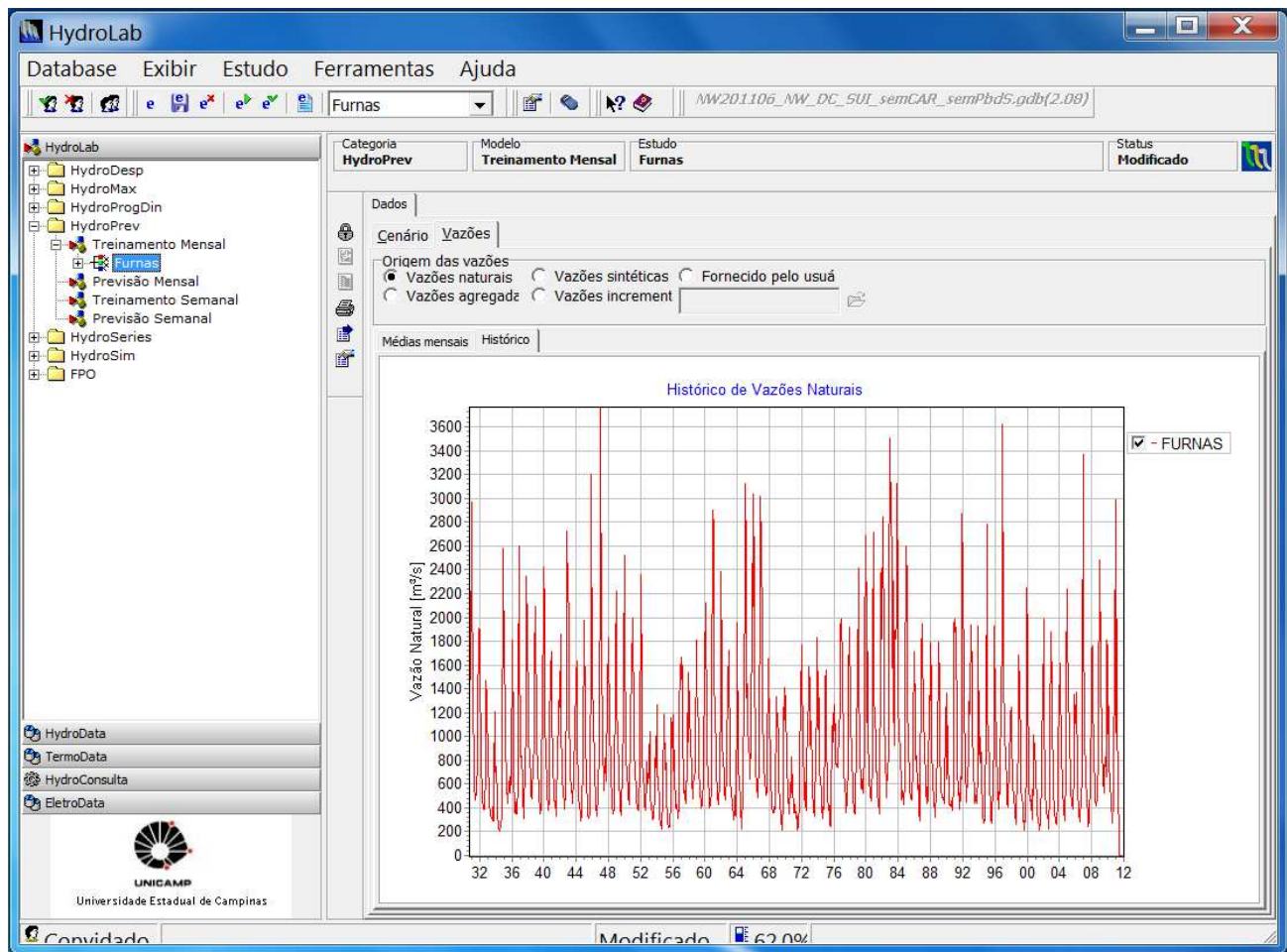


Figura 4.6: Histórico das vazões

Dentre os recursos de análise dos resultados, destaca-se aquele que permite analisar as estatísticas individuais de cada modelo mensal ajustado. Esses resultados podem ser visualizados na aba **Rede**. Esse recurso permite ao usuário identificar, rapidamente, quais os modelos que precisam de reajuste da sua estrutura, em função dos erros de previsão encontrados nos conjuntos de treinamento e teste.

Por exemplo, na Figura 4.10, pode-se visualizar os resultados obtidos para a rede treinada para a previsão das vazões do mês de Julho. Note que, assim como no caso da aba **Geral**, a aba **Rede** também apresenta os erros de treinamento e de teste, além da trajetória de erros na aba **Trajatória**.

Na aba **Passo** o usuário pode comparar a evolução dos Erros Percentuais Absolutos médios (EPA), considerando um horizonte de previsão de um até doze passos à frente. Na Figura 4.11 pode-se ver, por exemplo os EPA médios para o mês de Julho. Já na aba **Desvio Padrão** pode-se visualizar a Raiz do Erro Quadrático Médio (REQM) mensal do período de teste sendo comparado ao desvio padrão mensal observado no histórico (Figura 4.12). Esta comparação permite verificar o poder preditivo do modelo tendo como referência o REQM obtido caso a média histórica (MLT) seja utilizada como estimativa da vazão. Observa-se ainda em **Desvio Padrão**, duas formas de visualização da REQM, **Passos atrás** e **Passos à frente**. A ilustração

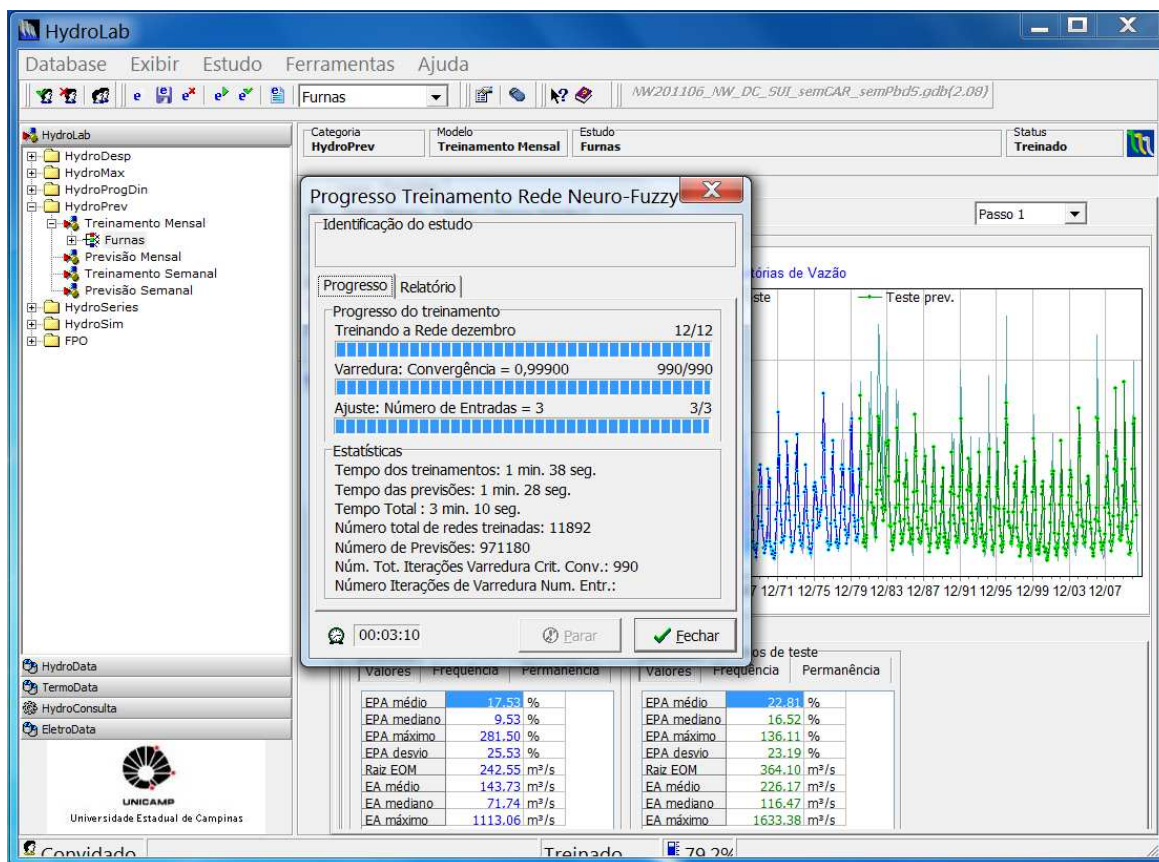


Figura 4.7: Execução do treinamento

Passos atrás exibe as previsões de um determinado mês cujo os dados de entrada são de um até dose meses passados. Enquanto **Passos à frente** exibe a REQM das previsões futuras cujo mês como dado de entrada seja o mês selecionado. Na Figura 4.13, por exemplo, observa-se que dado o mês de julho, tem-se a REQM da previsão dos meses consecutivos.

Além da visualização dos resultados na forma gráfica, o usuário pode usar a opção **Visualizar arquivos** definida na função **Estudos**. Se o usuário escolhe esta forma de visualização dos resultados, cria-se uma nova aba chamada **Arquivos** que contém quatro subdivisões. Cada uma dessas subdivisões tem o nome do estudo com extensão distinta segundo as informações contidas em cada aba (Figura 4.14).

Por exemplo, a aba Furnas.usi contém as informações gerais sobre o posto escolhido e Furnas.vaz, contém os dados de vazões. Já a aba Furnas.rhrp, contém os resultados dos ajustes dos modelos após a execução do módulo de treinamento, com a configuração final dos modelos, assim como as saídas da rede para os períodos de treinamento e de teste.

Recursos Avançados Alguns recursos avançados foram criados para auxiliar o usuário tanto na fase de ajuste dos modelos como na fase de visualização dos resultados. Esses recursos estão definidos entre as abas esquerda e direita no módulo HYDROLAB. Na aba **Resultados**, subdivisão **Geral** (Figura 4.8), foi verificado que os gráficos de resultados não oferecem elementos suficientes para a análise de detalhes das trajetórias, devido à extensão dos horizontes

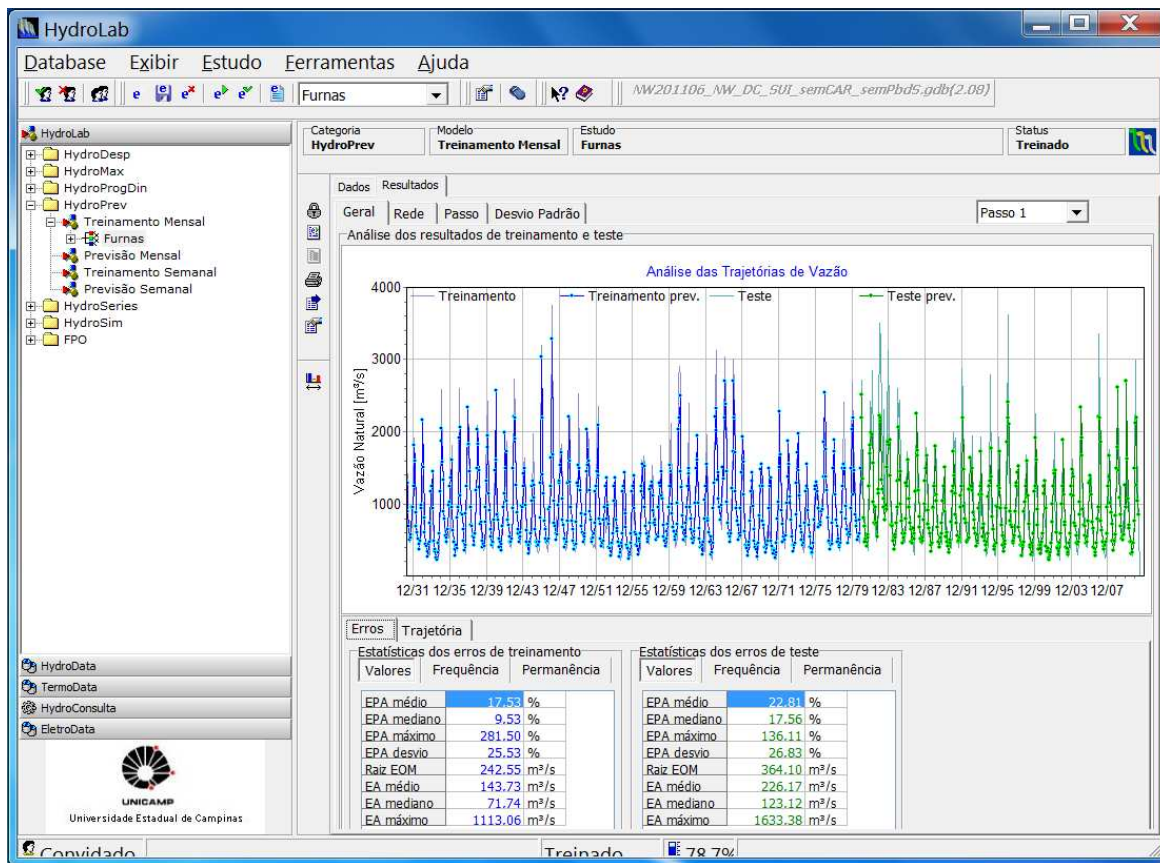


Figura 4.8: Resultado das previsões para treinamento e teste

encontrados em um estudo de previsão.

Para resolver esse problema uma função de *zoom* das escalas temporais foi implementada. Usando esta função, pode-se escolher a visualização dos dados em um determinado período definido previamente pelo usuário, ou então, a visualização pode ser feita por página. Neste caso, a visualização é dentro de um determinado ano previamente selecionado. Também nesta função, o usuário tem a opção de visualizar a média de longo termo (**Vazões médias**), como mostra a Figura 4.15.

Outro recurso avançado é a função *Propriedades*. Selecionado este recurso, vê-se a aba **Propriedades Treinamento** com duas subdivisões: **Geral** e **Redes** (Figura 4.16).

Na aba **Redes**, o usuário pode definir o limiar de convergência do algoritmo para cada rede. Feita as escolhas, o usuário deve executar o programa para então visualizar os resultados com a configuração escolhida.

Na aba **Geral**, tem-se a opção de escolha da função de normalização dos dados de entrada na rede **Transformações de vazões**, a escolha da previsão k passos à frente e a opção de varredura do critério de convergência que determina se novas regras serão incluídas no modelo, assim como a varredura do número de entradas para cada rede. A varredura é um procedimento de busca exaustiva que determina os parâmetros que melhor se adequam ao critério de desempenho selecionado.

Por exemplo, se for selecionada a opção de varredura, o usuário poderá escolher o intervalo

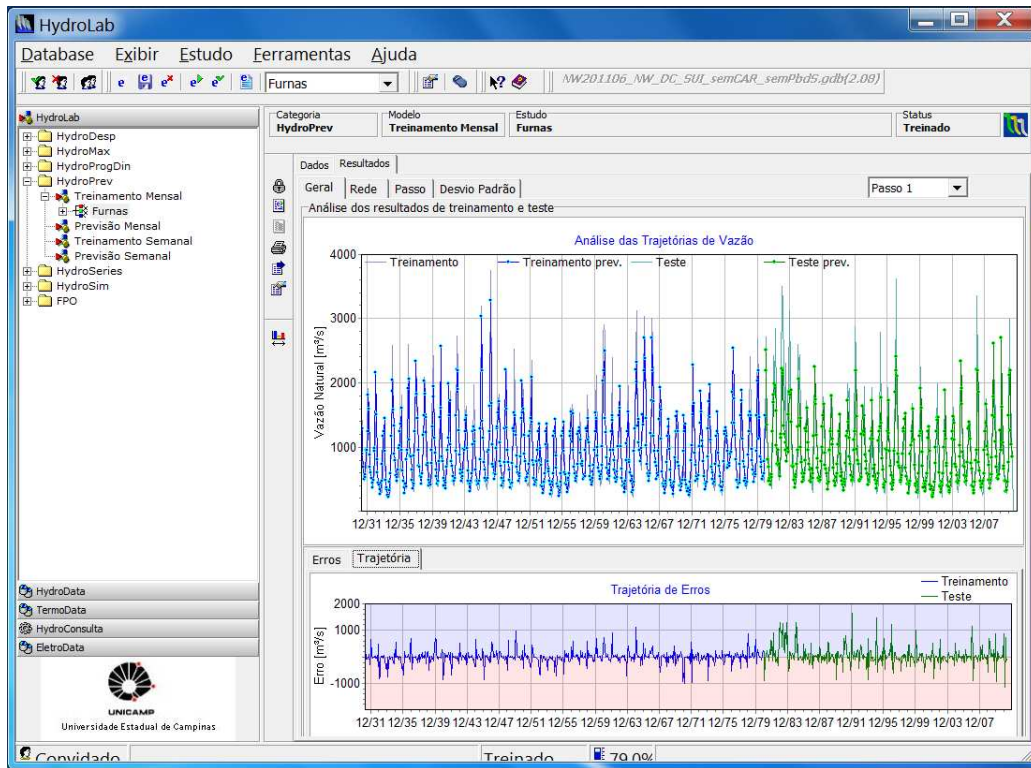


Figura 4.9: Trajetória dos erros de previsões

de varredura do parâmetro de convergência -hiperparâmetro da rede neurofuzzy necessário para o processo de otimização- e a discretização do intervalo. Feita esta opção, o usuário deverá treinar os modelos novamente, e no final o sistema retornará as saídas da rede para o valor do parâmetro que apresentar o melhor desempenho. Neste caso, durante a execução do estudo, na aba **Progresso Treinamento** e, na subdivisão **Progresso**, vê-se tanto as informações sobre o tempo de execução para o ajuste das redes como o tempo de execução para a escolha do parâmetro.

Módulo Previsão Mensal

O módulo Previsão Mensal permite ao usuário verificar o desempenho dos modelos ajustados no módulo Treinamento Mensal. Neste caso, o usuário tem a opção de verificar o desempenho do modelo nos mesmos períodos em que o modelo foi ajustado ou em um período diferente. No módulo Previsão Mensal o usuário também tem a opção de fazer previsões futuras a partir dos modelos ajustados.

Assim como no módulo Treinamento Mensal, o primeiro passo no módulo Previsão Mensal é criar um novo estudo. Na Figura 4.18, no quadro de conteúdo à direita, vê-se aba **Dados**. Nesta aba, o usuário deve adicionar o posto de vazão que será considerado (**Posto de Vazão**). Como este módulo utiliza os modelos já ajustados no módulo Treinamento Mensal, o usuário deve adicionar o estudo de ajuste para o posto em questão. Ainda nesta aba, o usuário pode escolher a origem das vazões e o período de previsão. Continuando o exemplo anterior, será feita a previsão de vazões mensais um passo à frente para o posto de Furnas. Como pode ser

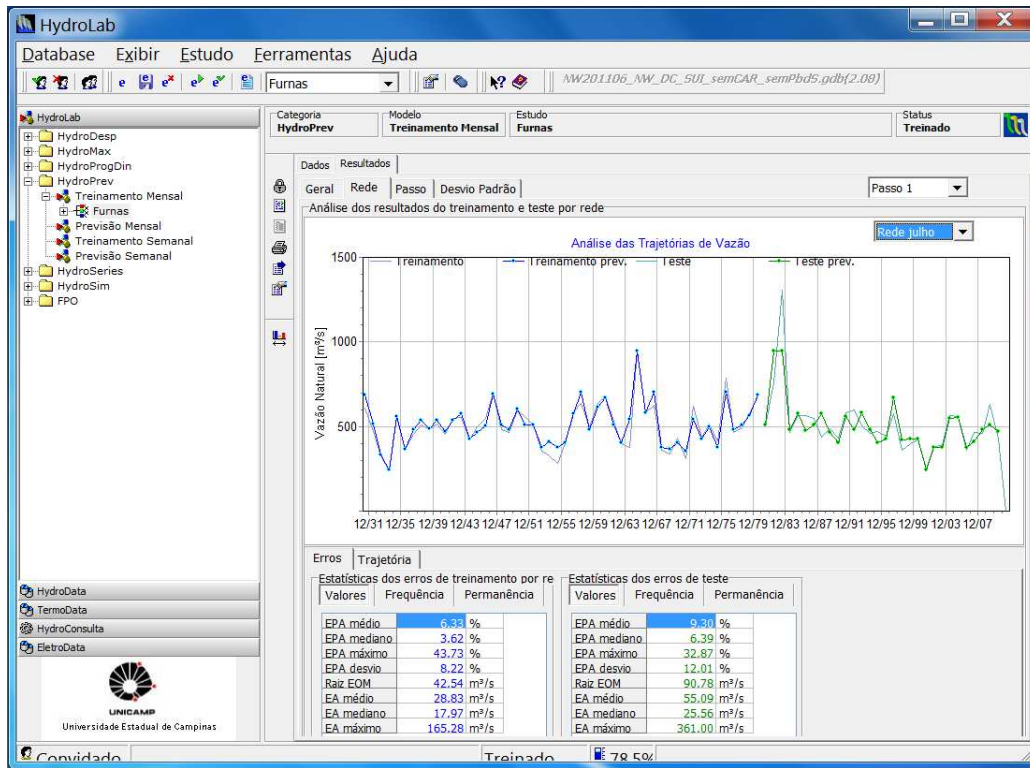


Figura 4.10: Resultados vistos por rede

verificado na Figura 4.19, foi adicionado o nome do posto e o nome do estudo de treinamento ajustado anteriormente. O período de previsão foi de 2001 a 2015. Como o histórico utilizado tem dados de vazões até dezembro de 2010, a partir deste mês e ano, e feita a previsão de valores futuros até dezembro de 2015.

A seguir, deve-se executar o estudo. Vale lembrar que neste módulo o modelo já foi ajustado e, portanto, a fase de execução equivale a fazer unicamente as previsões, sem o ajuste dos parâmetros ou seleção de entradas no modelo.

Durante a execução do estudo, vê-se a aba **Progresso Previsão Rede** a qual contém duas subdivisões: **Progresso** e **Relatório**. Na aba **Progresso** observam-se basicamente as informações referentes à execução para o ajuste das redes. Na aba **Relatório**, mostra-se um resumo sobre os resultados de previsão para os diferentes horizontes considerados e durante o período escolhido.

Após realizada a previsão, uma nova aba é criada. Esta aba é chamada de **Resultados** e subdivide-se em duas abas: **Geral** e **Rede** (Figura 4.20). A aba **Geral** contém os resultados mensais para todo o período de previsão. Já a aba **Rede** mostra as previsões para cada mês.

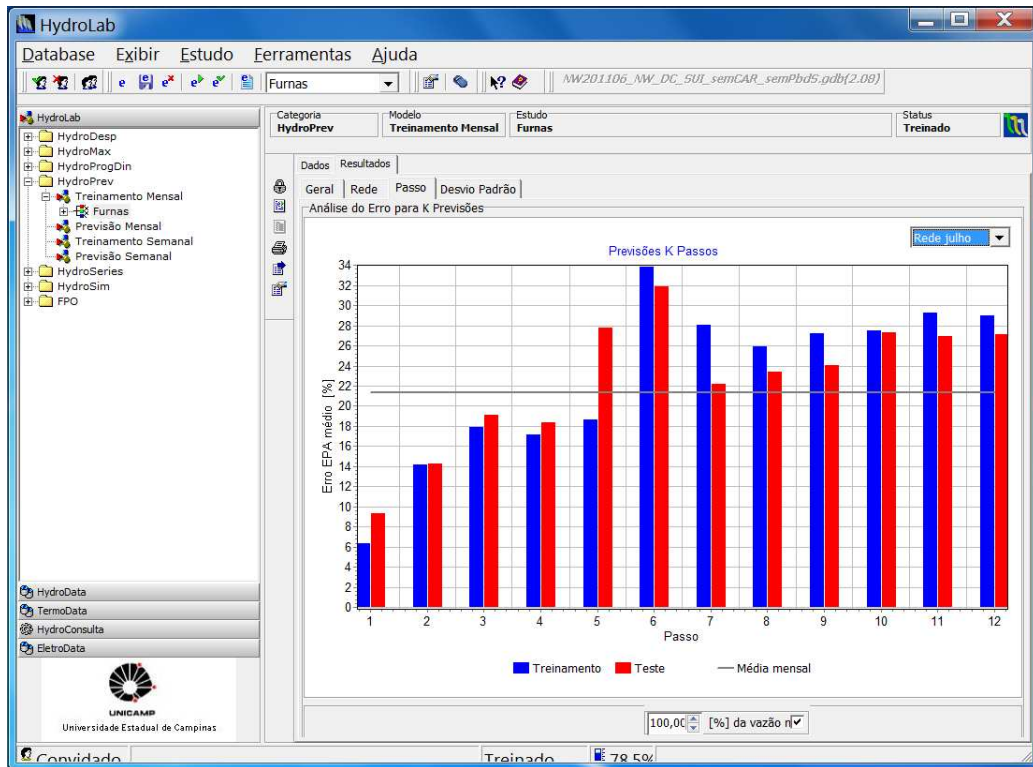


Figura 4.11: Erros percentuais absolutos médio de cada passo

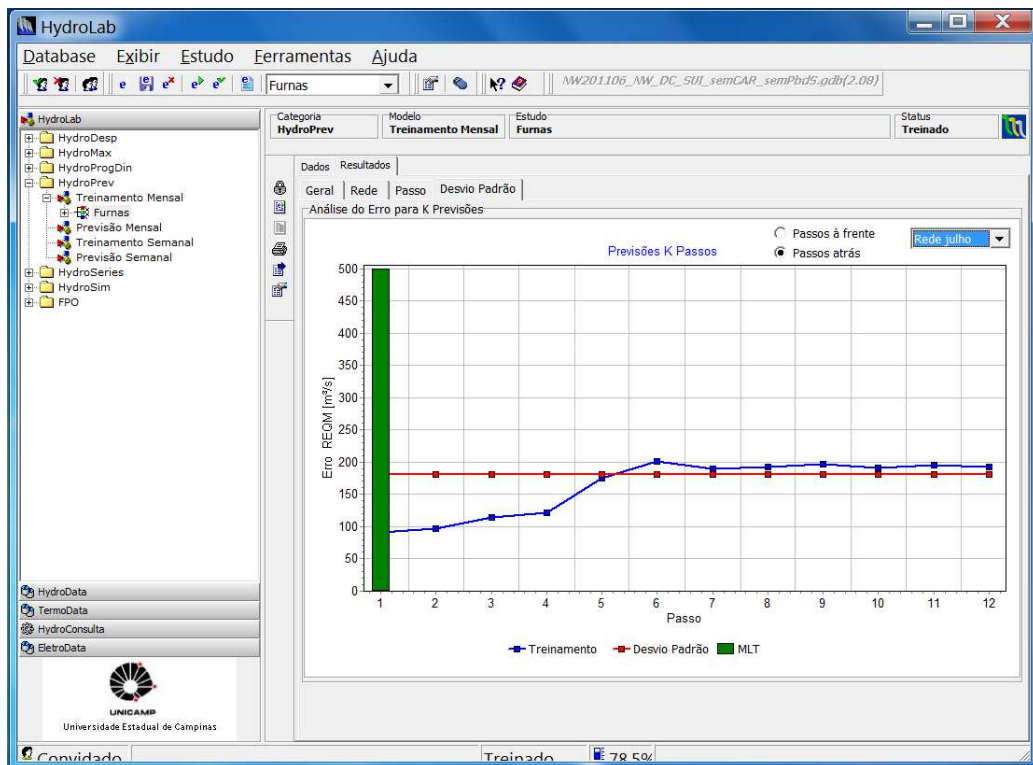


Figura 4.12: Desvio padrão do resultado previsto

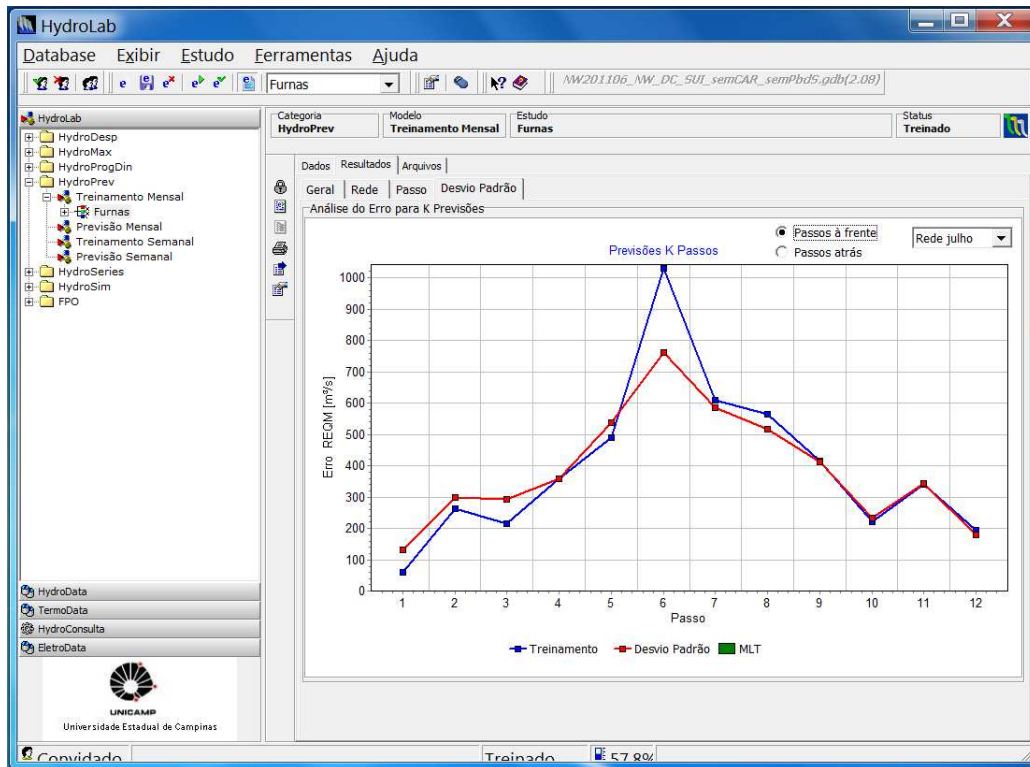


Figura 4.13: Desvio padrão - Passos à frente

```

1 // ARQUIVO DE CENÁRIO
2 // HydroPrev - Treinamento de Redes Neuro-Fuzzy
3
4 // Versão do tipo de dados contido neste arquivo.
5 [VERS]
6 1.5
7
8
9 // Dados de identificação do Estudo.
10 // Nome do estudo, nome do proprietário, data criação, status e bloqueado.
11 [ESTU]
12 Furnas
13 Convidado
14 08/07/2013 10:53:44
15 3
16 1
17
18
19 // Modelo escolhido.
20 [MODE]
21 0
22 Modelo RNF
23
24
25
26
27 // Identificação do posto de vazões da usina selecionada.
28 // Código e Nome do posto de vazões.
29 [PSTO]
30 61661000 FURNAS
    
```

Figura 4.14: Visualização dos resultados na forma de arquivo

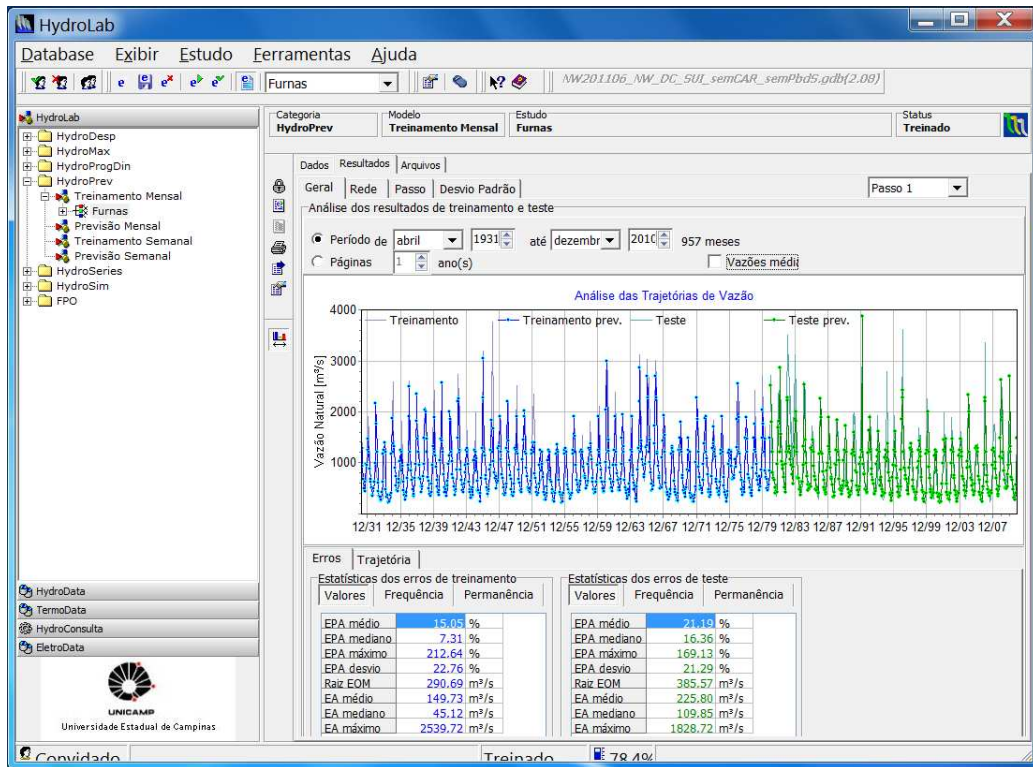


Figura 4.15: Análise dos resultados com o auxílio das escalas horizontais

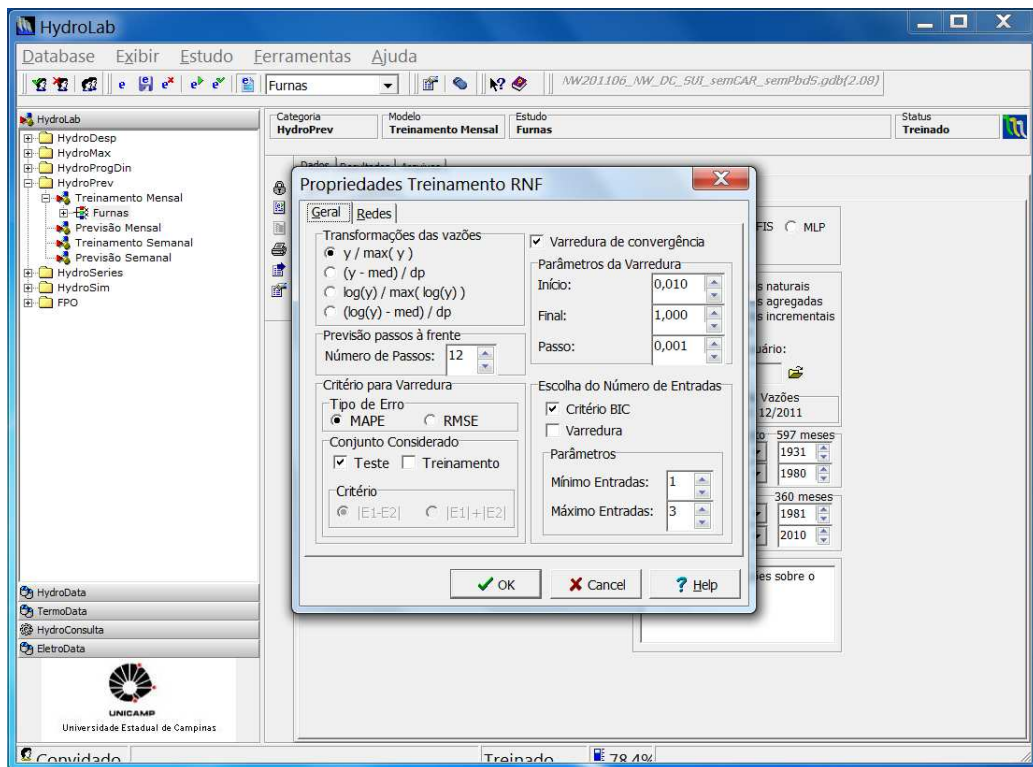


Figura 4.16: Recursos avançados - Propriedades aba Geral

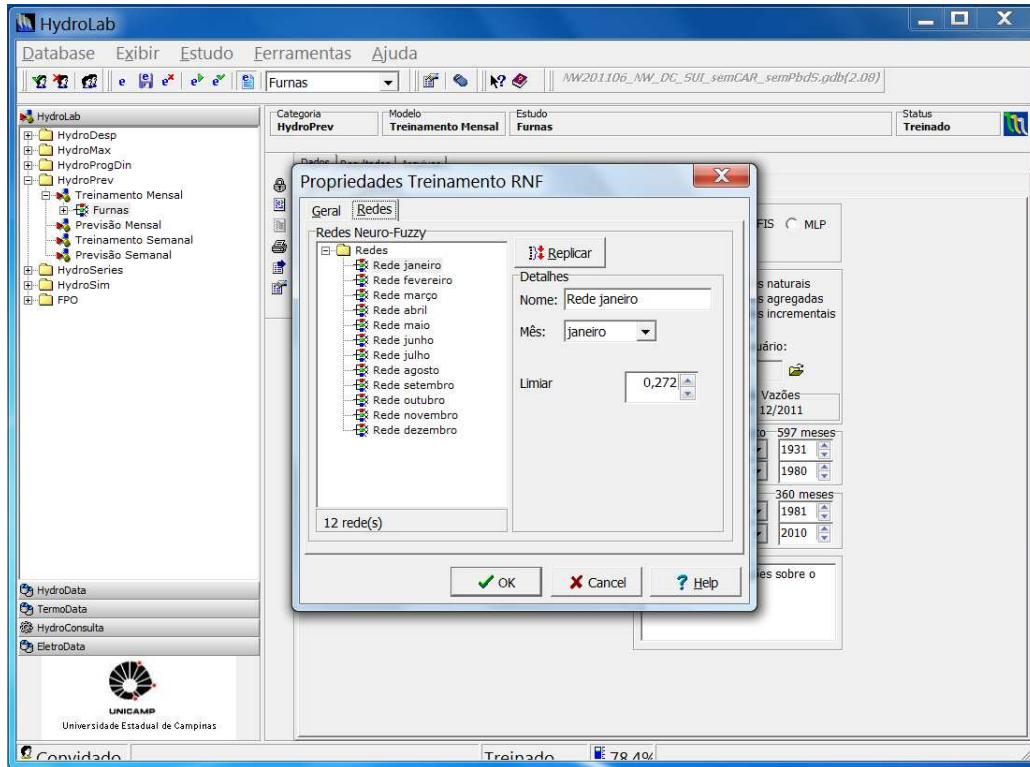


Figura 4.17: Recursos avançados - Propriedades aba Rede

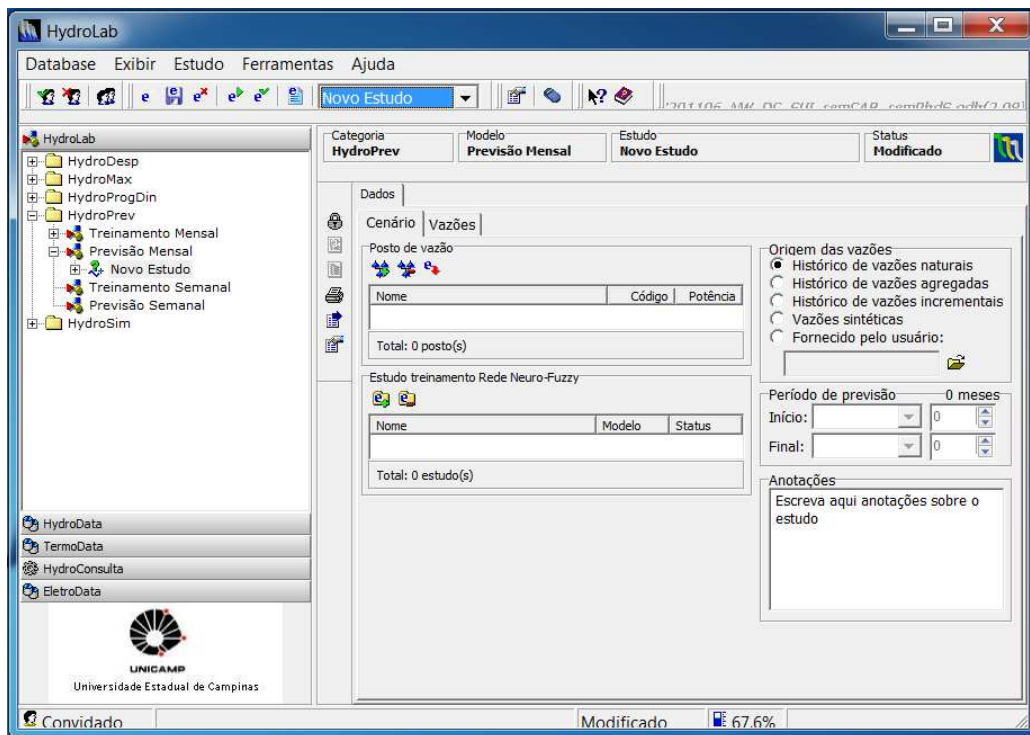


Figura 4.18: Módulo de Previsão Mensal

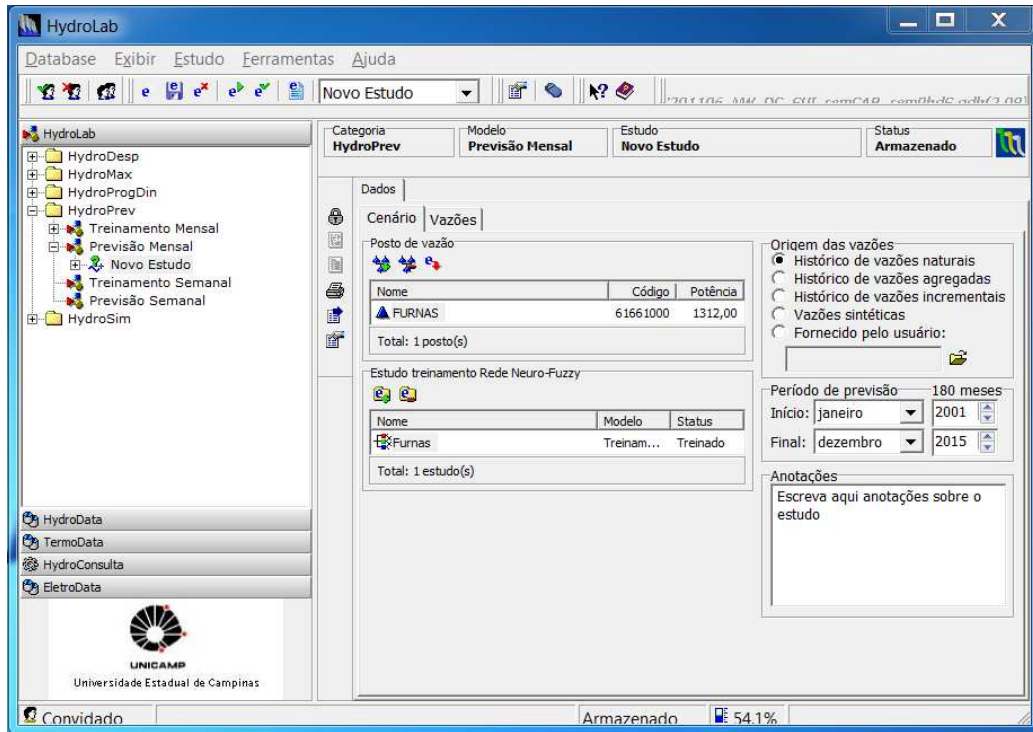


Figura 4.19: Módulo de Previsão Mensal

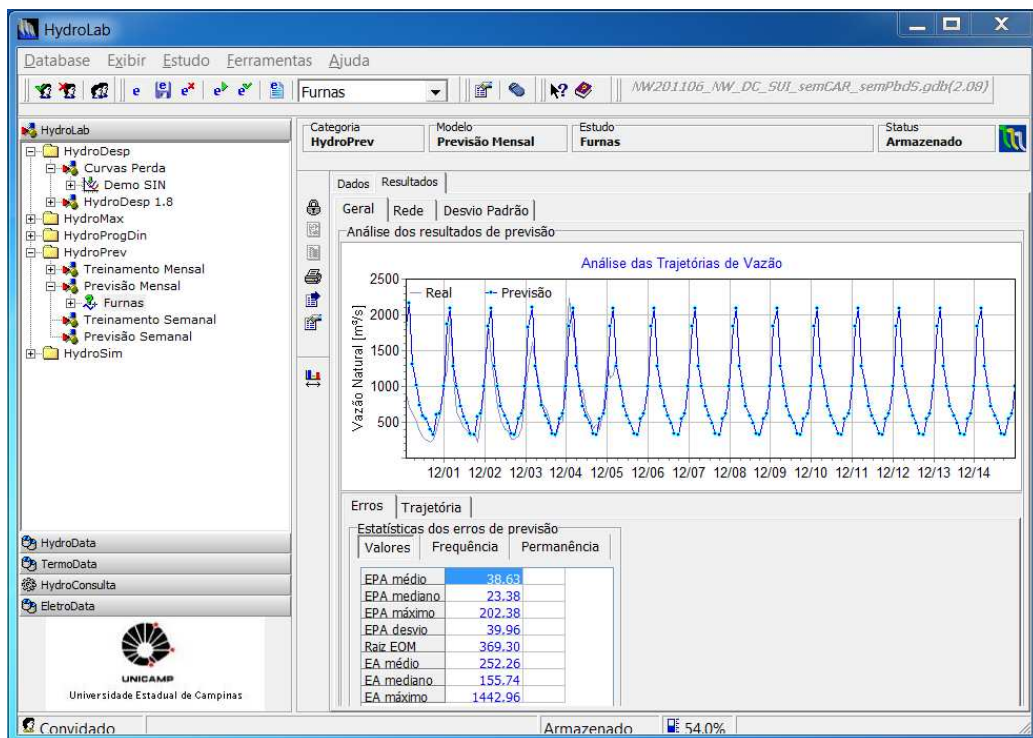


Figura 4.20: Vazões prevista pelo módulo Previsão Mensal

HydroSeries

5.1 Introdução

O HYDROSERIES é um sistema de geração de séries sintéticas que foi todo desenvolvido no decorrer deste trabalho e foi inserido no programa HYDROLAB como uma ferramenta que permite gerar séries temporais que poderão re-alimentar outros sistemas que também estão no HYDROLAB. Neste módulo, encontram-se implementados quatro modelos de geração de séries sintéticas denominados: FIS, MLP, PAR e RNF, anteriormente descritos.

Assim como para os outros módulos do HYDROLAB, a aplicação dos modelos facilmente estende-se para todas as usinas despachadas pelo Operador Nacional do Sistema (ONS). Ou seja, tem-se a possibilidade de gerar séries sintéticas para todas as usinas do SIN.

A seção seguinte descreve o funcionamento do módulo HYDROSERIES.

5.2 HydroSeries

A execução dos estudos é realizada a partir da criação de um novo estudo e a escolha da usina e ajustes do modelo escolhido. A tela principal que permite a geração do estudo é ilustrada na Figura 5.1, onde a partir do módulo Séries Sintéticas, obtemos a interface mostrada na Figura 5.2 Toda a parte de criação de um estudo acontece na aba **Dados** e os resultados são analisados na aba **Resultados**.

Como pode ser visto, este estudo é criado na função **Estudo** ou no botão **e***, como mostra a Figura 5.2.

Aba de Dados

A aba **Dados** é subdividida em duas abas, **Cenário** e **Vazões**, como pode ser visto na Figura 5.2. Todos os ajustes de configuração necessários para a construção de um estudo são realizados na aba **Cenário**.

Criado o estudo, inicialmente seleciona-se o posto de vazão desejado no campo **Posto de vazão**. Neste campo além de visualizar o posto selecionado, também mostram-se as informações do código e potência da usina. O campo **Origem das vazões** permite que o usuário escolha

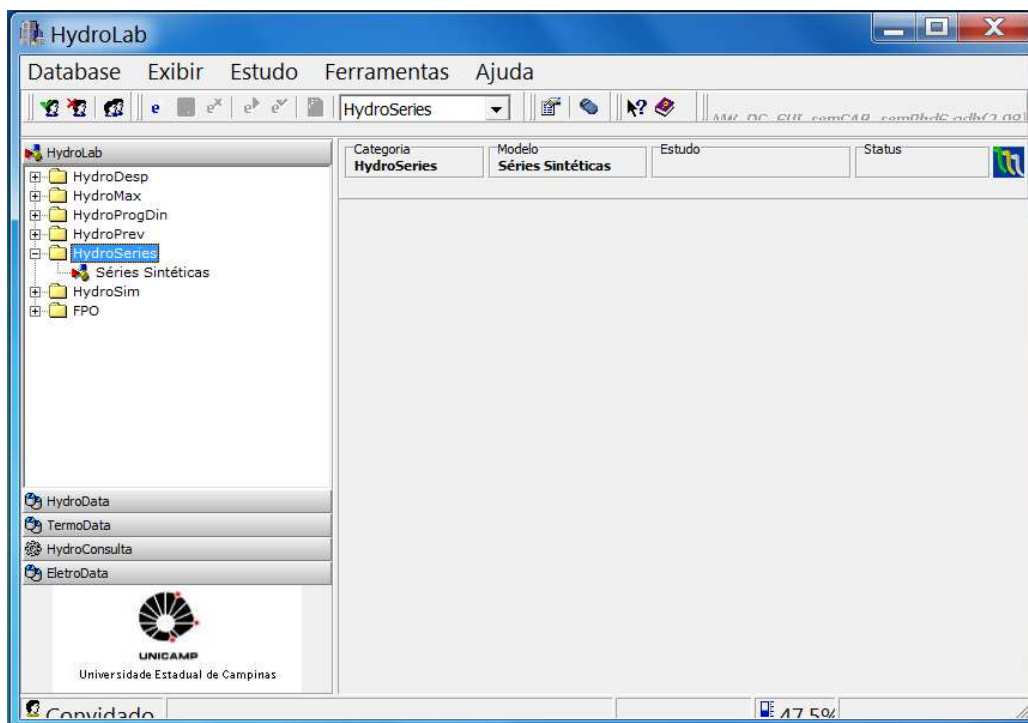


Figura 5.1: Módulo HYDROSERIES

um dos três históricos de vazões: naturais, agregadas e incrementais. Selecionado o histórico, o usuário pode visualizar a série temporal na aba **Vazões**, como ilustrado na Figura 5.4. Também é possível selecionar períodos específicos do histórico e ainda as médias mensais. Voltando na aba para a configuração do estudo, o campo **Modelo** permite a seleção de um dos modelos listados para a geração da série, assim como o período do histórico de vazões a ser utilizado na etapa de ajuste. No campo **Período de Geração** é necessário informar o número de anos da série que deseja-se gerar.

No caso do modelo escolhido ser a RNF, um novo campo aparece logo abaixo (Figura 5.5), onde é possível selecionar um estudo previamente ajustado via o módulo de treinamento mensal para o posto em questão, o qual será aproveitado para fins de geração de séries sintéticas. Ainda, visa-se a utilização de estudos de previsão previamente gerados no módulo de treinamento mensal para o caso dos modelos PAR, FIS e MLP. Em seguida, finalizado os ajustes dos parâmetros dos modelos, executa-se o estudo. Quando o processo de geração da série é encerrado, a aba **Resultados** é criada.

Aba de Resultados

A aba de resultados é subdividida em duas abas: **Geral** e **Estatísticas**. A aba **Geral** (Figura 5.6) exhibe a série que foi gerada enquanto a aba **Estatísticas** (Figura 5.7) traz os critérios utilizados para validação da série em base mensal. Esses critérios são medidas de probabilidade e estatísticas que são: média, desvio padrão, coeficiente de assimetria e curtose. A média é a soma de todos os valores observados dividido pelo número de observações. O desvio padrão mede a dispersão dos valores individuais em torno da média. Assim, o desvio padrão é a média

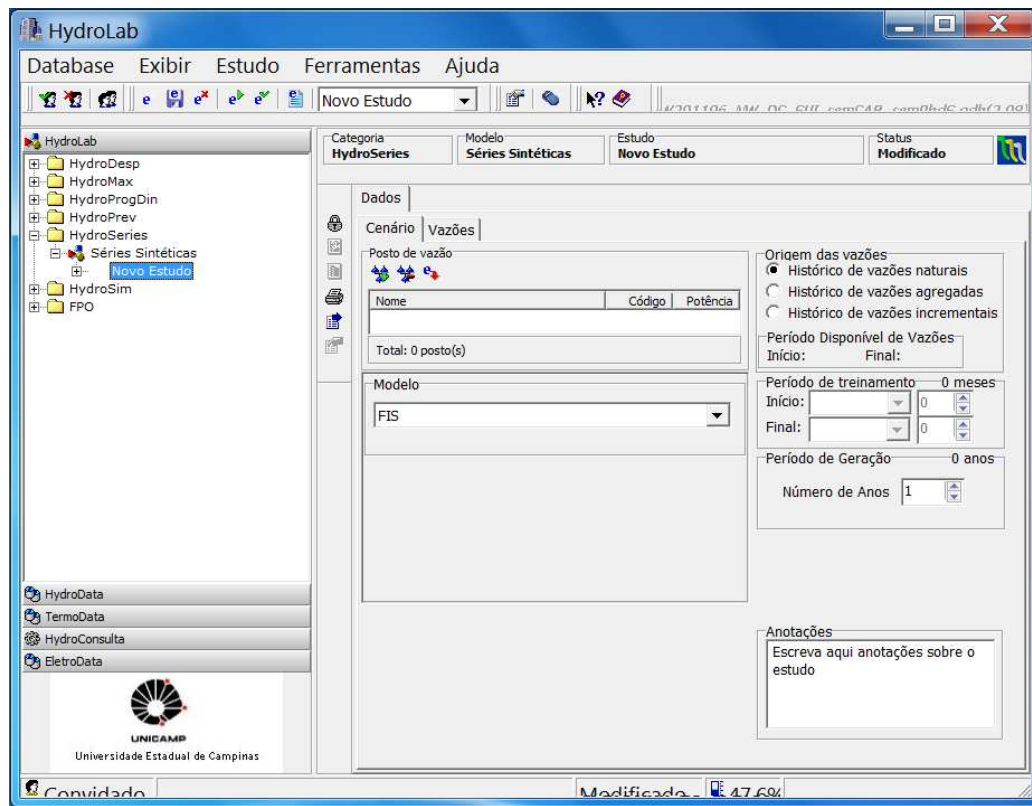


Figura 5.2: Módulo HYDROSERIES

dos desvios quadráticos. O coeficiente de assimetria indica o grau de afastamento que uma distribuição assimétrica apresenta do seu eixo de simetria. A medida de curtose determina o grau de achatamento da distribuição em relação à curva normal.

Além da visualização dos resultados na forma gráfica, o usuário pode usar a opção **Visualizar arquivos** definida na função **Estudos**. Se o usuário escolhe esta forma de visualização dos resultados, cria-se uma nova aba chamada **Arquivos** a qual contém quatro subdivisões. Cada uma dessas subdivisões tem o nome do estudo com extensão distinta segundo as informações contidas em cada aba (Figura 5.8). Por exemplo, a aba *Furnas.hss* traz informações sobre o estudo: período de treinamento, anos a serem gerados, o modelo e usina escolhidos. A aba *Furnas.usi* contém as informações sobre o posto escolhido e finalmente, a aba *Furnas.vaz*, contém os dados de vazões. Já a aba *Furnas.rhss*, contém os resultados da etapa de ajuste dos modelos, caso estes não tenham vindo de estudos anteriores, assim como a configuração dos modelos, a série gerada e os dados estatísticos observados no histórico utilizado e estimados sobre a série gerada.

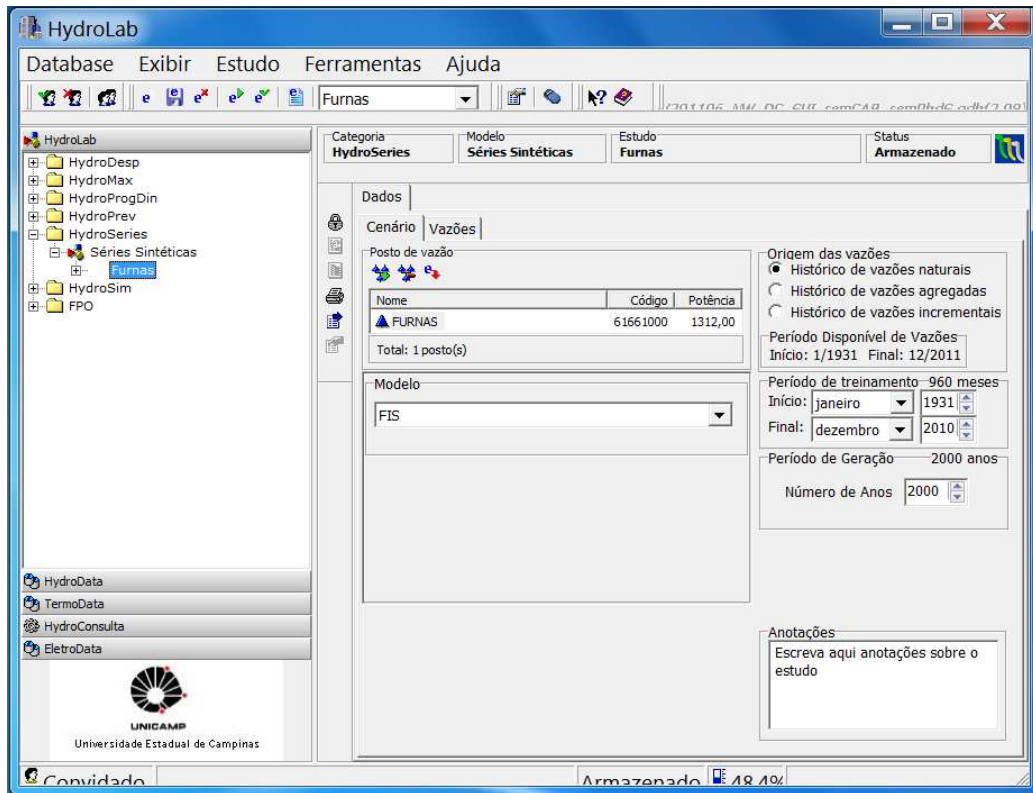


Figura 5.3: Dados de um estudo do módulo HYDROSERIES

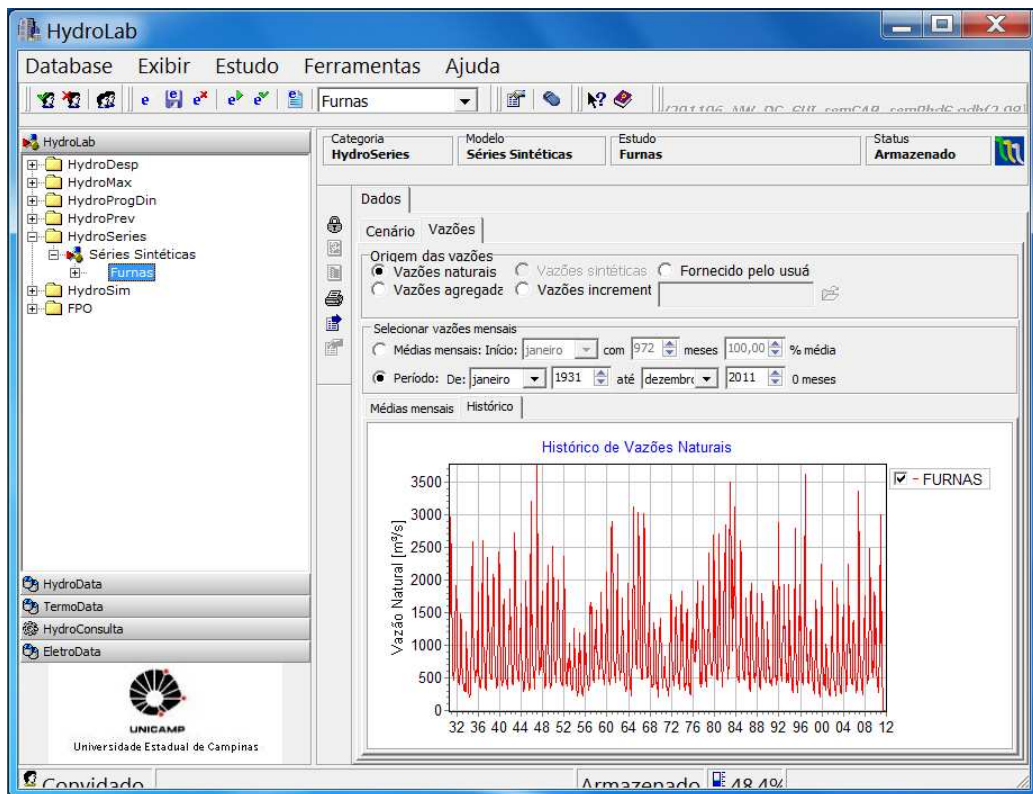


Figura 5.4: Aba para visualização do histórico de vazões

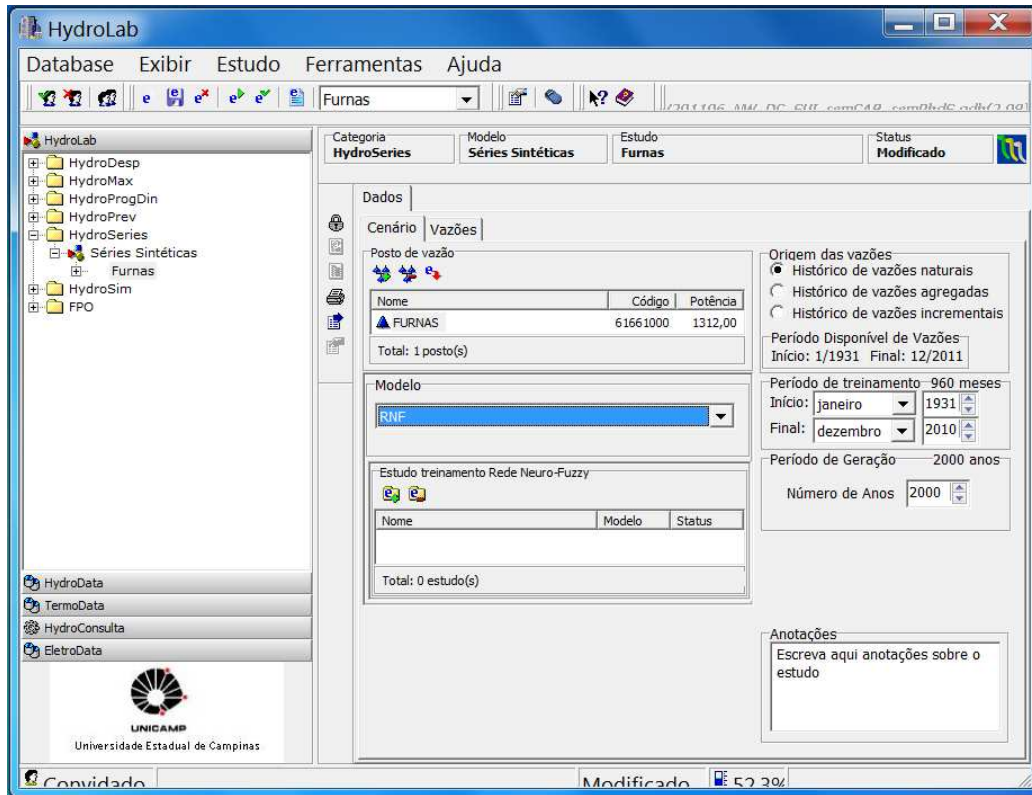


Figura 5.5: Escolha do modelo RNF

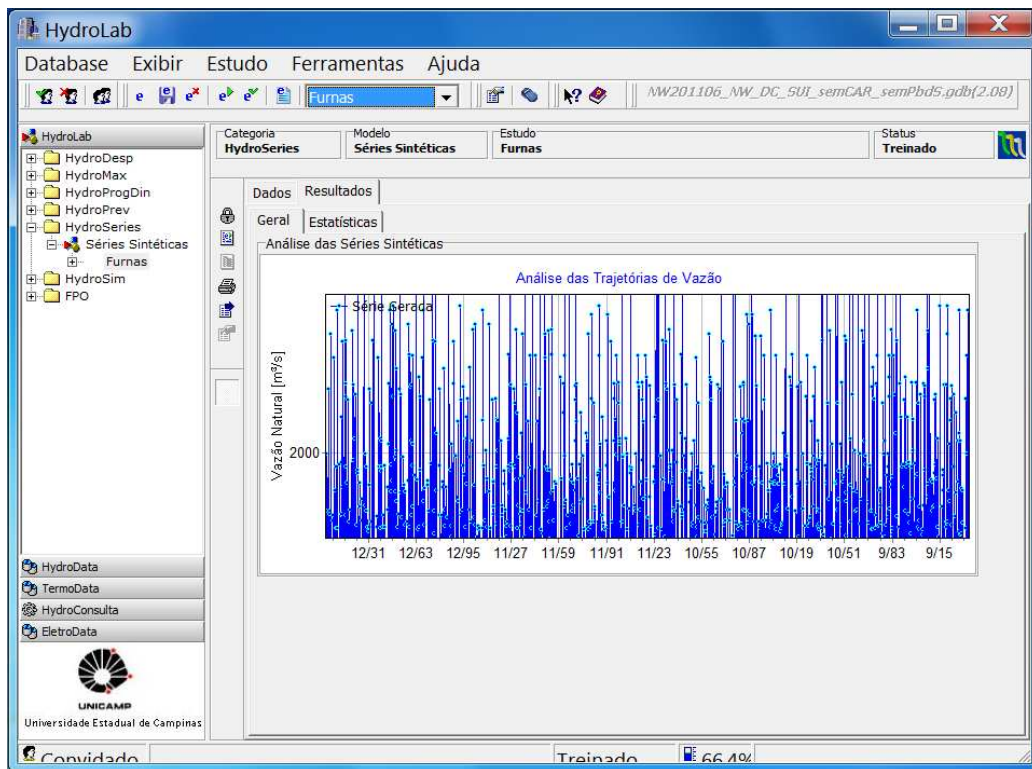


Figura 5.6: Série Gerada

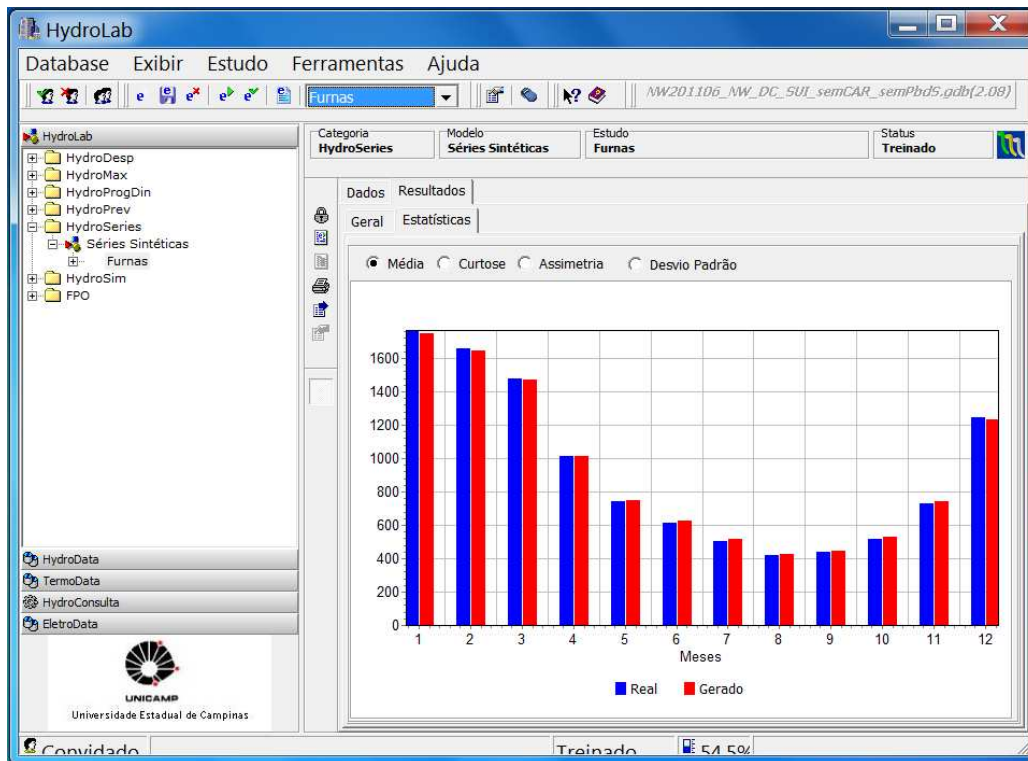


Figura 5.7: Aba para comparações estatísticas

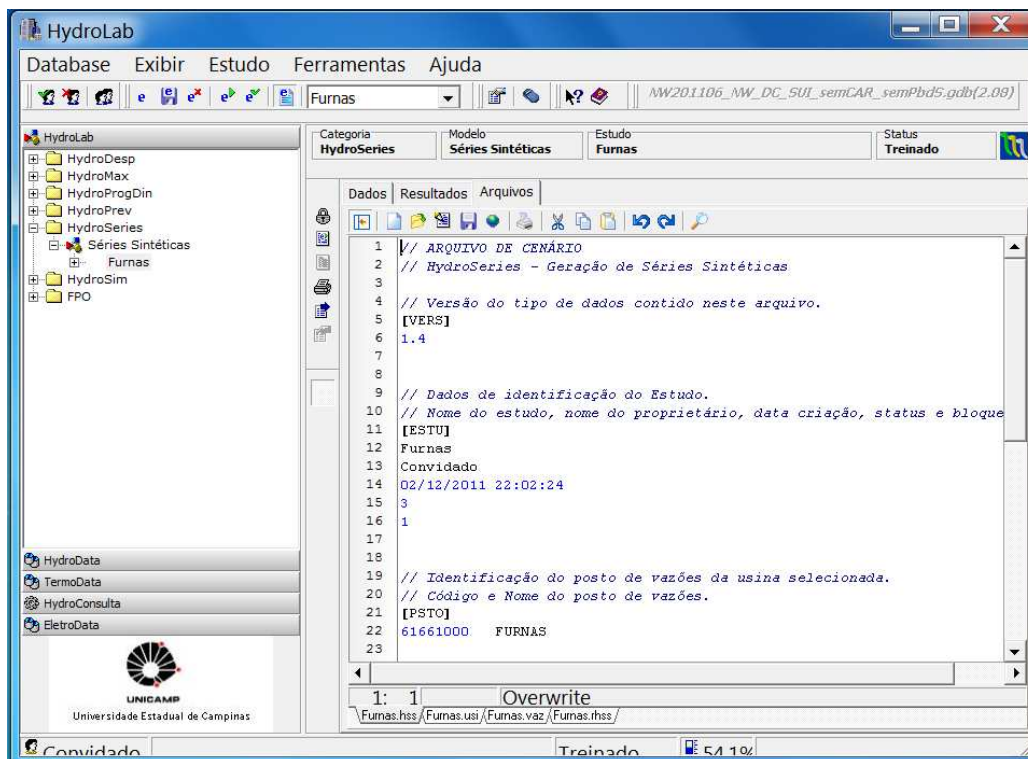


Figura 5.8: Visualização dos resultados na forma de arquivos

Estudos de Caso

Com a implementação do sistema de suporte para previsão e geração de séries sintéticas de vazões, possibilita-se a validação dos modelos utilizados para diferentes postos de vazões. A fim de apresentar os recursos que o sistema oferece, as seções a seguir mostram alguns estudos de caso.

Para avaliar o desempenho de cada modelo, escolheu-se os postos de Furnas, localizado na cascata do Rio Grande, região sudeste do Brasil. O histórico de vazões afluentes naturais médias mensais disponíveis são de janeiro de 1931 até dezembro de 2010.

Dessa forma, para ajustar tanto os parâmetros quanto o número de entradas dos modelos para um determinado mês foram calculados os critérios BIC variando o número de atrasos $p = 1, \dots, 6$. O valor de p que apresentou menor índice foi escolhido para representar as informações dos padrões de vazões.

6.1 Previsão de Vazões

A construção de modelos de previsão de vazões médias mensais tem por objetivo estimar as vazões futuras para um melhor planejamento da operação do SIN. Os resultados que serão apresentados a seguir correspondem a abordagem citada anteriormente para o planejamento de 12 meses. Devido ao tamanho reduzido das amostras disponíveis de vazões, dois conjuntos de estudos foram desenvolvidos estudos de previsão sobre o histórico de vazões e sobre as séries sintéticas de vazões.

Para o primeiro conjunto de estudos, a partir do histórico disponível, determina-se um conjunto de ajustes e um conjunto de validação. Neste caso, O conjunto de dados utilizados para o treinamento dos modelos é composto pelo histórico de janeiro de 1931 até dezembro de 1980 e o conjunto de teste, ou validação do modelo, é composto pelos dados a partir de janeiro de 1981 até dezembro de 2010.

Para a previsão sobre as séries sintéticas, o conjunto de dados considerados são séries sintéticas geradas pelos modelos univariados. Estas previsões serão analisadas na seção 6.3.

Em ambos os casos, os procedimentos de normalização de dados, identificação dos modelos e ajuste dos parâmetros são desenvolvidos de forma idêntica.

As principais medidas de seleção das variáveis de entrada dos modelos consideradas neste

estudo de caso são as raiz do erro quadrático médio REQM ou RMSE (do termo em inglês *Root Mean Squared Error*), e o critério de informação de Bayes BIC (Schwarz 1978).

Para a análise de desempenho dos modelos de previsão, serão considerados como referência dois meses do histórico de vazões: o mês de dezembro representando o período úmido e o mês de junho como representante do período seco.

O confronto entre os modelos foi realizado considerando a REQM como indicador de desempenho para os diferentes horizontes de previsão ($h = 1, \dots, 12$). Ainda, para verificar o poder preditivo dos modelos, utilizou-se como referência o desvio padrão observado na série, assim como a média de longo termo (MLT).

Como as séries de vazões apresentam sazonalidade, os erros de previsão vários passos à frente também apresentarão variações sazonais, o que dificulta a análise da degradação da previsão com o horizonte de previsão. Assim, para melhor análise dos resultados, optou-se por avaliar os erros de previsão vários passos atrás até um horizonte de 12 meses, casos em que o objetivo da previsão é sempre o mesmo mês, facilitando assim a análise dos resultados.

Como será visto, os estudos mostram que este critério de desempenho converge para um limite superior, de forma independente à natureza do modelo de previsão, isto devido à realimentação do erro e à consideração exclusiva do erro de previsão um passo à frente na função objetivo utilizada pelos algoritmos de otimização.

Estudo 1 Este estudo exemplifica o modelo RNF em base mensal com as configurações previamente ditas. A Figura 6.1 ilustra em azul o período de treinamento e em verde o período

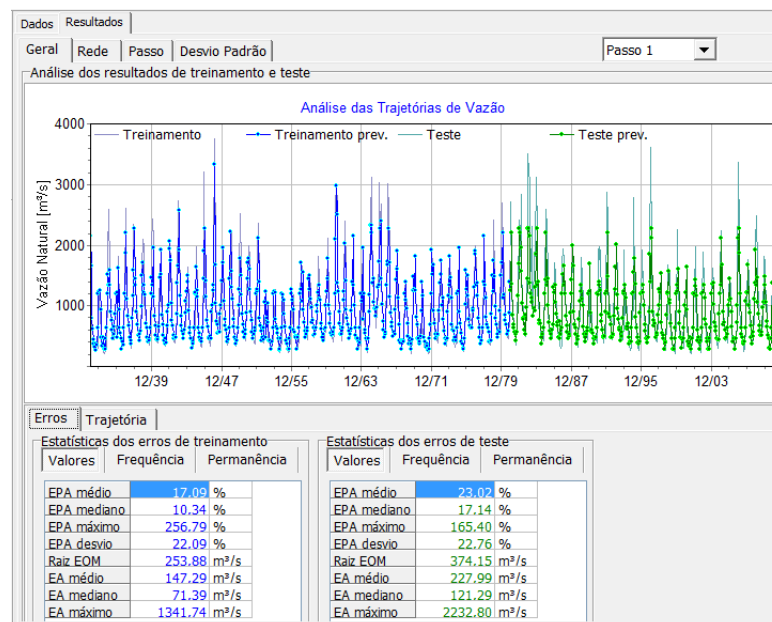


Figura 6.1: Resultado do treinamento do modelo RNF para Furnas

de teste. Nas tabelas abaixo do gráfico, pode-se analisar diferentes medidas de erro do modelo tanto para o período de treinamento quanto para o período de teste como: erro percentual absoluto (EPA) médio, mediano, máximo e do desvio padrão; a raiz do erro quadrático médio; erro absoluto (EA) médio, mediano e máximo.

As Figuras 6.2 e 6.3 mostra a evolução da REQM com o aumento do horizonte de previsão, além do desvio padrão mensal médio observado no histórico de vazões assim como a MLT. Observa-se que tanto para o mês de junho, representando o período seco, quanto para o mês de dezembro, a REQM tende ao limitante superior com o aumento do horizonte de previsão.

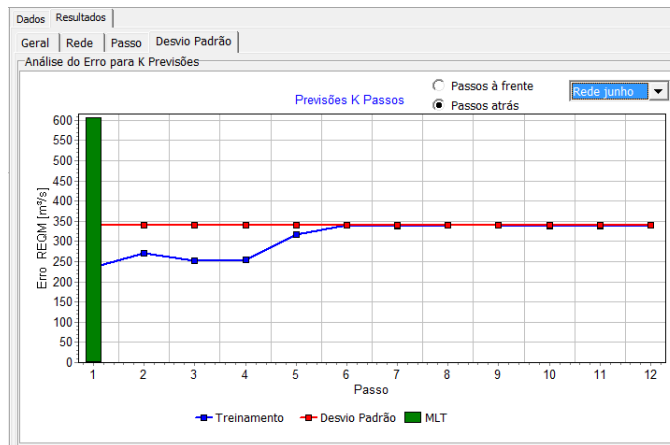


Figura 6.2: REQM do mês de junho para o posto de Furnas

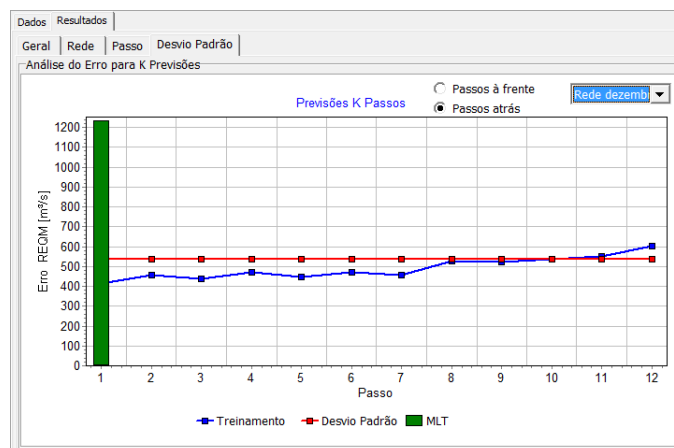


Figura 6.3: REQM do mês de dezembro para o posto de Furnas

Estudo 2 O estudo trata da comparação entre os modelos RNF mensal e a abordagem agregada, afim de ilustrar o poder preditivo do modelo agregado quando o horizonte de previsão é longo. A Figura 6.4 mostra uma previsão anual.

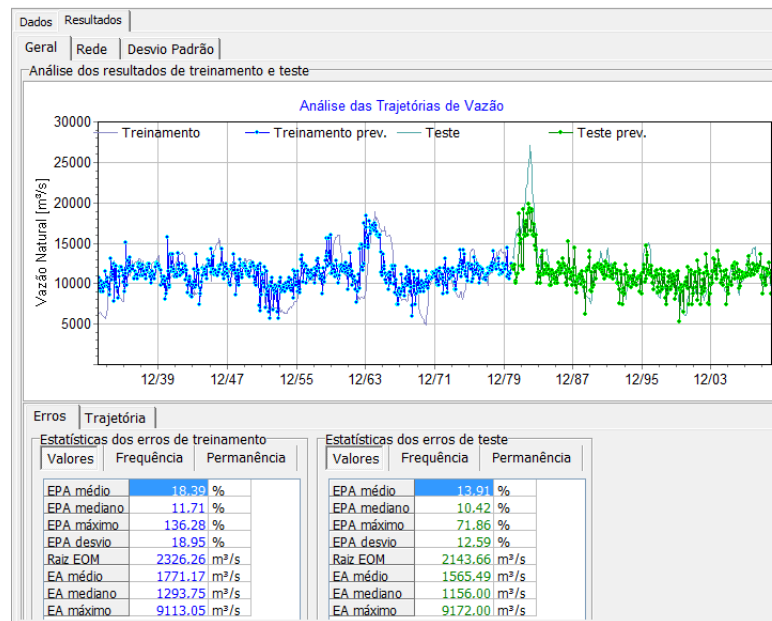


Figura 6.4: Resultado preditivo da abordagem agregada

Isto justifica, os altos valores de vazões que pode-se observar na figura. Ou seja, cada ponto no gráfico equivale a uma previsão anual assim como os erros são referentes a valores anuais. A partir dessa previsão, os valores são desagregados em doze valores mensais proporcionais à MLT.

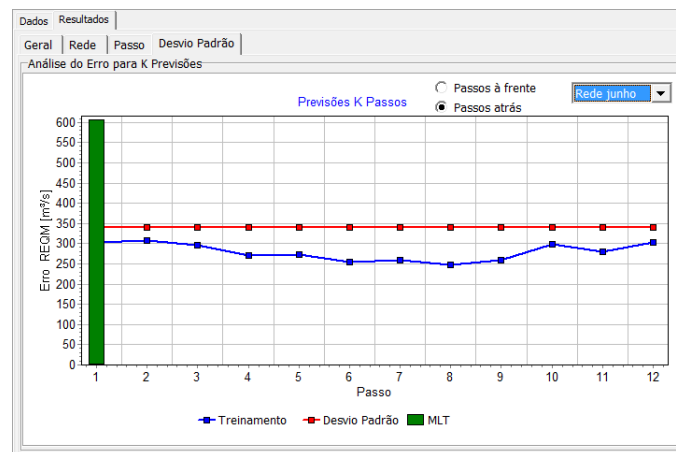


Figura 6.5: REQM das previsões da abordagem agregada para o mês de junho

Uma segunda observação que deve ser feita é em relação ao EPA médio menor da previsão anual em relação a previsão mensal. Na tabela de erros situada abaixo do gráfico na Figura 6.4, nota-se um erro de 13,91% enquanto para a previsão mensal do estudo anterior tem uma EPA médio de 23,02%. O mesmo acontece para a raiz EQM e para os erros absolutos, valores

também presentes nesta tabela, o que mostra o melhor poder preditivo da previsão anual em relação a previsão mensal.

Observa-se na Figura 6.5 e na Figura 6.6 que o poder preditivo do modelo agregado apresenta resultados satisfatórios para um longo período de previsão, permanecendo sempre abaixo do desvio padrão.

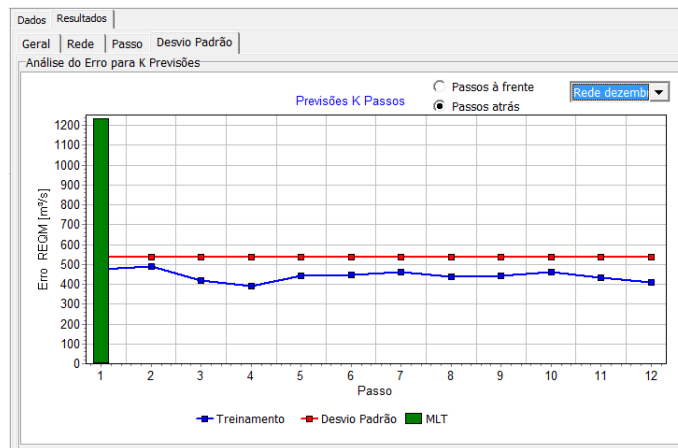


Figura 6.6: REQM das previsões da abordagem agregada para o mês de dezembro

Comparando a previsão mensal com a abordagem agregada, nota-se que no mês de junho (Figura 6.7) que para os primeiros 3 passos a previsão em base mensal mostra em média um melhor desempenho, visto que esta metodologia foi desenvolvida para ter bons resultados para uma previsão de curto prazo. Contudo, para horizontes de previsão superiores ao terceiro mês, os resultados obtidos a partir da desagregação das estimativas anuais apresentam valores de REQM em média inferiores aos obtidos pelos modelos mensais, atingindo níveis de erros inferiores aos desvio padrão observado e portanto, apresentando poder preditivo no longo prazo.

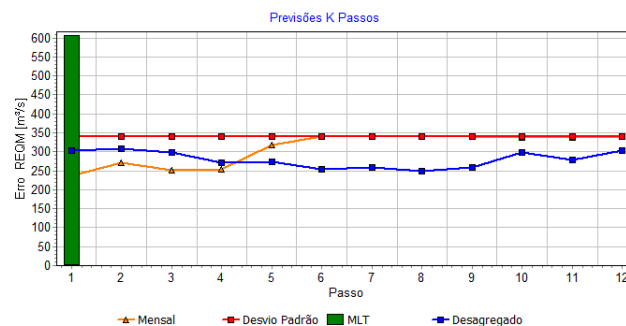


Figura 6.7: Comparação entre previsão mensal e abordagem agregada para o mês de junho

Já para o mês de dezembro (Figura 6.8) nos dois primeiros passos, tem-se em média um melhor desempenho nas previsões com base mensal, mas para uma previsão a partir do terceiro passo a previsão com base no histórico anual, apresenta em média, um desempenho melhor.

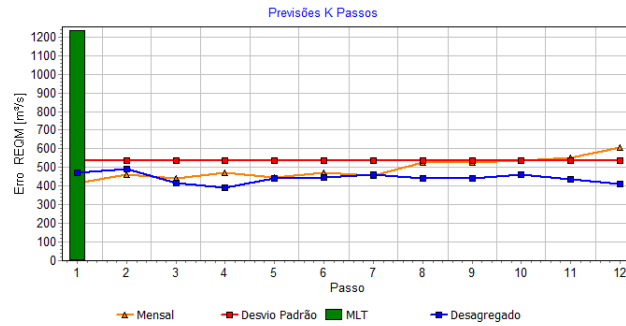


Figura 6.8: Comparação entre previsão mensal e abordagem agregada para o mês de dezembro

Estudo 3 Um terceiro estudo mostra o modelo PAR. Na Figura 6.9 observa-se a previsão deste modelo assim como a tabela com os erros do período de treinamento e teste. Tanto para

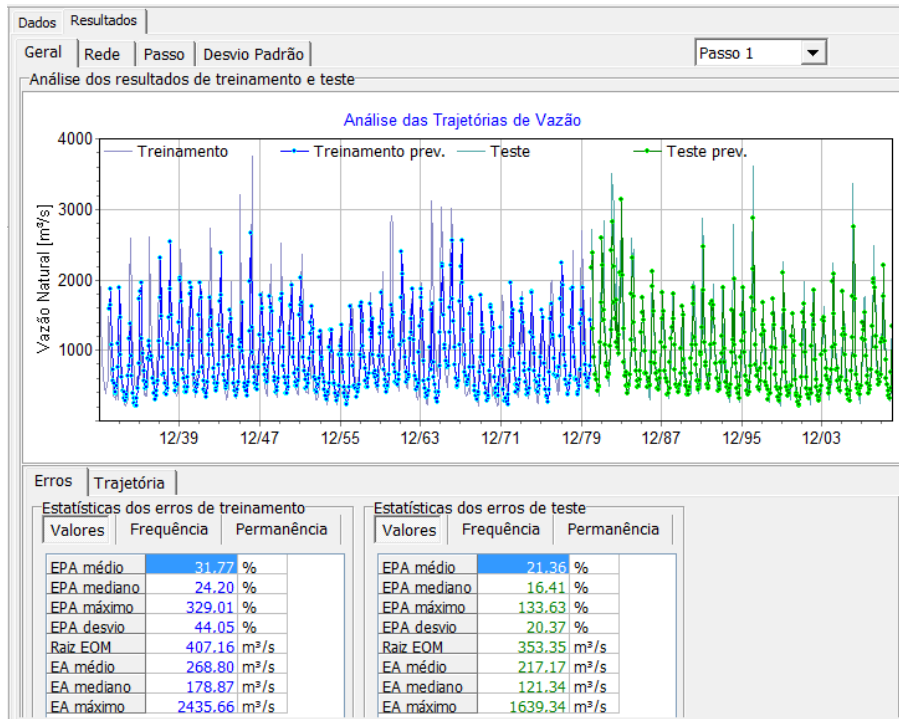


Figura 6.9: Previsão do modelo PAR

o mês de junho (Figura 6.10) como para o mês de dezembro (Figura 6.10), a convergência da REQM para o desvio padrão ao longo do horizonte de previsão novamente é verificada.

Comparando a capacidade preditiva do modelo PAR com as previsões do modelo RNF e abordagem agregada, para o mês de junho (Figura 6.12) observa-se uma melhor predição nos primeiros quatro passos para RNF e PAR. A partir do quarto passo, a abordagem agregada mantém menores REQM's enquanto os outros modelos tendem ao desvio padrão. Para o mês de dezembro (Figura 6.13), os valores do REQM dos modelos apresentam erros próximos, exceto o modelo RNF que apresenta uma ligeira desvantagem nos últimos dois passos por ultrapassar os valores do desvio padrão.

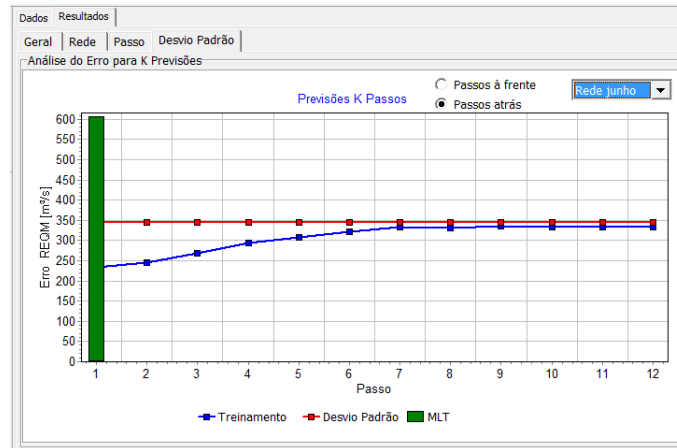


Figura 6.10: REQM das previsões do modelo PAR para o mês de junho

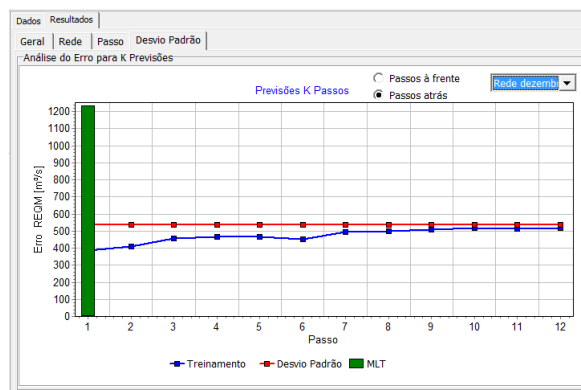


Figura 6.11: REQM das previsões do modelo PAR para o mês de dezembro

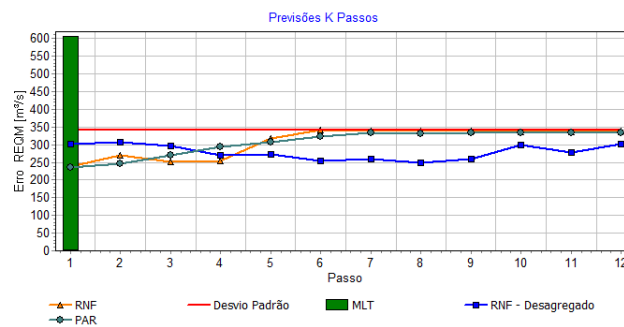


Figura 6.12: REQM das previsões do modelo PAR para o mês de junho

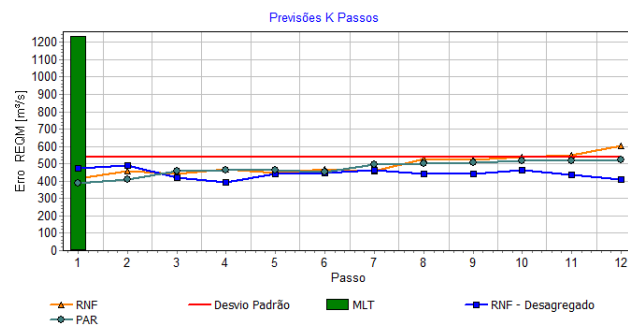


Figura 6.13: REQm das previsões do modelo PAR para o mês de dezembro

6.2 Geração de Séries Sintéticas

Para avaliar os modelos de geração de séries sintéticas, quatro estatísticas principais foram consideradas como métricas de validação: a média, o desvio padrão, a assimetria e a curtose, estimadas para cada mês. Além disso, as distribuições observadas e geradas também foram comparadas em níveis mensais e de uma forma mais geral ao longo de todo o histórico observado.

A inicialização da geração da série sintética considera as médias mensais do histórico de vazões como valores iniciais. São geradas séries sintéticas equivalentes a dois mil anos, sendo os primeiros cem anos desprezados para fins de eliminação do efeito da inicialização.

Para avaliar os modelos propostos, utilizou-se como estudo de caso a série de vazões naturais médias mensais da UHE de Furnas, com histórico de vazões de Janeiro de 1931 a Dezembro de 2010.

Estudo 1 A Figura 6.14 referente ao posto de Furnas apresenta os valores médios mensais das séries sintéticas geradas pelos modelos BCPAR, MLP, FIS e RNF. Os modelos FIS e PAR, em geral, foram os modelos que apresentaram valores mais similares ao observado. Nota-se que a série gerada pelo modelo RNF, replicou médias maiores que o observado para os meses de janeiro, fevereiro e maio, e médias menores para os meses de março, abril, junho, outubro e dezembro, mas mostrou resultados similares para os meses de julho e novembro. Já o modelo MLP, replicou uma média próxima das observadas, exceto para os meses de janeiro, setembro, novembro e dezembro que apresenta um média pouco maior em relação aos outros meses.

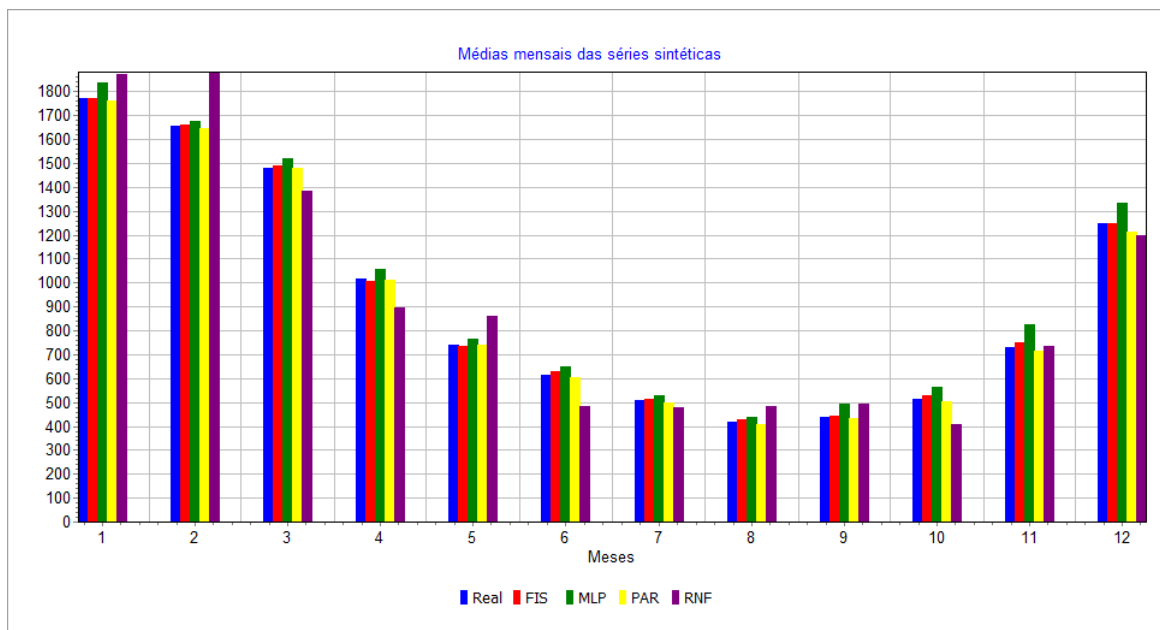


Figura 6.14: Médias mensais das séries sintéticas geradas

Os coeficientes de assimetria podem ser visualizados na Figura 6.15. O modelo FIS mostrou-se mais eficaz em relação aos outros modelos uma vez que consegue replicar o mais próximos dos valores observados em relação aos outros modelos para todos os meses. Os modelos MLP e RNF obtém resultados satisfatórios somente para o mês de fevereiro. Enquanto o modelo

BCPAR, que no HydroSeries é chamado de PAR, apresenta um coeficiente de assimetria mais próximo do observado somente para o mês de janeiro.

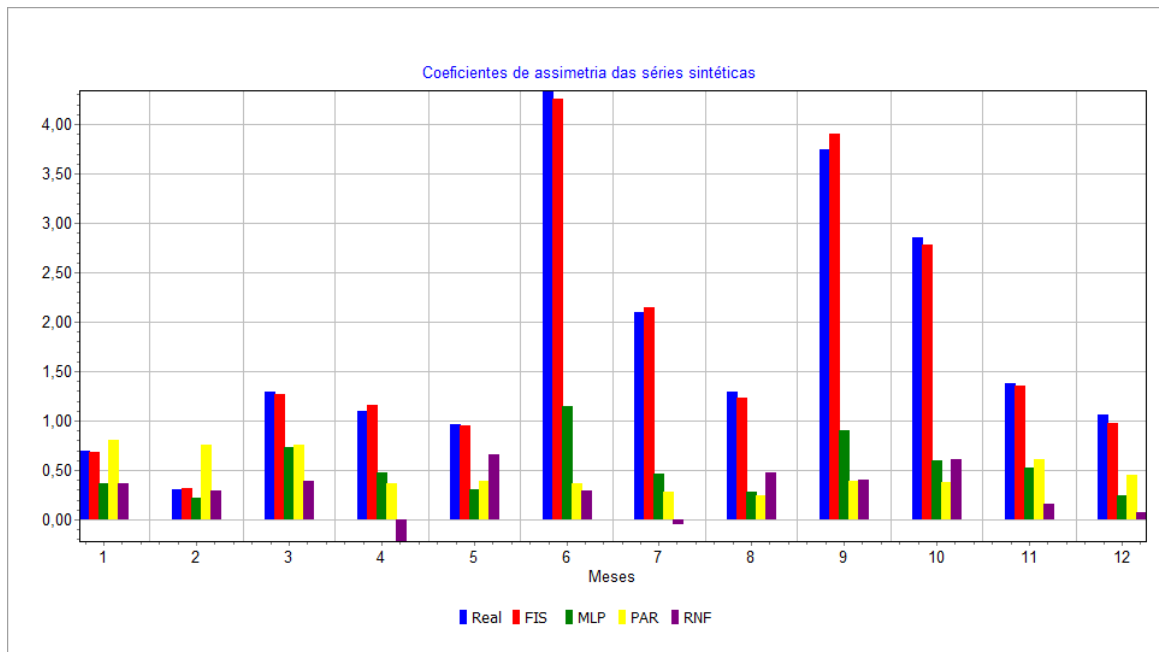


Figura 6.15: Coeficientes de assimetria das séries sintéticas geradas

Em relação a curtose, exibida na Figura 6.16, assim como com os coeficientes de assimetria, o modelo FIS em geral, mostrou-se melhor pelo fato de replicar valores das medidas de curtose próximo dos valores da série observada para todos os meses. Analisando os outros modelos, observa-se que os modelos PAR e MLP na maioria dos meses possuem medidas de curtose com pequenas diferenças, exceto no mês de junho que o modelo MLP possui valor maior. O modelo RNF apresenta valores próximos da medida de curtose da série observada principalmente para os meses de janeiro, fevereiro, março, abril e maio.

Quanto ao desvio padrão, de modo geral, os valores estimados são próximos da série observada, exceto o modelo RNF que nos meses de janeiro, fevereiro, abril, maio, junho, julho, agosto e dezembro apresenta desvio padrão maior em relação aos outros modelos. Pode-se observar na Figura 6.17 que os modelos FIS, MLP e PAR apresentam valores próximos ao da série observada, dando destaque para os meses de julho e agosto pela proximidade dos valores.

Em geral, as estatísticas descritivas para todos os meses foram devidamente replicadas. Por outro lado, os modelos apresentaram dificuldades em replicar os coeficientes de assimetria e curtose para alguns meses das séries de vazões. Deve ser ressaltado que os modelos FIS e MLP não consideram a hipótese de normalidade dos dados, como é feito pela maioria dos modelos de séries temporais. Os resultados forneceram evidências que apontam que os modelos utilizados são uma alternativa para a geração de séries sintéticas de vazões mensais e que reproduzem as principais características estatísticas das séries observadas.

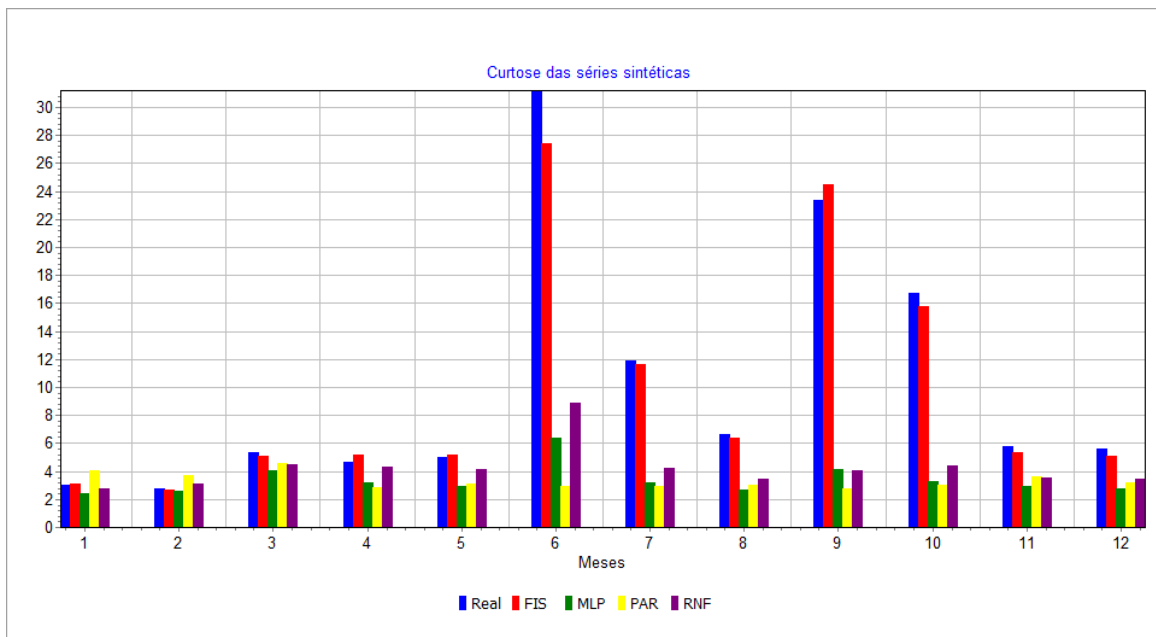


Figura 6.16: Medidas de curtose das séries sintéticas geradas

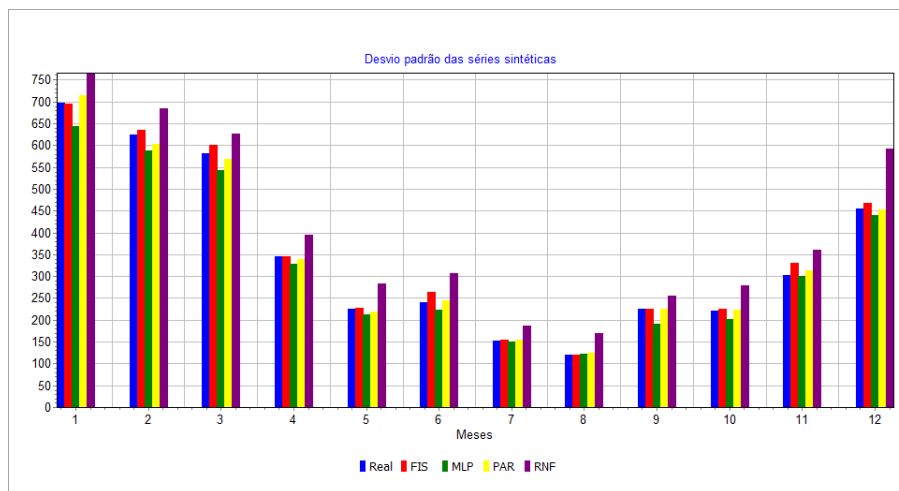


Figura 6.17: Desvio padrão das séries sintéticas geradas

6.3 Previsão de vazões sobre séries sintéticas

A seção 6.1 apresentou os resultados de previsão de diferentes modelos e abordagens utilizados neste trabalho, sobre as series históricas de vazões do posto de Furnas. De forma geral, observou-se uma capacidade limitada de previsão no períodos úmidos, e uma capacidade maior de previsão nos períodos secos. Contudo, o comportamento assintótico da REQM para os diferentes modelos apresentou alguns desvios devido principalmente ao tamanho da amostra utilizada. Assim, com intuito de consolidar os resultados de previsão e inferir de forma mais consistente sobre o desempenho dos modelos de previsão, procedeu-se à comprovação dos mesmos em uma amostra maior.

Nesta seção os modelos de previsão foram ajustados sobre as séries sintéticas de vazões naturais. Considerando a diferença pouco significativa entre o desempenho dos diferentes modelos, optou-se pelo ajuste do modelo RNF e do modelo MLP para fins de comparação. Lembrando que a tendência da REQM ao desvio padrão mensal observada na série, neste caso ao desvio padrão da série sintética, é um comportamento esperado para todo modelo de previsão, de forma independente à sua natureza (linear ou não linear) e devido ao efeito da realimentação do erro de previsão. Assim, as conclusões sobre estes estudos se estendem a todos os modelos de previsão.

Estudo 1 O primeiro estudo mostra resultados de previsão do modelo PAR utilizando uma série sintética gerada pelo modelo BCPAR. Observa-se que o REQM, tanto para o mês de junho (Figura 6.18) quanto para o mês de dezembro (Figura 6.19), tende rapidamente ao desvio padrão.

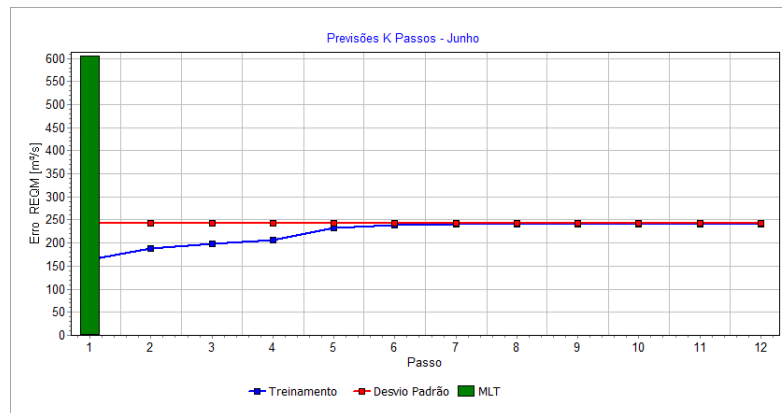


Figura 6.18: REQM das previsões do modelo PAR sobre séries sintéticas do mês de junho

Estudo 2 O estudo 2 trás um comparativo entre as previsões PAR e RNF sobre a série sintética MLP. Nas Figuras 6.20 e 6.21 pode-se observar o erro tendendo ao desvio padrão.

Finalmente, com base nos estudos de previsão sobre as séries sintéticas, podemos concluir que o estudo do desempenho dos previsores devido ao grande número de valores observados, permitiu comprovar a convergência assintótica do erro de previsão vários passos a frente para o desvio padrão incondicionado, fornecendo uma visão mais precisa do desempenho dos modelos

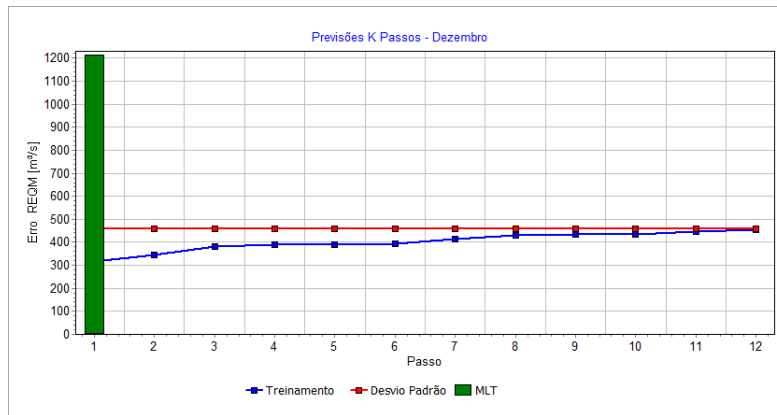


Figura 6.19: REQM das previsões do modelo PAR sobre séries sintéticas do mês de dezembro

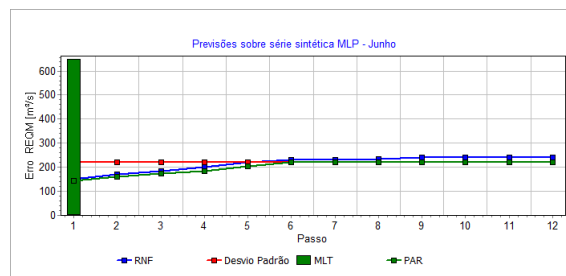


Figura 6.20: REQM das previsões do modelo PAR sobre séries sintéticas do mês de junho

de previsão de vazões em base mensal. A conclusão é de que, em função da baixa correlação médias das vazões, a REQM da previsão tende em poucos meses para o desvio padrão da série, sobretudo nos meses úmidos do ano quando as vazões são mais expressivas e as correlações são mais baixas.

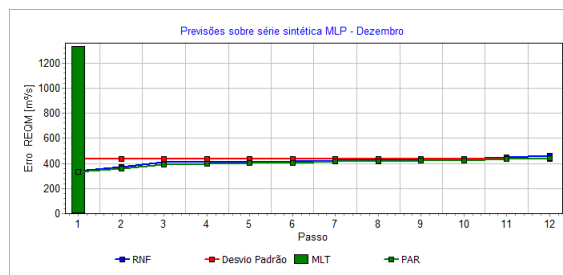


Figura 6.21: REQM das previsões do modelo PAR sobre séries sintéticas do mês de dezembro

Conclusões e Perspectivas

Conclusões

O sistema computacional apresentado neste trabalho possui recursos de grande valia perante as ferramentas disponíveis para o planejamento da operação existentes hoje no setor elétrico, visto que ele utiliza uma base dados composta pelas usinas hidrelétricas gerenciadas pelo Operador Nacional do Sistema Elétrico (ONS), destacando a possibilidade de visualização gráfica dos dados.

Com o acréscimo dos novos modelos de previsão de vazões no módulo HydroPrev, o sistema permite ao usuário comparar e analisar o poder preditivo dos modelos computacionais implementados no módulo de previsão para qualquer usina hidrelétrica despachada pelo ONS.

Tratando-se do módulo de geração de séries sintéticas, o sistema disponibiliza séries sintéticas mensais de vazões, amplamente utilizadas em simulações e técnicas de otimização uma vez que a não disponibilidade de históricos de vazões de tamanho suficiente dificulta a obtenção de resultados confiáveis. A inclusão do módulo HydroSeries possibilita que o usuário obtenha séries sintéticas de qualquer tamanho que desejar de qualquer usina hidrelétrica.

Perspectivas

Como trabalhos futuros, vislumbra-se a inserção de novos modelos no sistema computacional, tanto de previsão como de geração de séries sintéticas de vazões, e a avaliação comparativa dos diferentes modelos considerados. Também replicar os modelos apresentados para o planejamento semanal.

Outra sugestão seria “desautomatizar” os modelos MLP, PAR e FIS, uma vez que atualmente esses modelos utilizam a ferramenta de varredura e o critério BIC para determinar automaticamente o número de entradas e os hiperparâmetros de ajustes dos modelos. Assim, a exemplo do modelo RNF, além do modo “automático”, o usuário poderá determinar o número de entradas e o limiar do ajuste dos parâmetros dos modelos manualmente.

Publicações

- SBSE 2012: Maiana Santos Lopes, Ivette Luna, Rosangela Ballini, Secundino Soares Filho. Previsão de vazões mensais para o planejamento da operação energética do SIN. *IV Simpósio Brasileiro de Sistemas Elétricos*, Goiânia-GO.
- IEEE PES T&D LA 2012: Mônica S. Zambelli, Maiana Santos Lopes, Secundino Soares Filho. Long-Term Hydropower Scheduling using Model Predictive Control Approach with Hybrid Monthly-Annual Inflow Forecasting. *Sixth IEEE/PES Transmission and Distribution: Latin America Conference and Exposition (T&D-LA)*, Montevideo, Uruguai

Bibliografia

- Akbari, M., van Overloop, P. J. & Afshar, A. (2010). Clustered k nearest neighbor algorithm for daily inflow forecasting, *Water Resources Management* pp. 1–17.
- Angelov, P. P. & Filev, D. P. (2004). An approach to online identification of takagi-sugeno fuzzy models, *IEEE Transactions on Systems, Man and Cybernetics-part B* **34**(1): 484–498.
- Bakirtzis, A., Theocharis, J., Kiartzis, S. & Satsios, K. (1995). Short term load forecasting using fuzzy neural networks, *Power Systems, IEEE Transactions on* **10**(3): 1518 –1524.
- Ballini, R. (1996). *Redes Neurais para Previsão Chuva-Vazão*, Dissertação de Mestrado, Universidade de São Paulo, São Carlos, SP.
- Ballini, R. (2000). *Análise e Previsão de Vazões Utilizando Modelos de Séries Temporais, Redes Neurais e Redes Neurais Nebulosas*, Tese de doutorado, FEEC/UNICAMP, Campinas, SP.
- Ballini, R., Soares, S. & Andrade, M. G. (1998). An adaptive neural fuzzy network model for seasonal streamflow forecasting, *Proceedings of the IV Brazilian Symposium on Neural Networks–IEEE Computer Society*, pp. 215–219.
- Ballini, R., Soares, S. & Andrade, M. G. (1999). Comparison of seasonal streamflow forecasting models, *Proc. of the International Conference on Intelligent System Application to Power Systems - ISAP'99*, pp. 417–421.
- Bezdek, J. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, USA.
- Boswell, M. T., Ord, J. K. & Patil, G. P. (1979). *Statistical Distributions in Ecological Work*, International Cooperative Publishing House, chapter Normal and lognormal distributions as models of size, pp. 72–87.
- Box, G. & Cox, D. (1964). An analysis of transformations, *J. R. Statist. Soc. Serie B* **26**: 211–252.
- Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. (2008). *Time series analysis: forecasting and control*, 4th. edition edn, Holden-Day.

- Box, G., Jenkins, G. & G. C. Reinsel, G. C. (1994). *Time Series Analysis, Forecasting and Control*, 3rd edn, Holden Day, Oakland, California, EUA.
- Caminhas, W. M., Tavares, H. M. F. & Gomide, F. (1996). Pattern classification using neurofuzzy and neural networks, *Proceedings Information Processing and Management of Uncertainty on Knowledge - Based Systems - IPMU'96*, Vol. 1, pp. 175–180.
- Chiu, S. L. (1994). A cluster extension method with extension to fuzzy model identification, *Proc. Third IEEE Conf. Fuzzy Systems IEEE World Congress Computational Intelligence*, pp. 1240–1245.
- Chow, V. T. (1964). *Handbook of applied hydrology*, MacGraw-Hill Book Co., New York, EUA.
- Cicogna, M. A. (2003). *Sistema de Suporte a Decisão para o Planejamento e a Programação da Operação de Sistemas de Energia Elétrica*, Tese de doutorado, Faculdade de Engenharia Elétrica e de Computação, Unicamp.
- Cottrell, M., Girard, B., Girard, Y., Mangeas, M. & Muller, C. (1995). Neural modeling for time series: A statistical stepwise method for weight elimination, *IEEE Transactions on Neural Networks* **6**(6): 1355–1364.
- Crone, S. (2005). Stepwise selection of artificial neural network models for time series prediction, *Journal of Intelligent Systems* **14**(2-3): 99–122.
- Er, M. J. & Wu, S. (2002). A fast learning algorithm for parsimonious fuzzy neural systems, *Fuzzy Sets and Systems* **126**: 337–351.
- Evsukoff, A. G., de Lima, B. S. L. P. & Ebecken, N. F. F. (2011). Long-term runoff modeling using rainfall forecasts with application to the iguaçu river basin, *Water Resources Management* **25**: 963–985.
- Figueiredo, M. & Gomide, F. (1998). Fuzzy system design using neurofuzzy networks, *Proceedings of the 7th Information Processing and Management of Uncertainty on Knowledge - Based Systems - IPMU'98*, Vol. 2, pp. 1416–1422.
- Francelin, R., Ballini, R. & Andrade, M. G. (1996). Back-propagation and Box & Jenkins approaches to streamflow forecasting, *Latin-Iberian-American Congress on Operations Research and System Engineering - CLAIO, Simpósio Brasileiro de Pesquisa Operacional - SBPO*, Rio de Janeiro-RJ, pp. 1307–1312.
- Gladyshev, E. (1961). Periodically correlated random sequences, *Sov. Math.* 2 pp. 383–388.
- Haykin, S. (1994). *Neural Networks, A Comprehensive Foundation*, Macmillan Publishing Company.
- Haykin, S. (2001). *Redes Neurais, Princípios e Prática*, 2 edn, Bookman.

- Hazen, A. (1914). Storage to be provided in impounding reservoir for municipal water supply, *Trans. Am. Soc. Civ. Eng.* **77**: 1539–1640.
- Hong, Y.-S. T. & White, P. A. (2009). Hydrological modeling using a dynamic neuro-fuzzy system with on-line and local learning algorithm, *Advances in water resources* **32**: 110–119.
- Hush, D. & Horne, B. (1993). Progress in Supervised Neural Networks, *IEEE Signal Processing Magazine* **10**(1): 8–39.
- Jacobs, R., Jordan, M., Nowlan, S. & Hinton, G. (1991). Adaptive mixture of local experts, *Neural Computation* **3**(1): 79–87.
- Jardim, D. L. D. D., Maceira, M. E. P. & ao, D. M. F. (2001). Stochastic streamflow model for hydroelectric systems using clustering techniques, *IEEE Porto Power Tech Conference*, IEEE, Porto, Portugal, pp. 1–6.
- Kacprzyk, J. & Fedrizzi, M. (1992). *Fuzzy regression Analysis. Studies in Fuzziness*, Vol. 1, Physica-Verlag, Heidelberg.
- Kuo, R. J. & Xue, K. C. (1999). Fuzzy neural networks with application to sales forecasting, *Fuzzy Sets and Systems* **108**: 123–143.
- Lachtermacher, G. & Fuller, J. D. (1995). Backpropagation in time-series forecasting, *Journal of Forecasting* **14**(4): 381–393.
- Lapedes, A. & Farber, R. (1987). *Nonlinear Signal Processing Using Neural Networks: Prediction and System Modelling*, Los Alamos National Laboratory.
- Lawrence, A. J. & Kottegoda, N. T. (1977). Stochastic modelling of riverflow time series, *J.R. Statist. Soc. Serie A* **150**(Part 1): 1–47.
- Lee, C. C. & Lin, C. (1991). Neural-network-base fuzzy logic control and decision system, *IEEE Transactions on Systems, Man and Cybernetics* **40**(12): 1320–1336.
- Lin, Y. & Lye, L. M. (1994). Modelling long-term dependence based on cumulative departures of annual flow series, *Journal of Hydrology* **160**: 105–121.
- Lopes, M., Luna, I., Ballini, R. & Soares, S. (2012). Previsão de vazões mensais para o planejamento da operação energética do SIN, *Simpósio Brasileiro de Sistemas Elétricos*.
- Luna, I. (2003). *Sistemas neurais fuzzy aplicados em identificação e controle de sistemas*, Master's thesis, Universidade Estadual de Campinas, Campinas, São Paulo, Brasil.
- Luna, I., Ballini, R., Soares, S. & Silva Filho, D. (2011). Fuzzy inference systems for synthetic monthly inflow time series generation, in P. published by Atlantis Press (ed.), *7th conference of the European Society for Fuzzy Logic and Technology*, pp. 1–7.

- Luna, I., Soares, S. & Ballini, R. (2007a). A constructive-fuzzy rule based system for time series forecasting, *International Joint Conference on Neural Networks*.
- Luna, I., Soares, S. & Ballini, R. (2007b). A constructive-fuzzy system modeling for time series forecasting, *Proceedings of The International Joint Conference on Neural Networks*.
- Luna, I., Soares, S., Lopes, J. & Ballini, R. (2009). Verifying the use of evolving fuzzy systems for multi-step ahead daily inflow forecasting, *15th International Conference on Intelligent System Applications to Power Systems – ISAP ’09* pp. 1–6.
- Magalhães, M., Ballini, R., Gonçalves, R. & Gomide, F. (2004). Predictive fuzzy clustering model for natural streamflow forecasting, *Proceedings of the IEEE International Conference on Fuzzy Systems*, Budapest, Hungary, pp. 390–394.
- Magalhães, M., Ballini, R., Molck, P. & Gomide, F. (2004). Combining forecasts for natural streamflow prediction, *Proceedings of the International Conference of the North American Fuzzy Information Processing Society*, Banff, Canadá.
- Maier, H. R. & Dandy, G. C. (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, *Environmental Modelling & Software* **15**: 101–124.
- Marinho, M. H. N., Ballini, R. & Soares, S. (2005). Redes neurais na previsão de vazões afluentes vários passos à frente via agregação de vazões., *I Congreso Paraguayo de Recursos Hidricos*, Hernandarias, Paraguai, pp. 1–11.
- Martinez, L. (2001). *Políticas de controle malha fechada e malha aberta no planejamento da operação energética de sistemas hidrotérmicos*, Tese de doutorado, Universidade Estadual de Campinas.
- MecLeod, A. I. (1994). Diagnostic checking of periodic autoregression models with application, *J. of Time Series Analysis* **15**(2): 221–233.
- Morettin, P. A. & Toloi, C. M. C. (2004). *Análise de Séries Temporais*, 1 edn, Edgard Blücher Ltda, São Paulo, SP.
- Nadarajah, S. (2005). A generalized normal distribution, *Journal of Applied Statistics* **32**: 685–694.
- Ochoa-Rivera, J. C., García-Bartual, R. & Andreu, J. (2002). Multivariate synthetic streamflow generation using a hybrid model based on artificial neural networks, *Hydrology and Earth System Sciences* **6**(4): 641–654.
- Pagano, M. (1978). On periodic and multiple autoregressions, *The Annals of Statistics* **6**: 1310–1317.
- Pedrycz, W. & Gomide, F. (1998). *An Introduction to Fuzzy Sets: Analysis and Design*, MIT Press, Cambridge, MA, EUA.

- Pereira, M. V. F., Oliveira, G. C., Costa, C. C. G. & Kelman, J. (1984). Stochastic streamflow models for hydroelectric systems (brazil), *Water Resources Research* **20**(3): 379–390.
- Phien, H. N. & Ruksasilp, W. (1981). A review of single-site models for monthly streamflow generation, *Journal of Hydrology* **52**: 1–12.
- Rumelhart, D. & McClelland, J. (1986). *Parallel Distributed Processing*, Vol. 1, MIT Press, Cambridge.
- Schwarz, G. (1978). Estimating the dimension of a model, *Ann. Statist.* **6**(2): 461–468.
- Stedinger, J. R. (2001). Report on the evaluation of cepel's par models, *Technical report*, School of Civil and Environmental Engineering - Cornell University, Ithaca, New York.
- Stedinger, J. R. & Taylor, M. R. (1982). Synthetic streamflow generation. 1. model verification and validation (delaware river, new york state), *Water Resources Research* **18**(4): 909–918.
- Takagi, T. & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control, *IEEE Transactions on Systems, Man and Cybernetics* **1**: 116–132.
- Tang, Z., Almeida, C. & Fishwick, P. (1991). Time series forecasting using neural networks vs Box-Jenkins methodology, *Simulation* **57**(5).
- Thomas, H. A. & Fiering, M. B. (1962a). Mathematical analysis of streamflow sequences in analysis of river basins by simulation, *Design of Water Resource Systems* pp. 459–493.
- Thomas, Jr., H. A. & Fiering, M. B. (1962b). *Design of Water Resources Systems*, Harvard University Press, Cambridge, Mass., chapter 12, Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation, pp. 459–493.
- Troutman, B. M. (1979). Some results in periodic autoregression, *Biometrika* **66**: 219–228.
- Tucci, C. (1998). *Modelos Hidrológicos*, Editora da UFRGS - ABRH, Porto Alegre, Brasil.
- Vecchia, A. (1985a). Maximum likelihood estimation for periodic autoregressive-moving average models, *Technometrics* **27**(4): 375–384.
- Vecchia, A. V. (1985b). Periodic autoregressive moving average (parma) modeling with application to water resources, *Water Resources Bull.* **21**: 721–730.
- Wang, L. (1994). *Adaptive Fuzzy Systems and Control*, Prentice Hall.
- Weigend, A., Ruberman, B. & Rumelhart, D. (1992). Predicting sunspots and exchange rates with connectionist networks, in M. Casdagli & S. Eubank (eds), *Nonlinear Modeling and Forecasting*, Addison-Wesley, Redwood City, CA, EUA, pp. 395–432.
- Weigend, A. S., Rumelhart, D. E. & Huberman, B. A. (1990). Predicting the future: A connectionist approach, *International Journal Neural System* **1**: 193–209.

- Weigend, A. S., Rumelhart, D. E. & Huberman, B. A. (1991a). Generalization by weight-elimination application to forecasting, *Advances in Neural Information Processing Systems* 3, pp. 875–882.
- Werbos, P. (1974). *Beyond regression: New tools for prediction and analysis in the behavioral sciences*, PhD thesis, Harvard University, Cambridge, MA, EUA.
- Wilks, D. (1999). Simultaneous stochastic simulation of daily precipitation, temperature and solar radiation at multiple sites in complex terrain, *Agricultural and Forest Meteorology* 96: 85–101.
- Yager, R. & Filev, D. (1994). *Essentials of Fuzzy Modeling and Control*, Wiley Interscience, New York, EUA.
- Zambelli, M., Luna, I. & Soares, S. (2009). Long-term hydropower scheduling based on deterministic nonlinear optimization and annual inflow forecasting models, *IEEE Bucharest PowerTech* (4): 1–8.
- Zimmermann, H. (1987). *Fuzzy Sets, Decision Making and Expert Systems*, Kluwer-Nihoft, Boston.