Barriers to Increasing the Role of Demand Resources in Electricity Markets

Alberto J. Lamadrid
Department of Economics
Lehigh University
Bethlehem, PA, 18015
Email: ajlamadrid@lehigh.edu

Tim Mount, Wooyoung Jeon and Hao Lu
Dyson School of Applied Economics and Management
Cornell University
Ithaca, NY, 14853
Email: tdm2@cornell.edu

Abstract—The objective of this paper is to show that customers can benefit from a smart grid if they become more active participants in electricity markets by 1) relying more on deferrable demand (e.g. electric vehicles and augmenting space conditioning with thermal demand) to shift demand away from peak periods and buy more electricity when prices are low at night, and 2) selling ancillary services such as ramping capacity to mitigate the inherent uncertainty of wind generation. These two factors, coupled with the lower operating cost of wind generation compared to conventional generation from fossil fuels, have the potential for reducing the cost of electricity to customers. However, these benefits will not be realized unless the rates charged to customers reflect the true costs of supply. This paper compares how the bills charged to different types of customer are affected by different rate structures with and without the correct economic incentives. The main savings in operating cost come from the displacement of conventional generation by wind generation, and the main savings in capital cost come from reducing the amount of installed conventional generating capacity needed to maintain System Adequacy by 1) reducing the peak system load, and 2) by using deferrable demand to provide ramping services and reduce the amount of conventional generating capacity needed for operating reserves.

A new stochastic form of multi-period Security Constrained Optimal Power Flow is applied in a simulation using a reduction of the North Eastern Power Coordinating Council (NPCC) network for a representative summer day. This model treats potential wind generation as a stochastic input and determines the amount of conventional generating capacity needed to maintain reliability endogenously. The analysis assumes implicitly that all deferrable demand at a node is managed by an aggregator. If the rates are structured with the correct economic incentives (i.e. real-time nodal prices for energy, a demand charge determined by the demand during system peak periods, and compensation for providing ramping services), the results show that 1) the economic benefits for customers with thermal storage are substantial, and 2) the main benefits for customers with electric vehicles (without V2G capabilities in this application) come from buying less gasoline. In contrast, if customers pay conventional rates with a fixed price for energy and no demand charge, the economic incentives are perverse and customers with deferrable demand pay more and customers with no deferrable demand pay less.

I. INTRODUCTION

With high penetrations of variable generation from wind turbines, the system benefits of this relatively inexpensive source may be lower than expected because other system costs increase. Even though wind generation will typically displace more expensive generation from fossil fuel sources and reduce operating costs, the costs of the conventional generating units needed as reserve capacity to mitigate the uncertainty of wind generation and maintain reliability are likely to be higher. These costs include the direct operating costs of being available as reserves and the capital cost of building these units. Furthermore, there is growing evidence, particularly from Europe, that there are additional “ramping” costs caused by the higher maintenance costs associated with repeated changes in the dispatch points of the reserve units. The additional system costs associated with wind generation are reflected by the relatively low nodal prices paid to wind farms. When the system load is high, congestion on a network may effectively prevent wind generators from getting paid the high nodal prices in load pockets. In contrast, when the system load is low and there is little congestion, wind generation may be sufficient to meet most of the load throughout the network, and the low offer prices for this source will set the nodal prices.

The primary objective of this paper is to evaluate the role of installing storage capacity as a way to integrate more wind generation into a network and reduce system costs. Three different types of storage are considered. One uses utility-scale batteries collocated at the wind sites to deal with the variability of the potential wind generation. When wind speeds are high, some of the potential wind generation is stored on-site and the batteries are discharged when the wind speeds are lower. In this way, the variability of the generation dispatched at a wind site (direct generation plus discharging the batteries) is much lower than it is with no storage capacity.

The other two types of storage are examples of deferrable demand at load centers that decouple the purchase of electric energy from the delivery of the some energy services that customers want. Charging the battery in an electric vehicle is one example of deferrable demand, but the other example, thermal storage, is likely to have a larger impact on the grid. With the latter type of deferrable demand, ice is made at night, for example, when the price of electricity is low and melted to provide cooling services when cooling services are needed during the day. The overall effect of deferrable demand is to flatten the daily profile of purchases at the load centers from the grid. In this way, the amount of congestion on the grid and the amount of conventional generating capacity...
needed to maintain Operating Reliability are both reduced. If suitable incentive mechanisms are established, it would also be possible to use deferrable demand to mitigate the variability of wind generation in a similar way to using batteries at the wind sites. In fact, this type of mitigation occurs to some extent without explicit incentives for mitigating ramping because the nodal prices at load centers are affected by ramping costs as well as by the standard operating costs.

Traditionally, the system operators who manage operations on the grid have focused on optimizing different sources of supply and have treated the demand by customers as exogenous. The daily pattern of the aggregate demand from customers on a distribution network is predictable, and the operating criterion in a typical Security Constrained Optimal Power Flow (SCOPF) [1] is to minimize the cost of meeting a predicted pattern of demand subject to ensuring that a specified set of equipment failures (contingencies) can be covered. Treating demand as an exogenous input for planning expansion of the grid has resulted in a situation in which the peak system load grew faster than the annual demand for electric energy. Consequently, the average capacity factors of generators have decreased over time and some units are only used for a few hours each year. Supply systems in most regions in the US are designed to meet the summer peak load caused by the demand for air conditioning. There are, however, already signs that conditions are changing. In the 2010 Long-Term Reliability Assessment published by the North American Electric Reliability Corporation (NERC) [2], electric energy is forecasted to grow slightly faster than the peak demand over the next ten years due to the electrification of the transportation sector and increased demand-side management.

The empirical examples in this paper use a new stochastic form of multi-period SCOPF developed at Cornell (the second generation SuperOPF) and a reduced representation of the bulk power network in the Northeast Power Coordinating Council (NPCC) to evaluate the system effects of uncertain wind generation, ramping costs and storage. The most important features of the SuperOPF for this analysis are that 1) the stochastic characteristics of potential wind generation at multiple sites are incorporated, 2) the amount of conventional generating capacity, including reserves, needed to maintain Operating Reliability is determined endogenously, and it depends on how the stochastic characteristics of potential wind generation are represented, and 3) the additional ramping costs caused by the inherent variability of wind generation is incorporated into the objective function. As an example, if the use of on-site storage reduces the variability of wind generation, ramping costs are reduced and less conventional generating capacity is needed for reserves to maintain reliability. Since the capacity of the electric delivery system is designed to meet the peak system load, reducing this peak and the associated capital cost of equipment (e.g. peaking units with low capacity factors) is an important way to reduce the total system costs as well as reduce the amount of congestion on the grid.

This paper has the following structure. Section II briefly summarizes antecedents to this work. Section III presents a general description of the stochastic multi-period SCOPF followed in Section IV by a description of its specific features in the SuperOPF, such as the representation of dedicated utility storage, thermal storage, electric vehicles and the uncertainty of potential wind generation. The empirical results presented in Section V show that storage is an effective way to lower system costs by 1) flattening the daily pattern of dispatch by conventional generating units, 2) spilling less of the potential wind generation, 3) mitigating the ramping costs associated with wind uncertainty, and 4) reducing the amount of reserve capacity needed to maintain reliability. In addition, deferrable demand reduces the peak system load on the network, the corresponding total capacity of supply (conventional generation plus wind generation and discharging utility storage), and congestion on the network. The paper ends with the conclusions in Section VI that include our recommendations for the regulatory changes required to provide the necessary economic incentives for customers to make investments in deferrable demand.

II. LITERATURE REVIEW

The operation of the electricity system is based on a centrally planned solution to a SCOPF [1], often with heuristics-based hard constraints imposed to maintain the security of the system [3]. For longer time horizons in most of the deregulated markets in the US, this approach is coupled to a Unit Commitment (UC) problem that incorporates the startup and shutdown decisions for large thermal generators. There is an extensive body of literature analyzing the theoretical framework for this combined problem [4], [5], [6], [7], with recent developments analyzing the effect of integrating renewable sources of generation [8], [9]. These approaches contrast with market designs that leave the commitment decision to the individual generators [10]. Regardless of the final market design, there is some consensus about the effect of increasing the share of renewable generation on the system operations [11], [12], [13], especially regarding the secure operation of the system. The approach considered in this paper is germane to the approach in [1], with a robust determination of the reserves needed for contingencies, load following and ramping to mitigate the variability of wind generation.

III. FORMULATION OF THE ANALYTICAL MODEL

A new second-generation SCOPF, the SuperOPF, is used for the analysis. This model is an extension of the model proposed in [15] and [14] and was implemented using MATPOWER’s extensible architecture [16].

The objective criterion of the new SuperOPF is to maximize the expected sum of producer and consumer surplus over a twenty-four hour horizon for a set of contingencies, including uncertainty about the forecasts of potential wind generation. It also allows for storage and deferrable demand. Rather than using the standard criterion of minimizing cost subject

\footnote{A stochastic multi-period security constrained AC OPF with co-optimizing endogenous reserves to provide ramping to mitigate wind variability and cover a set of credible contingencies. [14].}
to covering physical contingencies, shedding load at a high Value of Lost Load (VOLL) is allowed if it is economically efficient to do so. This formulation determines the optimal dispatch of a set of previously committed generating units subject to their physical characteristics (e.g., rated capacity, cost and ramping capabilities) and the network’s topology (e.g., transmission line constraints). The model solves the expected cost for a number of high probability cases for stochastic wind generation (“intact” scenarios), as well as for a set of credible contingencies that occur relatively infrequently. The expected cost is determined for the intact scenarios and the contingencies using probabilities that reflect the relative likelihood of the different states of the system occurring. This formulation has the advantage of determining endogenously the amounts of different ancillary services (e.g., contingency reserve and ramping reserve to mitigate wind variability) needed to meet the load profiles and maintain the reliability of the delivery system. The optimum dispatch is determined in the spirit of a day-ahead contract, incorporating the best available information that the System Operator (SO) has at that time. A simplified formulation of the objective function for the problem is shown in (1) and the notation is defined in Table I.

\[
\min_{G_{itk}, R_{itk}, LNS_{itk}} \sum_{t \in T} \sum_{s \in T} \sum_{i \in I} \left\{ \sum_{k \in K} C_{G_{itk}}(G_{itk}) + \text{Inc}_{it}^+(G_{itk} - G_{itk}) + \text{Dec}_{it}^-(G_{itk} - G_{itk}) \right\} + \sum_{t \in T} \text{VOLL}_{lj} \text{LNS}(G_{itk}, R_{itk}) + \sum_{t \in T} \rho_t \sum_{i \in I} \left[ C_{R_{itk}}^+(R_{itk}) + C_{R_{itk}}^-(R_{itk}) + \text{Inc}_{it}^+(L_{itk}) + \text{Dec}_{it}^-(L_{itk}) \right] + \sum_{i \in I} \sum_{t \in T} \sum_{s \in T} \sum_{k \in K} \sum_{l \in L} \left( \text{Ramp}^+_l(G_{itk2} - G_{itk1}) + \text{Ramp}^-_l(G_{itk2} - G_{itk1}) \right) + f_s(p_{out}, p_{un})
\]

Subject to meeting Demand and all of the nonlinear AC constraints of the network.

The nodal levels of demand are fixed blocks for each time period and are modeled as negative injections with associated negative costs (VOLL). Since this specification allows for load shedding in some states of the delivery system, valued at VOLL, minimizing the expected cost, including load shedding as a cost, corresponds to maximizing the expected sum of consumer and producer surplus.

IV. MODEL SPECIFICATION

The calibration of input data was done using publicly available sources, and it encompasses the modification of the test network and the modeling of wind generation, deferrable demand and utility-scale Energy Storage Systems (ESS) collocated at the wind sites.

![Image](image1.png)

**TABLE I**

<table>
<thead>
<tr>
<th>A Simplified Formulation</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{F}$</td>
<td>Set of time periods considered, $n_t$ elements indexed by $t$.</td>
<td></td>
</tr>
<tr>
<td>$\mathcal{X}$</td>
<td>Set of scenarios in the system in period $t$, $n_x$ elements indexed by $t$.</td>
<td></td>
</tr>
<tr>
<td>$\mathcal{C}$</td>
<td>Set of contingencies in the system, $n_c$ elements indexed by $c$.</td>
<td></td>
</tr>
<tr>
<td>$\mathcal{J}$</td>
<td>Set of loads in the system, $n_j$ elements indexed by $j$.</td>
<td></td>
</tr>
<tr>
<td>$\mathcal{F}$</td>
<td>Probability of contingency $f$ occurring, in scenario $s$, period $t$.</td>
<td></td>
</tr>
<tr>
<td>$\text{VOLL}_{lj}$</td>
<td>Cost of providing ($c/s$) MW of load follow up.</td>
<td></td>
</tr>
<tr>
<td>$\text{Ramp}^+_l$</td>
<td>Cost of providing ($c/s$) MW of upward reserves.</td>
<td></td>
</tr>
<tr>
<td>$\text{Ramp}^-_l$</td>
<td>Cost of providing ($c/s$) MW of downward reserves.</td>
<td></td>
</tr>
<tr>
<td>$L_{itk}$</td>
<td>Cost of providing ($c/s$) MW of load follow down.</td>
<td></td>
</tr>
<tr>
<td>$G_{itk}$</td>
<td>Value of decreasing generation from contracted amount.</td>
<td></td>
</tr>
<tr>
<td>$G_{itk}$</td>
<td>Value of Lost Load (VOLL).</td>
<td></td>
</tr>
<tr>
<td>$LNS_{itk}$</td>
<td>Load Not Served (MWh).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Value of the leftover stored energy in terminal states.</td>
<td></td>
</tr>
</tbody>
</table>

A. The Test Network

Figure 1 is a one-line diagram of the network used in the case study. This is a New York and New England centric reduction of the Northeast Power Coordinating Council (NPCC) network [17], that has been modified to include very detailed information of the generating units at each bus obtained from the PowerWorld Corporation.

The total load of the system is around 138 GW, and the generation capacity available is 143 GW [17]. For the simulation, one day in a high demand period was calibrated (following historical load information from August 2008), distinguishing the profiles between urban and rural nodes. The coincident peak system load occurs at 3PM, caused mainly by the high demand at urban nodes. Table II has
TABLE II
SUMMARY OF GENERATION CAPACITY AND LOAD

<table>
<thead>
<tr>
<th>Capacity per Fuel Type (MW)</th>
<th>Total Cap</th>
<th>Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location (RTO)</td>
<td>coal</td>
<td>ng</td>
</tr>
<tr>
<td>ilona</td>
<td>1,840</td>
<td>9,219</td>
</tr>
<tr>
<td>mara</td>
<td>2,424</td>
<td>1,072</td>
</tr>
<tr>
<td>nyiso</td>
<td>4,097</td>
<td>18,185</td>
</tr>
<tr>
<td>ont</td>
<td>5,287</td>
<td>3,944</td>
</tr>
<tr>
<td>pjm</td>
<td>14,035</td>
<td>14,011</td>
</tr>
<tr>
<td>quebec</td>
<td>60 0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>28,562</td>
<td>18,530</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Location (RTO)</th>
<th>Total Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>ilona</td>
<td>143,707</td>
</tr>
<tr>
<td>mara</td>
<td>158,412</td>
</tr>
</tbody>
</table>

b Values shown are taken as peak values.

The cost of ramping services is consistent with our previous work (see [18]), assigned by fuel type using quadratic cost functions. The values set were comparatively high for baseload units and lower for peaking units, signaling the different generators’ willingness to be moved from their current operating point. This modeling approach implicitly presumes homogenous conditions for all generators of a given fuel type.

The loads available in the system were classified as either rural or urban, each one with a different profile over the day, with more pronounced peaks for the urban loads compared to the rural loads. The changes observed over one day were derived from 2008 historical data, to be consistent with available wind data, allowing for different changes in hour to hour demand according to the location of the load. In addition, the Value of Lost Load (VOLL) also depends on location, with a value of $10,000/MWh for urban areas and $5,000/MWh for rural areas.

B. Specifications for Stochastic Wind Generation

This study analyzes a case with a wind penetration close to 20% of the total system load. The setup of the wind specification is divided in two main tasks: specifying the locations and sizes of the wind resources on the network, and characterizing the variability of these wind resources.

The locations of the wind farms are derived from the National Renewable Energy Laboratory (NREL) Eastern Wind and Transmission Study (EWITS) data [12]). To match the data from NREL to the available buses in the NPCC network, a principal component analysis (PCA) was performed, leading to nine sites in New York and seven sites in New England that correspond to specific nodes on the network.2 To characterize the variability of the wind resources in spatial terms, a clustering analysis was implemented using a $k$-means++ methodology for scenario reduction [19].

The location of the wind farms is in the following buses: Orrington, Sandy Pond, Millbury, Northfield, Southington, Millstone, Norwalk harbor, Millwood, Newbridge, 9Mile Point, Leeds, Massena, Gilboa, Marcy, Niagara and Rochester.

determination of the clusters was done using the hourly wind speeds for different locations from the EWITS data. The wind speeds are then converted to the potential wind generation using a multi-turbine modeling approach [20]. The input data used for clustering represent the hourly values at 16 locations for a set of selected days that have similar characteristics in terms of wind speed. These daily profiles are then reduced to four scenarios (hourly profiles) and each day in the sample is assigned to the nearest cluster for each hour. This makes it possible to estimate the hourly probabilities of each scenario occurring and the corresponding transition probabilities of moving from one scenario to another scenario in the next hour. The overall objective is to model the variability of wind realistically in a way that captures geographic averaging and is consistent with the EWITS data from NREL. In this setup, there is a ‘high wind’ scenario, a ‘low wind’ scenario and two intermediate scenarios, with different profiles for each wind site.

C. Specifications for Deferrable Demand

The concept of using deferrable demand for improved system management dates back to work done in the late 80’s [21], [22] that laid the foundation for the work proposed in this paper. The specification of deferrable demand assumes that the timing of the purchase of electricity for specified percentages of the total hourly demand can effectively be decoupled from the timing of the energy services delivered. Examples include charging the batteries in electric vehicles and thermal storage for space conditioning (e.g. traditional central Air Conditioning (AC) systems can be augmented with ice batteries). In effect, there are now two hourly demand profiles. Conventional demand must be supplied in real time from the grid. Although deferrable demand must also be supplied in real time, the sources of supply can come from storage and/or the grid.

The specifications of the battery technology for Electric Vehicles (EVs) follow that of a GM Volt 2013. This type of battery is lithium-ion and the usable energy capacity is 10.8kWh. The total number of EV’s is set at 3,138,525, which is 20% penetration of total number of regular vehicles in NY and NE. This amounts to an aggregated energy capacity of 33.9GWh. The EVs are distributed in five major load centers, proportional to their load size. The average charging efficiency of lithium-ion batteries is 90% ([23], [24]). Two types of charging levels are considered using current technology. Level 1 chargers deliver up to 1.44 kW and level 2 chargers deliver up to 7.68 kW [25]. It is assumed that 70% of level 1 chargers and 30% of level 2 chargers are available in this network, which implies an average of 3.31 kW. The specified average driving distances for “rural,” “suburban,” and “center city” are 36.9 miles, 28.8 miles, and 27.2 miles, respectively [26]. This analysis specifies 27.2 miles, because EVs are located only in major demand centers. The driving pattern of commuters in this case is based on ‘Commuter Driving Profile’ [27]; the percentage of commuters at home, determining how many vehicles are connected to grid and available for charging, is
based on ‘Commuter-at-Home Profile’ [28]. This case assumes that EVs are connected to smart chargers as soon as they arrive at home, and stay connected until they leave for work. We assume that there is no charging station at work, so charging only takes place when EVs are at home. Vehicle to Grid (V2G) is not allowed in this case study, and the driving energy efficiency set at 0.25 kWh/mile.

The specification of thermo storage uses the same aggregated storage size of the EV case, 33.9GWh. This is approximately a 16.0% penetration level of the total potential cooling load based on the estimated temperature-sensitive load (TSL) on the chosen day. The technical characteristics are based on the products described in the reports by Evapco [29] and Calmac [30]. The hourly ice building power rate is 12% and the hourly ice melting power rate is 16.7% of the total storage capacity. These ice building and melting rates can vary by the number of chillers installed with thermal storage. The storage efficiency is 86% which is based on an average energy efficiency ratio (EER) of 8.8 of thermal storage, compared to an EER of 10.2 for average conventional AC. Thermal storages are also distributed in five demand centers. The total amount of deferrable demand (as a percentage of the total demand) set at a node are location specific. Deferrable demand accounts for 17% of the total demand for the New York City buses and for 18% at the Buffalo bus. For the Millbury bus and the Sandy Pond bus in New England, the values were set to 17% and almost 14% of the total demand, respectively. These values correspond to the estimated average values estimated econometrically from historical patterns of demand in the different regions for the years 2007 to 2010 [31].

D. Specifications for Utility-Scale Storage

To model Energy Storage Systems (ESS) collocated at the wind sites, special generators were specified with different charging and discharging efficiencies to represent the physical properties of the ESS. The energy available in any ESS can be used to provide energy in the different wind scenarios and to help support the grid in contingencies. The optimal use of storage is dependent in part on the value assigned to the stored energy. If it is valued at zero, then stored energy is always used in contingencies and in the last hour of the planning horizon. There is, however, an opportunity cost for discharging the ESS that provides a high threshold for discharging. If the nodal price in the terminal state is very low, for example, it would in reality be optimum to not discharge the ESS and wait until a later period when the price is higher than the high threshold. A similar argument can be made for charging the battery, and a low threshold provides the opportunity cost for charging. It is optimum to charge the ESS if the price is below the low threshold. If the price is between the two thresholds, the optimum is to do nothing and save the stored energy for use later. The formulation for ESS can also provide a robust estimate of the amount of reserves necessary to cover the extreme changes on demand and the stochastic inputs. For this paper, we adopted an implementation that uses an estimate on the expected value of the amount of storage at the beginning of the horizon considered.

This decision stems from the need to provide flexibility in the use of the ESS: by enforcing the limits necessary to respond to the worse likely changes, the ESS needs to be tightly dispatched at the beginning of the horizon, to provide time arbitrage opportunities. Once the bounds that limit the dispatch diverge, the ESS can be used for uncertainty mitigation. But this tightly controlled dispatch limits the possibility of controlling the uncertainty in early periods, leading to unexpected optimal dispatches, like spilling wind in high availability states while discharging the ESS. Thus, the use of the expected wind available leverages the information on the wind distribution and increases the flexibility in usage for uncertainty mitigation.

In the empirical analysis, the ESS are located at the same buses as the wind farms and the total capacity of the ESS is the same as the total energy of deferrable demand, 33.9GWh. This specification makes it easier to make comparisons between cases with ESS and deferrable demand. The maximum hourly power available per ESS is set to be 22% of the energy capacity. This is based on the assumption that 85% of level1 and 15% of level2 charging rates are available. Compared to 70/30 (level1/level2) for the EV case, this lower charging rate of 85/15 is assumed because many wind farms are located in rural areas, far from major demand centers and are connected via relatively low capacity transmission lines.

To determine the threshold price of the stored energy for discharging, the initial pattern of dispatch for generators and the initial amounts of stored energy, an iterative process is implemented in which the daily dispatch is simulated several times, using the same input specifications, until the differences in the threshold price and initial conditions are stable and below a tolerance level. These initial conditions can be considered, therefore, as a steady state solution for a series of identical days.

V. RESULTS OF THE CASE STUDY

The results in this section summarize the cost of serving a given demand profile for a 24-hour period for four different cases. The injections and exports from outside of the New York and New England region (NYNE) are fixed, to focus on this territory. For this reason, the results include information only for NYNE, and the locations of wind farms and storage are all in this region. The analysis assumes that the wholesale market is deregulated and run by an Independent System Operator (ISO). Many studies of the effects of renewable generation on system costs focus on the payments made by customers in wholesale markets and the associated decrease in the energy prices when renewable energy sources are available. We have argued in earlier research that this emphasis ignores the financial adequacy issue for conventional generators [32]. Since the offers submitted by renewable sources are effectively zero, average nodal prices are generally lower. Therefore, these new renewable sources displace fossil fuels and the conventional generators receive less net revenue to cover their capital expenses. To rectify this situation and still maintain
system reliability, generators are further compensated in capacity markets that help to provide the “missing money”. To avoid the distortions from evaluating a policy based solely on the wholesale payments from customers, the different cases are evaluated using measures that reflect the total system costs. The measures used for this analysis are 1) the actual operating costs incurred by conventional generators, 2) the amount of wind generation dispatched, and 3) the maximum conventional generation capacity needed to cover the peak demand and maintain system reliability.

The simulation starts at midnight and finishes at the end of the day. The main interest of the analysis is the management of stochastic wind generation and the provision of load following reserves. For this reason the time steps are hours, therefore abstracting from the provision of frequency regulation in real-time that require rapid changes in the dispatch patterns to balance demand and supply in response to the variability of both the intermittent renewable sources and demand.

A. The Effects of Storage Capacity on Total System Costs

The results in this section summarize the cost of serving a given demand profile for a 24-hour period for six different cases on a hot summer day using the network in Figure 1. The main purpose of the analysis is to determine how different types of storage interact with stochastic wind generation and how they affect system costs. System level results are presented for the following six cases:

1) Case 1: No Wind, Initial System
2) Case 2: 32GW of Wind Capacity at 16 locations.
3) Case 3a: Case 2 + 33.9GWh of Thermal Storage (TS) at 5 load centers.
4) Case 3b: Case 2 + 33.9GWh of batteries in Electric Vehicles (EV) at 5 load centers.
5) Case 3c: Case 2 + 33.9/2GWh of both TS and EV at 5 load centers.
6) Case 4: Case 2 + 33.9GWh of Energy Storage Systems (ESS) collocated at the 16 wind farms.

The wind capacity represents roughly 20% of the system load and the uncertainty of this resource requires the purchase of additional reserve capacity for “load following” (LF ramping reserves) as well as reserve capacity to cover contingencies. The specifications of Cases 3a and 3c distinguish Conventional Demand (CD) from the Deferrable Demand (DD) associated with TS, and two hourly demand profiles are used as inputs. CD must be covered each hour by purchasing electricity, and DD, representing the demand for cooling services, can be met by purchasing electricity or by melting stored ice made with previously purchased electricity. In Cases 3b and 3c, charging the batteries in EVs represents an increase of demand above the needs of CD and DD and the discharging corresponds to driving the EVs and does not affect operations on the grid.

In general, changing the dispatch points of thermal power plants to provide ramping services reduces their efficiency and causes damage that is accrued over time [33]. These costs are referred to as Ramping Costs (RC), and include lower performance (e.g. heat rate degradation for thermal generating units), equipment damage (e.g. creep damage, increases in Equipment Forced Outage Rates (EFOR)) and higher operating and maintenance costs (O&M). These costs are represented by $R^{\text{g}}_i()$ and $R^{\text{p}}_i()$ in the objective function in (1).

In order to schedule enough capacity for unforeseen inter-period changes in power requirements, a SO contracts with the conventional generators for Ramping Reserves (RR). These RR are similar to Contingency Reserves (CR) in their procurement i.e. they are paid in advance during the first settlement of the market (e.g. in a day-ahead market). The main difference in reserves is that the RR deals with relatively high probability events (e.g. intact states of the system associated with changes in load and wind generation) and CR deals with rare events associated with equipment failures (contingencies). In practice, reserve capacity can be used for both RR and CR and the CR is measured by the additional reserves needed to maintain reliability after determining the amount of RR. For RR, both up and down ramping reserves are needed, and the amount of down ramping reserves purchased determines how much of the potential wind generation has to be spilled when wind speeds are unexpectedly high. In contrast, CR is generally for up ramping to replace, for example, the failure of a large generating unit.

B. System Cost Analysis

The results for the daily composition of operating costs and the amounts of wind generation, conventional generation and reserve capacity needed to maintain reliability focus on the difference between Case 2 and Case 1, to assess the effects of adding wind generation, and the differences between Cases 3a - 4 and Case 2, to assess the effects of adding different types of storage. For each hour of the day, the model determines the optimum pattern of dispatch for 12 different states of the system, four corresponding to different levels of potential wind generation and eight to contingencies.

Adding wind capacity in Case 2 reduces the operating costs significantly, mainly by displacing fossil fuels and E[Generation Cost] is roughly 20% lower than it is in Case 1. However, the reduction in E[Total Operating Cost] is smaller than the reduction in E[Generation Cost] because of the total cost of ramping reserves is over three times as large as it is in Case 1. Comparing the four storage cases with Case 2, all four reduce the E[Generation Cost] by dispatching more wind generation, and all four reduce E[Ramping Costs] by providing some ramping services. In general, storage capacity makes it optimum to spill less wind, by charging the storage in the high wind states, and to commit less ramping reserves by mitigating the uncertainty of wind generation. Instead of using natural gas generating units to provide up and down ramping, storage plays this role instead at no cost.

The reductions in E[Total Operating Cost] are larger for ESS in Case 4 than either Case 3a or Case 3c with TS, and the reductions for Case 3b with EV only are very small. The reductions for the combination of TS and EV in Case 3c are between Case 3a and Case 3b. The basic reason for
the superior performance of ESS compared to TS is that the batteries are more efficient and can charge and discharge faster than TS. Even though an EV has a high quality battery, the poor performance of Case 3b is caused by 1) charging the batteries represents an addition to total demand, and 2) this example does not include vehicles with vehicle-to-grid capabilities.

The overall results for the E[Total Operating Cost] show that adding wind generation causes the biggest cost reduction (17%), and the four storage cases cause additional modest reductions of 8%, 2%, 6% and 11% for Cases 3a - 4, respectively. These results are summarized in Figure 2 below and have been discussed in more detail for a similar analysis in an earlier paper [34].

The next set of results presented in Table III show that a major economic benefit of storage is to reduce the amount of conventional generating capacity needed to maintain System Adequacy and the associated capital costs. Since the day chosen for the simulation is the hottest summer day, the sum of the maximum dispatch in any of the system states for each conventional generating unit is used to measure the installed generating capacity for System Adequacy.

<table>
<thead>
<tr>
<th>TABLE III</th>
<th>PEAK HOUR SUMMARIZED SYSTEM RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Outcome (MW)</td>
<td>c1</td>
</tr>
<tr>
<td>Conventional Generation</td>
<td>90,570</td>
</tr>
<tr>
<td>ESS Dischargeab</td>
<td>-</td>
</tr>
<tr>
<td>Deferrable Demand, TS</td>
<td>-</td>
</tr>
<tr>
<td>Deferrable Demand, EV</td>
<td>-</td>
</tr>
<tr>
<td>Capital Cost ($1000)</td>
<td>-</td>
</tr>
<tr>
<td>Conventional Generating Unitsb</td>
<td>104,844</td>
</tr>
<tr>
<td>ESS</td>
<td>-</td>
</tr>
<tr>
<td>Deferrable Demand TSd</td>
<td>-</td>
</tr>
<tr>
<td>Deferrable Demand EV</td>
<td>-</td>
</tr>
<tr>
<td>Total Capital Cost</td>
<td>104,844</td>
</tr>
</tbody>
</table>

ab Energy Storage System (ESS)

bc Annual capital cost for a peak of 300$/kWh/year allocated to 100 peak hours with 2 peak hours for this day

dBased on an installation cost of $500/kWh, an operating cost of $50/kWh-year and a 25 year life cycle

eBased on an installation cost of $500/kWh, an operating cost of $50/kWh-year and a 25 year life cycle

The results in Table III show the maximum amounts of conventional generation and storage charge/discharge and the differences of these maxima for the 12 different states of the system at the peak system load (3PM). The maximum Conventional Generation is lower when wind generation is available in Case 2 because the wind provides some capacity value, and it is even lower in the four other cases with storage. It is important to remember that the amount of electricity delivered to customers with TS in Cases 3a and 3c may be higher or lower than the fixed amount of conventional demand in Cases 1, 2 and 4. In Case 3b with EV, the optimum strategy is to avoid charging the vehicles during peak load periods. Although the biggest reduction of Conventional Generation occurs with ESS in Case 4, the most important implication is that the electric energy delivered to customers is lower with TS in Cases 3a and 3c than Case 4 because the TS provides customers with some cooling services. This reduces the level of congestion on the grid. In contrast, discharging ESS provides another source of energy for customers and the peak system load is unaffected.

The lower half of Table III shows the capital costs of Conventional Generating Units and the different types of storage. An explanation of how these costs are determined is provided in the footnotes of Table III. In simple terms, the batteries in ESS and EV are relatively expensive compared to Conventional Generating Units. In contrast, TS is relatively inexpensive because the cost represents an augmentation of an existing energy service, space cooling. In the same way with an EV, the cost of the battery is extra but the cost of the vehicle itself is already covered for transportation. The Total Capital Cost is lower for Case 2 and even lower for Case 3a with TS. Compared to Case 2, the Total Capital Cost is higher with EV in Cases 3b and 3c and about the same with ESS in Case 4.

![Fig. 2. A Comparison of Cost Savings for Different Types of Storage.](image)

Figure 2 combines the results for the operating costs and capital costs to provide an overall view of the additional savings in total cost for the four cases with storage compared to Case 2 with wind generation but no storage. The components of savings are the costs of Generation, Reserves, Capital of conventional generating units, and Gasoline and the capital cost of Storage is an extra expense. In all four cases, the savings are larger than the capital cost of storage, and the net savings for Cases 3a, 3b, 3c and 4 are $6.0, $8.8, $9.4 and $5.6 million, respectively. The savings in Gasoline are very important even though gasoline is relatively inexpensive by European standards. For Case 4 with ESS, the net savings will lower customers bills. For Cases 3a, 3b and 3c with deferrable demand, the savings in system cost will lower bills for customers and they will pay for the storage themselves and pay less for gasoline. Nevertheless, these savings will not be realized unless the bills paid by customers reflect the true system costs. This issue is addressed in the following subsection by comparing the typical electric bills for customers with different types of deferrable demand using Case 3c with both TS and EV.

C. Total Payments for Different Types of Customers

Given the optimization results from Case 3c with both EV and TS, it is possible to compute the typical bills paid...
by different types of customer. Customers were classified into the following four types: 1) customers who have no DD, 2) customers who own thermal storage, 3) customers who own an EV, and 4) customers who own both thermal storage and an EV. Customers who do not own thermal storage are assumed to have a conventional air conditioner, and customers who do not own an EV are assumed to have a conventional gasoline vehicle. To make comparisons easier, all customers are assumed to have identical hourly demand profiles for electric services (e.g. lighting), space cooling and transportation. Since the total number of vehicle owners in the region is 15,692,624, this is specified as the total number of customers. The specification of the number of EVs in Case 3c corresponds to 10% of this total; half of these customers are assumed to own an EV but no TS, and half own both TS and an EV. To make aggregate storage capacity consistent with Case 3c, another 5% of all customers own TS but not an EV. The remaining 85% of customers have no DD, have a conventional air conditioner and own a gasoline vehicle.

Figure 3 shows the expected hourly demand profiles for purchasing electric energy from the grid by four different types of customer. Customers with no DD have a standard demand profile for a hot summer day that is low during the night and high during the day. The profiles for customers with DD are strikingly different because they purchase significantly more energy at night from 11PM to 6AM to charge the EV battery and/or to make ice for TS. This strategy takes advantage of the low electricity prices at night and for TS avoids buying energy when prices are high in the afternoon. At night, the purchases for TS and EV complement each other and the peak demand for a customer with both TS and EV occurs at 2AM and is about 70% higher than the peak demand of customers with no DD that causes the peak system load at 3PM.

Another feature of Figure 3 is that the demand profiles for customers with DD, particularly TS, are not smooth because they are providing ramping services to mitigate wind uncertainty. For example, the increase of demand by customers with TS from 9 to 10 AM is in response to an increase in potential wind generation. After that, reducing the peak system load in the afternoon becomes a higher priority. For peak hours when prices are high, customers with TS melt stored ice to meet their demand for space cooling and reduce their use of conventional air conditioners and their purchase of expensive energy. By doing this, customers with TS also help to reduce the peak system load and the amount of conventional generating capacity needed to maintain System Adequacy. It is important for customers with TS to get rewarded for lowering the system costs, and every customer should pay a demand charge proportional to their own demand at the system peak. Consequently, a customer with TS will pay a lower demand charge than other customers.

Customers should also be paid if they use DD to provide ancillary services such as ramping. In practice, however, customers with DD (instructed demand) also have Conventional Demand (CD) (uninstructed demand) that causes ramping. All customers should pay for the ramping needed for their CD, get paid for providing ramping services with DD and their net payment to the grid for ramping may be negative. Treating CD and DD differently assumes implicitly that they are metered separately.

Table IV shows the composition of the economically efficient payments by each type of customer to their utility. In addition to purchasing energy, customers pay a demand charge proportional to their total demand at the system peak load (3 PM), CD pays for the ramping needed to change demand from hour to hour, and DD gets paid for providing ramping reserves (using the offer price of $5/MW for conventional reserve capacity from a peaking unit). For a customer with no DD, the energy and capacity payments are similar in magnitude and the ramping payment is very small to give a total payment of $7.20/day. The total payment for a customer with TS is two thirds of this total because more energy is purchased at night, demand at the peak is cut by one half and there is a small net payment to the customer for ramping. Customers with an EV pay more than customers with no DD because they buy a lot more energy, have slightly higher demand at the system peak and receive relatively low payments for providing ramping services. The differences in payment for customers with both

<table>
<thead>
<tr>
<th>$/day</th>
<th>No DD</th>
<th>TS only</th>
<th>EV only</th>
<th>TS and EV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Payment</td>
<td>3.60</td>
<td>3.36</td>
<td>4.11</td>
<td>3.87</td>
</tr>
<tr>
<td>Ramping Payment</td>
<td>0.0073</td>
<td>-0.3844</td>
<td>-0.2527</td>
<td>-0.6444</td>
</tr>
<tr>
<td>Payment by CD</td>
<td>0.0073</td>
<td>0.0056</td>
<td>0.0073</td>
<td>0.0056</td>
</tr>
<tr>
<td>Payment by DD</td>
<td>-</td>
<td>-0.3900</td>
<td>-0.2600</td>
<td>0.6500</td>
</tr>
<tr>
<td>Capacity Payment</td>
<td>3.59</td>
<td>1.78</td>
<td>3.67</td>
<td>1.87</td>
</tr>
<tr>
<td>Total Payment</td>
<td>7.20</td>
<td>4.76</td>
<td>7.53</td>
<td>5.10</td>
</tr>
<tr>
<td>Storage Cost</td>
<td>-</td>
<td>1.48</td>
<td>2.97</td>
<td>4.45</td>
</tr>
<tr>
<td>Gasoline Cost</td>
<td>5.44</td>
<td>5.44</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total Cost</td>
<td>12.64</td>
<td>11.68</td>
<td>10.50</td>
<td>9.55</td>
</tr>
<tr>
<td>Flat Payment</td>
<td>5.51</td>
<td>5.74</td>
<td>6.55</td>
<td>6.76</td>
</tr>
</tbody>
</table>
TS and EV compared to customers with TS only are similar to the differences between customers with EV and those with no DD.

![Efficient Payments for Different Types of Customer](image)

Fig. 4. The Composition of Efficient Utility Bills for Different Types of Customer.

The efficient utility bills for different types of customer, taken from Table IV are illustrated in Figure 4. For customers with TS, the importance of the capacity payment compared to a customer with no DD is obvious. The saving in energy payment is modest, and even though there is a net payment for providing ramping, it is still very small. However, it should be remembered that the ramping provided by TS and EV reduces the ramping reserves acquired from conventional generating units substantially. Adding an EV increases the amount of energy purchased and the total size of the utility bill. Table IV and Figure 4 also show the payments if customers paid a flat rate for purchasing energy and nothing else. This is the basic structure of the retail rates paid by most customers in the US. The flat price (¢13.5/kWh) is chosen to raise exactly the same total revenue as the efficient payments. The economic effects of the flat price are perverse for TS because adding TS results in a larger utility bill, and furthermore, customers with no DD or with an EV do not pay enough for their demand at the system peak.

Overall, the results for payments imply that the most important incentive for TS is to get the demand charge for capacity correct, followed by paying real-time prices for energy and getting paid for providing ramping. Table IV also shows the additional costs that customers have to pay for gasoline, if they do not have an EV, and the capital cost of storage if they acquire a TS or an EV. The customers with DD have lower total costs than customers with no DD. The reduction is only 8% for a customer with TS, but the large savings in gasoline purchases saves customers with an EV 17%, and customers with both TS and an EV save 25%. These are meaningful reductions.

VI. CONCLUSIONS

This paper presents a multi-period, stochastic form of Security Constrained Optimal Power Flow (SCOPF) and shows how this framework can be used to evaluate the system effects of adding stochastic sources of wind generation to a network. In general, stochastic sources require more reserve ramping capacity to maintain Operating Reliability. The economic effects of adding stochastic wind resources are illustrated in a simulation of operations for a hot summer day using a reduction of the NPCC network. The results presented in Section V demonstrate the beneficial effects of adding two different types of storage capacity: distributed storage (Deferrable Demand (DD) from Thermal Storage (TS) and Electric Vehicles (EV)) and Energy Storage Systems (ESS) collocated at wind farms. Both forms of storage can lower system costs by 1) flattening the daily pattern of dispatch by conventional generating units, 2) spilling less of the potential wind generation, 3) mitigating the ramping costs associated with wind uncertainty, and 4) reducing the amount of conventional reserve capacity needed for reliability. TS can further reduce system costs by reducing the peak system load (covered by conventional generation, wind generation, and discharging ESS) and the amount of congestion on the network.

Developing an electric delivery system that can accommodate high penetrations of renewable sources of generation effectively will require developing a smarter grid. Unfortunately, the publics reaction to the first small steps in building a smart grid for demand response, such as the Advanced Meter Initiative, has been largely negative. Smart meters are, however, essential for providing real-time price information for customers and the economic incentives needed to get the type of demand response exhibited in cases with DD. We have argued before that a successful smart grid must yield direct economic benefits for customers, i.e. paying lower bills while still meeting their energy needs. Our results show that the total system costs are reduced substantially by adding wind generation even though the cost of providing ramping services increases. An additional requirement with uncertain wind generation is to provide down-ramping reserves to avoid spilling potential wind generation for the rare system states with unexpectedly higher wind speeds. Total system costs are reduced further in the cases with storage by 1) spilling less of the potential wind generation, 2) providing ramping services, and 3) reducing the amount of installed generating capacity needed to maintain System Reliability. An important additional feature of TS is that it lowers the peak purchase of power from the grid, and thereby, reduces congestion on the network and the total amount of utility-owned capacity needed to maintain reliability.

When the costs of storage, and the savings in gasoline purchases for owners of an EV, are taken into account, the savings in operating costs and gasoline purchases are big enough to cover the capital cost of storage in all cases. The biggest savings are for the cases with EV caused mainly by the lower gasoline purchases. The main savings with TS is from reducing the amount of installed conventional generating capacity needed to maintain System Adequacy. ESS and TS affect the hourly dispatch of conventional generating units more than EV because they both lower the peak conventional
generation during the afternoon. Although the overall savings in operating costs are higher with ESS than with TS, the capital cost of ESS is also higher and the net saving is higher with TS.

To assess the relative benefits for different types of customer, economically efficient utility bills were computed for customers with 1) no DD, 2) TS only, 3) EV only, and 4) both TS and EV. Every customer is assumed to have exactly the same pattern of daily energy needs. Customers with an EV substitute electricity for gasoline, and customers with TS decouple their purchases of electricity from the delivery of space cooling services but they all receive the same energy services. The daily patterns of electricity purchases are strikingly different. Unlike customers with no DD, the purchases by customers with DD are higher at night than they are during the day. The lowest efficient payments are made by customers with TS by paying, primarily, lower demand charges for capacity. When the cost of storage and the relatively large savings in gasoline purchases are considered, the lowest total costs are for customers with an EV.

The final analysis demonstrates that the current widely used structure of flat retail rates for energy, with no demand charge and no consideration of ramping, provides perverse economic incentives for adopting DD capabilities. These flat rates are a very important barrier to adopting DD that do not reflect the correct economic incentives. For example, if customers only pay flat rates for energy, the non-adopters would be free loaders and pay lower bills than the adopters, because the latter have to cover storage inefficiencies to get the same amount of energy services delivered. Real-time pricing is essential to provide the correct economic incentives and reward adopters and penalize non-adopters. In addition, if customers are to get the correct economic benefit of reducing their demand at the peak system load, they should pay for their actual demand during peak periods. Most customers do not pay a demand charge at all, and when the level of demand is measured with a traditional meter, this level is the maximum demand over a billing period even if it occurs at night when the system load is low. Paying a lower demand charge is the main economic incentive for customers to invest in TS. Getting paid for providing ramping is also important for reducing operating costs. It is fair to say that unless the economic incentives for demand-side participation change to reflect the true system costs and benefits, it seems unlikely that customers will appreciate the potential economic benefits of the smart grid, and the utility industry will continue to depend on supply-side solutions for problems and assume that regulators will produce the correct economic incentives and reward adopters.

ACKNOWLEDGEMENTS

This research was supported by the US Department of Energy through the Consortium for Electric Reliability Technology Solutions (CERTS) and the Office of Electricity Delivery and Energy Reliability under the project “The Future Grid to Enable Sustainable Energy Systems”, and by the Power Systems Engineering Research Center (PSERC). The authors are responsible for all conclusions presented in the paper and the views expressed have not been endorsed by the sponsoring agencies.

REFERENCES
