

Print/Scan Recovery Frontier of PaperBack Data Pages

Dot density, printed fill, and recoverability on two office printers

This report brings together three weeks of controlled print-and-scan sweeps. It maps the boundary between reliable recovery and failure, compares Canon and HP output, and shows how scan resolution, dot density, and printed fill affect the data PaperBack can restore.

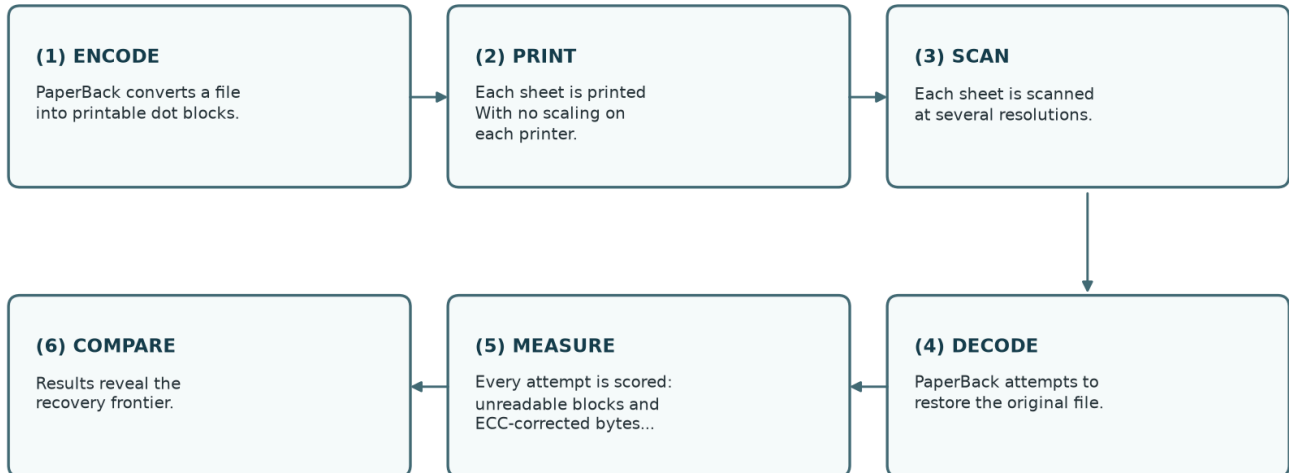
Data: 28 Sep · 17 Oct 2026 · Canon and HP sweep data
Chain: PaperBack 1.10 control build · standalone decode scoring
Print/Scan: Canon iR-ADV DX CS840i · HP LaserJet 4350dtn · Epson V700

Headline findings

- Below roughly 80 dpi, both printers have a broad recovery region.
- Near 100 dpi, printed fill and scan resolution become decisive.
- At 120 dpi and above, the clean region is narrow and printer-specific.

1 | What was tested

PaperBack encodes files as printed dot blocks and restores them from scanned images. We tested Canon & HP office printers with an Epson V700 scanner to measure how dot density, printed fill, and scan resolution affect recovery.



The same encode-print-scan-decode loop was repeated across the tested geometries and scan factors.

Terms used in the report

- Clean restore · the payload was restored with zero unreadable data blocks.
- Unreadable block · a block whose data could not be recovered by error correction.
- ECC corrections · bytes repaired inside blocks that remained readable; fewer corrections indicate greater margin.
- Fill · printed dot size as a percentage of dot spacing.
- Scan factor · scan resolution divided by nominal dot density; x2.5 at 100 dpi means a 250 dpi scan.

2 | Method and scoring

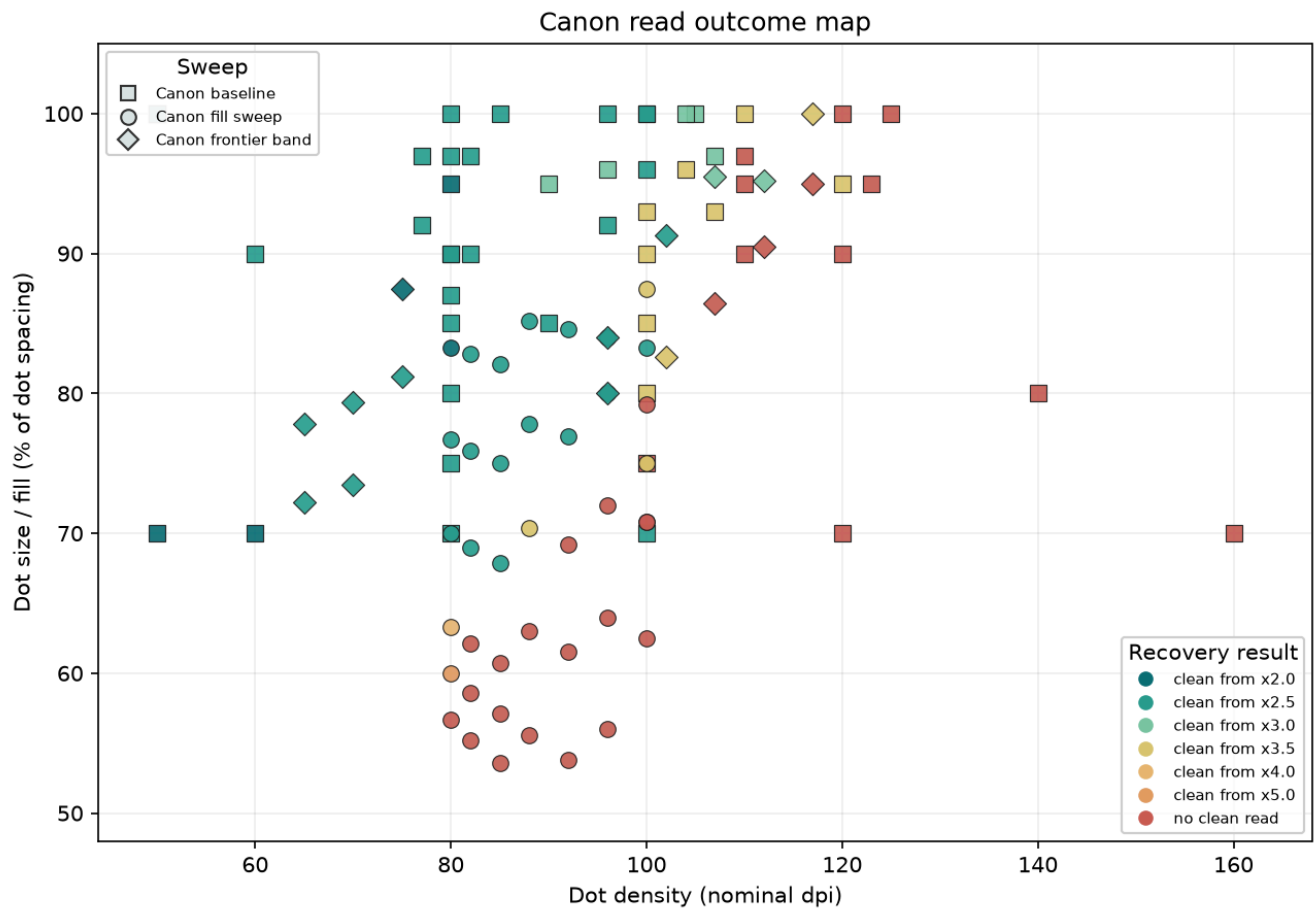
Sheets were printed at actual size with the highest available monochrome quality. Each band was scanned at factors from x2.0 to x5.0, then decoded as an independent attempt. The decoder slot was cleared between trials so a previous scan could not contribute blocks to the next result.

- Canon and HP candidates cover coarse, moderate, and fine dot pitches with multiple fill levels.
- Full-grid sweeps use per-family payloads sized to exercise the available block grid rather than a small shared test file.
- Campaigns affected by a page-order labeling error were relabeled from the composed XPS and re-decoded before inclusion.
- Outcome-map color records the lowest scan factor that produced a clean restore; red marks candidates that never restored cleanly.
- Frontier lines are conservative: at each density they sit above every failed fill, then are forced to rise monotonically with density.

Print	actual size · monochrome · best quality
Scan	grayscale · factors x2.0 / x2.5 / x3.0 / x3.5 / x4.0 / x5.0
Decode	independent trial · successful restore required
Quality	per-block ECC result + unreadable-block count
Redundancy	1:5 · maximal compression

3 | Canon read outcomes

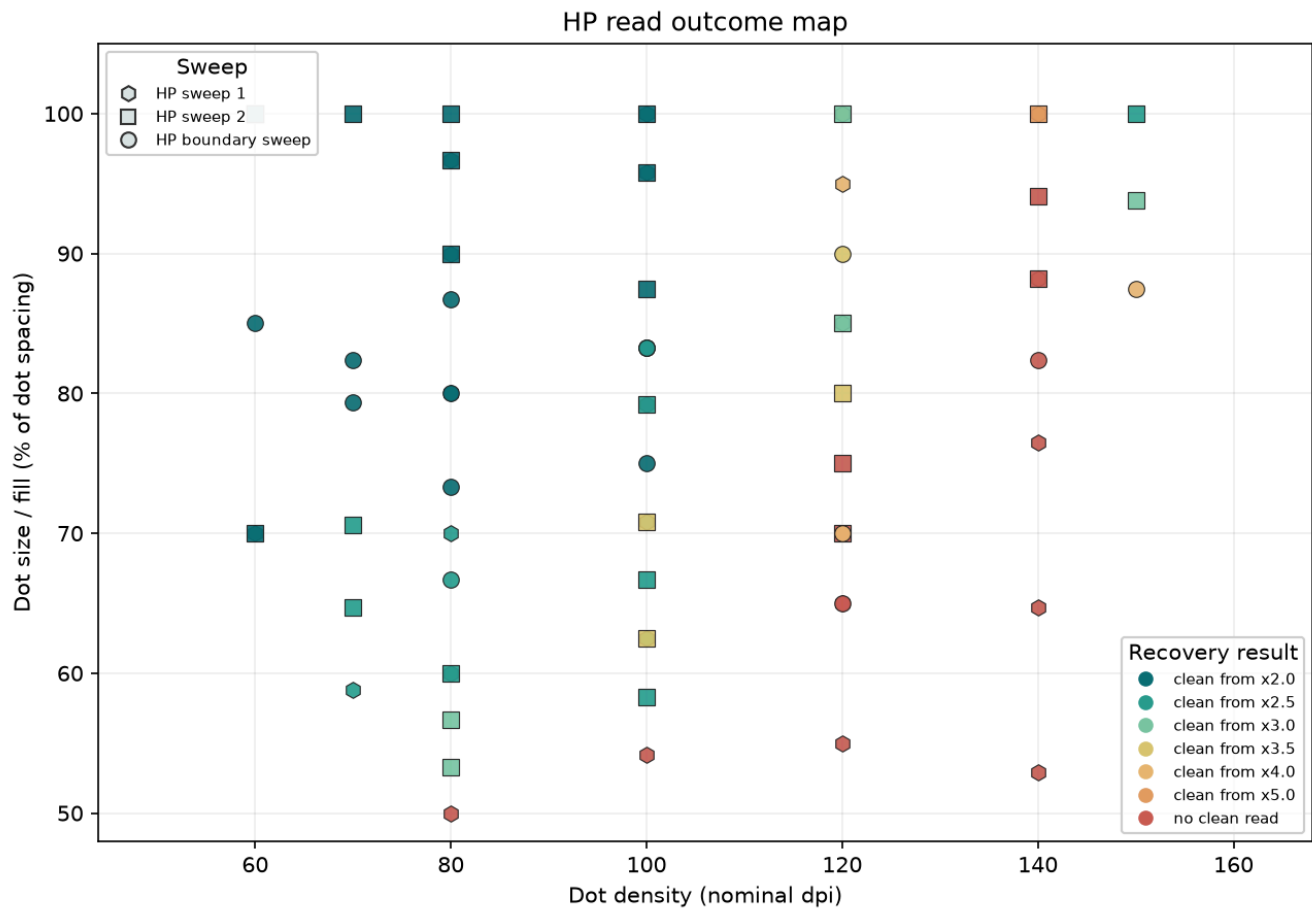
Each point is a tested combination of dot density and printed fill. Color shows the lowest scan factor that restored the payload; shape identifies the sweep. Red points did not restore cleanly at any tested factor.



Canon outcome map. The practical recovery region narrows as density rises; the frontier is broad below about 100 dpi and increasingly fill-sensitive above it.

4 | HP read outcomes

The HP map includes the original sweep, the full-payload sweep, and the boundary refinement sweep. Coarse pitches tolerate a wide range of fills; finer pitches need more fill and higher scan factors.

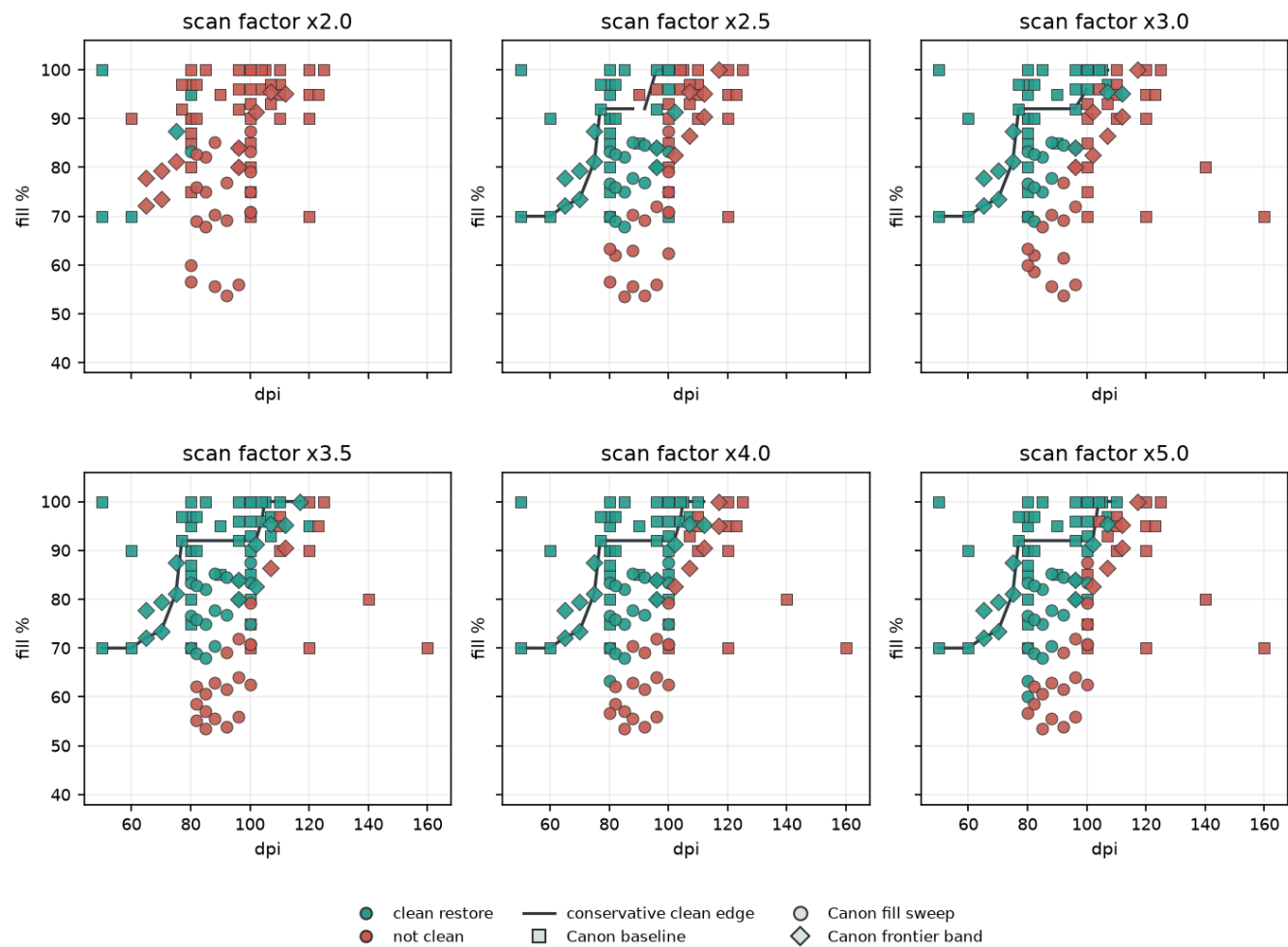


HP outcome map. Marker shape separates sweep 1, sweep 2, and the boundary sweep; color uses the same recovery scale as the Canon map.

5 | Canon frontier by scan factor

Separating the Canon data by scan factor shows how additional scan resolution expands the clean region. Sweep identity remains visible through compact marker shapes; no future or unmeasured candidates are shown.

Canon recovery frontier by scan factor

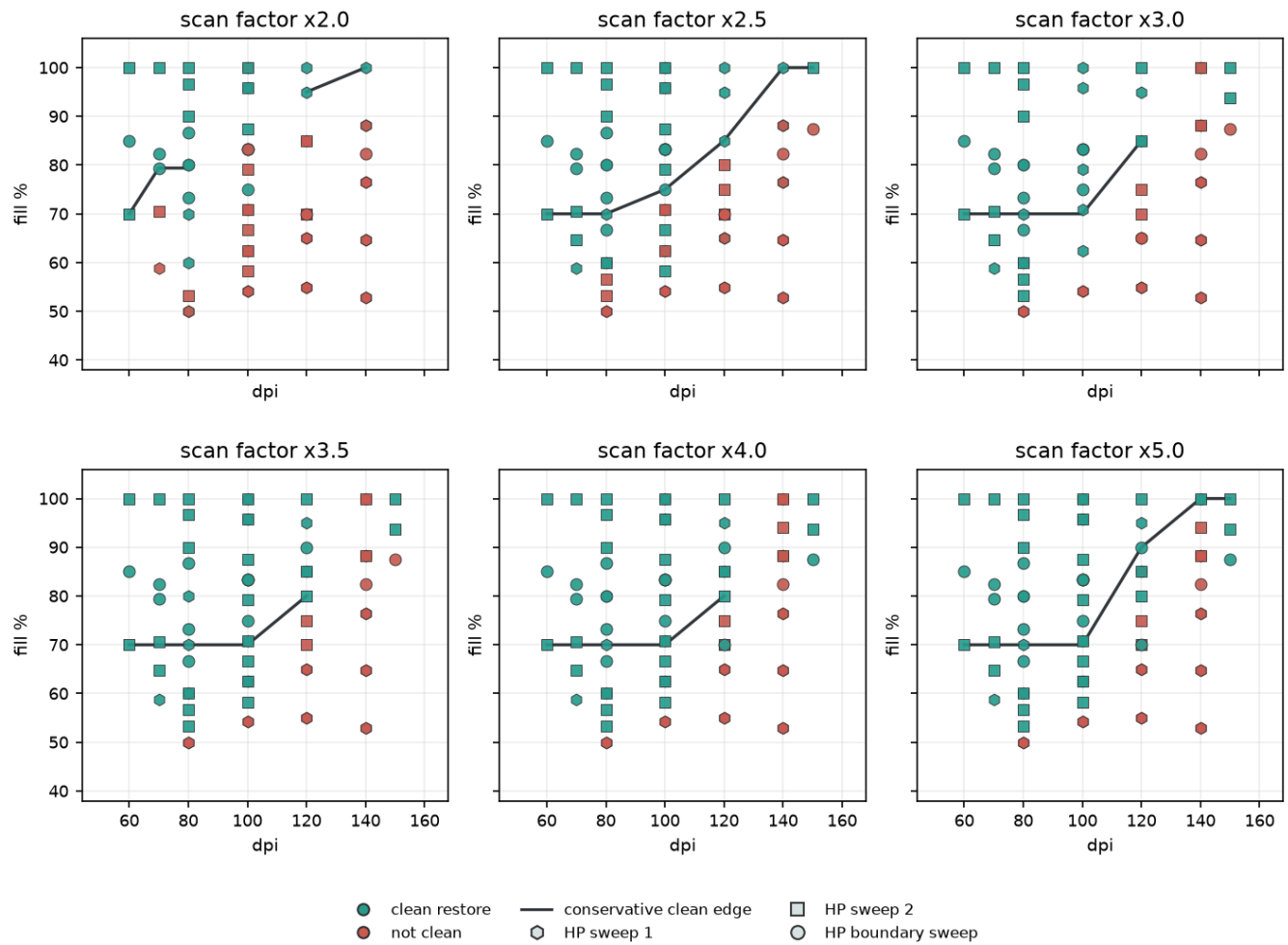


Teal points restored cleanly; rust points did not. The dark line is the conservative clean edge and breaks wherever no safe threshold is supported.

5.1 | HP frontier by scan factor

The HP slices include sweep 1 where no replacement sweep-2 result exists, plus the full-payload and boundary sweeps.

HP recovery frontier by scan factor

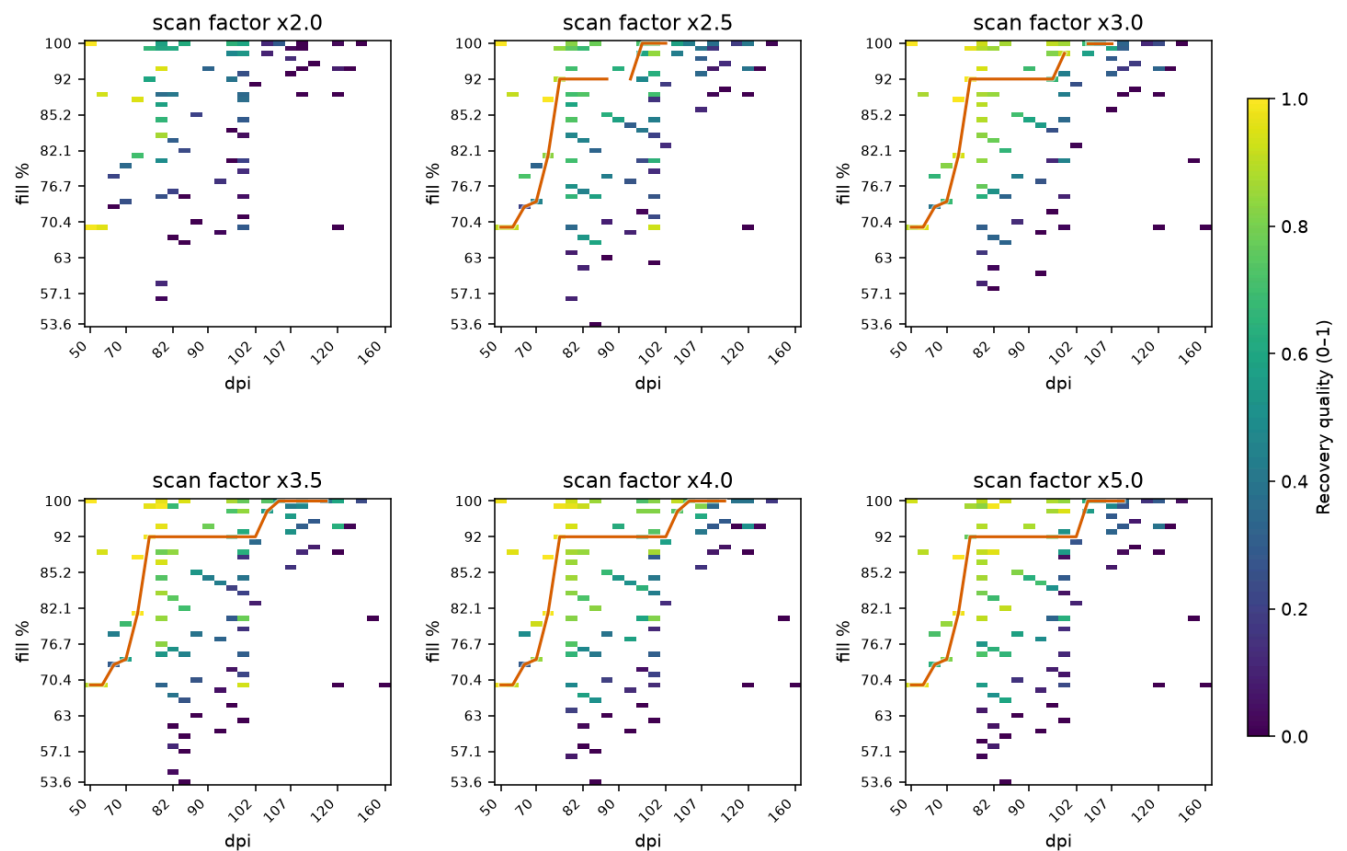


The clean edge remains above every observed failure and is not interpolated across unsupported densities.

5.2 | Canon recovery quality

The heatmaps summarize decoder quality across density and fill for each scan factor. Bright cells indicate stronger recovery margin. The orange line is the same conservative edge used in the frontier charts.

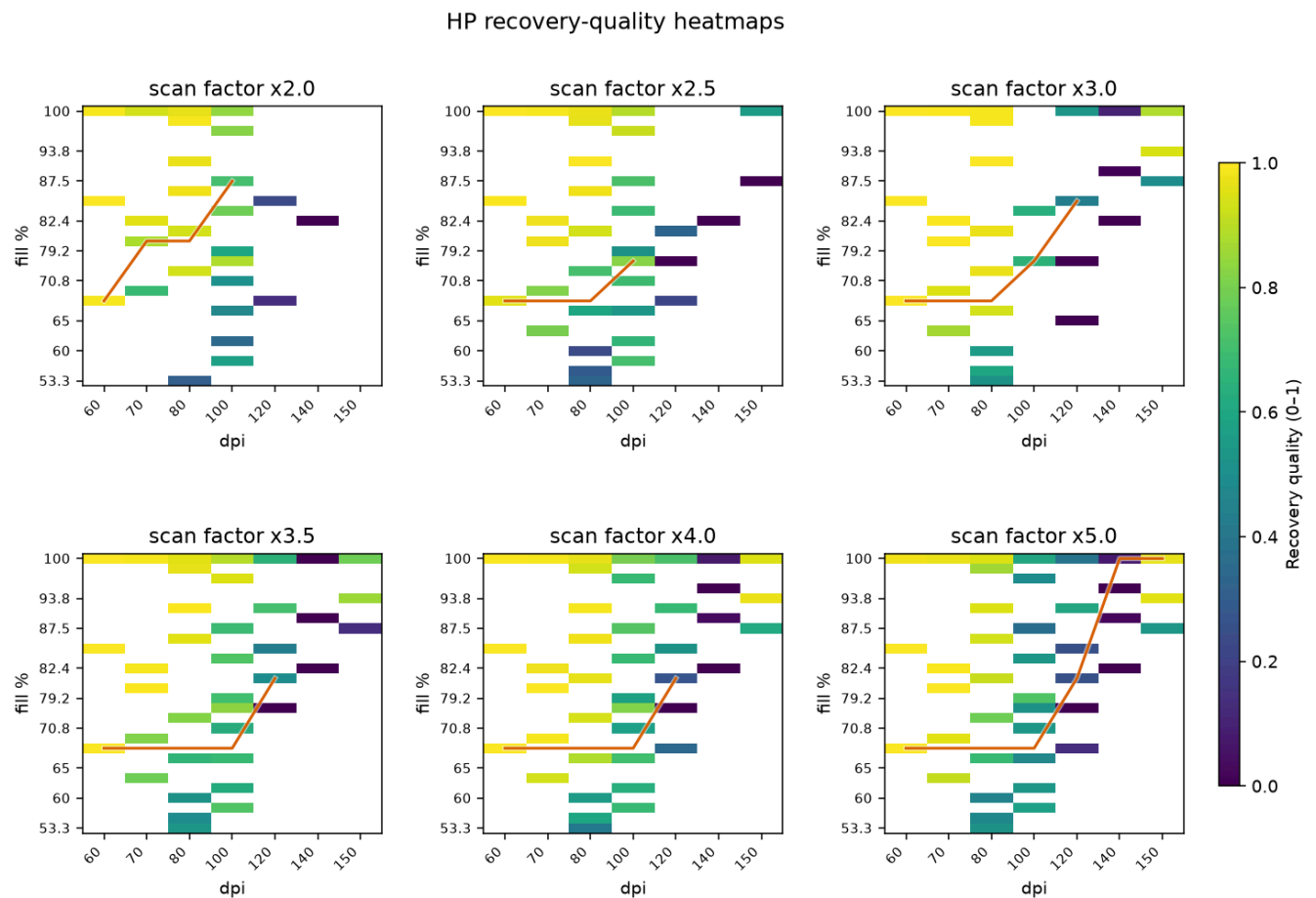
Canon recovery-quality heatmaps



Canon recovery-quality heatmaps. Blank cells indicate untested coordinates, not failures.

5.3 | HP recovery quality

The HP heatmaps use the corrected full-payload and boundary-sweep measurements. Sweep 1 appears in the outcome map and frontier slices; it is not included here because it was not normalized into the same quality metric table.



HP recovery-quality heatmaps. The narrow fine-pitch region is visible at 120-150 dpi.

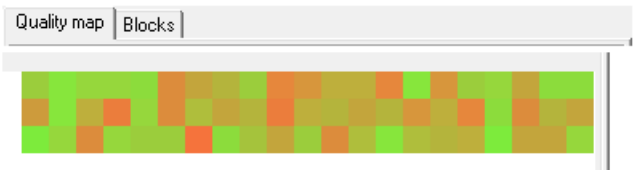
6 | PaperBack's per-block view

These are direct captures of PaperBack's Quality map tab after isolated decodes. Each trial ran in a fresh process with a unique save destination; the displayed map was checked against a fresh control-interface query before capture. Green blocks needed no correction, yellow and orange blocks used ECC, black blocks were unreadable, and white cells were outside the decoded data window.

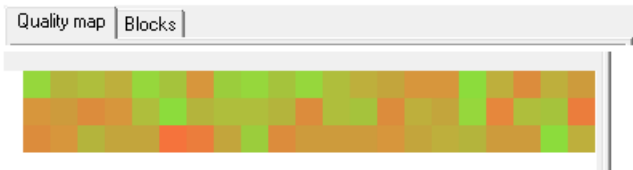
Canon CS840i · baseline · 80 dpi / 97% · x2.0 · 7 unreadable



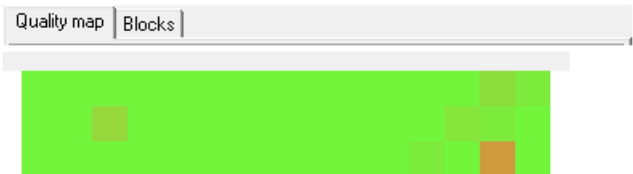
Canon CS840i · fill sweep · 100 dpi / 75% · x3.5 · clean



Canon CS840i · fill sweep · 100 dpi / 83.3% · x2.5 · clean



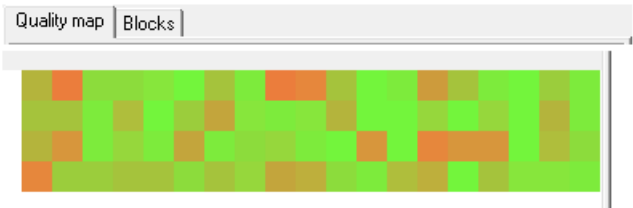
Canon CS840i · frontier band · 75 dpi / 87.5% · x2.0 · clean



HP 4350dtn · sweep 2 · 80 dpi / 90% · x2.0 · clean



HP 4350dtn · sweep 2 · 100 dpi / 79.2% · x2.5 · clean



HP 4350dtn · sweep 2 · 120 dpi / 75% · x3.5 · 57 unreadable



HP 4350dtn · boundary sweep · 140 dpi / 82.4% · x4.0 · 74 unreadable

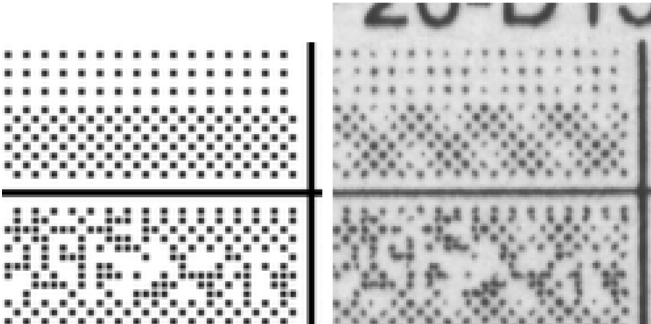


Eight software-native quality maps: four Canon trials and four HP trials, spanning clean, marginal, and failed pages.

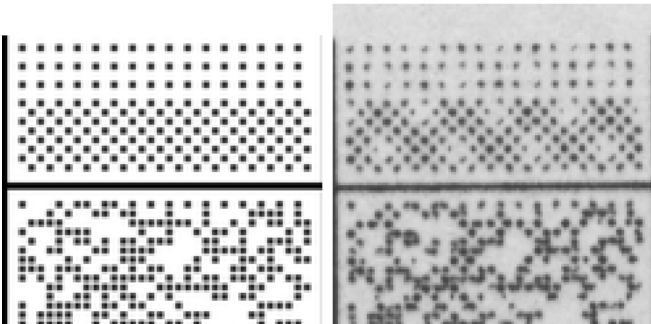
7 | Standout print-versus-scan blocks

Each pair shows the same data block at the same scan-pixel sampling. The 2400 dpi print bitmap was downsampled to the scan DPI, phase-registered to the scan, then locally correlation-checked using the decoder's exact block rectangle.

block (1,0) quality=17 PRINT|SCAN corr=0.66

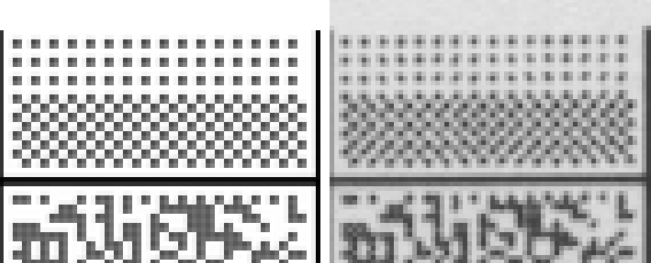


block (4,0) quality=16 PRINT|SCAN corr=0.77



HP sweep 2 · P20 / 75% · x3.5 · not clean. One unreadable block and one 16-byte correction block show how fine dots lose separation.

block (13,0) quality=5 PRINT|SCAN corr=0.75



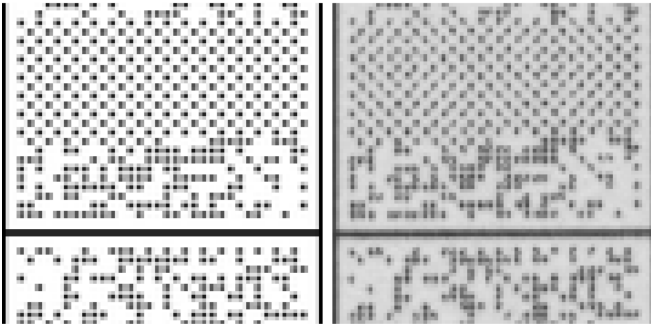
block (3,2) quality=0 PRINT|SCAN corr=0.76



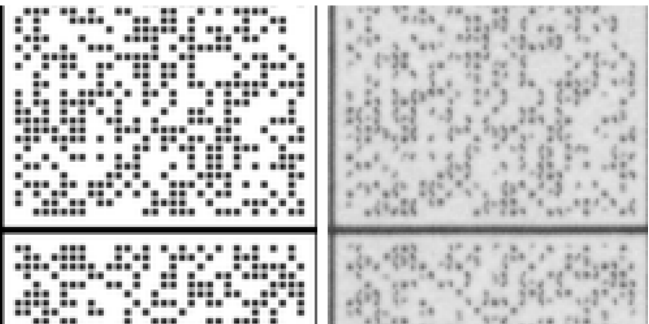
HP sweep 2 · P30 / 90% · x2.0 · clean. The print structure remains recognizable despite scanner blur and dot gain.

7.1 | Boundary and threshold cases

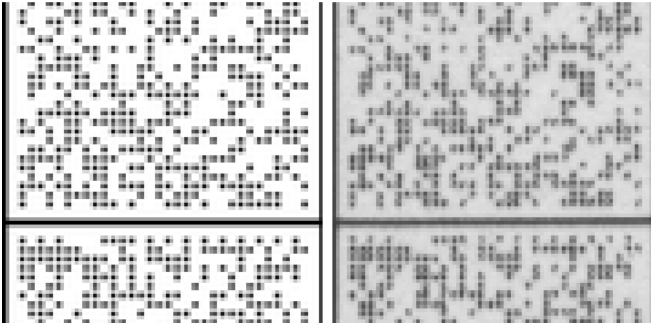
block (4,2) quality=15 PRINT|SCAN corr=0.71



block (7,2) quality=16 PRINT|SCAN corr=0.81



block (11,1) quality=0 PRINT|SCAN corr=0.80



HP boundary sweep · P30 / 66.7% · x2.5 · clean. This point sits near the lower-fill edge for P30.

Canon fill sweep · P24 / 75% · x3.5 · clean. The scan retains the same block pattern, but dot edges and background texture are visibly changed.

Registration details, source paths, block coordinates, shifts, and correlation scores are provided in cases-registration.json.

8 | Practical findings

- Below roughly 80 dpi, both printers provide a broad clean region; moderate fill often restores at x2.0·x2.5.
- Around 100 dpi, recovery depends strongly on fill and scan factor. Conservative use favors higher fill and at least x2.5·x3.5 scanning.
- At 120 dpi and above, the usable region narrows sharply. Near-full fill and higher scan factors are required, and some candidates remain unreliable.
- Canon and HP share the same overall frontier shape, but individual threshold points differ enough that printer-specific validation remains important.
- Per-block maps show why a page can be close to recovery without restoring: many blocks remain readable while a small cluster exceeds ECC capacity.

Conservative operating regions

Region	Suggested use	Scan guidance
60-80 dpi	reliability-first paper storage	x2.0·x2.5
around 100 dpi	capacity with tested fill margin	x2.5·x3.5
120 dpi and above	experimental / printer-specific	x3.5·x5.0

Capacity perspective

A P24 page at 100 dpi carries about 57 usable block positions at full fill, or roughly 5 KB before redundancy at about 90 payload bytes per block. Lower-fill P24 bands carry about 4 KB. P30 settings trade capacity for substantially wider recovery margin.

9 | Tooling, data, and limitations

The control build is a patched version of Ole Yuschuk's PaperBack 1.10 · an expanded command set plus a JSON control interface for agentic access to the software · driven over a local named pipe. It exports per-block ECC results, decoder geometry, and file-restoration status, and it feeds the quality heatmaps in section 5.3. A second build from the same patched sources reproduced the reference binary's exact size.

Distributed data

- canon-hp-quality-metrics-v2.1-20261013.csv · normalized Canon and HP quality records.
- outcome-map-canon-all-20261019.csv / .xlsx · point data and an Excel-native chart workflow.
- chart-manifest.json · figure counts and the conservative frontier rule.
- cases-registration.json · source images, exact block rectangles, registration shifts, and correlation scores.
- capture-manifest.json · software quality-map trials, save destinations, dimensions, and counts.
- PaperBak-Sweeps-Archive · scans, print masters, manifests, decode logs, and correction records.

Limitations

- Untested cells remain blank; charts do not imply measurements where none exist.
- The conservative frontier suppresses isolated optimistic results when a higher-fill failure exists at the same density.
- HP sweep 1 contributes verdicts to the outcome map and slices but is absent from the normalized quality heatmaps.
- Case comparisons show selected blocks, not every block in a page.

Generated from the durable sweep archive. Figure and report generators are included with the deliverable.