Forecasting Energy Product Prices

M. E. Malliaris
School of Business Administration, Assoc. Prof.
Loyola University Chicago
Chicago, IL
E-mail: mmallia@luc.edu

S. G. Malliaris
Department of Mathematics, Student
Massachusetts Institute of Technology
Cambridge, MA
E-mail: sgmallia@mit.edu

Abstract. **Five inter-related energy products are forecasted one month into the future using both linear and nonlinear techniques. Both spot prices and data derived from those prices are used as input data in the models. The models are validated by running data from the following year. Results show that, even though all products are highly correlated, the prediction results are asymmetric. In forecasts for crude oil, heating oil, gasoline and natural gas, the nonlinear forecasts were best while for propane, the nonlinear model had the largest average absolute error.**

I. INTRODUCTION.

Crude oil, heating oil, gasoline, natural gas, and propane are five energy products whose prices are inter-related. No one product is independent of all others in either usage or production. Crude oil and natural gas are both raw energy products; they come out of the ground. The former can be refined to produce heating oil and gasoline, and one component of gasoline is propane. The latter (natural gas) can also be processed to give propane, but only in limited amounts. Overall, propane is sourced in approximately equal amounts from 1) gasoline derived from crude oil refining and 2) natural gas processing [9]. Crude oil, the most actively traded commodity in the world, breaks specifically into the following products (Figure 1, [6]). Similarly, the natural gas breakdown is shown in Figure 2.

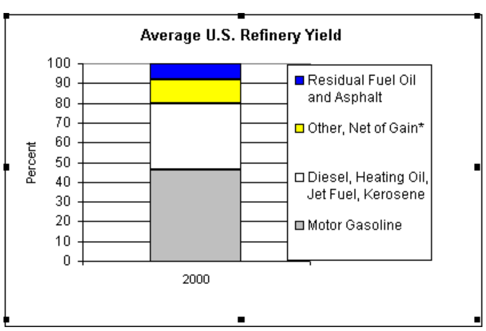

Figure 1. Yield from a barrel of crude oil[1]

At present natural gas accounts for about a quarter of US energy use, but large consumers of energy (factories, power plants) often absorb an up-front cost to implement systems allowing, with minimal change-over cost, the use

TYPICAL NATURAL GAS COMPONENTS

COMPONENTS	PERCENT BY VOLUME
Methane	89.5 - 92.5
Ethane	5.1 - 2.0
Propane	2.1 - 0.7
Butanes	1.6 - 0.5
Nitrogen	0.6 - 1.9
Carbon Dioxide	1.1 - 2.4
Total	100.0 - 100.0

Figure 2. Components of natural gas[2]

of either fuel oil or natural gas [17] as their main energy source as a hedge against fluctuating market prices and uncertain availability. Because of this substitute relationship, demand for either energy source is more elastic than one would otherwise predict—and, as one would expect, one of the primary determinants of the price of natural gas is the price of oil [18]. Table 1 shows the high correlations between the above-mentioned energy markets.

TABLE 1. Correlation between energy product prices.

	CL	HO	PN	HU	NG
CL	1.00	-	-	-	-
HO	0.96	1.00	-	-	-
PN	0.84	0.88	1.00	-	-
HU	0.96	0.93	0.85	1.00	-
NG	0.67	0.73	0.68	0.66	1.00

Academic and industry research and position papers differ as to their explanations for the causes and effects of both absolute energy prices and the relationships therein. Asche [1] found a long-run relationship between crude oil and gas oil prices, but no similar relationship between crude oil and fuel oil. Brown [3] found that rising oil prices hurt economic activity and preceded nine of the ten recessions

[1] Note: Processing gain is the volume increase that results as denser molecules (e.g., residual fuel oil) are split into less dense ones (e.g. gasoline). The processing gain in U.S. refineries is equal to about 6%. See: http://www.eia.doe.gov/pub/oil_gas/petroleum/analysis_publications/oil_market_basics/graphs_and_charts.htm

[2] Roanoke Gas Company, http://www.roanokegas.com/aboutus/origin.html

0-7803-9048-2/05/$20.00 ©2005 IEEE

following World War II. The U.S. government uses over 100 regression equations to form a system of forecasting equations [11] for prices of various forms of energy. Kasprzak [14] has had better results forecasting jet fuel prices with a neural network than with regression. McMenamin [15] has shown similar success of neural networks over regression in short term energy forecasting. However, Serletis has shown that the natural gas market is chaotic and thus prediction is unlikely. Current measures indicate that the potential for an energy crisis is historically high [23]. OPEC, often a target of criticism about rising prices in energy products, acknowledges a relationship between gasoline, heating oil, and crude oil, but believes the relationship is neither direct nor proportional, and claims that rising prices are due to taxes rather than OPEC policies [13]. Thus, there is much uncertainty about energy prices in the future [5]. This uncertainty in energy prices commands a great deal of foreign and domestic political attention, and facilitates an active market ranking second only to financial products[3] in amount of trading on futures contracts. As a result, energy commodity price risk has a dominant role in the energy industry [7,9]. Herard and Taylor [12] emphasize that a first step in minimizing earnings volatility in industry is to stabilize margins; a major reason for the margin volatility is the price change in natural gas. Therefore, by accurately forecasting prices into the future, we can create an effective way of managing this risk. This paper looks at a short-term forecast (one month) and compares the ability of two models to handle that forecast. More specifically, we use information from five energy markets to forecast each of those markets 21 trading days from the day the forecast is made. Data was specifically kept simple as we wanted to focus on the effect of recent price change on pricing one month away.

II. Data, Variables, Data Sets

A. Data.

The data consists of daily spot prices for Crude Oil (CL), Heating Oil (HO), Gasoline (HU), Natural Gas (NG), and Propane (PN). All the data are original data, provided by Barchart at www.barchart.com, in absolute values. HO is in cents per gallon, e.g. for 1/3/1994 the price was $ 0.453/gallon. PN is in cents per gallon, e.g. for 1/3/1994 the price was $ 0.2500/gallon. NG was measured in dollars per MMBtu, e.g. the price for 1/3/1994 is $2.050/MMBtu. HU was in cents per gallon, e.g. for 1/3/1994 the price is $0.3970/gallon. Finally, CL is in dollars per barrel, e.g. for 1/3/1994 the price was $14.54/barrel. The initial data set contained daily spot prices for the period between Jan 3, 1994 and Dec 31, 2002. Inspection of the data over time showed that the relationship between the five markets changed significantly between 1994 and 2002. They became more correlated in movement. Anticipating that energy markets will continue to become more rather than less inter-related, and attempting to build a stable model, we used data from December 1997 through November 2002 for this study.

B. Input and Output Variables

The input variables for each model consisted of the daily closing spot price, percent change in daily closing spot price from the previous day, standard deviation over the previous 5 days, and standard deviation over the previous 21 days, for each of the five markets.

For the Crude Oil variables, for example, these variables were labeled as CL, CLchg, CLStdDv5, and CLStdDv21. This gave us a total of 20 initial input variables. In addition, the neural network model used a cluster indicator. This is a non-numeric variable indicating the cluster group to which each row belonged. The output variable was the daily spot price one month into the future (21 trading days). Again using Crude Oil as an example, this variable was labeled CLplus21. Models were built for each of the five markets. The table below shows the correlation of each energy commodity price today with each price 21 trading days into the future. These values range from a low of .6599 to a high of .9293.

TABLE 2. Correlation of today's spot prices with prices 21 trading days away.

	CL	HO	PN	HU	NG
CLplus21	0.93	0.90	0.80	0.90	0.64
HOplus21	0.92	0.92	0.82	0.89	0.65
PNplus21	0.82	0.83	0.88	0.81	0.57
HUplus21	0.89	0.87	0.93	0.85	0.64
NGplus21	0.72	0.77	0.73	0.68	0.66

Notice that most of the correlations of one product's prices today with another product's prices 21 days away are slightly lower than those shown between those same products in Table 1, with the exception of natural gas. For example, the correlation between HO and CL is .9597, while the correlation between HO and CLplus21 is .9012 and between CL and HOplus21 is .9242. Though slightly lower, they are still good enough to encourage us to attempt a model.

C. Training and Validation Sets

The data set ran from December 1997 through November 2002, and contained 25 columns. This was split into disjoint training and validation sets. The training set contained four years of data, from December 1997 through November 2001, and had 1001 rows. The remaining year of data, December 2001 through November 2002 was used for the validation set. This set contained 247 rows. Twenty of the columns were used for input. Of the remaining five columns, one was selected at a time as the output for each model. The remaining four were unused for that model. For the non-linear model, another input

[3] US Department of Energy statistic

column was generated by the K-Means algorithm, a cluster identifier.

The models were judged by their results on validation data that they had not seen during training and which occurred after the end of the training period.

III. Models and Methodology

A. Models

Two model types, one linear, and one non-linear were built for the forecasts. A multiple regression model was built for each of the five energy markets with the spot price 21 trading days into the future as the dependent variable. As variables indicated lack of importance to the model, or multicollinearity, they were deleted and the model was re-run. A final model for each market was then used for the validation set forecasts. The variables used in each of the multiple regression models are shown in Table 3.

TABLE 3. Variables used in each final regression model.

Variables	**CL**	**HO**	**HU**	**NG**	**PN**
CL	x	x	x		x
CLchg	x	x			x
CLStdDv5		x			x
CLStdDv21	x	x	x	x	
HO		x		x	x
HOchg	x				
HOStdDv5	x		x	x	x
HOStdDv21		x	x		x
HU		x	x		
HUchg	x	x			x
HUStdDv5	x	x		x	x
HUStdDv21				x	
NG	x	x	x	x	x
NGchg			x		
NGStdDv5					x
NGStdDv21	x	x	x		x
PN	x	x		x	
PNchg			x		x
PNStdDv5				x	x
PNStdDv21		x	x	x	x

The number of variables used in each of the final regression models ranged from nine to fourteen. Only NG appeared in every model. The variables appearing in at least four of the five models were CL, CLStdDv21, HOStdDv5, HUStdDv5, NGStdDv21, and PNStdDv21. Five out of these six variables are based on movement within the market the past week or month.

Following the development of each of the regression models, the SPSS data mining package Clementine was used to develop a second set of models. These models were non-linear and used two processes in sequence. The data was first run through a K-Means clustering algorithm. In the clustering algorithm, the numerical records were grouped by similarity and a center for the group is calculated. This center is the arithmetic mean of the records in the cluster and distance refers to the distance between cluster means. The K-Means algorithm in Clementine generated 5 distinct clusters. Three of these clusters were fairly small, and the remaining two clusters were large. The number of rows of the training set per cluster are shown in Table 4. The distance between cluster means, also called cluster proximity, is shown in Table 5.

TABLE 4. Cluster centers and population.

Cluster	CL#1	CL#2	Cl#3	CL#4	CL#5
Size	482	17	384	58	60

TABLE 5. Distance between clusters.

	CL#1	CL#2	Cl#3	CL#4
CL#1	-			
CL#2	2	-		
Cl#3	0.92	1.42	-	
CL#4	1.25	1.4	0.64	-
CL#5	1.38	0.96	0.69	0.74

Once clusters have been developed on the training data, they are used for all new data. That is, a new row is fed thru the trained K-means model and a cluster number is assigned to it. The resulting cluster assignment becomes an additional input into the neural network model. A neural network model was developed for each of the five markets. Neural networks have been successful in forecasting because of their ability to forecast difficult types of relationships well, given the correct inputs and training set. See, for example, [14, 15, 19, 21, 22, 24] for papers using neural networks in forecasting problems.

Each neural network model used twenty-one inputs (the 20 original fields, plus the cluster identifier), one hidden layer with twenty nodes, and one output node. When training ended, Clementine also generated a sensitivity analysis that listed the variables in their order of importance to the model. These results are shown in Table 6 with the variables in each column in decreasing order of importance to that model. Note that in four of the models, heating oil price was a prominent variable. If we look at the top five variables in each of the models, we find little agreement among them. In particular, the CL model has no variable in the top five in common with any other model. HO is in the top five of each of the other models. PN and CL are in the top five of three of the models. Thus, they are using different variables sets to generate the best forecast.

B. Methodology

A regression model and neural network model was developed for each of the five energy commodities. Data was used, for training the models, from December 1997 through November 2001, for a total of 1001 rows of data.

TABLE 6. Variable Significance in decreasing order in Neural Network Models

CL	HO	HU	NG	PN
Cluster	HO	CL	HO	HO
PNStdDv5	NGStdDv5	HO	PN	CL
Huchg	NGStdDv21	NGStdDv21	CL	NGStdDv5
NGchg	PN	CLStdDv5	NG	PN
HU	PNStdDv5	Cluster	HUStdDv5	NGStdDv21
HUStdDv5	CL	NG	PNStdDv5	PNStdDv5
NGStdDv21	Cluster	PNStdDv5	PNStdDv21	HUStdDv21
CL	PNStdDv21	NGStdDv5	NGStdDv21	CLStdDv5
NG	HU	PNStdDv21	Cluster	HU
HOStdDv21	CLStdDv5	HUStdDv21	HUStdDv21	Cluster
PNStdDv21	Huchg	HU	CLStdDv21	Huchg
CLStdDv21	HOStdDv21	HOStdDv21	CLStdDv5	HUStdDv5
CLchg	HUStdDv5	PNchg	NGStdDv5	NGchg
HUStdDv21	CLStdDv21	NGchg	Huchg	Hochg
Hochg	HUStdDv21	PN	NGchg	PNchg
HOStdDv5	Hochg	Huchg	HOStdDv5	PNStdDv21
NGStdDv5	HOStdDv5	CLchg	HU	HOStdDv21
HO	NG	CLStdDv21	Hochg	CLStdDv21
CLStdDv5	CLchg	Hochg	PNchg	HOStdDv5
PN	PNchg	HUStdDv5	CLchg	CLchg
PNchg	NGchg	HOStdDv5	HOStdDv21	NG

The regression models were refined until they showed no multicolinearity. The non-linear approach used a K-Means clustering algorithm, then a neural network to develop a forecasting model. A neural network model was trained and saved for each of the five products. Next, a validation set, comprised of a year's worth of data (247 rows) that occurred in time after the training data, was run through each of the trained models. This validation set data, for the nonlinear model, ran first through the trained clustering algorithm where each row was assigned a cluster value. The data then flowed into the trained neural network and a forecast was generated. Regression and neural network models were compared by looking at the mean absolute error and the mean squared error of each model. These statistics from the validation sets were used as the measures of ability of the linear and nonlinear models to forecast the five markets. In addition, a simple forecast using no model was generated by assuming no change between today a month from now. This simplistic approach was added just for comparison purposes. The same statistics were calculated for this simplistic view.

IV. MODEL RESULTS

A. Results on Training Sets

The regression statistics on the training data for each of the models are shown in Table 7. As can be seen in the table, all models did well on the training set, with the lowest R Square going to Propane. Table 8 gives the coefficients of each of the regression equations. Table 9 shows the relative importance of each of the variables in the non-linear model. With a neural network, the higher the number attached to an input variable, the greater the importance of that variable in determining the value of the output variable. Variables whose importance are greatest in the neural network forecast (over .1) are shown in bold-face type in the table. Notice that NG, while important in every regression model, is more important in only two of the neural network models. Also, in the CL model, the cluster affinity variable in much greater in significance than any other variable in the model. The neural network results on the training data are shown in Table 10. This table gives the estimated accuracy of the model along with typical statistics generated by the network. The estimated accuracy for each model compares well with the regression results. On the training data, both models do rather well.

TABLE 7. Regression Statistics

	CL	HO	HU	NG	PN
Multiple R	0.95	0.95	0.92	0.91	0.88
R Square	0.90	0.90	0.84	0.83	0.78
Adjusted R Square	0.90	0.90	0.84	0.83	0.78
Standard Error	2.20	0.07	0.08	0.66	0.07

TABLE 8. Regression Coefficients.

Variable	CL	HO	HU	NG	PN
Intercept	1.28	-0.03	0.06	-0.48	0
CL	0.77	0.02	0.01		0.01
CLChg	-6.22	-0.29			-0.16
CLStdDv5		0.03			-0.01
CLStdDv21	1.15	0.02	0.02	-0.35	
HO		0.37		3.81	0.42
HOChg	5.25				
HOStdDv5	29.3		0.74	-4.01	0.92
HOStdDv21		-0.51	0.92		-0.56
HU		-0.22	0.27		
HUChg	7.85	0.25			0.15
HUStdDv5	-32.4	-0.75		5.37	-1.04
HUStdDv21				9.94	
NG	0.62	0.02	0.05	0.74	0.02
NGChg			-0.08		
NGStdDv5					0.09
NGStdDv21	-4.35	-0.14	-0.23		-0.24
PN	5.01	0.14		-1.63	
PNChg			0.14		0.09
PNStdDv5				-2.82	0.29
PNStdDv21		-0.5	-0.29	-13	-0.25

TABLE 9. Relative importance of neural network inputs.

Input	**CL**	**HO**	**HU**	**NG**	**PN**
CL	0.03	**0.16**	**0.40**	**0.15**	**0.25**
CLchg	0.02	0.02	0.03	0.00	0.01
CLStdDv21	0.02	0.05	0.03	0.06	0.03
CLStdDv5	0.01	0.09	**0.17**	0.06	**0.12**
HO	0.01	**0.36**	**0.38**	**0.24**	**0.36**
Hochg	0.02	0.02	0.02	0.02	0.06
HOStdDv21	0.03	0.07	0.08	0.00	0.05
HOStdDv5	0.02	0.02	0.01	0.04	0.02
HU	0.04	**0.14**	0.08	0.02	**0.12**
Huchg	0.06	0.08	0.04	0.05	0.10
HUStdDv21	0.02	0.04	0.09	0.08	**0.14**
HUStdDv5	0.04	0.07	0.02	**0.12**	0.09
NG	0.03	0.02	**0.14**	**0.14**	0.00
NGchg	0.04	0.01	0.07	0.04	0.07
NGStdDv21	0.04	**0.27**	**0.30**	0.09	**0.17**
NGStdDv5	0.01	**0.28**	**0.12**	0.06	**0.23**
PN	0.00	**0.23**	0.06	**0.19**	**0.23**
PNchg	0.00	0.01	0.07	0.01	0.05
PNStdDv21	0.02	**0.14**	0.10	**0.11**	0.05
PNStdDv5	0.06	**0.18**	**0.14**	**0.11**	**0.14**
Cluster	**0.20**	**0.14**	**0.15**	0.08	**0.11**

TABLE 10. Training Data Results with Neural Network

	CL	HO	HU	NG	PN
Minimum Error	-11.52	-0.20	-0.32	-2.46	-0.22
Maximum Error	10.40	0.40	0.48	4.75	0.39
Mean Error	-2.15	-0.04	-0.02	-0.16	0.01
Mean Absolute Error	4.73	0.10	0.10	0.64	0.06
Standard Deviation	5.18	0.11	0.12	0.98	0.09
Linear Correlation	0.77	0.86	0.79	0.79	0.82
Estimated Accuracy	93.76	95.44	96.08	97.06	95.68

B. Results on Validation Data

After the development of the models on the training set data, the next step was to run the validation set data through each of the models. The validation set data occurred in time after the end of the training data, so it is a good measure of the ability of each of the models to forecast into the future. As a measure for each regression and neural network model, we used the mean absolute error and the mean squared error. A third forecast was created using no model and was made by assuming that the market in one month would be the same as it was today. This forecast is labeled the Simple Model. The same error quantities were calculated for this model. Table 11 shows the mean absolute error and mean squared error for each of the models developed. With the exception of the models for Propane, the neural network models gave a forecast with smaller error over the year. For forecasting propane, the best results came from the Simple Model. Best model results for each product are shown in bold.

TABLE 11. Comparison of Simple Model, Regression and Neural Network Errors

	Avg. Absolute Error			Mean Squared Error		
	Sim	Reg	NN	Sim.	Reg	NN
CL	1.973	2.13	**1.12**	6.01	6.65	**2.269**
HO	0.051	0.06	**0.04**	0	0.01	**0.002**
HU	0.057	0.05	**0.03**	0.01	0	**0.001**
NG	0.388	0.41	**0.22**	0.24	0.24	**0.075**
PN	**0.041**	0.06	0.08	**0**	0.01	0.009

V. CONCLUSIONS, SUGGESTIONS FOR FUTURE RESEARCH

While the nonlinear models derived by the neural network provided superior forecasting in the majority of cases, there remain undeniable asymmetries in the predictive abilities of the three models examined. For crude oil, heating oil, gasoline, and natural gas, the neural network gave the best results, consistently boasting a mean squared error less than half that of the regression or simple predictions. However, with propane, the neural network gave the least accurate prediction. Also surprising were the results of the simple model, which predicted that the commodity price 21 trading days into the future would hold unchanged from the present day's price. Except in the case of gasoline, the mean squared error of the simple prediction was lower than that of the regression model.

One can draw a number of conclusions from these results. First, it is clear that in a number of—though by no means in all—cases, there is enough information contained in a set of price data to allow effective forecasting. That is to say, while the neural network had no extraneous knowledge (say, of news items regarding energy consumption, or foreign wars, or what season it might be), it was nonetheless able to make a reasonably accurate prediction for four out of five energy products.

Second, an ability to predict the price of a given source product does not necessarily imply an ability to predict the price of such a product's byproducts. This asymmetry is exhibited by the neural network model's inability to effectively price propane futures. Propane, as we recall from the introduction, is produced by processing or refining natural gas or crude oil; the model is generally capable of predicting prices for both these raw materials. But when we try to turn the same model to predicting the price of propane itself, not only does the model produce less-than-stellar results, but it would be bested both by a standard regression and by a trader following the simplistic idea that prices would remain constant.

Lastly, we observe that traditional statistical techniques for understanding and extracting information about trends are often less than ideal in market situations, despite their otherwise broad acceptance by the profession. Regression analysis might be a common tool in analysis of energy market movements [2,4,10,11,16], but the results show that in energy markets, regressions are not appreciably better at predicting future movements than models making the blanket assumption of a flat market. This paper has concerned itself with an examination of the performance of models simulating the actions of a trader concerned with only the most basic of information: present and historical commodity prices. As we have seen, while such simple information can often provide comparatively excellent predictions, some surprising asymmetries remain. Future work along these same lines might begin to explain the root causes of these asymmetries by examining more closely the relationships between these various energy products, considering specifics such as time-to-market, refining and processing costs, change-over costs for energy users switching from one fuel to another, and volume of usage and cyclical or seasonal factors.

REFERENCES

[1] Asche, F., O. Gjolberg, and R. Volker, "Price Relationships in the Petroleum Market: An analysis of crude oil and refined product prices", Discussion Paper #D-04/2001, Department of Economics and Social Sciences, Agricultural University of Norway.

[2] Baker, Jeff, "Leading Methods for Forecasting Energy Demand", Energy Business Solutions, www.new energyassoc.com/documents/forecasting_wp.pdf.

[3] Brown, S., M. Yucel, and J. Thompson, "Business Cycles: The Role of Energy Prices", *Federal Reserve Bank of Dallas*, Research Department Working Paper 0304.

[4] Craig, P., A. Gadgil, and J. Koomey, "What Can History Teach Us? A Retrospective Examination of Long-Term Energy Forecasts for the United States", Annual Review of Energy and the Environment, 2002, 27:83-118.

[5] Department of Energy, "Annual Energy Outlook 2004 with Projections to 2025", www.eia.doe.gov/oiaf/ aeo/economic.html

[6] Energy Information Administration, "A Primer on Gasoline Prices", www.eia.doe.gov/pub/oil_gas/ petroleum/ analysis_publications

[7] Energy Information Administration, "Derivatives and Risk Management in the Petroleum, Natural Gas, and Electricity Industries", U.S. Department of Energy, October 2002, SR/SMG/2002-01.

[8] Energy Information Administration, "Energy Price Impacts on the U.S. Economy", April 2001, www.eia.doe.gov/oiaf/economy/energy_price.html

[9] Energy Information Administration, "Propane Prices, What Consumers Should Know", tonto.eia.doe.gov/oog/info/hopu/hopu.asp.

[10] EIA, "Short-Term Energy Outlook, Energy Prices Model Description", www.eia.doe.gov/emeu/ steo/pub/document/ textpr.html.

[11] EIA, "Short-Term Integrated Forecasting System", www.eia.doe.gov/bookshelf/models2002/ stifs.html.

[12] Herard, Bruce and D. Taylor, "Reducing earning volatility", *Chemical Market Reporter*, Sept 8, 2003.

[13] Heselbarth, R, "Oil prices hitting homes", *Contractor*, July 2000.

[14] Kasprzak, M. "Forecasting Jet Fuel Prices Using Artificial Neural Networks", Master's thesis, Naval Postgraduate School, Monterey, Ca., 1995.

[15] McMenamin, J. and F. Monforte, "Short Term Energy Forecasting with Neural Networks", *The Energy Journal*, Vol. 19, No. 4, Oct. 98.

[16] New York State Energy Plan, "All Fuels Demand and Price Forecast Methodology", June 2002, www.nyserda.org/sep/sepappendix.pdf.

[17] NYMEX [New York Mercantile Exchange], "Risk Management with Natural Gas Futures and Options, www.nymex.com.

[18] NRG News Energy Shop, "What Drives Prices?", May 2004, www.energyshop.com/energyshop/ pricearticles.cfm

[19] Olson, Dennis, and C. Mossman, "Neural network forecasts of Canadian stock returns using accounting ratios", *International Journal of Forecasting*, Vol. 19, Issue 3, July-September 2003, pages 453-465.

[20] Serletis, A. and P. Gogas, "The North American Natural Gas Liquid Markets Are Chaotic", *The Energy Journal*, Vol 20, No. 2: 83-103, 1999.

[21] Smith, K, and J. Gupta, Neural Networks in Business: Techniques and applications, Idea Group Publishing, 2002.

[22] Trippi, R. and E. Turban, Neural Networks in Finance and Investing, revised edition, Irwin, 1996.

[23] Williams, J. and A. Alhajji, "The Coming Energy Crisis?", *Energy Economics Newsletter*, 2003, www.wtrg.com/EnergyCrisis/index.html.

[24] Zhang, G.P. (ed), Neural Networks in Business Forecasting, Information Science Publishing, 2003.