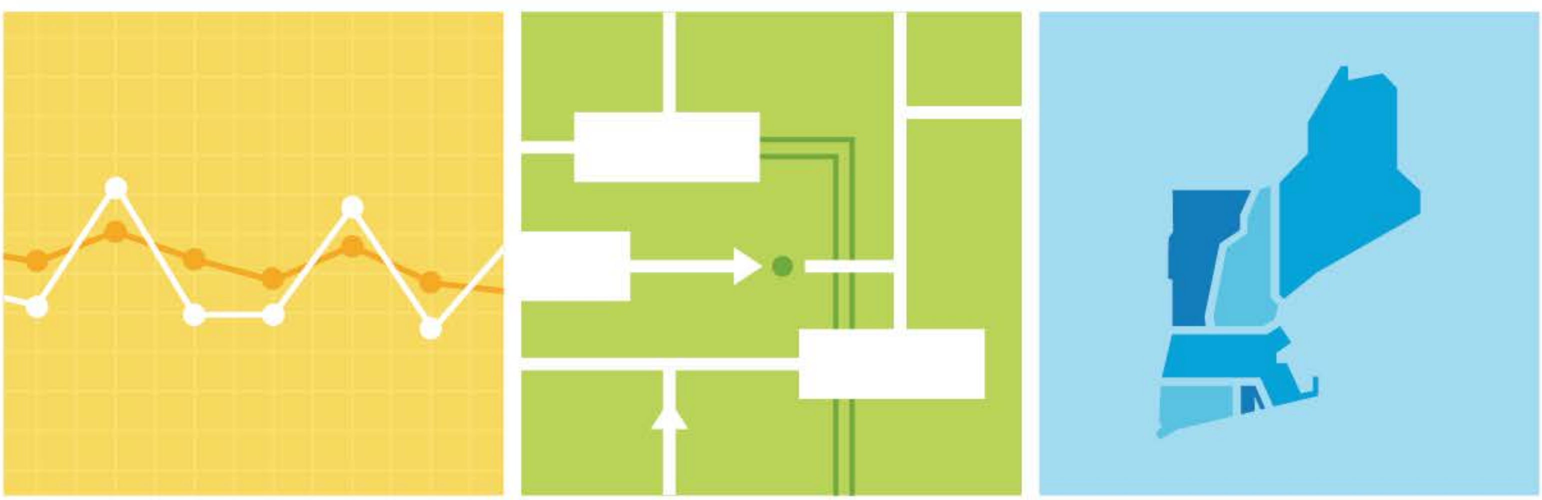
Forecast Modeling Procedure for the 2019 CELT Report: ISO New England Long-Run Energy and Seasonal Peak Forecasts

System Planning

APRIL 30, 2019



Contents

Figures

Section 1
Introduction

ISO-New England prepares forecasts of monthly energy and monthly and seasonal peak loads for the region as a whole and each of the six New England states. Forecast models are developed and estimated using econometric methods. Net energy for load (NEL) is modeled as a function of economic and other drivers. As such, it also is used to represent the underlying economic/demographic processes that influence peak load growth.

The peak load models were estimated with historical data from 2004 through 2018. The models were simulated with weather data from a 25-year historical period, generating 625 weekly observations encompassing the mildest to the most extreme weather conditions.

The energy models were estimated with historical data from 1992 through 2018. The economic forecast used in these models was Moody's Analytics' October 2018 release.

Section 2
Gross and Net Forecasts

"Gross" load forecasts are developed by first adding the energy savings from behind-the-meter photovoltaics (BTM PV), energy efficiency resources (EE), and active demand resources back into the historical NEL and daily peak load series before the models are estimated. The process of adding these savings back into the historical data is referred to as "reconstitution", and ensures the proper accounting of these resources in the development of the long-term load forecasting models. The reconstitution for EE is necessary because these resources are reflected on the supply side in the capacity market and they would otherwise be double counted. BTM PV reconstitution is needed because the annual long-term BTM PV forecast includes historical BTM PV that has reduced loads, and is therefore already embedded in the historical data. Consequently, the historical series needs to be "grossed up" to account for load reductions from BTM PV to similarly avoid double-counting. While gross forecasts are derived directly from the models developed with data after reconstitution, "net" forecasts are obtained by subtracting exogenous forecasts of BTM PV and EE[1] from the gross forecasts, and are therefore representative of the energy and loads expected to be observed in New England.

[1] The BTM PV forecast is provided by the Distributed Generation Forecast Working Group, and the EE forecast by the Energy Efficiency Forecast Working Group.

Section 3
Energy Forecasts

The 2019 energy forecast is produced using monthly models[2] of total energy consumption within the ISO-NE control area and the New England states. The goal is to specify forecasting models that predict electricity consumption as accurately as possible. To that end, regression models are designed to explain NEL in terms of variables believed to influence consumption.

3.1 The Regression Models[3]

NEL forecasts are prepared for the New England region and the New England states for each month. The models have the same fundamental structure, with some variation across months and states. The basic theoretical model is as follows:

$$Energy_{m,t} = f(Energy_{m,t-1}, Economy_t, Weather_{m,t}, X_t), \text{ where:}$$

Energy (NEL) = Monthly NEL in year t, with the impact of EE, BTM PV, and price-responsive demand (PRD) added back in.

Economy (RGSP) = Economic activity is represented by Gross Regional/State Product adjusted for inflation (RGSP). It is from Moody's Analytics. Gross domestic product is adjusted for inflation by Moody's price deflator series.

Weather = Weather is represented by two variables: Heating Degree Days (HDD) and Cooling Degree Days, depending on the season.

X = Unobservable variables that affect Energy Demand. Binary variables for specific years are included in some models. The relevant variables are determined by examining the residuals: (observed $NEL_{m,t}$ – modeled $NEL_{m,t}$). Large outliers are addressed by including a dummy variable for that year.

Sample Period: 1992-2018 for this forecast cycle.

3.2 Economic Input Data – Real Gross Regional/State Product

Real gross regional product represents overall economic activity in the energy models. Figure 1 presents historical and forecasted real gross regional product for New England used in the 2019 (NE_18) forecast, as compared to last year (NE_17).

[2] Monthly NEL models represent a new approach this year. In prior years, NEL was modeled annually.

[3] Inputs into the energy model are reported in the 2019 Forecast Details document, tab 8, on the ISO-NE website Load Forecasting page: https://www.iso-ne.com/system-planning/system-forecasting/load-forecast/

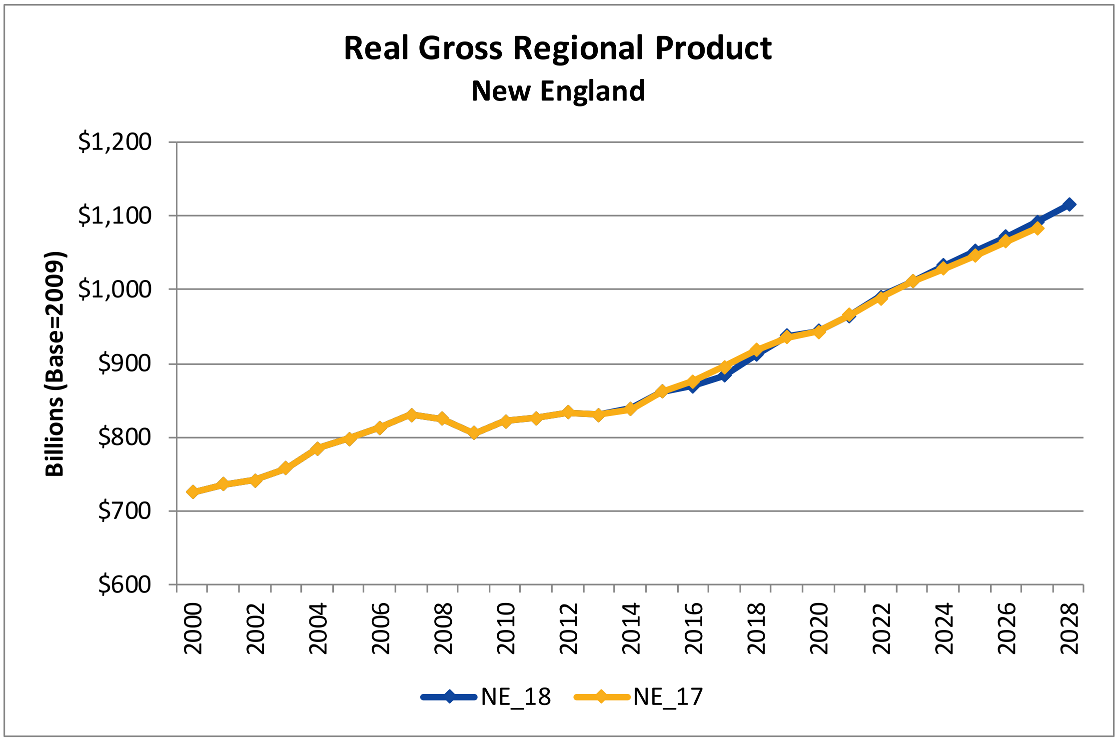


Figure 1: Real Gross State Product: 2000-2028

Section 4
Observations Regarding Historical and Projected Energy Consumption for New England[4]

As a result of a slightly higher economic forecasts from Moody's, the 2019 NEL forecast projects that weather-normalized energy demand will be marginally higher than expected in the 2018 forecast. For New England, gross NEL is expected to grow 1.1% over the forecast horizon. Last year, the forecasted growth rate for gross NEL was 0.9%.

New England's long-run net NEL growth rate is forecasted to be -0.4%, due to BTM PV and EE forecasts. This compares to -0.9% for last year's forecast.

[4] These results vary by state, depending on the individual economic forecasts. The forecasts by state are published on the ISO-NE website's Load Forecasting page in the *Forecast Data 2019* document located at: https://www.iso-ne.com/system-planning/system-forecasting/load-forecast/

Section 5
Peak Loads

5.1 Peak Load Forecast Distributions

Weekly peak load forecast distributions are developed by combining output from the daily peak load models with energy forecasts and weekly distributions of weather variables over 25 years[5]. Heating degree days (HDD) and an effective temperature correspond to the heating season (October-April), while the weighted temperature-humidity index (WTHI) and cooling degree days (CDD) are used for the cooling season (May-September).

5.1.1 Peak Load Distributions and Weather

The expected weather associated with the seasonal peak is considered to be the 50th percentile of the top 10% of the pertinent week's historical weather distribution. The monthly peak load is expected to occur at the weather associated with the 20th percentile of the top 10% of the pertinent week's weather distribution. The "pertinent week" is the week of the month or season with the most extreme weather distribution. For resource adequacy purposes, peak load distributions are developed for each week of the forecast horizon.

5.2 Daily Peak Load Models

Daily peak load models are estimated for the New England region as well as each of the New England states for each month. While the models have a common theoretical basis, they are individually adjusted for the unique characteristics of the region/state and the sample period.

Fundamental Drivers. Monthly NEL and daily weather variables comprise the foundation of the peak load models. Weather is the predominant observable cause of day-to-day variations in peak load, and also differentiates seasons. NEL serves as the base load, and represents underlying economic and demographic drivers.

Dummy Variables. The sample period comprises all days of the week, including holidays and weekends, while the monthly peak loads generally occur only on non-holiday weekdays. Including all days in the sample increases the sample size and reduces the number of "gaps" in the data. Significant gaps already exist due to the methodology of estimating separate models for each month. To accommodate the sample, dummy variables accounting for holidays and weekends are included in the models.

Sample Size. Because the forecast is based on "normal" weather, the estimation period must be long enough to capture significant variations in the weather; i.e. an abnormally warm or cool year cannot be allowed to unduly influence a long-run forecast. The sample period also must be short enough to assure reasonably consistent relationships between peak loads and the regressors.

Peak Load Model. The basic peak load model is a nonlinear function of energy and weather, expressed as:

$$Peak\ Load_{\mathrm{d}} = f(Monthly\ NEL_{m,t}, W_{h,d}, TW_d, D_{w,d}, D_{h,d}, X_d) \quad \text{where:}$$

[5] The 25-year weather distribution is new this year. Previously, a distribution over 40 years was used.

Peak $Load_d$ = Peak Load on day d

Monthly $NEL_{m,t}$ = Monthly NEL for month m in year t (see page 3 for details).

$W_{h,d}$ = Weather at the peak load hour on day d

= $(WTHI_h\text{-}55)^2$ for the months May-September

= Cooling Degree Days for the months May-September

= (Heating Degree Days)^2 for the months October-April

= Effective Temperature for the months October-April, where applicable

$WTHI_d$= 3-day weighted temperature-humidity index (THI) measured at the hour of the daily peak loads:

$$WTHI_d = \left\{\frac{[10*THI_d + 5*THI_{d-1} + 2*THI_{d-2}]}{17}\right\} - 55, \text{ and}$$

$$THI_d = 0.5 * DryBulbTemp_d + 0.3 * DewPointTemp_d + 15$$

Effective Temperature = Dry Bulb Temperature-[(65 Dry Bulb Temperature)/100]*wind speed

TW_d = Time Trend * $(WTHI_d\text{-}55)^2$ for the months May-September.

= Time Trend*HDD for the months October-April.

Time Trend is a year index e.g. 1992=1, 1993=2, etc.

$D_{w,d}$ = Dummy variable: 1 if day d is a weekend, 0 otherwise.

$D_{h,d}$ = Dummy variable: 1 if day d is a holiday, 0 otherwise. Holidays take precedence; if day d is both a weekend day and a holiday, $D_{w,d}$ = 0 and $D_{h,d}$ = 1.

X_d = Vector of other (unobservable) variables explaining daily peak loads

While NEL and Weather variables explain a substantial amount of the trend and variation in Peak Load, there are many other largely unknowable factors (X) that can be included in the model only by proxy, if at all.

The basic non-linear estimating equation with autoregressive error structure is specified as:

$$\text{Peak Load}_d = b_0 + b_1*NEL_{m,t} + b_2*W_{h,d} + b_3*TW_d + b4*D_{w,d} + b5*D_{h,d} + \hat{e}_d$$

$\hat{e}_d$ is the error term (residual), which follows an autoregressive process:

$\hat{e}_d = f(e_{d-1}, e_{d-2}, \dots, e_{d-n})$

Section 6
Observations Regarding the Projected Peak Forecast

Over the forecast horizon, the growth rate for New England's gross summer peak load is 0.7%, while it is -0.4% for the net summer peak load. The gross winter peak load is projected to grow by 0.6%, with a decline of -0.6% for the net winter peak.

The ISO-NE summer gross forecast for 2018 fell by 1.3% relative to last year's forecast for 2019. The gross forecast for the winter of 2018/2019 remained relatively flat.

Section 7
Forecast Model Evaluation and Testing

7.1 Peak Load Forecasting Methodology, Evaluation and Testing

The process for developing econometric-based peak load forecasting models is discussed in this section. The final equation used to forecast monthly peak loads is the result of iterations of the following steps.

(1) Informed by past years' models, a nonlinear econometric model with an autoregressive error structure of order (1) is specified and the parameters are estimated.
(2) The residuals are examined for extreme outliers.
(3) The residuals are examined to determine if they exhibit any trends or correlations.
(4) Influential observations are identified and removed if they cannot be explained.
(5) Proxy variables that might help explain the trends and influences in the residuals are evaluated.
(6) Statistical tests for goodness-of-fit and significance of the regressors are evaluated.

The modeling process begins with analysis of last year's models. The changes from CELT 2018 to Celt 2019 derive from:

(1) Changes in the sample period.[6]
(2) A re-specified weather distribution, from 40 years to 25 years.
(3) The effect of the economic forecast on the NEL forecast.
(4) Monthly NEL forecasts replace annualized monthly NEL forecasts.
(5) Additional weather-related variables are included.

7.1.1 Goodness-of-Fit and Statistical Significance[7]

It is important for the model to fit the historical data as closely as possible. Forecasts from an econometric model assume that historical relationships will continue into the future. A model that does not fit the sample data well introduces additional uncertainty into the forecast.

The traditional measure of how well the model fits the data is the R^2 statistic. The better the model fits the data, the higher will be the R^2 score. For New England's summer peak models, this statistic ranges from 0.94 to 0.97, with a median of 0.95. For the winter peak models, R^2 ranges between 0.87 and 0.95, with a median value of 0.92.

The t-test evaluates the statistical significance of each regressor, under the null hypothesis that the coefficient on the regressor is zero. A "rule of thumb" commonly used to indicate the explanatory power of a regressor is $t \geq 2$. However, if a variable is considered to be important in the initial design, careful consideration is given to retaining it in the equation regardless of reported significance. For example, the coefficients on the "Saturday" and "Sunday" variables do not always have t-values greater than 2, but they are important for identifying non-peak days.

---

[6] The sample period for the 2019 forecast is 2004-2018. Last year's sample period was 2003-2017.

[7] Relevant model statistics for all of ISO-NE and the state peak forecasting models are published as *2019 Regional and State Peak Model Details* on the ISO-NE website: https://www.iso-ne.com/system-planning/system-forecasting/load-forecast/

7.2 Energy Model[8]

Energy Model Methodology. There were several changes in the energy modeling methodology this year. The changes from last year include:

(1) A change in the sample period from 1990-2017 last year to 1992-2018 this year.
(2) Monthly models replace the annual models produced in previous years.
(3) Additional weather-related variables are included.

Energy Model Evaluation. The statistical model evaluation process for the energy models is similar to the peak load models.

Standard Statistical Tests. For New England's monthly energy models, the R^2 goodness-of-fit measure runs about .98 for all of the models. These high values are not unexpected, given the aggregate time-series nature of the data. The t-statistics on the coefficients of most regressors are over 2. Exceptions include some constant terms and an occasional variable considered *a priori* to be important for model fit.

[8] Relevant model statistics for all of ISO-NE and the state energy forecasting models are published as *2019 Regional and State Energy Model Details* on the ISO-NE website: https://www.iso-ne.com/system-planning/system-forecasting/load-forecast/