Data Quality: The Key to Storage Operation

Energy storage installations have been ramping up considerably in recent years. In 2017, only 288 MW / 645 MWh was added. Five years later, in 2022, 4.8 GW / 12.1 GWh was brought online. The 2022 total installations were nearly equal to the totals in 2020 and 2021 combined, as deployment rates continue to increase. The U.S. Energy Storage Monitor from the American Clean Power Association ("ACP") and the Wood Mackenzie indicates the U.S. may add as much as 75 GW of energy storage between 2023 and 2027 to meet the needs of an evolving grid.[1]

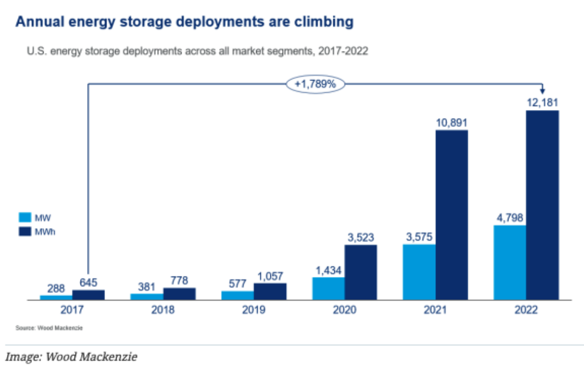

Figure 1 – Annual Energy Storage Deployments

Underlying this build-out is the asset's performance in its respective role, delivering multiple services to various markets. A Battery Energy Storage System ("BESS") will need to perform reliably over its lifetime to deliver value to the grid and help support the energy transition. BESS can play many roles and is starting to take centerstage. Whether contracted for energy delivery ($/MWh), capacity ($/kW-mo), or participating in merchant markets, an accurate view of its operational state is critical.

"If you can't measure it, you can't improve it" is a quote often attributed to Peter Drucker as a fundamental tenant of business management. However, to realize improvement, data quality is a necessary part of a measurement. Measurement quality, when overlooked, can lead to incorrect outcomes, and may impede improvement. Knowing what is being measured, its context, and the stack-up of potential error-inducing steps from the sensor to the final calculation will help achieve improvement goals.

BESS performance measurements are often buried at the end of a long chain of transformations originating from a black box that may not be well understood. Equipment suppliers may have intellectual property wrapped up in these black boxes (e.g., Battery Management Systems ("BMS"), plant controllers, etc.) so the details are not exposed or available for in-depth examination. A common BESS performance measurement, State of Charge ("SOC"), is a basic indicator of the BESS's status and is used in several important analytics including degradation calculations, compliance with wholesale market participation requirements, and charging/discharging limits. SOC determination typically consists of a calculation based on the measurement of battery cell voltage, which then gets transformed through data acquisition (e.g.,

[1] https://cleanpower.org/resources/u-s-energy-storage-monitor/; accessed on 5/31/23.

analog to digital conversion, etc.), data sampling with data aggregation (e.g., averaging, max/min, etc.), digital manipulation to change units of measurement or scaling, data transport protocols involving conversion of data types, and data cleaning to fill in missing datapoints or outliers (imputing and resampling). This is all before the actual calculation is executed to arrive at SOC, which is then compared with a model or use of other machine-learning techniques to apply statistical probability methods to improve the calculation. At each point along this chain of conversion and manipulation is a chance for error from one step to be compounded with further errors, known as error stack-up. This poor data quality can lead to incorrect results, or more simply, “garbage in = garbage out”.

In today’s complex energy markets, dispatch of a BESS asset is highly dependent on accurate data. Use-case models, bidding strategies, augmentation schedules, predictive maintenance, and operational decisions require a constant stream of high-quality data to provide insight to asset owner/operators regarding BESS performance, reliability, and safety. When the underlying data is questionable, the outcomes of critical estimations like lifetime modelling, optimization of key process indicators, and thermal safety cannot be relied upon, resulting in project funding and permitting challenges, asset transaction barriers, and market participation headwinds.

Battery cells within lithium-ion modules are the beginning of the BESS data chain and the most common analog signals used by the BMS are voltage, current, and temperature from individual cells or modules along with other signals like gas detection, water detection, or pressure. These signals form the basis of calculations utilized to operate the BESS like module balancing, thermal safety, capacity, state of health (“SOH”), and SOC. In the power conversion system (“PCS”), the inverter will have a similar set of low-level signals coming from components like the power blocks and voltage/current transducers to insure power limits, discharge/charge operation (i.e., direction of current flow), as well as phase detection, and synchronization. Higher functionality such as frequency response through internal droop curve calculations and coordination across module stacks and containers will also rely on the accuracy of these data streams. Ensuring calculations like SOC are accurate will allow optimization of BESS dispatch as well as maintaining operational safety. Data quality starts with low-level signal quality but enables overall BESS performance, reliability, and safety.

Data Quality Issues

Data quality issues along the measurement chain fall into the following categories:

- Sensors
- Process Understanding
- Data Transport
- Data Manipulation
- Data Interoperability

Looking at quality issues for a critical sensor, like cell voltage measurement, issues start with uncertainty in the measurement and needs to consider both accuracy and precision. The error in accuracy, or the alignment of a measurement to a known, verifiable quantity, along with the error in precision, or reproducibility of the measurement, are part of the sensor selection and are often buried deep inside the BESS. Issues like sensor hysteresis, sensitivity/resolution, and environmental effects will create the first challenges in achieving a quality measurement.

Process misunderstanding can lead to mismatched or incorrect sensors for the measurement where issues related to process response, radiated, conducted noise, and signal range can lead to error.

Data transport from the sensor to the first collection point may involve analog-to-digital conversion and differential signal processing, introducing errors related to sampling rate, resolution or quantization, and electrical noise.

Data manipulation is one part of the measurement chain where errors can be filtered out or amplified. Rounding or truncation errors along with scaling or other non-linear manipulation often involves digital filtering and processing, which may rely on a physical model not taking error stack-up into account and resulting in error propagation.

Data interoperability can often undo the quality of a measurement by introducing offsets from communication/network mismatch, data storage mismanagement, or even synchronization errors with timestamps.

At any point along this measurement chain, small errors of fractions of a percent can be compounded throughout the stack-up resulting in large errors which can invalidate calculations based on the manipulated data. Assumptions errors are small or stochastic and they will cancel out over some period can lead to mis-identifying error sources and make error mitigation more difficult.

Figure 2 shows the signal error stack-up for a simple cell voltage measurement used for an SOC calculation. The SOC calculation is then used by the BMS and energy management system ("EMS") for BESS operation as well as providing the plant historian with information for further analytics. At each stage of the signal-to-data chain we can see some of the typical sources of error or mismatch.

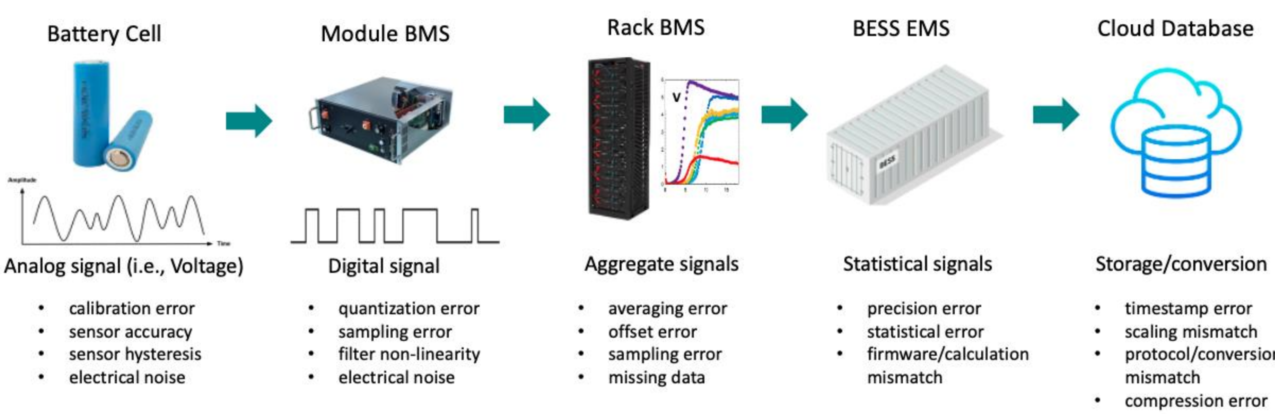

Figure 2 – BESS Signal Error Stack-up

Data Quality Management

Data quality management has been defined in many industries and multiple policies, procedures, and workflows have been established to realize it. For example, ISO 8000-1:2022 "Data Quality"[2] covers many aspects of organizational data quality management. Most data quality management approaches involve a process aligning to the following traits:

- Timeliness
- Accuracy
- Reliability
- Relevance
- Completeness

Looking at the previous example of an SOC calculation requiring a battery cell level voltage measurement, we can see these five traits apply to all the steps along the way.

Timeliness requires the voltage signal is sampled at a high rate (i.e., above the Nyquist frequency) to avoid aliasing effects in the data. For measurements of fast changing signals this sampling frequency can be in the kilohertz range.

Accuracy starts with calibrating sensors under the required conditions (e.g., temperature, humidity, etc.) but can include offsets and modelling curves applied to linearize the signal. Managing these methods for traceability (e.g., NIST Standard References[3]) and verification of assets in the field can be challenging, especially when the sensors are buried within proprietary systems. Equipment vendors can be helpful resources for ensuring proper calibration and monitoring, where such activities may be required due to contractual or warranty conditions.

Data quality does not exist just in the lab and needs to be addressed across the signal path to ensure signals are reliably managed from sensor to final calculation. Processes including testing and verification of standardized methods like data normalization, manipulation, and storage/transmission are more likely to catch data inconsistencies resulting from poor data quality practices. Calibration may only be good for a short time and won't be accurate if ignored for long periods of time. Data quality metrics including detection of data drift and other trends can help ensure standard processes are being followed.

Relevance of these data quality metrics is important to ensure the right indicators are being measured and tracked. Knowing a sensor is accurately calibrated and the measurement is normalized properly is imperative, but if the only measurement comes from the point of interconnection ("POI") and the plant has many different BESS containers, it is hard to guarantee plant health from a single measurement.

Data completeness can cover many different issues, from missing datapoints in time-series data to estimated values or constants used in calculations where measured values could or should be used. Assumptions about linear estimations or hysteretic behavior of measurements can

[2] https://www.iso.org/obp/ui/#iso:std:iso:8000:-1:ed-1:v1:en; accessed on 5/31/2023.

[3] https://www.nist.gov/standards; accessed on 5/31/2023.

also lead to issues when, for example, the asset under measurement is spending most of its time in a non-linear region.

Not all assets will have an interface to the granular levels required to validate signal integrity along the complete data chain. Most BESS owner/operators will have access to data down to the container level but may not be able to access individual cell voltages. Additionally, they may not have access to the calibration schedule for a current sensor buried deep within the BESS. Other equipment suppliers may only provide plant level data or data aggregations defined by the plant controller. One piece of the data chain will probably come with a documented calibration procedure as well as a NIST certificate for accuracy, the revenue meter or POI meter. Using an accurate (and traceable) meter as a reference standard for other measurements is one place to start with data quality validation.

The elements of a data quality management system can take many forms depending on the specific workflows and available documentation from the construction phase of the project:

- **Quality Control Requirements:** Comprehensive documentation regarding internal quality processes and standards for equipment installation, operation, and maintenance.
    - This includes change management for hardware and software.
    - Key instrumentation is defined.
- **Minimum Accuracy and Repeatability Standards**: Operator and equipment supplier define minimum accuracy and repeatability requirements for key instruments.
- **Reliability:** Reliability requirements are detailed for key instruments along with targets for the measurement system. The communication methods, environmental factors, and other aspects affecting reliability should be considered.
- **Calibration and Field Verification Standards:** Calibration and field verification standards are defined for key instruments and measurements. This includes schedules for performing calibration and field verifications, documentation of results, operating ranges, adjustments, scaling, and other modifications.
- **Data Storage and Transmission:** Methods, protocols, and other definitions needed for the transmission and storage of data are detailed.
    - Data characteristics such as data storage architecture, archival requirements, correction options, and other needs are detailed.
    - Data service and quality metrics are defined.
- **Data Transformation:** Key measurements taken from data transformations are detailed along with formulas, signal processing, sampling

rates, fill rules, corrections, and other data manipulations.

- **Interoperability and Data Integration:** A method for synchronizing key instruments and measurements is detailed. This includes how to pass time stamps through additional data aggregator systems to track latency and other issues affecting the end use of the measurement.
- **Taxonomy and Units of Measure Standardization**: Methods and requirements are defined by the operator and equipment supplier including key measurements and instruments passed through from other data systems from third party key measurements.

The BESS owner/operator will benefit from working closely with the equipment supplier to implement a data quality assurance program, ideally addressed as early as possible in the project lifecycle to avoid potential retrofitting exercises. A review of available data should be part of the contractual negotiation process and should include an overview of expected calibration and maintenance activities to support data quality.

Additionally, the data quality assurance program should align with the data usage/ownership sections of applicable agreements, especially those related to performance and warranty guarantees. For example, if the capacity is defined by a test procedure, as it often is, there will be conditions with requirements such as maintaining a certain SOC. Without calibration, this level could be off by several to tens of percent. One major equipment supplier has published calibration guidelines for SOC (reported by its software) claiming values can have up to a 50% error if the calibration procedure is not performed on a regular schedule.

Data Quality Myths

1. *"You need 100% data quality for good analytics."*
   -By its very nature, data analytics is designed to work with messy data. The goal is to provide insight from data that may be unstructured, uncorrelated, incomplete, and not normalized. Forbes reported[4] data scientists spend up to 60% of their time cleaning and preparing data prior to its use in any analytics or modeling/machine learning. At NexESS, our objective is to detect and remove anomalies prior to use in any algorithms so we are confident non-relevant information is not affecting analytical or model outcomes. That way, we ensure we are using the most appropriate datasets. Additionally, we employ up-sampling/down-sampling methods to align the datasets with the desired frequency and interpolate across data gaps using intelligent techniques. The result of

[4] Press, Gil. "*Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says*"; accessed 11/09/2020 on Forbes.com.

this data cleaning is what we call "curated data". This allows NexESS to bring big data to life so our machine learning algorithms can identify key events and trends which can be further mined for additional insight into performance and reliability. 100% data quality is a good goal to strive for, but don't fall into the trap where perfect is the enemy of good.

2. *"All of your data is a business asset."*
   -Data is a business asset when it is managed, secured, normalized, and utilized. It is a liability when its volume grows beyond standard means of access, its complexity becomes limiting to timely and repeatable cleaning techniques, security is compromised, or it becomes financially unrealistic to utilize it for operational or strategic purposes. NexESS believes in the power of big data when it is supporting business initiatives like performance, reliability, and safety. One second resolution data may be important for real-time frequency analysis, but it can probably be moved to secondary storage (i.e., long-term and cheaper) after a reasonable amount of time, so it is no longer part of "database bloat". NexESS uses a combination of techniques like micro-partitioning, columnar datastore, and archiving to keep database access nimble.

3. *"All data is important and valuable."*
   -Data value is dependent on several variables like context, lifecycle status, and operational/regulatory impact. Context drives value by focusing on big data's "4 V's"[5]: Velocity, Volume, Veracity and Variety. The right data, at the right time, in the right format, and verifiable is fundamental to analytics providing a return in the form of performance, reliability, and safety. Lifecycle status refers to what stage the data is currently in. When data is captured, the value is smaller as it may not be used for anything or may be strictly compliance based and warehoused. Once data moves beyond capture to integration, manipulation, and decision support (i.e., it's used for insight and improvement), it adds more value as it creates a more competitive edge. Stakeholders are key to identifying the key process indicators that will drive better decision-making.

4. *"Data quality is expensive."*
   -It's probably more correct to say bad data quality is expensive as more resources go into cleaning it up. Good data quality is the foundation for good analytics, which drives better decision-making and results in better business outcomes. Forrester found businesses relying on data driven decisions were

[5] Cerniauskas, Julius, *"Understanding the 4 V's of Big Data"*; accessed on 6/30/2023 at forbes.com.

almost three times more likely to achieve double-digit growth.[6]

Data Quality in Practice

The elements of a data quality assurance program may seem complicated at first, but quickly becomes reasonable when compared to what is left on the table if data quality is ignored. From calculation agents who ensure offtake payments are correct to asset managers who are making portfolio and project operational decisions to the trading desk (or algorithm) making decisions to participate in a merchant market - across the energy storage landscape, data quality is foundational to supporting the grid today and tomorrow. Investing in data quality early in the project lifecycle will pay dividends immediately and for the lifetime of the asset.

Revisiting the SOC example, we can see that a small error at the BMS level can propagate to a larger differential in revenue or in the worst-case scenario could lead to penalties/liquidated damages. SOC is a fundamental calculation used from pro-forma to dispatch, leading to widespread consequences when errors are introduced due to missed or non-existent calibration procedures. Operation of a complex asset like a BESS depends on the details, including calibration and periodic maintenance, to ensure that an asset continues to perform to expectations.

Data quality not only affects day to day operations but can also have effects across the project lifecycle. For example, when energy storage assets transact, a solid history with normalized data that can be accessed to demonstrate viability will reduce complexity in decision making.

Optimizing energy storage assets for performance, reliability, and safety is a proven method to reduce overall project risk and contributes to differentiating your assets in today's energy economy. Starting with a solid foundation of quality data is one of the many pieces to make sure your BESS does not become a case study in missed opportunity.

[6] Evelson, Boris, "*Data, Analytics and Insights Investments Produce Tangible Benefits – Yes, They Do*"; accessed 6/20/2023 at forrester.com.

About the Authors and NexESS

Sean Halloran, CTO/Co-Founder: Sean Halloran is a co-founder of NexESS Analytics and a veteran of energy projects around the world.

Mr. Halloran has served in a variety of engineering and leadership positions over his career, from a successful thin-film solar startup eventually acquired by General Electric to a global energy services company. He has hands-on knowledge of engineering complex products, systems, and transactions from energy storage applications to technology M&A.

Todd Tolliver, CEO/Co-Founder: Todd Tolliver is an energy industry professional with a combined experience of over 20 years in the fields of photovoltaics, energy storage, and component manufacturing.

Mr. Tolliver has managed projects with scope including research and development, process scale-up, product manufacturing, project finance due diligence, and technical advisory services. He has experience in renewable energy and energy storage due diligence, including product design and performance, supporting more than 2 GW of solar projects and more than 11 GWh of energy storage projects. Prior to the founding of NexESS Analytics, he developed and led ICF's independent engineering and owner's advisory energy storage practice.

NexESS Analytics: NexESS Analytics is a privately owned company offering BESS consulting services in the form of pre-COD advisory support, commissioning support, and post-COD periodic performance reporting and analytics. NexESS leverages its greater than 25 years of storage experience to provide advisory services ranging from early-stage project development support through BESS commissioning. Post-COD performance reporting is driven by NexESS's BESSiq™ analytics platform, which analyzes BESS operating data to understand the effect of dispatch on short- and long-term reliability. Additionally, the NexESS team evaluates performance against warranties, guarantees, offtake requirements, and market participation expectations. The NexESS team has experience with major BESS equipment suppliers, a variety of lithium-ion chemistries, alternative storage technologies, and wholesale market requirements.

