Using ensemble weather forecasts to manage utilities risk

Mary G Altalo, SAIC, and **Leonard A Smith**, London School of Economics

Errors in the forecast temperature are a major cause of short-term electricity load forecast error and thus grid instability. Weather forecast errors can account for 40% to 90% of demand forecast error, with the remainder attributed to load model error[1]. This weather-induced load error can be on the order of 8–10% of the load – and can therefore have significant financial consequences.

An end-to-end demonstration project[2] was conducted for a US electricity grid operator, the California Independent Systems Operator (Cal ISO)[3], which manages the state's 25,000-mile power transmission system. Cal ISO aggressively analyses its weather-induced load forecast error and reduces error through continuous operational improvements. It currently contracts several weather service providers for day-ahead regional weather forecasts which may be combined with regional "weighting" factors to produce the official daily forecast used to commit the electrical generation to supply the grid.

We show below that weather is the major contributor to load forecast error and associated costs during peak demand periods; that weather forecast accuracy can be significantly improved using a multi-model ensemble; and that use of probabilistic information represents reduction of costs of over 50% to the load forecaster.

'High impact' weather costs

Particular weather patterns can give rise to large errors in the day-ahead forecast, such as prolonged heat spells, fronts, marine boundary layers, sea breezes, Santa Annas and thunderstorms. These industry-relevant weather phenomena, or "high impact" weather, result in severe load imbalances with substantial cost. Forecasts which mitigate the impacts of these phenomena are called "high impact" forecasts, even if the weather involved is not severe. Temperature errors can produce either under-forecasts or over-forecasts of electricity demand.

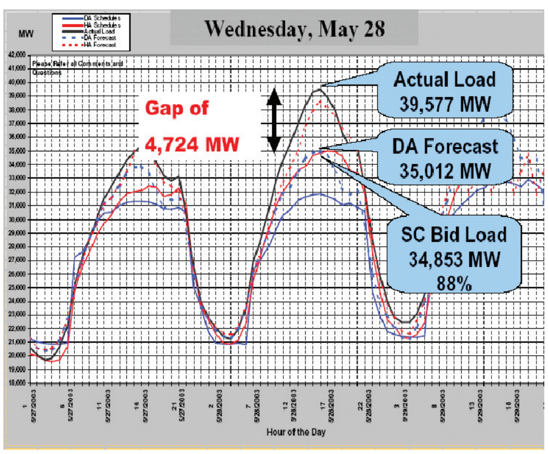
For example, on 28 May 2003, large temperature forecast error resulted in a 7% under-forecast of electricity demand (approximately 4,800MW). Meeting demand required buying on the spot market, costing millions of dollars for that one day. This gap between bid load and actual is shown in Figure 1. On 24 September 2003, another major temperature forecast error resulted in a large over-forecast of the next-day electricity demand. Generation had been scheduled and committed when the demand precipitously dropped due to more moderate temperatures. The costs associated with forecast error of this type, although high, are less costly than the under-forecasting load, since system stability, reliability and reputation are not at stake.

Because the load forecaster has knowledge of the asymmetry in the cost curve (see Figure 2, blue curve) she will put reserve on the grid, particularly if there are major uncertainties in the weather forecast. Probabilistic information is most valuable at these times. The cost curve is asymmetric at all absolute temperatures, but this is more pronounced at "threshold" temperatures (see Figure 2, red curve) when the demand for air conditioning peaks. Presently, a load forecaster uses her "experience" to determine how much "additional" generation to put on the grid to hedge this weather risk. Operational probabilistic information allows the forecaster to determine this safety margin more precisely.

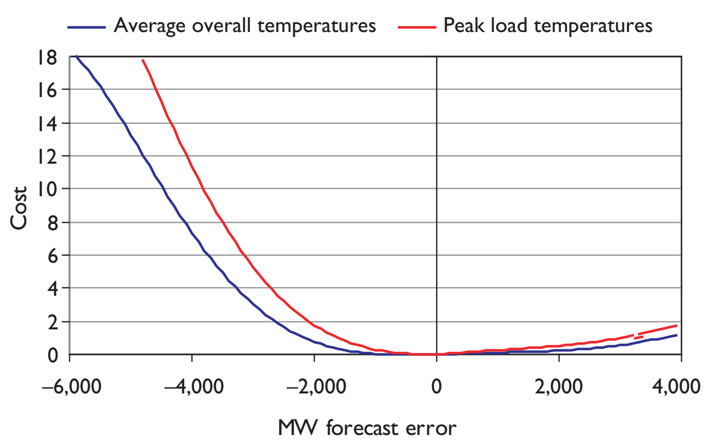
Temperature forecast error

Temperature forecasts from AVN, ETA, AVN MOS and ETA MOS systems serve as the basis for the Cal ISO weather forecast. The daily error of these forecasts varies significantly. For example, Cal ISO has reported significant error in forecasting load and temperature using the AVN MOS Tmax model. Analysis of 2003 summer forecasts results for the Bay Area shows consistent under-forecasting when temperatures rise and over-forecasting when temperatures fall. Under-forecasts of large increases of temperature (4 degrees) are common, and could cause as much as 1,800MW under-forecast. Clearly, a more accurate Tmax

1. Weather error induced gap in bid load and actual load (taken from the Cal ISO public web site



2. Relative cost curve for under/over-forecasting from Cal ISO showing the asymmetry



product with probabilistic information is needed.

Temperature forecast sources

National meteorological and hydrological services (NMHS) produce temperature forecast products based on high-resolution numerical model runs. Some charge for their product. Others do not. Numerous commercial vendors also produce a "tailored" product based on the NMHS product; these are usually a "reformatting" of the model forecast for spatial co-ordinates or delivery specifications. Because utilities have no way to judge the "pedigree"/"accuracy" of forecasts, many obtain multiple forecast products from vendors and average ("consensus" forecasts) or weight them to hedge against weather risk.

Temperature forecast types

All "day-ahead" forecasts are not equal:

- A "straight point forecast" is the output of a numerical prediction model taken as a single value which the demand forecaster ingests into his load model. This is the product which most weather vendors sell.
- A "distribution forecast" is a straight point forecast to which the historical averages/bias have been incorporated providing distributions. The spread in the distribution would allow an experienced operator to get a "feel" for the certainty of the forecast. Few operators currently deal with this uncertainty in a rigorous way.
- An "enhanced distribution forecast" adds additional information such as another piece of code to look at anomalies in the rate of temperature rise that may alert the operator to a sudden temperature spike.

The above products could be generated from a single model run. Ensemble products, however, extract additional information available only from multiple runs of one model, or many models. An "initial conditions ensemble" results from a single model, run repeatedly with different initial conditions to generate a forecast product with a probability distribution from that model. These ensembles have been produced operationally for over a decade from weather forecast offices in Europe, the United States and Canada. A "multi-model ensemble" is one in which a "collection" of numerical weather prediction models runs are united to produce a forecast product with probabilistic information. These are rarely available from any one office, since the ensemble products used are obtained from competing operational centres. This issue needs to be resolved through global forecasting cooperation.

Why are ensembles better?

Every numerical model has certain aspects of the atmosphere that it captures well, and others that it does not. If you follow only one model, or one particular initial condition, a critical weather pattern affecting your utility may be missed. Ideally, you would want to follow the ensemble that best captures "high impact" weather forecasts for your operation. Particular models can be given more weight in the ensemble forecast. There are a number of issues that an ensemble can better address, eg, better forecasting across time scales. In short-term forecasting (a few days), the ensemble reflects the chance of a weather pattern developing that was different from the pattern initially predicted. On the longer-term or seasonal scale, the ensemble yields information as to how this autumn is different from the average autumn – useful for planning purchases, scheduling maintenance or capital improvement plans.

Ensemble forecast performance at the Cal ISO

To determine whether an ensemble-based probabilistic forecast would outperform the forecast products currently used by the Cal ISO, several (five) one-day-ahead multi-model ensemble forecast products were constructed. The newly created forecast information product was "piloted" in the Cal ISO Load Forecast Operation to determine the "performance" of the improved probabilistic product. The approach was a re-analysis of the past load and weather error (summers of 2002, 2003) and a retrospective analysis of the financial impact of load imbalance. Two sets of performance metrics are used here. The first is for the technical performance or skill of the weather forecast model. The second is for the financial performance of the ensemble as it is transformed into a business forecast through the load forecasting process.

A full mathematical description and analysis of the method used to create the ensembles is published elsewhere[4]. In brief, for each day of the summers of 2002 and 2003, a collection of numerical weather prediction (NWP) station forecasts were obtained[5] and translated into probabilistic forecasts for maximum regional temperature (Tmax) in each of the four Cal ISO regions. We created five probabilistic forecasts, each under a slightly different method. The new ensemble-temperature forecasts were pushed through the load forecasting procedure and analysed for the financial impact on balancing the load. The economic costs for the 2002 and 2003 summer seasons were evaluated.

First, a baseline is determined for estimating the component of demand error related to weather forecast error. This is done by putting "perfect weather" data (observations) into the demand model. Different weather forecasts are then put through the same demand model, allowing the determination of the weather component of demand error and its cost.

Using the "best" of our (cross-validated) probabilistic forecasts resulted in difference of $5.4 million over having the perfect forecast information. The "consensus" forecast does significantly worst (up to $15 million in loss), while our probabilistic package based only on the AVN MOS does well and outperforms any other single operational forecast product for that area. Comparing the two years, the same modelling approach gives slightly better results for 2002, suggesting that the summer of 2002 may have been intrinsically easier to forecast than the summer of 2003.

The economic advantage of the probabilistic forecast package is due in part to a forecast median that is closer to the observed value, with a large impact coming from a few key days in September 2002. The significant economic savings reported for 2003 is found to be even larger for 2002. The probabilistic forecast is shown to save about $15 million, with a loss of about $5 million over perfect weather information, compared to a loss of over $21 million when using the straight AVN MOS Tmax values.

Conclusion

This is one of the first "end-to-end" studies in which weather error is monetized in an operational setting, a multi-model ensemble weather forecast product is constructed and compared with other forecast products, and the newly created product is used in a retrospective analysis to show an enhanced financial performance by using the probabilistic information. The major conclusion is that there is significantly more information contained in the existing weather data stream that can be gleaned from the interpretation of a multi-model ensemble.

This new information can save up to 50% of the weather forecast error-related costs by working in combination with the cost curve. This allows a forecaster to hedge risk based on fact, providing utilities an additional way to mitigate weather-related risks inherent in the load forecasting process. As more evidence accumulates for the value of the ensemble approach, there is a need for operational forecasting of multi-model ensembles as well as more global ensemble forecasting research and coordination efforts such as THORPEX to enhance the observations as well as the model interpretation.

Mary G Altalo is the WHERE-BASED WHAT[job-title] at SAIC, a DEFINITION OF COMPANY
E-mail: mary.g.altalo@saic.com
Leonard A Smith is a UK-based WHAT at the London School of Economics
E-mail PLEASE INSERT DESIRED EMAIL

1 Northeast Energy Network Performance Analysis, (NOAA) 2003
2 US NOAA contract PO 1022-8562 to Scripps
3 Courtesy of Dennis Gaushell, Cal ISO.
4 L A Smith, Altalo & Ziehmann (2004), 'Predictive distributions from an Ensemble of Forecasts: Extracting Electricity Demand From Imperfect Weather Models', *PhysicaD* (in review)
5 Provided by Quantum Weather