

Measuring the Benefits of Delayed Price-Responsive Demand in Reducing Wind-Uncertainty Costs

Seyed Hossein Madaeni and Ramteen Sioshansi, *Senior Member, IEEE*

Abstract—Demand response has benefits in mitigating unit commitment and dispatch costs imposed on power systems by wind uncertainty and variability. We examine the effect of delays in consumers responding to price signals on the benefits of demand response in mitigating wind-uncertainty costs. Using a case study based on the ERCOT power system, we compare the cost of operating the system with forecasts of future wind availability to a best-case scenario with perfect foresight of wind. We demonstrate that wind uncertainty can impose substantive costs on the system and that demand response can eliminate more than 75% of these costs if loads respond to system conditions immediately. Otherwise, we find that with a 30-minute lag in the response, nearly 72% of the value of demand response is lost.

Index Terms—Power system economics, wind power generation, wind forecast errors, real-time pricing, unit commitment

NOMENCLATURE

A. Model Sets and Parameters

- T time index set,
- I conventional generator index set,
- W wind generator index set,
- $c_i^V(\cdot)$ generator i 's variable cost function,
- c_i^{NL} generator i 's no-load cost,
- c_i^{SU} generator i 's startup cost,
- K_i^- generator i 's minimum operating point,
- K_i^+ generator i 's maximum operating point,
- R_i^- generator i 's rampdown limit,
- R_i^+ generator i 's rampup limit,
- $\bar{\rho}_i^{SP}$ generator i 's spinning reserve capacity,
- $\bar{\rho}_i^{NS}$ generator i 's non-spinning reserve capacity,
- τ_i^- generator i 's minimum-down time,
- τ_i^+ generator i 's minimum-up time,
- $\bar{\omega}_{w,t}$ maximum generation available from wind generator w in time period t ,
- $p_t(\cdot)$ inverse demand function in time period t ,
- $\bar{\eta}_t$ total reserve requirement in time period t , and
- η_i^{SP} spinning reserve requirement (as a fraction of total) in time period t .

B. Model Decision Variables

- $q_{i,t}$ energy provided by generator i in time period t ,
- $\rho_{i,t}^{SP}$ spinning reserves provided by generator i in time period t ,
- $\rho_{i,t}^{NS}$ non-spinning reserves provided by generator i in time period t ,
- $u_{i,t}$ binary variable indicating if generator i is online in time period t ,
- $s_{i,t}$ binary variable indicating if generator i is started-up in time period t ,
- $h_{i,t}$ binary variable indicating if generator i is shutdown in time period t ,
- $\omega_{w,t}$ energy provided by wind generator w in time period t , and
- l_t load served in time period t .

C. Miscellaneous Parameters

- Ω_w nameplate capacity of wind generator w ,
- $\phi_{w,t}$ maximum generation available from wind generator w in time period t , as a fraction of nameplate capacity,
- ϵ_t wind availability forecast error in time period t ,
- ν_t innovation in forecast error of time period t ,
- μ autocorrelation coefficient of wind availability forecast error,
- D_t actual historical demand in time period t ,
- p^{ret} average retail price of energy in 2005,
- σ_α total annual social welfare in case α , and
- v_α total annual wind energy use in case α .

I. INTRODUCTION

INTEREST in the use of renewable electricity has increased lately. Wind is currently a leading renewable technology, due to its relative maturity and low cost. Although its marginal generation cost is near zero, wind can impose ancillary costs on the power system. Such costs are largely due to the variable and uncertain nature of real-time wind availability, which can require greater use of high-cost, fast-responding, flexible generation. Wind variability can create large ramps in the net (of wind generation) system load. Wind uncertainty can increase the need for fast-responding generators to accommodate sudden and unanticipated increases or decreases in wind availability. In extreme cases, the system may not have enough conventional generating capacity committed and available to respond to unanticipated decreases in wind availability. Studies place the cost of providing these types of services at about \$5/MWh of wind generated [1]–[3].

S. H. Madaeni was with the Integrated Systems Engineering Department, The Ohio State University, Columbus, OH 43210, USA. He is now with the Short Term Electric Supply Department, Pacific Gas and Electric Company, San Francisco, CA 94105, USA (e-mail: SHM8@pge.com).

R. Sioshansi is with the Integrated Systems Engineering Department, The Ohio State University, Columbus, OH 43210, USA (e-mail: sioshansi.1@osu.edu).

The opinions expressed and conclusions reached are solely those of the authors and do not represent the official position of Pacific Gas and Electric Company.

Wind variability and uncertainty can also be accommodated using demand response. Having electricity demand follow wind output reduces the need for fast-responding generation. Papavasiliou and Oren [4] study the use of load control, wherein deferrable loads are directly controlled and scheduled to follow wind availability. They develop two methodologies for load scheduling and estimate the value of such a scheme. Klobasa [5] examines the effects of demand response in a future German power system with 48 GW of wind, showing that it reduces wind-uncertainty costs to less than €2/MWh. Sioshansi [6] studies the Texas (ERCOT) system with 14 GW of wind and real-time pricing (RTP). He shows that RTP can eliminate up to 93% of wind-uncertainty costs, depending on the price-responsiveness of the demand. Dietrich *et al.* [7] examine the effect of demand shifting and peak shaving on wind integration, showing that these programs can reduce wind-uncertainty costs by up to 30%.

These analyses implicitly assume that demand responds to real-time signals immediately, without any latency. While this assumption may be reasonable for some forms of direct load control, it can be more tenuous for indirect price-based mechanisms, such as RTP. This is because there may be a lag between price signals being sent, consumers observing them, and adjusting their behavior in response. Automated controls may alleviate such latency, however, since they reduce the need for consumers to exert real-time control. Such latency can reduce the value of RTP in mitigating wind-uncertainty costs, since its benefit arises from load quickly responding to wind availability and reducing the need for generators to provide balancing energy. Thus, a shortcoming of this literature is that it does not account for such latency in estimating the benefits of demand response in mitigating wind-uncertainty costs.

We address this shortcoming by studying the effect of consumer delays in responding to price signals on the benefits of RTP in reducing wind-uncertainty costs. This paper has two main contributions. The first is that we adapt existing techniques, based on day-ahead unit commitment and real-time dispatch models, to simulate RTP with a time lag in demand responding to prices. Second, we use an ERCOT-based case study [6] to quantify the effects of a time lag on the value of RTP in reducing wind-uncertainty costs. This is especially valuable given the interest in using demand response to accommodate wind in power systems [4]–[8]. We demonstrate that having a 30-minute lag in the demand response reduces the value of RTP in mitigating wind-uncertainty costs by up to 63%, compared to an immediate load response. The remainder of this paper is organized as follows: Sections II and III describe our modeling approach and case study, respectively, Section IV summarizes our results, and Section V concludes.

II. MODELING APPROACH

Our analysis focuses on the impacts of uncertain and variable wind availability on short-run unit commitment and dispatch and the resulting costs. This is done by comparing the cost of operating the system if imperfect wind forecasts are used when scheduling generators to a counterfactual best-case, in which wind availability is known with perfect foresight.

Since our interest is in studying wind-uncertainty costs, wind availability is the only parameter modeled as being uncertain.

System operations are modeled in a rolling fashion one day at a time. This is done using day-ahead unit commitment and real-time dispatch models, both of which have 15-minute timesteps. Both models are formulated as mixed-integer linear programs in GAMS and solved using the branch and cut algorithm in CPLEX 9.0. The unit commitment model uses a point forecast of future wind availability. This model determines unit commitments for each day using a 36-hour optimization horizon. The additional 12 hours are included to ensure that sufficient generating capacity remains committed at the end of each day to serve the following day’s load.

The real-time dispatch model then rolls forward through each 15-minute time period to determine generator dispatch, taking present and future wind availability (through the end of the 36-hour horizon of the day-ahead model) into account. When determining the time- t dispatch, the real-time model uses actual time- t wind availability and forecasts of future wind availability. We assume that these forecasts are less accurate for time periods that are further in the future. Wind availability and forecasts are iteratively updated as the real-time model rolls forward through each 15-minute period. The real-time model holds the generator commitments fixed based on the day-ahead solution, but allows fast-start generators and generators that are off-line but providing non-spinning reserves to be started up in real-time, as necessary. This rolling process is repeated 96 times for each day (once for each 15-minute time period), at which point the model rolls forward to the next day and the process is repeated starting with the day-ahead unit commitment model.

Both the day-ahead and real-time models are deterministic. If it is based on a distribution that accurately characterizes wind-availability, a stochastic model can provide more robust commitment and dispatch decisions. These operational decisions can reduce wind-uncertainty costs compared to using a deterministic model with point forecasts of wind availability [9]–[11]. Since our analysis uses a deterministic model, it overestimates wind-uncertainty costs and the benefits of RTP in mitigating them. Madaeni and Sioshansi [12] study the value of RTP in reducing wind-uncertainty costs if deterministic and stochastic models are used. They demonstrate that the combination of stochastic optimization and RTP reduces wind-uncertainty costs relative to a case with fixed loads by between 22% and 66%. They further show that if a deterministic model is used, RTP alone achieves about 94% of the benefits of RTP and stochastic planning together. Thus, the bias introduced by our use of a deterministic optimization is relatively small.

A. Day-Ahead Unit Commitment Model Formulation

The day-ahead unit commitment model is formulated as:

$$\max \sum_{t \in T} \left\{ \int_0^{l_t} p_t(x) dx - \sum_{i \in I} [c_i^V(q_{i,t}) + c_i^{NL} u_{i,t} + c_i^{SU} s_{i,t}] \right\}, \quad (1)$$

$$\text{s.t. } l_t = \sum_{i \in I} q_{i,t} + \sum_{w \in W} \omega_{w,t}, \quad \forall t \in T; \quad (2)$$

$$\sum_{i \in I} (\rho_{i,t}^{SP} + \rho_{i,t}^{NS}) \geq \bar{\eta}_t, \quad \forall t \in T; \quad (3)$$

$$\sum_{i \in I} \rho_{i,t}^{SP} \geq \eta_t^{SP} \cdot \bar{\eta}_t, \quad \forall t \in T; \quad (4)$$

$$\bar{\eta}_t = 0.03 \cdot l_t + 0.05 \cdot \sum_{w \in W} \omega_{w,t}, \quad \forall t \in T; \quad (5)$$

$$K_i^- u_{i,t} \leq q_{i,t}, \quad \forall i \in I, t \in T; \quad (6)$$

$$q_{i,t} + \rho_{i,t}^{SP} \leq K_i^+ u_{i,t}, \quad \forall i \in I, t \in T; \quad (7)$$

$$q_{i,t} + \rho_{i,t}^{SP} + \rho_{i,t}^{NS} \leq K_i^+, \quad \forall i \in I, t \in T; \quad (8)$$

$$0 \leq \rho_{i,t}^{SP} \leq \bar{\rho}_i^{SP} u_{i,t}, \quad \forall i \in I, t \in T; \quad (9)$$

$$0 \leq \rho_{i,t}^{NS} \leq \bar{\rho}_i^{NS}, \quad \forall i \in I, t \in T; \quad (10)$$

$$R_i^- \leq q_{i,t} - q_{i,t-1}, \quad \forall i \in I, t \in T; \quad (11)$$

$$q_{i,t} - q_{i,t-1} + \rho_{i,t}^{SP} + \rho_{i,t}^{NS} \leq R_i^+ \quad (12)$$

$$\forall i \in I, t \in T;$$

$$\sum_{y=t-\tau_i^+}^t s_{i,y} \leq u_{i,t}, \quad \forall i \in I, t \in T; \quad (13)$$

$$\sum_{y=t-\tau_i^-}^t h_{i,y} \leq 1 - u_{i,t}, \quad \forall i \in I, t \in T; \quad (14)$$

$$s_{i,t} - h_{i,t} = u_{i,t} - u_{i,t-1}, \quad \forall i \in I, t \in T; \quad (15)$$

$$0 \leq \omega_{w,t} \leq \bar{\omega}_{w,t}, \quad \forall w \in W, t \in T; \quad (16)$$

$$l_t \geq 0, \quad \forall t \in T; \quad (17)$$

$$u_{i,t}, s_{i,t}, h_{i,t} \in \{0, 1\}, \quad \forall i \in I, t \in T. \quad (18)$$

Objective function (1) maximizes social welfare, which is defined as the difference between the integral (up to the amount of load served, l_t) of the inverse demand function and total generation costs. In cases without RTP, the l_t 's are held fixed meaning that the integral terms are fixed and welfare maximization is equivalent to cost minimization. In cases with RTP the inverse demand function is represented as a non-increasing step function, implying that the integral terms are concave piecewise-linear in the l_t 's. The variable generation costs, $c_i^V(q_{i,t})$, are modeled as convex piecewise-linear functions. These assumptions yield an objective function that is linear in the decision variables.

Load-balance constraints (2) require demand in each period to be exactly served by conventional and wind generation. Constraint sets (3) and (4) ensure that enough non-spinning and spinning reserves are available and constraint set (5) defines the reserve requirement. Our model only considers spinning and non-spinning reserves, ignoring frequency regulation. This is because the 15-minute timesteps assumed in our analysis do not capture the temporal resolution implicit in the deployment of regulation. We use the so-called '3+5' rule assumed in the National Renewable Energy Laboratory's Western Wind and Solar Integration Study [13]. This is a heuristic rule, which sets total reserve requirements in each period equal to 3% of the load plus 5% of scheduled wind generation. The 5% part of the rule is designed to schedule reserves in proportion to the amount of wind on which the

system relies, due to wind's inherent variability. We further assume that half of these total reserves must be spinning, meaning that $\eta_t^{SP} = 0.5$.

We only explicitly model upward reserves (*i.e.*, excess capacity to deal with a generation shortfall, for instance due to overestimating wind availability). This is because overestimation of wind is typically a greater threat to system stability than underestimation. Overestimating wind availability requires the output of conventional generators to increase or loads to decrease to balance supply and demand. Underestimated wind can typically be accommodated by curtailing the output of wind generators, although to the extent that it is technically feasible, conventional generator output can also be decreased to accommodate unanticipated wind. In heavily thermal systems, not providing downward reserves may create difficulties to offset forecast errors, leading to potential overproduction. Although we do not explicitly model downward reserves, the real-time dispatch model allows for the output of conventional generators to be reduced (relative to the day-ahead commitment solution), to accommodate underestimated wind. Ruiz *et al.* [14] and Papavasiliou *et al.* [15] discuss the advantages and disadvantages of using operating reserves and such heuristic rules, as opposed to more sophisticated stochastic optimization models, to accommodate wind uncertainty. One advantage of stochastic optimization is that reserves can be determined dynamically, based on the probability distribution of wind availability. Another is that the mix of generators committed may be more flexible, allowing for unanticipated wind to be accommodated by reducing conventional generator output, as opposed to curtailing wind. Our approach to modeling reserves is similar to the heuristic rule that Papavasiliou *et al.* [15] study.

Constraint sets (6) through (8) ensure that each conventional generator operates between its minimum and maximum generation points, and that it does not violate its upper-bound if it must provide additional energy due to reserves being called in real-time. Constraint sets (9) and (10) bound ancillary services provided by each generator based on its rated capability. Constraint sets (11) and (12) enforce each generator's ramping limits. Constraint set (12) further ensures that each generator can feasibly provide reserves without violating its ramping restriction. Constraint sets (13) and (14) impose each generator's minimum up- and down-times when they are started up and shutdown. Constraint set (15) defines the startup and shutdown variables in terms of changes in the online state variables. Constraint set (16) limits each wind generator's production based on forecasted wind availability. Since actual wind used can be less than wind available, this constraint allows for wind curtailment. Constraint sets (17) and (18) impose non-negativity and integrality restrictions.

Our model treats demand response as a dispatchable resource that the system operator (SO) can use to balance load. The SO determines the amount of load, l_t , to serve in each period, based on the economic tradeoff between the value of energy consumption, which is given by the inverse demand function, and the cost of generation. This implicitly assumes that consumers truthfully reveal their willingness to pay for energy and that they adjust their demand based on the socially

optimal value of l_t determined by the SO. Thus, we do not tackle the issue of generating market-clearing prices that ensure that suppliers and consumers have proper incentives to provide the socially optimal amount of generation and demand response. This is a theoretically difficult task, due to the non-convex nature of unit commitment [16].

B. Real-Time Dispatch Model Formulation

The real-time dispatch model has the same structure as the day-ahead unit commitment, consisting of objective function (1) and constraint sets (2) through (18). The commitments of the generators are fixed based on the solution of the day-ahead model, with the exception of fast-start generators and generators that are off-line but providing non-spinning reserves. These generators can be started up, if needed, in real-time. Moreover, the values of $\bar{\omega}_{w,t}$ in constraint set (16) are updated to reflect new wind availability forecasts being available. Specifically, when making time- t dispatch decisions, actual time- t wind is known and wind availability in subsequent hours is modeled using a point forecast. The accuracy of the forecast of time- s wind is decreasing in $s-t$, which is indicative of forecasts further in the future being less accurate.

Although the model optimizes dispatch decisions in all periods from time t forward, only the time- t dispatch is fixed based on this model. We hereafter refer to the model used to determine the time- t dispatch as the ‘time- t dispatch model.’ After solving the time- t dispatch model, we roll forward and solve the time- $(t+1)$ dispatch model, with updated wind availability data, to determine the time- $(t+1)$ dispatch. This is repeated 96 times (corresponding to each 15-minute time period) for each day.

C. Wind Modeling

Actual wind generation available in each time period is modeled as:

$$\bar{\omega}_{w,t} = \Omega_w \cdot \phi_{w,t}, \quad (19)$$

where Ω_w is the assumed nameplate capacity of wind plant w and $\phi_{w,t} \in [0, 1]$ is the fraction of this capacity available at time t . The wind forecasts used in the day-ahead unit commitment and real-time dispatch models are generated by including a wind forecast error. Thus, the right-hand side of constraint set (16) for time periods for which forecasts are used becomes:

$$\Omega_w \cdot (\phi_{w,t} + \epsilon_t). \quad (20)$$

III. CASE STUDY AND DATA

A. Data

We examine system operations and costs over a one-year period using the case study analyzed by Sioshansi [6]. This is based on the ERCOT system, using load, conventional generator, and weather data from the year 2005. It also includes 15 GW of wind, which corresponds to all of the plants that were planned to be installed by the end of 2011. We compare cases in which loads are fixed to cases with price-responsive loads and consider loads responding to prices immediately or with a 15- or 30-minute lag.

Conventional generation costs are modeled using heat rates and historical fuel and SO_2 permit prices, which are obtained from Platts Energy and Global Energy Decisions. Conventional generator constraint data are obtained from Global Energy Decisions. Actual wind availability is modeled using the Western Wind Resources Dataset (WWRD) for the year 2005 [17]. The WWRD contains modeled historical wind generation data at 10-minute intervals for a number of sites across the western United States and is generated by 3TIER as part of the Western Wind and Solar Integration Study [13]. We associate the modeled wind plants to locations in the WWRD, based on geographic distance. The value of $\phi_{w,t}$ used in (19) is given by the associated data in the WWRD. We use linear interpolation on the 10-minute WWRD data to arrive at 15-minute wind generation data.

Based on the results of the California ISO’s renewable integration study [18], we assume that the forecast errors in (20) have a first-order autocorrelated structure of the form:

$$\epsilon_t = \mu \cdot \epsilon_{t-1} + \nu_t. \quad (21)$$

We fix the value of μ so the autocorrelation between ϵ_t and ϵ_{t-1} is 0.60. The innovations, ν_t , are assumed to have a truncated Gaussian distribution with mean zero. The innovations have a standard deviation of 0.15 in the day-ahead unit commitment. Standard deviations of the innovations in the real-time dispatch model range between 0.05 and 0.15, depending on the amount of time into the future wind availability is being forecasted. Specifically, when modeling the time- t dispatch, we assume a standard deviation of 0.05 for time $t+1$ and that the standard deviations of the wind forecast errors increase by 0.02 for every additional 15-minutes into the future, up to a maximum of 0.15.

Wind forecast errors exhibit serial and spatial correlations. Serial correlation means that the time- $(t+1)$ error is statistically dependent on recent (*e.g.*, time- t) forecast errors. Put another way, if wind availability is under- or over-forecast at time t , it is also likely to be at time $(t+1)$. Spatial correlation means that contemporaneous errors at two locations are correlated. Spatial correlation is typically inversely proportional to the geographic distance between locations.

Our use of an autocorrelated error structure explicitly captures serial correlation. We implicitly assume perfect spatial correlation of 1, since the same ϵ_t value is used for each location modeled. This assumption overestimates wind-uncertainty costs for a given standard deviation of the ν_t ’s, since spatial correlations are typically lower, especially over a large geographic area such as the ERCOT system. Explicitly modeling spatial correlation is difficult in our case, as we do not have access to historical wind forecast data with which to estimate a spatial correlation function. Alternatively, one can capture the effect of spatial correlation by adjusting the standard deviation of the ν_t ’s. This is because lower spatial correlation reduces the standard deviation of the error in forecasting total system-wide wind availability (to intuitively see this, lower spatial correlation results in errors at individual wind plants that tend to cancel each other out). Sioshansi [6] and Madaeni and Sioshansi [12] show that RTP has very similar effects in reducing wind-uncertainty costs with different forecast error

standard deviations, thus RTP should have similar effects if spatial correlation is explicitly modeled. On the other hand, better capturing spatial correlation should lead to a lag in the demand response having less of an effect on RTP's benefit in reducing wind-uncertainty costs. This is because discrepancies between time- t and $-(t+1)$ wind forecasts should be smaller, meaning that demands based on outdated price data provide greater benefits. Overall, our wind forecast model closely resembles that which Makarov *et al.* [19] use.

Modeled loads and demand functions are based on historical 15-minute load data from the year 2005, obtained from the Public Utility Commission of Texas (PUCT). In cases without RTP, the l_t variables in the day-ahead and real-time models are fixed based on these historical data. Thus the integral terms in objective function (1) are fixed and welfare-maximization is equivalent to cost-minimization. In cases with RTP, we use an assumed elasticity and calibrate the inverse demand function in each period so the actual historical load in the period corresponds to the historical average retail price of electricity in 2005 [6], [12], [20]–[22]. Thus the time- t demand function has the property that:

$$p_t(D_t) = p^{ret}, \quad (22)$$

where D_t is the actual historical time- t demand and p^{ret} is the average retail price of electricity in 2005. In calibrating the demand function we only model own-price elasticities, assuming cross-price elasticities to be zero. This assumption typically understates the extent to which loads shift from on- to off-peak periods with RTP, since cross-price elasticities between periods can be negative [6], [21], [23]. Such load shifting is captured to some extent by our model, however. This is because on-peak loads tend to decrease due to high real-time prices while relatively low prices have the opposite effect during off-peak periods. We estimate p^{ret} using tariff filings with the PUCT. Because these retail prices include non energy-related charges (*e.g.* charges for distribution and retail services), we subtract these from the tariff to yield a price per MWh of energy consumed. We assume an own-price elasticity of -0.2 , which is consistent with empirical estimates of the short-term price-responsiveness of electricity demand [24]. We approximate each demand function as a non-increasing step function with 100 segments. This assumption means that the integrals in objective function (1) are concave piecewise-linear in l_t 's.

B. Cases Modeled

We estimate wind-uncertainty costs by modeling six different cases, which are summarized in Table I. These cases differ in terms of how future wind availability is forecast and the extent to which loads respond to price signals. The cases with perfect foresight assume that real-time wind availability is known day-ahead and represent a best-case scenario in which there is no wind uncertainty or related cost. These cases are modeled by solving the day-ahead unit commitment model only. The real-time dispatch model is not needed, since the day-ahead model determines a feasible commitment and dispatch schedule against actual wind availability. The day-ahead model is still optimized in a rolling fashion to determine

operations for a single day using a 36-hour planning horizon. The cases with imperfect wind forecasts are optimized as outlined in Section II, with the day-ahead and real-time models used in a rolling fashion.

TABLE I
CASES MODELED

Case	Wind Availability Modeling	Loads
1	Perfect Foresight	Fixed
2	Perfect Foresight	RTP With Instant Response
3	Imperfect Forecasts	Fixed
4	Imperfect Forecasts	RTP With Instant Response
5	Imperfect Forecasts	RTP With 15-Minute Lag
6	Imperfect Forecasts	RTP With 30-Minute Lag

In all of the cases with RTP, we assume loads can fully react to price signals in the day-ahead unit commitment. This assumption implies that the SO is willing to make day-ahead commitment decisions in anticipation of real-time demand response. Moreover, in cases 2 and 4, in which there is an instant price response, the time- t dispatch model allows the current and all future loads, l_s with $s \geq t$, to be adjusted. Conversely, in case 5, which assumes a 15-minute lag in the load response, the value of l_t in the time- t dispatch model is fixed to the value determined by the time- $(t-1)$ dispatch model. The model allows loads in subsequent periods, *i.e.*, l_s with $s \geq t+1$, to change in the time- t model, however. This captures the lag in the load response, since time- t demand reacts to the wind forecasts in and prices generated by the time- $(t-1)$ dispatch, but does not react to updated wind availability data in and prices generated by the time- t dispatch. Fig. 1 is a schematic of our rolling model structure when modeling a 15-minute lag in the demand response. We use the notational convention that a model ‘determines a variable’ if the value of that variable is fixed based on solving that model. For instance, the time-1 dispatch model treats $q_{i,t}$ and $w_{\omega,t}$ for $t \geq 1$ and l_t for $t \geq 2$ as variables that can be adjusted. However, only the values of $q_{i,1}$, $w_{\omega,1}$, and l_2 are fixed based on the solution of this model. Thus, we say the time-1 dispatch model determines the values of $q_{i,1}$, $w_{\omega,1}$, and l_2 only.

Case 6, which assumes a 30-minute lag in the load response, is modeled analogously. The only difference is that the values of l_t and l_{t+1} are fixed in the time- t dispatch model, based on the values found in the time- $(t-1)$ dispatch model. As with the 15-minute lag, we allow subsequent loads, *i.e.*, l_s with $s \geq t+2$ to be adjusted in the time- t model. This method of modeling the lag in the demand response assumes that consumers adjust their consumption perfectly to historical prices (*i.e.*, prices generated either 15 or 30 minutes prior). We assume that when the SO sets prices in real-time, these are based on its best estimate of future wind conditions. Thus, we preclude the possibility that an SO may want to set prices higher than the efficient level to reduce demand and the probability that it has insufficient capacity to serve the load in real-time.

C. Measuring Wind-Uncertainty Costs

Wind-uncertainty costs are typically measured as the change in operating costs between a case in which generation is sched-

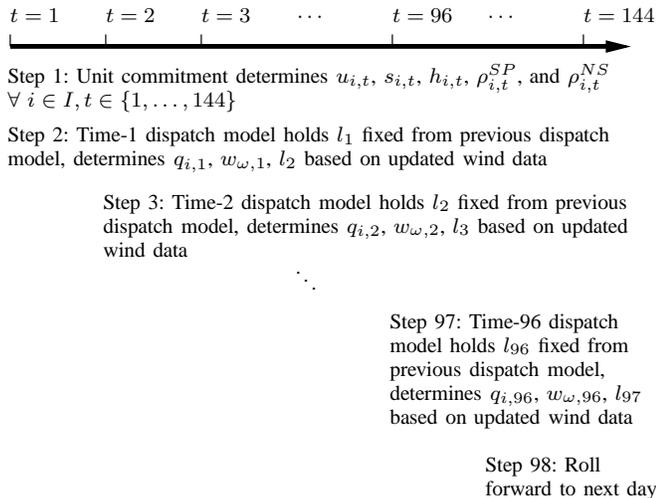


Fig. 1. Schematic of rolling solution technique used with 15-minute demand response lag.

uled with perfect foresight of wind to one in which forecasts are used instead [1]–[3], [25]. We follow this approach, except that social welfare changes are a more appropriate metric. This is because RTP causes demands to change, affecting consumer welfare. These consumer welfare changes, which impose real social costs and benefits, are not captured by generation costs alone. Changes in total annual social welfare, σ_α , between the cases listed in Table I better measure wind-uncertainty costs. We divide these welfare changes by the total amount of wind energy used, v_α , to arrive at a wind-uncertainty cost per MWh of wind energy used.

An issue in comparing welfare between cases with and without RTP is that demand response increases social welfare and system efficiency, independent of its mitigation of wind-uncertainty costs. This is because consumers making consumption decisions on the basis of the real-time marginal cost of energy yields allocative efficiency gains. The social welfare gain between cases 2 and 1, $\sigma_2 - \sigma_1$, measures these allocative efficiency gains, which are independent of the interactions between RTP and wind uncertainty. These welfare gains do include better accommodation of wind variability, however. This is because real-time prices are suppressed, increasing consumption during periods with high wind output, and *vice versa*. Welfare differences between case 2 and cases with wind uncertainty and RTP (cases 4 through 6) measure wind-uncertainty costs with RTP. Comparing these costs to welfare differences between cases 1 and 3, which measure wind-uncertainty costs with fixed loads, measure the benefits of RTP in reducing wind-uncertainty costs.

IV. RESULTS

Table II summarizes annual wind-uncertainty costs per MWh of wind energy used. For each case we compute a net wind-uncertainty cost in \$/MWh of wind and the wind-uncertainty costs averted by RTP (relative to the fixed-load case) in \$/MWh of wind and as a percentage of wind-uncertainty costs in the fixed-load case. The net wind-uncertainty cost in the four cases are computed as $(\sigma_1 -$

$\sigma_3)/v_3$, $(\sigma_2 - \sigma_4)/v_4$, $(\sigma_2 - \sigma_5)/v_5$, and $(\sigma_2 - \sigma_6)/v_6$, respectively. The averted wind-uncertainty costs in each of the three cases with RTP are computed as the difference between wind-uncertainty costs in the fixed-load and each of the RTP cases, or as:

$$\frac{(\sigma_1 - \sigma_3) - (\sigma_2 - \sigma_\alpha)}{v_\alpha}, \quad (23)$$

where α denotes the RTP case considered. The final column of the table gives the value in the third column as a percentage of the wind-uncertainty cost in the fixed-load case. The table shows that wind can impose significant external costs on the system, which is keeping with other wind integration analyses [1]–[3], [6], [25].

TABLE II
WIND-UNCERTAINTY COSTS

Load	Net Cost		Averted Cost	
	[\$/MWh of Wind]	[\$/MWh of Wind]	[%]	
Fixed	6.07			
RTP Instant	0.70	4.60	76	
RTP 15-Min	2.54	2.87	47	
RTP 30-Min	4.17	1.30	21	

Table II further shows that while RTP can mitigate wind-uncertainty costs, this benefit is sensitive to the latency of the demand response. With a 15-minute lag nearly 38% of the value of RTP in mitigating wind-uncertainty costs is lost. This number increases to 72% with a 30-minute lag. Table III summarizes annual generation costs and the breakdown of the load between conventional and wind generation with RTP and imperfect wind forecasts. It shows that increasing the latency of the demand response affects the system in two ways. One is that it reduces aggregate energy consumption, which decreases consumer welfare. The second is that it decreases wind generation, giving higher costs since more conventional generation is used. Both of these effects are due to interactions between the lagged demand response and imperfect wind forecasts. When real-time wind availability is underestimated, consumers have limited ability to adjust their consumption and use unanticipated wind. Although the real-time dispatch model allows conventional generator dispatch to be reduced (relative to the day-ahead solution) to accommodate more wind, technical constraints limit the extent to which this can be done [21], [26]. This inability to accommodate all unanticipated wind is exacerbated by our not explicitly modeling downward reserves in the unit commitment model. Thus, our model assumes that the SO is agnostic to the use of wind (other than wind having zero cost in the objective function). Some SOs prioritize wind, for instance through feed-in tariffs, and may explicitly commit more flexible units to minimize wind curtailment. Conversely, when wind availability is overestimated, the lagged demand response requires greater use of high-cost conventional generation to cover the generation shortfall.

These interactions between wind and lagged demand response are further investigated by examining load and generation patterns during select periods. Fig. 2 shows the aggregate system load during the evening of 2 October with an immediate demand response and if the demand response has a 15-minute lag. The loads in the lagged-response case are

TABLE III
TOTAL ANNUAL GENERATION COST AND LOAD SERVED WITH RTP AND IMPERFECT WIND FORECASTS

Case	Cost [\$ Billion]	Generation [TWh]		
		Total	Conventional	Wind
Instant Response	11.6	304.3	254.0	50.3
15-Minute Lag	11.9	303.7	254.4	49.3
30-Minute Lag	12.0	303.7	254.9	48.8

greater than those in the immediate-response case. This is due to wind availability being systematically overestimated during this evening, and the limited ability of consumers to respond to real-time wind availability in the lagged case. Although system demand is higher with the lagged response, wind use in both cases is equivalent—all of the 17 GWh of wind available during this three-hour period is used in both cases. Thus, all of the additional load in the lagged-response case is served using conventional generation.

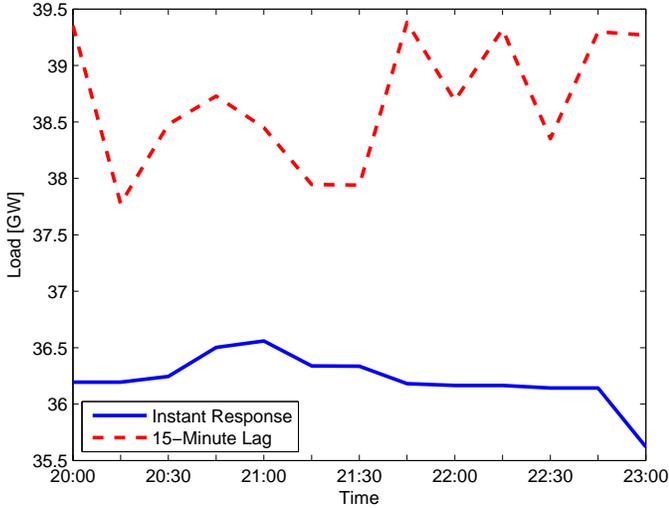


Fig. 2. Aggregate system load from 20:00 to 23:00 on 2 October with immediate and lagged demand response.

Fig. 3 shows the breakdown of conventional generation in the immediate- and lagged-response cases between slow and fast generators. We classify units that can be started up in the real-time dispatch model as fast generators and the remaining as slow. The figure shows more conventional generation in the lagged-response case, since loads cannot respond to real-time shortfalls in wind generation below the amount forecasted, as illustrated in Fig. 2. It also shows greater use of fast units, which typically have higher operating costs, to serve the wind shortfall in the lagged-response case. This contributes to the greater generation costs compared to having an immediate demand response.

As a second example, Fig. 4 shows the system load midday on 1 January in cases in which the load responds to prices immediately and with a 15-minute lag. Unlike the load profiles shown in Fig. 2, the lagged-response case results in lower system loads for most of the periods shown. This is due to wind availability being underestimated, as shown in Fig. 5, and the limited ability of consumers to use this excess wind in real-time. Although the SO ideally wants to maximize

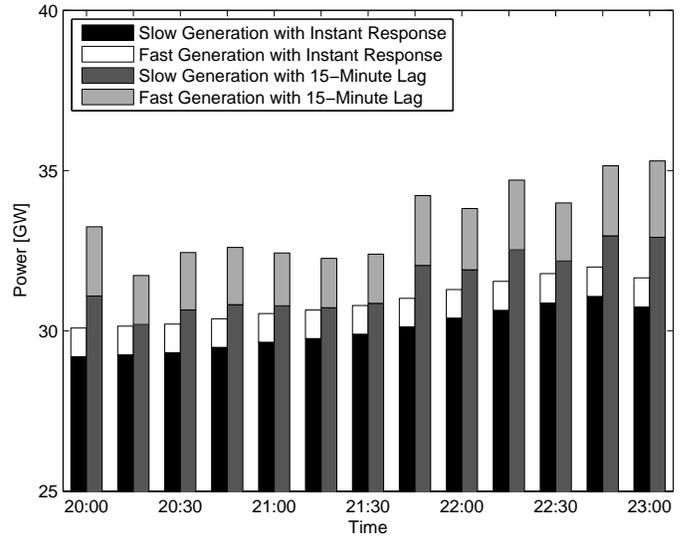


Fig. 3. Output of slow and fast generators from 20:00 to 23:00 on 2 October with immediate and lagged demand response.

wind use (since it has zero marginal cost), Fig. 5 shows that about 1 MWh and 3.4 GWh of curtailments occur in the immediate- and lagged-response cases, respectively. These curtailments are caused by conventional generator ramping and minimum load and up-time constraints, and are exacerbated by the high cost of cycling conventional generators on and off. For instance, conventional generation cannot drop below about 14 GW during the periods shown in Fig. 4 which have wind curtailment. The day-ahead wind forecast shown in Fig. 5 also illustrates the autocorrelated nature of the wind forecast errors modeled. Specifically, the overly conservative wind forecast for 11:15 am persists in the subsequent hours.

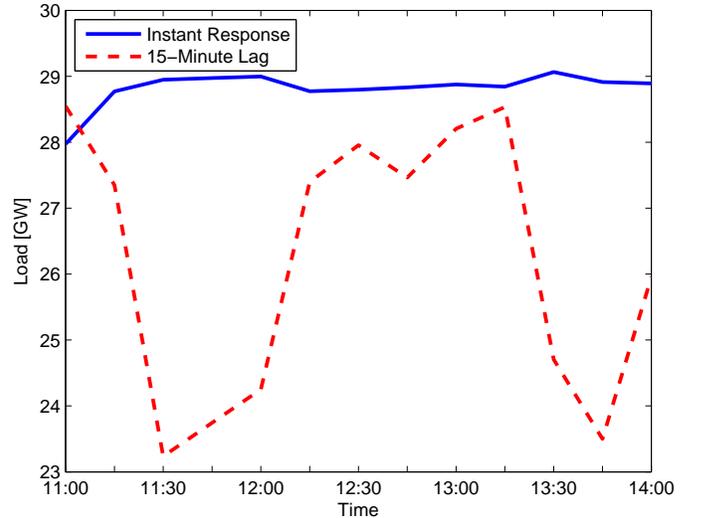


Fig. 4. Aggregate system load from 11:00 to 14:00 on 1 January with immediate and lagged demand response.

V. CONCLUSIONS

This study analyzes the value of demand response in reducing wind-uncertainty costs if there is a latency in the response.

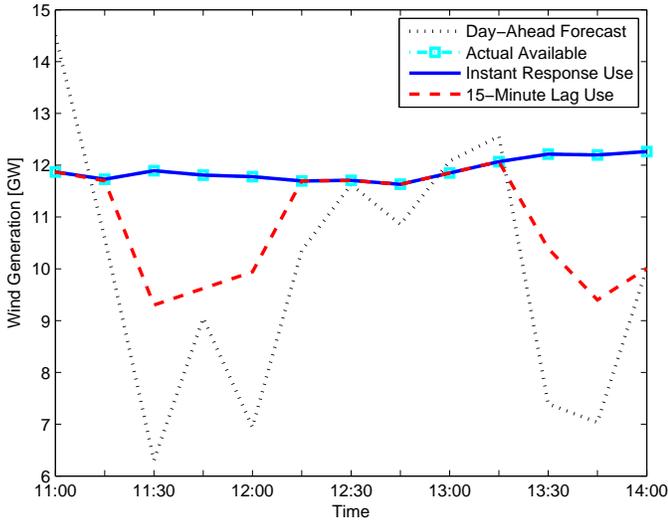


Fig. 5. Wind available and used from 11:00 to 14:00 on 1 January with immediate and lagged demand response.

We demonstrate that demand response has significant value in reducing wind-uncertainty costs, although much of this value is contingent upon loads responding to wind availability and system conditions immediately. More than 75% of wind-uncertainty costs are mitigated if loads respond to prices immediately. Less than a quarter of wind-uncertainty costs are averted if there is a 30-minute lag in the response. Even with such latency, however, demand response is a valuable tool in mitigating many adverse effects of wind. This is because loads can respond to more accurate wind forecasts that can become available intraday. Wind and demand response have other synergies shown in this study and in other analyses. This includes reducing wind curtailment, due to transmission, generator, or system operating constraints [21] and delivering greater emissions benefits [12]. Demand response may also help system planners with long-run capacity planning. An issue raised by increasing the penetration of wind is that it may leave the system with insufficient *installed* capacity to serve the load, if real-time wind output and installed conventional generation capacity are both sufficiently low. Demand response programs can encourage consumers to efficiently ration their consumption and shift it away from such periods, improving asset use and system reliability.

These benefits presuppose that the SO dynamically reoptimizes generator dispatch and prices in the rolling fashion that we assume. Otherwise, the value of demand response may be significantly less than our estimates suggest if prices are only set day- and hour-ahead, as is current practice in some markets. Similarly, the SO must update wind forecasts intraday to fully exploit the value of demand response. Conversely, our model assumes that generator commitments are not dynamically adjusted in real-time, and that only fast-start units and those offering non-spinning reserves can be started up in the real-time dispatch process. This assumption resembles the approach taken by Tuohy *et al.* [11] and Sioshansi [6]. If other generators have sufficiently short start-up times that they could be committed intraday, wind-uncertainty costs may

be lower than our model indicates. Our analysis also assumes that the SO adjusts its day-ahead commitment in anticipation of real-time demand response (since we model the l_t variables as being fully flexible in the day-ahead model). If the SO is not willing to make commitment decisions in such a fashion, the value of RTP may be significantly reduced. The method we use to calibrate the inverse demand functions assumes that all of the load is price-elastic. This is because the literature estimating demand elasticities computes them relative to the total load of the study participants. Our model can be easily adapted to only consider a portion of the demand as being price responsive. One approach is to adjust the demand elasticity itself—for instance, if only 50% of the load is price-responsive and has an elasticity of 0.2, then the elasticity of the overall total load is 0.1. Alternatively, one can define a fixed and a price-responsive load, as Sioshansi and Short [21] do in their analysis of RTP.

Another assumption of our model is that loads respond to price signals in a symmetric manner. The value of RTP is that loads respond to wind resource availability. In practice, demand may not respond symmetrically, since customers may respond more to price increases than to decreases. For instance, a consumer may turn off an appliance when electricity prices are high, but may not necessarily turn one when it is not needed simply because the price of electricity is low. While this type of asymmetric demand response may reduce some of the surplus gains from providing consumers with additional energy when actual wind availability is greater than forecast, much of the benefits of RTP stem from demand reductions when wind forecasts are too high. Thus, most of the benefits estimated here would be captured even with asymmetric demand response.

Although we focus on an indirect RTP mechanism, the results are likely applicable to other forms of demand response, such as direct load control. Such schemes may be preferable to price-based mechanisms, since there may be less latency and more predictability in the demand response. Our analysis is based on the ERCOT system and further analysis is needed to determine the synergies between demand response and wind in other systems. Much of the value of demand response stems from the generation mix in ERCOT. Since ERCOT has a mix of low-cost slow generators and higher-cost fast generators, demand response is particularly valuable since it reduces the use of the flexible generators. Demand response may have less value in mitigating wind-uncertainty costs in a hydroelectric-dominated system, such as Scandinavia, due to the abundance of low-cost flexible generation. Similarly, the recent decrease in the cost of natural gas in the United States can reduce the value of demand response. This is because much of the flexible generation in ERCOT is natural gas-fired, and lower gas prices reduce the cost of these resources relative to coal-fired plants. Another area of research is whether these synergies apply to other variable and uncertain renewables, such as solar. While demand response should have the same benefit of having loads follow solar supply more closely, solar and wind generation patterns are markedly different from one another. Whereas wind generation peaks overnight in some systems, solar peaks midday. Further analysis is needed to fully understand the

interactions between these technologies.

ACKNOWLEDGMENT

The authors thank A. Sorooshian, the editors, and five anonymous reviewers for helpful discussions and suggestions.

REFERENCES

- [1] E. A. DeMeo, W. Grant, M. R. Milligan, and M. J. Schuerger, "Wind plant integration," *IEEE Power and Energy Magazine*, vol. 3, pp. 38–46, November–December 2005.
- [2] J. C. Smith, M. R. Milligan, E. A. DeMeo, and B. Parsons, "Utility wind integration and operating impact state of the art," *IEEE Transactions on Power Systems*, vol. 22, pp. 900–908, August 2007.
- [3] E. A. DeMeo, G. A. Jordan, C. Kalich, J. King, M. R. Milligan, C. Murley, B. Oakleaf, and M. J. Schuerger, "Accommodating wind's natural behavior," *IEEE Power and Energy Magazine*, vol. 5, pp. 59–67, November–December 2007.
- [4] A. Papavasiliou and S. S. Oren, "Coupling Wind Generators with Deferrable Loads," in *Energy 2030 Conference*. Atlanta, GA, USA: Institute of Electrical and Electronics Engineers, 17–18 November 2008.
- [5] M. Klobasa, "Analysis of demand response and wind integration in Germany's electricity market," *IET Renewable Power Generation*, vol. 4, pp. 55–63, January 2010.
- [6] R. Sioshansi, "Evaluating the Impacts of Real-Time Pricing on the Cost and Value of Wind Generation," *IEEE Transactions on Power Systems*, vol. 25, pp. 741–748, April 2010.
- [7] K. Dietrich, J. M. Latorre, L. Olmos, and A. Ramos, "Demand Response in an Isolated System With High Wind Integration," *IEEE Transactions on Power Systems*, vol. 27, pp. 20–29, February 2012.
- [8] C. De Jonghe, B. F. Hobbs, and R. Belmans, "Optimal Generation Mix With Short-Term Demand Response and Wind Penetration," *IEEE Transactions on Power Systems*, vol. 27, pp. 830–839, May 2012.
- [9] F. Bouffard and F. D. Galiana, "Stochastic Security for Operations Planning With Significant Wind Power Generation," *IEEE Transactions on Power Systems*, vol. 23, pp. 306–316, May 2008.
- [10] J. García-González, R. M. R. de la Muela, L. M. Santos, and A. M. González, "Stochastic Joint Optimization of Wind Generation and Pumped-Storage Units in an Electricity Market," *IEEE Transactions on Power Systems*, vol. 23, pp. 460–468, May 2008.
- [11] A. Tuohy, P. Meibom, E. Denny, and M. O'Malley, "Unit Commitment for Systems With Significant Wind Penetration," *IEEE Transactions on Power Systems*, vol. 24, pp. 592–601, May 2009.
- [12] S. H. Madaeni and R. Sioshansi, "The Impacts of Stochastic Programming and Demand Response on Wind Integration," *Energy Systems*, vol. 4, pp. 109–124, June 2013.
- [13] "Western wind and solar integration study," National Renewable Energy Laboratory, Tech. Rep. NREL/SR-550-47434, May 2010.
- [14] P. A. Ruiz, C. R. Philbrick, E. Zak, K. W. Cheung, and P. W. Sauer, "Uncertainty Management in the Unit Commitment Problem," *IEEE Transactions on Power Systems*, vol. 24, pp. 642–651, May 2009.
- [15] A. Papavasiliou, S. S. Oren, and R. P. O'Neill, "Reserve Requirements for Wind Power Integration: A Scenario-Based Stochastic Programming Framework," *IEEE Transactions on Power Systems*, vol. 26, pp. 2197–2206, November 2011.
- [16] R. P. O'Neill, P. M. Sotkiewicz, B. F. Hobbs, M. H. Rothkopf, and W. R. Stewart, "Efficient market-clearing prices in markets with nonconvexities," *European Journal of Operational Research*, vol. 164, pp. 269–285, 1 July 2005.
- [17] C. W. Potter, D. Lew, J. McCaa, S. Cheng, S. Eichelberger, and E. Grimit, "Creating the Dataset for the Western Wind and Solar Integration Study (U.S.A.)," *Wind Engineering*, vol. 32, pp. 325–338, June 2008.
- [18] C. Loutan and D. Hawkins, "Integration of renewable resources," California Independent System Operator, Tech. Rep., November 2007.
- [19] Y. V. Makarov, P. V. Etingov, J. Ma, Z. Huang, and K. Subbarao, "Incorporating Uncertainty of Wind Power Generation Forecast Into Power System Operation, Dispatch, and Unit Commitment Procedures," *IEEE Transactions on Sustainable Energy*, vol. 2, pp. 433–442, October 2011.
- [20] S. Borenstein, J. B. Bushnell, and C. R. Knittel, "A Cournot-Nash Equilibrium Analysis of the New Jersey Electricity Market," New Jersey Board of Public Utilities, Appendix A of Review of General Public Utilities Restructuring Petition, Final Report, 1997, docket Number EA97060396.
- [21] R. Sioshansi and W. Short, "Evaluating the impacts of real-time pricing on the usage of wind generation," *IEEE Transactions on Power Systems*, vol. 24, pp. 516–524, May 2009.
- [22] S. H. Madaeni and R. Sioshansi, "Using Demand Response to Improve the Emission Benefits of Wind," *IEEE Transactions on Power Systems*, vol. 28, pp. 1385–1394, May 2013.
- [23] S. Borenstein and S. P. Holland, "On the Efficiency of Competitive Electricity Markets With Time-Invariant Retail Prices," *The RAND Journal of Economics*, vol. 36, pp. 469–493, Autumn 2005.
- [24] C. S. King and S. Chatterjee, "Predicting California Demand Response," *Public Utilities Fortnightly*, pp. 27–32, July 2003.
- [25] A. Fabbri, T. G. S. Román, J. R. Abbad, and V. H. M. Quezada, "Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market," *IEEE Transactions on Power Systems*, vol. 20, pp. 1440–1446, August 2005.
- [26] P. Denholm and R. M. Margolis, "Evaluating the limit of solar photovoltaics (PV) in traditional electric power systems," *Energy Policy*, vol. 35, pp. 2852–2861, 2007.



Seyed Hossein Madaeni holds the B.S. and M.S. degrees in electrical engineering from the University of Tehran and the Ph.D. in industrial and systems engineering from The Ohio State University.

He is a Senior Energy Supply Analyst in the Short Term Electric Supply Department at Pacific Gas and Electric Company, San Francisco, CA. His research focuses on electricity market design and mechanisms, renewable energy analysis, and power system operations.



Ramteen Sioshansi (M'11–SM'12) holds the B.A. degree in economics and applied mathematics and the M.S. and Ph.D. degrees in industrial engineering and operations research from the University of California, Berkeley, and an M.Sc. in econometrics and mathematical economics from The London School of Economics and Political Science.

He is an assistant professor in the Integrated Systems Engineering Department at The Ohio State University, Columbus, OH. His research focuses on renewable and sustainable energy system analysis and the design of restructured competitive electricity markets.