Limited-Knowledge Economic Dispatch Prediction using Bayesian Averaging of Single-Node Models

Kevin D. Smith
Electricity Infrastructure Group
Pacific Northwest National Laboratory
kevin.smith@pnnl.gov

Karen Studarus
Electricity Infrastructure Group
Pacific Northwest National Laboratory
karen.studarus@pnnl.gov

***Abstract*—A great deal of literature examines economic dispatch from the perspective of a grid operator, assuming full knowledge of generating unit costs and constraints, network topology, line capacities, etc. But others without access to the complete, data-intensive model may also wish to predict unit dispatch under various scenarios. This paper develops a Bayesian approach to predicting future economic dispatch, which relies only on historical dispatch observations, as well as general assumptions of operating costs. The approach uses a Markov-chain Monte Carlo method to create an ensemble of network-free models, which capture fundamental properties of economic dispatch (like merit order and marginal generator behavior) under particular "regimes" of grid operation (a particular range of load, set of congested lines, set of committed units, etc.). It then uses Bayesian averaging over many of these simple models to create an ensemble model, which approximates economic dispatch under general operating conditions. A case study using data from New York is used to verify the ensemble model.**

I. Introduction

The annals of power systems engineering contain many methods for grid operators to solve economic dispatch, using a variety of techniques, under a wide range of conditions (see [1] and [2] as two reviews on the subject). These economic dispatch formulations become arbitrarily complex, requiring large amounts of data—operating costs, power generation limits, ramping limits, transmission line constraints, spinning reserve constraints, etc. To researchers interested in predicting unit dispatch under various scenarios, but who have limited access to proprietary knowledge, these data requirements are a barrier. One example of a project facing this challenge is PNNL's Grid Project Impact Quantification Tool (GridPIQ). GridPIQ is a transparent, modular, and publicly available web-based screening tool that estimates the varied impacts of a growing range of specific smart grid technologies and project types, including coordinated electric vehicle charging, demand response, solar photovoltaic (PV) generation, and energy storage (ES) [3] [4]. A crucial step in the GridPIQ model is estimating the power outputs of known generating units in a region. But GridPIQ's transparency precludes the use of proprietary information, and a lightweight web calculator cannot run a power flow study.

Statistical modeling is a computationally inexpensive way to predict unit dispatch using only publicly available data. GridPIQ currently uses the EPA's AVoided Emissions and geneRation Tool (AVERT) [5], for instance, which estimates unit dispatch via statistical models of freely-available data from the Air Markets Database [6]. But statistical models are only valid so long as their underlying assumptions still apply. If a grid project is too large in scale, diverging from historical conditions by too great a degree, there is a risk that AVERTs predictions would no longer apply.

Approaches to dispatch prediction therefore fall on a spectrum of possibilities. On one side, "physical" models of the grid extend our knowledge of system dynamics into never-before-observed scenarios, for which historical data does not exist. Grid components obey well-understood rules, and power flow studies can predict grid dynamics (including dispatch) under many operating conditions. On the other side, statistical models can predict system dynamics from historical observations, even when knowledge of the system's inner workings are too limited for a physical model.

This paper lies somewhere in the middle. Our objective is to predict fossil fuel generator dispatch (i) under operating conditions not observed within a historical data set of interest, and (ii) without access to the computational power or the knowledge of network topology, dispatch costs, etc. needed for a power flow study. We rely only on historical observations of generator dispatch, assumptions about distributions of unit operating costs, and total load on fossil fuel generators in the system. With a lightweight model like this that is flexible enough to capture future energy scenarios, we can offer a scratch pad for stakeholders to explore important what-if scenarios and steer future research.

A. Notation

We consider a system with n generating units, with labels $i \in [n] = \{1, 2, \ldots, n\}$. For each time $t \in [T] = \{1, 2, \ldots, T\}$ in a known dataset X of historical dispatch, d_t denotes the total system demand, and p_{ti} denotes the true power output of each unit. The shorthand notation p_t represents the vector in $\mathbb{R}^n$ whose elements are the unit power outputs p_{ti}. The goal

Pacific Northwest National Laboratory is a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy under contract No. DE-AC05-76RL01830. This work was funded by the U.S. Department of Energy's Office of Electricity Delivery and Energy Reliability.

978-1-5386-3596-4/18/$31.00

is to predict the unknown unit dispatch $p_{t'}$ for some future times $t' \in \{T+1, T+2, \ldots, T+T'\}$, assuming demands $d_{t'}$ are known. For each unit $i \in [n]$ and time $t \in [T+T']$, these predictions are denoted by $\hat{p}_{ti}$, or by $\hat{p}_t$ to represent the vector of predicted dispatch for all units.

II. Model

Our dispatch model is an ensemble of N "single-node models," described in Section II-A. Individually, the single-node models are very simple, neglecting the effects that transmission constraints and unit commitment have on economic dispatch. But they are complex enough to reproduce merit order, the behavior of marginal units, and limits to unit power outputs. A single-node model can predict unit dispatch under a particular "operating regime" of the grid—a particular range of load, a particular set of congested lines, a particular group of committed units, etc.—but cannot predict dispatch under general operating conditions.

In order to predict unit dispatch under more general circumstances, we create an ensemble of single-node models using Bayesian Model Averaging, detailed in Section II-B. The ensemble model is an average of N single-node models, weighted by the posterior probability of each single-node model, where posterior probabilities are found using the Markov-chain Monte Carlo procedure outlined in Section II-C. By averaging over many single-node models, the ensemble model averages over unit dispatch predictions in many different operating regimes, weighted by the posterior probabilities of each respective model. Finally, in Section II-D, we present an algorithm to efficiently solve many instances of the specific economic dispatch problem required by the single-node model.

A. Single-Node Model

The fundamental building block of our analysis is the *single-node model*. The idea behind single-node models is familiar from introductory power systems texts [7]. It assumes a centrally-dispatched grid without line constraints, so that when there is a total demand d, the economic dispatch $p_i^*(d)$ of each unit $i \in [n]$ is given by a convex quadratic program:

$$\begin{aligned} \underset{p^*(d)}{\text{minimize}} \quad & \sum_{i=1}^{n} a_i p_i^*(d) + b_i p_i^*(d)^2 \\ \text{subject to} \quad & d = \sum_{i=1}^{n} p_i^*(d) \\ & p_i^{\min} \le p_i^*(d) \le p_i^{\max}, \ \forall i \in [n] \end{aligned} \tag{1}$$

Here a_i and $b_i > 0$ are effective linear and quadratic cost coefficients for each unit, and $p_i^{\min} \le p_i^{\max}$ are operating limits of each unit.

Of course, the single-node model is a highly idealized picture for power systems operation; there are many other factors (like transmission constraints) that influence real-world economic dispatch. It is rare that idealized economic dispatch $p_i^*(d_t)$ will equal true dispatch p_{ti} for given t, i, no matter how well the single-node model parameters $a, b, p^{\min}, p^{\max}$ are chosen. To characterize this inevitable error, it is desirable that the single-node model predict a probability distribution for unit dispatch, rather than yielding the single optimal vector $p^*(d_t)$. This analysis uses independent normal distributions for single-node model predictions $\hat{p}_{ti}$, with means $p_i^*(d_t)$ and variances σ_i^2, where σ_i is a parameter which varies by generating unit but is constant in time. Thus, for each unit $i \in [n]$ and time $t \in [T+T']$, the single-node model yields a probability distribution with the following density function to predict unit dispatch:

$$f_{\text{sn}}(\hat{p}_{ti}) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2\sigma_i^2}\left(p_i^*(d_t) - \hat{p}_{ti}\right)^2\right] \tag{2}$$

Hence the expected value of $f_{\text{sn}}(\hat{p}_{ti})$ is the idealized economic dispatch $p_i^*(d_t)$, computed by solving (1).

The choice of independent distributions for each unit may initially seem counterintuitive, as real-world generating units are certainly not operated independently. But the goal is to capture the interdependence of unit dispatch within the solution to (1), by an optimal choice of model parameters $a, b, p^{\min}, p^{\max}$. Independent distributions prevent a "lazy" model fit from summarizing covariances rather than fully exploring parameter space.

Furthermore, the use of normal distributions (rather than more complex, multi-model distributions) reinforces the role of a single-node model in the later ensemble model. The single-node model does not attempt to capture grid dynamics under every circumstance. Rather, it aims to model dispatch under a certain regime of operation—a certain range of load, a certain set of congested lines, a certain group of committed units, etc. It is the role of the ensemble model, rather than the single-node model, to account for all of the possible modes. But the normal distribution *does* allow the single-node model to be honest about its uncertainty. Because error scale varies by unit, the model can simultaneously make tightly-bounded predictions for units with a clear history of load following, and weaker predictions for units whose dispatch is less correlated with system demand.

The following definition summarizes the single-node model and its associated terminology:

Definition II.1. Let $\theta = \left(a, b, p^{\min}, p^{\max}, \sigma\right)$ be a vector of parameters for n generating units, with linear and quadratic cost vectors a and $b > 0$, operating limits $0 \le p^{\min} \le p^{\max}$, and error scales $\sigma > 0$. We say that θ is a *single-node model*. We also say that, given some total demand $\sum_{i=1}^{n} p_i^{\min} \le d_t \le \sum_{i=1}^{n} p_i^{\max}$ at time t, the model θ makes a *dispatch prediction* $\hat{p}_{ti}$ for each unit $i \in [n]$, a normal distribution whose density function $f_{\text{sn}}(\hat{p}_{ti})$ is given by (2).

With the proper choice of parameters, a simple, single-node model can capture many important aspects of dispatch behavior, including replicating merit order and identifying marginal generating units, under a certain regime of operation. But, to make the model robust under many different regimes of operation, we must combine insights from many single-node models—using an ensemble model.

B. Ensemble Model

Ensemble models are popular in statistics and machine learning to hedge against selecting the wrong model to describe a dataset. If the size of a training set is small compared to the space of possible models, then a candidate model's predictive accuracy is insufficient to determine whether that model is correct. For instance, several different models might make similar predictions for the training set but vary greatly in later predictions. To overcome this problem, ensemble models combine the predictions from many models, rather than relying on a single model [8]. We use an ensemble of single-node models to accommodate the absence of information that a more complex grid model would require. Without this contextual data, it is impossible to choose which single-node model is best at any given time. Instead, we hedge out bets, combining the predictions of many single-node models instead of trying to select the optimal model.

This analysis uses Bayesian Model Averaging (BMA) to create an ensemble of single-node models. BMA is a simple and intuitive ensemble method. It generates an ensemble prediction by averaging over the predictions of individual models, weighted by their posterior probabilities [9].

The posterior probability density $f_{\text{post}}(\theta \mid X)$ of a single-node model θ given the historical dataset X is calculated using Bayes' rule:

$$f_{\text{post}}(\theta \mid X) = \frac{f_{\text{lik}}(X \mid \theta) f_{\text{prior}}(\theta)}{f_{\text{prior}}(X)} \tag{3}$$

Here $f_{\text{prior}}(\theta)$ and $f_{\text{prior}}(X)$ are the density functions of our prior beliefs for the model and historical data, and $f_{\text{lik}}(X \mid \theta)$ is the likelihood of θ, i.e., the probability of observing the dataset X if θ were the correct model. Because the single-node model predictions are independent across units and time, the likelihood function is simply the product of the probability densities that the model θ assigns to each dispatch observation p_{ti} in X:

$$f_{\text{lik}}(X \mid \theta) = \prod_{t=1}^{T} \prod_{i=1}^{n} f_{\text{sn}}(p_{ti} \mid \theta) \tag{4}$$

The single-node prior $f(\theta)$ captures our existing knowledge of generating unit parameters, including the distributions of linear and quadratic costs that might be present in a real power system, and constraints like $p^{\min} \leq p^{\max}$ and $b > 0$. Finally, the dataset prior $f_{\text{prior}}(X)$ does not depend on the particular model θ in question, so it may be treated as a normalization constant which we need not evaluate. Together, equations (1) through (4) define a posterior density $f_{\text{post}}(\theta \mid X)$ over the space of single-node models, indicating our confidence in each single-node model (given our prior belief $f_{\text{prior}}(\theta)$ and the dataset X).

BMA uses these posterior probabilities to average over model predictions, thereby generating an ensemble prediction. Recall that a single-node model's dispatch prediction $\hat{p}_{ti}$ for a given unit and time is normally distributed, with the density function $f_{\text{sn}}(\hat{p}_{ti} \mid \theta)$ written in (2). Integrating over the space Θ of possible single-node models (i.e., marginalizing out θ) yields the BMA ensemble prediction:

$$f_{\text{bma}}(\hat{p}_{ti} \mid X) = \int_{\Theta} f_{\text{sn}}(\hat{p}_{ti} \mid \theta) f_{\text{post}}(\theta \mid X) \, d\theta \tag{5}$$

BMA hedges against the risk of using the wrong single-node model to make dispatch predictions by taking a weighted average of single-node models, within the integral in (5). But considering many models comes with a cost, as computing a single-node model prediction $f_{\text{sn}}(\hat{p}_{ti} \mid \theta)$ requires solving the quadratic program in (1). This hidden optimization problem makes the integral painful to evaluate. Instead, we approximate $f_{\text{bma}}(\hat{p}_{ti} \mid X)$, estimating the integral numerically with Markov-chain Monte Carlo sampling.

C. Markov Chain Monte Carlo Sampling

Rather than analytically integrating (5), we rely on Monte Carlo methods to approximate BMA ensemble predictions. Using basic Monte Carlo integration, we draw $N \gg 1$ samples of θ distributed according to $f_{\text{post}}(\theta \mid X)$, and approximate the model average by

$$f_{\text{bma}}(\hat{p}_{ti} \mid X) \approx \bar{f}_{\text{bma}}(\hat{p}_{ti} \mid X) = \frac{1}{N} \sum_{s=1}^{N} f_{\text{sp}}(\hat{p}_{ti} \mid \theta_s) \tag{6}$$

where θ_s are the single-model samples $s = 1, 2, \ldots, N$. This average is straightforward to compute. As a linear combination of independent normal distributions, $\bar{f}_{\text{bma}}(\hat{p}_{ti} \mid X)$ is itself a normal distribution, whose mean and variance are computed by averaging the means and variances of the single-node model samples θ_s.

The remaining challenge is to generate these random samples θ_s. Using equations (1) through (4), it is easy to evaluate the posterior probability density $f_{\text{post}}(\theta \mid X)$ of a particular single-node model, up to the unknown normalization constant $f_{\text{prior}}(X)$. Therefore, we can use Markov-chain Monte Carlo (MCMC) methods to generate random samples θ_s, which are distributed according to $f_{\text{post}}(\theta \mid X)$. A detailed exposition on MCMC methods is beyond the scope of this paper, but plenty of material exists on the subject, including [10]. The particular MCMC method we used is the adaptive Metropolis-Hastings algorithm [11].

D. Efficiently Solving Economic Dispatch

Generating many MCMC samples in a timely manner requires an efficient way to evaluate $p^*(d)$ for many total demands d. Because each evaluation of $p^*(d)$ involves solving a quadratic program, this task is not trivial. This section outlines an active set algorithm that solves for many d simultaneously.

Rather than labeling the active subset of constraints in (1) directly, it is more intuitive to define partitions of generating units that the active constraints induce. In particular, the active subset of constraints $p_i \leq p_i^{\max}$ and $p_i \geq p_i^{\min}$ partition the units into three sets:

Definition II.2. The *maxed set* is the set of units $X(p) = \{i \in [n] \mid p_i = p_i^{\max}\}$ that are operating at maximum power.

The *idling set* is the set of units $I(p) = \{i \in [n] \mid p_i = p_i^{\min}\}$ which are idling at minimum power. The *marginal set* consists of the remaining units $M(p) = [n] \setminus (X(p) \cup I(p))$ whose power level is not fixed by an active constraint.

When demand increases by a small amount from d to $d+\delta d$, only the units in $M = M(p^*(d))$ are capable of ramping up to meet the extra demand. This extra demand is, according to the Karush-Kuhn-Tucker conditions of (1), distributed so that the marginal costs of all marginal units are equal. From this property, it can be shown that $p^*(d+\delta d) = p^*(d) + (\delta d)r$, r is a vector of ramp rates given by

$$r_i = \left(\frac{b_i^{-1}}{\sum_{j \in M}^{n} b_j^{-1}} \right) 1_{i \in M} \tag{7}$$

assuming δd is sufficiently small that the marginal set does not change between d and $d + \delta d$.

In other words, $p^*(d)$ is piecewise-affine, with vertices corresponding to marginal set changes. If we can identify the vertices, then we can quickly solve $p^*(d_t)$ for every demand d_t in an arbitrarily large batch, using linear interpolation between the vertices.

The vertices can be computed recursively. The "base case" vertex is clearly $v_1 = p^{\min}$. Now, assuming knowledge of some vertex $v_t = p^*(d_t)$ with $t \geq 1$, it is possible to find d_{t+1}, v_{t+1} by finding the first marginal set change that occurs as demand increases from d_t. Two types of marginal set changes are possible. Either (i) a marginal unit $i \in M$ ramps up to $p^{\max}$, so the constraint $p_i \leq p_i^{\max}$ activates; or (ii) the marginal units' marginal cost surpasses marginal cost $a_i + b_i p_i^{\min}$ of some idling unit $i \in I$, so the constraint $p_i \geq p_i^{\min}$ is no longer active. This recursive computation is shown in more detail in Algorithm 1.

Algorithm 1 NextVertex

Input: $a, b, p^{\min}, p^{\max}, v \neq p^{\max}$
Output: v'

Find maxed, idling, and marginal sets
$X \leftarrow \{i \in [n] \mid v_i = p_i^{\max}\}$
$I \leftarrow \{i \in [n] \mid p_i = p_i^{\min}\}$
$M = [n] \setminus (X \cup I)$
Ensure the marginal set is nonempty
$c_i \leftarrow a_i + b_i v_i, \ \forall i = 1, \ldots, n$
if $M = \emptyset$ **then**
  $M \leftarrow \{i \in [n] \setminus X \mid c_i \leq \min_j \{c_j \mid j \in [n] \setminus X\}\}$
end if
Calculate rate at which units ramp up
$r_i \leftarrow b_i^{-1} / \left(\sum_{j \in M}^{n} b_j^{-1} \right), \ \forall i \in M$
$r_i \leftarrow 0, \ \forall i = [n] \setminus M$
Calculate system lambda and its rate of increase
$\lambda \leftarrow \sum_{j \in M}^{n} c_j r_j$
$r_\lambda \leftarrow \sum_{j \in M}^{n} 2 b_j (r_j)^2$
Calculate load increase until active set changes
$\Delta d_i^1 \leftarrow (p_i^{\max} - v_i) / r_i, \ \forall i \in M$
$\Delta d_i^2 \leftarrow (c_i - \lambda) / r_\lambda, \ \forall i \in I$
Place next vertex at first marginal set change
$\Delta d \leftarrow \min\{\min\{\Delta d_i^1\}, \min\{\Delta d_j^2\}\}$
$v_i' = v_i + r_i \Delta d, \ \forall i \in [n]$

III. NYISO Case Study

We tested the ensemble model with a case study of 20 fossil-fuel-burning generating units under the New York Independent System Operator (NYISO).

A. Experiment

In order to train the model, we used a dataset containing 240 observations of hourly dispatch, starting on January 1st, 2017. This data is from the EPA Air Markets Program database [6].

Prior probability distributions were selected as follows. The linear coefficient prior was normally distributed, with a mean of 35 \$/MW and standard deviation of 10 \$/MW. The quadratic coefficient prior was exponentially distributed, with a mean of 0.01 \$/MW^2. These economic priors were chosen to allow discussion of coefficients on a plausible scale. However, economic dispatch predictions depend only on the shape and width of these prior distributions, not on their locations. Generating limits priors for $p^{\min}$ / $p^{\max}$ were uniform between the minimum / maximum dispatch observations and the average observation, respectively, for each unit. Finally, the error scale prior was an inverse gamma distribution, with a mode of 5 MW and scale of 150. The inverse gamma prior forces error scales to be positive, while its steep decay (due to these mode and scale parameters) prevents the MCMC algorithm from getting stuck on large error scales (where posterior probabilities become insensitive to $a, b, p^{\min}, p^{\max}$).

We used the adaptive Metropolis-Hastings algorithm from [11] to generate 10^6 samples of the posterior distribution of single-node models. Because the randomly selected initial sample often falls within a region of low posterior probability, it is typical to discard a large number of MCMC samples as a "burn-in," so as to avoid biasing the samples. We found that discarding the first 10^5 samples was sufficient. Furthermore, it is common to downsample the MCMC chain to reduce autocorrelation between samples, often by a factor of 100 or more. We kept every 450th sample, resulting in $N = 2000$ remaining samples to perform the Monte Carlo integration in (6). (Keeping more samples for a larger N did not significantly affect the results.)

B. Results

Fig. 1 shows two typical comparisons between the model prediction and observed dispatch, in the 120 hours following the training window. Two categories of units emerge from the model. Some units, like Unit 6, don't tend to vary with total system load. These units are poor candidates for an economic dispatch model—no feasible single-node model is capable of reproducing its dynamics. For these units, the model sets a wide error scale, thereby making weak predictions. Other units, like Unit 12, show explicit load-following behavior. In these cases, the economic dispatch model has more predictive

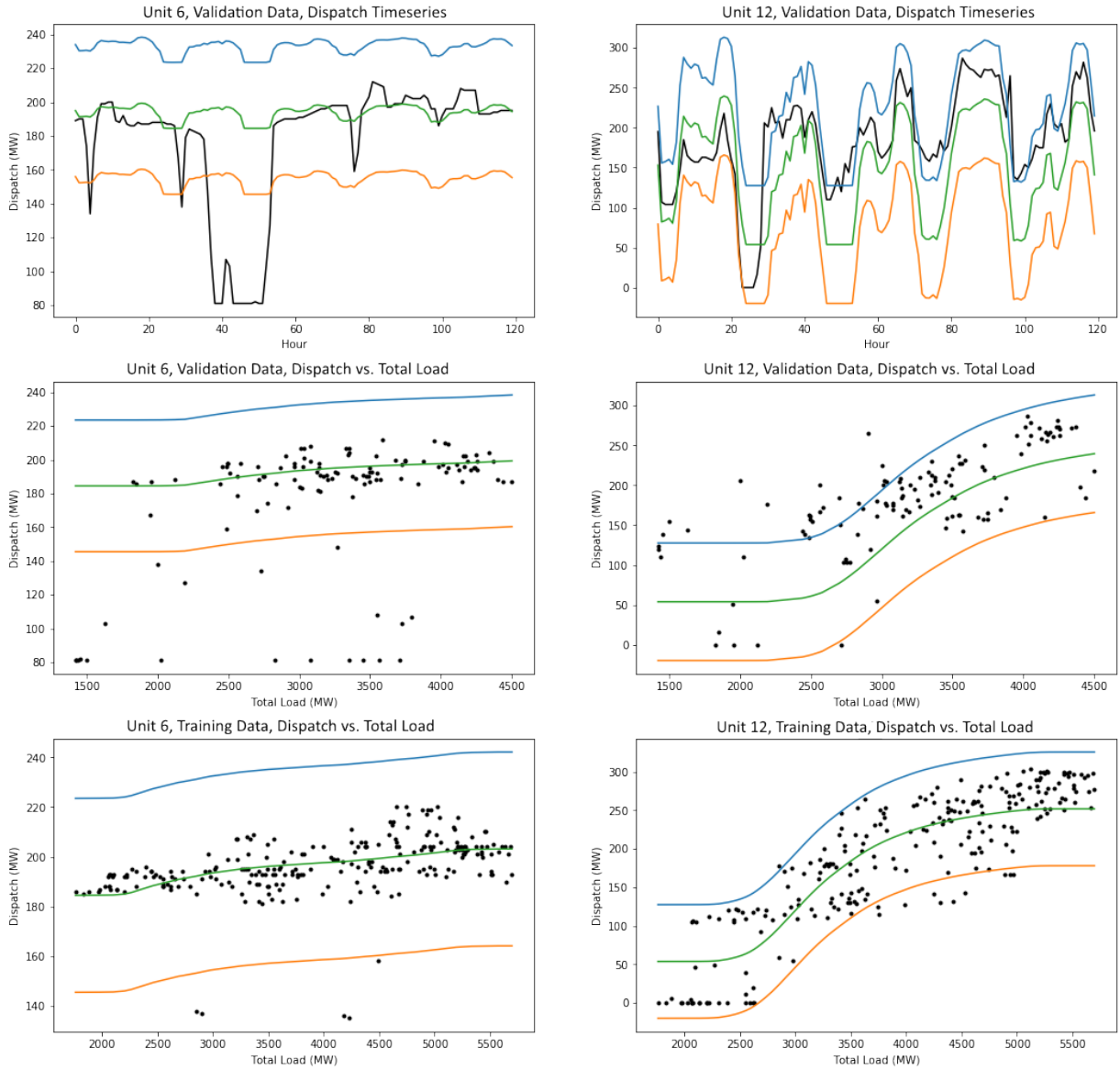

Fig. 1. The left column shows results for a typical unit whose dispatch is uncorrelated with total demand, while the right column shows results for a typical unit with explicit load-following behavior. The top row compares true dispatch (black) to model prediction (green) and the model's $+2\sigma$ and -2σ intervals (blue and orange) as a timeseries, over the 120 hours following the training set. The middle row also compares true dispatch and predicted dispatch during this period, but as a function of total load rather than time. The bottom row compares true dispatch and predicted dispatch during the 240 hour training period.

power. Then the model sets a tighter error scale, making stronger predictions.

C. Economic Predictions

Fig. 2 shows the marginal posterior distribution for parameters a, b, $p^{\min}$, and $p^{\max}$ corresponding to Unit 12. Generally, the marginal distributions for cost coefficients were close to the prior distributions in mean and variance. Unit 12, for instance, had a linear cost range between 30 \$/MW and 37.5 \$/MW, compared to the prior distribution with mean 35 MW and standard deviation 10 \$/MW. Readers hoping to use this analysis for economic espionage will likely be disappointed.

The marginal distributions for NYISO operating costs highlight the role of prior probabilities in the model. The degree to which fitted costs can be interpreted as real-world economic costs depends on the quality of the prior. If little prior information is available about operating costs, one could specify weak priors. The resulting posterior distributions may be far from real-world operating costs, despite yielding accurate dispatch predictions. In the other extreme, one could set strong priors to reflect great prior knowledge of operating costs. In this case, it would be more reasonable to interpret the resulting distributions as real (rather than merely functional) cost coefficients.

Unlike statistical models that look strictly at historical observation, this analysis can predict how economic dispatch would respond to changes in the grid. For example, Fig. 3

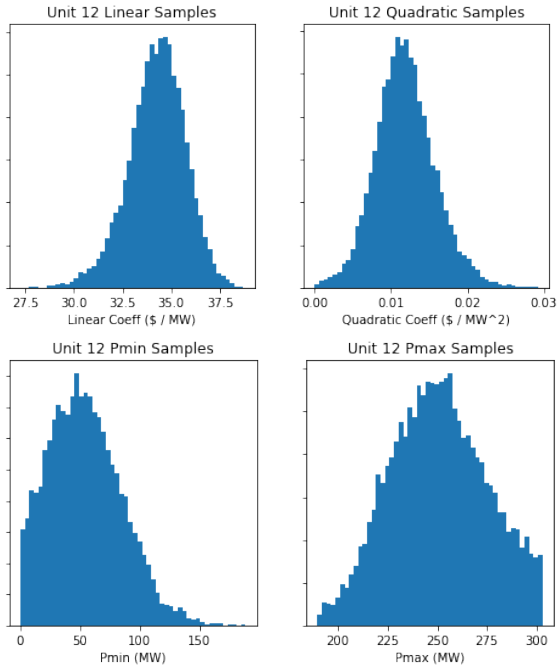

Fig. 2. Marginal posterior distributions of parameters for Unit 12. The upper left shows the linear cost term, the upper right shows the quadratic cost term, and the bottom two plots show $p^{\min}$ and $p^{\max}$.

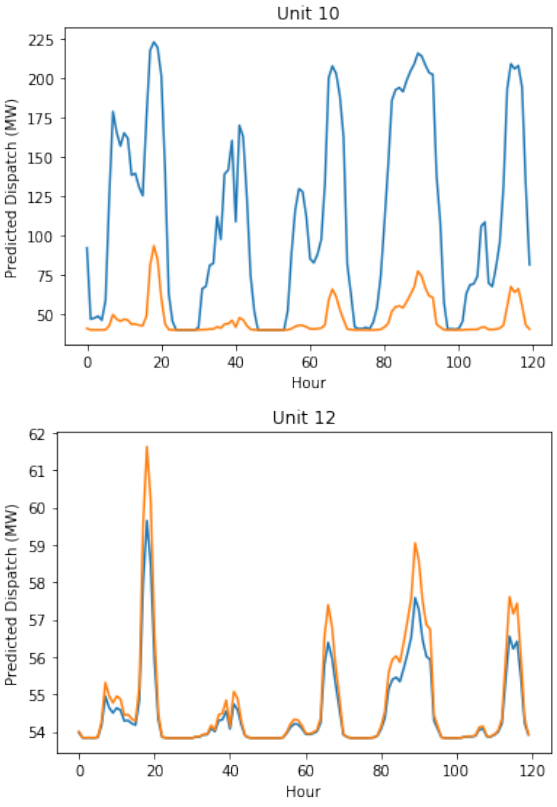

Fig. 3. Model predictions for change in dispatch if the linear cost of Unit 10 is increased by 20%. The top plot compares the predicted timeseries for Unit 10 dispatch without the cost increase (blue), and with the cost increase (orange). The bottom plot shows how Unit 12 is affected by this change, ramping up higher than usual to offset some of the load on Unit 10.

depicts how the model anticipates dispatch would be affected if the linear cost of Unit 10 were increased by 20%. The timeseries in the top plot matches intuition: with Unit 10 more expensive, it is ramped up less often, shifting the load to cheaper units like 12, shown in the bottom plot.

IV. Conclusion

This analysis studies generating unit dispatch through an ensemble average of simple physical models. In this paper, we recognize that the single-node model, despite its simplicity, captures important aspects of dispatch behavior. We show that a collection of many single-node models, using a simple method of averaging, can make meaningful predictions of generating unit dispatch. Furthermore, because the analysis is built from physical models rather than a statistical summary of historical observations, it can still make predictions when certain aspects of the system are changed.

As exploratory research, a theme of this analysis is simplicity: using a simple (Bayesian averaging) technique to combine simple (single-node) physical models whose predictions are simple (Gaussian) distributions. We believe that future work can improve this model's performance by increasing its complexity. A more-complex single-node model might capture the unit commitment process or incorporate ramping limits, and better ensemble models might use other machine learning techniques in place of BMA.

Acknowledgments

The authors gratefully acknowledge the leadership of Eric Lightner, director of the DOE Smart Grid Task Force, who conceived of the calculator and supports this work, and the guidance of the project's industry steering committee, which shapes its specifications and development.

References

[1] B. H. Chowdhury and S. Rahman, "A review of recent advances in economic dispatch," *IEEE Transactions on Power Systems*, vol. 5, no. 4, pp. 1248–1259, 1990.

[2] A. Mahor, V. Prasad, and S. Rangnekar, "Economic dispatch using particle swarm optimization: A review," *Renewable and sustainable energy reviews*, vol. 13, no. 8, pp. 2134–2141, 2009.

[3] Grid Project Impact Quantification Tool (GridPIQ). Pacific Northwest National Laboratory (PNNL). [Online]. Available: https://gridpiq.pnnl.gov/app and https://gridpiq.pnnl.gov/doc.

[4] K. Studarus, T. Hardy, B. Thayer, and R. Pratt, "Quantifying the emissions impacts of smart grid projects with a publically available web calculator," in *Power and Energy Society General Meeting (PESGM), 2016*. IEEE, 2016, pp. 1–5.

[5] EPA. AVoided Emissions and geneRation tool (AVERT). Available: https://www.epa.gov/statelocalenergy/avoided-emissions-and-generation-tool-avert and https://www.epa.gov/statelocalenergy/avert-user-manual.

[6] ——. Air markets program data. Available: https://ampd.epa.gov/ampd/.

[7] A. J. Wood and B. F. Wollenberg, *Power generation, operation, and control*. John Wiley & Sons, 2012.

[8] T. G. Dietterich *et al.*, "Ensemble methods in machine learning," *Multiple classifier systems*, vol. 1857, pp. 1–15, 2000.

[9] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, "Bayesian model averaging: A tutorial," *Statistical Science*, vol. 14, no. 4, pp. 382–401, 1999. [Online]. Available: http://www.jstor.org/stable/2676803

[10] W. R. Gilks, S. Richardson, and D. Spiegelhalter, *Markov chain Monte Carlo in practice*. CRC press, 1995.

[11] H. Haario, E. Saksman, J. Tamminen *et al.*, "An adaptive metropolis algorithm," *Bernoulli*, vol. 7, no. 2, pp. 223–242, 2001.