RACE for 2030 NETWORKS

RACE for Networks

# Maximising Solar ROI: Advanced Diagnostics for PV Systems

Progress Report

## Progress report

RACE for Networks

**Maximising Solar ROI: Advanced Diagnostics for PV Systems**

Project Code: 24.NT6.R.0758

3 September 2025

## Citations

Huang, H., Ibrahim, I., Severiano, B., Rispler, J., Roche, D., Martell, C., Duran, E., Meers, C., Yildiz, B., Bruce, A., Rougieux, F. (2025). Maximising Solar ROI: Advanced Diagnostics for PV Systems. Prepared for RACE for 2030.

## Project partners

Global Sustainable Energy Solutions

Green Peak Energy

University of New South Wales

University of Technology Sydney

## Project team


University of New South Wales

- Fiacre Rougieux
- Hanrong Huang
- Baran Yildiz
- Anna Bruce

University of Technology Sydney

- Ibrahim Ibrahim
- Bernardo Mendonca Severiano
- Jonathan Rispler
- David Roche

Global Sustainable Energy Solutions

- Chris Martell
- Earl Duran
- Clement Meers


## What is RACE for 2030?

Reliable, Affordable Clean Energy for 2030 (RACE for 2030) is an innovative cooperative research centre for energy and carbon transition. We were funded with $68.5 million of Commonwealth funds and commitments of $280 million of cash and in-kind contributions from our partners. Our aim is to deliver $3.8 billion of cumulative energy productivity benefits and 20 megatons of cumulative carbon emission savings by 2030. racefor2030.com.au

# Contents

# 1 Summary

Our progress report describes achievements across three core research areas for our project Maximising Solar ROI: Advanced Diagnostics for PV Systems. The project team has made substantial progress in automated shading detection, PV orientation optimisation, and distributed event correlation, directly supporting enhanced diagnostic capabilities and improved system characterisation.

The research team developed a two-level shading detection workflow within WP1 that reduces false positives by requiring both persistent temporal patterns and physics-consistent evidence. This dual tier approach combines AC level system behaviour analysis with DC level electrical signatures to confirm shading through bypass diode activation. Additionally, a shading aware re-baseline system was implemented to create realistic performance benchmarks for sites with confirmed, persistent shading by using a rolling window of recent clear sky days.

To address gaps or inaccuracies in installation metadata, WP2 established an automated method to estimate as-installed tilt and azimuth by fitting measured DC generation to theoretical PVLib models. This optimisation process uses Differential Evolution to simultaneously adjust orientation parameters and system loss factors, ensuring realistic performance expectations based on actual solar geometry rather than simplified assumptions.

Within WP4, a distributed framework was established using inverter fleets as a peer-to-peer sensing network to detect and characterise grid-side disturbances. Analysis of a major 13 February 2024 Victoria transmission line event revealed that system size is the dominant risk factor for tripping, with 39% of systems over 30 kVA experiencing unplanned disconnections compared to only one system at or below 30 kVA.

The second Industry Reference Group workshop confirmed strong stakeholder interest in these fleet-scale diagnostics and peer-to-peer event correlation. While the team encountered challenges regarding cross-vendor fault-code retrieval and data resolution, effective solutions including physics-based validation and site-adaptive thresholds have been implemented to ensure robust performance across diverse climate zones. These achievements position the project to deliver actionable outputs for OEMs and service providers while enhancing real-time grid observability.

# 2 Project Progress

## 2.1 Update on research activities

This quarter, the project progressed across three core research areas: automated shading detection and shading-aware re-baselining (WP1), PV tilt and azimuth optimisation (WP2), and distributed event correlation using peer-to-peer intelligence (WP4). These activities strengthen diagnostic capability and improve system characterisation across diverse PV installation configurations.

## 2.2 Preliminary findings

### Shading Detection and Re-baseline System (WP1)

We developed a two-level shading detection workflow that analyses (i) whole-system behaviour and (ii) per-MPPT electrical signatures. The approach is designed to reduce false positives by requiring both persistent temporal patterns and physics-consistent evidence of shading.

Tier 1: AC Level Detection

The Tier 1 analyser identifies system-wide problems by looking for repeated underperformance patterns on clear-sky days. Figure 2.1 illustrates the analysis workflow.

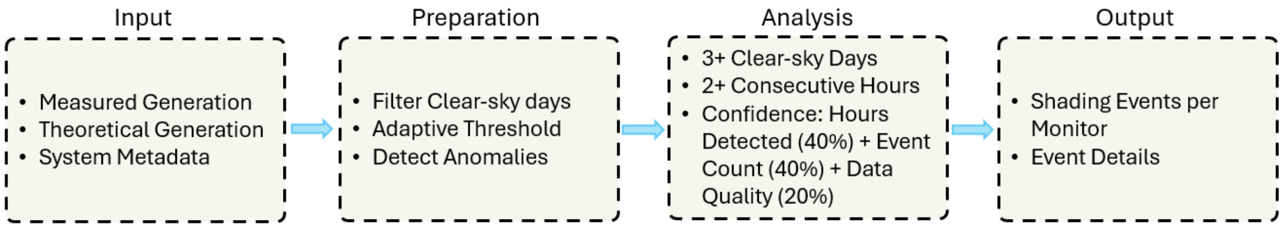


**Figure 2.1**: AC level shading detection flowchart.

A key innovation is the dynamic threshold that adapts to each site's characteristics. Sites that naturally underperform due to their specific setup or location have more relaxed thresholds, preventing the system from incorrectly flagging them as problematic.

The algorithm requires patterns to be consistent within individual days and persistent across multiple days before flagging an issue. This successfully filters out transient weather effects and single-day anomalies. The system also detects inverter tripping through sudden generation drops and power clipping during peak periods. See Appendix Figure A.1 for AC level shading detection example.

### Tier 2: DC Level Detection

The Tier 2 analyser examines voltage and current behaviour at each MPPT to identify shading through bypass diode activation. Figure 2.2 shows the DC analysis methodology.

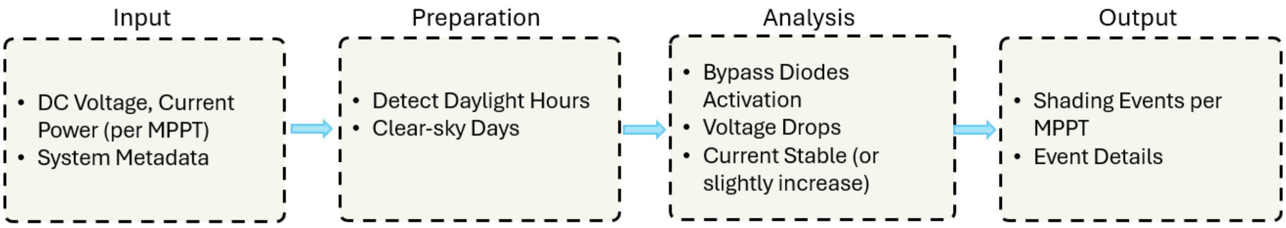


**Figure 2.2**: DC level shading detection flowchart.

When shading activates bypass diodes, voltage at the shaded substring drops significantly while current remains relatively stable. This physics-based approach identifies which specific MPPTs are affected and provides detailed event diagnostics. Refer to Appendix Figure A.2 for example detection cases and MPPT-level analysis.

### Shading-Aware Re-baseline

For sites with confirmed shading patterns, we implemented a re-baseline system that creates realistic performance benchmarks accounting for known shading. The system works by building adjustment patterns from recent clear-sky days, using a rolling window approach that captures how shading affects generation throughout the day.

The process scales actual generation patterns from clear-sky days to align with theoretical peaks, then smooths these patterns across a window of recent clear-sky days with outlier filtering to ensure reliability. These patterns are then applied chronologically: clear-sky days use their rolling window pattern whilst cloudy days reuse the most recent clear-sky pattern. This creates baselines that are less optimistic than pure theoretical calculations for sites with persistent shading. See Appendix Figure A.3 for shading re-baseline example.

### PV Tilt & Azimuth Optimisation System (WP2)

We developed an automated method to estimate as-installed tilt and azimuth by fitting measured DC generation to physics-based theoretical generation. This addresses common gaps or inaccuracies in installation metadata that can materially affect diagnostics and fault attribution.

The workflow identifies clear-sky days and estimates DC capacity for each MPPT through peak power analysis, providing the foundation for accurate theoretical generation calculations.

The optimisation process uses Differential Evolution to simultaneously adjust tilt angle, azimuth angle, and system loss factor. The algorithm searches for the combination that maximises correlation between measured DC power and theoretical generation calculated using PVLib's physics-based models. PVLib calculates solar position, clear-sky irradiance, and plane-of-array irradiance based on actual solar geometry and atmospheric conditions, ensuring realistic performance expectations rather than simplified assumptions.

The optimiser iteratively refines parameter estimates until convergence, producing comprehensive performance metrics and visualisations that show how well the optimised parameters match actual generation patterns. The system processes both individual sites and batch operations, automatically generating results files and diagnostic plots for validation. See Appendix Figure A.4 for optimisation results and correlation analysis.

### Distributed Event Correlation via Peer-to-Peer Intelligence (WP4)

We developed a distributed event correlation framework that uses inverter fleets as a peer-to-peer sensing network to detect and characterise grid-side disturbances. As presented in Figure 2.3, this methodology combines power-based event detection using 5-minute aggregated data with inverter fault logs to improve classification accuracy and causal insight. The framework was applied to a detailed showcase event in Victoria (Sydenham-Moorabool transmission line disturbance on 13 February 2024), enabling both fleet-level statistical analysis and device-level forensic investigation.

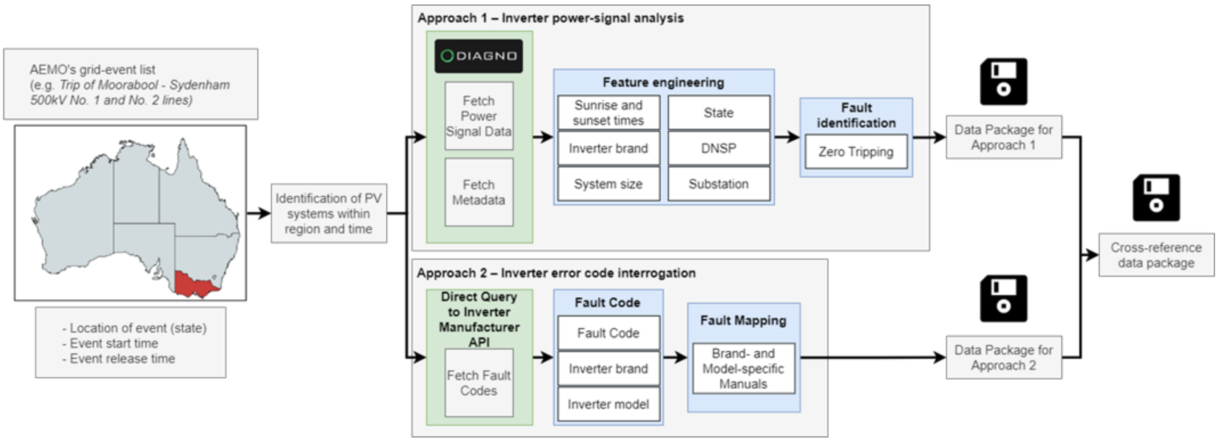


**Figure 2.3**: Schematic overview of the analysis pipeline for the inverter response to a grid event.

- **Power-Based Event Detection (Approach 1):** The first detection layer identifies potential tripping events by analysing sudden collapses in AC output toward zero within 5-minute intervals. This approach provides a scalable mechanism for identifying grid-related disturbances across large, distributed PV fleets.
- **Fault-Code-Based Detection (Approach 2):** The second detection layer analyses inverter-reported grid disturbance codes to directly identify voltage- and frequency-driven events. Fault log analysis confirmed a voltage-driven tripping mechanism where usable timestamped logs were available. Some devices recorded near-instant "Grid transient over-voltage" events followed by "Grid power outage" entries, aligning precisely with the 5-minute power-signal trip flags. Refer to Appendix


Figure A.5


Figure A.5 for an example of a 5-minute power signal fault-inference overlapped with a fault-code retrieval.

By quantifying the overlap between power-based and fault-code-based detection, we observed strong alignment for major tripping events. Logistic regression analysis of the showcase event revealed that system size is the dominant risk factor. During the VIC disturbance: 39% of systems >30 kVA tripped, with only 1 out of 60 systems ≤30 kVA tripping. Distance from the fault location and DNSP membership were not statistically explanatory once system size was controlled for. Large systems tripped both near and far from the disturbance, indicating that size-related protection behaviour is more predictive than geographic proximity. Combining 5-minute power-based detection with fault logs (Appendix Figure A.6) improves both event classification confidence and causal interpretation. While major trips are clearly visible in power data, fault logs frequently capture short-duration voltage excursions that may not fully suppress output within a 5-minute interval.

Ultimately, spatial clustering of similar fault codes across multiple inverters within narrow time windows demonstrated that inverter fleets can function as distributed sensors. This peer-level intelligence approach enhances situational awareness beyond individual site diagnostics and highlights the potential role of inverter fleets as real-time grid observability assets.

## 2.3 Issues and difficulties encountered, if any, during the course of the research activities

### Shading Detection Algorithm Challenges (WP1)

Developing a robust shading detection algorithm was challenging due to the diverse range of solar installations across the monitored fleet. Variations in panel orientations, tilt angles, inverter sizing ratios, and local environmental conditions meant the algorithm needed flexible parameters while still operating reliably in automated mode.

Distinguishing genuine shading from other system faults presented another complexity. Inverter clipping, system tripping, and equipment malfunctions can produce power reduction patterns that look similar to shading. The algorithm required careful tuning to differentiate between these scenarios by examining the physics of how power drops occur and analysing when patterns repeat over time.

Seasonal variations in sun path and shading patterns added further complexity, particularly across Australia's diverse climate zones. Northern tropical regions experience different shading behaviour compared to southern temperate areas, requiring adaptive approaches rather than fixed seasonal templates.

### Tilt and Azimuth Optimisation Bias Factors (WP2)

The optimisation process revealed systematic bias sources that can affect orientation parameter accuracy. Shading represents the largest source of error, as losses during morning or afternoon periods can be mistaken for orientation differences. This is particularly problematic for sites with one-sided shading patterns that look like azimuth variations.

Inverter clipping introduces bias by flattening peak generation periods, predominantly affecting midday high-power data. This can make the optimisation estimate tilt and azimuth parameters that appear less optimal than the actual installation, as the algorithm tries to fit theoretical curves to artificially capped measured data.

System tripping events cause irregular power drops unrelated to solar position. While tripping mainly degrades overall fit quality rather than systematically biasing orientation estimates, sites where trips consistently occur at specific times of day can introduce directional bias into the results.

## 2.4 Actions proposed or undertaken to overcome the issues difficulties

### Physics-Based Validation Approach (WP1)

Rather than relying on seasonal patterns that varied across climate zones, we shifted to physics-based system health validation. The algorithm now focuses on detecting system malfunctions through characteristic signatures: clipping through sustained flatness during peak periods, tripping through sudden complete generation loss, and shading through V-I decoupling at the DC level. This approach provides consistent performance regardless of geographic location.

The two-tier structure proved essential for distinguishing between fault types. AC-level analysis identifies whole-system issues and temporal patterns, while DC-level analysis confirms shading through bypass diode physics. This dual validation significantly reduces false positives by requiring both consistent temporal patterns and physical evidence of shading mechanisms.

### Dynamic Thresholding and Temporal Persistence (WP1)

We implemented site-adaptive thresholds that adjust based on each installation's baseline performance characteristics. This prevents incorrectly flagging systems that naturally underperform due to site-specific factors. The algorithm also requires multi-day persistence with strict within-day consistency before flagging issues, effectively filtering out transient anomalies and single-event outliers.

### Optimisation Quality Controls (WP2)

For the tilt and azimuth optimisation, we incorporated data quality screening to identify and exclude periods affected by clipping, tripping, or significant shading. The clear-sky day selection process provides the first layer of filtering; we also added additional filters to improve data quality.

We validate optimisation results against external references where available, such as Nearmap satellite images, to confirm parameter estimates align with physical installation characteristics. Sites showing poor fit quality or parameters inconsistent with expected ranges are flagged for manual review to investigate root causes.

### Data-Specific Challenges (WP4)

The interpretation of WP4 results is subject to several data-related constraints. The 5-minute temporal resolution can miss short-duration trips or round event onset, and the power-based classifier may under-detect partial, string-level, or masked trips.

Metadata gaps (e.g., AC ratings, IPD settings, commissioning standards) required inference from commissioning reports, introducing uncertainty. Spatial precision is limited by rounded coordinates, affecting distance and DNSP attribution. Aggregated multi-inverter sites limit intra-site behavioural resolution, while time zone and daylight-saving inconsistencies may leave minor residual timestamp misalignments.

Fault-code retrieval presented the main challenges. Only one brand, out of five analysed, returned complete, timestamped logs suitable for event-level alignment. Other OEMs were constrained by shallow historical retention (sometimes limited to less than 3 months retention), hourly rate limits, per-device authentication requirements, and varying timestamp integrity. Moreover, fault-code semantics differ across vendors and sometimes even inverter models. Intermittent empty payloads, and the absence of push-based mechanisms (e.g., webhooks) further reduced reliability. These limitations highlight the need for a harmonised cross-vendor logging framework to enable robust fleet-wide event correlation.

# 3 Outcome of second IRG and stakeholder workshops

The IRG showed strong stakeholder interest in the project's fleet-scale diagnostics. The team presented progress on fleet-scale diagnostics, including peer-to-peer event correlation for grid disturbances and evaluation of global irradiance data sources to improve theoretical generation baselines used in fault detection.

An important discussion point was the need to better distinguish inverter trips from central protection operation. IRG members noted that power-signal analysis alone cannot reliably separate inverter behaviour from central protection actions (including Secondary Mains Protection), and highlighted that SMP configuration and settings errors have been observed in practice. The group agreed that follow-on work should incorporate SMP make and applied settings where accessible, to strengthen root-cause discrimination and ensure recommended actions are appropriately targeted.

IRG members also requested clarification on WP5, which assessed global-coverage irradiance options to support theoretical generation baselines used in fault detection. Open reanalysis products (ERA5/MERRA-2) performed well in Australian benchmarking, with trade-offs in resolution and data latency; potential mitigation options (including complementary near-real-time sources and ensemble approaches) were discussed. Shading detection and re-baselining progress was presented, and stakeholders recommended clearly separating soiling from shading in reporting. Additional needs raised included improved visibility of soiling, improved integration of key site metadata (e.g., orientation/tilt where available), outputs that are actionable for OEMs and service providers, and ongoing attention to cybersecurity risks as connected diagnostics mature.

# APPENDIX

# A. Tables and figures

## 1. Tables

**N/A**

## 2. Figures

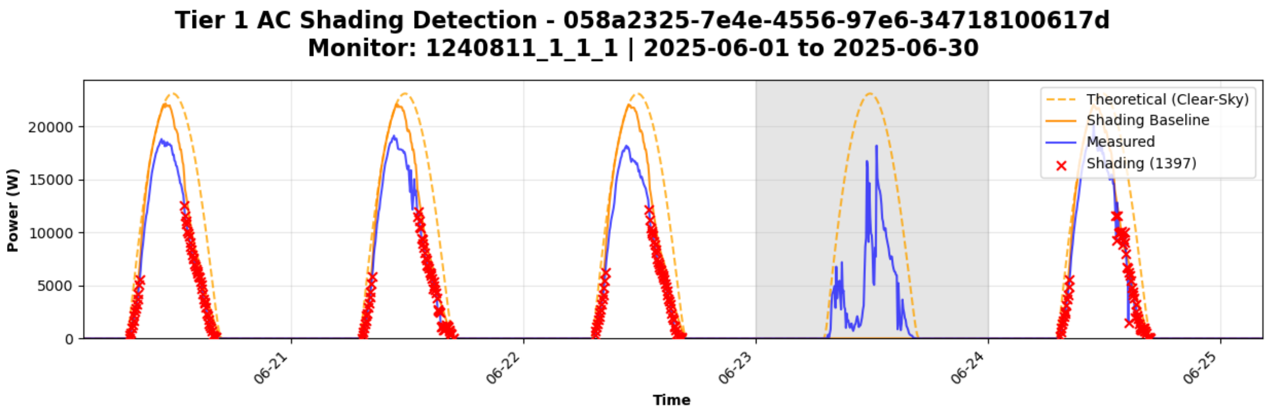


Figure A.1: AC-level shading detection example. Red markers show detected early morning and afternoon shading events across multiple clear-sky days.

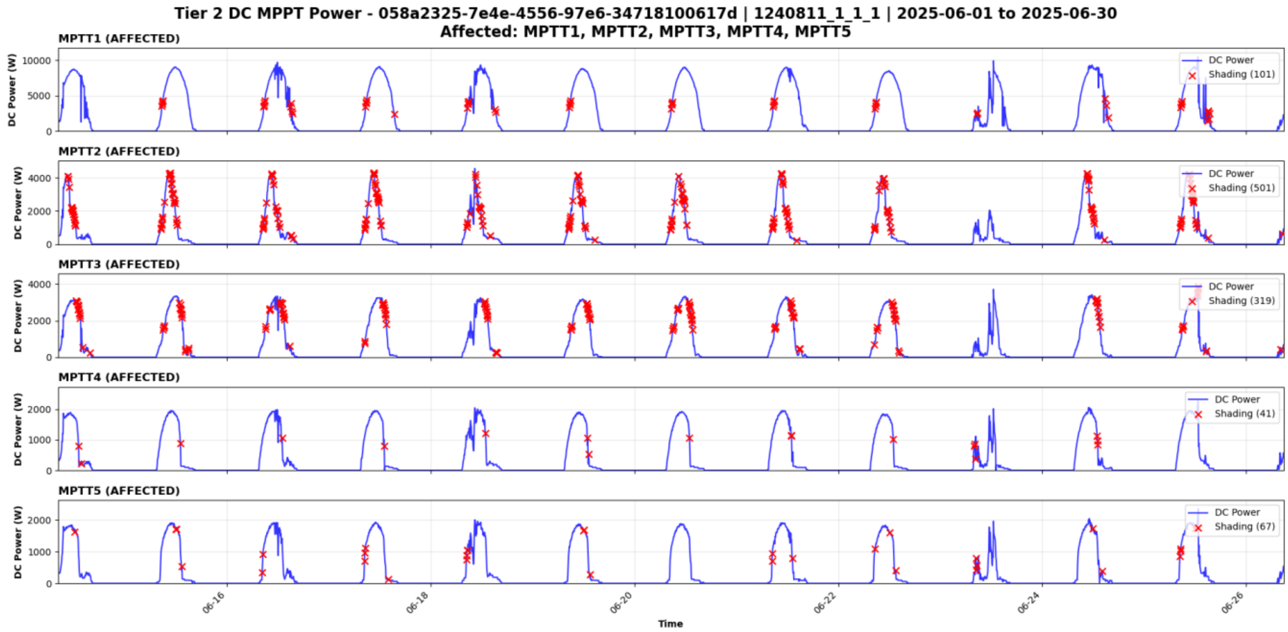


Figure A.2: DC-level shading detection at MPPT level. Red markers show detected shading events, identifying which specific MPPTs are affected.

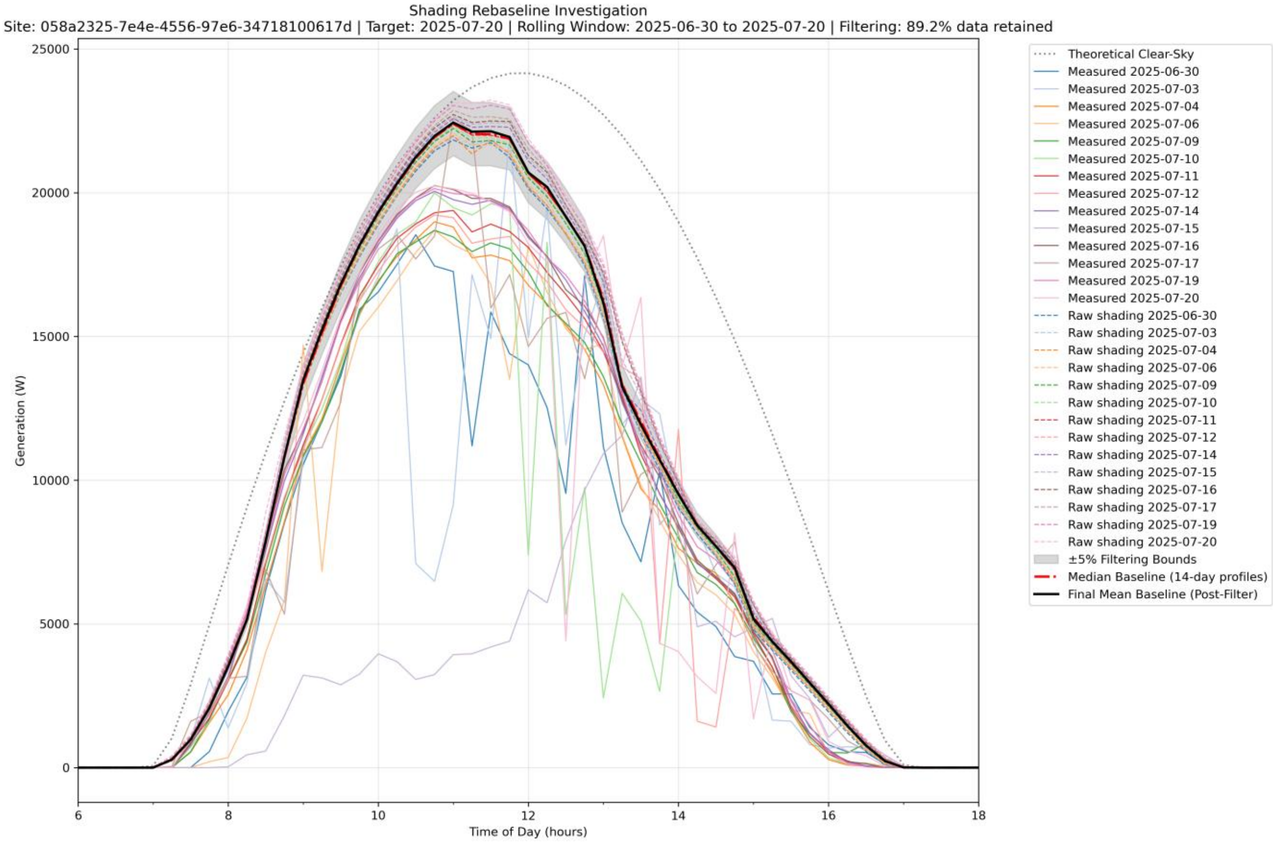


**Figure A.3:** Shading re-baseline example using 14-day rolling window of clear-sky days. Grey zone shows ±5% filtering bounds around the median; outliers are removed before calculating the final baseline.

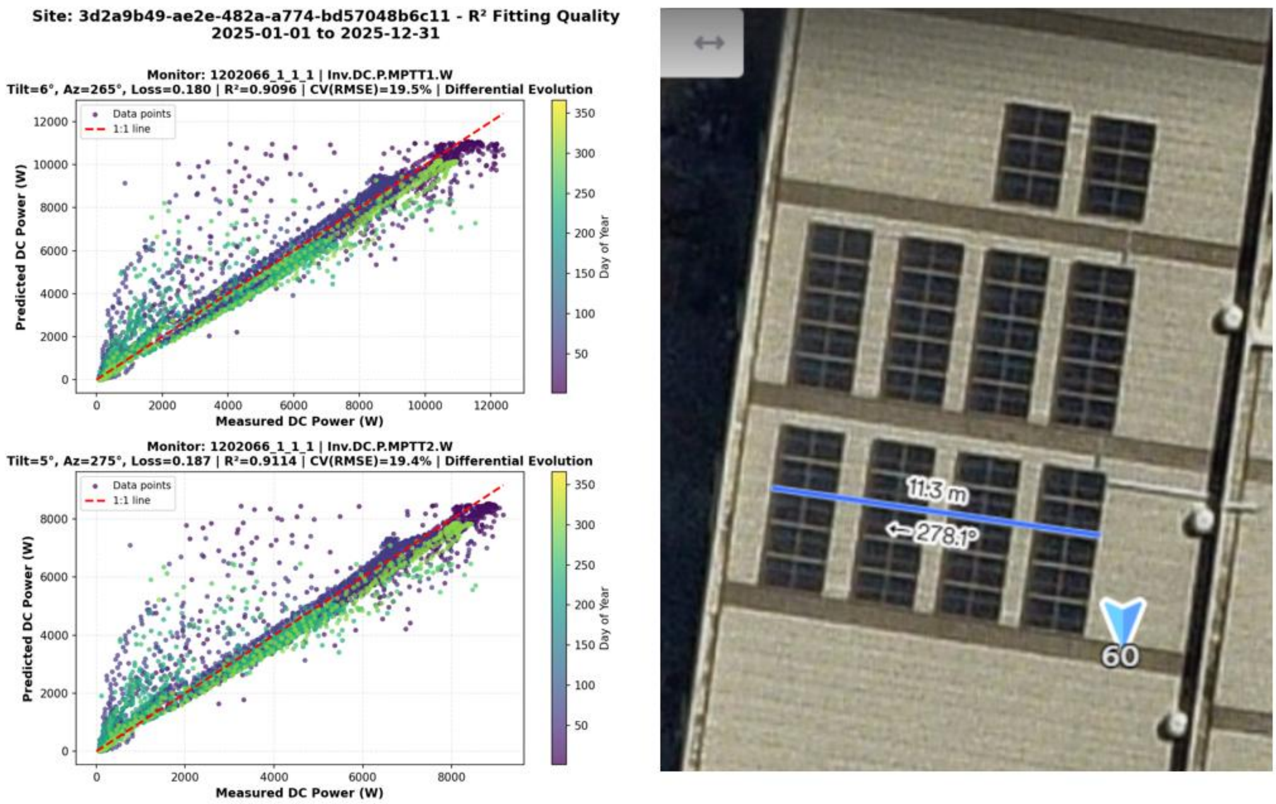


**Figure A.4:** Tilt and azimuth optimisation results validated against Nearmap image. Scatter plots show strong correlation between measured and predicted DC power for two MPPTs. Optimised azimuth values closely match the actual panel orientation visible in the satellite image (right).


**Figure A.5:** Dual-method event detection using 5-minute power data and inverter fault logs. The sudden drop in AC generation coincides with a grid undervoltage fault.

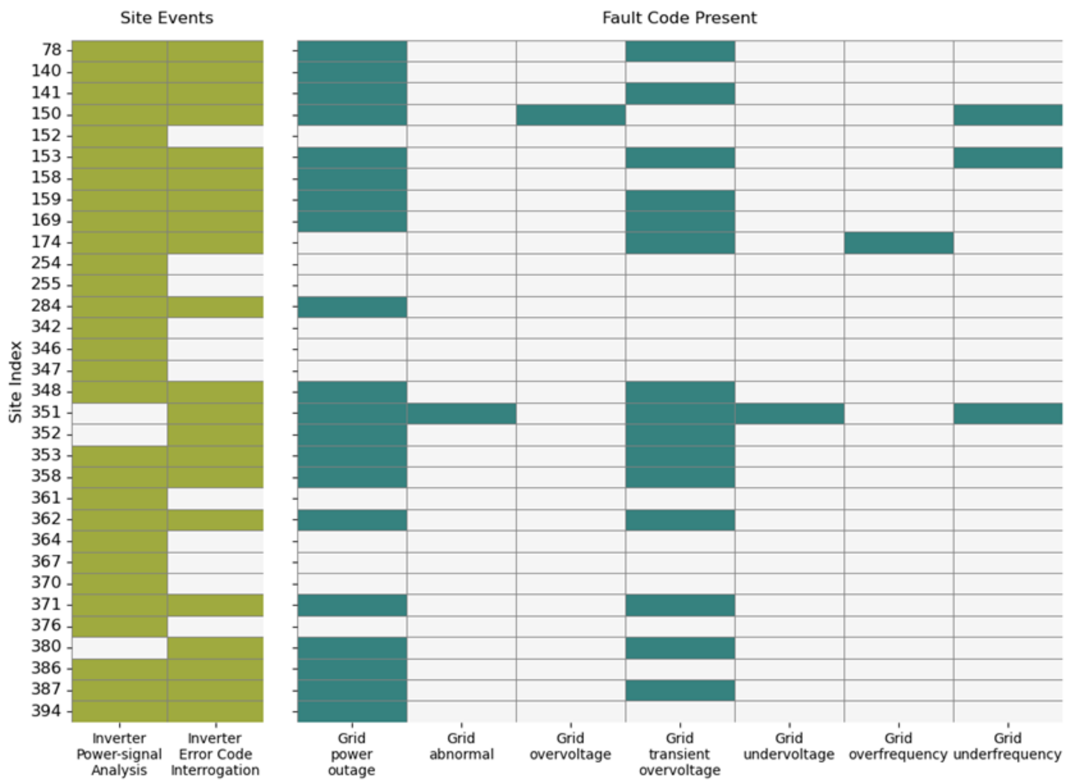


**Figure A.6:** Combined event matrix for 32 monitored sites. (Left: binary presence of inverter power-signal analysis (approach 1) and of any error code interrogation (approach 2); and Right: fault-category heatmap showing which sites recorded error logs in the 2 timestamps window following the event time). Of the 52 sites with available data, 20 did not exhibit either a generation drop to zero or a relevant fault log during the event window and are therefore not shown.

RACE for 2030

www.racefor2030.com.au

Australian Government
Department of Industry, Science and Resources

AusIndustry
Cooperative Research Centres Program