ORIGINAL RESEARCH Open access

A Theory of Multi-Objective Trade-Offs for Sustainable Materials Optimization with AI

Daniel Fischer[1*], Laura Meier[1], Thomas Braun[2], Stefan Koch[2], Felix Roth[1]

Abstract

Artificial intelligence (AI) is increasingly positioned as a design partner in materials optimization, enabling accelerated exploration of vast composition–processing–structure spaces under multiple, often conflicting, targets. Yet sustainability-centered materials design is not simply a larger version of multi-property optimization: it requires negotiating trade-offs across heterogeneous objective types such as performance, cost, safety, emissions, toxicity, circularity, and resource criticality, while accounting for lifecycle shifts and stakeholder-dependent priorities. Many current AI-enabled optimization workflows implicitly treat trade-offs as static Pareto-front problems with stable objective meanings and fixed feasibility boundaries. This conceptual manuscript argues that such assumptions are structurally incompatible with sustainable materials decisions, which involve trade-offs that are contextual, value-weighted, and regime-dependent. We introduce a novel theoretical framework—Trade-Off Sensitivity Theory (TOST)—which models sustainability optimization as a decision process governed by objective incompatibility geometry, lifecycle constraint migration, uncertainty-to-consequence coupling, and preference volatility. Rather than proposing algorithms or empirical evaluation, TOST provides a theoretical map linking Pareto efficiency to sustainability legitimacy through three layers: objective semantics, trade-off sensitivity, and action admissibility. The framework clarifies when AI outputs support responsible selection, when optimization is ill-posed, and how sustainable decisions can be justified under conflicting criteria.

Keywords Materials informatics, Multi-objective optimization, Sustainable materials, Pareto trade-offs, Life cycle assessment, Decision legitimacy

*Correspondence:
Daniel Fischer
daniel.fischer@gmail.com
[1] Department of Materials Science and Intelligent Systems, Faculty of Engineering, University of Freiburg, Freiburg, Germany
[2] Department of AI-Driven Materials Engineering, Faculty of Engineering, Karlsruhe Institute of Technology, Karlsruhe, Germany

Introduction

Materials innovation has historically been organized around performance improvement: increase strength, enhance conductivity, raise catalytic activity, reduce weight, improve durability, or minimize cost. In many conventional workflows, “optimization” is interpreted as a technical act—selecting a composition and processing route that maximizes a target property under known constraints. In materials informatics, AI has amplified this approach by accelerating property prediction, enabling rapid screening, and supporting high-dimensional exploration of candidate spaces [1–3]. However, sustainable materials optimization forces a deeper question: optimize for whom, against what harms, and under which lifecycle boundary conditions [4–7]?

Sustainability is not an additional objective that can be appended to legacy performance metrics without conceptual reframing. It changes the meaning of the optimization task because sustainability objectives differ from classical engineering objectives in at least four ways. First, they are inherently lifecycle-coupled: carbon footprint, toxicity, and circularity do not exist at a single point in the design space but emerge across supply chains, manufacturing routes, usage conditions, and end-of-life scenarios [4–6]. Second, they are often multi-stakeholder and value-dependent: what counts as "acceptable" depends on policy, regulatory constraints, corporate strategy, risk tolerance, and social priorities [6, 7]. Third, sustainability criteria are frequently heterogeneous in measurement and meaning: "low carbon," "non-toxic," and "recyclable" do not share common units, and they are not directly comparable without normative judgments [5, 7]. Fourth, they are regime-sensitive: a material choice that appears sustainable under one boundary condition (e.g., renewable electricity, local recycling infrastructure) can become unsustainable under another [4–6].

At the same time, multi-objective optimization (MOO) has become a standard formal language for handling competing design goals. Pareto efficiency provides a powerful conceptual tool: a solution is Pareto-optimal if no objective can be improved without worsening at least one other objective. In AI-enabled design, Pareto front reasoning supports trade-off navigation between performance and cost, performance and manufacturability, or strength and ductility [8–10]. Yet sustainable materials optimization differs from classical MOO because it often involves objectives that are not merely competing but structurally incompatible or semantically unstable across contexts [6, 11]. For example, optimizing for maximum recyclability can expand the allowable chemical space, reshaping performance feasibility; similarly, minimizing critical raw material dependence can transform processing options, altering both economic and environmental profiles [4–6]. The trade-off is not simply numeric—it is ontological, altering what design actions remain meaningful.

This manuscript argues that a central limitation in current "AI for sustainable materials optimization" thinking is the tendency to treat sustainability trade-offs as static and geometry-only: produce a Pareto front, choose a compromise point, and proceed. This approach often assumes that objectives are fixed, commensurable, and stable in meaning across deployment contexts [6, 7, 12]. However, sustainability decisions frequently require justification rather than mere selection. A designer must defend why a specific trade-off is legitimate under a stated lifecycle boundary, why certain harms were accepted, and why the chosen compromise remains stable under uncertainty and changing constraints [5–7]. This is a different kind of problem: it is a problem of decision legitimacy, not only optimization efficiency [6, 13, 14].

Recent scholarship in AI governance and responsible AI emphasizes that model performance alone is insufficient for trustworthy deployment; decisions must also be robust to shifts, transparent about their assumptions, and accountable for the consequences of error [13–16]. In materials contexts, these concerns are amplified because decisions may propagate into physical artifacts, industrial scaling, supply-chain lock-in, and environmental burdens that are expensive or impossible to reverse [4–7]. Therefore, a sustainability-optimized candidate is not merely a predicted winner; it is a claim of acceptability under a trade-off structure that must remain defensible beyond the dataset regime [6, 15, 16].

To address this conceptual gap, we propose Trade-Off Sensitivity Theory (TOST). This novel theoretical framework explains how sustainable multi-objective optimization should be understood as a governed decision-making process rather than a purely computational search. TOST does not propose an algorithm, training strategy, or empirical pipeline. Instead, it provides a conceptual architecture organizing sustainability trade-offs around four core mechanisms:

1. Objective incompatibility geometry (how objectives structurally conflict),

2. Lifecycle constraint migration (how feasible regions shift over the lifecycle),

3. Uncertainty-to- consequence coupling (why uncertainty becomes ethically and operationally weighted), and

4. Preference volatility (how stakeholders and policies reshape what "optimal" means over time) [4–7, 11, 13].

The contribution of TOST is theoretical: it reframes the Pareto front from an endpoint into an intermediate object whose meaning must be interpreted through legitimacy criteria. In short, the manuscript argues: Pareto-optimal is not equivalent to sustainability-acceptable. A sustainable decision requires reasoning about trade-off stability, semantic alignment of objectives, and consequence-weighted uncertainty under shifting constraints [6, 7, 15].

Theoretical Background & Literature Synthesis

AI in materials optimization: From prediction to design navigation

Materials informatics has matured rapidly as AI methods have become capable of representing materials structures, learning high-dimensional relationships, and supporting accelerated exploration of design spaces [1–3]. The role of AI has expanded from property prediction to candidate ranking, inverse design, and guidance of exploration strategies [2, 3]. In this evolution, the core promise is speed: replacing costly evaluation with model-guided navigation, while capturing complex interactions that are difficult to formalize analytically [1–3].

However, sustainability objectives complicate this picture, as optimization shifts from focusing on a single target to resolving conflicts. This conflict is familiar in mechanical property design (e.g., strength–ductility trade-offs), but sustainability trade-offs are broader and cross-domain, entangling environmental, economic, and safety constraints [4–7]. The literature increasingly recognizes that sustainable materials design demands integrating environmental assessment logics with technical performance reasoning [4–6].

Multi-objective optimization and the limits of pareto-front thinking

Multi-objective optimization is widely used to formalize design problems with competing objectives, and Pareto-front reasoning provides a principled way to define non-dominated solutions [8–10]. In materials design, MOO concepts are used across alloy optimization, functional materials discovery, and process optimization because materials inherently require balancing multiple properties [8, 9].

Yet Pareto-front reasoning can silently conceal critical assumptions. It typically treats objectives as stable in meaning, measurable without ambiguity, and mutually comparable through dominance relations [8–10]. Sustainability objectives often violate all three assumptions. For instance, “low toxicity” is not a single scalar property but a heterogeneous class of harms with distinct biological endpoints, exposure pathways, and uncertainties. Likewise, “circularity” is not an intrinsic material attribute but a system-level outcome that depends on infrastructure, regulation, and human behavior external to the material itself [5–7]. As a result, an optimization problem may be mathematically well-posed and solvable while remaining conceptually underspecified, with the Pareto front encoding implicit value judgments and modeling choices that are never made explicit.

Moreover, Pareto-optimal sets can be large and shallow, making them descriptive rather than prescriptive: they show what is possible, but they do not explain what is acceptable. Decision-making requires choosing among Pareto-optimal solutions, taking preferences, constraints, and risk considerations into account [10–12]. In sustainable contexts, these are not merely engineering preferences; they are governance-like commitments shaped by policy, stakeholder values, and long-term consequences [6, 7, 12].

Sustainability as lifecycle-dependent constraint structure

A central insight of sustainability science is that environmental burdens are not intrinsic properties of a material alone but outcomes of lifecycle processes and boundary definitions [4–6]. Life cycle assessment (LCA) formalizes this view by emphasizing system boundaries, functional units, and impact categories [4, 5]. For sustainable materials optimization, this implies that “objective functions” are not fixed objects; their values can change as boundaries change. A material optimized for low emissions under one electricity grid mix may no longer be optimal under another [4–6]. Likewise, a material with high recyclability potential may not realize circularity benefits without compatible collection and processing systems [6, 7]. **Figure 1** illustrates lifecycle constraint migration, showing how feasibility and Pareto-relevant trade-offs shift as design decisions move from lab conditions to scale-up and end-of-life regimes

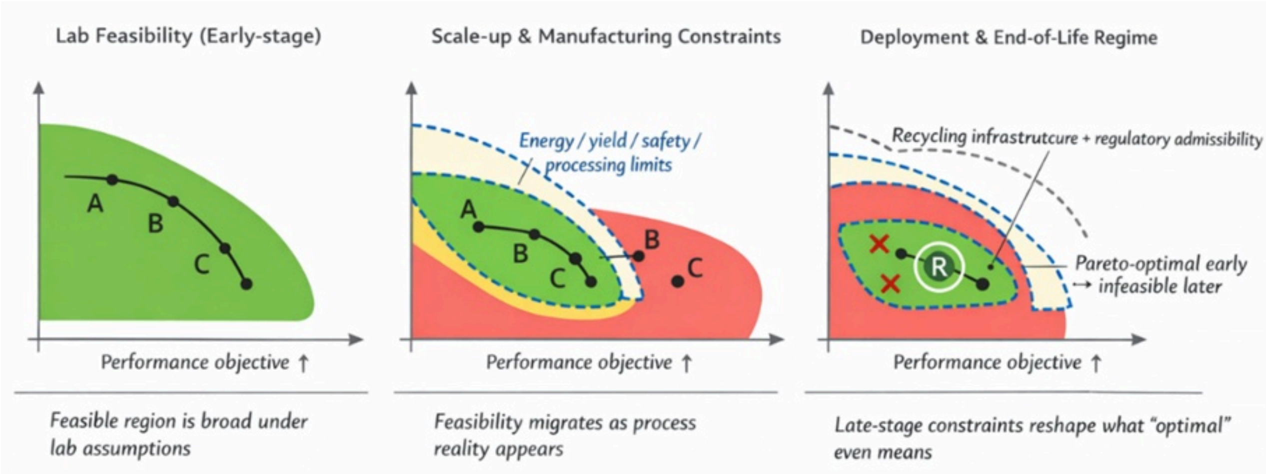

Figure 1. Lifecycle constraint migration: moving feasible regions across design stages

This observation generates an underappreciated structural consequence: the feasible region of material design is not static. Constraints migrate as one moves from lab-scale feasibility to industrial scale-up, from prototype performance to regulatory certification, and from use-phase behavior to end-of-life recovery [4–7]. Therefore, optimization must anticipate constraint migration rather than merely satisfy constraints at a single design stage.

Decision legitimacy and responsible AI: beyond "Best Compromise"

Responsible AI scholarship emphasizes that model outputs must be evaluated not only by performance but by robustness, transparency, accountability, and alignment to the decision context [13–16]. Especially relevant are concerns about distribution shift, hidden failure modes, and the consequences of errors when decisions influence physical systems [15, 16]. In materials contexts, AI-guided design can influence industrial choices that have long-term environmental impacts; therefore, errors are not merely statistical—they are consequential [4–7].

This shifts the central question from "which candidate is optimal?" to "which decision is defensible?" [6, 13, 14]. Defensibility requires more than Pareto optimality; it requires explaining how preferences were set, how uncertainty was handled, and why trade-offs were acceptable under a stated lifecycle and stakeholder frame [6, 7, 14]. This motivates a theory-first approach: a conceptual framework that treats trade-offs as objects requiring interpretation and justification.

Sustainability trade-offs as semantic, not only numeric

Many trade-offs in sustainable materials design involve objectives that are semantic constructs rather than direct measurable properties. For instance, "ethical sourcing," "responsible mining," or "socially acceptable risk" are not fully reducible to a scalar metric without assumptions and aggregation choices [6, 7, 12]. Even more "technical" terms like "low carbon" depend on choices about system boundaries, co-product allocation, and temporal accounting [4–6]. Thus, objectives carry meanings that must be stabilized before optimization can be meaningful.

This motivates the need for a conceptual layer that precedes optimization: objective semantics—the clarification of what each objective means, what is included or excluded, and what counts as improvement [6, 7, 12]. Without this, Pareto improvements can be illusory: an AI can optimize a proxy objective that appears sustainable while failing to achieve the intended sustainability meaning under realistic deployment [6, 13, 16].

Proposed conceptual framework

Trade-Off Sensitivity Theory (TOST): A decision-legitimacy model for sustainable materials optimization with AI

We propose Trade-Off Sensitivity Theory (TOST) as a novel theoretical framework for understanding and governing multi-objective trade-offs in sustainable materials optimization. TOST is not an algorithmic pipeline and does not prescribe models, datasets, experiments, or computational workflows. Instead, it defines the conceptual architecture required to interpret AI-supported trade-offs as defensible sustainability decisions.

TOST begins from a key claim: Pareto efficiency is necessary but not sufficient for sustainable acceptability. A solution can be Pareto-optimal yet still unacceptable due to toxicity risks, critical-material dependence, regulatory constraints, or lifecycle rebound effects [4–7]. Therefore, optimization must be coupled to legitimacy evaluation. TOST organizes this coupling through three governance layers:

Layer 1 — Objective semantics (Meaning stabilization)

Before trade-offs can be negotiated, each objective must be semantically stabilized. This includes clarifying boundaries (lifecycle stage, functional unit), proxy validity, and whether the objective represents a direct measurable quantity or a constructed indicator [4–6]. In TOST, unstable semantics create “false trade-offs,” where the optimization process appears rational but is built on inconsistent meanings.

Layer 2 — Trade-off sensitivity (Geometry + Context)

Traditional MOO interprets trade-offs geometrically via dominance and Pareto fronts [8–10]. TOST extends this by treating trade-offs as sensitivity objects: how sharply improvement in one objective forces degradation in another under specific lifecycle constraints. Sensitivity is high when small gains in performance require large increases in environmental burden or resource criticality; sensitivity is low when joint improvements remain feasible. This layer also explicitly accommodates constraint migration, recognizing that feasibility boundaries shift with scaling, processing realities, and deployment environments [4–7].

Layer 3 — Action admissibility (Decision legitimacy under consequences)

Finally, TOST introduces the notion of action admissibility: not all Pareto-optimal solutions warrant action. Admissibility depends on the interaction between uncertainty and governance constraints: if uncertainty interacts with high-stakes harms, design action may be inadmissible even if expected outcomes appear favorable [13–16]. This connects sustainable optimization to responsible decision-making: what matters is not only the predicted compromise but the defensibility of committing to it under uncertainty and stakeholder accountability [6, 14, 16-20].

The four core mechanisms of TOST

TOST explains sustainability trade-offs through four interacting mechanisms:

1. Objective incompatibility geometry (structural conflict among goals) [8–10]

2. Lifecycle constraint migration (feasibility reshaped across stages) [4–7]

3. Uncertainty-to-consequence coupling (risk emerges from uncertainty × stakes) [13–16]

4. Preference volatility (stakeholders and policies reshape priorities over time) [6, 7, 12]

Together, these mechanisms define trade-offs not as static compromise points but as governed decisions that must remain coherent across lifecycle boundaries and stakeholder frames. **Figure 2** operationalizes Trade-Off Sensitivity Theory (TOST) by showing how Pareto-efficient candidates must pass through objective semantics, trade-off sensitivity, and action admissibility filters before being interpreted as sustainability-legitimate design choices.

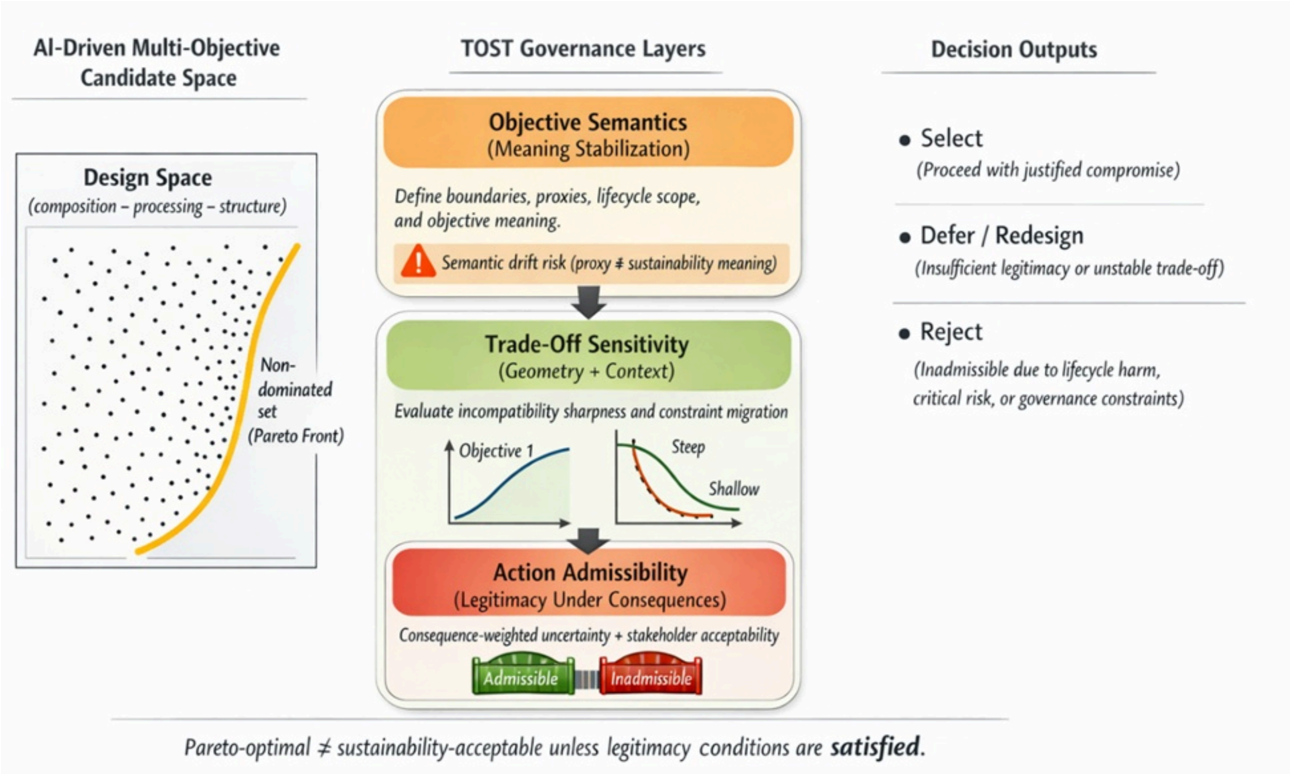

Figure 2. Trade-off sensitivity theory (TOST): from Pareto efficiency to sustainability legitimacy

Results and Discussion

Why “Pareto-Optimal” Is Not “Sustainability-Optimal”

A central implication of Trade-Off Sensitivity Theory (TOST**)** is that Pareto efficiency is an incomplete criterion for sustainable materials decisions. Pareto optimality certifies non-dominance with respect to a selected set of objectives. Still, it does not establish whether a compromise is acceptable across lifecycle boundaries, governance constraints, or stakeholder value commitments [1, 2]. In sustainable materials optimization, objectives are often heterogeneous (performance, embodied carbon, toxicity, recyclability, criticality), and the decision context frequently requires defensibility rather than mere selection [3, 4]. This is particularly relevant in AI-enabled workflows where rapid exploration can generate many plausible “winners,” creating an illusion of decisional clarity without a legitimate basis for commitment [5, 6, 21-25].

TOST therefore reframes optimization as trade-off justification rather than trade-off discovery. This framing aligns with the broader understanding that deploying AI in consequential decision contexts demands attention to robustness, accountability, and operational validity—not only predictive competence [7, 8]. In sustainable materials design, the question is not simply which point on the Pareto front is best, but which trade-off remains defensible when assumptions shift, constraints migrate, and harms cannot be undone [3, 4].

Objective Semantics: Why “Sustainability Objectives” Are Not Simple Scalars

TOST’s first governance layer—objective semantics—addresses a structural weakness common to sustainability optimization: sustainability goals are rarely stable single-property targets. Unlike mechanical strength or ionic conductivity, objectives such as “low carbon,” “circular,” or “non-toxic” often depend on system boundaries, impact category choices, allocation rules, and infrastructure assumptions [3, 9, 26-29]. Lifecycle assessment explicitly highlights this dependence: results can change when functional units, boundaries, electricity mixes, or end-of-life routes are altered [3, 9]. Consequently, AI systems can optimize sustainability proxies that are technically consistent yet semantically misaligned with the intended sustainability meaning [4, 10, 30-32].

This creates a theoretical risk of proxy sustainability, where optimization is “successful” by its own formal criteria but fragile to governance interpretation. Similar concerns appear in discussions of AI governance, where mis-specified targets and incentive-aligned metrics can yield outputs that are correct in a narrow sense but harmful in deployment [7, 8]. TOST frames this not as a computational problem but as a semantic precondition for legitimate optimization. Before any Pareto reasoning is meaningful, the optimization community must stabilize what the objectives mean and what counts as improvement [4, 10, 33, 34].

Trade-off sensitivity: Preference volatility and the fragility of compromises

Multi-objective optimization typically assumes that preferences exist (explicitly or implicitly) and that they can select among Pareto solutions by weights, utility functions, or decision rules [1, 2]. However, sustainability preferences are often volatile: they evolve with regulation, stakeholder pressure, supply shocks, corporate strategy, and risk tolerance [4, 11, 35]. Under such volatility, the apparent “best compromise” may be an artifact of unstable weights rather than a robust decision outcome.

TOST therefore introduces trade-off sensitivity as a new conceptual diagnostic: how easily does the preferred solution change under plausible preference perturbations? When sensitivity is high, the optimization is epistemically weak because small shifts in priority yield completely different “optimal” candidates, undermining decision stability [2, 4]. When sensitivity is low, compromise is robust and easier to justify. This maps to sustainability decision-making realities where long-horizon deployment requires choices that remain acceptable across changing contexts [3, 11]. In this view, the purpose of Pareto analysis shifts from selecting “the best” to identifying regions of the design space where preference-robust acceptability is achievable [1, 2].

Lifecycle constraint migration: Sustainability as a moving feasibility region

Another core implication of TOST is that feasibility boundaries in sustainable materials are not fixed. A candidate may be feasible in the laboratory but infeasible at scale due to energy requirements, yield constraints, safety limitations, or end-of-life incompatibilities [3, 9, 11]. Conversely, certain “unfavorable” options can become feasible when circular infrastructure improves or when manufacturing decarbonizes. This produces what TOST terms lifecycle constraint migration: the feasible region changes as the lifecycle stage changes.

LCA practice reinforces that sustainability evaluation depends on system parameters that can vary across deployment contexts and time [3, 9]. Therefore, optimization performed under one feasibility regime may no longer be meaningful under another. TOST uses this idea to argue that sustainable optimization must be judged not only by “current Pareto status,” but by how solutions behave under future constraints and realistic deployment boundaries [4, 11]. This reframing is essential for AI-guided sustainability decisions that must remain defensible beyond the training regime and beyond the short-term conditions under which the optimization was performed [5, 6].

Uncertainty-to-Consequence coupling: When uncertainty becomes decision-prohibitive

Materials AI increasingly recognizes the importance of uncertainty quantification, particularly when predictions guide downstream design and screening [12–15]. Yet uncertainty is not merely a statistical defect—it is decision-weighted. TOST formalizes this as uncertainty-to-consequence coupling: uncertainty becomes critical when paired with irreversible harms, regulatory non-compliance, or catastrophic failure modes. In high-stakes sustainability contexts, even moderate uncertainty can become decision-prohibitive if consequences are asymmetric or irreversible [7, 8].

Recent work evaluating uncertainty quantification for materials property prediction emphasizes that uncertainty estimates should inform trust and decision-making, not only model reporting [12, 15]. TOST extends this to sustainability: uncertainty must be interpreted through consequence categories (e.g., toxicity risk, emissions lock-in, supply criticality), rather than treated as a generic technical quantity [4, 11]. Thus, TOST provides a conceptual rule: a Pareto-leading solution is not action-licensed unless uncertainty is admissible under consequence weighting [7, 8, 12].

Practical use cases: What TOST changes (Without proposing methods)

Even as a non-algorithmic theory, Trade-Off Sensitivity Theory (TOST) reshapes how AI optimization outputs should be interpreted in sustainable materials workflows.

In **Use Case A — Circular design constraints**, TOST clarifies when circularity must be treated as an admissibility condition rather than a negotiable optimization objective. If a circularity threshold is defined by governance or regulatory mandate, it establishes a feasibility boundary rather than an objective to be compromised. Under such conditions, any attempt to trade circularity for short-term performance improvement violates the admissibility logic and renders the optimization outcome illegitimate [3, 9, 11].

In **Use Case B — Low-carbon optimization under boundary shifts**, TOST underscores the necessity of maintaining explicit **objective semantics**. The notion of “low carbon” must be anchored in stated life-cycle assessment (LCA) boundaries and system assumptions; otherwise, optimization results can be invalidated by mere redefinition of these boundaries, leading to semantic drift and the collapse of legitimacy [3, 9].

In **Use Case C — Substitution under criticality pressure**, TOST reconceptualizes **criticality** not merely as a secondary cost objective but as a dynamic governance constraint. Under geopolitical volatility or regulatory tightening, criticality can migrate from an optional trade-off variable to a dominant feasibility rule. TOST interprets this transition as a form of lifecycle-constraint migration interacting with preference volatility, revealing that the sustainability logic of material substitution itself must adapt under systemic stress [11].

Through these use cases, TOST demonstrates how conceptual sensitivity analysis can transform sustainable materials optimization from purely numerical problem-solving into a legitimacy-aware reasoning process that aligns AI outputs with the real conditions of sustainability governance.

Limitations and scope

TOST intentionally does not provide algorithms, performance benchmarks, or datasets. Its role is to define the interpretive governance layer required for sustainable trade-off decisions. One limitation is that TOST requires explicit articulation of objective meaning and consequence categories—an institutional challenge, as organizations often prefer to leave trade-offs implicit [4, 7]. Another limitation is that it does not resolve normative conflicts; it provides a language for locating where values enter and how they destabilize the legitimacy of optimization [4, 11]. This is appropriate because sustainability is irreducibly multi-value and cannot be “solved” purely as a technical optimization exercise [3, 4, 11].

Conclusion

This manuscript introduces Trade-Off Sensitivity Theory (TOST) as a novel theoretical framework for understanding multi-objective trade-offs in AI-enabled sustainable materials optimization. The central claim of the theory is that Pareto-optimality is not equivalent to sustainability acceptability. Sustainable decisions must remain legitimate amid variations in lifecycle boundaries, volatility in stakeholder preferences, and uncertainty weighted by the asymmetry of consequences.

TOST contributes three conceptual governance layers. The first is objective semantics, which seeks to stabilize what sustainability objectives actually mean and to prevent proxy optimization failures in which formal metrics drift away from the underlying societal or environmental intent. The second is trade-off sensitivity, which evaluates how fragile a proposed compromise is under plausible shifts in preferences, rather than assuming that preference weights are fixed or universally agreed upon. The third is action admissibility, which determines whether a Pareto-optimal solution is genuinely decision-licensed in the presence of uncertainty, risk asymmetry, and downstream consequences.

Taken together, these layers reposition multi-objective AI from a narrow exercise in compromise search toward a theory of defensible trade-off justification for sustainable materials design. Future conceptual work can build on TOST by developing standard reporting artifacts for objective semantics, sensitivity audits for Pareto-front decisions, and governance templates that formalize admissibility conditions across different materials classes and deployment regimes [3, 4, 7]. In this way, AI for sustainable materials can be evaluated not only by predictive performance, but also by the stability and legitimacy of the decisions it enables under real-world sustainability constraints.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 13 Jul 2023 Revised: 28 Sep 2023 Accepted: 27 Dec 2023
Published online: 18 July 2024

References

Deb K, Jain H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints. IEEE Trans Evol Comput. 2020;24(2):238–52.

Miettinen K, Ruiz F, Wierzbicki AP. Introduction to multi-objective optimization: interactive approaches. In: Multi-objective Optimization; 2020. p. 1–29.

Barbhuiya S, Barbhuiya S. Life Cycle Assessment of construction materials: A review. J Build Eng. 2023;74:106806.

Sala S, Benini L, Castellani V, Vidal-Legaz B. The need for integrated sustainability assessment in product and materials decision-making. Sustain Prod Consum. 2021;25:310–21.

Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science: Recent directions and sustainability-linked opportunities. npj Comput Mater. 2021;7:1–15.

Jablonka KM, Ongari D, Moosavi SM, Smit B. Big-data science in porous materials: machine learning and multi-objective design. Chem Rev. 2020;120(16):8066–129.

Dignum V. Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Springer; 2020.

Theodorou A, Dignum V. Towards ethical and socio-legal governance in AI. AI Soc. 2020;35:379–86.

Zhang P, Chen X, Shen L. Multi-objective optimization for sustainable design integrating LCA and decision-making. J Clean Prod. 2021;314:128095.

ISO. ISO 14040:2020 Environmental management—Life cycle assessment—Principles and framework. 2020.

Watari T, Nansai K, Nakajima K. Review of criticality assessment and implications for materials substitution and sustainability. Resour Conserv Recycl. 2020;162:105031.

Tavazza F, DeCost BL, Choudhary K. Uncertainty prediction for machine learning models of materials properties. ACS Omega. 2021;6(7):4374–82.

Musil F, Willatt MJ, Langovoy M, Ceriotti M. Fast and accurate uncertainty-aware machine learning for molecular and materials science. Chem Rev. 2021;121(16):9759–815.

Zhang Y, Ling C. A strategy to apply machine learning to small datasets in materials science. npj Comput Mater. 2020;6:25.

Chen C, Ye W, Zuo Y, Zheng C, Ong SP. Graph networks as a universal machine learning framework for molecules and crystals. Chem Mater. 2020;32(9):3564–75.

Moosavi SM, Jablonka KM, Ongari D, Smit B. Generative models for inverse design of materials. Nat Rev Mater. 2022;7:1–18.

Laakso J, Himanen L, Pouillon Y, Jäger MOJ, Foster AS. Updates to the DScribe library: New descriptors and derivatives. J Chem Phys. 2023;158(23):234802.

Himanen L, Jäger MOJ, Morooka EV, Federici Canova F, Ranawat YS, Gao DZ, et al. DScribe: Library of descriptors for machine learning in materials science. Comput Phys Commun. 2020;247:106949.

Shi B, Huang Y, Li X. Multi-objective optimization and its application in materials science: A review. Mater Genet Eng Adv. 2023;1:14.

Zhang P, Liu Y, Zhang J. Multi-objective optimization for materials design with improved NSGA-II. Mater Today Commun. 2021;29:102858.

Jain A, Ong SP, Hautier G, Chen W, Persson KA. Materials Project: accelerating materials design through open databases and AI (2020 update). npj Comput Mater. 2020;6:1–10.

Chen Y, Li Y, Zhang H. Multi-objective optimization method for sustainable material design using LCA and LCC. Sustainability. 2023;16(1):168.

Seifi M, Salem A, Satko D, Shao S. The role of sustainability metrics in materials and manufacturing decisions: A 2020–2023 synthesis. Mater Des. 2022;223:111119.

Himeur Y, Ghanem K, Alsalemi A, Bensaali F. Multi-objective decision-making under uncertainty for sustainability: A 2022 review. Renew Sustain Energy Rev. 2022;160:112287.

Allwood JM, Cullen JM. Sustainable Materials: With Both Eyes Open (updated 2020 edition). 2020.

European Commission Joint Research Centre. Product Environmental Footprint (PEF) method: 2021 update. 2021.

Hellweg S, Milà i Canals L. Emerging approaches to LCA under uncertainty and scenario dependence. Science. 2020;370(6517):eabc8747.

Guinée JB, Heijungs R, Huijbregts MAJ. Lifecycle impact assessment: recent methodological advances. Int J Life Cycle Assess. 2021;26:1–12.

Arvidsson R, Molander S. Prospective LCA for emerging technologies and materials: Pitfalls and principles. J Ind Ecol. 2020;24(4):820–32.

Marler RT, Arora JS. Survey of multi-objective optimization methods for engineering. Struct Multidisc Optim. 2020;62:1–18.

Keeney RL, Raiffa H. Decisions with Multiple Objectives: Preferences and Value Trade-offs (revisited 2021 commentary edition). 2021.

Roy B. Multicriteria decision aiding: Recent advances for sustainability decisions. Eur J Oper Res. 2020;283(2):409–23.

Floridi L, Cowls J. A unified framework of AI principles. Nat Mach Intell. 2020;2:370–3.

Raji ID, Smart A, White RN, Mitchell M, Gebru T, Hutchinson B, et al. Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. In: FAT* 2020 Proceedings; 2020. p. 33–44.

Tomašev N, Cornebise J, Hutter F, Mohamed S, Picciariello A, Connelly B, et al. AI for social good: Unlocking the opportunity for impact. Nat Commun. 2020;11:2468.