

## PREVISOR NEURAL DE CARGA ELÉTRICA BASEADO NA TRANSFORMADA WAVELET DISCRETA

Agnaldo J. Rocha Reis<sup>1</sup>, Alexandre P. Alves da Silva<sup>2</sup>

<sup>1</sup>Grupo de Engenharia de Sistemas, Instituto de Eng. Elétrica, Universidade Federal de Itajubá

Av. BPS, 1303 - Itajubá, MG, 37500-903, BRASIL

<sup>2</sup> PEE – COPPE - Universidade Federal do Rio de Janeiro

C.P. 68.504 – Rio de Janeiro, RJ, 21945-970, BRASIL

E-mails: agnreis@iee.efei.br, alex@coep.ufrj.br

### Abstract

*The importance of short-term load forecasting has been increasing lately. With deregulation and competition, energy price forecasting has become a big business. Bus-load forecasting is essential to feed analytical methods utilized for determining energy prices. The variability and non-stationarity of loads are becoming worse due to the dynamics of energy prices. Besides, the number of nodal loads to be predicted does not allow frequent interventions from load forecasting experts. More autonomous load predictors are needed in the new competitive scenario. This paper proposes two different strategies for successfully embedding the discrete wavelet transform into artificial neural network-based short-term load forecasting. Its main goal is to develop more robust load forecasters. Hourly temperature and load data for two whole years from a North-American electric utility was used in order to test the proposed methodology.*

### 1. Introdução

Redes Neurais Artificiais (RNA) têm sido aplicadas à Previsão de Carga a Curto Prazo (PCCP) com muito êxito [1]. Muitas concessionárias de energia elétrica que anteriormente empregavam ferramentas para a PCCP baseadas em técnicas estatísticas clássicas, estão atualmente utilizando programas para PCCP baseados em RNA [2], [3]. É um daqueles casos raros onde a ciência se transformou em tecnologia após um curto período de desenvolvimento. No entanto, essa tecnologia está longe de alcançar o seu limite.

O pré-processamento de dados é uma importante pré-condição para uma aplicação bem sucedida das RNA. Em [4], a implicação do pré-processamento de séries temporais para fins de previsão de carga foi investigado. Os resultados apresentados sugerem que técnicas tais como padronização de dados, diferenciação e filtros digitais são bastante úteis para enfatizar características importantes das séries de carga e remover *outliers*.

Outro requisito muito importante é uma certa regularidade dos dados [5]. Na maioria dos casos,

assume-se que uma dada série de carga é estacionária. Para séries de carga elétrica tal consideração de estacionariedade tem que ser descartada na maioria das vezes. Além disso, deve-se considerar a existência de diferentes tipos de não-estacionariedade.

Para lidar com o problema de não-estacionariedade, *wavelets* são utilizadas. Isso se deve ao fato das *wavelets* fornecerem uma boa representação local do sinal tanto no domínio do tempo quanto no domínio da frequência. Diferentemente das bases de Fourier, as *wavelets* podem ser mantidas em um intervalo fechado arbitrariamente pequeno [6]. Assim, a transformada *wavelet* é uma ferramenta muito poderosa para tratar com fenômenos transitórios. Além do mais, a análise de multiresolução via *wavelets* pode ser utilizada para fornecer uma descrição estatística de cargas elétricas [7].

Nesse artigo investiga-se estratégias de previsão de carga que exploram a propriedade de multiresolução da transformada *wavelet*. Dois problemas são abordados e tratam da extração de informação relevante de uma curva de carga contaminada com algum tipo de ruído, a saber: 1) Que informação está contida em um sinal, e quais partes dessa informação são úteis para fins de previsão? 2) Há a presença de vários tipos de ruído? É possível identificar cada um deles? Para responder a essas questões, decompõe-se uma série de carga via Transformada *Wavelet* (TW) em  $n+1$  componentes, sendo uma componente de baixa frequência (ou tendência) mais  $n$  componentes de detalhes (onde  $n$  é o número de níveis de detalhe). A partir dessa decomposição, duas abordagens são averiguadas. A primeira delas consiste em gerar um modelo de PCCP cujas entradas são compostas por sinais de carga e temperatura no domínio do tempo e sinais de carga no domínio de *wavelets*. A segunda abordagem consiste na modelagem da série de carga no domínio do tempo através da modelagem de cada uma de suas componentes no domínio de *wavelets*. Para testar a viabilidade das abordagens supraditas, dois anos completos de dados de carga e temperatura (tomados com base horária) de uma concessionária norte-americana são utilizados. As previsões realizadas são do tipo *off-line*, recursiva e de 1 a 24 passos à frente.

Na literatura especializada há alguns trabalhos publicados que tratam da combinação entre wavelets e PCCP via RNA, e.g. [8], [9]. Portanto, nossas principais contribuições são: (1) Aumento da capacidade de generalização de um predictor neural através da extração de características relevantes da série de carga via análise de multiresolução - Uma decomposição em diferentes componentes da série de carga pode ser produzida, o que reduz os esforços de aprendizagem durante o processo de treinamento e validação das RNA [4]; (2) Inclusão da variável temperatura em uma estratégia de previsão de carga que combina wavelets e RNA - Usualmente, apenas a série de carga tem sido considerada.

Esse artigo é dividido como se segue. A Seção 2 apresenta a teoria básica sobre a Transformada Wavelet Discreta (TWD). Na Seção 3, a análise em multiresolução e o algoritmo piramidal de Mallat são discutidos. A aplicação da TWD na PCCP é assunto da Seção 4. Os modelos de previsão propostos são descritos na Seção 5. Esses modelos são comparados através de simulações na Seção 6. Finalmente, a Seção 7 apresenta as principais conclusões desse artigo e algumas sugestões para trabalhos futuros.

## 2. A Transformada Wavelet Discreta

Wavelets são funções que satisfazem certas propriedades matemáticas, podendo ser suaves ou não, simétricas ou não, e ter expressões simples ou não. O fato básico sobre as wavelets é que elas são localizadas no tempo (ou espaço), contrariamente ao que ocorre com as funções trigonométricas. Este comportamento torna-as ideais para a análise de sinais não-estacionários, contendo transitoriedade e estruturas do tipo fractais [10].

Para a TW, os dados são representados via superposição de wavelets com diferentes posições e escalas, cujos coeficientes essencialmente quantificam a força da contribuição das wavelets naquelas posições e escalas. A análise via wavelets emprega uma função protótipo chamada “wavelet-mãe”. Essa função tem média zero e decai bruscamente de forma oscilatória. Matematicamente, a TWD de um dado sinal  $x(n)$  com respeito à wavelet-mãe  $g(\cdot)$  é genericamente definida como (1):

$$TWD(m, k) = \frac{1}{\sqrt{a_0^m}} \sum_n x(n) g\left(\frac{k - nb_0 a_0^m}{a_0^m}\right) \quad (1)$$

onde os parâmetros de dilatação e translação,  $a$  e  $b$ , respectivamente, são funções de um parâmetro inteiro  $m$ , i.e.,  $a = a_0^m$  e  $b = nb_0 a_0^m$ . Também,  $k$  é uma variável inteira que refere-se a uma dada amostra do sinal de entrada. O parâmetro  $a$  produz dilatações geométricas do tipo 1,  $a_0^{-1}$ ,  $a_0^{-2}$ , ... . Esse esquema de dilatação confere a TWD uma cobertura de frequência logarítmica. Para maiores detalhes, veja [10], [11], [12].

## 3. Análise em Multiresolução e o Algoritmo Piramidal de Mallat

A análise via wavelets deve sua eficiência ao rápido algoritmo piramidal. O algoritmo é usado para computar a TWD e teve sua origem em processamento de imagens [6].

A análise de multiresolução via algoritmo piramidal refere-se ao procedimento de se obter “aproximações” e “detalhes” de um dado sinal de interesse. Uma aproximação é uma representação de baixa frequência do sinal, enquanto que um detalhe é a diferença entre duas representações sucessivas de baixa resolução do sinal original (e.g,  $D_3=A_2-A_3$ ). Uma aproximação contém a tendência geral do sinal original enquanto que um detalhe exhibe os componentes de alta frequência do sinal de entrada.

O algoritmo piramidal apresenta dois estágios: um deles calcula a TWD e a outro calcula a Transformada Wavelet Inversa (TWI). A fase direta do algoritmo utiliza um banco de filtros lineares dos tipos passa-baixa (L) e passa-alta (H) na decomposição do sinal. Com referência a Fig. 1, define-se a análise em multiresolução como sendo (2):

$$D_j(n) = \sum_k h(k) A_{j-1}(n-k) \quad (2)$$

$$A_j(n) = \sum_k l(k) A_{j-1}(n-k)$$

onde  $D_j$  e  $A_j$  são os detalhes e aproximações na resolução  $j$ , respectivamente.  $A_{j-1}$  é a aproximação do nível imediatamente acima do nível  $j$ . Finalmente,  $k = 1, 2, \dots, K$ , onde  $K$  é o comprimento do vetor filtro. Também, como mostrado na Fig. 1, o algoritmo combina os filtros L e H com operações de decimação por 2, i.e., a cada duas saídas do filtro despreza-se uma delas. Essas operações conferem velocidade ao algoritmo pois o esforço computacional é reduzido geometricamente a cada iteração [12]. A outra fase, para o cálculo da TWI, simplesmente inverte o processo. Ela combina uma série de operações de filtragem linear com “upsampling” por 2. Essa operação adiciona zeros entre valores de dados, duplicando os dados a cada iteração (como cada zero inserido é multiplicado pelo coeficiente do filtro, os valores interpolados não aparecem no somatório final). Esse algoritmo requer o mesmo esforço computacional que o primeiro.

## 4. Aplicação da TWD na PCCP

Nessa seção encontra-se um exemplo de aplicação da TWD no problema de PCCP. Esse exemplo tem como objetivo mostrar a forma típica de saída da decomposição em wavelets de uma série de carga como uma análise de multiresolução. Para tanto, utilizou-se uma série de carga normalizada entre  $[-0,5;0,5]$ . Wavelets da família Daubechies foram utilizadas na análise da série. A escolha dessa wavelet-mãe foi

baseada em [7]. Também, três níveis de resolução foram considerados. Tal nível de detalhamento permite o estudo dos componentes do sinal para os quais o período é inferior a 8h. A Fig. 2 ilustra a série de carga considerada, a aproximação  $A_3$  e os níveis de detalhe  $D_3$ ,  $D_2$  e  $D_1$ . Com base nessa figura, nota-se que a aproximação  $A_3$ , correspondente a escala de tempo de 8 h, é uma boa aproximação do sinal. Ela captura a forma da curva de carga do período em análise. Já os níveis de detalhes  $D_1$  e  $D_2$  apresentam valores de baixa magnitude e alguns picos em determinados instantes. Esses sinais exibem um certo tipo de padrão e têm formas e valores médios similares. As irregularidades contidas nesses sinais podem ser causadas por erros de medição e outras fontes de ruído. Finalmente, o nível de detalhe  $D_3$  apresenta picos que permitem a localização temporal de eventos. Tal informação é bastante útil em se tratando de PCCP.

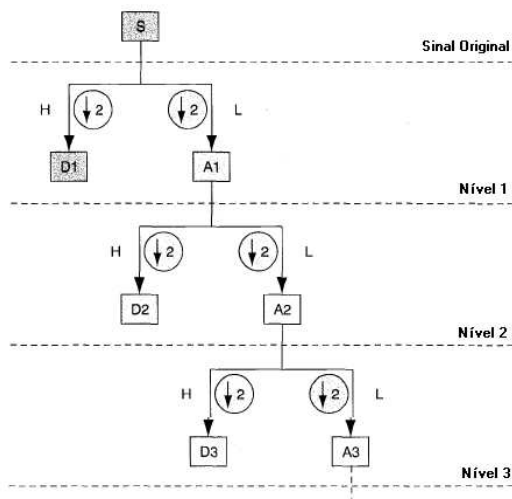


Figura 1: Representação esquemática do algoritmo piramidal ( $S=A_3+D_3+D_2+D_1$ ).

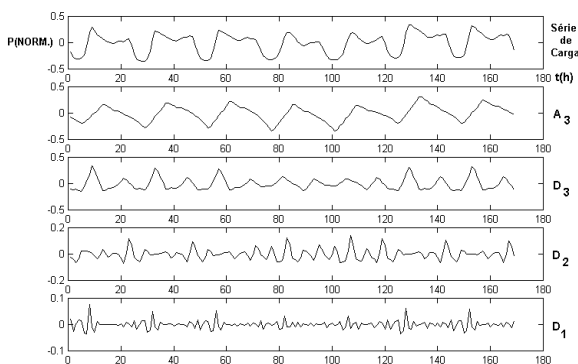


Figura 2: Decomposição em 3 níveis de resolução de uma curva de carga diária.

## 5. Modelos de Previsão de Carga

Essa seção descreve os modelos de previsão de carga considerados nesse artigo. Quatro modelos são comparados. O conjunto básico de variáveis de entrada corresponde aos valores em atraso da série de carga horária ( $P$ ) de 1h, 2h, 24h e 168h, aos valores em atraso da série de temperatura observada ( $T$ ) de 1h, 2h, 24h e 168h, a temperatura prevista para um passo à frente ( $\tilde{T}$ ), e duas entradas adicionais,  $HS(t)=\text{sen}(2\pi t/24)$  e  $HC(t)=\text{cos}(2\pi t/24)$ , codificando a hora do dia ( $t$ ). A saída corresponde a previsão de carga um passo à frente, que é usada recursivamente para estender o horizonte de previsão até 24 horas.

Em [13], os autores mostraram que as melhores variáveis de entrada para previsores de carga lineares não representam, necessariamente, o melhor conjunto de variáveis de entrada para modelos não lineares. Apesar disso, o presente trabalho continua utilizando as variáveis de entrada mais populares para fins de previsão de carga, pois o objetivo principal desse artigo é mostrar que a utilização de wavelets pode influenciar positivamente o desempenho de um predictor neural baseado em Perceptrons de Múltiplas Camadas (PMC).

As arquiteturas selecionadas para cada conjunto de entrada (Subseções 5.1-5.4) variam com respeito ao número de neurônios na camada oculta. Três ou quatro neurônios foram utilizados dependendo do modelo em questão (essa escolha foi feita de maneira empírica). Uma única camada oculta foi utilizada. As funções de ativação tangente hiperbólica e linear foram empregadas nas camadas oculta e de saída, respectivamente.

### 5.1. Modelo 1 (M1)

O modelo 1 é de certa forma bastante conhecido. Após definir-se a ‘janela’ de dados que será utilizada para fins de treinamento e validação, faz-se a normalização e a padronização da mesma. Esse tipo de pré-processamento faz com que os dados fiquem no intervalo  $[-0,5; 0,5]$ . Sua arquitetura consiste de 11 entradas, 4 neurônios na camada oculta e 1 neurônio na camada de saída. A Fig. 3 apresenta esse modelo de forma esquemática.

### 5.2. Modelo 2 (M2)

O modelo 2 é uma combinação de M1 com técnicas de diferenciação. Portanto, três séries são empregadas: carga padronizada, temperatura padronizada e carga diferenciada. A diferenciação é aplicada à série de carga padronizada. Para M2, apenas as diferenças de 1ª ordem ( $D$ ) são consideradas. A Tab. 1 apresenta o conjunto de entradas para M2. Para fins de treinamento, a saída da RNA é a potência ativa no instante de tempo  $t$  subtraída da potência ativa no instante de tempo imediatamente anterior (i.e.,  $D(t) = P(t) - P(t-1)$ ). A arquitetura da RNA

utilizada consiste de 15 entradas, 4 neurônios na camada oculta e um neurônio na camada de saída.

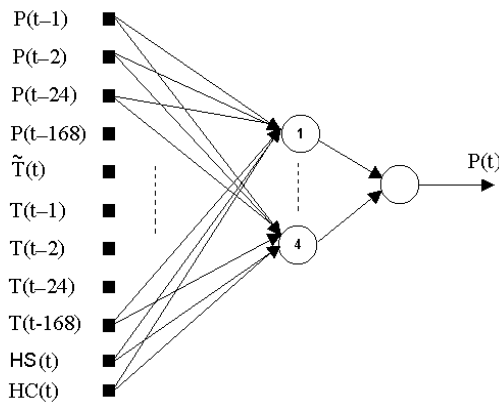


Figura 3: Previsor de carga com entradas básicas (M1).

Tabela 1: Variáveis de entrada para M2.

Entrada	Variável	Lags (h)
1-4	Carga (P)	1, 2, 24, 168
5-8	Carga Diferenc. (D)	1, 2, 24, 168
9	Temp. Prevista	0*
10-13	Temperatura (T)	1, 2, 24, 168
14	Hora seno (HS)	0*
15	Hora cos. (HC)	0*

Obs.: O “lag 0” representa a hora a ser prevista.

### 5.3. Modelo 3 (M3)

O modelo 3 combina as técnicas de normalização e padronização de dados, diferenciação e análise de multiresolução via wavelets da família Daubechies. Na verdade, M3 é uma combinação entre M2 e wavelets. Para esse modelo em particular, wavelets são usadas como uma ferramenta de pré-processamento. Existem duas razões para se fazer isso. Primeiramente, características importantes de uma série de carga podem ser enfatizadas. Em segundo lugar, a componente relativa ao ruído pode ser separada. A Fig. 4 mostra o diagrama esquemático de M3.

Inicialmente, a aproximação  $A_3$  e os três níveis de detalhes foram utilizados como entradas para o modelo. Com essa configuração, o modelo não apresentou um bom desempenho. De certa forma esse resultado era esperado já que as partes constituintes da carga (tanto as características principais quanto o ruído) foram fornecidas ao modelo. Com base na análise via wavelets apresentada na Seção 4, constata-se que as subséries mais informativas são  $A_3$  e  $D_3$ . Já as subséries  $D_1$  e  $D_2$  contêm algumas irregularidades que são provavelmente devidas a erros de medição e outras fontes de ruído aleatório. Em suma, essa análise sugere que somente as informações provenientes de  $A_3$  e  $D_3$  devem ser consideradas na modelagem. O conjunto de entradas para M3 pode ser encontrado na Tab. 2. Para fins de

treinamento, a saída da RNA é a mesma utilizada em M2 (i.e.,  $D(t) = P(t) - P(t-1)$ ). A arquitetura da rede é composta de 23 entradas, 4 neurônios na camada oculta e 1 neurônio na camada de saída.

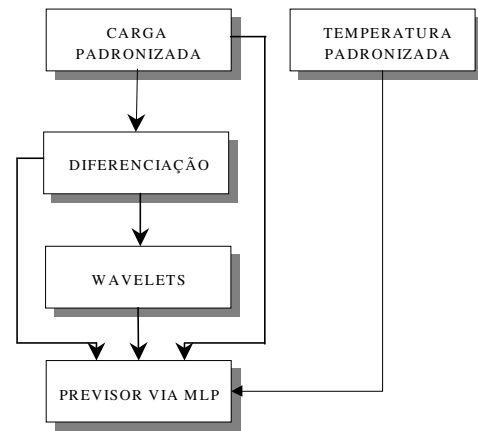


Figura 4: Diagrama esquemático de M3.

Tabela 2: Variáveis de entrada para M3.

Entrada	Variável	Lags (h)
1-4	Carga Dif.(D)	1, 2, 24, 168
5-8	Aprox. ( $A_3$ )	1, 2, 24, 168
9-12	Carga (P)	1, 2, 24, 168
13-16	Detalhe ( $D_3$ )	1, 2, 24, 168
17	Temp. Prev.	0
18-21	Temp. (T)	1, 2, 24, 168
22	Hora seno	0
23	Hora cos.	0

### 5.4. Modelo 4 (M4)

O modelo 4 apresenta uma abordagem diferente daquela empregada nos modelos anteriores. A idéia aqui é decompor a série de carga via wavelets e modelar a série de carga no domínio do tempo através da modelagem de cada uma de suas componentes no domínio de wavelets. Isto é, cada uma das componentes ( $A_3, D_3-D_1$ ) é modelada separadamente e a previsão final é computada através da combinação das previsões fornecidas por cada um dos submodelos. Notar que para M4, a “natureza” de cada componente é irrelevante. Tal fato implica que a aproximação  $A_3$  e todos os três níveis de detalhes devem ser necessariamente modelados. A Fig. 5 mostra o diagrama para esse modelo. As entradas e saídas de cada uma das RNA são apresentadas na Tab. 3. A arquitetura das redes (RNA1-RNA4) consta de 3 neurônios na camada oculta e 1 neurônio na camada de saída.

Analisando-se a Tab. 3, constata-se que somente RNA1 utilizou sinais de carga ( $A_3$ ) e temperatura como

entradas. Esse submodelo foi gerado após aplicar-se a ferramenta de correlação cruzada nos pares de séries A<sub>3</sub>-Temperatura, D<sub>3</sub>-Temperatura, D<sub>2</sub>-Temperatura, D<sub>1</sub>-Temperatura e verificar que somente o par A<sub>3</sub>-Temperatura apresentava índices elevados de correlação.

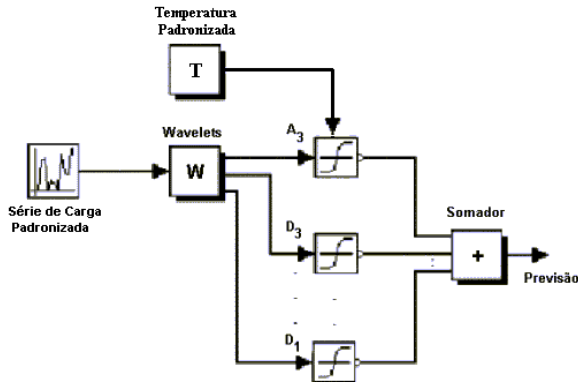


Figura 5: Diagrama esquemático do modelo 4.

Tabela 3: Variáveis de entrada para M4.

RNA	Entrada	Variável	Lags (h)	Saída
1	1-4	Aprox. (A <sub>3</sub> )	1, 2, 24, 168	A <sub>3</sub> (h)
	5-8	A <sub>3</sub> -diferenciada(DA <sub>3</sub> )	1, 2, 24, 168	
	9	Temp. Prevista	0*	
	10-12	Temperatura (T)	1, 2, 24, 168	
	13	HS	0*	
2	1-5	Detalhe (D <sub>3</sub> )	1, 2, 12, 24, 168	D <sub>3</sub> (h)
	6-10	D <sub>3</sub> -diferenciada(DD <sub>3</sub> )	1, 2, 12, 24, 168	
	11	HS	0*	
	12	HC	0*	
3	1-5	Detalhe (D <sub>2</sub> )	1, 2, 12, 24, 168	D <sub>2</sub> (h)
	6-10	D <sub>2</sub> -diferenciada(DD <sub>2</sub> )	1, 2, 12, 24, 168	
	11	HS	0*	
	12	HC	0*	
4	1-5	Detalhe (D <sub>1</sub> )	1, 2, 12, 24, 168	D <sub>1</sub> (h)
	6-10	D <sub>1</sub> -diferenciada(DD <sub>1</sub> )	1, 2, 12, 24, 168	
	11	HS	0*	
	12	HC	0*	

## 6. Testes

Realizou-se uma comparação entre os quatro modelos anteriormente mencionados. Para tanto, utilizou-se um histórico de dados de carga e temperatura observados por uma concessionária de energia elétrica norte-americana no período de janeiro de 1988 a outubro de 1992. Para fins de treinamento e validação dos modelos, janelas de seis semanas do ano corrente, mais janelas de seis semanas do mesmo período dos dois anos anteriores foram utilizadas (e.g., out./nov. 1988, out./nov. 1989 e out./nov. 1990). O agrupamento de dados foi feito de acordo com o dia da semana. Para cada dia da semana, um PMC foi treinado com uma variante do algoritmo “backpropagation” com validação

cruzada [14]. Diferentes partições para os conjuntos de treinamento e validação foram criadas aleatoriamente a cada 500 épocas (90% dos padrões disponíveis foram utilizados para treinamento). Durante o treinamento, nenhum tratamento especial para feriados foi aplicado e a temperatura prevista foi substituída pela temperatura observada. Isso se deve ao fato da concessionária não possuir as previsões de temperatura para o referido período.

Depois do treinamento para previsões um passo à frente, previsões recursivas de 1-24 passos à frente são computadas. Os previsores de carga são retreinados no fim do dia. Move-se a janela de treinamento um dia à frente e as previsões para as próximas 24 horas são realizadas (elas se iniciam à meia noite). Esse procedimento de teste é repetido para um período de dois anos completos. Os índices: Erro Percentual Absoluto Médio (MAPE), Erro Quadrático Médio (MSE), Erro Médio (ME) e Erro Máximo (MAX) foram utilizados para avaliar os modelos de previsão de carga.

As tabelas que se seguem apresentam os resultados obtidos. O MAPE horário da hora *t* para o conjunto de testes é mostrado na Tab. 4. Os testes realizados são conhecidos como “testes cegos”, já que os exemplares que formam esse conjunto não participaram das fases de treinamento ou validação dos modelos. A Tab. 5 apresenta uma síntese dos índices de avaliação para os quatro modelos.

Com base na Tab. 5, constata-se que o modelo 4, em média, foi o mais exato. Para múltiplos passos à frente M4 apresentou erros de previsão significativamente menores que os demais modelos (ver Tab. 4). É importante ressaltar que um horizonte de previsão maior não implica necessariamente em um erro de previsão maior. Isso depende da variabilidade dos dados para diferentes períodos do dia e do modelo de previsão adotado. Notar também que o índice MSE, na Tab. 5, aponta um número menor de erros elevados para M4 em relação aos demais. Exemplos gráficos de previsões de carga para um dia à frente são apresentados na Fig. 6.

## 7. Conclusões

Com o crescimento dos sistemas de potência e o aumento na sua complexidade, muitos fatores têm influenciado significativamente a geração e o consumo da energia elétrica. Portanto, o processo de previsão tem se tornado mais complexo, o que requer previsores de carga mais precisos no novo ambiente competitivo.

Nesse artigo, duas abordagens distintas para o problema de PCCP via RNA através da utilização da transformada wavelet discreta foram discutidas. Quatro modelos de previsão foram apresentados e tiveram seus desempenhos avaliados. As previsões realizadas foram do tipo recursiva, de 1 a 24 passos à frente, em modo off-line, realizadas por perceptrons de múltiplas camadas treinados com uma versão modificada do algoritmo de retropropagação de erros [14]. Dois anos

completos de dados reais de carga e temperatura foram utilizados para testar a metodologia proposta.

A modelagem da série de carga no domínio do tempo através da modelagem de cada uma de suas componentes no domínio de wavelets (M4), em média, foi a mais eficaz para o horizonte de 1-24 h. Observou-se um ganho médio de 21,77% em relação ao modelo 1 e 20,27% em relação a M2. Esse conjunto de resultados sugere que a análise de multiresolução via wavelets pode contribuir para uma melhoria considerável na exatidão das previsões de carga realizadas por arquiteturas neurais. Pesquisas futuras testarão outras famílias de wavelets e novos modelos de previsão. Por outro lado, a determinação de intervalos de confiança para as previsões obtidas também será considerada [15].

Tabela 4: MAPE horário para os quatro modelos.

Hora (h)	MAPE M1(%)	MAPE M2(%)	MAPE M3(%)	MAPE M4(%)
1	1,84	1,12	0,94	1,19
2	3,17	2,03	1,88	1,84
3	3,70	2,36	2,71	2,38
4	3,78	2,85	3,00	2,67
5	4,08	3,23	3,34	3,00
6	4,80	3,80	3,95	3,15
7	5,99	4,86	4,82	4,95
8	5,06	4,62	4,58	4,02
9	3,77	3,59	3,66	2,85
10	3,12	3,10	3,19	2,43
11	2,91	3,30	3,22	2,29
12	2,95	3,48	3,24	2,44
13	3,09	3,66	3,37	2,70
14	3,32	3,88	3,60	2,80
15	3,53	4,16	3,87	2,91
16	3,71	4,37	4,06	3,08
17	3,56	4,34	4,01	3,04
18	3,59	4,21	3,77	2,91
19	3,49	4,19	3,58	2,90
20	3,79	4,23	3,53	3,00
21	3,83	4,01	3,40	3,12
22	3,74	3,91	3,38	2,95
23	3,97	4,01	3,51	3,36
24	4,46	4,21	3,49	3,78

Tabela 5: Avaliação global dos quatro modelos (1-24 passos à frente).

Índice	M1	M2	M3	M4
MAPE(%)	3,72	3,65	3,42	2,91
MSE(MW <sup>2</sup> )	13681,25	13120,12	12449,28	8252,69
ME(MW)	-7,05	-0,30	-16,18	-1,62
MÁX.(%)	59,71	47,29	51,46	42,86

## Referências

[1] H. S. Hippert, C. E. Pedreira, R. C. Souza. Neural Networks for Short-Term Load Forecasting: A review and Evaluation. *IEEE Trans. Power Systems*, vol. 16, Fev. 2001, pp.44-55.

[2] A. Khotanzad, R. Afkhami-Rohani e D. Maratukulam. ANNSTLF – artificial neural network short-term load forecaster – generation three. *IEEE Trans. Power Systems*, vol. 13, n°. 4, Nov. 1998, pp. 1413-1422.

[3] H. Mori. State-of-the-art overview on artificial neural networks in power systems, In *A Tutorial Course on Artificial Neural Networks with Applications to Power Systems. IEEE Catalog no. 96TP112-0*, Edited by M.A. El-Sharkawi and D. Niebur, 1996, pp. 51-70.

[4] A. J. Rocha Reis & A. P. Alves da Silva. Pré-processamento de séries temporais para a previsão de carga a curto prazo via redes neurais. *Anais do XIII Congresso Brasileiro de Automática*, pp. 1138-1143, Florianópolis, SC, Brasil, 2000.

[5] D. W. Bunn. Forecasting loads and prices in competitive power markets. *Proceedings of the IEEE*, vol. 88, n°. 2, Fev. 2000, pp. 163-169.

[6] S. Mallat. A Theory for Multiresolution Signal Decomposition - The Wavelet Representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, n°. 7, Jul. 1989, pp.674-693.

[7] M. Misiti, Y. Misiti., G. Oppenheim, J.-M. Poggi. Ondelettes en Statistique et Traitement du Signal. *Revue de Statistique Appliquée*, vol. XLI, n°. 4, 1993, pp.33-43.

[8] B.-L. Zhang, Z.-Y. Dong. An adaptive neural-wavelet model for short term load forecasting. *Electric Power Systems Research*, vol. 59, pp. 121-129, 2001.

[9] C.-i. Kim, I.-k. Yu, Y. H. Song. Kohonen neural network and wavelet transform based approach to short-term load forecasting. *Electric Power Systems Research*, vol. 63, pp. 169-176, 2002.

[10] O. Rioul & M. Vetterli. Wavelets and Signal Processing. *IEEE Signal Processing Magazine*, Out. 1991, pp.14-38.

[11] C. H. Kim & R. Aggarwal. Wavelets transform in power systems - Part 1 – General introduction to the wavelet transform. *Power Engineering Journal*, vol. 14, n°. 2, pp. 81-87, Apr. 2000.

[12] G. Strang & T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Wellesley, 1996, 490 p.

[13] I. Drezga & S. Rahman. Input variable selection for ANN-based short-term load forecasting. *IEEE Trans. on Power Systems*, vol. 13, n°. 4, Nov. 1998, pp. 1238-1244.

[14] T. Masters. *Neural. Novel & Hybrid Algorithms for Time Series Prediction*. New York: Wiley, 1995, p. 514.

[15] A. P. Alves da Silva & L. S. Moulin. Confidence Intervals for Neural Network Based Short-Term Load Forecasting. *IEEE Trans. on Power Systems*, vol. 15, n°. 4, Nov. 2000, pp. 1191-1196.

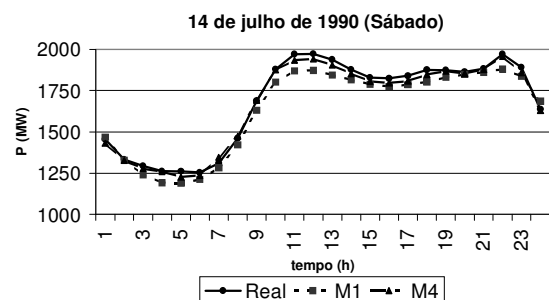


Fig. 6 – Exemplo de previsões de 1-24 passos à frente (M1 e M4).