A Robust Solution to the Load Curtailment Problem

H. P. Simão, H. B. Jeong, B. Defourny, *Member, IEEE*, W. B. Powell, *Member, IEEE*,
A. Boulanger, A. Gagneja, L. Wu, *Member, IEEE*, and R. N. Anderson, *Member, IEEE*.

Abstract—**Operations planning in smart grids is likely to become a more complex and demanding task in the next decades. In this paper we show how to formulate the problem of planning short-term load curtailment in a dense urban area, in the presence of uncertainty in electricity demand and in the state of the distribution grid, as a stochastic mixed-integer optimization problem. We propose three rolling-horizon look-ahead policies to approximately solve the optimization problem: a deterministic one and two based on approximate dynamic programming (ADP) techniques. We demonstrate through numerical experiments that the ADP-based policies yield curtailment plans that are more robust on average than the deterministic policy, but at the expense of the additional computational burden needed to calibrate the ADP-based policies. We also show how the worst case performance of the three approximation policies compares with a baseline policy where all curtailable loads are curtailed to the maximum amount possible.**

Index Terms— **Computer simulation, load management, power distribution, power system management, power system modeling, smart grids, demand response, optimization methods, mathematical programming, approximate dynamic programming.**

I. NOMENCLATURE

Sets

B Set of all customer loads b in the network

M Set of all nodes in the network (i.e., cable junctions, transformers, substations)

$M_{i\rightarrow}$ Set of all nodes j in the network that are connected to node i through a single, direct cable section $i \leftrightarrow j$ such that $j > i$

$M_{\rightarrow i}$ Set of all nodes j in the network that are connected to node i through a single, direct cable section $i \leftrightarrow j$ such that $j < i$

N Set of all transformers in the network ($N \subset M$)

L_b Set of all possible curtailment levels for load b ; if load is not curtailable, $L_b = \{0\}$

Parameters

l_b A possible curtailment level for load b ($l_b \in L_b; 0 \le l_b \le 1$)

T Length of the planning horizon (in number of discrete time steps)

β "Revenue" multiplier in the net revenue component of the objective function

λ "Cost" multiplier in the net revenue component of the objective function

$f_{i,j}^{\max}$ Rating of cable section $i \leftrightarrow j$, $\forall i \in M$, $\forall j \in M_{i\rightarrow}$

γ_1 Multiplier for tier-1 of the cable over-rating penalty in the objective function

γ_2 Multiplier for tier-2 of the cable over-rating penalty in the objective function

δ_γ Cable rating factor that serves as a threshold between tier-1 and tier-2 penalties

$f_k^{\max}$ Rating of transformer $k \in N$

ς_1 Multiplier for tier-1 of the transformer over-rating penalty in the objective function

ς_2 Multiplier for tier-2 of the transformer over-rating penalty in the objective function

δ_ς Transformer rating factor that serves as a threshold between the tier-1 and tier-2 penalties

κ_b Minimum curtailment notification time (in number of steps) required by customer load b

ρ_{b,l_b} Minimum time (in number of steps) load b needs to remain at curtailment level l_b

$X_{i,j}$ Reactance on cable $i \leftrightarrow j$, $\forall i \in M$, $\forall j \in M_{i\rightarrow}$

$R_{i,j}$ Resistance on cable $i \leftrightarrow j$, $\forall i \in M$, $\forall j \in M_{i\rightarrow}$

φ^{Loss} Linear factor for approximating power losses in a cable section

ψ_i^{PF} Power factor (i.e., real power divided by apparent power) at node (bus) $i \in M$

$\eta_{t,t',i,j}(\hat{q})$ = conversion factor in the power flow equation of cable $i \leftrightarrow j$, $\forall i \in M$, $\forall j \in M_{i\rightarrow}$ during $[t', t'+1)$, when planning at time $t \leq t'$, if $\hat{q}_{t,t',i,j} = 1$;

       = 0, if $\hat{q}_{t,t',i,j} = 0$

$A_{t,t',i,b}(\hat{q})$ = fraction of demanded power to be supplied by node (bus) i to load b during $[t', t'+1)$, when planning at time $t \leq t'$, if $\hat{q}_{t,t',i,i} = 1$;

       = 0, if $\hat{q}_{t,t',i,i} = 0$

Exogenous Information

$\hat{D}_{t,t',b}$ Apparent power (i.e., the absolute value of complex power) demanded at load b during $[t', t'+1)$, when planning at time $t \leq t'$

$\hat{q}_{t,t',i,j}$ = 1, if cable $i \leftrightarrow j$, $\forall i \in M$, $\forall j \in M_{i\rightarrow}$ is active during $[t', t'+1)$, when planning at time $t \leq t'$;

       = 0, if it is out of service

$\hat{q}_{t,t',k,k}$ = 1, if node k is active during $[t', t'+1)$, when planning at time $t \leq t'$;

       = 0, if it is out of service

W_t $= \left(\hat{D}_{t,t'}, \hat{q}_{t,t'} \right)_{t' \geq t}$

Decision Variables

y_{t,t',b,l_b} = 1 if load b should curtail to level l_b during $[t', t'+1)$, when planning at time $t \leq t'$;

       = 0, otherwise

z_{t,t',b,l_b} Time (in number of steps) load b will have been at curtailment level l_b by the end of $[t', t'+1)$, when planning at time $t \leq t'$

$p_{t,t',i}$ Real power injected at node (bus) i during $[t', t'+1)$, when planning at time $t \leq t'$

$f_{t,t',i,j}$ Real power flowing through cable $i \leftrightarrow j$, $\forall i \in M$ $\forall j \in M_{i\rightarrow}$ during $[t', t'+1)$, when planning at time $t \leq t'$

$f_{t,t',k,k}$ Real power flowing through transformer k during $[t', t'+1)$, when planning at time $t \leq t'$

$\theta_{t,t',i}$ Phase angle at node i during $[t', t'+1)$, when planning at time $t \leq t'$

u_t $= \left(y_{t,t'}, z_{t,t'}, p_{t,t'}, f_{t,t'}, \theta_{t,t'} \right)_{t' \geq t}$

General Notation Convention

$r_{t,t'}$ $= \left(r_{t,t',i,j} \right)_{\forall i, \forall j}$

II. INTRODUCTION

Several recent and ongoing developments are likely to transform the electricity distribution grid in the next decades. These developments include: (i) the widespread installation of smart meters, remote network monitoring equipment, and intelligent grid control systems; (ii) the incorporation of smart energy management technologies in buildings; (iii) the growing integration of time-intermittent renewable sources (like solar); (iv) the easier access to distributed power generation and storage devices; (v) the penetration of plug-in electric vehicles; and last but not least, (vi) the proliferation of various forms of demand response, load curtailment, and pricing programs. The presence of one or more of these features simultaneously in a distribution grid, combined with the uncertainty in demand and in the state of the electrical components in the network will create a more complex and challenging system, whose dispatch and control will require new procedures and computational tools.

We envision a short-term planning tool to be used by a utility dispatcher, particularly when a contingency in the distribution grid has already happened and/or one or more additional contingencies are likely to happen (*contingency* being defined here as the shutting down, or *failure*, of a whole section of the distribution grid). This system can be referred to as a *load and source optimization controller* (LSOC).

The main contributions of this paper are as follows. First, we propose a detailed, dynamic model of the load curtailment problem, with careful and accurate modeling of lagged information processes. Second, we propose and test two novel robust policies for making load curtailment recommendations based on the modeling and algorithmic framework of approximate dynamic programming. These policies are practical and computable, and can be used in a dynamic setting to provide guidance to human dispatchers to help prioritize curtailment decisions.

Distribution systems have been built with redundancy, particularly in large, dense urban areas. Branches can be disconnected in response to contingencies, at the cost of loading other lines and transformers. However, overloading remaining components increases risks of cascading failures.

We are proposing a different mode of operation under contingencies for the future, where the risk of having too many contingencies is mitigated by localized, preemptive actions at the load side, made on a voluntary basis in response to prior notification from the utility. We call this mode *proactive*. Since users such as building operators require advance notification of curtailments, we face the challenge of designing policies which carefully anticipate the possibility of not being able to meet demand, while minimizing unnecessary disruptions to daily activities.

To illustrate the application of the proposed methodology, we picked one particular action: the curtailment of loads. But this is just one of the many decisions that could be modeled within this framework. Other decisions involve when and how much charge to put in the batteries of electric vehicles, how much energy to take from solar panels, when and where to plug-in mobile generators in the grid, and so on.

We describe a stochastic optimization model comprised of a sequence of time-indexed sub-problems, solved successively in a rolling horizon fashion over a planning horizon. Each sub-problem is modeled as a mixed-integer programming problem solved over a shorter planning horizon. In order to deal with the short horizon of the sub-problems and the uncertainty in the problem, we implement and test two types of approximate dynamic programming approaches: a cost function approximation and a value function approximation (Chapter 6 in [1]), where the first is expected to be computationally easier to calibrate, but the second is expected to produce better results.

As already mentioned, in the application of the proposed modeling framework we focused on one particular feature: load curtailment in the presence of contingencies in the grid. The load curtailment problem involves determining a *robust* set of customer loads to curtail, and by how much, over the planning horizon, so as to maximize the *expected* value of a utility function. Here, we are using "robustness" to mean "works well on average over many outcomes," (see Mulvey et al. [30]) as opposed to "works well over all outcomes" (Bertsimas and Sim [31]). Equivalently, we seek to find a compromise between maximizing the amount of actual power provided to the customers and minimizing the likelihood of critically overloading the grid. As failures in sections of the grid (contingencies) do happen, or as their likelihood increases, it becomes increasingly likely that customers participating in the curtailment programs will be asked to pre-emptively curtail their loads within the next few hours. The current version of LSOC was designed to help with such short-term planning.

The optimization model presented in this paper can be seen as a type of unit commitment problem, where the control of load curtailment through binary variables resembles the control of generators [2]. Whereas the classic unit commitment problem focuses on planning generation (creating energy), our problem focuses on reducing load. Both result in integer programming problems planned on a rolling horizon basis. Saber and Venayagamoorthy [3] discuss the unit commitment problem in the presence of vehicle-to-grid (V2G) capabilities.

For an approach to control of (smart) grids using distributed devices or agents, see papers in [4] and [5]. Divan [6] and Divan and Johal [7] discuss a massively distributed control approach. Several authors have proposed the use of approximate dynamic programming methods to gain intelligence for the smart grid [8], [9], [10]. The use of stochastic programming to manage electric vehicle charging, V2G facilities and renewable sources in the context of distribution network congestion is discussed in [11] and [12].

Demand response programs have become increasingly important and popular in the power industry and research. A 2006 report published by the U.S. Department of Energy describes the benefits of demand response in electricity markets and provides several recommendations for achieving them [13]. Later, Spees et al. [14], the Federal Energy Regulatory Commission [15], and Goldman et al. [16] published assessment papers on demand response and energy efficiency.

One of the prerequisites of demand response is the capability to forecast short-term electrical load [17]. Several approaches of load and energy demand forecasting have been proposed since the early 1990s, including time series models such as ARMA (auto-regressive moving average) [18] and ARIMA (autoregressive integrated moving average) [19], neural networks [20], and support vector machines [21], [22]. In this paper, we use a deseasonalized exponential smoothing model, adapted from a demand forecasting model named damped trend multi-calendar (DTMC) exponential smoothing, first developed by Godfrey and Powell [23].

Another key factor in demand response is the reliability of the power grid, particularly with respect to failures of electrical components. Gross et al. have applied machine learning-based susceptibility analysis to electrical feeder failures [24]. Rudin et al. [25] have performed a comprehensive study on the application of machine learning techniques in the preventive maintenance of the power grid.

Among the approximate policies described in this paper, the rolling horizon look-ahead related policies were inspired by prior work on approximate dynamic programming done by some of the authors [1].

The remainder of the paper is organized as follows. Section III contains the description, the model and the solution approach to the load curtailment problem. Numerical experiments designed to show the robustness of the generated curtailment plans are described and presented in section IV. Conclusions are summarized in section V.

III. THE LOAD CURTAILMENT PROBLEM

We first describe the general setting of the problem. We then formulate it as a stochastic, sequential decision optimization problem. And finally, we present approximate policies to solve it.

A. Problem Description

The power distribution network used in our study is composed of (mostly) radial distribution feeders (27kV) – the *primary network* – connecting the substations to distribution transformers, which are in turn connected to a network (in the form of a mesh) of secondary low voltage lines (120V) – also known as the *secondary network*. Customers are by and large connected to the secondary network, but some (in general large load customers) may be connected directly to transformers in the primary feeders or to spot networks[1].

A contingency in this network is defined as the failure of a whole feeder. The shutting down of a feeder may also happen as a result of a planned outage. A subset of the customers in the grid has signed up to load curtailment programs and will thus be called curtailable loads. We assume that each customer

---

[1] A spot network is a small, isolated sub-network connecting a few customers to a few transformers.

may have a different curtailment contract, which specifies the discrete levels at which power can be curtailed (between 0 – no curtailment – and a maximum amount), and the minimum required curtailment notification time (if any). We assume also that customers will always comply when asked to curtail their loads.

The load curtailment problem can thus be stated summarily as follows. *Given:* (1) the distribution grid network and its known state at the initial time; (2) the set of customers served by the network, including the curtailable ones and their respective curtailment contracts; (3) a set of forecasts of the customer loads over the desired planning horizon; and (4) a set of estimates of the probabilities of failure of the primary feeders over the same horizon; *the goal* is to determine the set of loads to be curtailed, if any, and by how much, so as to maximize the expected value of a utility function that includes bonuses for the total amount of power withdrawn from the network and penalties for the amount of power flow above the ratings of the components in the distribution grid.

B. The Optimization Problem

We solve the load curtailment problem over a planning horizon of, say, the next 15 to 24 hours, by formulating it as a mixed-integer programming (MIP) problem, embedded in a sequential decision framework. We discretize time (typically we use hourly time steps). Integer variables are used to decide on the level of curtailment for each customer and to control how long a customer has been at a given level of curtailment. Continuous variables are used to describe the power flow through the links and the phase angles in the nodes of the distribution grid network. Linear constraints enforce that only one level of curtailment is active for each customer at each time and that customers stay at a level of curtailment for a minimum amount of time. The objective function to be maximized is a utility function that was designed to balance multiple goals: to serve as much demand with as little power injected and as little power above the ratings in the primary grid as possible.

The power flow in the network embedded in this optimization problem is solved through the DC optimal power flow approximation, with the addition of empirically estimated loss factors. There are several reasons why we chose the DC OPF approximation. As Stott et al [26] point out, its solution does not require iterations, the optimization problem remains linear and therefore computationally less complex (which is imperative, particularly when modeling uncertainty), and its data requirements are rather manageable. The DC OPF approach has been widely used in optimization models for transmission systems [27], and in many of those cases it has been shown to be reasonably accurate [28]. In distribution systems, however, due to the lower voltage levels and the higher resistance-to-reactance ratio of the branches, it is less accurate, and as a result should be recognized as a limitation of the model, as errors in the DC approximation could produce infeasibilities in a more accurate AC model. We envision our model as a tool embedded in a broader decision making process, involving an AC power flow model being run to

evaluate the actions recommended by LSOC, precisely to verify and adjust for inaccuracies in the power flow estimation.

Another related model feature worth mentioning is the use of soft penalties to enforce transformer and branch capacities. We use a two-tier set of linear penalties that was designed to allow for some violation of capacity, under a penalty, but to curb excessive violations (more than, say, 5% above the rating of a component). We believe that this is an effective way of modeling capacity constraints in an optimization system where the output decisions have the explicit goal of reducing any overloading of the grid components. Soft constraints allow the integer programming solver (used in the ADP algorithm) to search over solutions that might be infeasible (even if they do not appear in the optimal solution). Soft constraints also recognize the presence of errors in the representation of the network which may underrepresent capacity, as well as errors due to the DC approximation.

Ideally we would solve the problem as a single giant stochastic MIP over the whole planning horizon. Since this is impractical, we use a rolling horizon look-ahead procedure, where we decompose the planning horizon in shorter decision horizons (of, say, 4 hours), which overlap with each other, except for the first (hourly) time step of each sub-problem. We implemented and tested a deterministic rolling horizon look-ahead procedure and two types of approximate dynamic programming (ADP) procedures: (i) a rolling horizon look-ahead procedure with *cost function* approximation (CFA), and (ii) a rolling horizon look-ahead procedure with *value function* approximation (VFA).

In order to calibrate the ADP procedures, we run several iterations of the load curtailment problem. Each iteration consists of performing a rolling horizon look-ahead run over the planning horizon, with random events sampled after the solution of each sub-problem, before the clock advances to the next time step. For the CFA procedure, we simply accumulate a tally of the customers and the load amounts that were curtailed over the iterations, where we make decisions *after* learning of the failures. We use these *posterior* decisions to construct a lookup table of which customers to curtail. At the end, we round the final average curtailment fraction of a customer to the nearest acceptable level of curtailment, and this level becomes the curtailment policy for that customer. For the VFA procedure, as we solve each sub-problem t at each iteration, we estimate the marginal value of the level of curtailment of that customer at time t and we smooth this value into an average over all iterations. These smoothed marginal values are then added to the objective function, similar to how value function approximations are used in approximate dynamic programming.

A few relevant observations are in order with respect to the calibration of the ADP procedures (CFA and VFA). First, these calibrations can be done in two modes. One is called a *cold start* case, in which the calibration starts from a set of null policies/values. This case in general requires a larger number of iterations until the policies/values converge. The second mode is called a *warm start* case and it starts from an

existing set of policies/values. The latter usually requires just a few iterations before convergence. In a typical setting, *cold start* calibrations are performed only when the underlying characteristics of the problem suffer significant changes. Otherwise, *warm start* calibrations are performed repeatedly, as new exogenous information arrives (namely, updated demand forecasts and/or updated estimates of the probability of failure of components). In any case, however, calibrations are always done off-line, in the background, while the main system runs live in the foreground.

The second relevant observation is that there will be one set of calibrated policies/values for each contingency case. So, for instance, in a network with, say, 24 primary feeders, we would have 24 cases of N-1 contingencies, 276 cases of N-2 contingencies, and so on. In practice, though, the total number of contingencies for which policies/values need to be calibrated will be limited by the likelihood that a particular higher order contingency may ever occur.

In the remainder of this section we present the mathematical formulation of the optimization problem (subsection C) and the approximate policies used to solve it (subsection D).

C. The Optimization Model

Let S_t describe the state of the system at the discrete points in time $t \in \{0,1,...\}$, that is, the state of the distribution grid and all customer loads, and let u_t be the vector of all the decision variables (controls). Further let W_t be the vector of realizations of the exogenous information process (W_t is deterministic at time t), and assume we have a system model $S_{t+1} = S^M\left(S_t, u_t, W_{t+1}\right)$. The challenge is designing a policy $u_t = U_t^\pi\left(S_t\right)$ that provides robust performance and solves

$$\max_{\pi \in \Pi} \mathbb{E} \sum_{t=0}^{T-1} C\left(S_t, U_t^\pi\left(S_t\right)\right),$$

where $C\left(S_t, u_t\right)$ expresses the net contribution from the policy at time t.

Note that with the exception of time (and the integer variables, that are discrete by nature) no other variables need to be discretized, and neither do the policies, since nowhere in this optimization model do we need to enumerate states. Note also that the presence of the expectation in the objective function signals that the solution to this problem is expected to perform well, on average, over the whole state space, and thus its robustness.

We start with a simple, deterministic look-ahead policy (model predictive control). Note that throughout this paper we will be using the convention of a double indexing of time (t, t') in order to emphasize the distinction between the time at which the information to make a decision is *available* (t) and the time at which the information becomes *active* ($[t', t'+1)$, $t' \geq t$). This distinction is crucial for the proper representation of rolling horizon look-ahead formulations.

The objective function $F_t\left(S_t\right)$ for the look-ahead model at planning time t is expressed by maximizing the value of a utility function summed over all discrete decision time steps in the planning horizon $\{t,...,t+T-1\}$:

$$F_t\left(S_t\right) = \max_{u_t} \sum_{t'=t}^{t+T-1} \left(R_{t,t'}\left(y_{t,t'}, p_{t,t'}\right) - P_{t,t'}\left(f_{t,t'}\right)\right). \tag{1}$$

The utility function in (1) has two components: a net revenue term and an over-rating penalty term. $R_{t,t'}\left(y_{t,t'}, p_{t,t'}\right)$ is a proxy estimate of the net revenue of serving all customer loads $b \in B$, given by

$$R_{t,t'}\left(y_{t,t'}, p_{t,t'}\right) =$$
$$\beta \sum_{b \in B} \left(\hat{D}_{t,t',b} \sum_{l_b} \left(1 - l_b\right) y_{t,t',b,l_b}\right) - \lambda \sum_{i \in M} p_{t,t',i}. \tag{2}$$

$P_{t,t'}\left(f_{t,t'}\right)$ is a penalty term designed to minimize the amount of power flow over the ratings of the primary distribution grid components (cable sections $i \leftrightarrow j$ and transformers k), given by

$$P_{t,t'}\left(f_{t,t'}\right) =$$
$$\sum_{i \in M} \sum_{j \in M_{i \to}} \left[\gamma_1 \left[\frac{|f_{t,t',i,j}|}{f_{i,j}^{\max}} - 1\right]^+ + \gamma_2 \left[\frac{|f_{t,t',i,j}|}{\delta_\gamma f_{i,j}^{\max}} - 1\right]^+\right] \tag{3}$$
$$+ \sum_{k \in N} \left[\varsigma_1 \left[\frac{|f_{t,t',k,k}|}{f_k^{\max}} - 1\right]^+ + \varsigma_2 \left[\frac{|f_{t,t',k,k}|}{\delta_\varsigma f_k^{\max}} - 1\right]^+\right],$$

where $\left[a\right]^+ = \max\left(a, 0\right)$.

The constraints of the optimization model are related to customer load satisfaction and the power flow in the primary distribution grid.

For every customer load b and every decision time step $t' \in \{t + \kappa_b, ..., t + T - 1\}$, the following constraints apply:

$$\sum_{l_b \in L_b} y_{t,t',b,l_b} = 1 \tag{4}$$

$$z_{t,t',b,l_b} = \begin{cases} z_{t,t'-1,b,l_b} + 1 & \text{if } y_{t,t'-1,b,l_b} = 1 \\ 0 & \text{if } y_{t,t'-1,b,l_b} = 0 \end{cases} \tag{5}$$

$$y_{t,t',b,l_b} = \begin{cases} 1 & \text{if } 1 \leq z_{t,t',b,l_b} < \rho_{b,l_b} \\ 0 \text{ or } 1 & \text{otherwise} \end{cases}. \tag{6}$$

Constraint (4) specifies that at a given time t' there should be only one level of curtailment l_b active for customer load b. Constraints (5) and (6) guarantee that once a customer load b enters a curtailment level l_b (including level 0, that is, no curtailment), it will stay at that level for the minimum required amount of time.

For every cable section $i \leftrightarrow j$ ($\forall i \in M, \forall j \in M_{i \to}$) in the primary network and every time step $t' \in \{t,...,t+T-1\}$ we have:

$$f_{t,t',i,j} = \left(\theta_{t,t',i} - \theta_{t,t',j}\right)\eta_{t,t',i,j}(\hat{q})\big/X_{i,j} \ . \qquad (7)$$

Constraint (7) relates the flow of real power on a cable section to the state of the cable at time t' (whether it is active or not) and to the phase angles at the adjoining nodes (through the DC optimal power flow approximation). Note that $\eta_{t,t',i,j}(\hat{q})$, if nonzero, will take on the appropriate voltage-based value.

For every node $i \in M$ in the primary network and every time step $t' \in \{t,...,t+T-1\}$ the flow conservation constraint for real power is given by:

$$p_{t,t',i} + \sum_{j \in M_{\to i}}\left(f_{t,t',j,i} - \varphi^{Loss}R_{j,i}\right) = \sum_{j \in M_{i\to}\cup\{i\}} f_{t,t',i,j} \ , \qquad (8)$$

where the left-hand term represents flow *into* the node (after subtracting for losses in the cables) and the right-hand term represents flow *out of* the node (the sign of the flows being relative to the conventional direction). Note that a transformer is actually represented in the network by a node with two "sides:" the *high voltage* side and the *low voltage* side. Constraint (8) applies to the high voltage side of a transformer node.

The flow conservation constraint at the low voltage side of transformer node $i \in N$, at time step $t' \in \{t,...,t+T-1\}$ is given by:

$$f_{t,t',i,i} = \psi_i^{PF}\sum_{b \in B} A_{t,t',i,b}(\hat{q})\hat{D}_{t,t',b}\sum_{l_b \in L_b} l_b y_{t,t',b,l_b} \ , \qquad (9)$$

where customer load b is connected to the primary network through node i at time t' if $A_{t,t',i,b}(\hat{q}) > 0$; otherwise, it is not. Note that in this model *all* customer loads are being connected directly to the transformers, rather than to the secondary network (where most would have been actually connected to). This modeling approximation is being used because incorporating the secondary network in this optimization model would significantly increase the computational burden.

Finally we have:

$$y_{t,t',b,l_b} \in \{0,1\}; \ z_{t,t',b,l_b} \geq 0, \text{ integer}; \ p_{t,t',i} \geq 0 \ . \qquad (10)$$

D. Optimization Policies

Assume for now that we *can* sequentially solve the optimization sub-problems described by (1), (4)-(10) for every $t \in \{0,1,...\}$, in a rolling horizon mode; we then take the partial solution $\{y_{t,t+\kappa_b,b,l_b}\}$ of each sub-problem t, and concatenate them into an approximate solution to the original stochastic problem over the desired planning horizon.

It turns out, however, that solving each MIP sub-problem described in (1), (4)-(10) over a typical planning horizon ($T \geq 15$ hours) is still impractical. As a result we propose three strategies to circumvent this issue.

The first strategy is obtained by simply reducing the length of the planning horizon over which (1) is defined. Let $H << T$ be the length of a much shorter *decision horizon* over which we will solve each sub-problem. Note that it is advisable that $H > \max_{b \in B}\kappa_b$ (κ_b is the minimum curtailment notification time of load b), so that every curtailable load may

have a chance to be curtailed in the solution of every sub-problem. The optimization sub-problem at time t now becomes:

$$F_t^{RH}\left(S_t\right) = \max_{u_t}\left[\sum_{t'=t}^{t+H-1}\left(R_{t,t'}\left(y_{t,t'},p_{t,t'}\right) - P_{t,t'}\left(f_{t,t'}\right)\right)\right] \qquad (11)$$

along with constraints (4)-(10), where every occurrence of T is replaced by H. We solve this reduced-size sub-problem for every $t \in \{0,...,T-1\}$. This means that in practice this strategy spans the time interval $\{0,...,T+H-2\}$.

The optimization sub-problems described by (11), (4)-(10), with T replaced by H, when solved sequentially for every $t \in \{0,...,T-1\}$ constitute the first of our three approximate policies to solve the load curtailment problem over the desired planning horizon. This policy is called a deterministic rolling-horizon look-ahead procedure, henceforth referred to as "RH" policy. Rolling horizon policies such as the one described above are popular in the engineering community, but they assume a single, deterministic future.

Our main interest, however, is to propose approximate policies that yield robust solutions to the original problem, that is, curtailment plans that will work well, on average, over an as large as possible set of realizations of demand and feeder failures in the future. In order to attain that, we take the reduced-size optimization formulation (11), (4)-(10), with T replaced by H, and embed it in a Monte Carlo simulation to adaptively learn policies that allow us to guard against possible failures and uncertainty in the demand. We do this by running several iterations of the simulation over the horizon $\{0,...,T-1\}$, each with a different set of random realizations, in order to calibrate the approximate policies. Once these are calibrated, we can then use them to produce robust solutions to the underlying stochastic problem. The research question that arises is how robust these solutions are when compared to the solution produced by RH, vis-à-vis the additional computational burden imposed by their calibration through iterative simulation. Before we address this question in section IV, though, let us formally present the two ADP-based policies.

The first ADP procedure, known as a *cost function approximation*, is based on rounding the fractions of customer loads that were effectively curtailed over the calibration iterations. For each customer b we record the curtailment level observed at each time t in each simulation iteration. Assume that the average curtailment level computed at the end of the calibration procedure is $\tilde{v}_{t,b}$. Let $\tilde{l}_{t,b}$ be the closest curtailment level to $\tilde{v}_{t,b}$ in the finite set L_b. The optimization policy at sub-problem t is to curtail every customer load for which $\tilde{l}_{t+\kappa_b,b} > 0$. Modifying (11) to incorporate this policy yields:

$$F_t^{PFA}\left(S_t\right) = \max_{u_t}\left[\sum_{t'=t}^{t+H-1}\left(R_{t,t'}\left(y_{t,t'},p_{t,t'}\right) - P_{t,t'}\left(f_{t,t'}\right)\right)\right.$$
$$\left. y_{t,t+\kappa_b,b,\tilde{l}_{t+\kappa_b,b}} = 1, \forall b : \tilde{l}_{t+\kappa_b,b} > 0\right] \ . \qquad (12)$$

Equations (12), (4)-(10), with T replaced by H, for

$t \in \{0,...,T-1\}$ constitute a rolling horizon look-ahead procedure with cost function approximation (referred to as a "CFA" policy).

The second proposed ADP procedure is based on the use of *value function approximations*. We approximate the value of curtailing a customer load b at a given time t by a linear function of the curtailment level. We estimate the coefficient $\overline{v}_{t,b}$ of the linear function, by smoothing in observations of the marginal value $\hat{v}_{t,b}$ of curtailing load b at time t, computed at each iteration of the calibration procedure. These marginal values are computed through numerical derivatives, which involve modifying the observed level of curtailment of a load up or down (whichever produces the largest change) and then resolving the sub-problem. Let $\Delta l_{t,b}$ be the imposed change in the level of curtailment of load b at time t and $e_{t,t,b}$ be the corresponding change in $y_{t,t}$. Then, $\hat{v}_{t,b}$ is given by:

$$\hat{v}_{t,b} = \left[F_t^{\text{RH}}\left(S_t \mid y_{t,t} + e_{t,t,b} \right) - F_t^{\text{RH}}\left(S_t \right) \right] \times l_b^{\max} / \Delta l_{t,b} \ , \quad (13)$$

where $F_t^{\text{RH}}\left(S_t \mid y_{t,t} + e_{t,t,b} \right)$ represents the solution of the modified sub-problem and $l_b^{\max}$ is the maximum possible curtailment for load b. After n iterations of the calibration process, $\overline{v}_{t,b}^n$ is given by:

$$\overline{v}_{t,b}^n = \left(1 - \alpha_n \right) \overline{v}_{t,b}^{n-1} + \alpha_n \hat{v}_{t,b}^n \ , \quad (14)$$

where α_n is determined by a suitably chosen stepsize rule.

The value function approximations are incorporated into (11) resulting in the following formulation:

$$F_t^{\text{VFA}}\left(S_t \right) = \max_{u_t} \sum_{t'=t}^{t+H-1} \left(R_{t,t'}\left(y_{t,t'}, p_{t,t'} \right) - P_{t,t'}\left(f_{t,t'} \right) \right. \\ \left. + \sum_{b \in B} \overline{v}_{t',b} \sum_{l_b \in L_b} l_b y_{t,t',b,l_b} \right) . \quad (15)$$

Equations (15), (4)-(10), with T replaced by H, for $t \in \{0,...,T-1\}$ constitute a rolling horizon look-ahead procedure with value function approximation, also known as a "VFA" policy.

IV. NUMERICAL EXPERIMENTS

Several experiments were designed to test the robustness of the solution to the load curtailment problem produced by the three approximate policies. We will report on two types of experiments, both involving uncertainty in the state of the components of the primary distribution grid, but not in the demand (we used point forecasts). The first set of experiments involved assuming that a known feeder had already failed before the starting time of the planning horizon (0) and that there would be at least one additional, unknown feeder surely failing at the first time step of the planning horizon. The second type of experiment involved assuming a known feeder failure before the starting time and a given likelihood of another known feeder failing at any time during the planning horizon ($t \in \{0,...,T-1\}$). In each experiment, we compared the solution produced by the deterministic rolling-horizon look-ahead procedure (RH) to those produced by the ADP-

based procedures (CFA and VFA) to that produced by a *baseline* solution (obtained by curtailing the loads of all curtailable customers to the maximum possible amount). The comparisons for each experiment were made by simulating the power flow for each solution over the set of most likely feeder failure scenarios in that experiment and computing a number of performance statistics for each scenario. Unlike the optimization problems which involved only the primary distribution network, these power flow simulations included the secondary network too. For each experiment, we report on the worst-case results and average results over all scenarios. The statistics include: (i) the value of the objective function (the overall utility function); (ii) the percentage of load curtailment; and (iii) the percentages of primary grid components (transformers and cable sections) whose flows exceed the ratings.

A. Experimental Setting

The distribution network used in the experiments reported in this paper was derived by combining actual and synthetic data for the distribution network of a section of a large city in the United States. The primary network is composed of 24 substations (and respective feeders), 725 transformers and 3562 cable sections, running at the 27kV voltage level. The secondary mesh, originally composed of 11496 nodes and 13245 links, was simplified to a reduced network with 3681 nodes and 4878 links. It runs predominantly at the 120V level. The primary and secondary networks are connected through the transformers.

The customer loads in the section of the city were aggregated by the areas in the neighborhood of each transformer (essentially because that was the level at which historical load data was available). This resulted in a pool of 688 aggregate loads, each associated to basically one primary transformer, but a few of them to more than one. One hundred of these aggregate loads were randomly selected to be curtailable. We used historical aggregate load data collected for the summer of 2010 and an adapted damped trend multi-calendar exponential smoothing model [29] to forecast hourly aggregate loads during the desired planning horizon. In this set of experiments we used point forecasts for the loads, thus eliminating uncertainty in the demand. We scaled up the forecasts by 35%, so that the power flow on the distribution grid became near capacity.

Three levels of curtailment ($l_b \in L_b$) were associated to each curtailable customer: 0% (no curtailment), 50% and 100% (total curtailment). Each customer was assumed to require a minimum curtailment notification time (κ_b) of 2 hours, and each load was expected to remain in a given state (curtailed or not) for at least 4 hours (ρ_{b,l_b}).

Since the optimization models included only the primary distribution network (substations, transformers and cable sections), the aggregate loads were connected directly to the respective primary transformers. For the loads that are in reality connected to the network through the secondary grid (that is, the majority), we developed an approximate algorithm

to re-aggregate a load to nearby transformers when one or more of its primary transformers goes out of service because a feeder is out (this corresponds to the computation of the parameters $A_{t,t',i,b}(\hat{q})$ in (8)). This algorithm uses empirical data from AC power flow simulations of 1-contingencies in the primary network. The algorithm was embedded in the generation of the optimization models, within the simulation. The load re-aggregation algorithm is *not* necessary when the secondary network is added to the model, since in this case most of the loads connect to the distribution network through the secondary mesh. This is the case, for instance, when we are evaluating the robustness of a given curtailment plan through the simulation of the power flow for different scenarios of primary feeder failures. This evaluation is performed on a distribution network that includes both the primary and the secondary grids.

We assumed the time between failures in a feeder to have a Weibull distribution with given expected value and standard deviation, modified by an empirically computed factor reflecting the proximity of other failed feeders. In other words, feeders in the vicinity of a failed feeder are more likely to fail than feeders farther away. We used an expected time between failures equal to 80 days and a standard deviation equal to one third of the expected value. In the experiments in which we wanted to simulate a higher likelihood of a given feeder failing, we modified the expected time between failures for that feeder to achieve the desired failure rate. We further assume that once a feeder fails, it remains out for the remainder of the planning horizon.

We chose the length of the planning horizon (T) to be 15 hours, with a decision time step of 1 hour. In the rolling horizon decomposition, the decision horizon of each sub-problem (H) was set at 4 hours, and thus greater than the minimum curtailment notification time of any customer.

Both the CFA and VFA policies require the estimation of parameters before they can be used to generate a curtailment plan at a given time t, for the planning horizon T. This estimation is performed through several iterations of a Monte Carlo simulation, each with a different set of realizations of feeder failures. We run 100 iterations to calibrate each CFA policy, and 50 iterations to calibrate each VFA policy (these values were chosen based on empirical observations of the convergence rates of the calibration of different policies).

Finally, the parameters in the objective function terms (2) and (3) were set at the following values:

$$\beta = 1.1; \quad \lambda = 1.0; \quad \gamma_1 = 140; \quad \gamma_2 = 420; \quad \delta_\gamma = 1.05;$$
$$\varsigma_1 = 140; \quad \varsigma_2 = 420; \quad \delta_\varsigma = 1.05 . \tag{16}$$

These values were chosen based on some limited experimentation. We recognize that a more formal sensitivity analysis would be recommended, but, given that the primary goal of this paper is to introduce a new methodology, we believe that such an analysis is beyond the scope of the paper.

The parameters φ^{Loss} in (8) and ψ_i^{PF} in (9) were empirically computed from AC power flow simulations of the primary network.

The MIP problems were solved using IBM CPLEX v.12. The maximum number of threads available to CPLEX was limited to 8 (with "deterministic" parallel mode). The tolerance for the integer gap was set at 10^{-4} and the integer precision tolerance was set at 10^{-6}.

B. Experimental Design

The goal of these experiments was to assess the robustness of the curtailment plans produced by three rolling horizon approximate policies: RH, CFA and VFA. Given that we cannot find the optimal solution to the underlying stochastic optimization problem, we propose a *most robust* policy, in terms of worst case performance, obtained by curtailing all curtailable customers by the maximum amount possible, as a benchmark to compare the plans obtained through the three approximation policies. We call this benchmark policy a baseline plan and henceforth refer to it as the "Base" policy.

We designed five experiments under two broad categories. In both categories we begin with an initial state (at time 0) of the distribution network in which it is known that a given primary feeder has failed (and will remain out throughout the planning horizon). In experiments #'s 1, 2 and 3, we assume further that at least one more feeder will fail at the very beginning of the planning horizon (time 1), but we have no indication of which feeder that might be. Experiments 1, 2 and 3 differ from each other in the feeder that is known to have failed beforehand. These feeders will be generically referred to as A, B and C, respectively, and they were chosen so as to cover different areas of the distribution network. Fig. 1 shows the basic layout of the optimization network, with feeders A, B and C highlighted.

Fig. 1. The optimization network: primary feeders depicted by the solid lines, and aggregate customer loads represented by the dots. Prominent feeders in the numerical experiments: A (bottom-right, in blue), B (top-left, in purple) and C (center, in green).

In experiments #'s 4 and 5, we still have a known feeder failure at time 0 (feeder A), but now, instead of having the information that an unknown feeder will fail at time 1, we are given the information that feeder B has a higher than usual likelihood of failing at some time $t \in \{1,...,T-1\}$ and we can

estimate this likelihood (in experiment 4 it is around 50% and in 5, around 100%).

In each one of the experiments, the curtailment plan corresponding to the approximate policies needs to be computed, before it can be evaluated for robustness against the baseline plan. For the RH policy, computing the plan is trivial and involves running a single iteration of the deterministic rolling-horizon simulation. The CFA and VFA policies require calibration first, through Monte Carlo simulation (though these calibrations can be performed offline). Once calibrated, the CFA and the VFA curtailment plans can be computed with a single deterministic iteration too.

In order to evaluate and compare the four curtailment plans for each experiment, we designed a simulation-based evaluation procedure. It consists of generating the most likely scenarios (realizations) of feeder failures over the planning horizon for each experimental setting, and then running a power flow simulation of each curtailment plan for each scenario. For each power flow simulation, we estimate the overall objective function value [F(S)], compute the total curtailed (or dropped out) power (as a percentage of the total demand) [% CURT], and compute the percentages of primary grid components (transformers and cable sections) for which the power flow exceeds the ratings (note that the latter is *not* an estimate of the amount of power in MW above the ratings, but an estimate of the number of components above the ratings) [% TRANSF and % CABLE OVER]. We then report both the worst case and the average performance for each of these statistics and for each of the approximate policies, over all the scenarios in each experiment. We generate the 50 most likely scenarios in each experimental setting by simulating feeder failures over the planning horizon.

C. Experimental Results

Table I shows the worst case performance results, while Table II shows the average case performance results for experiment #1, where feeder A is assumed to have failed before the start of the planning horizon, and at least one other unknown feeder will fail shortly after. The objective function values were normalized so that the Base results are equal to 100. As expected, the Base policy yields the most robust plan (highest objective function value and lowest percentage of transformers above the ratings) in terms of worst case performance, but the percentage of total demand that has been curtailed is also significantly higher.

It is also interesting to compare the approximate policies among themselves. VFA yields the overall most robust solution among the three, in the worst case performance, by striking a balance between the percentage of power curtailed and the percentage of transformers with power above the ratings. CFA curtails less, but has more transformers violating ratings in a way that makes it less "optimal." And finally RH underperforms both.

TABLE I
EXPERIMENT #1 – WORST CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	97.6	3.0	21.5	0.9
CFA	95.5	2.1	22.1	0.9
RH	94.7	2.0	22.1	0.9
BASE	100	12.5	18.1	0.9

In the average case performance (Table II), VFA, CFA and RH all outperform Base. This is consistent with the fact that those three policies are approximations of an optimization policy that maximizes the expected value of the objective function (that is, they are supposed to maximize average performance).

TABLE II
EXPERIMENT #1 – AVERAGE CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	109.9	1.7	13.1	0.2
CFA	108.9	0.7	14.0	0.2
RH	108.6	0.7	14.2	0.2
BASE	100	11.2	11.9	0.1

Note that the percentage of curtailed power also includes the loads that have been dropped out of service because they belong to the (small) group of loads that are exclusively connected to primary feeders or to spot networks, and not to the secondary network. Therefore, when the primary feeder(s) to which they are connected fail(s), they cannot be served. Given that worst case performances generally correspond to scenarios with more feeder failures, it is thus expected that, for any given policy, the percentage of curtailed power will be greater for the worst case performance than for the average one.

Tables III and IV present similar results for experiment #2 where feeder B is assumed to have failed before the start of the planning horizon.

TABLE III
EXPERIMENT #2 – WORST CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	94.3	2.8	24.0	1.0
CFA	91.5	2.3	24.3	1.0
RH	91.4	2.3	24.3	1.0
BASE	100	12.5	20.1	0.9

TABLE IV
EXPERIMENT #2 – AVERAGE CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	110.1	1.4	15.7	0.2
CFA	109.9	1.0	16.0	0.2
RH	109.7	1.0	16.0	0.2
BASE	100.0	11.1	14.2	0.1

Tables V and VI depict the worst case and the average case results for experiment #3, where feeder C is assumed to fail at the start of the planning horizon. Note that the worst case performance of the VFA policy is markedly better than those of the CFA and RH policies (Table V), but not on the average

case (Table VI).

The results of experiments #'s 1, 2 and 3 indicate that the VFA policy outperforms the CFA policy by a slight margin, and the latter outperforms the RH policy also by a slight margin, both in the worst and the average performance cases. In terms of the computational effort involved, the RH policy does not require any calibration, whereas both VFA and CFA do. Moreover, in the current implementation, the calibration of the VFA policy takes over ten times longer than that of the CFA policy, though, as indicated before, both calibrations can be performed offline.

TABLE V
EXPERIMENT #3 – WORST CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	84.2	4.6	23.2	1.1
CFA	72.6	3.6	24.3	1.2
RH	71.9	3.5	24.3	1.2
BASE	100.0	13.1	19.9	0.9

TABLE VI
EXPERIMENT #3 – AVERAGE CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	108.3	2.9	15.2	0.2
CFA	108.3	1.9	15.8	0.3
RH	108.1	1.8	16.0	0.3
BASE	100.0	11.4	13.9	0.1

Experiment #4 is reported in Tables VII and VIII, whereas experiment #5 is reported in Tables IX and X. In both experiments feeder A is assumed to have failed before the start of the planning horizon, and the likelihood of feeder B failing at some time during the planning horizon is higher than usual (about 50% in experiment #4 and about 100% in experiment #5).

TABLE VII
EXPERIMENT #4 – WORST CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	85.2	1.9	19.9	0.9
CFA	83.6	1.6	20.5	0.9
RH	83.2	1.6	20.5	0.9
BASE	100	12.1	17.5	0.5

TABLE VIII
EXPERIMENT #4 – AVERAGE CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	109.7	0.9	14.1	0.1
CFA	109.5	0.6	14.7	0.1
RH	109.4	0.6	14.7	0.1
BASE	100	11.1	12.4	0.0

One particularly noteworthy aspect is the poorer results of the three approximation policies in the worst case performance of experiment #5 (Table IX). In terms of the overall objective function, they are more than 30% below the Base results. This is the case where feeder B will almost surely fail (in addition to A) at some time during the planning horizon (these two feeders are in opposite sections of the distribution area). This observation may indicate that the tuning of the values in (16), used for the parameters in the objective function terms, may be dependent on the topology of the failing feeders. Equivalently, the observed performance may have been expected because the underlying stochastic optimization uses an unconditional expectation, whereas these experiments have been made using an expectation conditional on the events that one or more specific feeders will fail with certainty.

Overall, the results for experiments #'s 4 and 5 are consistent with the results from the previous experiments. Thus the observations drawn from those experiments about the relative performance of the three approximate policies still hold, including the issue of the trade-off between the increased robustness of the curtailment plans and the computational burden of calibrating the VFA and the CFA policies that produce them.

TABLE IX
EXPERIMENT #5 – WORST CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	68.0	2.2	20.7	1.0
CFA	66.7	1.8	21.2	1.0
RH	62.4	1.6	21.5	1.0
BASE	100	12.1	18.2	0.8

TABLE X
EXPERIMENT #5 – AVERAGE CASE PERFORMANCE

POL	F(S)	% CURT	% TRANSF OVER	% CABLE OVER
VFA	111.6	1.1	14.0	0.1
CFA	111.1	0.7	14.7	0.1
RH	109.8	0.5	15.0	0.1
BASE	100	11.0	12.8	0.0

V. CONCLUSION

We showed in this paper how to formulate the problem of planning short-term load curtailment in a densely populated urban area, in the presence of uncertainty in demand and in the state of the distribution grid, as a stochastic mixed-integer optimization problem. We proposed three rolling-horizon look-ahead policies to approximately solve the optimization problem, one of which is deterministic and two of which are based on approximate dynamic programming techniques. Finally, we demonstrated through numerical experiments (involving uncertainty in the grid only) that the ADP-based policies yield curtailment plans that are more robust on average than the deterministic policy, but at the expense of the additional computational burden needed to calibrate the ADP-based policies. The VFA policy outperforms the CFA policy by a small margin, but requires a significantly longer calibration effort. We also showed how the worst case performance of the three approximation policies compares with a baseline policy where all curtailable loads are curtailed to the maximum amount possible.

REFERENCES

[1] W. B. Powell, Approximate Dynamic Programming: Solving the Curses of Dimensionality, 2nd ed. New York: Wiley, 2011.

[2] N. P. Padhy, "Unit commitment-a bibliographical survey," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 1196-1205, May 2004.

[3] A. Y. Saber and G. K. Venayagamoorthy, "Intelligent unit commitment with vehicle-to-grid – A cost-emission optimization," *Journal of Power Sources*, vol. 195, no. 3, pp. 898-911, Feb. 2010.

[4] M. Pipattanasomporn, H. Feroze, and S. Rahman, "Multi-agent systems in a distributed smart grid: Design and implementation," in *Proc. of the 2009 IEEE/PES Power Systems Conference and Exposition*, pp. 1-8, March 2009.

[5] W. Jia, C. Kang, and Q. Chen, "Analysis on demand-side interactive response capability for power system dispatch in a smart grid framework," *Electric Power Systems Research*, vol. 90, pp. 11-17, Sep. 2012.

[6] D. Divan, "Smart distributed control of power systems," in Proc. of the 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, pp. 1-3, July 2008.

[7] D. Divan and H. Johal, "Distributed FACTS - A new concept for realizing grid power flow control," *IEEE Trans. On Power Electronics*, vol. 22, no. 6, pp. 2253-2260, Nov. 2007.

[8] J. A. Momoh, "Optimal methods for power system operation and management," in *Proc. of the 2006 IEEE PES Power Systems Conference and Exposition*, pp. 179-187, Oct. 2006.

[9] P. J. Werbos, "Putting more brain-like intelligence into the electric power grid: What we need and how to do it," in *Proc. of the 2009 International Joint Conference on Neural Networks*, pp. 3356-3359, June 2009.

[10] G. Venayagamoorthy, "Dynamic, stochastic, computational, and scalable technologies for smart grids," *IEEE Computational Intelligence Magazine*, vol. 6, no. 3, pp. 22-35, Aug. 2011.

[11] M. Caramanis and J. M. Foster, "Management of electric vehicle charging to mitigate renewable generation intermittency and distribution network congestion," in *Proc. of the 48th IEEE Conference on Decision and Control*, pp. 4717-4722, Dec. 2009.

[12] C. Battistelli, L. Baringo, and A. J. Conejo, "Optimal energy management of small electric energy systems including V2G facilities and renewable energy sources," *Electric Power Systems Research*, vol. 92, pp. 50-59, Nov. 2012.

[13] U.S. Department of Energy. (2006, Feb.) Benefits of demand response in electricity markets and recommendations for achieving them: report to U.S. Congress pursuant to section 1252 of the Energy Policy Act of 2005. Staff Report, Washington, D.C. Available: http://eetd.lbl.gov/ea/EMP/reports/congress-1252d.pdf

[14] K. Spees and L. B. Lave, "Demand response and electricity market efficiency," *The Electricity Journal*, vol. 20, no. 3, pp. 69-85, Apr. 2007.

[15] Federal Energy Regulatory Commission. (2009, June). A national assessment of demand response potential. Staff Report, Washington, D.C. Available: http://www.ferc.gov/legal/staff-reports/06-09-demand-response.pdf

[16] C. Goldman, M. Reid, R. Levy, and A. Silverstein, "Coordination of energy efficiency and demand response," Lawrence Berkeley National Laboratory, Berkeley, CA, LBNL Paper LBNL-3044E, June 2010.

[17] H. Hahn, S. Meyer-Nieberg, and S. Pickl, "Electric load forecasting methods: Tools for decision making," *European Journal of Operational Research*, vol. 199, no. 3, pp. 902-907, Dec. 2009.

[18] J. Y. Fan and J. D. McDonald, "A real-time implementation of short-term load forecasting for distribution power systems," *IEEE Trans. Power Syst.,* vol. 9, no. 2, pp. 988-994, May 1994.

[19] M. Y. Cho, J. C. Hwang, and C. S. Chen, "Customer short term load forecasting by using ARIMA transfer function model," in *Proc. of EMPD '95., 1995 International Conference on Energy Management and Power Delivery, 1995.,* pp. 317-322 vol.1, Nov. 1995.

[20] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Trans. Power Syst.,* vol. 16, no. 1, pp. 44-55, Feb. 2001.

[21] B.-J. Chen, M.-W. Chang, and C.-J. lin, "Load forecasting using support vector Machines: a study on EUNITE competition 2001," *IEEE Trans. Power Syst.,* vol. 19, no. 4, pp. 1821-1830, Nov. 2004.

[22] L. Wu, G. Kaiser, D. Solomon, R. Winter, A. Boulanger, and R. Anderson, "Improving efficiency and reliability of building systems using machine learning and automated online evaluation," in *Proc. 2012 IEEE Long Island Systems, Applications and Technology Conference (LISAT),* pp. 1-6, May 2012.

[23] G. A. Godfrey and W. B. Powell, "Adaptive estimation of daily demands with complex calendar effects for freight transportation," *Transportation Research Part B: Methodological,* vol. 34, no. 6, pp. 451-469, Aug. 2000.

[24] P. Gross, A. Boulanger, M. Arias, D. Waltz, P. M. Long, C. Lawson, R. Anderson, M. Koenig, M. Mastrocinque, W. Fairechio, J. A. Johnson, S. Lee, F. Doherty, and A. Kressner, "Predicting electricity distribution feeder failures using machine learning susceptibility analysis," in *Proc. of the 18th Innovative Applications of Artificial Intelligence Conf.*, vol. 2, pp. 1705-1711, July 2006.

[25] C. Rudin, D. Waltz, R. N. Anderson, A. Boulanger, A. Salleb-Aouissi, M. Chow, H. Dutta, P. N. Gross, B. Huang, S. Ierome, D. F. Isaac, A. Kressner, R. J. Passonneau, A. Radeva, and L. Wu, "Machine learning for the New York City power grid," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, pp. 328-345, Feb. 2012.

[26] B. Stott, J. Jardim, and O. Alsaç, "DC power flow revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290-1300, Aug. 2009.

[27] K. W. Hedman, R. P. O'Neill, E. B. Fisher, and S. S. Oren, "Smart flexible just-in-time transmission and flowgate bidding," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 93-102, Feb. 2011.

[28] T. J. Overbye, X. Cheng, and Y. Sun, "A comparison of the AC and DC power flow models for LMP calculations," in *Proc. 37th Hawaii Int. Conf. System Sciences*, 2004.

[29] H. B. Jeong, "Approximate dynamic programming for the stochastic load curtailment problem," Senior thesis, Dept. of Oper. Res. and Finan. Eng., Princeton Univ., Princeton, 2011.

[30] Mulvey, J. M., Vanderbei, R. J., & Zenios, S. A. (1995). Robust Optimization of Large-Scale Systems. *Operations Research*, *43*(2), 264–281.

[31] Bertsimas, D., & Sim, M. (2004). The Price of Robustness. *Operations Research*, *52*(1), 35–53. doi:10.1287/opre.1030.0065

[32]

Hugo P. Simão received the Ph.D. degree in civil engineering and operations research from Princeton University. Currently he is a Senior Operations Research Engineer and deputy director of the Princeton Laboratory for Energy Systems Analysis at Princeton University, in Princeton NJ.

Hyun Bin (Vince) Jeong received the B.S. degree in operations research and financial engineering from Princeton University. Currently he works as a business analyst at McKinsey & Company, in Boston MA.

Boris Defourny received the B.S. degree in electrical engineering and the Ph.D. degree in applied sciences from the University of Liege, in Belgium. He is now an Associate Professional Specialist in the Department of Operations Research and Financial Engineering at Princeton University, in Princeton NJ.

Warren B. Powell is a Professor in the Department of Operations Research and Financial Engineering at Princeton University, where he has taught for over 30 years. He is director of the Princeton Laboratory for Energy Systems Analysis, in Princeton NJ.

Albert Boulanger received the B.S. degree in physics from the University of Florida, Gainesville FL, and the M.S. degree in computer science from the University of Illinois, Urbana-Champaign IL. He is currently a Senior Staff Associate at the Center for Computational Learning Systems, Columbia University, in New York NY, where he has worked for the past 18 years.

Ashish Gagneja received the M.S. degree in computer science from Columbia University. Currently he works as a Staff Associate at the Center for Computational Learning Systems, Columbia University, in New York NY.

Leon Wu received the M.S and M.Phil. degrees in computer science from Columbia University, where he is currently a Ph.D. candidate in the Dept. of Computer Science. He is also a Senior Research Associate at the Center for Computational Learning Systems, Columbia University, in New York NY.

Roger N. Anderson received the Ph.D. degree from the Scripps Institute of Oceanography, the University of California at San Diego. He is a Senior Research Scholar at the Center for Computational Learning Systems, Columbia University, in New York NY.