Capacity Valuation of Demand Response in the Presence of Variable Generation through Monte Carlo Analysis

Andrew Klem
University of Vermont
Burlington, Vermont
Email: andrew.klem@uvm.edu

Gordon Stephen
National Renewable Energy Laboratory
Golden, Colorado
Email: gord.stephen@nrel.gov

***Abstract*—The goal of decarbonizing the power grid is quickly becoming a reality as improvements in variable renewable generation, communications, and grid infrastructure make this more achievable. This will lead to a replacement of conventional fossil fuels with much cleaner forms of energy. It is still important to ensure that renewables can be used to meet the peak demand with adequate capacity, as is currently done with conventional generation. Determining an accurate capacity value of a resource such as demand response and energy storage can be a challenge due to the required intertemporal considerations that must be made, though various techniques are becoming present in literature. This paper contributes to the field of capacity value by investigating the effect of customer participation in DR programs. This work investigates the contribution that shiftable loads can provide for capacity value in a smart, standalone system. The increasing number of smart devices and smart meters will allow for greater flexibility of load use that can ideally be used at other points in the day. This paper uses a Monte Carlo simulation to investigate the effect of DR on the loss of load probability in a sequential simulation framework. Also presented here is a load-shifting model for customer loads where time of use is based on need, priority, and the availability of generation. Results of this work present a quantitative value for the equivalent firm capacity of demand response programs for a specific test case that depends on the participation rate among consumers.**

***Keywords*—Capacity value, demand response, loss of load expectation, equivalent firm capacity, Monte Carlo**

I. Introduction

System planners must have a good understanding of the maximum load that may be experienced by a system to ensure enough capacity exists to meet this demand, while understanding the cost associated with shedding load and the added difficulty of unplanned outagages and stochastic nature of variable renewable generation (VRG). Conventional generators (CG) are easy to manage being dispatchable, able to deliver reliable power when needed. Determining the capacity value (CV), or rather the ability to adequately meet demand, can be determined quite easily [1]. System operators are willing to pay generators for their ability to serve load during peak hours in the capacity market. This drives the incentive to determine how valuable a resource is to balancing supply and demand in peak hours, particularly those that are readily available during high-load hours, like demand response (DR) and solar.

Higher penetrations of VRG are leading to many studies and methods to determine the CV of wind and solar power [2]–[8]. While CG is typically equally available at all hours of the day and days of the year, most VRG resources like wind and solar show great variation throughout the course of the day and certainly over the different seasons of the year. It is also interesting to note how increasing the amount of a VRG resource often demonstrates diminishing returns in terms of CV in the absence of energy storage, due to the reshaping of the load curve by distributed resources [9]. As penetrations of renewables increase, energy storage systems (ESS) are built for reliability and economic reasons, leading to work regarding the CV of ESS [10]–[12].

Additionally, improvements in communication and smart grid technology are increasing the viability of the use of DR for multiple roles in improving the power system, including energy arbitrage and ancillary services [13]–[15]. However, DR is already playing an enormous role in the capacity market. In PJM, approximately 90% of the DR revenue comes from the capacity market [16], [17]. Shedding or shifting loads during periods of inadequate supply can significantly improve the resource adequacy of a system, as indicated by studies associated with DR [11], [12], [18], [19].

This work investigates the CV of flexible loads in a smart standalone system. The system is based on the realistic and representavie data for demand and VRG from the Greater Los Angeles area, with generic CG data [20], [21]. The data contains the natural relationships between adjacent hours and a sequential simulation is necessary to maintain these inter-hour correlations.

II. Capacity Value

Resource adequacy is a measure of the ability of a system to adequately supply its demand. CV represents the importantance of a resource to a system supplying its peak demand effectively. It is important that balancing authorities ensure that enough generation has been purchased to meet all of the demand. An additional reserve, called the planning reserve margin (PRM), needs to be available as many factors can affect a generator's ability to supply its peak generation. Typically, a PRM of 15-20% is used [22], [23]. This includes downtime for maintenance, unplanned mechanical failures, and faults. Additionally, when larger percentages of VRG are present, the amount of nameplate capacity will need to increase further to make up for the uncertainty of VRG power output. Because these events are often unanticipated, backup generation must be scheduled to ensure that the load can be adequately met. The larger the load margin that is scheduled, the lower the probability of dropping load due to inadequate supply.

Multiple methods exist to determine CV including analytical probabalistic methods and Monte Carlo simulations. Analytical methods involve convolving probability density functions [1], [3]. It is also accepted that capacity metrics can be determined through simulation with and without a resource, where the difference represents the CV of the resource being tested [24].

As VRG is non-dispatchable and uncertain, it is difficult to adequately and accurately determine its CV. Factors such as variability, capacity factor, site location, weather patterns, and demand curve shape can all have an effect. Because this is so reliant on the unpredictable wind and solar patterns, the method for obtaining an accurate, agreed-upon value is quite up for debate, though the IEEE Task Forces have some recommendations [2], [3]. As wind patterns can change and are quite variable from year-to-year, is it accepted that between four to ten years' worth of wind data are necessary to obtain an accurate CV [6]–[8], though there are circumstances where using less data is acceptable [19].

DR can certainly have an affect on a system's ability to match supply and demand and it would be beneficial to system operators to understand the value that this resource would play in generation balancing, particularly load-shifting. Again, an incredible number of factors play a role, including the amount of shiftable load, customer participation rate, and appropriate "payback" time. This adds to the difficulty of CV determination for DR and storage, whose CV is context-dependent. Even load location can be a factor for a networked system with transmission constraints. Therefore, it is difficult to gain an understanding of the CV of DR as its value can be quite conditional.

A. *Capacity Value Metrics*

Multiple metrics exist to measure the resource adequacy of a system with a variety of resources. These include loss of load probability (LOLP), loss of load expectation (LOLE), and expected unserved energy (EUE). Measures of CV typically include effective load-carrying capacity (ELCC) and equivalent firm capacity (EFC). Some argue that certain metrics are more appropriate than others [25], while others suggest that different metrics show similar relationships and are all appropriate for use in defining CV [26]. In this work, LOLE in terms of unserved load hours per year is used to establish baseline values of comparison for the test cases, which are used to determine the EFC of the resource.

III. Standalone System Setup

For the purposes of this experiment, a single-area system was modeled, consisting of a fleet of CG, solar and wind VRG, and customer enduses, some of which are eligible for DR participation and can delay their energy requirement until future hours, should generation resources be unavailable at the present time. The effect of capacity limitations on a transmission network can certainly have an effect on resource adequacy, but as the purpose of this work is simply to demonstrate a methodology on a single-area, standalone system, transmission constraints are ignored here for simplicity.

A single-node system is simulated sequentially with hourly timesteps for the course of a year. The sequential simulation allows for more accurate representation of DR events, where intertemporal memory is required. Values of generation, load, and VRG are sampled and held at that value for a timestep. During this time, the linear optimization determines if enough resources exist to supply the load, and record when load must be dropped due to inadequate supply of generation. This will be further discussed in Section V. Additionally, a Monte Carlo (MC) sampling was used to run the simulation for 10,000 iterations of a yearly simulation, with 8784 hourly timesteps to ensure that a wide variety of possibilities were examined in this work to avoid monotonicity violations in the EFC curve due to sampling error. Ideally, a larger number of MC iterations would have been preferred, but limited resources and an apparent success with 10,000 iterations seemed to justify this number. The simulation was run on the Peregrine high-performance computer at the National Renewable Energy Laboratory (NREL).

A. *Dispatchable Generators*

CG data was obtained from the Reliability Test System - Grid Modernization Lab Consortium (RTS-GMLC) RTS-96 Bus test case [21] which is data publicly dispersed by the Department of Energy, NREL, and the Alliance for Sustainable Energy, LLC. The maximum power ratings, forced outage rate (FOR), and mean time to repair (MTTR) were extracted from the RTS-96 data set to obtain a collection of 96 CG units.

The FOR and MTTR values for each generator were used to create a state transition matrix to determine the ON/OFF status of each generator at each hour in the simulation. Each generator is initialized based on the steady-state probability from the transition matrix. At the start of each hour, the transition matrix is used to determine the next state for each generators. This is simply used to determine if a generator is available to supply power or offline for repairs, but does not mean it will necessarily be used.

Since the MTTR represents an integer number of hours, and thus timesteps in this case, the probability of generator i transitioning from the OFF to ON state can be given by the reciprocal of its MTTR, r [27].

$$p_{01,i} = \frac{1}{r_i} \tag{1}$$

The determination of the transition probably from ON to OFF requires the use of the FOR and MTTR of each generator to determine the expected number of hours that generator i typically remains online. From this, the probability can be shown as

$$p_{10,i} = \frac{1}{\frac{r_i}{f_i} - r_i} \tag{2}$$

where f_i is the FOR for generator i. The state transition matrix for each generator is given by

$$g_i = \begin{bmatrix} 1 - p_{01,i} & p_{01,i} \\ p_{10,i} & 1 - p_{10,i} \end{bmatrix} \quad (3)$$

where $p_{jk,i}$ is the probability of generator i transitioning from state j to state k. State 0 represents the case where the generator is off for maintenance and state 1 represents the case where the generator is able to provide full capacity. The cumulative generator capacity at each hour in the simulation is calculated by summing the maximum capacities of all of the generators that are in state 1 at that time. It is assumed that the 96 generators used is large enough to provide a realistic generation fleet and a low probability that a large percentage of the total generating capacity will be out of service at the same time.

B. Variable Renewable Generation

RTS-GMLC also contains renewable data for wind and solar farms that is averaged over the span of an hour to produce hourly time series data. This data includes four wind farms and 25 solar farms. However, the solar and wind capacities needed to be scaled accordingly to fit the load data being used. A wind scaling factor and solar scaling factor were introduced to scale all of the VRG farms such that the average VRG output, based on its capacity factor from the data, made up 20% of the total generating capacity on average. The scaling factor then scaled the maximum ratings of the VRG devices, such that they are reasonably-sized for this work.

$$\alpha \sum_{i=1}^{n_{gen}} g_i + \beta \sum_{i=1}^{n_{sol}} s_i + \gamma \sum_{i=1}^{n_{wind}} w_i = \mu L_{max} \quad (4)$$

Here, α, β, and γ, represent the scaling factors of CG, solar, and wind respectively. L_{max} is the maximum load expected for the course of the year. While system planners have different targets for LOLE, such as one day in ten years [26], the PRM in this work, μ, was adjusted to produce about 2.4 hours of lost load [28] at the base case, when no DR is used.

C. Demand Model

Load data was obtained from NREL's dsgrid [20], which provides realistic data for southern California, estimating consumer enduses based on the load profile and other data, such as weather patterns, for the county of Los Angeles. This hourly data is divided into categories of fans, pumps, heating, cooling, interior lighting, exterior lighting, water systems, interior equipment, heat rejection, industrial, and commercial. Note that dsgrid disaggregates industrial and commerical loads into greater detail, but that level of detail was not used in this work. For the purposes of this work, it is assumed that only residential heating and cooling can be used for DR participation and are called "controllable" loads. The other categories are grouped into "uncontrollable" loads. Fig. 1 shows samples of this load data.

IV. Demand Response Model

Since DR requires customer participation in order to be effective, it can be assumed that only loads that result in minor or negligible customer inconvenience can be effectively utilized for load-serving purposes. Based on the customer enduses assumed for this work (fans, pumps, heating, cooling, interior lighting, exterior lighting, water systems, and interior equipment), it seems appropriate to only use heating and cooling systems, which includes appliances such as air conditioners, refrigerators, electric hot water heaters, and space heaters, all of which have been shown to be effective for DR participation.

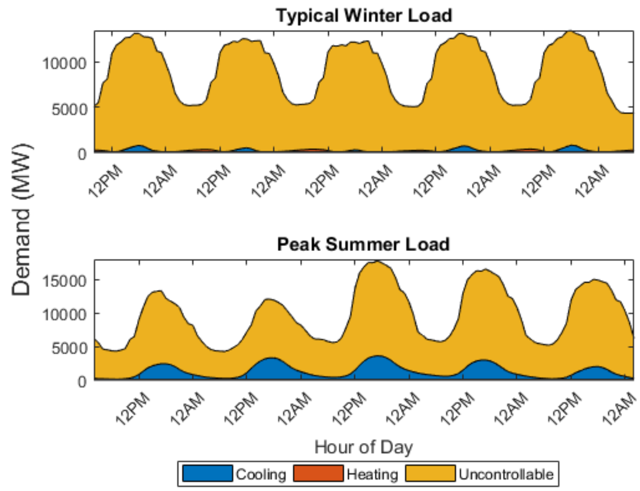

Fig. 1. Samples from the load data demonstrating how the amount of controllable loads (heating and cooling) vary throughout the day and year. It is important to note how small the percentage of controllable loads is in the winter and the absence of the more flexible heating loads in the summer.

TABLE I. DR Event List at Arbitrary Timestep

DR Event	Energy to Repay	Periods Remaining
Event 1	E_1	T_1
Event 2	E_2	T_2
⋮	⋮	⋮
Event n	E_n	T_n

Note that the events are sorted each timestep such that $0 \leq T_1 \leq T_2 \leq \ldots \leq T_n$

For this work, both heating and cooling loads were associated by a specific time for required payback. Since this experiment is discretized into hourly timesteps, an integer number of timesteps must be chosen for each load. For this work, it is assumed that cooling loads are more critical to customers and must be payed back sooner than heating loads. Thus, the optimization is set up in such a way that heating loads are the first to be shed in the event of a shortfall in generation. This provides maximum flexibility as there are more timesteps over which the DR energy can be repayed. Thus, when the system requires additional generation, cooling loads are the last to be utilized, and usually the first to be repayed if excess generation is available.

DR events are stored in a heap, as demonstrated in Table I from which events are extracted based on minimum payback time, to help ensure that the most urgent controllable loads are served first. Should a load's time limit expire prior to being served, a loss of load event will be recorded to represent the flexible demand that was unserved.

For this work, heating loads must be repayed within four time periods, while cooling loads need to be repayed in two time periods, as they are considered a higher priority. Note that DR cannot be used in advance during times of generation surplus in anticipation of shortage events. It is solely used to help bring supply and demand into balance during times when load shortage is otherwise inevitable.

V. Linear Optimization Setup

A linear program (LP) is used to determine the optimal dispatch of energy at each timestep; no intertemporal optimization is used. This was performed using Julia's MathProgBase package and CLP solver. Use of an LP was certainly unnecessary for a single-area system in the absence of transmission, but as this work represents the presentation of a simple methodology for a more complicated work to come, the use of LP seemed important. The additional simulation time due to LP was small enough to sanction its use. Eq. (5) shows the setup of the optimization used.

$$\min_{x} c^T [x_{gen} \; x_{DR,inj} \; x_{UL} \; x_{DR,rep}]^T \tag{5}$$

Subject to

$$\begin{aligned} x_{gen} + x_{DR,inj} + x_{UL} - x_{DR,rep} &= L_i \\ x_{DR,inj} + x_{DR,rep} &= L_{maxcont} \\ 0 \leq x_{gen} &\leq \overline{G}_i \\ 0 \leq x_{UL} &\leq L_i \\ 0 \leq x_{DR,inj} &\leq \overline{D}_{inj,i} \\ \underline{D}_{rep,i} \leq x_{DR,rep} &\leq \overline{D}_{rep,i} \end{aligned}$$

The components of x are the total amount of power used by the sum of CG and VRG (x_{gen}), total DR injection ($x_{DR,inj}$), total unserved load (x_{UL}), and total repayment of DR events respectively at each timestep ($x_{DR,rep}$). L_i is the total load to be served at time iteration i and $L_{maxcont}$ represents the maximum expected amount of controllable load over the course of a year. $\overline{D}$, $\underline{D}$, and $\overline{G}$ represent the bounds of the decision variables, determined based on generation and DR available, as well as the DR that is coming due for repayment.

c is a vector of costs, that were selected to set the order of importance of each of the resources supplying power. Since the value of the minimized objective function was of little importance, the specific value of the cost parameters does not matter, but rather the relative values to ensure that certain resources are chosen for use before others.

$$c_{DR,rep} < c_{gen} < c_{DR,inj} < c_{UL}$$

This relationship ensures that DR events are payed back prior to additional DR events being utilized, if possible, and prior to using unserved load. Note that unserved load is treated as a power injection for this work.

VI. Results

Since different power system layouts can have significant effects on the results of a CV study, multiple scenarios consisting of different penetrations of VRG and CG should be performed to fully gain an understanding of the capacity contribution of DR. However, due to size constraints for this paper, the results of only one scenario are presented here. For all of the results shown here, the total generation is comprised of 80% CG, 10% wind power, and 10% solar PV. Note that these percentages for VRG are based on capacity factors from previous years' data, rather than nameplate rating. Multiple trials are conducted, allowing λ, the DR participation rate, to vary.

TABLE II. Simulation Results

λ	LOLE (hours/yr)	EFC (MW)
0	3.27	0
0.25	1.11	643
0.5	0.33	1213
0.75	0.14	1549
1	0.06	1898

The main item of investigation is the effect of customer participation rate in DR programs on the reliability of the system. Of the enduse data used for this study, only residential heating and cooling loads are available for use as DR power injection, and only a fraction of these loads are used depending on the assumed DR participation rate. It is assumed that participation rate is applied to heating and cooling loads equally. The parameter λ was varied in increments from 0 to 1, to explore how the variability affects the resulting EFC.

At each value of λ, 10,000 MC iterations were performed to determine the mean number of loss of load hours. These data are presented in Table II. Next, the EFC was computed to quantify the CV of the DR resource. This value was determined by removing the DR capability and leaving all of the other resources unchanged, while including one additional CG unit, whose FOR is zero, so the unit is always available. The generating capacity of this additional unit is varied through multiple simulations. At each value, 10,000 iterations are performed to obtain a LOLE with this unit. The nameplate capacity at which the additional generator produces the same LOLE as the DR cases represents the EFC of the DR resource.

Fig. 2 shows the resulting simulations with the additional firm CG unit as circles, with an interpolation between the points as a line. A dashed horizontal line shows the LOLE at a particular value of λ. The crossing point is marked by an $\times$ to show the intersection point, and thus the EFC. Note the higher density of circles in particular areas to help ensure greater accuracy around those points. The EFC denoted by the intersection points are listed in Table II.

VII. Conclusion

This work investigates the effect of DR participation rates on the CV of shiftable DR loads. A simulation was performed with and without the use of DR resources in order to gain an understanding of how valuable the DR is to the system in terms of helping to ensure that all load is adequately supplied. The results of this work provide a CV to various levels of DR based on the percentage of customer participation for a particular system with 20% VRG penetration. Similar to other forms of VRG, there appears to a diminishing return as greater amounts of DR penetration provide less and less benefit. It is important to understand that this evaluation is very context-dependent and specific to a certain amount of

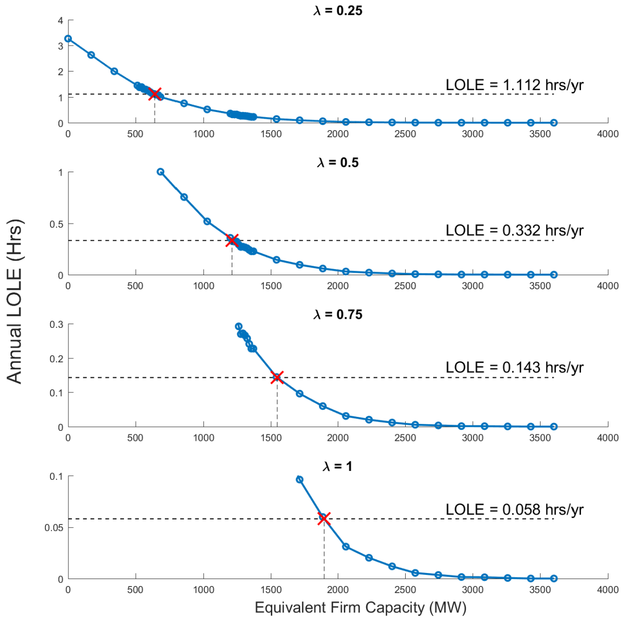

Fig. 2. Results of MC iterations using a "firm generator" in place of DR. The output of the iterations are show as circles, which are connection by an interpolation. The dashed lines indicate the LOLE determined using DR. Note that $\lambda = 0$ is not shown because its EFC is known to be zero. The same interpolation curve is shown in each, just zoomed in for detail. See Table II for intersection values.

VRG penetration and payback time requirement for DR. The relationship between the CV of DR in the presence of various penetrations of VRG will be investigated in future work.

ACKNOWLEDGMENT

The authors would like to thank Dr. Paul Hines of the University of Vermont as well as Madeleine McPherson and the Grid Systems Analysis team at NREL for their input and assistance with this work.

REFERENCES

[1] C. Dent and S. Zachary, "Capacity value of additional generation: Probability theory and sampling uncertainty," May 2013.

[2] A. Keane, M. Milligan, C. J. Dent, B. Hasche, C. D'Annunzio, K. Dragoon, H. Holttinen, N. Samaan, L. Soder, and M. O'Malley, "Capacity Value of Wind Power," *IEEE Trans. on Power Syst.*, vol. 26, pp. 564–572, May 2011.

[3] C. J. Dent, R. Sioshansi, J. Reinhart, A. L. Wilson, S. Zachary, M. Lynch, C. Bothwell, and C. Steele, "Capacity value of solar power: Report of the IEEE PES task force on capacity value of solar power," in *2016 Int. Conf. on Prob. Methods Applied to Power Syst. (PMAPS)*, Oct. 2016.

[4] B. Sigrin, P. Sullivan, E. Ibanez, and R. Margolis, "Representation of the solar capacity value in the ReEDS capacity expansion model," in *2014 IEEE 40th PV Specialist Conf. (PVSC)*, pp. 1480–1485, Jun. 2014.

[5] F. Fattori, N. Anglani, I. Staffell, and S. Pfenninger, "High solar photovoltaic penetration in the absence of substantial wind capacity: Storage requirements and effects on capacity adequacy," *Energy*, vol. 137, pp. 193–208, 2017.

[6] B. Hasche, A. Keane, and M. O'Malley, "Capacity value of wind power, calculation, and data requirements: The Irish power system case," *IEEE Trans. on Power Syst.*, vol. 26, no. 1, pp. 420–430, 2011.

[7] M. Milligan, B. Frew, E. Ibanez, J. Kiviluoma, H. Holttinen, and L. Söder, "Capacity value assessments of wind power," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 6, Jan. 2017.

[8] D. Gami, R. Sioshansi, and P. Denholm, "Data challenges in estimating the capacity value of solar photovoltaics," *IEEE Journal of Photovoltaics*, vol. 7, pp. 1065–1073, Jul. 2017.

[9] P. Denholm and R. Margolis, "The potential for energy storage to provide peaking capacity in california under increased penetration of solar photovoltaics," Mar. 2018.

[10] R. Sioshansi, S. H. Madaeni, and P. Denholm, "A dynamic programming approach to estimate the capacity value of energy storage," *IEEE Trans. on Power Systs*, vol. 29, pp. 395–403, Jan. 2014.

[11] Y. Zhou, P. Mancarella, and J. Mutale, "Modelling and assessment of the contribution of demand response and electrical energy storage to adequacy of supply," *Sustainable Energy, Grids and Networks*, vol. 3, pp. 12–23, 2015.

[12] Y. Zhou, P. Mancarella, and J. Mutale, "Framework for capacity credit assessment of electrical energy storage and demand response," *IET Gen., Trans. & Dist.*, vol. 10, pp. 2267–2276, Jun. 2016.

[13] M. Amini and M. Almassalkhi, "Investigating delays in frequency-dependent load control," in *2016 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)*, pp. 448–453, Nov. 2016.

[14] O. Ma, N. Alkadi, P. Cappers, P. Denholm, J. Dudley, S. Goli, M. Hummon, S. Kiliccote, J. MacDonald, N. Matson, D. Olsen, C. Rose, M. D. Sohn, M. Starke, B. Kirby, and M. O'Malley, "Demand Response for Ancillary Services," *IEEE Trans. on Smart Grid*, vol. 4, no. 4, pp. 1988–1995, 2013.

[15] A. Klem, K. Dehghanpour, and H. Nehrir, "Primary frequency regulation in islandec microgrids through droop-based generation and demand control," in *2017 19th Int. Conf. on Intelligent System App. to Power Systems (ISAP)*, pp. 1–6, Sep. 2017.

[16] D. Hurley, P. Peterson, and M. Whited, "Demand response as a power system resource," no. May, pp. 1–5, 2013. https://www.raponline.org/wp-content/uploads/2016/05/synapse-hurley-demandresponseasapowersystemresource-2013-may-31.pdf.

[17] S. Nolan and M. O. Malley, "Challenges and barriers to demand response deployment and evaluation," *Applied Energy*, vol. 152, pp. 1–10, 2015.

[18] M. McPherson and B. Stoll, "Demand response for variable renewable energy integration: a proposed implementation and its impacts," 2018. *in progress*.

[19] S. Nolan, O. Neu, and M. O'Malley, "Capacity value estimation of a load-shifting resource using a coupled building and power system model," *Applied Energy*, vol. 192, pp. 71–82, Apr. 2017.

[20] E. Hale, H. Horsey, B. Johnson, M. Muratori, E. Wilson, *et al.*, "The demand-side grid (dsgrid) model documentation," 2018. Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A20-71492. https://www.nrel.gov/docs/fy18osti/71492.pdf.

[21] C. Barrows and NREL, "Reliability test system - grid modernization lab consortium," 2018. https://github.com/GridMod/RTS-GMLC.

[22] NERC, "2016 Probabilistic Assessment," 2017. https://www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/2016ProbA_Report_Final_March.pdf.

[23] NERC, "2018 Summer Reliability Assessment," 2018. https://www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/NERC_SRA_05252018_Final.pdf.

[24] S. Nolan, M. O'Malley, M. Hummon, S. Kiliccote, and O. Ma, "A methodology for estimating the capacity value of demand response," *IEEE PES General Meeting*, Oct. 2014.

[25] R. Earle, E. P. Kahn, and E. Macan, "Measuring the capacity impacts of demand response," *Electricity Journal*, vol. 22, pp. 47–58, Jul. 2009.

[26] E. Ibanez and M. Milligan, "Comparing resource adequacy metrics," *13th Wind Integration Workshop*, Nov. 2014.

[27] R. Billinton, *Power System Reliability Evaluation*. Gordon and Breach, 1970.

[28] M. Evans, S. Tindemans, and D. Angeli, "Shortfall-Minimising Dispatch of Heterogeneous Stores and Application to Adequacy Studies," pp. 1–8, 2018.