Price Signal Analysis for Competitive Electric Generation Companies

James D. Nicolaisen, Charles W. Richter, Jr., Gerald B. Sheblé
Electrical and Computer Engineering Department
Iowa State University
Ames, Iowa 50011
Phone: (515) 294-3046, Facsimile: (515) 294-4263, E-mail: gsheble@iastate.edu

Abstract: Successful operation and bidding in the competitive electricity marketplace requires well-planned strategies. The appropriate strategy is dependent on the state of the system. Much data (including time series) is available, and a proper analysis of this data can provide insight in choosing the right strategies. Traditional data analysis techniques can be time consuming. Techniques that quickly analyze the data can assist in forecasting price and demand and identifying the present state of the market, which should help the savvy trader in reacting intelligently to the market before its competitors. Advanced data analysis techniques may reveal patterns in the data that may be very helpful in forecasting demand or price. This paper compares several techniques that may help in identifying useful patterns in relevant time series data. These patterns are keys or leading indicators of future electric utility price or demand of electricity. The importance of quickly identifying these signals will increase as competition increases. The techniques being investigated are Fourier and Hartley Transforms, Line Spectrum analysis using both Fourier Transforms and Hartley Transforms.

I. INTRODUCTION

Following successful deregulation efforts in the natural gas markets, airline industries, and telecommunications industries, the US has moved forward with efforts to deregulate electric power industry. The hope is that the industry as a whole with be more efficient. Power systems all around the world are moving from their former status as natural monopolies to something that resembles competition.

Open access allows entry of new generation companies (GENCOs) and energy service companies (ESCOs) into the market more easily than ever before. As in any competitive market, the market (formerly served entirely by one vertically integrated monopolistic utility) will be divided or shared among the participants.

Computer models allow research to be conducted without losing millions of dollars in "real world" experiments. Many models of the economies and agent interactions in various industries have been developed, and many of the principles learned from studying other markets apply to power systems. However, in an electricity market additional restrictions related to the physical properties of electricity (e.g., $P_G = P_D$, power generated equals power demanded) must be considered. Some of these additional considerations are: requirements for minimum production levels, startup and shutdown costs, inherent economies of scale of certain technologies, and the most efficient production level of equipment. In a competitive market, GENCOs strive to increase their profits (the positive difference between their selling price and their marginal cost). The consumers will attempt to purchase electricity at a price much below their marginal benefit. Strategic behavior is only one aspect of what makes market demand and prices difficult to forecast. There are additional aspects that must be considered in the model if we are to predict possible future market situations.

Most producers would benefit by knowing how much consumers actually need before it is purchased. During the last decade, companies have turned to just-in-time manufacturing to reduce costs. (Out of necessity the electricity industry has practiced just-in-time production since its inception.) Many theories of how to predict consumption *a priori* have been developed over the years. In the beginning, the most logical thing to do was record the amount consumed last time and produce at least that much again. For the most part, this method works well for a stable market with little deviation. However, the amount of electricity demanded by consumers is highly dependent on things like the weather (which tends to be not so predictable) and often varies wildly not only from day-to-day, but by the minute, hour, month, season, and year.

In a market driven by profit, the importance of accurately forecasting price and demand grows. Knowledge of market prices and demand can help in scheduling one's own production of electricity, and is important in identifying market arbitrage opportunities. Often there is much historical data available to help us predict the future, but it may not be a simple matter to extract the important information from the noise. Advanced methods are needed to improve the analysis of the recorded demand data. Artificial neural networks and other sophisticated techniques have come a long way, but there is room for improvement. Improving and combining techniques to improve forecasting and hence improve scheduling could save a single large GENCO millions of dollars per year.

This paper presents several signal analysis techniques and explains how they can be used in improving demand and price forecasts for the deregulated market. The remainder of the paper is organized as follows. Section II discusses how Fourier Transforms can be used to recognize patterns. Section III discusses the Hartley Transform. Section IV compares and contrasts the Fast Fourier Transform (FFT) and the Fast Hartley Transform (FHT). Section V deals with the value of time series analysis. Section VI discusses data analyses and artificial neural network forecasting. Finally, section VII presents some conclusions.

II. RECOGNIZING PATTERNS WITH FOURIER TRANSFORMS

An improvement over "produce what was produced last period" may be found by looking for patterns in the data, i.e. look up the demand yesterday at 4:00 and use that today. This technique is useful but it needs refinement. For example, one could look up the demand for August 19 of last year at 4:00 PM and use that for this year. Figure 1 is a plot of the typical hourly electricity demand for a particular week. Depending on the time frame of the prediction, comparisons of weekly data can be more useful in long term forecasting demand schedules than hourly daily data. This logic is carried out in some of the more advanced methods used today.

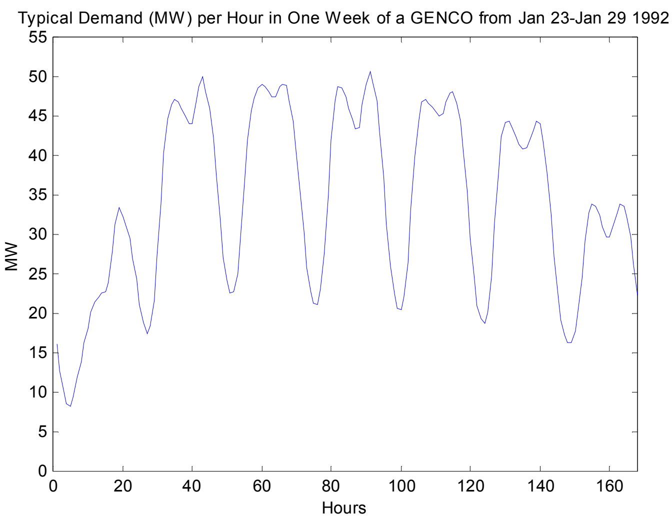

Figure 1 Demand of electricity per hour for one week

A number of procedures have been developed to re-present the data in such ways as to bring out the patterns normally not seen in the original plots. This data is considered time series data or discrete time series data because it's measured at constant time intervals. A popular approach to discover trends in time series data is to decompose into a set of sine waves of different frequencies. [17] The decomposition technique most often used is the Fourier Transform. To use the set of data obtained from the Fourier Transform, its magnitudes are squared and plotted versus frequency. An example of this is shown in Figure 2. This graph is called a periodogram, line spectrum, or power spectrum. It shows some interesting attributes of the data that might not be so obvious in the plain data plot. It can also confirm trends found by other time series analysis techniques, such as autocorrelation, partial autocorrelation, differences, and ARIMA (AutoRegressive Integrated Moving Average).
Much work has been done in forecasting using Artificial Neural Networks (ANN). Much of this work has produced very accurate forecasting methods. Forecasts made by neural networks can include factors, other than past data or patterns, such as the weather, which can greatly affect the electricity demand. The forecasting written about in this paper is for the times series analysis of hourly price or electric demand data.

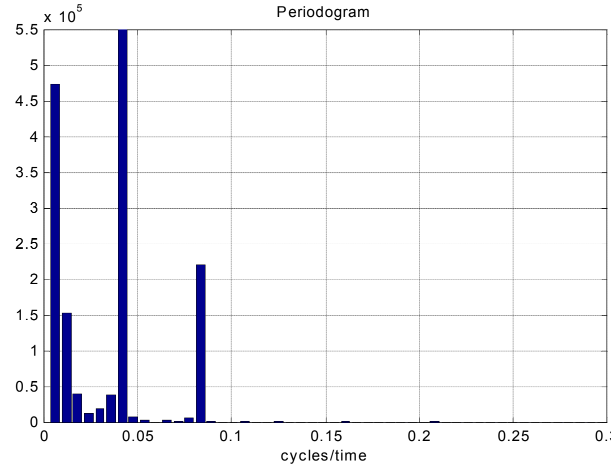

Figure 2 Periodogram using FFT

According to Makridakis [9], the main reasons for using periodograms are to help identify:

- randomness in the data series
- seasonality in a time series
- predominance of positive or negative autocorrelations

For positive autocorrelation low-frequency amplitudes should dominate, and for negative autocorrelation, high frequencies should dominate [9].

The Fourier Transform takes periodic data and finds the summation of sine waves at different frequencies that could be used to reconstruct the original data graph. Note in Figure 2 that only the first seven frequencies are needed to construct the original data plot for an acceptable degree of accuracy. The first fourteen is more than enough to match the real time series data. The result of the Fourier Transform comprises both real and imaginary numbers. Some computers and software programs have a very hard time dealing with imaginary numbers. The Fourier Transform switches the data from the time domain to the frequency domain. The complexity of the Fourier transformation of data results in excessive computation time and memory usage. Therefore, faster technique of calculating the Fourier Transform was developed to reduce this computer slow-down problem. It's called the Fast Fourier Transform (FFT). It is much faster because of the way it is calculated and it stores less information. The FFT function built into Matlab was used to generate Figure 2 from the data shown in Figure 1.

The periodogram, shown in Figure 2, confirms that the series is nonstationary because there is a dominance of low-frequency sine waves. A data series is stationary if its statistical properties are independent of the particular time period during which it is observed. According to Makridakis [9], the data from a stationary series fluctuates around a constant mean, independent of time, and the variance of the fluctuation remains essentially constant over time. Often enough, the times series data plot itself can convince a forecaster of it's stationarity but it can also be deceiving at times. The data in Figure 1 and the periodogram Figure 2 is clearly not random. The periodogram would have high amplitudes at random frequencies if the data were random.

Figure 2 also shows that the data is seasonal. Seasonality is a pattern that repeats itself over fixed intervals of time. Of course this data has the almost the same pattern every day for a week. The data plot easily exposes this concept. The periodogram can confirm the same conclusion but also provide the period for each season for all data. In Figure 2 there are seven amplitudes that dominate the low frequencies. Now, 168 data points divided by 7 gives the period of 24-hour seasonality. This 24-hour seasonality can be seen in Figure 5. The top plot in Figure 5 shows the period exactly at 24, which obviously corresponds to the day period of 24 hours. Again this was known a priori but maybe not if the data was electricity prices, which can fluctuate differently. These are some of the reasons way the Fourier Transform has been used for so long in time series analysis and forecasting.

III. HARTLEY TRANSFORMS

The Hartley Transform (**HT**) has been around since 1942. Instead of calculating a real and imaginary part, as does the Fourier, it remains in the real number set. This is done by the ***cas*** function, developed by Ralph Vinton Lyon Hartley [1]. It is defined as:

$$cas\,(vx) \equiv cos(vx) + sin(vx)$$

This kernel function eliminates the need for complex numbers, although it can be written with them. Many computers are not equipped to calculate with complex numbers and can cause serious problems. The extra memory used to keep track of the excess numbers for the complex calculations can significantly slow down processor speed. Bypassing the dilemma of imaginary numbers dramatically reduces the computational effort.

The Discrete Hartley Transform (DHT) uses the ***cas*** kernel function and is defined as follows:

$$H(v) = N^{-1}\sum_{\tau=0}^{N-1} f(\tau)cas\left(2\pi v / N\right)$$

The DHT requires only half the memory space for real data compared to the complex data. For large matrices, this can account for much of the lost space on hard drives.

An alternative way to calculate the Hartley Transform, assuming computation time is not a factor, is the real part of the Fourier Transform minus the imaginary part. This can make use of the FFT and gain some computer speed. A Fast Hartley Transform (FHT) was developed to reduce the time of computation even further. The existence of FHT is one of the reasons the why the Hartley class of transforms is currently being used in a vast number of applications including image processing, circuit analysis, power quality, signal processing, speech processing, fast convolutions, and multidimensional optics applications.

The Hartley Transform is completely reversible. The inverse of the transform is the exact data that was originally transformed. The Hartley Transform does not shift to the same frequency domain as the Fourier. There is a constant of $1/\sqrt{2\pi}$ included in the conversion to keep the transform reversible. This is why the Hartley Transform can be easily used as an alternative decomposition for time series data. A transform not required to use complex numbers is much easier to understand and learn. This is another reason why the Hartley Transform is starting to gain some momentum; it is being taught to beginning engineering students rather than teaching them the Fourier Transform. Figure 3 shows the periodogram of the demand data described earlier.

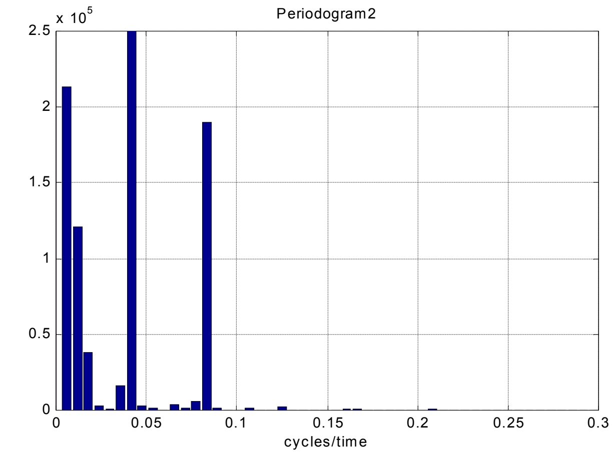

Figure 3 Periodogram using Hartley Transforms

IV. COMPARISON OF FFT AND FHT

As seen in Figure 3, the HT will produce a line spectrum very similar to the line spectrum produced by the FFT. The amplitude will not be the same but the general pattern is similar. Hence, the same trend information used for forecasting and modeling the systems can be obtained using the FHT. A comparison can be made from the plots presented together in Figure 4. As stated earlier, the amplitudes are not the same magnitude so the graphs have been scaled to show the details of each transform characteristics in their respective periodograms. The bars on the graphs match up very closely. The seasonality is nearly identical in the two transforms. Each has a season set of seven (the first set of amplitudes higher than their neighbors). The season for the data is 24 hours, as calculated earlier.

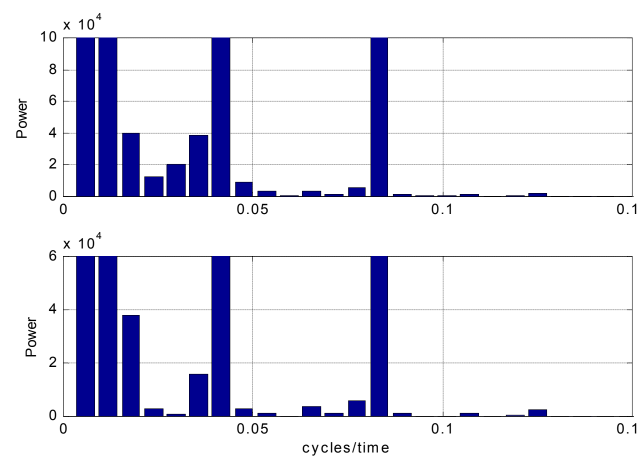

Figure 4 Comparison of FFT (top) and HT(bottom) periodograms

Another seasonality evaluation is shown in Figure 5. It shows a period exactly at the 24th mark in both the FHT version and in the FFT version. The other peak represents a smaller correlation to the 12-hour period. Of course, this can also be seen in Figure 4. After the initial seasonal set, the next peak is located at the 14th amplitude. The seasonal period corresponds to 12 hours, the same shown in Figure 5. When two or more weeks worth of data are fed into the FHT a season of 168 hours shows up. This is naturally expected and the precursor can easily be seen in Figure 5. The graph reaches the 168 mark but without more data it stop there. When a 2 or more weeks are involved a spike is shown as in the 12th and 24th.

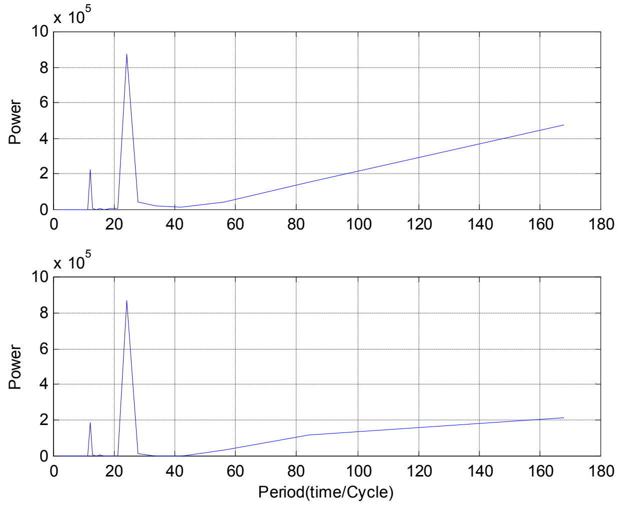

Figure 5 Season of the FFT periodogram (top) and HT (bottom)

There are many advantages that come with using the DHT instead of the DFT. The DHT performs fewer operations that may lead to less truncation and rounding errors from computer finite word length [4]. This can stem from different math co-processors and the version of the CPU. Obviously, when a math error occurs at the beginning of the calculations it will just become compounded as the calculations continue. A problem such as this can throw off any sort of accurate short-term forecasting. Long-term forecasting would develop trends that would not exist. With a competitive environment such mistakes can throw a company right into bankruptcy.

The compliment to the reduced-error argument is the complexity of the DFT as compared to the DHT. Not only is there a potential for errors because of rounding and truncation, there is a definite concern for the complexity of the DFT itself. Many people have a hard time learning the concept and algorithms associated with the Fourier transform. The switching from time domain to frequency domain develops problems for some engineers. Some people have a hard time understanding the real value with using imaginary numbers for practical purposes. The FHT is its own inverse which reduces the complication of reversing the transform considerable especially compared to the Fourier's extra algorithms. Forecasting should be easy enough for all employees to understand and use, not just the people with enough background and practice with Fourier transforms. In the competitive market, electricity price forecasting needs to be explained to the company executives, who may be inclined to go with the formula that does not include "imaginary numbers". For some applications such as circuit theory and power systems analysis there is an advantage of knowing the phase and magnitudes separately. The Hartley Transform has been developed to the point were it to can be used for such applications with the added advantage of the much faster FHT. These developments were due mostly because of the FHT. Its speed and reliability has given reason to the pursuit of other applications.

Another real advantage of the FHT is the speed of calculation compared to the DFT. The speed ratio between the DFT and the FFT convolutions is ρ/N, where N is the number of data points to be transformed and $N=2^{\rho}$. For example, if $N = 2^{10}$, the FFT would require less than 1% of the normal computing time [4]. "Timing studies have shown that for N greater than about 28, the FFT method is at least an order of magnitude faster" than a lagged products approach of calculating the same convolution [4]. Now consider that the FHT is quicker than the FFT--about twice as fast. The time saved in computation alone is a tremendous advantage over the competition.

The difference in memory is also a major consideration. The DFT physically has to use twice as much memory as the DHT because of the extra complex numbers. The DHT only calculates and stores the real numbers. The complex computations alone will slow down the computer but if storing to the hard drive is involved the machine will start to dramatically slow down. The more time wasted on excess calculation and storage means less time accurately analyzing the data. This time could have been more wisely spent on coming up with the bidding strategy as a result of the newly forecasted demand or price.

V. THE VALUE OF TIMES SERIES ANALYSIS

To use the FHT to in time series analysis is not a difficult process. The main purpose of this type of analysis is to find trends in the data. Depending on what type of data is being analyzed, it may discover trends on a monthly or weekly basis for base-load forecasts which can then be used as a signal to buy or sell given amounts of electricity. The FHT can be used to determine the load patterns of certain holidays throughout the year, a job for which it is difficult to train ANN because the holiday occurs only once a year leaving the analyst with little data. Another possible use is for the discovery of sudden changes in patterns meaning something has changed in the system. Such an event could mean that a competitor has brought a new unit online or has had changes in transmission capacity. The periodogram can be considered a fingerprint of the time period.

The steps in forecasting the demand next hour are as follows:

1. Plot the hourly demand for the week so far
2. Plot the demand for last week
3. Use the FHT to create the periodogram of each week
4. Find and plot the period and season of the data (the data should not be random)

5. If seasonal changes are minimal, if limited time has elapsed since the last measurement (ideally it is data from the previous period), and if the weather doesn't change much, then next period's electricity demand should be well-captured by the data.

As stated, these steps are quite general. The most time consuming part is usually to use a decomposition technique to analyze the time series data. With the FHT the task becomes much easier and quicker.

VI. DATA ANALYSES AND ANN FORECASTING

Artificial Neural Networks (ANN) have been studied and used for many years to help forecast events using a learned behavior of response to certain inputs. A common type of ANN is the fully connected feed-forward network as shown in Figure 6. The output of a node is high if the summation of the inputs is greater than some threshold value. If it is less than the threshold the output goes low. Often a sigmoid function is used to produce the output. In that case, the output may be somewhere between high and low depending on the strength of the inputs. The output of each node is carried to the next node via a connection. Each connection has a strength which by which the output is multiplied. Using many connections and many nodes allows the network to code a complex relationship that minimizes the sum-squared-errors over a set of input-output training pairs.

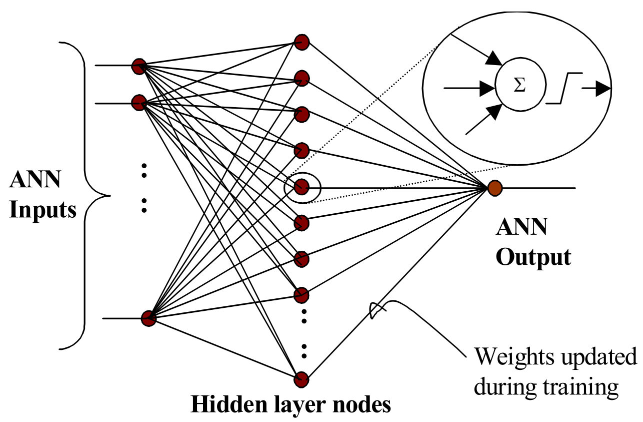

Figure 6 Fully connected MLP ANN

The training pairs should be similar to the data that will need to be forecasted later, and should contain sufficient information from which to draw conclusions about the functional mapping from the input space to the output space. The training period is used to adjust the weights (or how much it counts) via some learning technique. One common learning technique (during which weights are updated) is called back propagation, where connection weights are updated in proportion to the amount they contribute to the error. It is possible to over-train many ANNs, meaning that the network would only work well on the training data set, and would not perform well when presented with a new set of inputs. Following the presentation of input-output pairs, the training period is then stopped and the network should be ready for prediction on similar data patterns.

Providing the ANN with too much information (or the wrong information) can confuse the network in the beginning and it can settle on weights that are unable to handle variations of larger magnitude in the input data. When called upon to predict larger or smaller variations even when given many inputs about temperature (load forecasting) or currents trends (in price forecasting) the error will still increase. These increases in forecasting error can cost deregulated companies millions of dollars in fuel contracts or bilateral contracts. This is where the Hartley transform can be useful. It can be used to preprocess the data or be included in the training algorithm to adjust the weights of the ANN to speed up the process of finding the data that affects the prediction the most. As with the load data the correlation found at the periods of 12 and 24 are real seasons that affect the weights on the nodes in the neural network. These seasons will be found in pricing data and should be considered important signals for analysis.

There are a few ways of implementing the FHT into the neural network. The training data can be “filtered” by the FHT to find seasonalities and depending on the ranges can set the initial weights connecting nodes of the ANN. This is in lieu of initializing the weights randomly and updating them from scratch beginning with the first iteration of the training. This FHT filtering is a very quick process and will speed up the training process by cutting down on the number of iterations to reach the minimum absolute error in estimation.

One other possibility is to “filter” the load data through the FHT and have it identify the seasonality of the data. See Figure 7. These periods identify the hours having the more influence in the data. For example, the 24th hour and 12th hours can be used as heavier, weighted milestones to anchor the changing weights around. Using these hours as an input to the neural network as well as the load data may increase the accuracy of the forecast. This should focus the network by limiting the input data to those points which are most important and have it ignore less consequential data. This speeds up the process of the network closing in on the solution set.

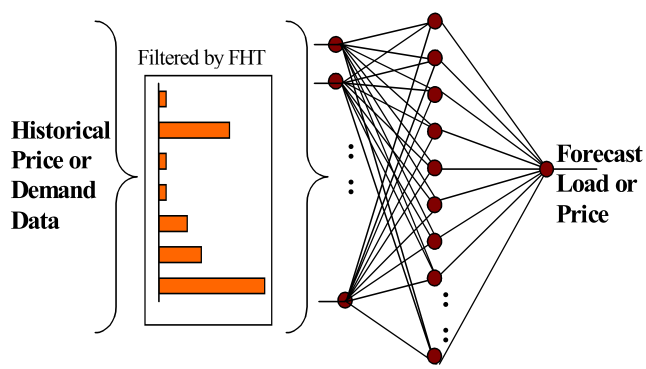

Figure 7 Using the FHT with an ANN to forecast demand or pices

VII. CONCLUSIONS

The Fast Hartley Transform is computationally less intensive and hence is much quicker than the Fourier Transform. The FHT is easier to understand because it avoids the complex arithmetic. With the FHT, the imaginary plane does not have to be considered when trying to model or forecast for time series demand or price data. The DHT will tend to be more error free with less complex calculations. The DHT is its own inverse, which reduces more complex arithmetic in switching between time and frequency domains. The FHT does everything the FFT does in times series analysis for forecasting and modeling electricity demand or market price of electricity. The FHT does what the DFT can do but with greater ease and speed. The FHT used in times series analysis can show what the competition might be doing to increase its market share. It is definitely an asset when quick accurate forecasting is essential in a competitive electric utility environment.

With the implementation of the FHT into neural network, the speed of the training will be increased as well as the reduction in prediction error. If the FHT can speed up the training and operation of an ANN which is employed continuously updating a forecasting model then the modeler would definitely have the advantage in a deregulated environment. Time is money, but with poor accuracy it's a detriment. The fastest and most accurate price signals will win the game.

VIII. ACKNOWLEDGEMENTS

The authors wish to acknowledge the National Science Foundation for their support in researching this topic.

IX. REFERENCES

[1] Bracewell R. N., 1986. The Hartley Transform. (New York: Oxford Press).
[2] Bracewell R. N., 1994. “Affine Theorem for the Hartley Transform of an Image,” Proceedings of the IEEE, Vol. 82, No. 3.
[3] Bracewell R. N., 1994. “Aspects of the Hartley Transform,” Proceedings of the IEEE, Vol. 82, No. 3.
[4] Hong J., Vetterli M., and Duhamel P., 1994. “Basefield Transforms with the Convolution Property,” Proceedings of the IEEE, Volume 82, No. 3.
[5] Hsu C. H. and Lin T. P., 1988. “Novel Approach to Discrete Interpolation Using the Subsequence FHT,” Electron. Lett. Vol. 24, pp 223-224.
[6] Jang, J-S. R., C-T. Sun, and Mizutani, E., 1997. Neuro-Fuzzy and Soft Computing: Computational Approach to Learning and Machine Intelligence.
[7] Khotanzad A., "Short-Term Load and Price Forecasting with Artificial Neural Networks" working paper, available from author.
[8] Loan C. V., 1992. Computational Frameworks for the Fast Fourier Transform. Philadelphia: (Society for Industrial and Applied Mathematics).
[9] Makridakis S., Wheelwright S. C., and Mcgee V. E, 1983. Forecasting: Methods and Applications. (New York: John Wiley and Sons).
[10] Millane R. P., 1994. “Analytic Properties of the Hartley Transform and their Implications,” Proceedings of the IEEE, Vol. 82, No. 3.
[11] Okankersoy, 1994. “A Comparative Review of real and complex Fourier-Related Transforms,” Proceedings of the IEEE, Vol. 82, No. 3, Mar.
[12] Olejniczak K. J. and Heydt G. T., 1994. “Scanning the Special Section on the Hartley Transform,” Proceedings of the IEEE, Vol. 82, No. 3, Mar.
[13] Olejniczak K. J. and Heydt G. T., 1991. “A Comparison of Alternative Transforms for Electric Power Engineering Applications,” Proc. 22nd Annual North American Power Symposium, pp. 84-93,
[14] Olejniczak K. J. and Heydt G. T., 1991. “Applications of the Hartley Transform for the Analysis of the Propagation of Nonsinusoidal Waveforms in Power Systems,” IEEE Trans. Power. Del., Vol. 6, pp 1862-1868.
[15] Papalexopoulos, A. D., Hao S., Peng T-M., 1994. "An implementation of a neural network based load forecasting model for the EMS." IEEE Transactions on Power Systems, Vol. 9, No. 4, Nov. 1994.
[16] Poularika A. D. editor, 1996. The Transforms and Applications Handbook. (Florida: CRC Press).
[17] Villasenor J. D., 1994. “Optical Hartley Transforms,” Proceedings of the IEEE, Vol. 82, No. 3.
[18] Welsh, D. E., 1993. "Training of artifical neural networks via interior point algorithms." Iowa State University Masters Thesis.

X. BIOGRAPHIES

James Nicolaisen received his BS in Physics from Morningside College in northwestern Iowa in 1998 and is presently studying power systems while pursuing an MS in EE at Iowa State University.

Charles W. Richter, Jr. received his BSEE from South Dakota State University in 1992. After working at a nuclear power plant and a consulting firm, he studied power systems at Iowa State University where he earned an MS (1996) and Ph.D. (1998) in power systems. Presently he is a temporary assistant professor at ISU. Professional interests include economics of power system operations, market modeling and complex adaptive systems.

Gerald B. Sheblé (M 71, SM 85, F 98) is a Professor of Electrical Engineering at Iowa State University, Ames, Iowa. Dr. Sheblé received his BS and MS degrees in Electrical Engineering from Purdue University, and his Ph.D. in Electrical Engineering from Virginia Polytechnic. His more than fifteen years of industrial experience include projects with public utilities, research and development, computer vendors and consulting firms. His research interests include power system.