

Toward Scalable Stochastic Unit Commitment

Part 1: Load Scenario Generation

Yonghan Feng · Ignacio Rios · Sarah M. Ryan · Kai Spürkel · Jean-Paul Watson · Roger J-B Wets · David L. Woodruff

Revised: November 15, 2014

Abstract Unit commitment decisions made in the day-ahead market and during subsequent reliability assessments are critically based on forecasts of load. Traditional, deterministic unit commitment is based on point or expectation-based load forecasts. In contrast, stochastic unit commitment relies on multiple load scenarios, with associated probabilities, that in aggregate capture the range of likely load time-series. The shift from point-based to scenario-based forecasting necessitates a shift in forecasting technologies, to provide accurate inputs to stochastic unit commitment. In this paper, we discuss a novel scenario generation methodology for load forecasting in stochastic unit commitment, with application to real data associated with the Independent System Operator for New England (ISO-NE). The accuracy of the expected scenario generated using our methodology is consistent with that of point forecasting methods. The resulting sets of realistic scenarios serve as input to rigorously test the scalability of stochastic unit commitment solvers, as described in the companion paper. The scenarios generated

Yonghan Feng · Sarah M. Ryan
Department of Industrial and Manufacturing Systems Engineering, Iowa State University
E-mail: {yhfeng,smryan}@iastate.edu

Ignacio Rios
University of Chile
E-mail: ignacio.rios.u@gmail.com

Kai Spürkel
University of Duisburg-Essen
E-mail: kai.s88@web.de

Jean-Paul Watson
Analytics Department, Sandia National Laboratories
E-mail: jwatson@sandia.gov

Roger J-B Wets
Department of Mathematics, University of California Davis
E-mail: rjbwets@ucdavis.edu

David L. Woodruff
Graduate School of Management, University of California Davis
E-mail: dlwoodruff@ucdavis.edu

by our method are available as an online supplement to this paper, as part of a novel, publicly available large-scale stochastic unit commitment benchmark.

1 Introduction

Because operational constraints on thermal generation units require them to be committed well in advance of when they may be needed, system operators daily solve unit commitment optimization problems for the day ahead. Typically, scheduling decisions for a day d are made on day $d - 1$, based on forecasts of uncertain quantities such as hourly load and renewables output; these quantities are generally aggregated across buses, for each of a system's load zones. In the context of traditional deterministic unit commitment procedures, such forecasts take the form of point or expected-value quantities – representing a single time series for each forecasted quantity. Uncertainty associated with such forecasts is addressed by maintaining a non-trivial level of generation reserves, which compensate for deviations from the predicted quantities as day d operations proceed.

In contrast, stochastic unit commitment procedures [16, 24, 25] assume the availability of a number of forecast scenarios, each representing a distinct potential time series of the forecasted quantities. Throughout, we use the term *scenario* in a narrow sense, representing a full specification of all random data required to instantiate a unit commitment problem, with associated probability of occurrence. In aggregate, the set of scenarios should represent the range of possible behaviors on day d . By explicitly representing forecast uncertainty through sets of scenarios, it should be possible to significantly decrease generation reserve margins and consequently reduce overall system operation costs [21]. However, the need for multiple scenarios imposes fundamentally novel requirements on forecasting technologies, which have yet to be adequately addressed.

Our goal in this paper is to present approaches and data sources for generating quantifiably accurate and realistic load scenarios for use in stochastic unit commitment. We focus on load, as opposed to renewables production, to tractably scope our study. The procedures described below extend to wind and solar plant production, which we will address in a future contribution. While stochastic bidding or trading [23] and market clearing formulations [2, 15] for the day ahead wholesale electricity market have also been proposed, we consider the reliability unit commitment (also called resource adequacy assessment) process conducted by system operators. Consequently, we generate scenarios specifying zonal or system-wide load rather than demand bids by load-serving entities. We demonstrate the ability of Regional Transmission Operators (RTOs) and other operator entities to generate accurate load scenarios for use with stochastic unit commitment procedures, using data that are readily available. Our experiments proceed in the context of publicly available data from the Independent System Operator of New England (ISO-NE). The resulting scenarios are then used in the companion paper [4] to rigorously test the scalability of a stochastic unit commitment solver. These scenarios and the resulting test cases are publicly available, filling a critical need for researchers investigating the scalability of stochastic unit commitment solvers. The value of stochastic unit commitment compared to deterministic or other approaches, in terms of cost savings or reliability improvement, is not addressed in either paper but is a subject for ongoing rigorous testing.

The remainder of this paper is organized as follows. Our methodology for data transformation and fitting of load regression models is detailed in Section 2, which also includes a discussion of related literature and approaches. Preliminary results of the proposed methodology for modeling summer data in a single ISO-NE load zone were reported in [7]. This paper describes enhancements to our load modeling method by accounting for the non-weather-dependent component of demand (Section 2.3), transforming data to aggregate across load zones as well as days of the week (Section 2.4), and improving the process for segmenting days by weather forecast (Section 2.5). We then describe our procedures for generating load scenarios in Section 3. Comprehensive experimental results for all zones and all seasons of a year are presented for data associated with ISO-NE in Section 4. We conclude with a summary of our results in Section 5.

2 Load Forecasting Methodology

When forecasting load, the information available to operators on day $d - 1$ includes weather forecasts for day d , historical records of previous weather forecasts, and historical actual system loads. Historical system load data exhibit temporal patterns that vary according to season of the year, day of the week, and hour of the day. While some temporal load patterns are predictable based on knowledge of business hours and diurnal light patterns, the portion of load derived from heating and cooling (both industrial and residential) depends strongly on weather. And while numerical weather prediction models have become increasingly accurate over the past several decades, there remains significant uncertainty associated with day-ahead weather forecasts. The challenge for system operators is to form an accurate and comprehensive picture of the day-ahead load, which not only includes point forecasts of the load in each hour, but also acknowledges the precision (or lack thereof) associated with those forecasts. To address this challenge, we introduce a novel optimization-based method to develop a stochastic model for the load on day d based on the weather forecast available on day $d - 1$. Optimization models are solved to fit functional regression models for expected temporal load patterns (Section 2.3), obtain transformations between load patterns for different day types (Section 2.4), and approximate forecast error density functions (Section 2.5). We note that our approach does not assume normality or other parametric forms for error density functions, as are often used in experimental analyses of stochastic unit commitment(e.g., see [12]).

2.1 Background

Common methods for short-term load forecasting can be categorized as either based on artificial intelligence or statistical techniques [10]. Methods from artificial intelligence, such as neural networks, are widely used but do not provide probabilistic information that could be used to generate multiple probability-weighted scenarios. Among statistical approaches, which can provide the required probabilistic information, the most prevalent methods are time series and regression models. Due to limited space, we do not provide a complete review but refer the

reader to recent surveys such as [6]. Instead, we now highlight samples of statistical approaches for load forecasting from the recent literature, focusing on achieved accuracy and limitations for purposes of scenario generation. Similarly, forecasting models for wholesale market prices have been assessed for their suitability in generating scenarios [13].

The weather variable most commonly used to predict load is temperature, due to its influence on heating and cooling requirements. Other variables considered include humidity and cloud cover, although their impacts on load are much smaller than that of temperature. Humidity increases load in the summer, again due to cooling requirements. In contrast, cloud cover increases load in the winter (due to increased lighting requirements) and reduces load in the summer.

Liu et al. [14] analyze the nonlinear relationship between temperature and load, using estimates derived from a nonparametric regression method. They fit a time series model to the residuals of the load-temperature regression and consider lags of 1, 24, and 168 hours in their day-ahead forecasting model. Using actual historical temperature and load data obtained from a US utility, they demonstrate an out-of-sample mean absolute percent error (MAPE) of 1.2% for their 24-hour-ahead forecasts. Hong et al. [10] develop a multiple linear regression model of load that considers temperature, hour, type of day, and month as independent variables; the model additionally contains a linear term trend, and terms to capture interactions among the independent variables. Using actual weather data to predict hourly loads for a US utility over a one-year time period, they obtain an out-of-sample MAPE of 4.6%. Black [1] also uses a multiple linear regression model to examine the influence of weather on load, but instead focuses on summer weekdays in the region served by ISO-NE. Time-of-day effects are captured through a separate regression model for each hour of the day, each considering temperature, humidity, and irradiance as independent variables. The out-of-sample MAPEs yielded by these models average 2-3% for the whole New England region and 3-4% for individual subregions such as Connecticut and Southeast Massachusetts.

While “hind-casting” studies of the type described above (which use *actual* historical weather data as input) are useful for identifying factors that influence hourly loads, they do not reflect the accuracy or precision of load forecasts available in practice – which necessarily rely on day-ahead weather forecasts, as opposed to actual quantities. Further, while load forecasting remains a topic of active investigation, considerable emphasis persists on identifying a single most accurate forecast trajectory [11]. For investigations into stochastic unit commitment, several approaches for generating multiple trajectories have been used. An ad hoc procedure was used in an early study to create scenarios of large increasing or decreasing load ramps [25]. A common approach is to perturb a single forecast with some stochastic error terms. In studies aimed at exploring the impacts of uncertainty on unit commitment, these errors have been sampled from a Gaussian process [3], densities derived from a jackknife procedure applied to historical forecast errors [26], a uniform distribution [28], and an autoregressive process [20]. Typically, the same process or distribution is used to generate perturbations in each hour of each day in the study, whereas the hourly load forecast error distributions estimated in this paper vary considerably. Perturbations that are independent across hours may result in trajectories that do not well reflect the underlying temporal patterns in the load – in particular, the ramps that pose the biggest challenge for operators may not be depicted accurately. Time series methods have been used to directly

generate multiple trajectories for load [5]. Because of weekly and daily patterns in the load as well as its persistence, the fitted processes contain autoregressive terms with multiple lags. The propagation of uncertainty through the lagged terms exposes a drawback of this approach when forecasting more than a few hours in advance. Scenarios for hourly load on day d generated in the early afternoon of day $d - 1$ depend on loads in the evening of day $d - 1$. Even if the forecasts a few hours ahead are highly reliable, the process of stepping through the forecast hours on day d accumulates increasing numbers of independent “random shocks,” so that the variance of the generated loads for the later hours of day d grows unrealistically large. See, for example, Fig. 3 in [5], where the dispersion about the mean in the generated set of scenarios appears to increase geometrically with the length of time ahead. In reality, while weather forecasts diminish somewhat in precision as they extend over a day, the ability to forecast load does not drop off as drastically as the ensemble of time series trajectories would suggest.

Because statistical scenario generation approaches require large sample sizes to obtain reliable solutions, scalability of solution procedures may be compromised. To mitigate this difficulty, a general method of generating scenarios based on approximating rather than sampling from error distributions was proposed and motivated by load scenario generation in [17]. Our contributions in this paper are (1) the development of methods to incorporate critical data such as seasons, weather patterns, day-of-the-week, and load zones; (2) description of the details necessary to actually generate scenarios; (3) in combination with the companion paper, optimization models that use them and tests on the efficacy of those scenarios; and (4) the scenarios themselves as a resource for the research community. Incorporating numerical weather prediction models in stochastic optimization methods [29, 22] may be a promising alternative to generating scenarios from weather forecasts in advance of scheduling. However, because – to the best of our knowledge – system operators do not yet have access to such models but rather rely on purchased weather forecasts, we believe our approach (starting with a single weather forecast for the day ahead) is more relevant to current and near-term practice. Another method for short-term load forecasting is to identify similar days within a historical database, where similarity is based on weather, day of the week, and time of year. For example, ISO-NE identifies up to five similar days drawn from the same season with the same day-type according to similarity of their actual temperatures to the forecast temperature of the given day as well as similarity of forecast loads in the last hour of the previous day [8]. Our method has some commonality with this approach, in that we create segments of days that are similar in some sense. Then, within each segment we employ a functional regression method to approximate the probability distribution of load in each hour of the day ahead.

2.2 Methodology Overview

We use a multi-step procedure to control for season and type of day, and then approximate the relationships between weather forecast data and the distribution of hourly load sequences within segments of similar days. Starting from a historical database of day-ahead hourly weather forecasts and corresponding actual loads, our load forecasting methodology proceeds as follows:

1. Identify *date ranges*, or “seasons,” in which the relationship between weather and load – disregarding day-of-week effects – is likely to be similar. This ad hoc characterization qualitatively accounts for diurnal light patterns, heating vs. air conditioning impacts, and sociological factors such as whether school is in session.
2. Within each date range, transform the data to account for day-of-the-week and zonal differences within the system. Then, segment the data into bands based on forecast temperature. Data segmentation can in principle proceed using multiple forecast quantities (e.g., temperature and humidity). However, our experiments (summarized very briefly in Table 4) indicate that these additional factors do not significantly improve load forecast accuracy.
3. Within each segment, approximate the relationship between weather and load via a regression function. Additionally, approximate the distribution of residuals associated with the resulting regression function.

Following completion of the segmentation and approximation steps, the procedure for generating load scenarios for a given day d (with associated weather forecast generated on day $d - 1$) is given as follows:

1. Identify the date range DR to which day d belongs. Within DR , identify the temperature segment to which the weather forecast for day d belongs.
2. Apply the regression function associated with the identified segment to the weather forecast for day d .
3. Generate forecasted load scenarios for day d using distributions of the forecast errors.
4. As necessary, perform inverse transformations of the load sequences to match the day of the week and the zone.

The segmentation and approximation steps are fully described in the remainder of this section. Details of the load scenario generation process are described subsequently in Section 3.

2.3 Estimating regression functions

The main idea to build the regression function is to consider the weather forecast from day $d - 1$ to day d and with this information build a regression function. For this we use second order epi-splines [27] that minimize the deviations from the observed load at the day d . These functions, which have piecewise constant second derivatives, provide great flexibility to fit varying temporal patterns. For each day d in a given historical segment it is assumed that the following information is available: the hourly load, l_h^d , and the weather prediction for day d made on day $d - 1$. We use the temperature for all months and in summer, we also add dew point.

We split the 24 hours of day d in N_R sub-intervals $(h_{k-1}, h_k]$ of length $\delta = 24/N_R$. This process determines the total number of coefficients that need to be estimated, i.e., $2 \cdot (N_R + 2)$. In the summer, our regression curves will be built by relying on two epi-splines of order 2, one associated with temperature, and a second one associated with dew point. Each epi-spline is specified by N_R constant values for the second derivative in sub-intervals and two integration constants.

The construction implicitly assumes that the curves we are fitting are twice differentiable, but not necessarily C^2 . In addition to the parameter N_R there is a maximum curvature parameter, κ . The impact of the parameters on the accuracy of the fit is explored in Section 4.2.

Further, we assume that the load can be represented as the sum of two components:

- A non-weather component, i.e., a component which does not depend on the weather forecast and which is related to the normal behavior of the consumers under conditions specified by each segment.
- A weather component, which depends on the weather forecast.

It is natural to think that the non-weather component depends on the segment considered: independently of the weather, people in winter use a different amount of energy than in the summer. Consequently, for each segment we estimate the baseline load, which is the average load for each hour in the segment.

Algorithmic details of regression using epi-splines are provided in [17]. That paper provides an example illustrating how to fit a regression function from the weather for day d to \mathbb{R}^{24} , using data from some set of days $\hat{\mathcal{D}}$ as

$$r(d; \hat{\mathcal{D}}).$$

In the present paper we include details of transformations needed to make the method work effectively in our application. For these descriptions, we use the notations d and d° when respectively using forecast and observed weather to estimate transformation functions. A more formal statement would include the regression control parameters, but we omit these in the interest of clarity.

2.4 Data Transformation

Within a date range, we use transformations to combine data from disparate day types and zones, prior to constructing regression and error distribution models. Inverse transformations are performed to create load forecasts for particular day types and zones. Without such transformations, there is typically insufficient historical data to yield accurate forecasts. For example, dividing a year into 4 seasonal date ranges yields approximately 12 samples for each week day type. Such small sample sizes can lead to model over-fitting and associated out-of-sample prediction inaccuracies. This problem is made more acute if clustering or other classification schemes are imposed on the data within a date range.

Suppose we are given observed load profiles $l^d = (l_1^d, \dots, l_{24}^d) \in \mathbb{R}^{24}$ for a range of dates $d \in \mathcal{D}$. A portion of the the load is dependent upon weather factors, but load profiles also depend on the *type* of the day, e.g., load patterns differ between weekends and weekdays. Based on discernible differences in the average daily load patterns [7], we consider six day types: one for each weekday and one representing weekend days and holidays. We denote the set of day types as \mathcal{T} . The set of all dates belonging to a day type $T \in \mathcal{T}$ is denoted by \mathcal{D}_T . Clearly, each date maps to a unique day type, such that $\bigcup_{T \in \mathcal{T}} \mathcal{D}_T = \mathcal{D}$ and $\mathcal{D}_T \cap \mathcal{D}_{T'} = \emptyset$ if $T \neq T'$.

While a regression model could be developed for each day type, this would decrease the amount of data available for fitting significantly. Instead, we compute a transformation for each day type to a standard reference day, which we somewhat

arbitrarily select as Wednesday. Midweek days have the highest average load and have load profiles that are very similar to each other, fairly similar to Monday, and similar to the start of Friday.

The transformation is easily inverted, so that observed loads can be transformed to “Wednesday” and forecast loads can be transformed back to the original day type. In our analysis, we consider linear transformations to allow for simple forward and inverse computation. We use observed weather to find a linear transformation from each day type to Wednesday, but the transformation itself is not based on observed weather so load forecasts and scenario generation can be based entirely on data that are available in advance.

For each $d \in \mathcal{D}_T$, let $w^d = (w_1^d, \dots, w_{24}^d)$ denote the expected loads for our reference day type (Wednesday) corresponding to the native day type loads $l^d = (l_1^d, \dots, l_{24}^d)$. We assume that for $\forall T \in \mathcal{T}$, $\mathcal{D}_T \neq \emptyset$. For each $d \in \mathcal{D}_T$ of day type T , w^d is computed as a regression $r^o(d^o; \mathcal{D}_{wed})$, where d^o denotes the observed weather for day d .

Our goal is then to find a 24×24 matrix A_T such that

$$A_T l^d \approx w^d \quad \forall d \in \mathcal{D}_T. \quad (1)$$

We formally characterize Equation 1 as an optimization problem in which the coefficients of A_T appear as variables:

$$\min_{A_T} \sum_{d \in \mathcal{D}_T} \|A_T l^d - w^d\|. \quad (2)$$

Note that this (very small) optimization problem might be linear or non-linear, depending on the choice of norm. In our experiments, we used the l_2 norm and employed the open-source non-linear solver *Ipopt* to solve the resulting optimization problems (see <https://projects.coin-or.org/Ipopt>).

Depending on the relative size of the available dataset $\{l^d \mid d \in \mathcal{D}_T\}$ and the number of coefficients in A_T , it may be necessary to introduce additional constraints and regularization terms to formulation (2). Another potential issue is that the resulting A_T may be ill conditioned, causing difficulties in the calculation A_T^{-1} . Finally, nonsingularity needs to be enforced in (2), which can be easily achieved by requiring that all coefficients above and below the diagonal of A_T equal zero; we use this simple approach in the experiments reported below. The linear transformation preserves pairwise correlations among hourly loads, which appear to be fairly consistent across day types [7]. In an application where these correlations differed substantially, the use of a nonlinear transformation could be explored instead.

In practice, a balancing area is typically divided into zones, for which loads are forecasted and reported. In order to increase the data available for regression and error distribution estimation, we additionally combine data from disparate zones in a fashion analogous to that described above for converting data associated with different day types to a reference day type. Data are converted to correspond to data from a reference zone in the same way that Wednesday is used as a reference day.

2.5 Segmentation

For each date range, we partition the weather data for the composite dates into distinct segments. The idea is to limit regression and error distribution inaccuracy by only considering data with similar response characteristics. We segment based on the forecast temperature, as inclusion of additional weather variables (e.g., via k -means clustering) failed to improve prediction accuracy in our experiments. For each date range, we form a temperature distribution over which we introduce cutting points that define the segments.

Let t_*^d be a scalar representation of the hourly forecast temperatures for day $d \in \mathcal{D}$, where \mathcal{D} is the set of days in the date range. In our experiments, we define $t_*^d = t_{12}^d$, although alternative metrics such as average hourly temperature can be substituted. We require scalar representations of hourly temperature vectors to prevent data for a given day to be mapped to multiple error categories (see Section 3).

We estimate the probability density function $f_{t_*}(\cdot)$ of the temperature scalar t_*^d by fitting an exponential epi-spline [18, 19]. We denote the corresponding cumulative distribution function by $\mathbb{F}_{t_*}(\cdot)$. To obtain N_S segments of equal probability, we introduce the break points $\{b_1, \dots, b_{N_S+1}\} = \{0, 1/N_S, \dots, (N_S - 1)/N_S, 1\}$ for $\mathbb{F}_{t_h}(\cdot)$ and then calculate the limit temperatures for each segment \mathcal{S}_i as

$$(\underline{t}_i, \bar{t}_i) = (\mathbb{F}_{t_*}^{-1}(b_i), \mathbb{F}_{t_*}^{-1}(b_{i+1})), \quad i = 1, \dots, N_S.$$

Finally, considering the limit temperatures of each segment $\mathcal{S}_i, i = 1, \dots, N_S$, we group the days in the date range according to the rule:

$$d \in \mathcal{S}_i \Leftrightarrow t_*^d \in [\underline{t}_i, \bar{t}_i).$$

3 Scenario Generation

To capture the notion that both the mean load response and associated error distributions vary during the day, we split the day into parts and then categorize each portion according to the relative error. A regression model is constructed for each of the resulting day parts and associated error categories. Scenarios are then constructed by constructing paths whose trajectories are determined by selecting a specific error category for each day part. This process is a specific instantiation of the general scenario generation methodology detailed in [17].

Let H be the set of hours that define a partition of the hours in a day, specified as follows:

$$H = \{H_i\}_{i=1}^{|H|}, \quad H_1 = 1, H_{|H|} = 24, H_i < H_{i+1}.$$

The elements H_i represent the partition end-points, e.g., the i -th part of the day is given by the set of hours $\{H_i, \dots, H_{i+1}\}$. For each partition boundary H_i , we compute the observed regression error ϵ_i^d for each day $d \in \mathcal{D}$ as:

$$\epsilon_i^d = l_{H_i}^d - r_{H_i}(d; \mathcal{D}).$$

We estimate the distribution of these errors by fitting an exponential epi-spline [18, 19]. This process is graphically illustrated in Figure 1.

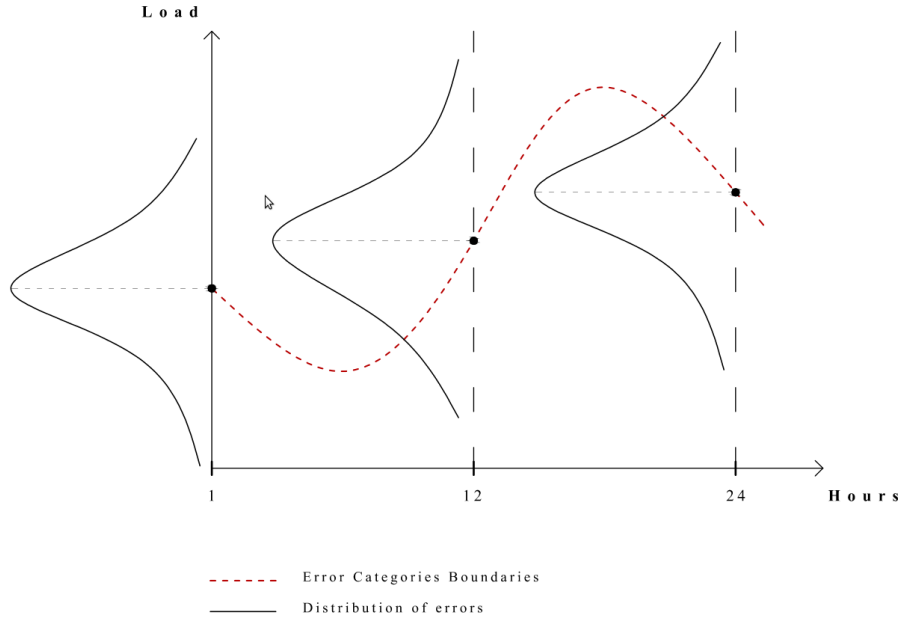


Fig. 1 Illustrative Load Regression Model and Error Distributions with Hour Partition $H = \{1, 12, 24\}$.

For each partition hour i , we denote the corresponding error probability density function by $f_{\epsilon_i}(\cdot)$. Categories within $f_{\epsilon_i}(\cdot)$ can then be defined through identification of break points of the associated cumulative distribution function $F_{\epsilon_i}(\cdot)$, as was performed for temperature segmentation. Specifically, to generate N_K equally sized categories, we select the break points $\{k_1, \dots, k_{N_K+1}\} = \{0, 1/N_K, 2/N_K, \dots, (N_K - 1)/N_K, 1\}$ in order to obtain equally weighted categories for each partition of the day. The resulting categories \mathcal{C}_i^k are then defined as $\mathcal{C}_i^k = [\mathbb{F}_{\epsilon_i}^{-1}(k_i), \mathbb{F}_{\epsilon_i}^{-1}(k_{i+1}))$, where $i \in \{1, \dots, |H| - 1\}$ denotes the day partition and $k \in \{1, \dots, N_K\}$ denotes the category.

Given an hour partition H and associated error categories \mathcal{C}_i^k , we sub-segment the days $d \in \mathcal{D}$ according to the observed regression model error at the corresponding partition i . Specifically, let \mathcal{D}_i^k denote the set of days in the segment for hour i and error category k , defined as follows:

$$d \in \mathcal{D}_i^k \Leftrightarrow \epsilon_i^d \in \mathcal{C}_i^k.$$

For each sub-segment \mathcal{D}_i^k , we fit a regression model $r(d; \mathcal{D}_i^k)$, from which a vector of predicted hourly loads $\hat{r}^{d,k}$ is extracted for each day $d \in \mathcal{D}_i^k$. Recall that each sub-segment has an associated regression function. Consequently, for boundaries in the middle of the day, there can be two points (i.e., two predicted loads, one from the “left” day part and one from the “right,” which might be a different error category.) We avoid discontinuities deriving a single regression curve per category, $\hat{r}^{d,k}$, by combining the predicted points at the limit hours. In particular,

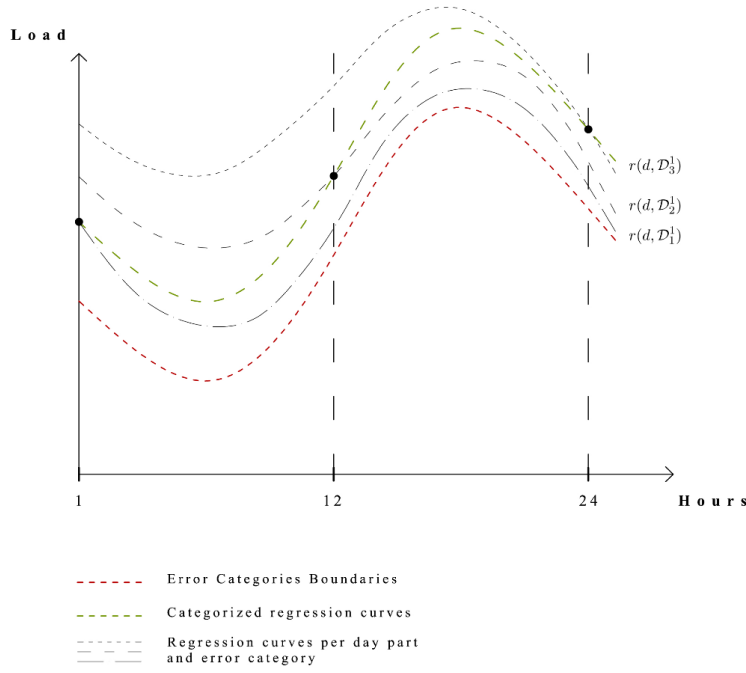


Fig. 2 Regression of the Error Category Construction.

for $h \in [H_i, H_{i+1}]$,

$$\hat{r}_h^{d,k} = \left(\frac{H_{i+1} - h}{H_{i+1} - H_i} \right) \cdot \hat{r}_{h,i}^{d,k} + \left(\frac{h - H_i}{H_{i+1} - H_i} \right) \cdot \hat{r}_{h,i+1}^{d,k}.$$

This process is illustrated in Figure 2.

Given regression models for each category \mathcal{C}_i^k , the observed load forecast errors are computed as

$$\varepsilon_{h,d}^k = l_h^d - \hat{r}_h^{d,k}.$$

We obtain a corresponding probability density function $f_{\varepsilon_i^k}(\cdot)$ by fitting an exponential epi-spline. An illustrative example of this step is shown in Figure 3, where $N_K = 2$.

The error densities $f_{\varepsilon_i^k}(\cdot)$ serve as the primary input to the scenario generation process. The first step in scenario generation involves the identification of a set of distribution cut points $C = \{c_z\}_{z=1}^{|C|}$ given as parameters, subject to $c_1 = 0.0$ and $c_{|C|} = 1.0$. For each partition i and category k , we then calculate the conditional expected value of the error in each interval defined by a pair of adjacent cutting

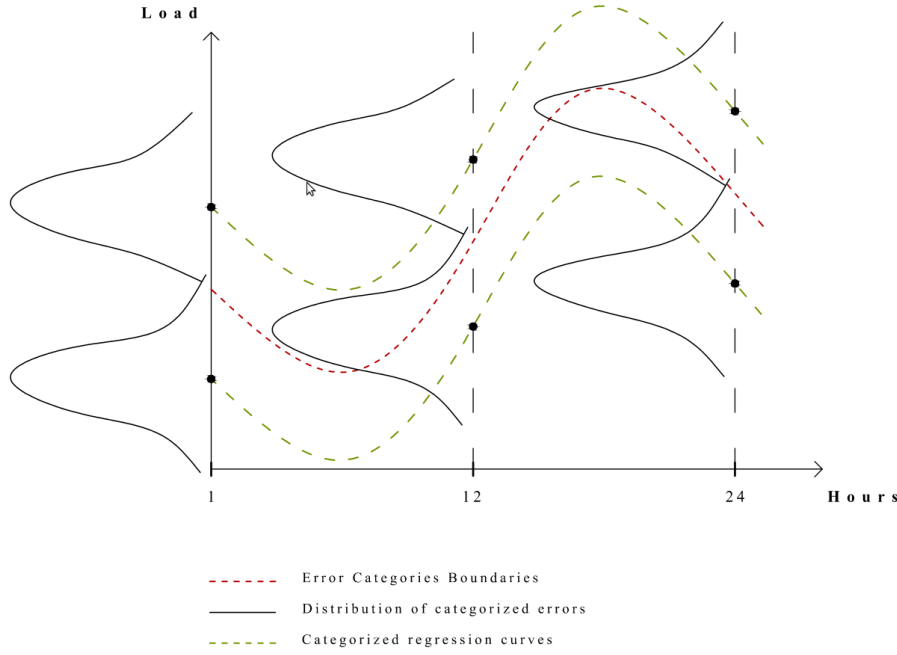


Fig. 3 Illustrative Error Category Regressions and Associated Error Distributions with Hour Partition $H = \{1, 12, 24\}$ and $N_K = 2$ Categories.

points:

$$\begin{aligned} \mathbb{E} \left[\varepsilon_i^k | \varepsilon_i^k \in [c_z, c_{z+1}] \right] &= \frac{\int_{c_z}^{c_{z+1}} x \cdot f_{\varepsilon_i^k}(x) dx}{\int_{c_z}^{c_{z+1}} f_{\varepsilon_i^k}(x) dx} \\ &= \xi_i^{k,z} \end{aligned}$$

where $\xi_i^{k,z}$ denotes the expected error in category k at cutting point z for hour i . The conditional probability associated with this error point is the width of the interval and the unconditional probability is obtained by multiplying by the width of the error category (which is $1/N_K$). The number of cutting points can vary per hour, as shown in Figure 4. In this illustrative example, $\mathcal{C} = \{0.0, 1.0\}$ for $i = 1$, but $\mathcal{C} = \{0.0, 0.5, 1.0\}$ for $i = 12, 24$. The cutting points also need not be symmetric, although they are in this example.

Given regression models $(\hat{r}_{h,i}^{k,d})$ for each hour partition boundary i and category k , we compute loads at the partition boundaries via:

$$l_{H_i}^{d,k,z} = \hat{r}_{H_i}^{d,k} + \xi_i^{k,z}.$$

For each hour H_i and each category C_i^k , this step yields $|C_i^k|$ forecast load samples. The final step in our scenario generation process is to connect these samples in

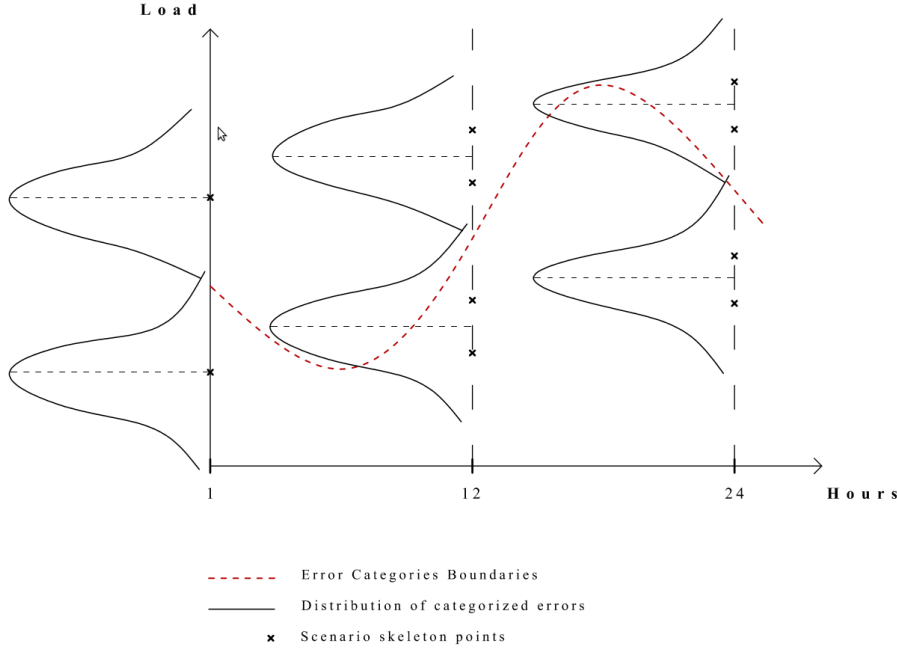


Fig. 4 Distribution Cutting Points for Scenario Generation when $H = \{1, 12, 24\}$ and $N_K = 2$. For $i = 1$, $C = \{0.0, 1.0\}$. For $i \in \{12, 24\}$, $C = \{0.0, 0.5, 1.0\}$.

order to construct a set of paths that approximates the stochastic process representing load for the full day. This is simply done by calculating the scenario loads at time $h \in [H_i, H_{i+1})$ by assuming that the deviation from the forecast varies between the deviation at hour H_i and hour H_{i+1} . This process is illustrated in Figure 5. Under this methodology, the number of paths (i.e., scenarios) generated depends on the parameters. For example, if the number of cutting points is the same value, $|C|$, for every hour partition boundary, then the number of scenarios would be

$$N_K \cdot (|C| - 1)^{|H|-1}.$$

The probability associated with each scenario must be computed as the product of the unconditional probabilities of the points used to construct it. Note that the generation process is deterministic, given a fixed set of historical input data.

4 Experimental Results

Although substantially better MAPEs can be achieved using standard leave-one-out validation, we have generated our forecasts and scenarios by simulating a rolling horizon as would be seen by a real-world system operator. In our experiments, we consider data for ISO-NE. We begin by fitting the models using only data from 2009 and 2010, and consider operations during 2011. As we simulate

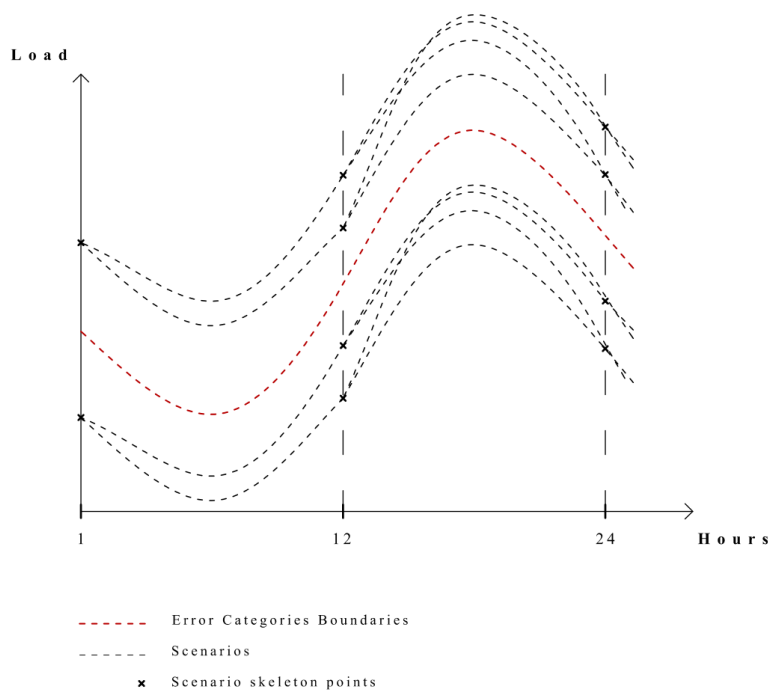


Fig. 5 Illustrative Scenario Paths Corresponding to the Cutting Points Shown in Figure 4.

the progression through the year, data from 2011 are added to the fit process as it becomes historical. Specifically, we execute the complete methodology outline in Section 2.2 for each simulated day. The entire procedure takes minutes of wall clock time to complete, and is therefore feasible in practice. Because our interest is in demonstrating load scenario generation methods, and not actually providing ISO-NE with load forecasts, we begin our simulation on January 2, 2011 and end on November 20, 2011. Excellent methods exist for dealing with the holiday season in the US [9], but their use is beyond the present scope. Further, we ignore August 28-30 of 2011, due to a hurricane event in the region. We partition the days of the year into the following date ranges: Winter (January 2 – March 31), Spring (April 1 – May 14), Summer (May 15 – September 14), and Fall (September 15 – November 20).

As reported in Table 1, ISO-NE is divided into 8 load zones. Weather data for each zone are taken from one or two weather stations. For zones with two stations, an aggregate weather forecast is computed by weighting the composite station data appropriately. Historical load data were obtained by ISO-NE through their public web site (http://www.iso-ne.com/markets/hstdata/znl_info/hourly/index.html). Hourly day-ahead temperature and dew point forecasts were provided directly by ISO-NE.

As indicated in Section 2.4, we aggregate data from disparate zones to make more data available to the fit process. Zone aggregation proceeds as follows, by employing two reference zones: CT and NEMASSBOST. The RI, SEMASS, and

Load Zone	Weather Stations	Weights
ME	PWM	1.0
NH	CON	1.0
VT	BTV	1.0
CT	BDL, BDR	0.13, 0.87
RI	PVD	1.0
SEMASS	PVD	1.0
WCMASS	BDL, ORH	0.5, 0.5
NEMASSBOST	BOS	1.0

Table 1 Load zones for ISO-NE, with weather stations and corresponding weights

WCMASS zones are aggregated with the CT zone, while the ME, NH, and VT zones are aggregated with the NEMASSBOST zone. This aggregation corresponds to a partition of ISO-NE in approximately northern and southern regions, which in turn share similar load characteristics. Within each partition, we select the zone with the greatest load as the reference zone, minimizing the total load transformed.

4.1 Forecast MAPEs

We quantify load forecast accuracy as the Mean Average Percent Error (MAPE), denoted by $\text{MAPE}(N_R, \kappa)$, as

$$\frac{\sum_{z \in \mathcal{Z}} \sum_{d \in \mathcal{D}} \sum_{h \in \mathcal{H}} \left(l_h^{z,d} - \mathbb{E}(l_h^{z,d} | (N_R, \kappa)) \right) / l_h^{z,d}}{|\mathcal{Z}| \cdot |\mathcal{D}| \cdot |\mathcal{H}|} \cdot 100 \quad (3)$$

where \mathcal{Z} , \mathcal{D} , and \mathcal{H} respectively denote the sets of load zones, dates, and hours under consideration. N_R and κ denote regression fit parameters. The aggregated MAPEs obtained for each date range in 2011 are reported in Table 2, considering $N_R = 24$, $\kappa = 500$, and a variable number of temperature segments N_S . The corresponding disaggregated (by zone) MAPEs are reported in Table 3. We observe that segmentation of the data by temperature does improve load forecast accuracy, although the benefits either stagnate or decrease once $N_S \geq 7$. Further, the data in Table 2 demonstrate load forecast accuracies that are consistent with those obtained by ISO-NE in practice, e.g., see http://www.iso-ne.com/support/training/courses/wem101/10_forecast_scheduling_callan.pdf. However, our method additionally provides estimates of forecast variability, represented through a collection of load scenarios.

We now briefly consider inclusion of dew point in addition to temperature in the segmentation process. Intuitively, dew point can influence load by impacting cooling requirements – particularly in the summer. To address this case, we segment forecast data by forming a scalar weather quantity based on a linear combination of temperature and dew point, using observations at $h = 12$. The resulting

Season	Segments (N_S)			
	1	3	5	7
Fall	5.45	4.66	4.20	3.99
Spring	3.10	2.88	2.67	2.73
Summer	10.25	4.82	4.14	4.19
Winter	5.25	3.32	3.29	3.47

Table 2 Aggregated ISO-NE MAPEs for 2011

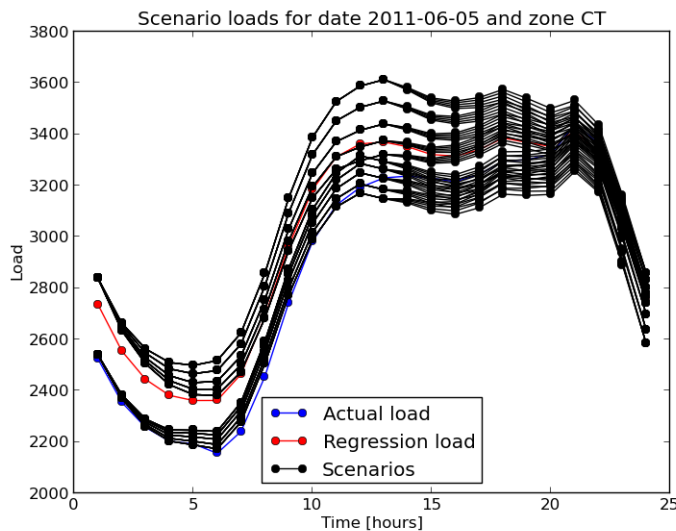


Fig. 6 Scenarios for Low Variance Load Forecast Day in 2011 for Zone CT

MAPEs, aggregated across all zones, are shown in Table 4, and indicate that inclusion of dew point quantities can marginally improve load forecast accuracy in specific contexts (e.g., in summer).

Finally, we provide exemplars of forecast load scenarios for the CT zone in ISO-NE, for early summer days in 2011. We have selected low and high variance load 50-scenario cases, respectively shown in Figures 6 and 7. We show only the first 24 hours of the time series, which repeat (for the purpose of stochastic unit commitment) for an additional 24 hours. With the exception of low-load hours 7 and 8, the scenarios shown in Figure 6 envelop the actual load. Further, during peak load for this day, the actual load closely mirrors the regression load. The situation is reversed in Figure 7, where the regression load mirrors the actual load in early hours of the day, but diverges after hour 18. These two particular cases serve as an integral component of test cases for assessing our stochastic unit commitment solver, described in the companion paper. All load scenario sets generated for ISO-NE for 2011, either at the zonal or aggregate level, are publicly available at <https://prescient.sandia.gov> for unrestricted use. We note that specific causes

Season	Zone	Segments (N_S)			
		1	3	5	7
Fall	NH	4.44	4.22	4.41	4.14
	VT	3.52	2.91	2.95	3.04
	ME	4.13	4.14	4.02	4.15
	CT	8.31	7.12	6.77	6.51
	RI	5.93	5.30	4.49	4.00
	SEMASS	5.75	4.99	4.40	3.84
	WCMass	7.20	6.60	6.34	6.66
	NEMASSBOST	4.82	4.63	4.19	4.07
Spring	NH	3.14	3.46	3.69	3.93
	VT	3.23	3.18	3.04	3.22
	ME	4.25	3.90	4.07	4.00
	CT	3.46	3.64	3.42	3.89
	RI	3.24	2.97	3.27	3.32
	SEMASS	3.05	2.93	3.17	3.24
	WCMass	3.74	3.50	3.80	3.98
	NEMASSBOST	3.24	3.41	3.35	3.48
Summer	NH	9.29	5.65	4.95	5.05
	VT	5.64	3.75	3.41	3.40
	ME	7.54	4.65	4.83	4.54
	CT	11.22	6.42	5.98	5.84
	RI	12.90	7.15	5.69	5.86
	SEMASS	12.72	6.76	5.60	5.74
	WCMass	9.47	5.51	4.57	4.63
	NEMASSBOST	11.34	6.17	5.78	5.32
Winter	NH	4.99	3.88	3.78	3.82
	VT	4.28	3.77	4.02	4.17
	ME	4.10	3.87	3.65	4.05
	CT	6.17	4.35	4.21	4.32
	RI	5.23	3.81	3.56	3.84
	SEMASS	5.34	4.05	3.76	4.11
	WCMass	5.46	3.69	3.84	3.95
	NEMASSBOST	5.42	3.54	3.62	3.67

Table 3 Disaggregated Zonal ISO-NE MAPEs for 2011

Condition	Segments (N_S)			
	1	3	5	7
Without DewPt	10.25	4.82	4.14	4.19
With DewPt	10.55	4.73	4.06	4.06

Table 4 Aggregated ISO-NE MAPE for Summer 2011, Using Segmentation Based on Temperature Both With or Without Dew Point

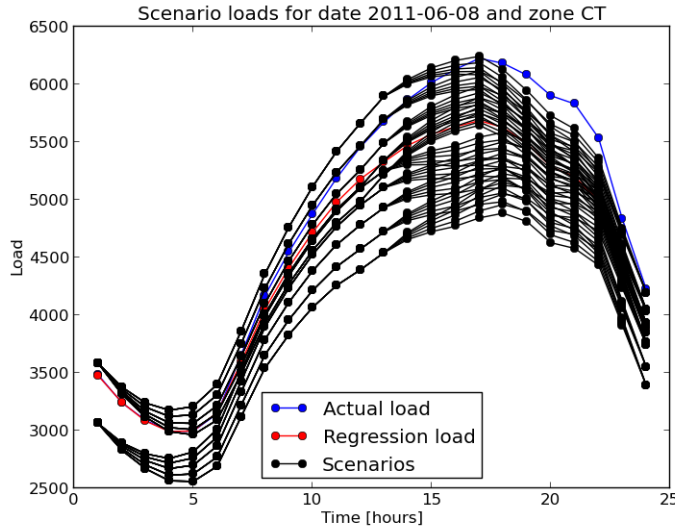


Fig. 7 Scenarios for High Variance Load Forecast Day in 2011 for Zone CT

of low versus high variance days are very difficult to identify. Rather, we simply observe that the differences are due to differences in the historical forecast errors associated with days of similar types and under similar forecast weather conditions.

4.2 Parameter Sensitivity for Model Fitting via Epi-Splines

To construct our regression models via epi-splines, it is necessary to specify values for the following key parameters:

- N_R : The number of intervals into which the hours of a day are sub-divided.
- κ : The maximum curvature of the regression model.

While it is necessary to specify a particular norm (e.g., L^1 or L^2) when solving the embedded optimization problems, in practice the choice of norm has almost no impact on the resulting load forecast MAPEs. In contrast, accuracy is more sensitive to the choice of N_R and κ , as we now demonstrate.

First, we consider the impact of N_R on load forecast accuracy. We compute the MAPEs for $N_R \in \{6, 12, 18, 24, 32, 48\}$, fixing $N_S = 5$ and $\kappa = 500$. The results

indicate that the MAPE changes from about 4.7% to 4.5% over the range, with values of 24 and above yielding nearly equal MAPEs.

Next, we consider the impact of κ on load forecast accuracy. Fixing $N_R = 24$ and $N_S = 5$, we compute MAPEs for $\kappa \in \{20, 40, 60, 80, 100, 150, 200, 500\}$. Extremely small values of κ do adversely impact the MAPEs by severely restricting the curvature. However, over the range of κ we tested, there is essentially no sensitivity.

5 Conclusion

In this paper, we introduced novel methods for obtaining distributions of forecast load in each hour of day d , based on weather forecasts available on day $d - 1$. Our goal is to estimate trends and error distributions from these data, in order to generate probabilistic scenarios for day-ahead load. We make use of transformations of day types and load zones to provide data for general, functional regression based on epi-splines that is conditional on the error category. We also fit the error distributions using epi-splines.

Our experiments indicate that our models:

- Can be generated using data readily available to system operators.
- Can be produced in minutes of run time for multiple years of input data.
- Produce errors that in expectation are competitive in the aggregate with load forecasting procedures used by industry.
- Obtain MAPE values for forecast load that are similar to those found in hind-casting studies, which eliminate weather forecast uncertainty by focusing strictly on actual or observed weather quantities.
- Explicitly model likely variability of realized load.

While we do not examine multi-stage stochastic unit commitment models in the companion paper [4], we note that our scenarios are tree-structured and consequently can be effectively used in that context.

We continue to refine our methods, specifically by focusing on reducing prediction errors associated with peak load periods and adapting the approach to times of the year where temperature is not as strong a predictor of the load. Our approach could also be extended to model price-responsive demand when such programs are widely implemented and data become available to estimate demand-price relationships.

The sets of scenarios generated by the methods we describe serve as input to rigorously test the scalability of stochastic unit commitment solvers, as described in the companion paper. Scenarios generated in a rigorous way from realistic input data are important for ongoing research in stochastic and robust unit commitment, and for simulation and related production cost models. The scenarios generated by our method using data from ISO-NE for 2011 are available as an online supplement to the paper.

Acknowledgements Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94-AL85000. This work was funded by the Department of Energy's Advanced Research Projects Agency - Energy, under the Green Energy Network Integration

(GENI) project portfolio. The authors would like to thank Dr. Eugene Litvinov and his research group at ISO-NE (in particular, Bill Callan) for their assistance with ISO-NE system and data sources.

References

1. Black, J.: Load hindcasting: A retrospective regional load prediction method using reanalysis weather data. Ph.D. thesis, Department of Mechanical and Industrial Engineering, University of Massachusetts Amherst (2011)
2. Bouffard, F., Galiana, F.D., Conejo, A.J.: Market-clearing with stochastic security - Part 1: formulation. *IEEE Transactions on Power Systems* **20**(4), 1818–1826 (2005)
3. Carpentier, P., Gohén, G., Culioli, J.C., Renaud, A.: stochastic optimization of unit commitment: A new decomposition framework. *IEEE Transactions on Power Systems* **11**, 1067–1073 (1996)
4. Cheung, K., Gade, D., Monroy, C.S., Ryan, S.M., Watson, J.P., Wets, R.J.B., Woodruff, D.L.: Scalable stochastic unit commitment, part 2: Solver performance. *Energy Systems* (Under Review)
5. Dupačová, J., Gröwe-Kuska, N., Römisch, W.: Scenario reduction in stochastic programming. *Mathematical Programming* **95**(3), 493–511 (2003). DOI 10.1007/s10107-002-0331-0
6. Feinberg, E., Genethliou, D.: Load forecasting. In: *Applied Mathematics for Restructured Electric Power Systems*, pp. 269–285 (2005)
7. Feng, Y., Gade, D., Ryan, S.M., Watson, J.P., Wets, R.J.B., Woodruff, D.L.: A new approximation method for generating day-ahead load scenarios. In: *Proceedings of the 2013 IEEE Power and Energy Society General Meeting* (2013)
8. Gould, S.: Create demand forecast. Tech. Rep. SOP-OUTSCH.0040.0010, ISO New England, Inc. (2011)
9. Hong, T.: Electric load forecasting with holiday effect. <https://sites.google.com/site/hongtao01/courses/holiday> (2012). Presented at AEIC Load Research Workshop 2012, Orlando, FL
10. Hong, T., Gui, M., Baran, M., Willis, H.: Modeling and forecasting hourly electric load by multiple linear regression with interactions. In: *Proceedings of the 2010 IEEE Power and Energy Society Meeting* (2010)
11. Hong, T., Pinson, P., Fan, S.: Global energy forecasting competition 2012. *International Journal of Forecasting* (2013). DOI 10.1016/j.ijforecast.2013.07.001
12. Huang, Y., Zheng, Q. P., Wang, J.: Two-stage stochastic unit commitment model including non-generation resources with conditional value-at-risk constraints. *Electric Power Systems Research* **116**, 427–438 (2014).
13. Klboe, G., Eriksrud, A.L., Fleten, S.E.: Benchmarking time series based forecasting models for electricity balancing market prices. *Energy Systems* (2013). DOI 10.1007/s12667-013-0103-3
14. Liu, J., Chen, R., Liu, L., Harris, J.: A semi-parametric time series approach in modeling hourly electricity loads. *Journal of Forecasting* **25**, 537–559 (2006)
15. Morales, J.M., Conejo, A.J., Liu, K., Zhong, J.: Pricing electricity in pools with wind producers. *Power Systems, IEEE Transactions on* **27**(3), 1366–1376 (2012)
16. Papavasiliou, A., Oren, S.: A stochastic unit commitment model for integrating renewable supply and demand response. In: *Proceedings of the 2012 IEEE Power and Energy Society Meeting* (2012)
17. Rios, I., Wets, R.J.B., Woodruff, D.L.: Multi-period forecasting and scenario generation with limited data. *Computational Management Science* (To Appear)
18. Royset, J., Wets, R.: Epi-splines and exponential epi-splines: Pliable approximation tools. Tech. rep., University of California Davis (2012)
19. Royset, J., Wets, R.: Fusion of hard and soft information in non-parametric density estimation. Tech. rep., University of California Davis (2012)
20. Ruiz, P., Philbrick, C., Sauer, P.: Modeling approaches for computational cost reduction in stochastic unit commitment formulations. *IEEE Transactions on Power Systems* **25**(1), 588–589 (2010)
21. Ruiz, P., Philbrick, R., Zack, E., Cheung, K., Sauer, P.: Uncertainty management in the unit commitment problem. *IEEE Transactions on Power Systems* **24**(2), 642–651 (2009)

22. Salazar, J.M., Diwekar, U., Constantinescu, E., Zavala, V.M.: Stochastic optimization approach to water management in cooling-constrained power plants. *Applied Energy* **112**(0), 12–22 (2013). DOI <http://dx.doi.org/10.1016/j.apenergy.2013.05.077>
23. Scharff, R., Egerer, J., Sder, L.: A description of the operative decision-making process of a power generating company on the nordic electricity market. *Energy Systems* **5**(2), 349–369 (2014). DOI [10.1007/s12667-013-0104-2](https://doi.org/10.1007/s12667-013-0104-2)
24. Siface, D., Vespucci, M.T., Gelmini, A.: Solution of the mixed integer large scale unit commitment problem by means of a continuous stochastic linear programming model. *Energy Systems* **5**, 269–284 (2014). DOI [10.1007/s12667-013-0107-z](https://doi.org/10.1007/s12667-013-0107-z)
25. Takriti, S., Birge, J., Long, E.: A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems* **11**(3), 1497–1508 (1996)
26. Takriti, S., Krasenbrink, B., Wu, L.S.Y.: Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem. *Operations Research* **48**(2), 268–280 (2000)
27. Wets, R.J.B., Bianchi, S.W.: Term and volatility structures. *Handbook of Asset and Liability Management: Theory and Methodology* **1**, 25 (2006)
28. Wu, L., Shahidehpour, M., Li, T.: Stochastic security-constrained unit commitment. *IEEE Transactions on Power Systems* **22**(2), 800–811 (2007)
29. Zavala, V.M., Constantinescu, E.M., Krause, T., Anitescu, M.: On-line economic optimization of energy systems using weather forecast information. *Journal of Process Control* **19**(10), 1725–1736 (2009). DOI [10.1016/j.jprocont.2009.07.004](https://doi.org/10.1016/j.jprocont.2009.07.004)