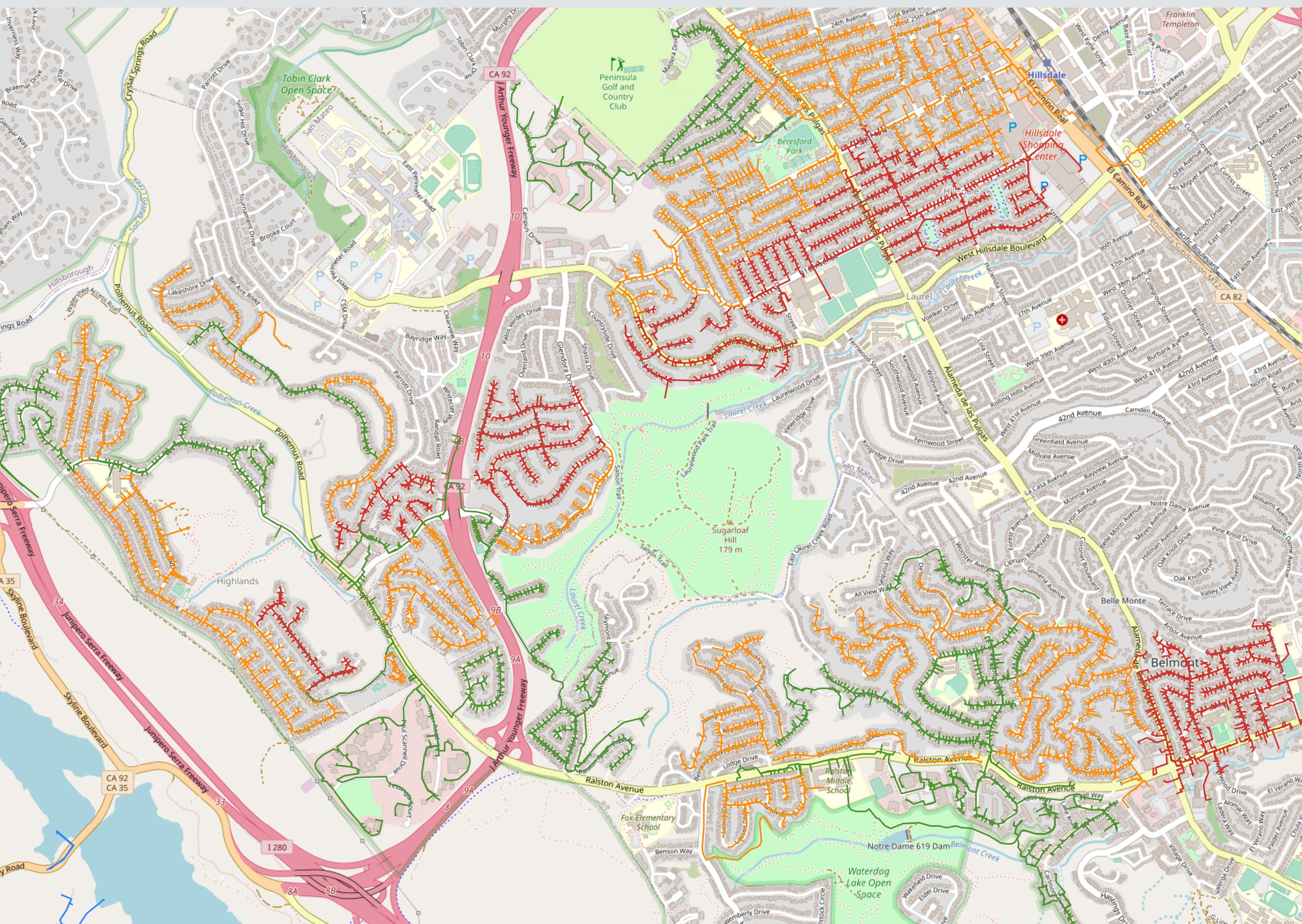
NREL Transforming ENERGY

IREC

Data Validation for Hosting Capacity Analyses



Data Validation for Hosting Capacity Analyses

Authors

Adarsh Nagarajan[1] and Yochi Zakai[2]

[1] National Renewable Energy Laboratory

[2] Shute, Mihaly & Weinberger LLP, attorney for Interstate Renewable Energy Council

APRIL 2022

SUGGESTED CITATION

Nagarajan, Adarsh and Yochi Zakai. 2022. *Data Validation for Hosting Capacity Analyses*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A40-81811. https://www.nrel.gov/docs/fy22osti/81811.pdf.

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Support for the work was also provided by the Interstate Renewable Energy Council, Inc. under Agreement SUB-2021-10440.

Preface

This report was authored by the National Renewable Energy Laboratory (NREL) and the Interstate Renewable Energy Council, Inc. (IREC).

NREL is a national laboratory of the U.S. Department of Energy (DOE) that specializes in the research and development of renewable energy sources, such as solar, wind, water, and geothermal. NREL is a lead in developing the future of sustainable and integrated energy systems with researchers who harness the power of data and high-performance computing, integrated testing and who focus on integrated solutions, delivering grid modernization and security. NREL has decades of experience providing leadership and novel research in distribution system analyses and planning. The research team that supported this project consisted of subject matter experts on the topic of hosting capacity analysis and has experience exceeding a decade. NREL regularly performs power flow analyses for various purposes, including analyses to further its research on advanced hosting capacity analyses and as a service to distribution utilities.[1]

IREC builds the foundation for rapid adoption of clean energy and energy efficiency to benefit people, the economy, and our planet. Its vision is a 100% clean energy future that is reliable, resilient, and equitable. IREC develops and advances the regulatory reforms, technical standards, and workforce solutions needed to enable the streamlined integration of clean, distributed energy resources. IREC has been trusted for its independent clean energy expertise for nearly 40 years, since its founding in 1982. IREC's Regulatory Team has been involved in numerous regulatory dockets and research projects associated with the development of distribution system plans and Hosting Capacity Analyses (HCAs).[2] IREC has published two papers and multiple in-depth blog posts about HCA design, which are available at: https://irecusa.org/our-work/hosting-capacity-analysis/.

[1] For more information, see "Advanced Hosting Capacity Analysis," NREL, https://www.nrel.gov/solar/market-research-analysis/advanced-hosting-capacity-analysis.html.

[2] CA Pub. Util. Comm., Dkt. R.14-08-013, Distribution Resources Plans; CA Pub. Util. Comm., Dkt. R.21-06-017, Rulemaking to Modernize the Electric Grid for a High Distributed Energy Resources Future; NV Pub. Util. Comm., Dkt. 17-08022, Rulemaking to Implement Senate Bill 146 (2017); NY Pub. Service Comm., Dkt. 14-M-0101, Reforming the Energy Vision; NY Pub. Service Comm., Dkt. 16-M-0411, Distributed System Implementation Plans; MN Pub. Util. Comm., Dkt. E999/CI-15- 556, Investigation into Grid Modernization; MN Pub. Util. Comm., Dkt. E002/M-15-962, Xcel Energy Biennial Report on Distribution Grid Modernization; MN Pub. Util. Comm., Dkt. E002/M-17-777, Xcel Energy 2017 Hosting Capacity Study; MN Pub. Util. Comm., Dkt. E002/CI-18-251, Xcel Energy Distribution System Planning; MN Pub. Util. Comm., Dkt. E002/M-18-684, Xcel Energy 2018 Hosting Capacity Study; MN Pub. Util. Comm., Dkt. E002/M-19-685, Xcel Energy 2019 Hosting Capacity Study; MN Pub. Util. Comm., Dkt. E999/CI-20-800, Grid And Customer Security Issues Related to Public Display or Access to Electric Distribution Grid Data; MN Pub. Util. Comm., Dkt. E002/M-20-812, Xcel Energy 2020 Hosting Capacity Analysis.

Executive Summary

Solar generation, energy storage, electric vehicles, and other distributed energy resources (DERs) are arriving on the electric distribution grid in fast-growing numbers, but it is not always clear how much incremental DER capacity the distribution system can accommodate. Clarity about grid capacity is of special importance to utilities, developers, and regulators, as well as customers, who are adding more DERs and require accurate, accessible, and trustworthy information.

Such information can be gathered in a hosting capacity analysis (HCA)—a process used by utilities and regulators in multiple states to determine the available capacity for new DERs without requiring expensive and time-consuming studies or grid upgrades. If performed properly, an HCA can streamline and add transparency to DER planning and interconnection processes.

However, some of the first-published HCAs included inaccurate data. For example, a published HCA result showed a feeder with zero capacity, but after an interconnection application was processed, it turned out the feeder actually could accommodate multiple megawatts. This undermined users' confidence in the HCA and raised doubts that the analysis accurately reflected real-world grid conditions. Without confidence in the HCA, users are unlikely to rely on the data and the HCA cannot fulfill its intended purpose. To improve the quality, accuracy, and trust in HCA data and to avoid the challenges found in early rollouts, the National Renewable Energy Laboratory (NREL) and the Interstate Renewable Energy Council (IREC) provide in this report a suite of best practices for HCA data validation.

Implementing this report's HCA data validation practices can increase trust in the HCA, making the results more useful for DER planning and interconnection processes. **The best practices recommended here could be useful for:**

- **Utilities**, to develop or refine their HCA data validation procedures
- **Regulators**, to inform their oversight of utilities' HCA data validation practices
- **Other stakeholders**, to evaluate the effectiveness of utility efforts.

For this work, NREL and IREC interviewed utilities, software vendors, U.S. Department of Energy national laboratories, regulatory commissions, and solar developers to identify common issues in HCA, and reviewed examples of HCA from utilities around the United States to understand current practices. From these findings, this report identifies both procedural and technical best practices for HCA data validation. Our goal is to reduce barriers and provide utilities, regulators, and all stakeholders with a replicable roadmap to help HCA deployments provide accurate, trustworthy, and reliable results from the day they are published.

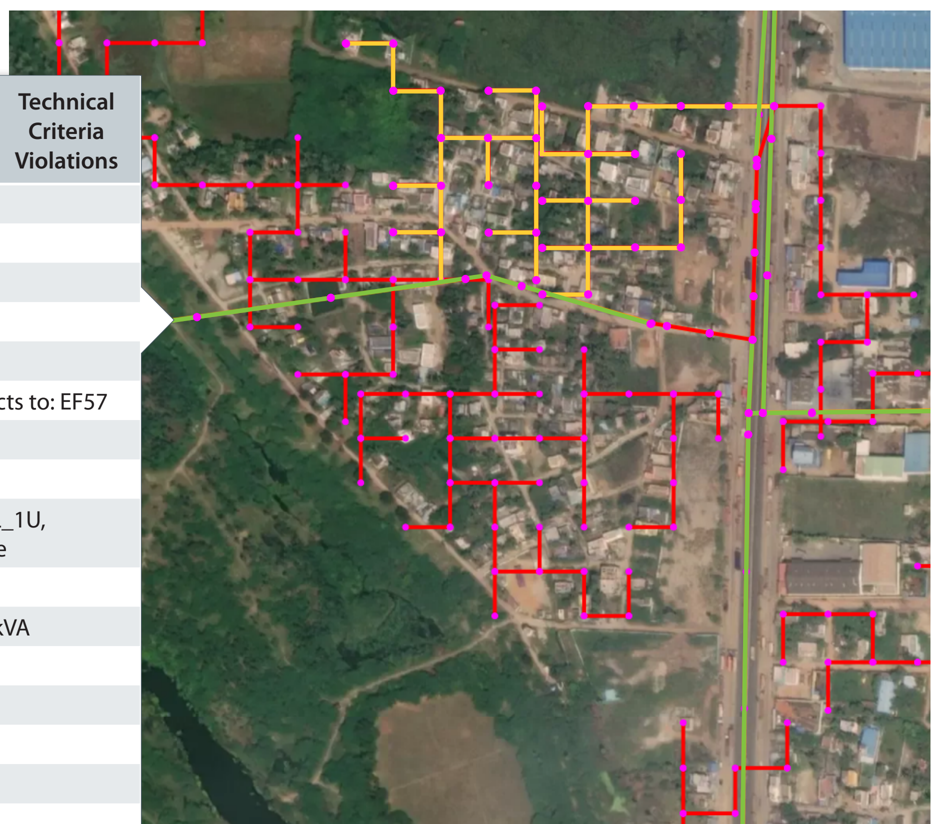
At a high level, HCA best practices include:

- **An appropriately resourced HCA team** that tracks metrics at each step of the process
- A **well-documented, repeatable process** for data validation, using suitable software to ensure digital feeder models reflect real-world grid conditions
- **Transparent and collaborative information sharing** for feedback and identification of errors.

To start, an HCA requires dedicated attention. Successful HCAs are managed by a specific HCA manager to oversee data validation and ensure their HCA team is well resourced. The role of the manager includes establishing and tracking metrics to assess the quality of HCA data in each step, as well as the quality of final results. The manager also works to ensure that each step of the HCA process functions efficiently altogether.

Effective data validation practices include developing, documenting, and following a standardized approach, so that the HCA team can efficiently identify errors and correct the failures in the thousands of nodes that comprise a typical service area. HCA data can have diverse origins involving different utility departments. HCA processes run most efficiently when errors identified by the HCA team are corrected in the source database, even when a different department is responsible for that database.

An HCA also involves building models of distribution feeders to simulate power flow, which is the most common root cause of errors. This report includes tables with examples of validation procedures for each step in the feeder model building process. It is a best practice



Line Section	Substation	Load Profile	Technical Criteria Violations
Section ID: 148104500			
Hosting Capacity: 5,000 kW			
Feeder ID: 23			
Feeder voltage: 12.47			
Number of phases: 3			
Which substation transformer the feeder connects to: EF57			
Feeder type: radial			
Feeder length: 6.213 miles			
Feeder conductor size and impedance: 1000_AL_1U, R1=0.517416 ohm/mile, X1= 0.267376 ohm/mile			
Service transformer rating: 100kVA			
Service transformer daytime minimum load: 35kVA			
Existing generation (weekly refresh rate): 50 kW			
Queued generation (weekly refresh rate): 20 kW			
Total generation (weekly refresh rate): 70 kW			
Currently scheduled upgrades: No			
Federal or state jurisdiction: State			

HCA map example by NREL

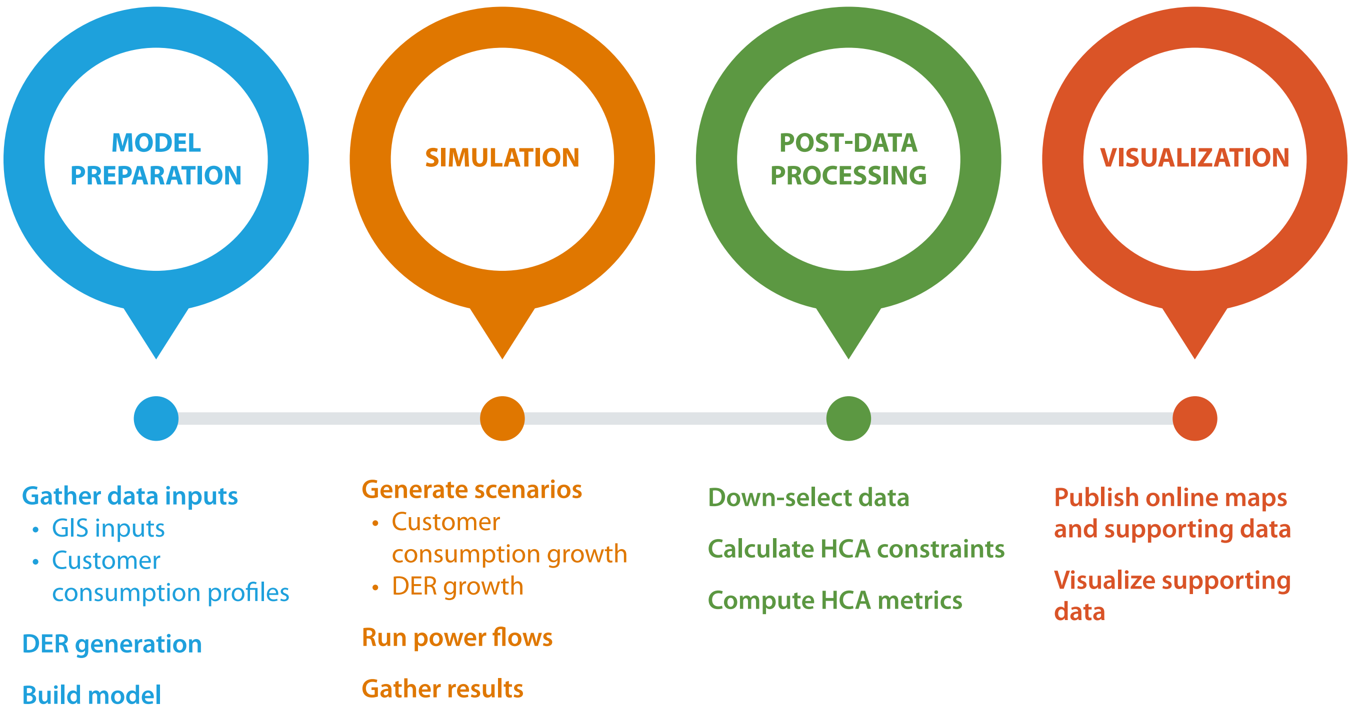


Figure 1. Steps in an HCA. *Illustration by Nicole Leon, NREL*

for utilities to standardize and document the steps in the feeder model building and validation process:

- Feeders that experience similar challenges would benefit from **being batched** so that engineers can easily develop solutions to common problems.
- **Scripting** can be used to automate error correction when building feeder models and to significantly accelerate decision making. Code base management tools are effective in preventing and resolving errors because they allow utilities to track the evolution of code and quickly revert to previous versions if needed.
- Using **actual—not estimated—customer consumption data** improves data accuracy. Existing commercial software, versus tools developed in-house, also provides an advantage in managing consumption profiles since they typically include helpful data validation features.
- Instead of attempting to perform the power flow simulations for an entire year and an entire service area at once, we propose examining a **prioritized set of load hours and a representative sample of feeders** first.

To maximize efficiency and effective public oversight, the HCA process can include measures that prioritize transparency and feedback to help catch errors and elevate confidence in HCA results. Suggested measures include a review process to flag irregularities before publication, as well as a mechanism to allow customers and HCA data users to offer feedback about user experience, identified errors, and usefulness of the HCA data.

Likewise, it is a best practice for regulators to provide transparency into the data validation process. This could be done by reviewing and requiring improvements to data validation plans, tracking the quality of HCA results over time with metrics that describe data quality, and requiring a root cause analysis for recurring problems in the HCA process.

The report identifies best practices for validation procedures, specific rules for identifying data errors, and suggestions for regulatory oversight. Using these processes, utilities and regulators can provide confidence that HCA results accurately reflect grid conditions. With that confidence, trusted HCA data can be used in modernized DER planning and interconnection processes.

Acknowledgments

We gratefully acknowledge the many people whose efforts contributed to this report. Many NREL and IREC colleagues reviewed and improved this report, including Jess Townsend, Killian McKenna, Peter Gotseff, Aadil Latif, Sky Stanfield, Midhat Mafazy, Radina Valova, and Gwen Brown. We thank the HCA experts listed in section 2.1 that took the time to sit down for an interview with us. As described in section 2.2, this report builds on the recommendations found in Quanta Technology's assessment of California's HCA data validation plans performed by Vic Romero, Stephen Teran, and Andrija Sadikovic. We also thank our DOE colleagues for providing critical review of interim and final work products, including Shay Banton, Michele Boyd, and Susanna Murley. This work was funded by the U.S. Department of Energy's Solar Energy Technologies Office under contract number DE-AC36-08GO28308. All errors and omissions are the sole responsibility of the authors.

List of Acronyms and Abbreviations

AMI	advanced metering infrastructure
CA	California
Comm.	Commission
DER	distributed energy resource
Dkt.	Docket
DOE	U.S. Department of Energy
GAT	grid analysis tool
GIS	geographic information system
HCA	hosting capacity analysis
ICA	integration capacity analysis
IREC	Interstate Renewable Energy Council
kVAr	kilovolt amps reactive
kW	kilowatts
MN	Minnesota
MW	megawatts
NREL	National Renewable Energy Laboratory
NV	Nevada
PG&E	Pacific Gas and Electric Co.
Pub.	Public
QA	quality assurance
QC	quality control
SCADA	supervisory control and data acquisition
SCE	Southern California Edison
SDG&E	San Diego Gas & Electric
Util.	Utilities
VAr	volt amps reactive

Table of Contents

List of Figures

List of Tables

1 Introduction

The usefulness of hosting capacity analysis (HCA) depends on users' confidence that the HCA results accurately reflect grid conditions. Confirming that the data used as inputs to the HCA are ready for use in sophisticated power flow simulations is essential to ensuring HCA results are accurate. This represents one of the most time-intensive parts of developing an HCA.

Failure to adequately validate HCA data before publication will produce an inaccurate representation of the distribution grid that users will not trust. Utility management and regulators can ensure trustworthy HCAs by overseeing the data validation process. This recommendation is rooted in the experience of the first utilities to perform HCA, some of whom published unvalidated results which users did not trust. For example, in January 2019, California utilities published their first systemwide HCA results.[1] Surprisingly, Pacific Gas and Electric's (PG&E's) first HCA showed that approximately 80% of PG&E's feeders had little or no hosting capacity for new solar available, and PG&E, San Diego Gas & Electric (SDG&E), and Southern California Edison's (SCE's) HCA showed 60%–70% of their distribution systems had little to no hosting capacity for new load.[2]

Though it is widely known that PG&E has higher solar deployment, it was highly unlikely that most of the PG&E system had no remaining capacity for new solar projects of any size. Similarly, it would be surprising if 60%–70% of California's grid could not support new distributed loads. These 2019 HCA results did not reflect the reality experienced by customers interconnecting projects and were met with immediate frustration and suspicion that the results were inaccurate.

Stakeholders pointed out to PG&E, SDG&E, and SCE that the results appeared anomalous, and subsequent discussions among stakeholders and regulators led to the conclusion that the results were erroneous.[3] As a result, PG&E implemented a concerted data validation effort that took about 15 months to produce validated results for solar generation.[4] PG&E's solar HCA results have largely been fixed, but problems remain with all three California investor-owned utilities' load HCA results. Over two years after the initial load HCA results were published, those results remain suspect and have yet to be validated. As a result, regulators decided to scrutinize utilities' data validation efforts more closely. The California Public Utilities Commission required each

[1] In California, HCA is called Integrated Capacity Analysis, but for consistency we use HCA in this document.

[2] CA Pub. Util. Comm. Dkt. R.17-07-007, Response of Pacific Gas and Electric Co. to Data Request 1 of the Interstate Renewable Energy Council, Clean Coalition, and California Solar and Storage Association, at p. 3 (Sept. 28, 2018) (Question 2 re "ICAOF" results); CA Pub. Util. Comm., Dkt. R.14-08-013, Reply Comments of The Interstate Renewable Energy Council, Inc. on Refinements to the Integration Capacity Analysis, Attachment 5: Sept. 9, 2019 Joint Investor-Owned Utility Presentation on Load ICA Methodology and Process, at p. 5 (Sept. 30, 2019) (IREC Reply Comments on ICA Refinements).

[3] See, e.g., IREC Reply Comments on ICA Refinements, at pp. 1–12; CA Pub. Util. Comm., Dkt. R.14-08-013, Administrative Law Judge's Ruling on Joint Parties' Motion for an Order Requiring Refinements to the Integration Capacity Analysis, at pp. 4–6 (Jan. 27, 2021) (CPUC ICA Refinements Order).

[4] PG&E implemented GridUnity's Network Model Management software beginning in Q1 2019 and reported that its maps included verified and published results on May 7, 2020. CA Pub. Util. Comm., Dkt. R.14-08-013, Pacific Gas & Electric's Integration Capacity Analysis Implementation Update, at p. 1 (May 7, 2020).

utility to file a data validation plan, invited stakeholders to comment on the plans, and then hired an independent technical expert to review the plans and suggest improvements.[5]

Concerns about HCA results have also been raised in Minnesota and Massachusetts. In Minnesota, developers reported that Xcel Energy's initial HCA results were unreliable and thus not used. For example, one developer reported that over half "of the locations we screened had more capacity indicated in the screen than the map. We no longer use the map as a result. One location that showed 0 capacity, had 14MW of capacity without upgrades when in final design with Xcel."[6] And in Massachusetts, a developer reported that a utility employee told the developer not to use the HCA because the results were unreliable.

As a result of the flawed HCA rollouts in California and Minnesota, as well as questions raised about the accuracy of first-published HCA maps in other states, the U.S. Department of Energy's Solar Energy Technologies Office agreed to support development of a guide on HCA data validation best practices. This report provides the findings and recommendations from that research. Specifically, it provides utilities with best practices for HCA data validation processes—and recommends that regulators and policymakers oversee the process—so that future HCA deployments provide useful and accurate data from the day they are published. Utilities can use the report to develop or refine their HCA data validation procedures. Regulators can use the report to inform their oversight of utilities' HCA data validation practices. And stakeholders can use it to evaluate the effectiveness of utility efforts.

[5] CPUC ICA Refinements Order, at pp. 4–6 (according to CPUC rules, stakeholders may comment on advice letters).
[6] MN Pub. Util. Comm., Dkt. E002/M-18-684, Fresh Energy's Comments on Xcel's 2018 Hosting Capacity Study, at p. 3 (Feb. 28, 2019).

2 Research Methodology

This project drew on four sources of information to identify HCA data validation procedures: interviews with HCA experts (Section 2.1); a review of relevant documents, including distribution system plans and HCA reports (Section 2.2); NREL's experience conducting power flow analyses and HCAs, and IREC's experience participating in regulatory proceedings that developed and refined HCAs.

Based on the data collected from these sources, NREL and IREC identified issues and errors that commonly occur when performing hosting capacity analyses, and procedures that can address these errors and enhance accuracy of HCA results. Based on the identified issues, errors, and procedures, NREL and IREC in this report identify the use of certain quality assurance (QA), quality control (QC), and regulatory best practices that can help produce HCA results that are accurate and trustworthy.

2.1 Interviews with HCA Experts

We interviewed individuals actively involved in HCA who work for utilities, software vendors, U.S. Department of Energy (DOE) national laboratories, regulatory commissions, and solar developers to identify issues and errors that commonly occur when performing HCAs, and procedures that can address these errors and produce sufficiently accurate HCA results. Before conducting interviews, NREL distributed surveys to each participant to learn about their role in the HCA process and to enable the project team to tailor the interview questions to the participant's experience. We then prepared interview questions based on the survey responses.

The interviews included HCA experts at two distribution utilities, four power flow simulation software vendors, two national laboratories, three regulatory commissions, three solar developers, and one nonprofit. Persons with the following roles and affiliations participated in the interviews, but please note that the report's recommendations are the work of the authors and do not necessarily reflect the views of the interview participants or their employers:

- Synergi Electric Principal Consultant, DNV
- CYME Power System Engineering Manager, Eaton
- Principal Engineer, Electrical Distribution Design, Distributed Engineering Workstation
- Lead Engineer for Distribution Operations and Planning, Electric Power Research Institute
- Principal Engineer, Pacific Gas and Electric Co. (PG&E)
- Manager Distributed Resources Engineering, Arizona Public Service Co.
- Distribution System Engineer, National Renewable Energy Laboratory
- Principal Member of Technical Staff, Sandia National Laboratory
- Staff, Colorado Public Utilities Commission
- Staff, Maryland Public Service Commission
- Staff, Nevada Public Utilities Commission
- Director, Sunrun
- Project Developer, Engie
- Engineer, Borrego Solar Systems
- Regulatory Engineer, IREC

2.2 Review of Distribution System Plans and Hosting Capacity Analysis Reports

We reviewed various documents from California, Minnesota, Nevada, and New York that provide descriptions of and recommendations for HCA data validation procedures.

As described above, the California Public Utilities Commission required utilities that perform HCAs to file a data validation plan and then hired an independent technical expert to review the plans and suggest improvements.[7] We reviewed multiple iterations of the data validation plans of PG&E, SCE, and SDG&E,[8] as well as the independent technical expert's assessment of those data validation plans.[9]

We also reviewed multiple years of Xcel Energy's HCA reports, NV Energy's Distribution System Plans, and New York utilities' HCA workshops, which describe the utilities' HCA data validation practices.[10]

---

[7] CPUC ICA Refinements Order, at pp. 4–6.

[8] See, e.g., CA Pub. Util. Comm., San Diego Gas & Electric Co. Advice Letter 3773-E-A, Improved Integrated Capacity Analysis Data Validation Plans, at 4 (Aug. 27, 2021) (SDG&E Data Validation Plan), https://tariff.sdge.com/tm2/pdf/3773-E-A.pdf; CA Pub. Util. Comm., Pacific Gas & Electric Co. Advice Letter 6212-E, Improved Integrated Capacity Analysis Data Validation Plan, Attachment 1: PG&E ICA Data Validation Plan (May 28, 2021) (PG&E ICA Data Validation Plan), https://www.pge.com/tariffs/assets/pdf/adviceletter/ELEC_6212-E.pdf; CA Pub. Util. Comm., Southern California Edison Co., Advice Letter 4508-E, Improved Integration Capacity Analysis Data Validation Plan (May 28, 2021) (SCE Data Validation Plan), https://library.sce.com/content/dam/sce-doclib/public/regulatory/filings/pending/electric/ELECTRIC_4508-E.pdf.

[9] Vic Romero and Stephen Teran, SDG&E Integration Capacity Analysis Data Validation Plan Assessment, Quanta Technology (June 24, 2021) (Quanta SDG&E Assessment), https://irecusa.org/wp-content/uploads/2022/02/QTech-SDGE-ICA-Data-Validation-Plan-Assessment-Report-6-28-21.pdf; Stephen Teran and Vic Romero, SCE Integration Capacity Analysis Data Validation Plan Assessment, Quanta Technology (June 24, 2021) (Quanta SCE Assessment), https://irecusa.org/wp-content/uploads/2022/02/QTech-SCE-ICA-Data-Validation-Plan-Assessment-Report-6-28-21.pdf; Andrija Sadikovic and Vic Romero, PG&E Integration Capacity Analysis Data Validation Plan Assessment, Quanta Technology (June 24, 2021) (Quanta PG&E Assessment), https://irecusa.org/wp-content/uploads/2022/02/QTech-PGE-ICA-Data-Validation-Plan-Assessment-Report_Redacted-6-25-21.pdf.

[10] MN Pub. Util. Comm., Dkt. E-002/M-20-812, Xcel Energy Hosting Capacity Analysis Report (Nov. 2, 2020) (Xcel Energy 2020 HCA Report); MN Pub. Util. Comm., Dkt. E-002/M-18-684, Order Accepting Study and Setting Further Requirements, Xcel HCA, at 4-5 (Aug. 15, 2019); Pub. Util. Comm. of NV, Dkt. 21-06-001, Nevada Power Co. Integrated Resource Plan Vol. 13, Narrative Distributed Resources Plan (June 1, 2021) (NVE 2021 DRP); Joint Utilities of New York, Hosting Capacity, https://jointutilitiesofny.org/utility-specific-pages/hosting-capacity/ (accessed Dec. 14, 2021).

3 Interview Findings

This section presents the interview findings about the common errors encountered in HCAs, and procedures that can address these errors and produce sufficiently accurate HCA results.

3.1 Steps in an HCA

After consulting with HCA experts at the DOE national laboratories, NREL identified four general stages to producing HCAs (Figure 1). Although the specific procedures used in each stage can vary considerably, every successful HCA will include these four stages. This remains true regardless of the use case or methodology selected for the HCA.

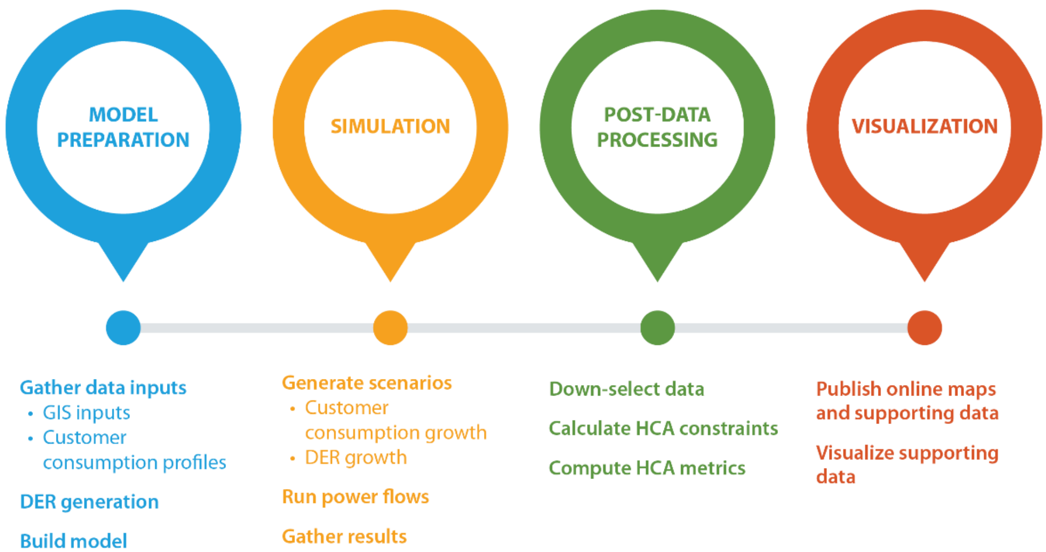


Figure 1. Steps in an HCA. ***Illustration by Nicole Leon, NREL***

The first step is model preparation. A feeder model is a digital representation of a part of the distribution system. The model is designed to match the characteristics of the feeder in the physical world. Ensuring the baseline feeder model accurately reflects the distribution feeder topology, and the load and DER profiles is a critical effort that can take considerable time and resources. Validating a distribution feeder model may require coordination between diverse teams within a utility.

In the second step, engineers and software developers create scenarios that test the ability of the distribution system to accommodate new DERs. The team performing the HCA must manage the status of each feeder as it travels between the various steps. For example, HCA experts we interviewed indicated that the presence of an error or inaccurate results in simulation could indicate the need for revisions to the baseline model, sending the feeder back to the first step.

Utilities often have hundreds of distribution system feeders, and HCAs commonly include hundreds of scenarios for each feeder. The power flow simulation software produces several output files for each scenario. These files are comprehensive and contain voltages, loads, hours, and generation profiles for each node in a feeder. A typical HCA team processes thousands of files.

In the third step, post-data processing often produces errors and inaccurate results that require investigation and correction. This step is critical to assessing accuracy of results at hand. Given the magnitude of data this step needs processes in place to evaluate key metrics. And in the last step, the HCA results are validated, visualizations representing available hosting capacity are produced and then published in various formats, including online maps, tabular files, geographic information system (GIS) files, and application programming interfaces.

3.2 Common Sources of Error

3.2.1 Model Input Errors and Lack of Process to Coordinate Across Utility Crosscutting Teams

The HCA experts we interviewed indicated that the most error-prone stage is feeder model preparation because:

- During the feeder model development and validation process, the HCA team is likely to encounter numerous feeder topology problems whose root cause is an error in the GIS database.
- Fixing GIS errors so that they do not repeat requires established processes, sufficient staffing resources, and cross-team coordination within utilities.
- GIS databases and software were not originally developed to support power flow modeling, and GIS is insufficiently integrated with power flow modeling software.

Power flow simulations assume the feeder model is correct and validated. Without a validated feeder model, a small error at this stage may be amplified and lead to clearly erroneous HCA results. In addition, investing insufficient upfront effort in the data validation process will ultimately cause more work for the HCA team.

3.2.2 Appropriate Use of Software

Power flow modeling software provides the HCA team automated tools that aid data validation efforts by flagging certain errors. Given the magnitude of data involved in an HCA, these tools help streamline and automate the HCA process. Software vendors we interviewed indicated that existing software tools provide basic QA/QC support. Table 1 lists selected capabilities in power flow modeling software that support QA/QC.

Table 1. Selected Capabilities in Power Flow Modeling Software That Support QA/QC

Error Flags	Checks Performed
Missing conductor types	Checks for conductor attributes with null values
Incorrect regulator settings	Checks of regulator controls; verifies whether controllers reach their limit and stop moving up or down
Transformer configuration	Checks for severely overloaded transformers
Inaccurate load data	Checks for anomalies in load values; flags load values far from average

Error Flags	Checks Performed
Mesh or loops in network topology	Checks for meshes and loops; identifies loops and alerts user with warnings
Inaccurate solar irradiance curves	Checks for solar irradiance curves; tools can provide clear sky irradiance based on the latitude and longitude of the solar facility
Zero hosting capacity result	Checks for large areas with zero hosting capacity

3.2.3 HCA Design Choices

In the interviews, several HCA experts we interviewed brought up HCA design choices that often lead to irrelevant HCA results. First, utilities brought up the need to select and clearly define a use case for the HCA. A clearly defined use enables utilities, developers, and regulators to focus their efforts on actions that provide customers the value described in the use case. Second, software vendors brought up the importance of carefully considering and vetting the limiting criteria and thresholds used in the HCA. Selecting improper limiting criteria or thresholds could have a significant impact on HCA results. These design choices are important to consider in the development of an HCA because the wrong choices can lead to irrelevant HCA results, even when accompanied by a data validation process. These design choices, however, are outside the scope of data validation procedures addressed in this report.[11]

3.3 Growing Prominence of HCA

HCA data are receiving additional attention and scrutiny from stakeholders and regulators. Some utilities envisioned an initial use case for their HCA data, but they now see stakeholders and regulators asking to use the data in decision-making processes in more consequential ways. Put another way, stakeholders and regulators may ascribe a higher value to HCA data, and conversely a higher cost to inaccuracies and errors in HCA data. As a result, stakeholders and regulators see significant value in investing in HCA data validation processes.

3.4 Regulatory Activities

We interviewed staff from regulatory commissions in Colorado, Maryland, and Nevada that oversee the publication of HCA data, and two DER developers that are active in markets with published HCAs.

These interviewees indicated a robust stakeholder engagement process produces more useful HCA data for customers. They noted that effective stakeholder engagement processes look to learnings from other states that provide HCA data to customers. As a part of the process, interviewees advised that regulators be open to learning from stakeholders as well as from utilities. Those developing a new HCA tool for the benefit of stakeholders should not presume utilities know the best HCA design. Interviewees indicated they are more likely to trust HCA results when stakeholders are provided the opportunity to make presentations at workshops

[11] For more information about these and other design choices, see Sky Stanfield, Yochi Zakai, Matthew McKerley, *Key Decisions for Hosting Capacity Analyses,* IREC (Sept. 2021), https://irecusa.org/resources/keydecisions-for-hosting-capacity-analyses.

discussing HCA program design, and stakeholders can provide their own proposals for HCA program design.

Regulatory staff are typically unaware of the quality of databases used as inputs into HCA for the utilities that they regulate, according to the regulatory staff we interviewed.

When developers believe HCA results are unreliable, they ask for and prefer to use basic distribution system data instead. Basic distribution system data are normally provided on the same website as the HCA data but are distinct from HCA results.[12] Developers would find HCAs more helpful if they were updated more frequently, used in the interconnection process, and used to identify areas for proactive upgrades. Developers support this recommendation by arguing that utilities are more likely to produce valid and accurate HCAs when utilities use the HCA results in their own interconnection and distribution planning decision-making processes.

[12] Basic distribution system data includes information about feeders and substations, including hourly load profiles, existing and queued generation, voltages, phases, type, length, transformer rating, and known constraints. See*, e.g.,* Key Decisions for Hosting Capacity Analysis, at pp. 11–12.

4 Best Practices

This report provides best practices for utilities and regulators for the design of data validation procedures to support HCAs. In Section 4.1, we discuss the business processes and types of resources necessary to support a robust HCA data validation process. Section 4.2 establishes validation procedures for each step in the feeder model development process. For each step, we include a table with examples of validation procedures. Though the tables are not comprehensive, they provide a starting point for the development of a complete validation process. Section 4.3 highlights ways to ensure results are valid after the utility completes its HCA but before the results are published. In Section 4.4, we discuss accepting feedback from users, and Section 4.5 discusses regulatory oversight of a data validation plan and regular reporting.

4.1 Business Processes

4.1.1 Identify Who Is Responsible for Managing and Improving the HCA and Verification Processes

HCA map generation involves using data generated or maintained across teams within a distribution utility. Successful HCAs involve appointing a specific HCA manager, supported by a team, who will be responsible for managing and improving the HCA processes by providing strategic direction, identifying specific objectives, and establishing a structure for the HCA and data validation activities.[13] The HCA manager, supported by their team, would have ultimate responsibility for data validation, performing the HCA, ensuring the accuracy of HCA results, and improving the efficiency of HCA processes.

HCA and data validation processes are complex and use significant utility resources. An HCA manager helps ensure they are completed in a manner that avoids waste and encourages continuous efficiency improvements. The HCA manager's specific responsibilities include, but are not be limited to: standardizing and documenting the HCA process, validating results, tracking and implementing identified needs for improvement, establishing a long-term strategy to maintain HCA results quality, and managing the processes described in the remainder of this section.

4.1.2 Establish Metrics to Track the Quality of Input Data and HCA Results Over Time

It is a best practice for HCA managers to establish and track metrics.[14] These metrics assess the quality of data used in each step of the HCA process, whether the HCA process is functioning efficiently, and the quality of results. Tracking these metrics over time will help identify trends

---

[13] Vic Romero and Stephen Teran, SDG&E Integration Capacity Analysis Data Validation Plan Assessment, Quanta Technology, at p. 2 (June 24, 2021) (Quanta SDG&E Assessment).
[14] Stephen Teran and Vic Romero, SCE Integration Capacity Analysis Data Validation Plan Assessment, Quanta Technology, at p. 3 (June 24, 2021) (Quanta SCE Assessment).

related to data quality and inform root cause analyses.[15] Some examples of metrics that the HCA manager could track include:

- The frequency of errors and issues for each HCA update[16]
- The frequency of each type of failed flag or check (as detailed in Sections 4.2 and 4.3) for each HCA update[17]
- The number of recurring problems in the model building process, and with which source database the problem is associated, if any[18]
- Whether the team completed its processes in the desired time frame and the HCA update was published on time[19]

4.1.3 Fix Identified Problems in the Source Database

Information from a utility's distribution system asset database, GIS database, load database, and generation profile database constitutes the primary inputs used to create the feeder models needed to perform power flow analyses.

Errors in the power flow analyses used to perform the HCA are often due to data quality and integrity problems in the source databases.[20] Efficient HCA processes fix identified errors in the source databases so HCA engineers are not required to fix the same errors each time they use the source database to update a feeder model. Otherwise, engineers often develop a script or another automated solution to fix the error each time they use the source database to update a feeder model. Though scripts remove the need for manual intervention, they are not the best solution because they require the continued use of computing resources and do not correct the problem for other users of the database.

It is a best practice for HCA managers to follow up with the source database owner when a root cause analysis shows the database includes inaccurate data or causes HCA errors.[21] Utilities are often large organizations, and the HCA staff may not interact regularly with the staff that maintain the source databases. We recommend that utilities overcome the challenges associated

---

[15] Andrija Sadikovic and Vic Romero, PG&E Integration Capacity Analysis Data Validation Plan Assessment, Quanta Technology, at p. 3 (June 24, 2021) (Quanta PG&E Assessment) ("While individual values for the metrics are informative (e.g., there are currently 100 nodes with zero hosting capacity), trends in the metrics can help identify emerging issues in the input data or process (e.g., the count of nodes with zero hosting capacity is not changing over time) or show improvements in quality (e.g., the count of nodes with zero hosting capacity is decreasing on feeders that have recently had limiting factors mitigated). The metrics should also be tracked to support analysis at various levels of system granularity (e.g., system-level, feeder-level, node-level, etc.) and troubleshoot potential data issues.").
[16] Quanta SCE Assessment at p. 14.
[17] Quanta PG&E Assessment at p. 12.
[18] Quanta SCE Assessment at pp. 12–14.
[19] Quanta SCE Assessment at p. 10.
[20] See, e.g., CA Pub. Util. Comm., San Diego Gas & Electric Co. Advice Letter 3773-E-A, Improved Integrated Capacity Analysis Data Validation Plans, at 4 (Aug. 27, 2021) (SDG&E Data Validation Plan), https://tariff.sdge.com/tm2/pdf/3773-E-A.pdf; Quanta SCE Assessment, at pp. 4–5, 12.
[21] Quanta PG&E Assessment at p. 4.

with siloing in large organizations and establish processes that fix identified errors in the source database.

4.1.4 Use Appropriate Employee and Computational Resources

The efficiency and accuracy of HCA processes depends on using appropriate resources, including skilled engineers and computational resources. Resources could be located within the utility or provided by external contractors.

Engineers working on an HCA should have experience with their utility's distribution engineering practices (including planning and operations), circuit models, and design standards.[22] They should also be familiar with the HCA methodology, their utility's implementation of the methodology, and the flow of data from source databases to the feeder model and then to the HCA software. All told, HCA requires skilled engineers with knowledge of the entire HCA process, from input data to the publication of the results.

Using appropriate computational technologies, such as high-performance computers or cloud computing, avoids unnecessary slowdowns due to hardware constraints and accelerates debugging.[23] Performing computationally intensive tasks without the appropriate resources (e.g., on a traditional laptop) will likely slow the entire HCA process, frustrate employees, and prevent HCA results from reaching customers in a timely manner. Providing employees access to the appropriate computing resources can increase the effectiveness and accuracy of the entire HCA process.

While portions of the HCA data validation program can be automated, engineers will always need to correct some problems. Effective HCA managers consider how to strike the right balance between automation and manual work. Although scripting is a powerful tool and commercial software tools are always improving, effective HCA managers monitor for the point at which increased automation provides diminishing efficiency returns.

4.2 Quality Control During the Feeder Model Development Process

The HCA experts we interviewed identified feeder model development as the most error-prone stage of HCA. To address this, recommendations for developing repeatable and streamlined processes to check and correct model input errors at each stage in the HCA process are presented in this section.

4.2.1 Create a Baseline Model and Validate Its Accuracy

A feeder model is a digital representation of a part of the distribution system. A baseline, or base case, digital feeder model is designed to match the characteristics of the feeder in the physical world. Ensuring the baseline feeder model accurately reflects the actual feeder is the first step in

---

[22] Quanta SCE Assessment at p. 3.

[23] See CA Pub. Util. Comm., Pacific Gas & Electric Co. Advice Letter 6212-E, Improved Integrated Capacity Analysis Data Validation Plan, Attachment 1: PG&E ICA Data Validation Plan, at p. 3 (May 28, 2021) (PG&E ICA Data Validation Plan) ("PG&E's current platform consists of 23 AWS servers with 18 cores (3GHz) processors each, that perform iterative ICA calculations. The platform is supporting ICA calculation of approximately 15% of PG&E circuits, on average, each month."), https://www.pge.com/tariffs/assets/pdf/adviceletter/ELEC_6212-E.pdf.

HCA data validation. Failing to take the time to confirm the accuracy of the initial baseline feeder will create unnecessary work for the HCA team later in the process. Table 2 lists procedures that can be used to validate the baseline model before running hosting capacity scenarios.

Table 2. Baseline Model Validation Procedures

Validation Checks	Validation Procedure
Voltage base	Check whether the feeder head voltage matches the real-world value.
Voltage at nodes	Check for nodal voltage violations at peak load allocation and minimum load allocation.
Loading check	Check for device overloads including transformers and conductor thermal ratings.
Equipment default settings	Check the settings of transformers, capacitors, and regulators in the model to ensure they match the settings of this equipment in the field. If the utility changes settings seasonally, check the model to reflect this seasonal change.
Short circuits	Check value of fault duty (i.e., the maximum current) on each node.
Circuit reactive power	Check the power factor at the feeder head at peak and minimum load.
Circuit losses	Check aggregate active power losses and check power losses as a percentage of load served at the feeder head.
Aggregate active power	Check whether the aggregate active power consumption at the feeder head matches the allocated peak load.

4.2.2 Develop, Document, and Follow a Standardized Approach to Resolving Errors

During the feeder model development and validation process, the HCA team is likely to encounter numerous issues and errors. To efficiently identify errors and correct the failures for the hundreds of feeders in a typical distribution system, the HCA managers develop, document, and follow a standardized approach. Without an organized approach to identifying and resolving problems with feeder models, the process will take longer, and human errors are more likely to occur.[24]

For example, once the baseline feeder models are created, the HCA team will perform the validation checks listed in Table 2. Some feeders will pass all the automated and manual checks, some feeders will produce an automatic warning that indicates review is recommended, and others that fail will require manual review. Effective HCA managers develop a tool that allows the HCA team to track which and how many circuits have passed, produced warnings, or failed at each milestone in the HCA process. Table 3 shows the labels one utility uses to track the progress of feeders through its HCA process. The data produced by this tracking tool can be used to create metrics that allow the HCA manager to monitor the team's progress.

[24] Quanta SCE Assessment at p. 6 ("A best practice to reduce potential human errors when manual intervention is required is using a standardized approach to identify and resolve issues with the distribution circuit models and the [HCA] process.").

Table 3. Statuses Distribution Engineers Can Use to Batch Distribution Circuits for Batch Processing

Status	Description
Completed	The circuit has successfully passed the stage.
Failed	A problem occurred that was serious enough to stop the workflow.
Completed with errors	Indicates an engineer should review the circuit results because the software raised a warning flag.
Stopped	A user chooses to stop a circuit.
In progress	Analysis is actively running.

Source: PG&E Integration Capacity Analysis (ICA) Data Validation Plan, at p. 8.

NREL recommends HCA managers, after identifying feeders with errors, batch feeders with similar challenges to allow engineers to develop efficient, systematic, and repeatable solutions to common problems. If individual engineers begin fixing feeder model errors before the HCA manager knows the number and type of errors found in a large area, other engineers could end up manually fixing the same problem multiple times when an automated solution would have been more appropriate, or different engineers could create automated solutions that are incompatible.

After batching feeders for review, an engineer is typically assigned to review the failed circuits and solve the problem. To streamline the root cause analysis of failures, it is important to document the procedures used to verify four categories of data: topology, equipment, conductor, and customer consumption and generation. The remainder of this section provides detailed recommendations for each category.

4.2.2.1 Topology Verification

A distribution circuit's geographical information is typically stored in a GIS database. This information is used to build the feeder model. When incorrect geographical information is used to build a feeder model, a feeder topology error will result. Errors stemming from incorrect geographical information should be corrected in both the feeder model and the source GIS database to avoid having to fix the same problem every time GIS data are transferred from the GIS database to the feeder model. We recommend adopting the procedures listed in Table 4 to identify errors in feeder topology.

Table 4. Topology Validation Procedures

Validation Check	Validation Procedure
Unintentional islands	Check for the presence of a cluster of nodes with voltages close to zero. This occurrence may indicate the presence of an unintentional island.
Unintentional meshes	Check for meshes in the feeder. Some radial feeders may erroneously mesh due to incorrect switching states. Erroneous meshes will produce incorrect hosting capacity limits.
Incorrect phase loadings	Checking for incorrect phase loadings is not always straightforward. One possible way is to check voltages at nodes during peak load. Incorrect phase loadings are often caused by errors in GIS data, phasing information, or loads.

Validation Check	Validation Procedure
Incorrect feeder switching states	Check if switch states match field data. If a switch is modeled in the incorrect position, the topology of the feeder model is also likely incorrect.
Incorrect phases for voltage correction equipment	Check the phasing of this equipment matches its placement in the real world. Voltage correction equipment such as regulators or capacitors can be single-phase equipment.
Location of existing DERs	Check whether the model places existing DERs on the correct feeder and in the correct phase.[25]

4.2.2.2 Equipment Verification

Distribution system asset databases are used to track equipment, including transformers, capacitors, and regulators. For an HCA, such regulating devices should be considered at their full operation range. The database typically includes the equipment's nameplate rating, configuration settings (e.g., voltage correction settings), and other information. When equipment information in a distribution system asset database is incorrect, the feeder model typically produces incorrect voltages. Table 5 provides a list of equipment validation procedures.

Table 5. Equipment Validation Procedures

Equipment to Check	Typical Issues Associated with Equipment
Substation data for default settings	Substation equipment with default setting (e.g., switches, regulators, reactors, and load tap changing transformers) need to be verified.
Substation regulator or load tap changer	This device sets the voltage of the feeder head and changes the voltage depending on time and the load it is serving. Incorrect voltage step band, time delays, and ratings need to be checked.
Line regulators	Line regulators should be checked for voltage band, time delays, phase, and control modes. With reverse power flows, control modes available in physical device may not have a match in the software.
Capacitors	Capacitors should be checked for kVAr ratings, voltage triggers, time delays, phase, control modes, and seasonal variations.

[25] CA Pub. Util. Comm., Southern California Edison Co., Advice Letter 4508-E, Improved Integration Capacity Analysis Data Validation Plan, at pp. 6–7 (May 28, 2021) (SCE Data Validation Plan) ("One area that has been particularly challenging is the complete and accurate modeling of DER projects in an automated fashion, especially in cases where multiple DER technologies are present at a single location, e.g. solar photovoltaic and battery storage. SCE is in the process of transitioning from multiple existing legacy DER databases to the Grid Interconnection Processing Tool, while making functionality enhancements in parallel to support accurate modeling of DERs. In the interim, SCE performs the following steps on a monthly basis to validate DER project records across multiple source systems to ensure DER projects are modeled accurately: . . . a. Check for DER project location updates: compare project location information from source systems with current customer connectivity to determine if the DER project's circuit has changed due to system reconfiguration
b. Verify equipment size modeled in Grid Connectivity Model (GCM) matches valid project list from source system
c. Verify DER profile information in GAT matches valid project list from source system
d. Verify final circuit modeling data in CYME gateway matches valid project list from source system
7. Compare current month's aggregate DER nameplate size by circuit to previous month's. Identify significant differences, validate and/or correct source data."), https://library.sce.com/content/dam/sce-doclib/public/regulatory/filings/pending/electric/ELECTRIC_4508-E.pdf.

4.2.2.3 Conductor Verification

A feeder model with incorrect conductor types will produce an incorrect reactive power mix, incorrect losses, and incorrect voltage profiles. HCA experts we interviewed indicated that identifying incorrect conductor types can be challenging. For instance, software tools may use different units to describe conductor lengths (e.g., mile, feet, or meters) and line capacitance (e.g., siemens or ohms), which may not align with the units used in the utility databases. Automating the verification of conductor types is not straightforward and thus typically requires a trained distribution engineer. Table 6 outlines some helpful conductor validation procedures.

Table 6. Conductor Validation Procedures

Validation Check	Validation Procedure
Reactive power	Check for the power factor at the feeder head. If a model shows higher or lower VAr or power factor at the feeder head, conductor inductance or capacitance may need validation. This may also mean incorrect power factor allocate for each load as well.
Circuit losses	Check for aggregate power losses of the feeder. Suppose these values are higher or lower than expected and validate line resistances.
Voltage drops at peak load	Check for voltage drop per mile for peak load allocation; higher or lower voltage drops for peak load allocation should indicate higher or lower circuit losses.
Short circuit currents	Check for short circuit currents; higher or lower short circuit currents at nodes may mean incorrect conductor impedances.

4.2.2.4 Customer Consumption (Load) and Generation Profile Verification

Customer consumption profiles consist of two parts: consumption from the grid (load) and behind-the-meter generation profiles.

Using actual consumption data in the HCA produces more accurate results than using estimated consumption data. For example, in its initial HCA rollout, Xcel Energy's HCA team did not have access to daytime minimum load data for almost half the feeders in Minnesota, so they estimated the daytime minimum load by multiplying the feeder's peak load by 20%.[26] After the Minnesota Public Utility Commission ordered Xcel to use actual instead of estimated daytime minimum load data in the HCA,[27] Xcel Energy reported a significant drop in the number of feeders inaccurately showing zero available hosting capacity.[28]

---

[26] MN Pub. Util. Comm., Dkt. E-002/M-18-684, Order Accepting Study and Setting Further Requirements, Xcel HCA, at pp. 4–5 (Aug. 15, 2019).
[27] MN Pub. Util. Comm., Dkt. E-002/M-18-684, Order Accepting Study and Setting Further Requirements, Xcel HCA, at 14 ("Xcel shall make the tracking and updating of actual feeder daytime minimum load a priority in 2019, and include those values in its 2019 hosting capacity analysis.").
[28] MN Pub. Util. Comm., Dkt. E-002/M-20-812, Xcel Energy Hosting Capacity Analysis Report, Attachment A, at 23 (Nov. 2, 2020) (Xcel Energy 2020 HCA Report) ("The number of feeders with zero maximum hosting capacity decreased by seven from the 2019 analysis, and this was likely the results of using more actual daytime minimum load data for feeders with SCADA in the 2020 analysis.").

Commercial software tools manage consumption data more efficiently and effectively than software developed in-house. Most utilities that perform HCAs use commercial software tools, as they typically include more features, including data validation features, and they are updated regularly.[29]

Consumption data are measured by supervisory control and data acquisition (SCADA) or advanced metering infrastructure (AMI) equipment. SCADA measures the net power flow at a few medium-voltage points, while each customer has their own AMI meter. If a utility has an operating AMI data system, using those AMI data for customer consumption is preferable. To date, utilities have used AMI data primarily to generate customer bills, suggesting it is likely more accurate and precise than SCADA data. Modern AMI equipment can record active power consumption from the grid, and in a separate channel, active power generation from a DER. Modern AMI equipment can also measure voltage and reactive power. AMI measurements are often taken every 15 minutes, and feeders typically have more AMI measurement points than SCADA measurement points. Therefore, there are likely more frequent AMI measurements, and more AMI measurement points, than with a traditional SCADA system. Selected validation procedures for reactive power and load power allocations are provided in Table 7.

Generation profiles should include the maximum potential export for each DER. Generation profiles should include the maximum possible DER output based on solar irradiance information, or a generation schedule for a particular DER. Some utilities maintain their own solar resource database for their service area, but most use open-source databases or third-party vendors.

Table 7. Customer Consumption and Generation Profile Validation Procedures

Validation Check	Validation Procedure
Reactive power allocation	Check whether the load power factor matches the customer class. Each customer class (residential, commercial, and industrial) consumes a different amount of reactive power. Utilities may match reactive power consumption to individual customers or typical power factors for each class. Other times utilities allocate a certain power factor to all loads on a feeder irrespective of customer class. Different assumptions can lead to different errors.[30]

[29] See, e.g., SCE Data Validation Plan, at p. 4 ("SCE has recognized the limited scope of profile validation [using an in-house tool to manage consumption data]. In partnership with SCE's Grid Modernization and Distribution System Planning (DSP) Teams, SCE is in the process of transitioning from [in-house software] to the Grid Analytics Tool (GAT), a commercially supported software tool."); Xcel Energy 2020 HCA Report, Attachment F, at p. 10 ("LoadSEER will allow us to probabilistically simulate DER adoption at a customer level based on system-wide adoption forecasts. This will allow us to study hosting capacity not only based on existing DER on the system, but also based on forecasted levels of DER that may be on the system in the future. Further, LoadSEER will allow us to export forecasted loads at a line section level directly to Synergi – one of the key systems involved in our HCA – which will decrease the amount of time required to allocate load in the Synergi model build process."); SDG&E Data Validation Plan, at p. 5.

[30] See, e.g., Robert Arritt & Roger Dugan, *Comparing load estimation methods for distribution system analysis*, 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013), pp. 1-4 (June 2013); Li Lin, et al., *Effect of load power factor on voltage stability of distribution substation*, 2012 IEEE Power and Energy Society General Meeting, pp. 1-4 (July 2012).

Validation Check	Validation Procedure
Load allocation	Check for accuracy of active power values (kilowatts or megawatts) either as a proportion of aggregate feeder consumption or as individual values. Raw AMI or SCADA load data must be validated and corrected before used in an HCA.[31] For example, load data from abnormal events (e.g., public safety power shutoffs) should be excluded.[32] Typical checks in customer consumption data include, but are not limited to, nonnumerical results, zeros, and blanks.[33]

4.2.3 Scripting and Versioning the Code Bases

Scripting involves developing a deterministic repeatable process that allows engineers to solve problems and make automations quickly and efficiently. These rule-of-thumb strategies shorten processing times, enhance decision-making, and allow teams to function without constantly stopping to think about their next course of action. Scripting is used to automate error correction in the feeder model building process, and it can significantly accelerate decision making. However, it is important to note that just because a script is developed to efficiently fix a data quality problem in the HCA process does not mean the HCA manager should not attempt to fix the root cause of the data quality problem in the source database.

Because feeders are not uniform, scripting code bases are often modified to reflect the needs of one or two unique feeder configurations. If not properly managed, scripts originally developed to be applied across an entire service area may get complex and customized for various feeder configurations, some of which are unique. Accordingly, using code base management tools ensures the HCA team knows who changed the code most recently, can track the evolution of code, and can revert to previous versions if needed. Using central code bases facilitates better script versioning and reduces misalignment in post-processing. For example, once stakeholders alerted SCE that approximately one-third of the HCA data was missing from its data portal, SCE acknowledged the need to improve its code base management. SCE began using a code base management tool to prevent this problem from reoccurring and now regularly publishes complete results.[34]

4.2.4 Prioritize the Screening Process

After baseline distribution circuit validation, the HCA team prepares scenarios. HCA experts we interviewed indicated that, even after an elaborate baseline feeder validation process, they expect the first set of power flow simulation scenarios to produce numerous errors. This is unsurprising, as HCA is an iterative process, where errors are first identified and resolved, and then eventually a useful result is produced. For this reason, it is important for HCA managers to develop and implement an efficient process for identifying and resolving errors.

Therefore, we propose the scenario simulation process begin by examining a prioritized set of load hours and a representative sample of feeders, rather than attempting to perform the power

[31] See, e.g., Pub. Util. Comm. of NV, Dkt. 21-06-001, Nevada Power Co. Integrated Resource Plan Vol. 13, Narrative Distributed Resources Plan, at pp. 36-37 (June 1, 2021) (NVE 2021 DRP) ("Invariably, not all the loading data for all of the feeders represented normal or accurate telemetry."); SCE Data Validation Plan, at p. 4.
[32] PG&E ICA Data Validation Plan at pp. 15-16; Quanta PG&E Assessment at p. 5.
[33] See, e.g., NVE 2021 DRP, at pp. 36-37; SDG&E Data Validation Plan, at pp. 5–6.
[34] SCE Data Validation Plan, at pp. 11-12.

flow simulations for the entire year and the entire service area at once.[35] The prioritized power flow simulations can help identify issues that are likely to be found throughout the distribution system. We recommend prioritizing critical load hours, including summer peak, summer minimum, winter peak, winter minimum, summer daytime peak, summer daytime minimum, winter daytime peak, and winter daytime minimum. In addition, we recommend selecting a set of representative feeders to prioritize. After errors in the selected hours and feeders are identified, the errors can be fixed throughout the system (in the source database or using automated scripts). This way, the manual intervention required to fix identified problems in later analyses is minimized.

4.3 Validation of HCA Results Before Publication

We recommend HCA managers establish a process to spot errors in HCA results and correct them before publication. As discussed in Section 1, utilities have published HCA results that, upon review by stakeholders, included clearly erroneous data. Therefore, we recommend that after the utility performs the power flow simulation, it establishes a process to flag irregularities that will trigger a review before publication. Table 8 provides a consolidated list of triggers and validation procedures. Most HCA experts we interviewed perform prepublication reviews, but they noted that the feeder model building process is most commonly the root cause of errors identified in the visualization and data publication processes.

Table 8. Consolidated List of Triggers to Validate HCA Results Before Publication

Validation Check	Validation Procedure
No (null) or invalid results	Check for null or invalid results. Implement rule-based screening for null or invalid results, for example:[36] • Are results present for all hours? • Are more than 20 node-hour results blank (null)? • Does the number of null results increase by more than 5% in the current HCA cycle compared to the previous HCA cycle? • Are results numeric? • Are there any null nodes in the final output map?
Zero hosting capacity available	Check for zero hosting capacity values. Implement rule-based screening of zero hosting capacity sections to identify potentially erroneous results. Trigger based on count of feeders or nodes: Most utilities check all HCA results for false negatives, manually reviewing a feeder model if the results show no hosting capacity remains on the entire feeder or when results for 10% or more nodes equal zero for each study criterion.[37]
Duplicate entries	Check for repeating or duplicate entries. Implement rule-based screening for duplicate entries, for example:[38] • Check for duplicate records in final output map.

[35] PG&E ICA Data Validation Plan at pp. 9–10.
[36] SCE Data Validation Plan, at pp. 9–10; Quanta SCE Assessment at pp. 7–8; PG&E ICA Data Validation Plan at pp. 10-11; SDG&E Data Validation Plan, at p. 3.
[37] SCE Data Validation Plan, at p. 9; Quanta SCE Assessment at p. 7; SDG&E Data Validation Plan, at pp. 6–7.
[38] PG&E ICA Data Validation Plan at pp. 10-11; SDG&E Data Validation Plan at p. 7.

Validation Check	Validation Procedure
	• Check for duplicate records in network section table. • Check whether the node has the same or repeating result in more than two hours.
Large discrepancy between previous HCA cycle results and current HCA cycle results	Check for variation between previous HCA cycle results and current HCA cycle results. There are multiple ways to do this, including:[39] • Review the most limiting result for all line segments and bin them into the ranges. Where the binned results for the current cycle differ by more than 5% from the binned results of the previous cycle, flag those results for review. • Review significant changes in the number of times a certain technical criterion is violated.
Random and spot checks	Check a certain number of randomly selected feeders or frequently error encountered circuits. Check for false positives and false negatives. False negatives are results that are not zero but are nonetheless incorrect (e.g.*,* the model produced a result of 100 kW, but the actual result should have been 500 kW).
Additional triggers	Additional miscellaneous checks include:[40] • Check for changes in results because of software upgrades or other changes (e.g., switching to a new version of power flow simulation software or switching to a new load database). • Check for differences in load profile variation and nodal results that could signal an error (e.g., if a load profile varies over time but the hosting capacity at a node does not).

4.4 Acceptance of Feedback from Customers and Users

As explained in Section 1, HCA data users can (1) identify errors that the utility is unaware of and (2) suggest improvements the utility would not conceive of on its own. Therefore, we recommend utilities provide a mechanism to allow customers and HCA data users to provide feedback about the user experience, any errors identified, and the usefulness of the HCA data provided.[41] Utilities can track the feedback they receive and report on any actions taken as a result of the feedback.

4.5 Regulatory Oversight of Data Validation Processes

Utilities can perform data validation independently or regulators can oversee these efforts, for example by requiring a utility to submit a data validation plan and periodic reports including metrics that track the quality of HCA results.[42] We recommend regulators provide transparency into the data validation process by reviewing and requiring improvements to HCA data

---

[39] Xcel Energy 2020 HCA Report, Attachment A, at p. 18; Quanta PG&E Assessment at p. 7; SCE Data Validation Plan at p. 10.

[40] Quanta SDG&E Assessment at p. 15; *id.* at p. 7.

[41] Quanta PG&E Assessment at p. 16.

[42] CA Pub. Util. Comm., Dkt. 14-08-013, Administrative Law Judge's Ruling on Joint Parties' Motion for an Order Requiring Refinements to the Integration Capacity Analysis, at pp. 4–6 (Jan. 27, 2021).

validation plans. For example, regulators could require submission of a draft data validation plan, accept feedback from stakeholders on the draft, and then require the submission of an improved data validation plan with periodic reporting on its implementation. HCA experts we interviewed indicate that stakeholder involvement in regulatory processes results in HCAs that provide more useful data to customers. For example, regulators can allow stakeholders to propose that the utility follow certain processes, and regulators can undertake an independent review of HCA processes in other states.

4.5.1 Require Data Validation Plans to Describe the Utility's Data Validation Processes

We recommend regulators require utilities to prepare a data validation plan that describes the utility's data validation process. The plan should identify the HCA manager and describe that person's responsibilities, as outlined in Section 4.1.1. Data validation plans should also describe the employee and computation resources devoted to HCA implementation and data validation. As a part of this description, we propose the utility describe how its plan balances the use of computerized and manual processes, as described in Section 4.1.4.

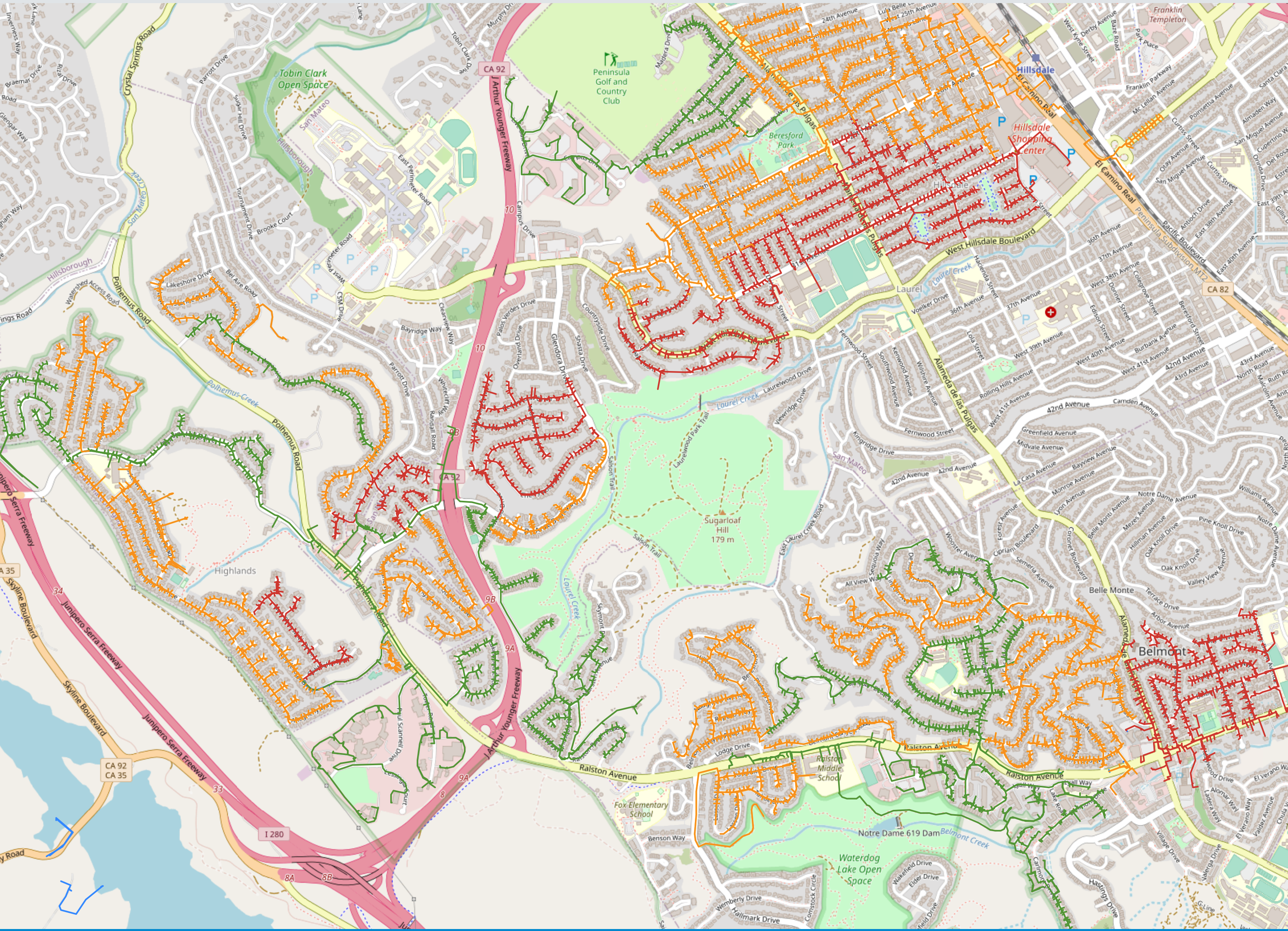
The data validation plans should describe the utility's standardized approach to resolving errors, as outlined in Section 4.2.2, including the use of a prioritized screening process as described in Section 4.2.4. This can include the processes used to verify the baseline model, feeder topology, equipment, conductors, load profiles, and generation profiles. Finally, the plan should describe how the utility validates HCA results before publication, as described in Section 4.3.

4.5.2 Require Periodic Reports to Track the Quality of the HCA Results Over Time

We recommend regulators require periodic reports of metrics that monitor the utility's performance and the accuracy of its HCA results. The reports could include summaries of how many circuits have passed, produced warnings, or failed at each milestone in the HCA process. The reports should include a root cause analysis for recurring problems in the HCA process and action plans with implementation timelines identifying improvements to fix recurring problems. If the action plans do not fix the problems at their source (e.g., a script that automatically corrects errors after importing data from the source database, rather than fixing the errors in the source database), the reports should identify the reasons problems could not be fixed at the source. Finally, we recommend the reports summarize feedback provided by customers concerning the user experience, problems identified, and usefulness of the HCA data (as outlined in Section 4.4) and either action plans for resolving issues identified or explanations why the utility cannot fix the problems or believes doing so is unnecessary.

5 Conclusion

This report is designed to provide utilities, regulators, and stakeholders a set of best practices so that future HCA deployments provide useful, trustworthy, and accurate data from the day they are published, thus avoiding the pain points experienced in early HCA deployments. These best practices include the use of QA, QC, and regulatory processes to help produce accurate and reliable results. We outline how these robust HCA data validation processes can be supported through the establishment of a standard HCA business processes and a well-resourced team. We provide tables with specific examples of validation procedures for each step of the feeder model development and validation process. Finally, we suggest ways to ensure HCA results are valid before they are published and propose that regulators oversee the data validation process by requiring a written plan and regular reporting on data quality metrics.



NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency and Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

NREL/TP-6A40-81811 • April 2022

National Renewable Energy Laboratory
15013 Denver West Parkway, Golden, CO 80401
303-275-3000 • www.nrel.gov

NREL prints on paper that contains recycled content.

Map on front and back cover from NREL