Proposed Incentive Goal #3: Load Forecast Accuracy

The Problem with 2002 Incentive Goal #3

The 2002 load forecast accuracy goal (incentive goal #3) measures the average of the Mean Absolute Percentage Errors (MAPE) of four zones, or superzones, in the New York market: West (zones A – E), East (zones F – I), J, and K. Everyday, the MAPE for each of these zones is calculated. The four MAPEs are added and divided by four to determine the average MAPE for that day. At the end of the year, the daily average MAPE will be accumulated and divided by 365 to find the annual average MAPE. This will be the metric for determining performance on incentive goal #3. An average annual MAPE of 2.8% is required for the threshold, 2.5% for the target, and 2.3% will be considered superior.

Through July, the MAPE stood at 3.28%. Reaching anything other than the threshold is practically impossible. Even that will require August – December accuracy of 2.18%, accuracy that has not be met in any month of 2002.

Throughout 2002, ISO staff has implemented a number of model improvements that have improved forecast results, but the results have not been sufficient to drive the actual MAPE to within sight of the incentive goal levels. Additional, more radical modeling changes are possible, but the price might be poorer accuracy in forecasting the NYCA peak itself. To see why this is so, a brief description of the day-ahead forecast model (DAFM) is required.

The DAFM produces a forecast of the NYCA peak load and loads for every other hour of the day. It also produces forecasts of the same concepts for each of the 11 NYCA zones. The NYCA forecast is produced independently of the 11 zonal forecasts (which are produced independently of each other, although a common set of weather variables, etc. is employed). The independent zonal forecasts are reconciled to the NYCA forecast by ratioing them up or down so that their total equals the NYCA forecast in each hour of the projected day. Therefore, in the final forecast, the sum of the load in each zone will equal the NYCA load for every hour of the day.

The most likely place to look for improvements in zonal forecast accuracy is to model each zone independently and then form the NYCA forecast by adding up the 11 zonal forecasts. This would eliminate the independent NYCA forecast and the step of reconciling the sum of the independent zonal forecasts to it. Doing this would allow enhancements in the zonal forecasting models to flow directly to the zonal forecasts, rather that being constrained by an independent NYCA forecast.

There are two serious problems with this approach, however. First, the load at the NYCA level contains less measurement error than the loads for the individual zones. The new approach would put more weight on the poorer data and less on the better data than the current approach does. Second, there is no guarantee that when the independent zonal loads are added up that the resulting NYCA load will result in an accurate forecast. In fact, it is almost certain that it will be a poorer forecast of the NYCA peak (and all NYCA hourly loads) than that presently produced by the independent NYCA models. So, gains in zonal forecast accuracy will come at the expense of NYCA forecast accuracy.

Is this trade-off desirable? Based on the expressed concerns of market participants, the answer is no. Every complaint about forecast accuracy has been about the NYCA peak. Not one has been about the forecasted load for a zone. Internally, Market Operations has expressed its consternation when the actual NYCA load comes in significantly above the NYCA forecast, but never when zonal forecasts are performing poorly.

To conclude, the focus on zonal forecast accuracy and the likely steps necessary to improve it may lead to an outcome that sacrifices a result that is actually more important to market participants and to system reliability.

It is for these reasons that the following alternative forecast accuracy goal is being proposed.

Alternative Load Forecast Accuracy Goal

Table 1
Proposed 2003 Load Forecasting Goal: 2 Parts, Each 5%

1) NYCA Peak Forecast MAPE	
Threshold	2.20%
Target	2.00%
Superior	1.80%
2) Days with 5% or greater absolute forecast error	
Threshold	21
Target	17
Superior	15

The proposed goal consists of two parts, each assigned 5% (assuming the load forecasting goal receives and overall weight of 10%). In the first, overall NYCA peak forecast accuracy is the focus. The target, 2.00%, is essentially what the NYCA peak forecast accuracy has been January, 2001 – July, 2002 (actually, 2.03%). Further reduction in this is problematical. Different versions of the DAFM have an in-sample MAPE of about 1.5%, meaning that even within the database on which the a model is estimated, the unexplained variation is about 1.5% of the peak. That represents a rock-bottom level of forecast error that will be exceeded once the model moves out of its estimation database (i.e., is put into production). From this perspective, the superior level of 1.80% looks to be very aggressive.

Table 2 shows the monthly performance of this metric.

Table 2

NYCA Peak Forecast Errors

	2001 Error	2001 Weath Adj Error	2002 Error	2002 Weath Adj Error
Jan	1.19%	1.61%	1.79%	2.08%
Feb	1.15%	1.57%	2.44%	2.40%
Mar	1.39%	0.98%	1.50%	1.46%
Apr	2.83%	2.62%	2.56%	1.87%
May	2.45%	1.96%	2.13%	1.84%
Jun	3.44%	1.74%	2.31%	1.51%
Jul	2.42%	3.02%	2.84%	2.07%
Aug	2.66%	2.34%		
Sep	2.09%	3.08%		
Oct	1.25%	1.35%		
Nov	1.19%	0.94%		
Dec	1.03%	0.90%		
Year to Date	1.92%	1.83%	2.22%	1.89%

Overall Averages:

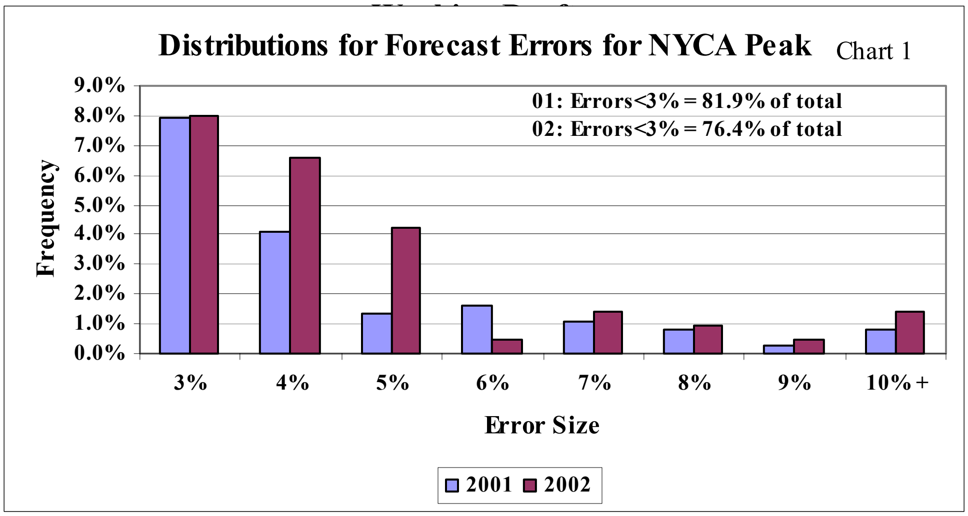
Error	2.03%
Weath Adj Error	1.85%

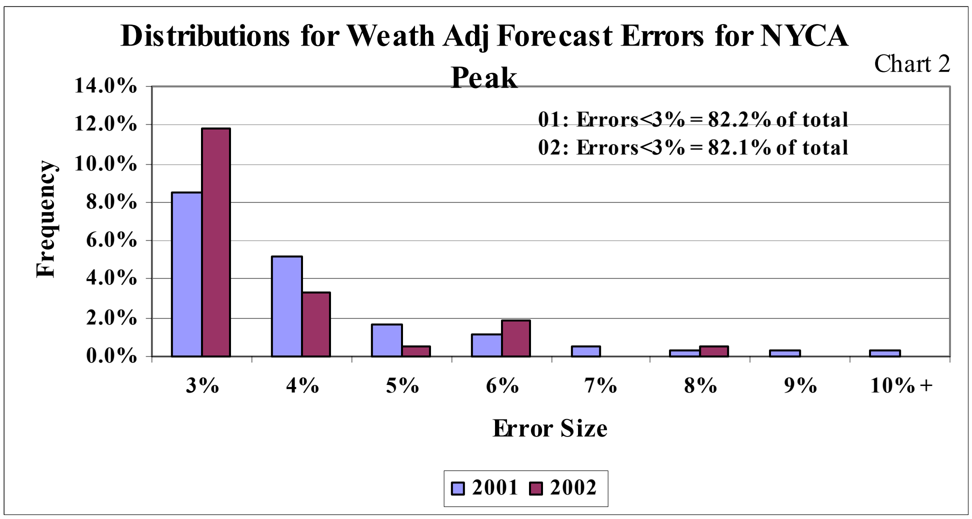
In-sample MAPE = 1.45%

(It is not clear, but 2001 appears to have been an unusually favorable year for forecast accuracy. Many of the monthly MAPEs are extremely low. January, February and December were all months that experienced no extreme winter weather. The performance levels set for forecast accuracy in 2002 are based on the 2001 experience, which at least partially explains why they have been so difficult to meet.)

The proposed target of 2.00% forecast accuracy for the NYCA peak represents a 10% improvement over the experience of 2002 so far.

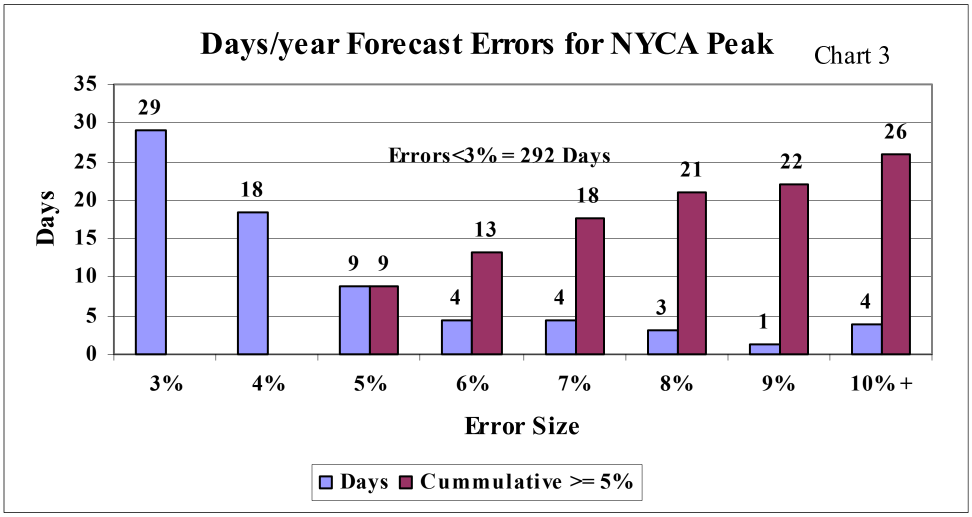
The second part of the proposed 2003 load forecasting goal addresses the most frequent concern of market participants and ISO operators regarding forecasts, large errors. The following charts show the distribution of daily NYCA peak forecast error for January, 2001 through July, 2003. Chart 1 shows total forecast error, Chart 2 weather adjusted error.

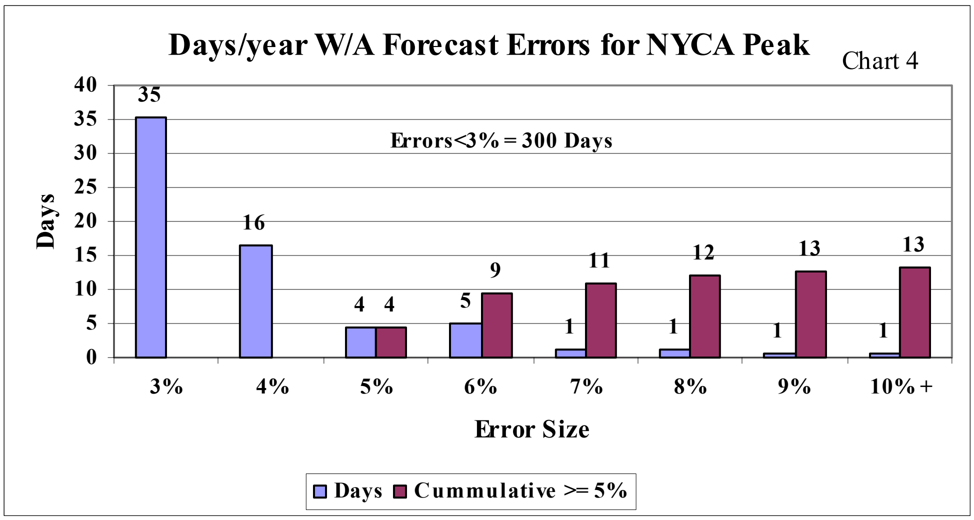




In both charts, the great majority of errors are small, less than 3%. Particularly in Chart 1, however, there is a significant frequency of large forecast errors and more so in 2002 than 2001. If we arbitrarily define 5% or greater errors to be large, we can see that there are a lot of days that fit into this category. If the errors on a significant number of these days can be reduced, the benefits to market participants and system reliability will be large.

Charts 3 and 4 translate the above information into what can be expected on a yearly basis.





Focusing on the non-weather adjusted errors in Chart 3, it shows that we can expect to see about 26 days per year where the forecast error is 5% or more. Four of these days will have 10%+ errors. On a weather-adjusted basis, the results are much better, as expected. Only 13 days have large errors, only one of which is 10%+.

The metric for the goal proposed to address this problem was developed as follows. The threshold performance level was set to reduce the number of days with large errors by 20%, the target represents a 1/3 reduction, and the threshold is a 40% reduction. In practice, a lot of progress can be made in reducing large forecast errors without it counting towards attaining this goal. For example, some days that would have 9% errors could wind up having 6% errors, which is an improvement, but it does not reduce the number of days with large forecast errors (defined as 5% or more). Therefore, meeting the standards for this goal represents a real challenge.

Conclusion

Refocusing the load forecasting goal on NYCA peak accuracy will direct ISO resources to where they will be most profitably employed. Addressing large forecast errors offers the opportunity to ameliorate the cause of most complaints about inaccurate forecasts and presents the best opportunity to save market participants real money.

For these reasons, the load forecasting accuracy proposed above should be adopted.