Connecticut Green Bank
Docket No. 26-08-05
Response to Interrogatory CAE-7
Date of Submission: July 9th, 2026
Witness: Connecticut Green Bank - Sergio Carrillo

Question:

CAE-7 Comment on whether any EDC DERMS data received by CGB contains material errors that affect data integrity or usability. If so, describe the processes and methodologies CGB employs to identify, clean, and validate such data.

Response:

As the only Program Administrator with access and oversight of data from the DERMS platforms, interconnection applications, and incentive applications, Green Bank has invested significant time and resources into building a robust system for synthesizing data from these various sources into a clean and validated "system of record" for telemetry and enrollment data for the Energy Storage Solutions Program.

Green Bank Data Cleaning Findings

Green Bank receives telemetry data from the residential and one EDC's C&I DERMS providers on a rolling basis, as available by the Original Equipment Manufacturer (OEM). Before any external reporting, all DERMS data passes through a multi-stage data warehousing process with schema normalization, deduplication, sign/timezone correction, canonical-name enforcement, referential joins to Green Bank enrollment data, and automated validation/audit steps.

Based on the Green Bank's review of the data submissions, it is possible to identify certain data issues that can be addressed through data cleaning, reconciliation, or validation, as well as other issues that cannot be fully corrected without additional information or action from the applicable OEM or DERMS provider. For example, some of those issues are the same ones identified by Customized Energy Solutions below but also include:

1. Format and structural errors: timezone inconsistencies, interval granularity, duplication in source feeds, sign-conventions
2. Missing intervals or intervals with missing key data points (namely the "Stored Energy" and "State of Charge" fields)
3. Schema drift: Uncommunicated changes in table formats, data format convention changes, inconsistent data validation on key fields (especially Utility Account number, which is essential for tying together DERMS data, interconnection data, and incentive enrollment data), etc.

Customized Energy Solutions Data Cleaning Findings

Connecticut Green Bank
Docket No. 26-08-05
Response to Interrogatory CAE-7
Date of Submission: July 9th, 2026
Witness: Connecticut Green Bank - Sergio Carrillo

Customized Energy Solutions noted in their Annual Report[1] that there are several data quality and standardization issues that required review, processing, and validation before the data could be relied upon for analysis. These activities were intended to improve consistency, address invalid or inconsistent values, and support the integrity of the resulting dataset.

- Aligning fields across datasets provided in different formats or from different sources;
- Correcting invalid data points related to available energy;
- Reviewing telemetry data for consistent sign conventions, including ensuring that discharging energy was represented as a negative value; and
- Collapsing data into uniform time intervals to support consistent analysis.

The report also identified broader data quality concerns that may affect the efficiency, comparability, and validation of DERMS-related data, including:

- Inconsistent or poor-quality state-of-charge and stored-energy fields, which are important because those fields may be used to assess potential upfront incentive clawbacks and provide normalized assessment of Active performance;
- Different OEM reporting formats, in part due to varying requirements across jurisdictions, which limits standardization.

The Annual Report also notes that, during Winter 2024/2025, there was a gap in DERMS provider coverage for certain C&I data. As a result, telemetry data for that period was obtained directly from OEM rather than through a single DERMS aggregator, which further increased the need for data reconciliation and quality review.

Green Bank Data Cleaning Processes and Methodologies

As mentioned previously, Green Bank maintains a dedicated data warehouse for use in Program management, Evaluation, Measurement, and Verification (EM&V), and public reporting. Every day, various Application Programming Interfaces (APIs) and Secure File Transfer Protocol (SFTP) sites are queried for the most recent data (by default the

---

1 https://www.dpuc.state.ct.us/dockcurr.nsf/8e6fc37a54110e3e852576190052b64d/b19880926e6dd18d85258d250066b224/$FILE/Docket%2025-08-05%20Order%203%20CT%20Energy%20Storage%20Solutions%20Annual%20Evaluation%20Report%20October%2016%202025.pdf

Connecticut Green Bank
Docket No. 26-08-05
Response to Interrogatory CAE-7
Date of Submission: July 9th, 2026
Witness: Connecticut Green Bank - Sergio Carrillo

prior 30 days, though further lookbacks occur regularly based on need) from the DERMS providers, and new or updated data is inserted into this data warehouse. Separate layers for data imports, data normalization/cleaning, and final reporting are utilized to catch issues early and alert Green Bank staff for investigation and correction. Stringent deduplication & idempotency rules, automated audit assertions, and use of database level referential integrity systems that prevent incorrect or excessive information from being shared, as well as regression testing practices and independent third-party data quality assessments (conducted by CES), all contribute to Green Bank's efforts to create a high-quality "telemetry of record" dataset for use in Program evaluation and reporting (though it is not used for Active Dispatch incentive payments currently). Data availability checks rely on a scaffold built from enrollment data, not from telemetry records, so that data availability is only tracked during relevant timespans. As noted in CAE-1 having data gaps in EDC PowerClerk reports impacts more than the self-enrollment process. Various internal dashboards, workbooks, and alert systems have been built to help support Green Bank staff in these efforts.

When issues with telemetry are discovered, such as missing telemetry or implausible battery energy storage system (BESS) measurements, Green Bank works with the relevant DERMS provider and OEM to understand the issue and, if possible, have it corrected in the DERMS source system via backfills from the OEM to the DERMS to ensure that all parties are working with the most recent versions of the data. In rare instances where the DERMS provider has been unable to accept a backfill, Green Bank has taken interval telemetry data directly from the OEM for integration into the telemetry of record for the Program.

Beginning with the Summer 2026 season, Green Bank shares Passive performance reports with OEMs and BESS Operators on a weekly basis. After Green Bank ingests and performs the initial QA/QC outlined above, an automated scoring process in Green Bank's data warehouse scores each Passive hour for each Passive project on a rolling basis using the Passive Performance Formula outlined in 6.2.1 of the Program Manual[2]. Green Bank filters by OEM and exports these tabulated scores to CSV files and shares each file with the respective OEM/BESS Operator responsible for performing Passive dispatch. These reports allow OEMs/BESS Operators to compare their source device telemetry against the telemetry Green Bank receives through the DERMS. Further, it gives OEMs/BESS Operators a framework to discuss performance issues and differentiate which performance issues may be related to device programming,

[2] https://energystoragect.com/wp-content/uploads/2026/05/Energy-Storage-Solutions-Program-Manual-Revised-for-05-19-2026.pdf

Connecticut Green Bank
Docket No. 26-08-05
Response to Interrogatory CAE-7
Date of Submission: July 9th, 2026
Witness: Connecticut Green Bank - Sergio Carrillo

enrollment issues, and data delivery issues between the OEMs/BESS Operators and the DERMS.

To aid in the discovery of potential telemetry issues, Green Bank and CES have developed a variety of month-level and event-level "flags" to surface issues in internal reporting tools. Some examples of these flags are:

- **"Has Telemetry" Flag:** Checks if telemetry is present during every expected 15-minute interval during the period evaluated
- **"Sign Reversal" Flag:** Checks for directional contradictions between the State of Charge (SOC) field and power/energy sums during 15-minute intervals using rolling window functions
- **"Expected Average Power" Flag**: Implemented against both SOC and Nameplate capacity, these flags check whether the device exported the expected amount of energy based on the discharge rating of the device during dispatch events, while taking into consideration event length and power vs energy limited systems
- **"Null/Missing Stored Energy" Flag:** Checks for events where the first interval of the event has no usable stored-energy reading

In certain cases, these flags are directly used to ensure downstream data quality in various internal and external reports. For example, when calculating hourly compliance scores for systems enrolled in Passive Dispatch, the presence of a Null/Missing Stored Energy Flag for a Passive Event will trigger the substitution of a known engineering rating, the device's nameplate energy capacity (minus 20%), into the Passive Dispatch formula to ensure that the missing data does not corrupt the Passive score for that day. This flag is then appended to the report to ensure any internal or external users are aware that missing telemetry data affected the calculation of that device's compliance score during that event.

While there are frequent occasions where data from OEMs and the DERMS providers has material errors in it, Green Bank has developed a robust system for analyzing, investigating, and correcting material errors in this data to ensure accuracy and consistency in the data used for program evaluation, passive clawback calculation, external reporting, and validation/auditing of performance-based incentive payments.