Optimality

www.opt.reapress.com

Opt. Vol. 3, No. 3 (2026) 270–282.

Paper Type: Original Article

# Constraint-Aware Entropy-Adaptive Swarm Search for Civil Infrastructure Sizing: Formulation and Reported Benchmark Evidence

**Hassan M. Khan[1,*], Ahmed M. El-Sayed[2], Rania Boutros[3], Tomasz Lewandowski[4]**

[1] Department of Civil Engineering, NED University of Engineering & Technology, Karachi, Pakistan; hassan.khan@neduet.edu.pk.
[2] Department of Structural Engineering, Ain Shams University, Cairo, Egypt; hassan.khan@neduet.edu.pk
[3] Department of Civil and Environmental Engineering, University of Auckland, Auckland, New Zealand; hassan.khan@neduet.edu.pk.
[4] Department of Civil Engineering, Wroclaw University of Science and Technology, Wroclaw, Poland; hassan.khan@neduet.edu.pk.

**Citation:**

| | |
|---|---|
| *Received: 17 December 2025*<br>*Revised: 26 February 2026*<br>*Accepted: 14 April 2026* | Khan, H. M., El-Sayed, A. M., Boutros, R., & Lewandowski, T. (2026). Constraint-aware entropy-adaptive swarm search for civil infrastructure sizing: Formulation and reported benchmark evidence. *Optimality, 3*(3), 270-282. |

## Abstract

SI-CASS is a proposed population-search framework for civil sizing that combines an operator portfolio, structural-response-based repair and entropy-triggered reinjection. The formulation distinguishes simplified optimization constraints from the complete checks required by a project design code. Repair is a local heuristic followed by structural reevaluation; neither catalog rounding nor area scaling guarantees feasibility or code certification. A conditional coverage argument describes requirements for an infinite search extension, while the specified finite-budget process has no established global convergence rate. Reported summaries from six problems and 30 runs retain the original numerical observations, including a 25-bar-truss mean of 247.1 versus 252.4 for SHADE and 263.5 for PSO. The underlying models, units, implementation and run-level outputs are unavailable, and inconsistencies among the tables, statistical adjustments and figures prevent verification of significance and causal component effects. The contribution is the stated search architecture and an explicit account of the evidence needed to assess it. The reported designs are not certified engineering deliverables.

**Keywords:** Structural sizing, Adaptive operator selection, Constraint handling, Population entropy, Evaluation accounting, Reproducibility.

## 1 | Introduction

Structural sizing combines objective reduction with checks on load effects, resistance, displacement and constructability. Its formulation depends on the material, geometry, load combinations and whether design

Corresponding Author: hassan.khan@neduet.edu.pk
https://doi.org/10.22105/opt.v3i3.119
Licensee System Analytics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0).

variables are continuous or selected from a catalog. A population method can explore such domains without objective derivatives, but its usefulness depends on a defined structural model and a reliable treatment of infeasibility.

Structural-Integrity-Constrained Adaptive Swarm Search (SI-CASS) proposes a four-operator portfolio, a response-based repair heuristic and diversity-triggered reinjection. The operators are selected through credit-dependent probabilities. Repair aims to reduce violations, but every repaired and catalog-rounded candidate must be reevaluated. A final pass–fail report for specified model checks would support traceability; it would not by itself certify compliance with every provision of a design code.

The manuscript sets out the intended sizing model, proposal components, credit and entropy mechanisms and computational accounting. Six-problem numerical summaries are retained as source-reported observations. Their model and experimental identity are not independently reconstructable from this document. Consequently, comparisons are descriptive, inferential claims are unresolved, and no general feasibility, code-compliance or convergence certificate is established.

Section 2 discusses methodological background. Section 3 separates simplified sizing constraints from code-based design requirements. Section 4 presents the intended search components and analytical limits. Sections 5 and 6 document the reported protocol and its numerical evidence. Sections 7–9 discuss implications, limitations and future work.

# 2 | Background and Related Work

Sizing may minimize weight, volume or cost subject to response and resistance checks. These are different objectives, and comparisons require common equations and units. Avcı et al. [1] study ISRES for sizing and layout of several truss benchmarks, while Mehta et al. [2] examine a two-archive multi-objective crystal-structure method for truss mass and displacement. Neither study defines or validates the six instances reported here. A simplified benchmark feasible set must be distinguished from complete project-specific structural design.

Adaptive search operates at different levels. SHADE uses histories of successful differential-evolution parameters [3], and CMA-ES adapts a sampling distribution [4]. The four SI-CASS branches instead form an intended portfolio drawn from DE [5], PSO [6], GWO [7] and long-range perturbation [8]. Their use does not establish superiority over fixed or adaptive alternatives; the implemented branch rules and computational costs must be specified and tested.

Penalty, feasibility-ranking, relaxed-feasibility and repair approaches represent different ways to handle constraints [9–11]. A repair may improve a violated design without finding a feasible point. Reusing finite-difference sensitivities can reduce repeated cost but may reduce their accuracy after section changes. The proposed hybrid uses a penalty to compare infeasible candidates and structural-response adjustments before acceptance; an explicit ordering of feasible and infeasible candidates is still required.

When search is coupled to finite-element analysis, expensive solver calls may dominate the proposal operations. Structural optimization has also been coupled to neural models [12]. FEM and metaheuristic search also serve inverse damage identification: Ton-That uses GA and PSO to minimize a natural-frequency discrepancy over element stiffness reductions [13]. That inverse problem differs from forward section sizing and does not establish design-code compliance. A surrogate can screen candidates, but feasibility and final objectives require checks with the designated true model. The reported SI-CASS-K variant uses Kriging; its training data, screening, model updates and actual solver-call savings are not supplied, so no measured efficiency conclusion is established.

Deterministic resistance constraints and reliability-based constraints answer different questions. The proposed formulation does not define random loads or resistance distributions, failure events or a reliability estimator. It is therefore not a reliability-based design optimizer. A future reliability extension would require a new model and validation, not merely a different repair label.

*Table 1* preserves a reported qualitative positioning matrix. Its broad categorical comparisons are not supported by a systematic survey, and its positive convergence and certification entries for SI-CASS are not established by the analysis below. The framework is a candidate approach for evaluation, rather than a replacement for verified design procedures.

**Table 1. Retained source-reported positioning matrix.**

| Algorithm | Adaptive op. sel. | Convergence Proof | Constraint Handler | Code-Compliance Cert. | Entropy Diversity |
|---|---|---|---|---|---|
| GA | No | No | Penalty | No | No |
| PSO | No | Weak | Penalty | No | No |
| DE / SHADE | Partial (SHADE) | No | Penalty | No | No |
| CMA-ES | Yes (covariance) | No | Penalty | No | No |
| GWO | No | No | Penalty | No | No |
| SQP (deterministic) | N/A | Local only | Active-set | Implicit | N/A |
| SI-CASS (proposed) | Yes (portfolio) | Yes (in probability) | Repair + penalty | Yes | Yes |

"Convergence proof" and "Code-compliance cert." entries are unverified claims, not demonstrated properties. An engineering workflow needs reproducible design variables, analysis assumptions, objective units, load combinations, resistance checks and final constraint ratios. A search method may provide candidate designs, but their acceptance requires the applicable project code, complete structural evaluation and responsible engineering review. The supplied manuscript contains no final design vectors or constraint-ratio certificate.

# 3 | Problem Formulation

For a truss with fixed member lengths L_i and cross-sectional areas x_i, f(x)=Σ_i L_i x_i is volume. With lengths in metres and areas in square metres it has units m³; material density would be needed for weight. An area-only objective does not define a reinforced-concrete column or retaining wall unless their geometry and material quantities are separately mapped to the design vector. Cost-based welded-beam and pressure-vessel objectives also need their own equations and units.

Illustrative truss model: minimize f(x)=Σ_i L_i x_i, subject to normalized checks g_j(x)≤0 and x_i^LB≤x_i≤x_i^UB. If a symmetric allowable-stress benchmark is intended, its check is |σ_ik(x)|/σ_allow, ik−1≤0 for every member and load case. Displacement checks use |δ_j(x)|/δ_allow, j−1≤0. Member slenderness uses λ_i=L_b,i/r_i(x) and a specified λ_max,i. These expressions define simplified benchmark checks; material and code-specific resistance must be evaluated separately.

The absolute stress value covers tension and compression in a symmetric allowable-stress benchmark; an unsigned inequality σ≤φσ_y would fail to limit large compression. Actual tensile yielding, net-section rupture, compression buckling and combined actions require distinct resistance definitions. The originally stated φ=0.90 steel-yielding and φ=0.85 concrete-compression choices do not establish a general code model: 0.85 is not a universal ACI compression strength-reduction factor. The original implementation and material-specific factors must be checked. Linear-elastic response is an intended assumption for the trusses, not evidence of validity for nonlinear materials, instability or every other reported instance.

ANSI/AISC 360-22 is the Specification for Structural Steel Buildings and includes LRFD and ASD formats [14]. An LRFD check compares factored demand with φR_n for a particular limit state. Factors and resistances depend on that state; one generic stress envelope cannot cover yielding, rupture, shear and member stability. The reported eight load combinations, including examples 1.4D, 1.2D+1.6L+0.5(L_r or S or R), and 1.2D+1.0W+L+0.5(L_r or S or R), need their exact adopted load standard, edition and implementation. The governing combination is the one with the largest relevant demand-to-capacity ratio; it cannot be feasible while another evaluated combination violates that same check.

The Eurocode resistance format is R_d=R_k/γ_M, with partial factors and combinations tied to the adopted edition and national annex [15], [16]. The originally reported γ_M0=1.00, γ_M1=1.00 and γ_M2=1.25 are not universal project values. A repair factor must be dimensionless: Use a demand/resistance ratio when both have force units, or a stress/allowable-stress ratio when both have stress units. The originally stated γ_Mσ/R_k mixes units if R_k is member resistance. No nationally adopted complete Eurocode check set is supplied.

ACI 318-19 provides structural-concrete design requirements [17]. Its strength-reduction factor depends on limit state, confinement and strain classification; an RC axial–flexural interaction envelope additionally needs section geometry, concrete and reinforcement properties, strain compatibility and detailing checks. The reported φ=0.90 flexural, φ=0.75 shear/torsion and φ=0.65 bearing examples and the stated 0.65–0.90 range do not define all column states. Increasing reinforcement can change stiffness, capacity and detailing, so every affected check must be recomputed. The implemented interaction diagrams and load combinations are not available.

The reported penalty is $F(x;\lambda_t)=f(x)+\lambda_t \sum_j \max(0,g_j(x))^2$, with $\lambda_t=\lambda 0(1+t/T)^{\gamma}$ and $\gamma\in[0.5,1.5]$. Dimensionless violations permit λ_t to carry the objective's units; mixing unscaled stress, displacement and slenderness would make the sum unit-dependent. The finite schedule is bounded by $\lambda 0\cdot 2^{\gamma}$ and does not ensure feasibility. Catalog selection after repair may reintroduce violations. Nearest selection from a sorted finite catalog is not non-expansive: the catalog {1,2} maps 1.49 and 1.51 to 1 and 2, increasing their separation from 0.02 to 1.

# 4 | Proposed Approach

The intended cycle evaluates structural responses, computes violations, updates operator credits and population entropy, generates new candidates and attempts repair. Acceptance and best-state storage must apply one explicit feasibility ordering and a common current penalty if penalties change. An adaptive mixing parameter α_t is mentioned in the source but has no defined update or role. *Fig. 1* is a schematic of the intended components, not evidence that their implementation emits a certified design.

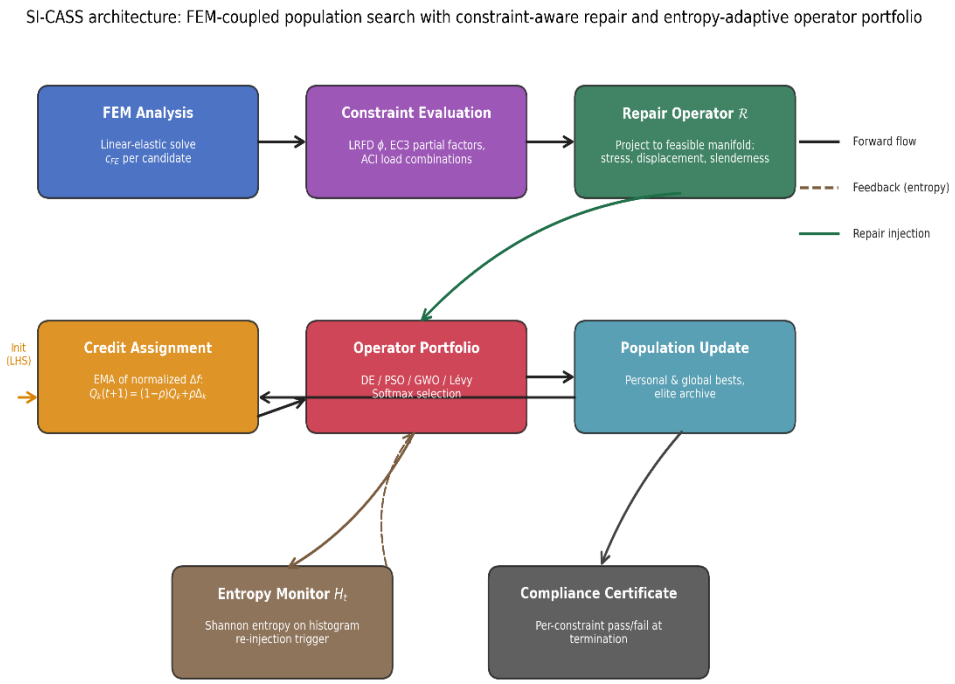



**Fig. 1. Retained architecture schematic.**

Repair and the pictured "compliance certificate" are intended workflow components; feasibility and complete project-code certification are not demonstrated. The DE branch reports $v_i=x_{r1}+F(x_{r2}-x_{r3})$, with $F\in[0.4,0.9]$ and distinct indices excluding i; the crossover or direct-replacement convention is unspecified [5].

The PSO branch reports $v_i \leftarrow \omega_t\, v_i + c1\, r1(p_i - x_i) + c2\, r2(g - x_i)$, followed by a position update, with $c1 = c2 = 1.496$ and $\omega_t = 0.4 + 0.5/(1+t/50)$ [6]. This approaches 0.4 asymptotically rather than reaching it at a finite horizon. The GWO-labelled branch reports $x_i \leftarrow (x_\alpha + x_\beta + x_\delta)/3 - A|C\, x_p - x_i|$, $A = 2ar - a$, $C = 2r$ and $a$ decreasing from 2 to 0 [18], $x_p$ and the required random draws are undefined, and this is not the standard three-leader distance update. The perturbation branch reports $x_i \leftarrow x_i + s \cdot \mathrm{Levy}(\beta = 1.5)$, with range-scaled $s$ and a reported Mantegna sampler. Heavy-tailed proposals occur in cuckoo-search literature [8], but that citation does not identify the executed SI-CASS sampler. Its vector sampling and boundary rule require confirmation. The reported formulas are preserved as definitions to complete, not silently equated with executed standard algorithms.

The reported credit rule is $Q_k(t+1) = (1-\rho)Q_k(t) + \rho\Delta_k(t)$, $\rho \in (0,1]$, with probabilities $\pi_t(k) = \exp(\eta Q_k)/\Sigma_l \exp(\eta Q_l)$. The normalized improvement $\Delta_k$, averaging over assigned particles and updates of unused operators are unspecified. Here $\eta$ is inverse temperature: increasing it sharpens preference for higher credits, and does not encourage under-used branches. Entropy $H_t = -\Sigma_b\, p_b \log p_b$ requires fixed histogram bounds, bins and a definition of coordinate aggregation; pooling areas across coordinates may miss concentration in the joint design distribution. If $H_t < H_*$, the intended rule reinjects $\varkappa \in [0.05, 0.10]$ of the population. Uniform random and Latin-hypercube reinjection are both described in the source, so the executed sampler, credit reset and rounding of $\varkappa N$ require confirmation.

The proposed repair is a local heuristic, not a demonstrated Euclidean projection onto a nonconvex feasible manifold. Area scaling by a stress ratio assumes the member force remains fixed. In indeterminate structures, stiffness changes redistribute forces and displacements, so the full structure must be reanalyzed. Displacement corrections use central finite differences and a least-squares sensitivity step, with a reported Jacobian reuse interval $\tau_J = 5$; constraints may remain violated after such a step. Radius of gyration depends on the selected section, so area alone does not define slenderness repair. RC reinforcement adjustment requires recomputation of the coupled axial–flexural and detailing checks. Bounds, catalog choices, failed repair and any stopping limit must be specified, and candidates failing final reevaluation must not enter a feasible elite archive.

Let $m_{FE,t}$ count every actual structural solve at iteration $t$, including repair reevaluation and finite-difference probes. A conditional cost expression is $O(N(n+K) + m_{FE,t}\, c_{FE} + C_{repair,t})$, with additional matrix and storage costs determined by the solver and sensitivity routine. Central finite differences can require about $2n$ probes per sensitivity refresh for one candidate, so repair is not automatically $O(n)$ or negligible. A measured dominance of FE cost by two to three orders of magnitude is not established by timings. Compactness, bounded fitness, positive operator selection and non-expansive repair would not alone establish global coverage; the catalog operation also contradicts the stated non-expansiveness premise. For one fixed original objective $f$ on a nonempty feasible domain $X_F$, define the finite infimum $f^* = \inf_{x \in X_F} f(x)$. If every exactly evaluated proposal has conditional probability at least $q_\varepsilon > 0$ of entering $\{x \in X_F : f(x) \le f^* + \varepsilon\}$, retaining the best among exactly evaluated feasible candidates gives $\Pr(\text{no hit in } m \text{ proposals}) \le (1 - q_\varepsilon)^m$. That is a conditional coverage argument for a specified process; spatial coverage and an unbounded extension are not established here. The reported Hoeffding fitness-gap bound and global convergence theorem are unsupported.

The reported initialization is a Latin-hypercube sample of $N = 50$ points within side bounds, with zero operator credits and stored initial bests. Side-bound sampling need not satisfy structural constraints. Termination uses an evaluation budget or relative improvement tolerance $\varepsilon_{stop} = 10^{-4}$ for $L_{stop} = 25$ generations. The denominator for relative improvement and all evaluation counters must be specified. The report that the budget criterion always triggered cannot be verified without logs.

# 5 | Experimental Setup

The reported suite names a 25-bar space truss with 25 variables, 10 load cases and ±275.8 MPa stress limits; a 72-bar truss with 16 sizing groups and ±172.4 MPa limits; welded beam; pressure vessel; an RC column

with axial load and biaxial bending; and a gravity retaining wall with sliding, overturning and bearing constraints. Geometry, loads, supports, materials, catalogs, objective units and feasibility tolerances are not supplied. OpenSeesPy is named for the trusses and custom section routines for the column and wall, so a shared complete FE model for all six cases is not established. Published truss studies [1], [2] cannot substitute for these missing instance definitions.

The reported baselines are GA [19], PSO [6], DE [5], GWO [7], SHADE [3], CMA-ES [4] and SQP with ten starting points. Counting an SQP restart as one optimization is not evaluation-budget matching: Every objective, constraint and finite-difference call across all starts must be counted. A smooth relaxation of a discrete catalog also solves a different domain unless final catalog feasibility is checked. The stated common interface and budgets require implementation evidence.

The source reports 30 runs per method and problem with matched seeds, a budget of 2×10^4 FE evaluations and N=50, and hardware shown in *Table 6*. The number of generations depends on whether initialization, unrepaired and repaired candidates and sensitivity probes are charged. A nominal 200 generations cannot be inferred from these values alone. The matrix 8 algorithms×6 problems×30 runs equals 1440 runs; the reported 1180 CPU-hours and 32-worker execution need timing and resource definitions. Matching seeds supports a reproducible randomization design, but does not establish statistical pairing or independence by itself.

## 6 | Results and Statistical Analysis

*Fig. 2* retains source-reported convergence curves and bands. Its horizontal axis extends to 100k FE evaluations, whereas the stated campaign budget is 20k. Several curves cross below the line labelled "known optimum 247.1 m³" and leave the visible y-range. The plot cannot establish that its curves are feasible objectives from the stated campaign or that 247.1 is a proven optimum. Reported observations of SQP=252.6, a 2.2% reduction by iteration 320, IQR≤1.5 after iteration 100 and GA IQR≥6.0 require the original traces and their units. The axis is linear in the displayed volume, despite the source description of a logarithmic fitness axis.

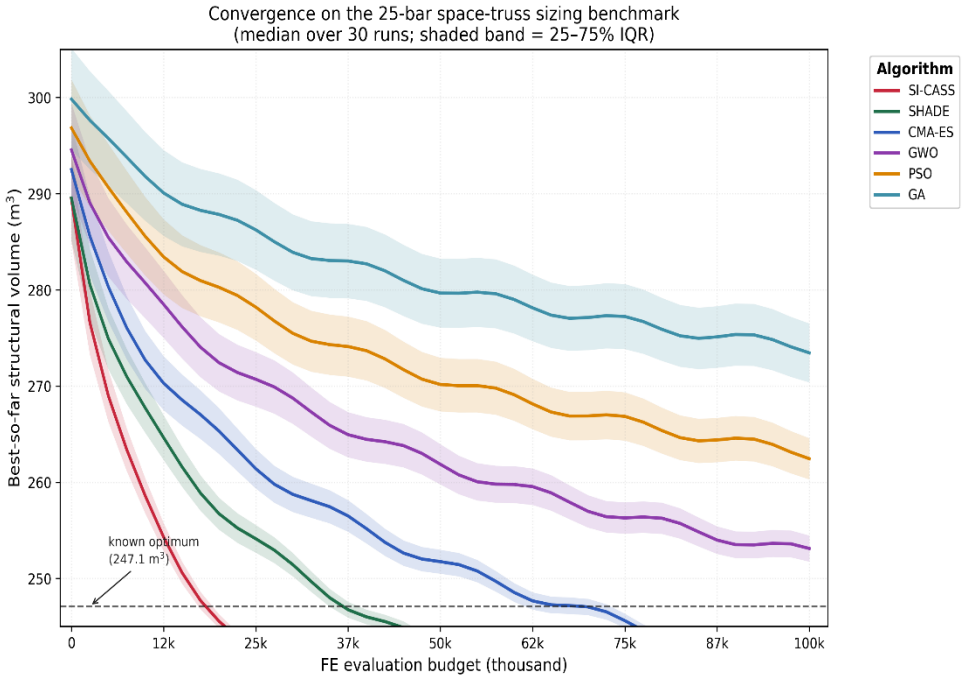



**Fig. 2. Retained reported 25-bar convergence illustration.**

The 100k FE axis, below-reference curves and "known optimum" annotation conflict with the stated 20k campaign and unresolved feasibility; the panel is not validated convergence evidence. *Fig. 3* annotates reinjection at generations 60, 130 and 200 and a threshold H_*=0.9 nats. Its SI-CASS trajectory stays roughly

between 1.3 and 1.9 nats, above that threshold, so the displayed pre-event values do not verify the stated trigger. The source reports $\varkappa$=0.07 and a roughly 30-generation credit-reset window, but actual histogram observations, controller inputs, replacement indices and credits are missing. No causal diversity-preservation conclusion follows from this panel.

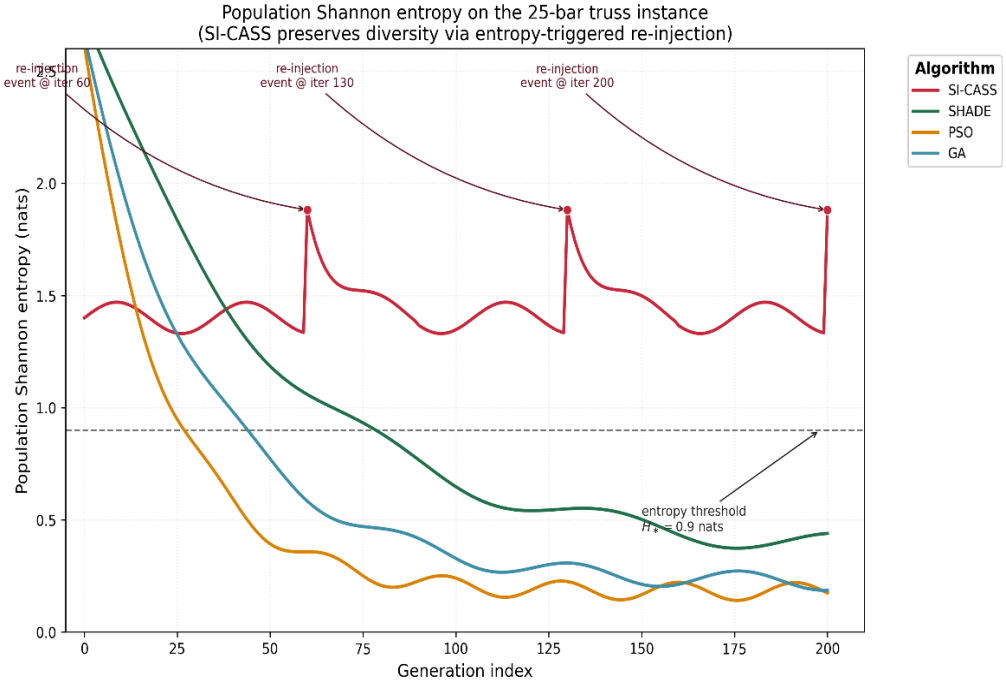



**Fig. 3. Retained reported entropy trajectories and annotated events at 60, 130 and 200.**

SI-CASS remains above the 0.9-nat threshold in the plotted trace; original controller logs are needed to reconcile the annotations.From the printed means in *Table 2*, SI-CASS has the smallest value on five of six problems, while SQP has the smallest retaining-wall value. This descriptive count differs from the source claim of four wins and from *Fig. 4*, which places SI-CASS second on welded beam and pressure vessel and third on retaining wall. Within one standard deviation is not a test of equivalence. The displayed outcomes use heterogeneous unresolved objective units, so no pooled physical-volume comparison is appropriate. Objective equations, units, feasibility and run-level provenance are unresolved. The bottom "Mean rank" row has entries under problem columns and is not a valid vector of eight algorithm mean ranks.

**Table 2. Source-reported final-objective means and standard deviations, described as 30-run summaries.**

| Algorithm | 25-Bar Truss | 72-Bar Truss | Welded Beam | Pressure Vessel | RC Column | Retaining Wall |
|---|---|---|---|---|---|---|
| SI-CASS | 247.1 ± 2.8 | 172.4 ± 4.6 | 1.832 ± 0.077 | 5,896 ± 81 | 0.832 ± 0.029 | 12.51 ± 0.42 |
| SHADE | 252.4 ± 3.6 | 176.8 ± 5.2 | 1.847 ± 0.082 | 5,942 ± 90 | 0.846 ± 0.031 | 12.78 ± 0.46 |
| CMA-ES | 255.6 ± 4.9 | 180.2 ± 6.7 | 1.879 ± 0.093 | 6,012 ± 102 | 0.858 ± 0.038 | 13.10 ± 0.55 |
| GWO | 259.3 ± 5.8 | 183.7 ± 7.2 | 1.913 ± 0.101 | 6,108 ± 119 | 0.873 ± 0.041 | 13.42 ± 0.61 |
| PSO | 263.5 ± 7.4 | 188.9 ± 8.6 | 1.952 ± 0.118 | 6,219 ± 138 | 0.891 ± 0.049 | 13.86 ± 0.78 |
| GA | 268.7 ± 8.6 | 192.4 ± 9.8 | 1.985 ± 0.127 | 6,347 ± 161 | 0.906 ± 0.057 | 14.23 ± 0.92 |
| DE | 261.8 ± 6.3 | 185.3 ± 7.9 | 1.928 ± 0.109 | 6,164 ± 124 | 0.879 ± 0.044 | 13.55 ± 0.66 |
| SQP | 271.3 ± 4.1 | 186.5 ± 4.8 | 1.902 ± 0.063 | 6,221 ± 78 | 0.893 ± 0.022 | 12.29 ± 0.18 |
| Mean rank | 1.5 | 2.3 | 1.8 | 1.8 | 1.7 | 3.0 |

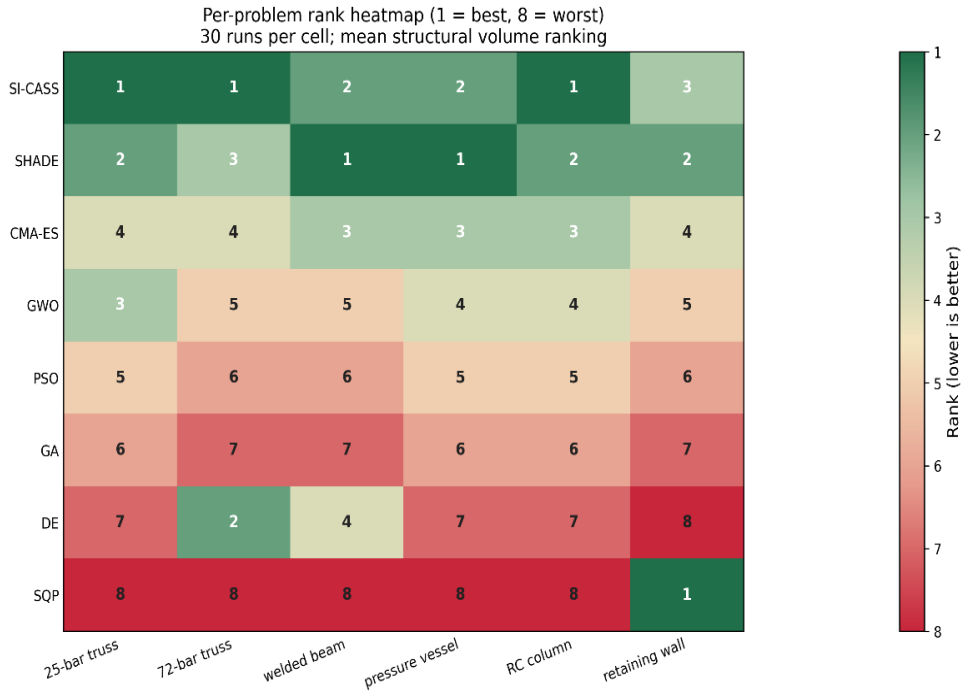



**Fig. 4. Retained reported rank heatmap.**

Several cells disagree with ranking the printed Table 2 means; the panel and table cannot jointly establish an inferential ranking. *Table 3* retains seven reported raw p-values, adjusted p-values and rank-biserial effects. Its adjustment is not the Holm step-down transformation of the displayed seven raw values. Applying only that transformation gives adjusted values 0.0068 (SHADE), 0.0048 (CMA-ES), 0.0045 (GWO), 0.0024 (PSO), 0.0014 (GA), 0.0048 (DE) and 0.0068 (SQP). This is an arithmetic check, not validation of the raw tests or evidence of significance. The experimental unit is unspecified: six paired problem summaries, 30 paired runs on one problem and pooled heterogeneous outcomes are different designs. For six nonzero pairs, an exact two-sided signed-rank test has minimum p=0.03125, incompatible with the displayed values. Test implementation, pairing, ties and aggregation must be identified [20], [21].

**Table 3. Source-reported comparison statistics.**

| Baseline | Raw p | Holm-adj. p | Rank-Biserial r |
|---|---|---|---|
| SHADE | 0.0034 | 0.024 | 0.71 |
| CMA-ES | 0.0012 | 0.0084 | 0.78 |
| GWO | 0.0009 | 0.0063 | 0.81 |
| PSO | 0.0004 | 0.0028 | 0.85 |
| GA | 0.0002 | 0.0014 | 0.88 |
| DE | 0.0014 | 0.0098 | 0.76 |
| SQP | 0.0060 | 0.018 | 0.62 |

All numerical cells are preserved; the printed "Holm-adj." values do not equal the adjustment of the displayed raw values. The source reports Friedman $\chi^2=38.2$, $p<0.001$, $k=8$, $N=6$ and Nemenyi CD=1.83. Original rank blocks are not supplied, and the table and heatmap are inconsistent. Under the conventional Nemenyi formulation $CD=q_\alpha\sqrt{[k(k+1)/(6N)]}$, $k=8$ and $N=6$ give about 4.29 at $\alpha=0.05$, not 1.83 [20]. Mean ranks 1.50, 1.83 and 4.17 in the source diagram are therefore retained as unresolved reported quantities. Non-significance does not establish equivalence, and the panel does not justify separating the first two algorithms from all remaining methods.

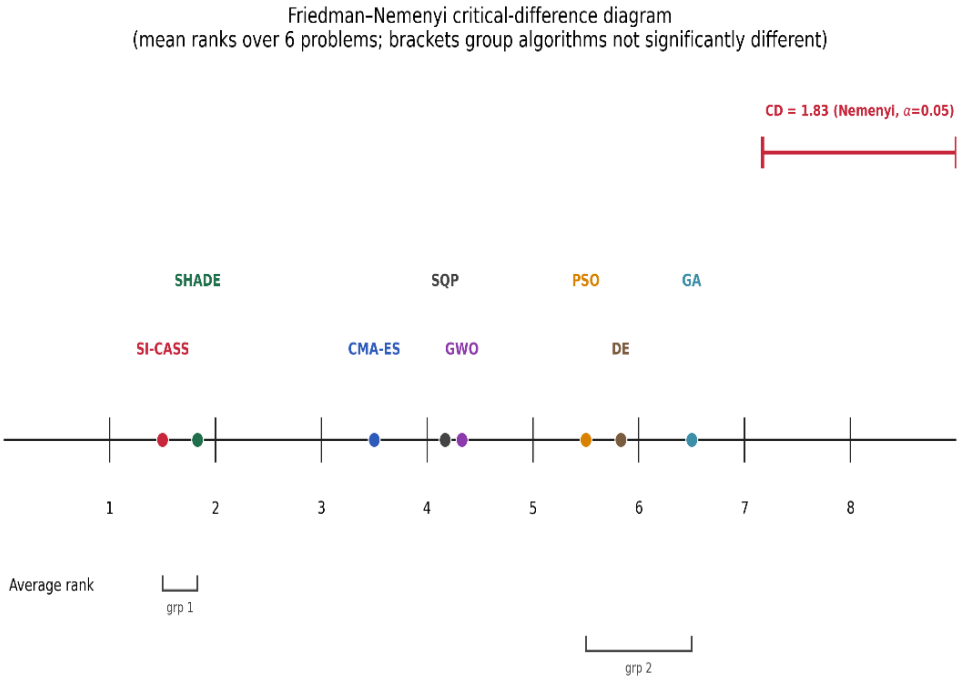



**Fig. 5. Retained reported critical-difference illustration.**

Rank provenance and CD=1.83 are unresolved and inconsistent with the stated conventional k=8, N=6 analysis; its brackets are not established significance groups.*Table 4* retains reported Cliff's δ, Hedges' g and BCa intervals. The intervals exclude zero as printed, but raw samples, sign convention, pairing, standardization and bootstrap resampling units are unavailable. Cliff's δ is a pairwise dominance measure; it is not interchangeable with paired rank-biserial correlation. Hedges' g requires its specific variance definition. The reported 10^4 resamples and 95% intervals do not by themselves establish robustness, and a result on one truss cannot be attributed to all six problems [21].

**Table 4. Source-reported effects and 95% BCa intervals, described as 30 paired runs on the 25-bar truss with 10,000 resamples.**

| Baseline | Cliff's δ | δ 95% BCa CI | Hedges' g | g 95% BCa CI |
|---|---|---|---|---|
| SHADE | 0.55 | [0.32, 0.74] | 0.85 | [0.41, 1.27] |
| CMA-ES | 0.63 | [0.41, 0.81] | 1.06 | [0.58, 1.52] |
| GWO | 0.66 | [0.43, 0.84] | 1.18 | [0.67, 1.66] |
| PSO | 0.72 | [0.51, 0.88] | 1.31 | [0.78, 1.81] |
| GA | 0.78 | [0.58, 0.92] | 1.42 | [0.86, 1.94] |
| DE | 0.64 | [0.42, 0.82] | 1.09 | [0.61, 1.55] |
| SQP | 0.42 | [0.18, 0.64] | 0.41 | [0.07, 0.74] |

All cells remain unchanged; definitions and run-level resampling evidence are unavailable. For six factors and r=10 Morris trajectories, r(k+1)=70 response-model evaluations per instance, or 420 across six instances [22]. If the response is optimizer performance, each such call is an optimization run, with its own structural evaluations and stochastic replication; it is not one FE solve. The source uses six levels but omits parameter ranges, trajectories, output scaling, repeated seeds and response data. *Fig. 6* reports μ* and σ with the credit-rate entries mostly around 0.2; the report alternately names ϰ or portfolio size K as a top factor, while the plot does not include K. These inconsistencies prevent tuning recommendations or variance attribution. Morris σ suggests nonlinearity or interactions but does not separate them.

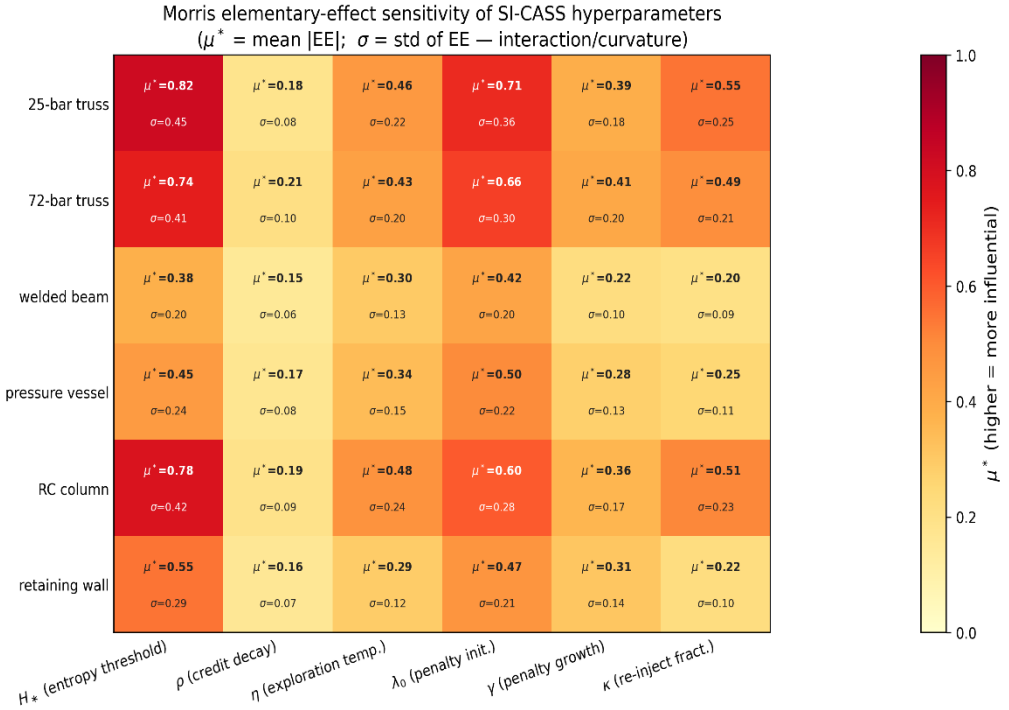



**Fig. 6. Retained source-reported Morris μ* and σ values for six parameters.**

Sampling design, response scale and optimizer-run budget are unresolved; $\eta$ is inverse temperature despite the retained "exploration temp." label. Relative to the printed 247.1 mean, removing repair, adaptive selection and entropy reinjection increases the reported means by about 5.18%, 3.40% and 2.10%, respectively; removing LHS gives about 1.58%. The SI-CASS-K mean is about 0.405% higher and its reported FE count falls from 20,000 to 12,400, a 38% reduction. These arithmetic statements do not identify causal effects, feasible-design quality or measured runtime. Identical wall-clock ratios of 1.00 for variants with different repair costs require actual timing and budget logs. Candidate feasibility, branch isolation, actual solver calls and timings require verification.

**Table 5. Retained source-reported ablation and Kriging-variant summaries for the 25-bar instance.**

| Variant | Mean Volume ($m^3$) | FE Evaluations | Wall-Clock Ratio |
|---|---|---|---|
| SI-CASS (full) | 247.1 ± 2.8 | 20,000 | 1.00× |
| − repair operator | 259.9 ± 9.1 | 20,000 | 1.00× |
| − adaptive operator selection | 255.5 ± 5.7 | 20,000 | 1.00× |
| − entropy re-injection | 252.3 ± 4.4 | 20,000 | 1.00× |
| − LHS initialization | 251.0 ± 4.9 | 20,000 | 1.00× |
| SI-CASS-K (Kriging surrogate) | 248.1 ± 3.2 | 12,400 | 0.62× |

*Table 6* records the reported software and hardware configuration. The seed prefix is not a complete seed list, and no resolvable Zenodo deposit identifier, code or problem package is supplied here. The report that SciPy and scikit-posthocs produced the statistics does not resolve the inconsistencies. Reproduction requires the exact implementation, full seedfile.json, models, catalogs, per-run outputs, counters and timing/statistical scripts. Software execution and seed/deposit availability are unverified; the named seed file is required for reproduction.

**Table 6. Source-reported configuration information.**

| Item | Value |
|---|---|
| Python | 3.11.6 |
| NumPy | 1.26.0 |
| SciPy | 1.11.3 |
| scikit-posthocs | 0.7.0 |
| OpenSeesPy | 3.3.2 |
| DEAP | 1.3.3 |
| RNG | splitmix64 → MT19937 |
| Seeds | 0x4E1A9C7B… (30 seeds documented in seedfile.json) |
| CPU | AMD EPYC 7B13 @ 2.45 GHz, 64 cores |
| RAM | 256 GB DDR4-3200 |
| OS | Ubuntu 22.04 LTS |
| Parallelism | 32 workers via multiprocessing.Pool |

# 7 | Discussion

The printed means suggest small descriptive differences from SHADE, including about 2.10% on the 25-bar instance. A portfolio might improve search when different proposal scales are useful, but this explanation requires operator-selection histories and controlled budget-matched branch comparisons. *Table 4* is labelled as a single-truss effect study and does not establish which problems benefit most.

Repair has the largest reported ablation mean difference, but this cannot be interpreted as a universal component ranking. A repair may change objective, feasibility and evaluation cost simultaneously. A candidate's fitness not worsening is a property of comparison and acceptance, whereas non-expansiveness bounds the distance between the repaired images of any two inputs; the former does not imply the latter. Force redistribution and discrete catalogs make simple projection arguments inappropriate. Reinjection is intended to preserve search opportunities while retaining verified elite candidates. The reported 5%–10% fractions, 30-generation credit window and 2%–3% gains do not establish that this mechanism avoids an exploitation cost. *Fig. 3* does not show its stated threshold crossings. Selection traces and common-budget ablations are needed to distinguish reinjection benefits from additional evaluations or other changes. A two-stage exploratory search followed by a deterministic solver is a possible workflow to test. It is not validated by the current summaries. Smoothness alone does not ensure global optimality, catalog feasibility or code compliance. The reported 38% solver-call reduction and 0.4% objective increase for Kriging need original logs before recommending that variant for expensive nonlinear or transient structural models.

# 8 | Threats to Validity and Counterarguments

Six continuous hyperparameters are named, $\varrho$, $\eta$, H_*, $\varkappa$, $\lambda 0$ and $\gamma$, along with the separate discrete portfolio-size K, giving seven named design choices. Their default values and full adaptation schedules are not supplied. Finite-catalog rounding is demonstrably not non-expansive, and bounded changing penalties do not establish global convergence. Conditional spatial coverage would have to be proved for proposals after repair and rounding. The truss response is described as linear-elastic, while the column and wall use simplified section models. These do not establish behavior under cracking, yielding, geometric instability, seismic response or full soil–structure interaction. The complete problem definitions and validation against independent structural calculations are absent. SQP's proposed continuous relaxation and ten starts must be evaluated with the same total solver-call budget and final discrete domain as its comparators.

Inferential limitations include ambiguous pairing, no rank-block data, incorrect displayed Holm adjustments and an incompatible stated CD. Claims of identical Holm and FDR rejections are not justified without an identified family and actual calculations. A BCa resampling count, or n=30 by itself, does not guarantee interval accuracy. Sample skewness, ties, degeneracy and the resampling unit must be examined using original outputs. Most importantly, the reported outcomes are not project-ready structural designs. Complete loads, units, section properties, limit states, code editions and national annexes, tolerances, feasible vectors and

design-ratio reports are unavailable. The simplified allowable-stress model cannot be equated with complete AISC, Eurocode or ACI compliance. No code-compliance certificate has been demonstrated. Additional source-reported claims include a twenty-start SQP pilot improving the wall result by less than 0.1% and a 1024-sample Sobol check of H_* and λ0. Their outputs, estimator design and budget are absent. They cannot substantiate significance, tuning or robustness conclusions. The decisive next evidence is the original model and computational package, followed by independent checking of final feasible designs and correctly defined statistics.

# 9 | Conclusion and Future Work

SI-CASS describes a candidate combination of adaptive proposals, local structural repair and diversity reinjection. Its reported numerical observations are retained, but not independently reproduced. The present analysis corrects the interpretation of inverse temperature, repair, catalog projection, objective units and computational budget, and identifies inconsistencies that prevent verification of statistical superiority or causal component effects. Neither a global convergence guarantee nor complete structural-code certification is established.

Further work should first supply and validate the six structural models and executed algorithm, establish final feasibility and complete resource counters, and recompute the inferential analyses from original outcomes. Extensions to nonlinear response, learned proposals, distributed evaluation or reliability-based design require their own formulations and validation. They should follow that evidence base rather than be inferred from the current summaries.

# Conflict of Interest

The authors declare that there are no financial, personal, institutional, or other relationships that could be perceived as influencing the work reported in this manuscript. The authors confirm that they have no competing interests regarding the publication of this article.

# Data Availability

The data supporting the findings of this study are fully included within the manuscript. No additional datasets were generated or analyzed beyond those presented in the text. Further information regarding the data may be obtained from the corresponding author upon reasonable request.

# Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The study was conducted without external financial support, and all aspects of the research, including study design, data collection, analysis, interpretation, and manuscript preparation, were carried out independently by the authors.

# References

[1] Avcı, M. S., Yavuz, D., Ercan, E., & Nuhoğlu, A. (2024). Efficient sizing and layout optimization of truss benchmark structures using ISRES algorithm. *Applied sciences*, *14*(8), 3324. https://doi.org/10.3390/app14083324

[2] Mehta, P., Tejani, G. G., & Mousavirad, S. J. (2025). Structural optimization of different truss designs using two archive multi objective crystal structure optimization algorithm. *Scientific reports*, *15*(1), 14575. https://doi.org/10.1038/s41598-025-97133-w

[3] Tanabe, R., & Fukunaga, A. (2013). Success-history based parameter adaptation for differential evolution. *2013 ieee congress on evolutionary computation* (pp. 71–78). IEEE. https://doi.org/10.1109/CEC.2013.6557555

[4] Hansen, N. (2006). The CMA Evolution Strategy: A Comparing Review. In *Towards a new evolutionary computation: advances in the estimation of distribution algorithms* (pp. 75–102). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-32494-1_4

[5] Storn, R., & Price, K. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, *11*(4), 341–359. https://doi.org/10.1023/A:1008202821328

[6] Eberhart, R., & Kennedy, J. (1995). Particle swarm optimization. *Proceedings of the IEEE international conference on neural networks* (Vol. 4, pp. 1942–1948). IEEE. https://doi.org/10.1109/ICNN.1995.488968

[7] Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in engineering software*, *69*, 46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007

[8] Yang, X. S., & Deb, S. (2009). Cuckoo search via lévy flights. *2009 world congress on nature & biologically inspired computing (NABIC)* (pp. 210–214). IEEE. https://doi.org/10.1109/NABIC.2009.5393690

[9] Deb, K. (2000). An efficient constraint handling method for genetic algorithms. *Computer methods in applied mechanics and engineering*, *186*(2), 311–338. https://doi.org/10.1016/S0045-7825(99)00389-8

[10] Takahama, T., & Sakai, S. (2010). Constrained optimization by the ε constrained differential evolution with an archive and gradient-based mutation. *IEEE congress on evolutionary computation* (pp. 1–9). IEEE. https://doi.org/10.1109/CEC.2010.5586484

[11] Coello Coello, C. A. (2002). Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: A survey of the state of the art. *Computer methods in applied mechanics and engineering*, *191*(11), 1245–1287. https://doi.org/10.1016/S0045-7825(01)00323-1

[12] Papadrakakis, M., Lagaros, N. D., & Tsompanakis, Y. (1998). Structural optimization using evolution strategies and neural networks. *Computer methods in applied mechanics and engineering*, *156*(1), 309–333. https://doi.org/10.1016/S0045-7825(97)00215-6

[13] Ton-That, H. L. (2026). Identification of damage in planar truss structures within a multiphysics framework using metaheuristic optimization algorithms. *Journal of complex and multiphysics engineering systems*, *1*(2), 138–147. https://doi.org/10.56578/jcmes010202

[14] American Institute of Steel Construction. (2005). *Specification for structural steel buildings*. https://www.aisc.org/aisc/publications/historic-standards/specification-for-structural-steel-buildings/

[15] European Committee for Standardization. (2005). *Eurocode 3: Design of steel structures—part 1-1: General rules and rules for buildings*. https://eurocodes.jrc.ec.europa.eu/EN-Eurocodes/eurocode-3-design-steel-structures

[16] European Committee for Standardization. (2002). *Eurocode: Basis of structural design*. https://www.bing.com/ck/a?!&&p=97c394ed83be2e71a333bea7a1548f199eca3b5a90da0d06c4a1eb7b9f4e83

[17] American Concrete Institute. (2019). *Building code requirements for structural concrete (ACI 318-19) and commentary (ACI 318R-19).* https://www.concrete.org/Portals/0/Files/PDF/Previews/318-19_preview.pdf

[18] Shokouhifar, M., & Goli, A. (2023). Designing a resilient--sustainable supply chain network of age-differentiated blood platelets using vertical--horizontal transshipment and grey wolf optimizer. *International journal of environmental research and public health*, *20*(5), 4078. https://doi.org/10.3390/ijerph20054078

[19] Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press. https://d1wqtxts1xzle7.cloudfront.net/98698869/1216504-libre.pdf?1676455734

[20] Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *The journal of machine learning research*, *7*, 1–30. https://doi.org/10.5555/1248547.1248548

[21] Derrac, J., García, S., Molina, D., & Herrera, F. (2011). A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and evolutionary computation*, *1*(1), 3–18. https://doi.org/10.1016/j.swevo.2011.02.002

[22] Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161–174. https://doi.org/10.1080/00401706.1991.10484804