Reliability Modeling Input Data – Two Case Studies

Peter Burgess, Ryan O'Connor. EDF power solutions

A good understanding of equipment reliability is an important factor for commercial success over the lifetime of solar and storage projects. It is essential to making informed decisions about maintenance strategies, resource allocation, and managing risk.

Achieving this objective of a complete and accurate understanding is dependent on the quality of input information. If input data is inaccurate or incomplete, the analyses which result will themselves lack accuracy, potentially leading to suboptimal allocation of resources.

We have found that the information needed to perform high quality assessments of equipment reliability are spread across multiple sources, each of which may lack for completeness or accuracy. We believe that these sources can be mutually reinforcing, layering up data sources to strengthen trust in data, and to provide a more complete picture by closing information gaps which would emerge if only one source were used. This is analogous the the Swiss cheese model as used in safety systems.

Two case studies are presented to convey examples of how a data layering approach improves the accuracy and completeness of the resulting information to serve as inputs to reliability models.

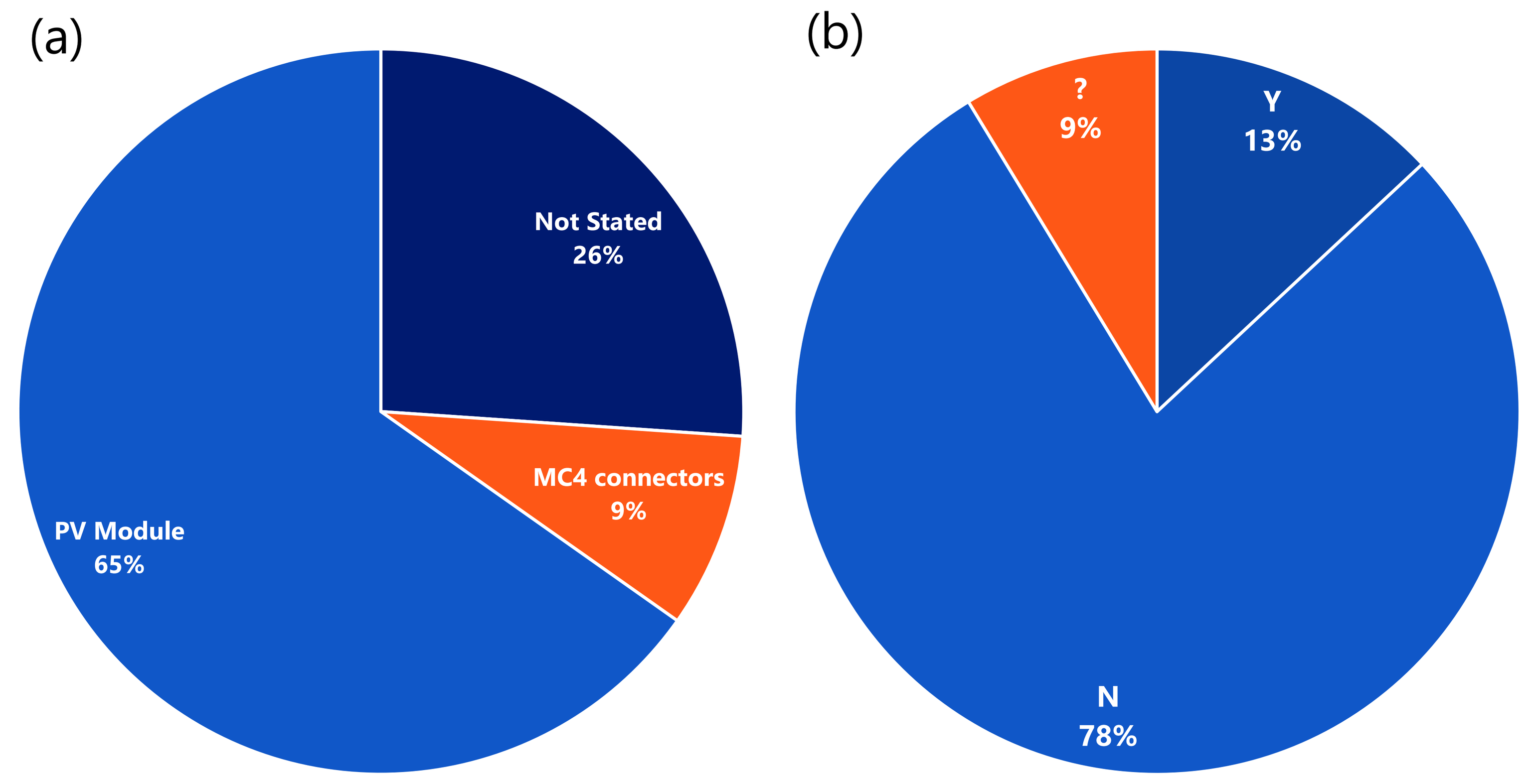
Case Study 1: All service orders for ground faults in 2025 that reference a load break disconnect or string isolation.

This case study gives a longer timeline for a specific fault type to give a flavor of completeness in the internal service order data. Expectations for best practices in documentation are that each service order will:

- Identify failed components
- Document parts consumption to show that they have been replaced
- Confirm that all isolation has ended on return to service
- Telemetry is used to confirm the failure dates in the service orders, and to determine the return to service date of the LBD when this is after the return to service of the parent inverter.

This case study illustrates the limitations of unstructured text data in establishing an asset's history.

(a) (b)



(a) Failed component category identified in service order.
(b) Fraction of service orders which are / are not complete.

Discussion

In almost three quarters of service orders the cause of the ground fault is identified as one or more failed components. However, in over three quarters of service orders there is not a clear indication that the failed components were replaced, and the system fully returned to service.

Telemetry confirms that in these cases a combiner has been left isolated as the initial containment action, for a median of 73 days.

Case Study 2: All inverter fault service orders for March 2026 where failed equipment was replaced.

Layering of internal service orders with OEM case management tickets, and telemetry data. For each internal service order, corresponding OEM case management tickets have been collected where they exist and start and end dates of downtimes have been extracted from telemetry.

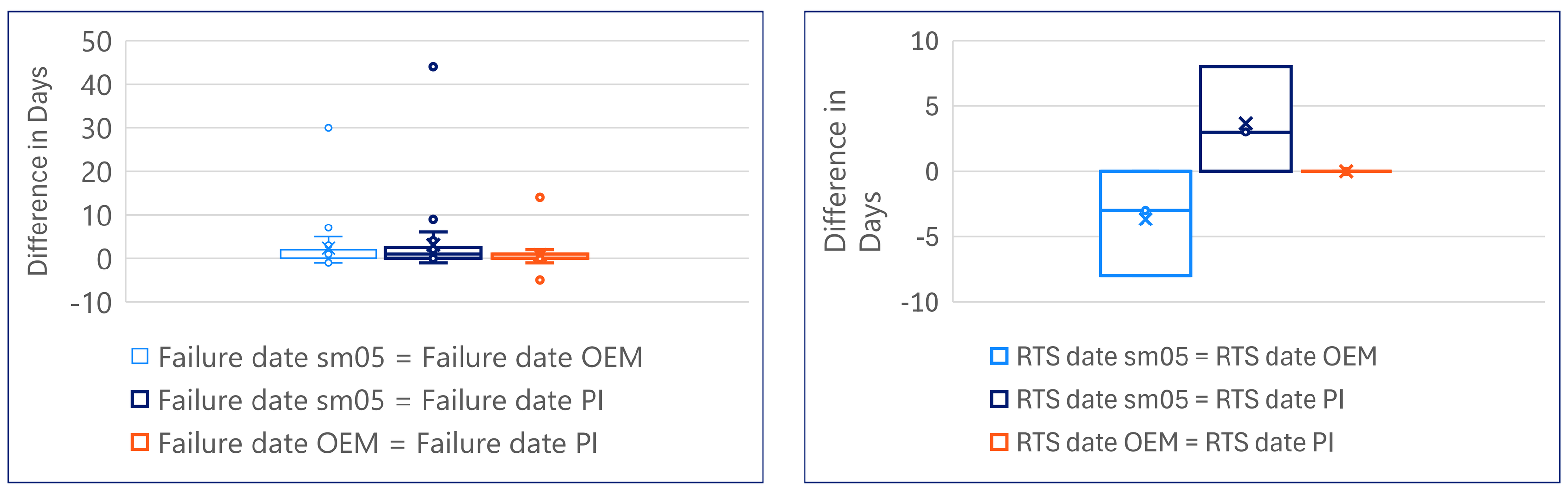
These three layers have been manually blended to establish how well they agree, and how complete each layer is individually, as well as the blended data as regards the data types required to calculate a reliability model. Conflicts between layers are discussed qualitatively.

The table below shows the data needed to conduct lifetime studies on components of PV systems, and whether they are available from each of the three sources considered.

Data Required	Internal Service Orders	OEM Case Tickets	Telemetry
Component's Initial Operating Date [Project COD is Available]	NO	NO	NO
Failure Date	YES	YES	YES
RETURN TO SERVICE Date	YES	YES	YES
Failed Component	YES	YES	MAYBE [If fault codes are unambiguous]
Failed Component Count	SOMETIMES	YES	NO
Type of Repair [Presumed to be with new, like-for-like]	MAYBE [Unless refurbished parts or compatible but not like-for-like parts are specifically identified]	NO	NO

Discussion

Boxplots of the differences in Failure and Return to Service Dates between internal service orders, OEM case tickets, and Telemetry shown below (positive values means internal dates are later) suggest that the internal service order data is prone to inaccuracy, given that there is a higher degree of agreement between OEM cases and Telemetry.



One difference between the internal and OEM records is qualitative accuracy of the component failure details. Internal service orders typically only imply counts based on singular or plural components which OEM records generally precise the quantity and use part numbers which remove ambiguity as to what components have been replaced.

Inventory management while under OEM warranty is a difficult area to manage since parts may be sourced by the OEM or the customer and may be drawn from construction spares. Unifying these records into a single source of truth can be challenging but is essential for lifetime studies since the source of the parts is of limited salience.

Conclusions

While no source of data is perfect, layering multiple sources can lead to better overall datasets.

Unless datasets are designed to be interoperable (e.g. does an internal service order have a field for capturing an associated OEM ticket number?), this is a manual process which lacks scalability and leads to underutilization of data.

There is a balance to be struck between lengthy data intake forms with many fields to complete, and a light touch form that is reliant on free text. In the case of the latter, setting expectations, training staff as to the necessary information to include, and diligent management confirming information is complete before closing service orders are critical. A particularly important gap identified in the course of preparing this work has been a lack of consistency in recording as found / as left conditions.

There are technology solutions to this, based on maximizing the use of cameras and of speech recording and analysis to ensure all necessary information is gathered each time in the field. Having three possible sources of data allows for cross-validation wherein agreement between two of three sources can enable error correction and increases the probability that at least one source will contain essential information.

EDF power solutions