US PowerGen

Report to the NYISO Load Forecasting Task Force

Observations and Recommendations

Prepared by Dr. Howard J. Axelrod on behalf of US PowerGen
4/15/2011

Table of Contents

Overview

US PowerGen (USPG) was an active participant in the 2010 Load Forecasting Task Force (LFTF) development of the official coincident peak load forecast for the year 2011. USPowerGen retained Dr. Howard Axelrod, an associate of M. J. Beck Consulting, LLC, to participate in the LFTF meetings on behalf of the company, to review each of the key processes and to offer suggestions as to ways that might improve the forecasting process.

Our primary objective as we stepped-up our involvement in the LFTF process was to listen, learn and hopefully contribute to the dialogue surrounding the enhancement of the LFTF process to better reflect more sophisticated econometric considerations, as well as other market conditions. Clearly, the higher the load forecast, the higher the ICAP and the higher the value of the capacity credit. However, while we tested many of the assumptions used to establish the 2011 coincident peak load, we made no effort via our participation to sway the process to drive a higher than expected peak demand for Zone J (i.e., the location of USPG's facilities), but rather to independently vet the process by which the final values were derived.

We recognize that the LFTF is an inclusive group made up of representatives from various constituencies, such as utilities, municipal authorities, independent power producers, large energy consumers and state regulatory staff. As a result, we found that the LFTF process, as explained to us by Arthur Maniaci of the NYISO staff, evolved over a number of years in an effort to reflect the inputs, positions and individual interests of each participant. Generally, we found the LFTF process to be refreshingly open to new participants, largely due to the willingness of Mr. Maniaci, who spent considerable time and effort to bring us up the learning curve and address each of our questions and concerns. As a result of our efforts, along with others, the LFTF has scheduled a follow-up meeting in April 2011 to review the current peak load forecasting process in order to solicit comments and suggestions for improvement, which might be incorporated in the next round of LFTF meetings scheduled for mid-2011. This report presents the observations and recommendations that we believe could serve as a basis for a constructive dialogue during that meeting to determine where and to what extent the LFTF process could be improved.

<u>**Observations**</u>

Key observations are described below. We will later offer what we believe are a reasonable set of recommendations to enhance the process in the future.

Observation 1: We found that the LFTF process provided all participants an opportunity to follow each step of the load forecast process including the development of weather normalized peak loads, the adjustments for load management and system losses, and the quantification of the three criteria used to assess each transmission owner's ("TOs") load growth forecast. However, it must be recognized that the coincident peak loads, are, in fact, developed by the TOs. There is no review, discussion or vetting of the TOs' short-term load forecasting methodology – the only information provided to the LFTF participants is the actual current year coincident peak load and the TO's projected one-year growth rate. A description of the forecast models, assumptions and historical databases are not disclosed to the participants.

Observation 2: The pursuing role of the NYISO staff, along with other participating parties in the LFTF process, is to assure that the TO's one-year load forecasts fall within the boundaries of at least two of the three criteria described in the NYISO's Load Forecast Manual. The three criteria used to approve the TOs' one-year peak load growth rate are intended to establish boundaries within which the TO's load growth rate must fall within two of the three, in order to achieve a state of acceptance by the NYISO. The three criteria are:

- A lower and upper boundary of annual peak load growth rates based on the prior five year annual load growth rates.
- A lower and upper boundary of ratios of annual historical peak load growth rates relative to the annual historical growth rate of one or more economic drivers.
- A statistically derived confidence interval of historical annual energy growth rates based on a Monte Carlo model developed by the NYISO staff. During the 2010 LFTF process an additional economic driver – state gross domestic product – was also included in the model.

We believe that the three criteria are not three truly independent perspectives from which to evaluate the TO's peak load forecasted growth rates. Furthermore, we believe that the methods used in at least two of the three criteria (i.e., Criteria 1 and 3) by which the upper and lower boundaries were derived failed to reflect the true confidence bands by which the historical data might demonstrate. Finally, we found that Criterion 2, the economic test, lacked substantive theoretical economic basis as the ratios were:

- Assumed to retain a constant ratio of coincident annual peak load growth to that of several tested economic drivers;
- Evaluated and then weighted by the relative degree to which each economic driver was correlated with peak load growth; and
- Derived on the assumption that the weighted ratio was a surrogate to a true consumption function which relates energy consumption to such exogenous socio-economic factors as growth in income, employment or households.

> **Observation 3:** Using the current methodology, the TO's projected growth rate will be used to derive the ICAP value as long as the projection falls within the boundaries of two of the three criteria. Given that two of the three criteria (Criteria 1 and 3) rely heavily on historical growth rates of prior yearly peak demand and energy consumption, only the second criterion exclusively addresses forecasted economic recovery. Thus, a TO's forecast, which may be developed using historical trends (an assumption that cannot be validated), could easily pass two of the criteria and yet be substantially below expected growth trends predicated on projected economic recovery. *To illustrate, if the TO's forecast is zero percent growth rate, and two of the three criteria suggest a plausible range of, say 0% to 3%, but the third criteria has a range of between 2% and 4%, the TO's 0% forecast would still be adopted for the ICAP* computation as it "passes" two of the three LFTF criteria. Yet, there is a preponderance of information developed during the LFTF process to suggest that the adopted growth rate should fall somewhere between the lower and upper boundaries. More importantly, as it currently stands, the ICAP committee would have no way of *knowing this information and/or considering it in its computation of the ICAP values.*

The recent economic recession, followed by the projected recovery over the next few years, could provide conditions where an underestimation of projected load by a TO might not be identified.

Recommendations

Based on these observations, the following is a summary of our primary recommendations. Subsequent sections expand upon the rationale supporting each and provide additional suggestions as to how the articulated concerns might be remedied.

Recommendation 1: Early in the LFTF process each TO should provide a detailed description of its load forecasting methodology, including data sources and forecast assumptions that it will be using to develop the one-year load forecast growth rate. Each forecast should also be presented in a form that provides the derived confidence band along with the base forecast.

Recommendation 2: The LFTF should perform a detailed review of the current methodology. For each criterion, the LFTF should immediately begin to assess the validity in terms of degree of independence, ability to accurately assess a reasonable range of forecast uncertainty, and concurrently balance the momentum of historical trends against the structural changes caused by socio-economic and technological shifts. In addition, rather than evaluating load in aggregate, the LFTF should consider a more “bottom up” approach. For example, load forecast could be evaluated by customer class (retail, commercial, industrial, etc.) or by certain usage characteristics.

Recommendation 3: The LFTF should evaluate additional analysis that would provide insight into the evaluation of a TO’s load growth projections. The LFTF could explore the development of a composite probability distribution (vs. an expected value or a banded approach) or a set of scenarios for load forecasts. For example, it might be insightful to develop a load forecast scenario based on the upper range of the forecasted economic drivers developed by the NYISO’s independent economic consultant.

Detail - Recommendation 1

Recommendation 1: *Early in the LFTF process each TO should provide a detailed description of its load forecasting methodology including data sources and forecast assumptions that it will be using to develop the one-year load forecast growth rate. Each forecast should also be presented in a form that provides the derive confidence band along with the base forecast.*

We found that while the LFTF process provided all participants an opportunity to follow each step of the load forecast process, the ultimate role of the non-TO participants was to vet the development of the weather normalized adjusted current year coincident peak load and the three criteria used to evaluate the TO's one-year load growth rate. We also found that little information was offered to the participants by the TOs as to the methodology, data sources and assumptions used to develop the one-year growth rate.

While the ideal situation might be to have the NYISO staff oversee the development of an independently derived forecast for each region with all stakeholders having equal opportunity to contribute to the process, this seems impractical from a time and resources perspective, particularly given the fact that the TOs must develop similar forecasts for other planning and regulatory requirements. However, we find that the current situation fails to provide enough information and opportunity to evaluate each forecast.

As a compromise suggestion, we propose that for the upcoming forecast of 2012 coincident peak loads, each TO presents to the LFTF, early in the process, its load forecasting methodology, data sources and forecast assumptions. Furthermore, the TO's peak load forecast should be presented as both a mean value and some form of statistical assessment of level of confidence. Such presentations could include (these also relate to our Recommendation #3):

- A high/low bandwidth based on 90% confidence
- The derived standard deviation
- A probability distribution representing the range of possible outcomes.

Detail - Recommendation 2

Recommendation 2**:** The LFTF should perform a detailed review of the current methodology. *For each criterion, the LFTF should immediately begin to assess the validity in terms of degree of independence, ability to accurately assess a reasonable range of forecast uncertainty, and concurrently balance the momentum of historical trends against the structural changes caused by socio-economic and technological shifts. In addition, rather than evaluating load in aggregate, the LFTF should consider a more "bottom up" approach. For example, load forecast could be evaluated by customer class (retail, commercial, industrial, etc.) or by certain usage characteristics.*

We believe that the three criteria are not three truly independent perspectives from which to evaluate the TO's peak load forecasted growth rates. Criterion 1 is based on historical annual peak load growth rates in weather normalized peak demand, while Criterion 3 is based on a trend of historical summer electric usage as measured by GWh. By using a Cooling Degree Day variable it, in effect, weather normalizes summer usage. As a result, peak loads and energy consumption follow similar trends as long as the system load factor remains relatively constant, which is typical over a short period of time.

We also found that the methods used in at least two of the three criteria (i.e., 1 and 3) by which upper and lower boundaries were derived failed to reflect the true confidence bands that historical data might demonstrate. For Criterion 1, the upper and lower boundaries were the single year growth rates for the second highest and second lowest annual growth rates derived from the previous six years of peak load data. Technically, the actual trend and volatility of annual percentage changes in coincident peak demands were ignored as the Criterion 1 boundaries were set by eliminating the lowest, middle and upper yearly growth rates.

For Criterion 3, the methodology for establishing the upper and lower boundaries was based on a two-stage model developed by the NYISO staff. First, a regression analysis was used to derive the coefficients that relate summer energy usage with a weather variable (cooling degree days (CCD)) and an economic variable (gross domestic product (GDP)). The forecasted summer usage was then derived using a Monte Carlo model into which the historical distribution of CDD and GDP (actually the difference between the actual GDP and its derived trend were used to estimate random GDP "shocks"[1]) were input. Comparing the medium forecast to the 2010 weather adjusted Summer GWh computation was used to derive an implied growth rate. The upper and lower boundaries were then derived by assuming a normal distribution and the boundaries based upon a +/- 25% interval.

There are two points that require discussion regarding the derivation of the upper and lower boundaries. First, it is assumed that the probability distribution of the derived summer usage value is normally distributed. This may be a poor assumption as the 30-year CDD mean value may actually be displaying a rising and more volatile pattern, and the GDP has no basis for being normally distributed as long term trends have generally been rising, interrupted only by periodic recessionary periods. Second, the Monte Carlo model also derives a standard deviation based upon the 500 simulations used to estimate the summer usage distribution. Instead of an arbitrary +/- 25% interval, a high low band based upon, say, 90% confidence could also be derived directly from the Monte Carlo model.

Finally, we found that Criterion 2, the economic test, lacked substantive theoretical economic basis as the ratios were:

- Assumed to retain a constant ratio of coincident annual peak load growth to that of several tested economic drivers,
- Evaluated and then weighted by the relative degree to which each economic driver was correlated with peak load growth, and
- Derived on the assumption that the weighted ratio was a surrogate to a true consumption function which relates energy consumption to such exogenous socio-economic factors as growth in income, employment or households.

Here again we found a few items worthy of discussion with regard to the computation of Criterion 2. First, the use of a correlation coefficient to serve as a weighted value in deriving the composite economic driver makes little sense. The correlation coefficient measures the degree to which one variable explains the variation of another. Two variables are perfectly correlated if the coefficient is 1.0. The lower the coefficient, the less the relationship can be explained between

[1] See Criterion 3 presentation by NYISO staff member Arvind Jaggi, December 3, 2010

the two variables. In other words, for a correlation coefficient of 0.5, 50% of the variation can be explained by the selected variable and 50% explained by other, unknown factors. From a practical perspective, a correlation coefficient of less than 0.7 or 0.8 reflects a poor method for predicting the value of one variable based on the value of another. Too much information remains unknown; hence the resultant forecast is unreliable. To use the derived correlation coefficients as a relative weighting scheme assumes that the lower the coefficient the less the relative value of that variable. However, this is an incorrect assumption as the relationship has no value as this statistical test demonstrates no meaningful relationship when its value falls below a reasonable range.

Second, the computation of a ratio of the peak load growth rate to the composite economic growth rate suggests a correlation that was not demonstrated by the NYISO staff's own analysis. Basically, the LFTF is using very low correlation coefficients between each economic driver and peak load growth to derive a weighted economic indicator that is then used to develop a ratio between the load growth rate and the weighted economic indicator. The 2011 peak load growth was then computed using the forecasted set of economic drivers. However, the relationship between the economic drivers and the annual growth in peak demand was not valid to begin with and its weighing does not negate the lack of statistical significance.

Third, the construction of the upper and lower boundary based on the selection of the second highest and second lowest ratio ignores the information that can be deduced from the total trend. Because only five years of data is used, it is difficult to assess the mean and standard deviation of the economic ratios. One approach that the LFTF might consider is to explore developing a regression model that equates annual summer usage to a set of statistically valid independent economic drivers. Then this model could then be used to forecast the following year's summer usage, which could then be converted to a coincident peak demand using an expected load factor. The upper and lower boundaries can be then derived by assessing the variation demonstrated from the historical trend as well as the uncertainty expected from the forecasted economic drivers.

Detail - Recommendation 3

Recommendation 3: *The LFTF should evaluate additional analysis that would provide insight into the evaluation of a TO's load growth projections. The LFTF could explore the development of a composite probability distribution (vs. an expected value or a banded approach) or a set of scenarios for load forecasts. For example, it might be insightful to develop a load forecast scenario based on the upper range of the forecasted economic drivers developed by the NYISO's independent economic consultant.*

While it is not the role of the LFTF to assess the impact of coincident peak load growth on electric system reliability, the submission of the TO's forecast (assuming it meets 2 of 3 criteria), provides the ICAP committee with limited information to assess and assign an ICAP value. Recommendation 3 would provide the ICAP committee with a wider range of information such that they could consider probability distributions or scenarios in addition to the expected values traditionally provided by the TOs.