Exergy-Guided Retrofit Selection for Heat-Exchanger Networks

George Weale
University of California, Santa Barbara
Santa Barbara, CA, USA
gweale@ucsb.edu

***Abstract*—Retrofit studies of heat-exchanger networks (HENs) often begin with energy targets and end with a ranked list of exchanger modifications. That workflow can obscure a central engineering distinction: equal heat duties moved at different temperature levels do not have equal thermodynamic value. This paper develops a reproducible method for selecting HEN retrofit actions through exergy-guided ranking while retaining capital, operability, fouling, and uncertainty constraints. The method starts from a validated stream table and an explicit representation of the as-operated network. It generates feasible actions, including area addition, match reassignment, bypass adjustment, utility-level substitution, and cleaning-policy changes; estimates each action's avoided utility exergy destruction; and ranks portfolios through robust multi-objective optimization. A common data contract, scenario ledger, baseline set, and engineering review sequence make the ranking auditable under declared temperatures, utility assumptions, and operational limits. The resulting decision framework distinguishes thermodynamic value from utility-duty magnitude and connects optimization outputs to physically reviewable retrofit choices.**

***Index Terms*—heat-exchanger networks, exergy, retrofit, pinch analysis, process integration, robust optimization**

I. PROBLEM FORMULATION

Heat-exchanger-network retrofit selection asks which feasible modifications to an existing network most reduce avoidable resource destruction under a finite capital and outage budget. The decision depends on the temperature grade of recovered heat, the physical constraints of installed equipment, and uncertainty in operating rate, fouling, utilities, and ambient conditions. A useful formulation must therefore connect thermodynamic screening to the as-operated network rather than rank abstract heat duties in isolation.

Pinch analysis remains a useful starting point because it exposes energy-recovery targets and temperature constraints [1], [2]. It does not, by itself, make every target-equivalent retrofit equally attractive. Exergy analysis provides a complementary measure of work potential relative to a declared reference environment [3], [4]. Combining the two perspectives makes it possible to compare modifications that achieve similar energy recovery at different temperature levels and with different operating burdens.

The formulation covers steady or quasi-steady operating windows with known stream duties and utility conditions. Candidate actions enter the ranking only after temperature, pressure-drop, materials, layout, fouling, and operability predicates are defined. Site-specific hazard analysis and management of change remain the authorization mechanisms for any physical modification.

II. RELATED WORK AND MOTIVATION

Pinch analysis supplies a disciplined language for heat recovery, temperature approach, and utility targeting [1], [2]. Its enduring value is that it converts a set of stream duties into thermodynamic targets before individual exchangers are selected. Subsequent synthesis work has shown why a target is not a retrofit design: exchanger matches, area, pressure drop, layout, controllability, and fixed charges create a difficult combinatorial problem [5], [9], [8]. Process-integration texts make the same practical point in different terms: an existing network carries historical tie-ins and operating practices that cannot be erased merely because a new network has a lower utility target [10], [11]. Exergy analysis adds a useful question to energy targeting. A unit of duty near ambient conditions and a unit of duty at a high temperature may have similar energy magnitude but different work potential relative to a stated environment [3], [4]. Exergy is therefore valuable as a transparent screening measure for the temperature grade of an avoidable utility demand. It is not a full economics model and it is not an environmental certificate. In particular, a maintenance-intensive exchanger replacement can be inferior to a smaller utility improvement if it introduces unacceptable outage, fouling, or control risk. The contribution is an auditable retrofit-selection method that begins with the as-operated graph, generates only physically reviewable actions, scores their temperature-grade consequence, and tests portfolios over a declared uncertainty ledger. The method makes each ranking falsifiable: a reviewer can change an ambient reference, utility boundary, fouling range, or outage assumption and observe whether the ordering survives. Figure 1 summarizes the information structure linking the network, candidate actions, utility grades, and decision criteria.

III. SYSTEM REPRESENTATION

A. Stream and Network Data Contract

Let $\mathcal{H}$ and $\mathcal{C}$ denote hot and cold process streams. Each stream i has supply and target temperatures (T_i^s, T_i^t), heat-capacity flow rate $C_{P,i}$, allowable pressure drop, material information, and an operating envelope. Temperatures are expressed in kelvin in every thermodynamic calculation. A stream-table record is accepted only if its duty is reconciled

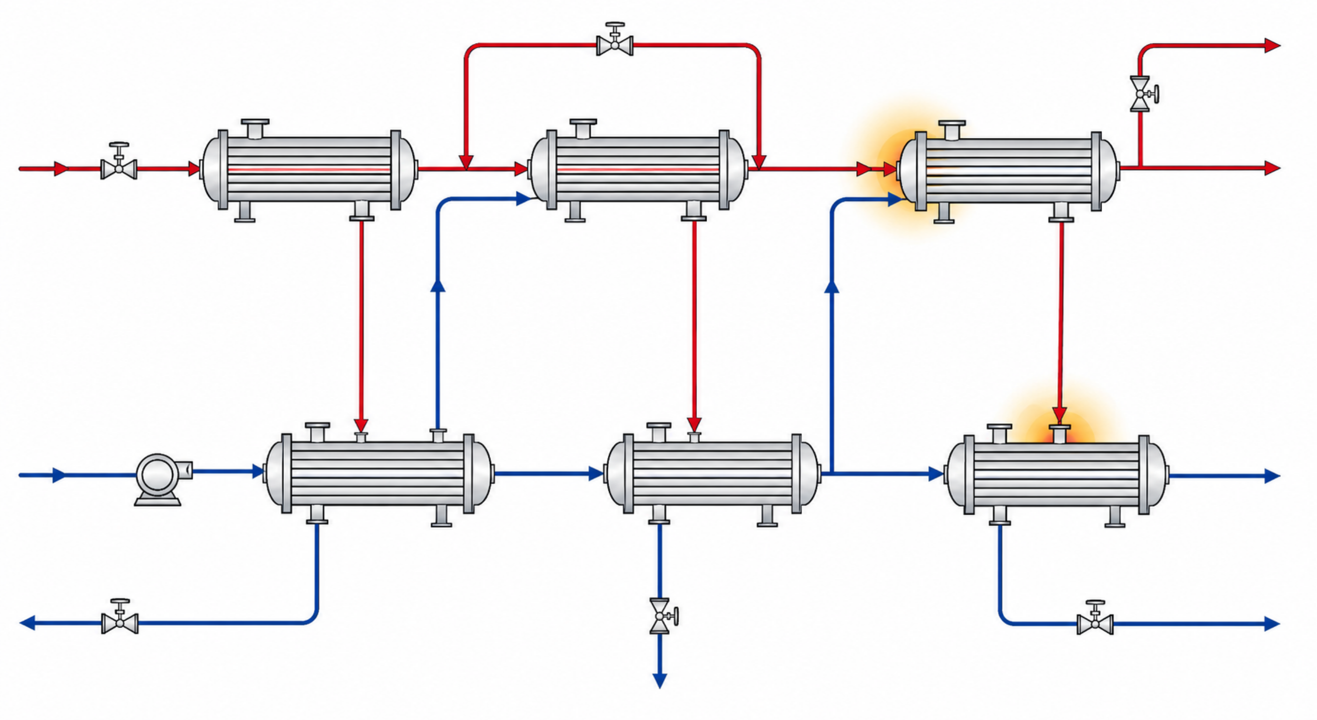

Fig. 1. Exergy-guided HEN review structure. Red and blue process paths identify locations for evaluating temperature-grade loss, bypassing, and utility substitution.

with the corresponding equipment or balance closure to a declared tolerance ϵ_Q.

The as-operated network is a graph $G_0 = (V, E_0)$ whose edges include process-process matches, heaters, coolers, bypasses, and mixers. For an exchanger e, the record includes heat duty Q_e, terminal temperatures, estimated UA_e, area A_e, fouling resistance range, and the observed or designed pressure-drop limit. Missing values are not silently imputed: they are represented as bounded uncertain parameters or mark the associated action as non-rankable.

For sensible segments, the data record carries the enthalpy change implied by the reconciled heat-capacity flow rate and terminal temperatures. For a phase-change or strongly temperature-dependent stream, the representation instead uses an enthalpy-temperature curve $H_i(T)$ and interval discretization. This avoids treating a nonlinear enthalpy curve as a fictitiously constant C_P.

B. Exergy Accounting

Let T_0 be the explicitly declared ambient reference temperature. For a constant-C_P stream cooled from T_a to T_b, the physical-exergy decrease available to the network is

$$\Delta B = C_P \left[(T_a - T_b) - T_0 \ln \left(\frac{T_a}{T_b} \right) \right]. \tag{1}$$

Equation (1) is a screening expression, not a complete chemical-exergy calculation. The analysis records the boundary conditions and excludes chemical exergy unless composition data and a compatible reference-state model are provided.

For a utility supplying Q_u at an effective boundary temperature T_u, a first-order avoided-utility exergy proxy is

$$B_u = Q_u \left(1 - \frac{T_0}{T_u} \right). \tag{2}$$

The proxy is supplemented, where available, by a site utility model that maps fuel, electricity, steam level, condensate return, and boiler efficiency into marginal resource consumption. No claim of environmental benefit is made from Eq. (2) alone.

IV. RETROFIT ACTION SPACE

The candidate-action set $\mathcal{A}$ is generated from the validated graph rather than from an unconstrained synthesized network. Each action $a \in \mathcal{A}$ has a physical change, a required outage class, an implementation-cost range, and a set of feasibility predicates. Table I defines the initial action taxonomy.

For a candidate network $G(a)$, heat-transfer feasibility is enforced for every exchanger and discretized interval:

$$Q_e \le UA_e(a)\,\Delta T_{\mathrm{lm},e}(a), \tag{3}$$

$$\Delta T_{e,k}(a) \ge \Delta T_{\mathrm{min},e}, \tag{4}$$

$$\Delta P_e(a) \le \Delta P_e^{\max}. \tag{5}$$

When a terminal-temperature crossing makes the logarithmic mean temperature difference undefined, that candidate is infeasible; it is not repaired by a numerical absolute value.

TABLE I

CANDIDATE ACTION CLASSES AND MANDATORY FEASIBILITY CHECKS

Class	Example action	Required checks
Area and UA	Add surface, alter channel allocation, or replace an exchanger	$\Delta T_{\min}$, pressure drop, materials, fouling, layout
Topology	Add a permitted match, split a stream, or alter bypass routing	Temperature feasibility, controllability review, tie-in access
Utility level	Shift a duty from high-pressure to lower-grade utility	Utility availability, condensate and control constraints
Reliability	Change cleaning interval or isolate a fouling-prone bypass	Production loss, inspection evidence, maintenance feasibility

V. ROBUST RANKING MODEL

A. Scenario Set

Uncertainty is represented by a finite scenario set Ω containing operating rate, feed temperature, ambient temperature, utility price, utility marginal-emission factor, fouling resistance, and exchanger availability. The scenarios must be constructed from a recorded source: bounded design envelope, historian segments, or a deliberately stated stress-test distribution. Historical frequency must not be inferred from a handful of selected snapshots.

For portfolio $x \in \{0,1\}^{|\mathcal{A}|}$ and scenario ω, define annualized utility exergy use $B(x,\omega)$, annualized cost $C(x,\omega)$, and unavailable-production proxy $L(x,\omega)$. The scalar screening score is

$$J(x,\omega) = w_B B(x,\omega) + w_C C(x,\omega) + w_L L(x,\omega), \tag{6}$$

where weights are declared before ranking. The default report retains the full Pareto set instead of presenting Eq. (6) as an objective truth.

The robust selection problem is

$$\begin{aligned} \min_{x} \quad & \mathbb{E}_\omega[J(x,\omega)] + \lambda \, \mathrm{CVaR}_\alpha(J(x,\omega)) && (7)\\ \text{s.t.} \quad & \sum_a \mathrm{Capex}_a x_a \le \overline{K}, && (8)\\ & G(x,\omega) \text{ satisfies Eq. (5)}, \ \forall \omega \in \Omega, && (9)\\ & x_a + x_b \le 1 \quad \forall (a,b) \in \mathcal{I}, && (10) \end{aligned}$$

where $\mathcal{I}$ contains incompatible actions. The conditional value at risk term makes fragile, high-downside portfolios visible. It does not replace detailed operability or safety review.

B. Exergy-Guided Marginal Value

To make the ranking inspectable, each action also receives a scenario-specific marginal indicator

$$\eta_a(\omega) = \frac{B(0,\omega) - B(e_a,\omega)}{\mathrm{Capex}_a + \rho \, \mathrm{OutageCost}_a}, \tag{11}$$

where e_a selects only action a and ρ converts the declared outage proxy into comparable units. This indicator is not used to greedily assemble a final portfolio because action interactions can invalidate additivity. It is used to explain why a candidate appears on the Pareto frontier.

TABLE II

METRICS TO REPORT WITHOUT SELECTIVE SCENARIO OMISSION

Metric	Definition
Utility duty	Heating and cooling duties by utility level and scenario
Avoided utility exergy	Difference in Eq. (2) or site utility model
Feasible-scenario rate	Fraction of scenarios satisfying all declared constraints
Downside score	CVaR_α of cost-plus-loss objective
Decision stability	Rank correlation under one-at-a-time parameter perturbations
Audit completeness	Fraction of actions with traceable source, assumption, and reviewer status

VI. EVALUATION FRAMEWORK

A. Baselines and Comparators

The evaluation compares the exergy-guided method with: (i) no retrofit, (ii) minimum-utility targeting followed by a duty-based ranking, and (iii) a capital-only ranking. All alternatives use the same data contract, action set, and feasibility predicates. A comparison is invalid if one method receives undocumented matches, relaxed $\Delta T_{\min}$ constraints, or a different outage assumption.

Leave-window-out validation applies when historian windows are available: uncertain UA or fouling ranges are calibrated on earlier windows, and rank stability is tested on disjoint later windows. When only design data exist, the evaluation is identified as scenario analysis. Numerical performance, savings, and comparative claims are tied to the full scenario ledger and preserved baseline outputs.

B. Reproducibility Artifacts

An executable release contains the stream-table schema, network graph, unit conventions, uncertainty ledger, action generator, solver version and termination status, infeasible-candidate log, and decision notebook. Each reported figure is reproducible from an immutable input manifest. Proprietary stream values may be masked only if a synthetic, balance-consistent surrogate and the masking transformation are supplied.

VII. REVIEW WORKFLOW AND ENGINEERING HANDOFF

The decision workflow has four review moments. First, a process engineer confirms the stream-table boundary, signs, temperature bases, and reconciliation tolerance. Second, a mechanical or reliability reviewer confirms the equipment constraints attached to every candidate action: allowable pressure drop, materials, maintenance access, cleaning history, and outage class. Third, a utilities reviewer confirms that a candidate change does not simply transfer a burden to an omitted steam, condensate, refrigeration, or electrical boundary. Fourth, an operations reviewer examines the surviving portfolios for start-up, turndown, bypass, and control implications. This review sequence is part of the technical method, not project administration. A HEN retrofit often fails at the interface between a mathematically feasible exchange and a physically maintainable plant. The framework therefore retains three decision classes for every candidate: SCREENED, REVIEWABLE, and REJECTED. A screened action satisfies the automated energy and temperature predicates. A reviewable action has enough equipment information for human review. A rejected action has a retained reason such as an impossible tie-in, missing pressure-drop data, an unacceptable outage, or a cross-boundary accounting issue. Rankings are computed only from reviewable actions. The final decision artifact is a portfolio card rather than a single savings number. It records the selected changes, relevant scenario percentiles, utility-level shift, capital range, outage assumption, infeasible alternatives, and open engineering questions. This keeps the work legible to readers who do not operate the optimizer. It also makes later disagreement productive: a reviewer can challenge a named assumption or predicate rather than debate a black-box score.

A. Uncertainty Ledger Construction

The scenario ledger is built from explicit uncertainty records rather than a single undifferentiated random sample. Operating-rate and ambient-temperature ranges may be obtained from selected historical windows or a declared design envelope. Fouling resistance and UA uncertainty identify whether the range reflects measurement error, expected degradation, or an unknown equipment condition. Utility price and marginal resource factors identify their temporal basis and whether they are used for economics, resource screening, or both. A scenario inherits the provenance of its inputs so that a stress test remains distinguishable from a forecast. Correlations are stated where they affect a conclusion. For example, a high ambient temperature may coincide with higher cooling demand; a higher throughput may simultaneously change stream duty and available utility capacity. Where the dependence is unknown, the framework uses named compound stress cases rather than simulated precision. The report must show which action rankings are stable across every scenario, which are stable only under an independence assumption, and which are dominated by uncertainty. This is more decision-useful than reporting a mean exergy benefit without the conditions that produced it. A final sensitivity analysis varies the reference temperature, minimum approach temperature, capex range, outage multiplier, and fouling bounds one at a time and in selected combinations. Its purpose is to determine whether a conclusion survives reasonable disagreement about the boundary. A portfolio that is attractive only at one narrow value is classified as hypothesis-generating rather than decision-ready.

VIII. APPLICATION AND FALSIFIABILITY

Application proceeds in two stages for a documented utility system. The first constructs a data package from PFDs, exchanger sheets, utility balances, and independently selected operating windows, with review for balance closure and for the boundary between process heat recovery and site utilities. The second runs the fixed candidate generator and scenario framework in shadow mode while operations and maintenance reviewers assess whether each retained action is physically implementable. Field changes remain subject to the site's engineering authorization process. The method is weakened if the ranked portfolio changes under small, defensible changes to T_0 or fouling ranges; if a high-ranked action fails a basic layout or pressure-drop predicate; if the apparent exergy reduction arises only from moving a duty outside the chosen boundary; or if a duty-based baseline produces the same robust ordering with less information. These falsification criteria distinguish a useful decision method from a relabeled utility-target calculation.

IX. THREATS TO VALIDITY AND BOUNDARIES

Exergy is sensitive to the chosen reference environment and to the utility boundary. A reduction in high-temperature utility duty may be valuable even when the inferred exergy gain is modest, and the converse can occur when an action creates a maintenance or control burden. Fouling models, heat-transfer coefficients, and future operating distributions are often poorly observed. The scenario framework exposes those uncertainties but cannot make them disappear.

The framework complements rather than replaces lifecycle assessment, equipment-integrity analysis, relief design, process-safety review, and commercial evaluation. Its methodological claim is that a retrofit recommendation states its energy, temperature-grade, uncertainty, and feasibility assumptions in a form that can be checked.

AUDIT AND REPRODUCIBILITY RECORD

Implementation creates an immutable decision record before any ranking is circulated. The record begins with an input manifest containing stream identifiers, units, supply and target temperatures, duty-reconciliation status, exchanger records, utility definitions, and the source of every uncertain range. Each manifest entry receives a revision identifier, data steward, date of extraction, and a statement of whether it is measured, design-derived, estimated, or synthetic. This prevents a later spreadsheet update from silently changing the evidence base for a recommendation. Every candidate action must also retain a machine-readable rationale. The action generator records which topology rule, area opportunity, utility substitution, or cleaning-policy change produced the candidate. The feasibility evaluator records terminal temperature checks, pressure-drop checks, missing attributes, and the first rule that rejected an action. The optimizer then records solver version, objective weights, scenario identifiers, termination status, and the unrounded numerical values used to create each portfolio card. A reviewer should be able to reproduce a reported ordering from the manifest without relying on an analyst's private notebook. The final audit packet separates assumptions from decisions. Assumptions include reference environment, utility boundary, fouling ranges, capital ranges, outage treatment, and scenario correlations. Decisions include which portfolios were reviewed, which were rejected by engineering authority, and which questions remain open. Any later change to an assumption creates a new scenario ledger and a diff against the earlier ranking. The resulting evidence trail supports accountable engineering review of safety, feasibility, and economics.

X. CONCLUSION

This paper establishes an exergy-guided method for HEN retrofit selection. It combines a reconciled as-operated network, physically constrained action generation, scenario-based robust selection, and an explicit engineering review sequence. The framework shows how temperature-grade information, economic constraints, uncertainty, and implementability can be evaluated within one auditable decision record, avoiding reliance on a single favorable utility target.

REFERENCES

[1] B. Linnhoff and E. Hindmarsh, “The pinch design method for heat exchanger networks,” *Chemical Engineering Science*, vol. 38, no. 5, pp. 745–763, 1983, doi: 10.1016/0009-2509(83)80185-7.

[2] I. C. Kemp, *Pinch Analysis and Process Integration: A User Guide on Process Integration for the Efficient Use of Energy*, 2nd ed. Oxford, U.K.: Butterworth-Heinemann, 2007.

[3] T. J. Kotas, *The Exergy Method of Thermal Plant Analysis*. London, U.K.: Butterworths, 1985.

[4] A. Bejan, *Advanced Engineering Thermodynamics*, 4th ed. Hoboken, NJ, USA: Wiley, 2016.

[5] S. A. Papoulias and I. E. Grossmann, “A structural optimization approach in process synthesis–II. Heat recovery networks,” *Computers & Chemical Engineering*, vol. 7, no. 6, pp. 707–721, 1983.

[6] I. E. Grossmann, “Enterprise-wide optimization: A new frontier in process systems engineering,” *AIChE Journal*, vol. 51, no. 7, pp. 1846–1857, 2005, doi: 10.1002/aic.10617.

[7] R. T. Rockafellar and S. Uryasev, “Optimization of conditional value-at-risk,” *Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.

[8] C. A. Floudas, *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*. New York, NY, USA: Oxford Univ. Press, 1995.

[9] T. Gundersen and L. Naess, “The synthesis of cost optimal heat exchanger networks: An industrial review of the state of the art,” *Computers & Chemical Engineering*, vol. 12, no. 6, pp. 503–530, 1988, doi: 10.1016/0098-1354(88)87002-9.

[10] R. Smith, *Chemical Process Design and Integration*, 2nd ed. Chichester, U.K.: Wiley, 2016.

[11] J. J. Klemes, Ed., *Handbook of Process Integration (PI): Minimisation of Energy and Water Use, Waste and Emissions*. Oxford, U.K.: Woodhead Publishing, 2013.