Stochastic Unit Commitment at ISO Scale: An ARPAe Project

Kwok Cheung, *Member, IEEE,* Yonghan Feng, Dinakar Gade, *Member, IEEE,* Youngrok Lee, Cesar Silva Monroy, *Member, IEEE,* Ignacio Rios, Sarah M. Ryan, *Member, IEEE,* Fabian Rüdel, Jean-Paul Watson, *Member, IEEE,* Roger J-B Wets, and David L. Woodruff, *Member, IEEE*

***Abstract*—We describe a multi-faceted effort to demonstrate the practical implementation of stochastic unit commitment at RTO scale. The inputs for the scheduling engine consist of multiple hourly trajectories of load and variable generation availability together with probabilities that accurately describe their likelihood of occurrence given information available on the day ahead. Preliminary computational results indicate that the methods have promise.**

***Index Terms*—Stochastic Unit Commitment, Stochastic Forecasts**

I. Introduction

THE restructured electric power industry brings new challenges and concerns for the secured operation of stressed power systems. As renewable energy resources, distributed generation, and demand response become significant portions of overall generation resource mix, smarter or more intelligent resource commitment and system dispatch technology is needed to cope with new categories of uncertainty associated with those new energy resources. The need for a new commitment and dispatch system to better handle the uncertainty introduced by the increasing number of new energy resources becomes more and more inevitable.

In North America, almost all Regional Transmission Organizations (RTO) such as PJM, Midwest ISO or ISO New England, are fundamentally reliant on wholesale market mechanisms to optimally commit generating resources and dispatch energy and ancillary services of generation resources to reliably serve the load in the large geographical region. Traditionally, at the heart of a Market Management System (MMS) is a mixed integer security-constrained unit commitment (SCUC) optimization algorithm operating in tandem with a security-constrained economic dispatch (SCED) optimization algorithm. Together these algorithms determine which generation resources will service load, hour by hour, considering costs of start up, shut down, operations, and security constraint reliability criteria.

The uncertainty of generation requirements for maintaining system balancing has been growing significantly due to the penetration of renewable energy resources such as wind power. To deal with such uncertainty, RTOs require not only more accurate point forecasts for demand and variable generation for longer-term prediction beyond real-time, but also descriptions of the uncertainty associated with those forecasts. In this project a new dispatch and commitment engine was proposed to account for these uncertainties in the commitment and dispatch processes.

This project has developed a new probabilistic SCUC software platform that utilizes probabilistic inputs to account for inaccurate forecasts in wind, solar, load use, energy storage, electric vehicles, aggregated demand response or committed generation that fails to fully deploy. The computational platform – the PySP package for stochastic programming [1] that is part of the Coopr optimization software package (https://software.sandia.gov/trac/coopr) – in which we implemented our new probabilistic SCUC formulation is flexible, modular, and extensible.

Devising solution methods that are scalable to ISO-size, managing the number of scenarios, and adapting the model to account for emerging market structures has presented algorithmic and computational challenges. In this paper, we describe some of those challenges and how we have addressed them.

Day-ahead scheduling means that decisions for a day d are made on day $d-1$ based on forecasts of uncertain quantities such as hourly load and renewables output that, in turn, are based on forecasts of more fundamental quantities, primarily weather. Traditional, deterministic unit commitment optimization relies on point forecasts or expected-value quantities – representing a single time series for each forecast quantity. Uncertainty associated with such forecasts is dealt with in part by maintaining a significant level of generation reserves, which enable compensation as inevitable deviations from predicted quantities occur during day d.

Stochastic unit commitment models (see [2], [3], [4], [5] for a representative sample), on the other hand, assume the availability of a number of forecast scenarios, each representing a distinct time-series. Throughout, we use the term *scenario* in a narrow sense, representing a full specification of all random data required to specify 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 the generation reserve margins and consequently reduce overall system operation costs [6]. While previous research has

Kwok Chueng is with Alstom Grid, in Redmond, Washington. Dinakar Gade is with Sabre Holdings. Sarah Ryan, Yonghan Fen, and Youngrok Lee are with Iowa State University, in Ames, Iowa. Cesar Silva Monroy and Jean-Paul Watson are with Sandia National Laboratories in Albuquerque, New Mexico. Roger Wets and David Woodruff are with the University of California Davis, in Davis, California. Fabian Rüdel is with the University of Duisburg-Essen. Ignacio Rios is with the University of Chile.

Manuscript received TBD 2013; revised TBD.

provided compelling evidence for the promise of cost savings, it has not addressed some of the practical issues involved in implementing stochastic unit commitment day-to-day at ISO scale. This project is aimed at bridging that gap.

II. Simulation at ISO Scale

Our goal in this project is to estimate the potential for cost savings if stochastic unit commitment is used at ISO scale. We are not concerned with market design issues, but rather with assessing how much money could be saved without regard to how (e.g., to which participants) the savings is allocated. To achieve this objective, we simulate – as closely as possible – the year 2011 as experienced by ISO New England (ISO-NE). We consider only the so-called reliability unit commitment problem, in conjunction with hourly economic dispatch. However, in order to test scalability and to estimate potential total cost savings, we treat all generators as being available for commitment.

Before executing the simulation, our software computes parameters for load forecast models as a function of forecasted weather variables and the errors associated with load forecasts, using data from 2009 and 2010. As the simulation through 2011 proceeds, our load forecast models are updated using data from days in 2011 that have already been simulated, i.e., days that are now historical. For each day, we make use of weather forecasts for the next day to compute (1) a forecast expected load for the next day and (2) a set of scenarios with associated probabilities representing possible load time-series for the next day. The former is used as input to a deterministic UC model, while the latter are used as input to a stochastic UC model. The actual loads observed by ISO-NE in 2011 are used as the simulated scenario, and costs are computed as a function of reliability unit commitment fixed costs and actual production costs as computed by iteratively solving an economic dispatch given the observed load time-series.

Because our goal is to estimate cost savings, and not to produce an operational forecasting and scenario generation methodology to be directly applied at an ISO, 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 [7], but their use is beyond the present scope. Additionally, we also omit August 28-30, 2011, due to the impact of a hurricane.

Load scenario generation is performed using methods described in [8], which extends an earlier version presented in [9] and is a specific implementation of general methods described in [10].

Optimization is performed using a baseline deterministic model that is an extension of the Carrion and Arroyo unit commitment model [11], which has been validated against an Alstom Grid test case. For baseline simulation, we simply solve the unit commitment model using forecasted load and wind quantities, with additional static reserve requirements. In simulations involving stochastic unit commitment, a two-stage stochastic programming model based on our deterministic unit commitment model is solved, using generated load and wind scenarios as input. As we discuss subsequently in Section III, solution of large-scale stochastic unit commitment problems presents a serious computational challenge. We have addressed this challenge using an implementation of Rockafellar and Wets' progressive hedging algorithm [12], to achieve solutions to stochastic unit commitment problems with reasonable numbers of scenarios in tractable run-times. Our solution methods are fully described in [13], which extends earlier work reported in [14].

Using publicly available data and engineering knowledge about the physical characteristics of thermal generators, we constructed data that approximates a description of the generator fleet at ISO-NE including heat-rates, ramping characteristics, minimum up/down time requirements, and startup costs. Our instance consists of 326 thermal generators.

Wind power must be simulated in a different way. In the case of load we have access to sufficient data to construct our own forecast technology based on weather forecasts and we have the data to use it in a simulation of 2011. Furthermore, there was not significant wind power in the ISO-NE region in 2011, so we must use some other source of stochastic wind power data.

The Bonneville Power Administration provides some data at bpa.gov that we make use of.. We used data from "Wind Power Forecasting Data" /www.bpa.gov/Projects/Initiatives/Wind/Pages/Wind-Power-Forecasting-Data.aspx as forecast data. For actuals we used the "Data for BPA Balancing Authority Total Load, Wind Gen, Wind Forecast, Hydro, Thermal, and Net Interchange" transmission.bpa.gov/Business/Operations/Wind/default.aspx. We "pretend" that this wind power was generated in the ISO-NE region in 2011. We scale it up or down depending on how much wind penetration we seek to simulate.

For this work, a subset of the load forecasting methodology is used. We are given forecasts $r_{h,d}$ of total wind power for each hour h of each day d as well as actual, observed total wind power $w_{h,d}$. When generating load scenarios we are able to create forecasts conditional on the error category, but in the case of wind, we are not able to create our own forecasts from forecasts of fundamental weather data.

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

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

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 $\{H_i, \cdots, H_{i+1}\}$. For each partition boundary H_i, we compute the error observed regression error ϵ_i^d for each day $d \in \mathcal{D}$ as:

$$\epsilon_i^d = w_{H_i,d} - r_{H_i,d}$$

We estimate the distribution of these errors, $f_{\varepsilon_i}(\cdot)$, by fitting an exponential epi-spline [15], [16]. Figure 1 provides a stylized illustration.

The error densities $f_{\varepsilon_i}(\cdot)$ serve as the primary input to the scenario generation process. Scenario generation begins with the specification of a set of distribution cutting points $C = \{c_z\}_{z=1}^{|C|}$, subject to $c_1 = 0.0$ and $c_{|C|} = 1.0$. For each partition i we then calculate the expected value of the error in each

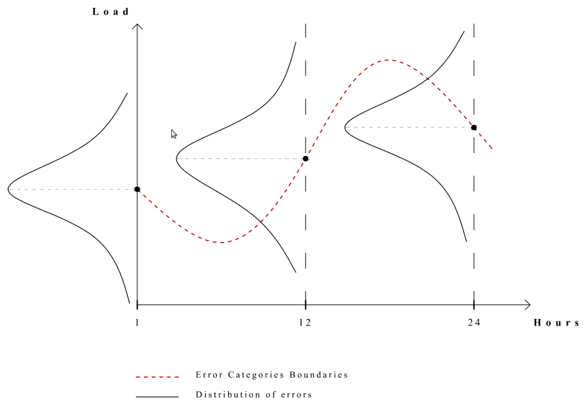

Fig. 1. Illustrative Forecasts and Error Distributions with Hour Partition $H = \{1, 12, 24\}$.

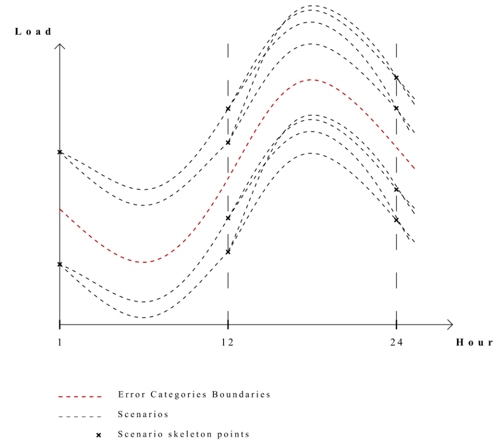

Fig. 2. Illustrative Scenario Paths.

interval defined by a pair of adjacent cutting points:

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

so ξ_i^z is the expected error for cutting point z at hour i. The number of cutting points can vary per hour

Given forecasts $(\hat{r}_{h,i}^d)$ for each hour, day, and partition boundary, we compute wind power at the partition boundaries via:

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

For each hour H_i, this step yields $|C|$ forecast load samples. The final step in our scenario generation process is to connect these samples in 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 2. Under this methodology, the number of paths (i.e., scenarios) generated is equal to

$$(|C| - 1)^{|H|-1},$$

such that the number of scenarios is dictated by the values of the referenced parameters. Further, the generation process is deterministic, given a fixed set of historical input data. The product of the cutting points gives the probability of the scenario.

The processes described above allow the user to generate any specified number of load and variable generation scenarios, to approximate the underlying stochastic processes as closely as desired. The resulting set of scenarios may include some redundancies when viewed from the perspective of the unit commitment problem. Because the computational effort for solving the unit commitment problem is almost directly proportional to the number of scenarios considered, we have developed a custom scenario reduction method that selects a representative subset of scenarios based on their impact on the unit commitment decisions and resulting dispatch costs [17].

III. Preliminary Results

Our experiments are executed on a commodity high-end workstation, consisting of eight 8-core AMD Opteron 6278 2.4GHz processors with 512GB of RAM. Such a workstation is representative of the type of resource that is likely to be currently available or available in the near term to typical utilities and ISOs, and can be purchased for less than $20K USD. The platform allows for modest-scale parallelism, specifically in the executing of our PH solution algorithm. Our deterministic unit commitment model is coded in the Pyomo [18] algebraic modeling library. Our stochastic unit commitment models are expressed in PySP [1], which also provides the base PH implementation used in our studies.

We first report on solve times associated with the PH algorithm on stochastic UC instances associated with our ISO-NE test case. Mirroring the results we previously reported for the simpler WECC-240 test case (with $\approx$ 100 thermal generators) [13], we are able to obtain solutions to our ISO-NE test case in less than 30 minutes of wall clock time in most cases. The exceptional cases involve larger than expected numbers of PH iterations, which can be mitigated by various techniques [19]. Our tests involve 50 or 100 scenarios for each stochastic UC solve. Run-times and general PH behavior are largely independent of scenario features such as variability. Bounds are obtained using a novel extension of the PH algorithm, as reported in [20]. Mirroring our prior results on WECC-240, the optimality gaps on PH solutions to our ISO-NE test cases do not exceed 2%.

We have performed limited simulations of the ISO-NE system for May and June of 2011, focusing strictly on stochastic load. As indicated above, we consider 50 load scenarios for each day-ahead data set. Empirical evidence indicates this quantity is sufficient to ensure out-of-sample stability of the solution, largely due to our ability to approximate – as opposed to sample – the underlying stochastic process. Relative to the simulation considering deterministic reliability unit commitment, we observe approximately 1.2% savings in terms of energy production costs when using stochastic unit commitment. This quantity increases to over 2.3% when forced outages (sampled based on their empirical forced outage rates) are considered in the load scenarios. These results, while preliminary, are consistent with those previously reported by [6] on smaller test cases, with significantly smaller numbers of scenarios.

More extensive investigation, particularly during peak summer months, is required before comprehensive conclusions can be drawn. However, the preliminary evidence is promising, indicating that stochastic unit commitment – when coupled with solvers that achieve tractable run-times – can achieve significant cost savings in practice. Further, we observe that we are not factoring in the cost of ancillary services, which will further inflate the cost savings associated with stochastic unit commitment – due to the reductions in reserve requirements afforded by the approach.

Our experiments concerning cost savings with wind scenarios are more limited, due to increased difficulty of creating stochastic process models of wind power. Our preliminary experiments have focused on the use of the BPA wind power scenarios previously described in Section II, which we scale to represent various hypothesized penetration levels at ISO-NE. For cases representing approximately 20% penetration levels, we are observing 3% cost savings relative to the deterministic approach for stochastic unit commitment. However, these savings are computed on a pair of weeks in the summer of 2011, and therefore represent a biased and incomplete sample. We are in the process of more comprehensive explorations, including extended time periods and variable penetration levels. Finally, we note that we have presently focused on load and wind power uncertainty independently, and will ultimately investigate ”crossed” scenarios when the univariate simulations are complete.

IV. Conclusion and In-Progress Research

We have described a multi-faceted effort to demonstrate the practical implementation of stochastic unit commitment at RTO scale. The inputs for the scheduling engine consist of multiple hourly trajectories of load and variable generation availability together with probabilities that accurately describe their likelihood of occurrence given information available on the day ahead. A custom scenario reduction method identifies a minimal set of trajectories to consider in the scheduling process. The engine efficiently exploits parallel processors to quickly identify a near-optimal schedule, together with a bound on its deviation from optimality. The cost savings have been rigorously assessed by simulating this process over a year of operation and comparing the cost performance of the daily commitment schedules obtained by stochastic unit commitment with the corresponding schedules obtained from the usual deterministic optimization with fixed reserves.

Acknowledgment

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, and by Sandia's Laboratory Directed Research and Development program.

References

[1] J.-P. Watson, D. Woodruff, and W. Hart, “Pysp: Modeling and solving stochastic programs in python,” *Mathematical Programming Computation*, vol. 4, no. 2, 2012.
[2] S. Takriti, J. Birge, and E. Long, “A stochastic model for the unit commitment problem,” *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1497–1508, 1996.
[3] Q. Zheng, J. Wang, P. Pardalos, and Y. Guan, “A decomposition approach to the two-stage stochastic unit commitment problem,” *Annals of Operations Research*, To Appear.
[4] P. Ruiz, C. Philbrick, and P. Sauer, “Modeling approaches for computational cost reduction in stochastic unit commitment formulations,” *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 588–589, 2010.
[5] A. Papavasiliou and S. Oren, “A stochastic unit commitment model for integrating renewable supply and demand response,” in *Proceedings of the 2012 IEEE Power and Energy Society Meeting*, 2012.
[6] P. Ruiz, R. Philbrick, E. Zack, K. Cheung, and P. Sauer, “Uncertainty management in the unit commitment problem,” *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 642–651, 2009.
[7] T. Hong, “Electric load forecasting with holiday effect,” https://sites.google.com/site/hongtao01/courses/holiday, 2012, presented at AEIC Load Research Workshop 2012, Orlando, FL.
[8] Y. Feng, I. Rios, S. M. Ryan, K. Spurkel, J.-P. Watson, R. J.-B. Wets, and D. L. Woodruff, “Scalable stochastic unit commitment, part 1: Load scenario generation,” *IEEE Transactions on Power Systems*, Under Review.
[9] Y. Feng, D. Gade, S. M. Ryan, J.-P. Watson, R. J.-B. Wets, and D. L. Woodruff, “A new approximation method for generating day-ahead load scenarios,” in *Proceedings of the 2013 IEEE Power and Energy Society General Meeting*, 2013.
[10] R. J.-B. Wets, I. Rios, and D. Woodruff, “Multi-period forecasting and scenario generation with limited data,” University of California Davis, Davis CA 95616, Tech. Rep., 2013.
[11] M. Carrion and J. Arroyo, “A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem,” *IEEE Transactions on Power Systems*, vol. 21, no. 3, 2006.
[12] R. Rockafellar and R. J.-B. Wets, “Scenarios and policy aggregation in optimization under uncertainty,” *Mathematics of Operations Research*, pp. 119–147, 1991.
[13] K. Cheung, D. Gade, C. S. Monroy, S. M. Ryan, J.-P. Watson, R. J.-B. Wets, and D. L. Woodruff, “Scalable stochastic unit commitment, part 2: Solver performance,” *IEEE Transactions on Power Systems*, Under Review.
[14] S. Ryan, R. Wets, D. Woodruff, C. Silva-Monroy, and J. Watson, “Toward scalable, parallel progressive hedging for stochastic unit commitment,” in *Proceedings of the 2013 IEEE Power and Energy Society General Meeting*, 2013.
[15] J. Royset and R. Wets, “Epi-splines and exponential epi-splines: Pliable approximation tools,” University of California Davis, Tech. Rep., 2012.
[16] ——, “Fusion of hard and soft information in non-parametric density estimation,” University of California Davis, Tech. Rep., 2012.
[17] Y. Feng and S. M. Ryan, “Scenario reduction for stochastic unit commitment with wind penetration,” in *Submitted to the Proceedings of the 2014 IEEE Power and Energy Society General Meeting*, 2013.
[18] W. Hart, J. Watson, and D. Woodruff, “Pyomo: Modeling and solving mathematical programs in python,” *Mathematical Programming Computation*, vol. 3, no. 3, 2011.
[19] J. Watson and D. Woodruff, “Progressive hedging innovations for a class of stochastic mixed-in teger resource allocation problems,” *Computational Management Science*, vol. 8, no. 4, 2011.
[20] D. Gade, G. Hackebeil, S. Ryan, J. Watson, R. Wets, and D. Woodruff, “Obtaining lower bounds from the progressive hedging algorithm for stochastic mixed-integer programs,” *Under Review*, 2013.