

## Redes Compostas por Blocos de Regressões Sigmóides Não-lineares: Uma eficiente rede de alta ordem com aplicações na previsão de séries temporais

Mêuser Valença<sup>1</sup>, Teresa Ludermir<sup>2</sup>

<sup>1</sup>Chesf/UNIVERSO-Recife

Rua Gaspar Peres 427/104, CEP 50670-350, CDU

<sup>2</sup>Universidade Federal de Pernambuco

Cx. Postal 7851 CEP 50732-970

E-mails: meuser@elogica.com.br, tbl@cin.ufpe.br

### Abstract

*This paper developed a new class of higher order feedforward neural networks, the Non-linear Sigmoidal Regression Blocks Networks (NSRBN), which is able to approximate any continuous function defined over a compact set. These networks are constructive since they deploy an constructive incremental learning algorithm. This algorithm is responsible for the selection of the optimal complexity model, that is, for the network's architecture definition.*

*Simulation results of forecasting was emphasized, specially in the area of water resources, which was the main motivation of this paper.*

### 1. Introdução

Aplicações realizadas com redes neurais artificiais têm apresentado desempenho satisfatório em diversas áreas de pesquisa, tais como: classificação, memória associativa, processamento de séries temporais [1] [2] [3].

Atualmente, no sistema elétrico brasileiro, as previsões de vazões e outras séries temporais, como consumo e demanda de energia são realizadas com modelos estatísticos lineares, especialmente os modelos de Box-Jenkins [4]. Todo planejamento e operação do sistema eletroenergético baseiam-se nessas previsões, daí sua importância.

O principal objetivo deste trabalho é analisar e desenvolver modelos de redes neurais, através do uso de um algoritmo de abordagem construtiva, de tal forma que usuários (não especialistas em redes neurais) possam fazer uso dessa poderosa ferramenta sem ter que se preocupar com a definição da arquitetura da rede neural, o que nem sempre é uma tarefa fácil. Como maior contribuição deste trabalho, apresenta-se uma nova classe de redes neurais, chamadas de redes com blocos de regressões sigmóides não-lineares (NSRBN) [5], que se constitui num aproximador universal de função e de arquitetura. Elas podem ser vistas como um caso de mapeamento polinomial multivariado, tendo

como base o teorema de aproximação de Weierstrass [5] e a funcionalidade da função sigmóide.

As redes NSRBN são eficientes, uma vez que elas utilizam blocos de sigmóides não-lineares como função de ativação. Este fato é contrário às redes neurais de alta ordem que tomam como base polinômios multivariados e apresentam uma explosão no número de parâmetros livres necessários para aproximar uma dada função.

Com o objetivo de avaliar a metodologia proposta, várias aplicações foram realizadas; tendo-se procurado abranger uma grande variedade de dados. Para cada uma das aplicações realizadas, fez-se uma análise comparativa com outras técnicas em uso.

A seguir na seção 2, introduz-se a metodologia das redes NSRBN e seu algoritmo de aprendizagem. Resultados das Aplicações são descritos na seção 3. Finalmente, na seção 4, as conclusões são apresentadas.

### 2. Rede Neural Construtiva - NSRBN

A arquitetura de uma rede neural, de acordo com a possibilidade de mudanças no seu tamanho durante o processo de treinamento, pode ser classificada em estática ou dinâmica. Uma rede é dita de arquitetura estática quando o número de camadas e unidades de processamento permanecem constante desde a concepção da rede até a finalização do seu treinamento. Por outro lado, uma rede de arquitetura dinâmica tem como característica principal um processo de mudanças no tamanho durante o treinamento.

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

A proposta do modelo NSRBN é de se construir a rede por blocos de polinômios homogêneos, utilizando uma função de ativação adequada (tangente hiperbólica, por exemplo), para os neurônios escondidos. Quanto aos neurônios de saída, estes têm uma ativação mais complexa, uma vez que realizam uma regressão logística não-linear com relação à saída dos neurônios escondidos. Desta forma, essas redes

NSRBN podem aproximar uniformemente qualquer função contínua definida em um conjunto compacto limitado e também lidar com a realização de funções *booleanas*, o que não é possível com os algoritmos tradicionais do grupo GMDH – Group Method Of Data Handling (combinatorial e *multilayer*) [5].

O primeiro passo para a construção de um modelo combinatorial consiste em dividir a função  $f(x)$  em blocos compostos de termos de igual grau, como apresentado na Figura 2.1.

$$f(x) = \tilde{f}_1(x) + \tilde{f}_2(x) + \dots + \tilde{f}_d(x) \quad (2.0)$$

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

$$f_p(x) = \sum_{i_1=1}^N \sum_{i_2=1}^N \dots \sum_{i_N=1}^N a_p(i_1, i_2, i_3, \dots, i_N) \cdot x_1^{q_{p1}(i_1)} \cdot x_2^{q_{p2}(i_2)} \cdot \dots \cdot x_N^{q_{pN}(i_N)} \quad (2.1)$$

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

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

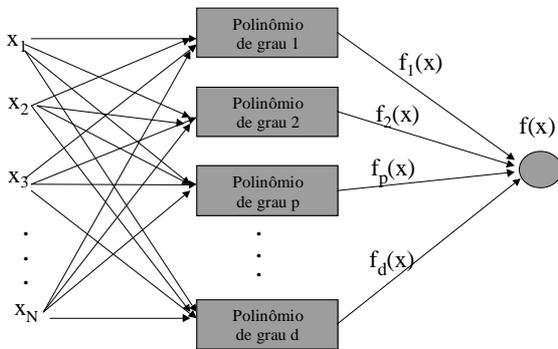


Figura 2.1 – Arquitetura combinatorial

Uma forma simples seria realizar todos os produtos das variáveis de entrada procurando aproximar diretamente a hipersuperfície não-linear, entretanto isto não é prático, a menos que  $d$  e  $N$  sejam relativamente pequenos (modelo combinatorial). Logo, para se ter um modelo eficiente, faz-se necessário que ele possua funções de ativação que sejam capazes de realizar vários termos dos polinômios homogêneos de forma otimizada, isto é, sem que ocorra um crescimento explosivo no número de parâmetros necessários para ajustar a função desejada. Com esse objetivo, propõe-se uma nova classe de redes neurais que tenha por base o algoritmo combinatorial (composta pela soma de

blocos) com uma estrutura de blocos similar a uma rede MLP, sendo que, diferentemente destas, os neurônios de saída realizam a técnica estatística de regressão logística não-linear [6], com relação aos neurônios escondidos.

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

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

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

$$f(x) = (\sigma_{net(h)} + \theta_h) \quad (2.2)$$

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

$$\sigma_{net(h)} = \frac{e^{net(h)} - e^{-net(h)}}{e^{net(h)} + e^{-net(h)}} \quad (2.3)$$

Poder-se-ia utilizar outras funções, como é o caso da função recíproca representada na Equação 2.5, isto é,

$$f_r(x) = \frac{net(h)}{1 + |net(h)|} \quad (2.4)$$

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

A estrutura proposta para cada bloco é similar à da Figura 2.2, na qual apresenta-se o mapeamento para um bloco de grau  $p$  (a ordem dos neurônios escondidos  $h$  variará de 1 até  $p$ ), para  $i$  variáveis de entrada (onde  $i=1,\dots,N$ ) e uma única unidade de saída.

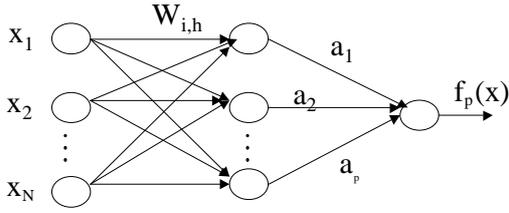


Figura 2.2 – Arquitetura de um bloco polinomial

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

Seja  $\mathbf{x} = [1, x_1, x_2, x_3, \dots, x_N]^T$  o vetor coluna das variáveis de entrada, no qual o elemento  $x_i$  denota um elemento qualquer dele. As entradas são ponderadas pelos vetores de pesos  $\mathbf{w} = [w_{0h}, w_{1h}, w_{2h}, \dots, w_{Nh}]^T$ ,  $h=1, \dots, p$ , nos quais  $p$  é a ordem do bloco. Portanto, utilizando-se a função de ativação estabelecida na Equação 2.2, a saída de cada unidade escondida,  $f(\text{net}(h))$ , desse bloco de grau  $p$  pode ser escrito como:

$$f(\text{net}(h)) = (\sigma_{\text{net}(h)} + \theta_h) \quad (2.5)$$

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

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

$$f_p(x) = a_1(f(\text{net}(1)))^1 + a_2(f(\text{net}(2)))^2 + a_3(f(\text{net}(3)))^3 + \dots + a_p(f(\text{net}(p)))^p \quad (2.6)$$

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

Fazendo-se uso da Equação 2.5, a função  $f_p(x)$  pode ser escrita da seguinte forma:

$$f_p(x) = a_1(\sigma_{\text{net}(1)} + \theta_1) + a_2(\sigma_{\text{net}(2)} + \theta_2)^2 + \dots + a_p(\sigma_{\text{net}(p)} + \theta_p)^p \quad (2.7)$$

ou

$$f_p(x) = \sum_{h=1}^p a_h (\sigma_{\text{net}(h)} + \theta_h)^h \quad (2.8)$$

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

De acordo com a Figura 2.1, a arquitetura de uma rede NSRBN é composta pela soma de blocos com grau variando de 1 a  $d$ . Aqui, está sendo utilizado o bloco  $p$  (Figura 2.2) para representar qualquer um desses blocos.

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

$$f(x) = \sigma_{\text{net}(o)} \left( \sum_{p=1}^d f_p(x) \right) \quad (2.9)$$

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

Por outra lado, a construção das redes NSRBN é feita de forma incremental, de maneira que pode-se representá-la em uma forma similar a de um bloco (Equação 2.7), conforme apresenta-se a seguir. Considere o passo do algoritmo para o qual se adiciona-se à rede o bloco de ordem  $d$ ,  $f_d$ , tendo-se anteriormente adicionado os blocos  $f_0, f_1, \dots, f_{d-1}$ . Fazendo-se  $f_0 \equiv 0$ , uma função  $f$  pode ser representada por:

$$f(x) = \sigma_{\text{net}(o)} \left( \sum_{p=0}^{d-1} f_p(x) + f_d(x) \right) \quad (2.10)$$

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

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

$$f(x) = \sigma_{\text{net}(o)} \{ c_0 + c_1(\sigma_{\text{net}(1)} + \theta_1) + c_2(\sigma_{\text{net}(2)} + \theta_2)^2 + \dots + c_d(\sigma_{\text{net}(d)} + \theta_d)^d \} \quad (2.11)$$

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

A Equação 2.8 é, portanto, idêntica a uma regressão logística não-linear quando a função de ativação utilizada é a sigmóide logística tradicional.

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

## 2.1. Algoritmo Construtivo para as NSRBN

Esse algoritmo construtivo tem como base um método de aprendizagem construtiva (também chamado de auto-organização na teoria das redes GMDH). O princípio construtivo pode ser formulado como segue: quando um modelo incrementa gradualmente sua complexidade, certos critérios, que são chamados de

critérios de seleção ou função objetivo e que têm a propriedade de complemento externo, passam através de um valor mínimo global. O encontro deste mínimo indicará a existência de um modelo de complexidade ótima.

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

$$f(x) = g(\hat{f}_1(x) + \hat{f}_2(x) + \dots + \hat{f}_p(x) + \dots + \hat{f}_d(x)) \quad (2.12)$$

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

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

$$f(x) = g(f_1(x)); \quad (2.13)$$

$$f(x) = g(\hat{f}_1(x) + f_2(x)); \quad (2.13)$$

...

$$f(x) = g(\hat{f}_1(x) + \hat{f}_2(x) + \dots + f_p(x)) \quad (2.14)$$

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

### 3. Aplicações

#### 3.1. Predição de vazões semanais

As empresas geradoras de energia elétrica realizam mensalmente a elaboração do Programa Mensal de Operação (PMO) que define a geração de cada unidade da empresa e a comercialização de intercâmbio de energia entre elas.

Uma condição de fundamental importância para que o atendimento à demanda de energia seja realizada de forma otimizada é que se disponha de um sistema de previsão de vazões eficiente, uma vez que a capacidade de geração futura do sistema é influenciada pelas futuras vazões afluentes cuja natureza é intrinsecamente estocástica

Portanto, sendo o setor elétrico possuidor de grandes reservatórios de regularização, o seu plano de operação pode ter grandes benefícios com a utilização de novos modelos de previsão hidrológica que venham a melhorar a qualidade das estimativas realizadas.

A avaliação da viabilidade do uso de Redes Neurais na previsão de vazões semanais concentrar-se-á em ajustar os 52 melhores modelos auto-regressivo média-móvel periódico – PARMA(p,q) e comparar os resultados com aqueles obtidos para outros modelos de Redes Neurais – Redes *Perceptron* de Múltiplas Camadas, Redes Recorrentes de Elman e Jordan, Redes GMDH e as Redes NSRBN. O uso de vazões médias semanais deve-se ao fato de que já se dispunha de resultados de estudos anteriormente realizados por Valença e Iudermir, 1998 [7] cujo trabalho mostrou que as redes *Perceptron* Multicamadas (MLP) apresentaram melhores resultados do que os modelos de Box-Jenkins, para aquele conjunto de dados analisados.

A estrutura experimental para o treinamento e teste dos modelos envolveu a série de vazões médias semanais incrementais afluentes ao reservatório de Sobradinho. Os valores referentes aos anos de 1929 a 1988 foram empregados na calibração dos modelos e no treinamento da rede neural. O conjunto desses dados corresponde, portanto, a um total de 3120 valores sendo 52 para cada ano (o ano foi dividido em 52 semanas) No período de 1989 a 1997, foi utilizado para avaliar o desempenho dos modelos e da rede neural (totalizando um conjunto de 449 valores semanais).

Como elemento de avaliação do desempenho dos modelos ARMA(p,q) periódicos e os modelos de rede neural foram utilizados o erro médio absoluto (EMA) definido na Equação 3.1; o erro relativo percentual médio absoluto (EPMA), definido na Equação 3.2; e o erro padrão de predição (EP), definido na Equação 3.3.

$$EMA = \frac{1}{N} \left[ \sum_{p=1}^N |Z_p - Z_o| \right] \quad (3.1)$$

$$EPMA = \frac{1}{N} \left[ \sum_{p=1}^N |Z_p - Z_o| / Z_o \right] \cdot 100 \quad (3.2)$$

$$EP = \left[ \frac{1}{N} \sum_{p=1}^N (Z_p - Z_o)^2 \right]^{0.5} \quad (3.3)$$

em que  $Z_p$  - é o valor previsto;  $Z_o$  - é o valor ocorrido e  $N$  - é o número de valores de verificação.

Foram inicialmente realizados inúmeros testes com redes MLP variando-se o número de camadas intermediárias e o número de elementos de processamento nessas camadas. Um resumo dos melhores resultados obtidos de forma que se permita uma análise comparativa será apresentado.

A janela de entrada nos primeiros testes foi de 11 elementos de processamento: sendo 5 valores passados das variáveis e 6 elementos de processamento para caracterizar a sazonalidade (6 bits combinados).

Nesses testes, efetuou-se também uma análise da sensibilidade quanto à mudança na taxa de aprendizagem e no termo momento, verificando-se a variação da taxa de momento entre 0,1 e 0,001 com o termo momento fixo 0,1. Não foi obtido nenhum ganho de precisão, mas alterava significativamente o tempo de aprendizagem, ou seja, um valor para a taxa de aprendizagem pequeno 0,001 duplicava o tempo de treinamento.

A janela de entrada foi em seguida considerada com 5 elementos de processamento, sendo 5 valores passados da variável, uma vez que se passou a trabalhar com a série padronizada, procurando se representar melhor a sazonalidade. Esta janela foi estabelecida através de uma análise das funções de autocorrelação e autocorrelação parcial [1][5], analisadas e também em função da experiência adquirida com o comportamento do sistema.

As redes Recorrentes foram então utilizadas com a série padronizada.

Dentre os vários algoritmos do tipo GMDH testados as Redes GMDH do tipo multiplicativo-aditivo com neurônio ativo [1][5], foi a que forneceu melhor desempenho. As Redes GMDH tem como elemento de processamento a ser fornecido apenas a janela de entrada que, nesse caso, é de 5 valores passados.

Nos vários testes realizados, com o algoritmo multiplicativo-aditivo com neurônio ativo foram basicamente testados: o melhor número de variáveis de entrada a serem combinadas para gerar um sobrevivente, o número de sobreviventes nas camadas construídas e o critério de seleção.

As Redes NSRBN tem como elemento de processamento a ser fornecido apenas a janela de entrada que, nesse caso, é de 5 valores passados. Nos vários testes realizados foram basicamente testados: a sensibilidade aos pesos iniciais e os critérios de seleção.

Os resultados que serão apresentados utilizaram o critério de mínima tendência para treinamento e melhor predição para a escolha da complexidade ótima.

Para esse caso, a rede NSRBN apresenta excelente resultado com apenas 2 blocos, isto é, 3 unidades escondidas são suficientes para realizar a aproximação desejada.

Na Tabela 3.1, vê-se um resumo dos melhores resultados conseguidos para a previsão de vazões.

Tabela 3.1 – Resultados com os diversos modelos

Modelos De Previsão para 1 semana à frente	EPMA (%)	EMA (m <sup>3</sup> /s)	EP (m <sup>3</sup> /s)
MLP	13,4	292	509
Elman	14,0	302	535
Jordan	13,4	292	508
GMDH	13,3	287	492
Box-Jenkins	14,1	304	518
<b>NSRBN</b>	<b>13,0</b>	<b>283</b>	<b>482</b>

Na Figura 3.1 fornece-se um zoom para o ano de 1992 (2<sup>a</sup> maior cheia do histórico).

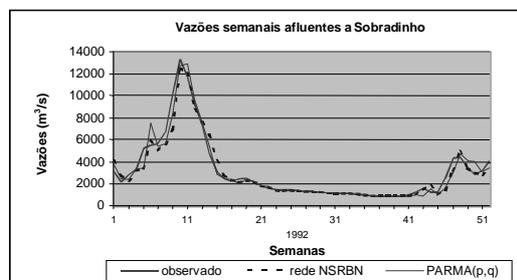


Figura 3.1 – Comparação entre a Rede NSRBN e o ARMA(p,q), periódico (ano de 1992)

### 3.2. Predição na área de demanda de energia.

A previsão de demanda de energia máxima mensal é de fundamental importância para uma econômica e segura operação de um sistema de geração de energia hidroelétrica. Funções básicas de operação tais como: configuração mais econômica para suprir uma determinada demanda, segura avaliação do sistema elétrico e entrada de novas obras (subestações, linhas de transmissão), são alguns dos fatores que requerem uma boa previsão de carga.

A importância dessa aplicação pode ser vista sob dois enfoques: o primeiro está relacionado com a aplicação da técnica NSRBN num problema de operação real, no qual os resultados previstos estão sendo comparados com os ocorridos, o segundo está ligado à inovação da técnica utilizada, ou seja uma rede NSRBN, que permite a seleção de um modelo de complexidade ótima. Valença, 1997[8] e Valença e Ludermir, 1998 [9] apresentam diversas aplicações com redes neurais para subestações da Chesf.

Para a subestação de Fortaleza, foi considerada a série de demanda máxima de energia mensal. Os valores referentes aos anos de 1990 a 1996 foram empregados na calibração do modelo de Box-Jenkins e

no treinamento e teste das redes neurais, correspondendo a 84 valores. O período de janeiro 1997 a setembro de 1997 foi utilizado para avaliar o desempenho, perfazendo um conjunto de 9 valores.

Como elemento de avaliação dos modelos foram utilizados o erro médio absoluto (EMA), definido na Equação 3.1; o erro relativo percentual médio absoluto (EPMA), definido na Equação 3.2; e o erro padrão de predição (EP), definido na Equação 3.3.

**Predição baseada nas redes NSRBN** - As redes NSRBN apresentam bons resultados com apenas dois blocos (quando se considera três valores passados como janela de entrada), para as três subestações.

**Predição baseada nas redes MLP** - A rede MLP usada para previsão das séries temporais apresenta uma estrutura com uma única camada escondida com 6 neurônios (função de ativação: tangente hiperbólica).

**Predição baseada nas redes GMDH com neurônio ativo** - Com os três valores passados como entrada, a rede GMDH converge na segunda camada [9].

A Tabela 3.2 resume os resultados conseguidos, utilizando-se as redes e o modelo de Box -Jenkins.

Tabela 3.2 – Demanda máxima de energia

Erros	NSRBN	MLP	Box Jenkins	GMDH
EPMA	1,43	1,60	3,20	1,44
EMA	8,60	8,62	9,71	8,63
EP	10,18	10,31	11,30	10,20

#### 4. Conclusões

Neste trabalho, vários modelos para predição de séries temporais e classificação foram apresentados. Mais precisamente, procurou-se abranger desde modelos estatísticos tradicionais e redes neurais tradicionais (redes com abordagem dedutiva) até redes neurais construtivas, tendo-se apresentado uma descrição e aplicação das mesmas.

A principal contribuição deste trabalho consiste no desenvolvimento das redes NSRBN. Estas redes são simples por possuírem uma única camada escondida, entretanto apresentam uma ótima capacidade de aproximação de funções por realizarem uma regressão sigmóide não-linear nas unidades de saída.

O objetivo das redes desenvolvidas neste trabalho foi pois de reduzir o envolvimento humano no processo, especialmente o envolvimento do usuário final, uma vez que o elemento humano muitas vezes é um fator responsável pela introdução de erros e decisões indesejáveis.

As redes NSRBN aqui desenvolvidas foram implantadas e testadas com essa finalidade, isto é, eliminar a participação do usuário final no estabelecimento da estrutura e escolha da melhor rede.

No que diz respeito às previsões de vazões, os resultados comprovam a eficácia das redes neurais, especialmente das redes NSRBN, tendo todos os tipos de redes consideradas aqui apresentado resultados superiores aos modelos estocásticos para as unidades métricas de erro utilizadas para avaliar o seu desempenho

Área de demanda de energia - os resultados encontrados nessa aplicação também foram significativos e comparáveis aos da área de recursos hídricos.

#### Referências

- [1] M. J. S. Valença and T. B. Ludermir. Multiplicative-Additive Neural Networks with Active Neurons. International Joint Conference on Neural Networks (IJCNN), IEEE, Book of Summaries (2073), Washington, DC, July 1999.
- [2] M. J. S. Valença and T. B. Ludermir. Self-organization Sigmoidal Blocks Networks. International Joint Conference on Neural Networks (IJCNN), IEEE, Book of Summaries (2086), Washington, DC, July 1999.
- [3] M. J. S. Valença and T. B. Ludermir. Uma Rede Neural Construtiva com Atualização dinâmica dos Pesos. IV Congresso Brasileiro de Redes Neurais, pages 114-117, ITA, São José dos Campos, Julho 1999.
- [4] G.E.P Box & G.M. Jenkins, Time Series Analysis - Forecasting and Control. Holden-Day Inc, California, 1976.
- [5] M. J. S. Valença. Analysis and Design of the constructive neural networks for complex systems modeling (in portuguese). Ph.D These, UFPE, Brazil, 1999.
- [6] B. J. T. Morgan. The cubic logistic model for quantal assay data. Applied Statistics, 34: 105-113.
- [7] M. J. S. Valença and T. B. Ludermir. Uma nova rede neural polinomial com aplicação na previsão de vazões. V Simpósio Brasileiro de Redes Neurais, . 2: 273-278, Belo Horizonte 1998.
- [8] M. J. S. Valença. Redes Neurais na Previsão de demanda Máxima Mensal. Third Latin-American Congress: Electricity Generation and Transmission. Campos do Jordão São Paulo, pages 346-351, 1997.
- [9] M. J. S. Valença and T. B. Ludermir. Previsão de demanda máxima mensal utilizando um modelo auto-organizável. V Simpósio Brasileiro de Redes Neurais, 2: 311-314, Belo Horizonte 1998.