Supplementary Material S1

Computational verification for manuscript jmse-4556064. This supplement documents reporting and event-capacity corrections, matched enumeration, repeated sensitivity experiments, and the independent-arrival stress test. Sections S7 and S8 provide the second-round correction reconciliation and exact handling-duration rule.

S1 Reporting and capacity corrections

The original PSO object could retain an initialization-best fitness but leave its returned schedule empty. Reporting code converted that empty record to zero pairs. Re-decoding the final best vector reproduces the stored objective and recovers the actual pairs without changing search. Original-code replay reproduced all nine published scale-mean objectives to their displayed precision. It recovered mean pair counts of 1.8, 1.8, and 1.6 for S1, S2, and S3.

A separate event audit found Q = 9 over 46.091470–46.955816 h in the original S2 seed-3 schedule, although Q = 8 was available. Integer-hour samples missed that interval. The revised evaluator checks all starts and completions using half-open occupancy. The FCFS constructor uses the same event test. Source files were not overwritten; an explicitly versioned implementation records the corrections. The source PSO velocity update, objective, crane rule, pairing rule, and budgets remain unchanged. All 2,864 retained event-version job outputs passed the event-capacity check; this is aggregate resource feasibility, not crane-level safety certification.

The sign-test reconciliation treats objective differences with absolute magnitude no greater than 1e-9 as ties. This avoids interpreting floating-point roundoff as an improvement or deterioration. In the 45-instance FCFS comparison, six improvements and 39 ties give the corrected one-sided unadjusted p = 0.015625. This is distinct from the nine-block multiplicity-adjusted analysis.

S2 Objective component audit

Table S1. Means over five instances for each small configuration after event-capacity correction. Port time has unit weight; delay and indirect transfer are weighted objective components. Displayed rounding can cause a 0.01 discrepancy in sums. Per-instance schedules and decompositions are included in the source-data package.

Case	Scheme	Port time	Delay	Indirect transfer	Total
S1	FCFS-SLSM	49.79	3.36	229.40	282.56
S1	PSO-DLSM	45.22	0.00	91.00	136.22
S2	FCFS-SLSM	56.26	10.73	135.80	202.79
S2	PSO-DLSM	54.24	0.00	16.60	70.84
S3	FCFS-SLSM	67.58	8.03	118.80	194.41
S3	PSO-DLSM	61.65	0.00	0.00	61.65

S3 Matched decoder enumeration

Table S2. Every six-vessel instance has 64 berth vectors. Feasible vectors use the identical deterministic decoder and event-capacity evaluator as PSO. No independent crane, pairing, sequence, or timing choices are enumerated.

Case	Seed	Feasible vectors	Best enumerated	PSO	Gap (%)
S1	1	64	76.00	76.00	0.00
S1	2	48	204.00	204.00	0.00
S1	3	48	83.00	83.00	0.00
S2	1	64	49.00	49.00	0.00
S2	2	64	97.00	97.00	0.00
S2	3	56	104.91	104.91	0.00
S3	1	48	55.00	55.00	0.00
S3	2	40	69.23	69.23	0.00
S3	3	64	57.00	57.00	0.00

S4 Repeated sensitivity experiments

Table S3. Five M2 instances at every setting; 20 independent runs per cell. Between SD is the sample SD of the five cell means. Within SD is the mean of five cell-specific sample SDs and is reported separately, not pooled as additional independent instances. There are 700 efficiency and 600 penalty runs. Repeats use common NumPy seeds across settings, 1000000 + 10000s + r, with s = 1,...,5 and r = 0,...,19. The unchanged parameter uses its baseline value.

Factor	Value	Mean Z	Between SD	Within SD	Gain (%)	Pairs
eta	0.70	153.83	55.33	5.01	44.50	1.92
eta	0.75	153.01	55.45	5.00	44.82	1.92
eta	0.80	151.60	54.21	5.23	45.24	1.92
eta	0.85	137.27	32.60	5.57	48.72	2.10
eta	0.90	136.87	32.12	5.57	48.83	2.10
eta	0.95	136.30	32.00	5.57	49.06	2.10
eta	1.00	135.35	32.03	5.61	49.38	2.10
penalty	0.20	117.57	16.14	0.89	21.16	2.00
penalty	0.50	125.00	22.03	3.08	35.39	2.11
penalty	1.00	137.27	32.60	5.57	48.72	2.10
penalty	2.00	156.76	51.15	10.41	62.13	2.14
penalty	5.00	217.59	115.01	26.14	74.92	2.15
penalty	10.00	300.35	229.52	37.02	81.36	2.18

The raw cell table supplies best, mean, sample SD, worst, CV, runtime, and pair count. The plots use between-instance SD. No significance threshold or terminal economic break-even point is inferred from these descriptive sensitivity curves.

S5 Independent feeder arrival stress test

Each M2 base instance retains its vessel characteristics, transshipment connections, and departure allowance. Feeder arrival times are independently sampled from U(0,80) hours using RandomState(700000 + s). This removes the mainline-window conditioning but is not a fitted historical arrival model. Both PSO modes use the same 20 NumPy seeds per instance as the sensitivity analysis; FCFS modes are evaluated once. Five instances, not the 100 runs, are the scenario units.

Table S4. Independent-feeder-arrival stress test. PSO columns and pair counts are means over 20 runs; gain compares PSO-DLSM with PSO-SLSM within each instance.

Seed	FCFS SLSM	FCFS DLSM	PSO SLSM	PSO DLSM	Gain (%)	Pairs
1	198.92	198.92	198.92	198.92	0.00	0.00
2	394.77	394.77	394.06	302.78	23.16	1.00
3	197.73	197.73	192.43	180.74	6.08	0.70
4	220.71	220.71	220.71	220.71	0.00	0.00
5	314.11	314.11	314.31	314.74	-0.14	0.00

The mean relative reduction compared with PSO-SLSM is 5.82%. The absent gains and one small deterioration are retained, not excluded. These outcomes show sensitivity to overlap opportunities; five hypothetical instances cannot establish external validity.

S6 Reproducibility and scope

The package contains immutable original-model copies, the event-validated model, execution wrappers, raw schedules and objectives, input vessels and berths, recorded seeds, per-run runtimes, environment metadata, and statistical/figure scripts. The 45-instance ablation uses 10 runs per PSO scheme; the nine representative seed-3 instances use 20. Seed rules for these two analyses are unchanged from the previous experiments. Q sensitivity uses 390000 + r across all five capacities. Timings were collected with four concurrent worker processes and must not be interpreted as controlled serial benchmarks.

The principal nine-block mean relative improvement is 49.27% with a percentile 95% bootstrap interval of 44.24%–54.60% (10,000 block resamples; seed 221009). Full pairwise tests, normality diagnostics, multiplicity adjustment, and effect sizes appear in paired_tests.csv. Historical-terminal validation, a closed full-model exact benchmark, joint variable search, and individual-crane safety constraints remain outside the completed experiments.

S7 Complete Table 5 correction reconciliation

Table 5 reports one PSO-DLSM run per instance, averaged over the five instance seeds in each configuration. It is separate from the ten-run ablation and twenty-run stability experiments. All 45 old and revised records were compared using identical vessel inputs, berth inputs, parameters, and post-generator NumPy state. Only S2 seed 3 and M1 seed 2 changed objective values. Every FCFS-SLSM objective remained unchanged.

Table S5. The two changed single-run PSO-DLSM objectives and within-instance improvements. The denominator is the unchanged FCFS-SLSM objective for that instance. Twelve-vessel M1 seed 2 uses twelve cranes; six-vessel S2 seed 3 uses eight.

Instance	FCFS	Old PSO	Revised PSO	Old gain (%)	New gain (%)	Reason
M1-2	483.793775	199.334307	212.938192	58.797670	55.985752	Search-path change
S2-3	265.856389	94.951837	104.908530	64.284538	60.539399	Old final schedule overloaded

For M1 seed 2, the original final objective was 199.334306598 and the revised objective was 212.938191662. Its within-instance improvement changed from 58.797670% to 55.985752%, while FCFS remained 483.793775241. The other four M1 objectives were unchanged. The five-instance PSO mean therefore increased by 13.603885064/5 = 2.720777013, from 223.528036596 to 226.248813609. The mean of the five within-

instance improvements decreased by 2.811918168/5 = 0.562383634 percentage points, from 50.903322534% to 50.340938901%. These are not improvements computed from the ratio of configuration-mean objectives.

The M1 change is attributable to the event-capacity correction affecting the search path. A controlled replay verified identical initial positions, velocities, random-number inputs, 40 particles, and 150 iterations. Along the old path, 87 candidates accepted by integer-hour sampling violated event capacity. The first such candidate, initial particle index 12 (zero-based), used 13 cranes over [42.982157682, 43.000000000) h when 12 were available. The old and revised best-fitness histories first diverged at iteration 4. The old final vector itself remained event-feasible with objective 199.334306598. Thus, the corrected run returned a poorer feasible solution within the finite search budget; the old final M1 result must not be called infeasible or invalid. No new seed was selected. The replay JSON retains candidate records, final vectors, objective histories, and hashes of initial positions, initial velocities, and NumPy states.

For S2 seed 3, the old final schedule demanded nine cranes when eight were available over [46.091469898, 46.955815550) h. Its PSO objective changed from 94.951837103 to 104.908530102, and improvement from 64.284538% to 60.539399%. The other four S2 objectives remained unchanged. This changes the configuration means from 68.846035132 to 70.837373731 and from 65.435251279% to 64.686223351%.

Table S6. All changed displayed cells in the original Table 5. SD is the population SD of the five within-instance improvement percentages, as used in the unchanged Table 5 aggregation. Fields not listed, including every FCFS objective and all M2, M3, L1, L2, and L3 cells, are unchanged.

Configuration	Field	Original	Revised
S1	DLSM pairs	0.0	1.8
S2	PSO-DLSM	68.85	70.84
S2	Improvement	65.44%	64.69%
S2	SD	5.97%	6.30%
S2	DLSM pairs	0.0	1.8
S3	DLSM pairs	0.4	1.6
M1	PSO-DLSM	223.53	226.25
M1	Improvement	50.90%	50.34%
M1	SD	6.04%	5.37%

The recovered DLSM pair counts changed only for S1, S2, and S3. They reflect final-best-vector recovery, including initialization-best cases formerly reported as empty schedules, and not added pairing variables or altered pairing rules. The complete per-instance objective, component, gain, capacity, and pair reconciliation is in table5_all_instance_reconciliation.csv. The archived original and event-validated JSONL records supply schedules and vessel inputs.

S8 Exact handling-duration rule

The code applies int(np.ceil(workload / rate)). For a mainline vessel the rate is its assigned conventional-crane count multiplied by the per-crane rate. For an SLSM feeder the rate is the per-crane rate. For a paired outer feeder the rate is the cross-over efficiency multiplied by the per-crane rate. Each positive workload thus produces a positive integer duration, consistent with Table 1 and Equation (23). Equations (8)-(10) are workload lower bounds; the constructive scheduler chooses the smallest positive integer satisfying the applicable bound. Start times are not rounded, and completion time is start time plus the integer duration.

The rounded cross-over duration is computed before the outer feeder is tested for completion within the mainline service interval. For example, a workload of 140 TEU at efficiency 0.85 and rate 30 TEU/h requires ceil(140/25.5) = 6 h, not 5 h. This is an illustration of the implemented arithmetic, not an added experimental

observation. An independent audit of all 35,184 stored service entries in the 2,864 event-validated job records found no duration discrepancy at tolerance 1e-9 h. Ceiling rounding and event-capacity checking are separate rules; the latter uses continuous start/completion events with half-open occupancy.

Accompanying files are Supplementary_Material_S1_round2.docx and jmse-4556064_round2_code_data.zip. The archive provides a README, file-level SHA-256 manifest, the full verified experimental record, the original records needed for reconciliation, and the round-two controlled replay. Historical superseded outputs are labeled and are not sources for the current manuscript results.