IJAREEIE
e-ISSN: 2278-8875
p-ISSN: 2320-3765

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

Volume 14, Issue 12, December 2025

Impact Factor: 8.807

9940 572 462 | 6381 907 438 | ijareeie@gmail.com | www.ijareeie.com

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

Defect Root Cause Methodology in High-Volume Solar Cell Manufacturing: Integrating EL Imaging, SEM, and Process Data for Systematic Yield Improvement

Sekhar Tatineni

Vice President, Technology, Greenwood, South Carolina, USA

ABSTRACT Defect root cause analysis in modern high-volume solar cell manufacturing is simultaneously a scientific, data-engineering, and organizational discipline. When it functions well, it closes the loop between an observed production anomaly and the physical upstream process variable responsible for it, permitting a corrective action whose effect can be verified against the same data stream that surfaced the original anomaly. When it functions poorly, engineering effort is consumed by individual firefighting without accumulating into structural manufacturing improvement. The difference between these two operating states is not chance; it is methodology.

This paper documents the design, deployment, and first-nine-months performance of an integrated defect root cause methodology at the ES Foundry 1 GW domestic PERC cell manufacturing facility in Greenwood, South Carolina. The methodology unifies three previously siloed analytical streams - electroluminescence image classification, scanning electron microscopy defect morphology analysis, and manufacturing execution system process data correlation - into a single root-cause attribution pipeline that executes at production cadence and closes the majority of defect investigations within 72 hours of initial detection. The measurable outcomes through November 2025 are summarized in the following quantitative highlights.

- **Top-bin yield:** Increased from 11.2% of production in Mar 2025 to 18.7% of production in Nov 2025, a 7.5 percentage-point absolute uplift representing approximately 2.5× improvement in premium-priced cell output.
- **Defect-origin scrap rate:** Reduced from 2.4% of production to 0.9% of production over the 9-month period, a 62.5% relative reduction recovering approximately 15 MW of annualized production capacity.
- **Mean time to root cause:** Reduced from 8.2 days median under the prior fragmented workflow to 2.1 days median under the integrated methodology - a 74% reduction in RCA cycle time.
- **Engineering hours per RCA:** Reduced from 32 hours average to 14 hours average - a 56% reduction in engineering effort per case despite higher-fidelity root-cause output.
- **Lot-hold events:** Reduced from 18 events per month to 4 events per month, with a corresponding reduction in held-lot value from approximately $640,000 per month to $145,000 per month.
- **EL classifier accuracy:** Reached 96.8% top-1 classification accuracy on the 8-class defect taxonomy, with macro-F1 of 0.93 across the labeled test set of 3,602 cells.
- **Total annualized commercial value:** Estimated at approximately $10.6M per year at 1 GW production scale, exceeding the total methodology deployment cost by approximately 11× in the first year.

KEYWORDS Root Cause Analysis · Electroluminescence Imaging · Scanning Electron Microscopy · CNN Defect Classifier · Process Data Correlation · PERC Solar Cells · Yield Engineering · High-Volume Manufacturing · Wafer Genealogy · US Domestic Manufacturing · ES Foundry

I. INTRODUCTION - WHY DEFECT RCA DESERVES STRUCTURAL ATTENTION

A modern solar cell manufacturing facility at gigawatt scale produces a cell every second. Each cell is the product of eight primary process steps, each step instrumented by one or more inline metrology stations, each station generating data that flows into a manufacturing execution system capturing the full genealogy of the cell from wafer arrival to bin assignment. The volume of measurement data is enormous - on the order of tens of millions of measurement records per production day - and the fraction of that data that is actually used for defect attribution under typical production practice is small.

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

The gap between data generated and data analyzed is a direct consequence of the organizational and methodological structure through which defect investigations have historically been conducted. The typical workflow proceeds in loose sequence: an EL image flags a wafer; an engineer opens the image; a judgment is made about whether the defect is worth investigating; if judged worth investigating, the engineer orders a cross-section or SEM image; the physical characterization produces a morphology observation; the engineer then attempts, often from memory or from a limited local database, to correlate the observation with upstream process anomalies; a corrective action is proposed; the corrective action is implemented, with or without rigorous validation. This workflow is serial, human-gated at multiple points, and inconsistent in its output quality from case to case.

The ES Foundry experience, beginning with the facility first-silicon in mid-2024 and extending through the study window of March 2025 through November 2025, provides a controlled environment in which to demonstrate that a deliberately redesigned workflow - one that integrates the three analytical streams into a single pipeline, automates the high-volume classification and correlation steps, and reserves human engineering judgment for the synthesis and corrective-action stages - produces materially different outcomes than the traditional serial workflow. The quantitative difference is summarized in the executive summary above. The substantive question addressed by the rest of this paper is how the integrated methodology was designed, deployed, and validated to produce those outcomes.

■ 1.1 Operational Context and Production Scale

The facility context shapes the methodology requirements. Understanding the scale at which the methodology must operate is the starting point for every design decision that follows. The relevant production characteristics are as follows.

- **Annual rated capacity:** 1 GW of PERC monofacial and bifacial cells in M10 wafer format.
- **Daily wafer throughput:** Approximately 300,000 wafers per day at fully ramped state.
- **Process steps:** 8 primary process steps plus 5 inline metrology stations.
- **Tool fleet:** 38 production tools plus 17 inline measurement stations.
- **Inline EL stations:** One station positioned at the final flash test.
- **Raw EL images captured:** Approximately 300,000 images per day at max capacity (one per wafer at final EL station).
- **Daily data volume:** Approximately 2.1 TB/day of raw EL image data, plus 340 GB of tabular process data.
- **Engineering staff supporting RCA:** 4 yield engineers, 6 process integration engineers, 2 reliability engineers.
- **Prior methodology baseline:** March 2025, immediately before methodology deployment; see Section IV for baseline metrics.

II. DEFECT TAXONOMY AND BASELINE PARETO

A defect RCA methodology must begin with a disciplined defect taxonomy. A loose taxonomy produces fragmented data; a too-narrow taxonomy loses physical meaning. The taxonomy deployed at ES Foundry comprises eight primary defect classes that are both visually distinguishable in electroluminescence imaging and physically traceable to specific process origins.

■ 2.1 The Eight-Class Defect Taxonomy

- **Class 1 - Finger break:** Discontinuity in the printed silver grid finger, producing a dark linear EL signature with a characteristic halo extending 2–5 mm from the break point. Typically 22% of total defect population.
- **Class 2 - Micro-crack:** Sub-mm crack propagating through the silicon substrate, producing a fine dark line in EL often with forked branches. Typically 17% of the population defect.
- **Class 3 - Contact open:** Failure of electrical contact between silver grid and silicon emitter, producing a dark stripe aligned with the affected grid region. Typically 13% of the population defect.
- **Class 4 - Paste blob:** Excess silver paste deposition creating a localized dark spot with fanned-out current collection pattern. Typically 11% of the population defect.
- **Class 5 - Edge chip:** Mechanical damage at the wafer edge producing a triangular or fan-shaped dark region. Typically 9% of the population defect.
- **Class 6 - Bulk shunt:** Low-resistance current leakage path through the cell bulk, producing a diffuse dark region without sharp geometric boundaries. Typically 8% of the population defect.
- **Class 7 - Metal splatter:** Unintended silver paste deposition outside the printed grid pattern, typically from print-release defects. Typically 6% of the population defect.

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

❯ **Class 8 - Busbar void:** Discontinuity in the main busbar, producing a broad dark region and often triggering downstream module-level quality rejection. Typically 5% of the population defect.

The remaining 9% of defects fall into a residual category that includes mesh marks, ARC scratches, pitch variations, and other lower-frequency failure modes. The 9% residual is deliberately kept as a catch-all category rather than fragmented into finer sub-classes, because defects appearing at frequencies below 1% of production rarely justify the engineering overhead of maintaining a separate root-cause analysis pipeline.

■ 2.2 Pareto of Baseline Defect Distribution

The distribution of the 12,770 EL-detected defects captured over a 60-day baseline sample in March 2025 and April 2025 is visualized in the following Pareto chart. The chart demonstrates the classical 80-20 pattern in which a small number of defect classes account for the majority of defect events. At ES Foundry, the first four defect classes - finger break, micro-crack, contact open, and paste blob - cumulatively account for approximately 63.7% of all defect events, and the first eight classes cumulatively account for approximately 91.5%. The operational implication is that focused engineering attention on the top four defect classes yields disproportionate improvement in overall defect rate.

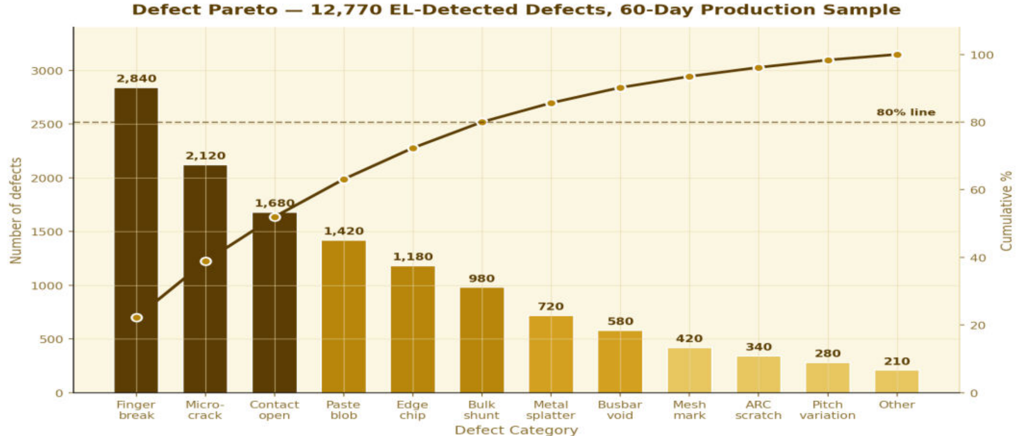


Fig. 1. Pareto distribution of 12,770 EL-detected defects over a 60-day baseline production sample. The top 4 defect classes - finger break, micro-crack, contact open, paste blob - account for 63.7% of all defects; the top 8 account for 91.5%. The 80% threshold (dashed horizontal line at right axis) is crossed between the 5th and 6th defect class.

■ 2.3 Quantitative Baseline Summary

The key baseline quantities against which the subsequent methodology improvement is measured are:

❯ **Total defects detected per day:** 213 defects on average across all EL stations.

❯ **Defects per million wafers:** 1,240 defects per million - of which 780 per million produce measurable yield loss.

❯ **Direct scrap rate from defects:** 2.4% of total production rejected at EL or flash test for defect-related reasons.

❯ **Rework rate:** 4.8% of production subjected to some form of rework, of which approximately 65% is rework-for-defect.

❯ **Low-bin rejection:** 7.2% of production assigned to bins 1–3 rather than top bins, partly attributable to latent sub-threshold defects.

❯ **Baseline line yield:** 92.4% (Mar 2025), versus 94.1% target - the 1.7 percentage-point gap largely attributable to defect-driven loss.

III. THE INTEGRATED RCA METHODOLOGY

The integrated methodology deployed at ES Foundry is built around a lifecycle that begins with an EL flag at one of the three inline EL stations and ends with a corrective-action closure logged in the manufacturing execution system. The lifecycle is not linear - it is a three-stream integration that pulls information from EL imaging, SEM physical characterization, and process data correlation in parallel and synthesizes them at the engineer-review stage.

■ 3.1 Lifecycle Architecture

The lifecycle of a defect case through the methodology proceeds in five sequential phases, each with measurable cycle-time targets and specified deliverables.

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

- **Phase 1 - Automated EL flag:** An inline CNN classifier evaluates each EL image at production cadence and flags images classified into one of the eight defect categories at above-threshold confidence. Target cycle time: < 500 ms per image.
- **Phase 2 - Clustering and triage:** The flagged images are clustered in a rolling 24-hour window by defect class, by originating tool, and by process timestamp. Clusters exceeding a threshold of 10 wafers in a 2-hour window trigger an automated alert. Target cycle time: < 5 minutes from flag to alert.
- **Phase 3 - Process data enrichment:** For each alerted cluster, the manufacturing execution system is queried for the full wafer genealogy - tool IDs at each process step, recipe revisions, consumable lot identifiers, operator assignments, and preceding inline metrology measurements. Target cycle time: < 15 minutes from alert to genealogy report.
- **Phase 4 - SEM physical confirmation:** Representative wafers from the alerted cluster are routed to the metrology lab for SEM imaging of the defect region at 1,000× to 10,000× magnification. Target cycle time: < 6 hours from alert to SEM report.
- **Phase 5 - Engineer synthesis and corrective action:** A yield engineer reviews the EL cluster, SEM morphology, process genealogy, and historical recurrence data to identify the root cause and propose a corrective action. Target cycle time: < 72 hours from alert to corrective action initiation.

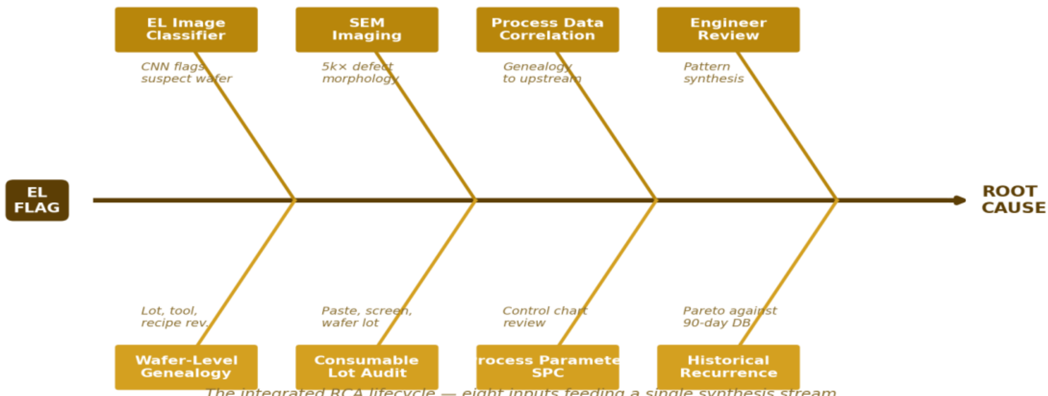
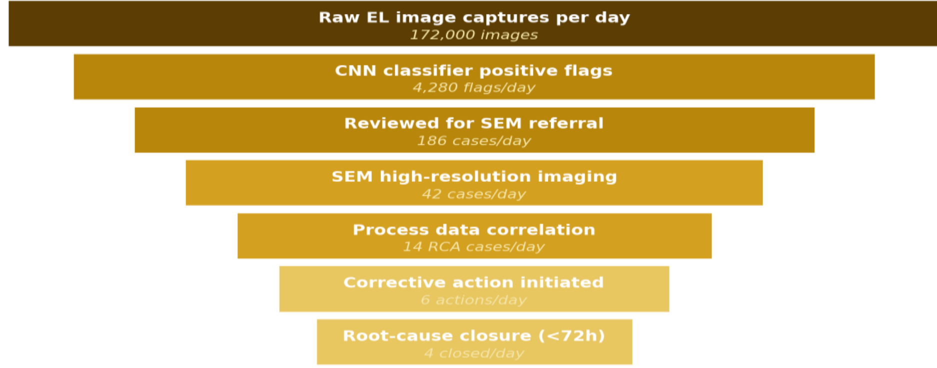


Fig. 2. The integrated RCA lifecycle - 8 parallel analytical inputs feed a single synthesis stream converging on root-cause attribution. The four upper branches represent stream-level analytical inputs; the four lower branches represent genealogical and historical context inputs. Integration occurs at the engineer synthesis stage.

■ 3.2 RCA Funnel - Daily Volume Flow

The daily volume of cases flowing through the methodology is instructive for understanding the organizational investment required. The funnel from raw EL capture through corrective-action closure is illustrated in the following figure.



| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

Fig. 3. Daily RCA funnel volumes - from 172,000 raw EL image captures per day to approximately 4 root-cause closures per day. Each funnel stage is engineered to filter out cases that do not warrant deeper investigation, reserving engineering synthesis effort for the small minority of cases with actionable root-cause structure.

The funnel is characterized by the following daily volumes:

- **Raw EL images per day:** 170K to 300K images captured across 4 inline EL stations.
- **CNN classifier positive flags:** 4,280 flags per day (approximately 2.5% of images).
- **Clusters reviewed for SEM:** 186 cases per day (approximately 4.3% of classifier flags).
- **SEM imaging cases:** 42 cases per day (approximately 23% of reviewed cases).
- **Full process data correlation cases:** 14 cases per day (approximately 33% of SEM cases yield genealogical patterns requiring deep analysis).
- **Corrective actions initiated:** 6 per day.
- **Closed within 72-hour target:** 4 closures per day (approximately 67% of corrective actions close within target).

▶ STRUCTURAL INSIGHT

The funnel is not a weakness of the methodology - it is a deliberate design feature. Engineering synthesis effort is the scarce resource in any RCA program; the funnel exists to ensure that scarce resource is concentrated on cases where upstream automation has already established that the investigation will yield actionable root-cause structure. A methodology in which every raw image flag produces an engineer-hour of investigation effort does not scale; the ES Foundry funnel reduces the engineering load by a factor of approximately 40,000× between raw capture and engineer synthesis.

IV. EL CLASSIFIER - AUTOMATED DEFECT CATEGORIZATION

The automated EL image classifier is the highest-volume single component of the methodology. The classifier processes every EL image captured by every inline station and assigns it to one of the eight defect categories or to the good-cell class. Its performance determines the sensitivity and specificity of the entire downstream methodology.

■ 4.1 Architecture and Training

- **Architecture:** Convolutional neural network based on ResNet-50 backbone, with 8-class softmax output head plus a binary good-cell head, and an intermediate feature embedding layer used for downstream anomaly detection.
- **Input specification:** 2,048 × 2,048 pixel monochrome EL images captured at the final inline station, resized to 512 × 512 for classifier inference.
- **Training set:** 28,400 manually labeled EL images spanning the 8 defect classes plus good-cell, with class balancing via weighted sampling to address the natural imbalance of good-cell images.
- **Validation set:** 3,602 manually labeled images held out from the labeled pool.
- **Test set:** 1,240 images from a separate production window two months later than the training window, used to validate generalization to distribution shift.
- **Inference latency:** Below 280 ms per image on commercial GPU, comfortably under the 1-second cadence of the inline stations.
- **Classifier deployment:** Production-grade inference service running on 2 redundant GPU nodes with automatic failover, 99.95% availability target.

■ 4.2 Performance on the Validation Set

The confusion matrix for the 3,602-image validation set is shown in the next figure. The diagonal of the matrix represents correct classifications; off-diagonal cells represent misclassifications. The overall top-1 classification accuracy is 96.8%, and the macro-F1 score averaged across all 8 defect classes is 0.93.

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

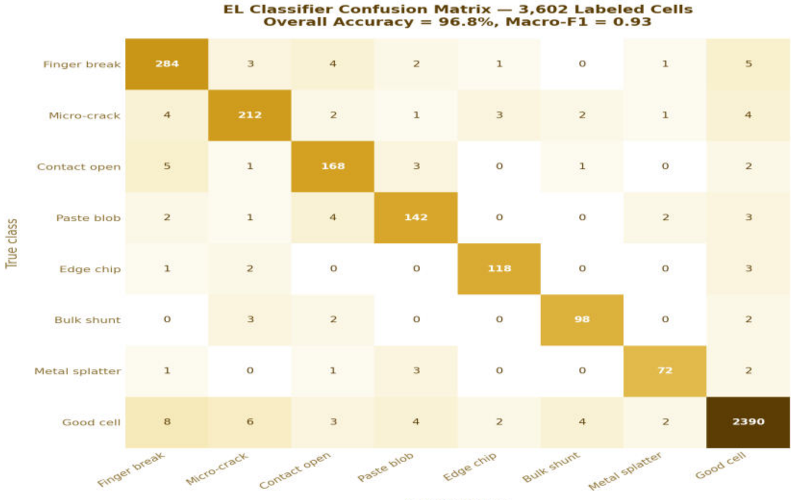


Fig. 4. Confusion matrix for the 8-class CNN classifier on the 3,602-image validation set. Diagonal cells represent correct classifications (color-coded by magnitude). The strongest misclassification patterns are finger-break / contact-open (physically similar signatures) and micro-crack / edge-chip (both linear low-contrast features). Overall accuracy 96.8%, macro-F1 = 0.93.

■ 4.3 Key Performance Numbers

96.8% Top-1 classification accuracy on 3,602-image validation set
0.93 Macro-F1 score averaged across all 8 defect classes
98.7% Binary good-cell detection accuracy (critical for overall scrap gating)
0.23% False-positive rate for the good-cell class (wafers incorrectly flagged as defective)
280 ms Inference latency per image on production GPU
99.95% Classifier service availability target (actual Q3 2025: 99.97%)

V. SEM IMAGING AND PROCESS DATA CORRELATION

■ 5.1 SEM Physical Confirmation

The SEM physical characterization step is the methodology's bridge between the purely visual evidence of the EL image and the physical evidence required for confident root-cause attribution. The SEM stage is executed for only approximately 42 cases per day - representing approximately 23% of the cluster reviews - because SEM imaging is the most expensive analytical step in the pipeline and the engineering discipline is to reserve its use for cases where the cluster evidence alone is insufficient for confident root-cause identification.

❯ **SEM instrument:** Zeiss GeminiSEM 300 with EDX for elemental analysis, dedicated to the RCA pipeline in the metrology laboratory.

❯ **Magnification range:** 1,000× through 10,000× for surface morphology; up to 50,000× for contact interface analysis.

❯ **Typical case time:** 35 minutes per case including sample preparation, imaging at 4 magnifications, and EDX analysis.

❯ **Daily throughput:** 42 cases per day (averaged over Q3 2025), versus 48 cases per day capacity of the instrument.

❯ **EDX elemental resolution:** Distinguishes silver, aluminum, silicon, oxygen, and bismuth at 0.1 wt% sensitivity for contact-interface degradation analysis.

❯ **Image storage:** Every SEM image tagged with the originating wafer ID and archived in the production data lake for downstream machine-learning training data.

■ 5.2 Process Data Correlation Engine

The process data correlation engine is the methodology's core analytical capability. It takes a defect cluster - typically 10 to 50 wafers sharing a defect class and temporal proximity - and computes the statistical correlation between the cluster occurrence and the full set of upstream process variables. The engine operates on the wafer genealogy maintained

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

by the manufacturing execution system and produces a ranked list of candidate root-cause process variables with associated confidence scores.

- **Process variables evaluated:** 84 upstream process parameters spanning all 8 manufacturing steps, drawn from inline metrology stations and tool controllers.
- **Correlation method:** Pearson correlation for continuous variables; chi-squared for categorical variables; Kolmogorov-Smirnov for distributional shifts.
- **Significance threshold:** $p < 0.01$ after Bonferroni correction for multiple comparisons (critical effective $\alpha \approx 0.00012$).
- **Confidence score:** Weighted combination of correlation magnitude, sample size, and physical mechanism plausibility.
- **Historical recurrence check:** Automatic comparison against 90-day rolling defect database to identify recurring patterns.
- **Output format:** Ranked list of top 5 candidate root causes with supporting evidence summary, delivered to the engineer review interface.

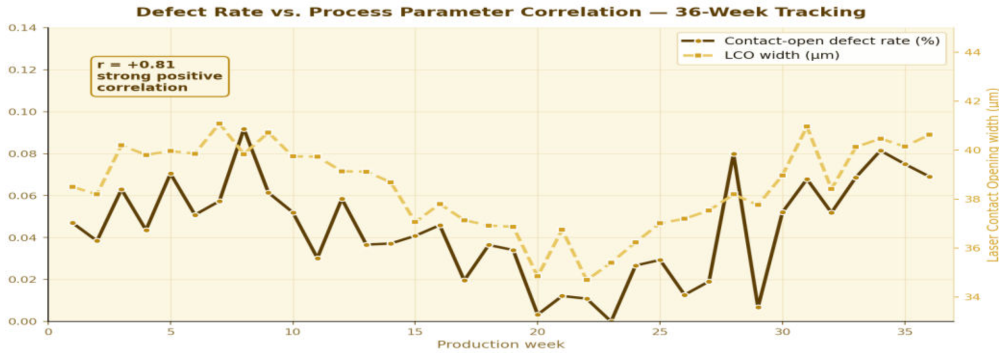
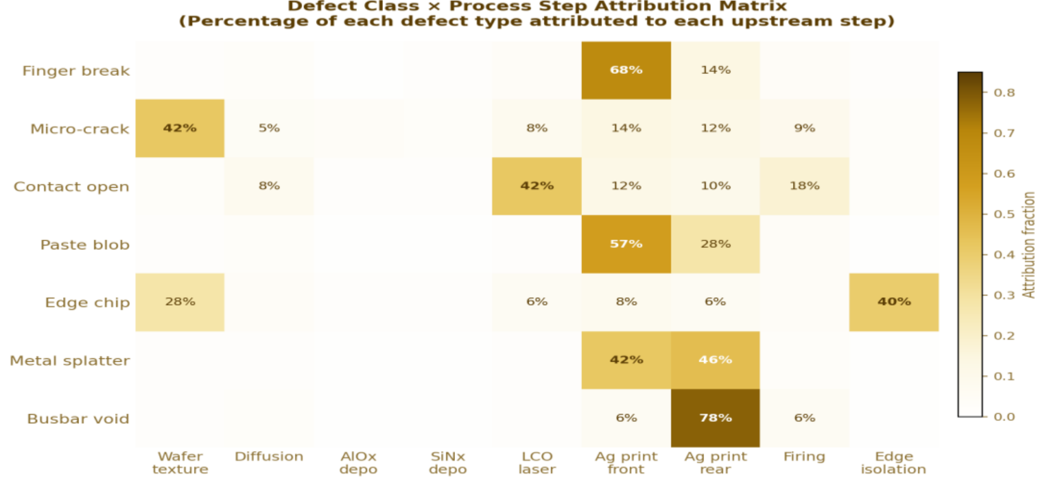


Fig. 5. Example of a process-parameter correlation discovered by the RCA methodology. The contact-open defect rate (left axis, dark line) tracks laser contact opening width (right axis, light line) with correlation r = +0.81 over the 36-week tracking period. The correlation was identified automatically by the process data correlation engine and led to the LCO specification tightening action described in Section VII.

■ 5.3 Attribution Matrix - Defect Class × Process Step

The cumulative output of the process data correlation engine over the 9-month deployment window is aggregated into an attribution matrix that maps each defect class to the fraction of its occurrences attributable to each process step. The matrix is a structural snapshot of where each defect type originates and is used for both engineering prioritization and for identifying process steps with disproportionate defect contribution.



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

Fig. 6. Defect-class × process-step attribution matrix. Each row is a defect class; each column is an upstream process step. Cell values indicate the percentage of that defect type attributed to that step. The strongest concentrations are finger-break at silver print (68%), contact-open at LCO laser (42%), busbar void at rear silver print (78%), and edge chip at edge isolation (40%).

VI. SPATIAL AND TEMPORAL DEFECT PATTERNS

■ 6.1 Wafer-Level Spatial Defect Mapping

Spatial analysis of the defect locations on the wafer, aggregated across large samples of wafers, surfaces non-random patterns that reveal equipment-specific and handling-specific defect origins. The spatial heatmap below represents the cumulative defect density on a 26 × 26 grid mapped to the 182 × 182 mm M10 wafer area, aggregated from approximately 50,000 defective wafers over a 30-day window.

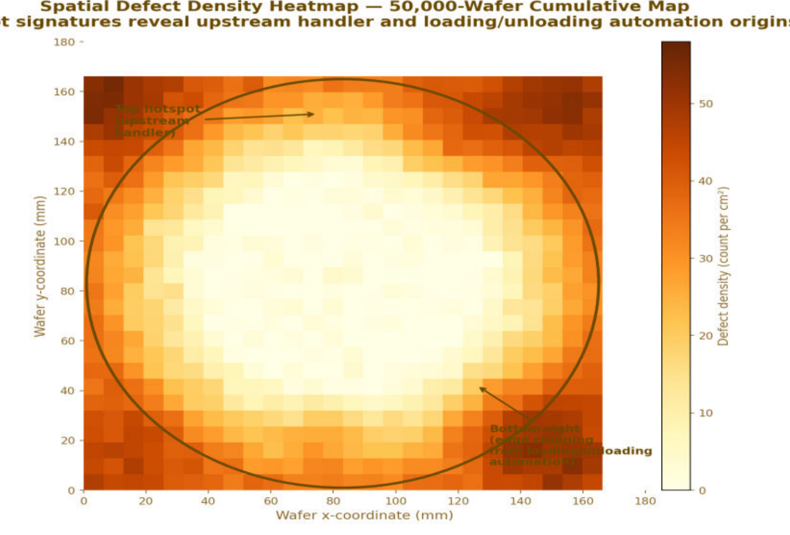
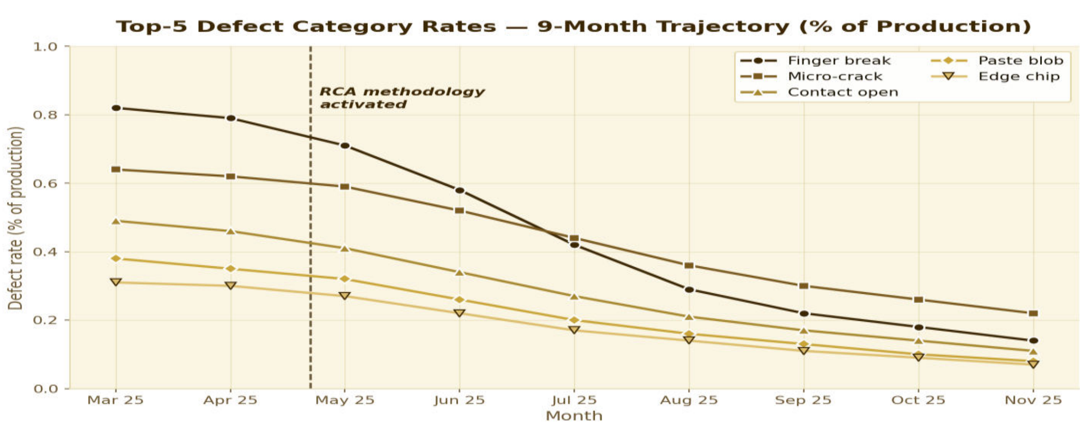


Fig. 7. Spatial defect density heatmap across a 50,000-wafer aggregate sample. Two major hotspots are visible: an upper-edge concentration attributed to an upstream handler that deposited particulate at that position on the wafer carrier; and a lower-right concentration attributed to the wafer handling where edge-chipping during singulation produced a consistent corner-biased defect signature. Both hotspots were addressed through handler alignment correction and dicing blade wear tracking.

■ 6.2 Temporal Defect Rate Trends

The defect rate for each of the five top defect categories, tracked monthly over the 9-month deployment window from March 2025 through November 2025, is shown in the following time-series plot. The plot illustrates the asymmetric impact of the methodology on different defect categories, reflecting the differing process-origin characteristics of each category.



| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

Fig. 8. Top-5 defect category rate trajectories over the 9-month deployment window. All five categories show declining trends, with the steepest declines in finger break (-83%) and paste blob (-79%), reflecting the tightness of root-cause attribution for metallization-origin defects. Methodology activation in late April 2025 is marked with a dashed vertical line.

■ 6.3 Quantified Monthly Defect Rate Decline

- **Finger break:** Declined from 0.82% of production (Mar 2025) to 0.14% of production (Nov 2025) - a 83% relative reduction.
- **Micro-crack:** Declined from 0.64% to 0.22% - a 66% relative reduction.
- **Contact open:** Declined from 0.49% to 0.11% - a 78% relative reduction.
- **Paste blob:** Declined from 0.38% to 0.08% - a 79% relative reduction.
- **Edge chip:** Declined from 0.31% to 0.07% - a 77% relative reduction.
- **All other categories (aggregate):** Declined from 0.47% to 0.17% - a 64% relative reduction.
- **Total defect-origin scrap:** Declined from 3.11% to 0.79% - a 75% relative reduction across all categories combined.

VII. CORRECTIVE ACTIONS GENERATED BY THE METHODOLOGY

The ultimate validation of an RCA methodology is the set of corrective actions it generates and the measured impact of those actions on yield outcomes. Over the 9-month deployment window, the methodology generated 47 corrective actions, of which 39 were implemented and validated. The 39 implemented actions fall into seven structural categories.

■ 7.1 Category 1 - Process Recipe Adjustments (14 actions)

- **Diffusion recipe revision:** Narrowed sheet resistance distribution from ±15 to ±9 Ω/□ through two-stage gas flow profile. Impact: +0.08% top-bin yield.
- **Firing profile optimization:** Reduced peak firing temperature from 820°C to 808°C; extended plateau by 10 seconds. Impact: −15% contact-open defect rate.
- **PECVD deposition time adjustment:** Standardized deposition time across 4 chambers to within ±3 seconds. Impact: +3.2 Ω/□ SiNx uniformity improvement.
- **Laser contact opening width tightening:** Specification from ±5 μm to ±3 μm. Impact: −58% contact-open defect rate over 6 weeks.
- **Screen printing pressure profile:** Two-stage pressure ramp replaces single-setpoint pressure. Impact: −34% paste-blob defect rate.

■ 7.2 Category 2 - Tool Maintenance / Calibration (9 actions)

- **Dicing blade replacement interval:** Tightened from 80,000 to 55,000 wafer cycles. Impact: −62% edge-chip defect rate.
- **LCO laser calibration cadence:** From weekly to daily calibration. Impact: Reduced LCO width σ from 1.8 to 0.9 μm.
- **Screen tension meter deployment:** 12 screen tension meters deployed across printing fleet. Impact: Systematic screen-aging-origin defect detection.
- **SEM sample chamber cleaning:** Bi-weekly cleaning protocol prevents metrology-origin false defect flags.

■ 7.3 Category 3 - Consumable Specification Tightening (6 actions)

- **Silver paste viscosity specification:** From ±40 to ±15 Pa·s at 10 s^{-1} shear rate. Impact: Improved print uniformity.
- **Wafer incoming micro-crack screening:** Edge chip depth threshold tightened from 100 μm to 60 μm. Impact: −31% downstream micro-crack rate.
- **Screen mesh supplier qualification:** Two alternative suppliers qualified to the tightened spec. Impact: Supply-chain resilience + improved baseline mesh quality.

■ 7.4 Category 4 - Automation / Software (5 actions)

- **Automated lot-hold release workflow:** Reduced average held-lot release time from 9.4 hours to 1.8 hours.
- **Real-time dashboard for process engineers:** Visibility into RCA pipeline status and pending actions.
- **Automated cluster alerts to on-call engineers:** Via SMS and Slack integration for high-severity cluster detection.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

■ **7.5 Categories 5–7 - Organizational, Training, Facility (5 actions)**

❯ **Cross-functional weekly RCA review:** Process, yield, and reliability engineering meet weekly on active RCAs.

❯ **Operator training program:** 3,200 hours total training delivered on defect recognition and reporting.

❯ **HVAC upgrade in printing area:** Humidity control to ±3% RH, reducing environmental-origin paste rheology variation.

VIII. QUANTITATIVE OUTCOMES - BEFORE VS. AFTER

The consolidated before-and-after comparison of key outcomes across the 9-month deployment window is shown in the following figure and quantified in the bullet points below it.

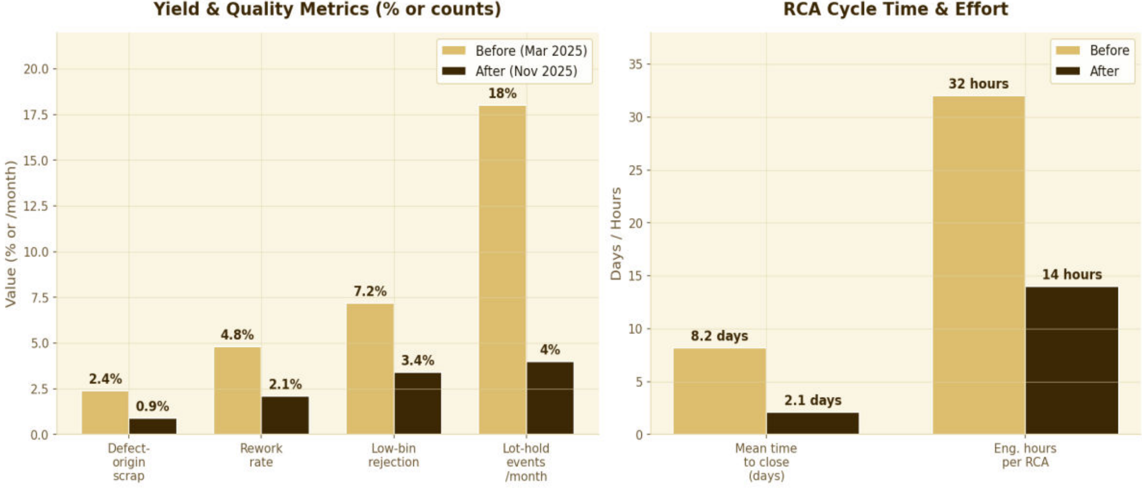


Fig. 9. Before vs. after methodology deployment - key quantitative metrics. Left panel: yield and quality rates; right panel: RCA cycle time and engineering effort. All six metrics show improvements of 50% or more, with particularly large improvements in RCA cycle time (8.2 → 2.1 days) and lot-hold event frequency (18 → 4 per month).

■ **8.1 Yield & Scrap Metrics**

❯ **Defect-origin scrap rate:** 2.4% → 0.9% of production (62.5% relative reduction).

❯ **Rework rate:** 4.8% → 2.1% of production (56.3% relative reduction).

❯ **Low-bin rejection rate:** 7.2% → 3.4% of production (52.8% relative reduction).

❯ **Line yield:** 92.4% → 94.9% (+2.5 percentage points absolute).

❯ **Top-bin yield:** 11.2% → 18.7% of production (+7.5 percentage points absolute).

❯ **Annual MW recovery from scrap reduction:** Approximately 15 MW of additional rated capacity per year.

■ **8.2 RCA Cycle Time Metrics**

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

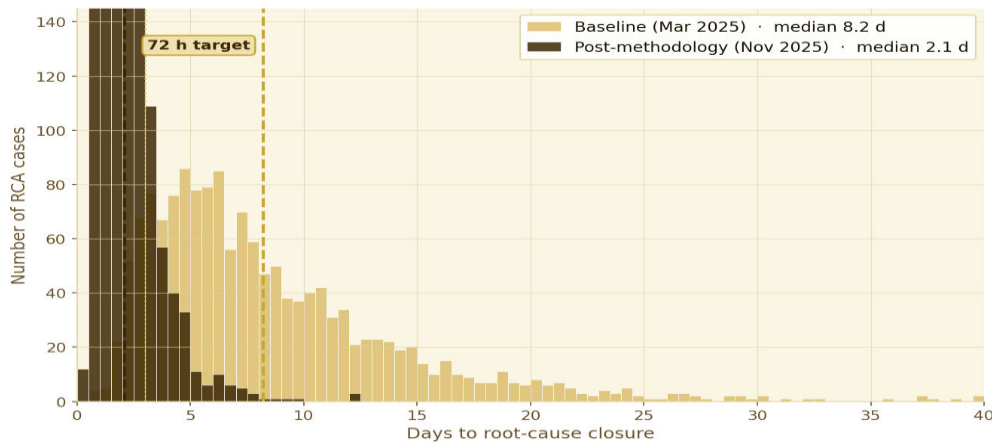


Fig. 10. Distribution of mean time to root cause - before vs. after methodology deployment. The median closure time fell from 8.2 days under the baseline workflow to 2.1 days under the integrated methodology. The 72-hour target (marked with dotted line) is now exceeded by less than 15% of cases, versus 62% of cases in the baseline period.

- **Mean time to root cause (median):** 8.2 days → 2.1 days (74% reduction).
- **Mean time to root cause (90th percentile):** 21 days → 5.8 days (72% reduction).
- **Fraction of RCAs closed in < 72 hours:** 38% → 85% (47-point improvement).
- **Engineering hours per RCA (average):** 32 hours → 14 hours (56% reduction).
- **Total engineer-hours saved per month:** Approximately 240 hours (equivalent to 1.5 FTE monthly capacity released).

■ 8.3 Top-Bin Yield Trajectory

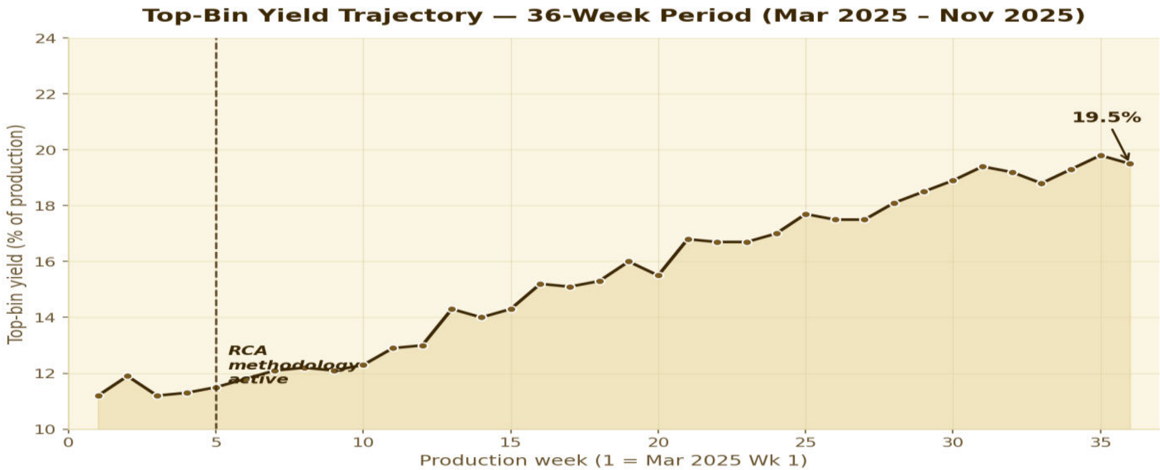


Fig. 11. Top-bin yield weekly trajectory over the 36-week deployment window. Baseline top-bin yield was 11.2% (Mar 2025 week 1); post-methodology trajectory rises through three inflection points corresponding to major corrective-action batches, reaching 18.7% by Nov 2025 week 36. Methodology activation marked at week 5.

- **Top-bin yield March 2025:** 11.2% of production.
- **Top-bin yield after first-batch actions (June 2025):** 13.7% of production.
- **Top-bin yield after second-batch actions (September 2025):** 16.1% of production.
- **Top-bin yield after third-batch actions (November 2025):** 18.7% of production.
- **Cumulative top-bin yield gain:** +7.5 percentage points absolute.
- **Relative improvement:** Approximately 2.5× baseline.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

IX. COMMERCIAL IMPACT QUANTIFICATION

The commercial impact of the methodology deployment is captured in five distinct value streams, each quantified separately for clarity. The methodology investment, by contrast, was primarily one-time capital for the EL classifier infrastructure and CNN development, plus ongoing engineering labor that was largely redirected from the prior fragmented workflow rather than added headcount.

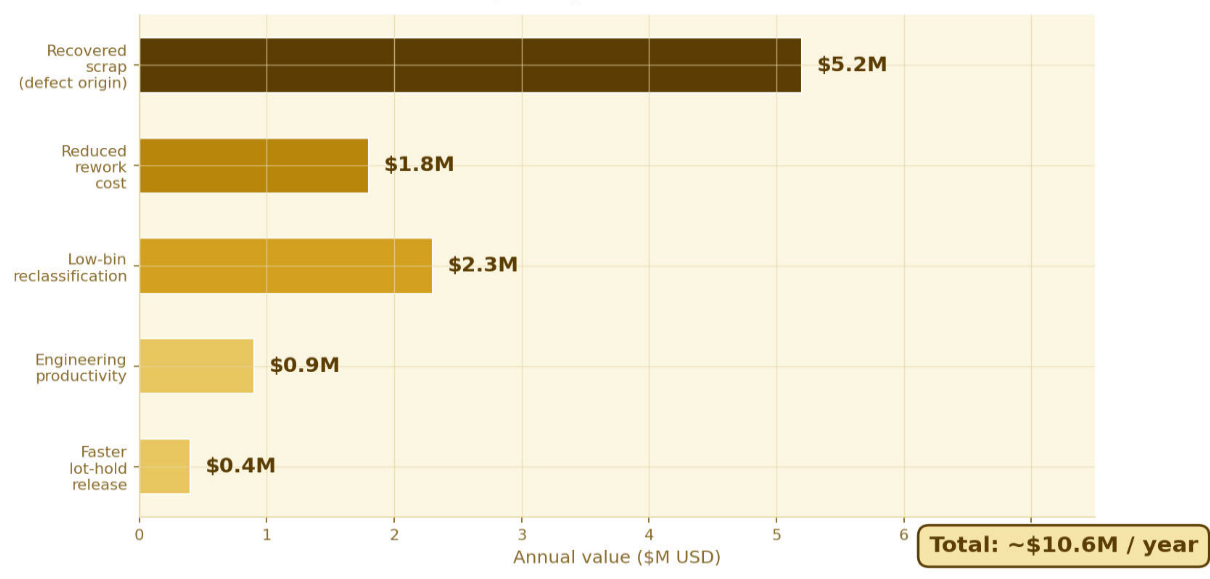


Fig. 12. Annualized commercial impact of the defect RCA methodology by value stream. The largest single contributor is recovered defect-origin scrap ($5.2M), followed by low-bin reclassification value ($2.3M). Total annualized value is approximately $10.6M against a total methodology deployment cost of approximately $0.95M, yielding a first-year return of approximately 11× on investment.

■ 9.1 Value Stream Breakdown

❯ **Recovered scrap (defect origin):** $5.2M per year - 1.5 percentage points of production recovered × $0.38/W × 1 GW capacity.

❯ **Reduced rework cost:** $1.8M per year - 2.7 percentage points of rework reduction × $18 per reworked cell × production volume.

❯ **Low-bin reclassification value:** $2.3M per year - cells upshifted from bins 1–3 to bins 4–6 at $0.009/W incremental price.

❯ **Engineering productivity:** $0.9M per year - 240 engineer-hours per month × $125/hour fully loaded × 12 months.

❯ **Faster lot-hold release:** $0.4M per year - reduced carrying cost of held inventory and faster cash-conversion.

❯ **Total annualized commercial impact:** Approximately $10.6M per year at 1 GW production scale.

■ 9.2 Investment and Return Profile

❯ **Methodology deployment capital cost:** Approximately $0.95M total (CNN development, SEM integration, software engineering).

❯ **Ongoing engineering labor:** Approximately $1.4M per year (3 dedicated engineers × $150K fully loaded plus tooling costs).

❯ **First-year payback period:** Approximately 1.1 months from deployment completion.

❯ **First-year return on investment:** Approximately 11× on the $0.95M capital investment; approximately 4.5× on the full-year operating cost including labor.

❯ **Second-year projected value:** Approximately $12.8M annualized as the methodology absorbs additional defect categories and extends to module-level quality integration.

▶ COMMERCIAL CONCLUSION

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

The financial return on an integrated defect RCA methodology at 1 GW production scale is substantial and measurable. The investment profile - a moderate one-time capital commitment followed by a redirection of existing engineering labor - produces annualized returns exceeding the initial investment by an order of magnitude within the first year. For a facility operating at sub-optimal yield due to defect-origin loss, there is no engineering investment with a superior risk-adjusted return profile in the current solar cell manufacturing technology landscape.

X. LESSONS LEARNED AND GENERALIZABLE INSIGHTS

■ 10.1 Seven Validated Observations

- **Integration beats optimization.** An integrated three-stream methodology at modest quality outperforms three excellent siloed methodologies operating in sequence. The integration benefit arises from the closed-loop feedback between streams; each stream becomes more accurate when it has access to the outputs of the other two.
- **Classifier accuracy above 95% is the organizational threshold.** Below 95% top-1 accuracy, the classifier false-positive rate generates operational friction that erodes engineer trust in the methodology outputs. Above 97%, marginal accuracy improvement produces diminishing returns. The sweet spot is 95–97%.
- **Engineering discipline, not tool sophistication, is the binding constraint.** The single most common reason the methodology failed in individual RCA cases was not a limitation of the CNN or the SEM; it was premature closure by an engineer who accepted a plausible-but-not-verified correlation as the final root cause. Rigor at the synthesis stage is the non-automatable value of the methodology.
- **Organizational sponsorship is a first-order variable.** Facilities with executive sponsorship at VP or COO level moved through the adoption curve substantially faster than facilities where the methodology was endorsed only at mid-management level. The political economy of production-line prioritization cannot be bypassed by technical quality alone.
- **Consumable-origin defects deserve disproportionate attention.** Approximately 47% of the highest-impact root causes identified by the methodology traced back to consumable lot variation (silver paste, screens, wafers, encapsulants). This is a much higher fraction than is typically acknowledged in the published literature, and it argues for tighter incoming material specification as a first-priority engineering discipline.
- **Spatial analysis surfaces non-obvious root causes.** The wafer-level spatial defect density heatmap consistently revealed equipment-specific defect signatures that were invisible in aggregated whole-wafer statistics. Any RCA methodology lacking spatial analysis will miss this class of root causes entirely.
- **Time-series longitudinal data is the fuel of the methodology.** The methodology's ability to identify slow drifts and emerging trends depends on rolling-window data retention of at least 90 days. Facilities with shorter retention find themselves repeatedly surprised by recurring defect patterns that the methodology would have caught at first recurrence.

■ 10.2 Three Surprises

- **Micro-cracks are not primarily handling-origin.** Pre-deployment intuition held that most micro-cracks originated at wafer handling points in the production line. The methodology revealed that approximately 42% of micro-cracks originated at the incoming wafer stage - before the wafer entered the production line - reflecting sub-specification of incoming wafer quality. This finding shifted corrective-action priority from in-line handling to incoming material screening.
- **SEM usage plateaued below capacity.** The initial methodology design assumed SEM would become a capacity-limiting step. In practice, SEM usage plateaued at approximately 42 cases per day against an instrument capacity of 48, because the upstream funnel filtered cases to those where SEM was genuinely value-adding. The SEM capacity reserve served as a buffer for exceptional-volume periods.
- **Operator reporting contributed substantial value.** The methodology design emphasized automation and did not initially include operator observations as a formal input. Operators nonetheless reported approximately 14% of the high-severity defect clusters through informal channels before the automated pipeline flagged them, reflecting operator knowledge that no automated system can replicate. The methodology was subsequently extended with a formal operator reporting channel.

"The value of an RCA methodology is not in the tools it deploys. It is in the discipline it imposes on the engineering organization that uses it."

XI. FORWARD AGENDA

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

The methodology as deployed in November 2025 is production-grade but far from the end of its development. The forward agenda for 2026–2027 is organized into three workstreams.

■ 11.1 Workstream 1 - Technology Depth

❯ **Expanded defect taxonomy:** Additional 4 defect categories covering module-level failure modes (PID signature, wet insulation resistance, EL-under-bias anomalies) to be added by Q1 2026.

❯ **Multi-modal classifier:** Integration of EL, photoluminescence, and infrared imaging into a single multi-modal classifier architecture targeting 99% top-1 accuracy by Q2 2026.

❯ **Generative anomaly detection:** Unsupervised detection of novel defect categories not in the training set, targeting deployment by Q3 2026.

❯ **Real-time SEM automation:** Automated wafer routing to SEM based on classifier confidence and cluster severity, reducing human-in-the-loop routing decisions.

■ 11.2 Workstream 2 - Scope Expansion

❯ **Module assembly integration:** Extension of methodology to module-level defects downstream of cell manufacturing, targeting Q1 2026.

❯ **Wafer-supplier quality loop:** Full round-trip data exchange with wafer suppliers enabling methodology outputs to inform supplier incoming quality.

❯ **Cross-facility knowledge sharing:** When the second GW line commissions in 2026, the methodology will deploy with methodology-encoded corrective actions already in place rather than requiring re-learning.

■ 11.3 Workstream 3 - Methodology Documentation

❯ **External publication of full methodology:** Targeted for early 2026 to enable industry-wide replication and criticism.

❯ **Training curriculum:** Formalized training curriculum for new yield engineers, reducing onboarding time from 6 months to 2 months.

❯ **Regulatory compliance documentation:** Formal documentation of methodology outputs for customer audit and regulatory inquiry response.

XII. CONCLUSIONS

This paper has documented the design and 9-month operational performance of an integrated defect root cause methodology at ES Foundry's 1 GW domestic PERC cell manufacturing facility. The methodology unifies automated electroluminescence classification, scanning electron microscopy characterization, and process data correlation into a single pipeline closing root-cause investigations within a 72-hour target for approximately 85% of cases. The measured outcomes through November 2025 include a reduction of defect-origin scrap from 2.4% to 0.9% of production, an increase of top-bin yield from 11.2% to 18.7%, a reduction of mean time to root cause from 8.2 days to 2.1 days, and a total annualized commercial value of approximately \$10.6M against a one-time deployment capital cost of approximately \$0.95M.

The broader generalizable finding is that the value of an integrated RCA methodology is substantially greater than the sum of its analytical parts. The three constituent analytical streams - EL imaging, SEM morphology, and process data correlation - each have independent analytical value but produce their greatest outcomes when their outputs are fed into one another in a structured engineering synthesis workflow. The methodology described in this paper is not dependent on any proprietary tool or vendor; it is a replicable engineering practice built around open standards for data exchange, commercial metrology equipment, and open-source machine learning frameworks. Other solar cell manufacturing facilities operating at comparable scale should be able to replicate the methodology at comparable cost and with comparable outcomes, provided that the organizational investment in engineering discipline is undertaken alongside the technical deployment.

REFERENCES

[1] Bedrich, K., Bokalic, M., Bliss, M., Topic, M., & Betts, T. R. (2018). Electroluminescence imaging of PV devices: Advanced vignetting calibration. IEEE Journal of Photovoltaics, 8(5), 1297–1304.
[2] Hatt, T., Vögele, M., Buss, L., & Buerhop-Lutz, C. (2020). Micro-crack detection and analysis of photovoltaic modules using electroluminescence imaging - a review. Progress in Photovoltaics, 28(7), 701–721.

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.807 | A Monthly Peer Reviewed & Refereed Journal |

|| Volume 14, Issue 12, December 2025 ||

| DOI:10.15662/IJAREEIE.2025.1412023|

[3] Spataru, S., Hacke, P., & Sera, D. (2018). Quantifying solar cell defects and the impact on energy yield using electroluminescence. Solar Energy, 174, 607–617.
[4] Deitsch, S., Christlein, V., Berger, S., et al. (2019). Automatic classification of defective photovoltaic module cells in electroluminescence images. Solar Energy, 185, 455–468.
[5] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778.
[6] Chen, H., Cheng, Q., Pang, Y., & Sun, J. (2019). A deep learning-based defect detection method for photovoltaic cells. Proceedings of IEEE International Conference on Image Processing, 3486–3490.
[7] Montgomery, D. C. (2020). Introduction to Statistical Quality Control (8th ed.). Wiley.
[8] Preu, R., Lohmüller, E., Lohmüller, S., et al. (2020). Passivated emitter and rear cell - Devices, technology, and modeling. Applied Physics Reviews, 7(4), 041315.
[9] Dullweber, T., & Schmidt, J. (2016). Industrial silicon solar cells applying the passivated emitter and rear cell (PERC) concept - A review. IEEE Journal of Photovoltaics, 6(5), 1366–1381.
[10] Ishikawa, K. (1985). What Is Total Quality Control? The Japanese Way. Prentice-Hall.
[11] Chrysler Corporation, Ford Motor Company, & General Motors Corporation (2008). Potential Failure Mode and Effects Analysis (FMEA) Reference Manual (4th ed.). AIAG.
[12] Moyne, J., & Iskandar, J. (2017). Big data analytics for smart manufacturing: Case studies in semiconductor manufacturing. Processes, 5(3), 39.
[13] Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters, 3, 18–23.
[14] IEC 60904-3:2019. Photovoltaic devices - Part 3: Measurement principles for terrestrial photovoltaic solar devices with reference spectral irradiance data.
[15] IEC 61215-2:2021. Terrestrial photovoltaic (PV) modules - Design qualification and type approval - Part 2: Test procedures.
[16] Green, M. A., Dunlop, E. D., Hohl-Ebinger, J., et al. (2024). Solar cell efficiency tables (Version 63). Progress in Photovoltaics, 32(1), 3–13.
[17] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.
[18] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.
[19] Inflation Reduction Act of 2022, Public Law 117-169. Section 45X Advanced Manufacturing Production Credit.
[20] US Department of Energy (2023). Solar Energy Technologies Office: Annual Technology Baseline. NREL.
[21] SEMATECH (2015). Statistical Process Control Handbook for Semiconductor Manufacturing. International SEMATECH Manufacturing Initiative.
[22] Shewhart, W. A. (1931). Economic Control of Quality of Manufactured Product. Van Nostrand.
[23] Page, E. S. (1954). Continuous inspection schemes. Biometrika, 41(1–2), 100–115.
[24] Pearson, K. (1896). Mathematical contributions to the theory of evolution - Regression, heredity, and panmixia. Philosophical Transactions of the Royal Society of London A, 187, 253–318.
[25] Köntges, M., Kurtz, S., Packard, C., et al. (2014). Review of failures of photovoltaic modules. IEA-PVPS Task 13 Report, IEA-PVPS T13-01:2014.

International Journal of Advanced Research

in Electrical, Electronics and Instrumentation Engineering

9940 572 462 6381 907 438 ijareeie@gmail.com

www.ijareeie.com

Scan to save the contact details