Supplementary material: Evaluating Neural Network Repairs in ACAS Xu: Static Fidelity and Off-Target Encounter Effects

Vimuth Fonseka and Ronnie Marlon Abeykoon

CEAS Aeronautical Journal | Corresponding author: Vimuth Fonseka, Independent researcher, Colombo, Sri Lanka; vimuth_fonseka@hotmail.com

Online Resource 1 accompanies the CEAS Aeronautical Journal manuscript. It supplies the complete candidate counts and version-specific reproduction inventory. Sections S1–S9 preserve the original and broader-study records; Section S10 adds the executed full-cohort stopping control and command-delay sensitivity. PASS (P), FAIL (F) and UNKNOWN (U) retain the manuscript's numerical meanings. Whole-population transitions include original nonexposed outcomes and must not be mistaken for eligible conditional denominators.

S1. Scope, estimands and artefact identity

There are twelve unchanged checkpoints: three configurations of each of N2,1 through N5,1. A candidate replaces one controller component and retains the other forty-four. Baseline-defined exposure E_n fixes each network's inclusion set. Regression concerns original PASS → FAIL; recovery concerns original FAIL → PASS. Known-failure change counts all transitions, including F → U and U → F, so it is not universally regressions minus recoveries. Target-domain exposure is zero and does not yield a target-effect estimate.

Specrepair commit 261eacc661ba8128a3e2fc576c42bfd90a1ea33f and simulator commit 502eddbd7269f5f84c7d6d3387135980bf9997ad control executable identity. Software versions, checkpoint selection and executable configurations are documented in the Zenodo Python/data package identified in S9. Standalone Python reproduction files, plain data/results and figures are publicly available in Zenodo, version 1.0.0, at https://doi.org/10.5281/zenodo.23202447. The additional supporting archive supplies numerical data and checks. Full trajectory databases require regeneration using the documented programs.

The original 278,528 draws in four populations, later 28,781-case extension and 238-case replication remain separate. The detailed original N3 investigation is post hoc. The broader protocol fixes the same comparison after the original results, before new outcomes; it is not public preregistration or a blind new hypothesis. The stopping-rule and response-delay protocol was fixed locally on 9 October 2026 before its additional simulations; it reuses existing inputs and is not an external preregistration.

S2. Detailed published/reconstructed static measures and original anchors

The agreement and MAE values retain the source table's displayed precision. Exact stored MAE values are also given in S5. Controller-path N3 agreement is a distinct measurement in manuscript Table 4.

Table S1. Complete static measures and original-anchor outcomes.

Network	Repair	Equal-point static agreement	Published MAE	MAE rank	Anchor recoveries	Anchor new failures	Known net failure change
N2,1	optimal	99.573447963%	0.098801196194	1	137	130	−7
N2,1	early_exit	99.363402401%	0.151783243238	3	160	137	−23
N2,1	deep_opt_early_exit	99.595999435%	0.108055449204	2	154	136	−18
N3,1	optimal	98.996785115%	0.216561075842	3	39	107	+68
N3,1	early_exit	99.309685080%	0.139732146761	1	54	30	−24
N3,1	deep_opt_early_exit	99.366928399%	0.171255599697	2	14	95	+81

Network	Repair	Equal-point static agreement	Published MAE	MAE rank	Anchor recoveries	Anchor new failures	Known net failure change
N4,1	optimal	99.831864135%	0.103448762235	3	12	2	−10
N4,1	early_exit	99.810543140%	0.071017688256	1	8	2	−6
N4,1	deep_opt_early_exit	99.806744784%	0.098856192168	2	7	2	−5
N5,1	optimal	99.654538967%	0.124275810484	2	1	0	−1
N5,1	early_exit	99.534254972%	0.125778607592	3	1	0	−1
N5,1	deep_opt_early_exit	99.627757064%	0.089642287250	1	1	0	−1

S3. All original population/candidate transitions: forty-eight whole-population and forty-eight exposed rows

The four populations below each retain twelve whole-population and twelve exposed-only rows, with every numerical state. Whole-population rows retain original nonexposed outcomes; exposed-only rows use their own denominators. Every row sums to its stated denominator. Conditional eligible PASS counts remain those of manuscript section 3.1 and Table S10.

Whole-population tables use the full population denominator; exposed-only tables use their stated inclusion counts. In the artefact labels, O, E and D denote optimal, early_exit and deep_opt_early_exit; the two following indices identify the network. Nonexposed cases remain on the diagonal only after the original-replacement and nonexposure controls pass. All nine PASS/FAIL/UNKNOWN cells are shown. A failure-difference range permits either binary classification of UNKNOWN; it is a numerical-classification envelope, not a confidence interval. Never subtract source regression and anchor recovery percentages with different populations or conditioning.

Source-law population

Table S2. Source-law whole-population transitions.

Artifact	Population n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U	Failure-difference count range
O·2·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
E·2·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
D·2·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
O·3·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
E·3·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
D·3·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
O·4·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
E·4·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
D·4·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]

Artifact	Population n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U	Failure-difference count range
O·5·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
E·5·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]
D·5·1	262144	262141	0	0	0	3	0	0	0	0	[0, 0]

Exposed-only paired cells, with separate denominators:

Table S3. Source-law exposed transitions.

Artifact	Exposed n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U
O·2·1	4221	4220	0	0	0	1	0	0	0	0
E·2·1	4221	4220	0	0	0	1	0	0	0	0
D·2·1	4221	4220	0	0	0	1	0	0	0	0
O·3·1	5504	5502	0	0	0	2	0	0	0	0
E·3·1	5504	5502	0	0	0	2	0	0	0	0
D·3·1	5504	5502	0	0	0	2	0	0	0	0
O·4·1	3534	3533	0	0	0	1	0	0	0	0
E·4·1	3534	3533	0	0	0	1	0	0	0	0
D·4·1	3534	3533	0	0	0	1	0	0	0	0
O·5·1	1291	1289	0	0	0	2	0	0	0	0
E·5·1	1291	1289	0	0	0	2	0	0	0	0
D·5·1	1291	1289	0	0	0	2	0	0	0	0

Anchor mixture

Table S4. Anchor whole-population transitions.

Artifact	Population n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U	Failure-difference count range
O·2·1	8192	2100	130	0	137	5824	0	0	0	1	[-8, -6]
E·2·1	8192	2093	137	0	160	5801	0	0	0	1	[-24, -22]

Artifact	Population n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U	Failure-difference count range
D·2·1	8192	2094	136	0	154	5807	0	0	0	1	[-19, -17]
O·3·1	8192	2123	107	0	39	5922	0	0	0	1	[67, 69]
E·3·1	8192	2200	30	0	54	5907	0	0	0	1	[-25, -23]
D·3·1	8192	2135	95	0	14	5947	0	0	0	1	[80, 82]
O·4·1	8192	2228	2	0	12	5949	0	0	0	1	[-11, -9]
E·4·1	8192	2228	2	0	8	5953	0	0	0	1	[-7, -5]
D·4·1	8192	2228	2	0	7	5954	0	0	0	1	[-6, -4]
O·5·1	8192	2230	0	0	1	5960	0	0	0	1	[-2, 0]
E·5·1	8192	2230	0	0	1	5960	0	0	0	1	[-2, 0]
D·5·1	8192	2230	0	0	1	5960	0	0	0	1	[-2, 0]

Exposed-only paired cells, with separate denominators:

Table S5. Anchor exposed transitions.

Artifact	Exposed n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U
O·2·1	5818	1608	130	0	137	3943	0	0	0	0
E·2·1	5818	1601	137	0	160	3920	0	0	0	0
D·2·1	5818	1602	136	0	154	3926	0	0	0	0
O·3·1	2637	554	107	0	39	1936	0	0	0	1
E·3·1	2637	631	30	0	54	1921	0	0	0	1
D·3·1	2637	566	95	0	14	1961	0	0	0	1
O·4·1	3860	1068	2	0	12	2778	0	0	0	0
E·4·1	3860	1068	2	0	8	2782	0	0	0	0
D·4·1	3860	1068	2	0	7	2783	0	0	0	0
O·5·1	4477	1305	0	0	1	3170	0	0	0	1
E·5·1	4477	1305	0	0	1	3170	0	0	0	1
D·5·1	4477	1305	0	0	1	3170	0	0	0	1

Target-speed challenge

Table S6. Target-speed whole-population transitions.

Artifact	Population n	P → P	P → F	P → U	F → P	F → F	F → U	U → P	U → F	U → U	Failure-difference count range
O·2·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
E·2·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
D·2·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
O·3·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
E·3·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
D·3·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
O·4·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
E·4·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
D·4·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
O·5·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
E·5·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]
D·5·1	4096	4096	0	0	0	0	0	0	0	0	[0, 0]

Exposed-only paired cells, with separate denominators:

Table S7. Target-speed exposed transitions.

Artifact	Exposed n	P → P	P → F	P → U	F → P	F → F	F → U	U → P	U → F	U → U
O·2·1	62	62	0	0	0	0	0	0	0	0
E·2·1	62	62	0	0	0	0	0	0	0	0
D·2·1	62	62	0	0	0	0	0	0	0	0
O·3·1	79	79	0	0	0	0	0	0	0	0
E·3·1	79	79	0	0	0	0	0	0	0	0
D·3·1	79	79	0	0	0	0	0	0	0	0
O·4·1	51	51	0	0	0	0	0	0	0	0
E·4·1	51	51	0	0	0	0	0	0	0	0
D·4·1	51	51	0	0	0	0	0	0	0	0

Artifact	Exposed n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U
O·5·1	16	16	0	0	0	0	0	0	0	0
E·5·1	16	16	0	0	0	0	0	0	0	0
D·5·1	16	16	0	0	0	0	0	0	0	0

Uncontrolled close-approach challenge

Table S8. Close-approach whole-population transitions.

Artifact	Population n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U	Failure-difference count range
O·2·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
E·2·1	4096	4091	1	0	0	4	0	0	0	0	[1, 1]
D·2·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
O·3·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
E·3·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
D·3·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
O·4·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
E·4·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
D·4·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
O·5·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
E·5·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]
D·5·1	4096	4092	0	0	0	4	0	0	0	0	[0, 0]

Exposed-only paired cells, with separate denominators:

Table S9. Close-approach exposed transitions.

Artifact	Exposed n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U
O·2·1	1979	1976	0	0	0	3	0	0	0	0
E·2·1	1979	1975	1	0	0	3	0	0	0	0
D·2·1	1979	1976	0	0	0	3	0	0	0	0

Artifact	Exposed n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U
O·3·1	2159	2157	0	0	0	2	0	0	0	0
E·3·1	2159	2157	0	0	0	2	0	0	0	0
D·3·1	2159	2157	0	0	0	2	0	0	0	0
O·4·1	1030	1029	0	0	0	1	0	0	0	0
E·4·1	1030	1029	0	0	0	1	0	0	0	0
D·4·1	1030	1029	0	0	0	1	0	0	0	0
O·5·1	539	536	0	0	0	3	0	0	0	0
E·5·1	539	536	0	0	0	3	0	0	0	0
D·5·1	539	536	0	0	0	3	0	0	0	0

These are family-adjusted research precision bounds; none is an operational acceptable failure rate.

Table S10. Source-law regression bounds.

Network, three repairs each	Regressions per repair / exposed baseline PASS	Family-adjusted one-sided upper bound
N2,1	0 / 4,220	0.129789%
N3,1	0 / 5,502	0.099562%
N4,1	0 / 3,533	0.155007%
N5,1	0 / 1,289	0.424283%

S4. Original N3 anchors and all fourteen families

Table S11. Original N3 nine-cell outcomes.

Repair	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U
optimal	2,123	107	0	39	5,922	0	0	0	1
early_exit	2,200	30	0	54	5,907	0	0	0	1
deep_opt_early_exit	2,135	95	0	14	5,947	0	0	0	1

Rec/reg means FAIL→PASS/PASS→FAIL. The remaining shared UNKNOWN is retained.

Table S12. Original N3 family contributions.

Seed	Draws	Exposed PASS	Exposed FAIL	Exposed UNKNOWN	Optimal rec/reg	Early rec/reg	Deep rec/reg
57305	590	0	0	0	0/0	0/0	0/0
174596	605	37	567	1	5/0	4/0	5/0
191434	608	2	0	0	0/0	0/0	0/0
207659	556	0	0	0	0/0	0/0	0/0
341255	573	12	0	0	0/0	0/0	0/0
468582	584	0	38	0	0/0	0/0	0/0
872977	576	53	523	0	5/0	4/0	2/0
891092	557	39	0	0	0/0	0/0	0/0
1084018	602	188	414	0	0/100	0/30	0/95
1182535	586	0	41	0	0/0	0/0	0/0
1339442	636	0	7	0	0/0	0/0	0/0
1379646	580	203	377	0	29/7	46/0	7/0
1380106	555	127	0	0	0/0	0/0	0/0
1381252	584	0	8	0	0/0	0/0	0/0

The direct early/deep pairing contains 2,148 PASS/PASS, 106 early-PASS/deep-FAIL, one early-FAIL/deep-PASS, 5,936 FAIL/FAIL and one shared UNKNOWN. The known-failure contrast is +105 for deep relative to early; conservative numerical envelope [104,106]. Resampling fourteen family totals with replacement 100,000 times (PCG64 seed 20261003) gives central 95% endpoints [−1,262], the 2.5th and 97.5th percentiles; mass 0.04949 at nonpositive values is descriptive and is not reported as a p-value. Omitting both seeds 1084018 and 1379646 leaves early/deep recoveries eight/seven, no regressions and a +1 known-failure contrast.

S5. All planned margin descriptors, exact stored MAE and stopping controls

The following retained tables provide twelve exact stored MAE rows, seventy-two margin-descriptor rows across six population/stratum combinations, and twenty-four stopping-control rows. Heading levels have been adapted for this supplement; numerical cells are unchanged from the named restored-analysis source.

These are existing main-study values copied from the executed analyses. Full-precision margin descriptors, paired outcomes and stopping controls are retained in data/recorded-original-results.json in the Zenodo package identified in S9. Exact labels and all strata are retained. MAE precision is the published stored string; margin descriptors are rounded to six decimal places here.

S5.1. Published MAE and static-agreement ranks

Published MAE is not a new equal-point/controller-path MAE measurement. Lower MAE is preferred. Source agreement ranks are higher-is-better and retain their within-network order under the later reconstruction.

Table S13. Published MAE and ranks.

Network	Repair	Published MAE, exact stored value	MAE rank	Published agreement rank
N2_1	optimal	0.09880119619383966	1	2
N2_1	early_exit	0.15178324323840378	3	3
N2_1	deep_opt_early_exit	0.10805544920386169	2	1
N3_1	optimal	0.2165610758422038	3	3
N3_1	early_exit	0.13973214676116677	1	2
N3_1	deep_opt_early_exit	0.17125559969732382	2	1
N4_1	optimal	0.10344876223457149	3	1
N4_1	early_exit	0.07101768825574706	1	2
N4_1	deep_opt_early_exit	0.09885619216751698	2	3
N5_1	optimal	0.12427581048437528	2	1
N5_1	early_exit	0.12577860759234813	3	3
N5_1	deep_opt_early_exit	0.08964228724955388	1	2

For N3,1 MAE favours early, while agreement favours deep. For the other three source networks those MAE choices and observed anchor net counts do not create a universal predictive ranking. This small related-candidate panel is not a correlation study over repair algorithms.

S5.2. Complete planned margin descriptors

Change is repaired minimum separation minus original minimum separation, in feet. Lower 1% quantiles retain the main-study calculation. The ≥100-ft loss and any-advisory-change columns are case counts with their own row denominator. `all_baseline_exposed` is baseline-defined network exposure, not target-domain exposure. Zero denominators and missing statistics, if any, are shown as recorded rather than assigned an effect. Population names are exact saved labels; `straight_line` is the uncontrolled close-approach challenge.

source / all_baseline_exposed

Table S14. Source exposed margin descriptors.

Network	Repair	n	Mean Δ ft	Min Δ ft	Lower 1% Δ ft	Loss ≥100 ft	Any advisory changed
N2_1	optimal	4221	+6.877120	-13336.413599	-1126.803087	479	1245
N2_1	early_exit	4221	-110.084009	-16275.687903	-2247.723903	884	1665
N2_1	deep_opt_early_exit	4221	-25.776412	-9102.472675	-1148.487439	612	1286
N3_1	optimal	5504	+432.111643	-8105.817346	-1201.026029	391	3275

Network	Repair	n	Mean Δ ft	Min Δ ft	Lower 1% Δ ft	Loss ≥100 ft	Any advisory changed
N3_1	early_exit	5504	+175.077634	-7925.837466	-1064.935861	384	2582
N3_1	deep_opt_early_exit	5504	+125.155920	-9784.801287	-1166.053957	434	2447
N4_1	optimal	3534	-32.254564	-17576.005151	-1051.049327	53	100
N4_1	early_exit	3534	-36.532582	-17576.005151	-1865.869889	52	107
N4_1	deep_opt_early_exit	3534	-23.208644	-17576.005151	-180.343249	40	95
N5_1	optimal	1291	+4.917489	-2279.004757	-257.825646	33	101
N5_1	early_exit	1291	+0.371826	-2279.004757	-342.139285	42	135
N5_1	deep_opt_early_exit	1291	-5.252893	-2279.004757	-313.802195	32	105

source / baseline_PASS

Table S15. Source baseline-PASS margin descriptors.

Network	Repair	n	Mean Δ ft	Min Δ ft	Lower 1% Δ ft	Loss ≥100 ft	Any advisory changed
N2_1	optimal	4220	+6.878749	-13336.413599	-1126.948901	479	1245
N2_1	early_exit	4220	-110.110095	-16275.687903	-2247.791563	884	1665
N2_1	deep_opt_early_exit	4220	-25.782521	-9102.472675	-1148.489376	612	1286
N3_1	optimal	5502	+432.268401	-8105.817346	-1201.220048	391	3274
N3_1	early_exit	5502	+175.140958	-7925.837466	-1065.196900	384	2581
N3_1	deep_opt_early_exit	5502	+125.205548	-9784.801287	-1166.066099	434	2446
N4_1	optimal	3533	-32.263693	-17576.005151	-1053.119060	53	100
N4_1	early_exit	3533	-36.542922	-17576.005151	-1865.910638	52	107
N4_1	deep_opt_early_exit	3533	-23.215213	-17576.005151	-181.037007	40	95
N5_1	optimal	1289	+4.925119	-2279.004757	-258.640089	33	101
N5_1	early_exit	1289	+0.372403	-2279.004757	-342.318036	42	135
N5_1	deep_opt_early_exit	1289	-5.261043	-2279.004757	-314.412047	32	105

anchor / all_baseline_exposed

Table S16. Anchor exposed margin descriptors.

Network	Repair	n	Mean Δ ft	Min Δ ft	Lower 1% Δ ft	Loss ≥100 ft	Any advisory changed
N2_1	optimal	5818	-6.367925	-4124.769114	-980.899502	371	2932
N2_1	early_exit	5818	+11.820744	-2903.681296	-1117.300188	393	3400
N2_1	deep_opt_early_exit	5818	+14.934247	-2128.362233	-914.388277	316	2830
N3_1	optimal	2637	-14.185441	-463.487644	-443.970981	101	1239
N3_1	early_exit	2637	+2.901168	-446.970015	-272.248197	36	996
N3_1	deep_opt_early_exit	2637	-16.467837	-458.428962	-430.539215	120	671
N4_1	optimal	3860	+1.254429	-292.168396	-4.201673	2	289
N4_1	early_exit	3860	+0.557601	-292.168396	-7.469948	2	190
N4_1	deep_opt_early_exit	3860	+0.781851	-292.168396	+0.000000	2	211
N5_1	optimal	4477	+0.054237	-277.772150	-1.717640	5	130
N5_1	early_exit	4477	+0.261006	-277.772150	-1.444456	6	192
N5_1	deep_opt_early_exit	4477	-0.056182	-277.772150	-11.492038	5	202

anchor / baseline_FAIL

Table S17. Anchor baseline-FAIL margin descriptors.

Network	Repair	n	Mean Δ ft	Min Δ ft	Lower 1% Δ ft	Loss ≥100 ft	Any advisory changed
N2_1	optimal	4080	+27.232755	-281.691622	-134.224599	148	2043
N2_1	early_exit	4080	+38.002961	-285.980404	-179.268440	189	2397
N2_1	deep_opt_early_exit	4080	+33.277029	-266.664120	-134.224599	149	2006
N3_1	optimal	1975	+2.653246	-96.610733	-93.109165	0	909
N3_1	early_exit	1975	+5.564748	-110.734195	-94.679210	5	723
N3_1	deep_opt_early_exit	1975	+2.658141	-292.135604	-75.131427	1	425
N4_1	optimal	2790	+1.673529	-31.027384	-4.177032	0	239
N4_1	early_exit	2790	+0.681199	-31.027384	-7.526483	0	145
N4_1	deep_opt_early_exit	2790	+0.994626	-31.027384	+0.000000	0	161

Network	Repair	n	Mean Δ ft	Min Δ ft	Lower 1% Δ ft	Loss ≥100 ft	Any advisory changed
N5_1	optimal	3171	+0.232496	-277.772150	+0.000000	4	82
N5_1	early_exit	3171	+0.355439	-277.772150	+0.000000	6	126
N5_1	deep_opt_early_exit	3171	+0.071980	-277.772150	-7.636375	4	139

target_speed / all_baseline_exposed

Table S18. Target-speed exposed margin descriptors.

Network	Repair	n	Mean Δ ft	Min Δ ft	Lower 1% Δ ft	Loss ≥100 ft	Any advisory changed
N2_1	optimal	62	+469.715695	-969.643895	-937.963480	11	39
N2_1	early_exit	62	-456.780420	-2306.193942	-2257.877044	25	29
N2_1	deep_opt_early_exit	62	+260.454651	-969.643895	-936.815892	10	34
N3_1	optimal	79	+200.950494	-1005.037944	-1002.693926	10	57
N3_1	early_exit	79	-316.975157	-3155.817086	-2830.569668	23	56
N3_1	deep_opt_early_exit	79	+67.148277	-1294.581927	-1068.737620	15	56
N4_1	optimal	51	-128.725938	-4735.845943	-3096.592611	3	3
N4_1	early_exit	51	-128.725938	-4735.845943	-3096.592611	3	3
N4_1	deep_opt_early_exit	51	-128.725938	-4735.845943	-3096.592611	3	3
N5_1	optimal	16	+11.478939	+0.000000	+0.000000	0	2
N5_1	early_exit	16	+11.478939	+0.000000	+0.000000	0	2
N5_1	deep_opt_early_exit	16	-100.582889	-1701.009031	-1445.857677	1	2

straight_line / all_baseline_exposed

Table S19. Close-approach exposed margin descriptors.

Network	Repair	n	Mean Δ ft	Min Δ ft	Lower 1% Δ ft	Loss ≥100 ft	Any advisory changed
N2_1	optimal	1979	+1.927280	-12598.981066	-1000.784261	255	689
N2_1	early_exit	1979	-127.039363	-13338.649541	-2189.231215	396	775
N2_1	deep_opt_early_exit	1979	-15.658032	-8837.739246	-999.359713	273	634
N3_1	optimal	2159	+507.671809	-2549.305919	-1386.856365	167	1478

Network	Repair	n	Mean Δ ft	Min Δ ft	Lower 1% Δ ft	Loss ≥100 ft	Any advisory changed
N3_1	early_exit	2159	+215.450708	-3972.358407	-1075.442061	171	1177
N3_1	deep_opt_early_exit	2159	+165.736811	-4904.911344	-1197.646393	194	1116
N4_1	optimal	1030	-37.804433	-5392.076452	-2180.907478	27	65
N4_1	early_exit	1030	-32.910094	-4623.974689	-1389.351917	27	71
N4_1	deep_opt_early_exit	1030	-9.676223	-4623.974689	-309.341859	17	56
N5_1	optimal	539	+22.658502	-1296.877577	-311.150850	24	66
N5_1	early_exit	539	+22.948078	-1296.877577	-280.106622	30	86
N5_1	deep_opt_early_exit	539	+27.912177	-1296.877577	-252.965430	15	65

S5.3. Complete default-horizon stopping controls

Default horizon is 150 seconds. Early stopping is disabled for these controls. Each source row uses the first 128 baseline-PASS/exposed cases; each anchor row uses the first 128 baseline-FAIL/exposed cases, per network in baseline order. Original baselines are shared across three repairs. Status/minimum change columns compare the original early-stop convention with the full-horizon version, separately for original and repaired traces. Nine-cell outcomes retain UNKNOWN cells explicitly; every unlisted transition count is zero only where the complete row below shows zero.

Table S20. Part 1: implementation and outcome descriptors. The shared first four columns identify the same records in both parts.

Network	Population	Repair	n	Original status changes	Original min changes >0.01 ft	Repair status changes	Repair min changes >0.01 ft	Max abs change in paired Δ ft
N2_1	source	optimal	128	0	0	0	0	0.0
N2_1	source	early_exit	128	0	0	0	0	0.0
N2_1	source	deep_opt_early_exit	128	0	0	0	0	0.0
N2_1	anchor	optimal	128	0	0	0	0	0.0
N2_1	anchor	early_exit	128	0	0	0	0	0.0
N2_1	anchor	deep_opt_early_exit	128	0	0	0	0	0.0
N3_1	source	optimal	128	0	0	0	0	0.0
N3_1	source	early_exit	128	0	0	0	0	0.0
N3_1	source	deep_opt_early_exit	128	0	0	0	0	0.0
N3_1	anchor	optimal	128	0	0	0	0	0.0
N3_1	anchor	early_exit	128	0	0	0	0	0.0
N3_1	anchor	deep_opt_early_exit	128	0	0	0	0	0.0

Network	Population	Repair	n	Original status changes	Original min changes >0.01 ft	Repair status changes	Repair min changes >0.01 ft	Max abs change in paired Δ ft
N4_1	source	optimal	128	0	0	0	0	0.0
N4_1	source	early_exit	128	0	0	0	0	0.0
N4_1	source	deep_opt_early_exit	128	0	0	0	0	0.0
N4_1	anchor	optimal	128	0	0	0	0	0.0
N4_1	anchor	early_exit	128	0	0	0	0	0.0
N4_1	anchor	deep_opt_early_exit	128	0	0	0	0	0.0
N5_1	source	optimal	128	0	0	0	0	0.0
N5_1	source	early_exit	128	0	0	0	0	0.0
N5_1	source	deep_opt_early_exit	128	0	0	0	0	0.0
N5_1	anchor	optimal	128	0	0	0	0	0.0
N5_1	anchor	early_exit	128	0	0	0	0	0.0
N5_1	anchor	deep_opt_early_exit	128	0	0	0	0	0.0

Table S20. Part 2: paired transitions. The shared first four columns identify the same records in both parts.

Network	Population	Repair	n	P → P	P → F	P → U	F → P	F → F	F → U	U → P	U → F	U → U
N2_1	source	optimal	128	128	0	0	0	0	0	0	0	0
N2_1	source	early_exit	128	128	0	0	0	0	0	0	0	0
N2_1	source	deep_opt_early_exit	128	128	0	0	0	0	0	0	0	0
N2_1	anchor	optimal	128	0	0	0	3	125	0	0	0	0
N2_1	anchor	early_exit	128	0	0	0	4	124	0	0	0	0
N2_1	anchor	deep_opt_early_exit	128	0	0	0	4	124	0	0	0	0
N3_1	source	optimal	128	128	0	0	0	0	0	0	0	0
N3_1	source	early_exit	128	128	0	0	0	0	0	0	0	0
N3_1	source	deep_opt_early_exit	128	128	0	0	0	0	0	0	0	0
N3_1	anchor	optimal	128	0	0	0	4	124	0	0	0	0
N3_1	anchor	early_exit	128	0	0	0	5	123	0	0	0	0
N3_1	anchor	deep_opt_early_exit	128	0	0	0	0	128	0	0	0	0

Network	Population	Repair	n	P→P	P→F	P→U	F→P	F→F	F→U	U→P	U→F	U→U
N4_1	source	optimal	128	128	0	0	0	0	0	0	0	0
N4_1	source	early_exit	128	128	0	0	0	0	0	0	0	0
N4_1	source	deep_opt_early_exit	128	128	0	0	0	0	0	0	0	0
N4_1	anchor	optimal	128	0	0	0	1	127	0	0	0	0
N4_1	anchor	early_exit	128	0	0	0	0	128	0	0	0	0
N4_1	anchor	deep_opt_early_exit	128	0	0	0	0	128	0	0	0	0
N5_1	source	optimal	128	128	0	0	0	0	0	0	0	0
N5_1	source	early_exit	128	128	0	0	0	0	0	0	0	0
N5_1	source	deep_opt_early_exit	128	128	0	0	0	0	0	0	0	0
N5_1	anchor	optimal	128	0	0	0	0	128	0	0	0	0
N5_1	anchor	early_exit	128	0	0	0	0	128	0	0	0	0
N5_1	anchor	deep_opt_early_exit	128	0	0	0	0	128	0	0	0	0

All 24 rows preserve classifications and minima within 0.01 ft. This does not mean every repair passes; the anchor rows include retained FAIL cases. These 1,024 network/case selections and 3,072 candidate/case comparisons are part of the earlier main-study controls, not additions to its main 278,528-draw population or its 10,752 overall control count. A case can appear in several network selections, so a distinct-encounter total is not inferred by summing row denominators. The selected 300-second diagnostic is documented in Section S7 and Online Resource 2, data/duration-outcomes.csv.

S6. All ninety-six broader cohort/candidate/subset transition rows

All twelve candidates are retained in each cohort and centre/jitter subset; no secondary candidate replaces the frozen N3 comparison after viewing the outcomes. For N2 optimal the extension's F→U transition explains why its known-failure change is −180 rather than 352−531=−179. Numeric envelopes and resampling sensitivity target different objects.

Table S21. Part 1: paired transitions. The shared first four columns identify the same records in both parts.

Cohort	Candidate	Subset	n	P/P	P/F	P/U	F/P	F/F	F/U	U/P	U/F	U/U
extension	optimal_2_1	whole	28781	10233	352	0	531	17663	1	0	0	1
extension	optimal_2_1	baseline_exposed	22693	7478	352	0	531	14330	1	0	0	1
extension	optimal_2_1	centre_only	1693	0	0	0	56	1637	0	0	0	0
extension	optimal_2_1	jitter_only	27088	10233	352	0	475	16026	1	0	0	1
extension	early_exit_2_1	whole	28781	10142	443	0	812	17383	0	0	0	1

Cohort	Candidate	Subset	n	P/P	P/F	P/U	F/P	F/F	F/U	U/P	U/F	U/U
extension	early_exit_2_1	baseline_exposed	22693	7387	443	0	812	14050	0	0	0	1
extension	early_exit_2_1	centre_only	1693	0	0	0	110	1583	0	0	0	0
extension	early_exit_2_1	jitter_only	27088	10142	443	0	702	15800	0	0	0	1
extension	deep_opt_early_exit_2_1	whole	28781	10224	361	0	586	17609	0	0	0	1
extension	deep_opt_early_exit_2_1	baseline_exposed	22693	7469	361	0	586	14276	0	0	0	1
extension	deep_opt_early_exit_2_1	centre_only	1693	0	0	0	79	1614	0	0	0	0
extension	deep_opt_early_exit_2_1	jitter_only	27088	10224	361	0	507	15995	0	0	0	1
extension	optimal_3_1	whole	28781	10165	420	0	212	17983	0	0	0	1
extension	optimal_3_1	baseline_exposed	18168	6533	420	0	212	11003	0	0	0	0
extension	optimal_3_1	centre_only	1693	0	0	0	34	1659	0	0	0	0
extension	optimal_3_1	jitter_only	27088	10165	420	0	178	16324	0	0	0	1
extension	early_exit_3_1	whole	28781	10278	307	0	395	17800	0	0	0	1
extension	early_exit_3_1	baseline_exposed	18168	6646	307	0	395	10820	0	0	0	0
extension	early_exit_3_1	centre_only	1693	0	0	0	50	1643	0	0	0	0
extension	early_exit_3_1	jitter_only	27088	10278	307	0	345	16157	0	0	0	1
extension	deep_opt_early_exit_3_1	whole	28781	10296	289	0	391	17804	0	0	0	1
extension	deep_opt_early_exit_3_1	baseline_exposed	18168	6664	289	0	391	10824	0	0	0	0
extension	deep_opt_early_exit_3_1	centre_only	1693	0	0	0	50	1643	0	0	0	0
extension	deep_opt_early_exit_3_1	jitter_only	27088	10296	289	0	341	16161	0	0	0	1
extension	optimal_4_1	whole	28781	10542	43	0	89	18106	0	0	0	1
extension	optimal_4_1	baseline_exposed	14084	5659	43	0	89	8292	0	0	0	1

Cohort	Candidate	Subset	n	P/P	P/F	P/U	F/P	F/F	F/U	U/P	U/F	U/U
extension	optimal_4_1	centre_only	1693	0	0	0	14	1679	0	0	0	0
extension	optimal_4_1	jitter_only	27088	10542	43	0	75	16427	0	0	0	1
extension	early_exit_4_1	whole	28781	10551	34	0	100	18095	0	0	0	1
extension	early_exit_4_1	baseline_exposed	14084	5668	34	0	100	8281	0	0	0	1
extension	early_exit_4_1	centre_only	1693	0	0	0	15	1678	0	0	0	0
extension	early_exit_4_1	jitter_only	27088	10551	34	0	85	16417	0	0	0	1
extension	deep_opt_early_exit_4_1	whole	28781	10554	31	0	99	18096	0	0	0	1
extension	deep_opt_early_exit_4_1	baseline_exposed	14084	5671	31	0	99	8282	0	0	0	1
extension	deep_opt_early_exit_4_1	centre_only	1693	0	0	0	16	1677	0	0	0	0
extension	deep_opt_early_exit_4_1	jitter_only	27088	10554	31	0	83	16419	0	0	0	1
extension	optimal_5_1	whole	28781	10581	4	0	44	18151	0	0	0	1
extension	optimal_5_1	baseline_exposed	17038	5983	4	0	44	11007	0	0	0	0
extension	optimal_5_1	centre_only	1693	0	0	0	9	1684	0	0	0	0
extension	optimal_5_1	jitter_only	27088	10581	4	0	35	16467	0	0	0	1
extension	early_exit_5_1	whole	28781	10562	23	0	73	18122	0	0	0	1
extension	early_exit_5_1	baseline_exposed	17038	5964	23	0	73	10978	0	0	0	0
extension	early_exit_5_1	centre_only	1693	0	0	0	11	1682	0	0	0	0
extension	early_exit_5_1	jitter_only	27088	10562	23	0	62	16440	0	0	0	1
extension	deep_opt_early_exit_5_1	whole	28781	10573	12	0	52	18143	0	0	0	1
extension	deep_opt_early_exit_5_1	baseline_exposed	17038	5975	12	0	52	10999	0	0	0	0
extension	deep_opt_early_exit_5_1	centre_only	1693	0	0	0	8	1685	0	0	0	0

Cohort	Candidate	Subset	n	P/P	P/F	P/U	F/P	F/F	F/U	U/P	U/F	U/U
extension	deep_opt_early_exit_5_1	jitter_only	27088	10573	12	0	44	16458	0	0	0	1
replication	optimal_2_1	whole	238	52	5	0	5	176	0	0	0	0
replication	optimal_2_1	baseline_exposed	225	46	5	0	5	169	0	0	0	0
replication	optimal_2_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	optimal_2_1	jitter_only	224	52	5	0	5	162	0	0	0	0
replication	early_exit_2_1	whole	238	52	5	0	4	177	0	0	0	0
replication	early_exit_2_1	baseline_exposed	225	46	5	0	4	170	0	0	0	0
replication	early_exit_2_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	early_exit_2_1	jitter_only	224	52	5	0	4	163	0	0	0	0
replication	deep_opt_early_exit_2_1	whole	238	53	4	0	5	176	0	0	0	0
replication	deep_opt_early_exit_2_1	baseline_exposed	225	47	4	0	5	169	0	0	0	0
replication	deep_opt_early_exit_2_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	deep_opt_early_exit_2_1	jitter_only	224	53	4	0	5	162	0	0	0	0
replication	optimal_3_1	whole	238	55	2	0	0	181	0	0	0	0
replication	optimal_3_1	baseline_exposed	92	25	2	0	0	65	0	0	0	0
replication	optimal_3_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	optimal_3_1	jitter_only	224	55	2	0	0	167	0	0	0	0
replication	early_exit_3_1	whole	238	56	1	0	0	181	0	0	0	0
replication	early_exit_3_1	baseline_exposed	92	26	1	0	0	65	0	0	0	0
replication	early_exit_3_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	early_exit_3_1	jitter_only	224	56	1	0	0	167	0	0	0	0
replication	deep_opt_early_exit_3_1	whole	238	55	2	0	0	181	0	0	0	0

Cohort	Candidate	Subset	n	P/P	P/F	P/U	F/P	F/F	F/U	U/P	U/F	U/U
replication	deep_opt_early_exit_3_1	baseline_exposed	92	25	2	0	0	65	0	0	0	0
replication	deep_opt_early_exit_3_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	deep_opt_early_exit_3_1	jitter_only	224	55	2	0	0	167	0	0	0	0
replication	optimal_4_1	whole	238	57	0	0	0	181	0	0	0	0
replication	optimal_4_1	baseline_exposed	110	20	0	0	0	90	0	0	0	0
replication	optimal_4_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	optimal_4_1	jitter_only	224	57	0	0	0	167	0	0	0	0
replication	early_exit_4_1	whole	238	57	0	0	0	181	0	0	0	0
replication	early_exit_4_1	baseline_exposed	110	20	0	0	0	90	0	0	0	0
replication	early_exit_4_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	early_exit_4_1	jitter_only	224	57	0	0	0	167	0	0	0	0
replication	deep_opt_early_exit_4_1	whole	238	57	0	0	0	181	0	0	0	0
replication	deep_opt_early_exit_4_1	baseline_exposed	110	20	0	0	0	90	0	0	0	0
replication	deep_opt_early_exit_4_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	deep_opt_early_exit_4_1	jitter_only	224	57	0	0	0	167	0	0	0	0
replication	optimal_5_1	whole	238	57	0	0	0	181	0	0	0	0
replication	optimal_5_1	baseline_exposed	135	41	0	0	0	94	0	0	0	0
replication	optimal_5_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	optimal_5_1	jitter_only	224	57	0	0	0	167	0	0	0	0
replication	early_exit_5_1	whole	238	57	0	0	0	181	0	0	0	0
replication	early_exit_5_1	baseline_exposed	135	41	0	0	0	94	0	0	0	0

Cohort	Candidate	Subset	n	P/P	P/F	P/U	F/P	F/F	F/U	U/P	U/F	U/U
replication	early_exit_5_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	early_exit_5_1	jitter_only	224	57	0	0	0	167	0	0	0	0
replication	deep_opt_early_exit_5_1	whole	238	57	0	0	0	181	0	0	0	0
replication	deep_opt_early_exit_5_1	baseline_exposed	135	41	0	0	0	94	0	0	0	0
replication	deep_opt_early_exit_5_1	centre_only	14	0	0	0	0	14	0	0	0	0
replication	deep_opt_early_exit_5_1	jitter_only	224	57	0	0	0	167	0	0	0	0

Table S21. Part 2: implementation and outcome descriptors. The shared first four columns identify the same records in both parts.

Cohort	Candidate	Subset	n	Exposed	Original target	Repair target	Advisory changed	Mean Δmargin ft	Min Δmargin ft
extension	optimal_2_1	whole	28781	22693	0	0	7681	5.806986071	-8583.722373338
extension	optimal_2_1	baseline_exposed	22693	22693	0	0	7681	7.364864324	-8583.722373338
extension	optimal_2_1	centre_only	1693	1380	0	0	466	37.341460267	-285.027202790
extension	optimal_2_1	jitter_only	27088	21313	0	0	7215	3.836081433	-8583.722373338
extension	early_exit_2_1	whole	28781	22693	0	0	9096	36.633121854	-6076.440127206
extension	early_exit_2_1	baseline_exposed	22693	22693	0	0	9096	46.460929806	-6076.440127206
extension	early_exit_2_1	centre_only	1693	1380	0	0	557	100.903851903	-370.169419197
extension	early_exit_2_1	jitter_only	27088	21313	0	0	8539	32.616201226	-6076.440127206
extension	deep_opt_early_exit_2_1	whole	28781	22693	0	0	7897	19.736148780	-8583.722373338
extension	deep_opt_early_exit_2_1	baseline_exposed	22693	22693	0	0	7897	25.030894903	-8583.722373338
extension	deep_opt_early_exit_2_1	centre_only	1693	1380	0	0	480	65.292133293	-367.396216284
extension	deep_opt_early_exit_2_1	jitter_only	27088	21313	0	0	7417	16.888899748	-8583.722373338
extension	optimal_3_1	whole	28781	18168	0	0	7847	-10.117568567	-8281.842748302
extension	optimal_3_1	baseline_exposed	18168	18168	0	0	7847	-16.027836907	-8281.842748302
extension	optimal_3_1	centre_only	1693	1048	0	0	464	37.100137527	-289.137423380
extension	optimal_3_1	jitter_only	27088	17120	0	0	7383	-13.068675198	-8281.842748302

Cohort	Candidate	Subset	n	Exposed	Original target	Repair target	Advisory changed	Mean Δmargin ft	Min Δmargin ft
extension	early_exit_3_1	whole	28781	18168	0	0	8254	28.241500544	-8281.842748302
extension	early_exit_3_1	baseline_exposed	18168	18168	0	0	8254	44.739026153	-8281.842748302
extension	early_exit_3_1	centre_only	1693	1048	0	0	488	67.446112386	-265.314694709
extension	early_exit_3_1	jitter_only	27088	17120	0	0	7766	25.791212304	-8281.842748302
extension	deep_opt_early_exit_3_1	whole	28781	18168	0	0	6848	29.414398401	-8281.842748302
extension	deep_opt_early_exit_3_1	baseline_exposed	18168	18168	0	0	6848	46.597082803	-8281.842748302
extension	deep_opt_early_exit_3_1	centre_only	1693	1048	0	0	411	70.034888293	-257.348358211
extension	deep_opt_early_exit_3_1	jitter_only	27088	17120	0	0	6437	26.875617782	-8281.842748302
extension	optimal_4_1	whole	28781	14084	0	0	1332	4.118608558	-3840.509357132
extension	optimal_4_1	baseline_exposed	14084	14084	0	0	1332	8.416477770	-3840.509357132
extension	optimal_4_1	centre_only	1693	793	0	0	82	12.528090351	-121.870907209
extension	optimal_4_1	jitter_only	27088	13291	0	0	1250	3.593015946	-3840.509357132
extension	early_exit_4_1	whole	28781	14084	0	0	1115	7.405934992	-3566.910320250
extension	early_exit_4_1	baseline_exposed	14084	14084	0	0	1115	15.134210096	-3566.910320250
extension	early_exit_4_1	centre_only	1693	793	0	0	68	13.305358217	-43.899652584
extension	early_exit_4_1	jitter_only	27088	13291	0	0	1047	7.037221040	-3566.910320250
extension	deep_opt_early_exit_4_1	whole	28781	14084	0	0	1171	7.144729911	-3566.910320250
extension	deep_opt_early_exit_4_1	baseline_exposed	14084	14084	0	0	1171	14.600431098	-3566.910320250
extension	deep_opt_early_exit_4_1	centre_only	1693	793	0	0	74	14.420090272	-121.870907209
extension	deep_opt_early_exit_4_1	jitter_only	27088	13291	0	0	1097	6.690019889	-3566.910320250
extension	optimal_5_1	whole	28781	17038	0	0	1376	2.227452562	-2556.885494547
extension	optimal_5_1	baseline_exposed	17038	17038	0	0	1376	3.762666520	-2556.885494547
extension	optimal_5_1	centre_only	1693	1047	0	0	72	6.269431661	-81.908476675
extension	optimal_5_1	jitter_only	27088	15991	0	0	1304	1.974828868	-2556.885494547
extension	early_exit_5_1	whole	28781	17038	0	0	1883	2.789730073	-2556.885494547
extension	early_exit_5_1	baseline_exposed	17038	17038	0	0	1883	4.712479237	-2556.885494547
extension	early_exit_5_1	centre_only	1693	1047	0	0	108	7.735818461	-81.908476675

Cohort	Candidate	Subset	n	Exposed	Original target	Repair target	Advisory changed	Mean Δmargin ft	Min Δmargin ft
extension	early_exit_5_1	jitter_only	27088	15991	0	0	1775	2.480599549	-2556.885494547
extension	deep_opt_early_exit_5_1	whole	28781	17038	0	0	1522	3.093587902	-2556.885494547
extension	deep_opt_early_exit_5_1	baseline_exposed	17038	17038	0	0	1522	5.225763200	-2556.885494547
extension	deep_opt_early_exit_5_1	centre_only	1693	1047	0	0	86	5.890505337	-81.908476675
extension	deep_opt_early_exit_5_1	jitter_only	27088	15991	0	0	1436	2.918780562	-2556.885494547
replication	optimal_2_1	whole	238	225	0	0	97	-25.105806635	-2334.418348580
replication	optimal_2_1	baseline_exposed	225	225	0	0	97	-26.556364351	-2334.418348580
replication	optimal_2_1	centre_only	14	13	0	0	6	-11.269292982	-102.466382016
replication	optimal_2_1	jitter_only	224	212	0	0	91	-25.970588738	-2334.418348580
replication	early_exit_2_1	whole	238	225	0	0	99	-39.268796496	-2334.418348580
replication	early_exit_2_1	baseline_exposed	225	225	0	0	99	-41.537660293	-2334.418348580
replication	early_exit_2_1	centre_only	14	13	0	0	7	-12.901321628	-102.466382016
replication	early_exit_2_1	jitter_only	224	212	0	0	92	-40.916763675	-2334.418348580
replication	deep_opt_early_exit_2_1	whole	238	225	0	0	88	-14.490738320	-1148.724716272
replication	deep_opt_early_exit_2_1	baseline_exposed	225	225	0	0	88	-15.327980979	-1148.724716272
replication	deep_opt_early_exit_2_1	centre_only	14	13	0	0	6	-11.269292982	-102.466382016
replication	deep_opt_early_exit_2_1	jitter_only	224	212	0	0	82	-14.692078654	-1148.724716272
replication	optimal_3_1	whole	238	92	0	0	34	-2.531478094	-456.052439796
replication	optimal_3_1	baseline_exposed	92	92	0	0	34	-6.548823766	-456.052439796
replication	optimal_3_1	centre_only	14	5	0	0	1	0.270346604	0.000000000
replication	optimal_3_1	jitter_only	224	87	0	0	33	-2.706592138	-456.052439796
replication	early_exit_3_1	whole	238	92	0	0	31	-1.979580139	-251.045668080
replication	early_exit_3_1	baseline_exposed	92	92	0	0	31	-5.121087751	-251.045668080
replication	early_exit_3_1	centre_only	14	5	0	0	1	0.270346604	0.000000000
replication	early_exit_3_1	jitter_only	224	87	0	0	30	-2.120200561	-251.045668080
replication	deep_opt_early_exit_3_1	whole	238	92	0	0	19	-3.800936211	-431.220654840
replication	deep_opt_early_exit_3_1	baseline_exposed	92	92	0	0	19	-9.832856720	-431.220654840

Cohort	Candidate	Subset	n	Exposed	Original target	Repair target	Advisory changed	Mean Δmargin ft	Min Δmargin ft
replication	deep_opt_early_exit_3_1	centre_only	14	5	0	0	1	0.000000000	0.000000000
replication	deep_opt_early_exit_3_1	jitter_only	224	87	0	0	18	-4.038494724	-431.220654840
replication	optimal_4_1	whole	238	110	0	0	13	0.147108220	-10.744749425
replication	optimal_4_1	baseline_exposed	110	110	0	0	13	0.318288693	-10.744749425
replication	optimal_4_1	centre_only	14	6	0	0	0	0.000000000	0.000000000
replication	optimal_4_1	jitter_only	224	104	0	0	13	0.156302483	-10.744749425
replication	early_exit_4_1	whole	238	110	0	0	5	0.009322057	-10.744749425
replication	early_exit_4_1	baseline_exposed	110	110	0	0	5	0.020169542	-10.744749425
replication	early_exit_4_1	centre_only	14	6	0	0	0	0.000000000	0.000000000
replication	early_exit_4_1	jitter_only	224	104	0	0	5	0.009904686	-10.744749425
replication	deep_opt_early_exit_4_1	whole	238	110	0	0	7	0.094730863	-10.744749425
replication	deep_opt_early_exit_4_1	baseline_exposed	110	110	0	0	7	0.204963140	-10.744749425
replication	deep_opt_early_exit_4_1	centre_only	14	6	0	0	0	0.000000000	0.000000000
replication	deep_opt_early_exit_4_1	jitter_only	224	104	0	0	7	0.100651542	-10.744749425
replication	optimal_5_1	whole	238	135	0	0	5	0.201401824	-1.894760881
replication	optimal_5_1	baseline_exposed	135	135	0	0	5	0.355063956	-1.894760881
replication	optimal_5_1	centre_only	14	8	0	0	1	3.559171069	0.000000000
replication	optimal_5_1	jitter_only	224	127	0	0	4	-0.008458754	-1.894760881
replication	early_exit_5_1	whole	238	135	0	0	6	0.292858387	0.000000000
replication	early_exit_5_1	baseline_exposed	135	135	0	0	6	0.516298490	0.000000000
replication	early_exit_5_1	centre_only	14	8	0	0	1	3.559171069	0.000000000
replication	early_exit_5_1	jitter_only	224	127	0	0	5	0.088713845	0.000000000
replication	deep_opt_early_exit_5_1	whole	238	135	0	0	7	0.101394919	-26.037496870
replication	deep_opt_early_exit_5_1	baseline_exposed	135	135	0	0	7	0.178755487	-26.037496870
replication	deep_opt_early_exit_5_1	centre_only	14	8	0	0	1	3.559171069	0.000000000
replication	deep_opt_early_exit_5_1	jitter_only	224	127	0	0	6	-0.114716090	-26.037496870

The Zenodo package identified in S9 contains all 20,484 centre/candidate transition rows in data/families.json, each centre's primary comparison and recorded patched-invocation diagnostics in data/recorded-broader-results.json, and all 377,247 case/system outcomes in data/case-results.csv. The same package supplies the 29,019 input vectors and resampling draws. Full trajectories require regeneration using its programs.

S7. Numerical controls and execution chronology

The original converted-ONNX attempt is stopped after an advisory mismatch; its repaired outcomes do not enter the revised comparison. A dated amendment switches repairs to direct scalar float32 CPU PyTorch with `normalize=False`, retaining original ONNX inference for untouched components. Matched N3 grid reconstruction later covers all 24,299,779 retained inputs, with rare exact minimum ties recorded. The conservative agreement-count advantage for deep is [13,903,13,911]. Stored trajectory-union and control samples are implementation diagnostics, not independent validation populations.

The source-law Clopper–Pearson regression family and the separate anchor-recovery family each use their own nominal error allocation. Their intervals are not combined into a joint operational guarantee. Numerical UNKNOWN envelopes allow binary assignments to unresolved states; paired jitter and centre-composition resampling instead concern the frozen artificial sampling design. The broader jitter and centre-composition analyses use 2,000 and 10,000 paired resamples, respectively; their central 95% sensitivity intervals use the 2.5th and 97.5th percentiles. Neither supplies fleet prevalence or aircraft certification evidence.

Stopping controls retain twenty-four rows with no classification or >0.01-ft minimum-separation change. The 300-second diagnostic has 717 selected cases and 2,868 system traces and retains their original classifications; recovery cases were selected using preceding outcomes. The main threshold checks cover 401,054 records. The broader checks reproduce 377,247 case/system outcome rows and compare all 5,265,103 patched invocations with a separately loaded same-backend checkpoint. This verifies wiring/convention rather than independent floating-point correctness.

The broader worker was stopped after a disk-monitor race. The preserved completed groups, saved baseline content and patched invocations were checked; only the missing 6,009 repair records and eight controls were completed under the storage-monitor amendment. Per-record verification spans the complete result; continuation global diagnostic counters apply only to that suffix. Population, horizon, thresholds, estimands and execution limit were not changed. Frozen execution/aggregation totals 1,386.323 seconds, including the continuation, against the fixed 7,200-second execution limit. The supporting ZIP retains the continuation details and recorded numerical checks; the Zenodo package supplies runnable simulation and analysis programs.

S8. Reference versions used

The manuscript gives published bibliographic identities. The following exact software, preprint and article versions supply the source-specific support for this assessment. Version-specific section and page pointers remain distinct from publisher pagination.

Table S22. Reference versions used.

Reference	Source version	Version-specific support
1	Pinned Specrepair commit	Property implementation, checkpoints and static calculator; commit governs executable identity.
2	Pinned simulator commit	Planar dynamics, advisory selection, original runtime and public unsafe seed catalogue.
3	arXiv 1810.04240v1, 9 Oct 2018, 13 pages	Section VII/Table II, PDF pages 10–11; the 2019 journal identity includes “Systems” in its title.
4	32-page author-hosted lopez2022jat.pdf	Abstract, II.B–C and IV; author-file page numbers are not the 17-page publisher pagination.
5	arXiv 2201.06626v1, 17 Jan 2022, 28 pages	Original arXiv title/version; the NFM 2022 publication has a distinct title.
6	arXiv 2112.02807v1, 6 Dec 2021, 12 pages	Repair evaluation in section 7.4; the ISSTA 2022 final uses the shorter MDPFuzz title. The comparison uses the specified author version.

Reference	Source version	Version-specific support
7	Complete 28-page author final	CoMPAQt evaluation, section 6, pages 340:22–24. The publication identity matches the source version.
8	ICML/PMLR final and artefact v1.0.0	Section 5.1, Appendix B.2.2 and Table 7; pinned checkpoint/repository identities are separate from artefact metadata.
9	arXiv 2106.01917v5, 12 May 2022, 27 pages	Table 1 and sections 3–5; the extended author version supports the text, with SPIN identity cited separately.
10	Complete TOSEM publisher article	Repairs/breaks prediction and ACAS-specific repair-method selection; 34(4), article 94, pages 1–42.
11	Complete supplied 13-page BIRDNN publisher-layout article, Neural Networks 183 (2025), 106949	Sections 3–5, pp.3–11; explicit repair/preservation objectives, ACAS pointwise improvement/drawdown, layer/budget tests and repair-guarantee limitation. Final Tables 7/9 and Figure 2(c) govern the current comparison.
12	Complete 12-page NCCDA publisher final	Distinct from the separately retained 14-page author file and its extra appendix.
13	Complete supplied five-page SPGR proceedings article	Distinct from the earlier four-page repository experiment note.
14	Complete supplied 14-page CCBR publisher-layout article, ESWA 257 (2024), 125082	Sections 2.3–3.4, pp.4–8; Section 4.2/Table 1, p.8; Section 5.1/Table 3, p.9. Counterfactual localisation, NSGA-III repair and pointwise ACAS property/accuracy metrics.
15	Focused NSRepair publisher HTML	Sections 2.1.4, 3, 4.1–4.2 and 5.2/Table 3; the cited passages support the comparison.
16	Kyushu dissertation author chapter	Complete Chapter 7, printed pages 130–151 / PDF pages 149–170; GECCO final-paper identity/equivalence is not inferred.
17	Focused ISAR publisher HTML	Section 2.2 and sections 3.1–3.3; region verification and sampled behaviour remain distinct.
18	Focused Stefani et al. CEAS publisher reading	ODD-based aviation scenario generation and FlightGear ACAS-related evaluation; the source ledger retains the cited passages.
19	arXiv 1702.01135v1, 3 Feb 2017, 31 pages	Section 6 and Appendix E, PDF pages 16–17 and 30, including φ2. CAV final metadata is cited separately.
20	Complete eight-page Sheikhi–Bak publisher benchmark article (2023)	Sections 2–4, pp.2–7: Dubins/F-16 closed-loop ACAS Xu models and unrepaired-controller simulations. The paper and current pinned executable remain separate source versions.
22	Nine-page SESAR Innovation Days 2025 paper 49; arXiv:2503.23518v1 used before the new runs	Section II-B, proceedings pp. 2–3: fixed one-/four-second response settings; Section III-B describes command shifting. Different controller and aircraft evaluation from the present early neural prototype.

S9. Reproduction files and access

The research materials are supplied in two places. Table S23 lists the public Zenodo Python/data package. Table S24 lists the additional numerical data and checks supplied in the accompanying supporting ZIP.

S9.1 Published Zenodo reproduction package

Zenodo version 1.0.0, cited as Reference 21 in the main article, is available at https://doi.org/10.5281/zenodo.23202447. Its file, 2026-10-06-acas-zenodo.zip, contains eleven standalone Python programs, fifteen plain data/configuration/result files, three vector figures, requirements.txt and README.md. Table S23 accounts for all 31 archive entries; folder rows group their contents for readability.

Table S23. Published Zenodo reproduction package.

File or folder	Contents and purpose
README.md	Installation and run instructions.
requirements.txt	Python dependencies.
reproduce.py	Runs the supplied-data analysis and recreates Figures 1–3.
analyse.py	Recalculates the broader tables and resampling summaries from saved outcomes.
draw_figures.py	Draws the three figures from the supplied definitions and numerical tables.
prepare_models.py	Downloads the public controller code and released network checkpoints.
run_encounters.py	Replays original and broader encounters using the documented study conventions.
simulation.py	Study measurement and scalar checkpoint functions.
main_helpers.py	Frozen encounter-generation and analysis helpers.
bootstrap.py	Loads the CPU inference dependencies and pinned public source modules.
check_threshold.py	Cross-checks the continuous-separation threshold.
static_fidelity.py	Reconstructs the frozen source-grid static metrics.
matched_static.py	Checks scalar N3 fidelity using the encounter inference convention.
data/	Fifteen JSON/CSV files: encounter inputs and outcomes, study designs, centre records, resampling draws, static/duration summaries and figure inputs.
figures/	Three vector PDFs: Figure_1_acasxu_controller.pdf, Figure_2_static_and_encounters.pdf and Figure_3_cohort_differences.pdf.

Extract the public ZIP and follow README.md. The command python reproduce.py recalculates the supplied-data analysis and draws the figures. Optional simulation and static-metric commands first obtain the pinned public software and checkpoints with prepare_models.py. Full trajectory databases require regeneration; they are not included in the public ZIP. Original, extension and replication populations retain their separate denominators.

S9.2 Additional data and checks supplied with the journal submission

Online Resource 2 (ESM_2.zip) contains one README, three Python programs and twelve plain data files. Table S24 lists all sixteen files. These supplement the public Zenodo package with the detailed static, inference, duration and original-family checks and the new stopping-rule and response-delay study.

Table S24. Additional data and numerical checks.

File	Content
README.md	Instructions and descriptions of the supporting files.
verify_supporting.py	Recalculation and consistency checks using the supplied data.
data/static-filter-batches.csv	All 5,936 original-network grid/filter batch records.
data/static-repair-batches.csv	All 17,808 repair batch records for the twelve checkpoints.

File	Content
data/matched-grid-batches.csv	All 4,452 N3 controller-path grid batch records.
data/inference-diagnostic-batches.csv	All 19,698 batch/scalar and normalisation comparison records.
data/duration-outcomes.csv	All 5,736 selected case/system/horizon outcomes at 150 and 300 seconds.
data/original-family-analysis.json	Original fourteen-family analyses, family omissions and four representative encounters.
data/inference-diagnostics.json	Recorded inference comparisons, discrepancy examples and matched-grid totals.
data/verification-records.json	Recorded static, original, duration and broader execution checks and continuation details.
strengthen.py	Runs the complete anchor stopping control and all centre delay conditions using the public package.
analyse_strengthening.py	Recalculates the new tables, family contributions and resampling summaries.
data/anchor-outcomes.csv	49,152 original-stop/full-horizon case/system outcome records for all 8,192 anchors.
data/delay-outcomes.csv	15,363 centre/system/delay records for the separate 1,693-centre extension and fourteen-centre replication.
data/strengthening-summary.json	Complete tables, paired transitions, family contributions and sensitivity summaries.
data/strengthening-checks.json	Recorded completion, path-equivalence, numerical, inference and runtime checks.

Extract the supporting ZIP. Run python3 verify_supporting.py with Python 3.9 or later to check the original supporting data without additional packages. Run python analyse_strengthening.py in the published NumPy 2.2.6 environment to reproduce the new tables and family resampling from the supplied CSV files. To rerun the new simulations, copy strengthen.py into an extracted copy of the Zenodo package, prepare its public models and run python strengthen.py --stage anchor and python strengthen.py --stage delay. The supporting README gives the complete short commands. The original populations, extension and replication keep separate denominators. No training or new repair is required.

S10. Complete stopping control and command-delay sensitivity

Two additional checks were specified before running their new outcomes on 9 October 2026. They use the same pinned public source revisions, unchanged N3,1 early_exit/deep_opt_early_exit checkpoints, one-thread CPU scalar inference, one-second updates, planar fixed-speed dynamics, straight intruder, $\tau=0$ and guarded continuous 500-ft criterion. The archive's twelve checkpoint identities match the original experiment record. These controls reuse the existing inputs; no new method, repair or training is introduced.

For the stopping control, all original anchor indices 1000000–1008191 are regenerated with the published seed and family-preserving jitter procedure. Every input is run for 150 seconds under the baseline, early and deep controllers. The source's first increasing-separation stopping prefix is extracted from each full path; direct calls to the unchanged source routine on the first sixteen cases and the known UNKNOWN case validate the shortcut for all three controllers. Their 51 prefixes and 51 zero-delay full paths match exactly, including continuous minima. Complete historical baseline and nine-cell repair transitions also reproduce. No baseline-nonexposed case is omitted from the full replay.

All 24,576 full-horizon trajectories preserve their corresponding original-stop classification and continuous minimum exactly. Every case continues after its original stopping time. Each controller gains 718 N3-exposed cases, increasing exposure from 2,637 to 3,355. There is no repaired-property-domain exposure. Table S25 retains both sets of totals; full paired cells and case records are supplied in Online Resource 2. Each repair's one UNKNOWN is the same original boundary case, not an additional independent observation.

Table S25. Complete original-anchor stopping control; 8,192 cases per row.

Protocol	Controller	PASS	FAIL	UNKNOWN	N3 exposure
Original stop	Baseline	2230	5961	1	2637
Original stop	Early	2254	5937	1	2637
Original stop	Deep	2149	6042	1	2637
Full 150 s	Baseline	2230	5961	1	3355
Full 150 s	Early	2254	5937	1	3355
Full 150 s	Deep	2149	6042	1	3355

The full-horizon early/deep direct table is 2,148 PASS/PASS, 106 PASS/FAIL, one FAIL/PASS, 5,936 FAIL/FAIL and one UNKNOWN/UNKNOWN, with every other cell zero. Thus the +105 contrast and conservative [104,106] numerical envelope remain. Table S26 shows the complete family contributions. The repeated 100,000-resample PCG64 analysis (seed 20261003, fourteen family totals drawn with replacement) gives the same central 95% endpoints [−1,262] and mass at or below zero 0.04949 under both stopping conventions. Removing families 1084018 and 1379646 leaves +1. These are composition sensitivities, not fleet intervals.

Table S26. Family contributions to the deep-minus-early known-failure difference.

Original family seed	Cases	Original stop	Full 150 s
57305	590	0	0
174596	605	-1	-1
191434	608	0	0
207659	556	0	0
341255	573	0	0
468582	584	0	0
872977	576	2	2
891092	557	0	0
1084018	602	65	65
1182535	586	0	0

Original family seed	Cases	Original stop	Full 150 s
1339442	636	0	0
1379646	580	39	39
1380106	555	0	0
1381252	584	0	0

The command-delay sensitivity uses every unchanged centre and no perturbations: 1,693 extension centres and fourteen replication centres. It executes 15,363 full trajectories, comprising three controllers at each of zero, one and four seconds of delay. Each controller issues a command from the current geometric state and previous issued advisory at each integer second. Applied motion uses the command issued d seconds earlier, with clear-of-conflict before second d. The previous-issued-advisory memory is retained for network selection; the delayed physical command does not replace that memory. This is a pure command queue, not a pilot reaction or roll-dynamics model. The one- and four-second values were motivated by Reference 22's author version, Section II-B, and also appear in its proceedings version, pp. 2–3.

The fixed centre list and all delay conditions were specified before their new results; no condition was selected to favour a repair. Every centre/controller zero-delay trajectory matches the preceding full-horizon runner exactly (5,121 paths), and every corresponding status and minimum matches the released centre record. At nonzero delay, each repair is compared with a fresh baseline at that same delay. All centres are retained whether or not the baseline visits N3. Issued commands are used for inference validation and exposure; applied commands are used for motion and separation checking. Table S27 gives the complete counts; Table S28 gives direct early/deep pairs.

Table S27. Complete delay counts and paired changes against the same-delay baseline.

Centre set	Delay (s)	Controller	PASS	FAIL	UNKNOWN	Recovery	Regression
Extension	0	Baseline	0	1693	0	0	0
Extension	0	Early	50	1643	0	50	0
Extension	0	Deep	50	1643	0	50	0
Extension	1	Baseline	374	1319	0	0	0
Extension	1	Early	392	1301	0	36	18
Extension	1	Deep	371	1322	0	22	25
Extension	4	Baseline	1049	644	0	0	0
Extension	4	Early	1053	640	0	10	6
Extension	4	Deep	1046	647	0	4	7

Centre set	Delay (s)	Controller	PASS	FAIL	UNKNOWN	Recovery	Regression
Replication	0	Baseline	0	14	0	0	0
Replication	0	Early	0	14	0	0	0
Replication	0	Deep	0	14	0	0	0
Replication	1	Baseline	2	12	0	0	0
Replication	1	Early	3	11	0	1	0
Replication	1	Deep	2	12	0	0	0
Replication	4	Baseline	7	7	0	0	0
Replication	4	Early	7	7	0	0	0
Replication	4	Deep	7	7	0	0	0

Table S28. Direct early/deep centre outcomes; early is the first status and deep the second.

Centre set	Delay (s)	P/P	P/F	F/P	F/F	Deep−early FAIL
Extension	0	28	22	22	1621	0
Extension	1	360	32	11	1290	21
Extension	4	1044	9	2	638	7
Replication	0	0	0	0	14	0
Replication	1	2	1	0	11	1
Replication	4	7	0	0	7	0

All remaining direct cells involving UNKNOWN are zero. The extension's zero-delay tie changes to +21 and +7 deep-minus-early failures at one and four seconds. The corresponding replication contrasts are 0, +1 and 0. Baseline N3 exposure is 1,048/1,095/1,191 extension centres and 5/4/8 replication centres at zero/one/four seconds; each repair has the same exposure count in its condition. Target-domain exposure is zero throughout. No delay conditions or centre sets are pooled.

The extension baseline failures decrease from 1,693 to 1,319 and 644 as the delay changes. Selection of these centres as unsafe under immediate response makes this conditional behaviour unsuitable for claiming that delay improves safety. The centre speed relationships, straight-intruder restriction and ideal turn-rate response are unchanged. The result establishes sensitivity within the stated model; it does not establish a feasible aircraft envelope, a calibrated response distribution, an operational failure probability or a generally preferable repair.

The two stages check 64,515 outcome records with the original independent threshold procedure: 49,152 anchor protocol records and 15,363 centre-delay records. No resolved classification disagrees. Maximum source-transition/analytic minimum discrepancy is 8.01×10^{-12} ft (rounded upwards); 400,194 issued repaired-network invocations match independently loaded same-backend checkpoints, with zero advisory mismatches. These checks establish implementation consistency, not independent floating-point proof or physical validation. The interval-pruning optimisation uses the relative-speed Lipschitz bound and the sampled minimum to skip only segments that cannot attain the minimum; all 5,172 checked full-path minima equal the original unpruned routine exactly.

Execution uses Python 3.12.14, NumPy 2.2.6, SciPy 1.16.1, PyTorch 2.8.0, ONNX Runtime 1.22.1 and ONNX 1.19.0 on the same CPU platform, with one inference thread per process. The stopping and delay stages complete in 113.42 and 131.37 seconds respectively, running concurrently. No incomplete batch contributes to a result. The supplied analysis reconstructs every aggregate, paired table and family-resampling summary directly from the new case CSV files; the published Zenodo data and original study tables are preserved.