MODELING AND FORECASTING SHORT-TERM ELECTRICITY LOAD USING REGRESSION ANALYSIS

Jennifer Hinman and Emily Hickey

Illinois State University

Fall 2009

Abstract

The objective of this study is to describe a parsimonious forecasting model for the hourly electricity load in the area covered by an electric utility located in the Midwest of the United States that performs well in out-of-sample forecast evaluation. This study proposes using an autoregressive moving average model with exogenous weather variables (ARMAX) to forecast short-term electricity load using hourly load data from Commonwealth Edison Company (ComEd). The proposed model treats each hour's load separately as an individual daily time series. This approach avoids modeling the complicated intraday pattern (load profile) displayed by the load, which varies through the week as well as through the seasons. To date, no published study to our knowledge has taken an ARMAX modeling approach to forecast short-term electricity load in ComEd's territory. The importance of accurate short-term forecasting is greatest for utilities operating in a restructured environment, such as ComEd in Illinois.

TABLE OF CONTENTS

I. INTRODUCTION

Electricity load forecasting has been an important risk management and planning tool for electric utilities ever since the conception of forecasting. Load forecasting is necessary for economic generation of power. Load serving entities use load forecasts for system security, to schedule generator maintenance, to make long-term investments in generation, and to plan the most cost-effective merit order dispatch. Over the last decade, as electricity markets have deregulated, the importance of load forecast accuracy has become even more evident. Without an optimal load forecast, utilities are subject to the risk of over- or under- purchasing in the day-ahead market. While an entity can buy or sell power in the real time market to correct for forecast inaccuracy, it comes at the expense of higher real time prices. A one percent reduction in the average load forecast error has the possibility of saving hundreds of millions of dollars for a utility. Thus, the financial costs of forecast errors are so high that much research is focused on reducing the error even by a fraction of a percentage point (Weron and Misiorek 2004). Today, load forecasting has become an integral part of planning for more than just utilities; regional transmission organizations (RTO), energy suppliers, financial institutions, and participants in the generation, transmission, and distribution of electricity have a vested interest in load forecast accuracy.

As the value of accurate load forecasting has grown throughout the years, the literature has expanded to consider a wide array of empirical and analytical tools including, adaptive forecasting techniques (Gupta 1985), neural network models (Hippert et al. 2001), various types of Box-Jenkins time-series approaches (Box and Jenkins 1970), Seasonal Integrated Autoregressive Moving Average models (SARIMA) (Soares and Souza 2006), Two-Level Seasonal Autoregressive models (TLSAR) (Soares and Medeiros 2005, 2008), Autoregressive Fractional Integration Moving Average models (ARFIMA) (Soares and Souza 2006), Dummy-Adjusted Seasonal Integrated Autoregressive Moving Average models (DASARIMA) (Soares and Medeiros 2005, 2008), Smooth Transition Periodic Autoregressive models (STPAR) (Amaral et al. 2008), cointegration analysis (Chen 1997), and many other techniques[1]. Generally speaking, electricity load forecasts can be classified in terms of the duration of the planning horizon: short-term load forecasting (STLF) (one hour to one week), medium-term load forecasting (MTLF) (one week to one year) and long-term load forecasting (LTLF) (longer than a year) (Feinberg and Genethliou 2005).

The objective of this study is to describe a parsimonious forecasting model for the hourly electricity load in the area covered by an electric utility located in the Midwest of the United States that performs well in out-of-sample forecast evaluation. This study proposes using an autoregressive moving average model with exogenous weather variables (ARMAX) to forecast short-term electricity load (24-hours ahead) using hourly load data from Commonwealth Edison Company (ComEd)[2]. The proposed model treats each hour's load separately as an individual daily time series. This approach avoids modeling the complicated intraday pattern (load profile) displayed by the load, which varies through the week as well as through the seasons. To date, no published study to our knowledge has taken an ARMAX modeling approach to forecast short-term electricity load in ComEd's territory. Our findings indicate that a multi-equation regression approach to forecasting short-term electricity load performs best when weekdays are modeled separately and when a separate model is fitted for the night- versus the day-time hours[3].

The importance of accurate short-term forecasting is greatest for utilities operating in a restructured environment, such as ComEd in Illinois (Wang 2004). Illinois' own restructuring experience has made

---

[1] See Bunn and Farmer (1985b), Alfares and Nazeeruddin (2002), and Weron (2006) for a good review of load-forecasting methods.

[2] ComEd is a utility that provides service to approximately 3.8 million customers across Northern Illinois, or 70 percent of the state's population. ComEd's service territory borders Iroquois County, Illinois to the south (roughly Interstate 80), the Wisconsin border to the north, the Iowa border to the west and the Indiana border to the east.

[3] Although we do not investigate the following in this study, we believe that performance is likely to be enhanced if a model is fitted for each of the twenty-four hours separately, rather than only two hours (one night and one day) and applied to the rest of the hours as is done in this study.

load forecasting a critical component of power procurement[4]. In 2007, the state of Illinois passed the Illinois Power Agency Act, which created the Illinois Power Agency (IPA), among other things. The IPA procures electricity on the behalf of ComEd and relies on load forecasts provided by the utility to devise its procurement plan.

The organization of the paper is as follows. The next section reviews the electricity load forecasting literature. Section III explains the theoretical model. Section IV describes the data used in this study. Section V discusses the econometric method utilized to fit the load demand as well as the forecasting results. Section VI offers some concluding remarks.

II. LITERATURE REVIEW

The literature on load forecasting extends as far back as the mid-1960s (Heinemann et al. 1966; Hahn et al. 2009). While Kalman filter and state space methods dominated the literature early on, artificial and computational intelligence methods and econometric techniques have largely dominated literature that is more recent. The choice of the appropriate technique for load forecasting depends largely upon the forecast horizon. Short-term load forecasts (STLF), which forecast one hour to one week ahead, have become one of the most important load forecasts performed by utilities (Hahn et al. 2009; Feinberg and Genethliou 2005). Medium-term load forecasts (MTLF) typically range from one week to one year ahead; whereas, forecasts that aim to predict load beyond one year are considered long-term load forecasts (LTLF). While long- and medium-term forecasts are useful for guiding planning and operational decisions, they have little practical use estimating day-to-day load flows, which are vital to utility short-term planning. Forecasting peak electricity loads has also been a topic among much of the load forecasting literature (e.g., Engle et al. 1992). Interestingly, cointegration analysis has been used to

[4] With respect to procurement, long-term load forecasting (LTLF) is of interest.

forecast long-term peak electricity load (Chen 1997). Recall the stated objective of this paper is to forecast short-term electricity load; accordingly, only those models and studies that perform short-term forecasts are considered. Within the short-term load forecasting literature, a number of models varying in the complexity of functional form and estimation techniques have been developed. Some of the most successful modeling techniques have utilized either classical time-series regression analysis (i.e., econometric techniques) or artificial and computational intelligence methods.

A. Classical Approach (Time Series and Regression)

Time-series techniques have been extensively used in load forecasting for decades and are among the oldest methods applied in forecasting (Hahn et al. 2009; Bunn and Farmer 1985a, 1985b; Weron 2006; Kyriakides and Polycarpou 2007). Two overarching classes of time-series regression models have emerged to address the time-scale issues in different ways. Amaral et al. (2008) contend the two broad classes of conceptual models include: (1) single-equation models and (2) multi-equation (vector) models. This distinction between single-equation models and multi-equation (vector) models is important because we utilize the multi-equation approach in our estimation. The articles that utilize the multi-equation conceptual approach include Fiebig, Bartels, and Aigner (1991); Peirson and Henley (1994); Ramanathan et al. (1997); Cottet and Smith (2003); Soares and Medeiros (2005, 2008); and Soares and Souza (2006). Comparatively speaking, the single-equation approach has been relatively dominant in the literature; however, recent load forecasting efforts have largely gravitated toward the multi-equation approach. The multi-equation model described by Ramanathan et al. (1997) was actually the winner of an electricity load-forecasting competition, so we find the lack of multi-equation approach load-forecasting literature quite surprising[5].

[5] To the best of our knowledge, only three multi-equation load-forecasting articles have been published since the Ramanathan

Irrespective of which approach is adopted, single equation or multi-equation, when constructing a time-series model there are four components that must be taken into consideration: trend, cyclicality, seasonality, and a random white noise error. Consequently, the time-series literature can be envisioned in terms of how it has addressed each time-series component. In general, accounting for both cyclicality and seasonality has been extensively covered in the literature whereas trend, while addressed, is typically not the focus of the analysis.

1. Modeling Trend

Within the literature, load data has occasionally been found nonstationary. Some, Darbellay and Slama (2000) for example, first difference the data to account for nonstationarity. Other studies, however, find that fitting a deterministic trend is more appropriate. Soares and Medeiros are highly critical of authors' tendency to first difference without first testing for a unit root or even considering a linear trend (2008). Soares and Medeiros (2008) point out that when the trend is in fact deterministic, taking the first difference will introduce a non-invertible moving average component, which in turn, will cause serious estimation problems. Upon examining hourly load data for Rio de Janeiro, Soares and Medeiros (2008) find that the data display a positive linear trend. Using the Phillips-Perron unit root test, the authors find that inclusion of a linear trend is necessary to make the data stationary for all hourly series considered (Soares and Medeiros 2008). Similarly, Ramanathan et al. (1997) fit both a linear trend and the reciprocal of that for their multi-equation estimation of hourly load for the Puget Sound Power and Light Company for the period 1989 to 1990[6].

Alternatively, there are studies that find the load series to be stationary. Taylor et al. (2006) look at twenty weeks of hourly data and they find no evidence of a unit root, and thus no justification for first

---

et al. (1997) article was issued. However, we do not know whether utilities have adopted the winning multi-equation forecasting approach.

[6] The model described by Ramanathan, Engle, Granger, Vahid-Arahi, and Brace (1997) was the winner of a load forecast competition.

differencing. Accordingly, it would appear that the existence of a trend may largely depend on both the length of the time period examined as well as location-specific factors[7].

2. Modeling Cyclicality

Autoregressive moving average models have been extensively applied in the load forecasting literature (Weron 2006; Taylor et al. 2006; Kyriakides and Polycarpou 2007; Feinberg and Genethliou 2005; Pappas et al. 2008). Of those studies that perform short-term load forecasting, the most popular time-series techniques that have been adopted are some formulation of Autoregressive Moving Average (ARMA) or Autoregressive Moving Average with exogenous variables (ARMAX) models (Feinberg and Genethliou 2005). When choosing between applying a univariate (ARMA) or multivariate (ARMAX) time-series model, the time horizon and data availability give some indication as to which technique is feasible. Weron (2006) notes that while moving average (MA) models are not particularly useful (besides their filtering properties) in forecasting electricity load, the combination of the moving average process with an autoregressive model (an ARMA process) provides a very powerful load-forecasting tool. Taylor et al. (2006) and Hahn et al. (2009) assert that univariate models are typically used for very short-term load forecasts, while Hahn et al. (2009) emphasize multivariate methods are typically applied to all time horizons.

a. ARMA Models

Amjady (2001) uses an ARIMA model to forecast load for four different types of days[8], which simultaneously accounts for intraday seasonality. In total Amjady (2001) estimates 16 ARIMA models, one for each type of day, and a hot- and cold-days model within each day-type model. Using data from the Iranian national grid and an in-sample forecast period from 1996 to 1997, Amjady (2001) finds the Mean Absolute Percentage Errors (MAPEs) to range from 1.48% (Sunday to Wednesday, hot) to 1.99%

---

[7] High industrial sector growth over a period in a specific location may impact the stationarity of a load series.
[8] Saturday, Sunday to Wednesday, Thursday and Friday, and holidays.

(public holidays, cold). Pappas et al. (2008) use daily load data as opposed to hourly loads from the Hellenic power market and after accounting for seasonality, they find that an ARMA(2,6) model successfully fits the data.

Soares and Medeiros (2008) and Soares and Souza (2006) utilize the multi-equation approach and apply univariate ARMA models to forecast electricity load in Rio de Janeiro. Soares and Souza (2006) propose a stochastic model that employs generalized long memory (by means of Gegenbauer processes) to model the seasonal behavior of load, while Soares and Medeiros (2008) propose a Two-Level Seasonal Autoregressive (TLSAR) model. As pointed out by Soares and Souza (2006), forecasting errors are generally quite high during the summer due to the influx of air conditioning; thus, including temperature or other exogenous variables would help resolve this issue. Poor data availability is the primary reason cited as to why temperature is excluded from the model; where temperature data are available, both studies advocate its inclusion (Soares and Medeiros 2008; Soares and Souza 2006).

b. ARMAX Models

It would appear that while univariate ARMA models are sufficient for short-term load forecasting, the literature agrees that including exogenous variables like temperature can potentially improve forecasting performance (Soares and Medeiros 2008; Soares and Souza 2006; Ramanathan et al. 1997; Taylor et al. 2006; Darbellay and Slama 2000; Carpinteiro et al. 2004). Darbellay and Slama (2000) do both univariate modeling using an ARIMA model and multivariate modeling using an ARMAX model that incorporates temperature data. Using hourly load data from the Czech Republic, Darbellay and Slama (2000) find the ARMAX model to be superior.

(i) Weather Variables

Variations in weather are largely regarded as important factors in modeling electricity demand (Feinberg and Genethliou 2005). Of the exogenous variables considered in ARMAX models, weather related variables (e.g., temperature) have by far been the most popular and at times, the most complex to account for (Ramanathan et al. 1997).

The nonlinear relationship between load and temperature is well documented in the literature and is accounted for in various ways (Ramanathan et al. 1997; Hor et al. 2005; Darbellay and Slama 2000; Cottet and Smith 2003; Kyriakides and Polycarpou 2007). Hor et al. (2005) propose including heating degree-days (HDD) and cooling degree-days (CDD) as a method to cope with the nonlinear relationship between load and temperature. Alternatively, Ramanathan et al. (1997) simply add the square of temperature to account for the nonlinear relationship between temperature and load. Bruhns et al. (2005) decompose the load model into a weather dependent and a weather independent part wherein the nonlinearity is addressed by differentiating between a heating part (temperature rises above a threshold) and cooling part (temperature falls below a threshold) of the weather sensitive load. Cottet and Smith (2003) note temperature seems to be the most important meteorological factor in most locations and that the relationship between load and temperature is approximately “V”-shaped and is known to vary depending on the time of day. In some locations, humidity is thought to have a relationship with electricity load. Cottet and Smith (2003) resolve the nonlinearity issue by interacting temperature and humidity. Peirson and Henley (1994) consider the dynamic specification of the relationship between load and temperature and find it to be quite important; in fact, they report that a static specification can suffer badly from serial correlation beyond the first order, which will yield biased and inefficient estimates of regression coefficients.

What can be taken away from the literature that considers the effect of temperature on load is that there tends to be agreement that load and temperature have a nonlinear relationship and that this nonlinearity must, in some way, be addressed in the model. While there is generally a consensus that accounting for temperature and other weather components will improve forecasting performance, data availability may preclude the inclusion of such variables. Moreover, Fidalgo and Matos (2007) declare that including weather variables in the model for load really depends on the region being studied and its climatic conditions.

3. Modeling Seasonality

Among those studies that have modeled electricity load using ARMA and ARMAX models, there has been a consensus that load data suffers from multiple seasonality. According to Hahn et al. (2009), time-series load data contain three seasonal patterns: intraday (daily), weekly, and annual. The intraday seasonal pattern reflects a peak (hours of high demand) and off-peak (hours of low demand) load pattern. The weekly pattern reflects the variation in load on weekdays versus weekends where load from the industrial sector is dramatically reduced. The specific weekday pattern can vary among regions and seasons (Hippert et al. 2001). Moreover, load can vary depending on the presence of a holiday or other exceptional event; several approaches exclude such exceptional cases by replacing load values on those days (Hippert et al. 2005; Taylor and McSharry, forthcoming).

Many techniques have been adopted to address the complexity of seasonality in the load forecasting literature. Deterministic weekly seasonality can be accounted for through the inclusion of day-of-the-week dummy variables (Gupta 1985). Soares and Medeiros (2008) include a dummy variable for each day of the week in addition to holiday dummies; in total 15 different binary variables are included to account for weekly and holiday seasonality. Similarly Cottet and Smith (2003) include 13 dummy variables; one for each day of the week and six to account for public holidays. Rather than include multiple dummy variables to account for holidays, Hippert et al. (2005) replace the atypical loads experienced on any holiday, by the load observed on that specific day of the week from the previous week. Cottet and Smith (2003) find that the estimates of the dummy variable coefficients were similar for workdays Monday through Friday with a slightly lower load on Friday afternoon and Monday morning. Moreover, Cottet and Smith find that holiday loads can vary, depending on the holiday, but generally resemble the load profile seen on Sundays.

Alternatively, other studies use seasonal differencing to account for stochastic seasonality (Taylor et al. 2006; Darbellay and Slama 2003). Taylor et al. (2006), for example, include both a 24th difference (to account for intraday seasonality) and a 168th difference (to model weekly seasonality). An alternative to 24th differencing, which accounts for intraday seasonality, is to model each individual hour separately thus allowing the parameters to vary based on intraday effects.[9] It has been pointed out that this approach may not be appropriate when the dataset being used is not sufficiently large (Hippert et al. 2001).

Annual and other types of seasonality have been addressed in the literature through the application of a Fourier decomposition[10] (Soares and Medeiros 2008; Cottet and Smith 2003; Schneider et al. 1985; Weron 2006). Soares and Medeiros (2008), for example, model the annual cycle as a sum of sines and cosines, like a Fourier decomposition technique where the number of trigonometric functions is determined by the Schwarz Information Criterion (SIC) (Schwarz 1978). Similarly, Cottet and Smith (2003) use the Fourier decomposition to account for annual seasonality. Besides Fourier decomposition, annual seasonality can be addressed through inclusion of monthly dummies (Ramanathan et al. 1997). It is worth noting that when a temperature variable is added to a model, the annual seasonality may go away if temperature is largely responsible for the annual load fluctuations (Weron 2006).

B. ARTIFICIAL INTELLIGENCE-BASED METHODS

Artificial intelligence-based methods or AI-based (non-parametric) techniques as they are commonly

[9] See Fiebig et al. 1991; Peirson and Henley 1994; Ramanathan et al. 1997; Cottet and Smith 2003; Soares and Medeiros 2005, 2008; Soares and Souza 2006.

[10] In mathematics, a Fourier series decomposes a periodic function into a sum of simple oscillating functions, such as sines and cosines.

referred to, have been a widely studied and applied load forecasting technique. While they are beyond the scope of this study, any literature review of load forecasting techniques would be incomplete without an overview of AI-based techniques.

AI-based techniques are very flexible nonlinear methods and no prior modeling experience is required for a decent load forecast. The algorithms utilized automatically classify the input data and associate it with the respective output values (Weron 2006). AI-based methods are often coined as "black-box"-type tools, though the empirical evidence from practical everyday use suggests they perform reasonably well (Weron 2006). Some examples of AI-based techniques include artificial neural networks (ANN), fuzzy logic, expert systems, and support vector machines (Weron 2006). The use of artificial neural network models has been a "widely studied electric load forecasting technique since 1990" (Feinberg and Genethliou 2005, 278). ANN models have received the most attention and seem to be the most popular of the AI-based techniques. Accordingly, we will limit our discussion of AI-based methods to those of artificial neural networks.

Artificial neural network (ANN) models are flexible nonlinear models that can be used in the electric utility industry for short-term forecasting. In practically applying a neural network technique to load forecasting, one must decide upon the number of architectures (e.g., Boltzmann machine, Hopfield, backpropagation), connectivity of layers, and bi-directional or uni-directional links among inputs and outputs (Kyriakides and Polycarpou 2007). After designing the neural network model to be used for time-series prediction, the next steps involve training the neural network, and then testing the trained network, as a form of neural network validation (Kyriakides and Polycarpou 2007).

Neural network models are "well suited to the forecasting task when there are several explanatory variables and when there are important nonlinearities and variable interactions" (McMenamin 1997, 22). Neural network models are analogous to econometric techniques in many ways. A general neural network forecasting model can be written as follows: $Y = F[H_1(X), H_2(X), \dots, H_N(X)] + u$ where the dependent variable (output) is Y, X is a set of explanatory variables (inputs), the F (output layer activation function) and H's (hidden layer activation functions) are neural network functions, and the u is the error term for the model (McMenamin 1997), the terms in parentheses represent the neural network language. The parameters of the model are termed "connection strengths or network weights. Constant terms are called biases, and slopes are sometimes called tilts. The sample period data are called the training set, and the estimation process is called training" model (McMenamin 1997, 18). The training process is similar to that of typical regression estimation; the goal is to find the network weights that minimize the model errors. A general linear regression model can be written as follows (in econometric language, Y is the dependent variable, X is a vector of explanatory variables, β is a vector of regression coefficients, and u is the model error term): $Y = X\beta + u$. McMenamin (1997) describes this regression model in neural network terms, "this is a single output feed forward system with no hidden layer and with a linear activation function at the output layer" (17). McMenamin reasons that "in this sense, the linear regression model is a severely limited special case of the neural network framework" (17).

While ANN proponents may put down regression analysis, econometricians and forecasters often view ANNs as "black boxes" (Weron 2006; Hippert et al. 2005). Wang (2004) states it best,

> Too much reliance on the automatic feature of the modeling software may cover up problems when they arise. The more sophisticated the automatic modeling software is, the more the model will work like a black box, and the more it will be out of the forecasters' control. Forecasters should always be in charge of their models, not the other way around. (15)

Hippert et al. (2005) investigate to what extent artificial neural network (ANN) models' over parameterization affects their performance in a practical forecasting sense. They examine this issue by comparing the performance of conventional regression forecasting methods to large neural networks to see which performs better in forecasting load profiles. They find that the ANN models perform at least as well as the conventional regression forecasting methods.

C. Conclusions from the Literature

As the load forecasting literature has evolved over the years, some approaches have proven to be superior. More recently, the literature has gravitated toward time-series forecasting and neural network modeling. The two main differences between a neural network model and a linear regression model are that the regression model is linear in parameters and there are no hidden layer functions as there are in neural network models (McMenamin 1997). While AI-based techniques have proven to be a useful tool, it has been cited that training the model can take a great deal of time (Alfares and Nazeeruddin 2002). Moreover, some remain skeptic as to the performance of ANN and whether they truly outperform standard forecasting methods (Weron and Misiorek 2004). Two shortcomings of ANN that bring into question the credibility of studies that utilize ANN include complaints that ANN techniques overfit the data and that the models were not systematically tested. Alternatively, time-series techniques have been widely recognized as a useful load-forecasting tool and have been met with relative success. Regression methods are advantageous in that they are relatively easy to implement, easy to interpret, and allow for relatively easy performance assessments (Hahn et al. 2009).

Darbellay and Slama (2000) directly compare the forecast accuracy of an ARMAX model and a neural network model and find that the ARMAX model is superior. Due to the relative ease and superior performance of regression analysis in the load forecasting literature, this study employs time-series multi-equation regression techniques to forecast short-term electricity load.

III. Theoretical Analysis

The literature comes to a consensus that temperature, past load, calendar events, and seasonal factors (intraday, weekly, and annually) have a significant impact on load in any given region. The driving factors behind hourly, weekly, and annual variations in load within ComEd's service territory resemble those experienced elsewhere and largely reflect consumer lifestyle choices as well as industrial and commercial business activity. According to Ramanathan et al. (1985), electricity usage in any given hour is determined by variables related to both the day of the week and the time of the day. These variables, in turn, reflect the household's respective lifestyle choices and responses to the environment. Lifestyle choices include the household's work and leisure patterns while potential environmental determinants include the temperature, humidity, cloud cover, and wind chill. While the aforementioned variables reflect household decisions and responses to the environment that change quickly from hour-to-hour and from day-to-day, another group of variables that can affect hourly load are those that change slowly over time (e.g., family size, income, appliances). In the context of a short-run load forecast; however, only short-run factors that can affect hourly load should be considered. According to Ramanathan et al. (1997), "Slowly changing variables such as increases in population, industrial growth, global warming, and so on are of little or no relevance in a model that is attempting to forecast from one day to the next" (164).

In the STLF literature discussed in the preceding section, various functional forms have been applied to address the seasonal, cyclical, and exogenous factors that can affect load. One increasingly popular form that has been adopted by Fiebig et al. (1991), Peirson and Henley (1994), Ramanathan et al. (1997), Cottet and Smith (2003), Soares and Medeiros (2005, 2008), and Soares and Souza (2006) involves modeling each hour as a separate time-series equation. Such an approach can result in a better fit of the data in addition to removing the complexity of modeling intraday seasonality. Some studies that have adopted a multi-equation approach have applied the same model to all twenty-four-hours of load data (Soares and Souza 2006). While this approach allows the parameters to vary based on intraday effects, it may not result in the most parsimonious model. This is due in part to the fact that load patterns during certain hours can display both unique and complex dynamics. For the sake of this study, we investigate whether fitting a single model for all twenty-four hourly series is a viable forecasting approach for the dataset at hand; specifically, we consider if a single model approach results in white noise residuals for all hours whilst satisfying the parsimony principle.

Any model of electricity load, irrespective of whether the data series is for the entire series or for a smaller subset, will contain the following components:

$$\boldsymbol{Load_{h,d} = Deterministic_{h,d} + Temperature_{h,d} + ARMA_{h,d} + \varepsilon_{h,d}} \quad \textbf{(1)}$$

where *h* indicates the hour of the day and *d* indicates the daily observations and $\varepsilon_{h,d} \sim WN(0, \sigma^2)$.

The deterministic component contains variables that are perfectly predictable. Dummy variables for specific days of the week and federal holidays are all included in the model:

$$\boldsymbol{Deterministic_{h,d} = \alpha + \lambda Holiday + \sum_{i=1}^{6} \beta_i Day_{i,d}} \quad \textbf{(2)}$$

Where Holiday indicates observed Federal holidays and Day indicates day-of-the-week dummy variables.

For the purposes of modeling all twenty-four hours through a single model, one day of the week is excluded and accounted for through the inclusion of a constant. Generally, the coefficients on the weekday dummies are expected to be greater than a coefficient on a weekend dummy; however, the magnitude and potentially the direction of the effect may vary depending on which day is excluded. Finally, holidays must be taken into consideration; while these variables can be expected to have a pronounced effect on load during peak hours, their effect on load during evening hours may be negligible. Regardless of the statistical significance, the coefficient on the holiday term is expected to be negative, which would reflect the significant decrease in industrial activities on those days as compared to a non-holiday.

The temperature component, which is sometimes neglected due in part to the elusiveness of data, is one of the most significant drivers of load. The nonlinearity of temperature has been documented extensively in the literature[11]. In addition to the square of temperature, lags of temperature and interactions with summer months may also be relevant; a lagged temperature variable would reflect the fact that consumption decisions today are influenced by temperature experiences yesterday. If the day prior was particularly hot, for example, a consumer may increase their air conditioning usage in the present ceteris paribus. Alternatively, interactions with summer months also help explain the complex relationship between load and temperature. In particular, they account for the non-constant effect of temperature across some months. Theory would suggest that all else constant, temperature has a greater effect on load in summer months such that the sign on the interaction terms should be positive. Temperature is expected to significantly influence load regardless of whether that load occurs during an off-peak or peak hour.

The ARMA component represents the appropriate model of cyclicality for off-peak and peak hours. The autoregressive (AR) component captures the fact that high load in hour *i* for any given day is a good indication that load will be higher in hour *i* on the following day(s). In other words, the load is assumed a linear combination of load from previous periods. The autoregressive component is more persistent than the moving average component which captures whether a shock in hour *i* persists the following day(s) in that hour. In other words, the autoregressive component is much better at modeling the

[11] See Ramanathan et al. 1997; Hor et al. 2005; Darbellay and Slama 2000; Cottet and Smith 2003; Kyriakides and Polycarpou 2007.

dynamic behavior of load as compared to including only a moving average component in the model (Weron 2006). The ARMA (p,q) representation can be specified for any given hour of load as follows:

$$\boldsymbol{ARMA_{h,d} = \alpha + \sum_{i=1}^{p} \varphi_i\, Load_{h,d-i} + \sum_{i=1}^{q} \theta_i\, \varepsilon_{h,d-i} + \varepsilon_{h,d}} \qquad \textbf{(3)}$$

While theory suggests that an ARMA model is appropriate, empirical testing is necessary to determine the exact ARMA order. The final model, developed with a single model framework in mind[12], is as follows:

$$\boldsymbol{Load_{h,d} = \alpha + \lambda Holiday_d + \sum_{i=1}^{6} \beta_i Day_{i,d} + \delta_{tmp} TMP_{h,d} + \delta_{tmp2} TMP_{h,d}{}^{2} + \delta_{ltmp} TMP_{h,d-1} + \delta_{ltmp2} TMP_{h,d-1}{}^{2} + \delta_{m6tmp} JUN * TMP_{h,d} + \delta_{m7tmp} JUL * TMP_{h,d} + \delta_{m8tmp} AUG * TMPh,d+i=1p\varphi iLoadh,d-i+i=1q\theta i\varepsilon h,d-i+\varepsilon h,d} \qquad \textbf{(4)}$$

where all variables will be defined in the following data section.

IV. DATA

We consider a dataset containing hourly loads from May 1, 2004 through September 30, 2009. The data from the period from May 1, 2004 through April 30, 2008 are used for estimation purposes (in-sample), and the data from the period May 1, 2008 through September 30, 2009 are left for out-of-sample forecast evaluation. We believe an estimation window of four years to be adequate for estimation[13]. Load data is obtained from the "Historical Load Data" report from the PJM website[14] and it

[12] A single model in the sense of applying it to all twenty-four hours (equations).

[13] The four-year estimation window results in 1,461 observations for the full-week dataset, and 1,043 observations for the dataset that excludes weekends. Exelon: Commonwealth Edison Company joined the PJM regional transmission organization

summarizes the megawatt-hour net energy for load as consumed by the Commonwealth Edison Company service territory. It represents the best quality level of metered load within the ComEd zone. We analyze two separate datasets: one consisting of the entire dataset and the other consisting of data from weekdays[15] only[16]. The datasets are then separated into twenty-four subsets[17], each containing the load for a specific hour of the day. Thus, we estimate the load for the first hour of the day with one equation and the load for the second hour of the day from a different equation, and so on for all twenty-four hours of a day (Ramanathan et al. 1997). This approach avoids having to model the complicated intraday patterns in hourly load (i.e., the load profile) and allows each hour to have a distinct weekly pattern[18].

A. DEPENDENT VARIABLE – LOAD

Summary statistics for the load variable, measured in megawatt hours (MWh) are presented in table 1 for the full in-sample data series and table 2 for the weekday in-sample data series[19]. As can be seen from figure 1 and figure 2, the data series appear to be stationary and they display clear daily, weekly,

---

(RTO) May 1, 2004. Thus for convenience of acquiring data on our dependent variable from one source, we chose this date as the starting point for our estimation period.

[14] http://www.pjm.com/markets-and-operations/compliance/nerc-standards/historical-load-data.aspx

[15] For the weekday estimation we consider a dataset containing hourly loads from May 3, 2004 through September 30, 2009. The data from the period from May 3, 2004 through April 30, 2008 are used for estimation purposes (in-sample), and the data from the period May 1, 2008 through September 30, 2009 are left for out-of-sample forecast evaluation.

[16] Ramanathan et al. (1997) estimate weekends and weekdays separately thus requiring 48 separate equations to forecast a full day. Although we do not estimate weekends separately, we are interested in determining whether modeling business days (Monday through Friday) separately alleviates any estimation problems associated with modeling the full-week. We will discuss specifics in the empirical method section.

[17] Actually 48 subsets if one considers the weekday and full-week data separately.

[18] Hippert et al. (2001) note that the difficulties associated in modeling the load profile are common to almost all of the load forecasting papers they review.

[19] Both series underwent corrections for daylight savings time. Spring's daylight savings occurred on the following dates within our entire sample: 4/3/2005, 4/2/2006, 3/11/2007, 3/9/2008, and 3/8/2009. During the spring's daylight savings time, an hour is lost. We replace the null value for the H1AM2AM with an average of the load from H12AM1AM AND H2AM3AM. Fall's daylight savings occurred on the following dates within our entire sample: 10/31/2004, 10/30/2005, 10/29/2006, 11/4/2007, and 11/2/2008. During the fall's daylight savings time, an hour is gained. We choose the value associate with the original H12AM1AM value and exclude the "gained" hour's value. The reasoning behind this treatment is that electricity consumers are likely to treat the original H12AM1AM hour as they normally do, and the "gained" hour for extra sleeping time.

and annual seasonality. We tested each series (hour) for unit roots using the augmented Dickey-Fuller and Phillips-Perron unit root tests, and both tests indicated the load series are stationary in levels and logs for both the weekday and full-week data (Dickey and Fuller 1979, 1981; Phillips and Perron 1988). The daily seasonality is apparent when comparing the evening hours to the daytime hours. The annual seasonality may largely be attributed to temperature. The peaks in the summer months occur during the hottest time of the year when electric air conditioners are powered on. The smaller peaks in the winter months can be attributed to the fact that only about one-tenth of Illinois households actually use electricity as their primary energy source for home heating (EIA 2009). This annual seasonality is apparent even after the exclusion of weekends, as is apparent from figure 2. The influence of holidays is apparent by the sharp drops in demand during the daytime hours, which is most easily seen in the weekday series in figure 2. Summary statistics regarding the variables we use to model these factors are given in table 3 and are discussed later in this section.

For estimation purposes, we take the natural logarithm of the MW-hour net energy for load as consumed by the Commonwealth Edison Company service territory[20]. We chose to work with logarithms of the load series because as Soares and Souza (2006) point out it allows the weekly seasonality and the holiday effect to be modeled additively, whereas they are multiplicative in the original series. Soares and Souza (2006) note that these "effects are believed to be multiplicative as consumption tends to vary proportionally with the number of consumers" (23).

B. Calendar Effects

The daily pattern or seasonality is removed by modeling each hour of the day separately (Fiebig et al. 1991; Peirson and Henley 1994; Ramanathan et al. 1997; Cottet and Smith 2003; Soares and Medeiros

[20] Although the results are not given in this paper, we did experiment with applying the same models to the original series and the results were essentially the same.

2005, 2008; Soares and Souza 2006). We include seasonal dummies to model the weekly pattern and the influence of holidays. We group several workdays together as one dummy variable, namely Tuesday, Wednesday, and Thursday, because we feel it appropriate to model these as a single type of day and it allows for a more parsimonious estimation. We model the other days-of-the-week as individual dummy variables and exclude one in each estimation. For the weekday model, we include one holiday dummy variable that reflects public holidays[21] for Federal employees as established by Federal law (5 U.S.C. 6103) and published on the U.S. Office of Personnel Management website[22].

C. WEATHER DATA

For modeling the weather component, we use data obtained from the National Climatic Data Center[23] for the Chicago O'Hare International Airport[24]. We obtained the average daily[25] temperature in degrees Fahrenheit and the average daily wind speed in knots. A wind speed adjusted temperature (WWP)[26] was utilized in some of the estimated equations to better model the winter weather component of the load series. The WWP is a measure of cold stress in winter and is widely used[27] by electric utilities (PJM

---

[21] Public holidays include New Year's Day, Birthday of Martin Luther King, Jr., Washington's Birthday, Memorial Day, Independence Day, Labor Day, Columbus Day, Veterans Day, Thanksgiving Day, and Christmas Day. Since most Federal employees work on a Monday through Friday schedule, when a holiday falls on a nonworkday such as a Saturday or Sunday, the holiday usually is observed on Monday (if the holiday falls on Sunday) or Friday (if the holiday falls on Saturday). The actual day the holiday is "observed" (as opposed to the actual holiday) is included in the holiday dummy variable in our weekday model.

[22] http://www.opm.gov/Operating_Status_Schedules/fedhol/2009.asp

[23] http://www.ncdc.noaa.gov/oa/land.html The National Climatic Data Center (NCDC) is part of the National Oceanic and Atmospheric Administration (NOAA) and the U.S. Department of Commerce.

[24] Weather data for Chicago O'Hare International Airport were thought to be a good proxy for the weather in the ComEd service territory because most of ComEd's load stems from Chicago and its surrounding areas, plus the data were available for the entire sample period.

[25] To our dismay, we had to use daily averages rather than the hourly series because the hourly data that NCDC provided was inadequate for estimation purposes (e.g., random hourly observations missing).

[26] The average daily wind speed in knots was converted to miles per hour (mph) by multiplying the series by 1.15077945.
WWP=TEMP-(0.5*(WDSPmph-10)) if WDSPmph>10mph; WWP=TEMP if wind≤10mph.
TEMP is the mean temperature for the day in degrees Fahrenheit to tenths.
WDSP is the mean wind speed for the day in knots to tenths.
WDSPmph=WDSP*1.15077945 converting knots to mph.

[27] Also widely used is the Temperature-Humidity Index (THI), though we were not able to find quality controlled humidity level data.

Resource Adequacy Planning 2009; Feinberg and Genethliou 2005; Weron 2006).

We investigate the relationship between load and temperature by examining scatter plots of the two series. The scatter plots of temperature and load at various hours indicate that indeed the relationship is nonlinear, which is consistent with the vast majority of the load forecasting literature (Ramanathan et al. 1997; Hor et al. 2005; Darbellay and Slama 2000; Cottet and Smith 2003; Kyriakides and Polycarpou 2007). In fact, the relationship varies depending on the time of day, which further supports the multi-equation specification used as the estimation procedure, which allows the temperature coefficient to vary depending on the hour of the day (and we allow for the month of the year in some cases). It appears there is a possibility of structural instability in the relationship for the full-week series. Examining the scatter plots of load versus temperature for the full-week series (figure 3a and figure 3c) and the weekday series (figure 3b and figure 3d), this instability seems to vanish once the weekends are removed.

V. Econometric Method and Forecasting Results

Some studies that have employed a multi-equation approach have applied the same model to all twenty-four hours of load data (Soares and Souza, 2006). We test whether such a specification falls short in terms of producing the optimal and most parsimonious model for all hours considered. The antithesis would be to create a unique model each hour; however, such an approach is somewhat cumbersome since many hours during the day and throughout portions of the week reflect the same patterns. For the sake of this study, we begin our investigation by exploring whether fitting a single model for all twenty-four hourly series is a viable forecasting approach for the dataset at hand. Accordingly, a parsimonious model is developed and fit to the entire series (i.e., all twenty-four equations). Then, after analyzing the residuals and the out-of-sample forecasting performance, the full-week data is disaggregated and modeled in a way that results in the most parsimonious forecasting model that not only has superior out-of-sample forecasting performance, but also is easily constructed and applied.

A. Full-week

1. Specification and Estimation Method

In choosing the best model, an analysis of the autocorrelation function (ACF), the partial autocorrelation function (PACF), the Schwarz Information Criterion (SIC) and the Akaike Information Criterion (AIC) for the 3AM4AM model was taken into account (Schwarz 1978; Akaike 1974, 1987; Diebold 2007; Makridakis 1998). This hour in particular was chosen to construct the full-week model due to its comparatively simple dynamics. We modeled the weekly seasonality first by including dummy variables for Monday, Friday, Sunday, and grouped Tuesday, Wednesday, and Thursday together[28]. The temperature component is the most complicated part of the model since temperature variables are very important in modeling short-run fluctuations in electricity consumption, thus we allow for a rich specification of the component (Ramanathan et al. 1997). We are able to estimate the model with the actual temperature, although in reality a forecaster would insert the forecasted temperature that the weather service provides. The current temperature as well as its lag (e.g., temperature from the day before) are included as well as their squares (i.e., allowing a nonlinear quadratic relationship). The temperature effect is not constant across months, thus for the summer months (i.e., June, July, and August) we include interaction terms between the monthly dummy variable and temperature[29]. Finally,

[28] Cottet and Smith (2003) find that the estimates of the dummy variable coefficients were similar for workdays Monday through Friday with a slightly lower load on Friday afternoon and Monday morning, "reflecting spillover from the weekends in NSW [New South Wales] working patterns. Loads on Saturday are higher than those on Sunday up until around 17:00, which reflects retail trading during Saturday. Loads on the public holidays vary depending on the holiday type, although they follow a profile similar to that of Sunday" (843).

[29] As noted previously we were unable to obtain a measure of humidity for the area, which would likely have proved useful in modeling the temperature effect in the summer. Other papers have included interaction terms between some months and temperature as well (Ramanathan et al. 1997). Cottet and Smith (2003) find that "for mid-summer, the seasonal effect on load

we determine the appropriate lag length for the autoregressive moving average (ARMA) terms by analyzing the ACF and the PACF of the residuals after taking the daily and weekly patterns as well as the weather component into account. We determine an ARMA(2,1) model to be most appropriate for the 3AM4AM model after examination of the SIC and the AIC (Schwarz 1978; Akaike 1974, 1987). The full-week model used for the estimation of each hour is presented below. We omit subscripts for the hours because the specification is the same. Let the hourly load for day *t* be represented by $Load_t$ and let $Y_t = ln(Load_t)$.

$$\boldsymbol{Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \beta_f F + \beta_{su} SU + \delta_{tmp} TMP_t + \delta_{tmp2} {TMP_t}^2 + \delta_{ltmp} TMP_{t-1} + \delta_{ltmp2} {TMP_{t-1}}^2 + \delta_{m6tmp} JUN * TMP_t + \delta_{m7tmp} JUL * TMP_t + \delta_{m8tmp} AUG * TMP_t + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1} \varepsilon_{t-1} + \varepsilon_t} \quad \textbf{(5)}$$

The estimated parameters and the corresponding adjusted R^2 for each hour of the day are illustrated in table 4[30]. The vast majority of estimated coefficients are significant at the one percent level. The table also shows the p-value of the Ljung-Box[31] test for no error serial autocorrelation of order five and six (Ljung and Box 1978). The Q-statistic is often used as a test of whether the series is white noise[32]. However it should be noted that Dezhbakhsh (1990)[33] finds the application of the Ljung-Box test to be

is highest during the period 10:30-17:00, probably due to preprogrammed office air-conditioning" (844).

[30] Possible heteroskedasticity is taken into account in statistical inference using the Newey-West correction for heteroskedasticity and autocorrelation (Newey and West 1994). Though the standard errors are not listed in this table in order to conserve space, these can be provided for the interested reader by emailing the authors.

[31] The Ljung-Box Q-statistic at lag k is a test statistic for the null hypothesis that there is no autocorrelation up to order k (EViews 2008; Ljung and Box 1978). $Q_{LB} = T(T+2)\sum_{j=1}^{k} \frac{\tau_j^2}{T-J}$ where τ_j is the j-th autocorrelation and T is the number of observations (EViews 2008).

[32] Choosing the order of lag to use for the test is a practical problem. By choosing too small a lag, the test may not detect serial correlation at higher-order lags. By choosing too large a lag, the test may have low power since the significant correlation at one lag may be diluted by insignificant correlations at other lags (EViews 2008). Thus, we decide to report the p-values of the Q-statistics for two different lag lengths, 5 and 6 for the full-week model.

[33] Dezhbakhsh (1990) evaluates the performance of several tests using Monte Carlo experiments. Dezhbakhsh reports that

inadequate when applied to linear models with lagged dependent variables and exogenous regressors (e.g., ARMAX models) (EViews 2008; Ljung and Box 1978). However, since these statistics are still reported in most of the recent load forecasting literature, we report them with the estimation results (Soares and Souza 2006; Soares and Medeiros 2005; 2008). Using the Ljung-Box test, there appears to be remaining autocorrelation for several hours during the day, namely hours 11 through 23. Increasing the lag order of both the autoregressive and moving average terms does not attenuate the problem[34], thus indicating the possibility of misspecification in the full-week model. Nevertheless, we still report the out-of-sample forecasting results for the full-week model in the next section.

2. Forecasting Results

Table 5 reports the one-step ahead out-of-sample forecasting results for each hour for the May 1, 2008 through September 30, 2009 period using the model described in the previous section and the parameter estimates reported in table 4 (i.e., the model is not re-estimated during the out-of-sample period). Although the most popular measure of forecasting accuracy in the load forecasting literature is the mean absolute percentage error (MAPE) (Soares and Souza 2006; Darbellay and Slama 2000), we report the MAPE[35] as well as the root mean squared error[36] (RMSE) and the mean absolute error[37]

"the results warn against using the popular DW [Durbin-Watson statistics] and *Q* tests and provide support for using Durbin's *h* and, in particular *m* tests" (1990, 127). Unfortunately, Durbin's m-test was not easily accessible, so we use the Ljung-Box Q-statistic as is still widely used in the literature (EViews 2008; Ljung and Box 1978; Soares and Souza 2006; Soares and Medeiros 2005; 2008).

[34] Besides changing the ARMA specification, it was investigated whether the inclusion of a holiday dummy variable (the holiday dummy was originally excluded because we modeled the hour 3AM4AM and the holiday dummy did not seem to have an impact during this period), removal of insignificant temperature variables, as well as various other changes would generate white noise residuals; however, none seemed to solve the problem. These facts point to the possible misspecification of the full-week model.

[35] Mean Absolute Percentage Error (MAPE) = $100\sum_{t=T+1}^{T+h}\left|\frac{\hat{y}_t-y_t}{y_t}\right|/h$ where y_t is the actual value in period t and $\hat{y}_t$ is the forecasted value in period t, and the forecast sample is $j=T+1,\ T+2,\ \ldots,\ T+h$ (EViews 2008).

[36] Root Mean Squared Error (RMSE) = $\sqrt{\sum_{t=T+1}^{T+h}(\hat{y}_t-y_t)^2/h}$ where y_t is the actual value in period t and $\hat{y}_t$ is the forecasted value in period t, and the forecast sample is $j=T+1,\ T+2,\ \ldots,\ T+h$ (EViews 2008).

[37] Mean Absolute Error (MAE) = $\sum_{t=T+1}^{T+h}|\hat{y}_t-y_t|/h$ where y_t is the actual value in period t and $\hat{y}_t$ is the forecasted value in period t, and the forecast sample is $j=T+1,\ T+2,\ \ldots,\ T+h$ (EViews 2008).

(MAE). Note that the RMSE and the MAE depend on the scale of the dependent variable[38]. All three statistics agree that the forecasting model performs the best for hour 4 (3AM4AM) and the worst for hour 17 (5PM6PM). Given that hourly loads on the weekend display a much different pattern than during the weekdays, removing weekends from the dataset may affect which hours have the best and worst forecasting performance. Specifically, it is of interest to see whether the best and the worst forecasting performance occur during the same hours[39].

B. WEEKDAYS

1. Specification and Estimation Method

In choosing the best model for the weekday dataset, an analysis of the autocorrelation function (ACF), the partial autocorrelation function (PACF), the Schwarz Information Criterion (SIC) and the Akaike Information Criterion (AIC) for the 5PM6PM model[40] was taken into account (Schwarz 1978; Akaike 1974; Diebold 2007; Makridakis 1998). We modeled the weekly seasonality first by including a dummy variable for Monday as well as the Tuesday, Wednesday, and Thursday group dummy. We also include a dummy variable to model observed holidays. We include the wind speed adjusted temperature (WWP)[41] as well as its squared value, to model the winter weather component of the load series. Finally,

[38] Though we estimate our model using the natural log of load, we forecast the actual value of the series rather than the natural log of the value.

[39] Since we alter the dataset by excluding the weekends for the weekday model, we will not be able to compare the actual MAPE, RMSE, or MAE across models. Soares and Medeiros (2008) state that several "authors achieve MAPEs as low as 2% when predicting the total daily load," but they point out that "the results of different models cannot be compared on different datasets because the differences among load curves in different countries" (639). Soares and Medeiros (2008) further note that "if different datasets are used, the same model(s) must be used, and the comparison should be made among datasets and not models. If the researcher wants to compare the performance of different models, the same data with the same forecasting period must be used" (639).

[40] We were most interested in analyzing the 5PM6PM model (hour 17) first because of the inability to obtain white noise residuals in the full-week estimation for this hour as well as the fact that it performed the worst in the out-of-sample forecasting for the full-week model.

[41] The average daily wind speed in knots was converted to miles per hour (mph) by multiplying the series by 1.15077945.
WWP=TEMP-(0.5*(WDSPmph-10)) if WDSPmph>10mph; WWP=TEMP if wind≤10mph.
TEMP is the mean temperature for the day in degrees Fahrenheit to tenths.
WDSP is the mean wind speed for the day in knots to tenths.

we determine the appropriate lag length for the autoregressive moving average (ARMA) terms by analyzing the ACF and the PACF of the residuals after taking the daily and weekly patterns as well as the holiday and weather component into account. We determine an ARMA(2,1) model to be most appropriate for the 5PM6PM model after examination of the SIC and the AIC (Schwarz 1978; Akaike 1974). The weekday model used for the estimation of each hour is presented below. We omit subscripts for the hours because the specification is the same. Let the hourly load for day *t* be represented by $Load_t$ and let $Y_t = ln(Load_t)$.

$$Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \delta_{wwp} WWP_t + \delta_{wwp2} WWP_t{}^2 + \lambda_{hol} H + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1}\varepsilon_{t-1} + \varepsilon_t \qquad \textbf{(6)}$$

The estimated parameters and the corresponding adjusted R^2 for each hour of the day are illustrated in table 6. Possible heteroskedasticity is taken into account in statistical inference using the Newey-West correction for heteroskedasticity and autocorrelation (Newey and West 1994), thus the heteroskedasticity and autocorrelation robust standard errors are in parentheses. The vast majority of estimated coefficients are significant at the one percent level. The table also shows the p-value of the Ljung-Box[42] test for no error serial autocorrelation of order five, six, and seven (Ljung and Box 1978). Using the Ljung-Box test, there appears to be remaining autocorrelation of order five for several hours during the evening, namely hours 23, 24, and 1 through 5; because of this, we decide to estimate the night hours using a different model, which is very similar[43] to the model chosen for the full-week

WDSPmph=WDSP*1.15077945 converting knots to mph.

[42] The Ljung-Box Q-statistic at lag k is a test statistic for the null hypothesis that there is no autocorrelation up to order *k* (EViews 2008; Ljung and Box 1978). $Q_{LB} = T(T+2)\sum_{j=1}^{k}\frac{\tau_j^2}{T-J}$ where τ_j is the j-th autocorrelation and T is the number of observations (EViews 2008).

[43] Even though this is the weekday only model, the lagged temperature variable represents the temperature from the day before, even if the day before is a Sunday.

estimation. Let the hourly load for day *t* be represented by $Load_t$ and let $Y_t = ln(Load_t)$.

$$\boldsymbol{Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \beta_f F + \delta_{tmp} TMP_t + \delta_{tmp2} {TMP_t}^2 + \delta_{ltmp} TMP_{t-1} + \delta_{ltmp2} {TMP_{t-1}}^2 + \delta_{m6tmp} JUN * TMP_t + \delta_{m7tmp} JUL * TMP_t + \delta_{m8tmp} AUG * TMP_t + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1} \varepsilon_{t-1} + \varepsilon_t} \quad (7)$$

The estimated parameters using this model for the night time hours and the previous model for the daytime hours and the corresponding adjusted R^2 for each hour of the day are illustrated in table 7. The heteroskedasticity and autocorrelation robust standard errors are in parentheses (Newey and West 1994). The vast majority of estimated coefficients are significant at the one percent level. The table also shows the p-value of the Ljung-Box[44] test for no error serial autocorrelation of order five, six, and seven (Ljung and Box 1978). Using the Ljung-Box test, the residuals for all hours of the day and night exhibit white noise, thus we proceed to evaluating the out-of-sample forecast performance. We are in agreement with Soares and Medeiros (2008) when they conclude that a "point that deserves attention is the fact that the final model specification differs across hours, corroborating our view that different hours need to be modeled separately because they have different structures and dynamics" (637-8).

2. An Example for Hour 17 (5PM6PM)

Explanatory variables for hour 17 in the weekday model include dummy variables for Monday (*M*), one for Tuesday, Wednesday, and Thursday (*TWT*), and one for Federal holidays (*H*). The explanatory variables also include a wind speed adjusted temperature (*WWP*)[45] as well as *WWP* squared[46] to better

[44] The Ljung-Box Q-statistic at lag k is a test statistic for the null hypothesis that there is no autocorrelation up to order k (EViews 2008; Ljung and Box 1978). $Q_{LB} = T(T+2)\sum_{j=1}^{k} \frac{\tau_j^2}{T-J}$ where τ_j is the j-th autocorrelation and T is the number of observations (EViews 2008).

[45] The average daily wind speed in knots was converted to miles per hour (mph) by multiplying the series by 1.15077945. WWP=TEMP-(0.5*(WDSPmph-10)) if WDSPmph>10mph; WWP=TEMP if wind≤10mph.

model the winter weather component of the load series. The *WWP* is a measure of cold stress in winter and is widely used by electric utilities (PJM 2009; Feinberg and Genethliou 2005; Weron 2006). ARMA(2,1) components were also included in the estimation. Let the 5PM6PM hourly load (hour 17) for day *t* be represented by $Load_{17,t}$ and let $Y_{17,t}=ln(Load_{17,t})$.

$$Y_{17,t} = \alpha_0 + \beta_m M + \beta_{twt} TWT + \delta_{wwp} WWP_t + \delta_{wwp2} WWP_t{}^2 + \lambda_{hol} H + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1}\varepsilon_{t-1} + \varepsilon_t \quad (8)$$

Hour[a]	$\hat{\alpha}_0$	$\hat{\beta}_m$	$\hat{\beta}_{twt}$	$\hat{\lambda}_{hol}$	$\hat{\delta}_{wwp}$	$\hat{\delta}_{wwp2}$	$\hat{\varphi}_{ar1}$	$\hat{\varphi}_{ar2}$	$\hat{\theta}_{ma1}$	Adj. R^2	$Q_{LB}(5)$[b]	$Q_{LB}(6)$	$Q_{LB}(7)$
17	9.6641	0.0267	0.0244	-0.1195	-0.0157	0.0002	1.2792	-0.2955	-0.7905	0.84	0.467	0.250	0.359
	(0.0391)	(0.0054)	(0.0039)	(0.0173)	(0.0014)	(0.0000)	(0.0535)	(0.0508)	(0.0355)				

Notes: Newey-West heteroskedasticity and autocorrelation robust standard errors are in parentheses (Newey and West 1994). Estimation sample includes the period 05/05/2004 – 04/30/2008, *n*=1041 and excludes weekends.

Plugging the estimated coefficients into equation (8), the estimated model for hour 17 can be written compactly as in equation (9).

$$Y_{17,t} = 9.66 + 0.03M + 0.02TWT - 0.02WWP_t + 0.0002WWP_t{}^2 - 0.12H + 1.28Y_{17,t-1} - 0.3Y_{17,t-2} - 0.79\varepsilon_{17,t-1} + \varepsilon_{17,t} \quad (9)$$

Equation (9) was estimated by ordinary least squares[47] and the Newey-West heteroskedasticity and autocorrelation robust standard errors were analyzed when determining the statistical significance of the estimated parameters (Newey and West 1994). All of estimated parameters are statistically significant at

TEMP is the mean temperature for the day in degrees Fahrenheit to tenths.
WDSP is the mean wind speed for the day in knots to tenths.
WDSPmph=WDSP*1.15077945 converting knots to mph.
[46] As can be seen in figures 3b and 3d, the relationship between load and temperature is not a linear one.
[47] EViews 6 was the software program used to estimate the equations.

the one percent level. The coefficient of determination indicates that roughly 84% of the variation in the 5PM6PM hourly load can be explained by all the explanatory variables included in the estimation taken together. The p-values of the Ljung-Box Q-statistic at lags 5, 6, and 7 fail to reject the null hypothesis that there is no autocorrelation up to order 5, 6, or 7, indicating the residuals are likely white noise.

Equation (9) for the 5PM-6PM (HOUR17) period shows the typical form of all of the equations. The dependent variable is the natural log of the hourly load for the 5PM-6PM hour over the period May 5, 2004 through April 30, 2008 (*n*=1041). Friday is the day of the week that is excluded from the estimation; therefore, Friday is considered part of the "reference" group. All of the estimated coefficients exhibit their expected signs. Since this is a semi-logarithmic equation the interpretation of the dummy variable coefficients as well as the *WWP* coefficients involves the formula: $100 * (e^{\widehat{\beta}_m} - 1)$, where $\hat{\beta}_m$ is the estimated coefficient (Wooldridge 2009; Halvorsen and Palmquist 1980). For dummy variables, this formula derives the percentage effect on hourly load of the presence of the factor represented by the dummy variable. Thus, we utilize this formula and proceed to interpretation of the estimated equation.

The estimated coefficient for Monday (*M*) in equation (9) indicates that the 5PM-6PM hourly load is 2.7%[48] higher on Monday (than Friday, the reference group). This is as expected since businesses are likely to be in full swing by 5PM-6PM at the beginning of the workweek and residential customers just getting home may engage in energy intensive activities (e.g., lighting and cooking), thus consuming more energy than at the end of the workweek (Friday) at the same time of day (5PM-6PM) when businesses may close early before the weekend and residential customers may decide to eat out at a later time. The estimated coefficient for Tuesday, Wednesday, and Thursday (*TWT*) in equation (9) indicates

[48] $e^{0.0267} \approx 1.027; 100 * (e^{0.0267} - 1) \approx 2.7\%$

the 5PM-6PM hourly load is 2.5%[49] higher on Tuesday, Wednesday, and Thursday (than Friday, the reference group). This is as expected since businesses are likely to be in full swing during the middle of the workweek and residential customers just getting home may engage in energy intensive activities (e.g., lighting and cooking), thus consuming more energy than at the end of the workweek (Friday) at the same time of day (5PM-6PM) when businesses may close early before the weekend and residential customers may decide to eat out at a later time.

The interpretation of the impact that the wind speed adjusted temperature variable (WWP) has on the 5PM-6PM hourly load must be analyzed with care. Starting at zero degrees Fahrenheit, as it warms up outside by one degree, there is a reduction in the 5PM-6PM hourly load by 1.56%[50], ceteris paribus. Taking the partial derivative of $Y_{17,t}$ with respect to WWP; then setting the equation equal to zero; and finally solving for WWP, we find that after 39.25[51] degrees Fahrenheit, an increase in temperature begins to increase the 5PM-6PM hourly load. In going from 39 degrees to 40 degrees Fahrenheit, load is predicted to decrease by 0.01%[52]. In going from 40 degrees to 41 degrees Fahrenheit, load is predicted to increase by 0.03%[53]. The mean temperature in the sample is about 51 degrees Fahrenheit. In going from 51 degrees to 52 degrees Fahrenheit, load is predicted to increase by 0.47%[54]. In going from 75 degrees to 76 degrees Fahrenheit, load is predicted to increase by 1.44%[55]. Thus, below 39.25 degrees Fahrenheit, the impact that temperature has on load is that load decreases at a decreasing rate as temperature increases. Above 39.25 degrees Fahrenheit, the impact that temperature has on load is that

[49] $e^{0.0244} \approx 1.0247; 100 * (e^{0.0244} - 1) \approx 2.47\%$
[50] $100 * (e^{-0.0157} - 1) \approx -1.56\% \approx \Delta\hat{Y}_{17,t}$
[51] $-0.0157WWP_t + 0.0002WWP_t{}^2$
$-0.0157 + 2 * 0.0002 * WWP_t$
$0.0004 * WWP_t = 0.0157$
$WWP_t = \frac{0.0157}{0.0004}$
$WWP_t = 39.25\ degrees\ Fahrenheit$
[52] $-0.0157 + 2 * 0.0002 * (39) \approx -0.0001;\ e^{-0.0001} \approx 0.9999;\ 100 * (e^{-0.0001} - 1) \approx -0.01\%$
[53] $-0.0157 + 2 * 0.0002 * (40) \approx 0.003;\ e^{0.003} \approx 1.0003;\ 100 * (e^{0.003} - 1) \approx 0.03\%$
[54] $-0.0157 + 2 * 0.0002 * (51) \approx 0.0047;\ e^{0.0047} \approx 1.0047;\ 100 * (e^{0.0047} - 1) \approx 0.47\%$
[55] $-0.0157 + 2 * 0.0002 * (75) \approx 0.0143;\ e^{0.0143} \approx 1.0144;\ 100 * (e^{0.0143} - 1) \approx 1.44\%$

load increases at an increasing rate as temperature increases.

The estimated coefficient for Federal holidays (*H*) in equation (9) indicates the 5PM-6PM hourly load is 11.3%[56] lower on Federal holidays (than non-holidays). This is as expected since many businesses close on Federal holidays, thus reducing their demand for energy.

A 1% increase in the 5PM-6PM hourly load the previous day ($Y_{17,t-1}$) is expected to have a positive impact on the 5PM-6PM hourly load today ($Y_{17,t}$) by about 1.28%[57], ceteris paribus. A 1% increase in the 5PM-6PM hourly load 2 days before ($Y_{17,t-2}$)[58] is expected to inversely impact the 5PM-6PM hourly load today ($Y_{17,t}$) by about 0.3%, ceteris paribus. A 1% increase in the 5PM-6PM hourly load disturbances (i.e., shocks to the system) ($\varepsilon_{17,t-1}$) from the previous day is expected to inversely impact the 5PM-6PM hourly load today ($Y_{17,t}$) by about 0.79%, ceteris paribus.

3. Forecasting Results

Table 8 reports the one-step ahead out-of-sample forecasting results for each hour for the May 1, 2008 through September 30, 2009 period using the models described in the previous section and the parameter estimates reported in table 6 and table 7 (i.e., the model is not re-estimated during the out-of-sample period). We compare the forecasting performance from the weekday model estimated in table 6 to that estimated in table 7. Estimating the weekday and weeknights using separate models proves to be superior[59] as the lower values of the RMSE, MAE, and MAPE indicate[60]. All three statistics agree that the forecasting model performs the best for hour 4 and the worst for hour 17, which is the same conclusion we reached in the full-week model. Thus the hours which exhibited the best and the worst out-of-sample forecasting performance are likely to be a function of the volatility experienced during

[56] $e^{-0.1195} \approx 0.887; 100 * (e^{-0.1195} - 1) \approx -11.26\%$
[57] Since the $Y_{17,t}$ is actually the natural log of load.
[58] Note that Y_{t-1} and Y_{t-2} actually represents the load for weekdays only, thus if Y_t is a Monday, then Y_{t-1} and Y_{t-2} represent load from Friday and Thursday, respectively.
[59] Although the RMSE chooses the weekday model (table 6) for hour 5.
[60] Comparing the RMSE, MAE, and MAPE for hours 23, 24, and 1 through 5 across models in table 8 shows these statistics to be lower when the model from table 7 is used.

those hours (e.g., hour 4 has the smallest standard deviation as can be seen in table 1 and table 2). It may also be warranted to use a more sophisticated modeling approach for hour 17 (5PM6PM)[61]. A measure of daylight hours may help improve forecasting performance in that it affects public lighting during different times of the year; however, we did not have access to this type of data.

VI. Summary and Conclusions

As electricity markets have deregulated over the last decade, accurate load forecasts have become a vital part of a utility's long-, medium-, and short-term generation and procurement planning. An inaccurate load forecast can have severe consequences for ComEd's customers in the form of higher rates. The objective of this paper was to forecast short-term electricity load using ARMAX models with hourly load data for the Commonwealth Edison Company (ComEd). To date, no published study to our knowledge has used ARMAX modeling to forecast electricity load in ComEd's territory.

Given the success with multi-equation models in the literature, each hour of the day is constructed as a separate series. Constructing a single model to apply to all of the twenty-four series, a model that excludes weekends, and a separate nighttime model reveals that it is necessary to model peak and off-peak hours separately. We found that applying a peak and off-peak model to the corresponding peak and off-peak hour's works relatively well. Ultimately, this study provides further evidence regarding the importance of modeling off-peak and peak hours separately. Our approach results in a parsimonious forecasting model that not only has superior out-of-sample forecasting performance, but also is easily constructed and applicable for day-to-day load forecasts for territories similar ComEd's.

Future research could consider those hours that exhibit greater complexity (e.g., 5PM6PM) and examine the possibility of time-varying volatility. Measuring any improvement in forecasting

[61] However, this was beyond the scope of the present study which is focused on parsimony and ease of use.

performance from modeling each hour separately versus modeling one peak and one off-peak hour would be useful.

VII. Tables

Table 1. Summary statistics: Load for each full-week hour from May 1, 2004 through April 30, 2008

Hour		*Mean*[a]	*Median*	*Max.*	*Min.*	*Std. Dev.*
H1AM2AM	1	9,680	9,449	16,547	7,385	1,274
H2AM3AM	2	9,428	9,218	15,890	7,237	1,198
H3AM4AM	3	9,323	9,131	15,467	7,104	1,155
H4AM5AM	4	9,449	9,279	15,370	7,190	1,149
H5AM6AM	5	9,948	9,822	15,705	6,975	1,236
H6AM7AM	6	10,796	10,765	16,495	7,012	1,496
H7AM8AM	7	11,481	11,551	17,833	7,440	1,680
H8AM9AM	8	11,955	11,943	19,111	7,878	1,746
H9AM10AM	9	12,285	12,197	20,168	8,279	1,837
H10AM11AM	10	12,596	12,368	21,385	8,418	2,006
H11AM12PM	11	12,758	12,393	22,265	8,472	2,184
H12PM1PM	12	12,841	12,348	22,812	8,577	2,353
H1PM2PM	13	12,929	12,362	23,336	8,474	2,521
H2PM3PM	14	12,933	12,294	23,491	8,383	2,629
H3PM4PM	15	12,940	12,257	23,613	8,226	2,672
H4PM5PM	16	13,101	12,513	23,613	8,305	2,622
H5PM6PM	17	13,247	12,831	23,386	8,505	2,509
H6PM7PM	18	13,197	12,872	22,970	8,621	2,324
H7PM8PM	19	13,115	12,782	22,431	8,774	2,116
H8PM9PM	20	13,008	12,621	22,416	9,138	1,991
H9PM10PM	21	12,566	12,155	21,929	9,159	1,903
H10PM11PM	22	11,689	11,302	20,578	8,821	1,721
H11PM12AM	23	10,757	10,425	18,762	8,249	1,530
H12AM1AM	24	10,095	9,809	17,459	7,812	1,381

Notes: n=1,461

[a]Summary statistics are in megawatt hours (MWh).

TABLE 2. SUMMARY STATISTICS: LOAD FOR EACH WEEKDAY HOUR FROM MAY 1, 2004 THROUGH APRIL 30, 2008

Hour		*Mean*[a]	*Median*	*Max.*	*Min.*	*Std. Dev.*
H1AM2AM	1	9,781	9,523	16,547	7,385	1,294
H2AM3AM	2	9,540	9,333	15,890	7,237	1,211
H3AM4AM	3	9,455	9,274	15,467	7,104	1,159
H4AM5AM	4	9,638	9,460	15,370	7,190	1,131
H5AM6AM	5	10,287	10,136	15,705	7,055	1,131
H6AM7AM	6	11,378	11,262	16,495	7,139	1,188
H7AM8AM	7	12,202	12,083	17,833	7,446	1,241
H8AM9AM	8	12,663	12,498	19,111	7,878	1,350
H9AM10AM	9	12,948	12,627	20,168	8,285	1,526
H10AM11AM	10	13,250	12,766	21,385	8,976	1,744
H11AM12PM	11	13,403	12,772	22,265	9,154	1,966
H12PM1PM	12	13,488	12,753	22,812	9,172	2,166
H1PM2PM	13	13,612	12,773	23,336	8,980	2,339
H2PM3PM	14	13,610	12,713	23,491	8,786	2,464
H3PM4PM	15	13,597	12,714	23,613	8,670	2,520
H4PM5PM	16	13,726	13,009	23,613	9,014	2,473
H5PM6PM	17	13,825	13,295	23,386	9,072	2,367
H6PM7PM	18	13,726	13,347	22,970	8,965	2,192
H7PM8PM	19	13,609	13,202	22,431	8,931	1,993
H8PM9PM	20	13,464	12,954	22,416	9,414	1,886
H9PM10PM	21	12,959	12,441	21,929	9,461	1,822
H10PM11PM	22	11,987	11,516	20,578	8,999	1,666
H11PM12AM	23	10,854	10,484	18,762	8,249	1,568
H12AM1AM	24	10,187	9,893	17,459	7,812	1,411

Notes: Sample excludes weekends. n=1,043

[a]Summary statistics are in megawatt hours (MWh).

TABLE 3. SUMMARY STATISTICS: VARIABLES

Name	*Type*	*Definition*	*n*	*Mean*	*Median*	*Max.*	*Min.*	*Std. Dev.*
Full Week								
M	Dummy	equals 1 if Monday, otherwise equals 0	1,461	0.14	0.00	1.00	0.00	0.35
TWT	Dummy	equals 1 if Tuesday, Wednesday, or Thursday, otherwise equals 0	1,461	0.43	0.00	1.00	0.00	0.50
F	Dummy	equals 1 if Friday, otherwise equals 0	1,461	0.14	0.00	1.00	0.00	0.35
SU	Dummy	equals 1 if Sunday, otherwise equals 0	1,461	0.14	0.00	1.00	0.00	0.35
WWP	Continuous	Wind speed adjusted temperature[a] (°F)	1,461	50.71	51.90	88.99	-5.05	19.92
TMP	Continuous	Temperature in degrees Fahrenheit[b]	1,461	51.29	52.40	89.00	-3.70	19.65
JUN	Dummy	equals 1 if June, otherwise equals 0	1,461	0.08	0.00	1.00	0.00	0.27
JUL	Dummy	equals 1 if July, otherwise equals 0	1,461	0.08	0.00	1.00	0.00	0.28
AUG	Dummy	equals 1 if August, otherwise equals 0	1,461	0.08	0.00	1.00	0.00	0.28
Weekday								
WWP	Continuous	Wind speed adjusted temperature (°F)	1,043	50.95	52.14	88.99	-4.86	20.02
TMP	Continuous	Temperature in degrees Fahrenheit	1,043	51.56	52.70	89.00	-3.70	19.76
TMP_{t-1}	Continuous	Temperature in degrees Fahrenheit from the previous day	1,043	51.67	53.40	89.00	-3.70	19.77
H	Dummy	equals 1 if day is an observed holiday[c], otherwise equals 0	1,043	0.04	0.00	1.00	0.00	0.19

Notes: Summary statistics are for the in-sample estimation period.

[a]WWP=TMP-(0.5*(WDSPmph-10)) if WDSPmph>10mph; WWP=TMP if wind≤10mph. Where TMP is the mean temperature for the day in degrees Fahrenheit to tenths; WDSP is the mean wind speed for the day in knots to tenths; WDSPmph=WDSP*1.15077945 converting knots to mph.

[b]Temperature data were obtained for the Chicago O'Hare International Airport. http://www.ncdc.noaa.gov/oa/land.html The National Climatic Data Center (NCDC) is part of the National Oceanic and Atmospheric Administration (NOAA) and the U.S. Department of Commerce.

[c]Public holidays include New Year's Day, Birthday of Martin Luther King, Jr., Washington's Birthday, Memorial Day, Independence Day, Labor Day, Columbus Day, Veterans Day, Thanksgiving Day, and Christmas Day. Since most Federal employees work on a Monday through Friday schedule, when a holiday falls on a nonworkday such as a Saturday or Sunday, the holiday usually is observed on Monday (if the holiday falls on Sunday) or Friday (if the holiday falls on Saturday). The actual day the holiday is "observed" (as opposed to the actual day the holiday falls on) is included in the holiday dummy variable in our weekday model. http://www.opm.gov/Operating_Status_Schedules/fedhol/2009.asp

TABLE 4. PARAMETER ESTIMATES AND DIAGNOSTIC STATISTICS FOR THE FULL-WEEK MODEL

$$Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \beta_f F + \beta_{su} SU + \delta_{tmp} TMP_t + \delta_{tmp2} TMP_t{}^2 + \delta_{ltmp} TMP_{t-1} + \delta_{ltmp2} TMP_{t-1}{}^2 + \delta_{m6tmp} JUN * TMP_t + \delta_{m7tmp} JUL * TMP_t + \delta_{m8tmp} AUG * TMP_t + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1}\varepsilon_{t-1} + \varepsilon_t \quad (5)$$

Hour[a]	$\hat{\alpha}_0$	$\hat{\beta}_m$	$\hat{\beta}_{twt}$	$\hat{\beta}_f$	$\hat{\beta}_{su}$	$\hat{\delta}_{tmp}$	$\hat{\delta}_{tmp2}$	$\hat{\delta}_{ltmp}$	$\hat{\delta}_{ltmp2}$	$\hat{\delta}_{m6tmp}$	$\hat{\delta}_{m7tmp}$	$\hat{\delta}_{m8tmp}$	$\hat{\varphi}_{ar1}$	$\hat{\varphi}_{ar2}$	$\hat{\theta}_{ma1}$	Adj. R^2	$Q_{LB}(5)$[b]	$Q_{LB}(6)$
1	9.5170	-0.0348	0.0127	0.0129	-0.0469	-0.0122	0.0001	-0.0086	0.0001	0.0007	0.0009	0.0010	1.5426	-0.5471	-0.9005	0.89	0.314	0.509
2	9.4969	-0.0216	0.0189	0.0181	-0.0444	-0.0127	0.0001	-0.0078	0.0001	0.0006	0.0009	0.0009	1.5649	-0.5692	-0.9051	0.89	0.403	0.589
3	9.4856	-0.0077	0.0269	0.0248	-0.0444	-0.0130	0.0001	-0.0070	0.0001	0.0006	0.0008	0.0009	1.5959	-0.5997	-0.9113	0.90	0.276	0.415
4	9.4796	0.0167	0.0456	0.0418	-0.0496	-0.0129	0.0001	-0.0062	0.0001	0.0007	0.0008	0.0009	1.6189	-0.6225	-0.9160	0.90	0.119	0.230
5	9.4789	0.0646	0.0910	0.0842	-0.0651	-0.0125	0.0001	-0.0052	0.0001	0.0007	0.0008	0.0008	1.6157	-0.6192	-0.9152	0.90	0.098	0.200
6	9.4857	0.1274	0.1558	0.1450	-0.0898	-0.0119	0.0001	-0.0043	0.0001	0.0009	0.0008	0.0008	1.5419	-0.5457	-0.9081	0.89	0.283	0.453
7	9.4980	0.1519	0.1805	0.1693	-0.1103	-0.0117	0.0001	-0.0036	0.0000	0.0010	0.0010	0.0009	1.5017	-0.5061	-0.9157	0.89	0.168	0.312
8	9.5289	0.1385	0.1627	0.1523	-0.1148	-0.0122	0.0001	-0.0032	0.0000	0.0011	0.0010	0.0009	1.4887	-0.4937	-0.9198	0.89	0.122	0.236
9	9.5487	0.1224	0.1418	0.1319	-0.1121	-0.0128	0.0002	-0.0028	0.0000	0.0012	0.0011	0.0010	1.4871	-0.4923	-0.9201	0.89	0.097	0.197
10	9.5559	0.1196	0.1364	0.1259	-0.1054	-0.0135	0.0002	-0.0021	0.0000	0.0012	0.0012	0.0010	1.4970	-0.5024	-0.9253	0.90	0.084	0.170
11	9.5487	0.1217	0.1371	0.1251	-0.0923	-0.0142	0.0002	-0.0016	0.0000	0.0013	0.0013	0.0010	1.4879	-0.4940	-0.9235	0.90	0.021	0.045
12	9.5342	0.1283	0.1429	0.1289	-0.0785	-0.0147	0.0002	-0.0012	0.0000	0.0014	0.0014	0.0011	1.4821	-0.4888	-0.9233	0.90	0.022	0.040
13	9.5146	0.1453	0.1604	0.1448	-0.0623	-0.0151	0.0002	-0.0010	0.0000	0.0014	0.0014	0.0011	1.4649	-0.4728	-0.9198	0.90	0.019	0.023
14	9.4987	0.1484	0.1640	0.1470	-0.0522	-0.0153	0.0002	-0.0006	0.0000	0.0015	0.0014	0.0011	1.4545	-0.4635	-0.9131	0.90	0.014	0.009
15	9.4848	0.1473	0.1626	0.1431	-0.0447	-0.0153	0.0002	-0.0002	0.0000	0.0015	0.0014	0.0010	1.4403	-0.4506	-0.8994	0.90	0.005	0.001
16	9.4905	0.1414	0.1563	0.1348	-0.0357	-0.0152	0.0002	0.0004	0.0000	0.0016	0.0014	0.0010	1.4184	-0.4283	-0.8698	0.88	0.002	0.001
17	9.5154	0.1347	0.1478	0.1238	-0.0237	-0.0154	0.0002	0.0008	0.0000	0.0017	0.0013	0.0010	1.4033	-0.4117	-0.8466	0.87	0.000	0.000
18	9.5295	0.1260	0.1390	0.1120	-0.0162	-0.0154	0.0002	0.0012	0.0000	0.0017	0.0014	0.0010	1.4215	-0.4291	-0.8530	0.86	0.000	0.000
19	9.5357	0.1234	0.1351	0.1027	-0.0073	-0.0150	0.0002	0.0013	0.0000	0.0015	0.0012	0.0009	1.4376	-0.4447	-0.8704	0.85	0.000	0.000
20	9.5398	0.1169	0.1283	0.0917	-0.0040	-0.0143	0.0002	0.0013	0.0000	0.0012	0.0010	0.0009	1.4439	-0.4533	-0.8821	0.85	0.000	0.000
21	9.5177	0.1023	0.1141	0.0849	-0.0031	-0.0138	0.0002	0.0012	0.0000	0.0013	0.0011	0.0009	1.4321	-0.4443	-0.8758	0.85	0.000	0.001
22	9.4650	0.0766	0.0891	0.0712	-0.0069	-0.0137	0.0002	0.0012	0.0000	0.0013	0.0012	0.0010	1.4185	-0.4306	-0.8701	0.85	0.000	0.000
23	9.5919	-0.0668	0.0030	0.0050	-0.0605	-0.0108	0.0001	-0.0097	0.0001	0.0007	0.0009	0.0010	1.4730	-0.4790	-0.8832	0.89	0.029	0.068
24	9.5470	-0.0502	0.0065	0.0079	-0.0509	-0.0116	0.0001	-0.0092	0.0001	0.0007	0.0009	0.0010	1.5039	-0.5091	-0.8909	0.89	0.101	0.198

Notes: Estimation sample includes the period 05/04/2004 – 04/30/2008, *n*=1458. No highlight indicates parameter is statistically significant beyond the 1% level. Pink indicates parameter is statistically significant beyond the 5% level. Green indicates parameter is statistically significant beyond the 10% level. Blue indicates parameter is not statistically significant at conventional levels of significance.

[a]Hours are based on Central Standard Time (CST), e.g., 2=2AM-3AM CST

[b]Q_{LB} is the Ljung-Box test for autocorrelation, and the p-values for the Ljung-Box test for autocorrelation up to order 5 and 6 are reported. However it should be noted that Dezhbakhsh (1990) finds the application of the Ljung-Box test to be inadequate when applied to linear models with lagged dependent variables and exogenous regressors (e.g., ARMAX models). However, since these statistics are still reported in the recent load forecasting literature, we provide them here for readers. The Ljung-Box Q-statistic at lag k is a test statistic for the null hypothesis that there is no autocorrelation up to order k (EViews 2008; Ljung and Box 1978). $Q_{LB} = T(T+2)\sum_{j=1}^{k}\frac{\tau_j^2}{T-J}$ where τ_j is the j-th autocorrelation and T is the number of observations (EViews 2008).

TABLE 5. ONE-STEP[A] AHEAD OUT-OF-SAMPLE FORECASTING RESULTS FOR EACH HOUR FOR THE FULL-WEEK MODEL FOR MAY 1, 2008 THROUGH SEPTEMBER 30, 2009

Model	$Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \beta_f F + \beta_{su} SU + \delta_{tmp} TMP_t + \delta_{tmp2} {TMP_t}^2 + \delta_{ltmp} TMP_{t-1} + \delta_{ltmp2} {TMP_{t-1}}^2 + \delta_{m6tmp} JUN * TMP_t + \delta_{m7tmp} JUL * TMP_t + \delta_{m8tmp} AUG * TMP_t + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1} \varepsilon_{t-1} + \varepsilon_t$ (5)		
Hour[b]	***Root Mean Squared Error[c]***	***Mean Absolute Error[d]***	***Mean Absolute Percent Error[e]***
1	447.62	313.61	3.18
2	413.46	289.68	3.03
3	393.99	276.50	2.93
4	387.01	270.29	2.84
5	423.12	296.07	3.00
6	521.11	352.11	3.36
7	591.57	389.39	3.53
8	610.56	402.92	3.50
9	622.70	419.20	3.50
10	649.01	444.76	3.59
11	674.78	474.24	3.75
12	706.50	500.24	3.91
13	752.08	530.99	4.13
14	785.74	556.38	4.32
15	826.95	592.92	4.58
16	863.19	624.08	4.78
17	878.04	639.18	4.86
18	857.29	621.46	4.77
19	820.15	592.54	4.57
20	785.92	559.92	4.30
21	746.32	529.89	4.18
22	684.32	485.31	4.09
23	529.68	371.97	3.40
24	486.04	339.72	3.31

Notes: The adjusted estimation sample includes the period 05/04/2004 – 04/30/2008, *n*=1458; the forecast sample includes the period 05/01/2008 – 09/30/2009, *n*=518. EViews 6 software package was used for the estimation and forecasts. One-step ahead static forecasts were employed to forecast the level of the series. The static method was used be it calculates a sequence of one-step ahead forecasts, using the actual, rather than forecasted values for lagged dependent variables.

[a]One-step ahead refers to the sectional data, which is daily. Since the primary data are hourly, one must interpret it as 24-steps ahead, so that one-daily-step ahead actually corresponds to 24-hourly-steps ahead.

[b]Hours are based on Central Standard Time (CST), e.g., 2=2AM-3AM CST

[c]Root Mean Squared Error (RMSE) = $\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}$ where y_t is the actual value in period *t* and $\hat{y}_t$ is the forecasted value in period *t*, and the forecast sample is *j=T+1, T+2, …, T+h* (EViews 2008).

[d]Mean Absolute Error (MAE) = $\sum_{t=T+1}^{T+h} |\hat{y}_t - y_t| / h$ where y_t is the actual value in period *t* and $\hat{y}_t$ is the forecasted value in period *t*, and the forecast sample is *j=T+1, T+2, …, T+h* (EViews 2008).

[e]Mean Absolute Percentage Error (MAPE) = $100 \sum_{t=T+1}^{T+h} \left| \frac{\hat{y}_t - y_t}{y_t} \right| / h$ where y_t is the actual value in period *t* and $\hat{y}_t$ is the forecasted value in period *t*, and the forecast sample is *j=T+1,…, T+h* (EViews 2008).

TABLE 6. PARAMETER ESTIMATES AND DIAGNOSTIC STATISTICS FOR THE WEEKDAY MODEL

$$Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \delta_{wwp} WWP_t + \delta_{wwp2} WWP_t^{\,2} + \lambda_{hol} H + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1} \varepsilon_{t-1} + \varepsilon_t \qquad (6)$$

Hour[a]	$\hat{\alpha}_0$	$\hat{\beta}_m$	$\hat{\beta}_{twt}$	$\hat{\lambda}_{hol}$	$\hat{\delta}_{wwp}$	$\hat{\delta}_{wwp2}$	$\hat{\varphi}_{ar1}$	$\hat{\varphi}_{ar2}$	$\hat{\theta}_{ma1}$	*Adj. R^2*	*$Q_{LB}(5)$*[b]	*$Q_{LB}(6)$*	*$Q_{LB}(7)$*
1	9.4569	-0.0519	0.0007	0.0038	-0.0157	0.0002	1.3739	-0.3929	-0.8320	0.83	0.066	0.105	0.178
	(0.0258)	(0.0044)	(0.0037)	(0.0090)	(0.0011)	(0.0000)	(0.0618)	(0.0574)	(0.0395)				
2	9.4404	-0.0429	0.0014	-0.0032	-0.0156	0.0002	1.4020	-0.4196	-0.8354	0.83	0.064	0.098	0.160
	(0.0254)	(0.0042)	(0.0035)	(0.0084)	(0.0011)	(0.0000)	(0.0616)	(0.0575)	(0.0399)				
3	9.4365	-0.0346	0.0025	-0.0098	-0.0154	0.0002	1.4350	-0.4512	-0.8410	0.84	0.065	0.099	0.167
	(0.0251)	(0.0039)	(0.0032)	(0.0083)	(0.0011)	(0.0000)	(0.0609)	(0.0569)	(0.0401)				
4	9.4509	-0.0251	0.0040	-0.0264	-0.0148	0.0002	1.4650	-0.4802	-0.8448	0.84	0.048	0.077	0.134
	(0.0246)	(0.0037)	(0.0030)	(0.0080)	(0.0010)	(0.0000)	(0.0601)	(0.0562)	(0.0407)				
5	9.4949	-0.0148	0.0071	-0.0622	-0.0137	0.0001	1.4999	-0.5129	-0.8602	0.83	0.065	0.096	0.164
	(0.0240)	(0.0036)	(0.0029)	(0.0098)	(0.0010)	(0.0000)	(0.0563)	(0.0527)	(0.0388)				
6	9.5693	-0.0056	0.0112	-0.1184	-0.0126	0.0001	1.5014	-0.5117	-0.8802	0.81	0.384	0.473	0.642
	(0.0255)	(0.0039)	(0.0030)	(0.0138)	(0.0010)	(0.0000)	(0.0465)	(0.0438)	(0.0318)				
7	9.6125	-0.0008	0.0117	-0.1537	-0.0121	0.0001	1.4748	-0.4872	-0.8820	0.79	0.926	0.763	0.796
	(0.0245)	(0.0041)	(0.0030)	(0.0172)	(0.0009)	(0.0000)	(0.0465)	(0.0432)	(0.0314)				
8	9.6365	0.0031	0.0111	-0.1539	-0.0126	0.0001	1.4263	-0.4444	-0.8628	0.81	0.984	0.793	0.703
	(0.0230)	(0.0039)	(0.0028)	(0.0184)	(0.0009)	(0.0000)	(0.0513)	(0.0467)	(0.0356)				
9	9.6474	0.0064	0.0106	-0.1419	-0.0133	0.0002	1.3915	-0.4118	-0.8490	0.84	0.674	0.580	0.495
	(0.0236)	(0.0038)	(0.0027)	(0.0182)	(0.0010)	(0.0000)	(0.0522)	(0.0475)	(0.0371)				
10	9.6561	0.0091	0.0113	-0.1319	-0.0139	0.0002	1.3791	-0.4003	-0.8478	0.86	0.182	0.213	0.239
	(0.0246)	(0.0039)	(0.0028)	(0.0176)	(0.0010)	(0.0000)	(0.0516)	(0.0467)	(0.0379)				
11	9.6569	0.0114	0.0128	-0.1233	-0.0146	0.0002	1.3646	-0.3862	-0.8485	0.87	0.155	0.292	0.425
	(0.0259)	(0.0042)	(0.0030)	(0.0172)	(0.0011)	(0.0000)	(0.0506)	(0.0456)	(0.0373)				
12	9.6529	0.0143	0.0148	-0.1203	-0.0151	0.0002	1.3514	-0.3737	-0.8495	0.88	0.254	0.368	0.527
	(0.0269)	(0.0044)	(0.0032)	(0.0173)	(0.0012)	(0.0000)	(0.0503)	(0.0454)	(0.0366)				
13	9.6523	0.0163	0.0163	-0.1263	-0.0155	0.0002	1.3281	-0.3516	-0.8473	0.88	0.522	0.562	0.718
	(0.0279)	(0.0047)	(0.0035)	(0.0182)	(0.0012)	(0.0000)	(0.0511)	(0.0460)	(0.0363)				
14	9.6453	0.0177	0.0177	-0.1288	-0.0158	0.0002	1.3132	-0.3374	-0.8440	0.88	0.595	0.650	0.791
	(0.0292)	(0.0050)	(0.0037)	(0.0184)	(0.0013)	(0.0000)	(0.0515)	(0.0464)	(0.0362)				
15	9.6346	0.0209	0.0202	-0.1296	-0.0157	0.0002	1.2951	-0.3200	-0.8330	0.87	0.661	0.693	0.817
	(0.0304)	(0.0053)	(0.0038)	(0.0186)	(0.0013)	(0.0000)	(0.0524)	(0.0474)	(0.0365)				
16	9.6419	0.0232	0.0221	-0.1266	-0.0155	0.0002	1.2840	-0.3046	-0.8105	0.86	0.634	0.534	0.701
	(0.0342)	(0.0054)	(0.0039)	(0.0182)	(0.0014)	(0.0000)	(0.0519)	(0.0484)	(0.0352)				
17	9.6641	0.0267	0.0244	-0.1195	-0.0157	0.0002	1.2792	-0.2955	-0.7905	0.84	0.467	0.250	0.359
	(0.0391)	(0.0054)	(0.0039)	(0.0173)	(0.0014)	(0.0000)	(0.0535)	(0.0508)	(0.0355)				
18	9.6721	0.0292	0.0272	-0.1143	-0.0157	0.0002	1.2939	-0.3094	-0.7974	0.82	0.590	0.333	0.464
	(0.0382)	(0.0054)	(0.0038)	(0.0169)	(0.0014)	(0.0000)	(0.0550)	(0.0522)	(0.0378)				
19	9.6685	0.0351	0.0325	-0.1078	-0.0153	0.0002	1.3108	-0.3250	-0.8283	0.80	0.588	0.497	0.647
	(0.0363)	(0.0052)	(0.0037)	(0.0159)	(0.0014)	(0.0000)	(0.0543)	(0.0518)	(0.0341)				
20	9.6602	0.0392	0.0367	-0.1030	-0.0148	0.0002	1.3114	-0.3293	-0.8423	0.81	0.610	0.602	0.682
	(0.0321)	(0.0049)	(0.0034)	(0.0150)	(0.0014)	(0.0000)	(0.0562)	(0.0529)	(0.0345)				
21	9.6323	0.0299	0.0290	-0.0934	-0.0145	0.0002	1.2839	-0.3078	-0.8223	0.82	0.842	0.817	0.812
	(0.0305)	(0.0049)	(0.0034)	(0.0138)	(0.0014)	(0.0000)	(0.0611)	(0.0564)	(0.0386)				
22	9.5690	0.0157	0.0174	-0.0797	-0.0144	0.0002	1.2765	-0.3003	-0.8128	0.82	0.941	0.857	0.900
	(0.0306)	(0.0047)	(0.0034)	(0.0122)	(0.0014)	(0.0000)	(0.0632)	(0.0585)	(0.0390)				
23	9.5306	-0.0775	-0.0005	0.0128	-0.0152	0.0002	1.3203	-0.3428	-0.8236	0.82	0.098	0.159	0.242
	(0.0263)	(0.0048)	(0.0042)	(0.0093)	(0.0011)	(0.0000)	(0.0625)	(0.0573)	(0.0402)				
24	9.4857	-0.0633	-0.0001	0.0101	-0.0156	0.0002	1.3485	-0.3690	-0.8284	0.82	0.048	0.079	0.134
	(0.0260)	(0.0046)	(0.0039)	(0.0091)	(0.0011)	(0.0000)	(0.0623)	(0.0575)	(0.0397)				

Notes: Newey-West heteroskedasticity and autocorrelation robust standard errors are in parentheses (Newey and West 1994). Estimation sample includes the period 05/05/2004 – 04/30/2008, *n*=1041. No highlight indicates parameter is statistically significant beyond the 1% level. Pink indicates parameter is statistically significant beyond the 5% level. Green indicates parameter is statistically significant beyond the 10% level. Blue indicates parameter is not statistically significant at conventional levels of significance. Yellow indicates residuals are not white noise.

[a]Hours are based on Central Standard Time (CST), e.g., 2=2AM-3AM CST

[b]Q_{LB} is the Ljung-Box test for autocorrelation, and the p-values for the Ljung-Box test for autocorrelation up to order 5, 6, and 7 are reported. However it should be noted that Dezhbakhsh (1990) finds the application of the Ljung-Box test to be inadequate when applied to linear models with lagged dependent variables and exogenous regressors (e.g., ARMAX models). However, since these statistics are still reported in the recent load forecasting literature, we provide them here for readers. The Ljung-Box Q-statistic at lag *k* is a test statistic for the null hypothesis that there is no autocorrelation up to order *k* (EViews 2008; Ljung and Box 1978). $Q_{LB} = T(T+2)\sum_{j=1}^{k}\frac{\tau_j^2}{T-j}$ where τ_j is the j-th autocorrelation and T is the number of observations (EViews 2008).

TABLE 7. PARAMETER ESTIMATES AND DIAGNOSTIC STATISTICS FOR THE WEEKDAY AND WEEKNIGHT MODELS

$$Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \delta_{wwp} WWP_t + \delta_{wwp2} WWP_t^2 + \lambda_{hol} H + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1} \varepsilon_{t-1} + \varepsilon_t \quad (6)$$

$$Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \beta_f F + \delta_{tmp} TMP_t + \delta_{tmp2} TMP_t^2 + \delta_{ltmp} TMP_{t-1} + \delta_{ltmp2} TMP_{t-1}^2 + \delta_{m6tmp} JUN * TMP_t + \delta_{m7tmp} JUL * TMP_t + \delta_{m8tmp} AUG * TMP_t + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1} \varepsilon_{t-1} + \varepsilon_t$$

Hour[a]	$\hat{\alpha}_0$	$\hat{\beta}_m$	$\hat{\beta}_{twt}$	$\hat{\beta}_f$	$\hat{\delta}_{wwp}$	$\hat{\delta}_{wwp2}$	$\hat{\delta}_{tmp}$	$\hat{\delta}_{tmp2}$	$\hat{\delta}_{ltmp}$	$\hat{\delta}_{ltmp2}$	$\hat{\delta}_{m6tmp}$	$\hat{\delta}_{m7tmp}$	$\hat{\delta}_{m8tmp}$	$\hat{\lambda}_{hol}$	$\hat{\varphi}_{ar1}$	$\hat{\varphi}_{ar2}$	$\hat{\theta}_{ma1}$	Adj. R^2	$Q_{LB}(5)$[b]	$Q_{LB}(6)$	$Q_{LB}(7)$
1	9.47	-	0.05	0.05	-	-	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	-	1.40	-0.41	-0.85	0.87	0.20	0.35	0.51
	(0.04)		(0.00)	(0.00)			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.06)	(0.05)	(0.03)				
2	9.46	-	0.04	0.04	-	-	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	-	1.43	-0.44	-0.86	0.87	0.26	0.44	0.61
	(0.04)		(0.00)	(0.00)			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.05)	(0.05)	(0.03)				
3	9.46	-	0.04	0.03	-	-	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	-	1.47	-0.48	-0.87	0.87	0.24	0.42	0.58
	(0.04)		(0.00)	(0.00)			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.05)	(0.05)	(0.03)				
4	9.47	-	0.03	0.03	-	-	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	-	1.51	-0.52	-0.88	0.86	0.21	0.37	0.52
	(0.04)		(0.00)	(0.00)			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.05)	(0.04)	(0.03)				
5	9.52	-	0.03	0.02	-	-	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	-	1.53	-0.53	-0.89	0.84	0.18	0.31	0.44
	(0.04)		(0.00)	(0.00)			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.04)	(0.04)	(0.03)				
6	9.57	-0.01	0.01	-	-0.01	0.00	-	-	-	-	-	-	-	-0.12	1.50	-0.51	-0.88	0.81	0.38	0.47	0.64
	(0.03)	(0.00)	(0.00)		(0.00)	(0.00)								(0.01)	(0.05)	(0.04)	(0.03)				
7	9.61	0.00	0.01	-	-0.01	0.00	-	-	-	-	-	-	-	-0.15	1.47	-0.49	-0.88	0.79	0.93	0.76	0.80
	(0.02)	(0.00)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.04)	(0.03)				
8	9.64	0.00	0.01	-	-0.01	0.00	-	-	-	-	-	-	-	-0.15	1.43	-0.44	-0.86	0.81	0.98	0.79	0.70
	(0.02)	(0.00)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
9	9.65	0.01	0.01	-	-0.01	0.00	-	-	-	-	-	-	-	-0.14	1.39	-0.41	-0.85	0.84	0.67	0.58	0.50
	(0.02)	(0.00)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
10	9.66	0.01	0.01	-	-0.01	0.00	-	-	-	-	-	-	-	-0.13	1.38	-0.40	-0.85	0.86	0.18	0.21	0.24
	(0.02)	(0.00)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
11	9.66	0.01	0.01	-	-0.01	0.00	-	-	-	-	-	-	-	-0.12	1.36	-0.39	-0.85	0.87	0.16	0.29	0.43
	(0.03)	(0.00)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
12	9.65	0.01	0.01	-	-0.02	0.00	-	-	-	-	-	-	-	-0.12	1.35	-0.37	-0.85	0.88	0.25	0.37	0.53
	(0.03)	(0.00)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
13	9.65	0.02	0.02	-	-0.02	0.00	-	-	-	-	-	-	-	-0.13	1.33	-0.35	-0.85	0.88	0.52	0.56	0.72
	(0.03)	(0.00)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
14	9.65	0.02	0.02	-	-0.02	0.00	-	-	-	-	-	-	-	-0.13	1.31	-0.34	-0.84	0.88	0.60	0.65	0.79
	(0.03)	(0.00)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
15	9.63	0.02	0.02	-	-0.02	0.00	-	-	-	-	-	-	-	-0.13	1.30	-0.32	-0.83	0.87	0.66	0.69	0.82
	(0.03)	(0.01)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
16	9.64	0.02	0.02	-	-0.02	0.00	-	-	-	-	-	-	-	-0.13	1.28	-0.30	-0.81	0.86	0.63	0.53	0.70
	(0.03)	(0.01)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
17	9.66	0.03	0.02	-	-0.02	0.00	-	-	-	-	-	-	-	-0.12	1.28	-0.30	-0.79	0.84	0.47	0.25	0.36
	(0.04)	(0.01)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.04)				
18	9.67	0.03	0.03	-	-0.02	0.00	-	-	-	-	-	-	-	-0.11	1.29	-0.31	-0.80	0.82	0.59	0.33	0.46
	(0.04)	(0.01)	(0.00)		(0.00)	(0.00)								(0.02)	(0.06)	(0.05)	(0.04)				
19	9.67	0.04	0.03	-	-0.02	0.00	-	-	-	-	-	-	-	-0.11	1.31	-0.33	-0.83	0.80	0.59	0.50	0.65
	(0.04)	(0.01)	(0.00)		(0.00)	(0.00)								(0.02)	(0.05)	(0.05)	(0.03)				
20	9.66	0.04	0.04	-	-0.01	0.00	-	-	-	-	-	-	-	-0.10	1.31	-0.33	-0.84	0.81	0.61	0.60	0.68
	(0.03)	(0.00)	(0.00)		(0.00)	(0.00)								(0.02)	(0.06)	(0.05)	(0.03)				
21	9.63	0.03	0.03	-	-0.01	0.00	-	-	-	-	-	-	-	-0.09	1.28	-0.31	-0.82	0.82	0.84	0.82	0.81
	(0.03)	(0.00)	(0.00)		(0.00)	(0.00)								(0.01)	(0.06)	(0.06)	(0.04)				
22	9.57	0.02	0.02	-	-0.01	0.00	-	-	-	-	-	-	-	-0.08	1.28	-0.30	-0.81	0.82	0.94	0.86	0.90
	(0.03)	(0.00)	(0.00)		(0.00)	(0.00)								(0.01)	(0.06)	(0.06)	(0.04)				
23	9.51	-	0.07	0.07	-	-	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	-	1.33	-0.35	-0.84	0.87	0.19	0.33	0.49
	(0.04)		(0.00)	(0.00)			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.06)	(0.06)	(0.03)				
24	9.48	-	0.06	0.06	-	-	-0.01	0.00	-0.01	0.00	0.00	0.00	0.00	-	1.37	-0.38	-0.84	0.87	0.16	0.30	0.45
	(0.04)		(0.00)	(0.00)			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.06)	(0.06)	(0.03)				

Notes: Newey-West heteroskedasticity and autocorrelation robust standard errors are in parentheses (Newey and West 1994). Estimation sample includes the period 05/05/2004 – 04/30/2008, *n*=1041. No highlight indicates parameter is statistically significant beyond the 1% level. Pink indicates parameter is statistically significant beyond the 5% level. Green indicates parameter is statistically significant beyond the 10% level. Blue indicates parameter is not statistically significant at conventional levels.

[a]Hours are based on Central Standard Time (CST), e.g., 2=2AM-3AM CST

[b]Q_{LB} is the Ljung-Box test for autocorrelation, and the p-values for the Ljung-Box test for autocorrelation up to order 5, 6, and 7 are reported. However it should be noted that Dezhbakhsh (1990) finds the application of the Ljung-Box test to be inadequate when applied to linear models with lagged dependent variables and exogenous regressors (e.g., ARMAX models). However, since these statistics are still reported in the recent load forecasting literature, we provide them here for readers. The Ljung-Box Q-statistic at lag *k* is a test statistic for the null hypothesis that there is no autocorrelation up to order *k* (EViews 2008; Ljung and Box 1978). $Q_{LB} = T(T+2)\sum_{j=1}^{k}\frac{\tau_j^2}{T-j}$ where τ_j is the j-th autocorrelation and T is the number of observations (EViews 2008).

TABLE 8. ONE-STEP[A] AHEAD OUT-OF-SAMPLE FORECASTING RESULTS FOR EACH HOUR OF THE WEEKDAY FOR MAY 1, 2008 THROUGH SEPTEMBER 30, 2009

Model	$Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \delta_{wwp} WWP_t + \delta_{wwp2} {WWP_t}^2 + \lambda_{hol} H + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1} \varepsilon_{t-1} + \varepsilon_t$ (6)			$Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \delta_{wwp} WWP_t + \delta_{wwp2} {WWP_t}^2 + \lambda_{hol} H + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1} \varepsilon_{t-1} + \varepsilon_t$ (6) $Y_t = \alpha_0 + \beta_m M + \beta_{twt} TWT + \beta_f F + \delta_{tmp} TMP_t + \delta_{tmp2} {TMP_t}^2 + \delta_{ltmp} TMP_{t-1} + \delta_{ltmp2} {TMP_{t-1}}^2 + \delta_{m6tmp} JUN * TMP_t + \delta_{m7tmp} JUL * TMP_t + \delta_{m8tmp} AUG * TMP_t + \varphi_{ar1} Y_{t-1} + \varphi_{ar2} Y_{t-2} + \theta_{ma1} \varepsilon_{t-1} + \varepsilon_t$ (7)[c]		
Hour[b]	***RMSE***[d]	***MAE***[e]	***MAPE***[f]	***RMSE***	***MAE***	***MAPE***
1	522.84	377.41	3.81	**464.22**	**327.32**	**3.33**
2	477.16	345.18	3.58	**432.07**	**304.85**	**3.19**
3	453.96	328.00	3.44	**418.83**	**294.98**	**3.11**
4	441.78	316.96	3.27	**417.12**	**290.18**	**3.01**
5	**461.01**	334.28	3.28	464.22	**322.74**	**3.19**
6	527.78	379.52	3.42	527.78	379.52	3.42
7	574.86	410.12	3.46	574.86	410.12	3.46
8	590.51	417.20	3.37	590.51	417.20	3.37
9	606.00	429.79	3.34	606.00	429.79	3.34
10	637.61	453.17	3.41	637.61	453.17	3.41
11	672.36	481.94	3.57	672.36	481.94	3.57
12	713.00	514.31	3.76	713.00	514.31	3.76
13	757.39	544.54	3.95	757.39	544.54	3.95
14	791.41	566.67	4.11	791.41	566.67	4.11
15	832.21	596.49	4.33	832.21	596.49	4.33
16	873.28	621.40	4.47	873.28	621.40	4.47
17	888.29	632.85	4.54	888.29	632.85	4.54
18	875.49	623.29	4.54	875.49	623.29	4.54
19	833.53	598.44	4.41	833.53	598.44	4.41
20	794.28	567.26	4.18	794.28	567.26	4.18
21	759.24	540.57	4.11	759.24	540.57	4.11
22	692.53	495.09	4.05	692.53	495.09	4.05
23	643.76	462.34	4.20	**546.48**	**389.21**	**3.57**
24	578.96	413.56	4.00	**502.09**	**353.94**	**3.46**

Notes: The superior forecasting model (e.g., lowest MAPE) is indicated by the bold type. The adjusted estimation sample includes the period 05/05/2004 – 04/30/2008, *n*=1041; the forecast sample includes the period 05/01/2008 – 09/30/2009, *n*=370. EViews 6 software package was used for the estimation and forecasts. One-step ahead static forecasts were employed to forecast the level of the series. The static method was used be it calculates a sequence of one-step ahead forecasts, using the actual, rather than forecasted values for lagged dependent variables.

[a]One-step ahead refers to the sectional data, which is daily. Since the primary data are hourly, one must interpret it as 24-steps ahead, so that one-daily-step ahead actually corresponds to 24-hourly-steps ahead. Since we have removed weekends from the models estimated here, the one-step ahead forecasts made on a Friday will actually be forecasting Monday's load, thus in some sense three-days ahead (72-hours ahead) forecasts are being made.

[b] Hours are based on Central Standard Time (CST), e.g., 2=2AM-3AM CST

[c]The night load model is used to estimate hours 1-5, 23, and 24.

[d]Root Mean Squared Error (RMSE) = $\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}$ where y_t is the actual value in period *t* and $\hat{y}_t$ is the forecasted value in period *t*, and the forecast sample is *j=T+1, T+2, …, T+h* (EViews 2008).

[e]Mean Absolute Error (MAE) = $\sum_{t=T+1}^{T+h} |\hat{y}_t - y_t| / h$ where y_t is the actual value in period *t* and $\hat{y}_t$ is the forecasted value in period *t*, and the forecast sample is *j=T+1, T+2, …, T+h* (EViews 2008).

[f]Mean Absolute Percentage Error (MAPE) = $100 \sum_{t=T+1}^{T+h} \left| \frac{\hat{y}_t - y_t}{y_t} \right| / h$ where y_t is the actual value in period *t* and $\hat{y}_t$ is the forecasted value in period *t*, and the forecast sample is *j=T+1, T+2, …, T+h* (EViews 2008).

VIII. FIGURES

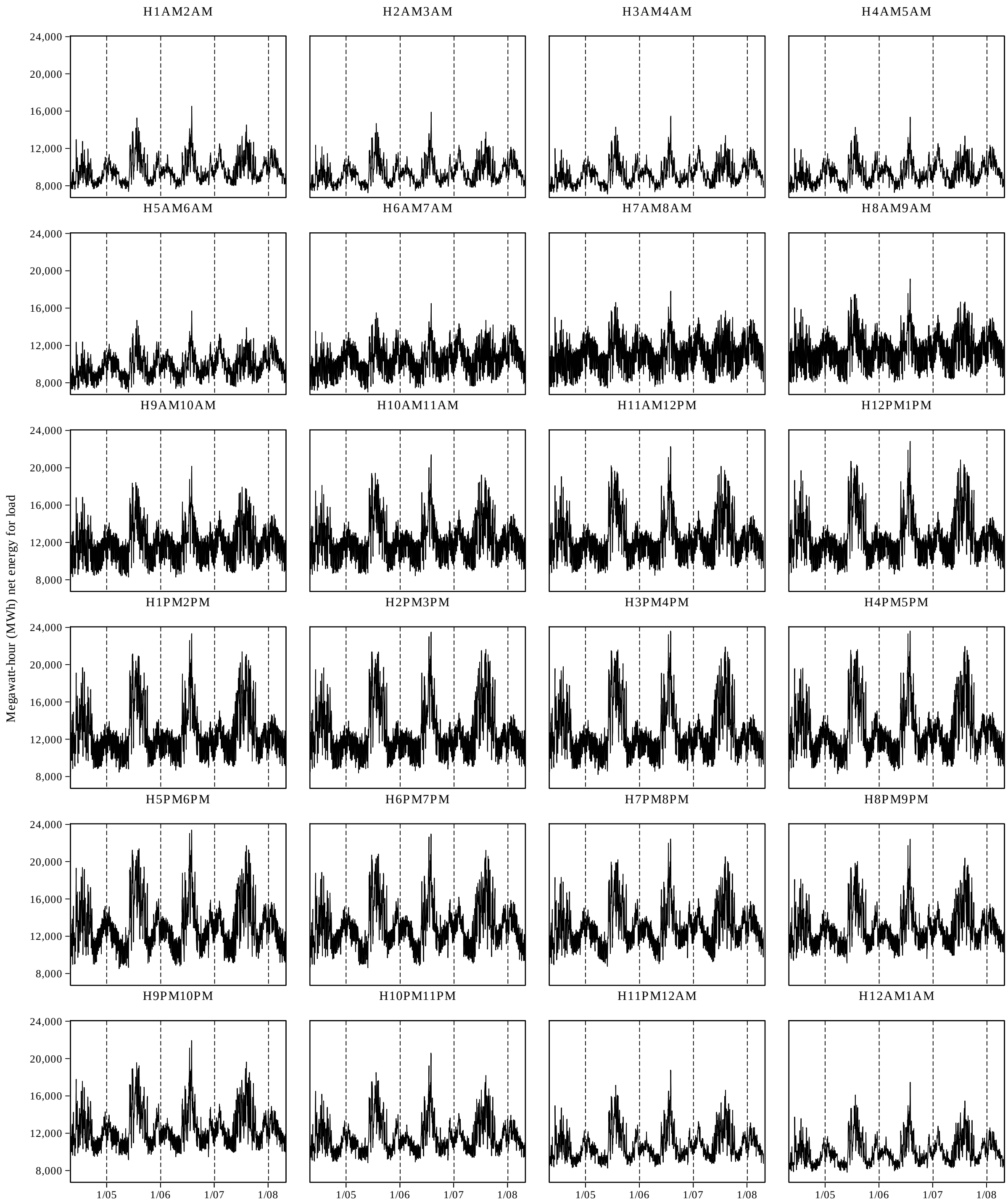

Figure 1. Load for each hour from May 1, 2004 through April 30, 2008 (in-sample period)

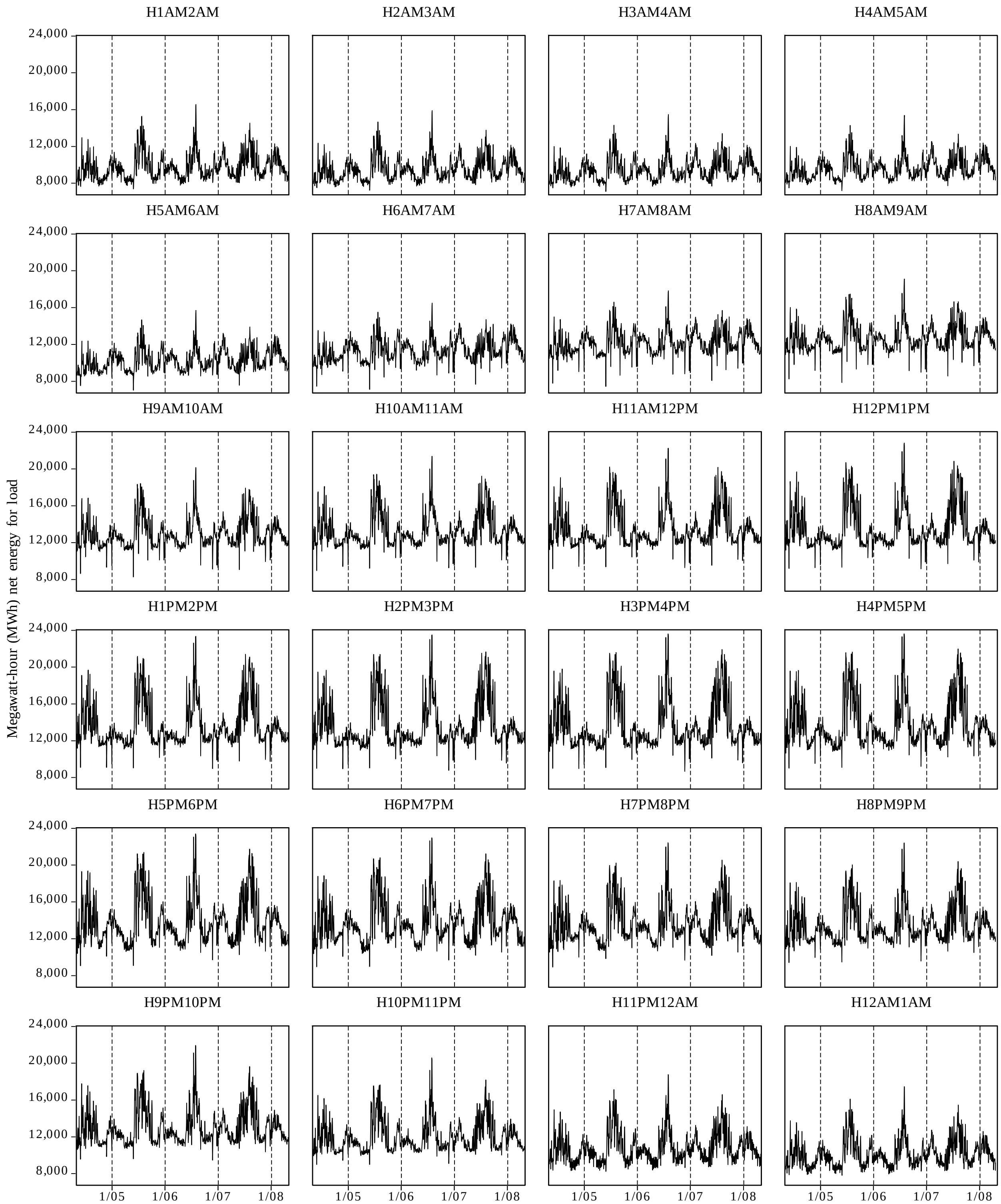

Figure 2. Load for each weekday hour from May 3, 2004 through April 30, 2008 (in-sample period)

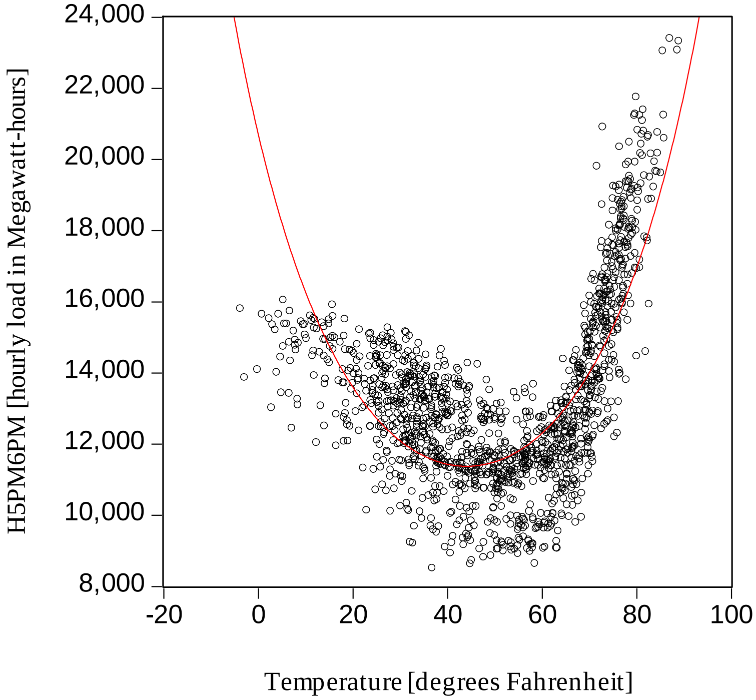

Figure 3a. Commonwealth Edison 5PM-6PM daily system-wide load versus average daily air temperature from May 1, 2004 through April 30, 2008 using full week data

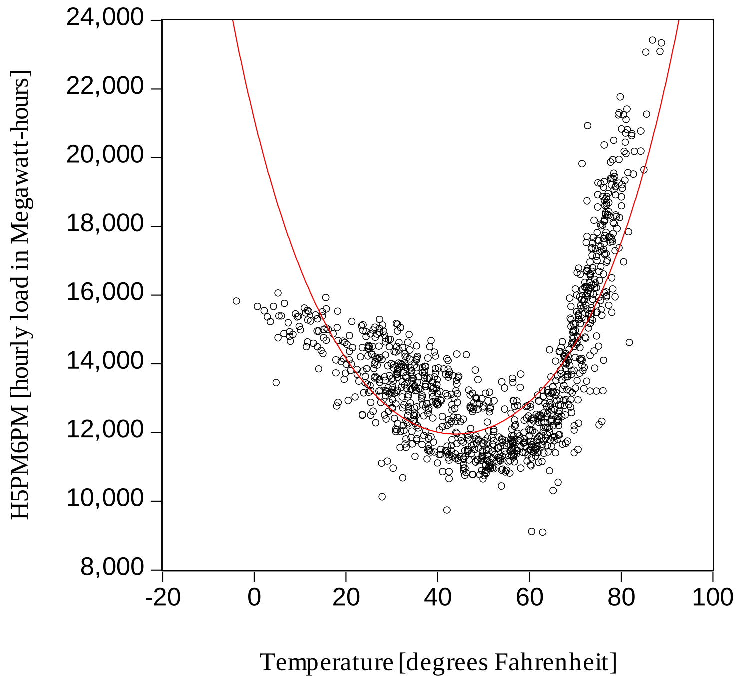

Figure 3b. Commonwealth Edison 5PM-6PM daily system-wide load versus average daily air temperature from May 3, 2004 through April 30, 2008 using weekday data

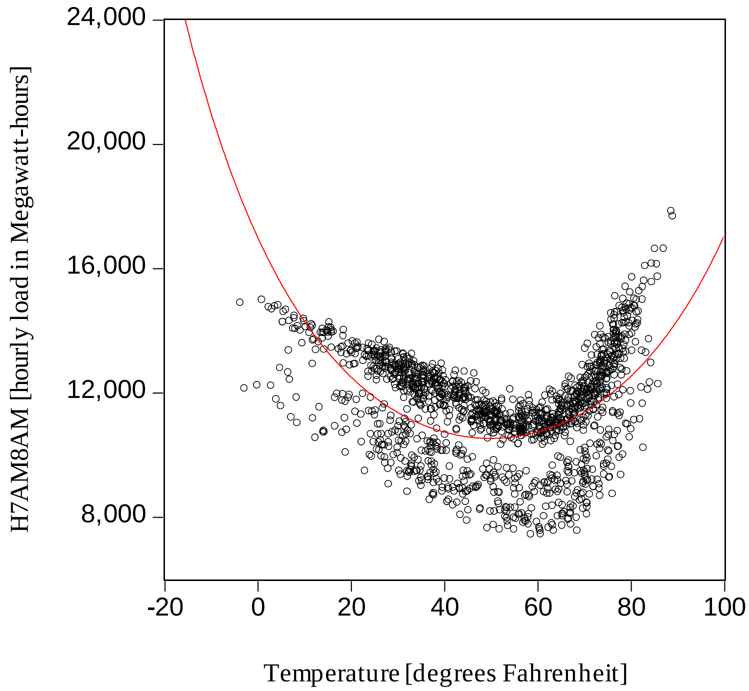

Figure 3c. Commonwealth Edison 7AM-8AM daily system-wide load versus average daily air temperature from May 1, 2004 through April 30, 2008 using full week data

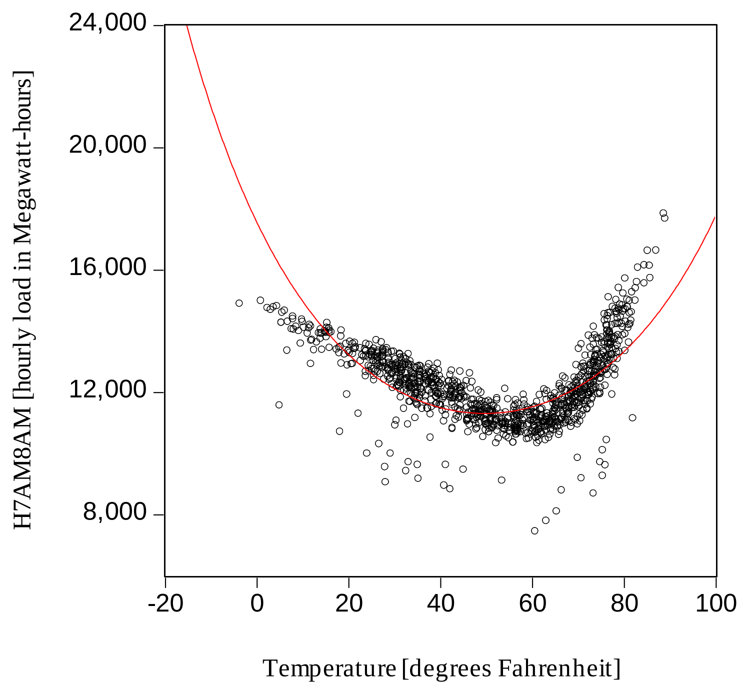

Figure 3d. Commonwealth Edison 7AM-8AM daily system-wide load versus average daily air temperature from May 3, 2004 through April 30, 2008 using weekday data

REFERENCES

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19:716-23.

Akaike, H. 1987. Factor Analysis and AIC. *Psychometrika* 52 (3): 317–32.

Alfares, H. K., and M. Nazeeruddin. 2002. Electric load forecasting: Literature survey and classification of methods. *International Journal of Systems Science* 33 (1): 23-34.

Amaral, L. F., R. C. Souza, and M. Stevenson. 2008. A smooth transition periodic autoregressive (STPAR) model for short-term load forecasting. *International Journal of Forecasting* 24:603-615.

Amjady, N., and A. Daraeepour. 2009. Mixed price and load forecasting of electricity markets by a new iterative prediction method. *Electric Power Systems Research* 79 (9): 1329-36.

Associated Press. 2007. The associated press stylebook: Stylebook and briefing on media law. 42nd ed., ed. N. Goldstein. New York: Basic Books.

Box, G. E. P., and G. M. Jenkins. 1970. *Time-series analysis: Forecasting and control*. Rev. ed. San Francisco: Holden-Day.

Bruhns, A., G. Deurveilher, and J. Roy. 2005. A nonlinear regression model for mid term load forecasting and improvements in seasonality. 15th Power Systems Computation Conference.

Bunn, D. W., and E. D. Farmer, eds. 1985a. Economic and operational context of electric load prediction. In *Comparative models for electrical load forecasting*, 3-11. New York: John Wiley and Sons.

Bunn, D. W., and E. D. Farmer, eds. 1985b. Review of short-term forecasting methods in the electric power industry. In *Comparative models for electrical load forecasting*, 13-30. New York: John Wiley and Sons.

Carpinteiro, O., A. Reis, and A. Silva. 2004. A hierarchical neural model in short-term load forecasting. *Applied Soft Computing* 4:405-12.

Chatfield, C. 2001. *Time-series forecasting*. Boca Raton, FL: Chapman and Hall/CRC Press.

Chen, T. 1997. Long-term peak electricity load forecasting in Taiwan: A cointegration analysis. *Pacific and Asian Journal of Energy* 7 (1): 63-73.

Chen, K., and L. Wang, eds. 2007. *Trends in neural computation.* Vol. 35 *of Studies in computational intelligence.* New York: Springer.

Cottet, R., and M. Smith. 2003. Bayesian modeling and forecasting of intraday electricity load. *Journal of the American Statistical Association* 98 (464): 839-49.

Darbellay, G. A., and M. Slama. 2000. Forecasting the short-term demand for electricity: Do neural networks stand a better chance? *International Journal of Forecasting* 16 (1): 71-83.

Dezhbakhsh, H. 1990. The inappropriate use of serial correlation tests in dynamic linear models. *Review of Economics and Statistics* 72 (1): 126-32.

Dickey, D. A., and W. A. Fuller. 1979. Distribution of estimator for autoregressive time series with a unit root. *Journal of American Statistical Association* 74 (366): 427-31.

Dickey, D. A., and W. A. Fuller. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49 (4): 1057-72.

Diebold, F. X. 2007. *Elements of forecasting*. 4th ed. Mason, OH: Thomson South-Western.

Energy Information Administration (EIA). 2009. State energy profiles: Illinois. http://tonto.eia.doe.gov/state/state_energy_profiles.cfm?sid=IL (accessed November 12, 2009).

Engle, R. F., C. Mustafa, and J. Rice. 1992. Modelling peak electricity demand. *Journal of Forecasting* 11:241-51.

EViews. 2008. Quantitative Micro Software. Version 6. CD-ROM.

Feinberg, E. A., and D. Genethliou. 2005. Load forecasting. In *Applied mathematics for power systems: Optimization, control, and computational intelligence,* eds. J. H. Chow, F. F. Wu, and J. A. Momoh, 269-85. New York: Springer.

Fidalgo, J., and M. A. Matos. 2007. Forecasting Portugal global load with artificial neural networks. In *Artificial neural networks, ICANN 2007*, eds. J. Marques de Sá et al., 4669 (2): 728-37. New York: Springer.

Fiebig, D. G., R. Bartels, and D. J. Aigner. 1991. A random coefficient approach to the estimation of residential end-use load profiles. *Journal of Econometrics* 50:297-327.

Granger, C. W. J. 1980. *Forecasting in business and economics*. New York: Academic Press.

Gupta, P. C. 1985. Adaptive short-term forecasting of hourly loads using weather information. In *Comparative models for electrical load forecasting,* eds. D. W. Bunn and E. D. Farmer, 43-56. New York: John Wiley and Sons.

Hahn, J., S. Meyer-Nieberg, and S. Pickl. 2009. Electric load forecasting methods: Tools for decision making. *European Journal of Operational Research* 199:902-7.

Halvorsen, R., and R. Palmquist. 1980. The interpretation of dummy variables in semi-logarithmic equations. *American Economic Review* 70 (3): 474-5.

Heinemann, G., D. Nordman, and E. Plant. 1966. The relationship between summer weather and summer loads: A regression analysis. *IEEE Transactions on Power Apparatus and Systems PAS* 85:1144-54.

Hippert, H. S., D. W. Bunn, and R. C. Souza. 2005. Large neural networks for electricity load forecasting: Are they overfitted? *International Journal of Forecasting* 21 (3): 425-34.

Hippert, H. S., C. E. Pedreira, and R. C. Souza. 2001. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on Power Systems* 16 (1): 44-55.

Hor, C., J. Watson, and S. Majithia. 2005. Analyzing the impact of weather variables on monthly electricity demand. *IEEE Transactions on Power Systems* 20 (4): 2078-85.

Kyriakides, E., and M. Polycarpou. 2007. Short term electric load forecasting: A tutorial. In *Trends in neural computation, studies in computational intelligence*, eds. K. Chen and L. Wang, 35:391-418. New York: Springer.

Levenbach, H., and J. P. Cleary. 2006. *Forecasting: Practice and process for demand management*. Belmont, CA: Duxbury Thomson Brooks/Cole.

Ljung, G. M., and G. E. P. Box. 1978. On a measure of lack of fit in time-series models. *Biometrika* 65:297-303.

Makridakis, S. G., S. C. Wheelwright, and R. J. Hyndman. 1998. *Forecasting: Methods and applications*. 3rd ed. New York: John Wiley and Sons.

McMenamin, J. S. 1997. A primer on neural networks for forecasting. *The Journal of Business Forecasting Methods & Systems* 16 (3): 17-22.

National Climatic Data Center (NCDC). 2009. Climate Data Online, Federal Climate Complex Global Surface Summary of Day Data Version 7. ftp://ftp.ncdc.noaa.gov/pub/data/gsod (accessed November 14, 2009).

Newey, W., and K. West. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61:631-53.

Pappas S. Sp., D. Ch. Ekonomou, G. E. Karamousantas, S. K. Chatzarakis, Katsikas, and P. Liatsis. 2008. Electricity demand loads modeling using autoregressive moving average (ARMA) models. *Energy* 33 (9): 1353-60.

Peirson, J., and A. Henley. 1994. Electricity load and temperature: Issues in dynamic specification. *Energy Economics* 16 (4): 235-43.

Phillips, P., and P. Perron. 1988. Testing for a unit root in time series regression. *Biometrika* 75:335-46.

Pindyck, R. S., and D. L. Rubinfeld. 1998. *Econometric models and economic forecasts*. 4th ed. Boston: Irwin McGraw-Hill.

PJM. 2009. Historical load data. http://www.pjm.com/markets-and-operations/compliance/nerc-standards/historical-load-data.aspx (accessed November 3, 2009).

PJM Resource Adequacy Planning. 2009. *PJM manual 19: Load forecasting and analysis*. Rev. 15. http://www.pjm.com/documents/~/media/documents/manuals/m19.ashx (accessed November 4, 2009).

Ramanathan, R., R. Engle, C. W. J. Granger, F. Vahid-Araghi, and C. Brace. 1997. Short-run forecasts of electricity loads and peaks. *International Journal of Forecasting* 13:161-74.

Schneider, A. M., T. Takenawa, and D. A. Schiffman. 1985. 24-Hour electric utility load forecasting. In *Comparative models for electrical load forecasting,* eds. D. W. Bunn and E. D. Farmer, 87-108. New York: John Wiley and Sons.

Soares, L. J., and M. C. Medeiros. 2005. Modeling and forecasting short-term electric load demand: A two-step methodology. Working Paper 495, Department of Economics, Pontifical Catholic University of Rio de Janeiro.

Soares, L. J., and M. C. Medeiros. 2008. Modeling and forecasting short-term electricity load: A comparison of methods with an application to Brazilian data. *International Journal of Forecasting* 24:630-44.

Soares, L. J., and L. R. Souza. 2006. Forecasting electricity demand using generalized long memory. *International Journal of Forecasting* 22 (1): 17-28.

Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6:461-4.

Taylor, J. W., and P. E. McSharry. Forthcoming. Short-term load forecasting methods: An evaluation based on European data. *IEEE Transactions on Power Systems.*

Taylor, J. W., L. M. De Menezes, and P. E. McSharry. 2006. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting* 22:1-16.

University of Chicago Press. 2003. *The Chicago manual of style*. 15th ed. Chicago: University of Chicago Press.

Wang, G. C. S. 2004. Forecasting practices in electric and gas utility companies. *The Journal of Business Forecasting Methods & Systems* 23 (1): 11-15.

Weron, R. 2006. *Modeling and forecasting electricity loads and prices: A statistical approach.* Chichester: John Wiley and Sons.

Weron, R., and A. Misiorek. 2004. Modeling and forecasting electricity loads: A comparison. In *Proceedings of the European Electricity Market EEM-04 Conference,* Łódź, 135-42.

Wooldridge, J. M. 2009. *Introductory econometrics: A modern approach*. 4th ed. Mason, Ohio: South-Western Cengage Learning.