Quantifying Losses and Failure Rates in Large-Scale PV Fleets with the SUPER Benchmarking Tool

Daniel Fregosi, Ph.D.

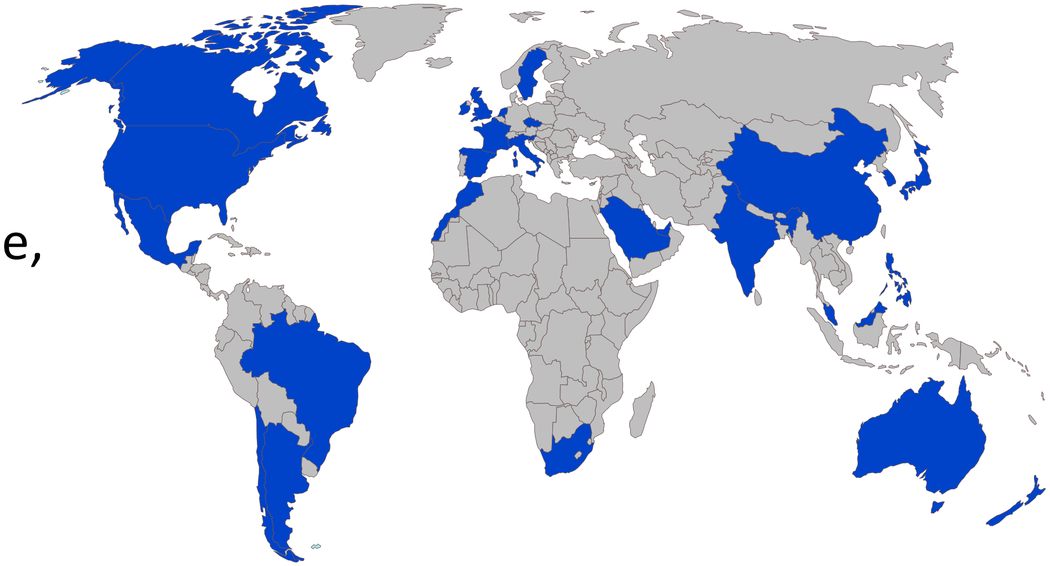
PVPMC
May 13, 2026

www.epri.com

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technology Office Award Number DE-EE0008976.

Who is EPRI?

- EPRI is a non-profit research institute focusing on clean, reliable, safe, affordable energy
- Our research is independent, for public benefit
- Funded primarily by membership
- Worldwide collaboration - over 60 entities in EPRI Renewables



EPRI Renewables – Worldwide Collaboration (60+ Entities)

PV Plants are underperforming financial P50 estimates by 5-8%*

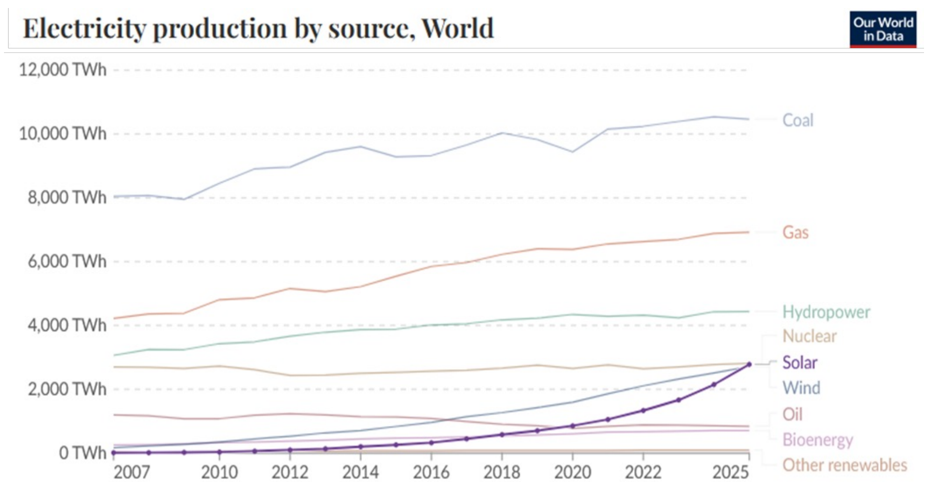
*Solar Risk Assessment 2025, kWh Analytics, DNV and others

- Why are plants underperforming?
  - **Possible** causes listed for underestimation: inverter downtime, DC health, shading, snow, curtailment
  - “monthly operating report data lacks sufficient detail to pinpoint root cause” – kWh Analytics

There’s a knowledge gap in industry-wide performance benchmarking

Why Benchmarking? As the solar industry rapidly expands, benchmarking enables feedback and improvement

- Industry-wide benchmarking answers:
  - What is typical **performance** and **reliability?**
  - How much is each sources of loss?
    - Weather
    - Inverter Downtime, Tracker Downtime, System Loss
    - Clipping, Curtailment, Irregular Performance
  - What are common component failure rates and costs?
- Quantify the impact of:
  - Upfront design decisions
  - OEM and EPC selection
  - Factors like technology, design, location, vintage, age, etc.
- Enable feedback loops with data
  - Inform upfront design decisions to reduce cost and improve energy
  - Inform maintenance strategy
  - Inform prediction models*

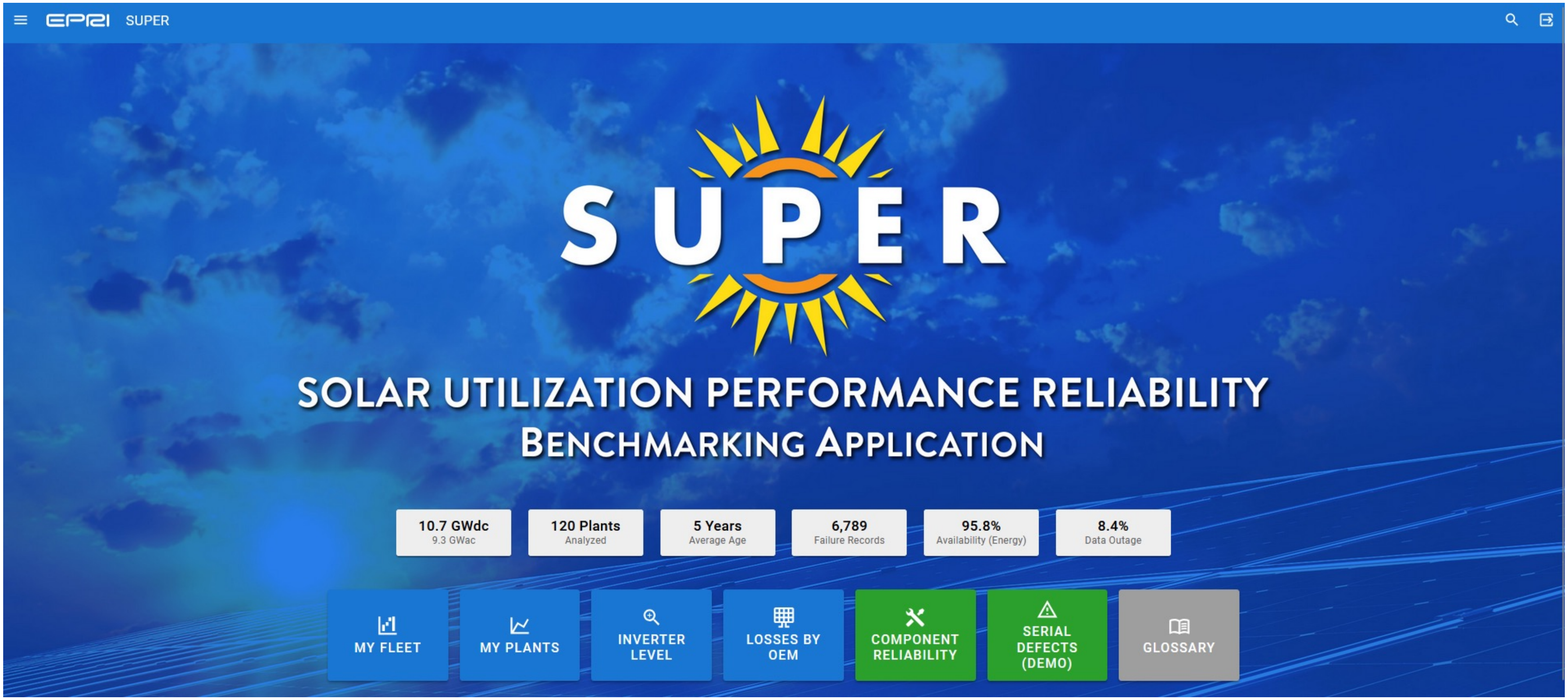


https://ourworldindata.org/grapher/electricity-production-by-source

* Performance Modeling Best Practices https://www.epri.com/research/products/000000003002024791

Benchmarking can answer these questions through….

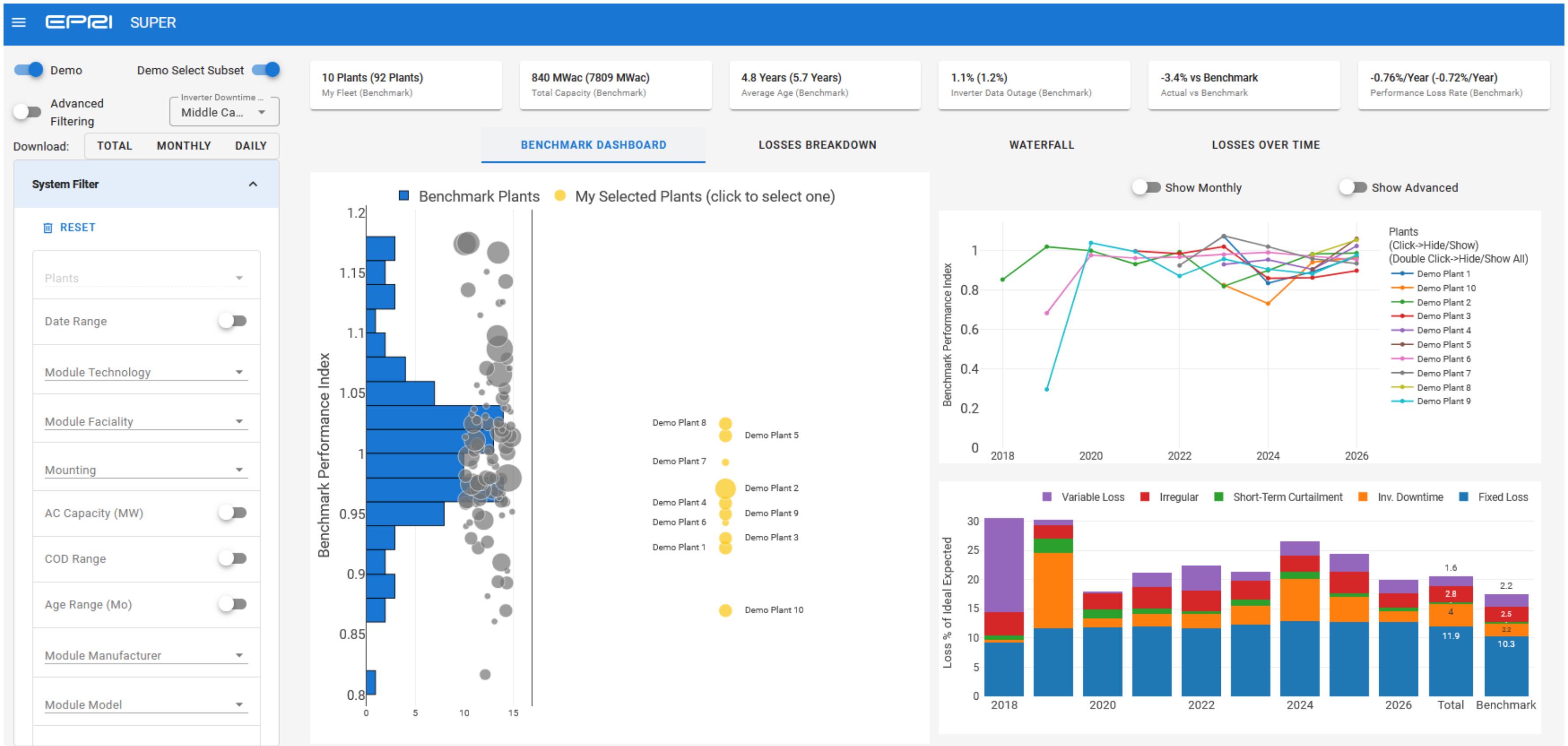
- Large (anonymized) datasets
  - Diversity of location, technology, manufacturers, EPC, O&M
  - Large samples sizes for better statistical conclusions
    - Ability to slice the data and make comparisons
- Consistency and transparency in analysis
  - Identical methods applied to all datasets (**using inverter-level, raw data**)
    - Quality control, satellite weather, models, KPI formulas
  - Methods shared with participants
- Collaboration through user group
  - Feedback and improvement
  - Industry knowledge sharing



SUPER.epri.com

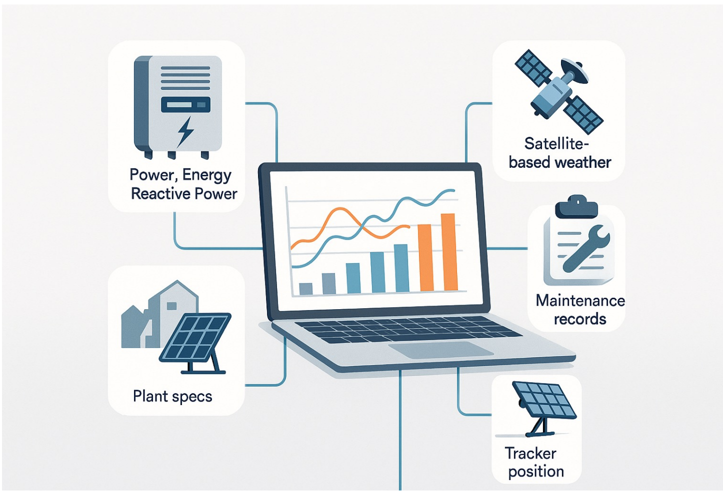
Overview Slides and Glossary available on public page

Compare your plants against anonymized database. Filter by OEM, technology, design, age, etc.



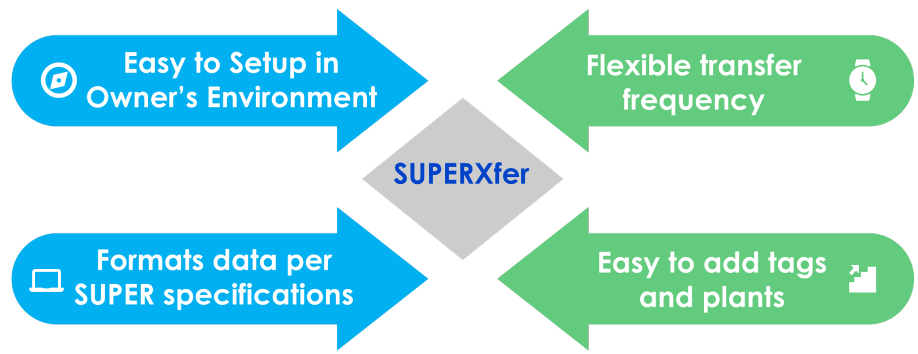
Methodology

Data



- Primary analysis is based on:
  - 5-minute SCADA data
    - Per inverter: Power, Energy, Reactive Power
  - Satellite-based weather data
  - Plant specs – design, manufacturer, model
- Additional data:
  - Maintenance records (CMMS)
  - Tracker position (SCADA)

Data transfer is made easier through an EPRI-provided python script that interfaces with the data historian API

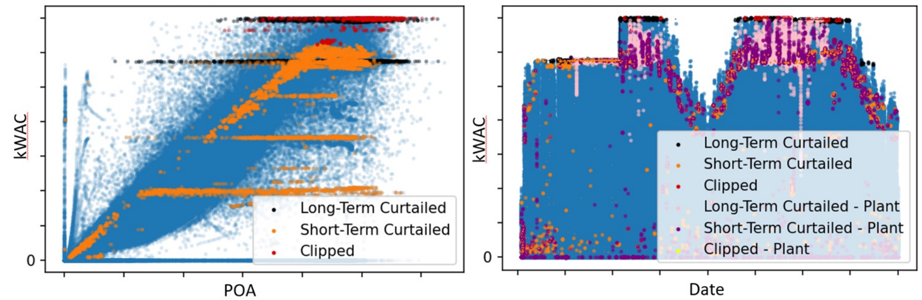


EPRI

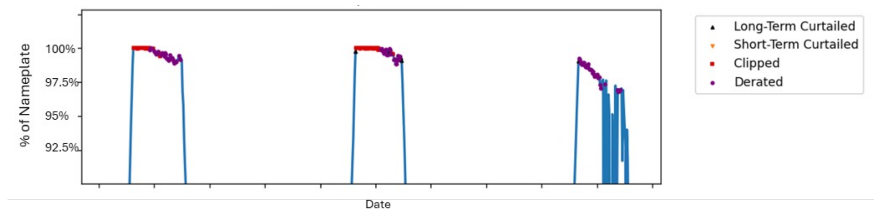
Data QC and Flagging

- All raw data is processed through the same quality control methods to:
  - Flag bad data
    - Stuck, interpolation, physically impossible
  - Identify unit errors, time zone errors
  - Verify Metadata
  - Label clipping/curtailment/temperature derating
  - Classify missing data as data outage or power outage
    - Using energy meter
    - Plant-level meter
- **More details on poster!**

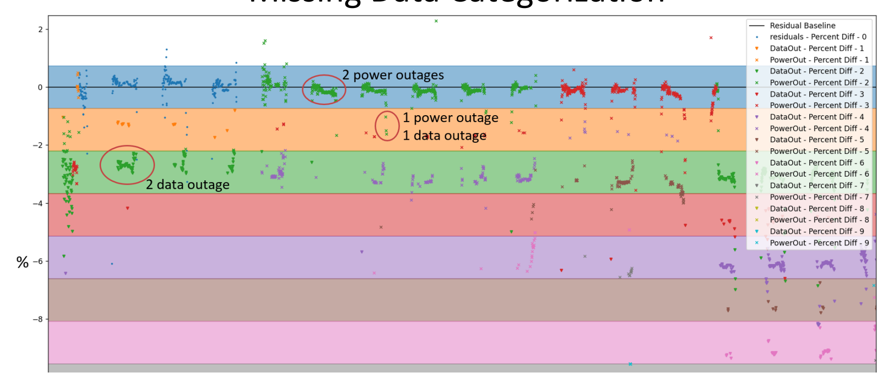
Clipping, Long-term and Short-term Curtailment



Inverter Temperature Derating

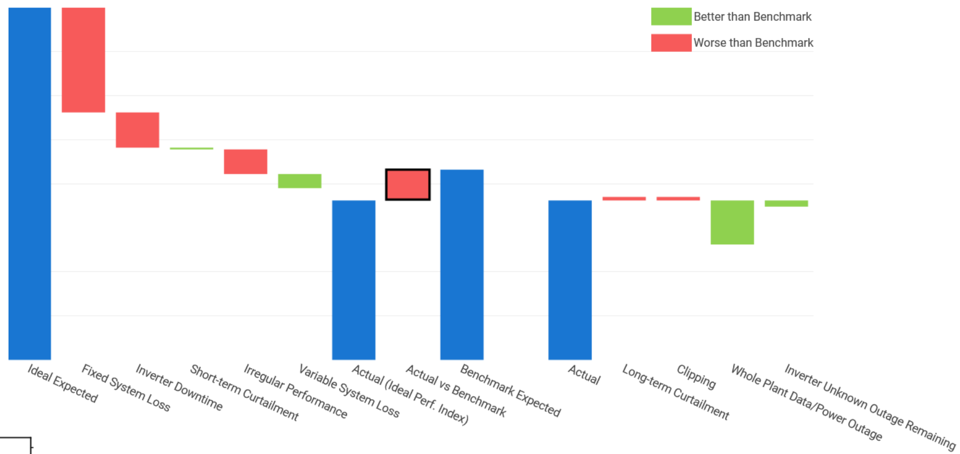


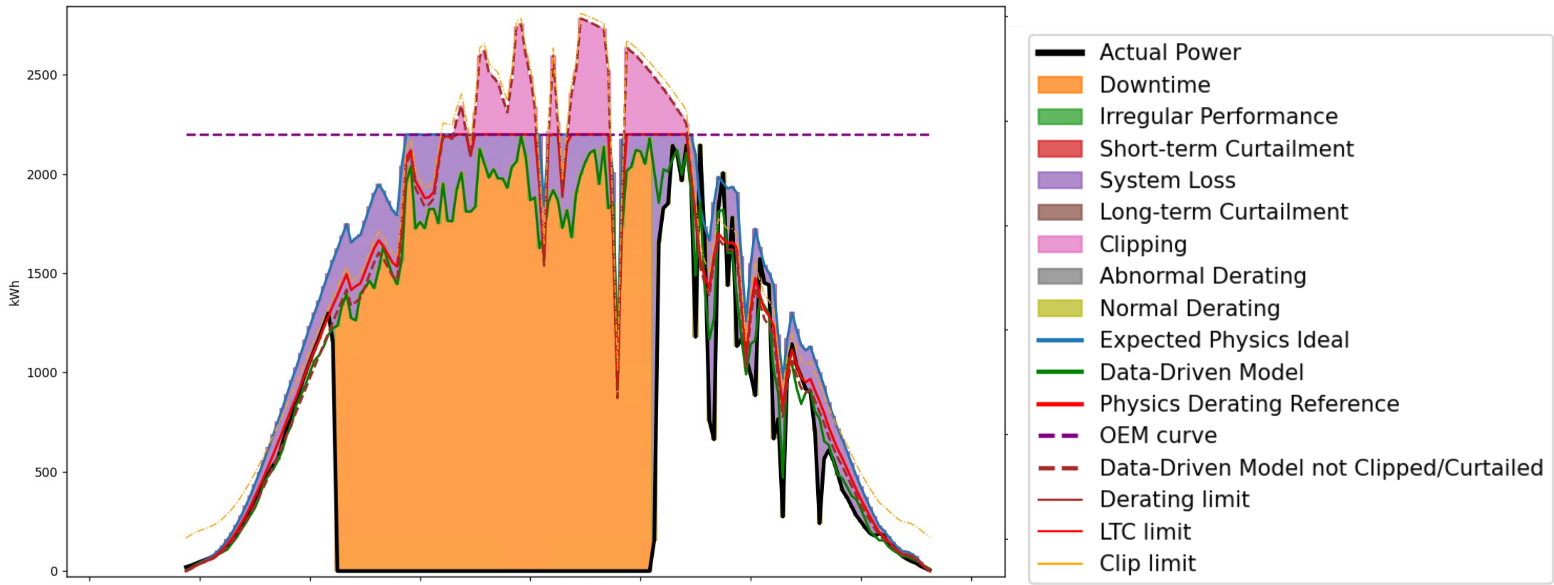
Missing Data Categorization



Loss Categorization

- From **Ideal** to **Actual** – losses are categorized and quantified
  - Using detected clipping/derating levels and pattern detection algorithms

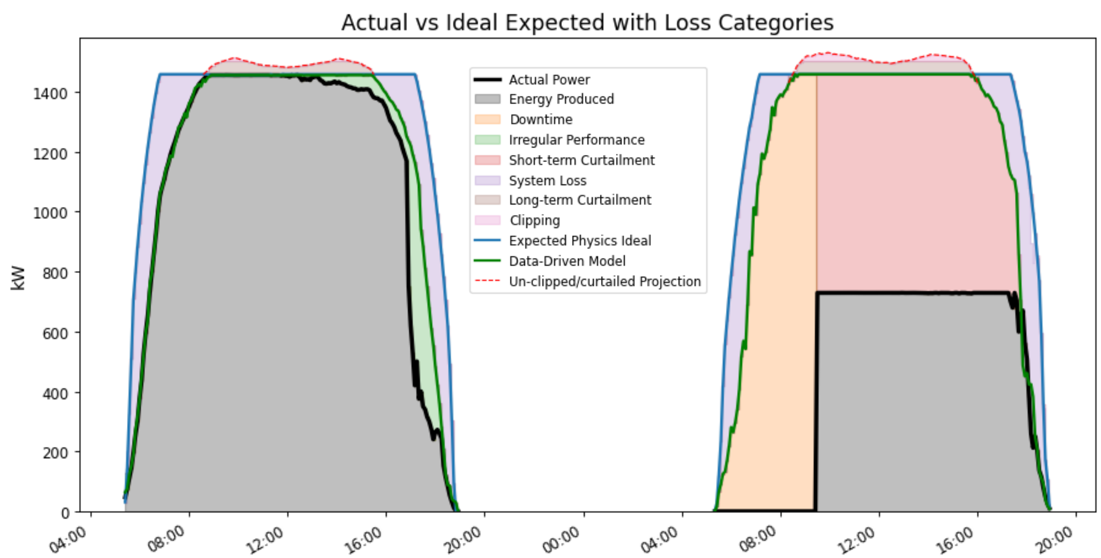




In development:

- Stuck-trackers
- Tracker-error
- Soiling
- Snow

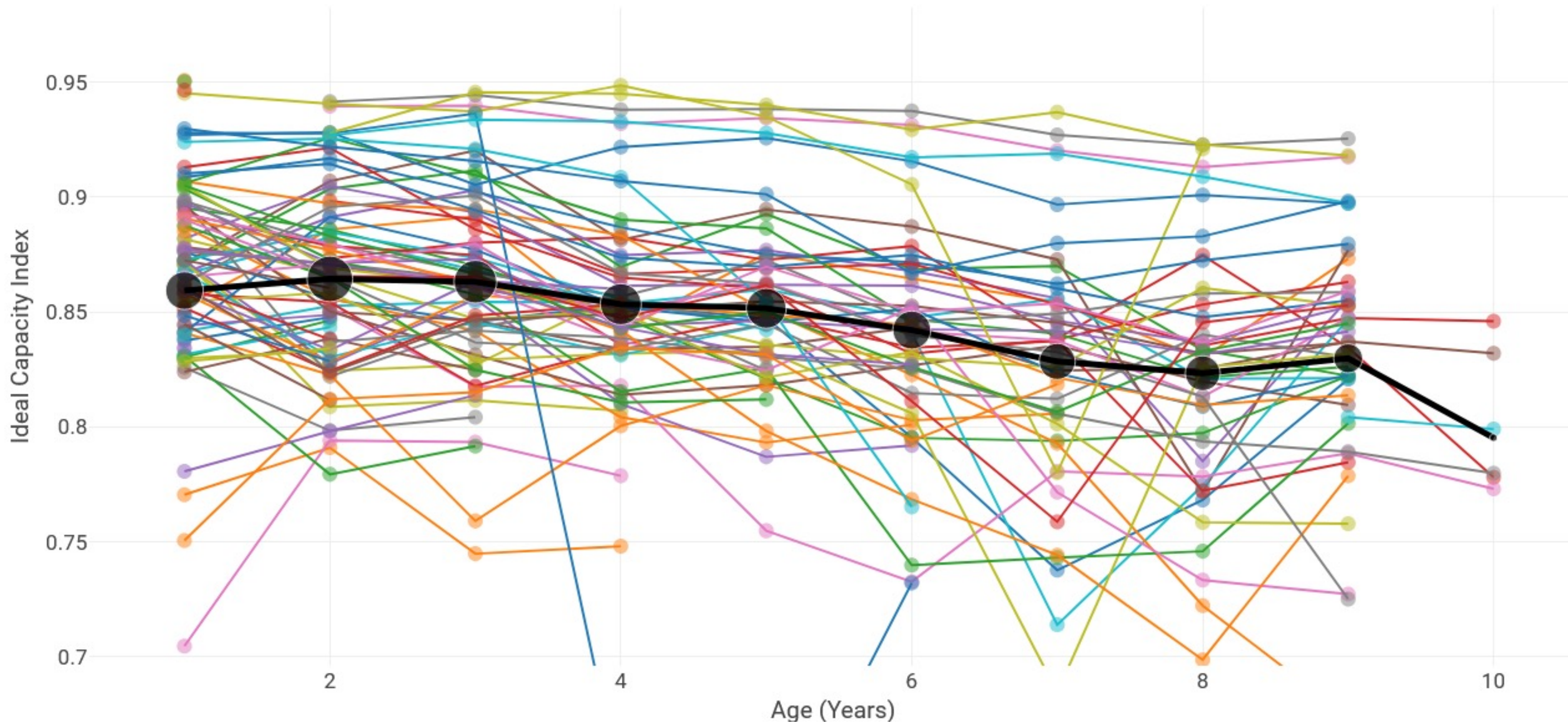
Definitions posted publicly on SUPER.epri.com



- **Ideal Expected Energy**: Theoretical energy produced given the plant's nameplate rating, design, and weather. Ideal expected energy is adjusted for the actual weather the plant experienced, including irradiance and temperature. A physics-based model of the plant is used to calculate expected power using plant specifications such as clipping/curtailment level, module temperature coefficient, module faciality, crystalline or thin film, ground coverage ratio, tilt, azimuth, and tracking. A generic inverter loss model simulates AC losses. No additional losses are included.
- **Fixed System Loss**: Efficiency loss inherent to the plant design and construction, not as a result of degradation over time. This baseline/constant loss is measured by average efficiency during the first 1-2 years of operation. Factors that may contribute to fixed system loss include module nameplate error, module mismatch, wiring and other ohmic loss, orientation error/uneven terrain, tracking error, self-shading, etc.
- **Variable System Loss**: Efficiency losses that accumulate over time while a plant is in operation. Variable system loss can come from module degradation, soiling, faults and failures of modules and wiring components, shade from vegetation, tracker controller drift, etc.
- **Inverter Downtime**: Energy lost due to inverter power outages.
- **Irregular Performance**: Energy lost due to a temporary and severe reduction in power. Potential causes include snow cover, offline combiners, severe shading, significant maintenance. Inverter outages and curtailment are quantified separately and not included in irregular performance.
- **Short-term Curtailment**: Energy lost due to a reduction of capacity by AC setpoint, lasting hours to days. Times of short-term curtailment are detected by flat peaks in the AC power at levels lower than the typical setpoint.
- **Benchmark Expected Energy**: Expected energy to be produced given the nameplate, design, weather, and **typical losses**. Ideal expected energy is used as the starting point for calculating benchmark expected. Then, typical values for each of the losses are subtracted from the ideal physics model to calculate benchmark expected energy. Typical loss values are calculated from the plants in the database after filtering selections are applied, and updated dynamically as filter selections change.
- **Long-term Curtailment**: Energy that theoretically could be gained by increasing the AC setpoint of the inverters to their hardware limit. This category quantifies energy that is curtailed either due to an interconnection limit or a derating of the hardware. The curtailment setpoint could be either plant-level or inverter-level. This differs from short-term curtailment in that the long-term curtailment settings are persistent across months to years. Unlike short-term curtailment, long-term curtailment does not count against the plant in overall KPI's (Actual vs Benchmark, Ideal PI, Ideal CI, or Benchmark PI) since it is part of the design of the plant. Therefore, the ideal expected and benchmark expected models are curtailed to the same long-term curtailment level as the actual plant.
- **Clipping**: Energy that is clipped due to the hardware limit of the inverters. Having clipped energy is a normal part of PV plant designs where the DC/AC ratio is chosen to optimize LCOE. Clipping does not count against Actual vs Benchmark, Ideal PI, or Benchmark PI since the ideal expected and benchmark expected models are curtailed to the same level as actual.

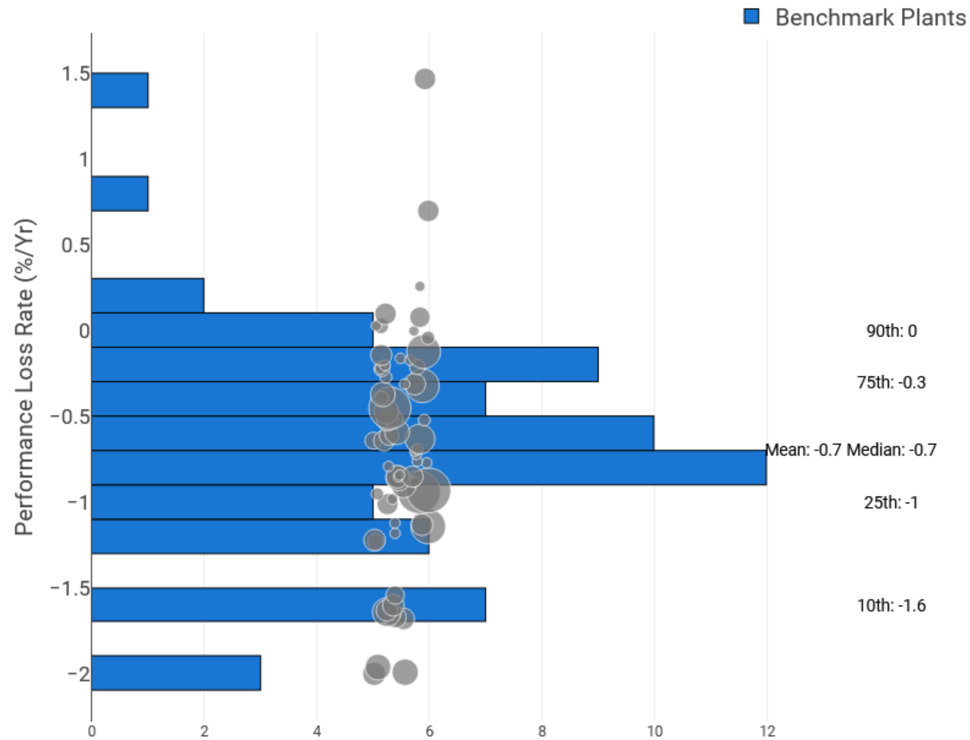
Performance Benchmarking Results

Capacity Index measures initial and ongoing capacity during uninhibited plant operation



Performance Loss Rate captures module and system degradation

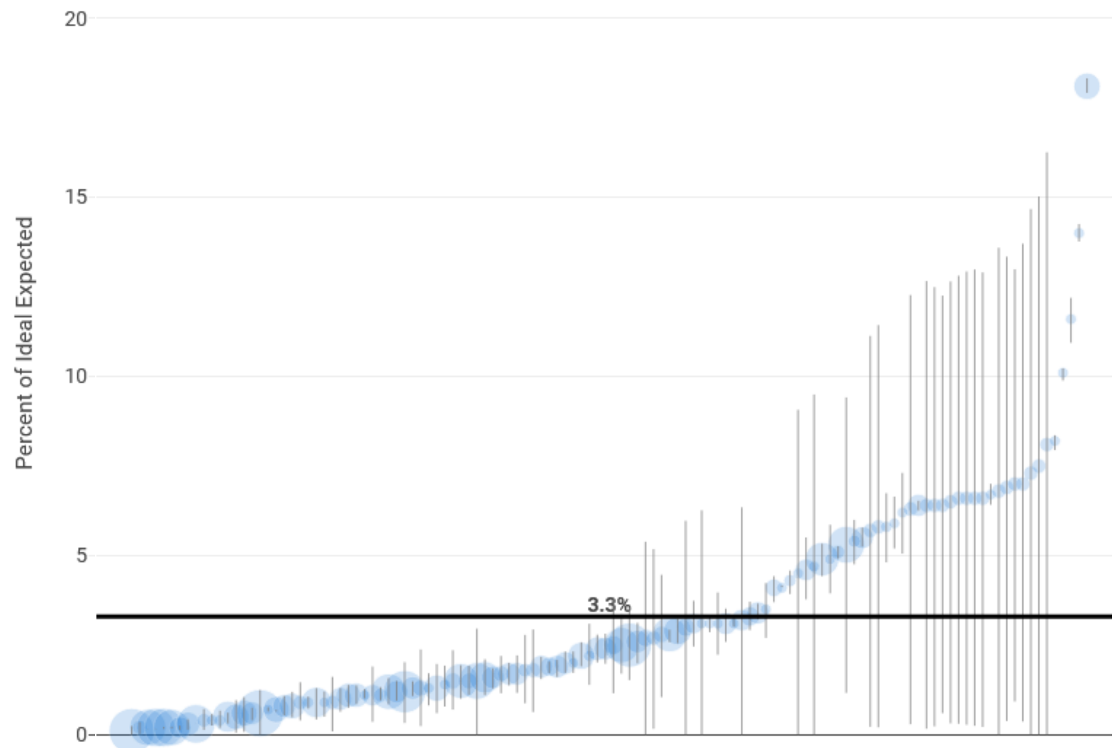
- -0.7%/yr mean/median
  - IQR: -0.3 to -1.0%/Yr



Inverter downtime has ±1.5% uncertainty due to missing plant/energy meter data

- Worst-case: 4.8%
- Middle-case: 3.3%
- Best-Case: 1.8%

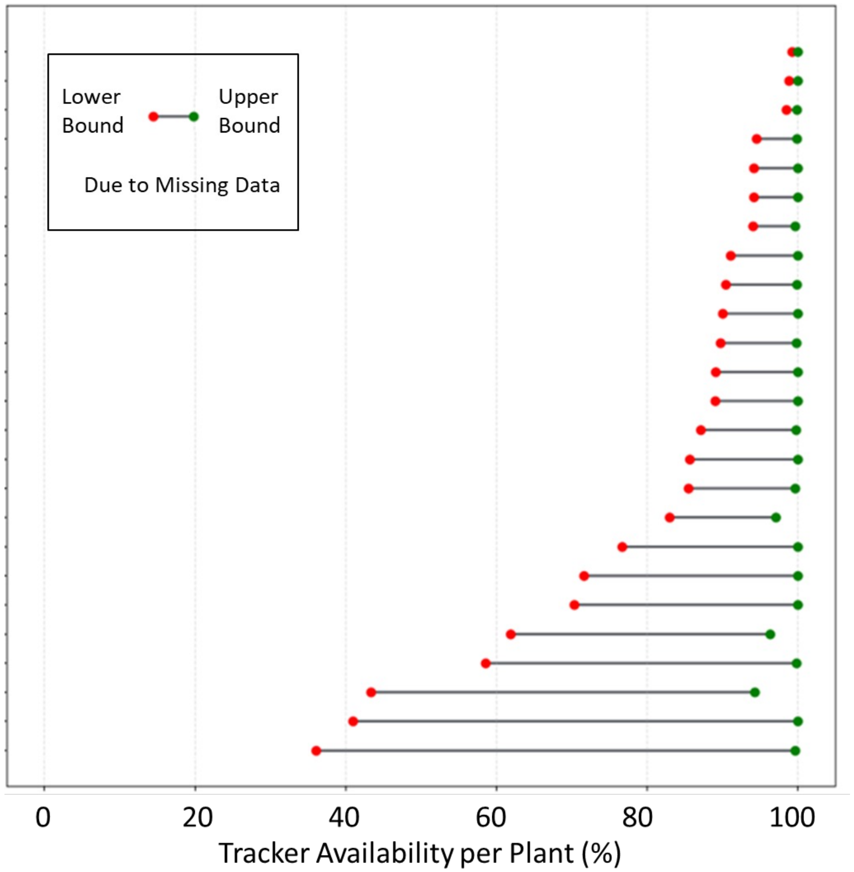
- Whole plant data outage: 8.4%



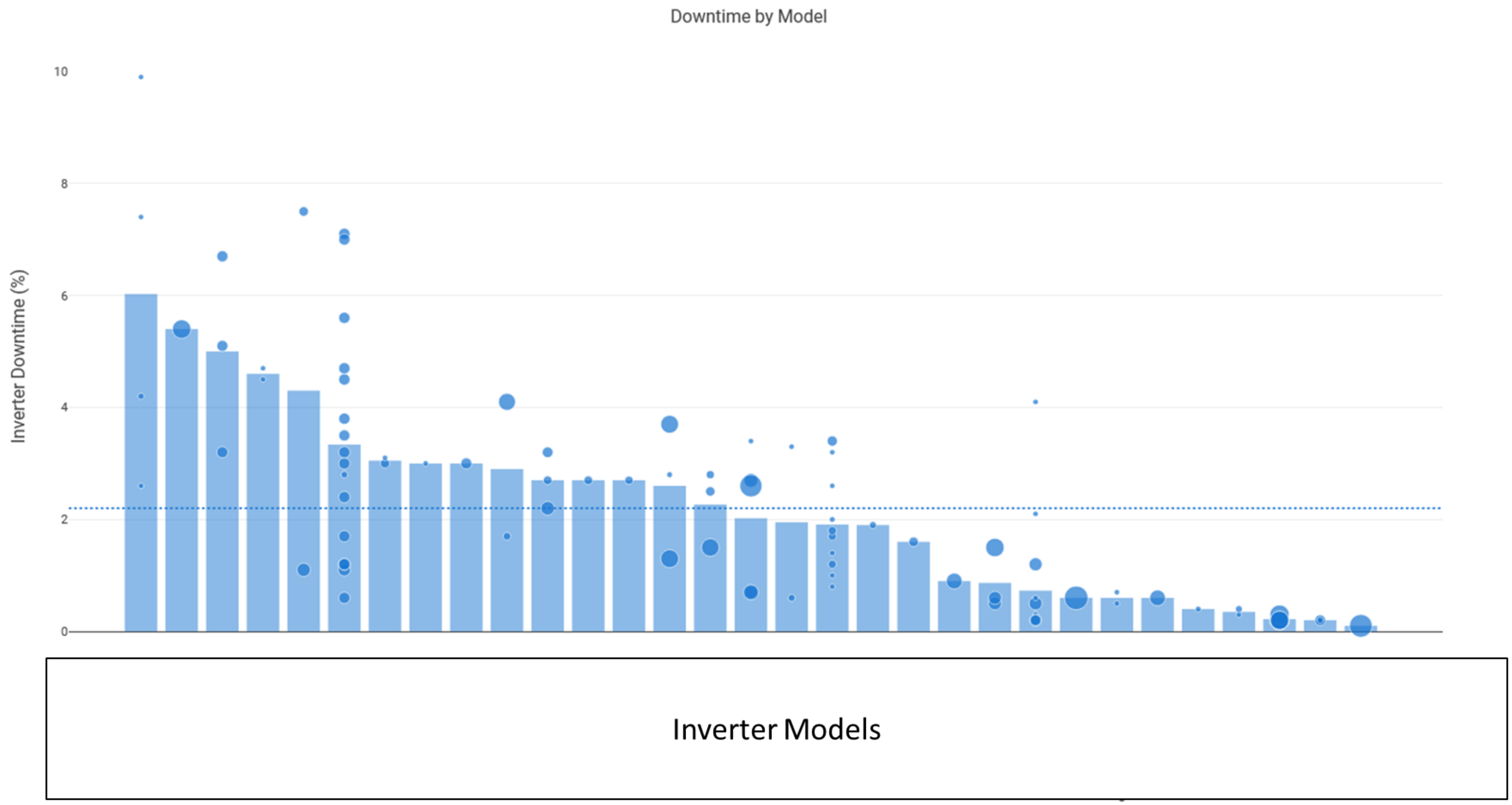
Energy and plant-level power data is used to distinguish communications and power outages whenever available. Uncertainty here is due to lack of these additional channels.

Tracker availability (stuck trackers) preliminary results

- Median ~90-95% availability
- Missing data leads to wider band than inverter downtime
- Still gathering data



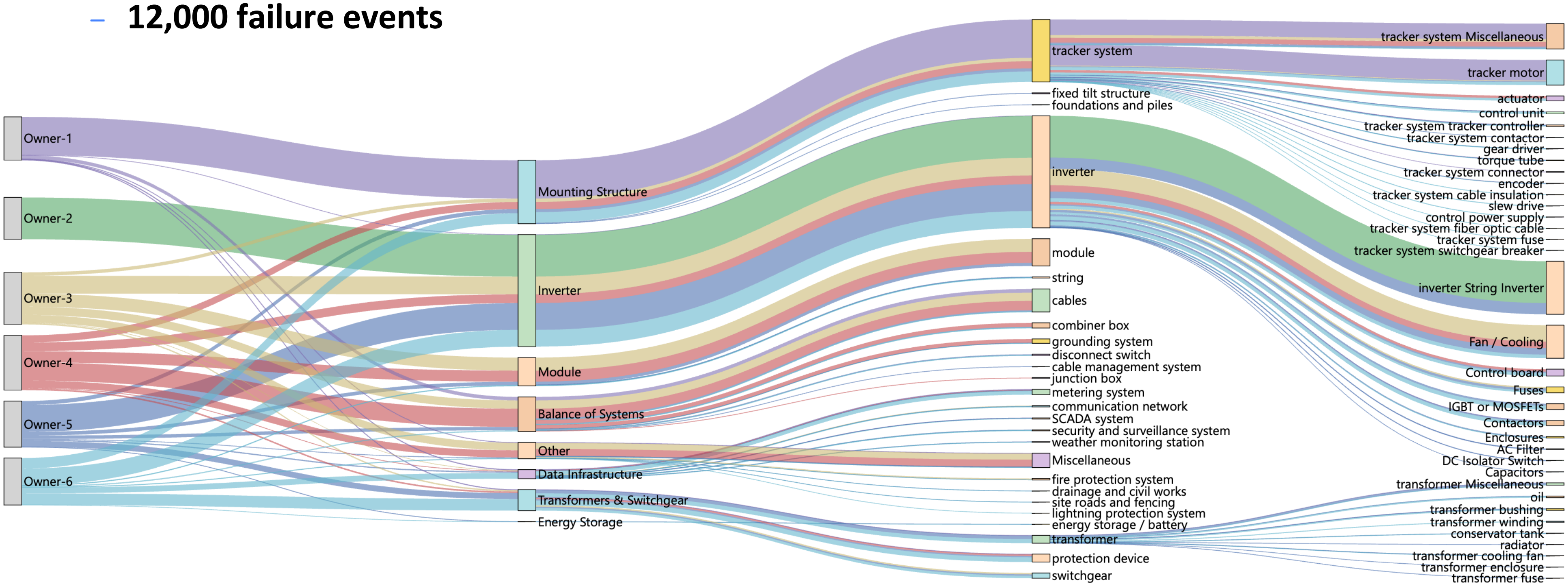
Losses are tracked by Inverter, Module, and Tracker Model



EPRI

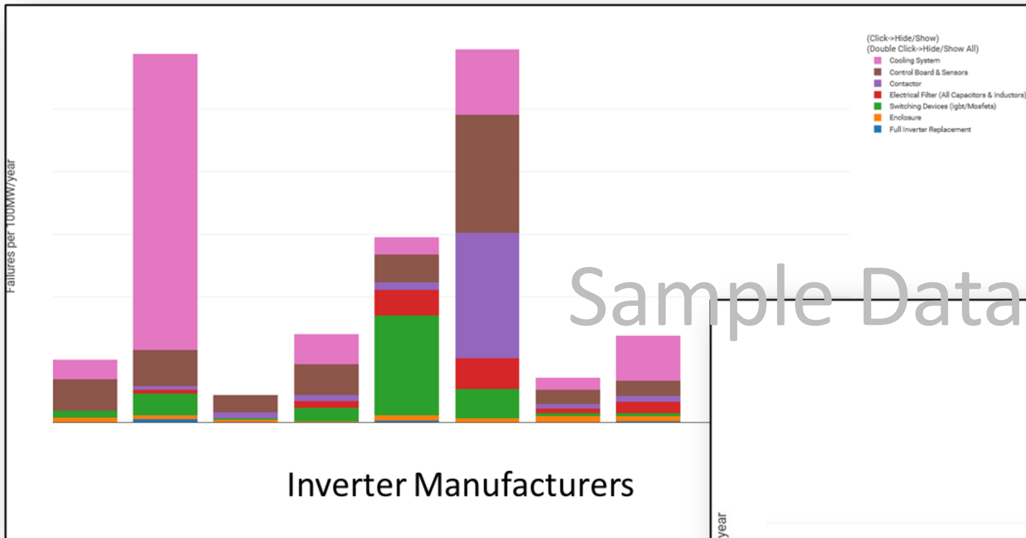
Reliability Benchmarking

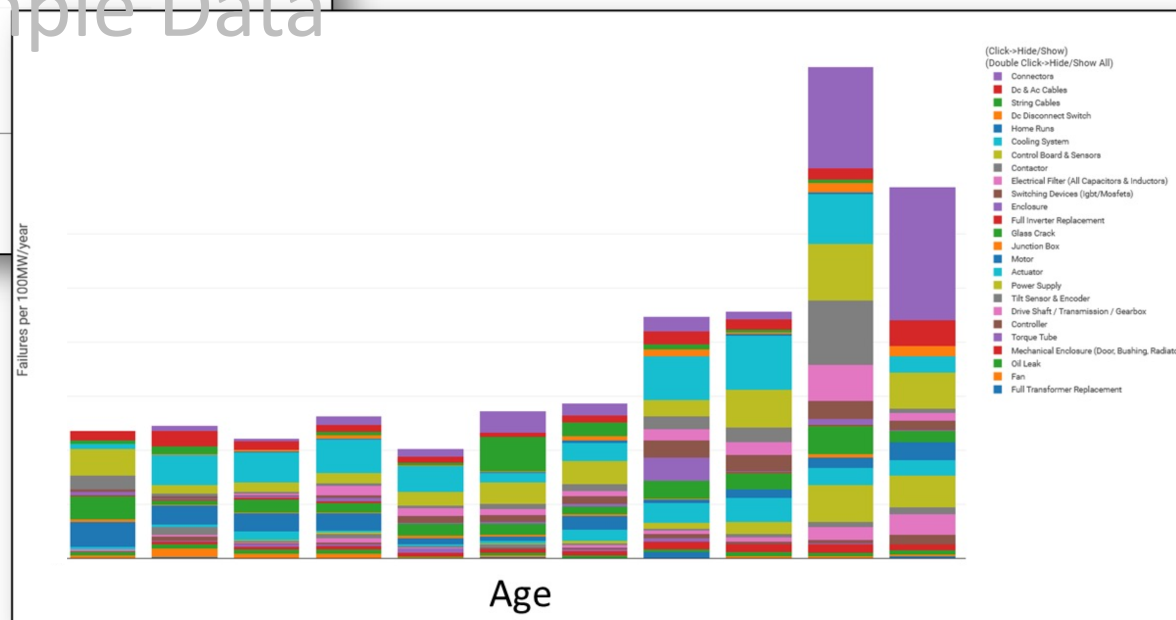
- Categorized over 125,000 maintenance records (57 plants) using an LLM
  - Standardized component, sub-component
  - Action: replace/repair, troubleshoot, inspect, other
- Focused on failures (replace/repair) for measuring component reliability
  - Excluded: scheduled/preventative maintenance, protection, extreme weather-related
  - **12,000 failure events**



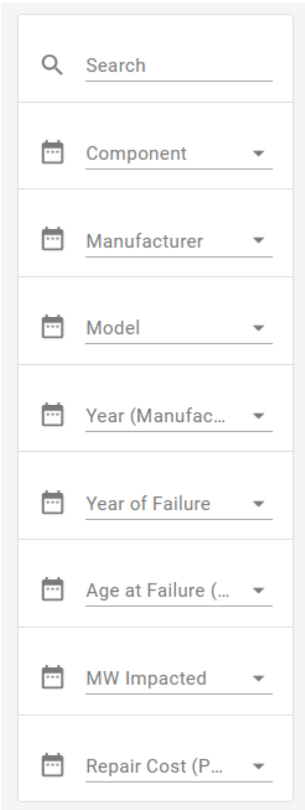
Normalized Failure Rates by Component/Subcomponent

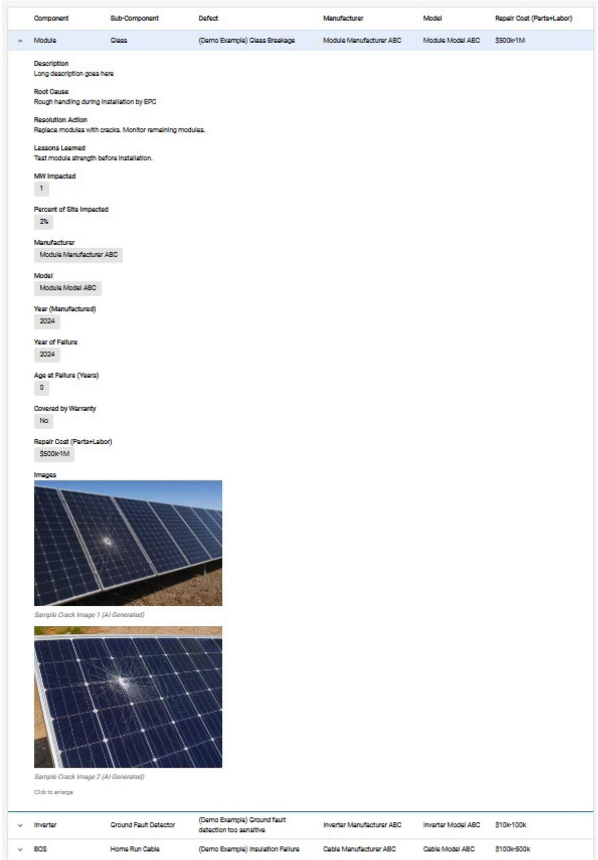
Breakdown by OEM, Age, and Plant, Weighted by Cost

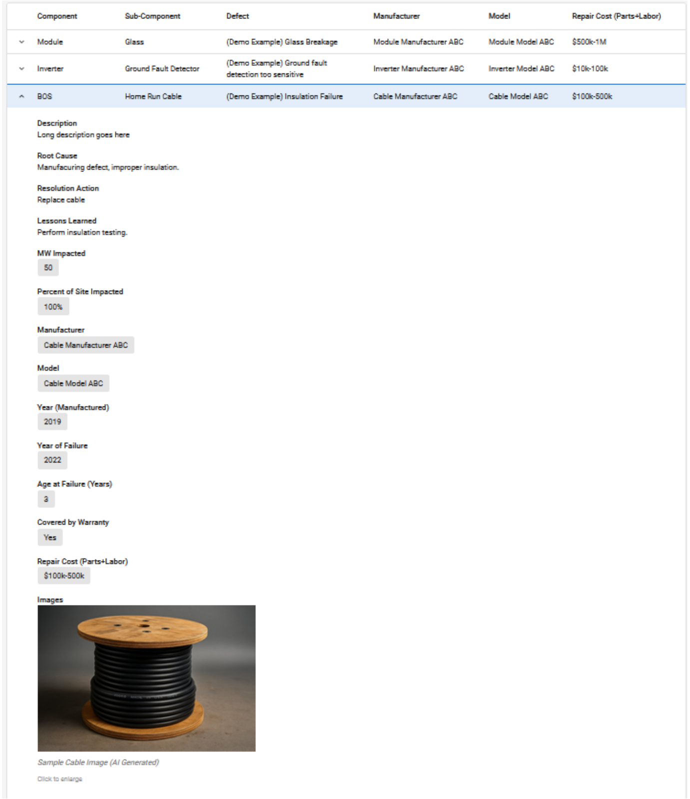




Library of Serial Defects: search and filter by component, manufacturer, etc.







TOGETHER…SHAPING THE FUTURE OF ENERGY®

www.epri.com