Title
Metrics for Evaluating Grid Service Provision from Communities of Grid-interactive and Efficient Buildings and other DER

Authors
MacDonald, Jason S
Regnier, Cynthia

Publication Date
2023-07-20

DOI
10.1109/pesgm52003.2023.10252606

Metrics for Evaluating Grid Service Provision from Communities of Grid-interactive and Efficient Buildings and other DER

1st Jason S. MacDonald
Energy Storage and Distributed Resources Division
Lawrence Berkeley National Laboratory
Berkeley, CA USA
jsmacdonald@lbl.gov

2nd Cynthia Regnier
Building Technologies and Urban Systems Division
Lawrence Berkeley National Laboratory
Berkeley, CA USA
cmregnier@lbl.gov

***Abstract*—Tapping into the flexibility of aggregations of Grid-interactive and Efficiency Buildings (GEBs) represent a large opportunity to cost-effectively improve operations in future low-carbon power grids. The US Department of Energy's Connected Communities program seeks to demonstrate collections of these GEBs at 10 different sites across the country. However, consistent metrics for evaluation across all projects are needed to build confidence in the approach. This paper presents the metrics relating to Grid Service Provision that will be computed at the 10 demonstrations. The metrics proposed quantify the magnitude of service offers, the consistency and quality of services provided by the community, and the individual DER contributions to the community-level service. The goal of this work is to present these metrics and describe some of their intended insights. The hope is that these insights will be useful and build confidence for grid operators, regulators, and aggregators and practitioners as they look to deploy these resources in grids of the future.**

***Index Terms*—Connected Communities, metrics, demonstration evaluation**

I. Introduction

Harnessing the flexibility of load in Grid-interactive and Efficient Buildings (GEBs) to provide power system services remains of interest to research and industry [1]. Aggregating GEBs in combination with other distributed energy resources (DER), in a community is seen as an opportunity to achieve scales large enough to participate in bulk-level service offers, as well as provide additional control points for advanced distribution utilities to support local services in the distribution network [2]. This type of collected resource supports a low-carbon future, enhances local resilience, and can provide access for low- or middle-income households to advanced energy technologies by lowering cost of ownership and approaching economies of scale. However, achieving these benefits not only requires the ability to scale the resource size, but also to build confidence in its deployment and operation by electricity system operators and planners. To enable this vision, The U.S. Department of Energy recently announced a portfolio of demonstration-focused projects looking to promote commercialization pathways and build confidence in these resources, called Connected Communities (CC) [3].

The work described herein was was funded by the U.S. Department of Energy's Building Technology Office in the Office of Energy Efficiency & Renewable Energy under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH11231.

The Connected Communities program awarded 10 demonstration projects across the US that represent a diverse set of building types, vintages, and environmental and regulatory contexts [4]. These buildings are coordinated with other distributed energy resources (DER) to provide a variety of grid services. Other examples of connected communities are being piloted in Alabama, California, Colorado, Georgia, New York and Washington [5]. DOE's CC effort represents a significant investment over the next 5 years with the goal to demonstrate the feasibility of various DER portfolios and control approaches that will unlock demand flexibility in regions across the country. To achieve that goal, it is critical that the metrics used and the data collected across the projects are consistent and informative. In the role of National Coordinator for the CC portfolio, researchers at Lawrence Berkeley National Laboratory have defined a set of metrics to be collected by all CC projects to aid this consistency. This paper aims to present the subset of those metrics corresponding to grid service provision and provide a qualitative discussion around each metric's purpose.

Quantifying grid services offered is typically done via methods from system operator's manuals and tariffs. Some literature has focused on generating resource specific metrics, such as Shed Intensity [6], but these have limited utility for grid operators. Other evaluation methods focus on long-term system-level benefits of DERs [7] rather than the event-based evaluation used for remuneration. This work extends the system operator methods to attempt to provide additional insight for all parties and advance the dialogue around how and what to measure for grid service provision such that CC resources can be relied on by electric system operators.

The paper's organization is as follows: Section II provides a short description of the potential grid services referenced in the CCs program. Section III proposes metrics to measure the quantity of service provided by a CC. Section IV proposes metrics to measure the consistency and quality of the grid service offered by a CC. Section V describes metrics to evaluate the contribution of individual DER to the community-level services provided. Section VI concludes.

II. Grid Services for Evaluation

Below is a brief description of possible grid services included for evaluation in DOE's CC program. The services are mapped below to be provided to system operators at either the transmission- or distribution-level.

A. Transmission-level Services:

Economic Energy Dispatch: This service is participation in energy markets and requires the ability to respond to a dispatch signal, which can be as often as every 5 minutes.

Flexible Ramp Energy: This style of service is a short term balancing capacity that is procured in 5 and 15 minute markets to ensure there is enough ramping capability available to meet forecast change and uncertainty in net demand. It may be procured as a symmetric or asymmetric (separately in ”up” and ”down” directions) product.

Capacity (Peak Hours): This is long term power capacity contracted to participate in the wholesale energy market at specific hours of the day for a monthly or seasonal period. It ensures adequate generation available to meet electricity demand in peak hours.

Capacity (Ramping Hours): This is long term power capacity with a specified minimum ramping capability that is contracted to participate in wholesale markets at specific hours of the day, monthly or seasonally.

Contingency Reserves: Contingency reserves are ancillary services that hold power capacity in reserve to be deployed if there is some kind of unexpected major event on the system that creates significant imbalance between supply and demand. These reserves require deployment to full capacity within 10 minutes, and the more valuable contingency reserves (”Spinning” or ”Synchronous”) require a resource to be synchronized with the grid and frequency responsive.

Frequency Regulation: This ancillary service is a balancing capacity deployed in real-time to manage area control error and grid frequency for the grid operator. Resources providing frequency regulation must respond to power commands and provide telemetry as often as every 2 seconds from/to the system operator’s Automatic Generation Control (AGC) system. This may be procured as a single symmetric or two separate (”up” and ”down”) capacity services.

Voltage Support: Is the ability to respond to requests from the system operator to increase or decrease voltage at the point of common coupling, often through changes in real and reactive power output.

Blackstart: Is the capability for a resource to serve load from an outage condition without relying on the wider transmission grid to provide power or reference AC waveforms.

B. Distribution-level Services:

Capacity Relief: This service provides local power capacity that may be deployed to limit power flow at a specific bus in the distribution network.

Emergency Load Transfer: This power capacity service supports the movement of load from one feeder to another during planned or unplanned events through coordination with distribution operators.

Outage Recovery: A capacity service that supports the recovery of distribution systems in outage conditions through coordination with distribution operators that may phase the networks restoration to maintain stability.

Voltage Management: A service that dispatches real and reactive power to maintain local voltage within ANSI specified levels (± 5% per unit).

Power Quality: This service broadly maintains voltage, current, frequency, and the AC power waveform so as not to interfere with normal operation of electrical equipment.

Phase Balancing: This is an active control service that maintains the power flow on each phase of a 3-phase power distribution network such that each is roughly equal.

III. Metrics for Quantifying Grid Service Offers

The physical grid service magnitude provisioned by a resource, like a connected community, is of critical interest to any evaluation of resource performance. These resource agnostic quantities are computed to evaluate the reliability and economic value of offers received by grid operators of distribution utilities and bulk-level systems, as well as for aggregators as they consider what to include in their portfolio. Beyond the physical quantity of service, an estimation of the financial value of the services offered is also necessary to guide investment in these technologies.

Metrics in this section quantify the energy and power capacity offers made by resources. The goals of the metrics are to provide grid operators and regulators with some expectation of the relative quantities of services that may be obtained, support the incorporation of these resource types into electricity planning practices, evaluate the energy consumption of providing services, and provide a sense of the value of the services offered. They should also give aggregators and community/building managers the information necessary for portfolio planning and making service offers in power systems. The quantities computed for these metrics will be based on both physical measurements (real/reactive power, energy, volts) and the service offers at the aggregate community level, and may be computed from statistical samples of measurements. Lastly, grid services from GEBs are often quantified as the absence of load. To compute these quantities requires that the measured real power must be compared to a counterfactual baseline (the building load without the grid service controls implemented). This work is agnostic to the baseline methodology used, but error in the counterfactual baseline computation may significantly impact the results and thus it is suggested that baseline error be estimated in addition to the metrics.

The list of metrics is below. Not every metric need be computed for each grid service. Table I maps the metrics to the grid services for which they may be useful.

A. Grid Service Quantity Metrics:

1) *Average Grid Service Capacity* [kW, kW-h, or kVAR]: This metric quantifies the average capacity offered to grid operators from the community level resource, either for long-term or ancillary service capacity. Long-term capacity may be contracted monthly, seasonally, or annually, and should be qualified with the hours of the day/week that the resource was obligated to provide it. Ancillary service capacity (contingency reserves and frequency regulation) is typically awarded hourly. An annual, seasonal, or monthly average (as appropriate) of the hourly reserves provided should be computed.
2) *Relative Grid Service Capacity* [kW/kW or kVAR/kW]: This metric normalizes the Average Grid Service Capacity by the average load in the community over the same interval to inform the amount of service achievable from the community and support portfolio planning.
3) *Annual/Seasonal Relative Grid Service Capacity* [kW/kW-peak or kVAR/kW-peak]: This metric

normalizes the Average Grid Service Capacity by the community peak load. The community peak load should be determined as the maximum average power, in kW, of all 15 minute intervals during periods (Annual, Seasonal, Monthly) consistent with the structure of capacity obligations in the region.

4) *Coincident Grid Service Capacity* [kW]: This final capacity metric identifies the contribution of the community's grid service capacity provided during the hour of peak consumption in the service territory to give grid operators a sense of how much this resource may be relied upon when capacity is stressed.
5) *Average Energy of Service* [kWh/hour]: The energy of the service is computed to understand the energy cost of providing a specific service, largely for aggregators and building managers. This energy is calculated for each hour in which service was provided by subtracting the measured community-level consumption from a counterfactual baseline energy for the community, and then averaging that across all hours in a period of interest (annual, seasonal, monthly).
   For reactive power services, the energy of service in a given hour is any real power curtailed in order to provide VARs or hold VAR capacity.
6) *Relative Energy of Service* [kWh/kWh]: The relative energy of service indicates the energy impact that service provision may have on the energy services received by the community. For this calculation, the average energy of the service is divided by the average baseline energy during the hours of service provision.
7) *Average Grid Service Mileage* [kW/hr or kVAR/hr]: Mileage measures how much the grid service requires a resource to change its output, and is used in most frequency regulation markets as an input to the performance payment calculation. Mileage is the sum of the absolute value of change in service setpoints sent to the resource. This aggregation-scale value is computed for each hour and reported as an average across all hours in which service was provided inside a period of interest.
   We also suggest a modified mileage calculation for reactive power support services to provide a sense of the physical quantity of support provided by resources within the network. As many reactive power support services are provided through autonomous control reacting to local measurements, instead of the change in setpoints, the absolute value of the change in reactive power output should be summed.
8) *Financial Value of Grid Service* [$/kW, $/kWh, or $/kVAR]: The value of a grid service provisioned by the community to the utility or wholesale market operator is useful for investment and other decision making. Price information should be gathered from the best available data depending on the service (wholesale markets, program incentives, utility regulatory filings and published costs, cost of replacement capacity, etc.).

IV. Metrics Evaluating Consistency of Grid Service Provision

It is important to understand the ability of a grid operator to rely on a resource to be available when needed, the repeatability of its performance, and the accuracy by which it follows dispatch instructions. The metrics described in this section will evaluate the performance of the community-level resource to ascertain how consistent it was in providing grid services. They will report distributions or histograms of services provided, both in absolute and relative terms, evaluate their ability to meet their capacity obligations, and calculate the accuracy of grid service provision. As not every metric presented can be applied to every service, the application of metrics to the grid services of interest are shown in Table II.

TABLE I
Applications of Service Quantity Metrics to Grid Services.

Grid Services	**Service Quantity Metrics**							
	1	2	3	4	5	6	7	8
Transmission								
Economic Energy Dispatch					•	•		•
Flexible Ramp Energy					•	•	•	•
Capacity (Peak Hours)	•	•	•	•				•
Capacity (Ramping Hours)	•	•	•					•
Contingency Reserves	•	•	•		•	•		•
Frequency Regulation	•	•	•		•	•	•	•
Voltage Support	•	•	•		•		•	•
Blackstart	•							•
Distribution								
Capacity Relief	•	•	•	•				•
Emergency Load Transfer	•	•	•		•	•		•
Outage Recovery	•							•
Voltage Management	•	•	•		•		•	•
Power Quality	•	•	•		•		•	•
Phase Balancing	•	•	•		•		•	•

In order to have a strong intuition for the repeatability of service provision, each metric listed is reported as a table of summary statistics and a histogram showing its distribution. The table should include values for mean, median, variance, 25th percentile, 75th percentile, minimum, maximum, and the total number of observations.

A. *Grid Service Consistency Metrics:*

1) *Hourly Grid Service Capacity* [kW or kVAR]: Similar to the Average Grid Service Capacity in section III, the hourly grid service capacity offered at the community scale should be recorded, and summary statistics and histogram presented for capacity services offered on an hourly or sub-hourly basis.
2) *Hourly Relative Grid Service Capacity* [kW/kW or kVAR/kW]: This metric normalizes the amount of grid service capacity offered by that hour's total community-level load to characterize a distribution of how controllable the total resource is from hour to hour. Summary statistics and histograms are presented to show the variation in this hourly value.
3) *Hourly Capacity Commitment Deficit* [kW or kVAR]: This metric presents a distribution of the amount of grid service capacity that was offered, but that the community resource was unable to deliver in real-time, to give a historical performance of its ability to meet its capacity obligations. For each hour of grid service capacity provision, the quantity recorded should be the capacity offered less the maximum power delivered during the hour. If the dispatch request during the entire hour was less than the grid service capacity offered, the deficit should

be recorded as the difference between the maximum dispatch and the corresponding power delivery. Hours in which no grid service dispatch is requested are ignored.

4) *Hourly Relative Capacity Commitment Deficit* [kW/kW or kVAR/kW]: This metric presents the capacity commitment deficit relative to the amount of capacity offered to give operators a sense of what percent of capacity can be counted on. To compute, each value for Hourly Capacity Commitment Deficit computed in the previous metric is then normalized by the capacity commitment for the grid service of interest for that hour.
5) *Energy of Service* [kWh/hour]: This metric provides an understanding of the variability in the Average Energy of Service metric described in the previous section. To compute the hourly energy of service, subtract the measured community-level consumption from the counterfactual baseline computed for the service. For reactive power, the hourly energy of services is any real power curtailed in order to provide VARs or hold VAR capacity.
6) *Relative Energy of Service* [kWh/kWh]: The relative energy of service metric normalizes the energy of service by the baseline energy to show the relative amount of energy services provided. To compute the values determined for the Energy of Service metric are divided by their hourly counterfactual baseline values.
7) *Hourly Grid Service Mileage* [kW/hr or kVAR/hr]: This is mileage metric provides an understanding of the hourly variability around the Average Grid Service Mileage described in section III by evaluating the metrics distribution across all hours.
8) *Accuracy of Service* [kW/signal/hour]: The accuracy of service metric identifies the success of the community control approach at managing their resources to the dispatch targets set by grid operators when providing grid services. This is computed at the community-level each hour as one minus the sum of the absolute value of the error between the dispatch signal and the measured response divided by the sum of the absolute value of the dispatch signal across the hour. Additionally, the dispatch signal should reflect only the power of the grid service, and not any power that may exist in the baseline; this is to ensure that the denominator of the accuracy metric is not dominated by power not offered as a grid service. The accuracy of service metric is documented as hourly averages, though a shorter interval may be used, up to as often as a dispatch command is received (5 minutes for electricity dispatch, 4 seconds for frequency regulation, etc). Summary statistics and histograms show the variation in accuracy across hours.
9) *Length of voltage violations* [minutes]: Voltage management services in the distribution system attempt to maintain voltage within a specific range of $\pm$ 0.05 per unit. Each deviation outside of that and its length should be recorded. Summary statistics and histograms show the success of the community's voltage management. Isolating specific locations within the grid to report statistics of violations independently may lead to more actionable information than community level reports.

TABLE II
APPLICATION OF CONSISTENCY METRICS TO GRID SERVICES.

Grid Services	**Consistency Metrics**								
	1	2	3	4	5	6	7	8	9
Transmission									
Economic Energy Dispatch					•	•		•	
Flexible Ramp Energy					•	•	•	•	
Capacity (Peak Hours)		•	•	•					
Capacity (Ramping Hours)		•	•	•					
Contingency Reserves	•	•	•	•	•	•		•	
Frequency Regulation	•	•	•	•	•	•	•	•	
Voltage Support	•				•		•		•
Blackstart		•	•	•					
Distribution									
Capacity Relief		•	•	•					
Emergency Load Transfer					•	•			
Outage Recovery		•	•	•					
Voltage Management	•	•	•	•	•		•		•
Power Quality	•	•	•	•	•		•		
Phase Balancing	•	•	•	•	•		•		

V. METRICS EVALUATING DER CONTRIBUTION TO AGGREGATE GRID SERVICE OFFERS

These metrics described in this section indicate the contribution of individual DER grid service capacity to the whole community-level grid service provided. The goal of these metrics is to provide aggregators and community managers an understanding of how much individual DER capacity is needed to make community-level offers to grid operators with certainty and to illustrate the differences across DER responses in the community. These metrics may also provide insight into the fairness of community control algorithms with regard to how often and at what magnitude individual DER are dispatch relative to each other.

The proposed metrics apply to the grid services of interest as shown in Table III. Statistical sampling to capture the variation of DER within the community may be employed to reduce the overall burden of data collection. The counterfactual baselines used in these metrics should be computed for individual DER in a manner consistent with the approach used for the community level metrics.

A. *Individual DER Contribution Metrics:*

1) *Average individual grid service capacity* [kW or kVAR]: This metric parallels the "Average Grid Service Capacity" metric in Section III, but is applied to individual resources in the community. It is computed for each resource contributing to a service, with the total capacity across all resources being greater than or equal to the grid service capacity offer from the community.
2) *Coincident individual grid service capacity* [kW]: Similar to the community-level coincident capacity, this metric attempts to capture the capacity delivered in the hours of peak electrical demand by each DER in the community. This may provide insight into specific resource types that could be used to improve the coincident offerings in the future.
3) *Average Energy for Grid Service Hours* [kWh/hour]: The average electrical energy provided by an individual resource during the hours in which the community

was providing grid services to indicate the individual resource's overall contribution to the community grid service provision. This energy is calculated for each service hour in a period (month, season, year), measured as a difference from the individual resource's baseline, and then averaged across all service hours in that period. For reactive power services, the energy calculated is any real power curtailed in order to provide VARs or hold VAR capacity. It is expected that some hours in which the community provides grid services, an individual resource may not be contributing to said services, which will drive down the individual average.

4) *Average Energy Delivery when providing Grid Services* [kWh/hour]: This metric reports the electrical energy provided by an individual resource averaged across the hours in which that resource provided the service. It gives a sense of the amount of service each resource provides when it is able to respond and is distinct from the previous metric in that the focus of the metric is on the individual's performance rather than its contribution to the whole. The computation performed for each service hour is same as the previous metric, it is only the hours included that differentiate it.
5) *Relative Average Grid Service Energy* [kWh/kWh]: This metric normalizes the amount of service provided by an individual resource by what was provided by the community as a whole. For each hour, the resource's total energy of providing grid service is divided by the total energy of grid service provided by the community. Like "Average Energy for Grid Service Hours", this value is averaged for all hours in which the community provided the relevant service.
6) *Average Individual Grid Service Mileage* [kW or kVAR]: This metric applies the community-level mileage metric to the individual resources providing the services. This will help to indicate which resources inside the community are being relied upon to provide service in real-time. The computation is the same as for the community level, but the signal is the setpoints sent to the individual resource by the local controller, not the system operator.
7) *Percentage of Participating Service Hours* [%]: The percentage of hours in which an individual resource contributed to the community's response to a grid service request within an appropriate period (month, season, year). The numerator is the number of hours in which a resource or site participated in a grid service dispatch response, the denominator is the number of hours the community provided that same service.

VI. Concluding Remarks

This work proposed a set of metrics that will be used to evaluate the grid services provisioned from demonstration projects in DOE's Connected Communities program. The goal is to contribute to the dialogue around what can and should be computed from demonstration projects in the future to enable grid operators and power systems regulators to gain confidence in these non-traditional resources for power systems planning and operations. Further, these metrics can support aggregators and community managers in developing business models for the successful implementation and management of communities of grid-interactive buildings and DER. Metrics were proposed that quantify the magnitude of service offers, the consistency and quality of services provided by the community, and the individual DER contributions to the community-level service, including novel applications of a mileage concept used for frequency regulation to reactive power services to help evaluate the burden of providing such a service on the resource. As the Connected Communities demonstrations collect data, further work will evaluate the efficacy of these metrics at generating the intended insight and alternative metrics may be proposed.

TABLE III
Grid Service Applications of DER Contributions Metrics.

Grid Services	**Indiv. DER Contribution Metrics**						
	1	2	3	4	5	6	7
Transmission							
Economic Energy Dispatch			•	•	•		•
Flexible Ramp Energy			•	•	•	•	•
Capacity (Peak Hours)	•	•					
Capacity (Ramping Hours)	•						
Contingency Reserves	•		•	•			
Frequency Regulation	•		•	•	•	•	•
Voltage Support	•		•	•		•	
Blackstart	•						
Distribution							
Capacity Relief	•	•					•
Emergency Load Transfer	•		•	•			•
Outage Recovery	•						
Voltage Management	•		•	•		•	•
Power Quality	•		•	•		•	•
Phase Balancing	•		•	•		•	•

References

[1] A. Satchwell, M. A. Piette, A. Khandekar, J. Granderson, N. M. Frick, R. Hledik, A. Faruqui, L. Lam, S. Ross, J. Cohen *et al.*, "A national roadmap for grid-interactive efficient buildings," Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), Tech. Rep., 2021.

[2] V. Olgyay, S. Coan, B. Webster, and W. Livingood, "Connected communities: A multi-building energy management approach," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2020.

[3] US Department of Energy, "Doe invests $61 million for smart buildings that accelerate renewable energy adoption and grid resilience," Available at https://www.energy.gov/articles/doe-invests-61-million-smart-buildings-accelerate-renewable-energy-adoption-and-grid (2022/11/02).

[4] D. Nemtzow, C. Regnier, K. LaCommare, N. M. Frick, and J. MacDonald, "If one geb is good, a community of gebs is better," in *Proceedings of the 2022 ACEEE Summer Study on Energy Efficiency in Buildings trivia conference*. Asilomar CA: ACEEE, 2022. [Online]. Available: https://escholarship.org/uc/item/5xh4m9jn

[5] Lawrence Berkeley National Laboratory, "Connected communities," Available at https://connectedcommunities.lbl.gov/projects (2022/11/18), 2022.

[6] J. Liu, R. Yin, L. Yu, M. A. Piette, M. Pritoni, A. Casillas, J. Xie, T. Hong, M. Neukomm, and P. Schwartz, "Defining and applying an electricity demand flexibility benchmarking metrics framework for grid-interactive efficient commercial buildings," *Advances in Applied Energy*, vol. 8, p. 100107, Dec. 2022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2666792422000257

[7] M. Shah, D. Cutler, J. Maguire, Z. Peterson, X. Li, J. Pohl, and J. Reyna, "Metrics and analytical frameworks for valuing energy efficiency and distributed energy resources in the built environment," in *Proceedings of the 2020 ACEEE Summer Study on Energy Efficiency in Buildings trivia conference*. Asilomar CA: ACEEE, 2020. [Online]. Available: https://www.nrel.gov/docs/fy21osti/77888.pdf