Journal of Intelligent and Robotic Systems **31:** 149–157, 2001.



# Neural-wavelet Methodology for Load Forecasting

RONG GAO and LEFTERI H. TSOUKALAS
*Appllied Intelligence Systems Lab, School of Nuclear Engineering, Purdue University, W. Lafayette, IN 47907, USA; e-mail: gao@purdue.edu, tsoukala@ecn.purdue.edu*

(Received: 15 March 2001; in final form: 15 March 2001)

**Abstract.** Intelligent demand-side management represents a future trend of power system regulation. A key issue in intelligent demand-side management is accurate prediction of load within a local area grid (LAG), which is defined as a set of customers with an appropriate residential, commercial and industrial mix. Power consumption is deemed to be unpredictable in some sense due to the idiosyncratic behaviors of individual customers. However, the overall pattern of a group of consumers is possible to predict. The developed neural-wavelet approach is shown capable of handling the nonlinearities involved and provides a unique tool for intelligent demand-side management. The paper presents the neural-wavelet approach and its implementation to load identification and forecasting.

**Key words:** load identification, load forecasting, neural-wavelet.

## 1. Introduction

Managing deregulated power systems endowed with enhanced information-processing capabilities (often referred to as the *intergrid*) presents extraordinary challenges due to the complexity of the grid itself and the multitude of information (about generation, transmission, distribution and protection functions) that need to be coordinated in a precise and effective manner. The conventional grid management strategy is generation-side management, that is, it focuses on the activities of generation and transmission-distribution. The underlying assumption in this type of strategy is that the behavior of customers is too complex to be predictable and hence demand can only be considered as an exogenous, noise-like variable. However, recent research shows that although the behavior of an individual power customer is unpredictable, the average pattern over a set of customers maybe quite predictable. This has led us to the concept of local area grid (LAG), defined as a set of residential, commercial, industrial customers with predictable characteristics, making possible anticipatory strategies for the intelligent management of the entire power system [1, 4].

The key issue in demand-size management is the prediction of power load with a reasonable degree of accuracy. Power load prediction is a difficult task because power consumption is determined by many factors, such as temperature, humidity, wind, vacations, economy, including ambient and the idiosyncratic habits of individual customers.

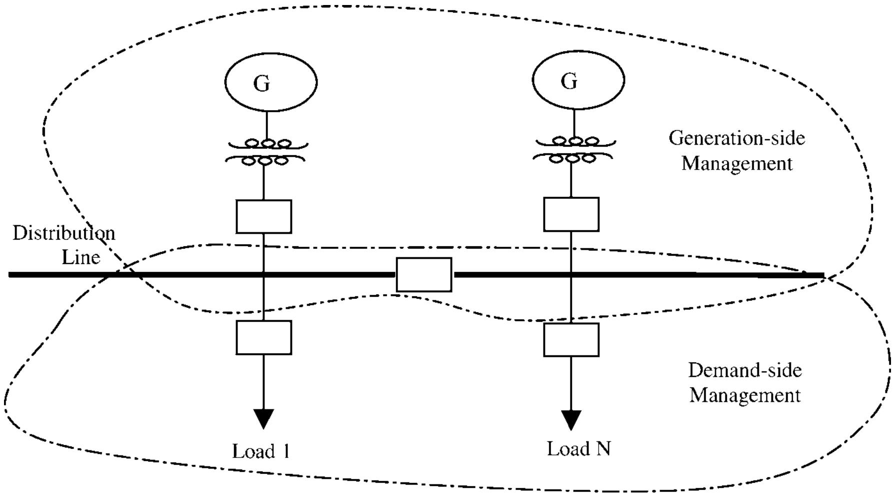



*Figure 1.* Generator-side and demand-side management.

A neural-wavelet approach is one of the most promising methodologies available to solve the load prediction problem because neural-wavelets provide excellent performance in predicting highly nonlinear behaviors. This paper presents a neural-wavelet approach and its implementation to power load identification and prediction.

The rest of the paper is organized as follows. Section 2 provides the basic formalism and theoretical foundations of the neural-wavelet approach and the relation between wavelet coefficients and the Lipschitz Exponent. The Lipschitz Exponent offers a measure of singularity (nonlinearity) of a demand signal. Section 3 briefly discusses the methodology. Section 4 focuses on load identification while Section 5 on load prediction. Load identification involves categorization of the power demand pattern in terms of previously identified patterns (e.g., residential customer, large industrial customer, small commercial industrial customer, etc.). Such categorization is essential for anticipatory demand side management strategies.

## 2. Theoretical Foundations

A wavelet transform decomposes the demand signal into one approximation group and several detailed groups. The continuous and discrete wavelet transforms are defined respectively as [2]:

$$C(a, b) = \int_R s(t)\frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right)\mathrm{d}t, \tag{1}$$

$$C(j, k) = \sum_{m\in\mathbb{Z}} s(n) g_{i,j}(n), \tag{2}$$

where $s(t)$ is the signal; $\Psi$ is the mother wavelet; $a$, $j$ and $b, k$ are scaling and shifting factors, respectively. The wavelet transform is a multi-resolution analysis technique, which is capable of identifying localized features. The multi-resolution analysis capability of wavelets provides us with advanced and powerful tools for signal processing. We see a graphical illustration of this in Section 5 (load prediction), where wavelets are combined with neural network techniques.

Another outstanding characteristics of wavelets is their capability of detecting singularities in a noisy signal. This can partially be explained by its close relation to a mathematic term that is used to describe the degree of singularity that is, the Lipschitz Exponent [3]. A low Lipschitz Exponent means high irregularity. The Lipshitz Exponent can be characterized using wavelet coefficients as follows,

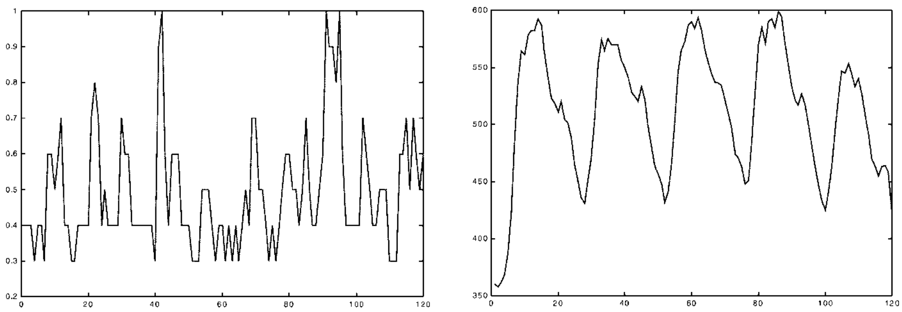
$$\log_2 \left| WT_{2^j} x(t) \right| \leqslant \log_2 K + j\alpha, \tag{3}$$

where $K$ is a constant, $\alpha$ is the Lipschitz Exponent (L.E.) and $j$ is the wavelet scale. Inequality (3) tells us that the maximum value of wavelet coefficients at a particular scale is constrained by the product of scale factor and Lipschitz Exponent. To be more applicable, this is rephrased as follows. If the L.E. of a signal is less than 0 (rather irregular), the maximum value of its wavelet coefficients decreases while the scale factor increases; if the L.E. is greater than zero (rather regular), the opposite situation happens. This characteristic can be utilized to identify the type of load.

## 3. Methodology

Two issues should be made clear before designing a successful prediction system. (1) For different types of load, a different predictor should be applied because different types of customers have a different consumption pattern. Figure 2 shows the hourly load (during weekdays) for two typical customers for a large commercial industrial and a residential customer. We can easily observe the difference between the two of them. It can be concluded that the demand of the large commercial industrial customers is more regular and therefore more predictable. Designing a universal predictor for all types of loads is impossible. This implies that, we should have a load type identification module in our system. (2) There should be a set of rules to take care of special events, like holidays, because the power consumption at theses special events is totally different from everyday behavior.

Taking these two points into account, we come up with the following methodological scheme which is illustrated in Figure 3. First historical and current data are fed to a neural-wavelet to determine the load type. Upon obtaining the result of load type, an appropriate neural network is used to predict the everyday power load demand for this particular customer. At the same time, a set of predefined rules in a fuzzy rule base system monitors the occurring conditions for special events. If any of the special events has occur, the corresponding rule will be fired to take



**Residential House (RLS)** **Large Commercial Industrial (LCI)**

*Figure 2.* Consumption patterns for a residential house and a large commercial industrial.

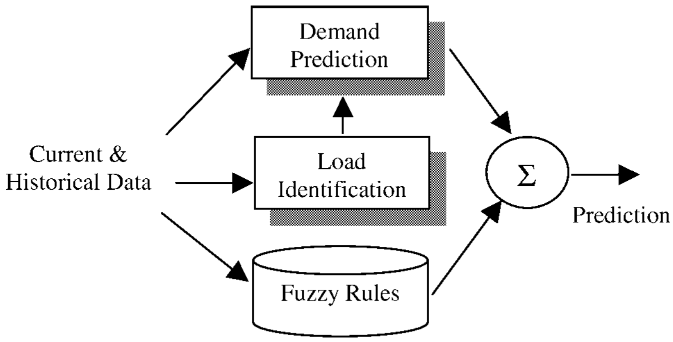


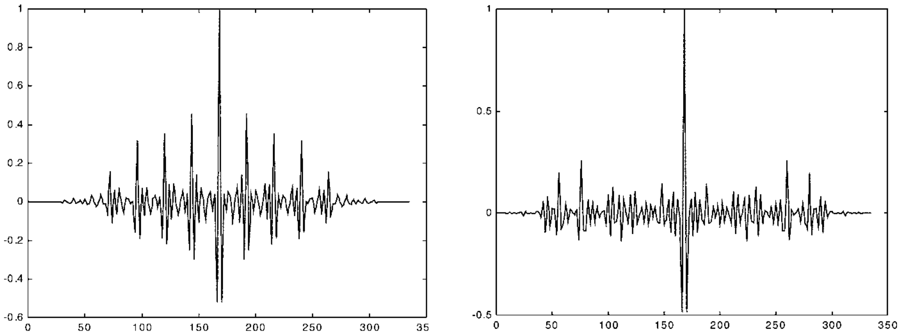

*Figure 3.* Electricity demand prediction.

into account the effect of this event on the power load. Thus, neural-wavelets are responsible for both demand prediction and load identification.

## 4. Load Identification

To examine the feasibility of identifying power loads using the neural wavelets approach, let us consider two types of power load: a typical large commercial industrial (LCI) and a typical residential customer (RLS), which are shown in Figure 2 for five week days (hourly power consumption). Conceptually, we can observe that the curve of LCI is more regular than that of RLS, which leads us to utilize the regularity factor we have discussed in Section 2. This regularity factor appears as a cross-scale feature in the wavelet decomposition tree.

First, we perform a wavelet decomposition for LCI and RLS signals of Figure 2. It is well known that for a smooth signal, the wavelet maximum does not change very much with the change of scale. The different cross-scale behaviors of these two types of load provide enough information to discriminate between them. We can compute the products of detail groups d2 and d3 of these two types of loads in the wavelet decomposition scheme and calculate their auto-correlations functions




**LCI** **RLS**

*Figure 4.* Autocorrelations for LCI and RLS.

respectively, as shown in Figure 4. In Figure 4, the auto-correlation functions are so different that we can distinguish between them using some sort of zero-crossing technique.

## 5. Load Prediction

Conventional neural networks encounter difficulties when trying to predict the mid/long term power demand due to existing high nonlinearities. The nonlinearity problem is the major obstacle that we have to overcome in order to make neural networks converge. Neural-wavelet networks can help. The key point is that generally a neural network is able to approximate any function. This function, from the viewpoint of wavelet transforms, can be represented as a series of wavelet functions. Let us assume that $N$ of these wavelet functions are able to cover the time-frequency space of the signal effectively (acceptable in terms of approximation). Then this signal is represented using these $N$ wavelet functions as:

$$f(t) = \sum_{i=1}^{N} c_i \Psi(a_i t - b_i) = \sum_{i=1}^{N} c_i \Psi_i(t). \tag{4}$$

In Equation (4), the variable $t$ can be intepreted as time. The original signal $f(t)$ is approximated by a linear combination of wavelet functions. The linear coefficients $c_i$ can be treated as connection weights in a neural network. A single input neural-wavelet network based on this concept is shown in Figure 5.

The output of the network shown in Figure 5 can be expressed as:

$$\hat{f}(t) = \sigma\left(\sum_{i=1}^{N} w_i \Psi_i(t)\right), \tag{5}$$

where $\sigma(\cdot)$ is the activation function of the output neuron.

Since the adjustable weights appear only at the output layer, any standard learning algorithm can be used. The structure shown in Figure 5 can be interpreted

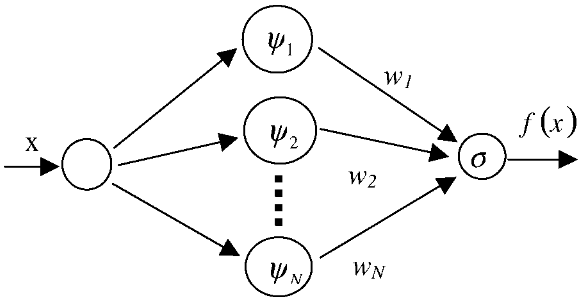



*Figure 5.* A simple neural wavelet network.

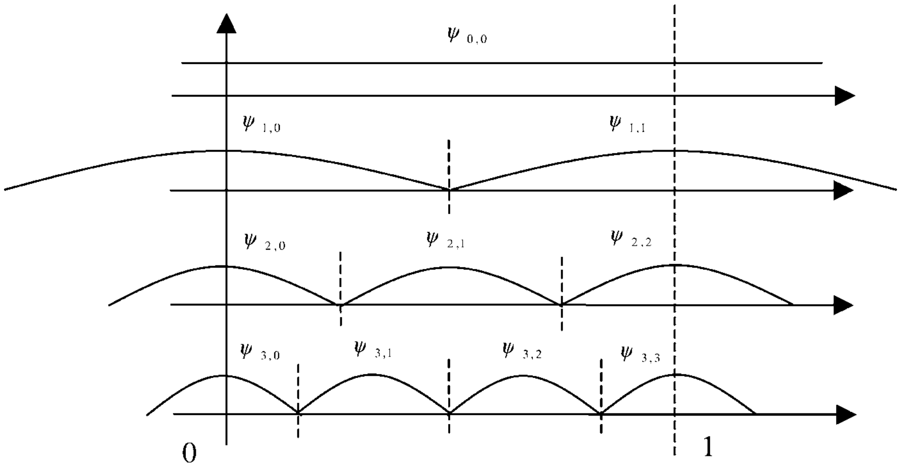



*Figure 6.* A family of non-orthogonal basic functions.

as using a wavelet synapse in an ordinary neural network. Although standard orthogonal wavelet functions can be used in a neural wavelet network, Yamakawa [5] has shown that a family of compactly supported non-orthogonal wavelets is more appropriate for function approximation. We have used a modified version of wavelets, shown in Figure 6 as presented by Yamakawa:

$$\Psi(t) = \begin{cases} \cos^2(\pi t) & \text{if } -0.5 \leqslant t \leqslant 0.5, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

The wavelet bases shown in Figure 6 are not orthogonal but simple and smooth, which are supposed to produce more stable results. Moreover, an over-complete number of bases is adopted to obtain improved results in terms of generality. It is shown that this structure gives better performance in predicting the chaotic behavior of nonlinear systems.

The integration of neural networks with wavelet elements is truly a very promising idea, which makes the networks converge much faster than ordinary feed forward networks.

Neural wavelets have been very successful in function approximation, which is basically a data interpolation problem. In prediction, we need to do extrapolation instead. Thus we need to manipulate the data in an anticipatory way so that an extrapolation problem can be solved by interpolation. Before doing that, let us consider the fact that the power consumption at a specific time in a given day highly correlates not only with those at previous hours in the same day but also with those time intervals at the same time in previous days.

As an example, let us suppose that the current time is 12 PM in day 3. Then it is natural to assume that the current demand has a close relationship with the power consumption at 11 AM and 10 PM in day 3 because they are in the same day. Moreover, the current demand is also similar to the power consumption at the same time intervals of day 2 and day 1 because they are in the same time. This dual temporal correlation is to be exploited in the next step. This can be accomplished by rearranging the data block. Thus, an interleaved arrangement of data is used and in this manner it is easier to make a model for short term or even long term prediction.

Figure 7 exemplifies the basic idea involved in solving our extrapolation problem by interpolation. The resultant data compound is not organized in increasing time base in the usual manners where the historical data (known) are put in the beginning and the data to be predicted (unknown) at the end. Instead, we mix the known and unknown data in such a way that the dual temporal correlation is exploited. In Figure 7, the data to be predicted are represented by question marks. The historical data from three days before, two days before and one day before are represented by circles, crosses and squares, respectively. The lower part of Figure 7 shows the way of mixing the data. Generally, in order to predict the load at a certain time, we use the load of the three previous days at this particular time interval. Moreover, the historical data at the next time step is utilized as the anticipatory information. This arrangement is an effective way to avoid over-fitting. An example of 24-hour ahead prediction is shown in Figure 8. As can be seen from the figure, the results are very promising. The prediction follows the overall trend of the actual

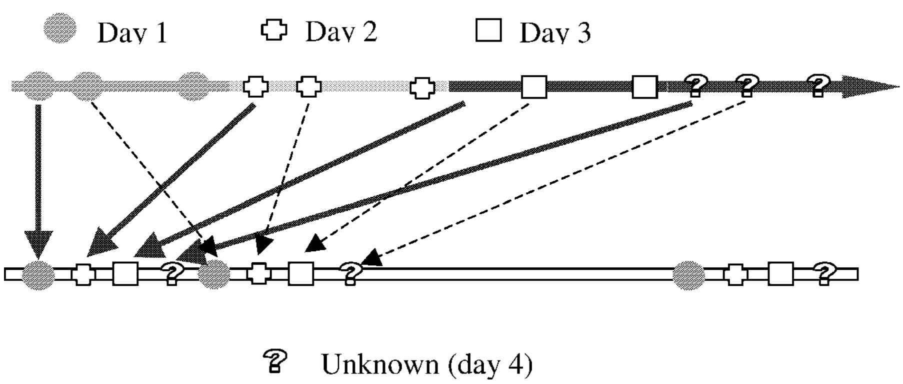



*Figure 7.* Data manipulation to exploit the dual correlations for neural wavelet predictor.

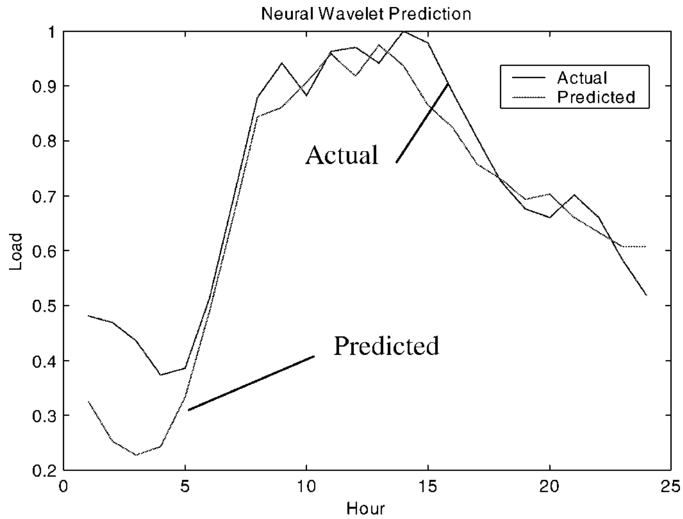



*Figure 8.* 24-hour ahead prediction result.

data and no observable over-fitting problem is present. The error appears in the small load region and is attributed by the use of mean square error as the driving function inside the neural network. Other alternatives such as relative percentage error may improve the results.

## 6. Discussion

An innovative neural wavelet approach has been developed. The approach is capable of both on-line categorization of demand characteristics (that is identifying whether a load behaves like a residential type or a mid-level commercial load, etc.) and prediction (especially mid/long term prediction). The wavelet transform is proven capable of capturing essential features of different types of loads and offers considerable promise to design an on-line wavelet-based discriminator.

It has been shown that the multi-resolution analysis capability of wavelets offers the neural-wavelet approach unique powers in function approximation. By appropriately selecting a wavelet base, the over-fitting problem can be effectively avoided. Moreover, we can carefully arrange the original data to take advantage of the neural wavelet approximator in performing mid-term prediction.

The accuracy and stability of the neural-wavelet prediction can be further improved by fine-tuning wavelet functions. More data and a greater number of loads will be used in order to test the predictive efficacy of the neural-wavelet approach. The methodological promise of both categorization and prediction is very interesting and could provide a useful tool for agents dedicated to predicting the behavior of consequential loads within a local area grid.

## Acknowledgement

This research has been partially supported by a grant from EPRI/DOD, No. 530 1290-1573 (Complex Interactive Networks/Systems Initiative).

## References

1. Amin, M.: Toward self-healing energy infrastructure systems, *Comm. IEEE Comput. Appl. Power* **14**(1) (2001) 20–28.
2. Daubechies, I.: Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* **41** (1988), 909–996.
3. Mallat, S. and Hwang, W.: Singularity detection and processing with wavelets, *IEEE Trans. Inform. Theory* **38** (1992), 617–643.
4. Tsoukalas, L. H.: Intelligent Management of the electric power grid through an innovative anticipatory, multi-agent, high-performance computing approach, First CIMEG Annual Report, 1999.
5. Yamakawa, T., Uchino, E., and Samatsu, T.: Wavelet neural networks employing over-complete number of compactly supported non-orthogonal wavelets and their applications, in: *IEEE World Congress on Computational Intelligence*, Vol. 3, 1994, pp. 1391–1396.