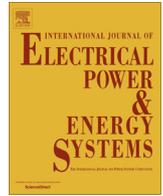




Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Stochastic simulation of power systems with integrated intermittent renewable resources



Yannick Degeilh, George Gross*

University of Illinois at Urbana–Champaign, 4052 Electrical and Computer Engineering Building, 306 North Wright Street, Urbana, IL 61801, USA

ARTICLE INFO

Article history:

Received 21 March 2014

Received in revised form 12 July 2014

Accepted 16 July 2014

Available online 15 August 2014

Keywords:

Monte Carlo/stochastic simulation
Transmission-constrained day-ahead markets
Production costing
Reliability
Emissions
Renewable resource integration

ABSTRACT

We report on the development of a comprehensive, stochastic simulation methodology that provides the capability to quantify the impacts of integrated renewable resources on the power system economics, emissions and reliability variable effects over longer periods with the various sources of uncertainty explicitly represented. We model the uncertainty in the demands, the available capacity of conventional generation resources and the time-varying, intermittent renewable resources, with their temporal and spatial correlations, as discrete-time random processes. We deploy Monte Carlo simulation techniques to systematically sample these random processes and emulate the side-by-side power system and transmission-constrained day-ahead market operations. We construct the market outcome sample paths for use in the approximation of the expected values of the various metrics of interest. Our efforts to address the implementational aspects of the methodology so as to ensure computational tractability for large-scale systems over longer periods include the use of representative simulation periods, parallelization and variance reduction techniques. Applications of the approach include planning and investment studies and the formulation and analysis of policy. We illustrate the capabilities and effectiveness of the simulation approach on representative study cases on a modified WECC 240-bus system. The results provide valuable insights into the impacts of deepening penetration of wind resources.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

The deepening penetration of intermittent renewable resources presents major challenges in power system planning and operations in light of their highly time-varying nature and their associated geographical and climatological sources of uncertainty. Indeed, unlike conventional resource outputs, wind and solar resource outputs cannot be controlled by the operator except to be curtailed. The high variability in wind speeds and insolation patterns, both temporal and spatial, results at times in intermittent wind and solar resource outputs [1]. A consequence is that the wind and solar outputs do not necessarily track the load pattern and thus cannot always contribute to serve the peak loads. There are also concerns about “spilling” of wind energy at night due to the insufficient load demand and the physical impossibility to shut down the base-loaded conventional units for short periods. While morning and mid-day solar power outputs are aligned with the loads, their quick decline after sunset occurs when the loads are still high. Both wind and solar resources therefore impose addi-

tional requirements on the conventional units to effectively manage the variability/intermittency and uncertainty effects. Further issues arise from the fact that the wind speed and insolation patterns show various degrees of spatial correlation, resulting in highly variable nodal power injections which may lead, at times, to congestion. These complications illustrate the critical need to appropriately represent the temporal and spatial correlations of the wind and solar resource outputs in the assessment of the power system performance. Such need implies that the various renewable resource outputs, as well as the demands and conventional generation available capacities, must be modeled by random processes (*r.p.s*) so as to capture the impacts of their variability across time and space. These requirements drive the need for a comprehensive simulation tool that can effectively quantify the expected economic, reliability and emission variable effects of power system with integrated renewable resources.

The conventional probabilistic simulation approach [2] and its extensions [3,4] cannot adequately provide the needed level of detail due to its inability to represent chronological phenomena such as the grid operations and their impacts on the day-ahead market (*DAM*) outcomes, as well as the time-dependent nature and temporal correlations of the demands and supply resources, particularly the renewable resources. A distinctly different

* Corresponding author.

E-mail addresses: degeilh1@illinois.edu (Y. Degeilh), gross@illinois.edu (G. Gross).

approach, which may be used to represent the uncertain *DAM* outcomes with the capability to explicitly represent the grid constraints, is the probabilistic optimal power flow (*P-OPF*), [5,6]. One drawback of the *P-OPF* approach, however, is that it requires a number of significant simplifying assumptions to render the problem into a solvable form. For instance, the representation of the power system evolution over time, including the temporal correlations among the system variables, requires the formulation of a *multi-period P-OPF*, whose tractability is questionable even for a small number of periods. Many renewable integration studies in the literature report the use of the Monte Carlo simulation to represent the power system and its sources of uncertainty. Most focus exclusively on the probabilistic modeling of a single renewable resource, generally, wind [7,8]. We are not aware of a comprehensive approach which integrates under a single umbrella the various sources of uncertainty that impact power system operations across time. In this paper, we report on the development of a comprehensive analytical framework and general Monte Carlo simulation approach with the capability to assess, over longer duration periods, the impacts on power grid variable effects of the variable demands, renewable generation outputs and conventional resource available capacities. While our approach can easily be adapted to incorporate various stochastic models, including those based on copulas, statistical transforms for multivariate dependence such as principal component analysis, time-series synthesis using many variants of *ARMA*-type schemes, numerical weather prediction methods, historical time-series re-sampling and hindcast [9–13], our objective is to construct a practical scheme based on models that require no calibration nor the use of complex transforms, unlike the various models just mentioned. As such, we construct appropriate stochastic models to capture the time-varying and uncertain behavior of multi-site renewable power outputs, with the cross-correlations between the sites and time correlations explicitly accounted for and to incorporate into a comprehensive stochastic simulation framework. Our approach, while relatively easy to implement, can handle any type of renewable output probability distribution, including non parametric distributions, as we require no assumptions on the shape of their joint cumulative distribution functions (*j.c.d.f.s*). The implementation, in fact, ensures the computational tractability of the approach for realistic sized power systems.

In the analytic framework, we represent the demands and supply resource outputs as discrete-time *r.p.s*. In particular, we model the multi-site wind (solar) power outputs over time as a discrete-time *r.p.* whose *j.c.d.f.* explicitly incorporates the spatial and temporal correlations among the wind (solar) power outputs at all the sites for all periods of interest. For concreteness, we assume that the power system described in this paper operates in a market environment. Our approach uses an hour as the smallest indecomposable unit of time and uses the realizations of the *r.p.s* at these subperiods. In addition, a snapshot representation of the grid is used to represent the impacts of the transmission constraints on the hourly day-ahead markets (*DAMs*) outcomes. Our Monte Carlo simulation methodology uses systematic sampling mechanisms to generate the realizations of the various *r.p.s* and to construct the so-called *sample paths* (*s.p.s*) [14, p. 97]. We note that such *s.p.s* embody the correlations among the constituent random variables (*r.v.s*) of the *r.p.s*. We use the *s.p.s* of the demand, multi-site renewable output and generator available capacity *r.p.s* as inputs into the emulation of the transmission-constrained hourly *DAMs*. We compute the market clearing results of the transmission-constrained hourly *DAMs* using the solution of the linearized optimal power flow (*OPF*) typically used in the *ISO* markets [15, p. 534]. The outcomes of the *DAMs* are used to construct the *s.p.s* of the so-called market outcome *r.p.s*. Clearly, these *s.p.s* capture

the correlations among the various market outcomes. We use the hourly realizations that constitute these *s.p.s* to compute the metrics of interest used to assess the performance of the power system and associated markets. These metrics include the hourly expected locational marginal prices (*LMPs*), revenues of the generators, total payments made by buyers in the *DAMs*, congestion rents, the system-wide CO_2 emissions, as well as the reliability indices *LOLP* and *EUE*. We note that these metrics are computed by explicitly accounting for the deliverability of the electricity. The methodology is also able to capture the seasonal effects in demands and renewable outputs, the impacts of maintenance scheduling and the ramifications of new policy initiatives. There is a broad range of applications of the simulation methodology to operations, resource planning studies, production costing issues, investment analysis, transmission utilization, reliability analysis, environmental assessments, policy formulation and to answer quantitatively various *what-if* questions. We also discuss the various implementational steps, such as parallelization, variance reduction and deployment of representative weeks, we used to improve the computational tractability of the proposed approach.

The paper contains four additional sections. In Section ‘The proposed simulation framework’, we describe the overall structure of the simulation approach by formally introducing the time frame in which the simulations are performed and the general Monte Carlo procedure. We also discuss in detail the modeling of the input *r.p.s*, the sampling schemes used to generate their *s.p.s*, and the mapping of these *s.p.s* into outcome *s.p.s* via the *OPF* solutions. In Section ‘Implementational aspects’, we discuss the steps implemented to improve the computational tractability of the approach. In Section ‘Illustrative studies’, we illustrate the capabilities of the approach with two representative case studies that focus on the impacts of deepening wind penetration and the substitutability of conventional resources by wind resources. We conclude with a summary in Section ‘Conclusion’ together with directions for future work.

The proposed simulation framework

The proposed simulation approach emulates the side-by-side power system operations and transmission-constrained markets, with the objective to evaluate, on average, the variable effects of the power system with integrated renewable resources. In this section, we describe the overall Monte Carlo procedure, including its time frame and the generation of the market outcome *s.p.s* from which we compute the power system metrics of performance.

The simulation is carried out for the specified study period \mathcal{T} . We decompose \mathcal{T} into non-overlapping simulation periods \mathcal{T}_i 's such that $\mathcal{T} = \bigcup_i \mathcal{T}_i$. We define each simulation period \mathcal{T}_i in such a way that the system resource mix and unit commitment, the transmission grid, the operating policies, the market structure and the seasonality effects remain unchanged over its duration. While there are many possible choices for a simulation period, we specify each simulation period to be of one-week duration. This choice captures the load patterns over the week and week-end days, and is able to incorporate the maintenance schedules. We further decompose each simulation period into subperiods, where a subperiod is the smallest indecomposable unit of time represented in the simulation. We assume that each variable is constant over the entire duration of a subperiod. The simulation, as such, ignores any phenomenon whose time scale is smaller than a subperiod. We choose to use subperiods of one hour duration as an acceptable compromise between the level of detail needed for a realistic representation of the power system and market operations and the computational tractability of the simulation. The subperiod selection is particularly appropriate as many existing *DAMs* are cleared on an hourly basis. We denote by h the index of the

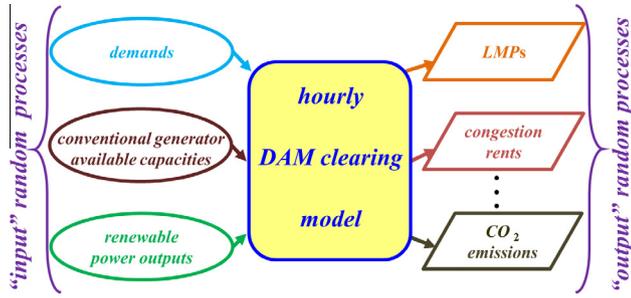


Fig. 1. Conceptual structure of the approach.

subperiods in a simulation period T_i such that $T_i = \{h : h = 1, \dots, 168\}$.¹

The simulation emulates the side-by-side power system operations and transmission-constrained market. Specifically, in each simulation period, we emulate the sequence of the 168 hourly DAM clearings where the market outcomes determine the contributions to the performance metrics of interest. We use the hourly discretized time axis in the construction of all the models and evaluate the metrics on an hourly basis. The modeling of the highly variable demands, multi-site renewable power outputs and conventional generator available capacities, all of which are uncertain and participate in the hourly DAMs, is in terms of discrete-time *r.p.s* which are collections of *r.v.s* indexed by the 168 hours of the simulation period. These *input r.p.s* are mapped by the DAM clearing mechanism into the *output* discrete-time *r.p.s*, also indexed by the 168 hours of the simulation period, that represent the market outcomes. We provide a conceptual representation of such a mapping in Fig. 1. We make use of Monte Carlo simulation techniques in the simulation approach to evaluate the performance metrics. Such procedure is formally described in what follows.

We denote by $\{X[h] : h = 1, \dots, 168\}$ a *r.p.* defined over 168 hours. We denote by $\{x[h] : h = 1, \dots, 168\} = \{x[1], x[2], \dots, x[168]\}$ a collection of time-indexed realizations – also known as a *sample path (s.p.)* – of the *r.p.* $\{X[h] : h = 1, \dots, 168\}$ [14, p. 97]. Note that a *s.p.* captures the time-dependencies of the individual realizations $x[h]$ in a manner that is consistent with the correlations that exist among the *r.v.s* $X[h]$, $h = 1, \dots, 168$. Our simulation approach makes use of the so-called *independent Monte Carlo* [16, p. 10], and requires the construction of multiple *independent and identically distributed (i.i.d.) s.p.s* for each output *r.p.* to evaluate the performance metrics. Note that in this context, the phrase “*i.i.d. s.p.s*” has the sense that the *s.p.s* constitute the realizations of independent identically distributed *r.p.s*. A simulation run is the basic process through which we construct a *s.p.* for each of the output *r.p.s*. It consists of sampling each one of the input *r.p. j.c.d.f.s* – or joint probability mass functions (*j.p.m.f.s*) depending on the *r.p.* model – in order to generate the *s.p.s* whose hourly realizations are used as inputs in the corresponding hourly DAMs. The *s.p.s* of the input *r.p.s* are mapped into *s.p.s* of the output *r.p.s* via the model of the market clearing mechanism for each hourly DAM of the simulation period. Thus, a *s.p.* for each output *r.p.* is obtained from the series of 168 cleared DAMs.

We carry out multiple simulation runs in order to create the output *s.p.s* from which we estimate the performance metrics for each market outcome of interest. We select our performance metrics to be the expected values of the time-indexed *r.v.s* making up the output *r.p.s* of interest.² Let M be the number of simulation runs;

¹ We point out that the proposed approach is sufficiently general to be applicable to other time scales.

² Other metrics may be defined along similar lines. It is also possible to approximate the output *r.p. j.c.d.f.s*.

we estimate, for an output *r.p.* $\{Y[h] : h = 1, \dots, 168\}$, the hourly sample mean point estimate $\bar{y}[h]$ of each *r.v.* $Y[h]$, $h = 1, \dots, 168$:

$$\bar{y}[h] = \frac{1}{M} \sum_{j=1}^M y^{(j)}[h] \quad (1)$$

where $y^{(j)}[h]$ is the realization of *r.v.* $Y^{(j)}[h]$ in simulation run j (note that the $Y^{(j)}[h]$'s are *i.i.d.*). The number of simulation runs M depends on the statistical reliability requirements specified for the estimation of the desired expected values. We define the statistical reliability of the hourly sample mean estimator $\bar{Y}[h] = \frac{1}{M} \sum_{j=1}^M Y^{(j)}[h]$ to be the length of the $100(1-\alpha)\%$ confidence interval with $0 < \alpha < 1$ for the true mean $\mu_{Y[h]}$ of *r.v.* $Y[h]$ [17, p. 82, p. 451].

According to the Central Limit Theorem, the sample mean estimator $\bar{Y}[h]$ is approximately normally distributed for large M [18, p. 78].

Thus, we can establish that $\mu_{Y[h]}$ lies in the interval

$$\left[\bar{Y}[h] - z_{(1-\alpha/2)} \frac{\sigma_{Y[h]}}{\sqrt{M}}, \bar{Y}[h] + z_{(1-\alpha/2)} \frac{\sigma_{Y[h]}}{\sqrt{M}} \right]$$

with a $100(1-\alpha)\%$ probability, where $\sigma_{Y[h]}$ is the standard deviation of *r.v.* $Y[h]$, and $z_{1-\alpha/2} = \Phi^{-1}(1-\alpha)$, with Φ^{-1} the inverse of the cumulative distribution function (*c.d.f.*) of the standard normal distribution $\mathcal{N}(0, 1)$. Note that the length of the confidence interval is a function in \sqrt{M}^{-1} . While it is possible to select M so as to set the confidence interval length and achieve the desired statistical reliability, in practice, a function in \sqrt{M}^{-1} decays very slowly for large M ; beyond a certain value of M , the improvement in statistical reliability is generally too small to warrant the extra computing-time needed to perform additional simulation runs.

We discuss the stochastic models and sampling procedures for the input *r.p.s*. For concreteness, the only renewable resource we consider in the rest of our analysis is wind.

Let \mathcal{B} the set of buyers in the hourly DAMs. For simplicity and clarity in the notation, we assume that each buyer $b \in \mathcal{B}$ submits a demand bid for a load located at one node and one node only. From the outset, we wish to capture the spatial and temporal correlations among the various buyer demands. Now, given that the cleared demands, as observed from historical load data, are seasonal and have a weekly cycle, we assume that, in each week of the same season, the buyer demands over a week period can be modeled by the discrete-time *r.p.* $\{\underline{D}[h] : h = 1, \dots, 168\}$, where

$$\underline{D}[h] = [D^1[h], \dots, D^{|\mathcal{B}|}[h]]^\dagger$$

and \dagger denotes the transpose. Such a *r.p.* is the collection of time-indexed random vectors $\underline{D}[h]$ for $h = 1, \dots, 168$, with each random vector $\underline{D}[h]$ in hour h containing the ordered collection of the buyer demand *r.v.s* for each buyer $b \in \mathcal{B}$. Such representation explicitly accounts for the correlations across buyer and time that exist among the hourly demands $D^b[h]$ of each buyer $b \in \mathcal{B}$. For clarity, we may represent all the jointly distributed *r.v.s* $D^b[h]$, $\forall b \in \mathcal{B}$, $h = 1, \dots, 168$ making up the buyer demand *r.p.* in the following array:

$$\begin{bmatrix} \underline{D}^1[1] & \underline{D}^1[2] & \dots & \underline{D}^1[168] \\ \underline{D}^2[1] & \underline{D}^2[2] & \dots & \underline{D}^2[168] \\ \vdots & \vdots & \ddots & \vdots \\ \underline{D}^{|\mathcal{B}|}[1] & \underline{D}^{|\mathcal{B}|}[2] & \dots & \underline{D}^{|\mathcal{B}|}[168] \end{bmatrix}$$

We now describe how to construct and sample, in practice, the discrete-time *r.p.* of the (hourly) buyer demands over a week. We gather weeks of *simultaneously-measured* hourly buyer demands from a seasonally disaggregated historical database so as to capture the cross-dependencies among the buyer demands across multiple time periods. We use these data to construct the sample space $\Omega_{\{D[h]: h=1, \dots, 168\}}$ of the buyer demand *r.p.* Note that each weekly set of simultaneously-measured hourly buyer demands constitutes a *s.p.* of $\{D[h]: h=1, \dots, 168\}$ and as such, contains realizations of each *r.v.* $D^b[h]$, $h=1, \dots, 168$, $b \in \mathcal{B}$. We assume the equi-probability of each one of the *s.p.s* retrieved from the historical data to approximate the *j.p.m.f.* of $\{D[h]: h=1, \dots, 168\}$. We proceed to discuss the sampling procedure of such *j.p.m.f.* The method is the multidimensional case of the procedure described in [18, p. 139], to generate a realization from a (discrete) *r.v.* “empirical” *c.d.f.* The procedure entails drawing one of the historical *s.p.s* making up the sample space $\Omega_{\{D[h]: h=1, \dots, 168\}}$ with probability one over the total number of *s.p.s* making up said sample space. The selected sample-path contains hourly realizations $d^b[h]$, $h=1, \dots, 168$, $b \in \mathcal{B}$ that are consistent with the correlations existing among the $D^b[h]$, $b \in \mathcal{B}$, $h=1, \dots, 168$, *r.v.s* making up the buyer demand *r.p.* Another way to restate this statement is to say that, since every historical *s.p.* has the cross-dependency information relating the hourly buyer demands embedded in it, so does the selected *s.p.*

We apply an analogous approach to the stochastic modeling of the multi-site hourly wind speeds. We denote a wind farm location by index $i \in \mathcal{I}$. For simplicity in the notation, we assume that each wind farm is a distinct seller in the hourly DAMs. We define a one-to-one and onto mapping between each wind farm location i and its seller $s \in \mathcal{S}^w$ in the market, where \mathcal{S}^w is the collection of the $|\mathcal{S}^w| = |\mathcal{I}|$ sellers at the nodes where the farms are located. We assume that each wind speed at each farm location is uniform for the entire farm. Furthermore, we assume that the wind speeds are seasonal and have a daily cycle. In a similar manner as with the hourly buyer demands, we seek to capture the spatial and temporal correlations of the wind speed *r.v.s* $V_i[h]$ across locations $i \in \mathcal{I}$ and hours of the day $h=1, \dots, 24$. Thus, we represent the multi-site hourly wind speeds by the discrete-time *r.p.* $\{V[h]: h=1, \dots, 24\}$, where $V[h] = [V_1[h], \dots, V_{|\mathcal{I}|}[h]]^\dagger$. We note that, similarly as for the buyer demand *r.p.*, the multi-site wind speed *r.p.* is the collection of time-indexed random vectors $V[h]$ for $h=1, \dots, 24$, with each random vector $V[h]$ in hour h containing the ordered collection of the wind speed *r.v.s* at the multiple sites represented in set \mathcal{I} . The construction and sampling procedures of such a discrete-time *r.p.* closely follow those of the buyer demand *r.p.* As such, a *s.p.* of the multi-site wind speed *r.p.* is a collection of hourly wind speed realizations at all the sites. Such a collection of hourly wind speeds is representative of the wind speed patterns at the multiple sites and so captures the existing cross dependencies. In the specific case of the multi-site wind speed *r.p.* however, we need to generate 7 daily *s.p.s* in order to construct the *s.p.* for the 7×24 hours in a week. Let us denote by $v^{(j)}$ the j th *s.p.* independently drawn from the *j.p.m.f.* of $\{V[h]: h=1, \dots, 24\}$. We sample and collect 7 independent *s.p.s* from the *j.p.m.f.* of $\{V[h]: h=1, \dots, 24\}$ to obtain the *s.p.* for the

week $\{\underline{v}^{(1)}, \underline{v}^{(2)}, \dots, \underline{v}^{(7)}\}$. Note that $\underline{v}^{(j)}$ does not necessarily have to represent the multi-site wind speeds of day j ; it can represent any arbitrary day in the week. However, for simplicity in the notation, we make it represent day j . Under this notation, $v_i^{(j)}[h]$ denotes the wind speed at wind farm $i \in \mathcal{I}$ in hour h of day j .

At this stage, we may convert the *s.p.* wind speeds into their corresponding power outputs. To do so, we make use of the wind farm characteristic power curves, using the procedure described in appropriate detail in [4]. As such, the power output of a particular wind farm is a piece-wise polynomial function of its wind speed. Note that by converting all the *s.p.s* making up the sample space of $\{V[h]: h=1, \dots, 24\}$, we obtain the corresponding multi-site wind power output *r.p.* $\{W[h]: h=1, \dots, 24\}$ in a straightforward manner. Now, exploiting the one-to-one and onto mapping that relates a wind farm location $i \in \mathcal{I}$ to its seller $s \in \mathcal{S}^w$, we may denote by $w_i^{(j)}[h] = (w^s)^{(j)}[h]$ the wind farm power output that is obtained from the conversion of $v_i^{(j)}[h]$. For convenience in the rest of the paper however, and to reflect the fact that we have constructed a multi-site wind power output *s.p.* for the week, we drop the dependency on exponent j and express h in terms of a week hour, so that $w^s[h]$ denotes the wind power output of seller $s \in \mathcal{S}^w$ in hour h of the week.

We introduce a simplifying assumption for the set \mathcal{S}^c of market participants who sell the outputs of the conventional generation resources: similarly as for the buyer demands, seller $s \in \mathcal{S}^c$ offers the output of only one conventional resource. As such, we shall index each conventional generation resource by $s \in \mathcal{S}^c$. We model each conventional resource as a multi-state unit with two or more states – outaged, various partially derated capacities and full capacity. We assume that each conventional resource is statistically independent of any other generation resource. As such, if $A^s[t]$ designates the available capacity of seller s unit at time t , then $A^s[t]$ and $A^{s'}[t]$, with $s \neq s'$, $s, s' \in \mathcal{S}^c$, are statistically independent. We use a discrete-event driven Markov Process model with the appropriate stochastic event-times distributions to represent the underlying *r.p.* governing the available capacity of each conventional resource. We assume statistically independent exponentially-distributed *r.v.s* to represent the transition times between the states. Such model allows us to explicitly represent the periods during which a conventional unit might be up, down, or running at derated capacities in the simulation.

In light of the statistical independence assumption, we construct individual *s.p.s* for each conventional resource. The methodology for simulating the available capacity of a conventional resource over time is well documented in the literature, and can be found under the names of *next-event method*, *state duration sampling*, or simply *sequential simulation* [19]. The procedure consists of simulating the sequence of available capacity states through which the conventional resource passes over time. We do so by sampling the appropriate transition-time exponential distribution as the conventional resource transitions from a state of the discrete-event driven Markov Process model to another [20, p. 59]. In terms of our approach, the state of seller s resource, i.e., its available capacity, is hence determined (after rounding the sampled transition-time to the nearest hour if necessary) for each hour h of the week. The collection of hourly realizations $\{a^s[1], a^s[2], \dots, a^s[168]\}$ constitutes a week-long *s.p.* of seller s resource available capacity.

We note that the random process-based models for the buyer demands, multi-site wind speeds (power outputs) and conventional generation resource available capacities are built and

sampled independently from one another.³ By doing so, we implicitly assume that the r.v. $V_i[h]$ that represents the wind speed at site i in hour h , the r.v. $D^b[h]$ that represents the buyer b demand in hour h' and the r.v. $A^s[h]$ that represents the available capacity of seller s conventional resource in hour h'' , are statistically independent of one another, with $h, h', h'' \in \{1, \dots, 168\}$. In other words, we assume that the behavior of buyer b demand in hour h has no impact on the behavior of the wind speed (power output) at site i (or the available capacity of any conventional resource s) in any hour and vice versa. We note that such statistical independence assumptions result from the same lines of reasoning as the widely-deployed assumption that the demand and available capacities of conventional resources are statistically independent.

The input *s.p.s* – as generated from the *j.p.m.f* of their respective *r.p.* – are key to the determination of the outcomes of the hourly DAMs in the simulation period. They serve to determine the maximum power outputs (withdrawals) that are offered (bid) into the DAM in hour h . Specifically:

- the sampled realization $d^b[h]$ of $D^b[h]$ determines the maximum demand quantity bid by buyer $\tilde{b} \in \mathcal{B}$ in hour $h \in \mathcal{T}_i$;
- the sampled realization $w^s[h]$ of $W^s[h]$ determines the maximum wind capacity offered by seller $\tilde{s} \in \mathcal{S}^w$ in hour $h \in \mathcal{T}_i$; and
- the sampled realization $a^s[h]$ of $A^s[h]$ determines the maximum power output offered by seller $\tilde{s} \in \mathcal{S}^c$ in hour $h \in \mathcal{T}_i$.

In this way, the sampled realizations serve to parametrize the model of the hourly DAM clearing mechanism in hour h . We refer the reader to the Appendix for the statement of the formulation of the hourly DAM clearing in terms of a linearized OPF problem, typically in use in ISOs. We use the shorthand $\mathcal{M}(\mathcal{S}, \mathcal{B})|_{[h]}$ to denote the optimization program associated with the hourly DAM clearing mechanism. At the optimum, the dual variables associated with the power balance equations constitute the LMPs in hour h at each node of the transmission grid. The LMPs are used as the nodal prices of the MWh commodity in hour h .⁴ The optimal solution to $\mathcal{M}(\mathcal{S}, \mathcal{B})|_{[h]}$ also yields the optimal resource dispatch, i.e., the power output (consumption) of each generation resource $s \in \mathcal{S}$ (demand $b \in \mathcal{B}$), as well as the power flows on the grid lines for the hour h . We use these optimal values to compute the hourly realizations of the output *r.p.s* of interest. Fig. 2 illustrates conceptually the mapping of the input *r.p.*' *s.p.s* into the output *r.p.*' *s.p.s* for a given simulation run.

In a simulation, we carry out multiple simulation runs, i.e., we collect multiple *i.i.d.* *s.p.s* of a given output *r.p.*, from which we compute the associated performance metrics, i.e., the hourly expected values that are assessed using Eq. (1). In the case of the reliability metrics, we note that a loss of load event occurs only in case the fixed demand is not met. Therefore, the evaluation of the LOLP and EUE reliability metrics refers purely to the fixed demand and is not meaningful for price responsive demands. In the simulation, we represent the fixed demands as price-sensitive demands with their willingness-to-pay set at the outage cost figures used in the evaluation of the VOLL [21]. In this way, the linearized OPF gives highest priority to serve the fixed demand in light of its higher willingness-to-pay. We note that a loss of load event may be due to a supply shortfall or lack of transmission transfer capability, and that the reliability metrics we compute reflect that

fact. At a node with loss of load, the unserved energy in hour h is the shortfall in supply that results in not meeting the total fixed demand at that node in that hour. We use the indicator function $i_{(0,+\infty)}(\cdot)$ to compute the system LOLP contribution in hour h . The function takes for argument the system unserved energy (i.e. the sum of all the nodal unserved energies) and return 1 whenever the system unserved energy is strictly positive, 0 otherwise. Let $U[h]$ the system unserved energy. Given that $LOLP[h] = E\{i_{(0,+\infty)}(U[h])\}$, we approximate LOLP[h] using Eq. (1) with the realizations of $i_{(0,+\infty)}(U[h])$ as shown in the following equation:

$$LOLP[h] \approx \frac{1}{M} \sum_{m=1}^M i_{(0,+\infty)}(u^{(m)}[h]). \quad (2)$$

Implementational aspects

We devote this section to a discussion of the implementational aspects to improve the computational tractability of the proposed simulation approach. A first step in the improvement of the computational tractability is the judicious selection of the number of simulation periods to be simulated. We take advantage of the fact that several weeks in a season have similar load shapes and wind patterns, and that certain resources are scheduled for planned outages in view of maintenance operations. In such cases, we select an appropriate representative week among them for simulation and weigh its results in the study period by the number of weeks it represents. In this way, we reduce the number of simulation periods to be run to cover the entire study period, thereby cutting down the computational efforts. Typically, for regions with four distinct seasons, 14–18 representative weeks suffice to cover the entire annual simulation, and so the overall simulation time is cut by about a third.

Another measure to reduce the computational burden is to systematically “warm-start” the market clearing optimizations [22]. The basic idea of a “warm start” is to provide the simplex algorithm used to solve the market clearing optimization formulated as a linear program (LP), as shown in the Appendix, with a “good” initial solution, i.e., a solution that is “close” to the desired optimal solution. The provision of such a “good” solution may reduce considerably the number of simplex pivots required to reach the optimal solution, thereby reducing the computational time requirements. In our approach, we make extensive use of our finding that the optimal solution to the LP in hour h is “close” to that of the LP in hour $(h - 1)$, as the demands, conventional generation available capacities and renewable resource outputs do not, typically, change much from an hour to the next. Thus, to solve the DAM in hour h , we provide our LP solver engine with an initial solution that is simply the solution to the DAM in hour $(h - 1)$. Our extensive testing shows that a reduction of 35% in simulation time is, typically, obtained with “warm start”.

We also have studied the application of a wide range of variance reduction techniques. Our findings indicate that only the control variate technique [18, p. 57], is effective in bringing about significant variance reduction. The use of the hourly aggregated available generation capacity, i.e., the sum of conventional resource and wind available capacities, as a control variate in each hour h can reduce computing times significantly for some of the metrics, in particular, the economic measures. For example, we have seen a 2 to 1 reduction in the evaluation of the average hourly total wholesale payments in the extensive testing we performed. On the other hand, the control variate scheme performs poorly in the evaluation of the reliability indices due to the weak correlation

³ For the purpose of the following discussion, any statement on the wind speed random process and its time-indexed random variables applies equally well to a statement on its associated wind power output random process and its time-indexed random variables, since the wind power output is a function of the wind speed.

⁴ In the simulation, if the fixed demand at node n cannot entirely be met (loss of load event), we set the LMP at the regulatorily specified cap price, typically, the marginal price of the most expensive conventional resource.

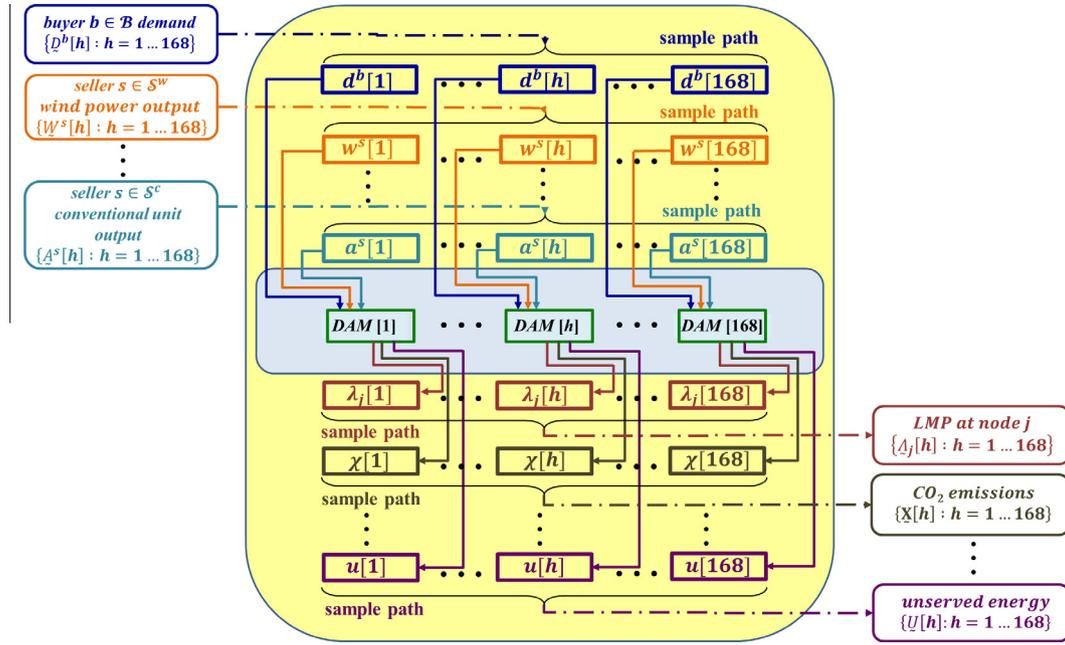


Fig. 2. Mapping of the input s.p.s into the output s.p.s for a given simulation run.

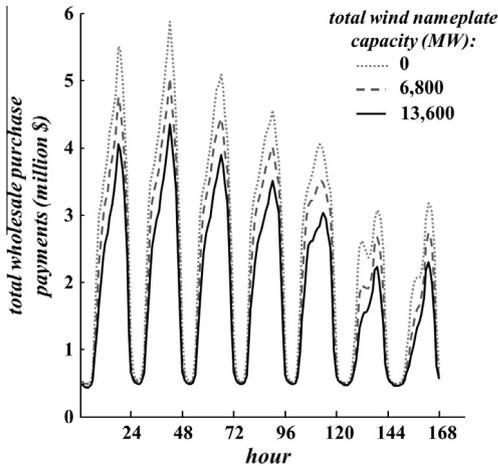


Fig. 3. Expected hourly total wholesale purchase payments over the “average week”.

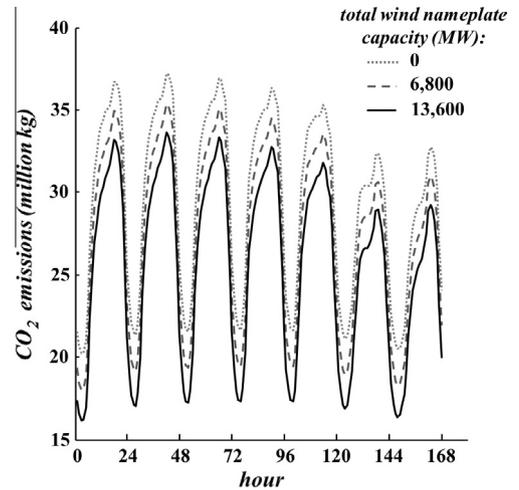


Fig. 4. Expected hourly total CO₂ emissions over the “average week”.

observed in practice among the control variate and the hourly total unserved energy. Such a result occurs due to the rarity of loss of load events. Consequently, the random variable representing the hourly total unserved energy is, in an arbitrary hour, very much akin to a constant equal to 0, save for the few positive outcomes – each with very low probability of occurrence – that quantify the unserved energy in the rare loss of load event cases. To put things in perspective, if one computes the correlation coefficient between a r.v. such as the hourly aggregated available generation capacity and a r.v. that is essentially equal to 0, such as the hourly total unserved energy, the correlation coefficient will be nearby 0. Experimental results are in line with this intuition: computed correlation coefficients among the hourly aggregated available generation capacity (the control variate) and the hourly total unserved energy (the random variable of interest) are very close to 0 in all hours, which is not practical for the control variate scheme that requires that the control variate and the random variable of interest be strongly correlated. We also make use of the Latin

Hypercube Sampling (*LHS*) technique in the systematic generation of the random numbers needed to sample the multivariate probability distributions of our input r.p.s. The application of the technique is along the lines of [23]. Our experience indicates, however, that the *LHS* technique does not significantly contribute to the variance reduction of our estimates and, consequently, to savings in the overall simulation time.

A significant improvement of the method’s computational tractability comes from the parallelization of the simulation of each representative week on dedicated cores/computers. In this way, the overall simulation time is dramatically reduced; indeed, the time reduction to essentially a single simulated week becomes possible whenever there are as many computers/cores as the number of representative weeks. We can take further advantage of parallelization from the fact that the simulation runs are statistically independent from one another. As such, we also parallelize the construction of the s.p.s. Such parallelization process can additionally reduce the overall computation times, with the reduction

depending on the number of dedicated cores available. As a result, the parallelization of a representative week simulation runs on a machine with X cores will divide the simulation time by X .

Illustrative studies

We performed extensive testing of the simulation approach on a broad range of applications, including resource planning, production costing issues, transmission planning, environmental assessments, reliability and policy analysis. We illustrate the application of the approach with two sets of representative studies carried out on a modified WECC 240-bus system [24]. The studies in this paper use scaled load data for the year 2004, historical wind data from the WECC geographic footprint [1] and offer data based on marginal cost information. In these case studies, we scale the load data so that the annual peak load is 81,731 MW. The 902 conventional generation units of the test system have a total nameplate capacity of 96,443 MW. The system incorporates from 1 to 4 wind farms at distinct Californian locations. We use for these farms the wind turbine characteristics, including power curves, described in the NREL wind integration studies [1]. We assume that each buyer bids his load as a fixed demand in each hourly DAM. Owing to the fact that wind power has no fuel cost, we assume that wind power is offered at 0 \$/MWh in each hourly DAM throughout the simulation period. For each study, we limit our analysis to a single year in order to focus on the insights into the nature of the results obtained. Taking into account the seasonality effects, as well as the maintenance schedules for the conventional generation units, we select 16 representative weeks for the simulation periods to quantify the variable effects over the 52 weeks of the year. For the test system, our extensive numerical studies indicate that beyond 100 simulation runs, there is too little improvement in the statistical reliability of the economic and emission metrics to warrant the extra computing time required for the execution of additional simulation runs. On the other hand, the computation of the reliability metrics required about 500 simulation runs, owing to the fact that our test system is relatively reliable and the loss of load events constitute rare occurrences.

In the first set of case studies, we examine the power system behavior under deepening wind penetration: from 0 MW total nameplate capacity in the base case to 13,600 MW in increments of 3,400 MW. The system has 4 wind farms with equal nameplate capacities. The locations of the wind farms are fixed and remain unchanged throughout this case study. All the case studies are with reserves margin set at 15%. Figs. 3 and 4 respectively show the average hourly total wholesale purchase payments and total CO₂ emissions for the “average week”, wherein the hourly values for each hour $h = 1, \dots, 168$ are averaged over all the representative weeks of the year with the appropriate weights. For clarity, we only display the base case (no wind) and the cases with 6,800 and 13,600 MW of total wind nameplate capacity at the 4 wind farms. In Table 1, we provide a summary of the annual average figures for the total wholesale purchase payments, the CO₂ emissions, as well as the annual reliability indices LOLP and EUE. In the third column, we give the annual MWh purchase price, which is computed by averaging the LMPs over nodes and time. Note that the

LMPs are weighted by the amount of load cleared at the associated nodes. These results clearly indicate that, as the wind penetration deepens, the wholesale purchase payments, average MWh purchase price and CO₂ emissions are reduced, while there are rather marked improvements in the system reliability indices. We note that the reductions and improvements are characterized by diminishing returns as the wind penetration deepens.

We further illustrate in Figs. 5 and 6 the impacts of deepening penetrations of integrated wind resources on the volatility in the hourly MWh purchase prices. Specifically, Fig. 5 shows the hourly expected values of the MWh purchase price over the “average week” while Fig. 6 shows its hourly standard deviations. We observe the more pronounced volatility – in terms of the larger standard deviation – in the hourly MWh purchase price during the peak-hours, when the electricity prices are, typically, the highest. However, we note that deeper penetrations of integrated wind

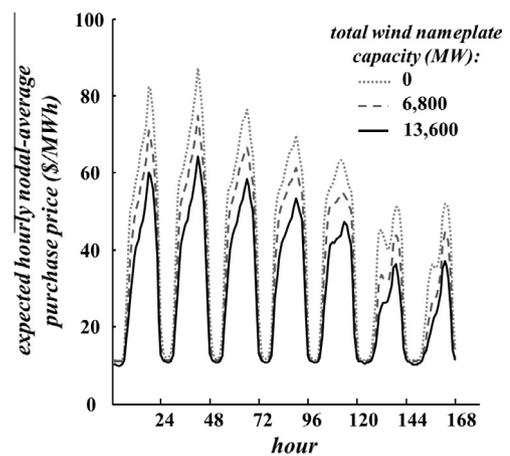


Fig. 5. Expected values of the hourly MWh purchase prices over the “average week”.

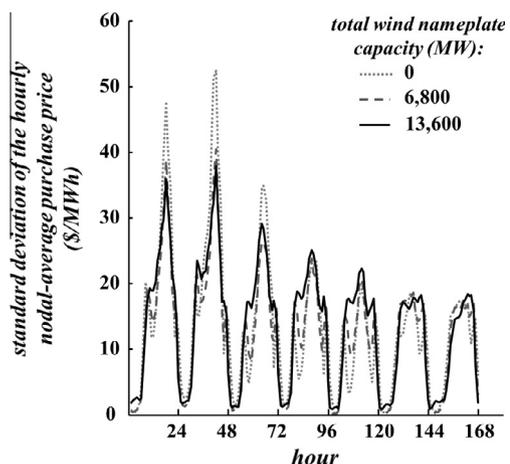


Fig. 6. Standard deviations of the hourly MWh purchase prices over the “average week”.

Table 1
Annual metrics of interest for the various wind penetrations.

Wind installed capacity (MW)	Wholesale purchase payments (10^9 \$)	MWh purchase price (\$/MWh)	CO ₂ emissions (10^9 kg)	LOLP	EUE (MWh)
0	21.99	41.66	257.82	0.0019	285
3400	20.28	38.37	248.81	0.0011	152
6800	18.63	35.20	239.95	0.0008	102
10,200	17.02	32.10	231.24	0.0006	79
13,600	15.52	29.24	222.79	0.0005	68

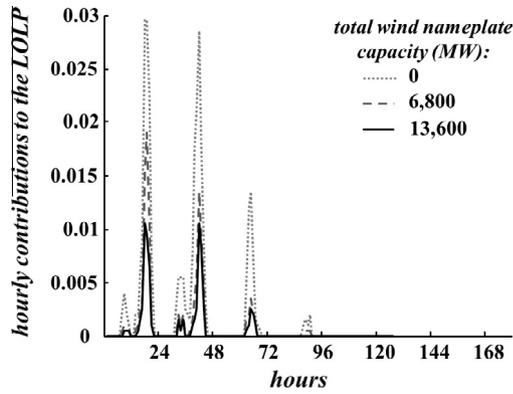


Fig. 7. Hourly LOLP contributions over the “average week”.

resources tend to exacerbate such volatility also on the days with lower peak loads, on the Thursday–Sunday period, but tend to reduce the volatility in the hours with the highest loads in the Monday–Wednesday period. An analysis of these results suggests that wind generation contributes to avert scarcity events in the hours of the week with the highest loads. The plot in Fig. 7 illustrates such phenomenon. The fewer scarcity events imply reduced occurrences of price spikes. As a result, the volatility in the MWh purchase prices at the peak load hours is reduced. On the other hand, wind generation tends to induce increased volatility in the electricity prices during the peak hours of the less-heavily-loaded days. In such cases, the variability of the wind outputs tends to change the set of generators that contribute to the determination of the LMPs in each simulation run. Due to these changes, the resulting LMPs have higher volatility.

In the second set of case studies, we investigate to what extent wind resources may substitute for conventional resources from purely a system reliability perspective. The base case has no installed wind capacity and uses 15% reserves margin to evaluate the system reliability. In the other sensitivity cases, the conventional resource mix is supplemented by the same 4 wind farms considered in the second set of case study with the 13,600 MW total nameplate capacity. We examine the impacts of retiring conventional resource capacity, thus leading to lowered reserves margin levels (we consider that the reserves margin is provided by the conventional resources only). Fig. 8 shows the LOLP and EUE as a function of the retired conventional capacity and resulting reserves margins after installation of the 13,600 MW of wind capacity, with the dashed line showing the associated reliability index for the base case with no wind and the 15% reserves margin. The

simulation results indicate that the 13,600 MW of installed wind capacity – about 16.6% of the annual peak load 81,731 MW – can substitute for about 4% of the weekly peak loads, on average over the year, in terms of retired conventional generation capacity, that is about 3,000 MW. In other words, from purely a system reliability perspective, wind resources constitute rather poor substitutes for conventional resources, since 13,600 MW of wind power nameplate capacity can substitute for only 22% of the conventional resource capacity on an annual basis.

Conclusion

In this paper, we present the comprehensive, stochastic simulation framework we developed to emulate the side-by-side behavior of power system and market operations over longer-term periods. Our approach makes detailed use of discrete-time *r.p.s* in the adaptation of Monte Carlo simulation techniques. As such, the framework can explicitly represent various sources of uncertainty in the demands, the available capacity of conventional generation resources and the time-varying, intermittent renewable resources, with their temporal and spatial correlations. In addition, the simulation methodology represents the impacts of the network constraints on the market outcomes. In this way, the simulation approach is able to quantify the impacts of integrated renewable resources on power system economics, reliability and emissions. The stochastic simulation approach has a broad range of applications in planning, operational analysis, policy formulation and analysis and to provide quantitative assessments of various *what if* case studies.

The representative results we present from the extensive studies performed effectively demonstrate the strong capabilities of the simulation approach. The results of these studies on a modified WECC 240-bus system, making use of scaled load data and historical wind data in the WECC geographic footprint clearly indicate that the integration of deepening levels of wind resources to a preexisting system may effectively drive the total wholesale purchase payments and CO₂ emissions down, as well as improve system reliability. There are, however, diminishing returns on these benefits as higher penetrations of wind power are achieved. Wind generation is also shown to exacerbate (reduce) the volatility in electricity prices in hours when the system is moderately (heavily) loaded. Simulation results indicate that, from a pure system reliability perspective, the wind resources constitute a fairly poor substitute for conventional resources.

Future work includes taking advantage of the tool design to gain insights into the integration of other intermittent, time-varying resources, such as solar, active demand response resources and

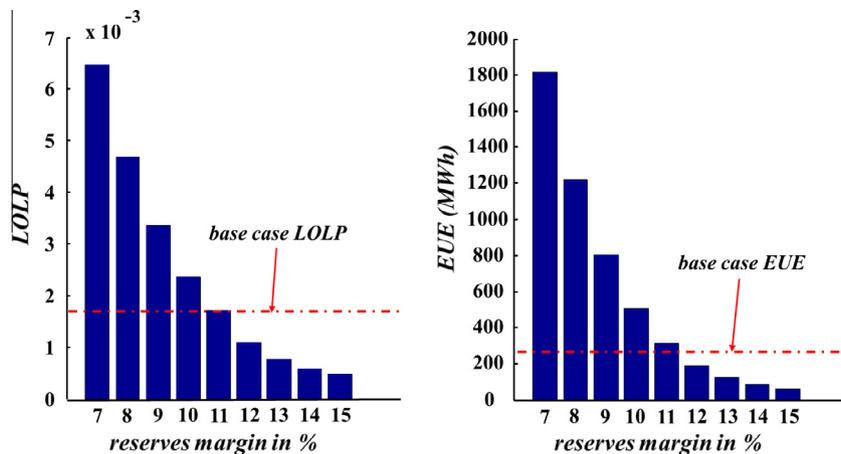


Fig. 8. Annual LOLP and EUE versus system reserves margins.

utility-scale storage units into the grid. Another topic of considerable interest is the analysis of the impacts of intermittent resource integration on ramping capability requirements.

Acknowledgements

Research performed was supported in part by the National Science Foundation under grant NSF ECCS-0925754 “Managing Intermittency in Planning and Operations of Power”, PSERC, the U.S. Department of Energy for “The Future Grid to Enable Sustainable Energy Systems” and Stanford University GCEP.

Appendix: Hourly DAM clearing model

We provide a model for the hourly DAM clearing mechanism. We make use of the lossless DC power flows to model the grid [15, p. 534], as is the practice in today’s ISO-run markets. We further assume that the sellers (buyers) submit piecewise linear offer (bid) functions, which we denote by $\gamma^s(\cdot)$ for seller $s \in \mathcal{S} = \mathcal{S}^c \cup \mathcal{S}^{sw}$ ($\beta^b(\cdot)$ for buyer $b \in \mathcal{B}$). Under these assumptions, the proposed OPF is a linear program.

Let $\mathcal{N} = \{n : n = 0, 1, \dots, (N-1)\}$ be the set of network buses with bus 0 being the slack bus, and $\mathcal{L} = \{l = 1, \dots, L\}$ the set of transmission lines. Let matrices \underline{A} , \underline{B}_d and \underline{B} designate the reduced branch to node incidence, the branch susceptance and the reduced nodal susceptance matrices, respectively. We denote by \underline{b}_0 the column vector of the augmented susceptance matrix corresponding to the slack node and by $\underline{\theta}$ the vector of voltage phase angles at the $|\mathcal{N}| - 1$ buses other than the slack bus. We denote by $\underline{f} = \underline{B}_d \underline{A} \underline{\theta}$ the vector of line flows, \underline{f}^m and \underline{f}^m the vectors of transmission line ratings in each flow direction. We specify $(\kappa^s)^m$ to be the minimum capacity of seller $s \in \mathcal{S}^c$ conventional resource. We also define the conventional generation (wind farm generation) power injection at node n in hour h as $p_n^c[h] = \sum_{s \in \mathcal{S}^c} \text{at node } n g^s[h]$ ($p_n^w = \sum_{s \in \mathcal{S}^{sw}} \text{at node } n g^s[h]$), where $g^s[h]$ is the output of seller s generation resource in hour h . The power consumption due to loads at node n in hour h is similarly denoted by $p_n^d[h] = \sum_{b \in \mathcal{B}} \text{at node } n \ell^b[h]$, where $\ell^b[h]$ is the cleared demand of buyer b in hour h . We use $\lambda_n[h]$ to denote the dual variable associated with the power balance equation at bus n in hour h . The linearized OPF in hour $h \in \mathcal{T}_i$ is formulated as:

$$\max_{\ell^b[h], g^s[h]} \left(\sum_{b \in \mathcal{B}} \beta^b(\ell^b[h]) - \sum_{s \in \mathcal{S}} \gamma^s(g^s[h]) \right) \quad (\text{A.1a})$$

subject to

$$\left(\underline{p}^c[h] + \underline{p}^w[h] \right) - \underline{p}^d[h] = \underline{B}\underline{\theta}[h] \leftrightarrow \underline{\lambda}[h] \quad (\text{A.1b})$$

$$\left(p_0^c[h] + p_0^w[h] \right) - p_0^d[h] = \underline{b}_0^t \underline{\theta}[h] \leftrightarrow \lambda_0[h] \quad (\text{A.1c})$$

$$-\underline{f}^m \leq \underline{B}_d \underline{A} \underline{\theta}[h] \leq \underline{f}^m \quad (\text{A.1d})$$

$$0 \leq \ell^b[h] \leq d^b[h], \quad \forall b \in \mathcal{B} \quad (\text{A.1e})$$

$$(\kappa^s)^m \leq g^s[h] \leq a^s[h], \quad \forall s \in \mathcal{S}^c \quad (\text{A.1f})$$

$$0 \leq g^s[h] \leq w^s[h], \quad \forall s \in \mathcal{S}^{sw}. \quad (\text{A.1g})$$

References

- [1] GE Energy. Western wind and solar integration study. NREL, tech rep; 2010. <<http://www.nrel.gov/docs/fy10osti/47434.pdf>>.
- [2] Balériaux H, Jamouille E, De Guertechin F Linard. Simulation de l’exploitation d’un parc de machines thermiques de production d’électricité couplées à des stations de pompage. *RevueE (édition SRBE)* 1967;5(7):3–24.
- [3] Zhang Y, Chowdhury A. Reliability assessment of wind integration in operating and planning of generation systems. In: Proceedings of the IEEE PES general meeting, Calgary; July 26–30 2009. p. 1–7.
- [4] Maisonneuve N, Gross G. A production simulation tool for systems with integrated wind energy resources. *IEEE Trans Power Syst* 2011;26(4):2285–92.
- [5] Verbic G, Canizares C. Probabilistic optimal power flow in electricity markets based on a two-point estimate method. *IEEE Trans Power Syst* 2006;21(4):1883–93. Nov.
- [6] Madrigal M, Ponnambalam K, Quintana V. Probabilistic optimal power flow. In: IEEE Canadian conference on electrical and computer engineering, vol. 1; 1998. p. 385–8.
- [7] Anderson C, Cardell J. The impact of wind energy on generator dispatch profiles and carbon dioxide production. In: 45th Hawaii international conference on system science (HICSS), Maui; January 04–07, 2012. p. 2020–6.
- [8] Vallee F, Lobry J, Deblecker O. System reliability assessment method for wind power integration. *IEEE Trans Power Syst* 2008;23(3):1288–97.
- [9] Haghi H, Bina M, Golkar M. Nonlinear modeling of temporal wind power variations. *IEEE Trans Sustain Energy* 2013;4(4):838–48.
- [10] Burke D, O’Malley M. A study of principal component analysis applied to spatially distributed wind power. *IEEE Trans Power Syst* 2011;26(4):2084–92.
- [11] Tastu J, Pinson P, Madsen H. Space–time scenarios of wind power generation produced using a gaussian copula with parametrized precision matrix. Technical University of Denmark, tech rep; 2013. <http://orbit.dtu.dk/fedora/objects/orbit:122325/datastreams/file_f6e2279f-5894-4d85-a4dd-e351e6048fda/content>.
- [12] Chen N, Qian Z, Nabney I, Meng X. Wind power forecasts using gaussian processes and numerical weather prediction. *IEEE Trans Power Syst* 2013;PP(99):1–10.
- [13] Rajagopalan B, Lall U, Tarboton D, Bowles D. Multivariate nonparametric resampling scheme for generation of daily weather variables. *Stoch Hydrol Hydraul* 1997;11:65–93.
- [14] Hajek B. Notes for ECE 534 – an exploration of random processes for engineers. Univ. of Illinois at Urbana–Champaign; 2009.
- [15] Wood A, Wollenberg B. Power generation, operation and control. NY: John Wiley and Sons, Inc.; 1996.
- [16] Fishman G. A first course in Monte Carlo. Duxbury; 2006.
- [17] Kleijnen J. Statistical techniques in simulation – Parts 1 and 2. New York: Marcel Dekker, Inc.; 1974.
- [18] Bratley P, Fox BL, Schrage LE. A guide to simulation. New York: Springer-Verlag; 1983.
- [19] Ubeda J, Allan R. Sequential simulation applied to composite system reliability evaluation. *IEE Proc Part C: Gener Transm Distrib* 1992;139(2):81–6.
- [20] Asmussen S, Glynn P. Stochastic simulation: algorithms and analysis. Springer; 2007.
- [21] Silva V. Value of flexibility in systems with large wind penetration. Ph.D. dissertation, Imperial College London; 2010. <http://tel.archives-ouvertes.fr/docs/00/72/43/58/PDF/VSilva_PhDAIChapters_v9_-_unlinked_-_test.pdf>.
- [22] Stott B, Hobson E. Power system security control calculations using linear programming. *IEEE Trans Power Ap Syst* 1978;PAS-97(5):1713–31.
- [23] Kowli A, Gross G. Quantifying the variable effects of systems with demand response resources. In: Proceedings of the IREP symposium – bulk power system dynamics and control, Buzios; August 01–06, 2010.
- [24] Price J, Goodin J. Reduced network modeling of WECC as a market design prototype. In: Proceedings of the IEEE PES general meeting, Detroit; July 24–28, 2011. p. 1–6.