Journal of Industrial and Management Optimization

JIMO, 22(12): 6248–6270.
DOI:10.3934/jimo.2026210
Received: 24 August 2026
Revised: 17 October 2026
Accepted: 26 October 2026
Published: 09 November 2026

https://www.aimspress.com/journal/jimo

Research article

Sparse LP rounding with resource-exchange descent for multi-item continuous-review inventory under dual shared-resource constraints

Qingbo Lu[1,*], Shuangcheng Huang[1], Mengzhen Huang[2] and Daojun Li[3]

[1] School of Intelligent Manufacturing and Equipment, Henan Technical Institute, Zhengzhou 450042, Henan, China
[2] Department of Mechanical and Electrical Engineering, Zhoukou Vocational and Technical College, Zhoukou 466000, Henan, China
[3] School of Automation and Internet of Things, Zhengzhou Technical College, Zhengzhou 450121, Henan, China

*** Correspondence:** Email: lqb0932@hati.edu.cn; (+86) 15038310825.

Abstract: Multi-item continuous-review inventory optimization couples integer replenishment policies, item-level service requirements, and shared average-resource limits. Sparse linear programming rounding with resource-exchange descent (LP-SERD) was developed for a finite single-objective planning model with full backordering. Mean inventory and one-time shortage costs were evaluated under an explicit uniform-inventory-position approximation. Service-feasible item policies were enumerated and dominance-pruned to form an exact multiple-choice master for that model. The standard LP rank bound limits fractionality to the independent active resource rows. Common within-item resource ordering guarantees a feasible positive-support corner and an additive support-span cost bound. Two independent descent paths refine this start using single-item moves and directional two-item exchanges. Across 15 instances with 500–2000 items, the pooled median LP-certified gap was 0.000701%, and the maximum was 0.004148%. On 35 instances, LP-SERD had 28 lower-cost, 4 tied, and 3 higher-cost endpoints than repaired maximum-weight rounding followed by single-item descent. At a 60 s solver limit, full MILP obtained lower costs on 14 of 15 instances and closed 3 optimality proofs. A paired correlation experiment did not establish a monotone association between resource-coefficient correlation and exchange gain. All 10 OIL main models had integral LP optima; MAN exposed both small optimization gaps and holdout-service shortfalls. The evidence supports a feasible construction with a finite-model certificate, not a general replacement for tractable exact optimization.

Keywords: multi-item inventory; continuous review; shared resources; LP rounding; multiple-choice knapsack; maintenance spare parts

Mathematics Subject Classification: 90B05, 90C10, 90C27, 90C59

1. Introduction

Maintenance systems balance spare-parts availability against inventory investment and storage capacity. In multi-item warehouses, replenishment policies are item-specific while these resources are shared. Continuous-review policies respond to inventory position, and supplier pack multiples and integer reorder points discretize their decisions. The resulting planning task is a resource-coupled nonlinear integer single-objective problem rather than a collection of independent item calculations.

Existing multi-item continuous-review models incorporate stochastic demand, backordering, budget and storage capacity constraints, and parameter ambiguity [1–4]. Data-driven robust formulations protect a decision against demand uncertainty [5], and maintenance-inventory lot-sizing models link stocking decisions to equipment condition and production planning [6]. Metaheuristics, goal programming, and commercial mixed-integer solvers have also been applied to resource-constrained inventory allocation [7,8]. These formulations motivate a computational question: under what inventory structure can a continuous relaxation be converted directly into a feasible integer policy, and what cost increase can be certified?

Studies of maintenance inventories highlight intermittent demand, part criticality, and policy-transition effects. Reviews identify the persistent gaps between classification of spare parts, forecasting, and inventory control [9]. Transition-control methods can minimize the impact of the transition to replace an incumbent policy [10], and stochastic mixed-integer formulations can model finite spare-parts capacity [11]. Joint maintenance–inventory and production–inventory models introduce additional (operative) couplings [12,13]. Whereas these extensions improve description fidelity, they broaden the discrete search space and exacerbate the need to distinguish optimization error from demand-model error.

Enumerating the possible policies of each item turns the system-level allocation problem into a multiple-choice multidimensional knapsack problem [14]. LP-guided reduce-and-solve and kernel-search methods use fractional values or reduced costs to define restricted integer subproblems, which are then solved by branch-and-bound or iterative kernel procedures [15,16]. Recent surveys describe exact, approximate, and heuristic techniques for this problem class [17,18], and comparative evidence shows that general-purpose MILP solvers can be competitive on appropriate instances [19]. Iterative semi-continuous relaxation combines rounding, repair, and local improvement [20]. Related studies on capacity acquisition and minimum-quantity lot sizing establish the value of discrete-quantity and capacity structure [21,22], while hybrid evolutionary methods have been applied to replenishment problems [23]. The question examined here is whether a checkable ordering of resource use within each fractional support can yield a feasible integer policy and a structural cost-loss bound. If the selected support contains no feasible integer corner, repair or a larger integer search is required.

The approach exploits the standard rank structure of the multiple-choice LP relaxation [24]: with only a few active shared-resource constraints, an optimal basic solution can be fractional in only a few item groups. Within the present inventory model, investment and storage for a given item are positive multiples of the same average-inventory measure. Their rankings are therefore identical within each item, even when investment-to-storage coefficient ratios differ across items. Selecting the minimum-inventory policy in each fractional support does not increase either resource use. This feasible corner provides a constructive link between the sparse LP solution and an implementable inventory policy.

The study makes three contributions. First, common within-item resource ordering establishes a feasible positive-support rounding construction and an additive cost-loss bound for the finite inventory-policy master. The standard LP rank argument explains why only a few item groups require rounding when few shared resources are active. Second, exact inventory-specific marginal changes connect this certified start to single-item and directional two-item local refinement. The generic exchange operator is established local search; the contribution is its integration with a checkable feasible construction and cost certificate under the stated inventory structure. Third, component comparisons, limited-time integer-search comparisons, and industrial-data studies examine the magnitude and limits of the resulting benefit. A separate paired experiment tests whether cross-item resource correlation predicts exchange gain. Common evaluators and explicit computation budgets are used throughout [25].

2. Problem formulation

2.1. System assumptions and policy measures

Consider a warehouse holding n maintenance spare parts under item-specific continuous-review (Q_i, r_i) policies. Unmet demand is fully backordered. Item-specific lead times are fixed, with no order crossover. The reference planning model assumes stationary demand increments and an inventory position uniformly distributed on $[r_i, r_i+Q_i]$, independently of lead-time demand [26,27]. Joint ordering, common-cause failures, and supplier-delay dependence are excluded. Normal and empirical lead-time-demand marginals approximate this reference model for intermittent or batch demand; integer decisions do not remove that process approximation. Shared capacities constrain long-run average inventory investment and equivalent storage.

For item i, D_i is annual demand, X_i is lead-time demand, A_i is cost per order, h_i is holding cost per unit per year, and p_i is a one-time penalty per newly backordered unit. Let G_i, μ_i, and σ_i denote the distribution function, mean, and standard deviation of X_i. The order quantity is a positive integer multiple of pack size m_i, and r_i is a nonnegative integer. Positive coefficients c_i and v_i convert average on-hand units into investment and equivalent storage; B and W are their system limits. The service requirement $G_i(r_i) \geq \beta_i$ concerns lead-time-demand coverage (cycle service in the reference model), not the demand fill rate. Φ and ϕ are the standard normal distribution and density functions.

Write $(x)_+ = \max(x, 0)$. The first-order loss function at threshold u is

$$N_i(u) = E[(X_i - u)_+] \tag{1}$$

Mean outstanding backorders are $\bar{B}_i(Q_i, r_i) = Q_i^{-1} \int_{r_i}^{r_i+Q_i} N_i(u)du$. The corresponding mean on-hand inventory is [26,27]

$$I_i(Q_i, r_i) = \frac{Q_i}{2} + r_i - \mu_i + \bar{B}_i(Q_i, r_i). \tag{2}$$

For a normal lead-time-demand approximation, $z_i = (r_i - \mu_i)/\sigma_i$ gives

$$G_i(r_i) = \Phi(z_i), \quad N_i(r_i) = \sigma_i[\phi(z_i) - z_i(1 - \Phi(z_i))]. \tag{3}$$

For an empirical distribution, the same expectations are finite averages over training lead-time-demand windows. The integral defining $\bar{B}_i$ can be evaluated from the difference of second-order loss moments. These expressions are exact for the specified marginal/uniform-position evaluator, not for every intermittent-demand process. As consistency checks, nonnegative deterministic demand $X_i = x$ gives $I_i = [(r_i + Q_i - x)_+^2 - (r_i - x)_+^2]/(2Q_i)$. Thus, $r_i \geq x$ recovers $I_i = Q_i/2 + r_i - x$ with zero

backlog, whereas $r_i+Q_i \leq x$ gives zero on-hand inventory. All measures satisfy $I_i - \bar{B}_i = Q_i/2 + r_i - \mu_i$.

2.2. Integer decisions and bounds

The design coordinates are the pack count k_i and reorder point r_i, with $Q_i = m_i\, k_i$. Write the decision vector as $y = (k_1, \dots, k_n, r_1, \dots, r_n)$. Quantity bounds represent supplier or engineering limits, while the service requirement tightens the reorder-point lower bound. They define

$$k_i^L = max\{1, ceil(Q_i^{min}/m_i)\}, \quad k_i^U = floor(Q_i^{max}/m_i) \tag{4}$$

$$r_i^L = max\{ceil(r_i^{min}), ceil(G_i^{-1}(\beta_i))\}, \quad r_i^U = floor(r_i^{max}). \tag{5}$$

The coordinates are independent: for each admissible r_i, every k_i in (4) is considered. No equation of the form $Q_i = Q_i(r_i)$ is imposed. A minimum order quantity bounds Q_i from below and need not be a multiple of the pack unless supplier data establish that interpretation. If either lower bound exceeds its upper bound, the specified domain is empty. In numerical studies, EOQ-based upper bounds define a finite search envelope; they are not proven lossless bounds for the unbounded inventory problem. Optimality and gap certificates therefore refer to the stated finite domain.

2.3. Objective and shared-resource constraints

The distinction between $N_i(r_i)$ and $\bar{B}_i$ is material: the first measures loss at one threshold, whereas the second averages the loss across the inventory-position interval. Since N_i is nonincreasing, replacing $\bar{B}_i$ with $N_i(r_i)$ overestimates on-hand inventory. The overestimate lies between zero and $Q_i[1 - G_i(r_i)]/2$. It therefore changes both shared-resource coefficients as well as annual cost.

Let $e_i(Q_i, r_i) = N_i(r_i) - N_i(r_i+Q_i)$. Under the reference model, e_i/Q_i is the stockout-time fraction. Interpreting $D_i\, e_i/Q_i$ as the annual rate of newly backordered demand additionally assumes independence of the demand arrival rate and stockout state [26]. With the one-time unit penalty p_i, the annual relevant cost is

$$F_i(Q_i, r_i) = \frac{A_i D_i}{Q_i} + h_i\, I_i(Q_i, r_i) + \frac{p_i D_i\, e_i(Q_i, r_i)}{Q_i}. \tag{6}$$

The single-objective constrained model is

$$min\, F(y) = \sum_{i=1}^{n} F_i\,(m_i k_i, r_i) \tag{7}$$

$$\sum_{i=1}^{n} c_i\, I_i(m_i k_i, r_i) \leq B, \quad \sum_{i=1}^{n} v_i\, I_i(m_i k_i, r_i) \leq W \tag{8}$$

$$k_i^L \leq k_i \leq k_i^U, \quad r_i^L \leq r_i \leq r_i^U, \quad k_i, r_i \in Z. \tag{9}$$

The capacities in (8) constrain long-run average resource use. Applications that require compliance with instantaneous peak capacities would need inventory-path or chance constraints and are outside the scope of the present formulation.

2.4. Exact multiple-choice reformulation

Because each item has a finite decision space, all service-feasible policies can be enumerated. Let P_i be the candidate set for item i. Candidate j has cost f_{ij}, average inventory a_{ij}, investment $b_{ij} = c_i\, a_{ij}$, and storage requirement $w_{ij} = v_i\, a_{ij}$. The binary variable x_{ij} is equal to one if policy j is selected. Therefore, the nonlinear integer model is precisely equivalent to

$$min \sum_{i=1}^{n} \sum_{j \in P_i} f_{ij} x_{ij} \tag{10}$$

$$\sum_{j \in P_i} x_{ij} = 1 \quad (i = 1, \dots, n) \tag{11}$$

$$\sum_i \sum_{j \in P_i} b_{ij} x_{ij} \leq B, \quad \sum_i \sum_{j \in P_i} w_{ij} x_{ij} \leq W, \quad x_{ij} \in \{0,1\}. \tag{12}$$

This reformulation limits the nonlinear computations to the generation of item-level policies and represents the system allocation as a multiple-choice binary linear programming problem. Because all service-feasible policies are enumerated before valid within-item dominance pruning, the pruned master preserves the optimal objective value without introducing unimplementable alternatives.

3. Sparse LP rounding and resource-exchange descent

3.1. Dominance pruning and LP relaxation

For two policies u and v of item i, policy v is dominated if $f_{iu} \leq f_{iv}$ and $a_{iu} \leq a_{iv}$, with at least one strict inequality. Let $P_i^P \subseteq P_i$ denote the nondominated policy set retained after dominance pruning. Since both shared-resource coefficients are positive multiples of a_{ij}, v cannot enhance the system objective or any resource constraint and can be excluded without affecting the optimal solution:

$$f_{iu} \leq f_{iv}, a_{iu} \leq a_{iv}, (f_{iu}, a_{iu}) \neq (f_{iv}, a_{iv}) \Rightarrow v \notin P_i^P. \tag{13}$$

The nondominated policies form an item-level Pareto set. Replacing the binary restrictions in (12) with $0 \leq x_{ij} \leq 1$ yields an LP relaxation with the optimum value L, which is the global lower bound on the objective value of every feasible integer policy. An optimal basic feasible solution is used because the number of item groups requiring rounding is determined by the dimension of its positive-support face.

3.2. Sparse extreme points

Let $\mathcal{F}$ denote the set of fractional item groups, i.e., those containing at least two positive x_{ij}. Lemma 1 indicates that if the LP relaxation has n item-selection equalities and q linearly independent shared-resource rows, which are active at the optimal basic solution, then $\sum_i (s_i - 1) \leq q$ and $|\mathcal{F}| \leq q$, where s_i is the positive-support size of item i, and q is the rank of the active resource rows. Inactive or linearly dependent resource constraints do not affect this bound. If the LP optimum is degenerate, the result holds for an optimal basic solution returned by the LP solver.

Proof. If there are only item-selection equalities, the face generated by the positive supports forms a Cartesian product of simplices with dimension $\sum_i (s_i - 1)$. At most q independent active resource equalities can be used to reduce the dimension at an extreme point. Therefore, the face dimension cannot exceed q. Each fractional group contributes at least one dimension, which supports the second statement.

With two shared resources, a fractional basic solution has either one group with at most three positive variables or two groups with two positive variables each. Thus, at most four integer combinations of the fractional positive supports need to be checked while the integral groups remain fixed. This bounds the rounding search independently of the total number of items.

3.3. Support-average-dominance rounding

For a fractional group i, let $S_i = \{j: x_{ij} > 0\}$. The item-selection equality means $\sum_{j \in S_i} x_{ij} = 1$. For resource k, the LP-weighted group average is $\bar{g}_i^k = \sum_{j \in S_i} x_{ij} g_{ij}^k$. The following theorem gives a feasible rounding solution.

Theorem 1. If every fractional group i contains a support policy j_i^* satisfying $g_{ij_i^*}^k \leq \bar{g}_i^k$ for every shared resource k, then choosing j_i^* in each fractional group while keeping all integral groups will yield an integer solution that is feasible for every shared-resource constraint.

Proof. For each resource and fractional group, the rounded usage is not greater than the LP-weighted usage. Summing up these inequalities and considering the unchanged integral groups guarantees the total usage does not exceed that of the feasible LP solution.

A common resource-minimizing policy within each fractional support provides a simple verification condition. Since $b_{ij} = c_i\, a_{ij}$ and $w_{ij} = v_i\, a_{ij}$, with $c_i, v_i > 0$, the supporting policy with the smallest average inventory minimizes both resource requirements. This policy satisfies Theorem 1 under the stated proportional structure. The construction establishes the existence of a feasible corner. The implementation enumerates all corners of the fractional positive supports, keeps integral groups unchanged, and selects the feasible corner with the lowest total cost. Support-average dominance may hold without a common componentwise minimum. Conversely, crossed resource rankings may leave a fractional support without a feasible integer corner even when the complete integer problem is feasible. The condition is sufficient, not necessary.

Let $\Delta_i = \max_{j \in S_i} f_{ij} - \min_{j \in S_i} f_{ij}$. If $\mathcal{F}$ is empty, the LP solution is integer and $F_R = L$. Otherwise, only fractional groups change, and their number is at most q, so the rounded cost F_R satisfies

$$0 \leq F_R - L \leq \sum_{i \in \mathcal{F}} \Delta_i \leq q \max_{i \in \mathcal{F}} \Delta_i. \tag{14}$$

Here, L is the optimal LP-relaxation value defined in Section 3.1. If the active-resource rank q is fixed, each item contributes at least $l > 0$ to L, and each fractional-support cost span is at most Δ, then

$$(F_R - L)/L \leq q\Delta/(n\ell). \tag{15}$$

Equation (15) provides a conditional scaling result rather than a distribution-free approximation ratio. The result does not apply if item costs have no positive common scale or if the fractional-support spans increase with n.

3.4. Lattice and resource-exchange descent

A unit-lattice move changes one coordinate of an item policy by $k_i \leftarrow k_i \pm 1$ or $r_i \leftarrow r_i \pm 1$ within its original bounds. Exact item-level differences in cost, investment, and storage are obtained from the enumerated policy measures. LP-SERD runs two deterministic descent paths independently from the same least-cost feasible support corner. The single-item path accepts the feasible unit move with the greatest strict cost reduction. Neighboring policies are evaluated in the original integer domain, including policies removed from the dominance-pruned LP master.

The exchange path considers both single-item moves and pairs of unit moves on distinct items. The representation $I_i = Q_i^{-1} \int_{r_i}^{r_i+Q_i} E\,[(u - X_i)_+]du$ is nondecreasing in either coordinate: shifting the interval upward increases its integrand, and extending its upper end increases the mean of a nondecreasing function. Consequently, decreasing a coordinate releases both resources, while increasing it consumes them. The scan pairs a releasing move with a consuming move and compares its net cost reduction with the best feasible single-item move. A releasing move may increase cost yet enable a larger reduction in another item. The complete evaluator verifies each accepted move.

Paired unit moves form a two-item local-search neighborhood. Their directional restriction has a specific scope. At a policy with no feasible improving single-item move, an improving feasible pair

cannot contain two releasing moves, since at least one component would itself reduce cost and be feasible. Nor can it contain two consuming moves, since feasibility of the pair would imply feasibility of a cost-reducing component. Every improving pair at such a policy therefore has opposite resource directions. This argument does not establish equivalence to unrestricted two-item best-improvement search at earlier iterates. The implementation allows single and directional-pair moves to compete throughout; multi-step and higher-order exchanges are not included.

Both paths accept only strict cost reductions on a finite integer domain. A path stops when its scanned neighborhood has no feasible improvement or its evaluation limit is reached. LP-SERD returns the lower-cost feasible terminal policy from the two paths. With $L > 0$, the reported LP-certified gap is $100(F - L)/L$ percent. The LP lower bound remains valid for either terminal policy. Figure 1 summarizes the complete procedure.

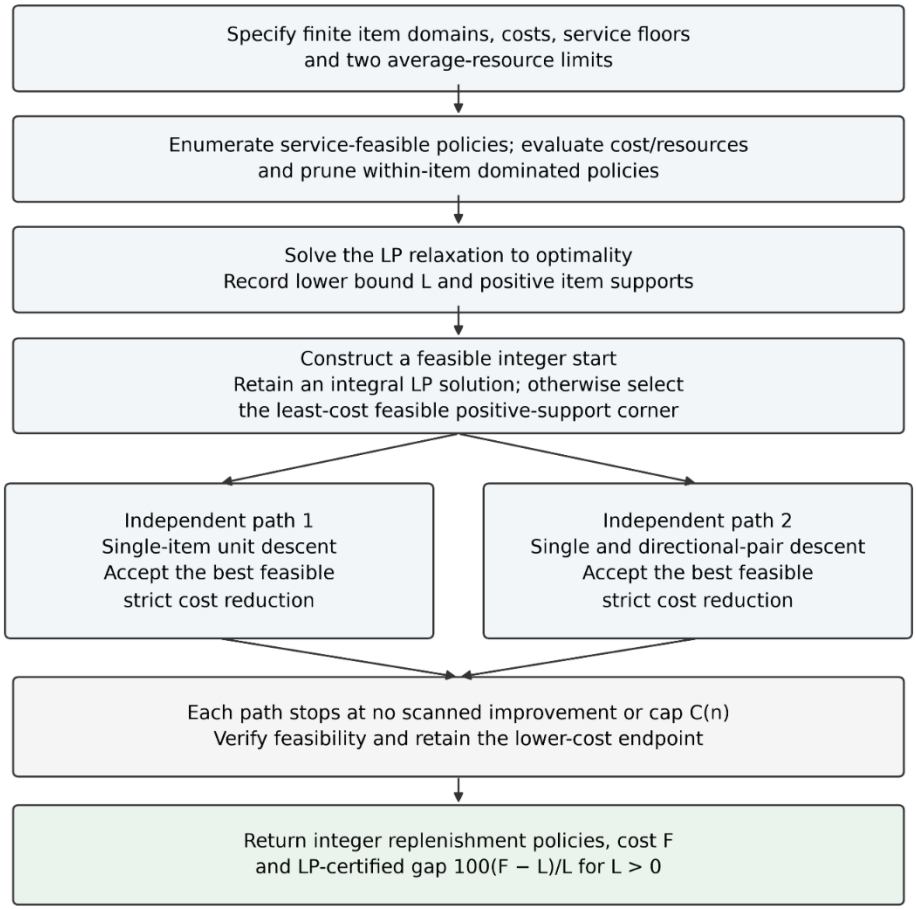


Figure 1. LP-SERD computational procedure. Complete MILP is an independent reference route. The support-rounded policy is integer and feasible before either descent path begins; the lower-cost endpoint is retained.

3.5. Applicability and computational burden

Item policies are enumerated and pruned after sorting by average inventory. The LP master contains one equality per item and a small number of shared-resource rows. The active-resource rank, rather than the total item count, bounds the number of fractional groups. Support-corner enumeration is therefore small, although policy enumeration and the LP solution still grow with the master size.

Local refinement also requires neighborhood scans; the two-path procedure incurs both paths' computational costs, and resource-pair scans can be more expensive than single-item scans.

LP-SERD applies to additive item costs, finite item-policy sets, and a small number of effective shared resources. Common within-item resource ordering supplies the feasible-rounding guarantee in the present model. Without that ordering, the support condition must be checked explicitly; support enlargement, repair, or integer search may be necessary. If the pruned complete MILP is tractable within the available time, exact optimization provides the final solution. Otherwise, LP-SERD offers a deterministic feasible construction with an LP-based gap certificate, rather than a general guarantee of superiority over longer integer search.

4. Computational study

4.1. Experimental design and common protocol

Synthetic instances examine accuracy, scalability, and method-selection boundaries, rather than represent a particular warehouse. Five fixed instances are used at each size $n \in \{20,50,100,200,500,1000,2000\}$. Annual demand ranges from 800 to 5000 units; lead-time-demand means and standard deviations range from 35 to 180 and from 8 to 35 units. Nominal ordering, annual unit holding and unit shortage penalties range from 45 to 180, 2 to 12, and 15 to 70 monetary units, respectively; all three are multiplied by the fixed adjustment 1.04. Cycle-service floors range from 0.90 to 0.98, and pack sizes are drawn from $\{5,10,20\}$. The quantity bounds are $Q_i^{min} = m_i$ and $Q_i^{max} = \max\{2m_i, 4\sqrt{2A_iD_i/h_i}\}$; reorder-point bounds are $r_i^{min} = \max\{0, \mu_i - \sigma_i\}$ and $r_i^{max} = \mu_i + 3.2\sigma_i$. Equations (4)–(5) impose pack and integer rounding and the service floor.

For the main conflict instances, a uniform draw $s_i \in [0,1]$ gives $c_i = 8 + 37s_i$ and $v_i = 0.2 + 1.6(1 - s_i)$. Each capacity is set 20% of the way from the minimum-inventory allocation to the independently cost-minimizing allocation, with both endpoints evaluated by (2) and (6). In the correlation experiment, a Gaussian copula generates four target resource correlations. Demand, costs, service requirements, packs, and policy domains remain fixed within each seed block; capacities are reconstructed after the resource coefficients change.

All algorithms use the same finite domains, policy measures, capacities, and final evaluator. A final policy is resource-feasible when the sum of the positive relative violations of the two resource limits is at most 10^{-12}. Numerical LP-support detection and internal solver tolerances are distinct from this final check. LP and MILP use identical dominance-pruned options, while local unit moves are checked in the original lattice. Reported workflow times sum policy-table construction, pruning, master assembly, LP or MILP calls, rounding or repair, verification, and method-specific search. Shared phases are measured once and added to each method's own phases; these are phase-summed complete-workflow estimates from model parameters, not independently repeated standalone wall-clock measurements. Input-file reading, training-data fitting, and scenario construction are excluded. Pair scans and both LP-SERD paths are included. Nominal solver limits apply only to solver calls. HiGHS uses LP limits of 30 s for 20–200 items and 90 s for 500–2000 items, with rounding only after optimal LP termination. The relative MILP gap tolerance is 10^{-9}.

Each independent descent path has the inherited safeguard $C(n) = \text{round}[0.1\max(3900,100n)]$ complete candidate evaluations, including its initial policy: 390 evaluations for $n \leq 39$ and $10n$ thereafter. The constants specify an operational stopping cap, not a theoretically optimal budget. Table lookups and rejected pair comparisons are excluded from this count but included in elapsed time. Reaching the cap precludes a local-optimality claim. The two-path method incurs both paths' work.

Restricted MILP retains kernel width 3; the LP-adapted TIKS implementation [16] retains a 20%/80% phase split and two buckets per phase. These inherited settings are fixed without retuning on the corrected-model test cases. Computations use an Intel Core i7-13700 under Windows 11, Python 3.14.7, NumPy 2.5.2, SciPy 1.18.0, and HiGHS 1.12.0 with default threading. Timed solver batches run serially without competing experiments.

To isolate rounding and local-refinement effects, six deterministic variants were evaluated on the same 35 instances: five each with 20, 50, 100, 200, 500, 1000, and 2000 items. Each instance used one common pruned master and one LP solution. MW-R selected each item's maximum-LP-weight policy and applied deterministic resource repair if necessary; this is the nearest one-hot choice in squared Euclidean distance before repair, not a globally nearest feasible solution. SR selected the least-cost feasible positive-support corner. MW-R-SD and SR-SD added the same single-item descent to the respective starts; SR-ED used the exchange path from the SR start. LP-SERD retained the better independently initialized SR-SD and SR-ED endpoints. No variant-specific tuning was performed. Each path used the cap in Section 4.1, and all final policies were checked by the same complete evaluator. The paired cost-comparison tolerance was the larger of 10^{-7} and 10^{-9} times the reference cost. These deterministic comparisons use different instances as the analysis units, not repeated stochastic runs.

During MW-R repair, lower-inventory alternatives are scanned across all items. The selected replacement minimizes incremental cost divided by normalized relief of the currently violated capacities. Relief is the inventory reduction weighted by the item's investment coefficient divided by the investment limit when that constraint is violated, plus its storage coefficient divided by the storage limit when the storage constraint is violated. Score ties across items favor greater relief; remaining ties follow fixed item and policy order. Replacements continue until both capacities are satisfied. This repair may change groups that were integral in the LP.

Pooled gap intervals use 10,000 bootstrap resamples within each fixed problem size, preserving five instances per size. In the correlation experiment, 10,000 bootstrap resamples retain whole seed blocks, and the two-sided permutation test uses 50,000 within-block permutations. These intervals describe variation under the specified instance design; the deterministic workflow times are not given stochastic-run confidence intervals.

4.2. Accuracy, component effects, and solver budgets

LP-SERD produced feasible policies on all 20 instances with 20–200 items. No optimal LP basic solution had more than two fractional groups. The median LP-certified gap was 0.0123% (bootstrap 95% interval 0.0079%–0.0291%), and the maximum was 0.1070%. Median phase-summed workflow time was 0.1816 s. Dominance pruning retained a median 3.94% of the original policies. Full MILP returned feasible incumbents on all 20 cases and closed 12 optimality proofs within the 60 s solver limit; limited-time incumbents are not identified with exact optima.

From the same support-rounded starts, median gaps were 0.1035% without descent, 0.0334% after single-item descent, and 0.0202% after the independently initialized exchange path. Taking the lower-cost terminal policy gave 0.0123%. The exchange path had 14 lower-cost, 3 tied, and 3 higher-cost endpoints than single-item descent. These are matched-start component comparisons, not sequential stages or independent repetitions.

For the 15 instances with 500–2000 items (five per size), the pooled median LP-certified gap was 0.000701%, and the maximum was 0.004148%. Median gaps at 500, 1000, and 2000 items were 0.002013%, 0.000701%, and 0.000343%, while median complete-workflow estimates were 2.6090 s,

6.6060 s, and 18.4795 s. The pooled median workflow estimate was 6.6060 s. Median-retained master sizes were 100,968, 201,457, and 400,562. No LP solution required rounding more than two groups, although enumeration and LP solution costs continued to grow (Figure 2).

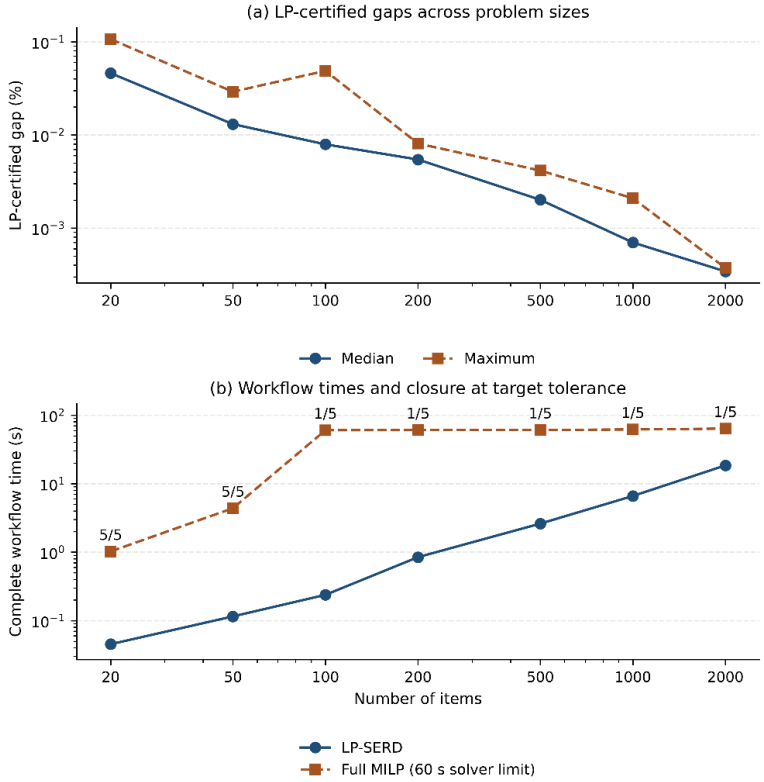


Figure 2. Scale-dependent solution quality and workflow cost. (a) Median and maximum LP-certified gaps, with five instances per size. (b) Median phase-summed workflow estimates. Labels on full-MILP points give proofs closed within the configured tolerance out of five; the nominal solver limit is 60 s. Both vertical axes are logarithmic.

Table 1. Large-instance solution quality and computational cost; 15 instances with 500–2000 items.

Method/target	Median LP gap (%)	Median workflow (s)	Lower cost than LP-SERD/15	MILP proofs/15
LP-SERD	0.000701	6.6060	—	—
Restricted/matched	0.000370	6.6316	13	—
TIKS/matched	0.000806	6.6510	9	—
TIKS/15 s	0.000379	15.0320	12	—
Full MILP/60 s	0.000305	61.7945	14	3

Table 1 compares the 15 large instances under the inherited solver protocols. At approximately matched workflow targets, LP-SERD had lower costs on 2 instances than restricted MILP and on 6 than the LP-adapted TIKS implementation. Their median phase-summed workflow estimates were 6.6316 and 6.6510 s, respectively, compared with 6.6060 s for LP-SERD. Restricted MILP had lower costs in 13 of 15 cases and a median LP-certified gap of 0.000370%, compared with 0.000701% for LP-SERD; TIKS at the matched target had lower costs in 9 cases. All final comparator policies were

feasible, including verified repair fallbacks; this does not imply that every internal integer subproblem produced an incumbent. With a fixed 15 s total target, TIKS had lower costs on 12 instances. This target is not uniformly longer than LP-SERD, whose median complete workflow at 2000 items exceeded 15 s. These comparisons show instance-dependent cost–computation trade-offs rather than uniform solver dominance.

Table 2 reports per-size median gaps for the six LP-based variants. Unrepaired maximum-weight rounding was feasible on 11 of 35 instances, whereas all 210 completed method outputs were feasible. Support rounding had 16 lower-cost, 16 tied, and 3 higher-cost starts than maximum-weight rounding with repair. Its value is a feasible positive-support construction with a loss certificate, not uniform cost superiority over a repair that can change originally integral groups. LP-SERD improved on repaired maximum-weight rounding in 34 instances and tied in one. Against the stronger MW-R-SD baseline, it had 28 wins, 4 ties, and 3 losses; median paired cost reduction was 0.002159%, and the largest disadvantage was 0.017109%.

Table 2. Median LP-certified gaps (%) for LP-rounding and descent variants; five instances at each size.

Items	MW-R	SR	MW-R-SD	SR-SD	SR-ED	LP-SERD
20	0.286250	0.272146	0.080256	0.071483	0.046142	0.046142
50	0.091687	0.100111	0.045664	0.013171	0.027356	0.013070
100	0.142257	0.142257	0.039230	0.042368	0.007928	0.007928
200	0.037962	0.037962	0.008738	0.008738	0.006412	0.005451
500	0.023238	0.013297	0.005316	0.005316	0.003020	0.002013
1000	0.013367	0.010662	0.002152	0.002688	0.001111	0.000701
2000	0.004218	0.003110	0.001082	0.000617	0.000343	0.000343

MW-R: maximum-weight rounding with repair; SR: positive-support rounding; SD: single-item descent; ED: exchange path (also admitting single-item moves). LP-SERD selects the better independent SR-SD/SR-ED endpoint.

The exchange path versus SR-SD gave 27 wins, 3 ties, and 5 losses, with a median paired reduction of 0.001257%. On the 15 large instances, LP-SERD improved on SR-SD in 13 cases and tied in 2; against MW-R-SD, it won 14 and lost 1. The portfolio is noninferior to SR-SD by construction; separate ED-versus-SD and external-baseline comparisons quantify the refinement benefit. No path reached its evaluation cap. Timing includes both paths and all pair scans. These deterministic, single-execution workflow estimates describe computational cost and do not establish equal-wall-time superiority.

Figure 3 separates full-MILP feasible-incumbent availability, objective quality, proof closure, and actual workflow cost. Independent calls with nominal solver limits of 1, 5, 15, and 60 s returned incumbents on 0, 10, 15, and 15 of the 15 instances and closed 0, 0, 1, and 3 proofs, respectively. The corresponding median complete-workflow estimates were 4.9362, 7.0348, 16.8625, and 61.7945 s. LP-SERD returned feasible policies on all 15 instances with a median workflow estimate of 6.6060 s and a maximum LP-certified gap of 0.004148%. Solver time limits do not include table construction and other workflow phases, and actual overruns are retained.

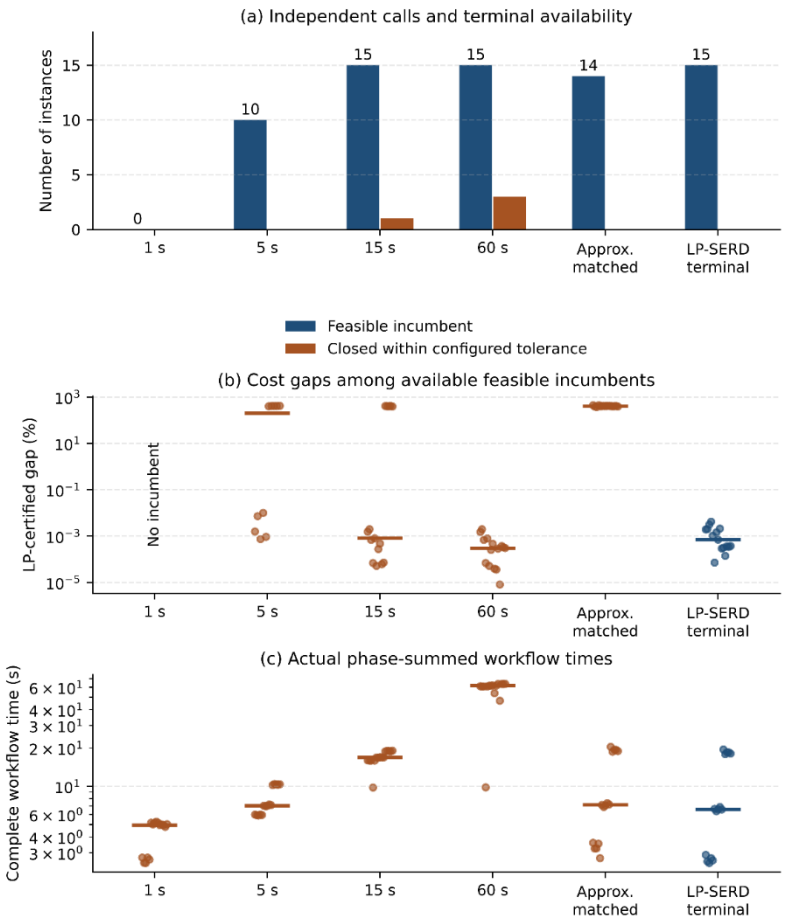


Figure 3. Finite-budget full-MILP results for 15 instances with 500–2000 items. (a) Feasible-incumbent availability and proof closure. (b) LP-certified gaps of available incumbents. (c) Actual phase-summed workflow times. Short horizontal segments denote medians. Missing incumbents have no objective value; objective populations therefore differ across limits. Numerical categories are nominal solver limits, while "Approx. matched" uses the LP-SERD workflow target. Each category is an independent solver invocation, not a point along one search trajectory. LP-SERD is shown at its own termination without a proof-closure claim. Panels (b) and (c) have logarithmic vertical axes.

At the approximately matched setting, full MILP returned 14 incumbents, of which 0 had lower costs, and 14 had higher costs than LP-SERD. Its median complete-workflow estimate was 7.1296 s; this setting is not an exact equal-wall-time experiment. At 60 s, full MILP had lower costs on 14 instances, including 11 of the 12 cases without a closed proof. Incomplete proof closure therefore did not imply poor incumbent quality. Each limit used a separate invocation, so the observations are not an any-time trajectory of one search.

4.3. Validity boundary of support rounding

The support condition was examined separately from the inventory evaluator using 150 random extreme-point instances with increasing violations of within-item resource ordering. Support-average dominance held in 30, 28, 23, 17, and 13 of the 30 instances at violation rates of 0, 0.133, 0.333, 0.533, and 0.667, respectively. Every sampled instance had a feasible support corner. A constructed crossed-resource

counterexample nevertheless has a feasible complete integer problem but no feasible integer corner in the selected LP support. Thus, the sample feasibility rate does not replace the structural condition.

For all 150 two-policy supports, support-average dominance and a common componentwise minimum gave the same classification: 111 satisfied both. A separate three-policy construction satisfies support-average dominance without a common minimum, showing that the former sufficient condition can be less restrictive. Figure 4 presents the two-policy sampling results and the constructed boundary cases. These abstract support diagnostics do not use the inventory-cost approximation.

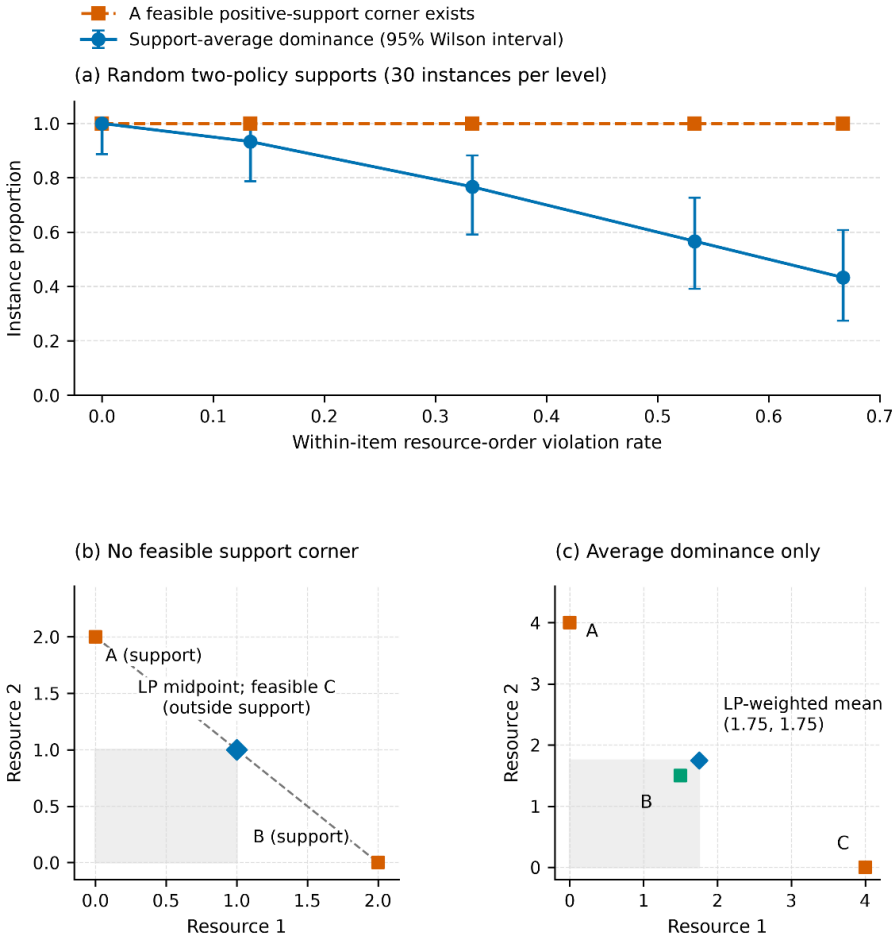


Figure 4. Boundaries of feasible support rounding, independent of the inventory evaluator. (a) Two-policy supports at five order-violation levels, 30 instances per level; error bars are 95% Wilson intervals. (b) Crossed supported resources A = (0,2), B = (2,0) with unit limits and equal LP weights: neither support corner is feasible, although the excluded C = (1,1) is feasible (costs 1, 1, and 2, respectively). (c) Three supported policies with weights (0.25, 0.50, 0.25); B uses no more of either resource than the support average (1.75, 1.75), without being a common resource minimum.

4.4. Resource-correlation sensitivity

The paired resource-structure experiment used 10 fixed seed blocks, each with 50 items and 4 target correlations. Within each block, demand, cost, service, packaging, and integer domains were held fixed while resource coefficients and corresponding capacity endpoints changed. All 40 LP-SERD

policies were feasible, and full MILP closed every optimality proof within the configured tolerance. The median and maximum LP-certified gaps were 0.0101% and 0.0558%, respectively.

The exchange path returned a lower terminal cost than the single-item path in 39 of the 40 cases and a higher cost in one. However, the association between realized coefficient correlation and exchange gain was not clearly resolved: Spearman correlation coefficient = −0.115, seed-block bootstrap 95% interval [-0.398, 0.180], and two-sided blocked permutation p = 0.4932. The interval resamples 10 whole seed blocks; the permutation test preserves those blocks. Figure 5 therefore documents the response distribution and the lack of clear evidence for this association, rather than establishing that lower coefficient correlation increases exchange gain.

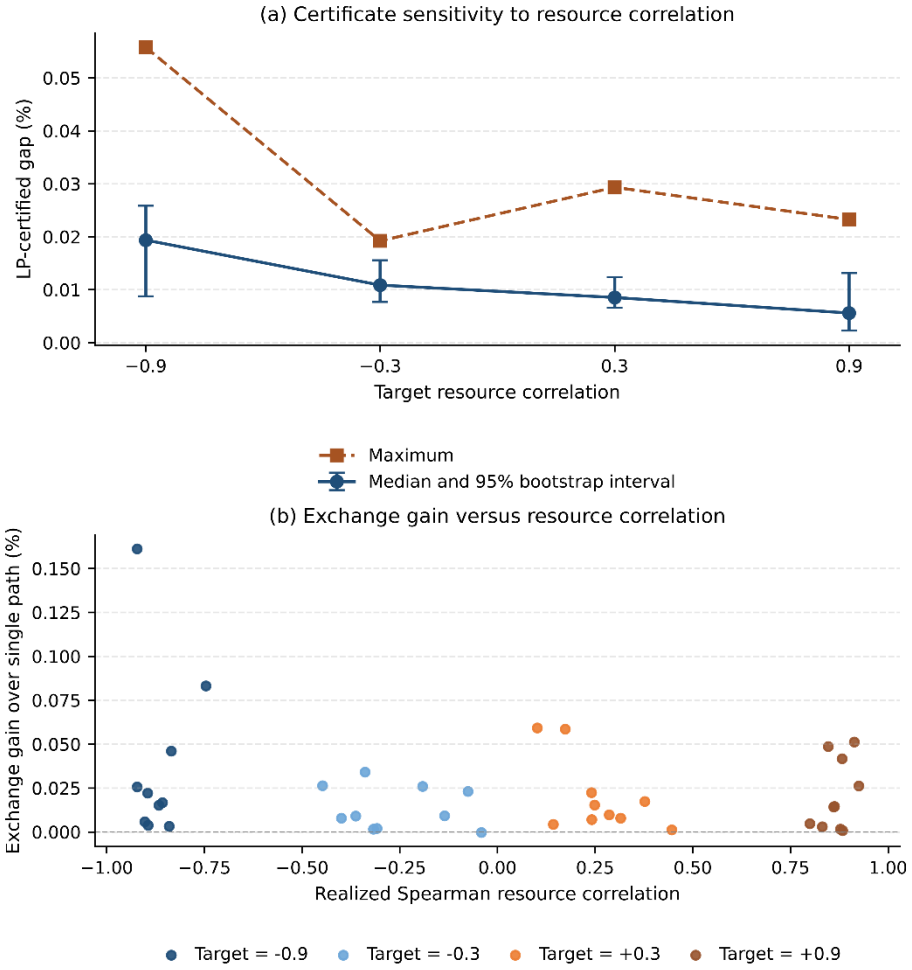


Figure 5. Resource-correlation sensitivity of the LP certificate and exchange gain. (a) Median certificates and 95% bootstrap intervals from 10 cases per target correlation, with maxima. (b) All 40 paired exchange-versus-single-path gains. The Spearman association is −0.115; the seed-block bootstrap 95% interval is [−0.398,0.180], and the two-sided blocked permutation p-value is 0.4932. No fitted trend is imposed.

5. Industrial-data studies

5.1. OIL refinery records

The OIL refinery dataset [28] contains 14,523 items, 66 attributes, and 56 monthly consumption records from January 1997 to August 2001. November 1997 is excluded in accordance with its

documented reliability problem. Missing entries are treated as months with no recorded issues, and negative returns or accounting reversals are excluded from outbound demand by setting them to zero. The 55 retained months are ordered chronologically. The first 43 months, January 1997–August 2000 excluding November 1997, form the training set; September 2000–August 2001 is the untouched 12-month holdout.

Eligibility and ranking use training observations only. Eligible items have positive lead time and price, nonnegative recorded minimum, maximum and current inventory, positive mean training demand, and at least six positive-demand training months. Ranking by annual demand value, with item code as the tie-breaker, defines nested sets of 20, 50, 100, 200, and 500 items. The normal and empirical models use the same selected items at each size. No holdout demand observation enters eligibility, ranking, parameter fitting, or optimization.

Normal lead-time-demand moments are estimated as $\mu_i = \bar{d}_i \ell_i$ and $\sigma_i = \max\{s_i\sqrt{\ell_i}, 0.05\}$, where ℓ_i is the recorded lead time in months (days divided by 30.4375). Empirical marginals use overlapping training lead-time windows, with proportional demand for a partial month. Windows crossing the excluded November 1997 are discarded rather than treating October and December as consecutive months. For M_i windows, the empirical generalized inverse uses ordered observation $\lceil M_i\beta_i \rceil$, followed by integer ceiling. Service floors are 0.99, 0.97, 0.95, and 0.90 for high, medium, low, and unclassified importance, respectively. Price supplies c_i; absent measured pack and volume data, $m_i = v_i = 1$ are assumptions, so the second capacity is an equivalent item-count limit, not cubic storage.

OIL cost coefficients require assumptions because ordering and shortage costs are not observed. With $Q_i^{obs} = \max\{1, S_i^{obs} - r_i^{obs}\}$, nominal holding cost is 20% of unit price, and nominal ordering cost is $\max\{1, 0.2c_i(Q_i^{obs})^2/(2D_i)\}$. Shortage penalties are one, two, four, and eight times the unit price from unclassified to high importance. A fixed factor 1.0375 is applied to these cost coefficients as an assumed adjustment, not an estimated loss. The quantity domain starts at one and ends at $\lceil \max\{2, Q_i^{obs}, 4\sqrt{2A_iD_i/h_i}\} \rceil$; the reorder-point upper bound is $r_i^U = \lceil \max\{r_i^{obs}, r_i^L, \mu_i + 3.2\sigma_i\} \rceil$. Common capacities are the componentwise maxima of resources used by the service-repaired recorded min–max policy under the two demand models. They are constructed comparison envelopes, not measured warehouse limits.

The 10 OIL models contained 179,889–4,138,525 candidate policies, of which 207–7310 remained after dominance pruning. The optimal LP solution returned for each model was integer, and its objective equaled the complete-MILP optimum; neither descent path accepted a move. The LP therefore solved these finite allocation models without requiring resource-exchange descent. The phase-summed workflow times were 0.0767–2.8001 s for LP-SERD and 0.0805–3.4058 s for complete MILP, including policy-table construction and pruning. Table 3 summarizes the results by demand model. The small timing differences in these cases do not establish a practical need for additional heuristic search.

Both shared-resource constraints remained slack in all 10 OIL solutions: capital utilization was 81.51%–91.21%, and equivalent-unit utilization was 76.60%–94.90%. The absence of accepted exchanges is thus consistent with a resource envelope that does not restrict the itemwise cost-minimizing allocation. These cases show the importance of inspecting the LP solution before invoking integer improvement procedures; they do not measure the advantage of paired exchange under binding resource constraints. The nested item sets describe changes with problem size and are not treated as independent experimental replications.

Table 3. Industrial optimization and marginal-service results by dataset and demand model.

Dataset/demand model	Max. LP-certified gap (%)	Full-MILP proofs/5	Median LP-SERD workflow (s)	Median full-MILP workflow (s)	Holdout service passes (n = 500)
OIL/Normal	0	5/5	2.2124	2.7272	451/500
OIL/Empirical	0	5/5	0.8684	1.3192	447/500
MAN/Normal	0.00799021	2/5	0.3230	60.0386	397/500
MAN/Empirical	0.00013357	5/5	0.2429	3.6159	357/500

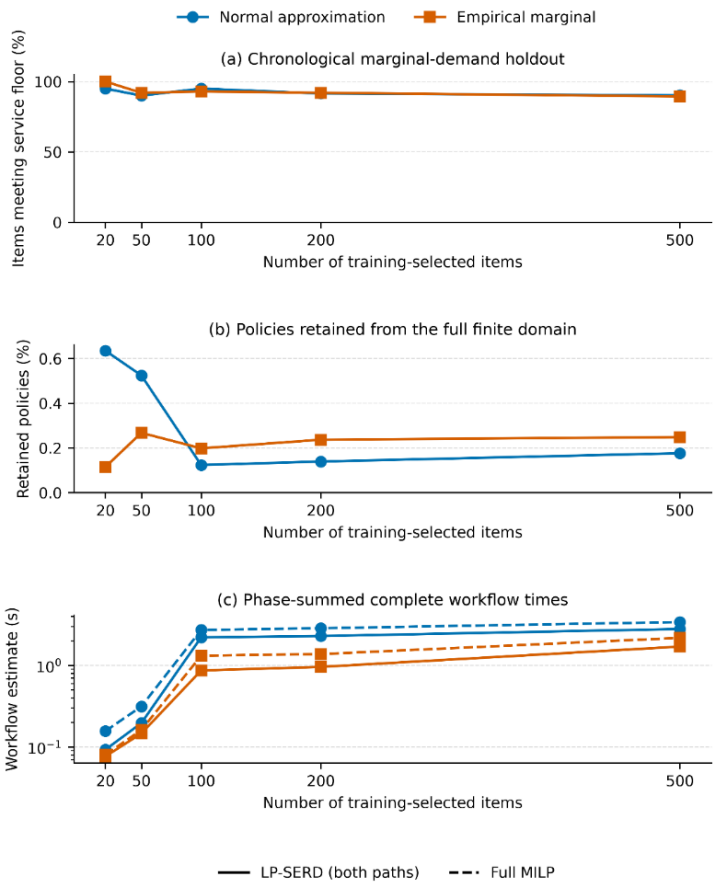


Figure 6. OIL optimization and temporal marginal-service assessment. (a) Holdout-service pass proportions on nested item sets, not independent replenishment cycles. (b) Policies retained after dominance pruning as a percentage of the full finite policy domain. (c) Phase-summed workflow estimates on a logarithmic scale. The empirical windows retain calendar continuity; pack size, equivalent storage, and unobserved costs follow the assumptions in Section 5.1.

The 20-item OIL models were also solved with resource-interpolation coefficients of 0.10, 0.20, 0.40, 0.60, and 0.80. Unlike the recorded-policy envelopes above, these limits interpolate between the minimum-inventory and independent minimum-cost allocations; componentwise maxima across the two demand models give common limits. All 10 solutions remained feasible. The largest LP-certified gap was 0.5050%, attained by the empirical model at 0.10, whereas the largest excess over a proved complete-MILP optimum was 0.3744%, attained by the normal model at 0.10. The empirical-model

solutions matched the complete-MILP objectives at all five settings despite their nonzero LP-certified gaps. This distinction reflects the integrality gap of the LP lower bound rather than an error in those integer solutions. A paired exchange was accepted only in the normal model at 0.40; tighter resources did not uniformly produce useful exchange opportunities.

In the 12-month holdout, 19 of the 20 normal-model policies and all 20 empirical-model policies met their item-specific marginal service requirements. Their mean service deficits were 0.27 and 0 percentage points, respectively; the corresponding maximum deficits were 5.33 and 0 percentage points. At 500 items, the counts were 451 and 447. The number of overlapping lead-time windows ranged from 3 to 12 per item at this scale, so these counts are not independent cycle-service observations or high-confidence service guarantees. Effective dates of the recorded policy and other static fields are undocumented; the temporal assessment therefore concerns the demand marginal under fixed planning inputs, not a complete historical decision replay. Figure 6 presents the pruning, computational, and holdout-service results.

5.2. MAN manufacturing records

The MAN manufacturing dataset [29] contains 3451 spare parts and 150 weekly observations from week 1 of 2012 to week 46 of 2014. Recorded attributes include sub-pack quantity, unit price, fixed order cost, holding cost, lead time, and minimum order quantity. Eligibility, annual-demand-value ranking, and fitting use the first 120 weeks; the final 30 weeks are reserved for holdout evaluation. Positive cost, lead-time, and minimum-order fields and at least eight positive-demand training weeks yield 911 eligible items. Item code breaks ranking ties. The same nested sizes and marginal-demand constructions as in OIL are used, with 52 weeks per year.

The recorded minimum order quantity imposes $Q_i \geq MOQ_i$, not divisibility by MOQ_i. Because the data do not establish an order-pack multiple or measured occupied volume, the quantity increment and equivalent storage coefficient are set to one. Recorded lead times are in weeks and require no unit conversion. Recorded order and annual holding costs are retained, c_i is unit price, $p_i = 0$ is assumed in the absence of measured shortage losses, and $\beta_i = 0.95$. The finite upper bound is $Q_i^U = \lceil \max\{2MOQ_i, 4\sqrt{2A_iD_i/h_i}\} \rceil$. The original reorder-point upper bound is $r_i^U = \lceil \max\{r_i^L, \mu_i + 3.2\sigma_i\} \rceil$. With zero shortage penalty, increasing r_i increases holding cost and both resources, so policies above the service-feasible lower bound are analytically dominated and removed before expansion. Both demand models share capacities constructed at the 0.20 interpolation between their minimum-inventory and independent minimum-cost allocations, taking componentwise maxima across models.

The 10 MAN models initially contained 17.10–806.71 million policies. With no recorded shortage penalty, the specified cost model sets this coefficient to zero; each reorder point then equals its minimum service-feasible value because increasing it raises both holding cost and resource use. This analytical reduction left 13,513–156,883 quantity policies before dominance pruning and 3244–36,041 afterward. The remaining decisions have unit quantity increments subject to MOQ lower bounds, rather than non-unit package multiples. LP-SERD was feasible in all 10 models; the maximum LP-certified gaps were 0.0079902% for normal demand and 0.00013357% for empirical demand. Complete MILP met the prescribed optimality tolerance in seven models. The 20-, 50-, and 200-item normal models reached the 60 s solver limit without closing their gaps; their incumbent costs are not treated as known optima. Phase-summed workflow times were 0.0366–0.8534 s for LP-SERD and 0.1741–60.1756 s for complete MILP. Paired moves were accepted in all 10 models, with 8 and 5 acceptances in the normal and empirical groups, respectively.

For the 500-item MAN models, changing the resource-interpolation coefficient from 0.20 to 0.10 or 0.40 preserved feasibility under both demand models. The maximum LP-certified gap across these four additional cases was 0.00007784%. Complete MILP met the prescribed optimality tolerance in three cases; the normal model at 0.40 reached the 60 s solver limit with a remaining primal–dual gap of 0.000005804%. Because the zero-penalty specification fixes reorder points at their service lower bounds, varying the resource limits changes quantities but not the marginal-service thresholds. Unchanged holdout counts across these settings therefore follow from the model structure rather than establish demand-model robustness.

At 500 items, 397 normal-model policies and 357 empirical-model policies met the 0.95 marginal-service threshold in the 30-week holdout. Across the five nested item sets, the corresponding proportions were 50.0%–79.4% and 70.0%–75.0%. Each item contributed 15–24 overlapping lead-time windows, not independent replenishment cycles. These service shortfalls coexist with small optimization gaps because computational optimality is assessed within the fitted demand model. Under the zero-penalty specification, further integer search cannot change the prescribed reorder-point thresholds; improved out-of-sample service requires reassessment of that demand and service specification. Figure 7 summarizes the pruning, optimization, and marginal-service outcomes.

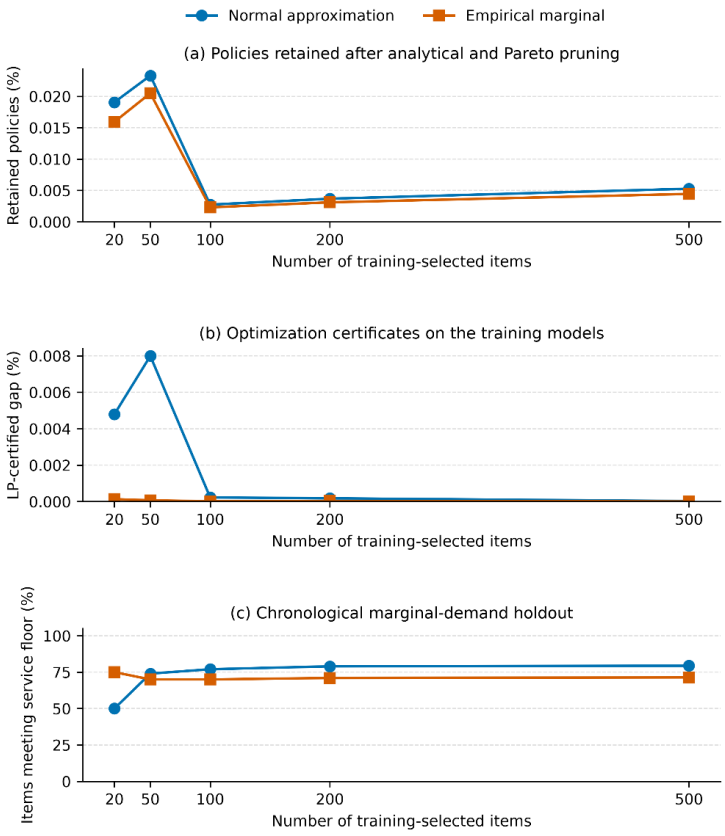


Figure 7. MAN optimization and temporal marginal-service assessment. (a) Retained policies as a percentage of the original quantity–reorder-point domain; analytical reorder-point reduction precedes dominance pruning. (b) LP-certified gaps, retaining zero values on a linear scale. (c) Holdout-service pass proportions for nested item sets. The quantity increment is one, MOQ is a lower bound, and the zero shortage-penalty assumption fixes reorder points at their service-feasible minima.

Each row summarizes the five nested sizes of 20, 50, 100, 200, and 500 items. Workflow times sum the measured table-construction, pruning, and solution phases; they are not independent runtime replications. Full-MILP proofs require the configured relative gap of 10^{-9} after integer-policy verification. The last column reports marginal-service passes in the 500-item holdout.

6. Discussion

6.1. When continuous relaxation is informative

The item-selection equalities and few active shared capacities yield a low-dimensional fractional support in a basic LP solution. Independent coordinate rounding may violate capacities, whereas support rounding retains the categorical structure of each fractional item group. A feasible support corner exists when each fractional group contains a policy using no more of either resource than its LP-weighted average. Selecting the least-cost feasible corner retains the support-span certificate without requiring a repair search outside the LP support.

The inventory model has this property because each item's resource requirements are positive multiples of average inventory. It need not hold when shelf position, hazard class, supplier quotas, or policy-dependent handling space induce crossed resource rankings. Sparse fractionality limits the initially ambiguous groups but does not guarantee that repair can be confined to them. A feasible repair may require support enlargement or changes to previously integral groups, followed by explicit feasibility verification.

6.2. Marginal compensation and its empirical limits

A shared-capacity boundary can block a cost-reducing single-item move because it consumes a resource with little remaining slack. Another item may release the required capacity at a smaller cost increase, making the pair feasible and beneficial. This is the usual compensation mechanism of the two-item local search. Here, common within-item resource ordering identifies releasing and consuming directions, while exact finite differences determine their joint objective and resource changes.

Useful compensation depends on the current resource slacks and the joint marginal costs and resource changes of the available moves. Cross-item coefficient correlation alone does not encode those quantities. The paired correlation experiment did not establish a clear relationship between correlation and exchange gain. This does not prove that resource geometry is irrelevant, nor does it invalidate the within-item common-ordering argument. The component comparisons also show that an exchange path can finish above single-item descent. Running the two paths independently from the same rounded start and retaining the lower-cost endpoint avoids accepting a worse exchange outcome. Directional coverage is conditional on the absence of an improving single-item move, and a cap-limited path has no local-optimality guarantee.

6.3. Method selection and limitations

Method selection depends on the current LP certificate, post-pruning master size, and available computation time. LP-SERD supplies a checkable feasible policy without requiring a general integer search, but the additional descent gains are small and instance-dependent. Full MILP at the 60 s limit obtained lower costs on 14 large instances, including 11 without closed proofs. A small gap to the LP lower bound does not establish that the policy is cheaper than another feasible solution. An integral LP optimum already solves the finite master; when complete MILP can close the proof within the

available time, it provides the final finite-model optimum. The decreasing relative gaps with size also reflect dilution of losses in at most two fractional groups over the total cost of many items, not necessarily stronger search capability. Temporal service shortfalls remain a separate demand-model issue even after optimization error is negligible.

The inventory evaluator assumes fixed lead times, no crossover, a uniform inventory position independent of lead-time demand, and a demand-arrival independence approximation for new-backorder penalties. Intermittent and batch demand may violate these assumptions. Non-demand attributes are treated as fixed scenario inputs because their effective dates are unavailable; the chronological holdout is not a historical decision replay. The industrial datasets do not establish physical volume, order-pack multiples, shortage losses, or actual shared-capacity limits. Their adopted unit grids, equivalent-unit storage, and constructed capacities support a computational study, not validation of warehouse operations. Gap certificates quantify optimization error only for the stated finite planning model. Demand calibration, nonstationarity, and sparse holdout windows can outweigh the residual optimization gap. Extensions require measured capacities, discrete-demand process models, concurrent failures, and nonstationary demand.

7. Conclusions

The dual-resource multi-item continuous-review inventory model was represented by an exact dominance-pruned multiple-choice master. The standard LP rank bound limits fractionality to the number of independent active resource rows. Common within-item resource ordering gives a constructive sufficient condition for feasible rounding within positive supports, and support cost spans bound the additive loss relative to the LP lower bound. Selecting the least-cost feasible support corner retains this certificate.

Independent single-item and directional two-item paths provided small local refinements of the same feasible rounded start. Against repaired maximum-weight rounding with single-item descent, LP-SERD had 28 lower-cost, 4 tied, and 3 higher-cost endpoints on the 35-instance set. The exchange path itself finished below single-item descent on 27 instances and above it on 5, with 3 ties. The pooled median LP-certified gap on the 15 instances with 500–2000 items was 0.000701%. Full MILP at a 60 s solver limit obtained lower costs on 14 instances and closed 3 proofs. The results establish the quality of the feasible construction on the tested set without uniform superiority over other local-search or integer-search routes.

The paired sensitivity experiment did not establish a monotone relationship between cross-item resource correlation and exchange gain. This limits that empirical interpretation without changing the common within-item ordering used in the rounding proof. All 10 OIL main models had integral LP optima and required no local refinement. MAN retained small finite-model optimization gaps, whereas several item policies missed their marginal-service targets in the holdout. The finite-domain certificate measures optimization accuracy, not warehouse validity or future service performance. Deployment requires measured capacity and cost inputs and validation of the demand and replenishment-process assumptions.

Author contributions

Qingbo Lu: Conceptualization, methodology, software, validation, formal analysis, and writing – original draft. Shuangcheng Huang: Methodology, investigation, data curation, visualization, and writing – review and editing. Mengzhen Huang: Validation, formal analysis, resources, and writing –

review and editing. Daojun Li: Supervision, project administration, resources, and writing – review and editing. All authors read and approved the final manuscript.

Use of generative-AI tools declaration

OpenAI Codex assisted with language and structural revision and numerical checks. The author is responsible for writing the initial Chinese draft, validating the mathematical model, calculating the results, providing explanations, compiling references, and producing the final draft. The author is fully responsible for the use of all AI tools.

Funding

This work was supported by the Science and Technology Research Projects of the Henan Province Science and Technology Development Plan (Grant Nos. 242102220107 and 262102221056) and the 2025 Key Research Projects of Higher Education Institutions in Henan Province (Grant No. 25B460026).

Data availability

The data generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflict of interest

The authors have no relevant financial or non-financial interests to disclose.

References

1. B. Ghalebsaz-Jeddi, B. C. Shultes, R. Haji, A multi-product continuous review inventory system with stochastic demand, backorders, and a budget constraint, *Eur. J. Oper. Res.,* **158** (2004), 456–469. https://doi.org/10.1016/S0377-2217(03)00363-1
2. K. Das, T. K. Roy, M. Maiti, Multi-item stochastic and fuzzy-stochastic inventory models under two restrictions, *Comput. Oper. Res.,* **31** (2004), 1793–1806. https://doi.org/10.1016/S0305-0548(03)00120-5
3. S. Y. Chou, P. C. Julian, K. C. Hung, A note on fuzzy inventory model with storage space and budget constraints, *Appl. Math. Model.,* **33** (2009), 4069–4077. https://doi.org/10.1016/j.apm.2009.02.001
4. X. Zhao, M. Qiu, J. Xie, Q. He, Computing (r,Q) policy for an inventory system with limited sharable resource, *Comput. Oper. Res.,* **39** (2012), 2368–2379. https://doi.org/10.1016/j.cor.2011.12.012
5. Y. Kou, Z. Wan, A new data-driven robust optimization approach to multi-item newsboy problems, *J. Ind. Manag. Optim.,* **19** (2023), 197–223. https://doi.org/10.3934/jimo.2021180
6. A. Darendeliler, D. Claeys, E. H. Aghezzaf, Integrated condition-based maintenance and multi-item lot-sizing with stochastic demand, *J. Ind. Manag. Optim.,* **19** (2023), 6908–6947. https://doi.org/10.3934/jimo.2022245

7. I. Saracoglu, S. Topaloglu, T. Keskinturk, A genetic algorithm approach for multi-product multi-period continuous review inventory models, *Expert Syst. Appl.,* **41** (2014), 8189–8202. https://doi.org/10.1016/j.eswa.2014.07.003
8. A. Azimian, B. Aouni, Multi-item order quantity optimization through stochastic goal programing, *Ann. Oper. Res.,* **346** (2025), 751–779. https://doi.org/10.1007/s10479-024-05903-y
9. S. Bhalla, E. Alfnes, H. H. Hvolby, F. Sgarbossa, Advances in spare parts classification and forecasting for inventory control: A literature review, *IFAC-PapersOnLine,* **54** (2021), 982–987. https://doi.org/10.1016/j.ifacol.2021.08.118
10. C. B. Haubitz, U. W. Thonemann, How to change a running system—controlling the transition to optimized spare parts inventory policies, *Prod. Oper. Manag.,* **30** (2021), 1386–1405. https://doi.org/10.1111/poms.13327
11. L. M. Johannsmann, E. M. Craparo, T. L. Dieken, A. R. Fügenschuh, B. O. Seitner, Stochastic mixed-integer programming for a spare parts inventory management problem, *Comput. Oper. Res.,* **138** (2022), 105568. https://doi.org/10.1016/j.cor.2021.105568
12. V. F. Yu, N. Y. Salsabila, N. Siswanto, P. H. Kuo, A two-stage genetic algorithm for joint coordination of spare parts inventory and planned maintenance under uncertain failures, *Appl. Soft Comput.,* **130** (2022), 109705. https://doi.org/10.1016/j.asoc.2022.109705
13. L. Zhang, Q. Deng, B. Miao, X. Liu, H. Shao, Parallel service mode of production and inventory for spare part inventory optimization, *Knowl. Based Syst.,* **241** (2022), 108282. https://doi.org/10.1016/j.knosys.2022.108282
14. P. Sinha, A. A. Zoltners, The multiple-choice knapsack problem, *Oper. Res.,* **27** (1979), 503–515. https://doi.org/10.1287/opre.27.3.503
15. Y. Chen, J. K. Hao, A reduce and solve approach for the multiple-choice multidimensional knapsack problem, *Eur. J. Oper. Res.,* **239** (2014), 313–322. https://doi.org/10.1016/j.ejor.2014.05.025
16. L. Lamanna, R. Mansini, R. Zanotti, A two-phase kernel search variant for the multidimensional multiple-choice knapsack problem, *Eur. J. Oper. Res.,* **297** (2022), 53–65. https://doi.org/10.1016/j.ejor.2021.05.007
17. V. Cacchiani, M. Iori, A. Locatelli, S. Martello, Knapsack problems—an overview of recent advances. Part II: Multiple, multidimensional, and quadratic knapsack problems, *Comput. Oper. Res.,* **143** (2022), 105693. https://doi.org/10.1016/j.cor.2021.105693
18. T. Szkaliczki, Solution methods for the multiple-choice knapsack problem and their applications, *Mathematics,* **13** (2025), 1097. https://doi.org/10.3390/math13071097
19. M. S. Song, Y. Lu, D. Rando, F. J. Vasko, Statistical analyses of solution methods for the multiple-choice knapsack problem with setups: Implications for OR practitioners, *Expert Syst. Appl.,* **262** (2025), 125622. https://doi.org/10.1016/j.eswa.2024.125622
20. I. Crévits, S. Hanafi, R. Mansi, C. Wilbaut, Iterative semi-continuous relaxation heuristics for the multiple-choice multidimensional knapsack problem, *Comput. Oper. Res.,* **39** (2012), 32–41. https://doi.org/10.1016/j.cor.2010.12.016
21. J. Guo, M. Zhang, On lot sizing with capacity acquisition, *J. Ind. Manag. Optim.,* **22** (2026), 3456–3479. https://doi.org/10.3934/jimo.2026126
22. M. Tang, F. Jing, X. Chao, A dynamic lot sizing model with production-or-outsourcing decision under minimum production quantities, *J. Ind. Manag. Optim.,* **16** (2020), 2389–2406. https://doi.org/10.3934/jimo.2019059

23. T. W. Leung, C. K. Chan, M. D. Troutt, A mixed simulated annealing-genetic algorithm approach to the multi-buyer multi-item joint replenishment problem: advantages of meta-heuristics, *J. Ind. Manag. Optim.,* **4** (2008), 53–66. https://doi.org/10.3934/jimo.2008.4.53
24. F. Grandoni, R. Zenklusen, *Approximation schemes for multi-budgeted independence systems, in Algorithms–ESA 2010,* Springer, Berlin, 2010, 536–548. https://doi.org/10.1007/978-3-642-15775-2_46
25. V. Beiranvand, W. Hare, Y. Lucet, Best practices for comparing optimization algorithms, *Optim. Eng.,* **18** (2017), 815–848. https://doi.org/10.1007/s11081-017-9366-1
26. P. Escalona, D. Araya, E. Simpson, M. Ramirez, R. Stegmaier, On the shortage control in a continuous review (Q, r) inventory policy using αL service-level, *RAIRO Oper. Res.,* **55** (2021), 2785–2806. https://doi.org/10.1051/ro/2021125
27. H. Zhang, A note on the convexity of service-level measures of the (r,q) system, *Manage. Sci.,* **44** (1998), 431–432. https://doi.org/10.1287/mnsc.44.3.431
28. E. Porras Musalem, R. Dekker, An inventory control system for spare parts at a refinery: An empirical comparison of different re-order point methods, *Eur. J. Oper. Res.,* **184** (2008), 101–132. https://doi.org/10.1016/j.ejor.2006.11.008
29. D. D. Haan, *Benchmarking spare part demand forecasting methods,* Erasmus University Rotterdam, 2021.



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)