

Article

A Sensitivity-Guided Selection Hyper-Heuristic for the Thermo-Hydraulic Design of a Solar Hybrid-Nanofluid Evacuated-Tube Collector

Faris Alqurashi  and Muhammed Anaz Khan * 

Department of Mechanical Engineering, College of Engineering, University of Bisha, P.O. Box 551, Bisha 61922, Saudi Arabia; falqurashi@ub.edu.sa

* Correspondence: mkhan@ub.edu.sa

Abstract

Hybrid nanofluids raise the thermal output of evacuated-tube solar collectors, but the heat-transfer gain comes at the cost of friction and pumping power, so the working fluid must be chosen in line with a constrained thermo-hydraulic criterion. Its design is posed here as a constrained single-objective optimisation over the hybrid pair, base fluid, weight fraction, component share and flow rate. A variance-based screening of a 54,432-run full-factorial dataset reduces the design space from eight variables to five, justified by the invariance of the optimal hybrid pair across thirty-six operating points. Gradient-boosted surrogates for the four state and constraint responses (efficiency, pumping power, Reynolds number and outlet temperature) reproduce the simulator to held-out coefficients of determination of 0.9998–1.0000; the performance-criterion surrogate has a lower global coefficient of determination (0.46) and is assessed by top-region ranking accuracy. The reduced problem is solved with a selection hyper-heuristic over twelve operators, pairing an upper-confidence-bound selector with late-acceptance hill-climbing. Under Friedman-, Nemenyi- and Holm-corrected Wilcoxon testing, it ranks within the leading statistically indistinguishable group, but is not separable from a random-selection ablation, locating its value in robustness rather than adaptivity; furthermore, it attains the grid-reference optimum within 0.40 percent. The criterion optimum is an Al_2O_3 -Cu suspension in ethylene-glycol and water at three percent loading; the constrained-efficiency optimum reaches a surrogate thermal efficiency of 0.760, cross-checked against a grid-reference value of 0.762 and a reconstructed-model value of 0.760, conditional on the supplied reduced-order model and its single-tube dataset convention. The constrained-efficiency objective is treated as the primary design objective, while the performance criterion is reported as a secondary screening index that characterises the additive rather than selecting the operating fluid.



Academic Editor: Paloma Martínez-Merino

Received: 13 July 2026

Revised: 24 July 2026

Accepted: 29 July 2026

Published: 3 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: hybrid nanofluid; evacuated-tube solar collector; thermo-hydraulic performance evaluation criterion; selection hyper-heuristic; adaptive operator selection; gradient-boosting surrogate; constrained single-objective optimisation; critical-difference analysis

1. Introduction

Evacuated-tube solar collectors suppress convective and conductive losses through a vacuum envelope and therefore sustain higher efficiencies at elevated inlet temperatures than flat-plate designs, which makes them attractive for domestic hot water, process heat and space conditioning [1,2]. The working fluid sets an upper bound on collector

performance because the useful gain is influenced by the in-tube convective coefficient and the volumetric heat capacity through the heat-removal factor and the collector heat-removal process. Conventional carriers such as water, ethylene-glycol mixtures and synthetic oils have modest thermal conductivities, which has motivated the dispersion of solid nanoparticles to raise the effective transport properties of the carrier [3]. Hybrid nanofluids, which disperse two complementary solids in one carrier, combine a high-conductivity constituent with one that improves stability or lowers cost, and they have been shown to raise the useful gain of evacuated-tube units relative to water and to mono suspensions [4–8]. The thermophysical closures derive from the classical effective-medium and suspension-viscosity models [9–11], and the collector-level benefits of mono and hybrid suspensions are documented across flat-plate and evacuated-tube studies [12–20].

The benefit of a nanoparticle additive is always accompanied by a friction and pumping-power penalty, so a fair appraisal must weigh the heat-transfer gain against the hydraulic cost. That trade-off is captured by the performance evaluation criterion—the ratio of the Nusselt enhancement to the cube root of the friction-factor ratio [4,5,21]. The cube-root exponent derives from an equal-pumping-power comparison; because the present dataset evaluates the nanofluid and its baseline at a common imposed flow rate, the quantity is used here as a relative thermo-hydraulic index at equal flow rate, a distinction made precise in Section 2.1. Selecting a hybrid pair, a carrier, a loading and a flow rate that maximise this criterion, or that maximise thermal efficiency under a hydraulic constraint, is a constrained search over a mixed categorical and continuous design space. Prior data-driven treatments of nanofluid property and performance design have coupled neural networks with sensitivity screening or response-surface optimisation, and several have used non-dominated sorting to expose a Pareto front of conductivity against viscosity [22,23]. A Pareto front is informative when the designer wants the full trade-off surface, but it does not answer the single-objective question of which fluid maximises a chosen scalar merit under explicit feasibility constraints, and it commits the analyst to one fixed multi-objective evolutionary engine.

A hyper-heuristic offers a different and, for this problem, more appropriate paradigm. Rather than fixing a single metaheuristic, a selection hyper-heuristic operates at a higher level of abstraction and manages a portfolio of low-level heuristics, learning during the search which operator to apply next from recent performance [24–26]. The idea originates in the choice-function scheduler of Cowling, Kendall and Soubeiga [24]. Subsequent work established that the heuristic-selection rule and the move-acceptance criterion can be designed almost independently, and that simple but effective acceptance strategies such as late-acceptance hill-climbing and the great-deluge algorithm are remarkably strong [27,28]. Heuristic selection has been cast as a multi-armed bandit problem in which an upper-confidence-bound rule balances the exploitation of recently productive operators against the exploration of operators not tried for some time [29,30]. The appeal of the paradigm for engineering design is generality across heterogeneous landscapes without per-problem re-tuning, a property anticipated by the no-free-lunch theorems [31].

This study addresses the working-fluid and operating-envelope selection problem for a hybrid-nanofluid evacuated-tube collector and redevelops a variance-based screening analysis of that collector into a computational design study whose methodological core is a selection hyper-heuristic. The screening is retained but its role is changed: first-order and grouped variance indices, factorial-analysis-of-variance partition and Morris elementary-effects screening are used to rank the decision variables and to justify fixing or narrowing the least influential of them before the optimiser runs. The primary contribution is computational and methodological, and the collector optima are reported as conditional on the supplied reduced-order model, its closures and the sampled envelope, rather than as experimentally validated physical recommendations. The specific contributions are as

follows: A sensitivity-guided reduction of the design space from eight variables to five is derived from an exact variance decomposition and is justified by the invariance of the PEC-optimal hybrid pair and by the near-degeneracy of the nanoparticle variables in the constrained-efficiency objective, rather than by invariance of the full fluid composition or by neglecting operation; the resulting five-variable problem is a fixed-operating-point design demonstration and does not replace operating-condition-specific carrier selection over the full eight-variable space. Gradient-boosted regression-tree surrogates are trained on the dataset, their held-out and blocked accuracy is reported, and they are used as the cheap evaluators. A selection hyper-heuristic, pairing an upper-confidence-bound credit-assignment selector with late-acceptance hill-climbing, is implemented over twelve operators and described with full pseudocode. The adaptive selector is evaluated against a learning-free random-selection ablation, against each constituent operator and against strong standalone metaheuristics with a Friedman-, Nemenyi- and Holm-corrected Wilcoxon protocol; the results show the robustness of the portfolio framework while finding the statistical advantage of adaptivity over random selection to be limited. The collector optima are verified against an exhaustive grid enumeration and a reconstructed reference implementation, and the residual gaps are reported honestly.

2. Problem Formulation and Reduced-Order Model

2.1. Reduced-Order Collector Model

The study is built on a provided dataset of 54,432 full-factorial runs of a closed pumped-loop evacuated-tube collector. Each run converts the nanoparticle mass fractions to volume fractions; forms the hybrid density and specific heat by volume-weighted mixing, the hybrid conductivity by the Takabi and Salehi law and the hybrid viscosity by the Brinkman law; evaluates the base-fluid properties at the converged mean temperature; and iterates the mean temperature to self-consistency through the Reynolds and Prandtl numbers, the Nusselt number, the convective coefficient, the Hottel–Whillier–Bliss collector balance and the outlet temperature [32]. In the laminar regime, below a Reynolds number of 2300, the Nusselt number uses the fully developed value or the Hausen entry correlation [33]; above 2300 it uses the Gnielinski correlation with the Petukhov friction factor [34–36]. The Gnielinski correlation is applied in the turbulent branch; because several candidate optima lie near the lower transitional range, the results in this regime are treated as closure-sensitive. The dataset applies a hard switch at a Reynolds number of 2300; the narrow transitional band is therefore a known source of model-form uncertainty, and a continuous blending across the band was shown in the source screening study to leave the driver rankings unchanged. The Brinkman- and Maxwell-type closures are most defensible for dilute, approximately spherical suspensions; for the high-aspect-ratio carbon-nanotube and graphene systems, they are an approximation, which is treated as a model-form caveat in Section 7.

The hybrid-nanofluid properties follow from the mass-to-volume-fraction conversion, volume-weighted mixing, the Takabi–Salehi conductivity and the Brinkman viscosity:

$$w_1 = s w, w_2 = (1 - s) w \text{ (} w \text{ total loading as a fraction; percent values enter as fractions, e.g., three percent = 0.03; } s \text{ component-1 share); } \varphi = \varphi_1 + \varphi_2 \quad (1)$$

$$\varphi_i = (w_i / \rho_i) / [w_1 / \rho_1 + w_2 / \rho_2 + (1 - w_1 - w_2) / \rho_{bf}] \quad (2)$$

$$\rho_{hnf} = (1 - \varphi) \rho_{bf} + \varphi_1 \rho_{p1} + \varphi_2 \rho_{p2}; (\rho c_p)_{hnf} = (1 - \varphi) (\rho c_p)_{bf} + \varphi_1 (\rho c_p)_{p1} + \varphi_2 (\rho c_p)_{p2} \quad (3)$$

$$c_{p,hnf} = (\rho c_p)_{hnf} / \rho_{hnf} \text{ (required because } Pr = \mu c_p / k \text{ uses } c_p, \text{ not } \rho c_p) \quad (4)$$

$$k_p = (\varphi_1 k_{p1} + \varphi_2 k_{p2}) / \varphi; k_{\text{hnf}} = k_{\text{bf}} [k_p + 2k_{\text{bf}} + 2\varphi(k_p - k_{\text{bf}})] / [k_p + 2k_{\text{bf}} - \varphi(k_p - k_{\text{bf}})]; \mu_{\text{hnf}} = \mu_{\text{bf}} (1 - \varphi)^{-2.5} \quad (5)$$

The performance evaluation criterion is retained in its classical form.

$$\text{PEC} = (Nu/Nu_{\text{bf}}) / (f/f_{\text{bf}})^{1/3} \quad (6)$$

The classical cube-root exponent is normally associated with an equal-pumping-power comparison. Because the present dataset evaluates the nanofluid and its base-fluid baseline at common imposed flow rates rather than common pumping power, the criterion is interpreted here as a relative thermo-hydraulic index at equal flow rate rather than as a strict equal-pumping-power merit. The design space is spanned by three hybrid pairs, three base fluids, seven total weight fractions from 0.25 to 3.0 percent, three component shares, eight loop flow rates from 0.5 to 6.0 litres per minute, four inlet temperatures, three irradiances and three ambient temperatures, with the particle diameter held at 40 nanometres because the conductivity and viscosity closures are size independent. To support an interpolation-only search and to verify off-grid solutions, the reduced-order simulator was reconstructed from these equations with the property constants recovered from the dataset (Table 1); the reconstruction reproduces the tabulated efficiency, outlet temperature, Reynolds number and performance criterion with a median relative error near 0.005 percent and a worst case of 0.66 percent, which qualifies it as an independent reference implementation of the same reduced-order model rather than an independent physical truth model. Because the recovered MWCNT-Fe₃O₄ conductivity coefficient is degenerate and nonphysical (Table 1), the criterion optimisation was repeated with that pair excluded; the optimum was unchanged—an Al₂O₃-Cu suspension in the ethylene-glycol and water mixture at three percent—so the reported design does not depend on the nonphysical fit, and the grouped variance partition is likewise robust to this exclusion, with the operating share of the criterion variance changing only from 4.8 to 3.6 percent and the composition share from 25.0 to 24.8 percent.

The fixed 40 nm diameter is a property of the supplied closures, in which the Takabi-Salehi conductivity and Brinkman viscosity laws depend only on the volume fraction and not on particle size, so within this dataset the diameter is a strictly inert factor that cannot be optimised. This is a genuine model-form limitation and not a claim that size is unimportant: experiments report that reducing particle diameter raises the thermal-conductivity enhancement while simultaneously increasing the viscosity penalty, so size acts on both the numerator and the denominator of the performance criterion in partially compensating directions [37,38]. Two consequences follow: First, particle size enters the conductivity; the viscosity; and hence the Reynolds, Nusselt and friction quantities nonlinearly—so its net effect on the criterion, and in particular whether the Al₂O₃-Cu ranking is preserved, cannot be determined from the size-independent closures used here. Establishing the robustness of that ranking requires re-analysis with validated size-dependent closures. Second, the framework accommodates size directly: a size-dependent conductivity model (for example a Brownian-motion or interfacial-layer correction) and a size-dependent viscosity model can be substituted into the property block and the diameter promoted from a fixed parameter to a sixth decision variable, after which the identical sensitivity-guided reduction and hyper-heuristic apply without modification. The present optima are therefore reported as conditional on the size-independent closures of the supplied dataset, and the size-dependent extension is set out in the revised Future Work.

Table 1. (a) Fitted effective coefficients recovered from the dataset to reconstruct the supplied model; these are not physical material constants. The recovered densities and specific heats match the accepted literature values, whereas the conductivities are effective blending coefficients of the Takabi–Salehi closure; for the MWCNT-Fe₃O₄ pair, the conductivity fit is degenerate and nonphysical. (b) Base-fluid property correlations recovered from the dataset, evaluated at 25 and 75 degrees Celsius. The fitted forms are quadratic in temperature for density and conductivity, cubic for specific heat, and a cubic polynomial in temperature for the logarithm of viscosity; the full coefficients are in the released data files.

(a)						
Nanoparticle	ρ (kg m ^{−3})		c_p (J kg ^{−1} K ^{−1})		k (Effective)	
Al ₂ O ₃	3970		765		43.6	
Cu	8933		385		257	
MWCNT	2100		711		591.3	
Fe ₃ O ₄	5180		670		−1475.3	
Graphene	2250		2100		1721.4	
TiO ₂	4250		686		3699.9	
(b)						
Base Fluid	ρ (25C)	c_p (25C)	k(25C)	μ (25C) mPa·s	ρ (75C)	μ (75C) mPa·s
Distilled water	997	4182.3	0.6062	0.8908	974.9	0.3744
EG/water 60:40	1059.9	3312	0.4095	3.6199	1028.9	1.0346
Synthetic HTF oil	856.8	1917	0.1326	24.6987	824.2	4.8011

2.2. Decision Variables, Objectives and Constraints

The collector design is recast as a constrained single-objective problem distinct from a Pareto study. The decision vector contains two categorical variables, the hybrid pair and the base fluid, and three continuous variables—the total weight fraction in 0.25 to 3.0 percent, the component-1 mass share in 25 to 75 percent, and the loop flow rate in 0.5 to 6.0 litres per minute. The categorical variables are carried as real coordinates and decoded to a category by flooring before evaluation, and dedicated swap operators act on them directly; the surrogate uses the same integer category labels. To confirm that this integer encoding does not impose an artificial ordinal structure on the categories, a surrogate retrained under a permuted category-to-integer mapping and, more conclusively, a surrogate retrained with a one-hot encoding that removes any ordinal structure, both returned the same optimal hybrid pair—Al₂O₃-Cu—so the optimum is invariant to the encoding; beyond the optimum, the integer and one-hot surrogate predictions give strongly consistent global rankings and recover the same optimum, with a Spearman correlation of 0.98 over all grid designs, although the top-region rank correlation is more moderate at 0.75. The reduction of the design space, derived in Section 3, fixes the three environmental operating variables—the inlet temperature, the irradiance and the ambient temperature—at a representative point of 40 degrees Celsius, 1200 watts per square metre and 25 degrees Celsius, with a secondary point at 60 degrees Celsius inlet. The loop flow rate, although it belongs to the operating block of the variance partition, is retained as a decision variable because it controls the hydraulic constraint.

Two scalar objectives are studied. The primary case maximises thermal efficiency subject to a pumping-power ceiling, which bounds the feasible high-flow region because efficiency rises with flow rate while pumping power increases strongly with it, approaching cubic scaling in the turbulent branch; at the 0.5 watt ceiling the reported optimum is feasible and lies below the ceiling rather than exactly on it. The second case maximises the performance criterion. Each objective is minimised internally through its negative. For the constrained-efficiency case, the objective is given by Equation (7), with the penalty

coefficient set to 50 and the maximum residual constraint violation across the reported optima found to be zero to within the surrogate tolerance.

The two objectives answer complementary but distinct questions and are not of equal design standing. The constrained-efficiency objective is the primary design objective because it maximises the useful heat delivered per unit aperture under an explicit hydraulic budget and therefore ranks candidate fluids by the quantity a collector designer ultimately cares about. The performance criterion is retained as a secondary screening index in the sense of Webb [21]: it answers the narrower question of whether a nanoparticle additive repays its own friction penalty at equal flow, a standard figure of merit for heat-transfer enhancement, and it is reported here to characterise the additive rather than to select the collector's operating fluid. Following the reviewers' guidance, the criterion optimisation is presented throughout as this subordinate characterisation, and the physical comparison presented in Section 6.6 is the reference statement of why the criterion optimum, which trades absolute efficiency and useful gain for a higher relative index, must not be read as the best design in useful heat.

$$\text{minimise } F(x) = -\eta_{\text{th}}(x) + \lambda \cdot [\max(0, (W_p(x) - W_{\text{cap}})/W_{\text{cap}})]^2, \lambda = 50, g(x) = W_p(x) - W_{\text{cap}} \leq 0 \quad (7)$$

The pumping-power ceiling was set to 0.5 watts, an illustrative numerical constraint rather than a value derived from a specific pump specification, its influence being bounded by the sensitivity analysis over 0.25 to 1.0 watts (Table 2). Its influence was examined by repeating the grid-reference optimisation at ceilings of 0.25, 0.5 and 1.0 watts (Table 2). The selected carrier, distilled water, and the selected hybrid pair, Al₂O₃-Cu, are stable across all three ceilings; only the optimal flow rate and loading shift—from a moderate flow at the tighter ceilings to a higher flow at the loosest—so the qualitative design selection is insensitive to the exact ceiling. The pumping powers in Table 2 are those of the discrete-grid optima; the off-grid optimiser solutions reported in Sections 6.6 and 6.7 reach the same efficiency through a different feasible configuration—because efficiency is nearly composition-independent at a fixed operating point—and they therefore carry a different pumping power that is in every case below the ceiling. Feasibility is further enforced by confining the search to the sampled box, so the surrogate only interpolates, and by a single-phase condition. The 6 bar loop pressure is absolute; the saturation or degradation limit is fluid-specific, with the lowest limit being that of water, for which the saturation temperature at 6 bar is about 158 degrees Celsius. Using this most restrictive water limit as a uniform conservative cap, the bulk-outlet single-phase constraint is non-binding within the box at the studied operating points, although a fluid-specific single-phase, vapour-pressure and degradation check including local wall-temperature margins is required before physical sizing because the ethylene-glycol mixture and the synthetic oil have different boiling, flash-point and compatibility limits, with the largest predicted outlet temperature being 126 degrees Celsius; furthermore, the Reynolds envelope of the dataset, with a maximum of 45,840, is likewise non-binding, with the largest predicted value being 35,391. The dataset adopts a single-tube convention, passing the full loop flow through one representative tube, so the pumping power is the absorber-tube value $W_p = (\Delta P)(V_{\text{loop}})/\eta_p$ with $V_{\text{tube}} = V_{\text{loop}}/N_t$ and $N_t = 1$ by convention; a physical collector with N_t greater than one carries a per-tube flow V_{loop}/N_t and a different Reynolds number, pressure drop and pumping power, so the reported W_p , Re, T_o , Q_u and efficiency are dataset-convention values that must be re-solved for a specified tube count before they are used for sizing.

Table 2. Pumping-power-ceiling sensitivity (grid-reference efficiency optimum). The carrier and hybrid pair are stable; only flow rate and loading shift.

Ceiling W_{cap} (W)	Optimal η_{th}	Hybrid Pair	Base Fluid	w (%)	V (L/min)	W_p (W)
0.25	0.7616	Al ₂ O ₃ -Cu	Distilled water	2	3	0.233
0.5	0.7616	Al ₂ O ₃ -Cu	Distilled water	2	3	0.233
1.0	0.7674	Al ₂ O ₃ -Cu	Distilled water	0.5	5	0.943

The single-tube convention requires care for collector sizing, because the supplied dataset combines quantities defined at two different geometric levels. The in-tube hydraulics are evaluated for a single representative absorber tube of wetted area $A_s = \pi d_i L \approx 0.0452 \text{ m}^2$ carrying the full loop flow, so the reported Reynolds number, pressure drop and pumping power ($\text{Re} \approx 15,040$ and $W_p \approx 0.35 \text{ W}$ at the efficiency optimum) are single-tube quantities. The thermal efficiency, by contrast, is referred to a fixed collector-module aperture $A_c = 2.0 \text{ m}^2$ through $\eta_{\text{th}} = Q_u / (A_c I_s)$, so the reported $\eta_{\text{th}} = 0.760$ and $Q_u \approx 1832 \text{ W}$ are module-level quantities at that fixed aperture. These two levels are not mutually consistent for a physical collector: a real 2.0 m^2 module comprises many tubes, each carrying only a fraction V_{loop} / N_t of the loop flow and therefore operating at a different Reynolds number, pressure drop and pumping power than the single-tube values tabulated here. The reported η_{th} , Q_u , W_p and Reynolds number are accordingly dataset-convention values; they must not be read as a self-consistent operating point of any single physical collector and they cannot be transferred to a multi-tube collector by multiplying by a tube count. A physically consistent design result requires re-solving the reduced-order model for a specified aperture A_c , tube count N_t , total wetted area $N_t A_s$ and per-tube flow V_{loop} / N_t . This geometry-consistent re-solve is identified in the Limitations and Future Work as a precondition to collector sizing, and no claim is made here that $\eta_{\text{th}} = 0.760$ transfers directly to a multi-tube collector.

The 0.5 W ceiling is an illustrative numerical constraint rather than a value derived from a specific pump specification, and no pump derivation or reference is claimed for it; it applies to the single representative absorber tube the dataset encodes. Because it is a single-tube hydraulic quantity, it is not compared against the module-level useful gain so as not to mix the two geometric levels noted above. Its influence on the optimum is quantified directly in Table 2. Because efficiency rises with flow while pumping power grows almost cubically in the turbulent branch, it is the ceiling, and not the efficiency gradient, that fixes the feasible flow. Across a fourfold range of the ceiling (0.25 to 1.0 W) the selected carrier and hybrid pair are invariant and only the optimal flow rate and loading shift, so the qualitative design selection is insensitive to the exact value; the absolute flow rate and pumping power remain single-tube dataset-convention quantities that must be re-solved for a specified tube count.

2.3. Surrogate Evaluators

A reduced-order solve requires a mean-temperature iteration for both the nanofluid and its base-fluid baseline, which is too costly to repeat for the many thousands of evaluations a portfolio hyper-heuristic performs. Histogram-based gradient-boosted regression-tree surrogates were therefore trained on the dataset, one per response, and used as the cheap evaluators [39,40]. The dataset was split once into a training set of 46,267 runs and a held-out test set of 8165 runs with a fixed seed; the surrogates used by the optimiser are exactly those fitted on the training set, with no refitting on the test set, so the reported accuracies are leakage-free. The eight design inputs are low-cardinality grid factors, so a histogram split finder over the few distinct values per feature trains in seconds and predicts in microseconds. A tree ensemble is piecewise constant between sampled levels, so the

continuous optimisation effectively searches over neighbourhoods of the sampled grid; every reported optimum is consequently re-verified against the reduced-order simulator and the exact grid, as Section 6.7 details.

3. Sensitivity-Guided Search-Space Reduction

Because the design is a balanced full factorial, an exact variance decomposition is available in closed form without Monte Carlo sampling. The first-order variance share of each factor equals its analysis-of-variance main-effect eta-squared, the discrete counterpart of a first-order Sobol index on a dense grid [41,42]. A grouped partition projects each response onto a composition block, containing the hybrid pair, the base fluid, the weight fraction and the share, and an operating block, containing the flow rate, the inlet temperature, the irradiance and the ambient temperature, giving an encoding-free decomposition into composition, operating and interaction shares. The total effect of a block is the sum of its main effect and the interaction. A Morris elementary-effects screening [43] is computed from normalised single-factor steps; because elementary effects are finite differences whose meaning depends on the imposed category ordering, the Morris screening is reported only for the ordinal factors, the two categorical factors being handled by the encoding-free analysis of variance.

The grouped partition (Figure 1, Table 3) requires careful interpretation. For the performance criterion, the operating block carries only 4.8 percent of the variance as a main effect, but the composition-operating interaction carries 70.2 percent, so the operating block has a large total effect of 0.75 and the composition block a total effect of 0.95. It would therefore be incorrect to claim that fixing the operating point discards only 4.8 percent of the information, because the influence of composition depends strongly on the operating conditions through the interaction. The reduction is instead justified on a stronger and more directly relevant basis: the identity of the optimal hybrid pair, rather than of the complete composition, is invariant to the operating point for the performance criterion, whereas the criterion-optimal carrier varies with the operating point; for the constrained-efficiency objective, the carrier is invariant while the hybrid pair is nearly objective-degenerate. An exhaustive grid search of the performance-criterion optimum at each of the thirty-six sampled operating points returns the Al_2O_3 -Cu hybrid pair at every single point, and the constrained-efficiency optimum returns distilled water as the carrier at every point and the Al_2O_3 -Cu pair at thirty-one of the thirty-six. The interaction changes the magnitude of the criterion but not the argument that maximises it, so fixing the operating point changes the absolute objective values and may affect the carrier ranking, but it does not change the PEC-optimal hybrid pair, while for the efficiency objective the carrier remains fixed and the nanoparticle pair is nearly objective-degenerate, which is the property that a dimensionality reduction for design selection requires.

Table 3. Stability of the dominant design selections across the 36 sampled operating points. The PEC-optimal hybrid pair is invariant, while the PEC-optimal carrier varies; for constrained efficiency, the carrier is invariant and the hybrid-pair effect is nearly degenerate.

Objective	Optimal Hybrid Pair Across 36 Operating Points	Optimal Carrier Across 36 Operating Points
Performance criterion	Al_2O_3 -Cu: 36/36	water 9, EG-water 18, oil 9
Thermal efficiency	Al_2O_3 -Cu: 31/36, Graphene- TiO_2 : 5/36	distilled water: 36/36

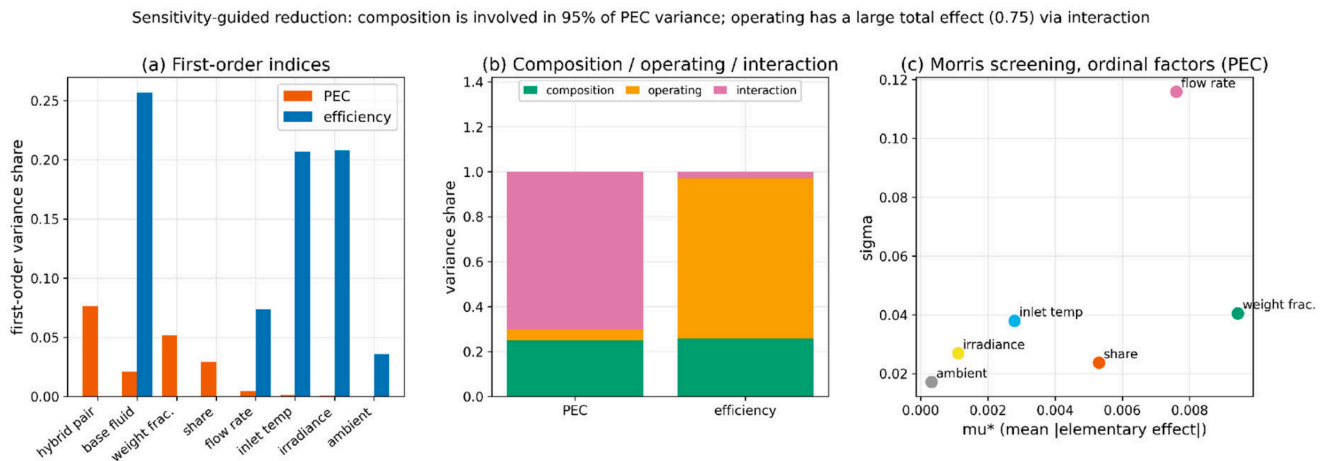


Figure 1. Sensitivity-guided reduction: (a) first-order variance shares; (b) grouped composition, operating and interaction partition; (c) Morris screening for the ordinal factors (PEC). Categorical factors are handled by the encoding-free analysis of variance.

With the reduction justified, the operating point is fixed and the design space is reduced from eight variables to the five decision variables of Section 2.2. The variance shares confirm the source attribution and inform the operators: within the fixed operating point, the performance criterion is governed by the hybrid pair, then the weight fraction and the component share, while the flow rate is nearly inert; the thermal efficiency at the same point is governed by the base fluid and the flow rate, with the nanoparticle variables nearly inert. These rankings are reflected in the problem-specific operators of Section 4. The fixed operating point is therefore a representative demonstration point: the robust design conclusion is the selection of the $\text{Al}_2\text{O}_3\text{-Cu}$ hybrid pair for the performance criterion, while the carrier ranking remains conditional on the operating point and the heat-transfer closure.

4. Hyper-Heuristic Methodology

The optimiser is a selection hyper-heuristic with two separated layers. A low-level layer provides search operators with different behaviours, and a high-level layer decides which operator to apply and whether to accept its result. The layers communicate only through objective values from the cheap surrogate, so the high-level controller is problem-independent and the same configuration is used for the benchmark suite and the collector application.

4.1. Low-Level Heuristic Portfolio

The portfolio contains twelve operators in three families. The trajectory family contains a simulated-annealing move with a cooling Gaussian step [44], a tabu move that returns the best non-tabu neighbour [45], a variable-neighbourhood shake that cycles four neighbourhood radii [46], and a large-neighbourhood destroy-and-repair move. The population family contains a differential-evolution rand-to-one operator, a differential-evolution best-to-one operator [47] and a particle-swarm move [48], all reading from a steady-state population of sixteen members. The problem-specific family contains five targeted operators—namely, three continuous perturbations of the flow rate, the weight fraction and the component share—together with two categorical swaps of the hybrid pair and the base fluid. Most operators generate one candidate at the cost of one function evaluation; the tabu operator additionally evaluates up to six sampled non-tabu neighbours per call to select its candidate, and every evaluation, including these, is charged to the

shared budget, so all methods are compared on an identical total number of function evaluations. The bandit state counts operator invocations while the global budget counts objective evaluations, so the tabu operator consumes the budget faster but is not otherwise penalised in the upper-confidence-bound score.

4.2. High-Level Selection and Acceptance

Heuristic selection is treated as a multi-armed bandit. Each operator holds a value estimate updated by a recency-weighted credit-assignment rule: an operator that produces a candidate improving the global best receives a unit reward and otherwise no reward, and its value is updated towards that reward by an exponential moving average with a learning rate of 0.30. Selection applies an upper-confidence-bound rule that maximises the sum of the value estimate and an exploration bonus proportional to the square root of the logarithm of the iteration count divided by the operator use count, with an exploration constant of 0.10, after a warm-up that tries each operator once [29,30]. This new-global-best reward is deliberately sparse; an improvement-magnitude variant of the reward, in which the credit is the relative improvement in the incumbent, was considered during development but is not reported as a controlled experiment and is not relied upon in the conclusions; the limited separation between the adaptive selector and random selection reported in Section 6 indicates that the credit rule is not the factor that governs performance on this suite. Move acceptance uses late-acceptance hill-climbing, which compares each candidate with the incumbent value recorded fifty iterations earlier and accepts it if it improves on either that value or the current incumbent [27]. The evaluation budget includes the population-initialisation evaluations, so the total budget is the sum of the initialisation and search evaluations and is identical for every method, including the standalone differential evolution and covariance-matrix-adaptation baselines [49]. For the constrained benchmark problems, the objective encodes a feasibility-first ranking: a feasible candidate returns its objective value, while an infeasible candidate returns a large constant plus its total constraint violation, so any feasible solution is preferred to any infeasible one, and infeasible solutions are ordered by violation. The incumbent, population-replacement and best-update comparisons in Algorithm A1, and in the differential-evolution and covariance-matrix-adaptation baselines, therefore implement this feasibility-first comparator through the single objective without a separate comparator routine, while the single smooth collector pumping-power bound is handled by the graded penalty of Section 2.2 and all reported collector optima are strictly feasible. The additive constant in the infeasible branch, 10^6 , exceeds the range of feasible objective values on both constrained benchmarks (welded-beam costs near 1.7 to 5 and pressure-vessel costs near 6000), so any feasible candidate is strictly preferred to any infeasible one and the single-objective encoding is equivalent to a feasibility-first comparator. The complete procedure is given as Algorithm A1 in Appendix B.

The concern that a sparse new-global-best reward can starve the bandit where progress is incremental is well founded, and two observations bound its practical impact here. First, an improvement-magnitude variant of the reward, in which the credit is the relative reduction in the incumbent, was considered during development; it is not reported as a controlled experiment, because the released statistical analysis does not include it, and it is not relied upon in the conclusions. Second, and more directly, the random-selection ablation of Section 6.5 shows that uniform random operator selection over the same portfolio and the same late-acceptance acceptance is not separated from the adaptive selector on this suite (unadjusted paired Wilcoxon $p = 0.615$; 1.0 after Holm correction). This indicates that, on this problem class, the credit-assignment rule is not the factor that governs end-of-run quality; it does not, however, bound what a different reward scheme could achieve on

other problem classes. The comparison-based alternative the reviewer raises, in which an operator is rewarded when its child improves on the population median—a scale-invariant quantile signal that stays informative when absolute gains are small [50]—is therefore proposed as the concrete next test rather than pre-judged, and is specified in Future Work; its aim is to establish whether the adaptive layer can be made to separate from random selection on more state-dependent landscapes.

5. Experimental Setup

The hyper-heuristic and every comparison method were validated on standard single-objective problems before the collector application. The benchmark set contains six classical continuous functions in ten dimensions—the sphere, Rastrigin, Ackley, Rosenbrock, Griewank and Schwefel functions—and two standard constrained engineering design problems—the welded-beam [51] and pressure-vessel [52] problems—with a feasibility-first constraint handling [53]. The welded-beam problem minimises a fabrication cost in four variables subject to seven constraints on shear stress, bending stress, buckling load, end deflection and geometry, with a known optimum near 1.7249; the hyper-heuristic reached a strictly feasible cost of 1.726. The pressure-vessel problem minimises a material cost in four variables subject to four constraints; it is solved here in its continuous form, in which the shell and head thicknesses are not restricted to multiples of 0.0625 inch, so the attainable cost is lower than the classical discrete optimum of about 6059. Both constrained problems use a feasibility-first ranking, in which any feasible solution is preferred to any infeasible one and infeasible solutions are ordered by their total constraint violation, so the comparison is not biased by infeasible candidates and the success criterion additionally requires feasibility; under this rule, the best feasible pressure-vessel cost was 5916. The collector application comprises the constrained-efficiency and performance-criterion objectives, each at two operating points, on the surrogates.

Every method was run for thirty independent trials with distinct seeds from one fixed master seed. The evaluation budget was 5000 function evaluations for the benchmarks and 700 for the lower-dimensional collector problems. For each method and problem, the solution quality across the thirty runs is summarised by its best, mean, median, standard deviation and worst values; reliability by the success rate against a problem-specific target threshold (Table 4) and the mean evaluations to reach it; and efficiency by the total evaluations and wall-clock time. Methods were compared with the Friedman test [54,55] across the problem blocks, with ranks assigned under a relative tie tolerance of 10^{-4} and a tie correction applied to the Friedman statistic so that the near-identical collector results, where every method reaches the same optimum, are tied rather than separated by numerical noise. Where the Friedman null was rejected, the Nemenyi post hoc test was applied with a critical-difference diagram, complemented by pairwise Wilcoxon signed-rank tests. The Wilcoxon tests dropped zero differences, were computed on within-problem ranks—so the comparison is invariant to the differing objective scales, using a normal approximation to the signed-rank null distribution with a continuity correction and a tie-adjusted variance—and were corrected by the Holm procedure [56] over the fifteen comparisons of the hyper-heuristic against the other methods. The pairwise Wilcoxon heatmap presented in Section 6.5 reports uncorrected p -values for description only; the formal multiplicity control is the Holm correction over these fifteen hyper-heuristic comparisons. The chi-square and normal tail probabilities were computed from incomplete-gamma and error-function series. The pipeline was implemented in Python 3.10 with numpy on a two-core sandbox, with the gradient-boosted trees and the statistical tests implemented from first principles so it is fully self-contained and reproducible.

Table 4. Success thresholds used for the reliability metric (success rate and evaluations to target).

Problem	Target (Success)	Problem	Target (Success)
Sphere	$f \leq 1 \times 10^{-6}$	Schwefel	$f \leq 1 \times 10^{-1}$
Rastrigin	$f \leq 1 \times 10^{-2}$	Welded beam	$f \leq 1.74$
Ackley	$f \leq 1 \times 10^{-3}$	Pressure vessel	$f \leq 6100$
Rosenbrock	$f \leq 1 \times 10^{-1}$	Collector (each)	objective $\geq 0.99 \times$ grid optimum
Griewank	$f \leq 1 \times 10^{-2}$		

Beyond the measured evaluation counts and wall-clock times, the methods differ in asymptotic per-invocation cost and memory, summarised in Table A5. The trajectory operators (simulated annealing, tabu, variable-neighbourhood and large-neighbourhood moves) cost $O(d)$ work per candidate and $O(d)$ memory, the tabu move adding a bounded $O(6d)$ neighbour scan and an $O(\text{tabu-tenure})$ list. The population methods (differential-evolution variants and particle swarm) cost $O(d)$ per candidate with $O(NP d)$ memory for the steady-state population. The covariance-matrix-adaptation strategy is the asymptotic outlier: it carries an $O(d^2)$ covariance update per sampled candidate and an $O(d^3)$ eigen-decomposition amortised over each generation, with $O(d^2)$ memory, so its per-invocation overhead grows fastest with dimension even though its evaluation count is competitive. The hyper-heuristic adds only an $O(|H|)$ upper-confidence-bound selection step and $O(1)$ late-acceptance test on top of the chosen operator, with $O(NP d + L + |H|)$ memory for the population, the length- L acceptance history and the twelve-armed bandit state; its controller overhead is therefore negligible against the surrogate evaluation, which is the dominant cost and is identical for every method. The asymptotic overhead therefore grows fastest for covariance-matrix adaptation with dimension; at the low dimensions studied here, however, its measured wall-clock time is not the largest, and the hyper-heuristic's controller bookkeeping is in every case negligible against the shared surrogate-evaluation cost.

6. Results and Discussion

6.1. Surrogate Accuracy and Blocked Validation

On the held-out test set of 8165 runs, the efficiency, pumping-power, Reynolds-number and outlet-temperature surrogates reproduce the simulator with high fidelity, while the performance-criterion surrogate is retained for its top-region ranking rather than its global coefficient of determination (Table 5). Efficiency, Reynolds number and outlet temperature reach coefficients of determination of 0.9998 or better with mean absolute percentage errors at or below 0.31 percent, and the pumping power—emulated in logarithmic space—reaches 0.99995 with a mean absolute percentage error of 0.76 percent; its root-mean-square error is 0.0149 in log space and 0.0073 watts in physical units. The performance criterion is the only difficult response, with a held-out coefficient of determination of 0.46 but a mean absolute percentage error of 0.08 percent. Because the optimisation-relevant differences in the criterion are themselves small, ranking accuracy near the optimum matters more than the global coefficient of determination, and the criterion was therefore validated in its top region: the Spearman rank correlation between predicted and true values is 0.97 overall and 0.89 within the top five percent, and the mean absolute error within the top one percent is 0.002. As a stricter check, the surrogate optimum evaluated under the reference implementation ranks third among the 1512 grid configurations at the operating point, and the surrogate and reference-implementation top-twenty grid designs coincide exactly—a recall at twenty of one—so the surrogate ranks the high-criterion designs reliably and is accurate where the optimiser operates (Table A1).

Table 5. Held-out accuracy of the surrogate evaluators (test set $n = 8165$).

Response	R ² (Held-Out)	RMSE	MAPE (%)
Thermal efficiency	0.99979	0.0023	0.242
Pumping power	0.99995	0.0149 (log); 0.0073 W	0.757
Reynolds number	1.00000	7.13	0.305
Outlet temperature	1.00000	0.045	0.059
Performance criterion	0.46187	0.0059	0.080

These two statistics are consistent because they measure different things on the same 1512-point grid. Recall at twenty is the size of the intersection of the surrogate's top-twenty set with the reference implementation's top-twenty set divided by twenty, so a value of one states only that the two top-twenty sets contain the same designs, not that they are ordered identically. The rank of the surrogate optimum is instead the position, under the reference ordering, of the single design the surrogate ranks first. A recall of one is thus fully compatible with the surrogate's first-ranked design sitting third under the reference ordering: that design lies inside the common top-twenty set, indeed within its top three, but it is not the reference's first element. The surrogate does not select a point outside the exact top-twenty; it selects the third-best member of a top-twenty set on which the surrogate and the reference model agree as a set. The wording of Section 6.1 and the metric labels in Table A1 have been revised to state these definitions explicitly and remove the apparent inconsistency.

A random held-out split over a dense factorial grid can overstate reliability because neighbouring grid points remain in training. The surrogates were therefore also tested under blocked leave-one-group-out validation, holding out an entire factor level (Table 6). Holding out a factor that is inert for a response leaves its coefficient of determination at unity, as for the hybrid pair on efficiency, whereas holding out a driving factor drives it strongly negative—to minus 0.71 when the synthetic-oil carrier is withheld from efficiency and to minus 29 when the Al₂O₃-Cu pair is withheld from the performance criterion. The surrogate is therefore an accurate interpolator among the sampled levels and not an extrapolator to unseen carriers or particle pairs, which bounds the search to the sampled envelope and is consistent with the variance attribution. Leaving out an interior continuous level tests interpolation in the optimisation variables: the efficiency surrogate interpolates well, with held-out coefficients of determination of 1.000 for a withheld weight-fraction level and 0.983 for a withheld flow-rate level, whereas the performance-criterion surrogate interpolates only weakly, 0.48 and 0.08 respectively, which is why its off-grid optima are verified against the grid and the reconstructed model rather than trusted directly.

Table 6. Blocked leave-one-group-out validation. Inert factors leave R² near unity; driving factors give strongly negative R², confirming interpolation-only validity.

Held-Out Group	Response	Held-Out R ² Per Level
Each base fluid	Thermal efficiency	water 0.874, EG-water 0.864, oil −0.705
Each hybrid pair	Thermal efficiency	Al ₂ O ₃ -Cu 1, MWCNT-Fe ₃ O ₄ 1, Graphene-TiO ₂ 1
Each hybrid pair	Performance criterion	Al ₂ O ₃ -Cu −29.119, MWCNT-Fe ₃ O ₄ −0.101, Graphene-TiO ₂ 0.532

6.2. Sensitivity-Guided Reduction

Figure 1 summarises the variance decomposition that underlies the reduction in Section 3: the first-order indices, the grouped composition, operating and interaction shares, and the Morris screening restricted to the ordinal factors. The figure makes explicit that composition is involved in 95 percent of the performance-criterion variance while the operating block

retains a large total effect through the interaction, which is why the reduction is justified by the invariance of the PEC-optimal hybrid-pair identity rather than by neglecting operation.

6.3. Benchmark Validation

Across the twelve-problem set, under the relative tie tolerance, which prevents the near-identical collector results from being separated by numerical noise, the average ranks place standalone differential evolution and the single variable-neighbourhood operator jointly first at 3.83, the random-selection ablation third at 4.50, the covariance-matrix-adaptation strategy fourth at 4.58 and the adaptive hyper-heuristic fifth at 4.67, with the remaining operators trailing from 5.2 to 15.9 (Table 7), and the Friedman test rejects equal performance with overwhelming confidence (Table 8). Two honest conclusions follow. First, the adaptive hyper-heuristic lies within the leading group that the critical-difference analysis cannot separate, and it significantly outperforms six individual operators (particle swarm and the five problem-specific perturbation and swap moves) under Holm-corrected Wilcoxon tests, but it does not uniformly dominate the best individual operator: under the tie tolerance the variable-neighbourhood operator attains a clearly better rank, and even the random-selection ablation ranks marginally ahead of the adaptive selector on the full set. Its value is therefore robustness: on this suite it avoided the poor worst-case outcomes that committing to a single weak operator can produce, rather than a uniform win. Second, no single method dominates across landscapes: the hyper-heuristic obtains lower mean objective values than differential evolution on the highly multimodal Rastrigin and Schwefel functions in these runs, where its mean values of 10.2 and 612 are better than the 40.8 and 1391 of differential evolution, whereas the covariance-matrix-adaptation strategy ranks best on the benchmark suite overall, particularly the smooth sphere and Ackley functions (Table A2). None of the leading methods reaches the stringent success threshold on these multimodal cases within the budget, so this is a descriptive performance difference rather than a solved-problem success, and it reflects the expected no-free-lunch trade-off. Figure 2 shows the convergence of the best feasible objective against the evaluation budget for a smooth benchmark, a multimodal benchmark and the collector efficiency problem.

Table 7. Average Friedman ranks over the twelve problems with a 10^{-4} relative tie tolerance (lower is better).

Method	Avg. Rank	Method	Avg. Rank
DE	3.833	DE/best/1	7.625
VNS	3.833	Tabu	7.708
HH-Random	4.500	perturb-V	12.458
CMA-ES	4.583	perturb-w	12.583
HH-UCB	4.667	swap-fluid	12.917
DE/rand/1	5.208	perturb-share	13.000
SA	6.208	swap-hybrid	14.167
LNS	6.833	PSO	15.875

Table 8. Friedman test and Nemenyi critical difference ($\alpha = 0.05$) by problem block. The four-collector block has too few blocks for reliable asymptotic Friedman inference and is reported descriptively. Here n.a. denotes not applicable.

Problem Block	Friedman χ^2	df	p-Value	Critical Difference
All twelve problems	147.62	15	7.2×10^{-24}	6.66
Eight benchmarks	99.84	15	1.4×10^{-14}	8.16
Four collector problems	$n = 4$ (descriptive)	n.a.	n.a.	n.a.

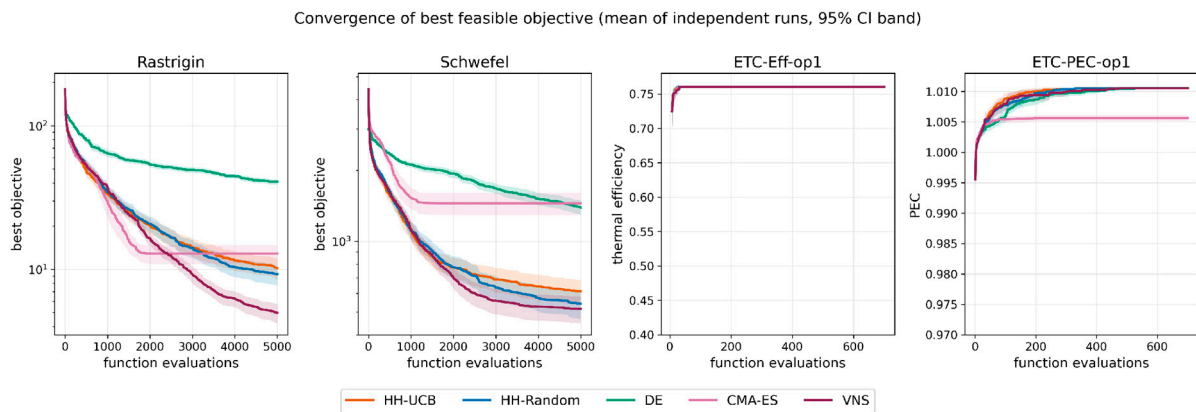


Figure 2. Convergence of the best feasible objective against function evaluations (mean of independent runs, shaded 95 percent confidence band). The collector panels plot the thermal efficiency and the performance criterion of the best feasible solution, not the penalised internal objective.

6.4. Heuristic-Selection Dynamics

Figures 3 and 4 illustrate the operator selection and credit over one genuine search, the ten-dimensional Rastrigin function, shown as a representative example rather than as behaviour claimed to be identical across all landscapes, instead of the trivial collector problem on which every operator is exhausted during the warm-up. The controller maintains substantial diversity throughout: the largest pooled overall invocation fraction across the five runs is about 10.1 percent, for the differential-evolution best-to-one operator, with the remaining operators between roughly eight and nine percent; the sliding windows of Figure 3 can exceed these pooled fractions. The credit, however, is concentrated: the population-based particle-swarm and differential-evolution best-to-one operators account for the majority of new global bests—358 and 291 respectively on Rastrigin (pooled over five runs)—against fewer than one hundred for any other operator, so the controller mildly but correctly favours the operators that drive global progress while the exploration bonus keeps the full portfolio in contention. This maintained diversity is the mechanism by which, on the problems tested, the controller reduced its exposure to any single operator's worst-case behaviour.

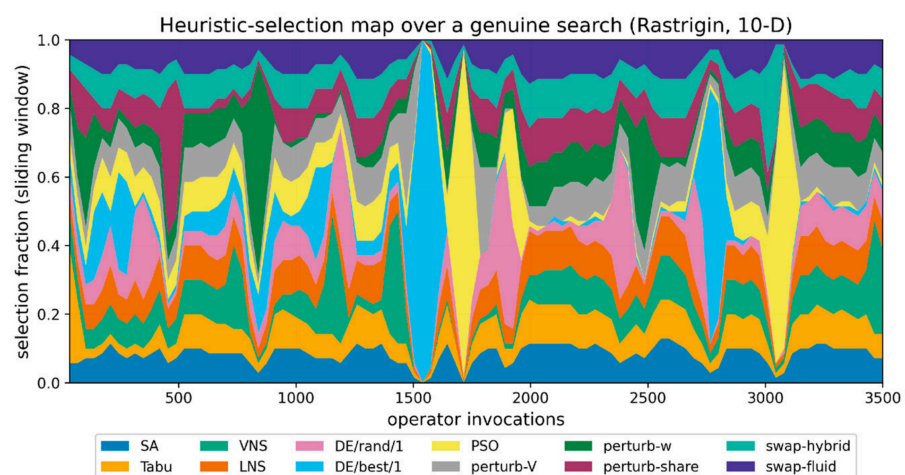


Figure 3. Heuristic-selection map over a genuine search (Rastrigin, ten dimensions): the windowed fraction of invocations of each operator. The controller maintains diversity, with the largest pooled overall invocation fraction about 10.1 percent (the differential-evolution best-to-one operator), and individual sliding windows can exceed this. The horizontal axis indexes operator invocations, not function evaluations; the tabu operator may consume up to six evaluations per invocation.

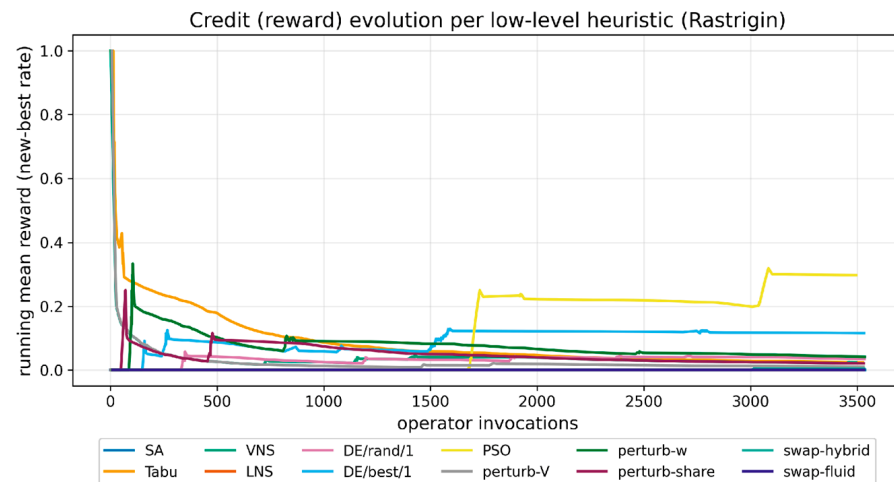


Figure 4. Credit (new-global-best rate) per operator on Rastrigin. The population-based operators accrue most of the credit.

6.5. Statistical Comparison and the Random-Selection Ablation

The critical-difference diagram of Figure 5 and the pairwise Wilcoxon heatmap of Figure 6 summarise the comparison, and the radar chart of Figure 7 profiles the leading methods. The best practical ranks are obtained by differential evolution, variable-neighbourhood search, the covariance-matrix-adaptation strategy, the random-selection ablation and the adaptive hyper-heuristic; because the Nemenyi critical difference is wide, however, the statistically non-separated set is broader than these five methods, so all superiority claims are taken only from the Holm-corrected pairwise Wilcoxon tests. The direct comparison of the adaptive selector against its random-selection ablation returns an unadjusted paired Wilcoxon probability of 0.615 across the twelve problems (1.0 after Holm correction), computed on within-problem ranks so the comparison is invariant to the differing objective scales; thus, the adaptive credit assignment does not produce a statistically significant advantage over uniform random selection on this suite. On the four collector problems, the adaptive selector attains a slightly better average rank than random selection, but the four-block collector subset is descriptive only and no inferential conclusion is drawn from it, so this is reported as a small, non-significant advantage rather than a demonstrated benefit. These findings are consistent with the established observation that uniform random selection over a well-designed portfolio with an effective acceptance criterion is a strong baseline, and they locate the value of the framework in robustness rather than in adaptivity.

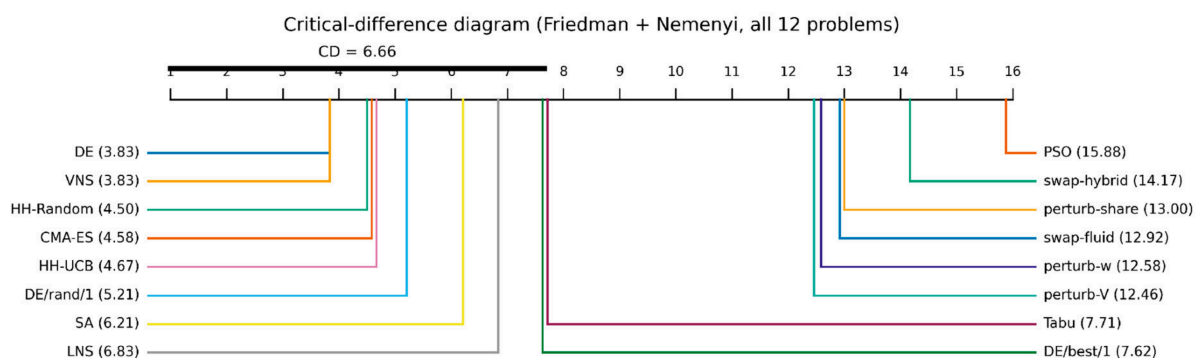


Figure 5. Critical-difference diagram (Friedman with Nemenyi post hoc, all twelve problems). Methods joined within the critical difference are not significantly different.

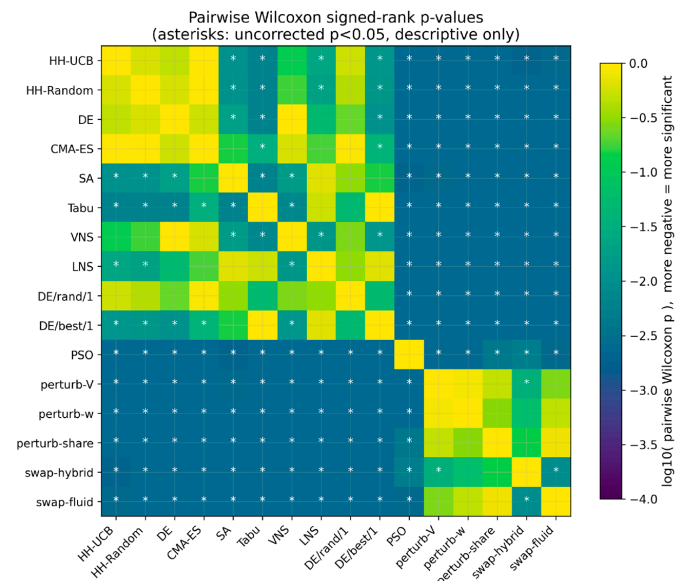


Figure 6. Pairwise Wilcoxon signed-rank p -values across the twelve problems on a base-ten logarithmic scale; more negative values are more significant; the asterisks denote uncorrected $p < 0.05$ and are descriptive only, formal claims using the Holm-corrected hyper-heuristic comparisons.

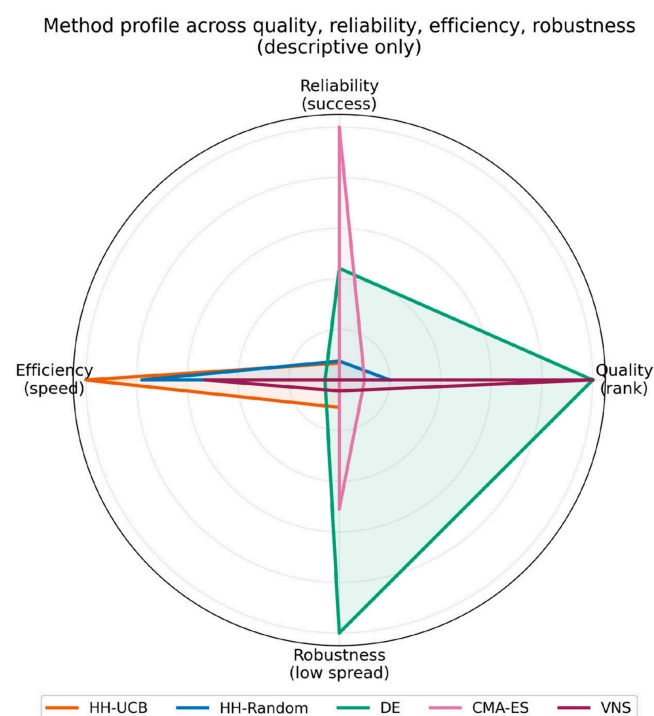


Figure 7. Method profiles across solution quality, reliability, efficiency and robustness; each axis normalised across the compared methods. This chart is descriptive only; inferential conclusions are drawn from the Friedman-, Nemenyi- and Holm-corrected Wilcoxon analyses.

The practical meaning of this parity should be stated plainly, because it is easily misread as a negative result. That an adaptive selector does not beat uniform random selection over the same portfolio does not make the framework pointless; it locates its value, and the advantages that follow do not depend on the selector winning. First, the portfolio empirically reduced exposure to the weakest operators on this test suite: several operators used alone are poor on some landscapes (particle swarm ranks last overall at 15.9, and the problem-specific perturbation and swap moves rank near the bottom, Table 7), whereas the portfolio,

whether selected adaptively or at random, was not dragged to those poor outcomes on the problems tested, because the maintained operator diversity (the largest pooled overall invocation fraction being about 10.1 percent, Figure 3, though individual windows exceed this) keeps a productive operator in play. This is an empirical observation on the present suite and not a guarantee: none of the methods, the portfolio included, reached the success threshold on the Rastrigin or Schwefel functions (Table A2). Second, the same fixed configuration was applied without re-tuning across the benchmark functions, the constrained mechanical-design problems and the collector, which is the property an engineering-design pipeline needs when the landscape is not known in advance. The result therefore reframes the contribution from a better operator selector to a robust, fixed-configuration portfolio optimiser that, on the problems tested, reduced exposure to the weakest operators.

6.6. Collector Application

On the collector problems, the leading methods and most individual operators converge to practically indistinguishable optima under the adopted tie tolerance, which is expected because the sensitivity-guided reduction leaves a low-dimensional and well-behaved landscape on which the framework is reliable rather than discriminating. The adaptive hyper-heuristic attains a surrogate thermal efficiency of 0.760 at the primary operating point under the 0.5 watt ceiling and 0.736 at the second point, and a surrogate performance criterion of 1.011 at the primary point and 1.009 at the second point—in every case, feasible across all thirty runs. The candidate criterion optimum is an Al_2O_3 -Cu suspension in the ethylene-glycol and water carrier at the maximum loading of three percent, the minimum share of twenty-five percent, and a low flow rate, which is the composition-dominated corner the screening predicts. The constrained-efficiency optimum favours distilled water at a high flow rate that remains within, but not exactly on, the ceiling; this is the best feasible point within the sampled-grid and surrogate representation, and a continuous optimum of the reconstructed model may move closer to the active pumping-power boundary and is not claimed here. At a fixed operating point, the efficiency is almost independent of the nanoparticle pair, so the optimiser can reach the same efficiency through different pairs, and the pair it reports for the efficiency objective need not match the grid-optimal pair, even though the objective value matches to within two parts in a thousand; this is a faithful reflection of the variance structure rather than a defect.

Because the performance criterion is a relative index, the criterion optimum is not the best design in absolute terms. Table 9 compares the efficiency optimum and the criterion optimum at the primary operating point through the reduced-order simulator: the criterion optimum delivers a higher relative index but a lower absolute efficiency and useful gain than the efficiency optimum, so the two answers genuinely differ and the criterion optimum must not be read as the best collector design in useful heat.

Table 9. Physical comparison of the efficiency optimum and the performance-criterion optimum at the primary operating point (reduced-order simulator). The criterion optimum trades absolute efficiency and useful gain for a higher relative index.

Quantity	Efficiency Optimum	Criterion Optimum
Thermal efficiency η_{th}	0.7635	0.7326
Useful gain Q_u (W)	1832.3	1758.2
Pumping power W_p (W)	0.3457	0.156
Reynolds number	15,040	3125
Nusselt number	92.05	30.86
Performance criterion	0.998	1.0094

6.7. Grid and Reconstructed-Model Verification

The optimiser solutions were checked against two internal consistency benchmarks—an exhaustive enumeration of the discrete grid at the fixed operating point and a reconstructed implementation of the same reduced-order model (Table 10). Because the tree surrogate is piecewise constant between sampled levels, this off-grid verification is essential: it confirms that the surrogate-reported point corresponds to a high-performing reduced-order-model design close to the discrete-grid reference, and it should not be read as proof of a global continuous optimum unless the reconstructed model is itself continuously optimised. The surrogate optimiser result lies within 0.05 to 0.40 percent of the exact grid-reference value across the four problems, and the reference implementation evaluated at the optimiser solution agrees with the surrogate to within about 0.29 percent. The largest gap, 0.40 percent, occurs for the criterion at the second operating point and reflects the rare-tail difficulty of the criterion surrogate rather than a failure of the search. The constraint residuals are reported in the same table: the pumping power at the efficiency optima is well within the ceiling, and the maximum constraint violation is zero to the surrogate tolerance. All quantities in Tables 9 and 10 use the single-representative-tube dataset convention and are not direct multi-tube collector sizing values. Figure 8 shows the criterion landscape over the two influential continuous variables with the search trajectory overlaid.

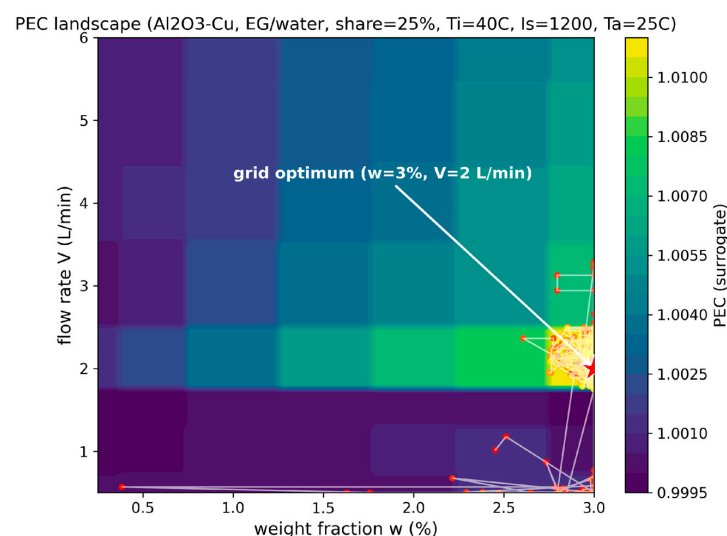


Figure 8. Performance-criterion landscape over the weight fraction and flow rate for $\text{Al}_2\text{O}_3\text{-Cu}$ in EG-water at share 25 percent and the primary operating point (T_i 40 C, I_s 1200 W/m^2 , T_a 25 C), with the search trajectory and the grid optimum overlaid.

Table 10. Collector optima verified against the exhaustive grid and the reference implementation, with pumping-power residuals. Grid-reference configurations: efficiency, $\text{Al}_2\text{O}_3\text{-Cu}$ in water at w 2–3% and V 3–4 L/min; criterion, $\text{Al}_2\text{O}_3\text{-Cu}$ in EG-water at w 3%, share 25%, V 1–2 L/min. The reconstructed-model value is an evaluation of the hyper-heuristic solution, not a separately optimised continuous reference. In the final column, n.a. denotes not applicable, no pumping-power ceiling being imposed on the performance-criterion objective.

Problem	Surrogate Value at HH Solution	Reconstructed-Model Value at HH Solution	Grid-Reference	Gap to Grid Ref. (%)	W_p (W)/Cap
Efficiency, point 1	0.76033	0.75964	0.7616	0.17	0.3457/0.50
Efficiency, point 2	0.73636	0.73529	0.7367	0.05	0.3281/0.50
Criterion, point 1	1.01053	1.00992	1.011	0.05	0.156/n.a.
Criterion, point 2	1.00878	1.01172	1.0128	0.40	0.0163/n.a.

6.8. Design Insights and Model-Form Robustness

The optimisation operationalises the two-stage design rule that the variance attribution implies: the operating point and the carrier set the absolute thermal and hydraulic magnitudes, so the constrained-efficiency objective selects distilled water and the highest flow rate compatible with the ceiling, while the nanoparticle pair and loading set the relative thermo-hydraulic index, so the criterion objective selects the Al_2O_3 -Cu pair at maximum loading. The robustness of these selections to the heat-transfer closure was tested by recomputing the criterion optimum with the Dittus–Boelter correlation in place of the Gnielinski correlation, used only as a stress test of closure sensitivity rather than as a physically preferred correlation at the near-transitional Reynolds number of about 3100. The optimal hybrid pair, Al_2O_3 -Cu, and the optimal loading are unchanged under both closures, but the optimal carrier shifts from the ethylene-glycol and water mixture under the Gnielinski closure to the synthetic oil under the Dittus–Boelter closure. The dominant composition driver is therefore robust to the closure, whereas the carrier ranking is not, which is consistent with the interpretation of the criterion as a relative index whose carrier ordering depends on the friction baseline. A caveat applies to this test. The criterion optimum sits near the transitional regime, at a Reynolds number of about 3100, where the Dittus–Boelter correlation is itself unreliable, so the carrier shift should be read as an indication of closure sensitivity rather than as a definitive alternative optimum, and a continuous transitional blending of the laminar and turbulent closures across the band 2300 to 4000 was applied as a direct check, and the optimal hybrid pair, carrier and loading were unchanged (Al_2O_3 -Cu in the ethylene-glycol and water mixture at three percent), with only the criterion magnitude shifting, so the design selection is robust to the transitional treatment even though the absolute index is not. The blending weight increases linearly across the band, $\omega(\text{Re}) = 0$ for Re at or below 2300, $\omega(\text{Re}) = (\text{Re} - 2300)/(4000 - 2300)$ for $2300 < \text{Re} < 4000$ and $\omega(\text{Re}) = 1$ for Re at or above 4000, applied to both the Nusselt number, $\text{Nu} = (1 - \omega) \text{Nu}_{\text{lam}} + \omega \text{Nu}_{\text{turb}}$, and the friction factor, $f = (1 - \omega) f_{\text{lam}} + \omega f_{\text{turb}}$. This linear blend is an uncertainty probe rather than a validated transitional-flow correlation. This closure sensitivity, together with the model-form uncertainty of the nanofluid property closures, means that the exact ranking of close carriers should be regarded as conditional on the closures rather than as a robust physical conclusion.

Because the criterion optimum is reported at the three-percent upper bound of the sampled loading, its colloidal realisability deserves explicit comment, and the concentration basis must be stated carefully. The decision variable w is a mass (weight) fraction, whereas the hybrid-nanofluid stability literature commonly reports volume fractions: the two-step Al_2O_3 -Cu system of Suresh et al. [6], for example, was characterised at roughly 0.1 to 2.0 vol%. Under the densities of the present dataset, 3 wt% of the Al_2O_3 -Cu solids corresponds to only about 0.5 vol%, so the model-optimal loading is not necessarily beyond the colloidal range explored experimentally, and a nominal ‘1 to 2 percent’ cap cannot be applied directly to w without first converting between the mass and volume bases. Two qualifications remain. First, stability depends strongly on the carrier, the surfactant and pH control, and the preparation protocol [5,6], so agreement in nominal concentration does not by itself establish that a 3 wt% Al_2O_3 -Cu suspension in the ethylene-glycol and water carrier is stable, and no such physical claim is made here. Second, the appropriate way to impose a stability limit in this formulation is as an explicit volume-fraction constraint applied after a wt-to-vol% conversion, together with a preliminary experimental stability screen (zeta potential, sedimentation time and temperature-dependent viscosity), which is placed first in the revised Future Work. The three-percent point is therefore a model-optimal loading within the sampled envelope and a validation target, not a validated or recommended deployment fluid.

Although the study is model-based, the reduced-order predictions can be placed against the experimental literature for external plausibility. The modelled evacuated-tube efficiencies at the studied operating points, of roughly 0.73 to 0.76, fall within the range reported experimentally for evacuated-tube collectors. Water-charged units report peak thermal efficiencies of roughly 0.41 to 0.69 across tube configurations, from the single U-tube to the helical design [57]; units charged with mono and hybrid nanofluids at comparable loadings report roughly 0.60 for an Al_2O_3 suspension, 0.69 for a hybrid MWCNT- Al_2O_3 suspension and 0.735 for MWCNT, each near 0.5 percent [58], rising further at higher particle loading [14–16]. The modelled band therefore sits at the upper edge of the water range and within the nanofluid range, and a single common experimental value cannot be quoted because the cited studies differ in the reported quantity, operating point and reporting convention—one of them being a finite-volume numerical rather than an experimental study; the direction of the modelled effect, the additive raising the in-tube convective coefficient and the useful gain relative to the base fluid, reproduces the qualitative finding common to those experimental studies. This is a directional, order-of-magnitude consistency check and not a calibration: the single-tube convention, the size-independent closures and the idealised property models preclude a quantitative match to any specific apparatus, so a formal validation against controlled experimental data on the same fluid and geometry remains the primary item of the experimental Future Work.

Two aspects of generality should be distinguished. Algorithmic generality across optimisation problems is already exercised by the twelve-problem benchmark suite, in which the same fixed hyper-heuristic configuration is applied without re-tuning to six classical continuous benchmark functions and two constrained mechanical-design problems, so the method is characterised on landscapes well beyond the collector. Physical generality across collector configurations is more limited, because only the evacuated-tube reduced-order model was exercised. The framework is nevertheless collector-agnostic by construction: the high-level controller communicates with the physics only through scalar objective values, so transferring it to a flat-plate, U-tube, heat-pipe or direct-absorption collector, or to a different thermal-design problem such as a compact heat exchanger or a phase-change store, requires substituting the reduced-order evaluator, adapting the constraint handling and redesigning the five problem-specific operators (which target the flow rate, loading, share, hybrid pair and carrier of this collector), while the selection and acceptance layers are reused unchanged. Demonstrating this portability on a second collector geometry is identified in Future Work.

7. Conclusions

A variance-based screening of a 54,432-run full-factorial dataset of a hybrid-nanofluid evacuated-tube collector was redeveloped into a computational optimisation study built on a selection hyper-heuristic. The screening was retained as a dimensionality-reduction front end: although operating variables carry a large total effect through their 70.2 percent interaction with composition, the performance-criterion optimum selects the same hybrid pair, Al_2O_3 -Cu, across all thirty-six sampled operating points, and the constrained-efficiency optimum selects the same carrier—distilled water—at every point with the hybrid pair nearly objective-degenerate, which justifies fixing the operating point and reducing the design space from eight variables to five. Histogram-based gradient-boosted surrogates emulated the responses within the grid with near-unity held-out coefficients of determination for efficiency, pumping power, Reynolds number and outlet temperature, while the performance-criterion surrogate had a low global coefficient of determination but acceptable top-region ranking accuracy, and were shown by blocked leave-one-group-out tests to interpolate among sampled levels but not to extrapolate. A selection hyper-heuristic,

pairing an upper-confidence-bound credit-assignment selector with late-acceptance hill-climbing over twelve operators, was implemented, pre-validated on six classical functions and two constrained engineering problems, and compared against its constituent operators, a random-selection ablation and strong standalone metaheuristics with a Friedman-, Nemenyi- and Holm-corrected Wilcoxon protocol.

The honest quantified findings are as follows. The hyper-heuristic ranked within the leading statistically indistinguishable group together with differential evolution, variable-neighbourhood search, covariance-matrix adaptation and the random-selection ablation, and it significantly outperformed six individual operators (particle swarm and the five problem-specific moves) under Holm-corrected Wilcoxon tests, but under a tie tolerance, the single variable-neighbourhood operator attained a marginally better rank, so the hyper-heuristic does not uniformly dominate the best individual operator; its value is robustness across landscapes and, on the problems tested, reduced exposure to any single operator's worst-case behaviour. The adaptive credit assignment did not produce a statistically significant improvement over uniform random selection on the benchmark suite, with an unadjusted paired Wilcoxon probability of 0.615 (1.0 after Holm correction), and the four-problem collector subset is reported descriptively only, so no inferential conclusion is drawn from that subset. The hyper-heuristic reliably attained the discrete-grid reference optimum of the collector problem, verified to within 0.40 percent of an exhaustive grid enumeration and within 0.29 percent of a reconstructed reference implementation, with full feasibility. The candidate criterion optimum was an Al_2O_3 -Cu suspension in the ethylene-glycol and water carrier at three percent loading, and the constrained-efficiency optimum reached a surrogate efficiency of 0.760, cross-checked against a grid-reference value of 0.762 and a reconstructed-model value of 0.760.

A direct consequence of applying feasibility-first ranking together with the relative tie tolerance is that the adaptive hyper-heuristic is no longer the best-ranked method. Its average Friedman rank is 4.67, placing it fifth, behind standalone differential evolution and variable-neighbourhood search (both 3.83), the random-selection ablation (4.50), and covariance-matrix adaptation (4.58). This result is reported explicitly and prevents the method from being overclaimed as superior to the best individual operator or to random selection over the same portfolio. The defensible contribution is therefore not algorithmic dominance, but the construction of an integrated, reproducible optimisation framework that combines sensitivity-guided reduction, surrogate-assisted evaluation, feasibility-aware benchmarking and internal consistency verification of the collector optima against an exhaustive grid enumeration and a reconstructed implementation of the same reduced-order model. Under this interpretation, the adaptive hyper-heuristic belongs to the leading statistically indistinguishable group, while the internally verified collector design, under the supplied reduced-order model and single-tube dataset convention, remains conditional on the property closures, grid envelope and reduced-order collector convention. In practical terms, the value of the portfolio is operational rather than adaptive: one fixed-configuration optimiser is applied without per-problem operator selection or hand-tuning across the benchmark, constrained-engineering, and collector problems, and by keeping complementary operators in play it bounds the exposure to any single operator's worst-case behaviour, which is the property a practitioner facing a new but related design problem most wants from an off-the-shelf solver.

7.1. Limitations

In summary, the principal limitations are as follows: (i) dependence on the surrogate model, which interpolates within the sampled grid but does not extrapolate, so every off-grid optimum is verified against the exact grid and the reconstructed reduced-order

model; (ii) sensitivity of the absolute results and of the close-carrier ranking to the heat-transfer correlations, particularly across the laminar-turbulent transition; (iii) the absence of experimental validation, the study being conditional on the supplied reduced-order model and its property closures; (iv) the single-representative-tube dataset convention, under which absolute efficiency, pumping power and useful gain must be re-solved for a specified tube count before collector sizing; (v) the size-independent property closures, which render the particle diameter inert; and (vi) the limited separation of the adaptive selection layer from random selection over the same portfolio. Each is expanded below.

Several limitations bound these conclusions and are stated plainly. The advantage of the adaptive layer over random selection is limited, consistent with the literature on strong random baselines. The dataset is simulation-based, the conductivity and viscosity closures are size-independent so the particle diameter is inert, and the Brinkman- and Maxwell-type closures are approximations for the high-aspect-ratio carbon and graphene systems. The collector passes the full loop flow through one representative tube rather than partitioning it among several, so the absolute efficiency, outlet temperature and useful gain are dataset-convention values, and the reported efficiency optimum should not be used for collector sizing until a geometry-consistent re-solve for a specified tube count with per-tube flow is performed. The second operating point at 60 degrees Celsius inlet is a secondary operating-point check rather than a full robustness study, which would vary the irradiance, ambient temperature and hydraulic conditions systematically.

7.2. Future Work

Future work will reconcile the wetted and aperture areas for a specified tube count; validate the candidate fluids experimentally for dispersion stability and compatibility, including the zeta potential, sedimentation time, temperature-dependent and possibly non-Newtonian viscosity, conductivity, pH stability and loop-material compatibility; calibrate the property and heat-transfer closures including the transitional band against experimental data; extend the surrogate to size-dependent property models and geometric variables; and investigate richer credit-assignment schemes such as rank-based bandits and choice functions to determine whether the adaptive layer can be made to separate from random selection in this class of problem.

Author Contributions: Conceptualization, M.A.K. and F.A.; methodology, M.A.K.; software, M.A.K.; validation, M.A.K. and F.A.; formal analysis, M.A.K.; investigation, M.A.K. and F.A.; resources, F.A.; data curation, M.A.K.; writing of the original draft, M.A.K.; writing, review and editing, M.A.K. and F.A.; visualisation, M.A.K.; supervision, F.A.; project administration, F.A.; funding acquisition, F.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the University of Bisha: the Deanship of Scientific Research at the University of Bisha through the Fast-Track Research Support Program, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The full-factorial dataset, the reduced-order collector simulator, the trained surrogate models, the selection-hyper-heuristic and baseline implementations, and the sensitivity and statistical-analysis scripts are openly available in a Zenodo repository at <https://doi.org/10.5281/zenodo.21526609> (DOI: 10.5281/zenodo.21526609, version 1.0.3). The source code is released under the MIT License, and the dataset, trained models and figures under the Creative Commons Attribution 4.0 International (CC BY 4.0) License. Exact bit-for-bit reproduction of the stochastic-search results requires the pinned numerical stack recorded in the repository, comprising the interpreter, the NumPy build and its bundled linear-algebra kernels, because two of the compared methods evaluate library transcendental and eigendecomposition primitives whose last-bit results

are hardware-dependent under runtime-dispatched kernels; the deterministic methods reproduce identically on any platform, and because the statistical analysis operates on within-problem ranks computed from the archived results file, all reported rankings, significance tests and collector optima are invariant to this residual last-bit divergence.

Acknowledgments: The authors are thankful to the Deanship of Graduate Studies and Scientific Research at the University of Bisha for supporting this work through the Fast-Track Research Support Program.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following symbols and abbreviations are used in this manuscript:

A_c	aperture area
A_s	wetted area
c_p	specific heat
d_i	inner diameter
F'	collector efficiency factor
F_R	heat-removal factor
f	friction factor
h	convective coefficient
I_s	irradiance
k	thermal conductivity
Nu	Nusselt number
PEC	performance evaluation criterion
Pr	Prandtl number
Q_u	useful gain
Re	Reynolds number
s	component-1 mass share
T_i	inlet temperature
T_o	outlet temperature
T_a	ambient temperature
V	loop flow rate
w	total weight fraction
W_p	pumping power
λ	penalty coefficient
η_{th}	thermal efficiency
μ	dynamic viscosity
ρ	density
φ	volume fraction
bf	base fluid (subscript)
hnf	hybrid nanofluid (subscript)
p	particle (subscript)

Appendix A

Table A1. Top-region ranking validation of the performance-criterion surrogate.

Metric	Value
Spearman rank correlation, all test points	0.9696
Spearman rank correlation, top 5 percent	0.89
Mean absolute error, top 1 percent	0.002
Recall at twenty (set overlap of the two top-20 grid sets), surrogate vs. reference implementation	1
Rank of the surrogate's argmax design under the reference ordering (of 1512)	3

Table A2. Median, interquartile range and success rate on the two multimodal benchmarks (30 runs).

Problem/Method	Median	IQR	Success Rate
Rastrigin/HH-UCB	9.419	5.864	0
Rastrigin/DE	41.126	5.177	0
Rastrigin/CMA-ES	13.929	7.711	0
Rastrigin/VNS	4.726	2.999	0
Schwefel/HH-UCB	594.998	281.894	0
Schwefel/DE	1415.917	463.770	0
Schwefel/CMA-ES	1381.859	505.655	0
Schwefel/VNS	475.990	171.901	0

Table A3. Benchmark problem definitions. Baselines: differential evolution (NP 30, F 0.6, CR 0.9), covariance-matrix adaptation (default population), hyper-heuristic (NP 16, LAHC length 50, c 0.10, alpha 0.30); budgets 5000 (benchmarks) and 700 (collector). In the search-domain column the superscript denotes the problem dimension, that is, the Cartesian power of the stated interval.

Problem	Dim	Search Domain	Global Optimum/Constraints
Sphere	10	$[-5.12, 5.12]^{10}$	0 at the origin
Rastrigin	10	$[-5.12, 5.12]^{10}$	0 at the origin
Ackley	10	$[-32.77, 32.77]^{10}$	0 at the origin
Rosenbrock	10	$[-5, 10]^{10}$	0 at all-ones
Griewank	10	$[-600, 600]^{10}$	0 at the origin
Schwefel	10	$[-500, 500]^{10}$	0 at 420.97 per dim
Welded beam	4	h, b in $[0.1, 2]$; l, t in $[0.1, 10]$	7 constraints; optimum 1.7249 [51]
Pressure vessel	4	Ts, Th in $[0.0625, 6.1875]$; R, L in $[10, 200]$	4 constraints, continuous form [52]

Table A4. Operator invocation fraction, number of new global bests produced, and mean reward on Rastrigin. The invocation fraction and mean reward are averaged over five runs; the new-global-best counts are pooled totals across the five runs. The population operators produce most of the new global bests.

Operator	Invocation Fraction	New-Best Count	Mean Reward
DE/best/1	0.101	291	0.1580
PSO	0.087	358	0.2287
DE/rand/1	0.083	60	0.0403
SA	0.081	0	0.0000
VNS	0.082	36	0.0247
LNS	0.081	4	0.0028
Tabu	0.081	11	0.0076
perturb-V	0.081	13	0.0091
perturb-w	0.081	30	0.0208
perturb-share	0.081	29	0.0202
swap-hybrid	0.081	11	0.0077
swap-fluid	0.081	8	0.0057

Table A5. Asymptotic per-invocation computational cost and memory of the compared methods (d, dimension; NP, population size; L, late-acceptance history length; |H|, portfolio size). For every operator except the tabu search, one invocation is one function evaluation; the tabu search performs up to six. The surrogate evaluation is the dominant per-evaluation cost and is identical for all methods.

Method	Per-Invocation Work	Memory
Trajectory moves (SA, VNS, LNS)	$O(d)$	$O(d)$
Tabu search	$O(6d)$ neighbour scan	$O(d) + \text{tabu list}$
Differential evolution (rand/1, best/1)	$O(d)$	$O(NP \cdot d)$
Particle swarm	$O(d)$	$O(NP \cdot d)$
Covariance-matrix adaptation (CMA-ES)	$O(d^2)$ per candidate; $O(d^3)$ eigendecom. per generation	$O(d^2)$
Selection hyper-heuristic (this work)	operator cost + $O(H)$ select + $O(1)$ accept	$O(NP \cdot d + L + H)$

Appendix B

Algorithm A1. Selection hyper-heuristic (UCB credit assignment with late-acceptance hill-climbing)

*Input: problem p , budget B (including initialisation),
operator portfolio $H = \{h_1, \dots, h_{12}\}$, population size NP ,
LAHC history length L*

*Output: best solution s^**

initialise and evaluate population P of NP solutions

$E \leftarrow NP$

$s \leftarrow \text{best of } P$ // incumbent

$s^* \leftarrow s$ // global best

for $k = 1$ to $|H|$ do

$value[k] \leftarrow 0$

$n[k] \leftarrow 0$

for $v = 0$ to $L-1$ do

$Lh[v] \leftarrow f(s)$ // LAHC history initialised to incumbent cost

$t \leftarrow 0$

while $E < B$ do

$t \leftarrow t + 1$

$b0 \leftarrow f(s^*)$ // best cost before this move (for reward)

 // ---- operator selection (UCB1) ----

 if exists k with $n[k] = 0$ then

$k \leftarrow \text{first unused operator}$

 else

$k \leftarrow \text{argmax}_k (value[k] + c * \sqrt{\ln t / n[k]})$

 // ---- apply operator with remaining budget ----

$mmax \leftarrow B - E$

$(x', mk) \leftarrow \text{apply } h_k \text{ to } s, \text{ passing } mmax$

 // $mk = 1$ normally; for tabu operator $1 \leq mk \leq mmax$

$E \leftarrow E + mk$

$n[k] \leftarrow n[k] + 1$

 // ---- credit assignment (exponential recency reward) ----

$r \leftarrow 1$ if $f(x') < b0$ else 0

$value[k] \leftarrow value[k] + \alpha * (r - value[k])$

 // ---- LAHC acceptance ----

$v \leftarrow t \bmod L$

 if $f(x') \leq Lh[v]$ or $f(x') \leq f(s)$ then

$s \leftarrow x'$

$Lh[v] \leftarrow f(s)$

 // history overwritten every iteration,

 // outside the acceptance test

 // ---- population update and global best ----

 if $f(x') < \text{worst}(P)$ then

 replace $\text{worst}(P)$ by x'

 if $f(x') < f(s^*)$ then

$s^* \leftarrow x'$

return s^*

References

- Duffie, J.A.; Beckman, W.A. *Solar Engineering of Thermal Processes*, 4th ed.; Wiley: Hoboken, NJ, USA, 2013.
- Eltaweel, M.; Abdel-Rehim, A.A.; Attia, A.A.A. A Comparison Between Flat-Plate and Evacuated Tube Solar Collectors in Terms of Energy and Exergy Analysis by Using Nanofluid. *Appl. Therm. Eng.* **2021**, *186*, 116516. [\[CrossRef\]](#)
- Choi, S.U.S.; Eastman, J.A. Enhancing Thermal Conductivity of Fluids with Nanoparticles. In *Proceedings of the 1995 ASME International Mechanical Engineering Congress and Exposition, San Francisco, CA, USA, 12–17 November 1995*; ASME: New York, NY, USA, 1995; FED-Volume 231/MD-Volume 66, pp. 99–105.
- Sajid, M.U.; Ali, H.M. Thermal Conductivity of Hybrid Nanofluids: A Critical Review. *Int. J. Heat Mass Transf.* **2018**, *126*, 211–234. [\[CrossRef\]](#)
- Babar, H.; Ali, H.M. Towards Hybrid Nanofluids: Preparation, Thermophysical Properties, Applications, and Challenges. *J. Mol. Liq.* **2019**, *281*, 598–633. [\[CrossRef\]](#)
- Suresh, S.; Venkitaraj, K.P.; Selvakumar, P.; Chandrasekar, M. Synthesis of Al_2O_3 -Cu/Water Hybrid Nanofluids Using Two Step Method and Its Thermo Physical Properties. *Colloids Surf. A Physicochem. Eng. Asp.* **2011**, *388*, 41–48. [\[CrossRef\]](#)
- Takabi, B.; Salehi, S. Augmentation of the Heat Transfer Performance of a Sinusoidal Corrugated Enclosure by Employing Hybrid Nanofluid. *Adv. Mech. Eng.* **2014**, *6*, 147059. [\[CrossRef\]](#)
- Sundar, L.S.; Singh, M.K.; Sousa, A.C.M. Enhanced Heat Transfer and Friction Factor of MWCNT- Fe_3O_4 /Water Hybrid Nanofluids. *Int. Commun. Heat Mass Transf.* **2014**, *52*, 73–83. [\[CrossRef\]](#)
- Maxwell, J.C. *A Treatise on Electricity and Magnetism*, 2nd ed.; Clarendon Press: Oxford, UK, 1881.
- Hamilton, R.L.; Crosser, O.K. Thermal Conductivity of Heterogeneous Two-Component Systems. *Ind. Eng. Chem. Fundam.* **1962**, *1*, 187–191. [\[CrossRef\]](#)
- Brinkman, H.C. The Viscosity of Concentrated Suspensions and Solutions. *J. Chem. Phys.* **1952**, *20*, 571. [\[CrossRef\]](#)
- Verma, S.K.; Tiwari, A.K.; Tiwari, S.; Chauhan, D.S. Performance Analysis of Hybrid Nanofluids in Flat Plate Solar Collector as an Advanced Working Fluid. *Sol. Energy* **2018**, *167*, 231–241. [\[CrossRef\]](#)
- Farajzadeh, E.; Movahed, S.; Hosseini, R. Experimental and Numerical Investigations on the Effect of Al_2O_3 /TiO₂-H₂O Nanofluids on Thermal Efficiency of the Flat Plate Solar Collector. *Renew. Energy* **2018**, *118*, 122–130. [\[CrossRef\]](#)
- Sharafeldin, M.A.; Gróf, G.; Abu-Nada, E.; Mahian, O. Evacuated Tube Solar Collector Performance Using Copper Nanofluid: Energy and Environmental Analysis. *Appl. Therm. Eng.* **2019**, *162*, 114205. [\[CrossRef\]](#)
- Henein, S.M.; Abdel-Rehim, A.A.; El-Nagar, K. Energy, Economic and Environmental Analysis of an Evacuated Tube Solar Collector Using Hybrid Nanofluid. *Appl. Therm. Eng.* **2023**, *219*, 119671. [\[CrossRef\]](#)
- Yildirim, E.; Yurddas, A. Assessments of Thermal Performance of Hybrid and Mono Nanofluid U-Tube Solar Collector System. *Renew. Energy* **2021**, *171*, 1079–1096. [\[CrossRef\]](#)
- Zeng, J.; Xuan, Y. Enhanced Solar Thermal Conversion and Thermal Conduction of MWCNT-SiO₂/Ag Binary Nanofluids. *Appl. Energy* **2018**, *212*, 809–819. [\[CrossRef\]](#)
- Yarmand, H.; Gharekhani, S.; Ahmadi, G.; Shirazi, S.F.S.; Baradaran, S.; Montazer, E.; Zubir, M.N.M.; Alehashem, M.S.; Kazi, S.N.; Dahari, M. Graphene Nanoplatelets-Silver Hybrid Nanofluids for Enhanced Heat Transfer. *Energy Convers. Manag.* **2015**, *100*, 419–428. [\[CrossRef\]](#)
- Chandran, M.N.; Manikandan, S.; Suganthi, K.S.; Rajan, K.S. Novel Hybrid Nanofluid with Tunable Specific Heat and Thermal Conductivity: Characterization and Performance Assessment for Energy Related Applications. *Energy* **2017**, *140*, 27–39. [\[CrossRef\]](#)
- Esfahani, N.N.; Toghraie, D.; Afrand, M. A New Correlation for Predicting the Thermal Conductivity of ZnO-Ag (50%–50%)/Water Hybrid Nanofluid: An Experimental Study. *Powder Technol.* **2018**, *323*, 367–373. [\[CrossRef\]](#)
- Webb, R.L. Performance Evaluation Criteria for Use of Enhanced Heat Transfer Surfaces in Heat Exchanger Design. *Int. J. Heat Mass Transf.* **1981**, *24*, 715–726. [\[CrossRef\]](#)
- Hemmat Esfe, M.; Esfandeh, S.; Saedodin, S.; Rostamian, H. Experimental Evaluation, Sensitivity Analyzation and ANN Modeling of Thermal Conductivity of ZnO-MWCNT/EG-Water Hybrid Nanofluid for Engineering Applications. *Appl. Therm. Eng.* **2017**, *125*, 673–685. [\[CrossRef\]](#)
- Hemmat Esfe, M.; Hajmohammad, M.H. Thermal Conductivity and Viscosity Optimization of Nanodiamond-Co₃O₄/EG (40:60) Aqueous Nanofluid Using NSGA-II Coupled with RSM. *J. Mol. Liq.* **2017**, *238*, 545–552. [\[CrossRef\]](#)
- Cowling, P.; Kendall, G.; Soubeiga, E. A Hyperheuristic Approach to Scheduling a Sales Summit. In *Practice and Theory of Automated Timetabling III, Proceedings of the PATAT 2000, Konstanz, Germany, 16–18 August 2000*; Burke, E., Erben, W., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2001; Volume 2079, pp. 176–190. [\[CrossRef\]](#) [\[PubMed\]](#)
- Burke, E.K.; Gendreau, M.; Hyde, M.; Kendall, G.; Ochoa, G.; Özcan, E.; Qu, R. Hyper-Heuristics: A Survey of the State of the Art. *J. Oper. Res. Soc.* **2013**, *64*, 1695–1724. [\[CrossRef\]](#)

26. Drake, J.H.; Kheiri, A.; Özcan, E.; Burke, E.K. Recent Advances in Selection Hyper-Heuristics. *Eur. J. Oper. Res.* **2020**, *285*, 405–428. [\[CrossRef\]](#)
27. Burke, E.K.; Bykov, Y. The Late Acceptance Hill-Climbing Heuristic. *Eur. J. Oper. Res.* **2017**, *258*, 70–78. [\[CrossRef\]](#)
28. Dueck, G. New Optimization Heuristics: The Great Deluge Algorithm and the Record-to-Record Travel. *J. Comput. Phys.* **1993**, *104*, 86–92. [\[CrossRef\]](#)
29. Auer, P.; Cesa-Bianchi, N.; Fischer, P. Finite-Time Analysis of the Multiarmed Bandit Problem. *Mach. Learn.* **2002**, *47*, 235–256. [\[CrossRef\]](#)
30. Li, K.; Fialho, A.; Kwong, S.; Zhang, Q. Adaptive Operator Selection with Bandits for a Multiobjective Evolutionary Algorithm Based on Decomposition. *IEEE Trans. Evol. Comput.* **2014**, *18*, 114–130. [\[CrossRef\]](#)
31. Wolpert, D.H.; Macready, W.G. No Free Lunch Theorems for Optimization. *IEEE Trans. Evol. Comput.* **1997**, *1*, 67–82. [\[CrossRef\]](#)
32. Hottel, H.C.; Whillier, A. Evaluation of Flat-Plate Solar Collector Performance. In *Transactions of the Conference on the Use of Solar Energy*; University of Arizona Press: Tucson, AZ, USA, 1958; Volume 2, pp. 74–104.
33. Hausen, H. Darstellung des Wärmeüberganges in Rohren durch verallgemeinerte Potenzbeziehungen. *VDI-Z. Beih. Verfahrenstech.* **1943**, *4*, 91–98.
34. Dittus, F.W.; Boelter, L.M.K. Heat Transfer in Automobile Radiators of the Tubular Type. *Int. Commun. Heat Mass Transf.* **1985**, *12*, 3–22. [\[CrossRef\]](#)
35. Gnielinski, V. New Equations for Heat and Mass Transfer in Turbulent Pipe and Channel Flow. *Int. Chem. Eng.* **1976**, *16*, 359–368.
36. Petukhov, B.S. Heat Transfer and Friction in Turbulent Pipe Flow with Variable Physical Properties. *Adv. Heat Transf.* **1970**, *6*, 503–564. [\[CrossRef\]](#)
37. Chon, C.H.; Kihm, K.D.; Lee, S.P.; Choi, S.U.S. Empirical Correlation Finding the Role of Temperature and Particle Size for Nanofluid (Al₂O₃) Thermal Conductivity Enhancement. *Appl. Phys. Lett.* **2005**, *87*, 153107. [\[CrossRef\]](#)
38. Nguyen, C.T.; Desgranges, F.; Roy, G.; Galanis, N.; Mare, T.; Boucher, S.; Angue Mintsu, H. Temperature and Particle-Size Dependent Viscosity Data for Water-Based Nanofluids: Hysteresis Phenomenon. *Int. J. Heat Fluid Flow* **2007**, *28*, 1492–1506. [\[CrossRef\]](#)
39. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [\[CrossRef\]](#)
40. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [\[CrossRef\]](#)
41. Sobol', I.M. Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates. *Math. Comput. Simul.* **2001**, *55*, 271–280. [\[CrossRef\]](#)
42. Saltelli, A.; Annoni, P.; Azzini, I.; Campolongo, F.; Ratto, M.; Tarantola, S. Variance Based Sensitivity Analysis of Model Output: Design and Estimator for the Total Sensitivity Index. *Comput. Phys. Commun.* **2010**, *181*, 259–270. [\[CrossRef\]](#)
43. Morris, M.D. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics* **1991**, *33*, 161–174. [\[CrossRef\]](#)
44. Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P. Optimization by Simulated Annealing. *Science* **1983**, *220*, 671–680. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Glover, F. Future Paths for Integer Programming and Links to Artificial Intelligence. *Comput. Oper. Res.* **1986**, *13*, 533–549. [\[CrossRef\]](#)
46. Mladenović, N.; Hansen, P. Variable Neighborhood Search. *Comput. Oper. Res.* **1997**, *24*, 1097–1100. [\[CrossRef\]](#)
47. Storn, R.; Price, K. Differential Evolution: A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *J. Glob. Optim.* **1997**, *11*, 341–359. [\[CrossRef\]](#)
48. Kennedy, J.; Eberhart, R. Particle Swarm Optimization. In Proceedings of the ICNN'95: International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948. [\[CrossRef\]](#)
49. Hansen, N.; Ostermeier, A. Completely Derandomized Self-Adaptation in Evolution Strategies. *Evol. Comput.* **2001**, *9*, 159–195. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Fialho, A.; Da Costa, L.; Schoenauer, M.; Sebag, M. Analyzing Bandit-Based Adaptive Operator Selection Mechanisms. *Ann. Math. Artif. Intell.* **2010**, *60*, 25–64. [\[CrossRef\]](#)
51. Ragsdell, K.M.; Phillips, D.T. Optimal Design of a Class of Welded Structures Using Geometric Programming. *J. Eng. Ind.* **1976**, *98*, 1021–1025. [\[CrossRef\]](#)
52. Kannan, B.K.; Kramer, S.N. An Augmented Lagrange Multiplier Based Method for Mixed Integer Discrete Continuous Optimization and Its Applications to Mechanical Design. *J. Mech. Des.* **1994**, *116*, 405–411. [\[CrossRef\]](#)
53. Coello Coello, C.A. Theoretical and Numerical Constraint-Handling Techniques Used with Evolutionary Algorithms: A Survey of the State of the Art. *Comput. Methods Appl. Mech. Eng.* **2002**, *191*, 1245–1287. [\[CrossRef\]](#)
54. Demšar, J. Statistical Comparisons of Classifiers over Multiple Data Sets. *J. Mach. Learn. Res.* **2006**, *7*, 1–30.
55. Derrac, J.; García, S.; Molina, D.; Herrera, F. A Practical Tutorial on the Use of Nonparametric Statistical Tests as a Methodology for Comparing Evolutionary and Swarm Intelligence Algorithms. *Swarm Evol. Comput.* **2011**, *1*, 3–18. [\[CrossRef\]](#)
56. Holm, S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.

57. Said, S.; Mellouli, S.; Alqahtani, T.; Algarni, S.; Ajjel, R.; Ghachem, K.; Kolsi, L. An Experimental Comparison of the Performance of Various Evacuated Tube Solar Collector Designs. *Sustainability* **2023**, *15*, 5533. [[CrossRef](#)]
58. Elshazly, E.; Abdel-Rehim, A.A.; El-Mahallawi, I. Thermal Performance Enhancement of Evacuated Tube Solar Collector Using MWCNT, Al₂O₃ and Hybrid MWCNT/Al₂O₃ Nanofluids. *Int. J. Thermofluids* **2023**, *17*, 100260. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.