



SPOLM 2007

ISSN 2175-6295

Rio de Janeiro- Brasil, 08 e 09 novembro de 2007.

PREVISÃO DE DEMANDA DE ENERGIA ELÉTRICA NO BRASIL UTILIZANDO REDES NEURAIS DE ELMAN

Fernanda Fidelis Paschoalino

Universidade Federal Fluminense
Rua Passo da Pátria 156, 22210-240, Niterói, RJ
fernandapaschoalino@gmail.com

Tabita Yaling Cheng Loureiro

Universidade Federal Fluminense
Rua Passo da Pátria 156, 22210-240, Niterói, RJ
tabis@globo.com

João Carlos Correia Baptista Soares de Mello

Universidade Federal Fluminense
Rua Passo da Pátria 156, 22210-240, Niterói, RJ
jcsmello@producao.uff.br

Luiz Biondi Neto

Universidade do Estado do Rio de Janeiro
Rua São Francisco Xavier 524, sala 5019B, 20550-013, Rio de Janeiro, RJ
lbiondi@uerj.br

Resumo

Este estudo aplica a técnica de inteligência computacional conhecida como Rede Neural Artificial em modelos de previsão do consumo de energia elétrica no Brasil. A modelagem utilizada foi Rede Neural de Elman associada a um algoritmo de aprendizado supervisionado (backpropagation). O principal objetivo foi analisar a performance de aprendizado da RNA frente ao comportamento não-linear e não-sazonal das séries históricas. A base de dados contemplou o consumo mensal de energia elétrica referente ao período de 1982 a 2003. Em um primeiro momento, os padrões de entrada foram os dados de 1982 a 2002, devidamente pré-processados. O ano de 2003 não fez parte do treinamento e serviu como referência para investigar a previsão com o objetivo de testar a generalização da rede e validar os resultados. No segundo momento, o ano de 2001 foi escolhido para validação dos resultados devido ao uso racional da energia nesse período. Por fim, foram calculados os erros médios quadráticos de cada modelo investigado, alcançando resultados na ordem de 0,02%.

Palavras-Chaves: Energia Elétrica; Redes Neurais de Elman; Redes Recorrentes; Previsão de Séries Temporais; Consumo de Energia Elétrica.

Abstract

This study applies computational intelligence as Neural Network to the Electric Power Consumption in Brazil. It was used the Elman network architecture based on backpropagation. The main purpose of this study was testing how Elman networks were capable to deal with non linearity and non stationarity in time series. The database comprises a period of monthly electric power consumption from 1982 to 2003. There were structured

two models. First, the patterns used as inputs for training the neural networks are correspondent to the period from 1982 to 2002. Since 2003 was not included in training, it was adopted as a parameter to generalization test and validation. In the following step, it was chosen the year of 2001, according to its particular demand - drop in the electric power consumption. At last, there were calculated the mean average errors in each model, which achieved results as 0.02%.

Keywords: Electric Power; Elman Neural Networks, Recurrent Neural Networks; Times Series Forecast; Electric Power Consumption.

1. Introdução

Este estudo propõe ao emprego da técnica de inteligência computacional conhecida como Rede Neural Artificial para previsão do consumo de energia elétrica no Brasil. A modelagem utilizada foi a Rede Neural de Elman (RNA) associada a um algoritmo de aprendizado supervisionado (backpropagation) na previsão de séries temporais do consumo de energia elétrica.

De acordo com Biondi Neto et al, 2005, as principais características das RNA's na análise de séries temporais são a capacidade de aprendizado na realização de previsões precisas; flexibilidade para previsões iterativas; capacidade de lidar com a não-estacionariedade e a não-linearidade das séries temporais investigadas e a possibilidade de oferecer predição paramétrica e não-paramétrica.

A base de dados para o presente estudo contempla o consumo de energia elétrica no Brasil referente ao período de 1982 a 2003, totalizando 264 elementos, deslocados mês a mês. Nas camadas de contexto, camadas ocultas que simulam a memória da rede, foi estruturado um atraso para que a rede recebesse os dados processados em períodos de um ano, defasados mês a mês.

No primeiro modelo, os padrões de entrada da rede foram os dados referentes ao período de 1982 até 2002. O ano de 2003 não fez parte do treinamento e pôde servir como referência para testar a generalização da rede e validar os resultados. Uma outra modelagem foi proposta utilizando o ano de 2001 como referência para validar os resultados e investigar como a rede neural de Elman iria se comportar frente a um outlier, já que esse foi o ano em que se intensificou o uso racional da energia no país.

Na fase de treinamento, diversos parâmetros foram balanceados. Alguns deles foram: o número de camadas escondidas; o número de neurônios artificiais por camada e o tipo de função de ativação.

A rede recorrente, característica da rede de Elman, pode apresentar uma ou mais malhas de realimentação, fazendo com que suas conexões permitam à rede memorizar as informações atuais na camada de contexto no momento seguinte à sequência investigada. Portanto essa estrutura é a mais indicada na aplicação de modelos de previsão de séries temporais. A arquitetura empregada na RNA baseou-se em uma rede parcialmente recorrente, onde a implementação foi feita usando-se o software MATLAB® versão 7.0.

Para este estudo, os dados do consumo de energia elétrica no Brasil abrangem o somatório de toda a composição setorial do consumo de eletricidade do país e foram obtidos do IPEADATA (Instituto de Pesquisa Econômica e Aplicada).

2. Características das Redes Neurais

As Redes Neurais Artificiais se destacam por apresentarem características semelhantes ao funcionamento do cérebro humano como capacidade de aprendizado, generalização, associação, procura paralela e endereçamento pelo conteúdo. Algumas áreas de aplicação são: previsão de séries temporais; controle de processos; robótica; classificação de dados; reconhecimento de padrões em linhas de montagem; filtros contra ruídos eletrônicos; análise de imagens, avaliação de riscos de financiamento, detecção de bombas, identificação de reservas de petróleo etc. (Braga et. al., 1998).

Uma Rede Neural Artificial é um modelo computacional que compartilha algumas propriedades do cérebro humano. Consiste em estruturas paralelas, baseadas em unidades processadoras bem simples (modelo artificial do neurônio) e densamente interconectadas.

As Redes Neurais Artificiais são semelhantes ao cérebro humano em dois aspectos basicamente: a experiência é a fonte do conhecimento adquirido e o conhecimento adquirido é armazenado nas sinapses. (Haykin, 2001)

O neurônio, célula constituinte do cérebro, é composto por um corpo celular, axônios e dendritos. O axônio, uma fibra nervosa de superfície lisa com poucas ramificações e maior comprimento é responsável pela transmissão da informação na comunicação com outros neurônios. Os dendritos, que têm aparência de árvores, possuem superfície irregular e muitas ramificações e atuam como receptores desta comunicação. A Figura 2.1 ilustra a estrutura de um neurônio e seus principais componentes.

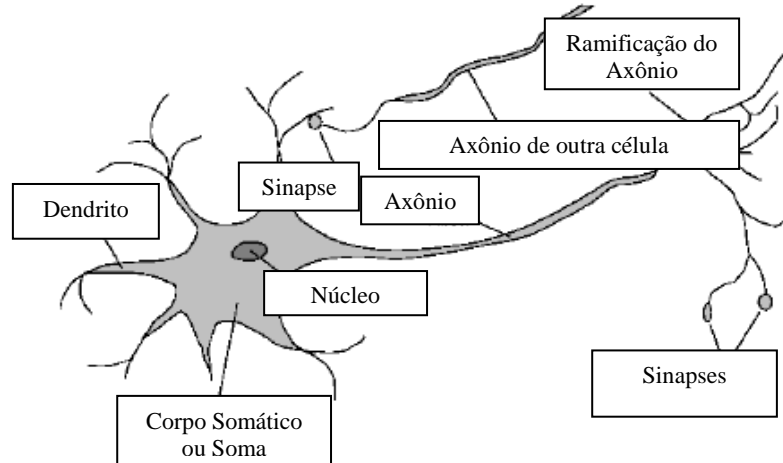


Figura 2.1 – Neurônio Biológico
(Fonte: Biondi Neto et al, 2005)

A transferência de informação entre neurônios biológicos é denominada sinapse, que se dá através de um processo químico no qual são liberadas substâncias transmissoras que se difundem pela junção sináptica entre neurônios, causando aumento ou queda do potencial elétrico do neurônio receptor. Pelos dendritos, o neurônio recebe sinais de outros neurônios a ele conectados, os quais são acumulados no corpo do neurônio. Quando a soma dos sinais ultrapassa um determinado limiar, um sinal é propagado no axônio para outros neurônios.

Paralelamente, a transferência de informação entre neurônios artificiais é semelhante à sinapse, que se dá através de uma função de ativação, simulando o processo químico no qual são liberadas substâncias transmissoras que são difundidas entre neurônios. No neurônio artificial, os pesos relacionados às entradas representam o processo que causa aumento ou queda do potencial elétrico do neurônio receptor. (Haykin, 2001). Esse processo pode ser visualizado na Figura 2.2 que traz uma comparação entre o neurônio biológico e o neurônio artificial.

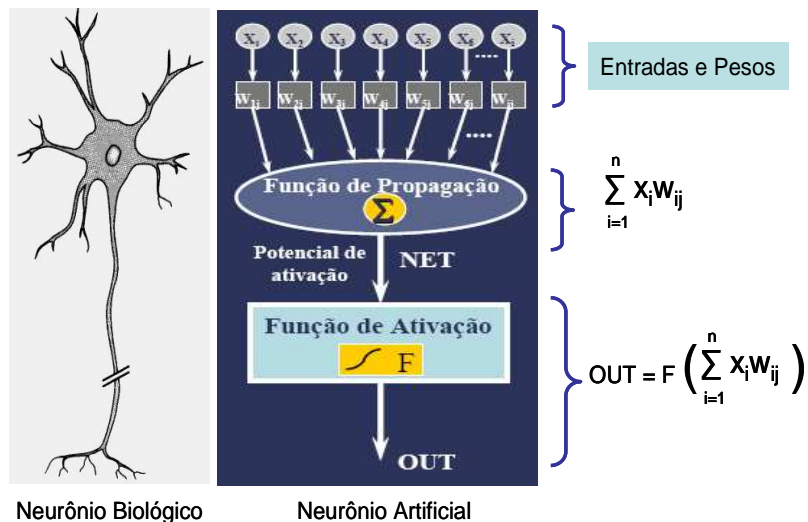


Figura 2.2 – Neurônio Biológico Versus Neurônio Artificial
(Adaptado de Biondi Neto et al, 2005)

2.1 O processamento da Rede Neural

As RNA's, diferentemente da programação tradicional, não têm execução direta, apresentando por característica duas fases distintas de processamento.

A primeira fase é conhecida por treinamento, sendo caracterizada pela apresentação de padrões de entrada com objetivo de mapear o conhecimento que a rede deve aprender. É nesta fase que os padrões são apresentados separadamente, em lotes, e o erro é medido pela diferença entre a saída obtida a cada iteração do treinamento e o alvo, que é o resultado que se deseja obter. Quando o erro atinge um valor satisfatório, o processo de treinamento é encerrado e o conhecimento obtido é armazenado em um arquivo sob forma de uma matriz numérica, conhecida também por matriz de pesos sinápticos.

Depois de treinada, a rede passa para a segunda fase, que é a fase de execução. Nesta etapa do processamento, são apresentados vetores de entrada à rede, também sob forma numérica, que normalmente não fizeram parte do treinamento. Dessa forma, o valor da função soma é o argumento de uma segunda função, denominada função de ativação. O resultado desta função é a saída processada por aquele respectivo neurônio.

2.2 A Função de Ativação

Cada neurônio realiza um processamento simples: recebe uma entrada e computa um novo nível de ativação (Russel e Norvig, 1995).

Este processamento é composto por duas etapas: na primeira, cada entrada x_i do neurônio é multiplicada pelo peso sináptico correspondente w_{ji} (peso da entrada i do neurônio j). O resultado de cada multiplicação é então somado. Na segunda etapa a soma é aplicada a uma função de ativação F , obtendo-se a saída do neurônio (Equação 2.2.1):

$$OUT = F\left(\sum_{i=1}^n x_i w_{ij}\right) \quad (2.2.1)$$

A função de ativação deve simular as características não lineares do neurônio biológico. As funções mais utilizadas são:

- **Função Linear**

É uma equação linear da forma $F(x) = ax$, representada no gráfico 2.2.1:

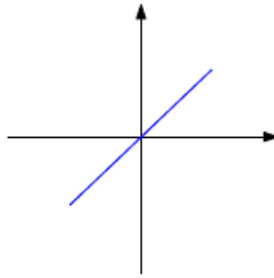


Gráfico 2.2.1 – Exemplo de Função Linear

▪ **Função Degrau**

É uma equação utilizada para valores binários e é da forma representada no gráfico 2.2.2:

$$F(x) = \begin{cases} 1, & \text{se } x > 0 \\ 0, & \text{se } x \leq 0 \end{cases},$$

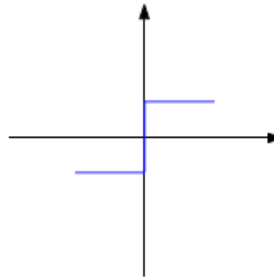


Gráfico 2.2.2 – Exemplo de Função Degrau

▪ **Função Sigmóide**

Também chamada de função logística, é uma função contínua que permite a transição gradual entre os dois estados (varia entre 0 e 1).

É dada por $F(x) = \frac{1}{1 + e^{-x}}$, representada no gráfico 2.2.3:

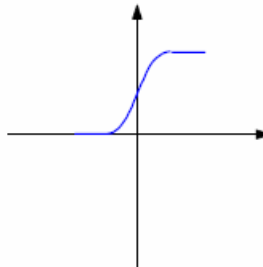


Gráfico 2.2.3 – Exemplo de Função Sigmóide

▪ **Função Tangente Hiperbólica**

É uma função sigmóide que varia entre -1 e 1.

É dada por $F(x) = \frac{1 - e^{-x}}{1 + e^{-x}}$, representada no gráfico 2.2.4:

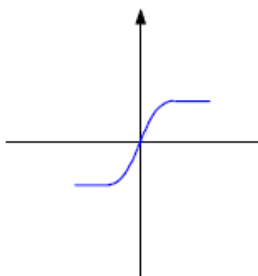


Gráfico 2.2.4 – Exemplo de Função Tangente Hiperbólica

2.3 A arquitetura das Redes Neurais Artificiais

As arquiteturas neurais são organizadas em camadas, em que as unidades processadoras, a começar pela camada de entrada, apresentam conexões com as unidades da camada seguinte. A Figura 2.3.1 representa uma rede multicamadas (Dhar & Stein, 1997).

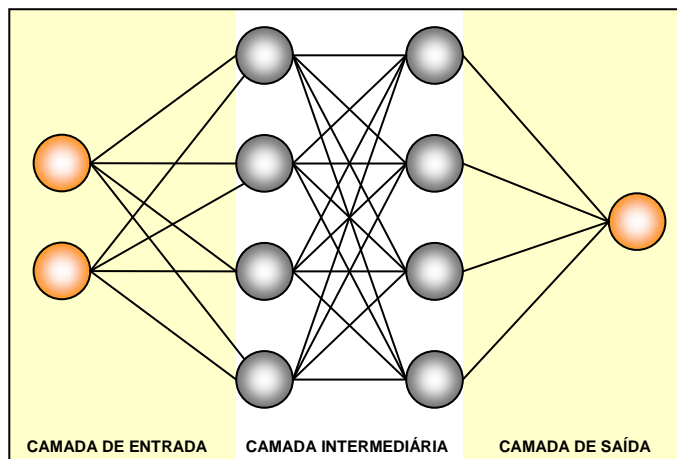


Figura 2.3.1 – Esquema da organização das camadas dos neurônios artificiais em uma rede neural (Fonte: DHAR & STEIN, 1997)

A organização das camadas pode ser classificada basicamente em três grupos:

- Camada de entrada, onde os padrões são apresentados à rede;
- Camada intermediária ou escondida, onde ocorre maior parte do processamento por meio de conexões ponderadas;
- Camada de saída, onde o resultado final é concluído e apresentado.

O comportamento inteligente de uma rede neural artificial vem das interações entre as unidades de processamento da rede. A maioria dos modelos de redes neurais possui alguma regra de treinamento, onde os pesos de suas conexões são ajustados de acordo com os padrões apresentados, ou seja, elas aprendem através de exemplos. (Yamamoto & Nikiforuk, 2000)

Uma rede neural é especificada, principalmente pela sua topologia, pelas características dos nós e pelas regras de treinamento. A seguir, serão analisados as características da topologia da RNA e os diferentes processos de aprendizado.

2.4 Topologia da Rede Neural

Há uma grande variedade de redes, cada uma produzindo diferentes resultados. Podem ser classificadas basicamente em redes alimentadas a frente e redes recorrentes. (Russell & Norvig, 1995)

2.4.1 Redes Alimentadas a Frente

Nas redes alimentadas a frente, também conhecidas por redes *Feed-forward*, o sentido de propagação dos sinais é único, sendo sempre da entrada para a saída da rede (Figura 2.4.1.1). Nessa topologia, os neurônios estão dispostos em camadas, podendo haver redes com uma única camada (ex. perceptron) ou com múltiplas camadas (ex. MLP – *multi layer perceptron*). No caso de redes multicamadas, as entradas de um neurônio são as saídas dos neurônios da camada anterior, portanto não há ligação entre neurônios de uma mesma camada.

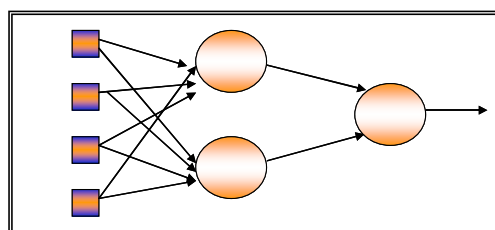


Figura 2.4.1.1 – Representação de Redes Alimentadas a Frente (Fonte: As autoras, 2007)

2.4.2 Redes Recorrentes

As redes recorrentes possuem laços de realimentação, ou seja, a saída de um neurônio pode ser entrada para outro de uma camada precedente ou, no caso de auto-realimentação, para o próprio neurônio (Figura 2.4.1.2). As redes recorrentes, chamadas de redes com memória, não possuem organização rígida e seus neurônios têm liberdade para se ligar a qualquer outro neurônio. (Abelem, 1994; Haykin, 2001; Russell & Norvig, 1995)
Exemplos: Redes de Elman; Redes de Hopfield; Redes de Jordan e NARX.

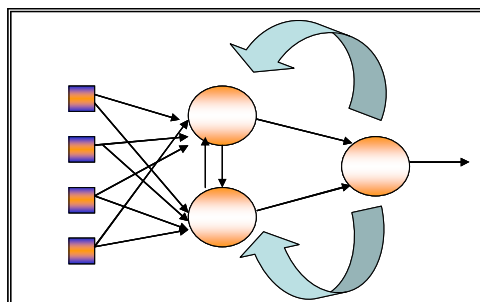


Figura 2.4.1.2 – Representação de Redes Recorrentes
(Fonte: As autoras, 2007)

2.5 Tipos de Aprendizado

De acordo com a apresentação de padrões de entrada à rede, o processo de aprendizagem ocorre através do ajuste dos parâmetros livres, pesos sinápticos, por estimulação do ambiente. Há dois tipos de aprendizado: supervisionado e não-supervisionado.

2.5.1 Aprendizado Não-Supervisionado

Depois de estabelecido um padrão de entrada, a rede permite que livremente escolha o padrão de saída a partir das regras de aprendizado adotadas. Não exige Matriz Saída Desejada (alvo). O sistema, auto-organizado, extrai as características do conjunto de padrões, agrupando-os em classes. Exemplo: Rede de Hamming e Rede de Kohonen. (Haykin, 2001). O aprendizado não-supervisionado pode ser representado na Figura 2.5.1.1.

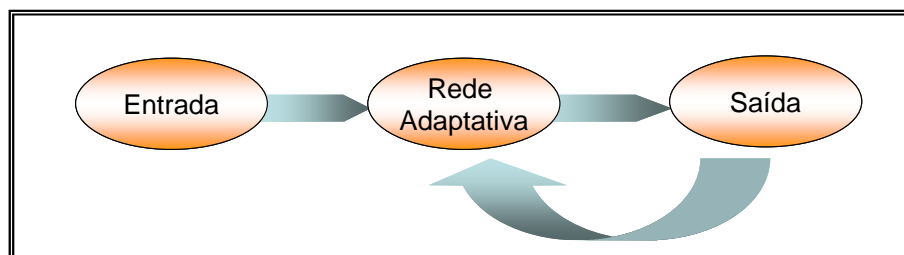


Figura 2.5.1.1 – Representação esquemática do Aprendizado Supervisionado
(Fonte: Biondi Neto et al, 2005)

2.5.2 Aprendizado Supervisionado

Esta forma de aprendizado é baseada em um conjunto de padrões de entrada-saída que é apresentado à rede. A partir da entrada, a rede realiza seu processamento e a saída obtida é comparada com a saída alvo. Caso não sejam iguais, é aplicado um processo de ajuste de pesos até obter um resultado satisfatório, ou seja, um erro aceitável (Figura 2.5.2.1). O algoritmo de aprendizado supervisionado mais comum é o Backpropagation. (Haykin, 2001)

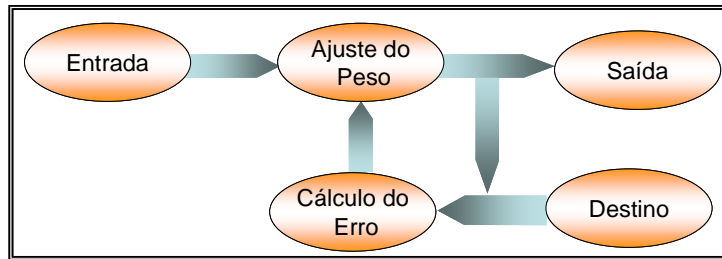


Figura 2.5.2.1 – Representação esquemática do Aprendizado Supervisionado
(Fonte: Biondi Neto et al, 2005)

3. Redes Neurais de Elman

A rede de Elman é uma estrutura de rede neural com características recorrentes (Elman, 1990), ou seja, pode apresentar uma ou mais malhas de realimentação, fazendo com que suas conexões permitam à rede memorizar as informações atuais na camada de contexto no momento seguinte à seqüência investigada. Por esse motivo, tal arquitetura é a mais indicada para aprender as características temporais de um sinal que represente a série histórica do consumo de energia elétrica no Brasil.

Neste estudo, a rede é parcialmente recorrente, pois os elos de realimentação ocorrem entre a saída e a entrada da primeira camada oculta (Figura 2.10). O primeiro padrão de entrada é processado como uma rede alimentada à frente, ou seja, o sinal propagado segue o sentido da entrada para a saída. Ainda na primeira iteração, os neurônios ocultos ativam os neurônios da camada de contexto e esses armazenarão a saída desta primeira iteração que será utilizada no próximo ciclo.

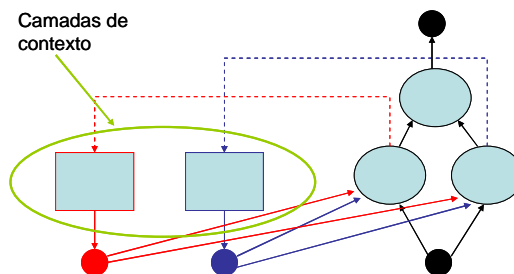


Figura 2.10 – Representação da primeira iteração
(Fonte: As autoras, 2007)

Na primeira iteração, a saída processada pelos neurônios da camada de contexto servirá de entrada para a próxima iteração (Figura 2.11).

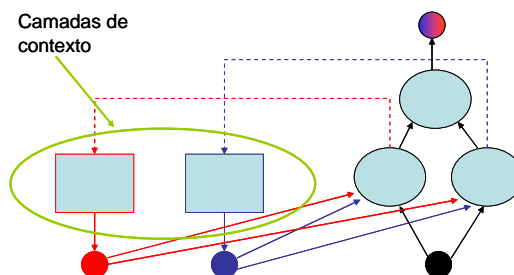


Figura 2.11 – Representação da segunda iteração
(Fonte: As autoras, 2007)

A recorrência da Rede de Elman ocorre através de uma unidade de contexto, normalmente uma estrutura de atraso do tipo z^{-1} , armazenando a saída da primeira camada oculta por um passo de tempo. Dessa forma, essa arquitetura permite a geração de padrões variáveis no tempo, habilitando essa configuração neural para aplicações envolvendo séries temporais.

Além da camada recorrente, a rede de Elman pode apresentar várias outras camadas formando uma tradicional MLP com uma ou mais saídas, sendo o algoritmo de treinamento da rede o *backpropagation*.

No caso da utilização em previsão de séries temporais é possível adotar soluções híbridas, fazendo-se com que as entradas sejam devidamente atrasadas e defasadas dentro de uma faixa de tempo constante. (Biondi Neto et al, 2005)

4. Modelagem

As etapas de modelagem, treinamento e execução das redes neurais foram elaboradas com o auxílio do toolbox de redes neurais do software MATLAB® versão 7.0 da MATHWORKS®.

A função de ativação dos elementos processadores pertencentes às camadas escondidas foi definida como tangente hiperbólica. Já os elementos processadores constituintes da camada de saída foram modelados com uma função de ativação do tipo $f(x) = \text{logsig}(x)$. O método numérico usado para a minimização do erro foi o do gradiente decrescente, com 10% de taxa de aprendizado, que minimiza as oscilações. A convergência variou em número de iterações, devido à gama de redes treinadas, sendo esses dados apresentados mais especificamente nos resultados. Utilizou-se apenas como variável de input os dados históricos mensais do Consumo de Energia Elétrica no período de 1982 a 2003, totalizando 264 elementos, deslocados mês a mês. O ano de 2003 não fez parte do treinamento e pôde servir como vetor alvo composto de 12 elementos representativos para investigar a previsão do consumo de energia elétrica neste período a fim de validar os resultados da rede.

Também foi utilizado o período de 1982 a 2000 para prever a demanda de energia elétrica no ano do racionamento de energia e analisar o comportamento da rede em face a esse outlier. (Paschoalino et al, 2007).

5. Resultados

No primeiro momento, os padrões de entrada da rede foram os dados referentes ao período de 1982 até 2002, devidamente pré-processados. O ano de 2003 não fez parte do treinamento e pôde servir como referência para investigar a previsão do consumo de energia elétrica nesse período

A arquitetura empregada na rede que resultou na melhor performance de aprendizado consiste em quatro camadas ocultas (200 100 50 25) e uma camada de saída com apenas um neurônio, e usando a função tangente hiperbólica como função de ativação.

O Gráfico 5.1 representa a sobreposição das curvas (Real versus Previsto), onde a curva (contínua e de cor azul) de consumo de energia faz referência ao período investigado para validação. Já a curva (pontilhada e de cor vermelha) é a curva de previsão obtida dentro do período de defasagem da rede.

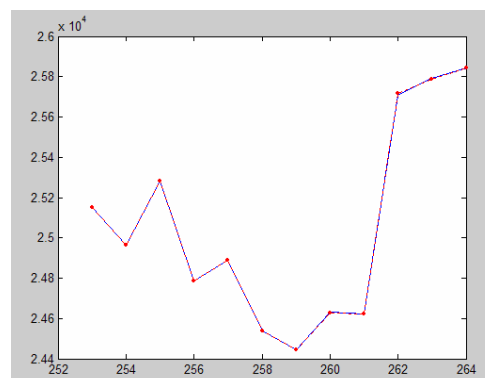


Gráfico 5.1 – Sobreposição da Curva de Consumo com a Curva de Previsão do Consumo

O resultado final obtido nesta pesquisa mostra que os erros percentuais da aplicação analisada foram menores que 0,2%, sinalizando resultado satisfatório comparado com modelos matemáticos e estatísticos de previsão (Gráfico 5.2).

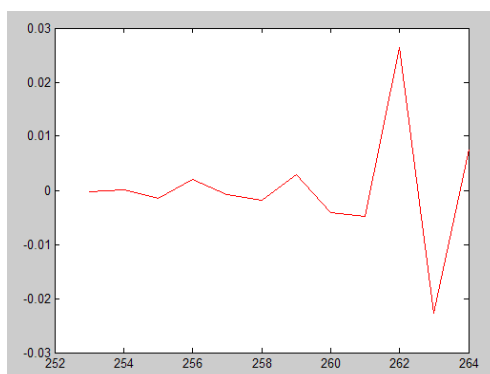


Gráfico 5.2 – Curva do Erro %

No segundo momento, buscou-se prever o ano de 2001, alvo do racionamento e da conseqüente queda no consumo de energia elétrica, a fim de investigar o comportamento da rede face a essa não linearidade (Gráfico 5.3).

A arquitetura empregada na rede que resultou na melhor performance de aprendizado consiste em três camadas ocultas (180 90 45) e uma camada de saída, com apenas um neurônio, e usando a função tangente hiperbólica como função de ativação.

O erro médio obtido nesta previsão foi de 0,5% pela modelagem por RNA (Gráfico 5.4). Observe que no período do segundo trimestre do ano, ou seja, o período em que foi iniciado o racionamento de energia, a rede não conseguiu aprender, configurando que o consumo deveria ser maior do que o ocorrido. O que na realidade, não deixa de ser um ponto positivo da rede. Possivelmente não existia capacidade instalada suficiente ou linhas de transmissão que pudessem levar a energia produzida para suprir a demanda prevista.

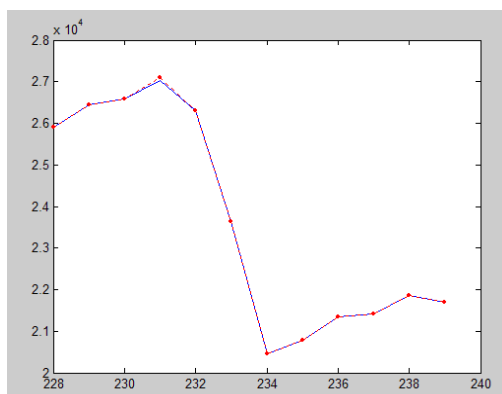


Gráfico 5.3 – Sobreposição da Curva de Demanda com a Curva de Previsão de Demanda

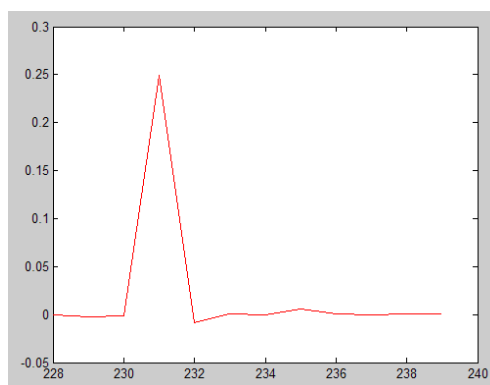


Gráfico 5.4 – Curva do Erro %

A taxa de aprendizado utilizada para ambos os modelos foi de 10%, sendo que o tempo de aprendizado para o segundo modelo investigado ultrapassou 7 horas de treinamento e mesmo apresentando resultado satisfatório, o modelo não convergiu dentro do número de 100.000 iterações estabelecido a priori. Em contrapartida, o primeiro modelo teve um tempo de pouco menos de 2 horas de treinamento e convergiu com aproximadamente 7.645 iterações. A Tabela 5.1 ilustra o resumo geral dos dois modelos treinados.

	Modelo Temporal 1	Modelo Temporal 2
Variáveis Utilizadas	Consumo de Energia Elétrica	Consumo de Energia Elétrica
Taxa de Aprendizado	0,1	0,1
Estrutura das Camadas	(180 90 45 1)	(180 90 45 1)
Função de Ativação	(tansig tansig logsig logsig)	(tansig tansig tansig logsig)
Número máximo de iterações	100.000	100.000
Meta de treinamento	1,00E-08	1,00E-08
Periodicidade	mensal	mensal
Horizonte de Estudo dos Dados	1982-2003	1982-2001
Referência para Validação	2003	2001

Tabela 5.1 – Resumo geral dos dois modelos investigados

6. Conclusões

Os resultados obtidos através da metodologia proposta neste estudo são satisfatórios e apontam que o emprego das redes parcialmente recorrentes de Elman como ferramentas de previsão de séries temporais podem trazer simultaneamente flexibilidade e confiabilidade para futuras análises. Flexibilidade porque as RNA's podem ser treinadas para diferentes tipos de séries históricas e até mesmo séries híbridas, como foi demonstrado. Confiabilidade porque as RNA's são capazes de prever o consumo de energia elétrica apresentando pequenas margens de erro.

No primeiro modelo investigado com série histórica do consumo como variável de input e com o ano de 2003 para validação dos resultados, o erro médio foi de 0,02%. Este resultado é muito menor quando comparado aos processos puramente estatísticos, o que destaca a superioridade da rede de Elman.

Já o modelo em que o ano de 2001 serviu como referência para validação dos resultados, apresentou um erro médio de 0,5%. A explicação para esse comportamento da rede é que no período do primeiro trimestre de 2001, houve uma campanha para o uso racional da energia no país, que acarretou na diminuição da demanda de energia elétrica. Essa repentina descontinuidade na tendência de crescimento de demanda configurou em um estágio delicado para a performance de aprendizado da RNA.

Em resumo, este projeto apresentou a utilização da inteligência computacional para analisar modelos de previsão com características não lineares, assumindo como inputs a série histórica de consumo de energia elétrica no Brasil. Foi obtido êxito na modelagem de redes simples com séries históricas e deixa-se a proposta de aprofundamento na coleta e análise de dados que configurem também sucesso na modelagem de redes híbridas.

7. Referências Bibliográficas

ABELÉM, A. J. G. “*Redes Neurais Artificiais na previsão de séries temporais*”. 1994. Tese (Mestrado em Engenharia Elétrica) - Curso de Mestrado em Engenharia Elétrica, PUC - Rio, Rio de Janeiro.

- BIONDI NETO, L., COELHO, P.H.G., CHIGANER, L., ANGULO MEZA, L., MONTEIRO, L.P., “*Previsão da Vazão afluyente mensal do rio São Francisco usando-se redes de Elman*”. *Engevista*, v. 7, n. 1, p. 63-72, 2005.
- BRAGA, A. P., CARVALHO, A. P. L. F., LUDERMIR, T. “*Fundamentos de Redes Neurais Artificiais*”. 11ª Escola de Computação. Rio de Janeiro, RJ
- CASTRO, N. J., FERNANDES, P. C. “A Expansão da Fronteira Elétrica do Brasil”. *Brasil Energia*, Rio de Janeiro: Editora Brasil Energia nº312, novembro de 2006.
- CICHOCKI, A., UNBEHAUEN, R., 1996. “*Neural Networks for Optimization and Signal Processing*”. New York, USA, John Wiley & Sons, Inc.
- DHAR, V., STEIN, R. “*Seven Methods for Transforming Corporate Data into Business Intelligence*”. Prentice-Hall, New York - 1997.
- ELMAN, J. L., “*Finding Structure in Time*”. 1990. In *Cognitive Science*, vol. 14, pp. 179-211.
- HAIR, J. et al. “*Multivariate Data Analysis*”. Editora Prentice Hall, New Jersey, 1998.
- HAYKIN, S. “*Redes Neurais - Princípios e prática*”. Porto Alegre, RS: Bookman, 2001.
- IPEADATA. <<http://www.ipeadata.gov.br>> Acesso em: 14 de novembro de 2006.
- PASCHOALINO, F. F., LOUREIRO, T. Y. C., SOARES DE MELLO, J. C. C. B., BIONDI NETO, L. (2007). “Previsão de Consumo de Energia Elétrica no Brasil baseada em Redes de Elman”. Pôster publicado no XXXIX SBPO, 2007.
- RUSSELL, S. J. e NORVIG. Peter. “*Artificial Intelligence: a modern approach*”. New Jersey: Prentice-Hall Inc., 1995
- YAMAMOTO Y, NIKIFORUK P N. “*A new supervised learning algorithm for multilayerd and interconnected neural networks*”. *IEEE Trans Neural Networks*, 2000.