Simulation-Driven Optimization of Load Sequencing and Explainable Hoist Dispatching in a Multi-Hoist Anodizing Line

Taha Mankır[1], Sanem Aydar[2]

[1]Department of Industrial Engineering, Faculty of Engineering, Çukurova University, Adana, Türkiye & Engineering Intern, R&D Department, Zahit Alüminyum San. ve Tic. A.Ş., Adana, Türkiye
[2]R&D Engineer, R&D Department, Zahit Alüminyum San. ve Tic. A.Ş., Türkiye

Abstract

Capacity-one baths and non-passing hoists can leave completed loads in an anodizing bath when downstream capacity or rail access is unavailable. This study addresses the limited separation of day-ahead sequencing from event-level hoist dispatch under stochastic processing and transport times. A field-informed discrete-event simulation represents four overlapping hoists, route-dependent processing, tank blocking, handovers, and shared-rail interference. Simulated annealing reorders the daily list, while the proposed explainable risk- and blockage-aware (RBA) hoist-dispatch rule ranks feasible tasks using process-risk, downstream-blockage, age, and guard information. Five historical production days containing 549 loads were evaluated with four paired seeds, yielding 40,000 simulated-annealing candidates and 80,000 equal-budget random and greedy comparators. Keyed common random numbers supported paired comparisons. Relative to the Current-rule baseline, RBA alone reduced net overstay by 16.386% while increasing makespan by 3.320%. Integrated RBA and simulated annealing reduced net overstay by 34.714%, 90th-percentile cycle time by 10.895%, and work in process by 10.728%, with a 1.784% makespan increase. Simulated annealing outperformed random search at the tested budget but did not universally dominate greedy search. A fit-excluded holdout replay of 100 anodizing loads produced 2.42% absolute relative error between observed and simulated aggregate means and 8.01% position-level symmetric mean absolute percentage error. Net overstay is an operational congestion-risk indicator; validation is limited to the process level rather than the full line.

Keywords: Anodizing; Discrete-Event Simulation; Hoist Scheduling; Simulated Annealing; Common Random Numbers; Explainable Dispatching

1. Introduction and Research Gap

1.1. Industrial Motivation

Anodizing lines move product loads through chemical and finishing stages, including sealing [1, 2, 3]. Route and duration in the studied plant depend on surface finish, color, and coating-thickness recipe. A load can leave a bath only when its next destination and an eligible hoist are available. Completion of nominal processing therefore does not necessarily release the occupied tank: a full destination, blocked handover, or neighboring hoist on the shared rail can keep a completed load in its source bath and propagate delay upstream.

The plant has four hoists serving overlapping zones on one rail. They cannot pass each other, every physical tank has capacity one, and six route families create different demands for pretreatment, anodizing, coloring, and sealing. Recipes of 7, 13, and 25 µm further change occupancy. Field records showed residence times materially exceeding expected processing durations even on the primary validation day, when no dip cleaning, tank transfer, equipment interruption, or other reported outage occurred. The observed bottleneck was therefore associated with the joint state of route mix, tank occupancy, and hoist accessibility rather than one permanently capacity-constrained station.

The practical question was whether congestion could be reduced before adding a tank or hoist. Two decisions coexist: the daily load list determines the sequence of recipe demands before production, whereas each hoist selects among physically feasible transport requests during operation. Since thousands of candidate daily orders cannot be tested on the line without disruption, a stochastic discrete-event simulation (DES) is used as a counterfactual evaluator. Congestion is measured by net overstay, defined as positive source-bath residence beyond assigned processing time after a normal handling allowance of 45 s. It is an operational occupied-capacity and process-risk indicator, not an observed defect probability or direct quality outcome.

1.2. Related Research

Movement among tanks under processing-time, resource, and transport constraints is commonly studied as the Hoist Scheduling Problem (HSP). HSP variants differ materially with physical layout, production policy, and process assumptions [4]. The present setting is multi-product, multi-hoist, predictive, and non-cyclic: it concerns a heterogeneous daily list rather than a short, indefinitely repeated cycle.

Industrial studies establish the value of simulation-based scheduling in surface-treatment lines. Basán et al. evaluated heuristic strategies for a complex multi-recipe hoist line [5]. Most closely, Alsaffar and Suliman modeled a real extrusion-anodizing line with five overlapping hoists and compared sequence rules and tank configurations [6]. Their emphasis was makespan and production capacity, with overlap represented by a simplified hoist-selection rule; they did not isolate an explainable risk-and-blockage-aware dispatch rule from day-ahead sequencing under matched stochastic realizations.

Cyclic and deterministic HSP research rigorously treats recipes, time windows, re-entry, capacity, collision avoidance, and hoist count [7, 8, 9, 10, 11, 12]. More recent work couples task sequencing and hoist scheduling in multi-variety electroplating [13], while fuzzy formulations represent immersion-time preferences and productivity–quality trade-offs [14]. These models clarify the coupling between sequence, transport, and time-sensitive processing, but their periodic or deterministic assumptions differ from the stochastic historical-day planning problem studied here.

Simulation optimization provides a general architecture for evaluating candidate schedules when analytic evaluation is impractical [15]. Fair comparison is harder when different sequences change event-call order. Common random numbers (CRN) reduce comparison noise [16, 17], but an ordinary shared random stream can lose matching when calls diverge. The experiments below bind remaining transport randomness to load, operation, and logical route step. Validation is likewise bounded by intended use [18]: process replay can support offline comparison without implying that every full-line performance measure has been validated.

1.3. Research Questions and Contributions

The unresolved need is to separate and then integrate day-ahead ordering of a heterogeneous, non-cyclic load mix and event-level selection of the next feasible hoist task under stochastic processing and transport times. Evidence is also limited on whether an interpretable risk-and-blockage rule adds value beyond sequence optimization under matched randomness and equal evaluation budgets. Four questions are addressed:

1. RQ1—Dispatch effect: How does RBA change net overstay and 90th-percentile (P90) cycle time relative to the Current rule under the same empirical order, and what makespan trade-off occurs?
2. RQ2—Sequencing effect: Within each dispatch rule, how much does simulated annealing (SA) improve the baseline sequence?
3. RQ3—Mechanism: Which RBA components—downstream blockage, process risk, task age, and guards—drive its effects?
4. RQ4—Search value: Under 1,000 candidate evaluations, how does SA compare with independent random search and greedy swap search?

The contributions are a field-informed DES of capacity-one tanks and four overlapping, non-passing hoists; an explainable RBA rule applied only to physically feasible tasks; a factorial integration of event-level dispatch with SA-based daily sequencing without added physical capacity; and an evidence design combining keyed CRN, equal-budget comparators, 540 RBA component/sensitivity replays, objective sensitivity, day-clustered uncertainty, reproducibility manifests, and process-level holdout validation. The study does not claim a new SA variant, global optimality, a complete digital twin, or direct defect prediction.

2. Industrial System and Data

2.1. Production Line and Decision Problem

The model boundary extends from entry to exit and includes degreasing, route-specific pretreatment, neutralization, anodizing, optional coloring, sealing, hot-water treatment, draining, and oven drying. The six route families are Bright, SAKEM, PAN, Natural, MAN, and AcidMat. A coating designation of 7, 13, or 25 µm is a recipe attribute rather than an observed quality result; for example, a load processed under a recipe of 13 µm may be accepted when its measured thickness falls within the customer's specification.

The four hoists retain physical order and overlap at rinse and handover stations. Hoist 1 covers entry and pretreatment, Hoist 2 neutralization and anodizing, Hoist 3 coloring and sealing, and Hoist 4 hot water, draining, ovens, and exit. The model enforces a minimum separation of 0.05 m. The plant has six installed anodizing tanks, but routine operation keeps at most five active and one available as a spare during dip cleaning; the DES therefore represents five simultaneously active anodizing resources.

Table 1 summarizes the field system and corresponding model boundary.

Table 1. Industrial System and Model Scope

Dimension	Field System and DES Representation
Product flow	Entry; degreasing; route-specific pretreatment; neutralization; anodizing; optional coloring; sealing; hot water; draining; ovens; exit
Route families	Bright, SAKEM, PAN, Natural, MAN, and AcidMat; recipes of 7, 13, or 25 µm
Transport	Four overlapping-zone hoists on one rail; fixed order; 0.05 m gap; shared handovers

Dimension	Field System and DES Representation
Parallel resources	Two degreasing and two neutralization tanks; six installed/five effective anodizing tanks; two tin and one nickel coloring tanks; three sealing tanks; three hot-water tanks; two ovens
Capacity and handling	One load per tank; destination reserved before pickup; draining while Hoist 4 holds the load
Decisions	Offline daily-sequence permutation and online selection among feasible transport tasks

2.2. Data Lineage and Stochastic Inputs

Field records were converted to source-cell-traceable long format without altering the raw workbook. Cleaning recovered 210 text-form durations and logged 297 parsing corrections. Anodizing records were checked against plant-confirmed capacity-one and five-active-bath constraints. Forty-four violating records remain traceable but were excluded from fitting and validation; three occurred on 31 March, leaving 100 physically consistent intervals for the primary holdout. Those exclusions affected only the observational validation set: the 31 March planning instance retained all 103 loads and assigned them modeled, frozen durations.

Recorded entry and exit denote physical bath residence, which includes intrinsic processing and post-completion handling. Directly treating the full interval as DES service time would double count transport delay. A mean lifting component of 25 s was therefore removed from anodizing and applicable physical-bath inputs. Rinse and deionized-water durations were not adjusted. Sealing—named Tespit in the implementation—was assigned the standardized recipe thickness in minutes.

Input modeling used a parametric–empirical hybrid. Positive-support parametric candidates were fitted and checked with 1,000-repetition composite-null Kolmogorov–Smirnov and Cramér–von Mises bootstraps [19]. Sparse or physically implausible groups used empirical resampling or an engineering formula. Anodizing was stratified by recipe; the heterogeneous 13 μm group was further split by route with a pooled empirical fallback, while loads with recipes of 7 and 25 μm used pooled models. The 31 March holdout was excluded from anodizing fitting, model selection, bootstrap testing, and plots.

Five historical days supplied load composition, route, recipe, and active-step masks and contained 122, 116, 108, 100, and 103 loads (549 total). One frozen vector of applicable process-duration realizations was generated for each load with seed 20260807, comprising 12,627 auditable load–operation assignments shared by Current, RBA, and all search seeds. Durations moved with loads during sequence optimization, while interarrival times remained attached to list positions (slot-fixed arrivals). Thus, an optimizer could change recipe order but could not gain an advantage by moving favorable arrival gaps.

3. Integrated DES–RBA–SA Method

3.1. Discrete-Event Simulation and Measures

The SimPy-based DES [20] advances from the first load arrival until every planned load reaches exit. Each station is a capacity-one resource with explicit location and occupancy. At process completion, a transport request is published; its destination is reserved before pickup, and its source is released at pickup. Travel depends on distance and measured hoist speed, while pickup and drop-off times are stochastic. Rail control preserves hoist order and blocks movement that would violate the safety gap or an adjacent active path.

For station visit v, with exit e_v, entry b_v, assigned processing time p_v, and normal handling allowance h=45\ s, net overstay is

$$O_v = \max\{0, (e_v - b_v) - p_v - h\} \quad (1)$$

Total net overstay sums O_v over completed visits. Other outputs are makespan from first arrival to last departure, completed-load count, the 90th percentile (P90) of load cycle time, and time-averaged work in process (WIP), calculated as the sum of completed-load cycle times divided by makespan.

3.2. Current and RBA Dispatching

Current denotes the plant's existing hoist-dispatch rule, whereas RBA denotes the proposed risk- and blockage-aware event-level hoist-dispatch rule. The implementation and archived experiment manifests retain the internal identifier S7 for the RBA rule. Both rules receive the same feasible candidate set after checks for hoist-zone eligibility, source occupancy, destination availability or reservation, rail order, path clearance, and capacity one. Current then ranks requests by effective priority, entry/passive urgency, primary-hoist ownership, hoist distance, ready time, and creation order.

RBA instead uses the explainable score

$$S_i = w_i O_i + B_i + 0.2A_i + C_i + H_i \quad (2)$$

where O_i is current net overstay in seconds, A_i is transport-queue age, $0.5 \leq w_i \leq 5$ is a process-risk weight, B_i is downstream-blockage priority, C_i is a critical-process boost, and H_i balances parallel hot-water tasks. The risk weights combine historical outcome mass and severity with an engineering mapping between defect categories and plausible process groups; they are a decision proxy, not a causal quality model.

After positive net overstay, B_i adds 150 points for an occupied base destination, 150 for no alternative buffer, 200 when another task waits in the same upstream hoist zone, and 250 for a predefined critical chain, capped at 800. C_i adds 300 for caustic or AcidMat tasks exceeding 30 s and retaining at least 40% of the best base score, and 200 for sealing tasks exceeding 60 s and retaining at least 50%. For no-wait-critical caustic and stripping sources, a guard restricts selection to protected tasks after 180 s when their score reaches 50% of the best score, or unconditionally after 600 s. H_i adds half of positive overstay above the eligible hot-water median, capped at 300. RBA chooses the highest score, with ready time and creation order as tie breakers. Components were evaluated through ablation rather than retuned on the five final days.

3.3. Simulated-Annealing Sequence Optimization

SA applies the Metropolis search principle [21] outside the DES. A solution is a permutation of the day's load rows. Each iteration swaps two sampled positions and evaluates the candidate with a complete DES run under Current or RBA. The candidate is compared with its unoptimized day–dispatch–seed baseline through

$$Z = 10^6 V_M + 10^5 V_T + 10^4 V_C + 100 R_O + 10 R_{P90} + R_W \quad (3)$$

Here, $V_M=\max(0,M/M_0-1)$ penalizes makespan deterioration, $V_T=\max(0,N_0-N)$ penalizes incomplete throughput, and V_C sums positive relative increases in net overstay across AcidMat, stripping, caustic, anodizing, tin, nickel, and sealing. R_O, R_{P90}, and R_W are candidate-to-baseline ratios for total net overstay, P90 cycle time, and WIP. The first terms are strong hierarchical penalties, not mathematical hard constraints. Where a critical-process baseline is zero, zero candidate overstay contributes zero; a positive value enters in seconds instead of an undefined ratio and is strongly penalized.

Improving candidates are accepted; a non-improving candidate with $\Delta=Z'-Z$ is accepted according to

$$P(\text{accept}) = \exp\left(-\frac{\Delta}{T_k}\right), T_k = 10\alpha^k, \alpha = \left(\frac{1}{10}\right)^{\frac{1}{1000}} = 0.997700 \quad (4)$$

Each run used 1,000 iterations, one candidate per iteration, cooling from 10 to 1 without early stopping. Initial evaluation and final confirmation replay were excluded from every method's candidate budget.

3.4. Keyed Common Random Numbers

Process-duration realizations were frozen at input. Remaining pickup, drop-off, and special oven-rule draws were deterministic functions of simulation seed, load identity, operation, logical route-step index, and context, using a SHA-256-derived uniform value and the required inverse distribution. This load-keyed CRN design applies established pairing principles [16, 17] without silently reassigning a favorable duration when event order changes. Seeds 42–45 provide stochastic replications under the same contract.

4. Verification, Validation, and Experimental Design

4.1. Verification and Process-Level Validation

Static contract checks, unit tests, deterministic replays, and dynamic smoke tests verified route order, optional branches, capacity-one behavior, reservation and release semantics, non-passing motion, safety separation, handover ownership, duration precedence and fallback, absence of overlapping visits, completion, and equality of fast and detailed KPI paths. All 33 checks passed. Claims were bounded by the intended decision-support use [18].

Validation was separated from fitting. The primary replay used the 31 March sequence, 100 physically consistent anodizing intervals, and 30 stochastic replications. Actual residence was the recorded bath entry-to-exit interval. Reported measures were bias, absolute relative error between aggregate means, load-position MAE and RMSE, symmetric mean absolute percentage error (sMAPE), and empirical 90% prediction-interval coverage.

A supporting validation froze duration banks before replaying 26, 27, and 30 March with 30 replications across 11 non-rinse process groups. Rinses and deionized water were excluded because field and model semantics were not sufficiently comparable. A 10% absolute mean-error review threshold was fixed before reporting, not used to remove observations. With no matched full-line departure, WIP, and synchronized event records, this is partial process-level validation rather than full-system KPI validation.

4.2. Experiments, Inference, and Reproducibility

The final experiment crossed five days, two dispatch rules, and four seeds, giving 40 SA runs and 40,000 candidate evaluations. Equal-budget benchmarking added 40 random-search and 40 greedy-swap runs. Random search sampled an independent uniform permutation at each evaluation; greedy search used the same two-position swap neighborhood as SA and accepted only strict improvements. Each method received 1,000 candidates per day–dispatch–seed cell, giving 120 runs and 120,000 candidates overall. Table 2 consolidates the frozen design, robustness checks, validation, and reproducibility evidence.

Table 2. Experimental Design and Evidence Structure

Element	Frozen Design
Planning instances	Five days; 122, 116, 108, 100, and 103 loads; 549 total
Dispatch comparison	Current versus RBA on identical daily input and seed

Element	Frozen Design
Search methods and budget	SA, independent random, and greedy swap; 1,000 candidate evaluations per cell
Stochastic design	Seeds 42–45; frozen load durations; slot-fixed arrivals; keyed transport/oven CRN
RBA robustness	27 scenarios and 540 keyed-CRN replays
Objective sensitivity	Nominal objective, 16 alternative weight structures, and one lexicographic comparator; 40,000 candidates plus 40 initial solutions
Validation	Anodizing holdout of 100 loads × 30; three production days and 11 supporting process groups × 30
Reproducibility	Manifests; input, output, and code hashes; seeds; logs; acceptance tests

Seeds are technical replications, not independent days. Contrasts were paired by day and seed, then averaged within five day clusters so each day had equal influence. Uncertainty used 10,000 day-cluster bootstrap resamples with seed 20260806 [19], and exact two-sided sign-flip tests used the five paired day-mean differences. The smallest attainable two-sided p-value is therefore 0.0625, so effect sizes, intervals, and directional consistency receive priority.

Machine-readable manifests store configurations, code and input hashes, seeds, budgets, and output hashes. Distributed shards were merged only after run-count, identity, hash, and corruption checks. A cross-package audit linked the keyed-CRN SA, comparators, ablation, objective sensitivity, input lineage, DES verification, and validation evidence; all 16 upper-level acceptance checks passed across 21 source artifacts and eight claim–evidence links.

5. Results and Robustness

5.1. Effects of Dispatching and Sequencing

Every baseline and best-found solution completed its assigned 100–122 loads. Table 3 reports day-equal means. Changing only dispatch from Current to RBA reduced net overstay by 16.386%, P90 cycle time by 6.714%, and WIP by 7.520%, while makespan increased by 3.320%. Net overstay improved in all 20 paired day–seed comparisons and all five day clusters; its day-clustered 95% bootstrap interval was −20.876% to −11.649%. RBA therefore reduced congestion-related measures but did not simply accelerate the line.

Table 3. Day-Equal Performance of the Sequencing and Dispatching Combinations

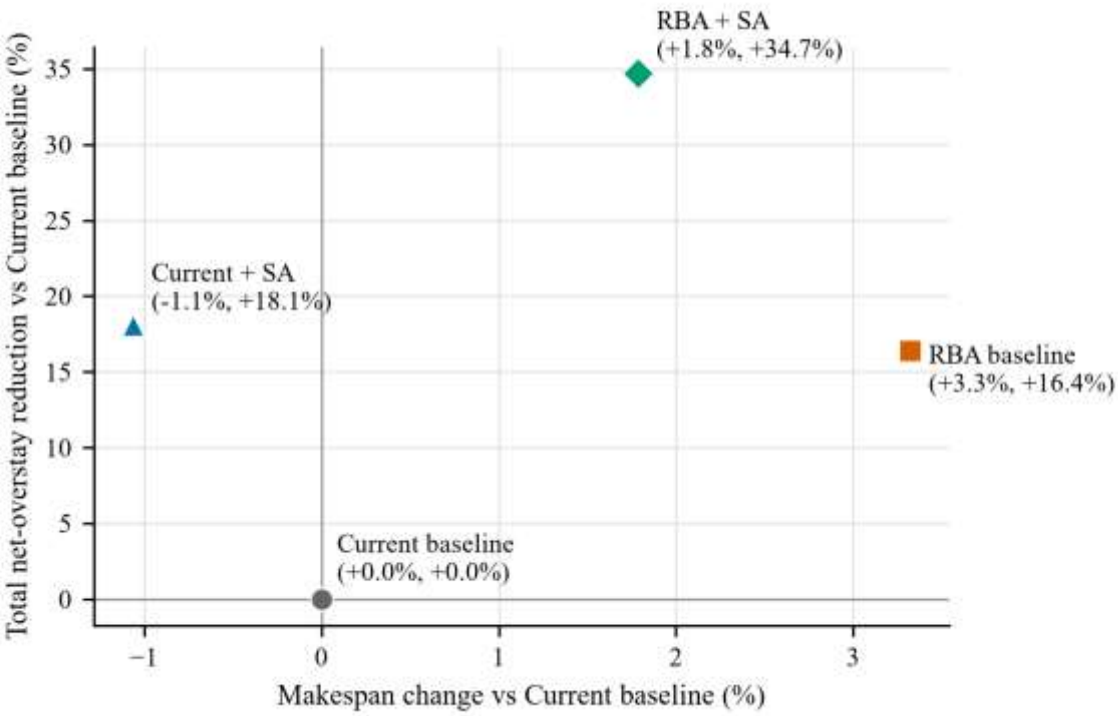
Configuration	Runs	Loads	Makespan (h)	Overstay (h)	P90 (min)	WIP	Δ Overstay
Current baseline	20	109.8	20.11	64.42	176.71	12.55	Reference
Current + SA	20	109.8	19.90	52.74	165.14	12.07	−18.079%
RBA baseline	20	109.8	20.79	53.20	164.36	11.60	−16.386%
RBA + SA	20	109.8	20.47	42.06	156.80	11.20	−34.714%

Note: Values are day-equal means across five days and four paired seeds per day. Net overstay excludes the first 45 s of post-process handling. Seeds are within-day stochastic replications, not independent field days.

SA improved both dispatch-specific baselines. Under Current, net overstay fell 18.079% (95% CI: −20.982% to −15.458%), P90 cycle time 6.255%, makespan 1.064%, and WIP 3.767%. Under RBA, the respective reductions were 21.451% (95% CI: −28.014% to −14.737%), 4.496%, 1.471%, and 3.447%. Net overstay improved in 17/20 runs and tied in three under each rule; none worsened.

Integrated RBA + SA gave the largest congestion reduction. Relative to Current baseline, net overstay fell 34.714% (95% CI: −37.876% to −31.300%), P90 cycle time 10.895%, and WIP 10.728%, while makespan increased 1.784%. Improvement occurred in 20/20 pairs and 5/5 days. Relative to Current + SA, RBA + SA reduced overstay a further 19.137% but increased makespan 2.882%. Figure 1 therefore shows both effects. With five independent day clusters, the consistent direction yields an exact two-sided sign-flip p-value of 0.0625, not a claim of significance at 5%.

Figure 1. Net-Overstay Reduction and Makespan Trade-Off Relative to the Current Baseline



5.2. Equal-Budget Search Comparison

Table 4 compares 40 solutions per method at 1,000 candidate evaluations per cell. SA achieved lower mean best objective Z than random search under both dispatch rules. Random-minus-SA was 9.058 under Current (95% day-clustered CI: 3.385–14.510) and 12.834 under RBA (6.022–21.682). SA was better in 5/5 days under both rules and in 15/20 and 16/20 day–seed pairs, respectively.

Table 4. Equal-Budget Comparison of Search Methods

Dispatch	Method	Improved	Best Z	Δ Overstay	Δ P90	Δ Makespan	Runtime (s)
Current	SA	17/20	92.258	−18.079%	−6.255%	−1.064%	528.8

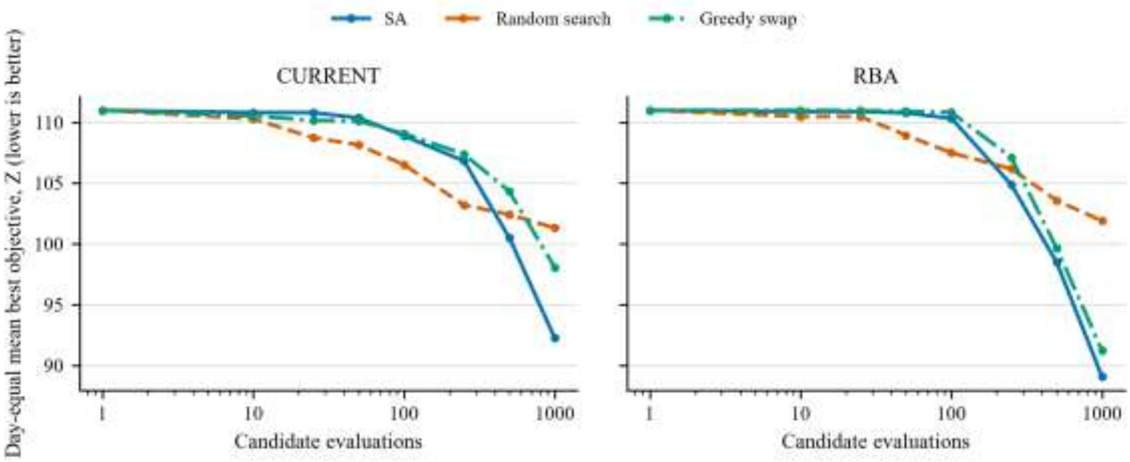
International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Dispatch	Method	Improved	Best Z	Δ Overstay	Δ P90	Δ Makespan	Runtime (s)
Current	Greedy swap	13/20	98.065	−12.500%	−4.087%	−0.731%	519.6
Current	Random search	12/20	101.317	−9.444%	−3.871%	−1.002%	550.2
RBA	SA	17/20	89.065	−21.451%	−4.496%	−1.471%	517.9
RBA	Greedy swap	14/20	91.226	−19.368%	−3.728%	−1.068%	498.2
RBA	Random search	10/20	101.900	−8.804%	−2.806%	−0.538%	529.4

Note: Lower Z is better. Changes are relative to the dispatch-specific baseline. "Runs improved" applies the frozen strict best-Z criterion and is a diagnostic count, not an effect size.

SA's advantage over greedy search was less uniform. Greedy-minus-SA was 5.807 under Current (95% CI: 1.108–8.979), with SA better in 4/5 days, but 2.161 under RBA (−1.204–5.526), with SA better in only 2/5. The evidence supports an advantage over random search at this budget, not universal dominance over greedy search. Runtimes were similar in magnitude, so no speed advantage is claimed. Figure 2 shows that most separation occurred later in the budget.

Figure 2. Best-So-Far Convergence Under the Equal Candidate-Evaluation Budget



5.3. RBA Ablation and Sensitivity

The 540 keyed-CRN replays covered 27 reference, leave-one-component-out, sensitivity, and interaction scenarios. Current and full-RBA reference cells matched the final baseline with a maximum absolute KPI difference of 1.17×10^{-10}. Table 5 summarizes the principal component removals.

Table 5. Selected RBA Component Ablations Relative to Full RBA

Variant	Δ Overstay (s/day)	Δ Makespan (s/day)	Δ P90 (s)	Δ Sealing (s)	Days B/W
Without downstream-blockage priority	+19,105	−404	+302	+864	0/5

Variant	Δ Overstay (s/day)	Δ Makespan (s/day)	Δ P90 (s)	Δ Sealing (s)	Days B/W
Without process-risk weighting	+4,185	−507	+117	+707	2/3
Without guards	+3,095	−44	+146	+688	1/4
Without age term	−6,561	+585	+128	+162	3/2

Not: Positive differences are increases relative to full RBA; B/W denotes days better/worse than full RBA. Sealing is named Tespit in the implementation. These are mechanism diagnostics on the same five days, not independent retuning experiments.

Removing downstream-blockage priority caused the largest overstay deterioration and was worse on all five days. Removing process-risk weighting or guards also increased mean overstay and P90, although makespan shortened. Removing age reduced aggregate overstay but increased makespan and sealing overstay, revealing starvation and protection trade-offs. A diagnostic no-age variant with downstream-blockage cap 400 reduced overstay further but increased makespan by 638 s/day and worsened one of five days; it was not promoted as a new rule without independent validation. RBA's direction also persisted at handling allowances of 30, 45, and 60 s, with overstay changes of −17.353%, −16.386%, and −16.326%.

5.4. Objective and Constraint Diagnostics

Rescoring 40,000 candidates and 40 initial solutions under the nominal weighted objective, 16 alternative weight structures, and one lexicographic comparator (18 scenarios total) left 32/40 selections unchanged throughout. Scaling the penalty tiers, the critical-process weight, or all three secondary-KPI weights together changed no selection; one-at-a-time changes to total overstay, P90, or WIP altered 1–4 selections, while the lexicographic rule altered five. Every selection satisfied load-completion and makespan-acceptance checks. This diagnoses the observed candidate pool; it does not represent 18 new SA trajectories.

All 120 final search solutions completed their loads and passed makespan checks. Because critical terms are strong penalties rather than hard constraints, 119/120 had zero critical penalty. One Current/random solution increased anodizing overstay by 6 s (0.032946%) while reducing total overstay by about 22,668 s. The exception is retained rather than removed by a post-hoc tolerance.

5.5. Process-Level Validation

The anodizing holdout of 100 loads and 30 replications had observed and simulated mean residence times of 32.67 and 33.46 min. Bias was +0.79 min, and the absolute relative error between aggregate means was 2.42%. At load-position level, MAE was 2.66 min, RMSE 3.50 min, sMAPE 8.01%, and empirical 90% interval coverage 81%.

Of 11 supporting non-rinse groups, six met the predefined 10% mean-error review target: degreasing (1.53%), supporting anodizing (9.68%), sealing (7.54%), hot water (8.66%), draining (3.30%), and oven processing (4.37%). Acid matting (47.16%), stripping (315.93%), caustic etching (12.64%), neutralization (21.47%), and coloring (48.68%) did not. Stripping residence definitions were not equivalent, and coloring had seven observations including one long event. Accordingly, the evidence supports partial process-level validation and paired offline comparisons, not validation of full-line makespan, throughput, cycle time, WIP, total overstay, or product quality.

6. Discussion

6.1. Interpretation of the Decision Layers

RQ1 shows that RBA changes which completed loads and occupied tanks are protected when transport demands compete. Lower overstay, P90, and WIP came with longer makespan; locally limiting post-process exposure can delay final completion of a daily batch. RQ2 shows that sequence and dispatch are complementary, not interchangeable. The permutation shapes when route- and recipe-specific demand reaches each zone, while the dispatch rule governs the line's response once feasible tasks compete. Their integration produced the largest overstay reduction, but its 1.784% makespan cost remains part of the operational decision.

For RQ3, downstream blockage was the strongest RBA mechanism, supporting the field diagnosis of propagation through occupied destinations, missing buffers, and adjacent zones. Process-risk weighting and guards offered additional protection. Age did not merely reduce overstay: it controlled starvation and redistributed protection, explaining why a same-sample diagnostic with lower aggregate overstay was not adopted. For RQ4, SA achieved a lower best-Z than independent random search on all five days under both dispatch rules, but its advantage over greedy swap varied by rule and day. SA is thus a useful tested search layer, not a universally superior optimizer.

6.2. Relation to Hoist-Scheduling Research

Prior HSP research offers cyclic formulations for time windows, tank capacity, re-entry, and collision avoidance [4, 7, 8, 10, 11, 12], industrial simulations of sequence and resource configuration [5, 6], and dynamic models for overlapping hoists and parallel units [9]. The contribution here is not the first use of anodizing, multiple hoists, DES, or SA. It is the integration and experimental separation of non-cyclic day-ahead sequencing and explainable event-level dispatch under stochastic inputs and fixed capacity, consistent with the wider simulation-optimization architecture [15]. Table 6 summarizes this positioning.

Table 6. Positioning Relative to Closely Related Hoist-Scheduling Research

Research Stream	Representative Studies	Main Emphasis	Distinction of This Study
Cyclic and exact scheduling	Qu et al. [7]; Feng et al. [8]; Ptuskin et al. [10]	Cycle time, windows, capacity, re-entry, collision-free cyclic motion	Non-cyclic daily mix; fixed plant; stochastic DES
Industrial simulation	Basán et al. [5]; Alsaffar and Suliman [6]	Heuristic sequencing, makespan, production, configuration	Separate sequence and explainable dispatch effects without added capacity
Dynamic multi-hoist scheduling	Ramin et al. [9]	Deterministic logic, overlapping hoists, parallel units	Risk/blockage dispatch, paired CRN, ablation, replay validation
Present framework	This study	Overstay, cycle time, WIP, makespan, mechanism robustness	Two decision levels, equal-budget benchmarks, bounded validation

The keyed CRN design is material because alternative sequences alter event order. Binding draws to load, operation, and route step implements established CRN principles [16, 17] without allowing a sequential stream to assign different transport realizations after trajectories diverge. Its contribution is an auditable application in this event-driven setting, not new variance-reduction theory.

6.3. Industrial Implications

The framework is intended for offline planning. Given a daily load set, it can compare entry orders without interrupting production, while RBA can expose why a task is protected through overstay, blockage, age, guard, and destination information. The simulated reductions required no additional nominal tank or hoist capacity, making policy change a reasonable intervention to evaluate before capital expansion.

Deployment should preserve operator authority and display the makespan–overstay trade-off rather than return an unexplained “optimal” sequence. A prospective pilot should freeze a next-day list, generate several high-performing feasible alternatives, expose RBA reasons, and compare the selected plan with Current operation. Since runtimes were similar across methods, feasibility depends on the available planning window and hardware, not a demonstrated speed advantage.

6.4. Limitations and Further Work

Only five purposively selected days form independent clusters; four seeds per day quantify simulation variation but do not create 20 field days. Broader product mixes, seasons, and prospective trials are needed before generalizing effect sizes. Five of 11 supporting processes missed the 10% validation target, and the nominal 90% anodizing prediction interval covered only 81% of observed positions. Matched full-line KPI records were unavailable; breakdowns, maintenance, rush orders, initial WIP, and unrecorded interventions were outside scope. Accordingly, the model supports purpose-specific offline comparison, not a complete digital-twin claim [18].

The process-risk term is an explainable proxy based on historical outcome severity and engineering mapping. It neither estimates causal defect probability nor represents a fuzzy immersion-time satisfaction function such as that used in productivity–quality scheduling [14]. Future work should link load-level process histories to matched quality outcomes and customer specifications, validate RBA on new days, and investigate broader neighborhoods, larger budgets, and multiobjective presentation before claiming defect or economic benefits.

7. Conclusion

A two-level DES framework separated and integrated daily sequencing with event-level hoist dispatch in a stochastic, non-cyclic anodizing line. RBA alone reduced net overstay by 16.386% with a 3.320% makespan increase. SA reduced dispatch-specific overstay by 18.079% under Current and 21.451% under RBA. Integrated RBA + SA reduced net overstay by 34.714%, P90 cycle time by 10.895%, and WIP by 10.728% relative to Current baseline, with a 1.784% makespan increase. Same-budget tests favored SA over random search but not universally over greedy search; ablation identified downstream blockage as RBA's dominant component and revealed age-related starvation trade-offs. Within five production days and partial process-level validation, the evidence supports prospective evaluation of sequence and dispatch policies before adding physical capacity, not global optimality or direct quality improvement.

References

1. U.S. Department of Defense. Anodic Coatings for Aluminum and Aluminum Alloy, MIL-PRF-8625F with Amendment 2. 23 November 2020. https://quicksearch.dla.mil/qsDocDetails.aspx?ident_number=7074

2. QUALANOD. Specifications for the QUALANOD Quality Label for Sulfuric Acid-Based Anodizing of Aluminium. Edition 01.07.2026. Zurich, Switzerland, 2026. https://www.qualanod.net/specifications.html
3. Ofoegbu, S. U., Fernandes, F. A. O., Pereira, A. B. "The Sealing Step in Aluminum Anodizing: A Focus on Sustainable Strategies for Enhancing Both Energy Efficiency and Corrosion Resistance." Coatings, 10(3), 226, 2020. https://doi.org/10.3390/coatings10030226
4. Manier, M.-A., Bloch, C. "A Classification for Hoist Scheduling Problems." International Journal of Flexible Manufacturing Systems, 15, 37–55, 2003. https://doi.org/10.1023/A:1023952906934
5. Basán, N. P., Pulido, R., Cóccola, M. E., Méndez, C. A. "Aerospace Manufacturing Industry: A Simulation-Based Decision Support Framework for the Scheduling of Complex Hoist Lines." Iberoamerican Journal of Industrial Engineering, 9(17), 100–117, 2017. https://periodicos.ufsc.br/index.php/IJIE/article/view/v9n1706
6. Alsaffar, A. J., Suliman, S. M. A. "Hoist Scheduling for an Anodizing Plant at an Extrusion Company." Arab Journal of Basic and Applied Sciences, 27(1), 93–104, 2020. https://doi.org/10.1080/25765299.2020.1729931
7. Qu, H., Wang, S., Xu, Q. "A New Method of Cyclic Hoist Scheduling for Multi-Recipe and Multi-Stage Material Handling Processes." Computers & Chemical Engineering, 90, 171–187, 2016. https://doi.org/10.1016/j.compchemeng.2016.04.014
8. Feng, J., Chu, C., Che, A. "Cyclic Jobshop Hoist Scheduling with Multi-Capacity Reentrant Tanks and Time-Window Constraints." Computers & Industrial Engineering, 120, 382–391, 2018. https://doi.org/10.1016/j.cie.2018.04.046
9. Ramin, D., Fraizzoli, D., Ballarino, A., Brusaferri, A. "Dynamic Hoist Scheduling for Multi-Recipe and Multi-Stage Production Lines: A Logical Framework." Computers & Industrial Engineering, 182, 109360, 2023. https://doi.org/10.1016/j.cie.2023.109360
10. Ptuskin, A., Levner, E., Kats, V. "Cyclic Multi-Hoist Scheduling with Fuzzy Processing Times in Flexible Manufacturing Lines." Applied Soft Computing, 165, 112014, 2024. https://doi.org/10.1016/j.asoc.2024.112014
11. Laajili, E., Lamrous, S., Manier, M.-A., Nicod, J.-M. "An Adapted Variable Neighborhood Search Based Algorithm for the Cyclic Multi-Hoist Design and Scheduling Problem." Computers & Industrial Engineering, 157, 107225, 2021. https://doi.org/10.1016/j.cie.2021.107225
12. Yan, P., Wang, G., Che, A., Li, Y. "Hybrid Discrete Differential Evolution Algorithm for Biobjective Cyclic Hoist Scheduling with Reentrance." Computers & Operations Research, 76, 155–166, 2016. https://doi.org/10.1016/j.cor.2016.06.011
13. Chen, X., Yang, B., Pang, Z., Zhou, P., Fu, G. "Coupled Optimization of Task Sequence and Hoist Scheduling for Electroplating Production Lines Based on an Improved Salp Swarm Algorithm." CIRP Journal of Manufacturing Science and Technology, 53, 34–47, 2024. https://doi.org/10.1016/j.cirpj.2024.07.002
14. Fargier, H., Lamothe, J. "Handling Soft Constraints in Hoist Scheduling Problems: The Fuzzy Approach." Engineering Applications of Artificial Intelligence, 14(3), 387–399, 2001. https://doi.org/10.1016/S0952-1976(01)00008-2
15. Ghasemi, A., Farajzadeh, F., Heavey, C., Fowler, J., Papadopoulos, C. T. "Simulation Optimization Applied to Production Scheduling in the Era of Industry 4.0: A Review and Future Roadmap." Journal of Industrial Information Integration, 39, 100599, 2024. https://doi.org/10.1016/j.jii.2024.100599

16. Nelson, B. L., Matejcik, F. J. "Using Common Random Numbers for Indifference-Zone Selection and Multiple Comparisons in Simulation." Management Science, 41(12), 1935–1945, 1995. https://doi.org/10.1287/mnsc.41.12.1935
17. Kleinman, N. L., Spall, J. C., Naiman, D. Q. "Simulation-Based Optimization with Stochastic Approximation Using Common Random Numbers." Management Science, 45(11), 1570–1578, 1999. https://doi.org/10.1287/mnsc.45.11.1570
18. Sargent, R. G. "Verification and Validation of Simulation Models: An Advanced Tutorial." Proceedings of the 2020 Winter Simulation Conference, 16–29, 2020. https://www.informs-sim.org/wsc20papers/002.pdf
19. Efron, B. "Bootstrap Methods: Another Look at the Jackknife." The Annals of Statistics, 7(1), 1–26, 1979. https://doi.org/10.1214/aos/1176344552
20. SimPy Team. "SimPy Documentation: Overview." SimPy 4.1 Documentation. Accessed 7 August 2026. https://simpy.readthedocs.io/en/latest/
21. Kirkpatrick, S., Gelatt, C. D., Jr., Vecchi, M. P. "Optimization by Simulated Annealing." Science, 220(4598), 671–680, 1983. https://doi.org/10.1126/science.220.4598.671