AN EVALUATION OF PJM'S PEAK DEMAND FORECASTING PROCESS

Prepared by

The Brattle Group
353 Sacramento Street, Suite 1140
San Francisco, California 94111

Prepared for

Capacity Adequacy Department
PJM Interconnection, LLC
Valley Forge, Pennsylvania

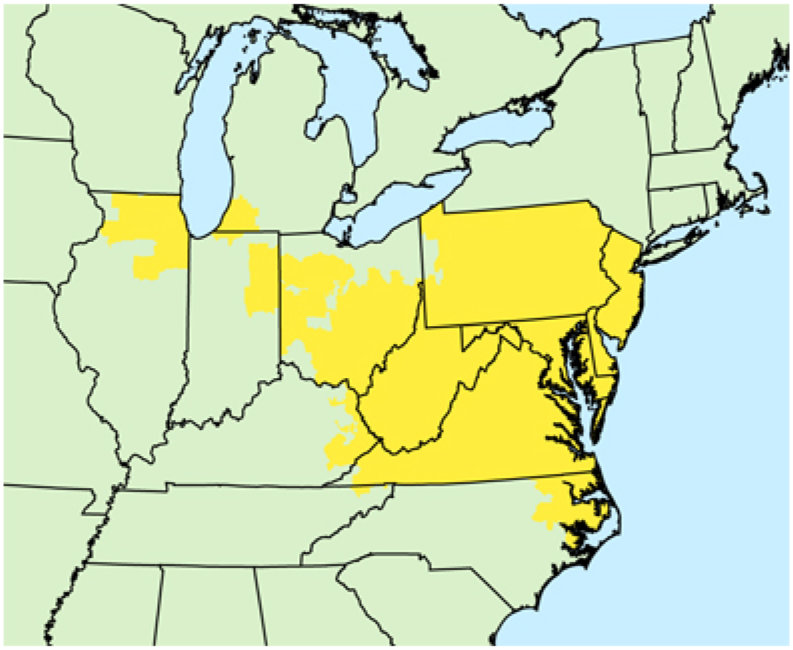
December 5, 2006

TABLE OF CONTENTS

EXHIBITS

1.0 THE FORECAST REVIEW PROCESS

PJM Interconnection, L.L.C. (PJM) is a regional transmission operator (RTO) serving a population of 51 million people living in 13 states and the District of Columbia. The RTO is comprised of 18 transmission zones, as shown in the following map.

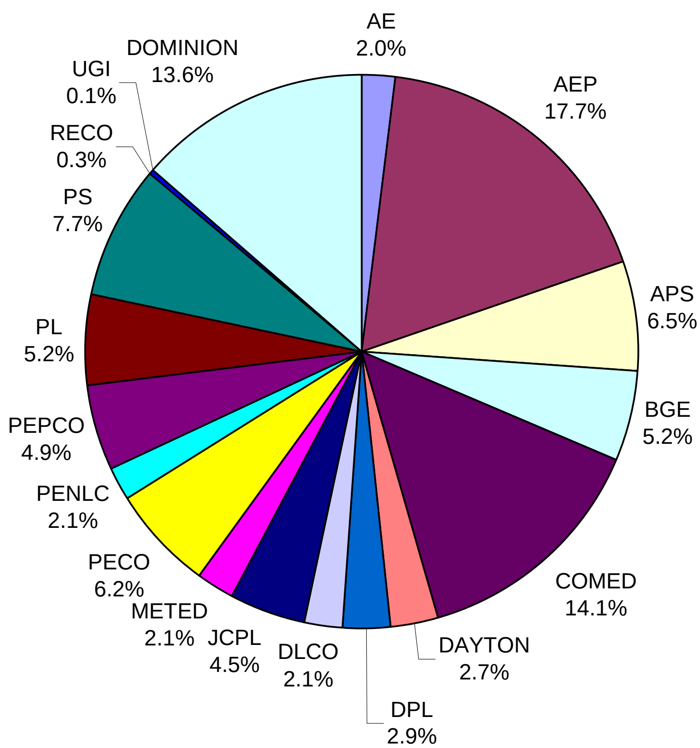


PJM's Capacity Adequacy Planning Department (CAP) is responsible for determining and monitoring the generation reliability requirements of PJM. This includes analyzing the growth of electrical peak load within the region.

Currently, CAP forecasts electrical peak load growth using a set of models and a simulation process that forecast load growth in 18 transmission zones that comprise the PJM RTO footprint. Each of the zones corresponds to an electric distribution company (EDC). These vary in size, as shown below in the following exhibit.

Exhibit 1

Percent of RTO Load by EDC Zone on 2005 Peak (7/26/2005)



Unusually hot weather was encountered in the summer of 2006 and this appeared to cause the PJM forecasting model to under-forecast peak demand. We were retained by PJM to perform an independent review of its modeling and simulation techniques to ensure there was no systematic bias in the forecasting process.

1.1 TERMS OF REFERENCE

Our review encompassed several facets of the forecasting process, including model specification, estimation data base, estimation results and forecasts. Specifically, we were charged with addressing the following fourteen issues:

1. Evaluate the CAP forecast methodology to determine the independence of the PJM RTO coincident peak (CP) forecast from the EDCs non-coincident peak (NCP) forecasts

2. Evaluate the reasonableness of the PJM RTO CP forecast with the sum of EDCs' NCP forecasts

3. Ensure that the sums of NCP forecasts for the three regions (Mid-Atlantic, Southern, Western) were consistent with the PJM CP forecast

4. Assess the applicability of the methodology on a zone, sub-zone or cluster (combination of zones or sub-zones) level

5. Evaluate the appropriateness of the major drivers of annual peak electrical demand within PJM and its regions (Mid-Atlantic, Southern and Western) used in the forecast models

6. Assess that the methodology achieves a reasonable balance between accuracy and ease of use

7. Assess that appropriate metrics are provided to measure forecast accuracy

8. Assess that the methodology correctly incorporates expected (50^{th} percentile) and extreme (90^{th} percentile/10^{th} percentile) forecasts of the drivers

9. Evaluate the accuracy of the methodology through out-of-sample forecasting

10. Assess that the methodology adequately addresses the issue of weather-normalization of data

11. Assess that the methodology accurately reflects the diversity across the entire RTO region

12. Evaluate the representation of RTO regional weather using zonal weather

13. Assess that appropriate weather variables are included in the model

14. Review the reasonableness of projected future model enhancements

1.2 MODEL SPECIFICATION AND FORECASTS

During the kick-off meeting that was held at PJM's offices, PJM staff provided a brief history of how the forecasting model evolved and its current specification. This information supplemented the White Paper that had accompanied the Request for Proposals.

PJM staff described the forecasting methodology that was developed and implemented in 2005. PJM's methodology addresses the different geographic regions, time differences and weather patterns within the RTO. It incorporates a simulation of historical weather patterns and regional diversities. And it was used to make the February 2006 forecast.

The PJM non-coincident peak (NCP) model specification consists of 55 independent variables, including calendar (day of the week, month of the year, holidays, amount of sunlight), weather, economic conditions and weather-regional economy interactions. This specification was estimated by Ordinary Least Squares (OLS) with the MetrixND software using non-coincident daily system peak data over the 1998 – October 2005 timeframe for the 18 electric distribution companies. It was chosen after reviewing several specifications based on in-sample and out-of-sample performance. In PJM's forecasting approach, while the parameter estimates do not vary by month, they do vary across the 18 electric distribution company zones. No correction is made for serial correlation, a commonly encountered problem in time series data that can lead to biased estimates of the standard errors of the parameters in the model.

Forecasts of annual peak demand (NCP) are made with this model for each zone using historical zonal weather for each year beginning with 1971 data. This produces a distribution of forecasted annual demand across yearly weather scenarios. A forecast of annual peak demand is developed using the median (50^{th} percentile), defined as normal or expected weather, of the historical weather distribution.

By summing up the 18 median zonal NCP forecasts, an estimate is derived of the aggregate NCP demand for the RTO. In addition, by applying historical coincidence factors to each of the zonal forecasts, CP forecasts are derived for each zone. An aggregate CP forecast for the RTO is then developed by summing up the zonal CP forecasts.

As a cross-check on this "bottom-up" CP forecast, a model is estimated for the RTO as a whole, using the RTO peak demand as the dependent variable. This model is then used to make a "top-down" forecast of RTO demand and is compared to the "bottom-up" aggregate CP forecast.

At the kick-off meeting, we reviewed the PJM model specification, estimation data base, estimation results and forecasts. We also reviewed projected model enhancements that had been identified by PJM. We suggested some alternative model specifications during the meeting and PJM staff ran them at the RTO level and for two of the zones during the meeting using the MetrixND software.

1.3 EVALUATION CRITERIA

Following the meeting, we established three types of model evaluation criteria: quantitative, qualitative and other. These are listed in the following table:

Quantitative Criteria	Qualitative Criteria	Other Criteria
In-sample goodness-of-fit (high R-squared) Signs, magnitudes and statistical significance of parameters Plausible "influence statistics" In-sample stability of parameters (Chow test) Lack of serial correlation in residuals (DW around 2) Lack of multicollinearity among regressors Out-of-sample goodness-of-fit (low MAPE)	In-sample goodness-of-fit (high R-squared) Signs, magnitudes and statistical significance of parameters Plausible "influence statistics" In-sample stability of parameters (Chow test) Lack of serial correlation in residuals (DW around 2) Lack of multicollinearity among regressors Out-of-sample goodness-of-fit (low MAPE)	Independence of forecasting process from EDC forecasting process General consistency of forecasting results with those from EDCs Ability to capture changes in diversity factors Ability to forecast load in different states of the world, not just to generate outcomes in a given state of the world

1.4 EXPLORATORY DATA ANALYSIS

The Brattle Group (TBG) acquired the econometric database and forecasting model from PJM. As a prelude to model testing and experimentation, we plotted out the dependent variable, daily peak demand, for the RTO as a whole and for each individual zone. Results for the RTO and two representative zones, COMED and PEPCO, are shown in the following exhibits. All the other zones are shown in the Statistical Appendix.

The data spans the time period from January 1, 1998 to October 31, 2005, or nearly eight years. Several things are evident in the plots. The data has periodicity, seasonality and a gradual trend. It also has some random movements that probably reflect weather variations.

Exhibit 2
Non Coincident Daily Peak Load RTO

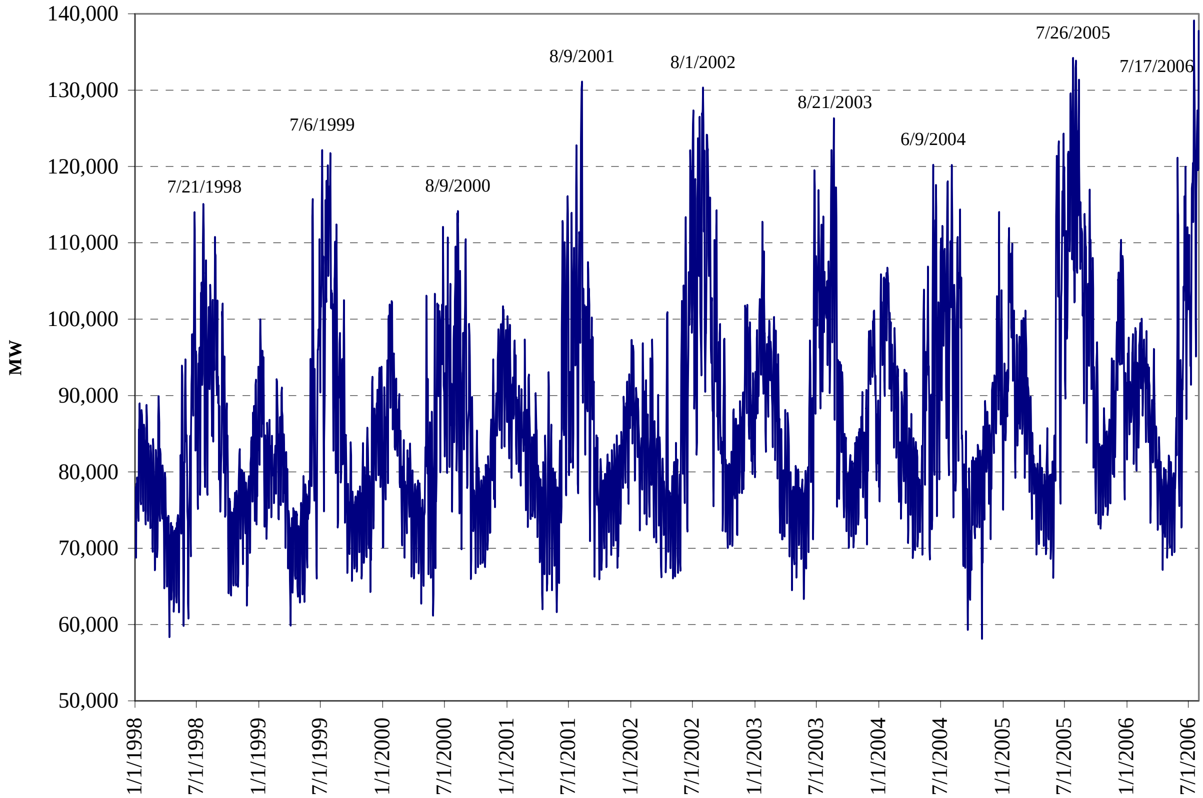


Exhibit 3
Non Coincident Daily Peak Load COMED

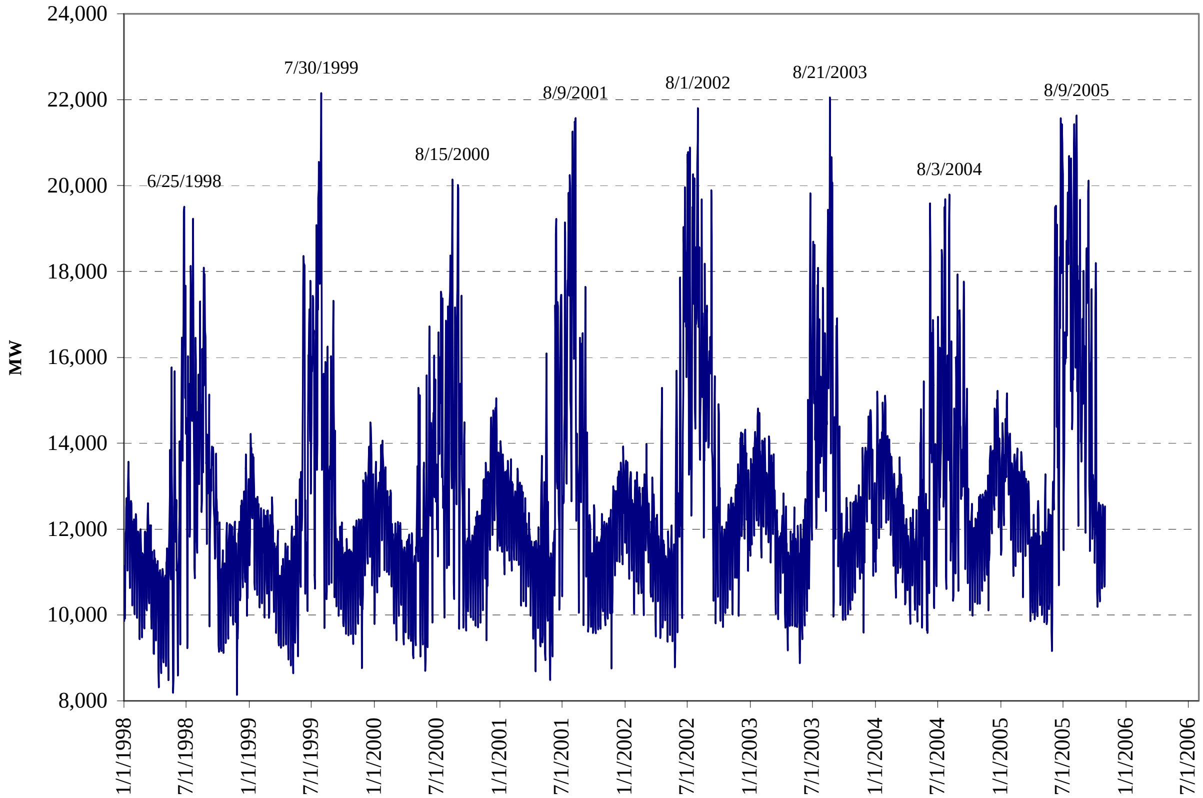
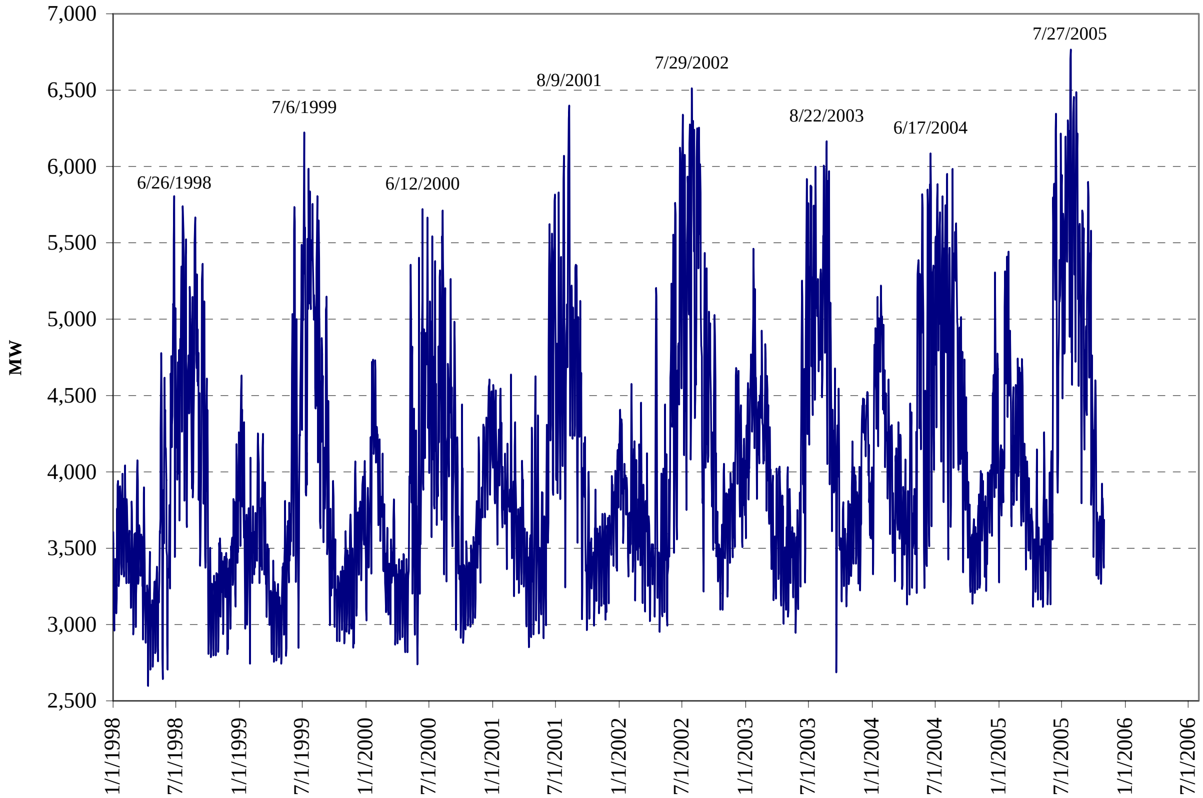


Exhibit 4
Non Coincident Daily Peak Load PEPCO



The data suggests that yesterday's peak demand is a major determinant of today's peak demand. We tested this by plotting the two against each other. Such scatter plots for the RTO and the two representative zones are shown in the exhibits that follow.

Exhibit 5
Non Coincident Peak Daily Load vs. Non Coincident Peak Daily Load Lagged 1 Day RTO

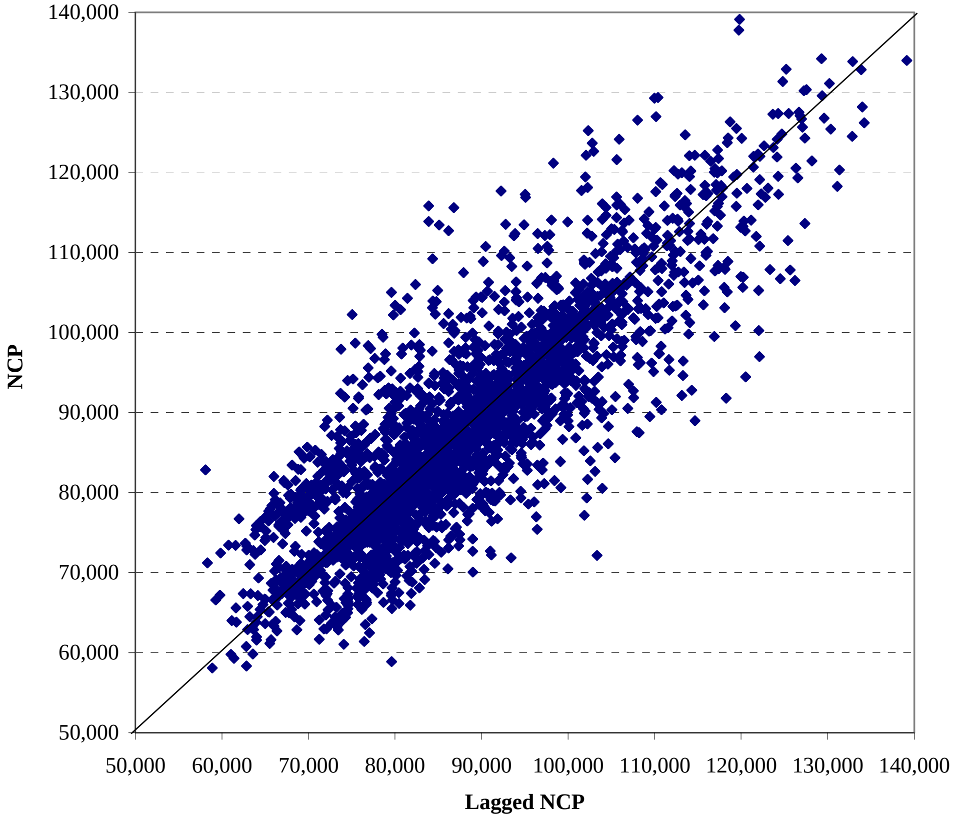


Exhibit 6
Non Coincident Peak Daily Load vs. Non Coincident Peak Daily Load Lagged 1 Day
COMED

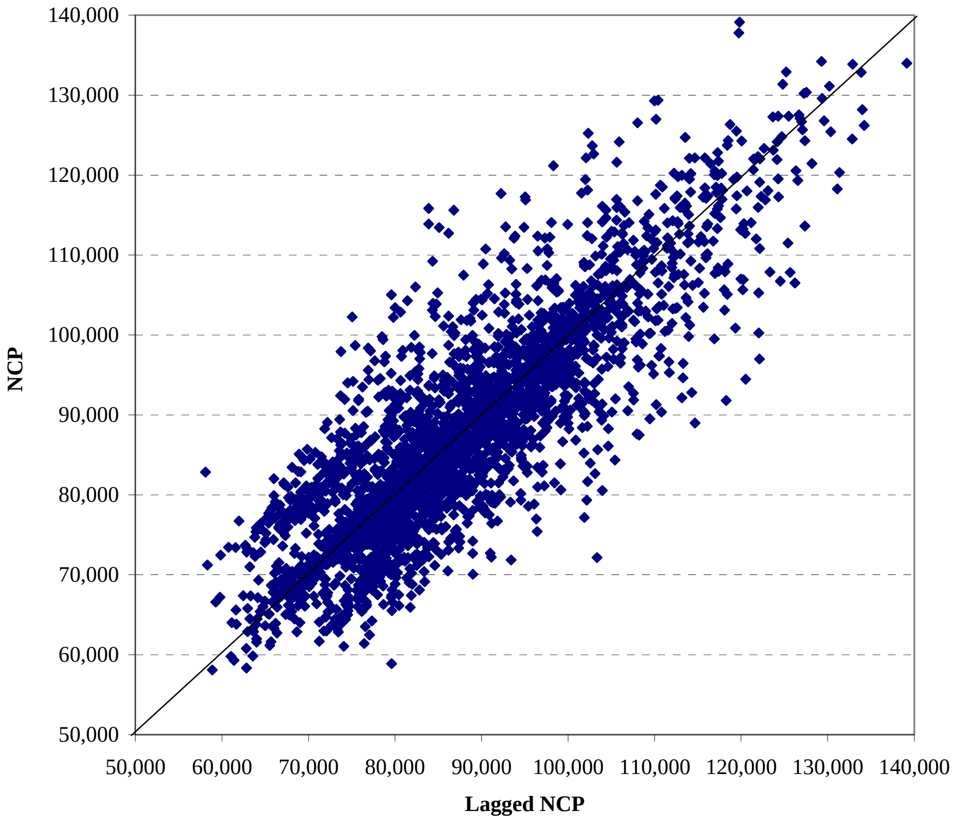
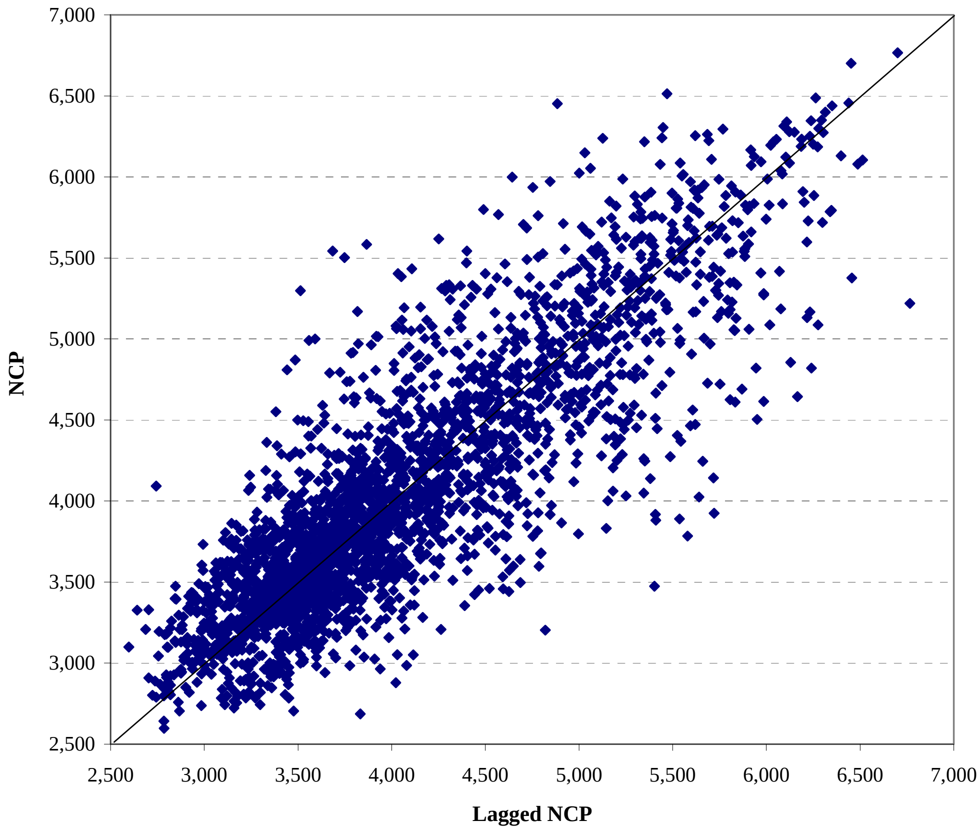


Exhibit 7
Non Coincident Peak Daily Load vs. Non Coincident Peak Daily Load Lagged 1 Day
PEPCO



The plots show a moderately strong relationship between today's peak demand, plotted on the vertical axis, and yesterday's peak demand, plotted on the horizontal axis. This relationship is there even though the data has not yet been adjusted for differences arising from day-of-the-week, month-of-the-year, and weather and economic conditions. Such adjustments are best performed through econometric analysis.

1.5 Model Testing and Experimentation

Initially, we re-ran the model specification from the February 2006 PJM forecast in the STATA econometric software package. Initially, we ran this on the RTO data set and on two representative zones to make sure that the STATA software was well calibrated to the MetrixND software. Once that was established, we proceeded to estimate other model specifications on the PJM database.

After having completed this calibration process, we set out to estimate some simple regression models to explain the behavior of daily peak demand, in order to identify the role played by key drivers.

- TBG1. To capture the periodicity in the data, we laid out a simple specification that just contained binary variables reflecting day of the week, month of the year, and holidays. This list was very similar to that contained in the PJM model except that we eliminated any variables that were labeled as non-binary since we felt that would make the resulting equation easier to grasp.

- TBG2. This built on the first one by adding cooling and heating degree day variables, which were a subset of the many weather variables included in the PJM specification.

- TBG3. This built on the second one by adding an economic variable (gross metropolitan product, GMP). The frequency of GMP is quarterly, not daily. Even then, it is useful to include in the specification. A priori, one would expect a growing economy to be accompanied by growing electric demand.

- TBG4. This built on the third one by allowing weather sensitivity to change as a function of regional economic growth. This captures the effect of rising penetration of central air conditioning and electric heating systems. During the eight years represented in the sample, all zones have seen growth in such penetrations, with some reporting a doubling in saturation rates.

We ran the four simple models and compared their goodness of it with the more comprehensive PJM model. At the RTO level, the simplest model explains 59 percent of the variation in daily peak demand. This is quite remarkable, since the model is limited to calendar variables. It suggests that six-tenths of the variation in daily peak demands stems from very predictable factors. When two simple weather variables are included in the second model, we are able to explain 75 percent of the variation in system peak demand. When the regional economy is included in the third model, the amount of explained variation rises to 82 percent. Finally, when we allow weather responsiveness to vary with economic growth, the amount of explained variation rises to 93 percent. This is just a percent shy of the amount of variation that is explained by the PJM model, which comes in at 94 percent. Generally similar results are observed by zone, as shown in the following exhibit.

Exhibit 8
Adjusted R-Squared by Model and by Zone RTO

	PJM Model	TBG-1 Model	TBG-2 Model	TBG-3 Model	TBG-4 Model
RTO	0.9434	0.5896	0.7538	0.8160	0.9332
AE	0.9599	0.6015	0.7224	0.7612	0.9242
AEP	0.9510	0.5863	0.7358	0.7620	0.9093
APS	0.9502	0.4889	0.6524	0.7814	0.9173
BGE	0.9620	0.4962	0.6483	0.7035	0.9175
COMED	0.9683	0.5695	0.7353	0.7691	0.9386
DAYTON	0.9047	0.5840	0.7073	0.7298	0.8732
DLCO	0.9504	0.5973	0.7447	0.7598	0.9091
DOMINIC	0.9610	0.4919	0.6364	0.7421	0.9262
DPL	0.9496	0.5316	0.6872	0.7550	0.9326
JCPL	0.9667	0.5251	0.6797	0.7460	0.9431
METED	0.9349	0.5254	0.6589	0.7808	0.9152
PECO	0.9626	0.5491	0.6993	0.7459	0.9381
PENLC	0.9391	0.6598	0.7619	0.8552	0.9187
PEPCO	0.9710	0.5721	0.7073	0.7698	0.9441
PL	0.9592	0.6042	0.7473	0.8101	0.9244
PS	0.9745	0.6350	0.7639	0.7972	0.9529
RECO	0.9499	0.6519	0.7665	0.7676	0.9193
UGI	0.9590	0.5449	0.7222	0.7894	0.9263

Having seen how the simpler models got progressively better with the inclusion of weather, economy and weather-economy variables, we focused henceforth on comparing the PJM model with TBG4 model.

Our next task was to assess the forecasting performance of the models. This cannot be captured by the goodness of fit calculations shown above. It requires that out-of-sample forecasts be generated, i.e., that some of the data be withheld from the model estimation database. We performed four such experiments.[1] Initially, we withheld the data for the 10 months of 2005 that were available and estimated the PJM and TBG-4 models over the 1998-2004 dataset. Subsequently, we withheld data for 2004-05, then for 2003-05 and finally for 2002-05.

All four tests were performed at the RTO level and for several of the zones.

After having performed these out-of-sample tests, we proceeded to conduct several other tests that are enumerated below:

Tested for the presence of serial correlation by re-estimating the model with a first-order autoregressive AR(1) process. We also introduced a lagged dependent variable in the

[1] In each case, we used actual weather and economic conditions to simulate the out-of-sample performance. This procedure is called ex post forecast and differs from ex ante forecasting. The latter requires the use of forecast weather and economic conditions. Since those were not available to us, we were constrained to ex post forecasting.

specifications, based on the relationships shown in Exhibits 5-7. However, this did not out-perform the AR(1) specification and we did not pursue it any further.

Estimated separate models for each of the 12 months and allowed each month to have its individual weather responsive profile

Estimated separate models for each of the four seasons and allowed each season to have its individual weather response profile

Estimated models where the dependent variable was transformed into natural logs, thus making it easier to compare parameter values across the zones and also eliminating some types of heteroscedasticity in the error term

Explicitly tested for the presence of heteroscedasticity in the error term using two different statistical tests and derived White's robust standard errors of the parameters that correct for the downward bias in LS estimates

Estimated models with alternative weather variables

The test results are discussed in the next section. Complete results are contained in the Statistical Appendix.

1.6 LAS INTERVIEWS

Finally, we interviewed half a dozen members of the PJM Load Analysis Sub-committee (LAS), who were associated with a representative cross section of EDCs that varied in size and were dispersed throughout the region. We obtained their impressions of the PJM model specification, estimation results and forecasts. In particular, we wanted to verify that the PJM models were independent of the zonal models but producing forecasts that were generally consistent with the zonal forecasts. These interviews were carried out over the phone and ran from 30-60 minutes.

2.0 MODEL ESTIMATION CONCLUSIONS

Based on the econometric tests and experiments described in the previous section, we have come to a number of conclusions about the PJM model estimation process. These are summarized in this section. Supporting tables are contained in the Statistical Appendix.

We have established that the PJM model has been specified and estimated independently of the models being used by the EDCs. We have arrived at this conclusion based on our telephone interviews with a sample of LAS members. We have found that typically EDCs don't forecast daily peaks but instead forecast annual or seasonal peaks. They use simpler model specifications than PJM's. In addition, the EDCs do not use specifications identical to PJM's.

We have concluded that the PJM forecasts for the summer of 2006 are generally consistent with the EDC forecasts for that period. Again, this information comes from our LAS interviews. Some of those interviewed said that PJM forecasts were closer to actual 2006 numbers than their own forecasts while others felt the opposite was the case.

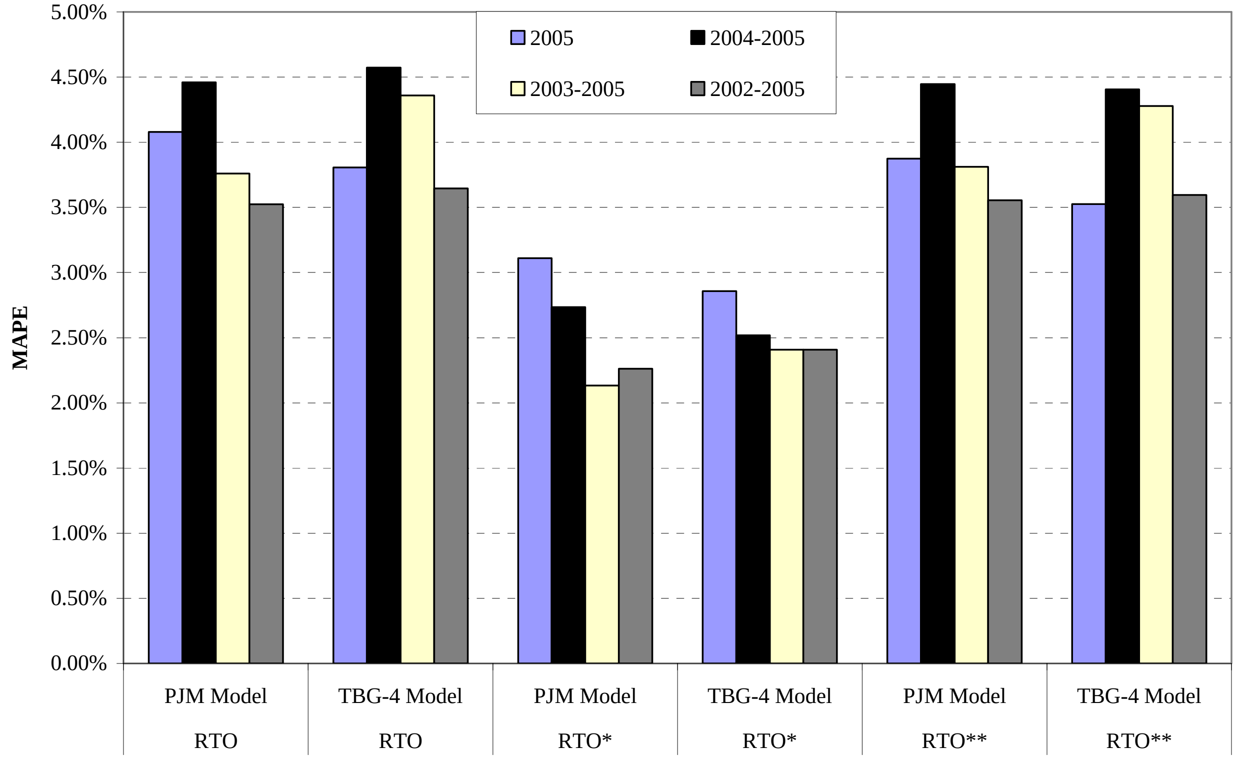
We have concluded that the PJM model specification does a reasonably good job of forecasting peak demands at the RTO and zonal levels. The mean absolute percent error (MAPE), a commonly used statistic for appraising forecast accuracy, at the RTO level for the year 2005 (comprised of 10 months) is 4.08 percent. This compares with 3.81 percent for the TBG-4 specification.

Better results are obtained at the RTO level when MAPEs are developed by adding up zonal results. We had to exclude one of the zones (RECO) since it only had data for four years. The MAPE for the PJM model is 3.11 percent and for the TBG-4 model is 2.86 percent.

When the out-of-sample MAPE is calculated for all 12 months of 2005, using the aggregate RTO model, the value is 3.87 percent for PJM and 3.53 percent for TBG-4.

Variation in results by forecast horizon is shown at the RTO level in the chart below. In general, the longer the time horizon of the forecast, the higher would be the MAPE. This pattern is not consistently observed in the chart, perhaps because the weather varies appreciably across the years, making some more recent years more difficult to forecast than some earlier years.

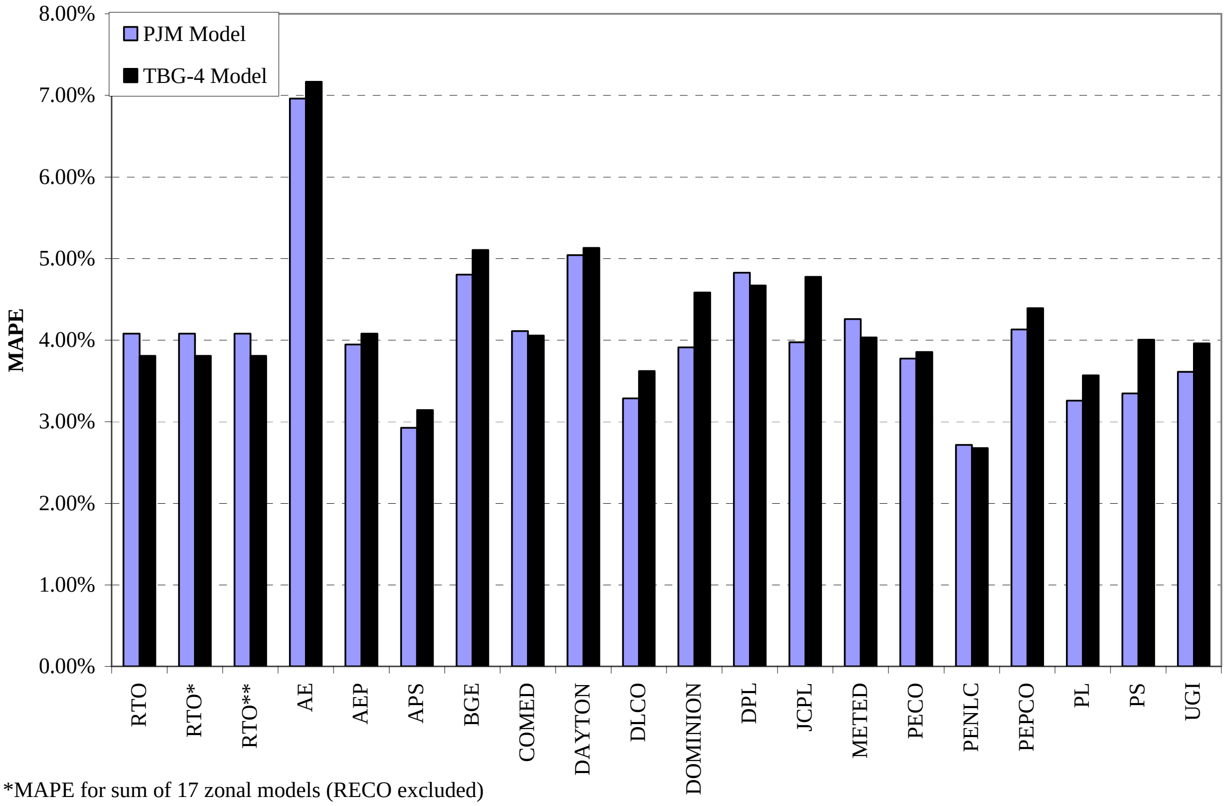
Exhibit 9
Comparison of Out of Sample MAPE for Varying Forecast Horizons RTO



*MAPE for sum of 17 zonal models (RECO excluded)
**Out of sample forecast extended to December of 2005.

Variation in MAPEs across the 18 zones for the two model types is shown in the following chart. The same general conclusions apply. However, some zones appear to have lower forecasting errors than others.

Exhibit 10
Comparison of 2005 Out of Sample MAPE for PJM and TBG Model 4



*MAPE for sum of 17 zonal models (RECO excluded)
**Out of sample forecast extended to December of 2005.

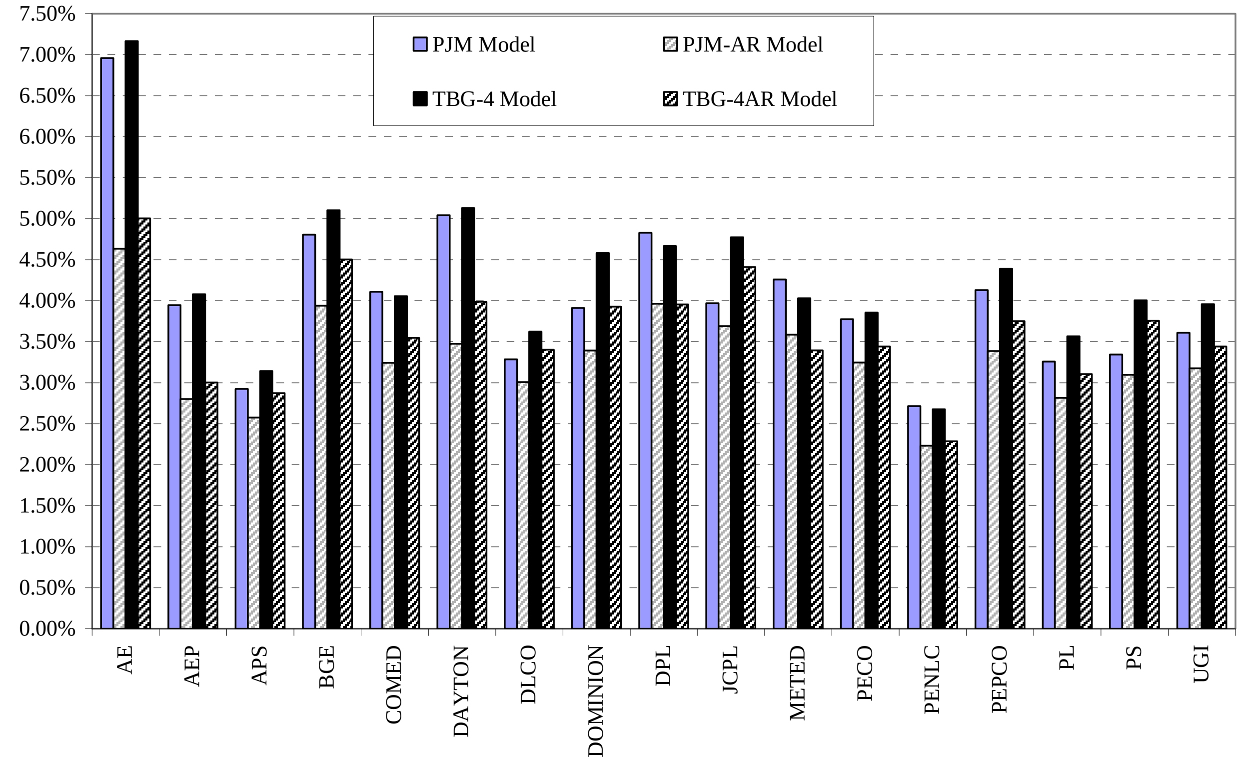
- There is ample evidence of serial correlation in both the PJM and TBG-4 models. Their Durbin-Watson statistics are closer to 1.0 than to 2.0, as shown in the regression results in the Statistical Appendix. This finding is not a surprise, since data on daily system peaks is highly trended, as seen earlier, and the independent variables in the regression specification are unable to eliminate it.

- When we ran both models with an AR(1) process, the serial correlation went away. Forecasting performance over multiple out-of-sample time horizons was markedly better with models that included an AR(1) process, i.e., when models were estimated with generalized least squares (GLS) rather than OLS.

- For the year first 10 months of 2005, the MAPE for the PJM model dropped from 4.08 percent to 2.92 percent. For all the months of 2005, it dropped from 3.87 percent to 2.70 percent. And using the sum-of-all-17 zones, the MAPE dropped from 3.11 percent to 2.28 percent. Similar improvements were observed with the TBG-4 model.

- The same results are found in the zonal models. Serial correlation is present and MAPEs improve when AR(1) process is used in the estimation. This is brought out in the following chart.

Exhibit 11
Comparison of 2005 Out of Sample MAPE for PJM and TBG Model 4 with and without AR(1) RTO Zones



*MAPE for sum of 17 zonal models (RECO excluded)
**Out of sample forecast extended to December of 2005.

- We estimated monthly and seasonal models that allowed weather responsiveness to vary across months and seasons respectively. In most cases, they improved forecasting performance compared to the annual model. The improvement in accuracy was more pronounced in 2006, which had unusually hot weather in certain months, than it was in 2005. The 2006 results are shown in the two following chars, once for monthly models and once for the seasonal models. Detailed results are shown in the Statistical Appendix.

Exhibit 12
Comparison of Monthly MAPE from Annual and Monthly TBG Model 4 with AR
RTO 2006

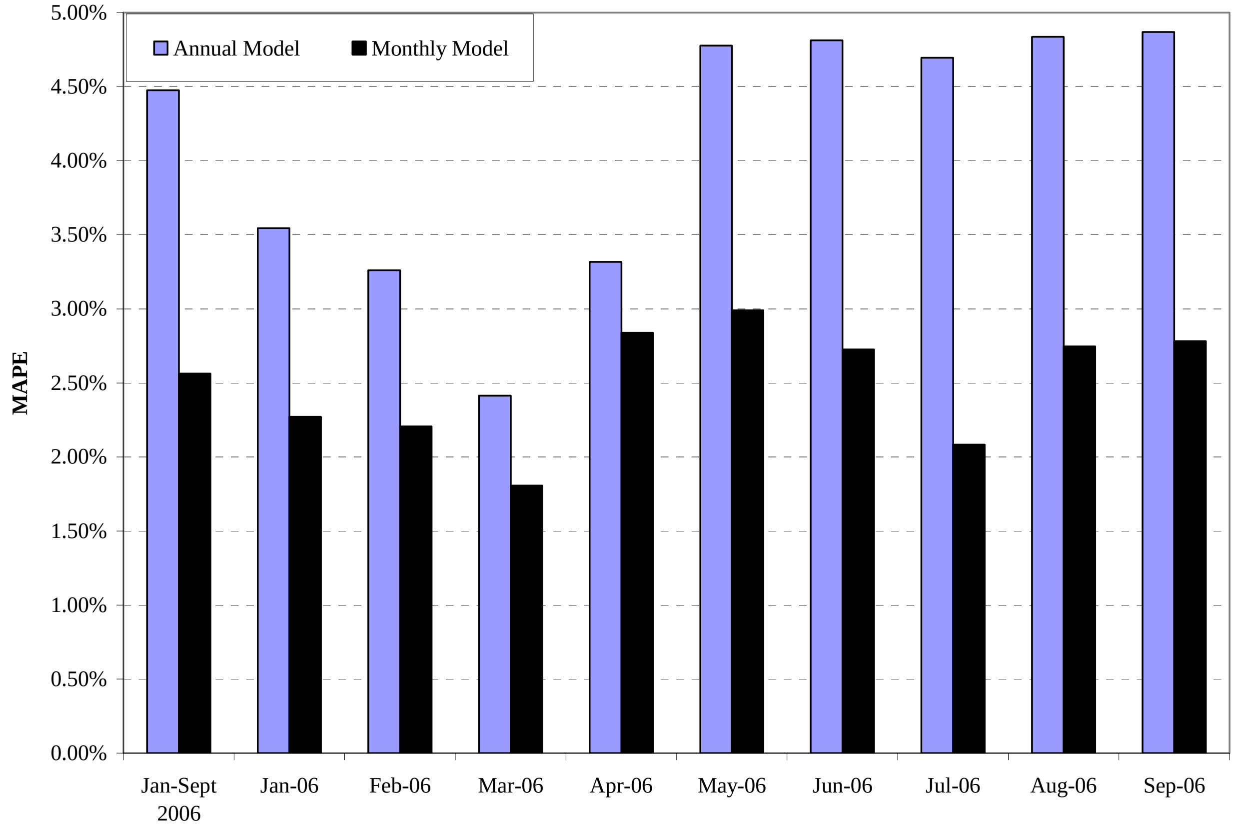
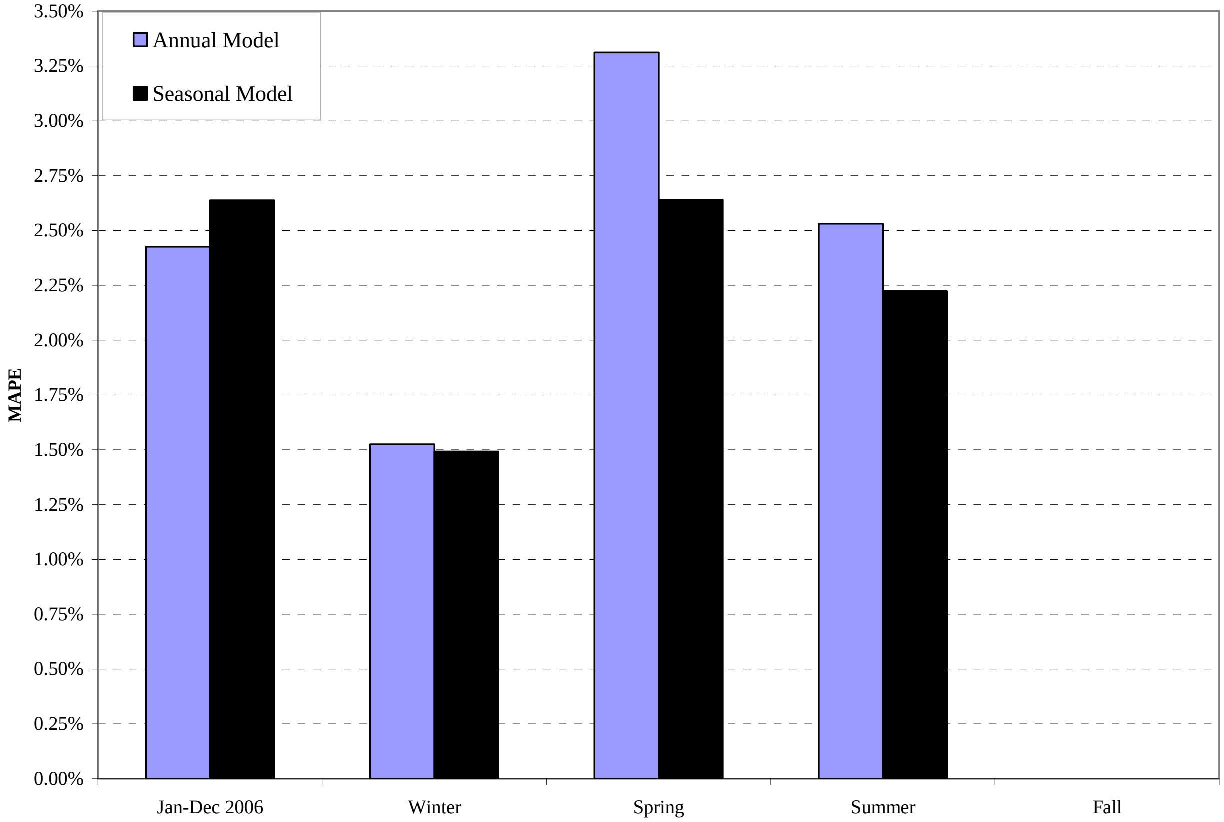
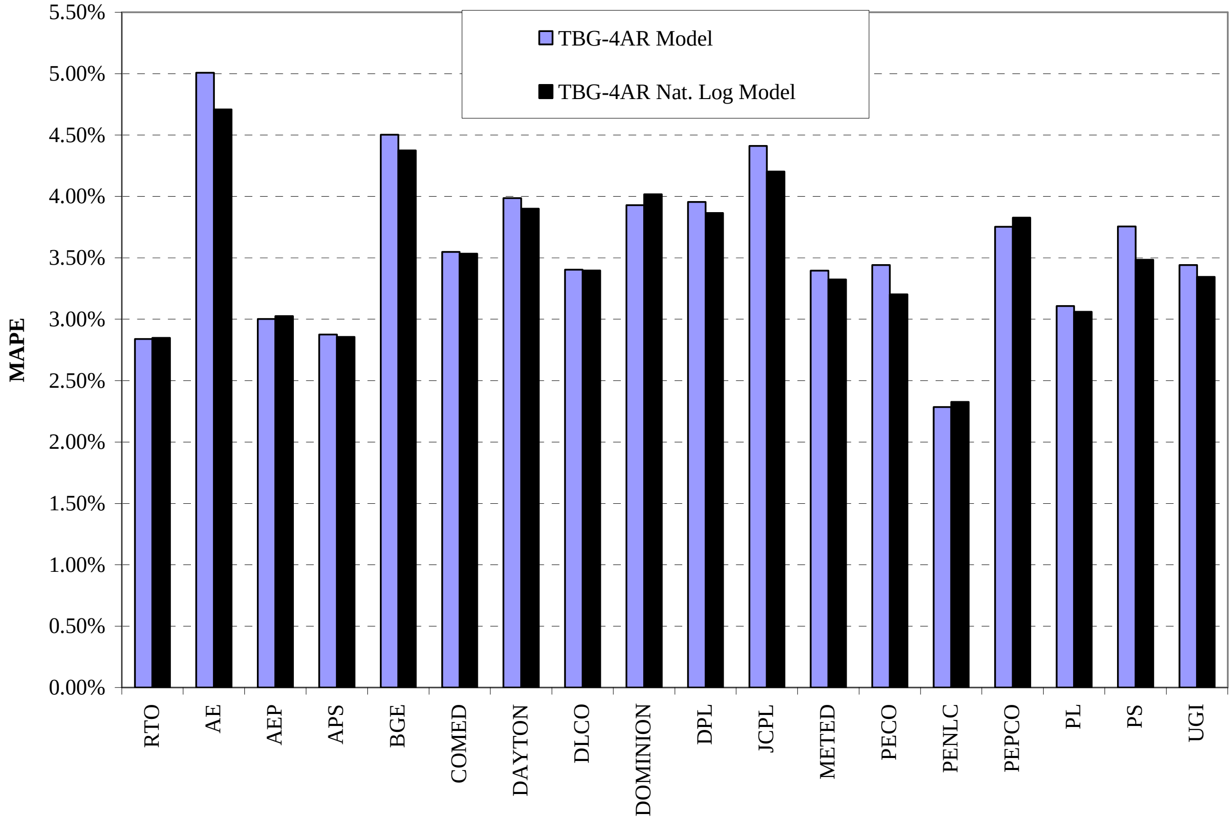


Exhibit 13
Comparison of Seasonal MAPE from Annual and Seasonal TBG Model 4 with AR RTO-2006



- We also found that transforming the dependent variables into natural logs moderately improved forecasting model performance for some zones and moderately worsened it in others. But it did make the model parameters directly comparable across the zones. Most importantly, it seems to have contributed to the constancy of the error variance. The effect on forecasting performance is shown in the following chart. The results pertain to the TBG4 model, which was estimated with and without natural logs. Additional details are contained in the Statistical Appendix.

Exhibit 14
MAPE from 2005 Out of Sample Forecast Using PJM with AR(1) and TBG Model 4 with AR(1) With Nat. Log of NCP as Dependent Variable



- The model adequately accounts for changes in weather sensitivity over time. This is accomplished by interacting the weather variables with GMP. The following chart shows how the sensitivity of summer loads to weather rises over time, for the RTO as a whole and for two of the zones.[2] While the results are shown for the TBG4 model, similar results would be expected from the PJM model. It is noteworthy that the AR(1) model showed slightly higher weather sensitivity than the regular model. We also ran tests with an alternative specification where GMP was replaced with a simple annual time trend. It yielded inferior results compared to the GMP specification.

[2] The variable being shown on the vertical axis is the incremental gain in peak demand (MW load) caused by an increment of a single cooling degree day, conditional on a given amount of GMP. It is computed as the sum of the coefficient on the CDD term and the product of the coefficient of the CDD term and GMP.

Exhibit 15
Estimation of CDD Weather Effect for TBG Model 4 with and without AR(1) RTO

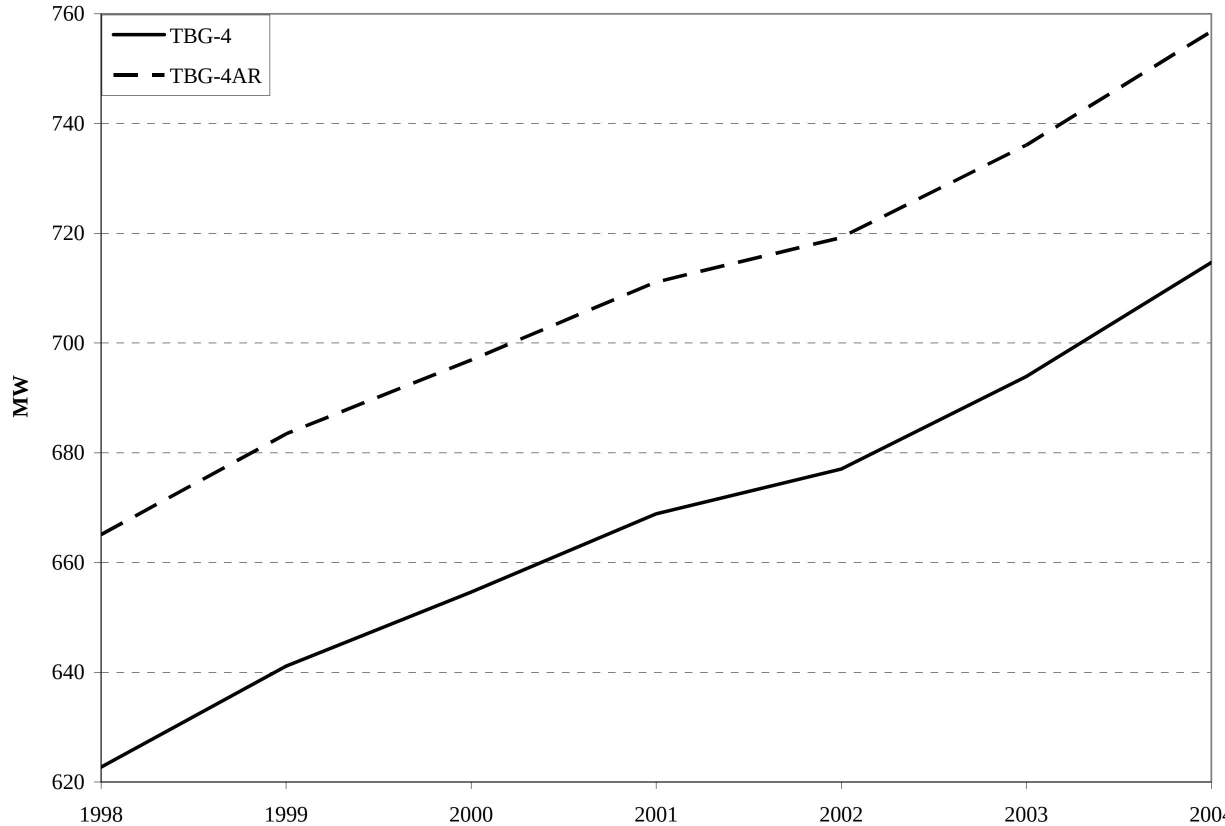


Exhibit 16
Estimation of CDD Weather Effect for TBG Model 4 with and without AR(1) COMED

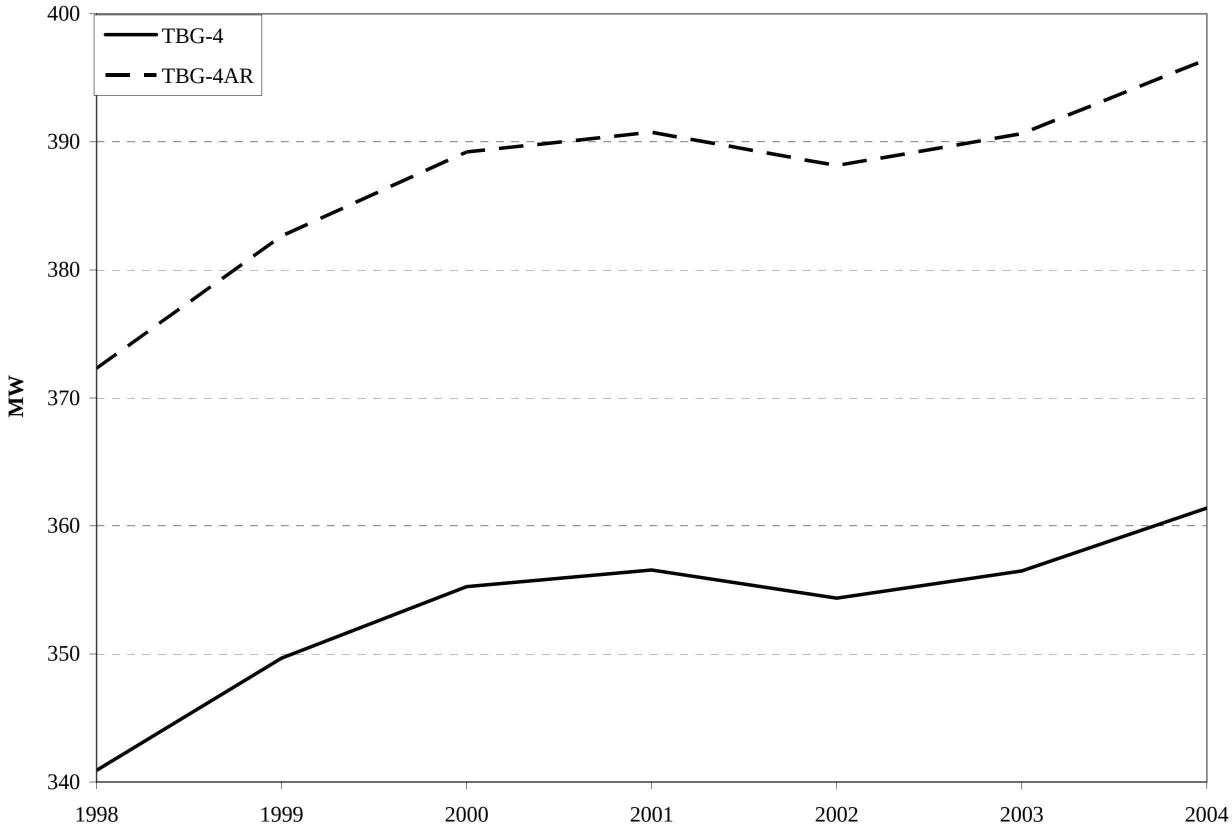
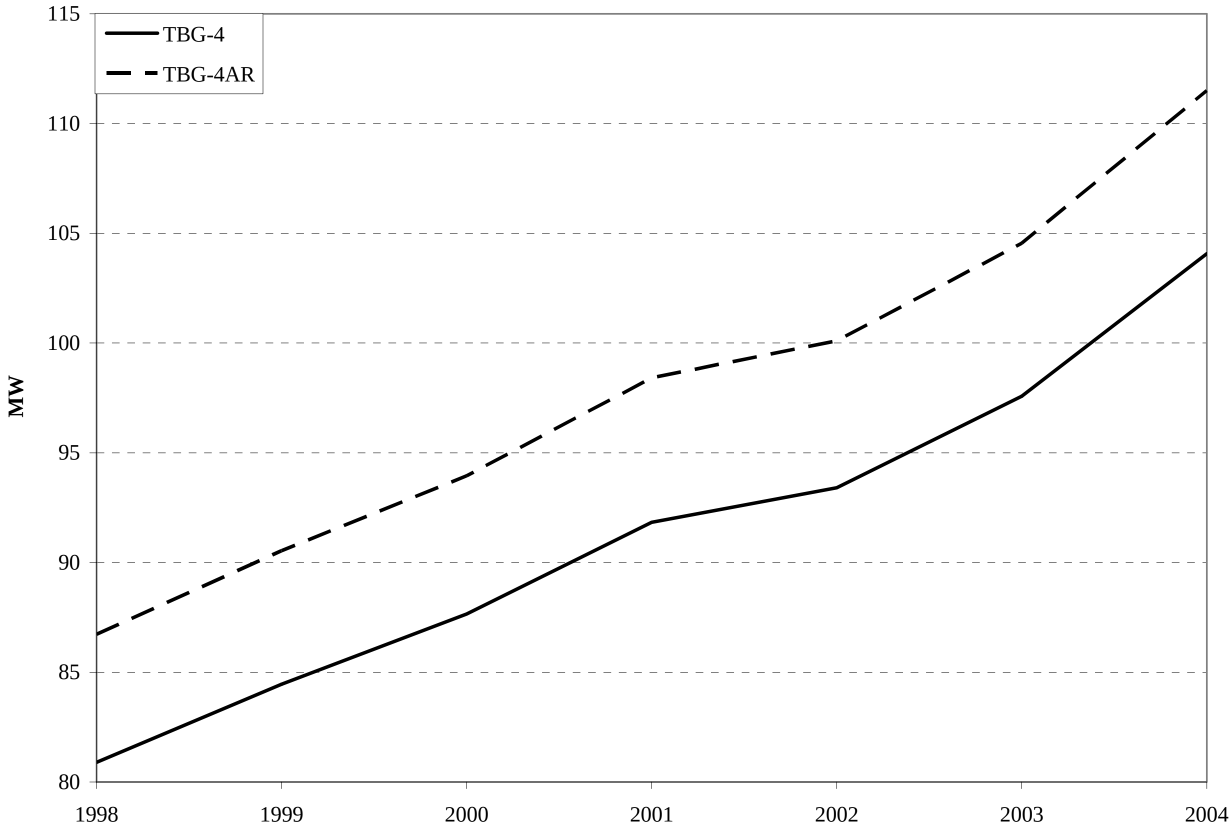


Exhibit 17
Estimation of CDD Weather Effect for TBG Model 4 with and without AR(1) PEPCO



3.0 MODEL SIMULATION AND FORECASTING CONCLUSIONS

3.1 ANALYSIS OF THE 2006 PJM FORECAST

It was noted earlier that the PJM official forecast for the year 2006, made in February 2006, fell short of the actual RTO peak that was observed on August 2, 2006. Specifically, the forecast of coincident peak demand was 133,500 MW and the actual coincident peak demand was 146,000 MW. This represents a forecast error of -9.36 percent.

This discrepancy could be due an inherent bias in the forecasting model, e.g., an inability to handle extreme weather conditions, and/or difficulty in forecasting weather, since unusually hot weather was observed on that day. A less likely reason is difficulty in forecasting the pace of economic activity.

In order to isolate the main cause of the problem, we asked the PJM staff to simulate their model with actual weather and actual economic conditions. When they ran the model under these conditions, the forecast demand came in at 147,000 MW, within 0.7 percent of actual, a very acceptable margin of error. This is especially remarkable since extreme weather conditions were encountered in 2006 and the model was able to model their impact on peak demand quite well.

Our conclusion is that the model is doing a good job of forecasting peak demand and the main source of error is weather. Details are shown in the following exhibit.

Solving the Models with Actual Weather

	Summer 2006	% Diference
2006 Load Report RTO Median(50/50) Forecast	133,500	
2006 Load Report Sum of Zones NCP Median(50/50) Forecast	137,400	
Metered RTO Peak - 8/2/2006	144,644	
Unrestricted RTO Peak - 8/2/2006	145,951	
Unrestricted Sum of Zone NCPs - Summer 2006	146,401	
RTO Model	146,438	0.33%
RTO Model	147,010	0.73%
Sum of Zone NCP Models	148,065	1.14%
Sum of Zone NCP Models	149,039	1.80%
Solved @ Actual Weather 8/2/2006 (100 Percentile) and 10/2005 Economic Forecast (50 Percentile)		
Solved @ Actual Weather 8/2/2006 (100 Percentile) and 9/2006 Economic Forecast (51Percentile)		

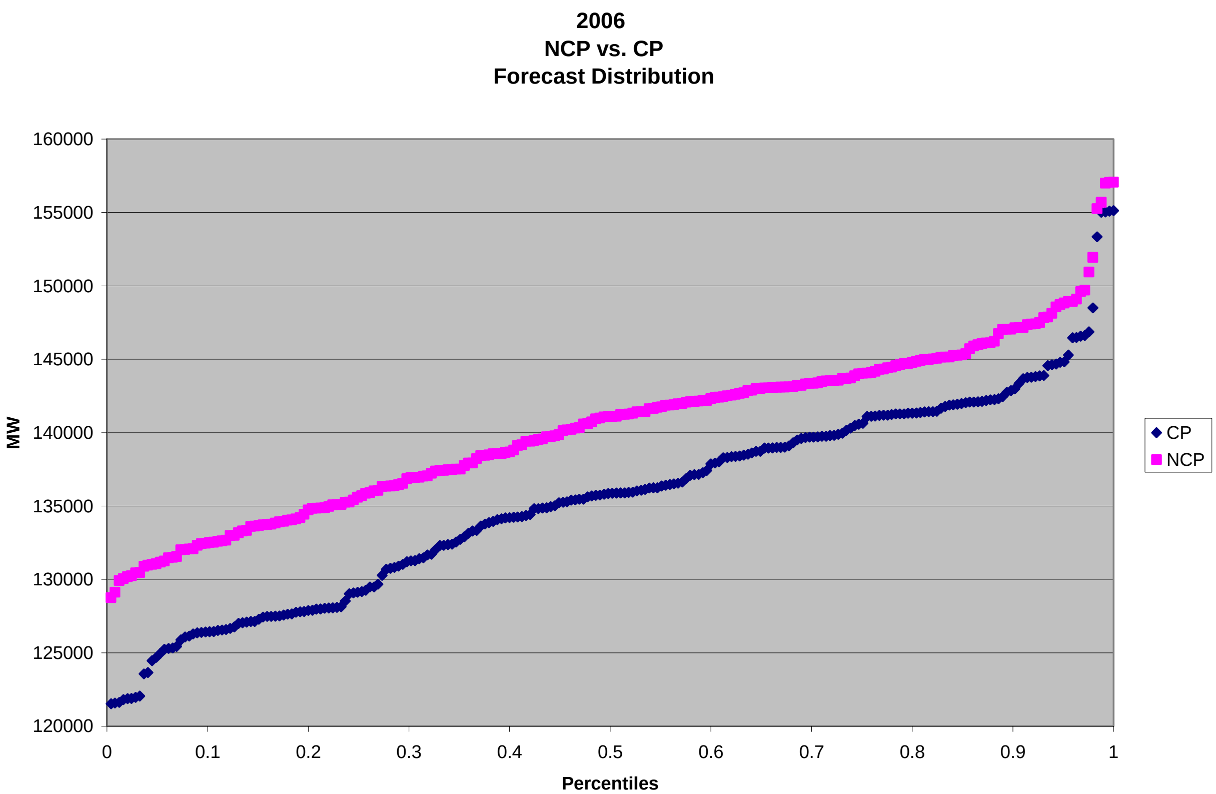
3.2 ANALYSIS OF THE DIVERSITY FACTORS

One of the issues identified by PJM staff at the kick-off meeting was the issue of modeling diversity across the zones. There was a concern that the existing approach kept diversity constant across alternative weather years. This issue also came up in discussions with LAS members. Some members felt that diversity had vanished on August 2, 2006 while the PJM model held diversity constant based on historical trends.

As part of a list of on-going enhancements PJM staff developed a series of zonal CP models, with the same specification for the independent variables as the zonal NCP models, but using the zonal CP as the dependent variable. The zonal NCP and CP models were solved with the weather simulations yielding a more accurate relationship between diversity and temperature conditions. This is shown in the following chart. As temperature rises, diversity (measured as the percent difference in non-coincident peak demand and coincident peak demand at the RTO level) declines.

Using this new approach to diversity will raise the median RTO forecast for 2006 by 825 MW, or by 0.62 percent from 133,500 MW.

Exhibit 18
2006/NCP vs. CP/Forecast Distribution



3.3 ANALYSIS OF SEASONAL VERSUS MONTHLY WEATHER HISTORIES TO GENERATE ANNUAL PEAKS

In our first meeting with PJM staff, they told us that they were planning to review other methods of simulating weather histories. This issue also came up in discussions with LAS members. Some felt that using the month of July to simulate weather conditions as input to the development of the annual peak forecast would tend to understate it. They argued that a seasonal approach should be undertaken to simulating weather conditions, where the summer season is defined as running from June to August.

Again as a part of the list of on-going enhancements, the PJM staff processed the solved weather simulations to obtain zonal seasonal peak demand instead of zonal monthly peak . Using a seasonal approach, the median RTO forecast for 2006 rose by 0.93 percent or 1,251 MW to 134,325 MW.

When combined with the diversity adjustment in the previous section, this yielded a new median RTO forecast of 135,576 MW, or an increase of 1.56 percent.

4.0 RECOMMENDATIONS

In this section, we present recommendations relating to model specification, simulation, forecasting and other issues.

4.1 MODEL SPECIFICATION

Serial Correlation. PJM should adjust for serial correlation by applying a first-order autoregressive process to the data. As shown earlier in this report, out-of-sample forecasting performance improved when this adjustment was applied in all cases. While this does not guarantee success in forecasting, since weather will always be difficult to forecast, it does ensure that the best possible model specification is being used. In addition, the estimated standard errors will not be biased in a downward direction, creating in some cases the illusion of statistical significance.

Seasonal or Monthly Models. PJM should replace its annual model specification with a monthly or seasonal specification, since the weather sensitivity coefficients vary appreciably by month and month. Holding them constant across the year causes both lack of fit and loss of forecasting accuracy.

Log Transformation. PJM should investigate the possibility of transforming the dependent variable into natural logs. This would minimize any heteroscedasticity that is present in the error term and may also improve forecasting performance. In addition, it would make it easier to compare parameter values across the zones.

Consider alternative weather variables. There is reason to believe that the weather effects during the summer may be nonlinear. This effect could be captured by breaking the current weather variables into pieces, so to speak, and permitting each of these new variables to have a separate coefficient. This would provide a linear approximation to any potential nonlinearities in the weather response function. In addition, PJM may wish to consider modifying the base from which cooling and heating degree days are defined.

Parsimony. PJM should consider using a parsimonious specification than one that uses 55 independent variables. As shown in the previous section, we were able to get MAPEs that were almost as good with a simpler model specification (TBG-M4).

Joint estimation. Explore joint estimation of the 18 zonal models to improve the efficiency of estimation using Zellner's seemingly unrelated regression (SUR) procedure adjusted for serial correlation (Parks' procedure). It may be useful to allow each of the zones to have its own intercept term in the regression, i.e., we recommend using the "fixed effects" model. The Zellner procedure also allows tests to be made on whether certain key parameters, such as the weather response ones, are similar across the RTO or within sub-regions.

4.2 MODEL SIMULATION AND FORECASTING

Correlating diversity with weather conditions. PJM should allow for diversity to vary inversely with the weather. Otherwise forecasts will be understated under extreme weather scenarios, when diversity is much less than at average weather conditions.

Using seasonal data to develop annual peak forecasts. PJM should develop weather distributions by aggregating the data over the summer reason rather than just relying on the month of July. Otherwise forecasts will be understated even under normal weather scenarios.

Using alternative weather histories. The traditional practice has been to rely on a 30-year history. This has been challenged by those who believe that greater greenhouse gas emissions are triggering climate change. It is unclear whether PJM's climate has changed. Our analysis of PJM's nighttime temperatures did not reveal any discernible trend in those temperatures. In addition, we did not find any change in precipitation in the PJM footprint, another potential effect. Thus, there is no strong supporting evidence to argue for local greenhouse gas induced meteorological change. However, recognizing the merits of the debate, we recommend that PJM use two weather histories: a long-one and a one short-one.

4.3 OTHER ISSUES

New technologies. Since the model is being used to make fairly long-term forecasts, it is advisable to introduce additional features in the specification that would allow impacts associated with new technologies (such as server farms, plug-in hybrids and HDTVs) and dynamic pricing programs to be measured.

Consistency with energy and load shape forecasts. To ensure consistency in resource planning, PJM should determine that its peak demand forecast is generally consistent with forecasts of energy consumption being made either by the EDCs or by PJM. The same applies to hourly load shape forecasts.

Structural changes. It will be important to address the issue of structural changes in the economy, such as shifts of employment from manufacturing to services, changing the housing mix toward larger houses and/or fewer multi-family units, rising levels of equipment efficiencies and so on. These will change the structure of load shapes and the growth rate of peak demand.

Bootstrapping. To quantify forecast uncertainty, it may be useful to consider the use of "bootstrapping" methods pioneered by D. Freedman. The bootstrap relies upon the empirical distribution function to derive estimates of uncertainty. In the context of the PJM forecasting models, one would derive the variability, i.e. uncertainty, in the non-coincident peak demand models, then use that result in forecasting peak demand. Let a non-coincident peak demand model for an arbitrary EDC zone z be denoted as

$$NCP_{zt} = constant_z + \underline{a_z} * \text{economic variables}_{zt} + \underline{b_z} \text{ weather variables}_{zt} + error_{zt}$$
$$\text{for } t= 1,\ldots, T$$

where $\underline{a}$ and $\underline{b}$ are vectors of parameters associated with the economic and weather variables, respectively. Let $e_{z1}, \ldots, e_{zT}$ denote the residuals from an ordinary least squares estimation of this equation with estimates constant^*_z, $\underline{a}^*_z$, $\underline{b}^*_z$. To bootstrap this model, for each of the observations of the estimated design matrix, i.e., $\text{constant}^*_z + \underline{a}^*_z$ * economic $\text{variables}_{zt} + \underline{b}^*_z$ * weather variables_{zt} add a draw from the residuals vector with replacement. This means that, for example, e_{z3} may be combined with the 10th and 21st observations, while some residuals may not be combined with any. This process creates a vector of NCP observations which, with the original independent variables can be used to produce new parameter estimates which, for the first replicate, we denote as $\text{constant}^{(1)*}_z$, $a^{(1)*}_z$, and $b^{(1)*}_z$. Let us collectively call the parameter vector $\underline{\text{parameters}}_z$, the initial estimate, $\underline{\text{parameters}}^*_z$, the n^{th} bootstrap replicate estimate $\underline{\text{parameters}}^{(n)*}_z$. Suppose we replicate this procedure N times. Let $\underline{\text{parameters}}^{(ave)*}_z$ be the average across the N replications. Then

$$(1/N)\sum_{n=1}^{N} \left(\underline{\text{parameters}}^{(n)*}_z - \underline{\text{parameters}}^{(ave)*}_z \right) \left(\underline{\text{parameters}}^{(n)*}_z - \underline{\text{parameters}}^{(ave)*}_z \right)^{\text{transpose}}$$

is an empirical estimate of the variance-covariance matrix of the parameters. The initial estimate and the associated bootstrap variance are normally distributed in samples as large as these.

The next step is to use the historical weather record in combination with the bootstrapped coefficient estimates to produce a distribution of potential non-coincident peak demands. Effectively, the goal is to convolve the bootstrapped distribution of coefficient estimates with the weather vectors. In the case of the weather, PJM must consider the weather that it uses for a particular date a single observation across all EDCs. Given the daily nature of the current weather data used in the model, PJM should consider using only selected percentile points of the implied distribution, say, the 5th, 10th, 50th, 90th, and 95th, as the basis for combining with the weather data, rather than a larger proportion of the distribution.

Scenario planning. It may be useful to also examine other ways of quantifying the uncertainty in the forecast. The current approach, where different weather histories are brought in to bound the uncertainty in the forecast, is equivalent to simulating different outcomes in a given state of the world. Over the long run, the state of the world itself becomes uncertain. An approach such as scenario planning would be able to characterize different states of the world within which alternative peak demand forecasts would exist.

Communications. Finally, given the increasing interest in load forecasts, it may be useful to explore new ways of communicating the forecast internally and externally to the various stakeholders.