Exploring a Direct Policy Search Framework for Multiobjective Optimization of a Microgrid Energy Management System

Amandeep Gupta	Mengwei Liu	David Gold	Patrick Reed	C. Lindsay Anderson
Cornell University	Cornell University	Cornell University	Cornell University	Cornell Univesrity
ag729@cornell.edu	ml2589@cornell.edu	dfg42@cornell.edu	patrick.reed@cornell.edu	cla28@cornell.edu

Abstract

With an increasing focus on integration of distributed energy resources, it is likely that microgrids will proliferate globally. These microgrid systems will be expected to achieve multiple stakeholder objectives, motivating the study of microgrid operations using a multiobjective framework. A multiobjective perspective has the potential balance the trade-offs implicit to efficient use of available resources. To address this challenge, this paper proposes a simulation based parametric approach for multiobjective optimization for microgrid energy management. The methodology generates a Pareto-approximate set of control policies, to provide a microgrid controller with diverse alternative strategies for utilizing resources to balance competing objectives. The policies also help to illustrate the complex relationships between the objectives, and the consequences of compromises across performance. The methodology is implemented on a test microgrid and the potential benefits are demonstrated with a set of illustrative case studies.

1. Introduction

Microgrids (MGs) are receiving increased attention as an effective mechanism for distributed energy and renewable resource integration, with additional benefits attainable through increasing coordinated use of the available resources in the system. Microgrids (MGs) are generally considered autonomous networks connected as single entities to the distribution or transmission grid, which are capable of buying and selling energy based on need[1]. These networks usually consist of local generators, storage units, renewable and traditional generation, and dispatchable/controllable loads. These distributed energy resources (DERs) can be operated for local electricity provision in an "islanded" mode, or with energy exchange capability between the microgrid and distribution/transmission, when the connection with the grid is active [1, 2].

A key research challenge for microgrids requires solving for the best set of sequential decisions (amount of generation or energy exchange with the grid at any given hour) over a planning horizon, which optimize predefined objectives while satisfying system constraints. The majority of existing research in this area has concentrated on minimizing the cost of energy, while maintaining stable operations in either grid connected or islanded mode. This approach neglects certain additional benefits available to MG users [3, 2]. As MGs become more sophisticated and ubiquitous, the stakeholders may want to utilize available DERs for more than minimizing cost of operation. As explored in [3, 4] MG energy management can be effectively utilized for different classes of objectives including capital and operations cost, environmental emissions, energy storage costs, and miscellaneous needs based objectives. Given that decision makers will want to use the MG for different purposes, it supports the assertion that MG operation should be studied in a multiobjective optimization (MO) context [3, 4].

To this effect there has been a shift towards optimizing multiple objectives when solving the energy management problem for MGs [5, 6, 7]. One MO approach explored in the literature includes weighted sums (or scalarization) based on assigning weights to each objective and converting the problem into a single objective formulation [5, 8]. The single-objective optimization is then repeated using different scalarization values [9] to obtain Pareto optimal solutions. This process assumes availability of a consensus understanding of the consequences of alternative preference of weighting schemes *a priori*, which is prone to ignoring complex relationships between the objectives[10, 2, 5].

Alternative approaches use iterative population-based solution tools such as evolutionary algorithms with the intent of approximating the full Pareto front, so that decisions can be made with improved knowledge of the implications of performance tradeoffs. This approach to multiobjective formulation

HⁱCSS

is termed a posterior analysis, as it does not assume prior information about objectives or stakeholder preferences. Emerging solution methods include genetic algorithms [6, 11], a cross entropy approach as described in [7], and variations of particle swarm optimization [12, 13, 14]. Even though the numerical techniques differ, multi-objective evolutionary algorithms (MOEAs) all evolve approximations to the full Pareto-front, which is a set of non-dominated solutions where improvement in performance in any single objective degrades performance in one or more of the remaining objectives [15, 2]. However, best practices for using meta-heuristic search methods, require a careful consideration of how decisions are represented during search (i.e., compact and effective abstractions of decision variables)[16].

Evolutionary multiobjective direct policy search (EMODPS) is a policy approximation abstraction of control problems[17]. The approach exploits MOEAs to search a parameterized space of candidate control policies in a manner that maximizes the efficiency and effectiveness of the algorithm's global search features [10]. EMODPS has been adopted in water resources literature to solve large-scale reservoir control problems[18]. This set of problems require solving for a sequence of decisions over time similar to MG control problem in power systems [10, 19]. The direct policy search (DPS) approach, utilized in EMODPS, falls under the umbrella of approximate dynamic programming class of problems[20]. Unlike traditional dynamic programming, DPS, solves for a parameterized control policy that maps state of the system to decision space, thus compacting and reducing the number of decision variables. Single objective formulations for DPS have been shown to perform well for an energy system with storage[21]. The EMODPS approach implements a multiobjective DPS formulation which can then be solved using MOEAs.

In this paper, a novel methodology is implemented, utilizing EMODPS, to solve a MG energy management problem to explore the potential benefits of the approach. The formulation is implemented on a small MG while considering three stakeholder objectives (revenue, emissions, demand response). A set of case studies have been performed to demonstrate the following potential benefits of the formulation:

- consideration of microgrid energy management problem in a multiobjective context to broaden the scope of performance concerns that can be considered

- an extensible control framework that can incorporate a wide array of information sources,

and

- provision a set of competing policy solutions which can provide a set of alternative control policies that can provide stakeholders with diverse set of strategies to utilize resources in the microgrid.

The paper is organized as follows: Section 2 provides an overview of the EMODPS approach, followed by a detailed description of the formulation for MG energy management in Section 3. The performance of the approach is illustrated in Section 4 followed by concluding remarks in Section 5.

2. Evolutionary Multiobjective Direct Policy Search

EMODPS is a simulation based optimization approach for exploring a parameterized space of candidate control policies, which has proven to be an efficient methodology for application in problems where regular stochastic dynamic programming approaches struggle. This methodology accomplishes this by; parameterizing the decision space which reduces curse of dimensionality [10], directly incorporates simulation models which allows for more flexible formulation [10], and allows the users to explore multiple objectives without *a priori* assumptions about preferences [22]. The EMODPS framework has two main components: (1) Direct Policy Search and; (2) Multiobjective Evolutionary Algorithms.

2.1. Direct Policy Search

Direct Policy Search (DPS) is a control strategy that directly searches in the solution space [23]. DPS is based on defining a parameterized function(policy) which maps the system state(s) to the decision variables, which is followed by exploration of the parameter space to find a policy that optimizes the objectives under consideration. In other words, the parameters are optimized rather than the decisions themselves. The result of the optimization, therefore, are not a set of hourly decisions as would be the case in traditional MG operation formulations but a control policy that can be used to guide those decisions. Using the example from [24], we can describe our system in a state S_t, from which we take an action x_t and then observe new information which takes us to a new state S_{t+1}. DPS can then represent the rule(or policy) for making this decision using the function $F(S_t)$. We, generally, also have a system model S^M that describes how the system evolves from S_t to S_{t+1}. The dynamics of our problem can then be described as

$$x_t = F(S_t)$$

$$S_{t+1} = S^M(S_t, x_t)$$

The function $F(.)$ is the operating policy approximating the relationship between the state and decision variables by learning from the feedback of the system simulations. It can vary from being a linear function for simple systems with single objective problems to more flexible mappings in complex systems (like Artificial Neural Networks) where high number of parameters might be required to avoid restricting the search for the optimal policy to a subspace of decision space [10]. Radial basis functions (RBFs) have proven to be efficient universal approximators and are widely adopted in many applications [15]. For the purpose of this paper, RBFs have been used as the approximating function with further detail in Section 3.5.

2.2. Multiobjective Evolutionary Algorithms

The EMODPS approach utilizes MOEAs to search the parameter space for the optimal values. MOEAs are iterative search algorithms that generate a Pareto-approximate set of solutions by mimicking the randomized mating, selection, and mutation operations that occur in nature [25]. A Pareto optimal set (or a non-dominated set) constitutes solutions for which improvement in any one objective would result in the degradation of one or more of the other objectives. The term Pareto-approximate set is the best known approximation of a Pareto optimal set when formal global convergence cannot be guaranteed. MOEAs provide a promising alternative to gradient-based optimization methods when solving multiobjective optimization problems, given their efficacy in dealing with multimodality, nonlinearity and stochasticity associated with such problems. There are different MOEAs available to be coupled with DPS approach, for the purpose of this study the Borg MOEA [25] has been utilized based on it's ability to obtain high quality tradeoff solutions [15, 19]. The Borg MOEA's success has been attributed to its use of ϵ-dominance archiving, ϵ-progress, and multiple self-adaptive search operators that combine to allow the algorithm to earn effective exploration strategies while solving challenging problems. [25, 15].

3. EMODPS for Microgrid Energy Management

A key research challenge for microgrids requires solving for the best set of sequential decisions (amount of generation or energy exchange with the grid at any given hour) over a planning horizon to optimize predefined objectives while satisfying system constraints. The majority of research in this area has concentrated on minimizing the cost of energy, while maintaining stable operations in either grid connected or islanded mode. These works, however, neglect certain additional benefits available to MG users [3, 2]. As MGs become smarter and more ubiquitous the stakeholders would want to utilize available DERs for more than minimizing cost of operation. As explored in [3, 4] MG energy management can be effectively utilized for different groups of objectives including capital and operations cost, environmental emissions, energy storage costs, and miscellaneous need based objectives. Given that decision makers would want to utilize the MG for different purposes, it stands to reason that MG operation should be studied in a multiobjective optimization (MO) context [3, 4].

In this study, we generate multiple MG energy management control strategies for generation and grid exchange decisions while balancing multiple stakeholder objectives. The proposed EMODPS methodology simulates daily operation of a MG over N stochastic samples of load, wind and solar while simultaneously optimizing three stakeholder objectives. This approach generates a Pareto-approximate set of control policies for conventional generation and grid exchange providing a diverse set of possible approaches for MG operation. These control policies use the battery storage levels to determine generation and buy/sell decisions with the grid.

For this demonstration, a test MG has been explored consisting of a battery, a diesel generator, critical and non-critical loads, and connection to the distribution grid based on the configuration details provided in Section 4.1. The MG is assumed to be community-owned with stakeholders that include residents and/or local businesses. In this case study, revenue, emissions and demand response use are included as objectives. The selection of revenue and emission objective represent two possible competing points of view for the stakeholders. Demand response is included as an objective in order to avoid setting a preference or value for DR *a priori*, but to explore the leverage available to the system with DR. In addition, DR can serve as a proxy for system reliability, in that the necessity of overuse of DR implies that they system may be under stress.

The following sections provide a comprehensive outline of the EMODPS approach starting with defining the objectives, followed by the constraints, the RBF policies and a description of the solution strategy.

3.1. Revenue Objective

The revenue objective is calculated as the expected value over N simulations of daily revenue as represented below.

$$O_1 = -\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} \bigg(C_{gen}(p_{i,t}) + \alpha_t(e_{i,t}) + C_{wt}(wt_{i,t}) $$
$$+ C_{pv}(pv_{i,t}) + C_{ES} |ES_{i,t} - ES_{i,t-1}| \bigg) \tag{1}$$

where $p_{i,t}$ and $e_{i,t}$ are generation and grid exchange decisions, $wt_{i,t}$ and $pv_{i,t}$ are the wind and solar generation, and $ES_{i,t}$ is the battery storage level for simulation i at time t. The exchange decision $e_{i,t}$ takes negative values when selling to the grid and positive when buying. C_{gen}, C_{wt} and C_{pv} are the cost for diesel, wind and solar generation while C_{ES} is the battery charging/discharging costs. α_t represents the real-time energy prices to buy or sell from or to the utility at time t.

3.2. Emissions Objective

Equation 2 defines the emissions objective which is the expectation of daily emissions over the N simulations. The formulation assumes emissions from diesel generation and any energy bought from the grid as shown below.

$$O_2 = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} \bigg(C_{ef}(max(e_{i,t}, 0) + p_{i,t}) \bigg) \tag{2}$$

where grid emissions are non-zero when $e_{i,t} > 0$ and C_{ef} represents the carbon emissions factor.

3.3. Demand Response

The objective is defined as reliability in the system under consideration and is utilized to explore the limits of demand response in a MG system. The structure of the objective is set to quantify load flexibility at each hour. The objective, O_3, is the expected percentage of hours in a day when demand response is required as shown below.

$$O_3 = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T} \sum_{t=1}^{T} \varphi_{i,t}$$
$$\varphi_{i,t} = \begin{cases} 1, & 1 - \frac{d_{i,t}^{DR}}{d_{i,t}} > d_{crit} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where $d_{i,t}^{DR}$ is the hourly flexible load for simulation i. This implies that any given hour in a day is considered reliable i.e. $\varphi_{i,t} = 1$ when $d_{i,t}^{DR}$ is less than a predefined critical threshold percentage, d_{crit}, of the hourly load $d_{i,t}$. For the purpose of this study d_{crit} was set to be 0.95 which implies that at any given hour a maximum of 5 percent of $d_{i,t}$ can be curtailed. .

3.4. Optimization Formulation

The three objectives are optimized simultaneously subject to constraints defined in equations (4) - (10). The optimization problem is formulated as described below.

$$\min_{\theta_p, \theta_e} \quad (-O_1, O_2, -O_3)$$
$$p = (p_{1,i}, p_{1,i}, ..., p_{T,i})$$
$$e = (e_{1,i}, e_{1,i}, ..., e_{T,i})$$

s.t.

$$ES_{i,t+1} = ES_{i,t} + wt_{i,t} + pv_{i,t} + e_{i,t} $$
$$+ p_{i,t} - (d_{i,t} - d_{i,t}^{DR}), \forall i \in N, t \in T \tag{4}$$

$$p_{i,t} - p_{i,t-1} \leq p_{rmp,up}, \forall i \in N, t \in T \tag{5}$$
$$p_{i,t-1} - p_{i,t} \leq p_{rmp,down}, \forall i \in N, t \in T \tag{6}$$
$$P_{min} \leq p_{i,t} \leq P_{max}, \forall i \in N, t \in T \tag{7}$$
$$E_{min} \leq e_{i,t} \leq E_{max}, \forall i \in N, t \in T \tag{8}$$
$$|ES_{i,t+1} - ES_{i,t}| \leq ES_{chg/dischg}, \forall i \in N, t \in T \tag{9}$$
$$ES_{min} \leq ES_{i,t} \leq ES_{max}, \forall i \in N, t \in T \tag{10}$$

where equation (4) is energy balance constraint for the system dependent on battery storage levels at any time t, equations (5) and (6) are the ramping constraints for the diesel generator, equations (7) and (8) are the upper and lower limits for the generator and energy exchange with the grid, equation (9) represents the limits on charging/discharging of the battery and equation (10) is the bounds on battery storage level. The formulation is solved to optimize the three objectives by finding the best set of parameters θ_p and θ_e for the operating policies p and e as explained in Section 3.5.

3.5. Formulation of Operating Policies

In this study cubic RBFs are implemented as the control policies described in section 2.1. The battery storage level and hour of the day t are the input variables representing the system state, which are then mapped to hourly diesel generation and buy/sell decisions as shown in equations (11) and (12).

$$p_{i,t} = \sum_{j=1}^{n} w_j \left(\left| \frac{ES_{i,t} - c_j}{r_j} \right| + x_t^2 + y_t^2 \right)^3, \ \forall t, i \tag{11}$$

$$e_{i,t} = \sum_{j=1}^{n} w_j \left(\left| \frac{ES_{i,t} - c_j}{r_j} \right| + x_t^2 + y_t^2 \right)^3, \ \forall t, i \quad (12)$$

where $x_t = sin(2\pi t/T - a_1)$ and $y_t = cos(2\pi t/T - a_2)$. $p_{i,t}$ and $e_{i,t}$ are the policy-prescribed generation and buy/sell decisions at hour t for sample i; $ES_{i,t}$ is the battery storage level at a given hour; n is the number of RBFs; x_t and y_t are phase shifted $sin(.)$ and $cos(.)$ functions for cyclic representation of time; w_j, c_j and r_j are the weights, centers and radii, respectively of jth RBF associated with the generation or buy/sell decision; and a_1, a_2 are the phase shifts on $[0, 2\pi]$.

As defined earlier, the goal of EMODPS is to find a non-dominated set of parameter vectors θ_p and θ_e minimizing the system objectives. Each of the parameter vectors are composed of the weights, centers and radii defining the RBF policies. For example, $\theta_p = [w_j, c_j, r_j]$ where j is the number of RBFs. In this study there is a set of parameter vectors each for the diesel generator and the utility buy/sell decisions. For each decision $n = 2$ RBFs, are used with parameter limits set as $w_j \in [0, 1]$, $c_j \in [-2, 2]$, $r_j \in (0, 2]$. The parameter vectors are then optimized with the help of MOEAs. These limits were selected after experimenting with multiple ranges and number of RBFs. The preferred ranges may differ they are dictated by the needs and the resources available to the operator. For example, setting a higher range translates into a larger decision space which will require more computation time to effectively explore, while setting a more restrictive range might result in less desirable results.

3.6.   Model Implementation

The EMODPS formulation begins by defining initial policies for the buy/sell and generation decisions, which are then used to simulate MG operation over N simulations. The resulting objective performance acts as a feedback for performance of the parametric policy. The BORG framework, as the MOEA solver, then uses this information to iterate over the parameter space of θ_p and θ_e. This process is repeated for multiple seeds and the final set of solutions is a compilation of non-dominated policies that represent best performance across the three objectives. The Borg MOEA was parameterized according to the user guidelines[25].

4.   Performance Evaluation

The goal of the following case studies is to demonstrate the utility of the multiobjective approach, through the EMODPS methodology to simulate and optimize microgrid control. Not only does this formulation provide the stakeholders with multiple control policies to achieve their objectives, but also allows for a deeper insight into the relationship between these objectives and the limits of the system. These insights then enable better control and utilization of the system resources.

4.1.   System Overview

To test the EMODPS approach, it is implemented on a test MG system consisting of a battery, a diesel generator, critical and non-critical loads, and connection to the distribution grid. The MG has a peak load of 220 kW with a battery size of 200 kW. The installed capcaity for wind and solar is set to be 20% and 10% with respect to the peak load. The load, wind, solar, and real-time market energy price distributions were obtained by projecting daily CAISO data on to the test system. For this study, load, wind, and solar are modeled as uncertain parameters with a maximum possible variation of 10% at any given hour. All other parameters used for the study are provided in the Appendix.

The DPS policies were optimized by simulating over $N = 1000$ randomly sampled scenarios of load, as well as wind and solar generation. The EMODPS approach generated 3000 pareto-approximate control policies that could be implemented to control the daily operations of the MG in consideration. To validate, each of these policies were then used to simulate MG operation over 1000 out-of-sample scenarios of load, wind and solar. For comparative purposes three separate energy management strategies were analyzed and validated as listed below.

- Perfect Information Strategy - Assuming availability of perfect forecast information about load, wind and solar. This strategy was used to provide a best performing baseline of hourly MG control decisions for the 1000 out-of-sample scenarios.

- Expected Forecast Strategy - This strategy optimized for a risk neutral energy management strategy based on the expected forecast distribution available for load, wind and solar, when determining the hourly control decisions.

- Conservative Strategy - This strategy decided hourly control decisions by assuming the worst-case scenario based on the available forecast.

The above mentioned control strategies were optimized as both single objective and scalarized multiobjective frameworks and compared with the DPS policies as explained in the following sections.

4.2. Need for Multiobjective Framework

This section demonstrates the need for considering multiple objectives by comparing performance of the EMODPS policies against the Perfect Information, Conservative and Expected Forecast strategies while optimizing the revenue objective only.

Figure 1 shows a projection of the revenue objective performance for the different formulations in consideration. Each point in the plot represents average objective performance over 1000 out-of-sample test scenarios. The figure shows performance of each of the 3000 DPS policies for the revenue objective in green while, those from the perfect information are colored pink, solutions from robust policy are colored blue, and those from expected policy, light-blue. As expected, the single-objective Perfect Information formulation dominates all others in Figure 1, followed by DPS, Expected Forecast and Conservative methods. The DPS policies provide a wide range of solutions for the objective although a majority of them are worse than the expected forecast strategy. Using the Perfect Information as a benchmark, this projection illustrates that while some of the DPS policies are doing better, in the single objective case the expected forecast strategy is the right approach, as it will allow the MG controller to follow a simple strategy with a small compromise in objective performance.
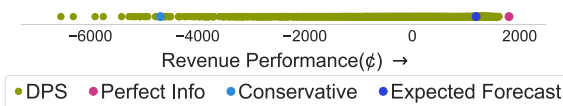

Figure 1. 1d comparison of solution performance for DPS policies, and single objective solutions for Perfect Information, Expected Forecast and Conservative strategies

The same solutions have been projected over two objectives in Figure 2. The x-axis represents the need for demand response (curtailable load) as percentage of total daily load and y-axis shows revenue while the arrows point towards direction of preference for each objective. Figure 2 provides more information and context than the one dimensional comparison. The perfect information solution still dominates all other formulations without any demand response requirement whereas the expected forecast strategy requires a relatively higher demand response to achieve it's revenue performance. Figure 2 provides a new perspective on the results from different formulations, as the stakeholders might no longer be satisfied with the performance of the expected forecast formulation,

and the conservative formulation may become a viable strategy as it does not require any demand response. However, the biggest improvement is for the DPS policies as there are multiple policies providing viable potential alternatives with high revenue performance and low demand response requirements.
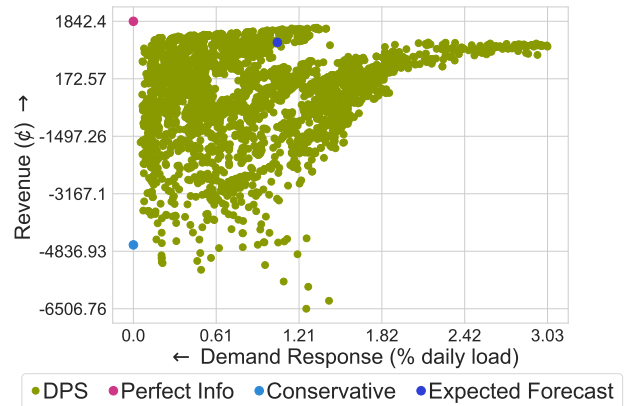

Figure 2. 2d comparison of solution performance for DPS policies, and single objective solutions for Perfect Information, Expected Forecast and Conservative strategies

Figure 3 depicts the solutions across the three objective space considered in the study. The x-axis represents emissions in kilograms, y-axis covers the need for demand response and z-axis is for revenue. The ideal point for the three objectives is near the upper left corner as annotated on Figure 3. It becomes apparent when looking at the figure that optimizing for only one objective provides myopic representation of the solution space. For example, the perfect information solutions that appeared so dominant in one- and two-objective views of the problem now lie on the extreme edge with very poor performance in the emissions objective. Similarly, the conservative solution results in extremely high emissions even though it doesn't require demand response. The single solutions occupy a small space and will not provide an accurate picture of possible solutions available to the decision maker. The DPS policies, however, represent a pareto approximate surface across the three objectives providing a diverse set of viable strategies across the solution space.

Figures 1, 2 and 3 demonstrate that solving a MG energy management in a multi-objective context could provide a better understanding of attainable system performance for the decision maker and introduce a broader suite of control strategies that could cause operator's preference among solutions to change dramatically.
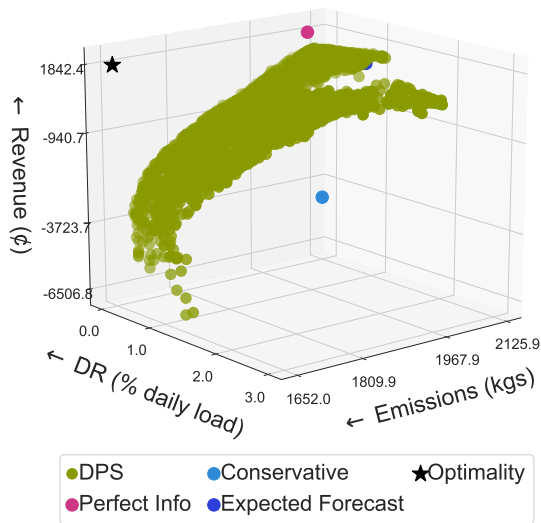
Figure 3. 3d comparison of solution performance for DPS policies, and single objective solutions for Perfect Information, Expected Forecast and Conservative strategies

with non-dominated solutions will not be a trivial task. In addition, it is apparent from Figure 4 that DPS approach is preferable among the formulations used, outperforming even perfect forecast approach in some cases.
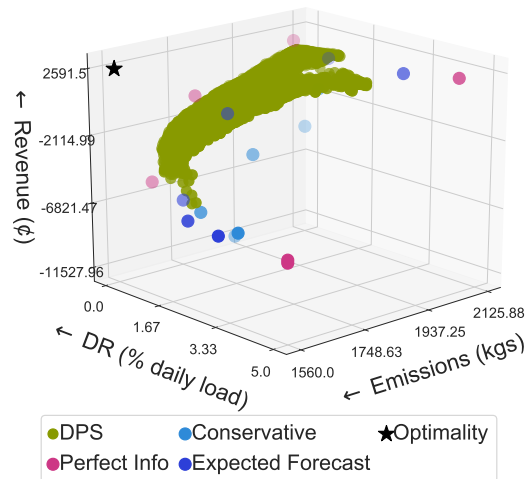

Figure 4. 3d comparison of solution performance for DPS policies, and scalarized multiobjective solutions for Perfect Information, Expected Forecast and Conservative strategies

4.3. Comparing Multiobjective Frameworks

Having demonstrated the utility of multiobjective approach, this section compares the EMODPS formulation against frequently used scalarization method. Scalarization is an approach wherein the problem is converted into a single objective framework by using valuation functions and weighting parameters for each objective. These parameters are either 1) chosen subjectively to obtain one optimal solution [8] or; 2) different parameters sampled randomly to obtain a set of solutions [9]. Both approaches result in inefficient performance, as the former assumes correct and complete prior knowledge about the system as well consensus among stakeholders, while the latter is unable to capture complex non-linear relationships between objectives [26]. The benefits of using an approach like EMODPS over scalarization are demonstrated in Figure 4. The figure compares performance of 3000 DPS policies against multiple scalarized objectives (with different weights) of the other three formulations. Ten different weighting parameters were chosen for each of the three formulations. Figure 4, however, shows only the six solutions per strategy as the other solutions were dominated by those presented in the figure. The scalarized solutions of the three formulations om Figure 4, are spread across the solution space and hardly any strategy seems worth exploring. This supports the earlier argument that sampling the weight parameters to obtain a Pareto-approximate surface

4.4. Aggregate Analysis of DPS Policies

For a more detailed analysis of solution performance, the policies optimized by the DPS approach are shown in Figure 5 on a parallel axis plot. This figure is a representation of trade-offs across the three stakeholder objectives for all the DPS policies. For ease of comparison, each axis has been oriented such that the top of the axis represents best performance for the objective. Each line segment is a policy that is shaded according to performance on the revenue objective across 1000 out-of-sample simulations shown by the colorbar. Examining Figure 5 shows the value of the DPS approach as the policies cover a wide range of possibilities for all three metrics. Such a diverse set of solutions will provide stakeholders with useful insight into the complex trade-offs between the objectives. Figure 5 shows a strong trade-off between the revenue and emissions objectives which can be attributed their definitions. Equations (1) and (2) show that increased diesel generation is required for higher revenue, leading to higher emissions and vice versa.

The trade-offs across the three objectives (especially for demand response) are more complex. For example, high revenue (or low emissions) policies can reside
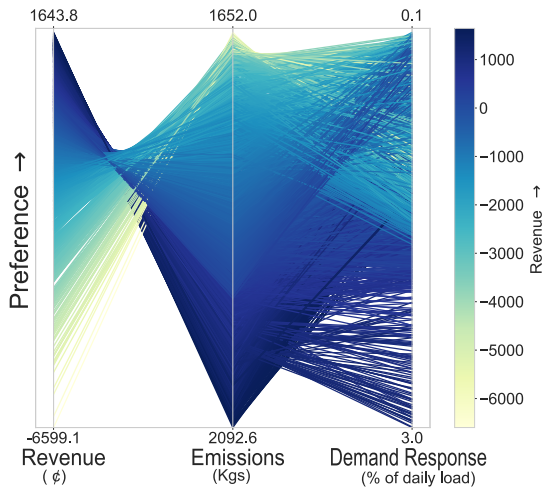
Figure 5. Parallel axis plot of the trade-off set for 3000 non-dominated DPS policies.
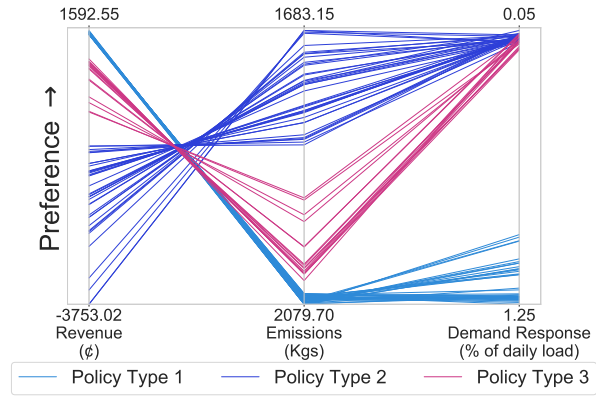

Figure 6. Parallel axis plot of the trade-off set for select policy types where Policy Type 1 are high revenue policies, Policy Type 2 are low emissions policies and Policy Type 3 are balanced policies.

on both low and high end of demand response spectrum based on the emissions (or revenue) strategies. Similarly, other policies portray behavior which would be difficult to predict and replicate *a priori* using a scalarized or single objective optimization. This figure provides stakeholders an overview of possibilities into the behavior of different performance objectives.

Figure 5 provides a comprehensive view of all solutions available to the stakeholders. However, in real-world decision support applications, stakeholders would have certain performance criteria for each objective, allowing them to narrow down to policies attuned to those requirements. For illustrative purposes, Figure 6 is based on three possible sets of criteria a stakeholder might require.

- High Revenue policies (Policy Type 1)– Policies where average revenue performance in the 70th percentile(higher than 70% of the policies).

- Low Emission policies (Policy Type 2)– Policies with average emissions are in the 30th percentile(lower than 70% of policies).

- Balanced policies (Policy Type 3)– Policies with average revenue in the 40th percentile, average emissions in 70th percentile and average demand response requirement in the 30th percentile.

Figure 6 shows the average performance of specific policy types across 1000 out-of-sample scenarios. Type 1 policies generate high revenues by sacrificing on emissions and demand response objectives. These policies show a wide range of demand response values which is not translated to either emission reduction or

revenue generation. Such information can be useful for stakeholders in understanding the limits of the system. Looking at policies of type 2 a wider range of leverage is available in both revenue and emissions objectives for a smaller variations in demand response requirements. While for policies of type 3 the available leverage for revenue and emissions objectives is much higher. Figure 6 points out a key pattern regarding stakeholder objectives in the test system, as the sensitivity of revenue and emissions objectives is decreasing with increasing value of demand response. A more detailed analysis of the interactions between the objectives will allow provide the stakeholders with understanding of the system allowing them to leverage the most value out of the MG resources.

4.5. Hourly Analysis of DPS Policies

This section examines the DPS policies across the 24 hour planning horizon. To accomplish this, a policy is selected from each policy type defined in the previous section, and hourly decision analyzed as shown in Figures 7, 8, and Table 1. As previously discussed, the EMODPS methodology used battery storage level as the system variable dictating the energy exchange(buy/sell) and generation decisions for these policies. Figures 7 and 8 illustrate the average storage level distribution across 24 hours for the three policies selected, with the colorbars representing the electricity buy/sell and generation decisions at those hours, respectively, with actual decision values provided in Table 1. The uppermost plot in both figures are the net load distribution for the horizon depicting the diurnal load requirements, and hence the electricity cost, for the MG system. The lower subfigures demonstrate the

different decisions taken by the three policy types to accommodate the pattern.
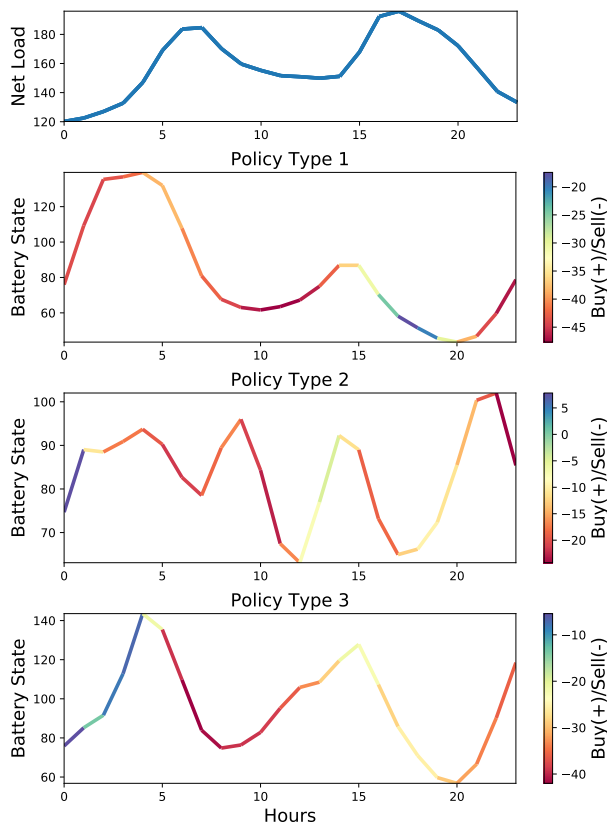

Figure 7. Sample control policies for electricity buy/sell decisions based on storage levels where (+) values represent buying and (-) values represent selling.

Policy type 1 prioritizes high diesel generation and selling to the utility; the battery is charged earlier in the day, and focus on sales results in reduced availability for load peak later in the day, leading to high load curtailment relative to the other policies as can be seen in Table 1. Conversely, policy type 2 utilizes regulation of the conventional generation to reduce emissions, as shown in in Figure 8 and Table 1. The resulting battery level is more variable than the other policy types. The deviation in the storage behavior occurs at peak load hours as both generation and stored electricity is required to meet the extra load requirements. Another interesting behavior from this policy type is the need for demand response in the middle of the day when generation is kept low to reduce emissions. Policy type 3, as expected, is the most balanced approach compared to the other two policies, maintaining a balance between revenue, emissions and demand response utilization. As a result the high amounts of diesel generation is used
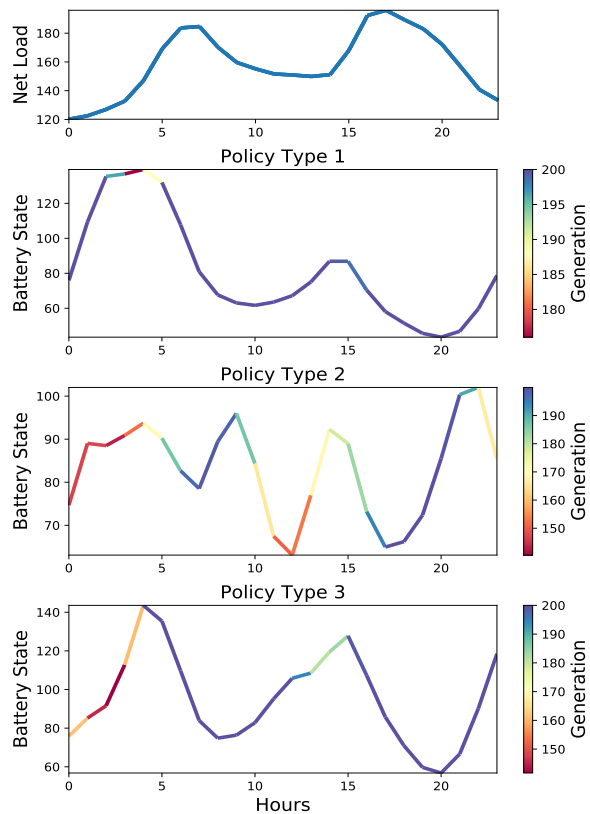

Figure 8. Sample control policies for electricity generations decisions based on storage levels.

to maintain a balance between battery storage levels and selling. While the generation isn't as constrained as the low emissions policy, it is reduced whenever possible to reduce emissions. This policy also requires very low levels of demand response, just after peak hours when the stored electricity in the system is less available. These policies are just three samples of the 3000 policies generated by the EMODPS approach, and each provides a different energy management strategy for the MG which in turn provide the decision-makers as many options. These results may also inform operators regarding the presence of excess conventional generation in the system given that all three policy types are selling to the grid frequently.

5. Conclusion

This study explores a parametric simulation based optimization approach to identify diverse control strategies for a stochastic microgrid system to address multiple objective decision-making problem. The utility of the approach is explored via case studies with promising results. Not only did the the

Table 1. Microgrid Control Decisions

Hr	Policy Type 1			Policy Type 2			Policy Type 3		
	Utility(kw)	Generation(kw)	DR(kw)	Utility(kw)	Generation(kw)	DR(kw)	Utility(kw)	Generation(kw)	DR(kw)
1	-43.73	200.00	0.00	7.82	147.11	0.00	-5.43	161.62	0.00
2	-44.06	200.00	0.00	-10.64	147.59	0.00	-13.65	145.51	0.00
3	-43.26	196.35	0.00	-17.13	143.61	0.00	-8.28	141.69	0.00
4	-41.80	176.03	0.00	-17.30	152.52	0.00	-7.01	161.02	0.00
5	-38.17	188.09	0.01	-19.20	169.16	0.00	-21.55	200.00	0.00
6	-37.98	200.00	0.00	-21.31	186.91	0.05	-38.92	200.00	0.00
7	-40.50	200.00	0.00	-20.80	196.92	0.01	-42.01	200.00	0.01
8	-42.32	200.00	0.08	-18.32	198.81	0.00	-41.17	200.00	0.02
9	-44.01	200.00	1.02	-16.66	197.68	0.00	-39.14	200.00	0.17
10	-46.37	200.00	2.00	-20.82	186.96	0.00	-38.87	200.00	0.10
11	-47.66	200.00	2.55	-22.53	165.97	0.20	-38.31	200.00	0.00
12	-47.64	200.00	2.53	-16.55	150.94	0.54	-36.04	200.00	0.00
13	-46.18	200.00	1.72	-7.43	153.34	1.11	-32.96	194.52	0.00
14	-42.41	200.00	0.43	-4.36	168.64	0.33	-28.52	181.14	0.00
15	-36.39	199.43	0.00	-11.33	178.38	0.00	-21.40	183.67	0.00
16	-30.64	198.52	0.13	-18.71	183.31	0.11	-22.50	198.90	0.00
17	-24.24	200.00	0.10	-18.21	194.58	0.19	-28.21	200.00	0.00
18	-17.43	200.00	1.41	-12.38	199.89	0.28	-25.56	200.00	0.02
19	-20.14	200.00	4.00	-9.56	199.99	0.20	-25.56	200.00	0.01
20	-29.61	200.00	8.33	-10.91	199.99	0.00	-29.02	200.00	0.82
21	-38.26	200.00	10.77	-13.89	199.25	0.00	-31.66	200.00	1.69
22	-43.87	200.00	7.83	-19.20	190.69	0.00	-33.64	200.00	0.25
23	-46.56	200.00	1.34	-24.26	166.70	0.00	-35.64	200.00	0.00
24	-47.63	200.00	0.00	-23.79	140.46	0.06	-38.35	200.00	0.00

EMODPS approach demonstrate the benefits of a multiobjective framework, it also showed potential value for stakeholders by providing diverse set of control strategies. Even though the framework is computationally expensive the policies generated could be used for long term decision making before retraining is required.

The control policies also provide information regarding the limits of the system as well the complex relationships between stakeholder objectives. A deeper analysis with help of the policies could provide useful insights to decision makers which among other things might drastically change their preferences. As EMODPS is a simulation-based method, it can include multiple state variables, multiple objectives as well as network constraints. The parameterized policies improve performance for MOEAs by reducing the curse of dimensionality. The current study should be explored in more detail to study the complex relationships between revenue/emissions and demand response. Initial experiments, in this direction, have shown that these interactions need to be studied in an empirical setup to be effectively utilized. The proposed framework provides such a framework and will be of great benefit in this line of questioning.

This methodology also opens up many interesting directions for future work. It could be used to perform life cycle analysis for microgrid configurations based on different stakeholder objectives. The simulation approach could be utilized to test systems under islanding conditions to better prepare control strategies. The objective of this work was to explore the applications of this approach on a test microgrid, however, future implementations could leverage physics based models for detailed grid representation in a hybrid formulation.

6. Appendix: Parameters

Cost parameters: $C_{gen} = 0.25$ ct/kWh, , $C_{wt} = 1.0$ ct/kWh, $C_{pv} = 2.0$ ct/kWh, , $C_{ES} = 0.30$ ct/kWh. Emissions factor: $C_{ES} = 0.437$ kg/kWh. Reliability criteria: $d_{crit} = 95\%$. System limits: $p_t \in [80, 200]$ kW, $ES_t \in [40, 160]$ kW. Buy/Sell limits: $e_t \in [-50, 50]$ kW. Ramp limits: $p_t \in [-50, 50]$ kW.

Acknowledgment

References

[1] Y. Levron, J. M. Guerrero, and Y. Beck, "Optimal power flow in microgrids with energy storage," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3226–3234, 2013.

[2] M. Ross, C. Abbey, F. Bouffard, and G. Jos, "Multiobjective optimization dispatch for microgrids with a high penetration of renewable generation," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1306–1314, 2015.

[3] A. A. Khan, M. Naeem, M. Iqbal, S. Qaisar, and A. Anpalagan, "A compendium of optimization objectives, constraints, tools and algorithms for energy management in microgrids," *Renewable and Sustainable Energy Reviews*, vol. 58, pp. 1664–1683, 2016.

[4] M. F. Zia, E. Elbouchikhi, and M. Benbouzid, "Microgrids energy management systems: A critical review on methods, solutions, and prospects," *Applied energy*, vol. 222, pp. 1033–1055, 2018.

[5] J. Tant, F. Geth, D. Six, P. Tant, and J. Driesen, "Multiobjective battery storage to improve pv integration in residential distribution grids," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 182–191, 2012.

[6] B. Zhao, X. Zhang, J. Chen, C. Wang, and L. Guo, "Operation optimization of standalone microgrids considering lifetime characteristics of battery energy storage system," *IEEE transactions on sustainable energy*, vol. 4, no. 4, pp. 934–943, 2013.

[7] L. Wang, Q. Li, R. Ding, M. Sun, and G. Wang, "Integrated scheduling of energy supply and demand in microgrids under uncertainty: A robust multi-objective optimization approach," *Energy*, vol. 130, pp. 1–14, 2017.

[8] A. Gupta and C. L. Anderson, "Statistical bus ranking for flexible robust unit commitment," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 236–245, 2019.

[9] R. Soncini-Sessa, E. Weber, and A. Castelletti, *Integrated and participatory water resources management-theory*. Elsevier, 2007.

[10] M. Giuliani, A. Castelletti, F. Pianosi, E. Mason, and P. M. Reed, "Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations," *Journal of Water Resources Planning and Management*, vol. 142, no. 2, p. 04015050, 2015.

[11] J. Sarshar, S. S. Moosapour, and M. Joorabian, "Multi-objective energy management of a micro-grid considering uncertainty in wind power forecasting," *Energy*, vol. 139, pp. 680–693, 2017.

[12] P. Li, D. Xu, Z. Zhou, W.-J. Lee, and B. Zhao, "Stochastic optimal operation of microgrid based on chaotic binary particle swarm optimization," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 66–73, 2015.

[13] A. A. Moghaddam, A. Seifi, T. Niknam, and M. R. A. Pahlavani, "Multi-objective operation management of a renewable mg (micro-grid) with back-up micro-turbine/fuel cell/battery hybrid power source," *Energy*, vol. 36, no. 11, pp. 6490–6507, 2011.

[14] M. Azaza and F. Wallin, "Multi objective particle swarm optimization of hybrid micro-grid system: A case study in sweden," *Energy*, vol. 123, pp. 108–118, 2017.

[15] J. D. Quinn, P. M. Reed, and K. Keller, "Direct policy search for robust multi-objective management of deeply uncertain socio-ecological tipping points," *Environmental Modelling & Software*, vol. 92, pp. 125–141, 2017.

[16] T. Bäck, D. Fogel, and Z. Michalewicz, "Introduction to evolutionary algorithms," *Evolutionary computation*, vol. 1, pp. 59–63, 2000.

[17] D. Bertsekas, "Reinforcement learning and optimal control." Online, 2019.

[18] M. Giuliani, J. D. Quinn, J. D. Herman, A. Castelletti, and P. M. Reed, "Scalable multiobjective control for large-scale water resources systems under uncertainty," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 4, pp. 1492–1499, 2017.

[19] J. D. Quinn, P. M. Reed, M. Giuliani, A. Castelletti, J. W. Oyler, and R. E. Nicholas, "Exploring how changing monsoonal dynamics and human pressures challenge multireservoir management for flood protection, hydropower production, and agricultural water supply," *Water Resources Research*, vol. 54, no. 7, pp. 4638–4662, 2018.

[20] W. Scott and W. B. Powell, "Approximate dynamic programming for energy storage with new results on instrumental variables and projected bellman errors," *Submitted to Operations Research (Under Review)*, 2012.

[21] D. R. Jiang, T. V. Pham, W. B. Powell, D. F. Salas, and W. R. Scott, "A comparison of approximate dynamic programming techniques on benchmark energy storage problems: Does anything work?," in *2014 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, pp. 1–8, IEEE, 2014.

[22] J. R. Kasprzyk, P. M. Reed, B. R. Kirsch, and G. W. Characklis, "Managing population and drought risks using many-objective water portfolio planning under uncertainty," *Water Resources Research*, vol. 45, no. 12, 2009.

[23] M. T. Rosenstein and A. G. Barto, "Robot weightlifting by direct policy search," in *International Joint Conference on Artificial Intelligence*, vol. 17, pp. 839–846, Citeseer, 2001.

[24] W. B. Powell, "What you should know about approximate dynamic programming," *Naval Research Logistics (NRL)*, vol. 56, no. 3, pp. 239–249, 2009.

[25] D. Hadka and P. Reed, "Borg: An auto-adaptive many-objective evolutionary computing framework," *Evolutionary computation*, vol. 21, no. 2, pp. 231–259, 2013.

[26] P. J. Fleming, R. C. Purshouse, and R. J. Lygoe, "Many-objective optimization: An engineering design perspective," in *International conference on evolutionary multi-criterion optimization*, pp. 14–32, Springer, 2005.