Summary of Energy Storage Control Performance Metrics

ESIC Energy Storage Controls Task Force

Project Manager
Andres Cortes

3002020109

December 2020

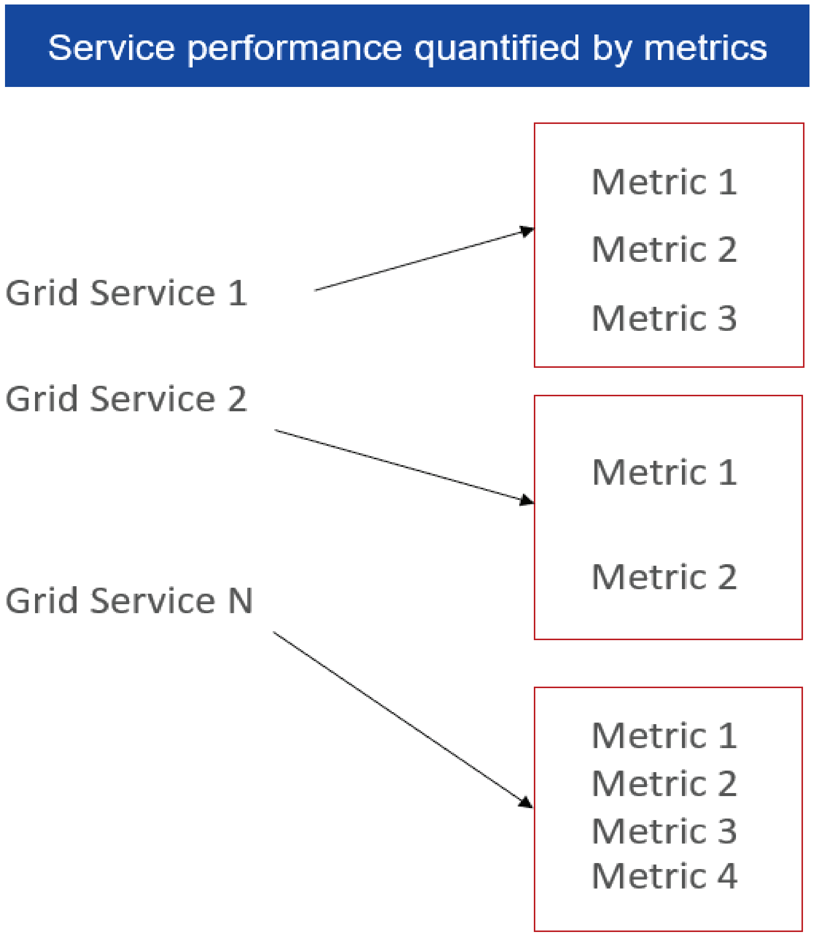
Overview

The value of energy storage is directly tied to the ability of the controller to meet the desired objectives. Therefore, it is important to be able to benchmark and compare controllers using a consistent set of metrics. This document introduces performance metrics that evaluate the performance of battery storage controls when providing different grid services. The objective is to create a living document that includes widely deployed applications of energy storage. This first publication includes use cases on capacity deferral, frequency regulation, and customer bill reduction. This ensures transparency and buy-in from all the stakeholders. The definitions and calculation methods presented have been vetted through EPRI's Energy Storage Integration Council (ESIC). Vendors, integrators, and other stakeholders have provided input and will serve as a resource for continuous refinement of these metrics and development of new metrics for additional use cases. Services like voltage support via Volt/VAR and primary frequency response via Frequency/ Watt, which are included in EPRI's common functions for smart inverters (EPRI Report 3002008217, 2016), IEEE 1547, and UL 1741, are covered in testing guidelines from the same standards. They will be included in a future release of this document.

Use Cases

In the context of this document, a use case of battery storage is a group of grid services that a storage system is designed to provide at its location in order to generate value to the power grid and to the owner/operator. The interaction between the grid and the storage system is also subjected to regulatory and technical requirements that might affect the way the system operates.

The operation of the storage system control can be evaluated by assessing all the technical objectives that must be attained to provide each service. These objectives are in turn associated with a mandatory action or economic metric that can be calculated based on variables that represent the system performance or a value. For example, to provide capacity services, a storage system may only need to reserve energy, without ever having to discharge. In this case, the metric that characterizes the performance must tie the availability of energy to when it is needed. Secondary metrics could address if the system was charged at a low cost or at a high cost to the operator, as well as how much degradation was caused on the system by the operation. Figure 1 shows the way in which the performance of a grid service can be characterized by a set of metrics.



Types of metrics for storage control performance

Non-economic metrics

Application technical compliance metrics: These metrics are tied to grid services for which the storage system is mandated to perform some specific action, e.g., discharge a given amount of power during certain time, maintain metered power below certain limit, respond to a droop control signal. Performance is characterized as pass/fail. The constraints might vary according to the system setup and use case. Violation of these constraints by a given percentage or range for a given duration might lead to issues in the grid, as well as regulatory penalties. For example, in the case of capacity deferral the net load must always be below the power limit. Violation of a constraint would be heavily penalized and might affect the system, therefore, the mandatory action metric cannot be monetized, because it must not be included as a business-as-usual action.

Application technical performance metrics: Application technical performance metrics characterize the performance of services that are not tied to an economic objective, but do not have a specific obligation to reach a performance threshold. As an example, variability of the power generated by a hybrid PV+ES plant can be an application technical performance metric, if there is no maximum variability requirement.

Economic metrics

These metrics represent the economics of the operation for a given use case. Operation of a storage system includes dispatch and power capacity reservations for some services. Economic metrics capture the costs and revenues that the system incurs given a specific operational profile. In the case of purely economic use cases, for example demand charge management, the customer bill represents the economic metric of performance of the system controller. Economic metrics are expected to be optimized, but there is no binding requirement to obtain any minimum or maximum score. It allows control developers to design strategies that consider different economic metrics to improve the economic performance of storage systems. Economic metrics can be connected to payments obtained from grid operators, savings with respect to a baseline operation, or operation cost, as is the case in the degradation-related metrics. Economic metrics are the ones that ultimately display the difference between the performance of a smart storage control and optimization strategy, and a poorly devised one, when both can satisfy the mandatory actions of the use case.

Metrics that apply to different services

There are performance metrics that apply to most grid services. The most important ones capture how effectively charging/discharging for State of Charge (SOC) restoration is being done, and how much is the storage system degrading due to operation. Table 1 introduces the description of the State of health reduction. It is a technology-agnostic metric. Note that storage availability will be captured by the ability of the system to perform grid services. For example, if the storage system fails, it might affect the ability to maintain the load below the system capacity.

Name	Definition	Calculation	Test duration	Benchmark
State of Health (SOH) reduction [Application technical performance metric]	Degradation of a storage system after operation during a determined time	SOH is defined here as a percentage variable, that goes from 100% when a storage technology is new, to 0% when the storage technology needs replacement. The replacement criterion can be set in terms of the loss of retention capacity, e.g., a battery needs to be replaced when it has reached 60% of its nameplate capacity. The loss of retention capacity can be measured for some battery technologies. The SOH definition here is a linear function of the loss of retention capacity.	From a technical standpoint, SOH can be metered for many technologies at any point in time. Therefore, SOH reduction could be metered within any timespan. The test duration will just include measuring at the beginning and the end of the test required for other metrics	A technology-agnostic benchmark would be zero degradation. The smaller the SOH loss, the more effective is the operation.

Peak Shaving

The nature of a capacity upgrade to be deferred can vary from case to case (e.g. feeder overload or transmission contingency). In all cases, the action of the approach being used for upgrade deferral is the same: maintain the power going through the overloaded component below a limit. The performance metrics listed below are sorted in order of priority. Since a grid capacity service is considered binding, maintaining the load below the limit value is the highest priority. State of health reduction and cost/revenue of energy take lower priority.

Notes: If the battery storage system is part of a fleet of resources (e.g., a virtual power plant (VPP)) providing the capacity upgrade deferral service, energy trade cost/revenue and constraint violation must be evaluated with respect to the performance of the entire fleet. This would also require evaluating metrics that are specific to other energy resources in the fleet: operational cost of fuel-based generation, running costs, degradation, etc.

Name	Definition	Calculation	Test duration	Benchmark
Component overload [Application technical compliance metric]	Integral of net load above thermal limit	$\int_0^T [Load_t - PV_t - ES_t - TL]_+ dt$ $Load_t$ is the load behind the overloaded component PV_t is any PV generation that can be measured separate from $Load_t$ ES_t is the energy storage dispatch TL is the rated limit of the component $[\cdot]_+$ represents the positive part of the expression in brackets	Depends on the time in which the constraint violation is analyzed. If tested with Hardware in the loop, the duration can span up to a whole year. If tested with real systems, it would require modifying the thermal limit to artificially create an overload, in which case it can be up to a few days.	Perfect case, no overload
SOH reduction	See description of SOH reduction in Table 1			
Cost/Revenue of energy trade [Economic metric]	Value of trading energy back to the grid during scenario duration	• $\sum_{t=0}^{T}\{Rates_{export}(t) * ([L(t) - PV(t) - ES(t) - themal(t)] < 0) - Rates_{import}(t) * ([L(t) - PV(t) - ES(t) - thermal(t)] > 0)\}$	Possible time frames ○ Few days ○ Month ○ Season ○ Year	Can be calculated using a valuation tool such as StorageVET, which provides a best-case scenario of the bill savings

Frequency Regulation Performance Metrics

Ability to perform in the frequency regulation market consists of two main components: following the Automatic Generation Control (AGC) commands and participating in the AGC capacity market. Following the AGC commands. When an energy resource is selected to provide AGC capacity, it is tasked with: 1. being synchronized with the grid 2. modifying its power dispatch according to the AGC commands, which are usually received every 3 – 6 seconds.

Name	Definition	Calculation	Test duration	Benchmark
Regulation accuracy [Application technical performance metric]	Regulation revenue based on accuracy, delay and precision. Reflected in payments for performance. A standard metric could be the mean squared error (MSE) between the actual dispatch and the AGC commands	Accuracy must be calculated according to the regulation performance formula for the balancing area. Refer to each individual ISO operations manual to see the exact metric of performance.	At least one hour.	The MSE for a system that has no delays and responds perfectly to the AGC signal is zero
Cost/Revenue of market trade [Economic metric]	Settlement of frequency regulation, DA and RT energy markets for the energy exchanged with the grid during frequency regulation	The calculation is provided in the ISO business practice manual. It includes the cost/revenue of energy throughput, the regulation capacity payments. Payment for performance is addressed separately.	The revenue must be estimated for at least one system operation cycle. For most regions it would be a day	A valuation analysis with a tool like StorageVET, with historical market prices could be used as a benchmark
SOH reduction	See description of SOH reduction in Table 1			

Customer Bill Reduction Performance metrics

Storage can be used to reduce customer's electricity bill. Storage can reduce the electricity bill by charging during times of low energy costs and discharging when energy costs are higher. Another way to reduce the bill is to manage the customer's maximum demand to reduce the monthly demand charge. In some cases, when the customer owns solar generation, but has interconnection constraints, e.g., no export constraints, a control system will try to store all the excess solar generation to use it at later hours, reducing the demand for energy from the grid.

Name	Speed requirements	Monitored variables	Comments
Customer bill savings (Energy and demand charges)	Demand charges are typically estimated on the maximum 15-minutes average demand of the month. Then, the monitoring rate for testing this service will be 15-minutes. Energy charges in a Time of Use (TOU) tariff would change every 30 minutes. Therefore, the minimum metering requirement would be 30 minutes	• Customer load • Storage dispatch • Customer net load at the point of interconnection	• Metering rate depends on metering devices • Generally, it does not need a high sampling rate • Possible sampling rates: 1 min – 1 hr
SOH reduction	See SOH reduction monitoring considerations in Table 2		

Voltage and Active/Reactive Power Control

Fast dispatch of active or reactive power according to local voltage measurements. The storage system must modify its power up or down according to the function used for voltage control. If the function is Volt/VAR, the storage system must modify its reactive power according to a Volt/VAR curve. If the function is Volt/Watt, the storage system must modify its active power according to a Volt/Watt curve. In this service, the storage control is concerned with guaranteeing that the Volt/VAR or Volt/Watt setpoints are followed, instead of whether the local voltage is kept at the adequate values. It is the responsibility of the grid operator to determine the parameters for the control curves to ensure appropriate voltage regulation.

Name	Definition	Calculation	Timeframe	Benchmark
Voltage control accuracy [Application technical performance metric]	Error measure between the power setpoint provided by a Volt/VAR (Volt/Watt) curve and the power delivered by the storage system	The control accuracy is characterized as the MSE between the active and reactive power setpoints or power factor, and the metered power (power factor)	The minimum duration of a voltage control test must be between a few minutes and a few hours, depending on whether real-time voltage measurements are used. The test might need to be longer in case of high voltage stability	The performance benchmark is to provide the power dictated by the setpoint immediately. In that case, the error should be zero

Frequency Droop

Fast dispatch of active or power according to local frequency measurements. The storage system must modify its power up or down according to a frequency/Watt curve. In this service, the storage control is concerned with guaranteeing that the frequency/Watt setpoints are followed, instead of whether the local frequency is kept at the adequate values. It is the responsibility of the grid operator to determine the parameters for the frequency/Watt curve to ensure appropriate frequency control.

Name	Definition	Calculation	Timeframe	Benchmark
Frequency control accuracy [Application technical performance metric]	Error measure between the power setpoint provided by a frequency/Watt curve and the power delivered by the storage system	Frequency control accuracy is calculated as the MSE between the power setpoint given by the frequency droop characteristic and the metered storage power	Typically, between a fraction of a second and a few seconds	The performance benchmark is to provide the power dictated by the setpoint immediately. In that case, the error should be zero

Renewable Smoothing

Renewable smoothing consists on actively charging and discharging a storage system to compensate for the variability of a renewable generator. It is expected that the net generated power by the renewable + storage presents a lower variability than that of the renewable alone, according to any of the existing variability metrics. In the following table, three common metrics are introduced. Each metric may be used to pose a different objective for the control system, but in general, it would be expected that if an algorithm reduces one metric, it will also reduce others.

Name	Definition	Calculation	Timeframe	Benchmark
Ramp rate excess [Application technical compliance metric]	Average ramp excess seen on the combined renewable + storage power during a test	Over a time-series of metered data collected faster than 10 seconds resolution, ramps are calculated as the difference between two samples separated by 1 minute. The ramp excess is calculated as the maximum between 0 and the difference between the ramp magnitude and the maximum allowed ramp. The average ramp excess is calculated as the average of ramp excess estimates seen in the signal.	It is desired that a renewable smoothing scenario be tested over days with different variability. A set of three days with different levels of variability, including a clear day can be used for the test.	The perfect performance in this case is a ramp rate excess of zero, meaning that all ramps are below the specified limit.
Variability index [Application technical compliance metric]	Arc length of the generation signal (movement) normalized by the arc length in a clear sky. Originally defined for solar irradiance, the normalization term can be dropped or replaced for a more suitable number.	For a time-series of metered generation data with N measurements: $\frac{1}{K}\sum_{i=2}^{N}\sqrt{(P_i - P_{i-1})^2 + \Delta t_i}$ where K is a normalization constant.		If normalized by generation in a clear sky, the perfect performance would be 1. For other normalizations, it may differ.
Maximum power sag [Application technical compliance metric]	Maximum power drop in kW/min with respect to the average from the last N generation samples.	For a time-series of metered generation data with N measurements, the maximum power sag is the maximum difference between a measurement and the average of the last N samples, divided by the sampling period in minutes		
SOH reduction	See description of SOH reduction in Table 1			

Renewable Firming

The renewable firming service consists in supporting a renewable generation to behave as a relatively dispatchable resource. The storage system must be able to maintain the power from the renewable + storage plant to be above a given threshold. Moreover, a smart controller would be expected to estimate how much power capacity can be delivered considering the renewable resource and the storage system size.

Name	Definition	Calculation	Timeframe	Benchmark
Power capacity under-delivery [Application technical compliance metric]	Given a time period, where at any moment, certain amount of power must be available for delivery, under-delivery is the integral of the power below pre-agreed value	$$\frac{1}{T}\int [P_G(t) - P(t)]_+ dt$$ $P(t)$ is the power metered from the combined storage + renewable at time t, $P_G(t)$ is the guaranteed power expected to be available at time t, and T is the number of intervals in the schedule of guaranteed power capacity	At least few hours to few days.	The perfect amount of under-delivered capacity is zero.
Value of delivered capacity [Economic metric]	Given a price of capacity in $/kW (MW) for any time interval during a time period, the value of delivered capacity is the aggregate value of providing an amount of promised capacity during the period.	The value of delivered capacity is calculated as: $$\sum_{i=1}^{T} P_G^i * c_i$$ where P_G^i is the guaranteed power capacity delivered during time interval i, and c_i is the price in $/kW to be accrued for delivering capacity during time interval i		The perfect amount of firm capacity that a renewable+ storage resource can offer could be estimated with an optimization model like DER-VET to be used as a benchmark.
SOH reduction	See description of SOH reduction in Table 1			

Arbitrage

Energy arbitrage consists in storing energy at times when there is excess of supply or low price, to release it when there is scarcity of supply or high price. Generally, arbitrage can be characterized according to a price signal in $/kWh which varies over time. Performing successful arbitrage implies being able to recover the cost of efficiency losses in the storage system, and in the case of Li-ion technologies, the cost of degradation associated with the charge/discharge cycle

Name	Definition	Calculation	Timeframe	Benchmark
Energy trading revenue [Economic metric]	Total revenue obtained from buying and selling energy at the settlement prices determined by the market that applies at the storage system location	The revenue is calculated as the sum of the hourly energy delivered multiplied by the settlement price in $/kWh for that hour: $$Revenue = \sum_t c_t * e_t$$ Where e_t is the energy traded and c_t is the energy price at time period t	At least one performance period, generally one day. If energy prices significantly, it is recommended to run tests for different day types	The benchmark for this economic benefit would be the perfect forecast analysis of energy arbitrage using StorageVET or a comparable benefit estimation tool
SOH reduction	See description of SOH reduction in Table 1			

Renewable Self-Consumption

Under some interconnection agreements of BTM DERs, it may happen that the utility requires that the customer does not export power to the grid. This measure is aimed at avoiding reverse power flows and reducing voltage issues that can occur with renewable generation. Depending on the size of the renewable system, there might be times during which the generated power is greater than the site load, forcing the renewable controller to curtail or shut power down. Energy storage can be used to recover the energy that would have been otherwise wasted had power been curtailed.

Name	Definition	Calculation	Timeframe	Benchmark
Curtailment reduction [Application technical performance metric]	Curtailment reduction can be captured as the amount of energy generated on-site recovered by the storage system that would have been curtailed otherwise. Alternatively, it could be the market value of such energy	The calculation is made using measurements of locally generated power, local load, and the storage power dispatch over the analysis period: $$\sum_t \min\{\max\{0, p_t^{gen} - L_t\}, p_t^{ES}\}\Delta t$$ Where p_t^{gen} is the power locally generated, L_t is the site load, and p_t^{ES} is the power charged into the storage system	At least one day	The benchmark for curtailment reduction is to recover all the solar power that would have been curtailed.
SOH reduction	See description of SOH reduction in Table 1			

Together...Shaping the Future of Electricity

Disclaimer of Warranties and Limitation of Liabilities

NOTE

For further information about EPRI, call the EPRI Customer Assistance Center at 800.313.3774 or

e-mail askepri@epri.com.

The Electric Power Research Institute, Inc. (EPRI, www.epri.com) conducts research and development relating to the generation, delivery and use of electricity for the benefit of the public. An independent, nonprofit organization, EPRI brings together its scientists and engineers as well as experts from academia and industry to help address challenges in electricity, including reliability, efficiency, affordability, health, safety and the environment. EPRI also provides technology, policy and economic analyses to drive long-range research and development planning, and supports research in emerging technologies. EPRI members represent 90% of the electricity generated and delivered in the United States with international participation extending to nearly 40 countries. EPRI's principal offices and laboratories are located in Palo Alto, Calif.; Charlotte, N.C.; Knoxville, Tenn.; Dallas, Texas; Lenox, Mass.; and Washington, D.C.

Together…Shaping the Future of Electricity

3002020109

Electric Power Research Institute

3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com