

Print/Scan Survival of PaperBack Data Pages

A fine-grain sweep of dot density, dot size, and scan factor

How well does a paper backup survive printing and scanning?

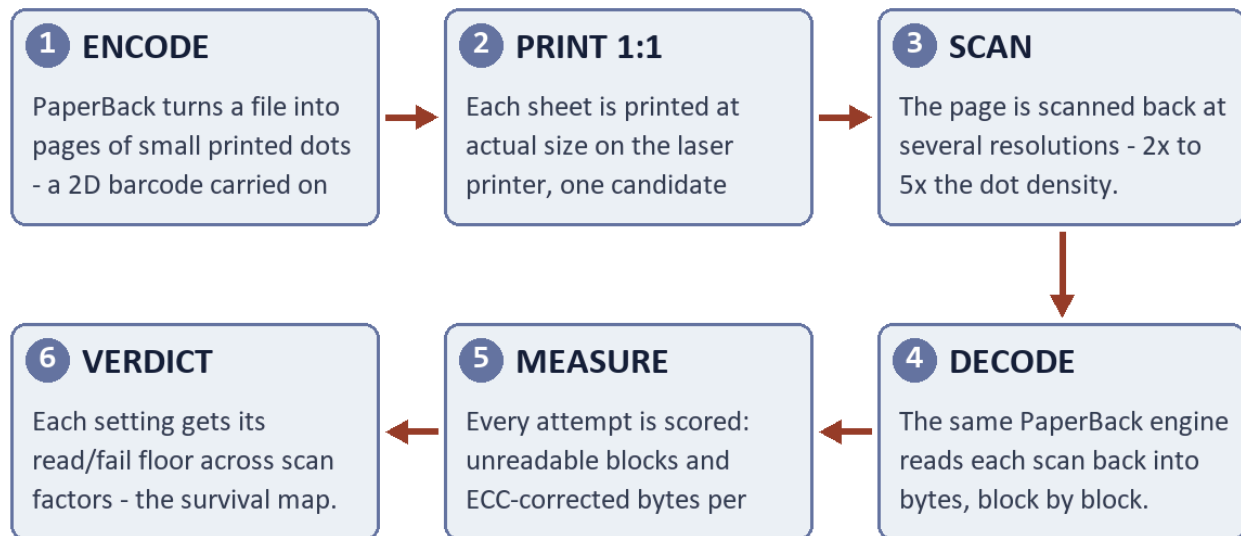
REVISED EDITION v5 - the October sweeps folded in (Oct 12-17)

This edition extends v4 with four October sweeps (101 new candidates): HP sweep-2-final and sweep-2b on the LaserJet 4350dtn, Canon sweep-3d (the first full-payload fill-floor map, P30-P24) and the Canon frontier-band. The big-picture outcome map is rebuilt to cover every Canon candidate across all five Canon campaigns (106 points), and the October finals - frontier-slices v3 and the Canon quality heatmap - carry a corrected clean criterion. The control binary reaches v8.2 and is proven rebuildable clean-room from durable sources.

Data: Sep 28 - Oct 17, 2026 (20 sweep campaigns)
Station: SCANNER01 / TECHCENTER15 - library workstations
Chain: PaperBack 1.10 - rebuilt from source, automation-enabled (ctl v8.2)
Print/Scan: Canon iR-ADV DX CS840i / HP LaserJet 4350dtn + Epson V700
Editions: v3 9/30 -> v4 10/9 -> **v5 10/19 (this PDF)**

1 | Start here: what was tested and how to read this report

PaperBack encodes a computer file into a two-dimensional matrix of small printed dots - a 2D barcode carried on paper. To retrieve the file, the page is scanned and decoded again. The catch: real printers and scanners are imperfect - fine dots print faint, scans blur, paper moves. This project measured which dot settings survive the whole chain, on two printers (Canon primary, HP added in October), across 20 sweep campaigns.



The workflow: a file is encoded into dot pages, printed 1:1, scanned at multiple resolutions, and decoded programmatically for measurement. Flow runs left-to-right across the top row, then drops into the bottom row and returns right-to-left (reverse-S).

A candidate is one (dot density, dot size) combination; October sweeps add fill level and per-family full-size payloads. Charts from the September editions (outcome map, ECC curves, quality maps, case standouts) are carried forward; the October finals (frontier slices, quality heatmaps, fill-floor curve) use a corrected clean criterion: a candidate counts as clean only on a successful standalone restore - no-grid and label-fail trials are never counted as clean.

How to read the numbers

- Unreadable blocks (nb) - data blocks the decoder could NOT recover. 0 = complete decode. A block holds about 90 payload bytes before 1:5 redundancy.
- Floor ("reads from xN") - smallest scan factor with a standalone clean raw decode. NG = no grid; LBL = capture label unreadable (a capture failure, not a decode verdict).
- ECC corrections (bytes) - bytes needing error-correction inside blocks that DID decode. Lower = more margin.
- Every verdict in this edition comes from standalone decodes: the decode slot is discarded before every trial, so no number is inflated by accumulation across scans.

2 | Equipment, constants, and the sweeps

Print chain: 1:1 prints ("Actual size", best quality) through the BW-8.5x11 EnvisionWare queue - Canon iR-ADV DX CS840i for the Canon sweeps, HP LaserJet 4350dtn for the HP sweeps (the color queue was retired 2026-10-13; all October sweeps print monochrome). Scan: Epson Perfection V700 Photo, 8-bit grayscale, no sharpening, factor ladder x2.0-x5.0 (factor x base dpi = scan dpi). Decode: PaperBack 1.10 rebuilt from source with a JSON-over-named-pipe control surface (section 11).

Compression	Maximal (bzip2 lvl = exact block size)	Redundancy	1:5
Encryption	off	Factors	x2.0 x2.5 x3.0 x3.5 x4.0 x5.0
Header/border	on	Payloads	309 B classic; 3.0-4.3 KB full (October)

Sweep design standards introduced in October

- px-exact geometry: dpi solved so $\text{floor}(2400/\text{dpi}) = \text{pitch}$; $\text{dot} = \text{floor}(\text{pitch} \times s/100)$; per-campaign px-validation audits.
- Labels nearest their own boxes (gap 14 DIU), "Page N of M" headers, pages packed to the 1004 DIU budget.
- Ground-truth scan boxes read the composed XPS page order - never re-derived (the 10/13-14 capture-label bug, section 6.4).
- Corrected clean criterion in every chart (frontier-slices and heatmaps, 20261017-v3).

Erratum (carried from v4) - the decode-slot accumulation error

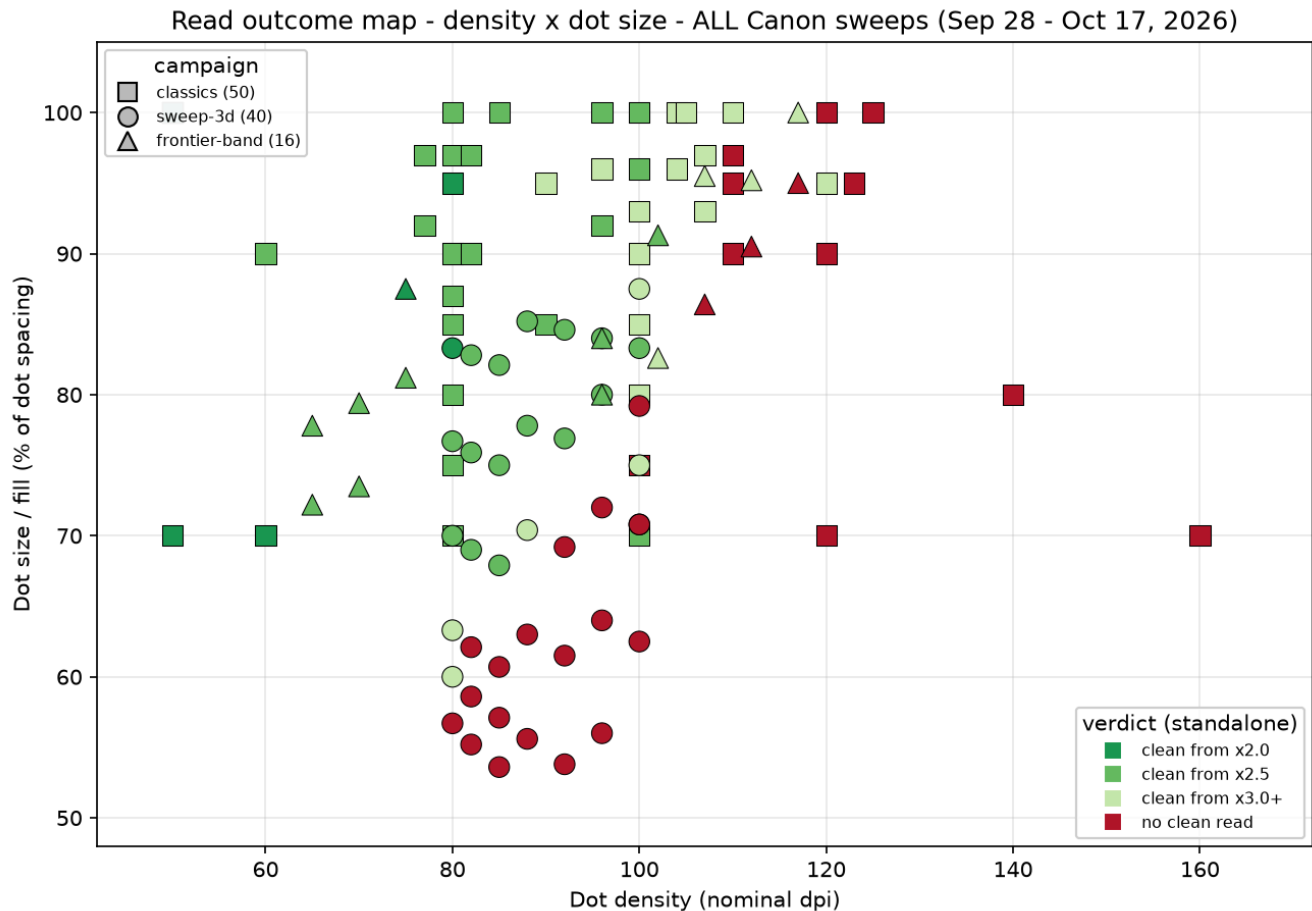
During HP-sweep work (2026-10-02) we found that PaperBack accumulates goodblocks/badblocks/restoredbytes across decodes of the same payload into the same decode slot. All 600 Canon trials were re-decoded standalone on 2026-10-02; one verdict changed (candidate 17: FAIL -> CLEAN). Sections 3-5 of the classic set are re-derived from that standalone re-audit.

Correction (new in v5) - the capture-label bug, 2026-10-13/14

The scan-box generator for HP sweep-2b and the Canon frontier-band re-implemented page packing instead of reading the composed print order, so every capture folder on three pages carried the wrong candidate label. Detected by block-count fingerprints vs. true pitch and non-monotonic floors; fixed by renaming from the printed XPS truth and re-decoding everything. All HP-2b and frontier-band tables in this edition are the corrected ones; the rule is now part of the method.

3 | The big picture: which settings read, which fail (all Canon sweeps)

Each marker below is one Canon candidate across all five Canon campaigns: horizontal = dot density, vertical = dot size (fill). Marker shape gives the campaign (squares = the September classics, circles = October sweep-3d, triangles = the October frontier-band); green shades give the floor (dark = clean already from x2.0, mid = from x2.5, light = only from x3.0 or higher), red = no clean standalone read. The frontier still slopes up to the right - finer dot grids need larger dots - and the density ceiling near 120 dpi is unchanged; the October sweeps add the fill dimension inside the 80-120 dpi window that the classics only probed at its edges.

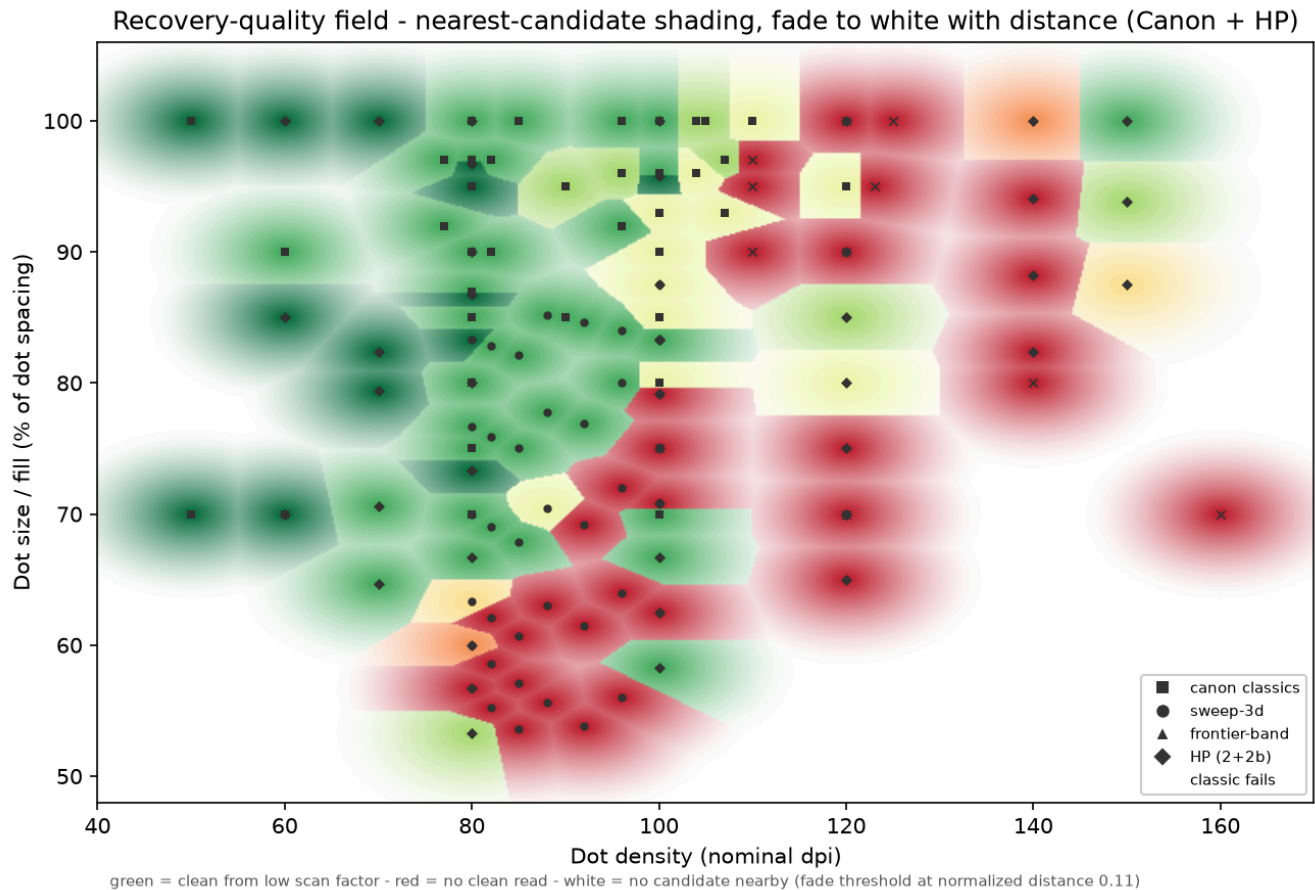


Read outcome map - density x dot size - all Canon sweeps (Sep 28 - Oct 17): 106 candidates = 50 classics, sweep-3d (40), frontier-band (16). The full point set ships as outcome-map-canon-all-20261019.csv for Excel re-plotting; the September-only historical version remains in the v4 edition.

An Excel-native companion ships alongside: outcome-map-canon-all-20261019.xlsx carries the same 106 points as a live, formatted scatter chart (data sheet + chartdata series columns + README describing the refresh workflow), so the map can now be maintained in Excel exactly like the September-30 original. The historical September-only version remains in the v4 edition.

3.1 | Recovery-quality field - nearest-candidate shading (new in v5)

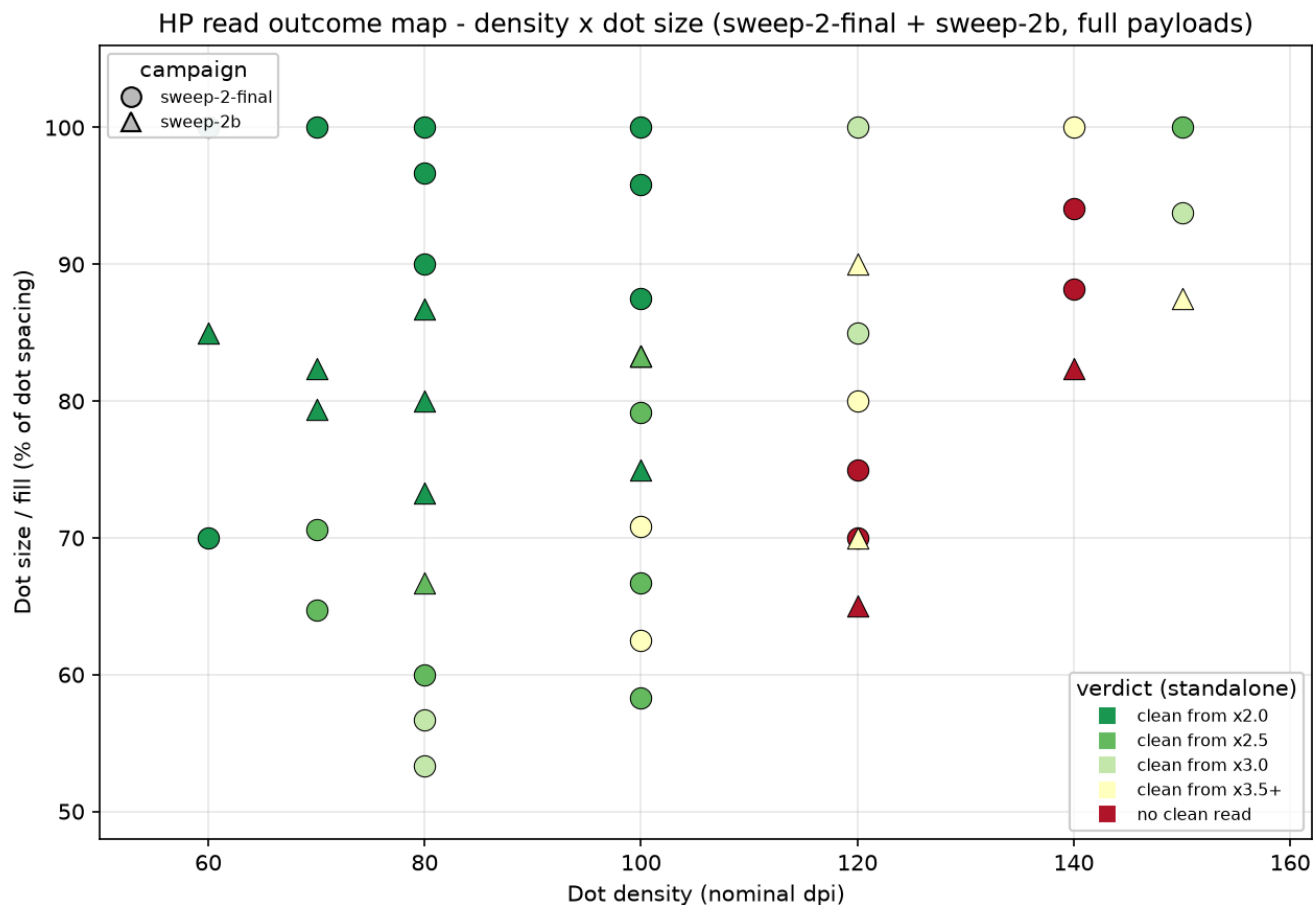
Owner idea, realized: instead of shading symbols, every pixel of the chart region takes the color of the NEAREST data vertex (nearest-neighbor shading on normalized axes), so the recovery landscape reads as a continuous field. Beyond a distance threshold the shade fades toward white - no fill where no candidate is nearby. Color encodes verdict quality (green = clean from a low scan factor, through yellow and red, to no clean read); small dark markers show the actual candidates and their campaign.



Recovery-quality field, Canon + HP combined (135 candidates). The white margins are honest emptiness - no sweep data there. The green plateau at 60-80 dpi / $\geq 85\%$ fill and the fine-pitch cliff at 120+ dpi read directly as regions rather than points.

3.2 | The HP big picture

The same view for the HP LaserJet 4350dtn, full-payload era: sweep-2-final (29 bands) and the corrected sweep-2b (16 bands). The frontier mirrors the Canon shape - coarse pitch reads from x2.0 at $\geq 85\%$ fill; P24 spans 66.7-100% fill; P20 needs $\geq 85\%$ fill and x3.0+; P17 reads only at 100% fill (x5.0). (Sweep-1, whose scans were lost, is superseded by these re-measurements; its decode data remains in the appendix sources.)



HP read outcome map: 45 candidates, full payloads, same shading grammar as the Canon map (green shades = floor, red = no clean standalone read).

4 | Master results - every classic candidate that reads cleanly (corrected)

How to read this table

Capacity (blocks) = usable data-block cells in the band as printed (~90 payload bytes per block before 1:5 redundancy).

Reads-from = smallest scan factor with a standalone clean raw decode. Best ECC = fewest corrected bytes at any clean factor.

Values: standalone re-decode set (2026-10-02).

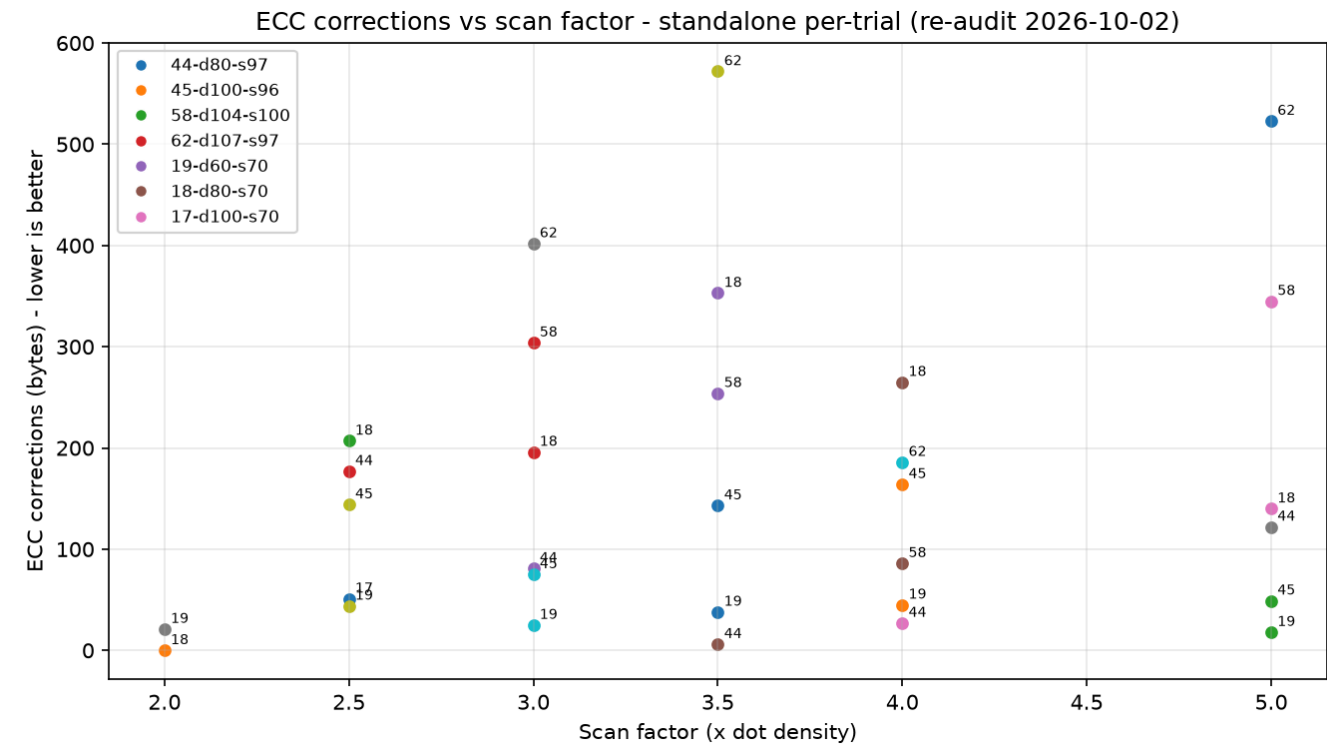
ID	dens	size	cap	reads-from	bestECC	note
20	50	70	27	x2.0	4	
65	50	100	27	x2.0	0	bulletproof fallback
19	60	70	33	x2.0	18	first success (classic)
27	60	90	33	x2.5	23	
52	77	92	42	x2.5	9	
53	77	97	42	x2.5	12	
18	80	70	45	x2.5	140	
33	80	75	45	x2.5	197	
34	80	80	45	x2.5	67	
35	80	85	45	x2.5	76	
43	80	87	45	x2.5	43	
25	80	90	45	x2.5	9	sweep2 champion (90%)
41	80	90	45	x2.5	64	
36	80	95	45	x2.0	23	
44	80	97	45	x2.5	6	record-low corrections
26	80	100	45	x2.5	25	
50	82	90	45	x2.5	29	
51	82	97	45	x2.5	7	
46	85	100	48	x2.5	40	
37	90	85	51	x2.5	295	
38	90	95	51	x3.0	100	
57	96	92	54	x2.5	219	
56	96	96	54	x3.0	114	
55	96	100	54	x2.5	82	
17	100	70	57	x2.5	51	NEW clean (was fail in v3)
29	100	80	57	x3.5	171	
30	100	85	57	x3.5	542	
23	100	90	57	x3.5	120	
54	100	93	57	x3.5	238	
45	100	96	57	x2.5	49	Zone B efficiency pick
24	100	100	57	x2.5	61	
42	100	100	57	x2.5	84	
59	104	96	60	x3.5	250	
58	104	100	60	x3.0	86	Zone B capacity pick
47	105	100	63	x3.0	256	
63	107	93	63	x3.5	492	
62	107	97	63	x3.0	186	Zone C pick
40	110	100	66	x3.5	450	
31	120	95	69	x3.5	405	

Classic candidates that never decoded cleanly (standalone)

ID	dens	size	closest standalone attempt
16	120	70	x2 conv: 69/69 unreadable (Page label)
21	160	70	x3 conv: 92/92 unreadable (Page label)
22	120	90	x2 conv: 69/69 unreadable (Page label)
28	100	75	x3.5 conv: 18/57 unreadable
32	120	100	x3.5 conv: 1/69 unreadable (complete)
39	110	90	x3.5 conv: 57/66 unreadable
48	125	100	x3.5 conv: 49/72 unreadable
60	110	97	x5 conv: 20/66 unreadable
61	110	95	x5 conv: 20/66 unreadable
64	123	95	x4 conv: 70/72 unreadable
66	140	80	x3 conv: 81/81 unreadable (Page label)

5 | Closer look: margins and failure modes

ECC corrections vs scan factor (standalone per-trial decodes): the reliable picks sit low and flat; density-edge candidates carry more corrections and more scan-to-scan variation. Candidate 17 - published as a failure in v3 - appears here as a clean x2.5 reader, tightening the 100-dpi fill frontier down to 70%.



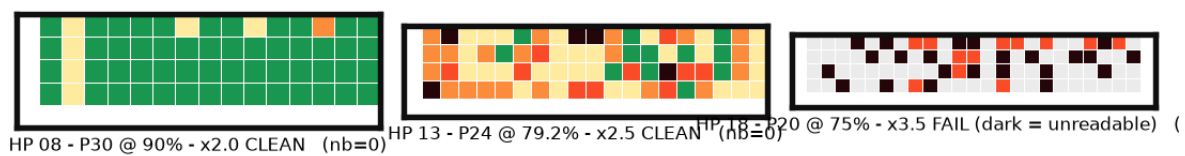
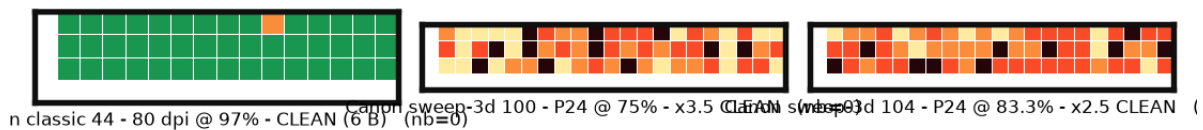
ECC corrections vs scan factor - canon classics (Canon iR-ADV DX CS840i), standalone per-trial (re-audit 2026-10-02).

The per-block quality maps - the decoder's own view of each page - follow in section 5.1; standout print-vs-scan cases from every campaign follow in 5.2.

5.1 | Per-block quality maps - the decoder's own view

For every decoded trial the control surface exports the quality map: one cell per block, colored by how many bytes ECC had to correct (0 = perfect block = green, yellow-to-red = corrections, dark = unreadable). Cells outside the data window render white - the software's page window - and the border is a fine frame, not a band. All six maps below are rendered from the raw quality-map exports with one renderer, so the printers are directly comparable.

Quality maps - the decoder's own view (green = clean, heat = bytes corrected, dark = unreadable, white = outside data window)

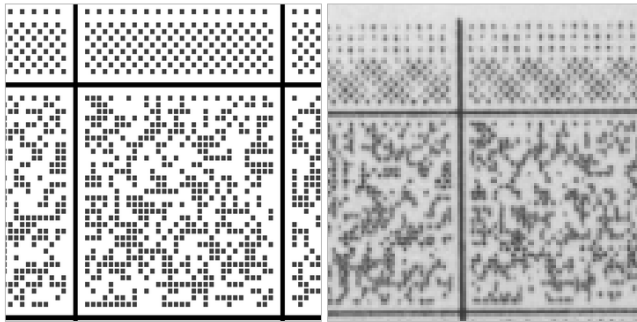


Top (canon classics + sweep-3d, Canon iR-ADV DX CS840i): classic 44 (80 dpi @ 97%, record-low 6 B), sweep-3d 100 (P24 @ 75%, x3.5), sweep-3d 104 (P24 @ 83.3%, x2.5). Bottom (HP sweep-2-final, HP LaserJet 4350dtn): 08 (P30 @ 90%, x2.0 clean), 13 (P24 @ 79.2%, x2.5 clean), 18 (P20 @ 75%, x3.5 FAIL - dark cells = unreadable blocks). The same exports drive the heatmaps in section 8.1.

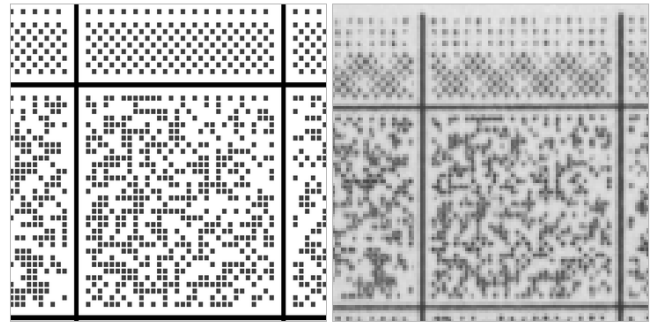
5.2 | Case standouts - block-level print vs. scan

Workflow for this edition: (1) standout case per sweep; (2) print bitmap + ground-truth-aligned scan crop; (3) a JSON mapping of every data block's position on BOTH images (gutter-phase self-calibration on the scan; print mapped by physical scale - cases-blockmap-20261019.json); (4) visually interesting scan blocks picked from the trial's quality map (unreadable / max-corrections / clean control); (5) pixel-level render: the SAME block from print and scan at IDENTICAL display size, scan upscaled nearest-pixel so scanner artifacts are directly observable. Two block pairs per case, two cases per page. Canon classics and frontier-band lack aligned sources (lost raw scans / no raw exports) and are represented by tables, maps and quality maps instead.

block (1,0) qm=17 PRINT | SCAN (same block, same size)

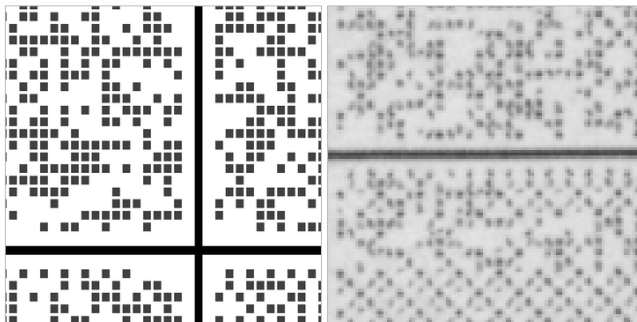


block (4,0) qm=16 PRINT | SCAN (same block, same size)

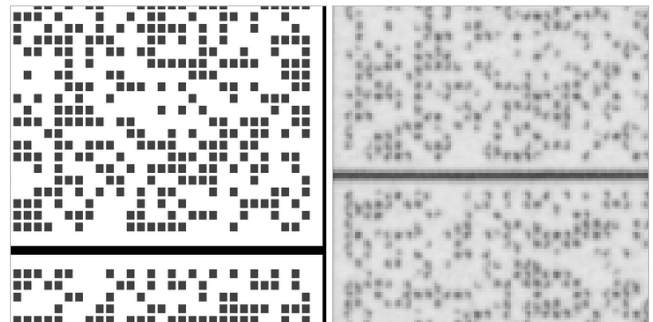


HP sweep-2-final 18-P20-D15 (120 dpi, 75% fill) @ x3.5 - block (1,0) UNREADABLE and block (4,0) 16 B corrected: dot rows survive but modulation is marginal - the fine-pitch failure mode (HP LaserJet 4350dtn).

block (18,2) qm=1 PRINT | SCAN (same block, same size)

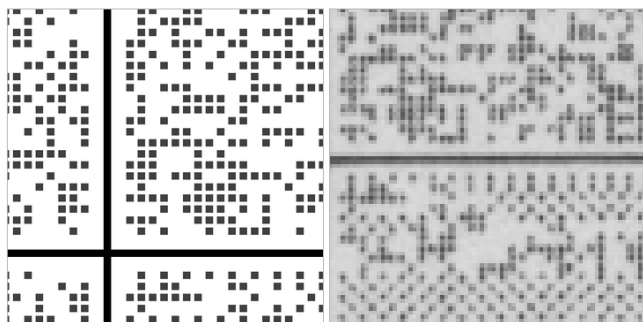


block (1,2) qm=1 PRINT | SCAN (same block, same size)

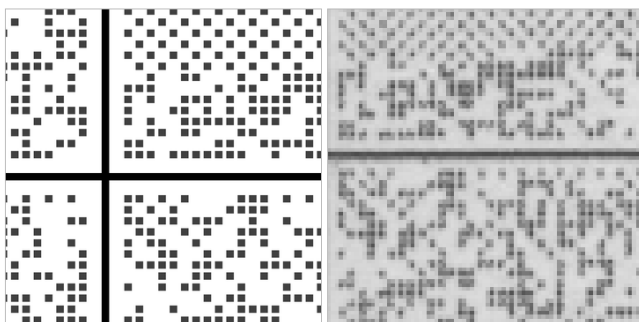


Canon sweep-3d 100-d100-s75 (100 dpi, 75% fill) @ x3.5 - low-correction blocks (1 B each): clean decode; note halftone texture and dot gain vs. the crisp render (Canon iR-ADV DX CS840i).

block (10,3) qm=0 PRINT|SCAN (same block, same size)

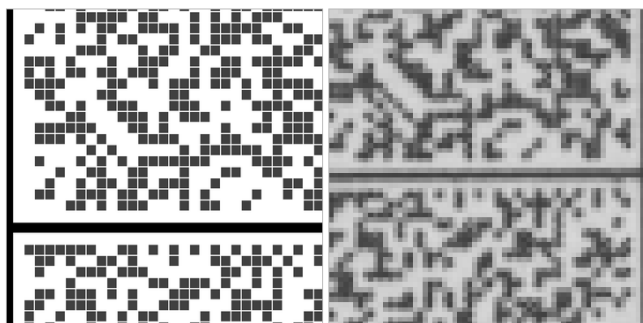


block (10,1) qm=0 PRINT|SCAN (same block, same size)

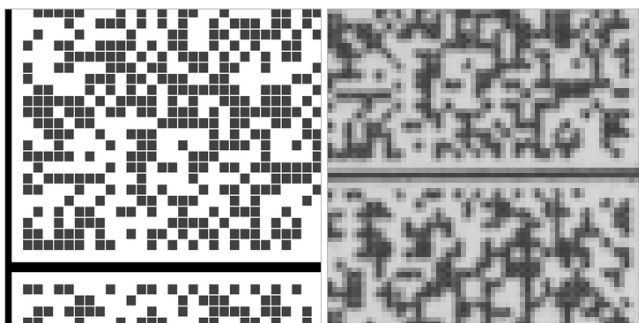


HP sweep-2b B16-P30-D20 (80 dpi, 66.7% fill) @ x2.5 - clean control blocks at the lowest clean P30 fill measured on HP (HP LaserJet 4350dtn).

block (8,2) qm=0 PRINT|SCAN (same block, same size)



block (8,3) qm=0 PRINT|SCAN (same block, same size)



HP sweep-2-final 08-P30-D27 (80 dpi, 90% fill) @ x2.0 - clean control blocks; the reliability pick (HP LaserJet 4350dtn).

6 | The HP LaserJet 4350dtn (extended in v5)

6.1 Sweep-1 (2026-10-02) - recap and status

The 24-band intensity sweep established the first HP floors; its raw scans were lost in a station cleanup - the decode CSVs, manifests and standalone verification survive and remain valid. Sweep-2-final + sweep-2b (below) re-measure every family with full payloads and supersede sweep-1 for floor claims.

6.2 Sweep-2-final - 29 full-payload bands (2026-10-12)

Six-page sheet, job B224555, per-family payloads sized to eliminate crosshair filler; 166 scans / 166 standalone decode trials. Floor = lowest factor with nb = 0.

#	Band	dpi	fill%	floor	#	Band	dpi	fill%	floor
01	P40-D28	60	70.0%	x2.0	16	P24-D24	100	100%	x2.0
02	P40-D40	60	100%	x2.0	17	P20-D14	120	70.0%	FAIL
03	P34-D24	70	70.6%	x2.5	18	P20-D15	120	75.0%	FAIL
04	P34-D34	70	100%	x2.0	19	P20-D17	120	85.0%	x3.0
05	P30-D16	80	53.3%	x3.0	20	P20-D20	120	100%	x3.0
06	P30-D17	80	56.7%	x3.0	21	P17-D15	140	88.2%	FAIL
07	P30-D18	80	60.0%	x2.5	22	P17-D16	140	94.1%	FAIL
08	P30-D27	80	90.0%	x2.0	23	P17-D17	140	100%	x5.0
09	P30-D29	80	96.7%	x2.0	24	P16-D16	150	100%	x2.5
10	P30-D30	80	100%	x2.0	25	P34-D22	70	64.7%	x2.5
11	P24-D14	100	58.3%	x2.5	26	P24-D16	100	66.7%	x2.5
12	P24-D15	100	62.5%	x3.5	27	P24-D17	100	70.8%	x3.5
13	P24-D19	100	79.2%	x2.5	28	P20-D16	120	80.0%	x3.5
14	P24-D21	100	87.5%	x2.0	29	P16-D15	150	93.8%	x3.0
15	P24-D23	100	95.8%	x2.0					

HP picks (updated by sweep-2/2b)

Reliability: P30-D27 @ 80 dpi (90% fill) - clean at every factor incl. x2.0.

Capacity: P24-D19 @ 100 dpi (79.2%) - x2.5 floor; or P24-D23 (95.8%) clean from x2.0.

Max density: P20-D20 @ 120 dpi (100%) - x3.0 floor (full-payload re-measure improved the v4 x4.0 claim). P17 @ 140 dpi remains the practical ceiling (100% fill, x5.0).

Avoid: P20 <=75% fill, P17 <100% fill, and any band below ~66% fill at fine pitch.

6.3 Sweep-2b - boundary refinement + P24 controls (2026-10-13)

16 candidates reusing the hp2 family payloads for comparability, plus P24 control twins. Corrected floors (post label-fix):

Band	dpi	fill%	floor	note
B01-P40-D34	60	85.0%	x2.0	
B03-P34-D28	70	82.4%	x2.0	
B02-P34-D27	70	79.4%	x2.0	
B16-P30-D20	80	66.7%	x2.5	P30 floor ~66.7% (down from 73%)
B04-P30-D22	80	73.3%	x2.0	
B06-P30-D26	80	86.7%	x2.0	
B05-P30-D24	80	80.0%	x2.0	
B14-P24-D20	100	83.3%	x2.0	control twin of B08
B08-P24-D20	100	83.3%	x2.0	clean x2.0-5.0 incl. x3.0
B07-P24-D18	100	75.0%	x2.0	75% threshold clean incl. x3.0
B15-P24-D20	100	83.3%	x2.5	x2.0 = 1 bad block
B11-P20-D18	120	90.0%	x3.5	
B09-P20-D13	120	65.0%	FAIL	below floor
B10-P20-D14	120	70.0%	x4.0	
B12-P17-D14	140	82.4%	FAIL	P17 needs higher fill
B13-P16-D14	150	87.5%	x4.0	x3.0 one-block near-miss

Cross-checks that matter: the B08/B14 twins agree (P24 @ 83.3% clean everywhere); B15's single x2.0 bad block shows the run-to-run edge noise; and HP agrees with Canon on the P24 75% fill threshold - band B07 (75.0%) is clean at every factor.

6.4 The capture-label bug, found and fixed (2026-10-13/14)

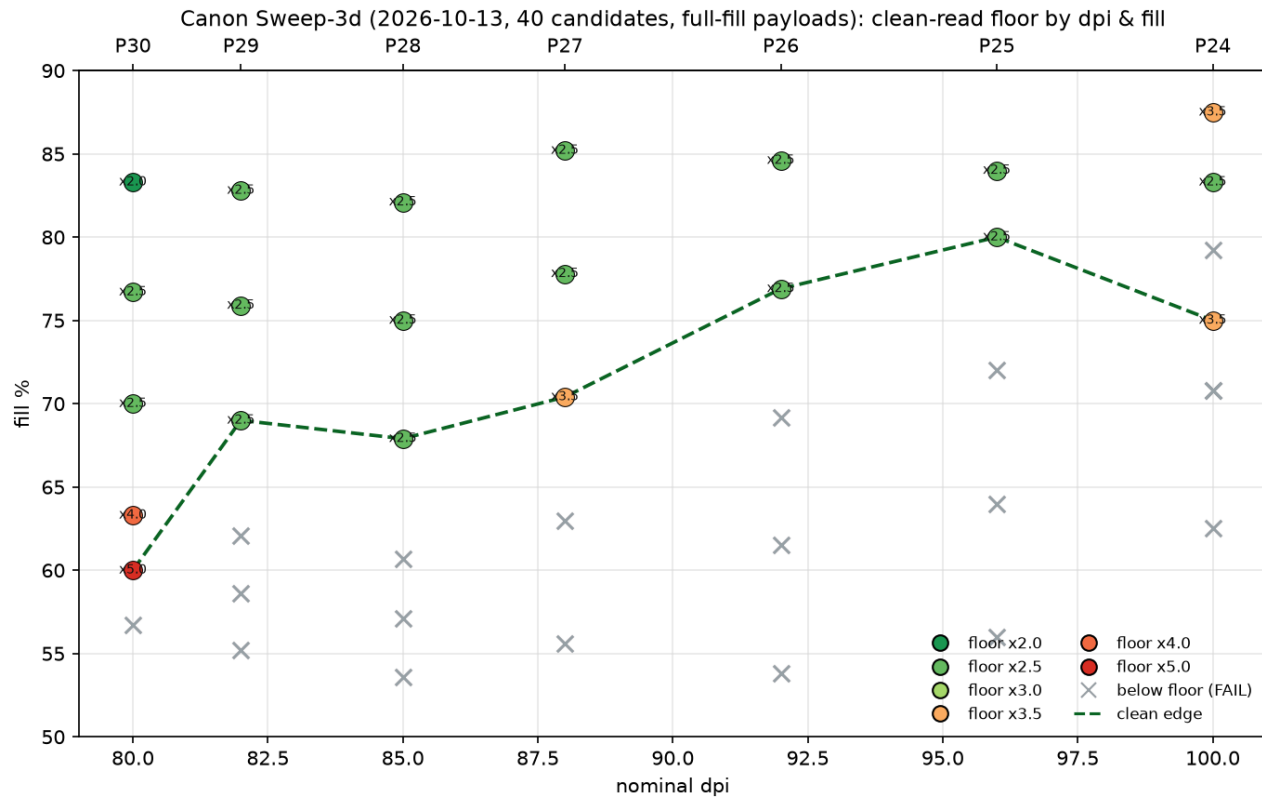
The scan-box generator for HP-2b and the Canon frontier-band re-implemented page packing instead of reading the composed print order - every capture folder on HP pages 1/2/4 and all three frontier-band pages carried wrong candidate labels. Detection: nblock fingerprints vs. true pitch plus non-monotonic floors. Fix: true page order taken from the printed XPS, trees renamed, all decodes re-processed; earlier sweeps verified unaffected (sequential compose). Rule added to the method: box-gen reads the composed XPS, never re-derives packing.

7 | Canon, October: the fill-floor map and the frontier edge

7.1 Sweep-3d - fill floors at full payloads, P30 -> P24 (2026-10-13)

40 candidates (ids 68-107): 7 pitches (P30 d80 -> P24 d100) x 5-7 fill levels + a same-render control pair (99/103) + ladder hole 104; 240 standalone decode trials. 21/40 clean on the standard ladder:

floor	candidates
x2.0	72
x2.5	70, 71, 75, 76, 77, 80, 81, 82, 86, 87, 91, 92, 96, 97, 104
x3.5	85, 100, 102
x4.0	69
x5.0	105
FAIL	68, 73, 74, 78, 79, 83, 84, 88, 89, 90, 93, 94, 95, 98, 99, 103, 106, 107



Fill-floor curve, P30 -> P24: minimum fill for a clean standalone read. P30 reads from ~66.7% fill; P24 reads from 75% fill (the full factor ladder from ~79%).

Control pair 99/103 (identical render, 70.8% fill) fails identically - same-render determinism confirmed; the d100 window remains the known noisy neighborhood (section 8.2).

7.2 Canon frontier-band - bracketing the rising edge (2026-10-13/14)

16 candidates at 65-117 dpi, 72.2-100% fill, chosen on the rising pass/fail edge between the coarse classic map and the d100 window. Corrected results (post label-fix):

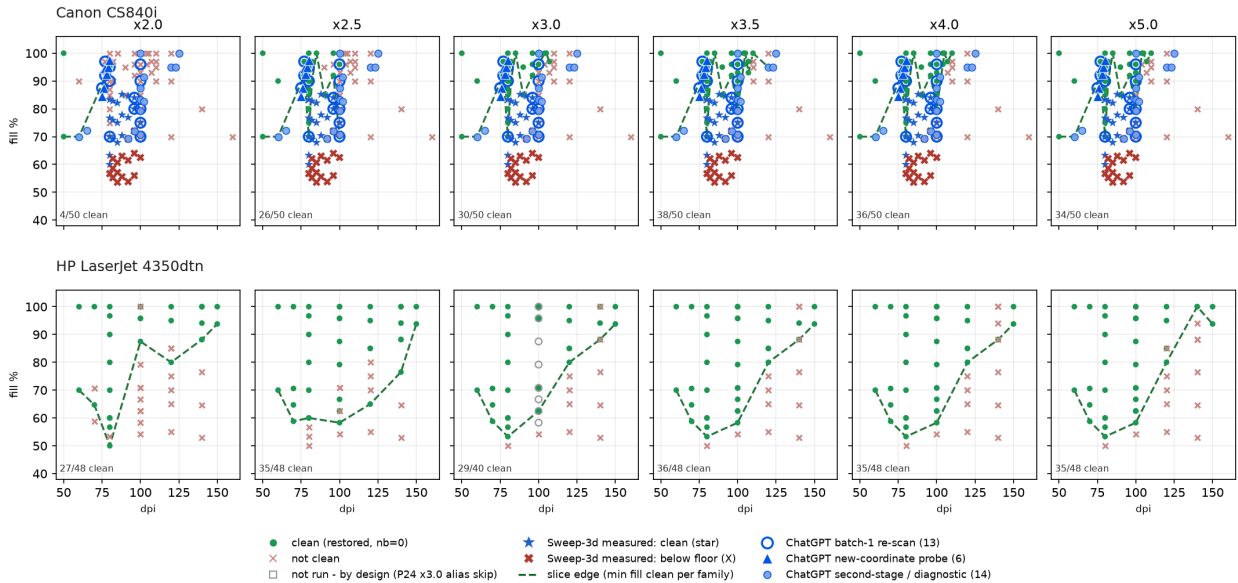
Band	dpi	fill%	floor	note
C01-d65-s73	65	72.2%	x2.5	
C02-d65-s78	65	77.8%	x2.5	
C03-d70-s74	70	73.5%	x2.5	
C04-d70-s80	70	79.4%	x2.5	
C05-d75-s82	75	81.2%	x2.5	
C06-d75-s88	75	87.5%	x2.0	cleanest coarse point
C07-d96-s80	96	80.0%	x2.5	
C08-d96-s84	96	84.0%	x2.5	
C09-d102-s83	102	82.6%	x3.5	
C10-d102-s92	102	91.3%	x2.5	
C11-d107-s87	107	86.4%	FAIL	
C12-d107-s96	107	95.5%	x3.0	
C13-d112-s91	112	90.5%	FAIL	
C14-d112-s96	112	95.2%	x3.0	re-scanned; corrected
C15-d117-s95	117	95.0%	FAIL	
C16-d117-s100	117	100.0%	x3.5	

Readings: the coarse knee (65-75 dpi) sits near 72% fill; 96 dpi behaves like a second coarse band (x2.5 floors at 80-84% fill); 102-117 dpi is threshold territory where x3.5 and band-specific fine structure decide. Non-monotonic x3.0 behavior at fine pitch matches the alias findings (7.2).

8 | Cross-printer: Canon vs. HP on the same question

With per-family payloads now used on both printers, comparisons are apples-to-apples within a payload generation. HP renders fine pitch slightly softer at the edge (P20 needs x3.5-x4.0 where Canon's 120 dpi @ 95% reads x3.5) but equals or beats Canon at P24-P30. Both printers agree on: fill as the master lever at coarse pitch; the P24 75% threshold; near-saturation modulation collapse; and the coarse knee near ~72% fill at 65-75 dpi.

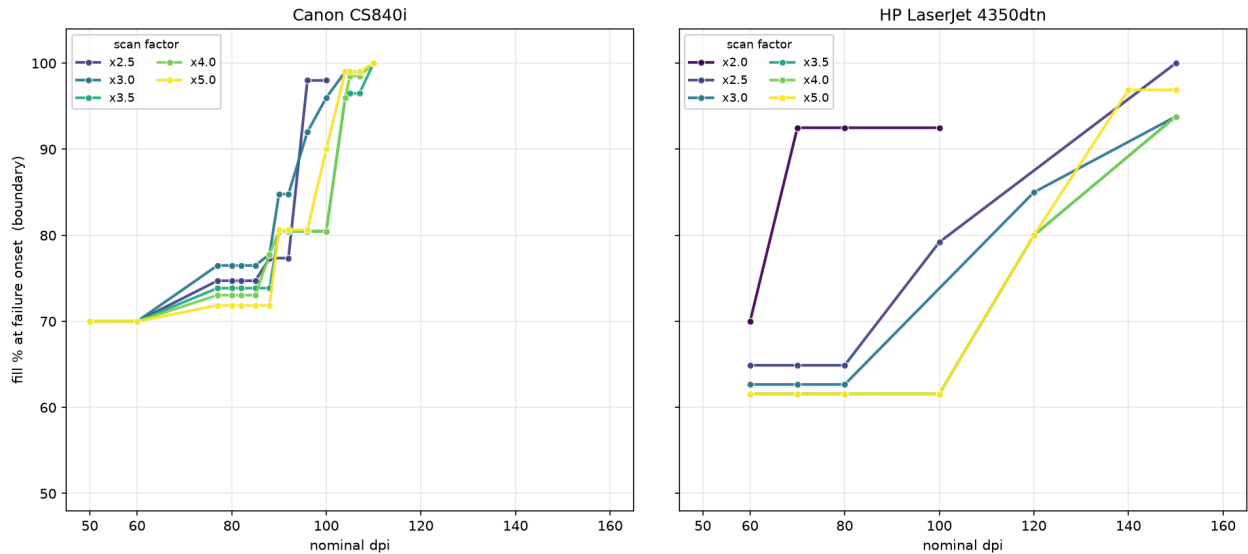
Frontier slices by scan factor - clean standalone read per tuple + ChatGPT recommended tuples (blue, Canon); dashed line = clean-read edge per slice



Canon = conv variant (clean requires successful restore); HP = sweep-2 primary + sweep-1 where sweep-2 absent. Blue = tuples recommended in "ChatGPT - Recommend PaperBack Settings.md" (2026-10-15), Canon only. Chart 2026-10-17 v3 (v2.2 base + ChatGPT recommendations).

Frontier slices, final (20261017-v3): per-band decode outcome across scan factors, both printers, corrected clean criterion - the October sweeps merged with the corrected classics. This is the report's hero chart.

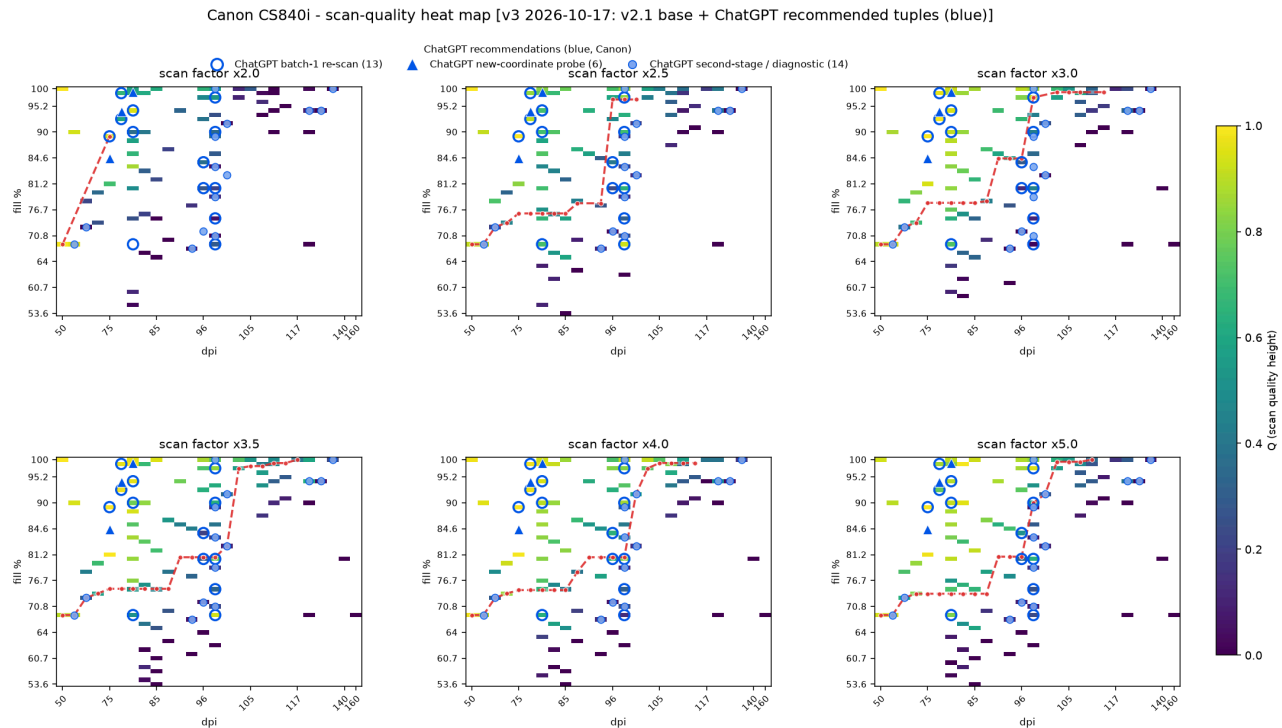
Pass/fail boundary spread - all 6 scan-factor boundary curves overlaid per printer
(boundary = lowest fill% in the top-anchored clean run, isotonic-monotone; rise = more fill needed before failures stop)



Curves higher/later to the right = harder points. Vertical spread between a printer's curves = how much scan factor moves the boundary. Source: canon-hp-quality-metrics-20261013.csv.

Pass/fail boundary spread: six outcome curves per panel, showing how sharply each family transitions from clean to fail as fill drops.

8.1 Quality heatmaps (corrected criterion)



Canon quality heatmap, final (20261017-v3): per-candidate decode quality across the (dpi, fill) plane - same corrected criterion as the frontier slices. The HP heatmap companion is produced by the same script family and ships with the site edition of this report.

8.2 Era comparability - why old d100 "failures" were retired

Canon-old d100 verdicts (309-B legacy payload, 57 blocks) conflict with sweep-3d at matched (dpi, fill): 28-d100-s75 FAIL vs. 100-d100-s75 CLEAN from x3.5. The discrepancy analysis concludes this is a comparison artifact, not a device change: identical fill % with different payload/block structure is not an identical sample, and the d100 window is intrinsically low-repeatability (adjacent-factor swings like nb = 43/1/10/0/0/0). The HP control (sweep-1 vs. sweep-2, 13 matched cells) flipped zero verdicts - its matched coordinates sat far from thresholds. Standing rule: supersede canon-old d100 verdicts with sweep-3d; never compare "same fill %" across payload generations without carrying payload/block metadata (now a column in the quality-metrics CSVs).

9 | Capacity math and redundancy

- Worked example (both printers agree): P24 @ 100 dpi, 100% fill = 57 usable data blocks, about 5.1 KB/page before redundancy; at 79.2% fill the same pitch carries 45 blocks (~4.0 KB - the sweep-3d P24 payload size).
- P30 @ 80 dpi, >=90% fill: ~3.1 KB with x2.0 floors - the reliability choice.
- P20/P17 gains shrink the page margin to nothing (labels + borders consume the budget) and demand x3.5-x5.0 scans; P16 @ 150 dpi reads only near 100% fill. The density ceiling near 120 dpi stands.
- Redundancy 1:5 is the printed geometry; content-addressed blocks make the copy distribution deterministic (redundancy sweep of 9/30 remains the reference).
- ctl v4 pinned compression semantics: None + bzip2 levels 1-9, level = exact block size (the old default 2 meant "Maximal" under the previous scale).

10 | Tooling - the ctl control binary at v8.2

Every sweep in this report was driven through a JSON-over-named-pipe control surface built into the rebuilt decoder: requests are marshaled onto the UI thread, so automation cannot diverge from interactive behavior; the qualitymap export (per-cell ECC correction counts) feeds the heatmaps in section 8.1.

Family	Version	Purpose
MinGW/TDM portability	ctl v2 base	Borland-era sources compile clean with TDM-GCC 10.3
JSON/named-pipe control	v2 (09-29)	headless driving; UI-thread marshaling
Render-accurate options	v3	dropdowns match real print metrics
Compression semantics	v4	None + bzip2 1-9, level = exact block size
Custom geometry / packing	v5/v6	geom.plan/geom.set; full-capacity pages
Stop/discard/margins	v7	operational safety for long runs
TWAIN control + headless	v8/v8.1	named sources; dpi/pixeltype/frame/raw-DIB capture
Robustness	v8.2	120 s dispatch budget; pure-match scan.source; graceful idempotent scan.close (fixes an upstream DSM-close bug); self-closing headless transfers

Reproducibility proof (2026-10-17)

The v8.2 binary (760,868 B, 2026-10-06) was rebuilt clean-room from durable sources: pristine 1.10 tree + full-file patch (17 files) + pinned aes/bzip2/sha objects, TDM-GCC 10.3. Result: identical 760,868-byte size; hash differs only in the PE link timestamp (non-deterministic linker), pei-i386 confirmed. Build kit, guide, patch notes and logs: Documents\PaperBack Build\ . Upstream: PaperBack 1.10 by Oleh Yuschuk - the ctl line is a fork of his work; credit accompanies every distribution.

11 | Appendix - data index, corrections, provenance

Primary sources (durable archive E:\PaperBack\PaperBak-Sweeps-Archive\): hp-sweep2-final (results\scan-20261012: ALL-PAGES-FLOORS.md, sweep2-decode-all-20261012.csv, manifests, ground-truth boxes); hp-sweep2b (notes: h2b-ALL-RESULTS.md, per-page decode logs); canon-sweep3d (manifest v3, floors/decode CSVs); canon-frontierband (cfb-ALL-RESULTS.md, label-fix record); canon-vs-hp-frontier (quality-metrics v2.1 CSV - 779 corrected trials, boundary spread, discrepancy analysis). Chart finals: PaperBack\chart-updates-20261017 (frontier-slices v3, canon heatmap v3, recommended-tuples CSV). Case standouts and quality maps for this edition: reportwork-v5 figures\basemap-20261019.png + qmaps-both-v2-20261019.png; recovery-quality field + HP outcome map: figures\quality-field-20261019.png, outcome-map-hp-20261019.png; Excel pipeline: outcome-map-canon-all-20261019.xlsx (data + live chart + workflow README).

Correction provenance: classic set re-audited standalone 2026-10-02 (one verdict change); capture-label bug corrected 2026-10-14 (trees renamed from printed-XPS truth, re-decoded); quality-metrics v2.0 stamped CORRUPTED DATA in-place, superseded by v2.1 (clean copies retained).

Archive integrity: PaperBak-Sweeps-Archive v2 reorganization verified by SHA256 pre-burn audit (9,446/9,446 matched, 2026-10-15). BD-R archival burn deferred by owner gate (≥ 120 uninterrupted minutes required); the durable USB mirror is the current archival medium.

Next steps: HP quality heatmap final + HP GUI-quality chart (same script family as the Canon v3 finals); the 37-tuple recommended rescan batch; optional C14 re-scan.

Compiled 2026-10-19 17:40 | generator: make_report_v5.py (lineage: make_report.py v2 -> v3 -> v4 -> this) | text source: paperback-sweep-report-20261019-v5.md