THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

JOHN AND WILLIE LEONE FAMILY DEPARTMENT OF ENERGY AND MINERAL ENGINEERING AND DEPARTMENT OF ECONOMICS

THE COST OF WEATHER FORECAST ERROR IN ELECTRICITY PRICES IN SOUTHEASTERN PENNSYLVANIA

ALECIA GRACE DECUOLLO
Spring 2012

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Energy Business and Finance and Economics
with interdisciplinary honors in Energy Business and Finance and Economics

Reviewed and approved* by the following:

George Young
Professor of Meteorology
Thesis Supervisor

Andrew Kleit
Professor of Energy and Environmental Economics
Thesis Supervisor

Jonathan P. Mathews
Assistant Professor of Energy and Mineral Engineering
Honors Adviser

James Tybout
Professor of Economics
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

With the structure of competitive electricity markets, there are several incentives to power producers and consumers to accurately forecast the price of electricity for a given area. Through the use of regional forecast weather data and historical load information, one could potentially predict the next-day load schedule, with the ultimate goal of predicting the price of electricity per kilowatt hour. However, errors in forecast temperature are a major cause of short-term electricity load forecast error, having significant financial consequences. The National Research Council estimates that 25% of the U.S. gross domestic product is weather- and climate-sensitive. This paper examines the effect of the margin of error between forecast and observation weather has on electricity prices and load schedule. Using locational marginal pricing along with historical load data for the PECO region along with forecast and observation weather data, this study conducted on an hourly basis from July 10th, 2010 to April 30th, 2011, will test the effectiveness of several weather attributes in accurately predict electricity prices. In my initial data review, I found significant price fluctuations occurred at extreme temperatures of less than 40°F and greater than 70°F. Additionally, I found that at the temperature deviates from my estimated temperature threshold of 55°F threshold, the margin of forecast error becomes more significant in electricity price.

Table of Contents

ABSTRACT i
Acknowledgements iii
Chapter 1: Introduction 1
Chapter 2: Sources for Study 3
2.1: Pennsylvania-New Jersey-Maryland Interconnection 3
Figure 1: Map of PJM 4
2.2: The National Weather Service 5
Chapter 3: Literature Review for Predicting Electricity Prices 6
Chapter 4: Description of Data 9
4.1: Locational Marginal Pricing 9
4.2: Hourly Load 10
Figure 2: Yearly Load Cycle 11
Figure 3: LMP versus Load 12
4.3: Global Forecast System Model Output Statistic (GFS MOS) 13
Figure 4: MOS Stations 13
Figure 5: Forecast Temperature Patterns 14
Figure 6: Observation Temperature versus LMP 15
Chapter 5: Model 16
5.1: Model Specifications 16
Figure 7: Simulated HDD Threshold versus Price 17
Figure 8: Simulated HDD Threshold versus Load 18
5.2: Significance of Forecast Weather to Predict LMP and Load 19
Equation 1: *Stepwisefit* Regression Model Forecast Weather versus Price 19
Table 1: 24 Hour Lead-time Results 19
Table 2: 12 Hour Lead-time Results 20
Equation 2: *Stepwisefit* Regression Model Forecast Weather versus Load 20
Table 3: 24 Hour Lead-time Results 21
Table 4: 12 Hour Lead-time Results 21
5.3: Forecast Error Relating to Price 21
Figure 9: Marginal Cost of Error 22
Equation 3: *Stepwisefit* Regression Model Marginal Cost of Forecast Error 22
Table 5: 24 Hour Lead-time Results 23
Table 6: 12 Hour Lead-time Results 23
Chapter 6: Additional Considerations and Conclusion 25
Appendix - 1: Observation versus LMP Statistics Summary 28
Appendix - 2: Observation versus Load Statistics Summary 29
Bibliography 30

Acknowledgements

I would first like to thank my parents for all of their support and encouragement over the past four years. I also wanted to thank my family for all of their unwavering support, and all of the advice they have given me. No matter what, you have always been there for me and have always supported my decisions. To my family, thank you so much!

I would also like to thank my thesis advisor, Dr. Young. Your encouragement and guidance throughout this whole process has always made the difference. Thank you so much for believing in me and my work. I would also like to thank Dr. Kleit for his support and guidance for my membership in the Honors College. Also, I would like to thank Dr. Mathews and Dr. Tybout for their additional support on my work. I truly appreciate all of the guidance and advice given, and the time taken to make this possible. To Dr. Young, Dr. Kleit, Dr. Mathews, and Dr. Tybout, thank you!

I feel so fortunate to have the opportunity to work with three departments, including the Departments of Earth and Mineral Engineering, Meteorology, and Economics. I was given the opportunity to learn so much, growing academically and professionally. I have had so many wonderful opportunities at Penn State, and I am so grateful for every moment!

Chapter 1: Introduction

Originating in Chile in the 1980s, the deregulation and restructuring of the electricity market was implemented in the United States by the Energy Policy Act of 1992 (Levin, 2011). Both the Environmental Protection Agency (EPA) and Federal Energy Regulatory Commission (FERC) led to the restructuring of vertically integrated electric utilities, along with the creation of the Independent System Operators (ISO) and the Regional Transmission Organizations (RTO). This deregulation allowed for competition within the wholesale power markets and the creation of power based financial derivative instruments, which have been introduced on several exchanges, including the Chicago Mercantile Exchange (CME) and the Intercontinental Exchange (ICE) (Levin, 2011).

With the creation of weather derivatives and greater competition within the wholesale power markets, there is an incentive for models that not only predict the load for a given region, but also the price of electricity (Nogales et.al 2002). Through the use of weather predictors including temperature and wind, along with historical load data for a given region, one can begin to estimate the price of electricity, which offers significant financial incentives. To predict these electricity prices, it is proposed that there should be a greater emphasis placed on the need to improve forecasting models of loads in electric power systems, created through the use of either forecast or observed weather attributes. Commercial and industrial sectors consume over half the total output of electric power, and fluctuating electricity prices pose a substantial financial risk for input costs (Levin, 2011).

Electricity, because of its distinct properties, makes it an interesting, yet challenging commodity. One of the unique characteristics of electricity is that it cannot be stored economically at scale. With traditional goods, we assume that inventories can be used to buffer differences in supply and demand, which can prevent instability in prices caused by market fluctuations (Cartea et al., 2008). Storage of electricity is geographically limited, such as pumped storage/hydroelectric generation in which kinetic energy can be harnessed from the movement of water, and be converted to electric power when needed. However, this form of power is geographically limited, and most undeveloped sites are not economically viable at this time (Bhanot, 2002). Because of the impracticality of storing electricity, supply and demand must be coordinated on an instantaneous basis and is not subject to the buffering effect provided by an inventory of this commodity. Electricity also has a distinctive property of limited transportability; electricity is grid-bound and restricted by not only location and extent of the power grid, but also by transmission line capacity limits, congestion, and marginal efficiency loss as distance increases from the generation site to the point of consumption. As a result, electricity is considered a geographically concentrated good and is subject to local supply and demand conditions (Wilkens et al., 2007).

Previous research on load forecasting has been performed using more traditional forecasting models, such as the breakdown of weather data, including by day, season, hour, workday and weekend. This paper will examine how the margin of error between forecast and observation weather data affects electricity load and locational marginal pricing. Errors in forecast temperature are a major cause of short-term electricity load forecast error, causing instability on the grid and large fluctuation in prices. Weather forecast errors can account for approximately 40 to 90 percent of demand forecast error. As a result, the weather induced load error can be on average 8 to 10 percent of the load, which can cause significant financial consequences (Altalo et al., 2004). The purpose of this study is to examine how electricity load and prices are affected as the lead time in forecast weather data increases. Using the GFS MOV forecast data from the National Weather Service, which extends to 66 hours of lead time, this paper will test the effectiveness of supporting the use of forecast data, including the value of "quality" versus "quantity" of collecting weather data. Should we place more funding into collecting more observations for a longer period of time? Or should we focus on improving the accuracy of weather forecasts along with observations? An understanding of the effectiveness of using forecast data to predict electricity prices has great value for many corporations. With more accurate electricity load predictions, this can help manage utilities risk and prevent large variations in prices.

Chapter 2: Sources for Study

2.1: Pennsylvania-New Jersey-Maryland Interconnection

This study utilizes electricity pricing and load data provided by Pennsylvania-New Jersey-Maryland Interconnection (PJM). PJM is a regional transmission organization (RTO) that coordinates the movement of wholesale electricity in parts of 13 Mid-Atlantic States and the District of Columbia. In 1927, PJM became the first continuing power pool, with the merger of the dispatch services of the Public Service Electric and Gas Company, the Philadelphia Electric Company the Pennsylvania Power & Light Company, the Baltimore Gas and Electric Company, and General Public Utilities. In integrating these major utility companies, the goal was to dispatch electricity to its end-users at the lowest cost. In 1997 the Federal Energy Regulatory Commission (FERC) approved PJM as the first independent system operator (ISO), which is responsible for the generation and transmission units in order to maintain the reliability of the power grid (PJM, 2012).

PJM became the nation's first RTO in 2001, and currently serves a 214,000 square mile area, including all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia, and the District of Columbia (PJM, 2012). The PJM Interconnection is a pool of 18 electric companies, including PECO, which was used in this study. The scope of PJM's operates the largest competitive wholesale electricity market in the world, as this RTO services a population of over 51 million residents, with a generating capacity of 185,600 megawatts and a peak demand of 163,848 megawatts. PJM makes an annual energy delivery of approximately 794 million megawatt-hours, with about 1,365 generating sources of diverse types, both renewable and non-renewables (PJM, 2012). Figure 1 is a map of the states serviced by PJM.

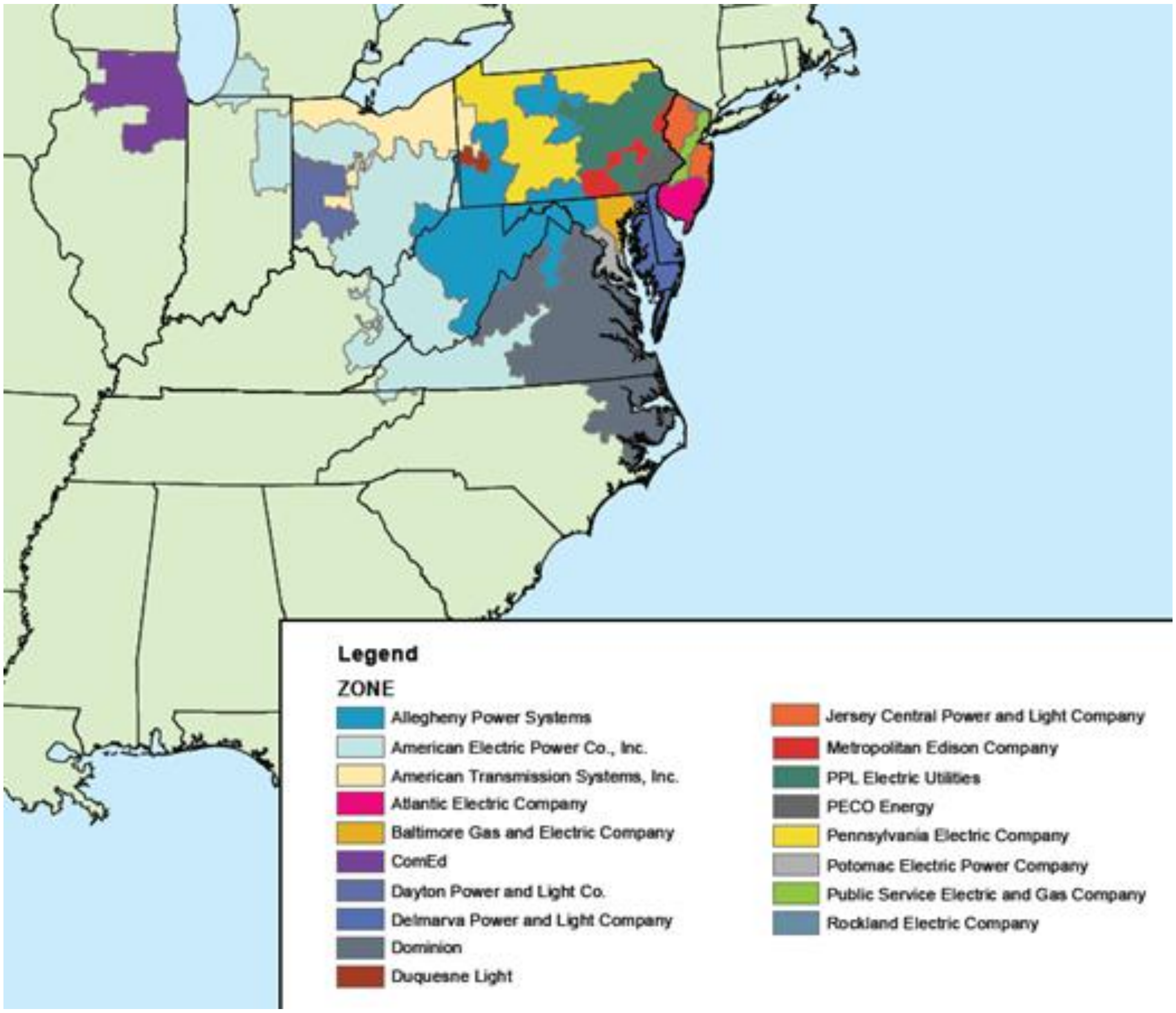

Figure 1: Above is a map of the states serviced by PJM (PJM, 2011).

The basic functions of PJM include monitoring the transmission grid for reliability, balancing the generation and load on an instantaneous basis, operate and support various markets within PJM and plan for transmission expansion on a regular basis. As an RTO, PJM is independent of all generation and power marketing interests, and has the exclusive responsibility for grid operations, short-term reliability, and transmission service for a given region. PJM does not own any transmission or generation assets, function as a publically traded company, perform actual maintenance on generators or transmission systems or directly serve retail customers (PJM, 2012). Like the stock exchange, PJM's wholesale electricity market establishes a market price by coordinating supply with demand. PJM has recently created online tools that enable participants to submit bids and offers for electricity by providing companies with real-time data. These participants are thus able to effectively respond to market fluctuations, enabling PJM to service their customers more efficiently because they are able to meet their demand (PJM, 2012).

2.2: The National Weather Service

The National Oceanic and Atmospheric Administration (NOAA) is a scientific agency within the United States Department of Commerce. This administration was formed in 1970 from the United States Coast and Geodetic Survey (1807), the Weather Bureau (1870), and the Bureau of Commercial Fisheries (1871) (NOAA, 2012). One of the six agencies within NOAA is the National Weather Service (NWS). As stated in the mission statement, NWS “proves weather, hydrologic, and climate forecasts and warnings from the United States, its territories, adjacent waters and ocean areas” (National Weather Service, 2011). Advancement of technology and the formulation of specific mathematical models, along with operation of weather satellites, radars, information processing and communication systems, automated weather observing systems, and super computers allow the NWS to provide weather decision support services, forecasting and climate services to the nation (National Weather Service, 2011).

Chapter 3: Literature Review for Predicting Electricity Prices

My research conducted here quantified how weather forecast error affects LMP prices and revealed extensive prior research on price prediction models. Because of the storage impracticality and coordination difficulties, there are a number of papers that discussed different models used to predict electricity demand. Traditional models to predict electricity demand include the Autoregressive Integrated Moving Average (ARIMA) model, the Generalized Autoregressive Conditional Heteroskedastic (GARCH) model, and the Artificial Neural Networks (ANNs) model. Before beginning my studies in the effects of forecast error on price, I first investigated current models used to predict electricity demand.

The ARIMA model is a times series model which is intended to describe the behavior of independent variables in terms of linear relationships with their past values. This can be broken down into two components. The first is the integrated component, which represents the amount of differencing to be performed on the series to make it stationary. The second component is the ARMA model, containing the autoregressive (AR) component that captures the correlation between the current value of the times series and the past values. The moving average (MA) component represents the duration of the influence of a random shock (Weisang, et al. 2008). There are distinct disadvantages to the ARIMA model due to the occurrence of the wide fluctuations in temperature error. In the findings of Garcia and Contreras (2005), the basic assumption of the ARIMA is that error terms could include zero mean and constant variation, or do not account for heteroskedasticity. Research supported the fact inclusion of heteroskedastic error specification is significant in hourly electricity price data, which improves forecasting day-ahead prices (Garcia, et al., 2005).

The GARCH model considers the moments of time series as variant, in which the error term, or real value minus the forecasted value does not require a zero mean and constant variance as with the ARIMA process. As a result, this model can measure the implied volatility of a time series due to price spikes (Garcia et al., 2005). Data, such as observation weather and electricity, in which variances of the error term are not equal, may be larger at some points than at others and is considered heteroskedastic (Engle, 2001). The major problem of this heteroskedastic data, or data with different variability, is that standard errors and confidence intervals estimated by conventional procedures may be too narrow, which gives a false sense of precision. From the study prepared by Garcia and Contreras in the California Energy Crisis in to predict day-ahead electricity prices , the GARCH model outperformed the traditional ARIMA model, which suggests that success can be attributed to the model's ability to better account for extreme price volatility (Garica et al., 2001).

With Artificial Neural Networks (ANNs), networks of nonlinear regressions are configured to find hidden patterns in a given data set and make generalizations of the patterns even in the presence of "noise" or missing information. Used in load forecasting, ANN can be used to measure the daily detailed variations in load schedule, given various weather and historic load schedules (Ortiz-Arroyo, et al., 2005). This model is an adaptive system-changes structure based on external or internal information that flows through the network. Specifically, the ANN model has been used in short-term load forecasting (1-24 hours), and is able to account for large fluctuations due to temperature. In an ANN model proposed by Hayati and Shirvany, the model specifically used basic weather predictors and calendar effect to predict the load schedule for a given region (Hayati, et al., 2007). One of the key advantages of this system is the model's ability to represent both linear and non-linear relationships, and for these relationships to be derived strictly from the data being modeled.

After reviewing the models used to predict load schedules and electricity prices, I decided that I wanted to investigate a different angle of electricity load schedule and price problem. I wanted to measure the amount of impact meteorological forecast error has on electricity prices. As weather forecast error increases, my study shows that this could have significant influences on the price of electricity, particularly during hot or cold spells. Electricity usage in a given region is very sensitive to temperature fluctuations when temperatures are extreme. Thus, when the predicted load schedule is over- or under-forecasted, we see large price variations. Especially, as we move from off-peak to peak sources of electricity generation, the price of electricity can significantly increase (Zareipour et al., 2010). Though the cost of over-forecasting load is significantly lower, since system stability and reliability is less at stake, the costs associated with under-forecasting and not meeting the demand of the consumers is high (Altalo et al., 2004). Additionally, improved weather forecasting can also lead to more efficient pricing of such commodities at natural gas and agricultural products, which are more supply-side and are sensitive to weather conditions in an area.

When initially reviewing my data, I found that significant price fluctuations occurred at extreme temperatures (40°F and 70°F), suggesting that forecast errors do indeed interact with demand to drive the more extreme price fluctuations. I wanted to test the significance of the impact of forecast error had on electricity prices, thereby quantifying the incentive to improve weather forecasts in the future.

Chapter 4: Description of Data

4.1: Locational Marginal Pricing

This study utilizes data drawn from the real-time energy markets, which reports hourly Locational Marginal Prices, broken down into marginal loss, transmission loss, the base LMP price, and the total LMP price for the PECO region (site ID: 51297). LMP is a pricing method used by PJM to price energy purchases and transmission congestion costs. Specifically, this study captures hourly LMP prices between July 10th 2010 and April 30th, 2011, so that the results from this study would not be influenced by significant inflation of prices. This study thus captures 295 days or 7080 hours. All pricing data can be found on the PJM website. The LMP prices are reported in Eastern Standard Time on an hourly basis.

LMP utilizes the physical, flow-based pricing system, in which electricity is priced by how it flows through the grid, and not based on contract paths (LMP, 2012). PJM provides a market place for power producers to submit generation bids and their corresponding bidding prices, along with consumption companies to submit consumption bids. This process creates a single round auction performed every hour to determine the resulting market clearing price (Nogales et al., 2002).

LMP is often referred to as "nodal pricing," as the purpose of this price is to determine the resulting delivered energy price at a specific location. A market manager will schedule the generators from the lowest bid to the higher bids, until the full forecasted load and losses in a particular region are satisfied. Generation is rescheduled to avoid transmission constraints in the most cost-effective manner. Generation in the transmission constrained areas will be relatively more expensive than the non-constrained area, which will be reflected in the LMP price (Delea et al., 2010). For this study, I have collected the real-time LMP prices; LMP values are based on generation dispatching process used for balancing the transmission system and alleviating congestion. LMP values are also calculated in the day-ahead market, which is priced by scheduled quantities of consumption for the various market participants, in the form of bilateral contracts, price bids, and the reflecting the impact of transmission congestion (Harvey et al., 2005).

LMP determines the energy price for each "node," or point on the electrical network, and includes the transmission congestion and marginal loss price settled at the specific node (Bo et al., 2009). The energy component price represents the optimal dispatch, is the same price across the network in PJM, and is calculated both in day ahead and real time (Harvey et al., 2005). If this transmission network is congested, the next increment of energy cannot be delivered from the least expensive unit without causing an overload to the system. This energy component cost is the cost to serve the next increment of demand at a given node or location that can be produced from the least expensive generating unit in the system that still has available capacity. The transmission congestion cost reflects the cost of constraints on the flow of power across the transmission system, and can be negative if the there is a greater amount of generation than demand. The marginal loss cost at a specific location represents the marginal cost to operate generation, the cost of delivery on the transmission system, and the loss of electrical energy due to the transport of electricity, usually through heat (Litvinov, 2004).

4.2: Hourly Load

This study also incorporated hourly load data from the PECO region. The hourly load data was collected for July 10th, 2010 to April 30th, 2011 in megawatt hours. PJM provides in the real-time energy market the hourly load data from the PJM Mid-Atlantic Region. An electric load can be defined as the "electric demand," as it reflects the power requirement of any device or equipment that converts electric energy into light, cooling, heat, or mechanical energy (Lijesen et al., 1971). Usually, the electric load for a given region is measured in megawatt hours, and varies widely with hourly, weekly, monthly, or annually. For this study, hourly load was reported in Eastern Standard Time on an hourly basis. Figure 2 is a graph created to show the general pattern of the electric load demanded between July 10th, 2010 and April 30th, 2011 by hour.

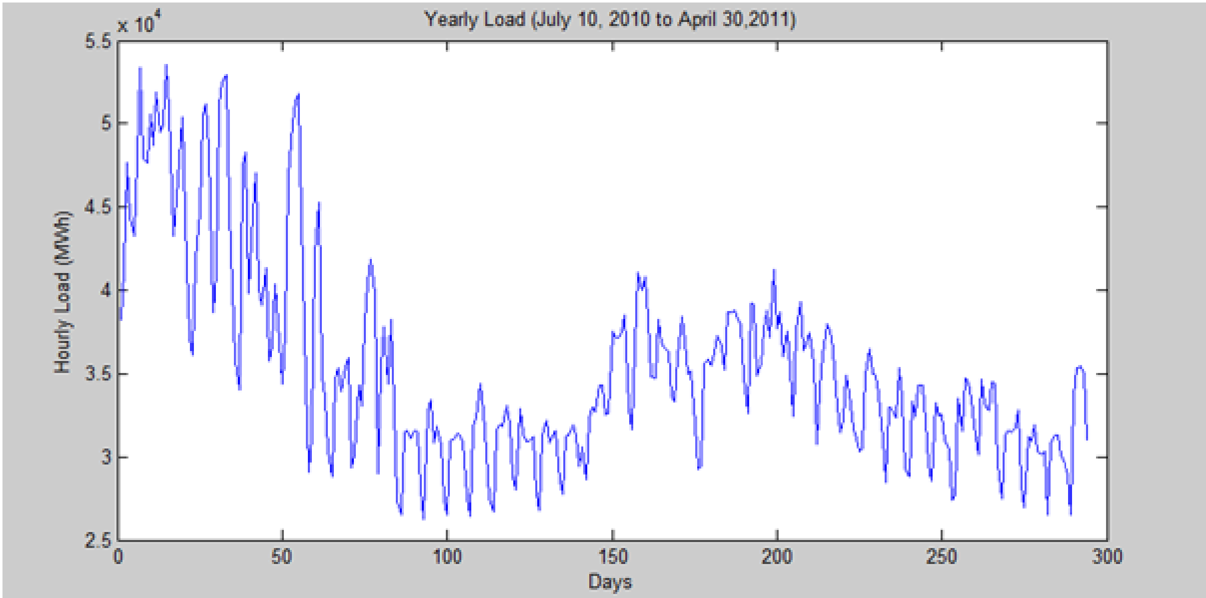

Figure 2: Load schedule for the PECO region, with Day 0 as July 10, 2010 and Day 295 as April 30, 2011.

Figure 2 shows there is distinct patterns in the load schedule for the PECO region. The first pattern to recognize would be that there is daily fluctuation of load demanded. Traditionally, the hours between 5 and 7 pm show the greatest demand for electricity, when there is little to no natural sunlight and there is still activity in the household, whereas the lowest demand generally falls between 3 to 5 am (Levin et al. 2011). The second pattern that is evident is the yearly fluctuations of electricity load patterns. From Figure 2, one can see that in the summer months (day 0-75), the amount of electric load per MWh experiences an increase. This can be attributed to mainly the increase in air conditioning throughout the Mid-Atlantic Region. During the winter months (day 166-263), one can again identify an increase in the load in the Mid-Atlantic, which can be attributed to both the increased use in heating systems and also lighting (as the sun goes down between 6-7pm in this region, sometimes as early as 5pm). During the fall and spring months, the load demands decrease due to the fact that there is less of a need for heating or air conditioning (Crowley, et al. 2005).

There is also a clear pattern that can be seen as one compares the price of electricity versus the electric load for the region, shown in Figure 3.

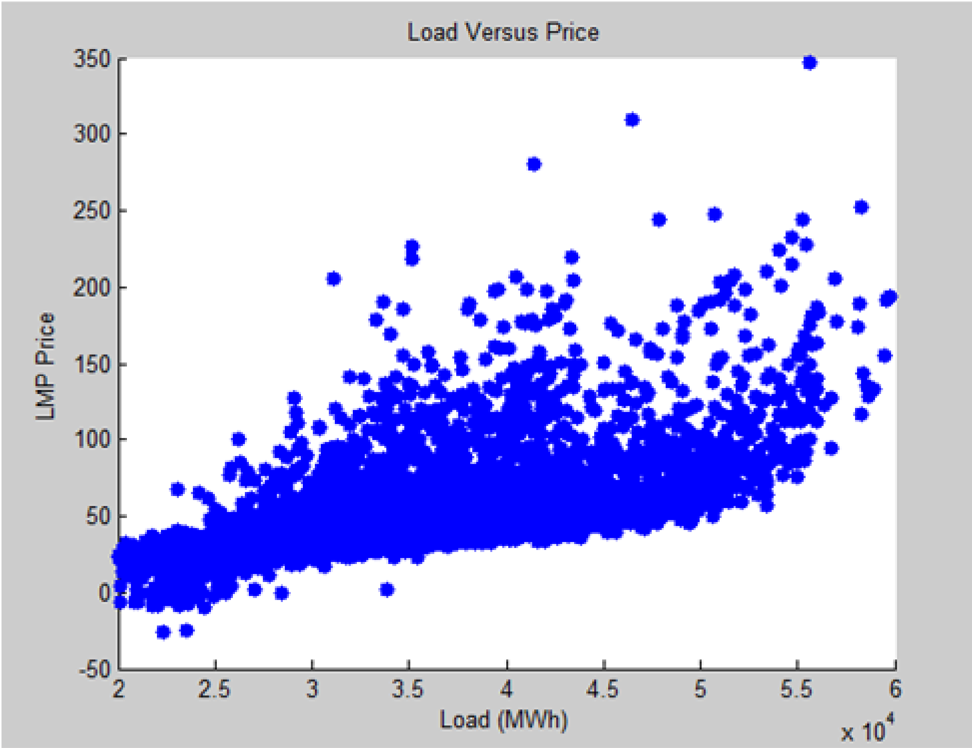

Figure 3: LMP versus Load for the PECO region.

From Figure 3, one can see that as the electric load demand increases, the price of electricity not only increases, but becomes volatile. Simple supply and demand theory can explain the increase in the demand for a commodity will cause the price of that commodity to rise. However, only with the understanding of the distinct characteristics of electricity can one begin to examine the existence of price volatility as the amount of electric demand increases. From this chart, one can see that the price per megawatt can rise to extreme values of over $500 per MWh, when the average price is $49.6444 per MWh. Because of this commodity's properties, when the supply is unstable, one can see extreme fluctuations in price.

4.3: Global Forecast System Model Output Statistic (GFS MOS)

This study also utilizes the Global Forecast System Model Output Statistic (GFS MOS) hourly weather forecast data for the Doylestown Airport (site ID: KDYL). MOS is a weather forecasting technique consisting of a statistically derived relationship between a predicted weather variable and various predictor variables which are forecast by a physically-based numerical model at specific projected times (Allen, 2001). MOS is considered a "backbone" for modern weather forecasting. Figure 4 is a map of the MOS sites throughout the United States. Screening regression is applied to the "weather related" output of a numerical model to generate statistical models for the prediction of surface wind, probability of precipitation, maximum temperature, cloud amount, and conditional probability of frozen precipitation (Glahn et al., 1972). Predictors for MOS include numerical weather prediction (NWP) model output interpolated to observing site, prior observations, and geoclimatic data (Maloney, 2005). MOS combines both numerical weather forecasting and pure statistical models, using complex numerical forecasts based on the physics of the atmosphere to forecast weather patterns. MOS then uses regression equations in to improve surface weather. Statistical post-processing is used to correct for systematic biases of the model, and account for local influences which are not resolved in the model (Glahn et al., 2008).

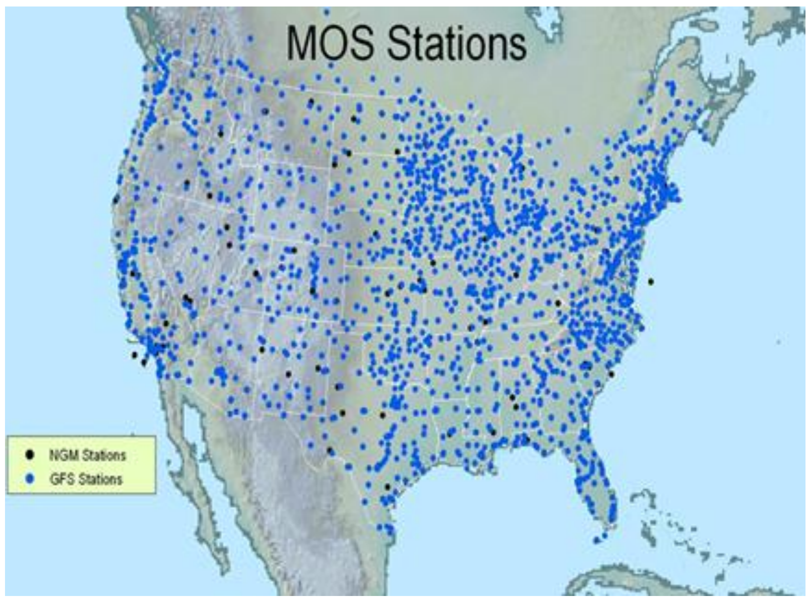

Figure 4: MOS stations throughout the United States (Glahn, 1972).

This study uses MOS forecasts of average temperature (°F), dew point (°F), and wind speed (knots). GFS MOS is provided by the National Weather Service with projections from 6 to 72 hours in advance, in 3 hour intervals, set in Greenwich Mean Time (GMT). This study uses the 1200 UTC (Coordinated Universal Time) forecast cycle, which is noon GMT. All GFS MOS forecasts are reported with valid times of 0000, 0600, 1200, and 1800 UTC (Glahn et al., 1972). For the purposes of this study, all forecast weather data was interpolated into hourly forecasts using cubic interpolation. Figure 5 is a graph that shows the seasonal cycle of temperature, with Day 0 as July 10, 2010 and Day 295 as April 30, 2011.

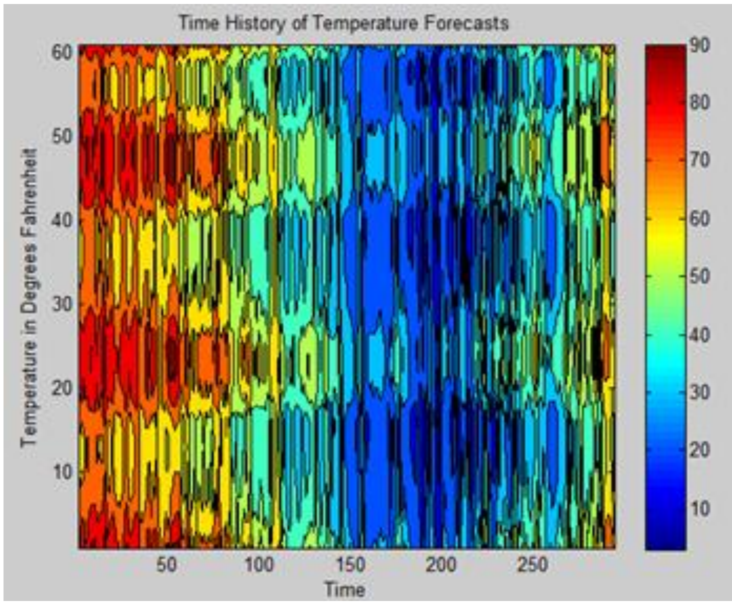

Figure 5: Yearly temperature patterns in the PECO region.

From Figure 5, we can see the temperature cycle throughout the year. Temperatures in the summer months (day 0-75) fall between 55-90°F, whereas the temperatures during the winter months (day 166-263) fall between 10-40°F. The temperature cycle occurring on the vertical axis represents the cycle of temperatures for the 66 hour forecasts.

Hourly observation weather data for this study was provided by The Pennsylvania State Climatologist, which provides climatological data for Pennsylvania (The Pennsylvania State Climatologist, 2011) . Observation data was initially collected at major airports by both the NWS and Federal Aviation Administration (FAA). The Automated Weather Observing System (AWOS) is operated and controlled by the FAA, and generally reports in 20 minute intervals. Various weather attributes are collected, including temperature, dew point, wind speed, wind direction, visibility, sky condition and cloud cover, and precipitation.

Figure 6 is a scatter plot comparing the observed air temperature at the Doylestown Airport, and total LMP for PECO during the studied time period:

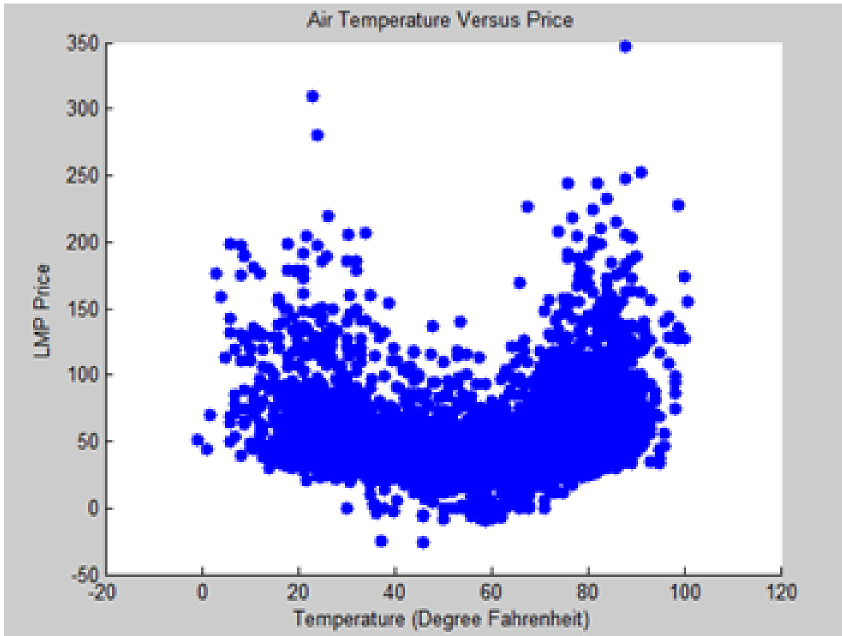

Figure 6: Observation temperature versus LMP in the PECO region.

From Figure 6, we can see that the LMP averages at $49.6444 per megawatt hour. However, prices can increase dramatically, to between $150 and $300 per megawatt hour. This price volatility could be initiated by changes in weather, resulting in the inability to meet load demand. Through the use of price forecasting models, the goals of this study would be to eliminate some of the higher prices, which can reach well over $400 from occurring. For this study, I would like to see how much price is influenced given errors in forecasted weather. From Figure 6, large price fluctuations occur at temperatures below 40°F and above 70°F.

One of the major challenges that I faced for this study was to coordinate all data sets. For simplicity, I coordinated all of my data in Greenwich Mean Time. I also found there was a significant amount of missing weather data, which limited the amount of predictors in this study. I would have liked to include six hour precipitation levels and cloud cover, however my data was limited.

Chapter 5: Model

5.1: Model Specifications

After observing the relationship between total LMP and the observation air temperature, I wanted to create a model that could incorporate all weather attributes in order to predict the electricity load and price throughout the entire year, as opposed to different models for specific seasons. To do this, temperature had to be replaced as a predictor by some quantity which would have the same relationship to load in both summer and winter, (hot summers and cold winters would both increase load, for example). This could be done by taking the absolute value of temperature minus the heating versus cooling threshold. To set this value I found the temperature threshold that yielded the highest correlation between this new metric LMP and load. In effect, I have created a weather derivative.

Created in 1997, weather derivatives, are financial instruments that are used to reduce risk associated with adverse or unexpected weather conditions (Zeng, 2000). Survey data suggests that 70% of the end users of weather derivatives are energy firms, and according to the National Research Council, it is estimated that 25% of the U.S. gross domestic product is weather- and climate-sensitive. As a result, the use of weather derivatives can be an important strategy for many firms to reduce their exposure to weather risk (Pérez-González, et al., 2010).

Key weather derivatives in the energy sector are heating and cooling degree days (HDD and CDD, respectively). HDD and CDD values track the temperature deviations below and above 65°F, respectively (Pérez-González, et al., 2010). Indices will take a specific location's daily average temperature, and this number is used to determine the day's average temperature deviation from the 65°F threshold (Considine, 2011). Essentially, the average temperature for that day determines the value of the weather derivative. Using a stepwise regression model, I first tested if the standard 65°F temperature threshold produced the highest correlation to predict electricity load and LMP. Taken together HDD and CDD have the same form as my metric. The *stepwisefit* regression model in Matlab uses the adjusted R^2 to determine the significance of each variable within the model. The variable with the largest R^2 can be assumed to account for the largest variation in the dependent variable (Lambert, 2007). I chose the *stepwisefit* regression model because of the limited data I had and I wanted to test the significance of each variable.

Using the *stepwisefit* function in Matlab, this regression method contains an algorithm for determining which variables have a significant relationship with the output. For my model, I selected the observation average temperature collected at the Doylestown Airport, between July 10^{th}, 2010 and April 30^{th}, 2011, specifically at 4pm EST. This threshold could vary, given the time of day chosen. HDD and CDD have a threshold of 65°F because it takes into consideration the daily average temperature.

From my study, I found that instead 56°F threshold produced a higher correlation between LMP, with a correlation of 0.2232, as opposed to 65°F with a correlation of 0.183 (refer to Appendix - 1), as shown in Figure 7.

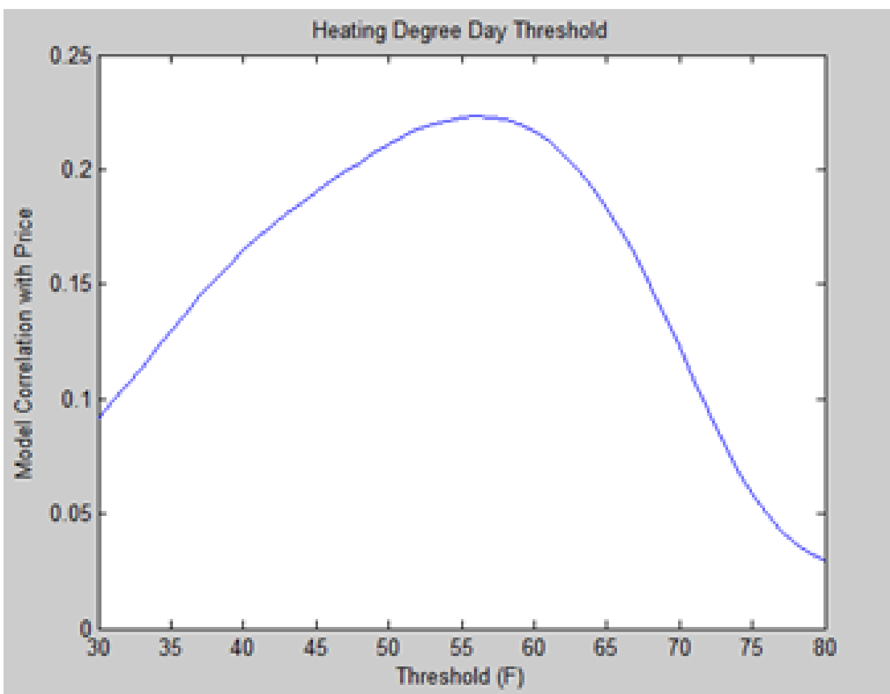

Figure 7: Simulated heating degree day threshold against price.

Similarly, I found that a threshold 55°F had the highest correlation of 0.4965 of predicting load for the Doylestown region, whereas the standard 65°F threshold had a correlation of 0.3697, shown in Figure 8 (refer to Appendix - 2).

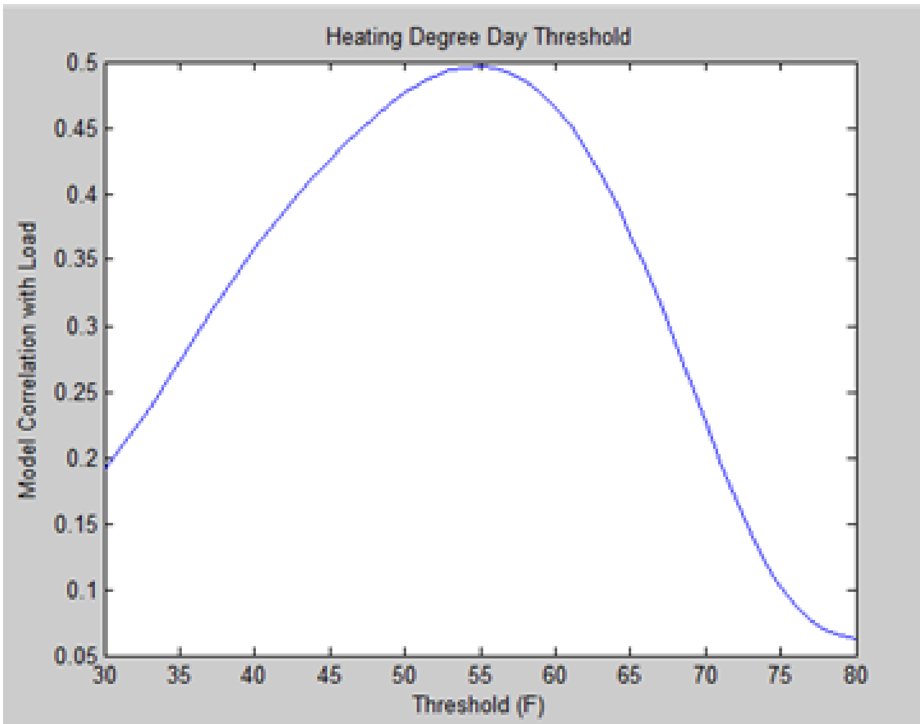

Figure 8: Simulated heating degree day threshold against load.

The current threshold for weather derivatives is 65°F, albeit for 24-hour average load rather than hourly load studied here. From my study of the PECO region, however, it appears that a more efficient threshold should be about 10°F lower when looking specifically at load at the time of maximum temperature. Perhaps improvement in home insulation or the increased cost of heating has also contributed to the lower threshold. On the other hand, the difference could result from looking at one hour throughout the day. An extension of this study would be to find the temperature threshold with the highest correlation for every hour of the day, and perhaps compare this with another site. The threshold temperature may on average be at 65°F, however by checking every hour we can find the most efficient threshold, especially during peak hours.

5.2: Significance of Forecast Weather to Predict LMP and Load

I first examined which weather predictors best predicted the LMP. My weather predictors were the absolute value of the difference between forecast temperature and my 55°F threshold (i.e. max(HDH,CDH)) or (X_{FT}), forecast dew point (X_{DP}), and forecast wind speed (X_{WS}). I chose to include the difference between forecast temperature and my threshold (HDH and CDH) so that my model could be used throughout the entire year. Essentially, the predictor is a measure of the deviation of temperature and its effect on price that is valid year round. In contrast, throughout my study of past research, I found that most models were fit on a seasonal basis, with one model each for the summer and winter season. My approach is more parsimonious. This *stepwise* regression model includes forecasts in degrees Fahrenheit, dew point in degrees Fahrenheit, and wind speed in knots for the Doylestown Airport from July 10th, 2010 and April 30th, 2011 at 4pm.

Equation 1:
LMP = f(abs(Temperature Forecast - 55°F Threshold) + Dew Point Forecast + Wind Speed Forecast)

$$LMP = \beta_o + abs[\beta_1 (X_{FT} - 55)] + \beta_2 X_{DP} + \beta_3 X_{WS}$$

From this regression, I found that the only significant predictor of LMP was the temperature forecast. After performing the regression, the only selected predictor was the absolute value of the different between forecast temperature and the 55°F threshold. It is evident that the primary predictor of LMP would be forecast temperature, shown in the model predicted in Table 1.

24 Hour Lead-time Stepwise Regression Model of Weather Predictors versus Price

Symbol	Variable	Coefficient	Standard Error	p-value
$(X_{FT} - 55)$	Abs(Temperature - 55°F Threshold)	1.3633	0.1706	3.1481e-14
X_{DP}	Dew Point Forecast	0.1810	0.0928	0.0520
X_{WS}	Wind Speed Forecast	-0.0450	0.4331	0.9173

Table 1: 24 Hour Lead-time *Stepwise* Regression model of the given weather predictors versus price.

To compare, I completed the same study again using price as the predictand, with the same predictors, however at a lead-time of 12 hours, or 4am EST. I found the *stepwise* regression found both the difference between the temperature forecast and my threshold, along with forecasted dew point to be significant predictors in this model, shown in Table 2.

12 Hour Lead-time Stepwise Regression Model of Weather Predictors versus Price

Symbol	Variable	Coefficient	Standard Error	p-value
$(X_{FT} - 55)$	Abs(Temperature - 55°F Threshold)	2.0171	0.2394	1.6515e-15
X_{DP}	Dew Point Forecast	1.1907	0.1381	4.0359e-16
X_{WS}	Wind Speed Forecast	0.8601	0.4424	0.0528

Table 2: 12 Hour Lead-time *Stepwise* Regression model of the given weather predictors versus price.

From my model, it is estimated that for every degree in temperature that deviates from the 55°F, LMP increases by $2.21 per MWh, and for every degree of dew point that deviates from the observation, the average price increase is $1.27 per MWh.

I also completed the same study, but used load schedule as a predictand for the regression. The load sample was taken from the PECO region, provided by PJM, and is reported in megawatt hours. Below is a model of my regression:

Equation 2:
Load = f(abs(Temperature Forecast - 55°F Threshold) + Dew Point Forecast + Wind Speed Forecast)

$$\boldsymbol{Load = \beta_o + abs[\beta_1 (X_{FT} - 55)] + \beta_2 X_{DP} + \beta_3 X_{WS}}$$

I completed this regression for a 24 hour lead-time (4pm EST) and a 12 hour lead-time (4am EST). Again, the difference in forecast temperature was a major predictor of load for the PECO region, but the regression results also show that dew point (humidity) is also a predictor of load. The model shows that at the 24 hour lead-time, for every degree variation there is a 453.9970 MWh increase in load. The *stepwise* regression did not consider wind speed a significant predictor of load, shown in Table 3.

24 Hour Lead-time Stepwise Regression Model of Weather Predictors versus Load

Symbol	Variable	Coefficient	Standard Error	p-value
$(X_{FT} - 55)$	Abs(Temperature - 55°F Threshold)	453.9970	18.5435	1.2192e-72
X_{DP}	Dew Point Forecast	86.0886	9.9724	3.9965e-16
X_{WS}	Wind Speed Forecast	19.0727	24.5892	0.4386

Table 3: 24 Hour Lead-time *Stepwise* Regression model of the given weather predictors versus load.

Similarly, at the 12 hour lead-time, for every degree deviation from the 55°F threshold, there was a 452.06 MWh, shown in Table 4. From this, we can see that temperature played a significant role in the load schedule for the PECO region.

12 Hour Lead-time Stepwise Regression Model of Weather Predictors versus Lead

Symbol	Variable	Coefficient	Standard Error	p-value
$(X_{FT} - 55)$	Abs(Temperature - 55°F Threshold)	452.0630	18.6704	1.0656e-71
X_{DP}	Dew Point Forecast	84.1940	9.9514	1.3158e-15
X_{WS}	Wind Speed Forecast	-10.1379	24.0864	0.6741

Table 4: 12 Hour Lead-time *Stepwise* Regression model of the given weather predictors versus load.

5.3: Forecast Error Relating to Price

The final test performed was to see how much LMP is affected by temperature forecast errors at any given temperature. For example, say one predicts the temperature to be 55°F and the actual temperature is 56°F. Given this 1°F forecast error, there will be a specific cost resulting from an incorrect forecast. At a more extreme temperature of say 70°F, when the forecast temperature is 71°F, this error will impose a greater cost than in the prior example because as we deviate from the threshold, the margin of error in forecast becomes more significant. As shown in Figure 9, as we deviate from the threshold temperature of 55°F, the marginal cost for each degree of temperature error will increase.

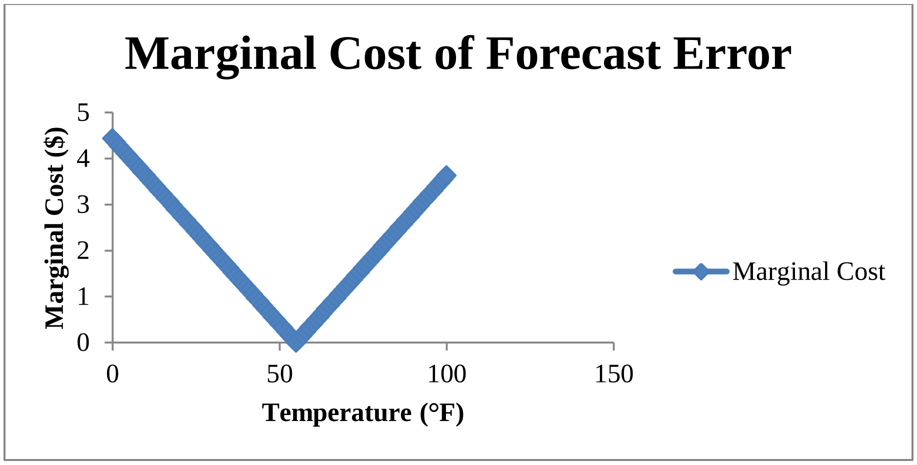

Figure 9: Graph depicts the marginal cost of forecast error increases as we deviate from the 55°F threshold.

This new model includes the same three predictors as the previous model, but includes an additional predictor that models the cost of the temperature forecast error ($(X_{FT} - 55)*X_E$). The absolute value of this variable was not taken because an over-forecast on a hot day has essentially the same effect as an under-forecast on a cold day. I will find the difference between forecast weather attributes and observation attributes, and find the total cost for these errors. Using the *stepwisefit* regression function in Matlab, I will be able to estimate a cost for errors in the forecast temperature.

Below is the stepwise function that I used to calculate the price for error in temperature forecasts:

Equation 3:

LMP = f((Temperature Forecast - 55°F Threshold)*Error(Temp) + abs(Temperature Forecast - 55°F Threshold) + Dew Point Forecast + Wind Speed Forecast)

$$LMP = \beta_o + \beta_1 (X_{FT} - 55)*X_E + abs[\beta_2 (X_{FT} - 55)] + \beta_3 X_{DP} + \beta_4 X_{WS}$$

From my results, shown in Table 5, at a lead-time of 24 hours (4pm EST), the *stepwisefit* model included three predictors, including the measure of cost for forecast error, the difference in forecast temperature and my threshold and, lastly, dew point. From these results, one can see that for every degree that temperature deviates from the 55°F threshold, the total additional cost is approximately $1.41 per

MWh and for every degree increase of humidity it is about \$0.21 per MWh. In this model, however, we have included a measure of forecast error cost as we deviate from our threshold. For this study, at a lead-time of 24 hours, we found that for every 12.41°F the actual temperature deviate from the normal temperature, and degree of forecast error has an effect of 1°F real change in temperature. For example, if the actual temperature were 70°F , the departure from 55°F would be 15°F so each degree of temperature error would have 15/12.41 = 1.21 times as much effect as a degree in actual temperature change.

24 Hour Lead-time Stepwise Regression Model to Measure the Marginal Cost of Forecast Error

Symbol	Variable	Coefficient	Standard Error	p-value
$(X_{FT} - 55)*X_E$	(Temperature - 55°F Threshold)*Error(Temperature)	-0.1136	0.0289	1.0766e-4
$(X_{FT} - 55)$	Abs(Temperature - 55°F Threshold)	1.4055	0.1718	9.2941e-15
X_{DP}	Dew Point Forecast	0.2103	0.0912	0.0218
X_{WS}	Wind Speed Forecast	-0.1027	0.4468	0.8183

Table 5: 24 Hour Lead-time *Stepwise* Regression model of the given weather predictors versus price.

I also completed this study for the 12 hour lead-time (4am EST), shown in Table 6. I found similar results, in which a degree deviation in temperature from the 55°F threshold costs approximately \$1.95 per MWh and a degree increase in humidity costs approximately \$1.04 per MWh. I also found that for every 15°F that the temperature is above or below 55°F, the price effect of temperature increases by one times that of real temperature deviation. From both studies we see the marginal cost for temperature error increases as we deviate from the given threshold of 55°F.

12 Hour Lead-time Stepwise Regression Model to Measure the Marginal Cost of Forecast Error

Symbol	Variable	Coefficient	Standard Error	p-value
$(X_{FT} - 55)*X_E$	(Temperature - 55°F Threshold)*Error(Temperature)	-0.1289	0.0315	5.4926e-5
$(X_{FT} - 55)$	Abs(Temperature - 55°F Threshold)	1.9452	0.2368	7.3607e-15
X_{DP}	Dew Point Forecast	1.0397	0.1407	1.6274e-12
X_{WS}	Wind Speed Forecast	0.7578	0.4422	0.0877

Table 6: 12 Hour Lead-time *Stepwise* Regression model of the given weather predictors versus price.

From the outcomes of this study, we see that temperature and dew point are the most significant predictors for electricity price and load schedule. My final test emphasized that as temperature forecasts deviate from the observed temperature, the price of electricity will increase, particularly when the observed temperature is already in the high-price regime far from 55°F.

In order to hedge against weather risk exposure, corporations will buy CDD and HDD futures contracts six to nine months in advance in order to reduce extreme weather exposure. This is where this study becomes important. We see that for every degree Fahrenheit that is over- or under-estimated has a specific cost (my study takes two hours for comparison, 4am and 4pm), there can be significant financial consequences. When temperature is underestimated, utility companies are forced to tap into other energy sources that are significantly more expensive to generate electricity (Considine, 2011). Generally, base load power schedules can be met using the relatively more inexpensive power sources, including coal, nuclear and hydropower. However, during peak hours, when there is a greater demand for electricity, the demand is met through relatively more expensive generating sources such as wind and natural gas (NEED, 2011). Though historical weather data is provided through NOAA, from this study, we see that seasons with significantly higher or lower averages in temperature force utilities to tap into peak demand energy sources, which have a higher variable cost than base-load coal.

Chapter 6: Additional Considerations and Conclusion

There are aspects of this model that could be improved or considered within this study. The structure of the data sets, along with the difficulty extracting various components of desired data, inhibiting archived variables I would have liked to have considered, including six hour precipitation measures along with cloud cover.

One of the first improvements that could be made to this study is the length of time period chosen. If I were to perform this experiment again, I would increase the time frame of this study, perhaps to 5 years, or even 10 years. Though there would be constraints to the extension of this time frame because load data found does not exist before 1998, I believe that an expansion of this variable would improve the predictability of price, and give more clarity to how the price or amount of electricity demanded has increased or decreased, perhaps due to improvements in technology or increases in population, especially for a highly congested area similar to southeastern Pennsylvania. However, if the time period is expanded, there would need to be appropriate considerations of price inflation, as prices for electricity would need to be adjusted for inflation within the market (the effects of market inflation could be more difficult, in light of the recession, which according to the NBER, fell from December 2007 to December of 2009).

The second change that I would make to this study is that I would extend this study in a geographic location, perhaps in the western region of PJM, or an area that is less congested or does not include a major city. It should be noted that the PECO region is heavily congested, as it contains the fifth most populous city in the country, with highly and densely populated suburbs that surround this region. If this study were to be conducted again, I would like to include a comparison of several regions throughout the PJM Interconnection. However, it is important to recognize that these studies would be conducted on an individual basis, and at no point would any of the weather predictors be averaged across sizes.

Throughout this experiment, the use of weather data was particularly important to the determination of load schedules, which will ultimately influence the LMP for the region. Extracting forecast and observation data from the National Weather Service, proved to be the most difficult task of this study. Both forecast and observation data structures have not been updated for several decades, and thus extracting necessary data to complete this study was both tedious and inefficient. The main argument becomes not to improve the availability of this data to the public, but to improve the amount of information and the quality of the forecasts. For this study, there were only four forecast variables that could be collected, including air temperature, dew point, wind direction and wind speed over a lead time span of 72 hours. As the forecast weather data consisted of collections every three hours, as temporal interpolation was needed to begin the process of comparing this data with other attributes on an hourly basis. Though only small pieces of this data were missing, I believe it would be interesting to see how this model would change if more variables were present. The observation data, however, had additional weather attributes collected, including total cloud cover and six hour accumulated liquid precipitation. Because the forecast data did not have these attributes calculated, I chose not to use them in the comparison of forecast versus observation study. However, these variables are important attributes that could have an effect on the LMP for PECO, and more importantly, yield better price prediction models.

An important consideration for determining electricity demand in a given region is also the extent of extreme weather has on a given area, including the occurrence of extreme weather over an extended period of time. For example, say that the temperature reaches levels above average for a period of five days. A household will be more like to run their air conditioning after this extended period of above average temperatures (thus demanding more electricity), than if the temperature had been above average for merely a day. If a more extensive study is conducted, another consideration would be to quantitatively measure the impact of above average temperatures for an extended period of time has on the consumption of electricity in a specific region.

Changes in energy regulation and legislation, along with improvements in energy technology have a significant effect on the price of electricity. Coincidently, recent policy changes that place more stringent environmental regulations on coal-fired plants has caused companies such as FirstEnergy Corporation, an energy company that services parts of New Jersey, Pennsylvania, Ohio, West Virginia and Maryland, to close six of its older coal fired plants by September of 2012. This is not only expected to cause a about 3% in company employees, but reduce FirstEnergy's total generating capacity of approximately 23,000 megawatts by about 2690 megawatts, which is over 10% (Stynes, 2012). As roughly 50% of the electricity in this country is generated from coal, significant changes to environmental and energy policy could greatly impact the price of electricity. Though recent changes to legislation my not directly eliminate a specific source of energy, it may however drive this resource out of the market, as stricter environmental regulation could increase the marginal cost of using a specific a resource. By increasing the regulations, the EPA estimates that the cost of carrying out new regulations on mercury regulations could cost upwards of $9.6 billion per year, as only 60% of the country's current coal-and-oil fired plants utilize technology that meet these standards (Silverstein, 2012). As political and economic forces continue to challenge and influence the types of resources used to produce energy, a more extensive study could quantify these changes to predict future electric loads.

This study shows the influence of different weather predictors on the price of electricity and load schedule for Southeastern Pennsylvania to be significant. I think an interesting extension of this work would be to find the most efficient threshold of temperature with the highest correlation to price and load schedule for every hour of the day for a given region. This could be valuable information, and allow for more efficient pricing of weather derivatives. It may also allow corporations with significant weather risk to properly hedge against extreme temperature conditions and avoid wide price fluctuations due to the over- or under-supply of electricity. However, this study successfully showed that weather forecast error does cause significant price fluctuations in LMP, and that at extreme temperatures the marginal cost of error increases, causing large fluctuations in electricity prices.

Appendix - 1: Observation versus LMP Statistics Summary

Temperature	Coefficient 1	Std. Err	p	Coefficient 2	Std. Err	p	Coefficient 3	Std. Error	p	Correlation
30	0.7275	0.0283	3.61E-141	-0.4146	0.0257	8.20E-58	0.5585	0.0666	5.54E-17	0.09144
31	0.7612	0.0279	1.71E-157	-0.4217	0.0249	1.31E-63	0.5526	0.0662	7.66E-17	0.09877
32	0.7912	0.0275	7.01E-174	-0.4235	0.024	1.43E-68	0.5495	0.0658	7.50E-17	0.1061
33	0.8211	0.0272	1.36E-191	-0.4224	0.0231	3.14E-73	0.5467	0.0654	7.01E-17	0.1139
34	0.8489	0.0268	1.03E-209	-0.417	0.0222	6.11E-77	0.5449	0.065	5.73E-17	0.1218
35	0.8741	0.0264	7.04E-228	-0.407	0.0214	2.40E-79	0.5455	0.0646	3.47E-17	0.1297
36	0.8962	0.026	9.48E-246	-0.3929	0.0205	3.28E-80	0.5484	0.0642	1.54E-17	0.1374
37	0.9158	0.0256	5.26E-263	-0.3749	0.0197	2.34E-79	0.5532	0.0638	5.25E-18	0.1448
38	0.9333	0.0252	1.49E-279	-0.3534	0.0189	1.03E-76	0.5607	0.0635	1.23E-18	0.1518
39	0.948	0.0249	3.05E-295	-0.3284	0.0181	3.54E-72	0.5703	0.0632	2.06E-19	0.1583
40	0.9618	0.0246	1.8786e-310	-0.301	0.0174	4.39E-66	0.582	0.0628	2.47E-20	0.1647
41	0.9727	0.0242	0.00E+00	-0.2705	0.0167	3.03E-58	0.5955	0.0625	2.15E-21	0.1705
42	0.9823	0.024	0.00E+00	-0.2378	0.016	4.10E-49	0.6113	0.0623	1.21E-22	0.1759
43	0.9893	0.0237	0.00E+00	-0.2028	0.0154	5.18E-39	0.6293	0.062	4.48E-24	0.1808
44	0.9969	0.0234	0.00E+00	-0.1666	0.0149	7.65E-29	0.6468	0.0618	1.59E-25	0.1856
45	1.0044	0.0232	0.00E+00	-0.1293	0.0144	3.53E-19	0.6648	0.0615	4.75E-27	0.1902
46	1.0141	0.0231	0.00E+00	-0.0912	0.014	7.28E-11	0.6838	0.0613	1.04E-28	0.1947
47	1.0257	0.023	0.00E+00	-0.052	0.0136	1.32E-04	0.705	0.0611	1.35E-30	0.199
48	1.0338	0.0221	0.00E+00	-0.0118	0.0133	0.3762	0.7406	0.0591	1.07E-35	0.2031
49	1.0555	0.023	0.00E+00	0.0295	0.0131	0.0239	0.752	0.0607	5.40E-35	0.2072
50	1.072	0.023	0.00E+00	0.0722	0.0129	2.10E-08	0.7762	0.0605	2.33E-37	0.2109
51	1.091	0.0232	0.00E+00	0.1164	0.0128	9.61E-20	0.8003	0.0603	8.63E-40	0.2144
52	1.1083	0.0233	0.00E+00	0.162	0.0127	9.04E-37	0.8246	0.0602	2.70E-42	0.2172
53	1.1267	0.0235	0.00E+00	0.2091	0.0128	2.61E-59	0.8485	0.0601	8.26E-45	0.2196
54	1.1442	5.0749	0.00E+00	0.2576	0.0129	9.82E-87	0.8731	0.06	1.94E-47	0.2214
55	1.1621	0.024	0.00E+00	0.3075	0.0131	6.37E-118	0.8981	0.0599	3.68E-50	0.2227
56	1.1789	0.0243	0.00E+00	0.3584	0.0134	2.94E-151	0.9232	0.0599	6.46E-53	0.2232
57	1.1934	0.0247	0.00E+00	0.4099	0.0138	9.45E-185	0.9482	0.0599	1.19E-55	0.2228
58	1.2088	0.0251	0.00E+00	0.4624	0.0143	2.54E-217	0.9744	0.06	1.47E-58	0.2218
59	1.2208	0.0255	0.00E+00	0.5147	0.0149	7.74E-247	1.002	0.0601	1.34E-61	0.2198
60	1.2292	0.0259	0.00E+00	0.5661	0.0155	7.85E-272	1.0304	0.0602	1.09E-64	0.2165
61	1.2335	0.0264	0.00E+00	0.6161	0.0163	1.77E-291	1.0598	0.0604	7.58E-68	0.212
62	1.2361	0.027	0.00E+00	0.6646	0.0171	1.12E-305	1.0898	0.0606	4.90E-71	0.2065
63	1.2336	0.0275	0.00E+00	0.7101	0.018	9.0375e-314	1.1207	0.0609	3.11E-74	0.1998
64	1.2268	0.0282	0.00E+00	0.7523	0.019	3.8456e-316	1.1531	0.0613	1.55E-77	0.192
65	1.215	0.0288	0.00E+00	0.7902	0.0201	9.7799e-313	1.1862	0.0617	8.33E-81	0.183
66	1.1948	0.0295	0.00E+00	0.8214	0.0212	1.30E-303	1.2186	0.0622	8.07E-84	0.1728
67	1.1682	0.0302	5.56E-303	0.8459	0.0224	1.18E-289	1.2485	0.0627	2.26E-86	0.1616
68	1.1327	0.031	7.30E-274	0.8615	0.0237	3.12E-271	1.275	0.0633	2.72E-88	0.1494
69	1.0871	0.0317	3.46E-243	0.8664	0.0249	5.52E-249	1.2971	0.064	1.85E-89	0.1363
70	1.029	0.0323	3.03E-211	0.8583	0.0262	1.78E-223	1.3114	0.0647	1.80E-89	0.1225
71	0.9609	0.033	8.64E-179	0.8374	0.0274	9.25E-196	1.3177	0.0654	2.40E-88	0.1082
72	0.8837	0.0335	1.44E-147	0.8045	0.0286	1.16E-167	1.3152	0.0662	3.73E-86	0.09432
73	0.8013	0.0341	5.52E-119	0.7617	0.0296	8.55E-141	1.3047	0.0669	4.01E-83	0.08139
74	0.7127	0.0345	7.43E-93	0.7085	0.0306	3.01E-115	1.2863	0.0676	2.56E-79	0.06941
75	0.6206	0.0348	4.24E-70	0.6476	0.0315	3.97E-92	1.2602	0.0682	6.48E-75	0.05885
76	0.5273	0.0349	7.59E-51	0.5811	0.0322	9.30E-72	1.2275	0.0688	4.07E-70	0.04984
77	0.437	0.035	1.71E-35	0.5132	0.0327	9.60E-55	1.1908	0.0692	2.81E-65	0.0426
78	0.3481	0.035	2.97E-23	0.4432	0.0332	2.19E-40	1.1501	0.0697	2.23E-60	0.0368
79	0.2639	0.0348	3.75E-14	0.3744	0.0335	7.36E-29	1.1077	0.07	1.14E-55	0.03249
80	0.1856	0.0346	8.31E-08	0.3083	0.0337	6.63E-20	1.0651	0.0703	2.93E-51	0.0295

Appendix - 2: Observation versus Load Statistics Summary

Temp	Coef. 1	Std. Err	p	Coef. 2	Std. Err	p	Coef. 3	Std. Error	p	Corr.
30	271.6124	7.0777	3.62E-299	-128.0179	6.4322	2.4156E-86	146.7108	16.6688	1.59E-18	0.1916
31	284.2181	6.9566	0	-130.7189	6.1935	1.29E-96	144.4999	16.4862	2.20E-18	0.2061
32	296.1434	6.8252	0	-131.8958	5.947	3.11E-106	142.9768	16.2968	2.04E-18	0.2213
33	308.2123	6.6868	0	-132.0576	5.6943	1.06E-115	141.4922	16.095	1.75E-18	0.2379
34	319.8608	6.5435	0	-130.8327	5.4399	4.93E-124	140.2509	15.886	1.25E-18	0.2552
35	330.9283	6.397	0	-128.0924	5.1861	1.69E-130	139.726	15.6702	5.72E-19	0.2731
36	340.9654	6.2494	0	-123.7383	4.9382	3.94E-134	140.0808	15.4537	1.51E-19	0.2909
37	350.1989	6.1065	0	-117.9071	4.6991	1.77E-134	141.1316	15.2391	2.48E-20	0.3085
38	358.8387	5.971	0	-110.7012	4.4698	3.55E-131	143.189	15.0262	1.99E-21	0.3259
39	366.4149	5.8376	0	-102.0512	4.2507	1.29E-123	146.1202	14.8186	7.99E-23	0.3426
40	373.4023	5.7142	0	-92.2236	4.0448	5.24E-112	149.9973	14.6178	1.43E-24	0.3587
41	379.1039	5.5945	0	-81.0328	3.8519	5.12E-96	154.6914	14.4294	1.17E-26	0.3735
42	384.2821	5.4832	0	-68.8456	3.6725	5.47E-77	160.3618	14.2485	3.40E-29	0.3877
43	388.5144	5.3752	0	-55.6671	3.5073	5.41E-56	166.9124	14.0776	3.40E-32	0.4009
44	392.7444	5.2775	0	-41.847	3.3568	2.19E-35	173.424	13.9134	2.24E-35	0.4136
45	396.7955	5.1893	0	-27.4324	3.222	1.94E-17	180.198	13.7577	7.48E-39	0.4256
46	401.559	5.1158	0	-12.6046	3.1022	4.88E-05	187.4601	13.6048	9.00E-43	0.4373
47	408.4268	4.7879	0	2.7468	2.9983	0.3596	192.6679	13.0608	1.07E-48	0.4483
48	412.6857	5.0181	0	18.6544	2.9106	1.54E-10	204.5964	13.325	1.46E-52	0.4584
49	419.431	4.9934	0	35.0189	2.8385	1.07E-34	214.1861	13.1955	1.92E-58	0.4681
50	426.0001	4.9806	0	51.9919	2.7843	2.04E-76	223.7976	13.0835	1.32E-64	0.4764
51	433.3388	4.9865	0	69.5408	2.7475	9.99E-137	233.4034	12.9845	4.84E-71	0.4838
52	439.7565	5.0017	0	87.6718	2.7312	1.81E-214	243.1226	12.9107	1.17E-77	0.4892
53	446.5171	5.0322	0	106.3155	2.7339	1.45E-306	252.6442	12.8522	2.56E-84	0.4935
54	452.6048	5.0749	0	125.4502	2.7583	0.00E+00	262.4395	12.8187	3.81E-91	0.4958
55	458.5345	5.1338	0	145.0172	2.8045	0.00E+00	272.3933	12.8074	4.79E-98	0.4965
56	463.7298	5.2088	0	164.863	2.8744	0.00E+00	282.3388	12.8251	9.77E-105	0.495
57	467.8704	5.2995	0	184.7742	2.969	0.00E+00	292.1517	12.8739	5.31E-111	0.4911
58	471.8584	5.4111	0	204.8033	3.0878	0.00E+00	302.4026	12.9467	2.70E-117	0.4854
59	474.2428	5.5381	0	224.5081	3.233	0.00E+00	313.062	13.0536	2.29E-123	0.477
60	474.9313	5.6849	0	243.6222	3.4061	0.00E+00	323.9237	13.198	6.24E-129	0.4656
61	473.7277	5.8502	0	261.8347	3.6069	0.00E+00	335.0203	13.3776	5.49E-134	0.4513
62	471.2796	6.0422	0	279.0006	3.8363	0.00E+00	346.1831	13.5893	2.02E-138	0.4344
63	466.7901	6.2497	0	294.7276	4.0921	0.00E+00	357.4502	13.828	2.50E-142	0.4152
64	460.279	6.4761	0	308.7612	4.3755	0.00E+00	369.0655	14.0935	6.93E-146	0.3937
65	451.4161	6.7242	0	320.6092	4.6872	0.00E+00	380.6419	14.3872	9.23E-149	0.3697
66	439.1813	6.9795	0	329.5311	5.0216	0.00E+00	391.5917	14.705	1.29E-150	0.3436
67	424.4396	7.2486	0	335.5373	5.3778	0.00E+00	401.3525	15.0381	2.97E-151	0.3161
68	406.135	7.523	0	337.8089	5.7518	0.00E+00	409.4304	15.3874	2.20E-150	0.2873
69	384.0611	7.7947	0	335.754	6.1344	0.00E+00	415.5239	15.7468	5.05E-148	0.2576
70	357.1282	8.0513	0	328.4303	6.517	0.00E+00	418.3303	16.1141	1.71E-143	0.2272
71	326.5974	8.3036	4.75e-313	316.2478	6.898	0.00E+00	417.873	16.4786	3.20E-137	0.1972
72	292.6784	8.5316	6.54E-243	299.3102	7.2624	0.00E+00	413.8569	16.8267	1.89E-129	0.1687
73	256.8162	8.7304	6.48E-182	278.5166	7.598	7.83E-275	406.6017	17.1452	9.44E-121	0.1231
74	219.4408	8.8957	3.48E-130	254.3618	7.899	9.48E-216	396.4757	17.4323	1.79E-111	0.1209
75	181.2307	9.0143	4.97E-88	227.629	8.1558	1.13E-164	383.5218	17.6805	8.12E-102	0.1023
76	143.2245	9.0909	3.31E-55	199.3022	8.3664	8.64E-122	368.3201	17.8873	3.73E-92	0.0875
77	106.2963	9.125	3.83E-31	170.3906	8.5305	6.13E-87	351.5314	18.0519	8.23E-83	0.0766
78	70.388	9.1215	1.32E-14	141.0953	8.6522	5.76E-59	333.4399	18.1792	7.70E-74	0.0691
79	36.2405	9.0855	6.69E-05	112.2649	8.7356	1.76E-37	314.6838	18.271	1.85E-65	0.0647
80	4.7156	9.0253	0.6013	80.6892	3.6972	5.16E-103	293.1226	17.4672	2.76E-62	0.0631

Bibliography

Allen, Rebecca L. (2001). Observational Data and MOS: The Challenges in Creating High-Quality Guidance. *Preprints 18th Conference on Weather Analysis and Forecasting*, Ft Lauderdale, American Meteorological Society.

Altalo, Mary G & Smith, Leonard. (2004, October). Using Ensemble Weather Forecasts to Manage Utilities Risk. *Environmental Finance*, pp. 48-49. ISSN 1468-8573.

Bhanot, Karan. (2002). Value of an Option to Purchase Electric Power – The Case of uncertain Consumption. *Energy Economics, 24(2)*, 120-135.

Bo, Rui. (2009). Congestion and Price Prediction in Locational Marginal Pricing Markets Considering Load Variation and Uncertainty. *The University of Tennessee, Knoxville. Trace: Tennessee Research and Creative Exchange*, 75-120.

Cartea, L., & P. Villaplana. (2008). Spot Price Modeling and the Valuation of Electricity Forward Contracts: the Role of Demand and Capacity. *Journal of Banking & Finance*, *32(12),* 2502.

Company Overview. *PJM*. 2012. Retrieved January 20, 2012 from http://www.pjm.com

Considine, Geoffrey. (2011). Introduction to Weather Derivatives. *Weather Derivatives Group, Aquila Energy*, 1-10.

Crowley, Christian & Joutz, Frederick L. (2005, December 15). Weather Effects on Electricity Loads: Modeling and Forecasting. *The George Washington University: Department of Economics*, 1-48.

Daily Real-Time LMP Files. (2011, April). *PJM*. Retrieved June 20, 2011 from http://www.pjm.com

Delea, Frank & Casazza, Jack. (2010). Understanding Electric Power Systems. *IEEE Press*. John Wiley & Sons, 289-291.

Electricity. (2011). *National Energy Education Development Project*, 1-7. Retrieved March 24, 2010.

Engle, Robert. (2001). GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives*, *15(4),* 157-168.

Garcia, Reinaldo C. & Contreras, Javier. (2005, May). A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices. *IEEE Transactions on Power Systems*, *20(2),* 867-874.

Glahn, H. R., K. Gilbert, R. Cosgrove, D. P. Ruth, and K. Sheets, (2008). Gridded MOS Guidance in the National Digital Guidance Database. *Preprints 19th Conference on Probability and Statistics*, New Orleans, American Meteorological Society, 11.3

Glahn, Harry & Lowry, Dale. (1972, July 27). The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *American Meteorological Society*, 1203-1211.

Harvey, Scott M. & Pope, Susan L. (2005, February 23). Comments on the California ISO: MRTU LMP Market Design. *LECG, LLC*. Prepared for California Independent System Operator. Retrieved February 10, 2012 from http://www.hks.harvard.edu

Hayati, Mohsen & Shirvany, Yazdan. (2007). Artificial Neural Network Approach for Short Term Load Forecasting for Illam Region. *World Academy of Science, Engineering and Technology*, 1-5.

Hourly Load Data. (2011, April). *PJM*. Retrieved December 15, 2011 from http://www.pjm.com

Introduction and PJM Governance. *PJM*. 2012. Retrieved January 18, 2012 from http://www.pjm.com

ISD-Lite. (2011). *NOAA*. Retrieved October 12, 2011 from ftp://ftp.ncdc.noaa.gov/pub/data/noaa/isd-lite/

Lambert. (2007, November). Stepwise Linear Regression Terms and Implementation. *Brigham Young University*. Retrieved March 5, 2012 from math.byu.edu/

Levin, Noah. (2011, May). The Impact of Weather Forecasts on Day-Ahead Power Prices. *Claremont McKenna College, 210*, 1-31.

Lijesen, D.P. & Rosing, J. (1971, March). Adaptive Forecasting of Hourly Loads Based on Load Measurements and Weather Information. *INTASA*, 1757-1767.

Litvinov, Eugene. (2004). Marginal Loss Modeling in LMP Calculation. *IEE Transactions on Power Systems, 19(2)*, 880-888.

LMP Overview. *PJM*. 2012. Retrieved January 20, 2012 from http://www.pjm.com

Maloney, Joe. (2005, July 26). Everything You Wanted to Know About MOS. *NOAA*, Retrieved March 1, 2012 from www.nws.noaa.gov

MDL's GFSMAV MOS Archives. (2011, June). *NOAA*. Retrieved July 20, 2011 from http://mdl.nws.noaa.gov

National Weather Service Mission Statement. (2011, February 23). *National Weather Service*. Retrieved February 28, 2012 from http://weather.gov/

National Weather Service Strategic Planning and Policy. (2011, June 20). *National Weather Service*. Retrieved March 1, 2012 from http://www.nws.noaa.gov/sp/

NOAA Legacy. (2012). *NOAA*. Retrieved January 28, 2012 from http://www.history.noaa.gov

Nogales, Francisco & Contreras, Javier. (2002, May). Forecasting Next-Day Electricity Prices by Time Series Models. *IEEE Transactions on Power Systems*, *17(2*), 342-348.

Observation Weather Data. (2011). *The Pennsylvania State Climatologist*. Retrieved December 12, 2011 from http://climate.met.psu.edu/www_prod/

Ortiz-Arroyo, Skov, Morten, & Huynh, Quang. (2005). Accurate Electricity Load Forecasting with Artificial Neural Networks. *Aalborg University*, 1-6.

Pérez-González, Francisco & Yun, Hayong. (2010, September). Risk Management and Firm Value: Evidence from Weather Derivatives. *Stanford University*, 1-49.

Silverstein, Ken. (2012, January 25) Obama Gives Green Light to New Mercury Rules. *Forbes*. Retrieved from http://www.forbes.com

Stynes, Tess. (2012, January 26). FirstEnergy To Close 6 Coal-Fired Plants, Raises 2011 View. *The Wall Street Journal*. Retrieved from http://online.wsj.com

Weisang, Guillaum & Awazu, Yukika. (2008). Vagaries of the Euro: An Introduction to ARIMA Modeling. *Bentley College,* 45-55.

Wilkens,Sascha & Jens Wimschulte. (2007) The Pricing of Electricity Futures: Evidence From the European Energy Exchange. *The Journal of Futures Markets*, *27(4)*, 387

Zareipour, Hamidreza, Cañizaries, Claudio, & Bhattacharya, Kankar. (2010, February). Economic Impact of Electricity Market Price Forecasting Errors: A Demand-Side Analysis. *IEEE Transactions on Power Systems, 25(1)*, 254-262.

Zeng, L. (2000). Pricing Weather Derivatives. *Journal of Risk Finance*, 72-78.

Current Address:
255 South Atherton St., Apt 406
State College, PA 16801

Alecia G. DeCuollo

aleciadecuollo@gmail.com
(267) 566-5045

Permanent Address:
4 Kenmore Road
Yardley, PA 19067

Education:

The Pennsylvania State University – Schreyer Honors College — Graduation: May 2012
Bachelor of Science in Energy, Business and Finance
Bachelor of Science in Economics
Interdisciplinary Honors in Energy Business and Finance and Economics
Departmental Honors in Economics
Honors Thesis: The Cost of Weather Forecast Error in Electricity Prices in Southeastern Pennsylvania
Thesis Supervisor/Honors Advisors: Dr. G. Young, Dr. A. Kleit, Dr. J. Mathews, Dr. J. Tybout

Work Experience:

Research Assistant – Departments of Earth and Mineral Engineering and Meteorology — Summer 2011

- Goal: Built statistical tests and time series models to show the effectiveness of using 63 hour lead-time forecast weather data to be used to predict electricity load and prices for Southeastern Pennsylvania
- Developed, organized and structured large data sets using Matlab for current and future research projects which have been presented to the National Weather Service for further review and future implementation
- Conducted statistical analysis on the movement of prices given the corresponding forecast/observational data

Teaching Assistant
The Pennsylvania State University

EBF 473: Risk Management in Energy Industries	Fall 2011
EBF 301: Global Finance for the Earth, Energy, and Materials Industries	Spring 2011
GEOG 160: Mapping Our Changing World	Fall 2009

Lifeguard/Swim Instructor — 2005-2009
Lower Makefield Township Community Pool

- Monitored and provided supervision of the pools in order to ensure patrons' safety
- Completed required certification in CPR, AED, First Aid and Life-guarding (American Red Cross)
- Provided group and private lessons which included swim stroke development, water skills for all skill levels (Level 1 – 7) and various ages (3 – 12 years of age)

Club Water Polo Team — 2008-2011

Energy Business and Finance Society — 2008-2012

Honors/Awards/Skills:

- Proficient in Microsoft Office, Matlab, STATA
- Africana Research Center Undergraduate Research Symposium, 1st Place Winner of Undergraduate Research Competition "Ideology, Race and Education: A Study of the Impact of Apartheid on the Education of Black South Africans –1948-94"
- Redmond Academic Merit Scholarship, 2009
- ConocoPhillips Scholarship, 2010-2011
- R. Wayne Atwell Scholarship, 2011-12
- Hess Scholarship, 2011-2012
- Mary Kathleen Kelley Memorial Scholarship, 2012