Further Advances in Forecasting Day-Ahead Electricity Prices Using Time Series Models

Hany S. Guirguis* and Frank A. Felder†

Abstract - Forecasting prices in electricity markets is critical for consumers and producers in planning their operations and managing their price risk. We utilize the generalized autoregressive conditionally heteroskedastic (GARCH) method to forecast the electricity prices in two regions of New York: New York City and Central New York State. We contrast the one-day forecasts of the GARCH against techniques such as dynamic regression, transfer function models, and exponential smoothing. We also examine the effect on our forecasting of omitting some of the extreme values in the electricity prices. We show that accounting for the extreme values and the heteroskedactic variance in the electricity price time-series can significantly improve the accuracy of the forecasting. Additionally, we document the higher volatility in New York City electricity prices. Differences in volatility between regions are important in the pricing of electricity options and for analyzing market performance.

Keywords: forecasting, electricity prices, GARCH, volatility, extreme values

1. Introduction

In many parts of the world, the electric power industry is using competitive markets to meet consumers' demand for electricity. Accurate forecasts of prices are critical for producers and consumers in planning their operations and managing their price risk.

Recent work summarized the literature on electricity forecasting and in particular the use of time-series analysis and provided some accurate and efficient price forecasting tools such as dynamic regression model (DRM) and transform function approach (TFA) [1]. We extend this work by employing the generalized autoregressive conditionally heteroskedastic (GARCH) method, among others, to forecast electricity prices in New York City (NYC) and Central New York State (CNYS). Additionally, we incorporate prices for oil and natural gas, two fuels used by marginal generation units, into our forecasting models. Finally, we calculate volatility estimates, which are important in pricing electricity options and have important implications for analyzing market performance.

Price volatility is one of several inputs in calculating the value of an option, e.g., by using the Black-Scholes equation. In addition, differences in volatility in subregions may indicate two separate electricity markets, perhaps due to transmission constraints and resulting higher production costs. Being able to identify when such separations occur is important because the separation of a generally competitive market into smaller markets provides the necessary but not sufficient conditions to exercise market power.

We selected two regions of New York State for analysis to compare our forecasting and volatility results between a subregion of the state that is transmission-constrained and may be subject to the exercise of market power (NYC) with a subregion of the state that has surplus generation and where the exercise of market power is not a serious concern (CNYS).

Conventionally, the econometric modeling such as DRM and TFA assumes a constant one-period forecast variance. Most of the economic time-series, however, violate the classical assumption of constant variance (homoskedasticity). Many time-series exhibit periods of large volatility followed by periods of relative tranquility. Thus, the variance at time t might depend on past information. As a result, if additional information from the past were allowed to affect the forecast variance, one might expect better forecast intervals [2]. In 1982, a new model with mean 0, and serially uncorrelated with nonconstant variances conditional on the past but with constant unconditional variance, was introduced [2].

Linear autoregressive conditional heteroskedastic (ARCH(q)) models consist of two equations. The first equation fits the time series to the best autoregressive moving average (ARMA) specification, whereas the second equation models the conditional variance as an AR(q) process where q is the square of estimated residuals calculated from the first equation. Since ARCH(q) requires long lags in the modeling of many applications, Bollerslev

* Dept. of Economic and Finance, School of Business, Manhattan College. (hany.guiriguis@manhattan.edu)
† Corresponding Author: Center for Energy, Economics & Environmental Policy, Edward J. Bloustein School of Planning and Public Policy, Rutgers, The State University of New Jersey. (ffelder@rci.rutgers.edu)
Received May 13, 2004 ; Accepted August 23, 2004

introduced a GARCH(p,q) model in which both autoregressive (p) and moving average components (q) in the heteroskedastic variance are included in the current conditional variance equation to allow for a more flexible lag structure [3]. More precisely, the ARCH model allows limited number of lagged shocks to affect the conditional variance, whereas the GARCH model allows all the lags to affect the conditional variance by incorporating both the lagged values of the squared errors and the lagged conditional variance.

Engle, Lilien, and Robins extended the basic ARCH and GARCH model to allow the mean of a sequence to depend on its own conditional variance and is called ARCH-M and GARCH-M [4]. Since the introduction of ARCH and its variations, there have been many applications utilizing such models in different economic and finance settings. Readers not familiar with ARCH, GARCH, and related variations are referred to [2] and [3].

2. Time-Series Analysis

A general description of the assumptions used in time-series models and the time-series modeling approach as applied to electricity markets is provided in Nogales et al., 2002.

2.1 The New York State Electricity Markets

New York State implemented wholesale electric markets on November 18, 1999. It has a day-ahead market and a real-time market for energy based on locational marginal pricing. The New York Independent System Operator (NYISO) collects start-up, no-load, and up to ten energy bids in dollars per megaWatt-hours ($/MWh) that span the output of each generation unit. The NYISO performs unit commitment based on these day-ahead bids and clears for each hour for the next day (the day-ahead locational prices), which will be paid to selected generators and charged to day-ahead load. A real-time market occurs within the day to accommodate any system changes.

Power flows are generally west to east and north to south as lower cost generation units outside of the greater New York City region export power to New York City and Long Island. The NYISO defines the NYC and CNYS zones. Differences in spot electricity prices between these two regions of the state are due primarily to transmission constraints that limit the ability to export cheaper electricity located in upstate New York to the load center in New York City.

The New York State power system is summer-peaking, but also has a substantial winter peak. New York City accounts for approximately 30% of the energy usage and 33% of the peak energy demand (NYISO, 1999). The CNYS zone consumes approximately 10% of the state's electrical demand; for example, on August 9, 2003 at 3 pm (according to publicly available data on the NYISO webpage), CNYS demand was 2073 megaWatt-hours and total demand in New York State was 31,036 megaWatt-hours. This power system has the expected hourly, daily, and seasonal trends with respect to demand and prices common to typical U.S. power systems in the Northeast and elsewhere.

2.2 Data

We constructed time-series of day-ahead, zonal, wholesale electricity prices in New York City and for Central New York State for 2 pm along with six input fuel prices (three oil and three natural gas price streams at different locations). Oil or natural gas is likely to be the marginal fuel at 2 pm. Fuel costs are the major component of a fossil fuel unit's variable costs. Given that electricity prices in New York State are the marginal cost of providing one additional megaWatt-hour of electricity and the marginal megaWatt is typically a fossil fuel unit, we would expect *a priori* a positive relationship between the prices of fossil fuels and electricity prices. However, due to transmission constraints, which may require the backing down of inexpensive units and the ramping up of expensive units, locational electricity prices are also affected by congestion. None of the oil price streams, and only one natural gas price stream, are statistically significant. The natural gas price stream that is statistically significant is denoted TRNY to indicate the Transcontinental Gas Pipe Line Corporation daily prices reported by DRI/McGraw Hill.

The electricity price time-series runs from December 28, 2000 through March 31, 2003. Due to data limitations, the input fuel price data series began on December 28, 2000 but ended at the start of December 2002 and were available only for "work days" (i.e., weekdays that are not holidays).

2.3 Forecasting Techniques

We begin our empirical study by testing whether our data sets are nonstationary with a stochastic trend. If our data are difference stationary, we transform them into stationary sets by differencing. To test for stationarity, we conduct the Dickey-Fuller and the Phillips-Perron unit root tests (with and without time trend) on the level of NYC, CNYS, and TRNY natural gas prices for the entire sample period. Table 1 displays the results of unit root tests for each variable where the appropriate number of lags in our tests is determined by Akaike information criterion (AIC) specified as follows:

$$AIC = T\ ln(residual\ sum\ of\ squares) + 2n \quad (1)$$

where n = number of parameters estimated (p + q + possible constant term); and T = number of useable observations.

Table 1 shows that our data set is stationary as indicated by Dickey-Fuller and Phillips-Perron tests at the 5% significance level.

Table 1 Dickey-Fuller and Phillips-Perron for Unit Root Test (December-28-00 to November-25-02)

	NYC	CNYS	TRNY
Dickey-Fuller			
Trend	-7.22503	-7.73473	-3.37069
No-Trend	-7.22995	-7.73372	-3.91138
Phillips-Perron			
Trend	-12.92253	-14.31921	-10.09450
No-Trend	-12.93464	-14.32340	-10.51210

Next, we adopt four different estimation techniques to forecast the electricity prices in NYC and CNYS where the estimated parameters are allowed to vary over time:

2.2.1 Dynamic Regression Model (DRM)

The most parsimonious specification (i.e., the model with the least number of estimated coefficients) with significant coefficients can be stated as follows:

$$P_t = \alpha_1 + \alpha_2 P_{t-1} + \alpha_3 TRNY_{t-1} + \varepsilon t \quad (2)$$

where the electricity price (P_t) is related to the values of its first lag and to the first lag of the natural gas prices ($TRNY_{t-1}$).

We find other fuel prices and lagged electricity prices to be statistically insignificant at the 5% significance level.

2.2.2 Transfer Function Approach (TFA)

The transfer function is an extension of the ARMA model, where the process of the dependent variable (P_t) is allowed to depend on other independent variables such as the energy prices. The most parsimonious specification with significant coefficients can be stated as follows:

$$P_t = \alpha_1 + \alpha_2 P_{t-1} + [(w_0 + w_1 L + \ldots\ldots + w_n L^n)/(1-\delta_1 L - \ldots\ldots - \delta_m L^m)]TRNY_{t-d} + \varepsilon_t \quad (3)$$

where (L) is the lag operator indicating the number of lags for each variable, the number of numerator lags (n) is zero, the number of denominator lags (m) is zero, and the delay period for the series (d) is one.

2.2.3 Exponential Smoothing (ES) with Trend and Seasonality Presentation

We employ the exponential smoothing method. For each period, we perform nine exponential smoothing techniques that exploit all the available combinations from trends (no, linear, exponential) and seasonal trends (none, addictive, multiplicative). For example, one of the combinations is a linear trend and a seasonal additive trend. We then choose the best-fitting model that minimizes the in-sample squared one-step forecast errors based on Schwarz criterion [5].

2.2.4 The Generalized Autoregressive Conditional Heteroskedastic (GARCH) Methd

Here we investigate whether electricity prices can be modeled to capture the volatility variations in the electricity market. We run Lagrange multiplier test for ARCH and GARCH disturbances [6]. The purpose of this test is to determine whether ARCH or GARCH are appropriate by evaluating the correlation of the square of the residuals (variance) by regressing the square of the residuals on a constant and on one lag value. First, we estimate the residuals from the DRM for the whole time series. Second, we regress the squared residuals on a constant and on their first lagged value. With a sample of T residuals, under the null hypothesis of no ARCH errors, the test TR^2 converges to a χ^2 distribution with one degree of freedom.

Table 2 Lagrange Multiplier Test for ARCH of GARCH Errors (December-28-00 to November-25-02)

	NYC	CNYS
Lagrange Multiplier	3.86	2.13
Significance Level	.049	.144

As indicated by Table 2, the null hypothesis that the squared disturbances are uncorrelated is rejected in favor of the alternative hypothesis of ARCH or GARCH errors for the electricity prices in NYC at the 5% significance level. In contrast, the Lagrange multiplier test does not indicate the presence of the ARCH or GARCH errors in case of the electricity prices in CNYS. However, as shown later, modeling the conditional variance of the electricity prices in CNYS tends to improve significantly the forecasting performance of our models.

We begin our analysis by searching for the most parsimonious ARMA specification of the electricity price equation where the energy prices are included. Next, we explore modeling the volatility of the electricity prices as an autoregressive conditional heteroskedasticity (ARCH) process, a generalized autoregressive conditional heteroskedasticity (GARCH) process or GARCH in mean (GARCH-M). We try different specifications and reach the final preferred model specified by GARCH(1,1) and ARMA(1,0) with one lag of natural gas prices (TRNY) whose coefficient is significant for values between 0.1 and 0.35. Our jointly estimated specification can be stated as follows:

$$P_t = \alpha_1 + \alpha_2 P_{t-1} + \alpha_3 TRNY_{t-1} + \varepsilon_t \quad (4)$$

$$h_t = \beta_1 + \beta_2 h_{t-1} + \beta_3 \varepsilon_{t-1}^2 \quad (5)$$

where the appropriate log likelihood function of equations 4 and 5 can be defined as follows:

$$\log L = -\frac{T-1}{2}\ln(2\pi) - 0.5\sum_{t=2}^{T}\ln h_t - 0.5\sum_{t=2}^{T}\frac{(\varepsilon_t^2)}{h_t} \quad (6)$$

Then, we utilize the BFGS (Broyden, Fletcher, Goldfarb, and Shanno) algorithm to maximize the log likelihood function with respect to the α_i's and β_i's for i = 1 to 4.

We initially estimate the electricity prices over the first 100 days extending from December 28, 2000 to May 21, 2001 using the four estimation techniques. Then, the one-day out-of-sample forecasting performance of the four estimation techniques is evaluated. Next, we add one day at a time to the ending date, and repeat the process of estimating and forecasting the electricity prices over the next 377 weekdays extending from May 23, 2001 to November 22, 2002. The main advantage of such rolling window estimates is that our forecasts are more sensitive to including observations from the dataset, which helps in locating any extreme observations that might mask the causality between the electricity prices and their determinants.

3. Numerical Results

We use four measures of the performance of the four forecasting techniques: mean forecasting error (MFE), mean absolute forecasting error (MAFE), root mean squared forecasting error (RMSFE), and the Theil U statistics. These measures are expressed in the following equations:

$$MFE = \sum_{i=1}^{377}\tfrac{1}{377}(P_i^f - P_i) \quad (7)$$

$$MAFE = \sum_{i=1}^{377}\tfrac{1}{377}(\left|P_i^f - P_i\right|) \quad (8)$$

$$RMSFE = \sqrt{\sum_{i=1}^{377}\tfrac{1}{377}(P_i^f - P_i)^2} \quad (9)$$

$$U = \frac{RMSFE}{\sqrt{\sum_{i=1}^{377}\frac{1}{377}[P_i^f]^2} + \sqrt{\sum_{i=1}^{377}\frac{1}{377}[P_i]^2}} \quad (10)$$

where P^f is the one-day forecasting price.

Tables 3 and 4 report the MFE, MAFE, RMSFE, and the Theil U statistics at a one-step for New York City and Central New York, respectively. The results reveal that the forecasts of the rolling GARCH outperform the forecasts of the other estimation techniques for the 377 days. Additionally, our rolling forecasts locate four extreme observations on August 7-10, 2002. The extreme observations were identified by the significant deterioration in the forecasting ability of our techniques and the unprecedented increase in the electricity prices. (We are not correcting for all the extreme values, which would require adopting a formal technique that is beyond the scope of this paper.)

Extremely high prices in electricity markets can occur for a variety of reasons. During the August 7-10, 2002 period, the prices reached levels of $1024.91 and $907.77 for NYC and CNYS, respectively. The electricity markets in New York State have a bid cap of $1000/MWh. Generation units cannot submit bids for energy above this cap, but electricity prices may reach higher levels due to locational marginal pricing. During these periods, there was either insufficient energy and reserves to meet load or there was an unusually pronounced ability to exercise market power.

Although such observations are of important value, their nature and the probability of their occurrence seem to be unique and non-repetitive. Therefore, including these observations with such unusual high values may produce bias in parameter estimates and hence may deteriorate the efficiency of our forecasts. There is a growing body of evidence suggesting that the efficiency of both parameters estimates and out-of-sample forecasts can be improved if extreme values are accounted for in GARCH. For example, [7] attribute the excess kurtosis of the estimated residuals from GARCH models to the additive outliers in the stock market returns. When they account for such outliers, the adjusted data are normally distributed and the out-of-sample forecasts of the GARCH improve significantly.

Thus, we account for the four extreme observations by replacing them with the electricity price on August 6, 2001 (omitting the four observations presents similar forecasting results). The GARCH model (the one most preferred from the previous step) is then re-estimated after replacing these extreme values. As Tables 3 and 4 indicate, the four statistics reveal a significant improvement when we account

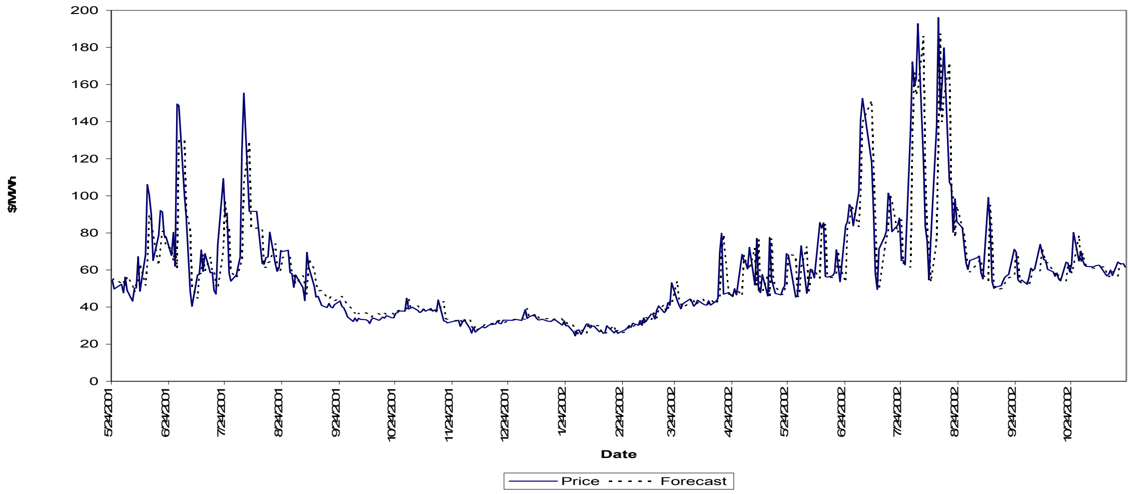

Fig. 1 Actual Versus Forecast Day-Ahead New York City Electricity Prices (May 24, 2001 Through November 22, 2002) Using GARCH, Omitting Extreme Values

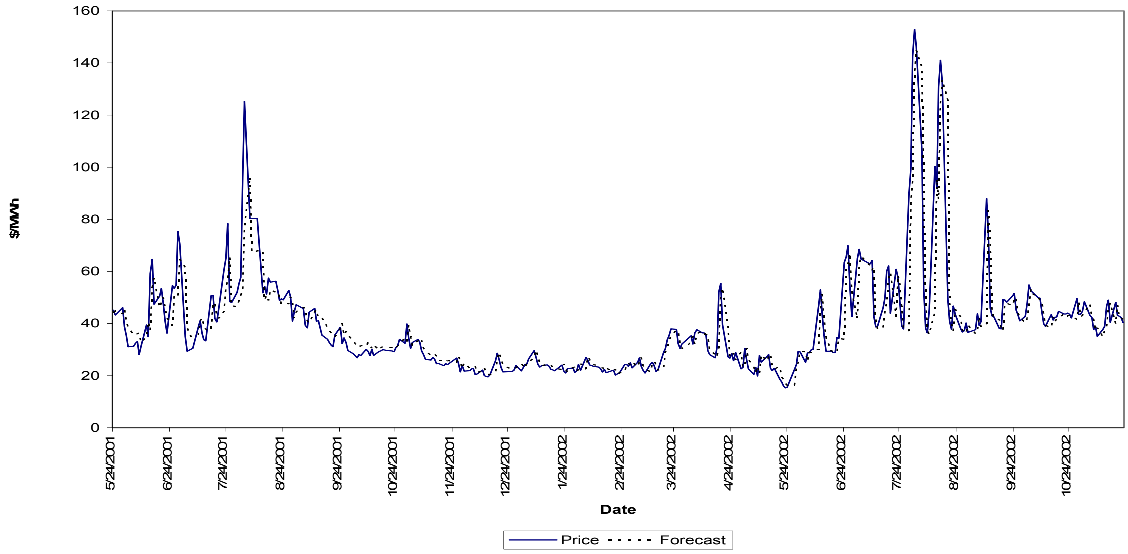

Fig. 2 Actual Versus Forecast Day-Ahead Central New York State Electricity Prices (May 24, 2001 Through November 22, 2002) Using GARCH, Omitting Extreme Values

Table 3 Comparison of Rolling One-Step Out-of-Sample Forecasts for New York City (May 23, 2001 to November 22, 2002)

	MFE	MAFE	RMSFE	U Theil
Dynamic Regression	-5.55965	22.35189	102.71036	0.49625
Transfer Function	-5.62115	22.33291	102.67120	0.49608
Exponential Smoothing	-14.7251	38.31004	230.66750	0.69260
GARCH	2.24775	12.68334	53.39037	0.33554
GARCH, Omitting Extreme Values	0.84005	8.44462	14.93854	0.11706

for the extreme values in the electricity prices between August 7, 2001 and August 10, 2001.

Table 4 Comparison of Rolling One-Step Out-of-Sample Forecasts for Central New York State (May 23, 2001 to November 22, 2002)

	MFE	MAFE	RMSFE	U Theil
Dynamic Regression	-8.21430	19.78906	133.60543	0.63770
Transfer Function	-8.16056	19.71594	133.51422	0.63749
Exponential Smoothing	-30.9230	55.52721	586.39677	0.88618
GARCH	1.43390	9.45427	43.20229	0.38025
GARCH, Omitting Extreme Values	0.57665	5.30377	10.34263	0.11747

Figs. 1 and 2 (located at the end of the paper) show the forecasts of rolling GARCH as compared to actual electricity prices at a horizon of one day for New York City and Central New York State, respectively. Electricity prices ($/MWh) are plotted on the y axis and the sample days on the x axis.

The content of the figures reveal the high accuracy of our forecasts, which capture the main movements in the electricity prices when accounting for the price volatility and the extreme values.

4. Comparison Of Price Volatility In New York City And Central New York State

Volatility is important in the electricity market for several reasons. First, it is critical in pricing electricity options, an important risk-management device commonly used by all types of market participants. Since the value and therefore the price of an option depend directly and significantly on volatility, accurate measures of volatility are critical.

Second, comparisons of volatility between subregions of a market are important not only for pricing electricity options with different delivery points but also for determining whether there are multiple markets within a region. Subregions with different levels of volatility indicate separate markets, which is critical information in a market power analysis.

We test whether the sum of the GARCH variance of the electricity prices in NY City (GV_{NY}) is statistically greater than that of Central New York State (GV_{CNYS}). In line with [8], [9], and [10] we use randomization tests to avoid making assumptions about the normality and statistical properties of the variance series. Our randomization test can be described as follows. First, we calculate the GARCH variance from equation (4) for New York City and Central New York State for the entire sample period extending from December 28, 2000 to November 25, 2002. We then calculate the historical ratio between the sum of GV_{NY} and GV_{CNYS}. Second, we perform a complete shuffle of all of the elements of a vector combining GV_{NY} and GV_{CNYS}. We then calculate the randomized ratio. Third, we repeat second and third steps 999 times. Finally, we calculate the p value of the historical ratio as the fraction of the randomized ratios greater than the original ratio. To account for the difference in the magnitude of electricity prices between the two regions, we define a normalized measure of the GARCH variance as follows:

$$(NGV_{NY})_t = \frac{(GV_{NY})_t}{(P_{NY})_t} \tag{11}$$

and

$$(NGV_{CNYS})_t = \frac{(GV_{CNYS})_t}{(P_{CNYS})_t} \tag{12}$$

where P_{NY} and P_{CNYS} are the electricity prices in New York City and Central New York State, respectively.

Figs. 3 and 4 (also located at the end of the paper) depict the GARCH variance and the normalized GARCH variance for the New York City and Central New York State for the entire period. The figures reveal the higher level of volatility in New York City. Additionally, Table 5

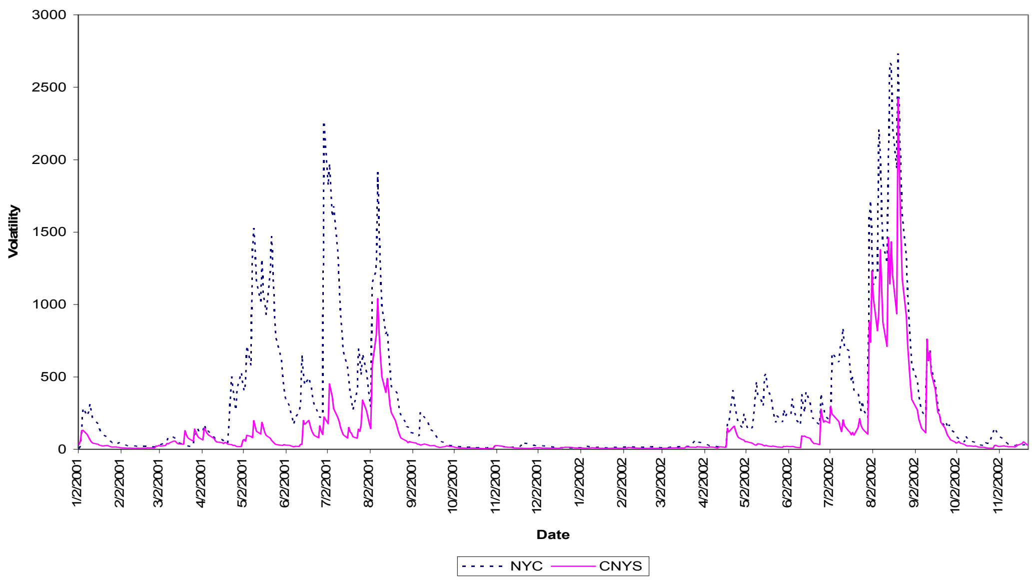

Fig. 3 Comparison of the Volatility of the Electricity Prices in New York City and Central New York State for the Period of January 1, 2001 to November 22, 2002)

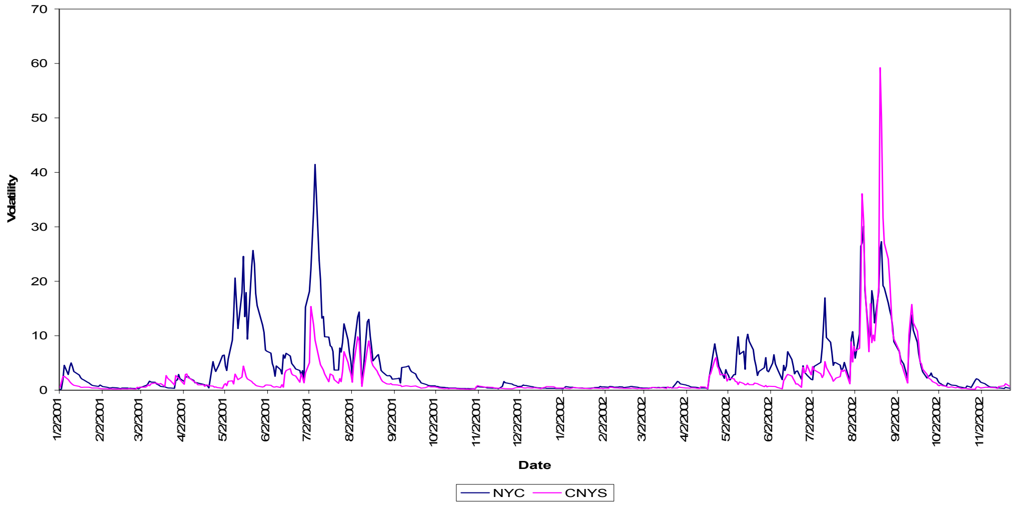

Fig. 4 Historical Ratios of the GARCH Variance of New York City to Central New York State and their Significance Level

provides a formal test for the variance ratio of the two regions. Table 5 shows that the probability of obtaining a ratio test greater than the historical, or unshuffled, ratio is .001. Thus, the results confirm that the electricity prices in New York City have a higher level of volatility.

Table 5 Historical Ratios of the GARCH Variance of New York City to Central New York State and their Significance Level

	$\sum \frac{GV_{NC}}{GV_{CNCS}}$	$\sum \frac{NGV_{NC}}{NGV_{CNCS}}$
Ratio Statistics	2.46	1.7
Marginal Sign Level	.001	.001

It takes slightly less than 60 minutes to run the entire analysis just described on a PC Pentium II with 128 Mb of RAM at 1 GHz. This is a short enough time to allow these techniques to be used in practice. We used Regression Analysis of Time Series (RATS) software to conduct most of the computations.

5. Conclusions

In line with other studies such as [11], we conclude that incorporating volatility into price forecasting via the GARCH process significantly improves the forecasting performance over the other techniques evaluated. In particular, the GARCH process performance is a substantial improvement over DRM, TFA, and ES. Elimination of the few extreme values further improves the performance of the GARCH process. Extreme values, however, are important to estimate in many applications. More research is needed that combines GARCH with techniques that forecast extreme values.

We also find that New York City has a larger daily volatility than does Central New York State. One possible explanation is that the market for electricity in New York City has more market power than does the upstate region. Higher volatility can be an indicator of collusive behavior or higher transmission constraints. More detailed analysis is required to distinguish between these two possible causes. We also notice that there are two peaks and pretty stable variance in between: additional investigation is needed to determine the cause.

Our work can be extended in several directions. First, a formal technique to identify and correct for outliers instead of the use of visual inspection can be incorporated into the analysis. Second, the electricity price forecasting and volatility analysis can be applied to other hours in the day-ahead market besides 2 pm, to the real-time market, and to other regions within New York State and elsewhere.

References

[1] F. J. Nogales, J. Contreras, A. J. Conejo, R. Espínola, “Forecasting Next-Day Electricity Prices by Time Series Models,” *IEEE Transactions on Power Systems*, vol. 17, no. 2, 2002, pp.342-348.

[2] R. Engle, “Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation,” *Econometrica,* 50, 1982, pp.987-100.

[3] T. Bollerslev, “Generalized autoregressive conditional heteroscedasticity,” *Journal of Economics,* 31, 1986, pp.302-327.

[4] R. Engle, D. Lilien, and R. Robins, “Estimating time varying risk premia in the term structure: The ARCH-M model,” *Econometrica*, 55, pp.391-407, 1986.

[5] NYISO Report, *New York Power Pool Zone*

Forecasting Models, February 10, 1999, p.1, available at nyiso.com.

[6] G. Schwarz, "Estimating the Dimension of a Model," *Annals of Statistics* 6, pp. 461-464, 1978.

[7] E. Walter, "Applied Econometric Time Series," New York: John Wiley & Sons, 1994, pp.146-148 & pp.233-243.

[8] P. Franses and H. Ghijsels, 1999, "Additive outliers, GARCH and forecasting volatility," *International Journal of Forecasting* 15, pp.1-9.

[9] E. W. Noreen, *Computer-Intensive Methods for Testing Hypotheses: An Introduction,* New York, NY: John Wiley & Sons, 1989.

[10] M. Kim, C. Nelson, and R. Startz. "Mean Reversion in Stock Prices? A reappraisal of the Empirical Evidence," *Review of Economic Studies* 58, 1999, pp.515-528.

[11] G. McQueen, "Long-Horizon Mean-Reverting Stock Prices Revisited," *Journal of Financial and Quantitative Analysis* 27, 1992, pp.1-18.

[12] H. Guirguis, C. Giannikos and R. Anderson, "The US Housing Market: Asset Pricing Forecasts Using Time Varying Coefficients," *Journal of Real Estate Finance and Economics*, forthcoming.

Hany S. Guirguis
He earned his Ph.D. in Economics from the University of Oregon in 1995. He is Assistant Professor of Economics and Finance at Manhattan College in Riverdale, New York. His research interests are in applied econometrics.

Frank A. Felder
He earned his Ph.D. in Technology, Management and Policy from the Massachusetts Institute of Technology in 2001. He is Assistant Research Professor with the Center for Energy, Economic and Environmental Policy at the Edward J. Bloustein School of Planning and Public Policy, Rutgers, The State University of New Jersey.