OPERATIONS RESEARCH
Vol. 61, No. 3, May–June 2013, pp. 578–592
ISSN 0030-364X (print) | ISSN 1526-5463 (online)

http://dx.doi.org/10.1287/opre.2013.1174

Multiarea Stochastic Unit Commitment for High Wind Penetration in a Transmission Constrained Network

Anthony Papavasiliou
Department of Mathematical Engineering, Center for Operations Research and Econometrics, Catholic University of Louvain, B-1348 Louvain la Neuve, Belgium, anthony.papavasiliou@uclouvain.be

Shmuel S. Oren
Department of Industrial Engineering and Operations Research, University of California at Berkeley, Berkeley, California 94720, oren@ieor.berkeley.edu

In this paper we present a unit commitment model for studying the impact of large-scale wind integration in power systems with transmission constraints and system component failures. The model is formulated as a two-stage stochastic program with uncertain wind production in various locations of the network as well as generator and transmission line failures. We present a scenario selection algorithm for selecting and weighing wind power production scenarios and composite element failures, and we provide a parallel dual decomposition algorithm for solving the resulting mixed-integer program. We validate the proposed scenario selection algorithm by demonstrating that it outperforms alternative reserve commitment approaches in a 225 bus model of California with 130 generators and 375 transmission lines. We use our model to quantify day-ahead generator capacity commitment, operating cost impacts, and renewable energy utilization levels for various degrees of wind power integration. We then demonstrate that failing to account for transmission constraints and contingencies can result in significant errors in assessing the economic impacts of renewable energy integration.

Subject classifications: unit commitment; stochastic programming; wind power; transmission constraints.
Area of review: Environment, Energy, and Sustainability.
History: Received July 2011; revisions received August 2012, October 2012, January 2013; accepted February 2013. Published online in *Articles in Advance* May 24, 2013.

1. Introduction

The large-scale integration of renewable energy resources such as wind and solar power in power systems is limited by two adverse characteristics of renewable power supply. Certain renewable sources are unpredictable and, in contrast to conventional generators, the energy source is not controllable. Because of the need for maintaining a constant balance between supply and demand in the system and also because of the prohibitively high cost of electricity storage, these two adverse characteristics greatly influence our ability to utilize renewable resources at a large scale by (a) imposing costs on power system operations,[1] (b) reducing the amount of renewable power that can be absorbed in the network and, (c) necessitating the deployment of reserves that can ensure the reliable operation of the system. The focus of this paper is on quantifying these impacts in the presence of transmission network congestion and contingencies through the use of a stochastic unit commitment model.

Because of the large-scale integration of renewable energy sources, the solution to the stochastic unit commitment problem is becoming highly relevant to the electric power industry. Given the current organization of electricity markets (e.g., the two-settlement systems in California, Texas, the Midwest, Pennsylvania-Jersey-Maryland, New York, the United Kingdom, the Nordic region, and New Zealand to name a few examples), stochastic unit commitment simulates an idealized two-settlement system and provides a good indication as to how to procure reserves in the system. Moreover, the model is relevant in an operational setting in the residual unit commitment process. Residual unit commitment takes place after the day-ahead market closes, and in this phase the independent system operator evaluates if the market has provided sufficient resources to meet the needs of the real-time electricity market in order to deal with supply uncertainty, demand uncertainty, and network element failures.

The paper contributes to existing literature in terms of both modeling and methodology. The modeling contribution of the paper relies on the use of authentic and detailed data for evaluating the cost of large-scale renewable energy integration and its impact on day-ahead reserve requirements while accounting for transmission constrains and component failures. Because of the fact that the model includes transmission constraints, it is necessary to model renewable power production in each region separately and capture the spatial correlations of renewable power

production, as well as numerous other features of the renewable power production process (temporal autocorrelation, nonlinearity in the conversion of wind speed to wind power, non-Gaussian distribution of wind speed data). The methodological contribution of the paper is the presentation of a computationally tractable approach for implementing stochastic unit commitment and comparing it to alternative day-ahead reserve commitment approaches. Computational tractability is achieved by the development of a scenario selection procedure inspired by importance sampling for reducing the representation of uncertainty, as well as the implementation of a parallel decomposition algorithm for solving the stochastic unit commitment problem. In the following section we review relevant literature regarding the modeling and computational methods for assessing the impacts of uncertainty on power system operations, and we elucidate how our paper contributes in each of these areas.

2. Literature Review

2.1. Modeling

The operation of power systems under uncertainty can be perceived as a multistage decision process, where resources are committed prior to an operating interval and decisions are updated as the operating interval approaches and conditions in the system are revealed. We follow the approach of Ruiz et al. (2009b) in modeling system operations as a two-stage decision process. The first-stage decisions represent the day-ahead commitment of generators based on demand forecasts. Subsequently, uncertainty is revealed and in the second stage the commitment of fast-responding units and the dispatch of all committed units is updated in order to respond to system conditions. In deregulated power systems (e.g., the two-settlement markets referenced earlier), the first and second stage of the model can be interpreted as simulating the day-ahead and real-time markets, respectively.

A stochastic unit commitment model is especially appropriate for quantifying the impacts of renewable energy integration on renewable energy utilization, operating costs, and day-ahead generating capacity requirements. Renewable power utilization is an explicit decision variable in the problem, operating costs are quantified in the objective function of the problem, and reserve capacity requirements are quantified indirectly by the fact that the commitment of generators is an endogenous decision variable in the problem. The estimation of day-ahead generating capacity requirements is an especially challenging aspect of our analysis. To commit reserves, system operators and analysts often resort to suboptimal ad hoc rules in a deterministic formulation of the unit commitment problem. These commitment policies are not calibrated to an environment with large amounts of renewable resources and may result in inefficiencies and miscalculation of the economic impacts of renewable energy integration. These drawbacks have motivated the work of Piwko et al. (2010) who propose adaptive reserve commitment approaches in order to overcome the drawbacks of current practice in an environment of large-scale renewable energy integration.

2.1.1. Transmission Constraints and Contingencies. Transmission constraints result in shedding power supply because of the fact that the system, often, cannot support power flows during periods of increased renewable energy supply. This phenomenon has been observed in previous studies by Sioshansi and Short (2009) and Morales et al. (2009b) and it may result in waste of renewable energy supply, thereby undermining the benefits of renewable energy integration. Moreover, transmission constraints greatly complicate the task of determining reserves in various locations of the system and render the use of stochastic unit commitment especially relevant. Arroyo and Galiana (2005) demonstrate that an ad hoc allocation of reserves in various locations of a transmission-constrained network results in suboptimal system performance. The complex influence of transmission constraints on locational reserve requirements is also demonstrated by Galiana et al. (2005), as well as Bouffard et al. (2005), on smaller scale systems.

Alternative formulations of the unit commitment problem have been recently proposed in order to address uncertainty. This work is motivated by the fact that system operators tend to operate the system so as to protect against worst-case outcomes, and also by the fact that the stochastic programming formulation requires detailed knowledge about the underlying uncertainty since it requires the generation and weighting of scenarios. Ozturk et al. (2004) formulate a chance-constrained optimization of the unit commitment problem without transmission constraints and ramping constraints. Jiang et al. (2012) use a robust optimization formulation of the unit commitment problem, where locational demand is assumed to obey the polyhedral and cardinal uncertainty models defined by Bertsimas and Sim (2004). The authors present a Benders' decomposition algorithm for solving the problem with transmission and ramping constraints. A similar formulation is proposed by Bertsimas et al. (2013). However, neither of these models accounts for contingencies explicitly.

The work presented in this paper extends the model of Papavasiliou et al. (2011) by introducing transmission network constraints as well as generator and transmission line failures. Both features are crucial in accurately assessing the economic impact of renewable energy integration. As we demonstrate in §7, ignoring transmission constraints and contingencies can result in a significant overestimation of the day-ahead generation capacity savings and operating cost savings of wind power. In line with existing deterministic and stochastic unit commitment literature, a DC lossless representation of the transmission network has been used that ignores reactive power, voltage limits, and line losses. The results of such optimization are, therefore, regarded as advisory and in practical applications further post-processing of such results is used to assure AC power flow feasibility and voltage stability.

2.1.2. Multiarea Wind Modeling. Although transmission constraints and contingencies cannot be ignored when analyzing the economic impacts of renewable energy integration, they also introduce certain modeling and computational challenges that are addressed in this paper. The modeling of wind power production as an aggregate resource located in a single bus is not adequate. Instead, it is necessary to develop a wind power production model that accounts for the nonlinear relationship between wind speed and wind power, the diurnal and seasonal characteristics of wind power supply, as well as the spatial and short-term temporal correlation of wind power supply.

Early work on the time series modeling of wind power production in order to address short-term temporal correlation and the nonlinear relationship between wind speed and wind power was performed by Brown et al. (1984) and later by Torres et al. (2005). In §5 we present a multiarea wind production model that captures both the temporal as well as spatial correlations of wind power production that are observed in the data set. The modeling and calibration methodology described in this section extends the approach of Morales et al. (2010) by also accounting for the diurnal and seasonal patterns of wind power supply.

2.2. Methodology

Recently, various studies have utilized unit commitment models for assessing the impacts of wind integration. A variety of modeling approaches and solution strategies are employed in these studies. Ruiz et al. (2009a) use the two-stage stochastic unit commitment model of Ruiz et al. (2009b) to analyze the impact of wind integration in the Colorado power system, without considering transmission constraints. A method employed in Ruiz et al. (2009b) that we also adopt in our paper is to classify generators as either "fast" or "slow" resources. Slow generators are committed in the first stage, and fast generators can be committed in the second stage of recourse. Sioshansi and Short (2009) study the impact of wind integration in the Electricity Reliability Council of Texas (ERCOT) system with a deterministic unit commitment model that includes transmission constraints. Wang et al. (2008) use a two-stage formulation that ensures a feasible dispatch in the second stage for all possible wind production outcomes given first-stage commitment decisions. Recent work by Constantinescu et al. (2011) also uses the two-stage stochastic programming formulation of Ruiz et al. (2009b). The authors test their model on a system with 10 generators without transmission constraints. Tuohy et al. (2009) develop a stochastic unit commitment model that accounts for load and wind uncertainty and perform a simulation of the Irish power system. Morales et al. (2009a) use a two-stage stochastic programming model for studying wind integration in the IEEE RTS 96 model of Grigg et al. (1999) that includes transmission constraints but not contingencies. The model assumes explicit bids for spinning and nonspinning reserves. A similar two-stage stochastic programming model is used by Bouffard and Galiana (2008) for a case study of wind integration in a test system with four generators without transmission constraints and contingencies. With the exception of Wang et al. (2008), none of the aforementioned wind integration studies employ decomposition techniques, which limits the size of the models that are studied. By comparison, our model uses a scenario selection technique inspired by importance sampling in conjunction with a dual decomposition algorithm implemented in a parallel environment in order to examine a reduced model of California with 130 generators, 375 lines, and 225 buses.

2.2.1. Scenario Selection. In §6 we address the issue of scenario selection. As in the case of locational wind power production, the introduction of generation contingencies requires identifying generator failures depending on where they occur in the network. This should be contrasted with modeling the loss of aggregate generation capacity which is a common approach in simpler models without transmission constraints such as Takriti et al. (1996) and Ruiz et al. (2009a). Because of the fact that wind production uncertainty is interleaved with contingencies, an enormous set of scenarios needs to be considered. To develop a systematic approach for selecting and weighing scenarios, we introduce a scenario selection algorithm inspired by importance sampling techniques that selects uncertain scenarios on the basis of their likelihood of occurrence and the severity of their impact on operating costs. The algorithm that we present formalizes the intuition used in the scenario selection algorithm of Papavasiliou et al. (2011) and is shown to outperform alternative deterministic and stochastic day-ahead commitment rules in all the case studies that are presented in §7.

2.2.2. Decomposition of the Stochastic Unit Commitment Problem. Even with a limited scenario set, the resulting problem requires the use of decomposition techniques in order to achieve computational tractability. The use of decomposition algorithms for solving the stochastic unit commitment problem was pioneered by Takriti et al. (1996), who use a multistage stochastic unit commitment formulation for studying load uncertainty and generator failures. The authors use the progressive hedging algorithm of Rockafellar and Wets (1991) to decompose the stochastic formulation to single-period subproblems. In Carpentier et al. (1996) the authors use the augmented Lagrangian algorithm to decompose a multistage stochastic program to single-generator subproblems. Nowak and Römisch (2000) develop a Lagrangian decomposition algorithm for optimizing the operations of a hydrothermal system under load uncertainty. Shiina and Birge (2004) develop a column generation algorithm for decomposing a multistage stochastic program into single-generator subproblems. In §4 we present an extension of the decomposition algorithm of Papavasiliou et al. (2011) for decomposing the stochastic unit commitment problem to individual scenarios and solving the more general model presented in this paper.

3. Unit Commitment and Economic Dispatch

In our study we consider a set of generation resources G that is partitioned among a set of slow generators G_s and fast generators G_f. Uncertainty is modeled as a discrete set of realizations S. In the stochastic unit commitment model, the commitment of slow generators is a first-stage decision that cannot be altered in the second stage, whereas the commitment of fast generators and the production level of all generators can be adjusted once a realization $s \in S$ is observed. Analogously, in the deterministic model, which follows Sioshansi and Short (2009), slow generators can only provide slow reserves whereas fast generators can provide slow as well as fast reserves.

3.1. Notation

We begin this section by introducing the notation that is used in the subsequent models.

Sets

G: set of all generators, G_s: subset of slow generators, G_f: subset of fast generators;
G_n: set of generators that are located in bus n;
S: set of scenarios, T: set of time periods, L: set of lines, N: set of nodes;
$LI_n = \{l \in L\colon l = (k, n), k \in N\}$, $LO_n = \{l \in L\colon l = (n, k), k \in N\}$;
IG: set of import groups, IG_j: set of lines in import group j.

Decision variables

u_{gst}: commitment, v_{gst}: startup, p_{gst}: production of generator g in scenario s, period t;
θ_{nst}: phase angle at bus n in scenario s, period t;
w_{gt}: commitment, z_{gt}: start-up of slow generator g in period t;
s_{gt}: slow reserve, f_{gt}: fast reserve provided by generator g in period t;
e_{lst}: power flow on line l in scenario s, period t.

Parameters

π_s: probability of scenario s
K_g: minimum load cost, S_g: start-up cost, C_g: marginal cost of generator g;
D_{nst}: demand in bus n, scenario s, period t;
P^+_{gs}, P^-_{gs}: minimum and maximum capacity of generator g in scenario s;
R^+_g, R^-_g: minimum and maximum ramping of generator g;
UT_g: minimum up time, DT_g: minimum down time of generator g;
T^{req}_t: total reserve requirement, S^{req}_t: slow reserve requirement in period t;
B_{ls}: susceptance of line l in scenario s;
TC_l: maximum capacity of line l;
FR_g: fast reserve limit of generator g;
IC_j: maximum capacity of import group j;
γ_{jl}: polarity of line l in import group j.

3.2. Stochastic Unit Commitment

The stochastic unit commitment (SUC) problem can be stated as follows:

$$\text{(SUC):}\quad \min \sum_{g\in G}\sum_{s\in S}\sum_{t\in T}\pi_s(K_g u_{gst} + S_g v_{gst} + C_g p_{gst}) \tag{1}$$

$$\text{s.t.}\quad \sum_{l\in LI_n} e_{lst} + \sum_{g\in G_n} p_{gst} = D_{nst} + \sum_{l\in LO_n} e_{lst}, \quad n \in N,\ s \in S,\ t \in T; \tag{2}$$

$$e_{lst} = B_{ls}(\theta_{nst} - \theta_{mst}), \quad l = (m, n) \in L, \quad s \in S,\ t \in T; \tag{3}$$

$$-TC_l \leqslant e_{lst} \leqslant TC_l, \quad l \in L,\ s \in S, \quad t \in T; \tag{4}$$

$$P^-_{gs} u_{gst} \leqslant p_{gst} \leqslant P^+_{gs} u_{gst}, \quad g \in G,\ s \in S, \quad t \in T; \tag{5}$$

$$-R^-_g \leqslant p_{gst} - p_{gs,t-1} \leqslant R^+_g, \quad g \in G,\ s \in S, \quad t \in T; \tag{6}$$

$$\sum_{q=t-UT_g+1}^{t} z_{gq} \leqslant w_{gt}, \quad g \in G_s,\ t \geqslant UT_g; \tag{7}$$

$$\sum_{q=t+1}^{t+DT_g} z_{gq} \leqslant 1 - w_{gt}, \quad g \in G_s, \quad t \leqslant |T| - DT_g; \tag{8}$$

$$\sum_{q=t-UT_g+1}^{t} v_{gsq} \leqslant u_{gst}, \quad g \in G_f,\ s \in S, \quad t \geqslant UT_g; \tag{9}$$

$$\sum_{q=t+1}^{t+DT_g} v_{gsq} \leqslant 1 - u_{gst}, \quad g \in G_f,\ s \in S, \quad t \leqslant |T| - DT_g; \tag{10}$$

$$z_{gt} \leqslant 1, \quad g \in G_s,\ t \in T; \tag{11}$$

$$v_{gst} \leqslant 1, \quad g \in G,\ s \in S,\ t \in T; \tag{12}$$

$$z_{gt} \geqslant w_{gt} - w_{g,t-1}, \quad g \in G_s,\ t \in T; \tag{13}$$

$$v_{gst} \geqslant u_{gst} - u_{gs,t-1}, \quad g \in G_f,\ s \in S, \quad t \in T; \tag{14}$$

$$\pi_s u_{gst} = \pi_s w_{gt}, \quad g \in G_s,\ s \in S, \quad t \in T; \tag{15}$$

$$\pi_s v_{gst} = \pi_s z_{gt}, \quad g \in G_s,\ s \in S,\ t \in T; \tag{16}$$

$$p_{gst}, v_{gst} \geqslant 0, \quad u_{gst} \in \{0, 1\}, \quad g \in G, \quad s \in S,\ t \in T; \tag{17}$$

$$z_{gt} \geqslant 0, \quad w_{gt} \in \{0, 1\}, \quad g \in G_s,\ t \in T. \tag{18}$$

The objective of the problem is to minimize expected operating costs, which consist of minimum load costs, start-up costs, and fuel costs. Equation (2) requires balancing the amount of power that flows into and out of each bus.

Equation (3) represents a linearized, lossless model of the power flow equations (Kirchhoff's law) according to which the power flow on a line l is proportional to the phase angle difference between the two end buses of the line. The proportionality factor B_{ls} is the susceptance of line l under scenario s, where $B_{ls} = 0$ for scenarios $s \in S$ in which line l is out of service, thereby forcing the flow in line l to equal zero. Constraints (4) define thermal capacity limits on the transmission lines. Constraints (5) impose minimum and maximum generator capacity limits. Similarly to the case of transmission line failures, $P^+_{gs} = 0$ and $P^-_{gs} = 0$ holds for those scenarios $s \in S$ in which generator g is out of service, thereby forcing power production for generator g to equal zero. The constraints defined by Equation (6) represent ramping constraints on the rate of change of generator output. Equations (7)–(10) represent the minimum up and down time constraints of both fast and slow generators. Constraints (11) and (12) place upper bounds on the start-up variables. Note that although start-up variables are in principle binary, following O'Neill et al. (2010) we are able to relax them as continuous variables since they are penalized in the objective function of the model, thus forcing them to an extreme value. In this way we are able to reduce the size of the branch-and-bound tree. The transition rule for generator start-up variables is imposed in Equations (13) and (14). Equations (15) and (16) are the nonanticipativity constraints that imply that, for slow generators, second-stage commitment and start-up (u_{gst}, v_{gst}) needs to be consistent with the first-stage commitment and start-up (w_{gt}, z_{gt}) under each scenario $s \in S$. Integrality and nonnegativity constraints are imposed in Equations (17) and (18). Note that the model determines the amount of generation capacity that needs to be committed in the day-ahead time frame by endogenously accounting for uncertainty within the model formulation. This motivates the use of the two-stage stochastic unit commitment model for quantifying the amount of day-ahead reserve requirements that are imposed by system operation uncertainty.

3.3. Deterministic Unit Commitment

A simpler approach to protect the system against uncertainty is to introduce exogenous requirements on excess generation capacity rather than explicitly modeling the ability of the system to observe and respond to uncertainty. More specifically, in contrast to modeling renewable production and load uncertainty, the deterministic formulation imposes ad hoc requirements on reserve supply, which is generation capacity that can be called upon in the case of renewable supply and load fluctuations as well as transmission line and generation outages. Likewise, rather than modeling contingencies explicitly, the deterministic model imposes exogenous import constraints that ensure that the system can withstand the failure of major generation resources within load pockets as well as the failure of major transmission interties. Since reserve requirements and import limits in practice have relied upon user experience, these procedures may need to be revised in light of new uncertain generation sources. For the deterministic formulation, we follow the model of Sioshansi and Short (2009):

$$\text{(DUC):}\quad \min \sum_{g\in G}\sum_{t\in T}(K_g w_{gt} + S_g z_{gt} + C_g p_{gt}) \tag{19}$$

$$\text{s.t.}\quad \sum_{l\in LI_n} e_{lt} + \sum_{g\in G_n} p_{gt} = D_{nt} + \sum_{l\in LO_n} e_{lt}, \quad n\in N,\ t\in T; \tag{20}$$

$$-TC_l \leqslant e_{lt} \leqslant TC_l, \quad l\in L,\ t\in T; \tag{21}$$

$$e_{lt} = B_l(\theta_{nt} - \theta_{mt}), \quad l=(m,n)\in L,\ t\in T; \tag{22}$$

$$p_{gt} + f_{gt} \leqslant P^+_g w_{gt}, \quad g\in G,\ t\in T; \tag{23}$$

$$p_{gt} + s_{gt} + f_{gt} \leqslant P^+_g, \quad g\in G,\ t\in T; \tag{24}$$

$$p_{gt} \geqslant P^-_g w_{gt}, \quad g\in G,\ t\in T; \tag{25}$$

$$p_{gt} - p_{g,t-1} + s_{gt} \leqslant R^+_g, \quad g\in G,\ t\in T; \tag{26}$$

$$p_{g,t-1} - p_{gt} \leqslant R^-_g, \quad g\in G,\ t\in T; \tag{27}$$

$$\sum_{g\in G_f} f_{gt} + \sum_{g\in G} s_{gt} \geqslant T^{\text{req}}_t, \quad t\in T; \tag{28}$$

$$f_{gt} \leqslant FR_g, \quad g\in G,\ t\in T; \tag{29}$$

$$\sum_{g\in G} s_{gt} \geqslant S^{\text{req}}_t, \quad t\in T; \tag{30}$$

$$\sum_{l\in IG_j} \gamma_{jl} e_{lt} \leqslant IC_j, \quad j\in IG,\ t\in T; \tag{31}$$

$$\sum_{q=t-UT_g+1}^{t} z_{gq} \leqslant w_{gt}, \quad g\in G, \quad t\geqslant UT_g; \tag{32}$$

$$\sum_{q=t+1}^{t+DT_g} z_{gq} \leqslant 1 - w_{gt}, \quad g\in G, \quad t\leqslant |T| - DT_g; \tag{33}$$

$$z_{gt} \leqslant 1, \quad g\in G,\ t\in T; \tag{34}$$

$$z_{gt} \geqslant w_{gt} - w_{g,t-1}, \quad g\in G,\ t\in T; \tag{35}$$

$$p_{gt},\ z_{gt},\ s_{gt},\ f_{gt}, \geqslant 0, \quad w_{gt}\in\{0,1\}, \quad g\in G,\ t\in T. \tag{36}$$

Note that decision variables in the resulting model are not contingent on scenarios, thereby reducing the size of the model. The maximum capacity constraint in Equation (23) is modified to account for the provision of fast reserves and Equation (24) is added to account for the provision of fast and slow reserves. The total reserve requirement is imposed in Equation (28) and an upper limit on the provision of fast reserves is imposed in Equation (29). The online slow reserve requirement is imposed in Equation (30) and the import constraints are imposed in Equation (31).

3.4. Economic Dispatch

The deterministic and stochastic unit commitment models represent two different methods for committing slow generation resources. These resources need to have their schedules determined in the day-ahead scheduling time frame because of their limited operating flexibility. Once slow resources are committed, the performance of the system is tested by performing Monte Carlo simulations of its response to net demand and contingency outcomes, given the unit commitment schedule of slow generators. The net demand outcomes are generated using the multiarea wind power production time series model presented in §5 and the generation of contingencies is described in §7.3. The economic dispatch of units for each outcome c requires solving the following problem:

$$(\mathrm{ED}_c): \quad \min \sum_{g\in G}\sum_{t\in T}(K_g w_{gt} + S_g z_{gt} + C_g p_{gt}) \tag{37}$$

$$\text{s.t.} \sum_{l\in LI_n} e_{lt} + \sum_{g\in G_n} p_{gt} = D_{nct} + \sum_{l\in LO_n} e_{lt}, \quad n\in N,\ t\in T; \tag{38}$$

$$e_{lt} = B_{lc}(\theta_{nt} - \theta_{mt}), \quad l=(m,n)\in L,\ t\in T; \tag{39}$$

$$P^-_{gc} u_{gt} \leqslant p_{gt} \leqslant P^+_{gc} u_{gt}, \quad g\in G,\ t\in T; \tag{40}$$

$$p_{gt} - p_{g,t-1} \leqslant R^+_g, \quad g\in G,\ t\in T; \tag{41}$$

$$w_{gt} = w^\star_{gt}, \quad g\in G_s,\ t\in T; \tag{42}$$

$$z_{gt} = z^\star_{gt}, \quad g\in G_s,\ t\in T; \tag{43}$$

$$(21),\ (27),\ (32),\ (33),\ (34),\ (35)$$

$$p_{gt},\ z_{gt}, \geqslant 0, \quad w_{gt}\in\{0,1\}, \quad g\in G,\ t\in T. \tag{44}$$

Equations (42) and (43) set the commitment of slow generators to the optimal solution of the unit commitment problem.

4. Decomposition Algorithm

By dualizing the constraints of Equations (15) and (16) we obtain the following Lagrangian:

$$\mathcal{L} = \sum_{g\in G}\sum_{s\in S}\sum_{t\in T}\pi_s(K_g u_{gst} + S_g v_{gst} + C_g p_{gst}) + \sum_{g\in G_s}\sum_{s\in S}\sum_{t\in T}\pi_s(\mu_{gst}(u_{gst} - w_{gt}) + \nu_{gst}(v_{gst} - z_{gt})). \tag{45}$$

We can then decompose the Lagrangian to one subproblem for each scenario, that determines optimal second-stage decisions:

$$(\mathrm{P2}_s): \quad \min \sum_{g\in G}\sum_{t\in T}\pi_s(K_g u_{gst} + S_g v_{gst} + C_g p_{gst}) + \sum_{g\in G_s}\sum_{t\in T}\pi_s(\mu_{gst} u_{gst} + \nu_{gst} v_{gst}) \tag{46}$$

$$\text{s.t.} \quad (2),\ (3),\ (4),\ (5),\ (6),\ (9),\ (10),\ (12),\ (14)$$

$$p_{gst}\geqslant 0, \quad v_{gst}\geqslant 0, \quad u_{gst}\in\{0,1\}, \quad g\in G,\ t\in T. \tag{47}$$

Note that the constraint $v_{gst}\leqslant 1,\ g\in G_s$ of Equation (12), although redundant for the model (SUC), is necessary for bounding $\mathrm{P2}_s$ when applying the decomposition algorithm. We also obtain a single subproblem that determines optimal first-stage decisions:

$$(\mathrm{P1}): \quad \min -\sum_{g\in G_s}\sum_{s\in S}\sum_{t\in T}\pi_s(\mu_{gst} w_{gt} + \nu_{gst} z_{gt}) \tag{48}$$

$$\text{s.t.} \quad (7),\ (8),\ (11),\ (13)$$

$$w_{gt}\in\{0,1\}, \quad z_{gt}\geqslant 0, \quad g\in G_s,\ t\in T. \tag{49}$$

The dual variables are updated as follows:

$$\mu^{k+1}_{gst} = \mu^k_{gst} + \alpha_k \pi_s(w^k_{gt} - u^k_{gst}), \quad g\in G_s,\ s\in S,\ t\in T; \tag{50}$$

$$\nu^{k+1}_{gst} = \nu^k_{gst} + \alpha_k \pi_s(z^k_{gt} - v^k_{gst}), \quad g\in G_s,\ s\in S,\ t\in T, \tag{51}$$

where w^k_{gt} and z^k_{gt} are the optimal solutions of (P1) at iteration k; u^k_{gst} and v^k_{gst}; are the optimal solutions of $(\mathrm{P2}_s)$ at iteration k; and α_k is the step size at iteration k. We could have relaxed only the nonanticipativity constraint on the commitment variables. The advantage of also relaxing the nonanticipativity constraint on the start-up variables is that $(\mathrm{P2}_s)$, $s\in S$, is a smaller problem, since the constraints on the unit commitment of the slow generators are a part of (P1). An additional advantage of this choice of decomposition is that, at each step, the slow generator unit commitment solutions of the first subproblem can be used for generating a feasible solution to the original problem by solving an economic dispatch problem (ED_s), Equations (37)–(44), for each scenario $s\in S$. As a result, at each step of the algorithm we get an upper bound on the optimal solution that can be used for terminating the algorithm, as well as a feasible schedule. This should be contrasted with the case where we would have chosen to relax only the nonanticipativity constraints on the unit commitment variables, and not the start-up variables.

An important feature of the proposed algorithm is the fact that the second-stage subproblems $(\mathrm{P2}_s)$ and the economic dispatch problems (ED_s) can be solved in parallel. As we discuss in §7.5, we have implemented a parallel algorithm for the problem in a cluster of 1,152 nodes, with eight CPUs per node at 2.4 GHz and 10 GB per node. The step size rule follows Fisher (1985) and Held et al. (1974) and is presented in Papavasiliou et al. (2011).

5. Multiarea Wind Power Model

Because of the highly nonlinear but static relationship between wind speed and wind power production it is common in the wind power modeling literature to model wind speed rather than wind power with time series models and then convert speed to power using a static conversion curve. The task of modeling wind speed consists of removing seasonal and daily patterns, fitting the data to a parametric or nonparametric distribution and fitting an appropriate time series model to the underlying "noise" in order to capture the strong temporal dependency of wind speed.

Given a multiarea data set y_{kt}, where k indexes location and t indexes time period, the first step is to remove hourly and seasonal patterns by subtracting the hourly mean and dividing by the hourly standard deviation in order to obtain a stationary data set y^S_{kt} for each location.[2]

We then filter the data set in order to obtain an approximately stationary Gaussian data set y^{GS}_{kt}. Brown et al. (1984), Torres et al. (2005), and Morales et al. (2010) use this approach for transforming Weibull-distributed wind speed data to Gaussian data, and Callaway (2010) uses a nonparametric transformation. Since no single parametric distribution provides a close fit for the observed data in all locations of our data set, we fit an empirical distribution $\hat{F}_k(\cdot)$ to the data for each location k. The resulting data set y^{GS}_{kt} can be modeled by an autoregressive process:

$$y^{GS}_{k,\,t+1} = \sum_{j=0}^{p} \hat{\phi}_{kj} y^{GS}_{t-j} + \hat{\omega}_{kt}, \tag{52}$$

where $\hat{\omega}_{kt}$ is the estimated noise and $\hat{\phi}_{kj}$, $j \in \{1, \ldots, p\}$ are the estimated coefficients of the autoregressive model. The calibration process is summarized in the following steps:

Step (a). Remove systematic seasonal and diurnal effects

$$y^S_{kt} = \frac{y_{kt} - \hat{\mu}_{kmt}}{\hat{\sigma}_{kmt}},$$

where y_{kt} is the data, y^S_{kt} is the transformed data that is stationary, and $\hat{\mu}_{kmt}$ and $\hat{\sigma}_{kmt}$ are the sample mean and standard deviation, respectively, for location k, season m, and hour t.

Step (b). Transform the data in order to obtain a Gaussian distribution on the data set

$$y^{GS}_{kt} = N^{-1}(\hat{F}_k(y^S_{kt})),$$

where y^{GS}_{kt} is the transformed data that is stationary Gaussian distributed, $N^{-1}(\cdot)$ is the inverse of the cumulative distribution function of the normal distribution, and $\hat{F}_k$ is the cumulative function of the fit for the data in location k.

Step (c). Use the Yule-Walker equations (Box and Jenkins 1976) to estimate the autoregressive parameters $\hat{\phi}_{kj}$ and covariance matrix $\hat{\Sigma}$ of the estimated noise in the autoregressive model of Equation (52). We have observed that the inclusion of spatial correlations in the covariance matrix $\hat{\Sigma}$ improves the fit of the model, compared to using a diagonal covariance matrix that only captures the variance of the noise in each location but ignores its correlation with the noise in other locations.

Once the parameters of the statistical model are estimated, the inverse process can be performed for simulating wind speed. The relationship between wind speed and wind power typically depends on wind turbine size, manufacturer design, and various other factors, and an aggregated power curve can be used to represent a group of turbines. To simulate wind power production, we use a piecewise linear approximation of such an aggregate power curve for each location. We provide details about the data used in our model and the goodness of fit in §EC.2 of the electronic companion (available as supplemental material at http://dx.doi.org/10.1287/opre.2013.1174).

6. Scenario Selection

The challenge of selecting scenarios for the stochastic unit commitment problem is to discover a small number of representative outcomes that properly guide the stochastic program to produce a unit commitment schedule that improves average costs, as compared to a schedule determined by solving a deterministic unit commitment model. The basic trade-off that needs to be balanced in dispatching fast reserves is the flexibility that fast units offer in utilizing renewable generation versus their higher operating costs. Fast generators typically incur higher marginal fuel costs. In addition, the start-up and minimum load costs of these units are similar to those of slow units, however their capacity is smaller; hence, their start-up and minimum load cost per unit of capacity is greater than that of slow generators. The advantage of largely relying on fast units is that the system is capable of discarding less renewable power, which results in significant savings in fuel costs. Unlike fast generators that can shut down on short notice in the case of increased renewable power generation, slow generators can only adjust their output level but cannot deviate from their day-ahead commitment schedule, as indicated by the nonanticipativity constraints of Equations (15), (16), because of limitations on their operational flexibility. As a result, slow units cannot back down from their minimum generation levels and therefore, in the case of oversupply, require the waste of excess renewable energy in order to stay online. We note that the dispatch of the system is performed based on the objective of minimizing operating costs, as indicated in the models of §3. There are no must-take or priority dispatch requirements for renewable resources, or a penalty that must be paid for curtailing renewable power injections.[3]

The introduction of transmission constraints complicates scenario selection considerably because of the fact that the fast reserves are not readily accessible when certain transmission lines are congested and the availability of

renewable resources may be limited because of transmission constraints. The further introduction of composite outages introduces the complication of protecting the system against very low probability outcomes that can severely impact system reliability.

The stochastic unit commitment literature has relied extensively on the scenario selection and scenario reduction algorithms proposed by Dupacova et al. (2003) and their faster variants that were proposed by Heitsch and Römisch (2003). The effectiveness of these algorithms in the stochastic unit commitment problem was first demonstrated by Gröwe-Kuska et al. (2002) who apply the algorithms for scheduling hydro and thermal units in a German utility. Morales et al. (2009a) propose a variant of the scenario reduction algorithm of Heitsch and Römisch (2003) that removes the scenarios that cause the least change in the second-stage costs of the optimization problem. These scenario selection and scenario reduction techniques rely on clustering scenarios in such a way that the transformed measure of the stochastic program is perturbed minimally. General scenario clustering techniques have been used previously in the stochastic optimization literature by Pflug (2001) and Latorre et al. (2007).

Despite the theoretical justification of the scenario reduction algorithm of Dupacova et al. (2003) and Heitsch and Römisch (2003), as dicussed in Papavasiliou et al. (2011) these algorithms perform poorly in a unit commitment model without transmission constraints and contingencies, where wind power production is the only stochastic input. This is attributed to two reasons. Firstly, the algorithms of Dupacova et al. (2003) and Heitsch and Römisch (2003) are not guaranteed to preserve the moments of hourly wind generation. Because of the predominant role of fuel costs in the operation of the system, the accurate representation of average wind supply in the case of large-scale wind integration is crucial for properly guiding the weighting of scenarios. Moreover, the modeler cannot specify certain scenarios that are deemed crucial. For example, the realization of minimum possible wind output throughout the entire day needs to be considered explicitly as a scenario. Otherwise, there is the possibility of under-committing resources and accruing overwhelming costs from load shedding in economic dispatch. To overcome these drawbacks, Papavasiliou et al. (2011) generate a large number of samples from the statistical model of the underlying process and select a subset of samples based on a set of prescribed criteria that are deemed important. The authors then assign weights to each scenario such that the first moments of hourly wind output are matched as closely as possible. Scenario selection that strives to match certain statistical properties of the underlying process has been proposed in previous literature, e.g., by Hoyland and Wallace (2001) and Hoyland et al. (2003).

The introduction of transmission constraints and contingencies complicates the task of scenario selection considerably, as it is not clear how to extend the scenario selection algorithms of Dupacova et al. (2003) and Papavasiliou et al. (2011) to networks with multiarea wind production and how to account for network component failures. Assuming independence among net load outcomes and contingencies, which is a very reasonable assumption, one natural approach for generating scenarios would be to decouple the selection of contingencies from the selection of net load scenarios. The previously discussed algorithms could then be adapted for selecting multiarea net load scenarios (resulting from random scenarios of wind generation), and we can take the cross product of these net load scenarios with a set of significant contingencies. Such an approach is described in §EC.3 of the electronic companion. One natural question that this approach raises is which contingencies to select and how to weigh them relative to each other. For example, if the failure of any given generator over an entire day has a likelihood of 1% (Pereira and Balu 1992), in a network with 130 generators the chances of a single-generator failure are approximately 35.6%.[4] The question arises, then, if a scenario includes the failure of a single generator, how should that scenario be weighed against other scenarios? Which generator failures should we include in the scenario set? Should we consider composite failures, for example, the failure of multiple generators, multiple lines, or generators and lines in the same scenario? It becomes clear that a methodical approach for scenario selection is needed.

6.1. Importance Sampling

The scenario selection algorithm that we propose in this section is inspired by importance sampling. In contrast to scenario generation techniques that aim to match the moments of the underlying stochastic process such as those proposed by Hoyland and Wallace (2001) and Hoyland et al. (2003), the proposed scenario selection algorithm aims at selecting scenarios that best represent the average cost impact of uncertainty on the problem. Importance sampling is a statistical technique for reducing the number of Monte Carlo simulations that are required for estimating the expected value of a random variable within a certain accuracy. For an exposition see Mazumdar (1975) and Infanger (1992). As Pereira and Balu (1992) report, this technique has been used in reliability analysis in power systems with composite generation and transmission line failures, where the estimated random variable is a reliability metric (e.g., loss of load probability or expected load not served).

Given a sample space Ω and a measure p on this space, importance sampling defines a measure q on the space that reduces the variance of the observed samples of the random variable C, and weighs each simulated outcome ω by $p(\omega)/q(\omega)$ in order to unbias the simulation results. The measure q is ideally chosen such that it represents the contribution of a certain outcome to the expected value that

is being computed, i.e.,

$$q^\star(\omega) = \frac{p(\omega)C(\omega)}{\mathbb{E}_p C}. \tag{53}$$

Of course, it is not possible to determine this measure since $\mathbb{E}_p C$ is the quantity we wish to compute. Nevertheless, the intuition of selecting samples according to their contribution to the expected value can be carried over to scenario selection. For example, in Papavasiliou et al. (2011) the authors include the wind power production outcome with the lowest aggregate production over the entire day. Although the likelihood of this outcome is very low, its impact on system costs can be extremely high, making its contribution to expected cost $p(\omega)C(\omega)$ significant.

6.2. Proposed Algorithm

The extension of the intuition of importance sampling to the case of scenario selection is straightforward: if the ideal measure $q^\star$ of Equation (53) were closely approximated by a measure q, then selecting a small number of outcomes according to this measure and weighing them according to $p(\omega)/q(\omega)$ would provide an accurate estimate of the expected cost. Therefore, samples selected according to q can be interpreted as representative scenarios that need to be weighted according to $p(\omega)/q(\omega)$ relative to each other in order not to bias the result.

We proceed by generating an adequately large subset of the sample space $\Omega_S = \{\omega^1, \ldots, \omega^M\}$ and we calculate the cost of each sample against a deterministic unit commitment policy $C_D(\cdot)$. Since $\bar{C} = \sum_{i=1}^{M}(C_D(\omega_i)/M)$ provides an accurate estimate of expected cost, we interpret the sample space of the system as Ω_S and the measure as the uniform distribution over Ω_S, hence $p(\omega) = M^{-1}$ for all $\omega \in \Omega_S$. We then obtain $q(\omega_i) = C_D(\omega_i)/(M\bar{C})$, $i = 1, \ldots, M$, and each selected scenario is weighed according to $\pi_s = p(\omega)/q(\omega)$, hence $\pi_s/\pi_{s'} = C_D(\omega^{s'})/C_D(\omega^s)$ for each pair of selected scenarios $\omega^s, \omega^{s'} \in \hat{\Omega}$. Hence, the proposed algorithm selects scenarios with a likelihood that is proportional to their cost impact, and discounts these scenarios in the stochastic unit commitment in proportion to their cost impact in order not to bias the stochastic unit commitment policy. We therefore propose the following algorithm:

Step (a). Define the size N of the reduced scenario set $\hat{\Omega} = \{\omega^1, \ldots, \omega^N\}$.

Step (b). Generate a sample set $\Omega_S \subset \Omega$, where $M = |\Omega_S|$ is adequately large. Calculate the cost $C_D(\omega)$ of each sample $\omega \in \Omega_S$ against the best deterministic unit commitment policy and the average cost $\bar{C} = \sum_{i=1}^{M}(C_D(\omega_i)/M)$.

Step (c). Choose N scenarios from Ω_S, where the probability of picking a scenario ω is $C_D(\omega)/(M\bar{C})$.

Step (d). Set $\pi_s = C_D(\omega)^{-1}$ for all $\omega^s \in \hat{\Omega}$.

6.3. Discussion

In contrast to the scenario selection algorithms that build the scenario set sequentially by evaluating the impact of candidate scenarios on the existing scenario set (see, for example, Dupacova et al. 2003, Gröwe-Kuska et al. 2002, and Latorre et al. 2007), the proposed algorithm selects and weighs scenarios in one shot (as in, for example, Hoyland and Wallace 2001). The proposed algorithm ensures that, as long as the cost impacts of all selected scenarios are of the same order of magnitude, which is commonly the case, then so are the weights in the stochastic unit commitment formulation. As a result, each scenario influences the first-stage decisions. This should be contrasted with the case where the probabilities of certain scenarios are very small compared to the probabilities of other scenarios. In that case, as we can see in Equation (48), the influence of scenarios with very small probability is minimal in the objective function of (P1), and consequently these scenarios will not tend to influence the optimal solution of (P1) and therefore the first-stage decision, i.e., the commitment of slow generators. In addition, from Equations (50) and (51) we see that scenarios with small probability exhibit very small changes in the values of their dual multipliers, which further supports the argument that these scenarios do not influence (P1). In fact, we observed that the stochastic unit commitment policy remained completely unaffected by scenarios that were 100 times less likely to occur than their competing scenarios in the stochastic unit commitment formulation. Therefore, the inclusion of these scenarios introduced superfluous computational load. This is prevented with our proposed scenario selection algorithm.

Commonly occurring scenarios are most likely to populate the original set of candidate scenarios Ω_S in step (b), and are therefore also most likely to populate the reduced scenario set $\hat{\Omega}$ in step (c). Thus, highly likely outcomes are represented. At the same time, adverse scenarios are accounted for (by including them in the scenario set) without being overemphasized (by discounting their probability weighting).

An additional appealing feature of the proposed algorithm is that it selects a rich set of multiarea net load outcomes. This should be contrasted to the case where we would multiplex net load scenarios selected according to the scenario selection algorithms of Dupacova et al. (2003), Gröwe-Kuska et al. (2002), and Papavasiliou et al. (2011) with contingency scenarios. Moreover, in contrast to the scenario selection method proposed in Papavasiliou et al. (2011), the scenario selection algorithm proposed in this paper does not depend on the judgement of the modeler for specifying criteria that are deemed important. The proposed procedure can be applied to a broad setting of problems under uncertainty in a straightforward fashion. In §7 it is shown that the resulting stochastic unit commitment policy outperforms the approach of Dupacova et al. (2003) as well as common deterministic rules for four case studies of uncertainty, one of which is the case study that was performed in Papavasiliou et al. (2011).

7. Results

In this section we present simulation results for a reduced model of the California power system with 225 buses, 130 generators, and 375 transmission lines. The model includes a sparse representation of the entire Western Electricity Coordinating Council Western interconnect outside of California, which is necessary in order to capture loop flow effects as well as imports and exports into and out of California. A schematic description of the model showing the wind sites, zones, and import constraints of the network is presented in §EC.1 of the electronic companion. The model is also used in Yu et al. (2010) and Papavasiliou et al. (2011). We study three levels of wind integration, the zero wind integration case as well as a case of moderate and deep wind integration corresponding to the 2012 and 2020 renewable energy integration targets of California, respectively. Ex post we have observed that the moderate integration case corresponds to approximately 7% wind energy penetration, and the deep integration case corresponds to approximately 14% wind energy penetration. In the rest of the paper we refer to these cases as moderate and deep integration, respectively. We also perform the deep integration case study for the case where transmission constraints and contingencies are not accounted for, in order to quantify the impact of these effects on the analysis. The latter case is denoted as deep simple.

7.1. Day Types

We focus our analysis on eight day types that represent weekdays and weekends for each season. We assume that the day-ahead forecast of wind power production for each day type is equal to the hourly average wind profile of the day type, as estimated by the available data. Since the system operator is facing the same source of uncertainty from day to day within days of the same type, we assume that unit commitment decisions are identical for days of the same type. This determines initial conditions of the system at the start of each day for the unit commitment models. Following the approach of Bertsimas et al. (2013), in the economic dispatch model of §3.4 the initial conditions of the system are ignored.

By assuming an identical unit commitment policy for each day type, and by ignoring boundary conditions in the economic dispatch model, we are able to parallelize computation in the evaluation phase of the Monte Carlo simulation since each day is treated independently and can be simulated by a different processor. As a result, we are capable of simulating a total of 8,000 days (1,000 days for each day type), by contrast to Sioshansi and Short (2009), Ruiz et al. (2009a), and Tuohy et al. (2009), who only simulate one year of operations on a rolling horizon basis. As a result, we are able to obtain more accurate results regarding cost performance. The ability to parallelize computation through our assumptions comes at the price of ignoring boundary conditions in the economic dispatch model and transitions between different day types, e.g., from a weekend to a weekday or from a weekday to a weekend. In future research our intention is to quantify the impact of these assumptions by comparing our batch processing model with a serial model based on rolling planning.[6]

7.2. Data

We do not use the wind production data from Yu et al. (2010) since our wind production model is more detailed. As we describe in §EC.2, we are using 2006 wind production data from the National Renewable Energy Laboratory database. We also use load data from the same year, which is publicly available at the California ISO Open Access Same-Time Information System, CAISO (CAISO 2011). The average load in the system is 27,298 MW, with a minimum of 18,412 MW and a peak of 45,562 MW.

We use a more general model for thermal generators within CAISO, with 124 generators, compared to the model in Yu et al. (2010), which uses 23 aggregated thermal generators. The value of lost load is set to 5,000 $/MW-h. The number of generators and the capacity for each fuel type are shown in Table 1. The last two rows of Table 1 describe how the fossil fuel generation mix is partitioned into fast and slow generators. The entire thermal generation capacity of the system is 30,231.5 MW. The total capacity of the system, not including wind power resources, is 53,665.5 MW.

7.3. Relative Performance of Policies

In this section we discuss the relative performance of the stochastic and deterministic unit commitment policies. The results are obtained by running the economic dispatch model against 1,000 Monte Carlo outcomes of wind power production and contingencies for each day type, with a probability of generator failure of 1% (Pereira and Balu 1992) and a probability of transmission line failure of 0.1% (Grigg et al. 1999). We assume that network element failures occur over the entire day. As we clarify in §7.1, by parallelizing the Monte Carlo simulations we are able to simulate 8,000 days of operations and obtain more accurate results regarding average cost performance. In §EC.4 we

Table 1. Generation mix for the test case.

Type	No. of units	Capacity (MW)
Nuclear	2	4,499
Gas	94	20,595.6
Coal	6	285.9
Oil	5	252
Dual fuel	23	4,599
Import	22	12,691
Hydro	6	10,842
Biomass	3	558
Geothermal	2	1,193
Wind (moderate)	5	6,688
Wind (deep)	10	14,143
Fast thermal	88	11,006.1
Slow thermal	42	19,225.4

provide information about the variance of our results. Wind production outcomes, generator failures, and transmission line failures are assumed to be independent. We consider two deterministic policies. The first policy places a total reserve requirement T_t^{req} in Equation (28) equal to a fraction of peak forecast net load for the day and an online slow reserve requirement S_t^{req} in Equation (30) equal to half of the total reserve requirement. Forecast net load, here, refers to forecast load minus forecast wind, scheduled imports and scheduled non-wind renewable resources. To determine the best policy of this type, we find the fraction of peak load that yields the best performance. The other deterministic policy that we consider is a variant of a reserve commitment policy that was recently proposed in Piwko et al. (2010). The authors propose a heuristic approach for committing spinning reserves, "the 3+5 rule," which requires the system to carry hourly spinning reserve no less than 3% of hourly forecast load plus 5% of hourly forecast wind power. This rule is adapted in our model by setting this as the online slow reserve requirement and setting the total reserve requirement at twice the level of the online slow reserve requirement. The results are shown in Table 2, where the best peak-load policy is highlighted in italic font and the best deterministic policy is highlighted in bold font. We note that the best peak-load policy outperforms the $3+5$ rule for all case studies.

To validate our proposed scenario selection algorithm that is inspired by importance sampling, denoted as SUC-IS, we compare it to an alternative scenario selection algorithm. The alternative stochastic unit commitment policy, denoted as SUC2, is based on crossing the most severe system contingencies with wind production outcomes produced by the scenario reduction algorithm of Dupacova et al. (2003). The details of the alternative scenario selection algorithm are described in §EC.3 of the electronic companion. We also consider a perfect foresight policy that commits and dispatches resources under perfect forecast of uncertain outcomes. The perfect foresight policy bounds the attainable cost of any unit commitment rule.

The relative performance of the proposed scenario selection algorithm based on importance sampling with respect to the deterministic policies, the alternative stochastic unit commitment policy SUC2, and the perfect foresight policy for the four case studies is presented in Figure 1. The results are presented in terms of the relative cost of each policy compared to the proposed stochastic unit commitment policy for each of the eight day types. In the last three rows of Table 3 we present the absolute cost of the proposed scenario selection algorithm SUC-IS as well as its gains over the best deterministic policy and the alternative stochastic unit commitment policy SUC2. We note that the benefits of the proposed stochastic unit commitment policy SUC-IS range between \$145,261 to \$244,226 per day relative to the best deterministic policy, and between \$5,022 to \$52,622 per day for the alternative stochastic unit commitment policy SUC2. We note that the proposed scenario selection algorithm SUC-IS outperforms SUC2, although the majority of the potential benefits is captured by either stochastic formulation. Moreover, the introduction of transmission constraints affects the gains of stochastic unit commitment, which supports the argument that stochastic unit commitment is especially valuable for the determination of locational capacity requirements.

7.4. Operating Costs, Committed Conventional Capacity, and Renewables Utilization

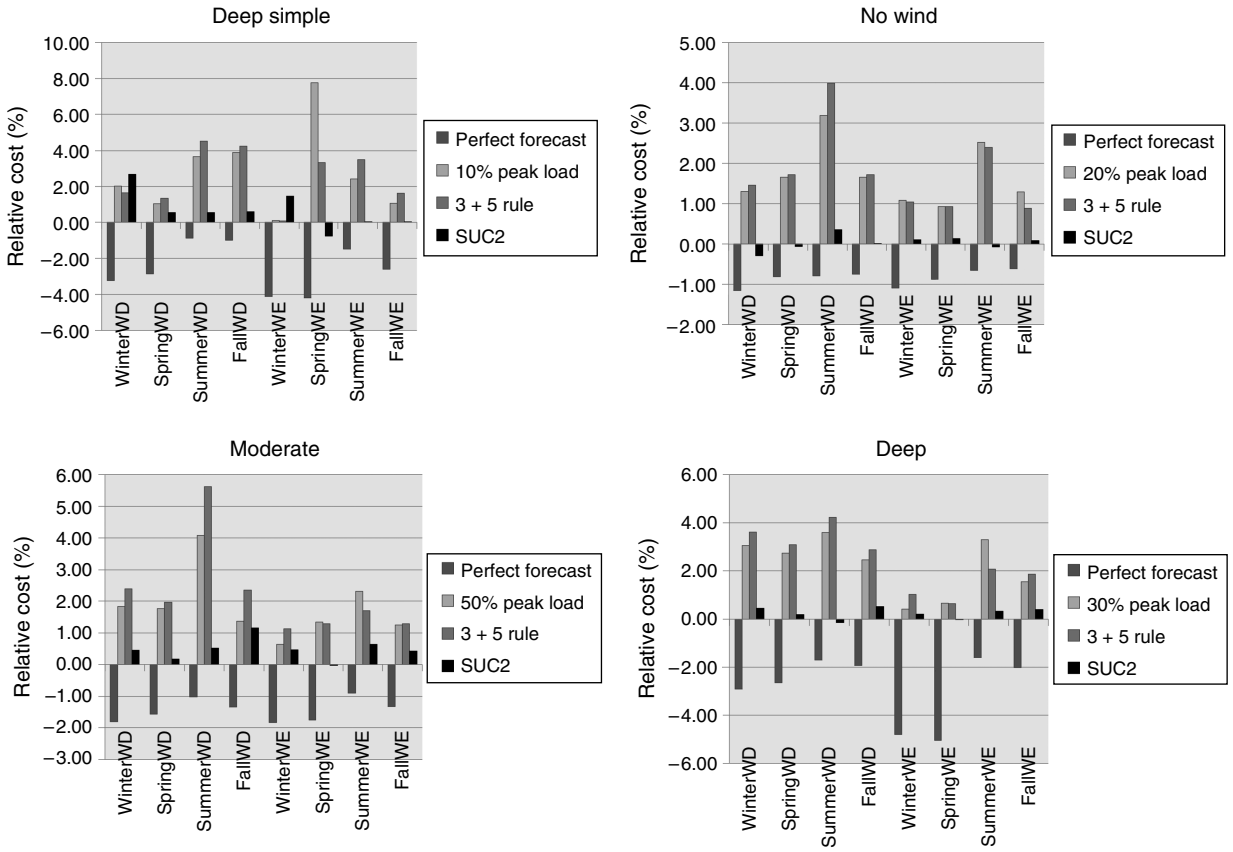
In Table 3 we present summary results for renewable energy waste, operating costs, and committed conventional capacity for the four case studies under consideration. Renewable energy losses range between 0.2% of total renewable energy production in the case of the deep integration study without transmission constraints and contingencies to 2.5% of total renewable energy production when transmission constraints and contingencies are accounted for. Operating costs decline steeply as the level of renewable power penetration increases, due to the decrease in fuel costs, which are the predominant cost in the system. Further details regarding operating costs are provided in §EC.4 of the electronic companion. By comparing the cost of column 2 (deep simple) to that of column 5 (deep), we note that failing to account for transmission constraints and contingencies results in an underestimation of operating costs by 31.0%. The significant cost increase resulting from transmission constraints can be attributed to the operating cost impacts of contingencies but also to the reduced flexibility of dispatching units in the system.

The commitment of conventional generation capacity, which is the most important factor in analyzing the economic of renewable energy integration, presents the most interesting results. Committed capacity is computed as the sum of slow generation capacity committed in the day ahead, plus the total amount of available fast capacity (11,006.1 MW, as indicated in Table 1) that is readily available in real time. The result reported in Table 3 therefore represents the optimal choice of conventional generation capacity that is required for reliably operating the system

Table 2. Deterministic policy cost comparison.

Case	10%	20%	30%	40%	50%	3+5
Deep simple	***5,251,279***	5,267,340	5,305,363	5,334,489	5,381,318	5,270,797
No wind	11,507,776	***11,505,126***	11,525,722	11,511,003	11,595,341	11,528,686
Moderate	9,575,616	9,573,424	9,575,465	9,592,592	***9,539,965***	9,605,016
Deep	7,638,328	7,645,104	***7,612,689***	7,661,245	7,659,490	7,637,464

Figure 1. In reading order: Cost comparison for the deep integration case without transmission or contingencies, the zero wind integration case, the moderate integration case, and the deep integration case.



under uncertainty, given the set of available candidate resources and without accounting for capital investments in new capacity, which are assumed to be a sunk cost. Note that, since committed slow unit capacity is the only first-stage decision in the stochastic unit commitment model, the third line of Table 3 (capacity (MW)) indicates the change in first-stage decisions across different day types.

We note that the moderate integration case reduces average conventional committed capacity by a mere 840 MW, which represents 12.6% of the 6,688 MW of installed wind capacity. Average conventional committed capacity for the deep integration case is reduced by 1,670 MW, which represents 11.8% of the 14,143 MW of installed wind capacity. Most importantly, we note that failing to account for transmission constraints in the deep integration study results in an underestimation of the committed capacity by 25.9% of installed wind capacity, relative to the estimated 11.8% reduction when these features are accounted for. This strongly supports the argument that the inclusion of transmission constraints and contingencies is crucial for accurately assessing the impact of large-scale renewable energy integration.

7.5. Computational Performance

The stochastic unit commitment problem of §3.2 for 42 scenarios has 909,216 continuous variables, 173,376 binary variables, and 2,371,824 constraints. The stochastic unit commitment algorithm was implemented in the Java callable library of CPLEX 12.4, and parallelized using the message passing interface (MPI). The algorithm was implemented on a high performance computing cluster in the Lawrence Livermore National Laboratory on a network of 1,152 nodes, 2.4 GHz, with eight CPUs per node and 10 GB

Table 3. Summary results for each case study.

	Deep simple	No wind	Moderate	Deep
RE daily waste (MWh/%)	163 (0.2%)	0 (0%)	877 (1.9%)	2,346 (2.5%)
Capacity (MW)	19,958	23,619	22,779	21,949
Cost ($M)	5.106	11.283	9.329	7.405
Daily savings relative to best det. ($)	145,261	221,854	244,226	207,698
Daily savings relative to SUC2 ($)	39,681	5,022	52,622	15,574

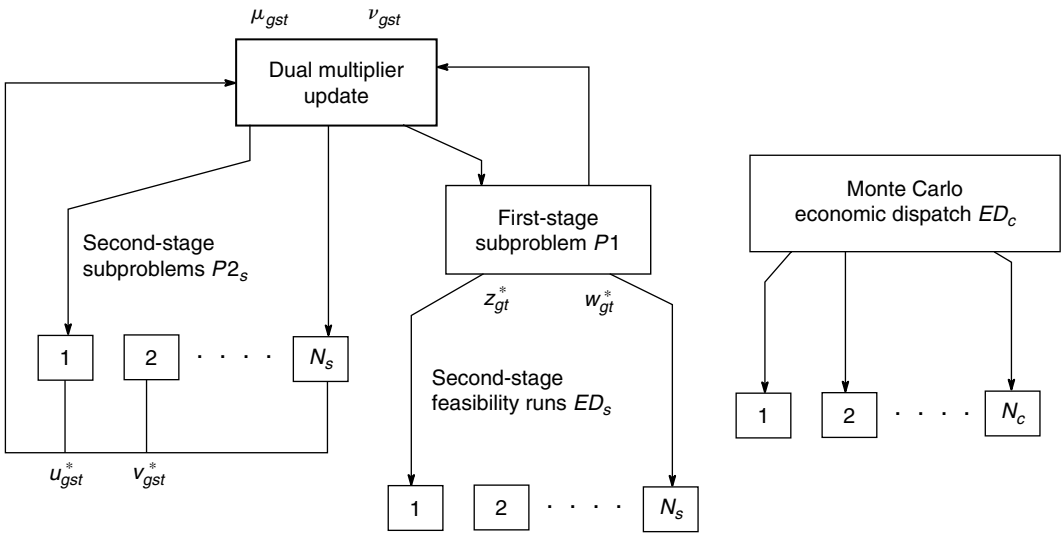
per node. The parallel implementation of the Lagrangian relaxation algorithm and Monte Carlo simulations is shown in Figure 2. Problems (P1) and ($P2_s$), $s \in S$ were run for 120 iterations. For the last 40 iterations, (ED_s) was run for each $s \in S$ in order to obtain a feasible solution and an upper bound for the stochastic unit commitment problem. The average elapsed time for the 42-scenario problem on 20 machines was six hours, 47 minutes. The mixed integer programming (MIP) gap for (P1) and ($P2_s$), $s \in S$ was set to $\epsilon_1 = 1\%$, and the MIP gap for obtaining a feasible schedule from (ED_s) was set to $\epsilon_2 = 0.1\%$. Note from Table 3 that the daily savings of SUC-IS relative to SUC2 fall within the MIP gap of the economic dispatch problem for the zero wind integration study. For all other cases (moderate, deep, and deep simple), the benefits of the SUC stochastic unit commitment policy are guaranteed to reflect a superior scenario selection approach. The sum of the optimal solutions of the first and second subproblem yield a lower bound LB on the optimal cost, whereas the optimal solution of the feasibility run results in an upper bound UB. The average gap, $(\mathrm{UB} - \mathrm{LB})/\mathrm{LB}$, that we obtained is 0.77%. However, to estimate an upper bound on the optimality gap it is also necessary to account for the MIP gap ϵ_1 that is introduced in the solution of (P1) and ($P2_s$), $s \in S$. The average upper bound on the optimality gap, $(\mathrm{UB} - (1-\epsilon_1)\mathrm{LB})/((1-\epsilon_1)\mathrm{LB})$, is 1.79%.

8. Conclusions and Perspectives

In this paper we present a two-stage stochastic unit commitment model that can be used for assessing the impact of wind power integration on operating costs, committed day-ahead generation capacity, and renewable energy utilization. We present a scenario selection algorithm inspired by importance sampling that is shown to outperform alternative stochastic and deterministic unit commitment approaches for four case studies, and a subgradient algorithm for solving the resulting stochastic program. The subgradient algorithm has been parallelized in a high performance computing cluster in order to solve a moderately sized system within an acceptable time frame. The daily operating cost benefits of stochastic unit commitment are shown to range between \$145,261 to \$244,226 relative to deterministic commitment rules and between \$5,022 to \$52,622 relative to an alternative stochastic unit commitment model. Renewable energy waste is negligible across all case studies. Failing to account for contingencies and transmission results in an underestimation of operating costs by 31.0%. Conventional committed capacity decreases by 11.8%–12.6%s of installed wind capacity and failing to account for transmission constraints and contingencies results in an underestimation of committed capacity by 25.9% of installed wind capacity relative to 11.8% when these features are accounted for.

There are various extensions of the present model that are intended in future research. In practice, system forecasts can and will be updated, leading to the opportunity for individual units to be committed or shut down as required. The fact that forecasts and dispatch decisions are revised during the day can be represented through a multistage formulation of the stochastic unit commitment problem. Moreover, the seasonal correlation of load with renewable production can affect capacity requirement calculations and requires an extension of the time series modeling framework presented in §5. The model can also be extended to an optimal investment model, where the first stage is interpreted as investment in new generation capacity. The inclusion of investment decisions on transmission lines in order to integrate increased amounts of renewable resources also represents an exciting area of future research. As we describe in §7.1 our analysis focuses on eight representative day types and ignores the boundary conditions of slow units

Figure 2. Parallel implementation of the Lagrangian relaxation algorithm and Monte Carlo simulation.



in order to obtain a representative set of unit commitment schedules and parallelize the Monte Carlo simulations. This is justified on computational grounds and also by the fact that the California ISO day-ahead market currently operates with a 24-hour look-ahead and ignores boundary conditions, unless resources are self-dispatched or endogenize their on/off status by modifying their start-up bids. In future research we look forward to quantifying the implications of this simplifying assumption and the trade-offs in terms of computational effort.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/opre.2013.1174.

Endnotes

1. Increased operating costs can arise from the wear and tear of conventional generators that adjust their output in order to balance the supply of renewable power. Balancing generators are also often required to operate at inefficient operating levels. In addition, the minimum load levels of balancing resources result in increased emissions and fuel costs. In addition to the cost of corrective actions, the unpredictability of renewable resources results in suboptimal commitment of units in the day-ahead time frame. Quantifying the additional operating costs resulting from the large-scale integration of renewable energy sources has recently emerged as an active area of research (see Sioshansi and Short 2009, Ackermann 2005, van Hulle 2005).

2. The seasonal and diurnal systematic patterns are removed during calibration in order to obtain an approximately stationary data set, which is a necessary condition for the Yule-Walker equations that are used for estimating the autoregressive parameters to be valid. The systematic effects are added back to the data set in the simulation phase. By removing systematic effects we are able to exploit the full year of available data for estimating a time series model of underlying noise.

3. Such restrictions are imposed in many systems in order to facilitate the penetration of renewable energy and the reduction of carbon emissions.

4. What we are trying to capture is the probability of having a unit down in an operating day, conditioned on the unit being available during the closure of the day-ahead market. We approximate this quantity with the forced outage rate (FOR) of a unit. Academic as well as real-world data sets (e.g., Pereira and Balu 1992, Canadian Energy Association (CEA) 2006, Grigg et al. 1999) indicate that the range of FOR varies widely. We use a uniform FOR for all units, which is contained within the range of FORs that occur in practice and in academic research, although the proposed methodology can be applied for varying FORs without any excess computational requirements.

5. This is the probability of obtaining one success in 130 independent trials for which the probability of success is 0.01.

6. Ignoring boundary conditions between days and treating the unit commitment schedule of each day independently is also justified by the current ISO market structure where the day-ahead market is cleared each day separately, whereas the state of resources offered into the market can only be accounted for implicitly through the start-up and no-load bids of the units or through self-dispatch of units offered as price takers. There are strong incentives, however, against such behavior since self-dispatched units are not assured cost recovery like centrally dispatched resources. In addition, the inter-temporal optimization (medium-term scheduling) of hydro resources is performed at the individual resource level and not as part of the central unit commitment, and these resources are subsequently bid into the day-ahead market as single-day price-taking resources.

Acknowledgments

This research was funded by the National Science Foundation [Grant IIP 0969016], by the U.S. Department of Energy through the Future Grid initiative administered by the Power Systems Engineering Research Center, by the Lawrence Livermore National Laboratory, and by the Federal Regulatory Energy Commission.

References

Ackermann T (2005) *Wind Power in Power Systems* (John Wiley & Sons, Hoboken, NJ).

Arroyo JM, Galiana FD (2005) Energy and reserve pricing in security and network-constrained electricity markets. *IEEE Trans. Power Systems* 20(2):634–643.

Bertsimas D, Sim M (2004) The price of robustness. *Oper. Res.* 52(1):35–53.

Bertsimas D, Litvinov E, Sun XA, Zhao J, Zheng T (2013) Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Trans. Power Systems* 28(1):52–63.

Bouffard F, Galiana FD (2008) Stochastic security for operations planning with significant wind power generation. *IEEE Trans. Power Systems* 23(2):306–316.

Bouffard F, Galiana FD, Conejo AJ (2005) Market-clearing with stochastic security—Part II: Case studies. *IEEE Trans. Power Systems* 20(4):1827–1835.

Box GEP, Jenkins GM (1976) *Time Series Analysis: Forecasting and Control* (Holden-Day, San Francisco).

Brown BG, Katz RW, Murhpy AH (1984) Time series models to simulate and forecast wind speed and wind power. *J. Climate Appl. Meteorology* 23:1184–1195.

CAISO (2011) Oasis California Independent System Operator website. Accessed May 2, 2013, http://oasis.caiso.com/mrtu-oasis/home.jsp.

Callaway D (2010) Sequential reliability forecasting for wind energy: Temperature dependence and probability distributions. *IEEE Trans. Energy Conversion* 25:577–585.

Carpentier P, Cohen G, Culioli J-C, Renaud A (1996) Stochastic optimization of unit commitment: A new decomposition framework. *IEEE Trans. Power Systems* 11(2):1067–1073.

CEA (2006) Forced outage rates. Accessed May 2, 2013, http://www.science.smith.edu/~jcardell/Courses/EGR325/Readings/FOR%202006Update.pdf. Canadian Energy Assocation, Ottawa, Ontario, Canada.

Constantinescu EM, Zavala VM, Rocklin M, Lee S, Anitescu M (2011) A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation. *IEEE Trans. Power Systems* 26(1):431–441.

Dupacova J, Gröwe-Kuska N, Römisch W (2003) Scenario reduction in stochastic programming: An approach using probability metrics. *Math. Programming* 95(3):493–511.

Fisher ML (1985) An applications oriented guide to Lagrangian relaxation. *Interfaces* 15(2):10–21.

Galiana FD, Bouffard F, Arroyo JM, Restrepo JF (2005) Scheduling and pricing of coupled energy and primary, secondary, and tertiary reserves. *Proc. IEEE* 93(11):1970–1983.

Grigg C, Wong P, Albrecht P, Allan R, Bhavaraju M, Billinton R, Chen O, et al. (1999) The IEEE reliability test system—1996. *IEEE Trans. Power Systems* 14(3):1010–1020.

Gröwe-Kuska N, Kiwiel KC, Nowak MP, Romisch W, Wegner I (2002) Power management in a hydro-thermal system under uncertainty by Lagrangian relaxation. Greengard C, Ruszczyński A, eds. *Decision Making Under Uncertainty: Energy and Power*. IMA Volumes in Mathematics and Its Applications, Vol. 128 (Springer-Verlag, New York), 39–70.

Heitsch H, Römisch W (2003) Scenario reduction algorithms in stochastic programming. *Comput. Optim. Appl.* 24(2–3):187–206.

Held M, Wolfe P, Crowder HP (1974) Validation of subgradient optimization. *Math. Programming* 6(1):62–68.

Hoyland K, Wallace SW (2001) Generating scenario trees for multistage decision problems. *Management Sci.* 47(2):295–307.

Hoyland K, Kaut M, Wallace SW (2003) A heuristic for moment-matching scenario generation. *Comput. Optim. Appl.* 24:169–185.

Infanger G (1992) Monte Carlo (importance) sampling within a Benders decomposition algorithm for stochastic linear programs. *Ann. Oper. Res.* 39(1):69–95.

Jiang R, Zhang M, Li G, Guan Y (2012) Robust unit commitment with wind power and pumped storage hydro. *IEEE Trans. Power Systems* 27(2):800–810.

Latorre JM, Cerisola S, Ramos A (2007) Clustering algorithms for scenario tree generation: Application to natural hydro inflows. *Eur. J. Oper. Res.* 181:1339–1353.

Mazumdar M (1975) *Importance Sampling in Reliability Estimation, Reliability and Fault Tree Analysis* (SIAM, Philadelphia), 153–163.

Morales JM, Conejo AJ, Perez-Ruiz J (2009b) Economic valuation of reserves in power systems with high penetration of wind power. *IEEE Trans. Power Systems* 24(2):900–910.

Morales JM, Minguez R, Conejo AJ (2010) A methodology to generate statistically dependent wind speed scenarios. *Appl. Energy* 87:843–855.

Morales JM, Pineda S, Conejo AJ, Carrion M (2009a) Scenario reduction for futures trading in electricity markets. *IEEE Trans. Power Engrg.* 24(2):878–888.

Nowak MP, Römisch W (2000) Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty. *Ann. Oper. Res.* 100(1–4):251–272.

O'Neill RP, Hedman KW, Krall EA, Papavasiliou A, Oren SS (2010) Economic analysis of the $N-1$ reliable unit commitment and transmission switching problem using duality concepts. *Energy Systems* 1:165–195.

Ozturk UA, Mazumdar M, Norman BA (2004) A solution to the stochastic unit commitment problem using chance constrained programming. *IEEE Trans. Power Systems* 19(3):1589–1598.

Papavasiliou A, Oren SS, O'Neill RP (2011) Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Trans. Power Systems* 26(4):2197–2206.

Pereira MVF, Balu NJ (1992) Composite generation/transmission reliability evaluation. *Proc. IEEE* 80(4):470–491.

Pflug GCh (2001) Scenario tree generation for multiperiod financial optimization by optimal discretization. *Math Programming B* 89:251–271.

Piwko R, Clark K, Freeman L, Jordan G, Miller N (2010) Western wind and solar integration study. Technical report, National Renewable Energy Laboratory. http://www.nrel.gov/docs/fy10osti/47781.pdf.

Rockafellar RT, Wets RJ-B (1991) Scenarios and policy aggregation in optimization under uncertainty. *Math. Oper. Res.* 16(1):119–147.

Ruiz PA, Philbrick RC, Sauer PW (2009a) Wind power day-ahead uncertainty management through stochastic UC policies. *Power Systems Conf. Exposition* (IEEE Power and Energy Society, Piscataway, NJ), 1–9.

Ruiz PA, Philbrick RC, Zack E, Cheung KW, Sauer PW (2009b) Uncertainty management in the unit commitment problem. *IEEE Trans. Power Systems* 24(2):642–1651.

Shiina T, Birge JR (2004) Stochastic unit commitment problem. *Internat. Trans. Oper. Res.* 11(95):19–32.

Sioshansi R, Short W (2009) Evaluating the impacts of real time pricing on the usage of wind power generation. *IEEE Trans. Power Systems* 24(2):516–524.

Takriti S, Birge JR, Long E (1996) A stochastic model for the unit commitment problem. *IEEE Trans. Power Systems* 11(3):1497–1508.

Torres JL, Garcia A, De Blas M, De Francisco A (2005) Forecast of hourly wind speed with ARMA models in Navarre (Spain). *Solar Energy* 79(1):65–77.

Tuohy A, Meibom P, Denny E, O'Malley M (2009) Unit commitment for systems with high wind penetration. *IEEE Trans. Power Systems* 24(2):592–601.

van Hulle F (2005) Large scale integration of wind energy in the European power supply: Analysis, recommendations and issues. Technical report, European Wind Energy Association, Brussels.

Wang J, Shahidehpour M, Li Z (2008) Security-constrained unit commitment with volatile wind power generation. *IEEE Trans. Power Systems* 23(3):1319–1327.

Yu N-P, Liu C-C, Price J (2010) Evaluation of market rules using a multi-agent system method. *IEEE Trans. Power Systems* 25:470–479.

Anthony Papavasiliou is an assistant professor in the Department of Mathematical Engineering at the Catholic University of Louvain, and a member of the Center for Operations Research and Econometrics. His research interests are focused on energy systems operations, planning and economics, and optimization under uncertainty. He has served as a researcher for Pacific Gas and Electric and the Federal Energy Regulatory Commission and the Palo Alto Research Center.

Shmuel S. Oren is a professor of industrial engineering and operations research at the University of California at Berkeley. He is also the Berkeley site director of the Power System Engineering Research Center. He was an adviser to the Market Oversight Division of the Public Utilities Commission of Texas and to the Energy Division of the California Public Utilities Commission and is a member of the CAISO Market Surveillance Committee. He is a Fellow of INFORMS and IEEE.

Electronic Companion for "Multi-Area Stochastic Unit Commitment for High Wind Penetration in a Transmission Constrained Network", by Anthony Papavasiliou and Shmuel S. Oren

EC.1. The WECC Model

In Figure EC.1 we present a schematic diagram of the Western Electricity Coordinating Council (WECC) model that is studied in Section 7. The dashed boxes represent load and generation pockets. The thick solid lines represent the import constraints that are defined in Equation (31). Each thick solid line intersects a set of transmission lines IG_j over which the total amount of power cannot exceed a certain limit IC_j. These constraints limit the total flow of power into a load pocket in order to prevent load shedding in the case of generator failure within a load pocket, and also limit the total amount of power flow over combinations of inter-ties that connect the California ISO system to neighboring states. The wind generators of Table EC.1 are located in the five buses that are depicted as solid black circles. In order of appearance from top to bottom, these wind sites are Solano, Altamont, Tehachapi, Clark and Imperial. The net load profile for each day type, which needs to be served by thermal generators and wind power, is shown in Figure EC.2.

EC.2. Data Set for the Multi-Area Wind Model

As we discuss in Section 5, the relationship between wind speed and wind power typically depends on wind turbine size, manufacturer design, and various other factors, and an aggregated power curve can be used to represent a group of turbines. In order to simulate wind power production, we use a piecewise linear approximation of such an aggregate power curve for each location (see the lower right panel of Figure EC.3). We use wind speed and wind power production data from the 2006 data set of the National Renewable Energy Laboratory (NREL) Western Wind and Solar Integration Study (WWSIS) database described in Potter et al. (2008). We study two wind integration cases. The first represents a moderate energy integration level for wind power corresponding to the 2012 renewable integration targets of California, and the second case represents a deep integration level corresponding to the 2020 targets. Ex post we have observed that the moderate integration case corresponds to approximately 7% energy penetration, while the deep integration case corresponds to approximately 14% energy penetration. In the rest of the paper we refer to these cases as moderate and deep integration respectively.

In order to collect data for each case, we examined the interconnection queue of the California ISO until 2020 (CAISO (2010)), and placed individual wind generators in our model by matching the geographical locations of planned wind power installations with the corresponding wind park

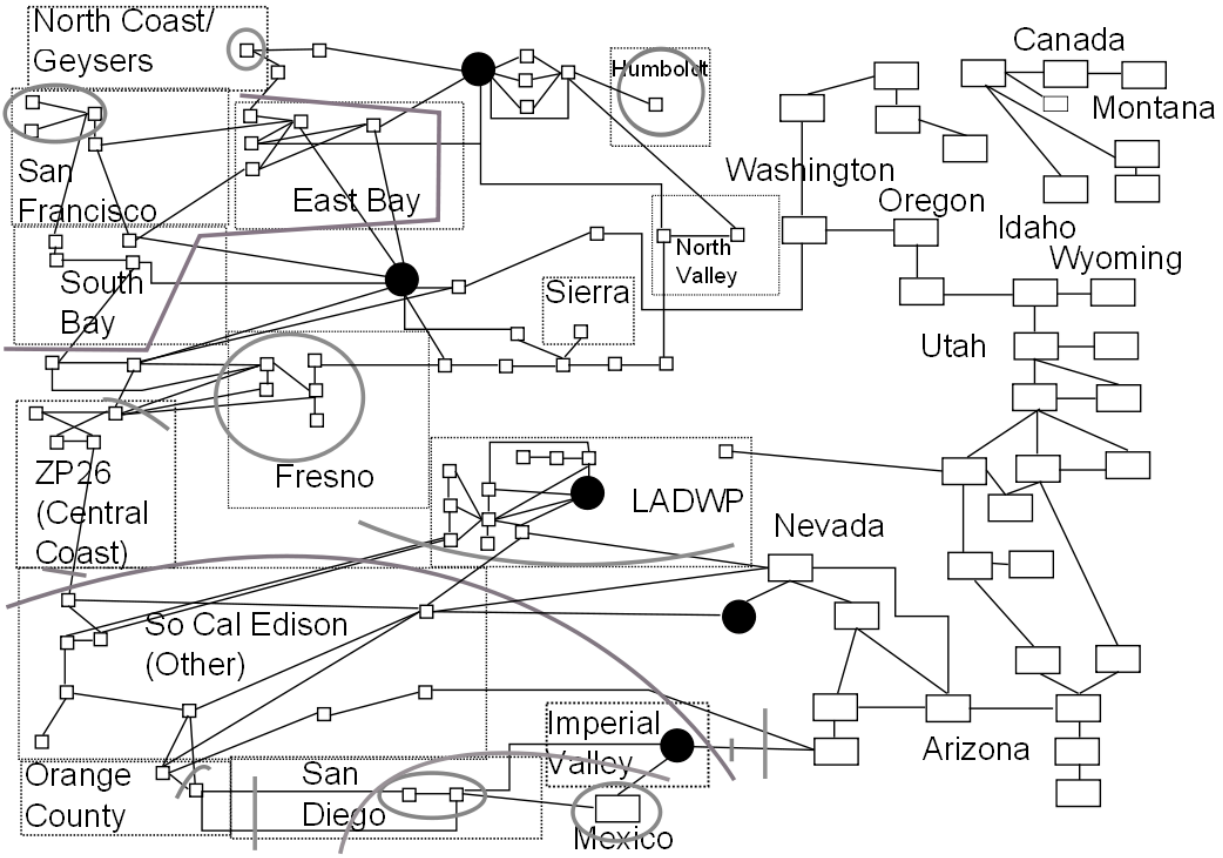


Figure EC.1 A schematic of the WECC model studied in the Results section.

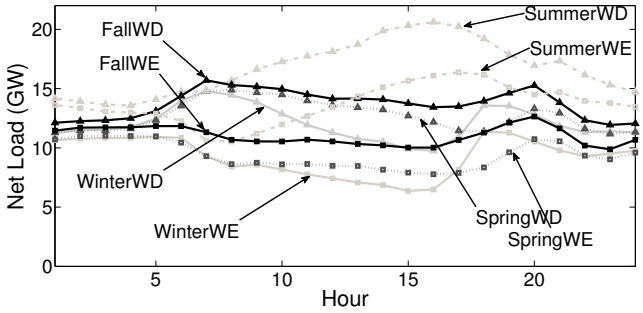


Figure EC.2 Net load for each day type without accounting for wind power production.

data in the WWSIS data set. In Table EC.1 we present the locations of existing wind generation capacity, as well as capacity for the moderate and deep integration cases. In Figure EC.3 we compare the inverse cumulative distribution function of our calibrated model to the raw data set for the deep integration case.

A wind power production time series model with more areas represents the geographical location of each resource more accurately. Neither the stochastic unit commitment model nor the scenario

Table EC.1 Current and projected capacity of wind power installations (MW)

County	Existing	Moderate	Deep
Altamont	954	954	1,086
Clark	-	-	1,500
Imperial	-	-	2,075
Solano	348	848	1,149
Tehachapi	1,346	4,886	8,333
Total	2,766	6,688	14,143

selection algorithm requires more computation for a wind power production model that considers more locations. The choice of five locations is therefore not computational. Instead, the specific choice is the result of trial-and-error with the time series model described in Section 5.

Figure EC.3 indicates that the complementary cumulative probability distribution deviates from the data mostly for high wind output levels for the Tehachapi area. From the power curve of the Tehachapi area we note that the scatter plot of wind speed to wind power exhibits a significant spread. This is due to the fact that Tehachapi covers a wide geographic area with wind parks located in most regions of the area. As a result, the power curve cannot reproduce the high-power results observed in the data. In order to alleviate this problem, we have experimented with further partitioning the Tehachapi area in smaller regions. However, this introduces greater inaccuracy to the model due to the higher dimensions of the correlation matrix $\hat{\Sigma}$ in step (c) of Section 5. As a result, we choose to model five areas as the best compromise between capturing locational dependencies and retrieving marginal wind speed distributions at each location.

The aforementioned discrepancy between the model and data is acceptable in the context of the present study, which focuses on day-ahead reserve requirements. Renewable supply variability affects day-ahead reserve requirements most strongly due to severe renewable shortages and wind power production ramping. Regarding renewable supply shortages, we note from the complementary cumulative probability distributions of Figure EC.3 that the wind power production data is modeled accurately at low wind production levels. In order to account for wind power production ramping, we isolate monthly and diurnal patterns and use a time series model for wind speed.

EC.3. Competing Scenario Selection Policy

In Section 7 we validate our proposed scenario selection algorithm inspired from importance sampling by comparing it to an alternative scenario selection algorithm, referred to as SUC2. The alternative scenario selection algorithm selects wind outcomes based on the scenario reduction algorithm proposed by Dupacova et al. (2003). The aggregate wind time series are allocated in each of the five regions of the model in proportion to the annual share of wind energy in each region, as estimated by the available data set. These multi-area wind scenarios are then multiplexed with

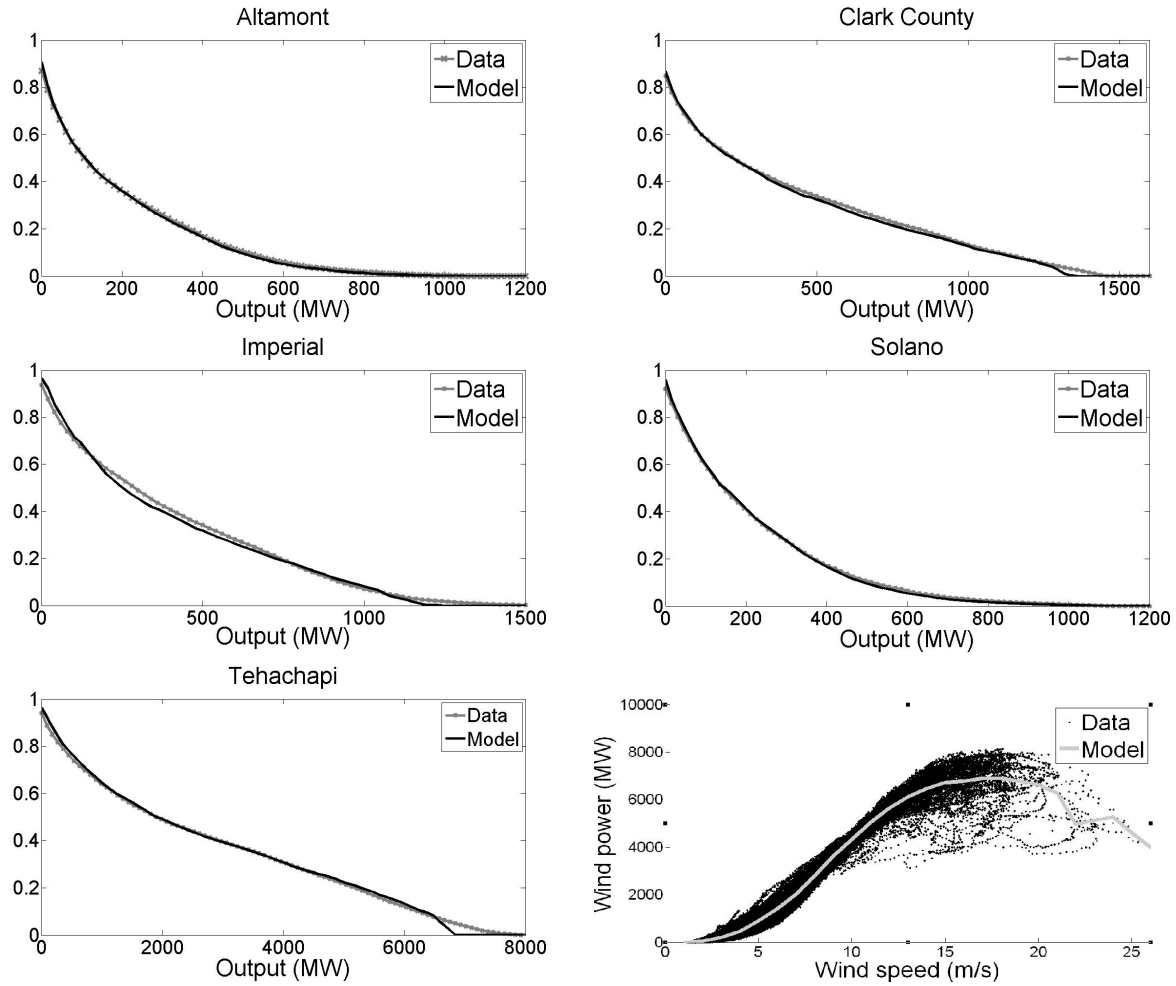


Figure EC.3 In reading order: inverse cumulative distribution functions for Altamont, Clark County, Imperial, Solano and Tehachapi, and power curve at the Tehachapi area for deep integration.

the most severe contingencies by taking the Cartesian product of each wind scenario with each contingency. In particular, for the moderate and deep integration study we choose 7 wind scenarios and multiplex them with the 5 worse contingencies (of which 2 represent generator failures and 3 represent line failures) and the no-contingency case. For the deep-simple integration study we use the scenario reduction algorithm of Dupacova et al. (2003) for selecting 42 single-area wind power production outcomes. For the zero wind case study, we are multiplexing 42 contingencies (including the no-failure contingency) with a single wind realization of zero wind in all regions.

As we discuss in Section 7.3, we assume a probability of generator failure of 1% (Pereira and Balu (1992)) and a probability of transmission line failure of 0.1% (Grigg et al. (1999)). The probability of losing a single generator and no lines is then computed to be equal to 0.2443, the probability of losing a single line and no generator is 0.0698, and the probability of no contingency is 0.1861. For the moderate and deep integration studies, the probability of each scenario is obtained as

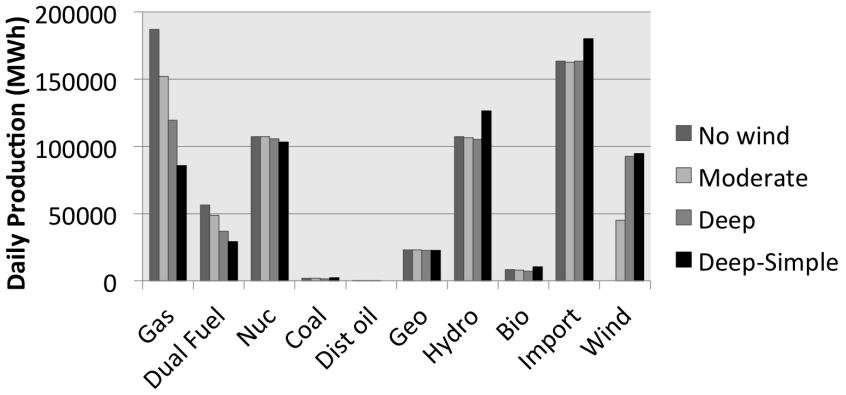


Figure EC.4 Fuel mix of the system for each case study.

the product of the failure probability of each element and the probability of each wind scenario (assumed equal across wind scenarios). By assigning weights in this way, we maintain a relative weighting of scenarios with the same wind outcome that is consistent with the relative probability of the contingencies associated with each scenario. We also maintain a relative weighting across scenarios with the same contingency that is consistent to the relative probability of the wind outcomes associated with each scenario. The obtained probabilities are then normalized in order to sum to unity.

EC.4. A Detailed Exposition of Operating Costs

In this section we provide a detailed exposition of the results presented in Section 7. The fuel mix of the system is presented in Figure EC.4. As wind is increasingly integrated in the system, it primarily displaces gas units and dual fuel units. Nuclear units are also displaced slightly. In Table 1 we note that only six coal units exist in the system, totaling a capacity of 285.9 MW. The contribution of coal to the fuel mix is therefore minimal. In the case where transmission constrains and contingencies are ignored, it is possible to rely more heavily on imports, hydro resources, wind and biomass. This results in a further displacement of gas, dual fuel and nuclear units and a further reduction of operating costs.

In Tables EC.2 - EC.5 we present the total daily production, committed capacity and number of units provided by fast and slow resources respectively, for each case study and each policy. It can be observed, in all cases involving wind integration, that the deterministic policy commits a larger number of fast and slow units, and that the total power output of conventional units in the deterministic unit commitment model exceeds that of the alternative policies. This indicates that the deterministic policy is performing poorly in terms of utilizing resources that are already committed, as well as utilizing non-conventional resources such as hydro, biomass, imports and

Table EC.2 Energy supply, capacity and number of fast and slow generators for the deep integration study.

	SUC-IS	SUC2	Det 30%	Foresight
Fast MWh	65,229	64,989	72,856	65,947
Slow MWh	198,479	198,601	191,961	197,257
Fast MW	175.5	174.7	173.9	177.8
Slow MW	1459.2	1458.1	1489.4	1480.6
No. fast	33.2	33.1	37.2	32.3
No. slow	20.4	20.6	22.2	19.6

Table EC.3 Energy supply, capacity and number of fast and slow generators for the deep-simple integration study.

	SUC-IS	SUC2	Det 30%	Foresight
Fast MWh	57,952	68,061	76,899	57,307
Slow MWh	162,344	152,187	188,261	162,581
Fast MW	167.8	167.8	172.6	172.0
Slow MW	1657.0	1740.8	1506.5	1660.0
No. fast	34.0	39.9	40.6	32.1
No. slow	14.1	12.5	21.0	14.3

Table EC.4 Energy supply, capacity and number of fast and slow generators for the moderate integration study.

	SUC-IS	SUC2	Det 30%	Foresight
Fast MWh	81,957	83,530	92,347	82,958
Slow MWh	227,887	226,351	218,815	226,368
Fast MW	177.4	175.9	174.6	177.9
Slow MW	1326.5	1335.4	1352.6	1340.4
No. fast	36.4	38.8	41.4	36.0
No. slow	21.7	21.4	23.3	21.5

Table EC.5 Energy supply, capacity and number of fast and slow generators for the no-wind integration study.

	SUC-IS	SUC2	Det 30%	Foresight
Fast MWh	94,702	94,058	97,190	96,258
Slow MWh	257,979	258,883	256,688	256,594
Fast MW	175.2	176.0	175.0	175.1
Slow MW	1219.8	1219.0	1215.3	1226.5
No. fast	38.7	39.1	38.0	39.8
No. slow	24.1	23.6	26.2	23.0

wind power. An important benefit of stochastic unit commitment is the ability of the model to commit locational resources in response to uncertainty, whereas deterministic reserve rules commit resources globally.

In Table EC.6 we present the standard deviation of cost for each day type for the stochastic unit commitment policy. As we explain in Section 7.1, we are able to obtain these results by parallelizing the Monte Carlo simulation and running 1,000 days of operations for each day type. Note the very low standard deviation in the no-wind case, despite the fact that this case has the highest operating costs. This confirms the degree of uncertainty introduced by wind power supply variability.

Table EC.6 Cost and standard deviation of cost in million dollars for each day type for the proposed stochastic unit commitment policy SUC-IS.

		WWD	SWD	SuWD	FWD	WWE	SWE	SuWE	FWE
Deep	Cost	6.007	5.996	12.058	8.739	3.690	3.374	8.800	5.808
	St. Dev.	2.073	1.930	1.586	1.808	1.680	1.509	1.399	1.629
Deep-Smpl	Cost	3.711	3.589	9.580	6.097	3.223	1.196	6.158	3.462
	St. Dev.	1.737	1.695	1.443	1.677	1.547	889	1.259	1.454
Moderate	Cost	8.284	8.384	13.828	10.374	5.499	5.239	10.400	7.295
	St. Dev.	1.337	1.297	1.071	1.181	1.238	1.244	946	1.116
No wind	Cost	10.217	10.869	15.645	12.136	7.301	7.503	12.060	8.936
	St .Dev.	0447	0.447	0.602	0.435	0.423	0.405	0.465	0.404