

ORIGINAL RESEARCH

Open access

# Perspective: From Prediction to Design — A Position on Optimization-Centric Workflows in Materials ML

Mohammed Al-Farsi<sup>1\*</sup>, Salim Al-Harthy<sup>1</sup>, Nasser Al-Rawahi<sup>2</sup>

## Abstract

Materials machine learning has achieved remarkable success in building accurate predictive models for properties such as formation energy, band gap, and mechanical strength. Yet the ultimate purpose of the field is not prediction but design: the discovery of new materials that deliver target properties under real-world constraints. This perspective argues that the community must shift from prediction-centric to optimization-centric workflows. Prediction-centric approaches focus on minimizing mean absolute error on randomly held-out test sets, while optimization-centric workflows aim to maximize (or minimize) a target property value within practical budgets, constraints, and sequential decision-making. These are fundamentally different objectives that demand different methods, evaluation metrics, and research priorities. A prediction-centric model may achieve low error on interpolation tasks yet fail catastrophically when asked to guide the search for extreme or out-of-distribution materials. In contrast, optimization-centric workflows treat the machine-learning model as a surrogate within an iterative loop that actively selects the next experiment or simulation. Four core principles underpin this shift: explicit goal specification that includes both the objective and all relevant constraints; robust constraint handling that distinguishes hard feasibility requirements from soft trade-offs; uncertainty awareness, where every prediction is accompanied by well-calibrated uncertainty estimates essential for balancing exploration and exploitation; and closed-loop integration that connects the model directly to automated or high-throughput experimentation and simulation. Adopting these principles will require new evaluation protocols that measure the best material discovered, sample efficiency, regret, constraint satisfaction, and extrapolation distance rather than isolated accuracy metrics. The implications extend to model development, benchmark design, journal standards, and funding priorities. Only by embracing optimization-centric workflows can materials machine learning fulfill its promise of accelerating discovery and delivering materials that solve pressing societal challenges.

**Keywords** Active learning, Bayesian optimization, Inverse design, Closed-loop discovery, Optimization-centric workflows, Self-driving labs

\*Correspondence:

Mohammed Al-Farsi  
[mohammed.alfarsi@gmail.com](mailto:mohammed.alfarsi@gmail.com)

<sup>1</sup> Department of Intelligent Materials Systems, Faculty of Engineering, Sultan Qaboos University, Muscat, Oman

<sup>2</sup> Department of Computational Materials Analytics, Faculty of Engineering, German University of Technology in Oman, Muscat, Oman

## Introduction

Materials machine learning stands at a crossroads. For more than a decade the community has invested heavily in developing ever-more-accurate predictive models, benchmarking them primarily on their ability to minimize mean absolute error or root-mean-square error on randomly partitioned test sets. These prediction-centric efforts have produced impressive results and valuable public datasets. Yet the stated motivation for most of this work has always been materials

design—the discovery of novel compounds or microstructures that achieve unprecedented properties for energy storage, catalysis, electronics, or sustainability.

Prediction and design, however, are not the same task. A model that excels at interpolation on a fixed test set may be useless—or even misleading—when the goal is to navigate a vast, high-dimensional search space toward an optimum that lies far from existing data. The community must therefore make a deliberate shift from prediction-centric to optimization-centric workflows. In an optimization-centric workflow the machine-learning model is not an end product but a decision-support tool embedded in an iterative process whose success is measured by the quality of the materials it helps discover within a finite experimental or computational budget [1, 2].

**Position Statement** Materials machine learning has focused on prediction accuracy: minimizing MAE on held-out test sets. But the ultimate goal of materials ML is design—discovering materials with target properties. Prediction and design are not the same. A model with low MAE on a test set may be useless for optimization, because optimization requires extrapolation, handling of constraints, and uncertainty awareness. The community must shift from prediction-centric to optimization-centric workflows, where the objective is not accuracy but finding the best material under constraints.

This position matters now because prediction benchmarks have largely saturated for many standard tasks, while the rate of new material discovery remains far below what society needs [3, 4]. Bayesian optimization, active learning, inverse design, and closed-loop self-driving laboratories have already demonstrated their ability to accelerate discovery in targeted domains [1, 2, 5]. Yet these optimization-centric methods remain under-represented in mainstream materials machine-learning literature, which continues to reward models primarily for low test-set error.

## Prediction-centric vs. Optimization-centric

The distinction between prediction-centric and optimization-centric workflows is not merely technical; it is philosophical. Each paradigm answers a different question, relies on fundamentally different data-generation logics, and evaluates success through incompatible criteria. A prediction-centric framework asks how accurately a model can reproduce known data, whereas an optimization-centric framework asks whether the model can guide the discovery of previously unknown, high-performing materials under real-world constraints.

A prediction-centric workflow begins with a fixed dataset that is randomly partitioned into training and test subsets. The model is trained to minimize average error on unseen data drawn from the same distribution, and its performance is evaluated using scalar metrics such as mean absolute error, root-mean-square error, or the coefficient of determination ( $R^2$ ). Uncertainty quantification is often treated as optional, and many models report only point estimates. Once the test error reaches an acceptable threshold, the model is considered successful, implicitly assuming that interpolation accuracy translates into practical utility. This assumption, however, embeds a critical limitation: the evaluation setting is static, distributionally aligned with the training data, and disconnected from the conditions under which materials discovery actually occurs.

An optimization-centric workflow reinterprets the same dataset as an initial condition rather than a complete representation of the problem space. The objective is not to approximate known values but to identify materials that maximize or minimize a target property—such as ionic conductivity, formation energy, or thermoelectric efficiency—while satisfying constraints related to stability, cost, or synthesizability. Evaluation is therefore outcome-driven and resource-aware, focusing on the best property value discovered within a finite budget, the number of iterations required to reach a predefined threshold, or improvements along a Pareto frontier in multi-objective settings. Data acquisition is inherently sequential and adaptive: each new experiment or simulation is selected by an acquisition function that balances predicted performance against uncertainty. As a result, the model is continuously updated, becoming locally accurate in precisely those regions of the design space that matter for achieving the optimization objective [1, 2, 6].

This contrast reveals a deeper structural divergence. Prediction-centric workflows assume that the data distribution is fixed and that generalization can be assessed through random sampling. Optimization-centric workflows instead operate under distributional shift by design, deliberately steering exploration toward regions that are poorly represented in the initial dataset but potentially rich in high-performing candidates. Uncertainty, which is peripheral in prediction settings, becomes central because it governs the exploration–exploitation trade-off. Similarly, extrapolation is not an error mode to be minimized but a necessary capability to be managed and exploited.

The divergence extends to the definition of success. In prediction-centric paradigms, success is equated with low aggregate error across a test set, regardless of whether the model identifies any materially superior candidates. In optimization-centric paradigms, success is defined by the quality of the best material discovered within a constrained number of evaluations, making efficiency and decision quality primary concerns. Consequently, a model that performs exceptionally well under prediction-centric metrics may still fail to contribute meaningfully to discovery if it cannot guide the search process toward optimal regions of the design space.

Recent work in Bayesian optimization and active learning demonstrates how this alternative paradigm operates in practice [1, 7–9]. By treating the model as a dynamic component within a sequential decision system, these approaches prioritize information gain and performance improvement over static accuracy. Closed-loop discovery platforms extend this logic further by integrating machine learning with automated experimentation, enabling continuous updating and rapid iteration [2, 5, 10, 11]. The persistent gap between prediction-centric success and real-world discovery outcomes can therefore be understood as a consequence of objective misalignment: the field has historically optimized for interpolation accuracy rather than for the identification of optimal materials under constraints.

## Why Prediction-Centric Metrics Fail for Design

The misalignment between prediction-centric metrics and design goals manifests in at least six concrete ways.

First, average error does not predict best-case performance. A model may achieve a low mean absolute error across the test distribution yet systematically underestimate the extreme values that designers care about most. In materials discovery the “best” material is often an outlier; an error of 0.05 eV/atom on average can translate into a 0.5 eV/atom underestimate of the true optimum, rendering the model useless for guiding synthesis decisions [3].

Second, random train-test splits test interpolation, whereas design requires extrapolation. When data are split randomly, test points are statistically similar to training points. In a real discovery campaign the most promising candidates lie far from the initial dataset. Models that perform well on random splits frequently collapse when asked to extrapolate, a failure mode documented across multiple inverse-design studies [12, 13].

Third, uncertainty is not optional for optimization. Bayesian optimization and active learning rely on well-calibrated uncertainty estimates to decide where to sample next. Most prediction-centric models provide only point predictions; adding uncertainty post-hoc is rarely as effective as training models that are explicitly designed to quantify it [1, 6, 14].

Fourth, constraint handling is ignored. Real materials design is never unconstrained. Cost, stability, toxicity, and synthesizability act as hard or soft filters. A prediction-centric model that ranks a hypothetical material highest may recommend something that is impossible to synthesize or too expensive to scale. Optimization-centric workflows must incorporate these constraints directly into the acquisition function [7, 15].

Fifth, sequential decision-making is not evaluated. Prediction-centric evaluation assumes a single, static test set. Design is a sequence of choices: which experiment to run next, which simulation to prioritize. The quality of those sequential decisions cannot be captured by a one-time accuracy score [2, 16].

Sixth, closed-loop integration is missing. Design workflows in practice integrate simulation, experiment, and machine learning in a continuous feedback loop. Static prediction models remain disconnected from this loop and therefore cannot adapt when new data reveal that the initial assumptions were incomplete [5, 10].

These six failures are not theoretical. They explain why many published high-accuracy models have not translated into accelerated discovery rates. The field has been measuring the wrong thing.

**Table 1** formalizes the structural misalignment between prediction-centric evaluation logic and the requirements of discovery-oriented optimization in materials machine learning.

**Table 1.** Structural Misalignment between Prediction-Centric Metrics and Discovery-Oriented Objectives in Materials Machine Learning

| Analytical Dimension      | Prediction-Centric Assumption     | Optimization-Centric Requirement                 | Resulting Failure Mechanism | Consequence for Materials Discovery                     |
|---------------------------|-----------------------------------|--------------------------------------------------|-----------------------------|---------------------------------------------------------|
| Objective Definition      | Global average error minimization | Extremal property optimization under constraints | Objective mismatch          | High-accuracy models fail to identify optimal materials |
| Data Regime               | IID random sampling               | Sequential, adaptive sampling                    | Distributional rigidity     | Poor performance in unexplored regions                  |
| Evaluation Logic          | Static test-set evaluation        | Iterative performance tracking                   | Temporal blindness          | No assessment of decision quality over time             |
| Uncertainty Treatment     | Optional or absent                | Central to decision-making                       | Overconfidence bias         | Inefficient or misleading exploration                   |
| Constraint Representation | Ignored or post hoc               | Integrated into acquisition                      | Feasibility blindness       | Selection of unrealizable candidates                    |
| Extrapolation Handling    | Avoided                           | Required and incentivized                        | Interpolation bias          | Failure to reach novel high-performing regions          |
| Success Criterion         | Low aggregate error               | Best feasible material discovered                | Metric misalignment         | Illusory progress without discovery gains               |

## Principles for Optimization-Centric Workflows

An optimization-centric workflow rests on seven interlocking principles that together reorient materials machine learning toward discovery.

**Table 2** consolidates the functional architecture of optimization-centric workflows, highlighting the interdependence of components required for efficient materials discovery.

**Table 2.** Integrated Architecture of Optimization-Centric Workflows: Functional Roles and Interdependencies

| Architectural Component | Functional Role | Dependency Structure | Failure if Isolated | Contribution to Discovery Efficiency |
|-------------------------|-----------------|----------------------|---------------------|--------------------------------------|
|                         |                 |                      |                     |                                      |

|                              |                                          |                                              |                                |                                             |
|------------------------------|------------------------------------------|----------------------------------------------|--------------------------------|---------------------------------------------|
| Goal Specification           | Defines objective and constraints        | Upstream of all components                   | Undefined search direction     | Aligns search with real-world targets       |
| Uncertainty Quantification   | Enables exploration–exploitation balance | Coupled to acquisition strategy              | Deterministic bias             | Reduces wasted evaluations                  |
| Constraint Handling          | Enforces feasibility and trade-offs      | Interacts with acquisition and evaluation    | Infeasible candidate selection | Ensures practical viability                 |
| Acquisition Strategy         | Selects next candidate                   | Depends on model + uncertainty + constraints | Random or greedy sampling      | Accelerates convergence to optima           |
| Sequential Adaptive Sampling | Updates knowledge iteratively            | Requires feedback from evaluation            | Static model stagnation        | Improves local accuracy in relevant regions |
| Closed-Loop Integration      | Connects model to experiment/simulation  | System-level integration                     | Disconnected prediction        | Enables autonomous discovery                |
| Extrapolation Awareness      | Drives search beyond known data          | Linked to uncertainty and distance metrics   | Local optimum trapping         | Enables discovery of novel materials        |
| Multi-objective Optimization | Handles competing objectives             | Extends acquisition and evaluation           | Single-objective bias          | Expands Pareto-optimal solutions            |

Effective materials discovery through machine learning demands an explicit formulation of the optimization objective—whether maximizing ionic conductivity, minimizing band gap, or optimizing a composite figure of merit—alongside all relevant constraints such as cost thresholds or stability criteria, defined at the outset together with clear success metrics including iteration limits.

This precision in goal specification must be paired with rigorous uncertainty quantification, whereby every prediction carries a well-calibrated estimate that distinguishes epistemic from aleatoric components, thereby enabling acquisition functions such as expected improvement or upper confidence bound to balance exploration and exploitation meaningfully [1, 14, 15].

A related implication arises in constraint handling, which requires integrating both hard feasibility requirements and soft trade-offs directly into the acquisition process through penalty mechanisms or constrained Bayesian optimization, ensuring navigation of viable regions within complex materials spaces [7, 15].

Such integration naturally extends into sequential adaptive sampling, in which active learning or Bayesian optimization intelligently selects each subsequent candidate, with the model retrained after every new measurement to progressively refine its guidance [2, 6, 16].

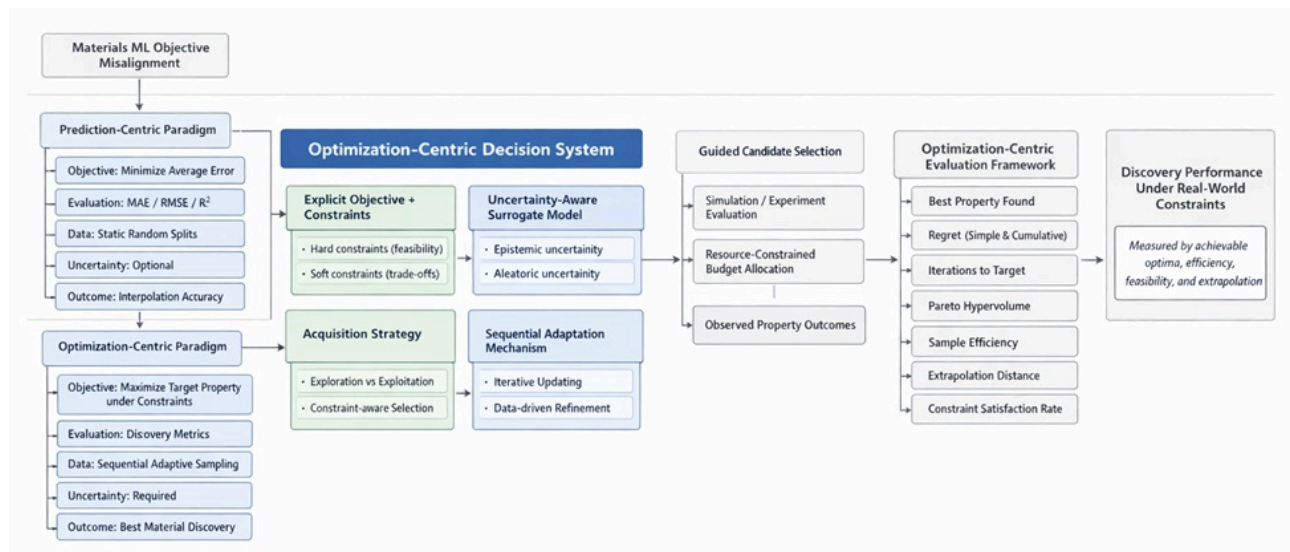
This adaptive process reaches its full potential only through closed-loop integration, whereby the predictive model interfaces directly with automated experimentation or high-throughput simulation platforms, creating a seamless predict–suggest–measure–update cycle that removes human latency and enables genuine autonomous discovery, as realized in emerging self-driving laboratory systems [5, 10, 17].

Beyond immediate operational concerns, effective extrapolation remains essential, since promising materials frequently reside outside the initial training distribution; acquisition functions must therefore incorporate deliberate bias toward high-uncertainty regions, augmented by explicit metrics of distance to the training manifold [12, 13].

Under these conditions, multi-objective optimization further demands maintenance of an evolving Pareto frontier approximation, with performance tracked through hypervolume improvement rather than isolated single-objective gains [7, 15].

When pursued in concert, these interlocking elements elevate materials machine learning from passive modeling to a robust engine of active scientific discovery [1, 2].

**Figure 1** illustrates that optimization-centric materials machine learning constitutes a structured sequential decision architecture in which objectives, uncertainty, constraints, and adaptive sampling jointly transform prediction into constrained discovery outcomes.



**Figure 1.** Optimization-Centric Materials Machine Learning as a Sequential Decision Architecture for Constrained Discovery

## Evaluation Metrics for Optimization-Centric Workflows

Adopting optimization-centric workflows in materials discovery demands evaluation frameworks that prioritize genuine advancement over interpolation fidelity. Central to this reorientation is the need to track the highest (or lowest) target property value attained within a fixed evaluation budget, visualized through learning curves that reveal progressive improvement [1, 2].

A related implication arises in the quantification of regret, where simple regret captures the gap between the best discovered material and the known global optimum, while cumulative regret aggregates this shortfall across iterations, both serving as direct indicators of optimization shortfall that must be driven toward zero [1, 16].

This shift also introduces practical considerations for resource-constrained settings, such as measuring the number of evaluations required to surpass a critical performance threshold, thereby exposing the true efficiency of proposed methods under realistic laboratory conditions [2, 10].

Beyond these single-objective concerns, multi-objective scenarios necessitate assessment of Pareto hypervolume, which quantifies coverage of the trade-off surface relative to a chosen reference and thereby reflects the workflow's capacity to navigate complex design spaces [15].

Sample efficiency further sharpens this evaluation by expressing the gain in best-found property per experimental evaluation, offering a transparent metric of return on investment that underscores the economic realities of discovery campaigns [1, 6].

Under these conditions, meaningful extrapolation must be verified by computing the feature-space or embedding-space distance between the leading candidate and the nearest training-set example, ensuring that reported successes extend beyond mere recall of known patterns [12, 13].

Finally, constraint satisfaction rates illuminate how effectively the optimization process internalizes feasibility boundaries, with higher fractions signaling robust navigation of the admissible region rather than reliance on post-hoc filtering [7, 15].

Optimization-focused contributions should therefore foreground these indicators in place of standalone prediction errors such as MAE, while benchmark suites must incorporate known global optima to enable rigorous regret computation. Random search, grid search, and expert heuristics remain essential baselines, without which claims of superiority lack substance. Only through such standards can the field establish credible measures of progress toward accelerated, reliable materials discovery [3, 16].

## Objections and Responses

Any call for a paradigm shift inevitably provokes resistance, yet sustained progress requires confronting the most persistent objections head-on.

Skeptics often maintain that high prediction accuracy remains indispensable for effective optimization. While undeniably necessary, accuracy alone proves insufficient: a model achieving low mean absolute error on a random test split may still falter in guiding discovery, precisely because optimization hinges on extrapolation beyond the training distribution, explicit constraint navigation, and well-calibrated uncertainty estimates [1, 3]. In practice, strong interpolation performance offers no assurance that the model will surface superior materials within realistic evaluation budgets.

A closely related misconception equates optimization with prediction on an alternative test set. This view overlooks a fundamental distinction. In conventional prediction tasks the test distribution stays fixed and independent of the model; in optimization-centric workflows, however, the model actively shapes the sequence of evaluated points, rendering the “test set” dynamic and co-constructed through sequential decision-making [2]. What appears as mere evaluation thus becomes the core challenge of adaptive, model-driven exploration.

Beyond these conceptual clarifications, practical concerns surface regarding scalability, particularly the claim that Bayesian optimization proves too slow for high-dimensional materials spaces. Naïve implementations can indeed struggle, yet established remedies—trust-region strategies, random embeddings, and multi-fidelity surrogates—have already demonstrated effective scaling to representative materials problems [7, 18]. The true bottleneck lies less in algorithmic feasibility than in widespread adoption of these mature techniques.

Uncertainty about the true global optimum presents another frequent objection. Yet empirical benchmarks grounded in well-characterized material families already permit rigorous computation of regret and sample efficiency, while comparisons against random search, grid search, or expert heuristics furnish reproducible baselines even when analytic optima remain unavailable [10, 17]. Such practices anchor evaluation in observable performance rather than unattainable perfection.

Finally, the assertion that certain models serve prediction rather than design carries weight only when claims remain appropriately scoped. Research explicitly positioned as advancing materials discovery cannot evade optimization-centric scrutiny; while prediction-centric contributions retain value in supplying reliable surrogates, the community must resist conflating low test-set error with substantive design success [6, 19].

Taken together, these considerations reveal the proposed reorientation not as a dismissal of prediction but as a necessary realignment in which predictive power is subordinated to—and ultimately judged by—its capacity to accelerate genuine discovery.

## Relation to Other Positions

This perspective builds upon and complements several existing positions in the literature while sharpening the focus on optimization as the central objective.

It aligns closely with calls for physics-constrained machine learning. Physics-driven frameworks such as PAL 2.0 embed domain knowledge directly into Bayesian optimization loops [8], demonstrating that prior physical insight accelerates discovery when paired with explicit optimization objectives. The present position extends that idea by arguing that physics constraints must be integrated not only into the model but also into the acquisition function and evaluation protocol.

It also resonates with positions advocating uncertainty separation. Bayesian optimization studies have shown that distinguishing epistemic from aleatoric uncertainty enables more effective exploration–exploitation trade-offs [1, 14]. The current argument makes uncertainty quantification mandatory rather than optional, because acquisition functions such as expected improvement rely on it to drive sequential sampling.

Active learning is a core component rather than a separate paradigm. Earlier work on adaptive sampling using uncertainties for targeted design has laid essential groundwork [6, 20]. This perspective incorporates active learning as the engine of sequential adaptive sampling within a broader optimization-centric workflow that also includes constraint handling and closed-loop integration.

Finally, the position relates directly to generative models for inverse design. Machine learning-based inverse design methods and generative deep neural networks for backpropagation and active learning already operate in an optimization-centric mode [12, 21–23]. The present framework supplies the missing evaluation principles—regret, hypervolume, extrapolation distance—that allow the community to compare generative approaches on discovery performance rather than reconstruction error alone.

By synthesizing these related positions under a unified optimization-centric banner, the field can move from fragmented methodological advances toward a coherent discovery methodology.

## Recommendations for the Community

To realize the shift to optimization-centric workflows, coordinated action is required across stakeholders.

For researchers: replace or supplement prediction MAE with optimization metrics. Every paper claiming design relevance should report best property found, regret curves, sample efficiency, and constraint satisfaction rate. Adopt Bayesian optimization or active learning baselines as standard comparators [1, 15, 16]. When publishing surrogate models, release them with uncertainty quantification and ready-to-use acquisition functions so others can embed them in closed-loop campaigns.

For benchmark designers: create and maintain optimization-centric benchmarks that include known or best-available optima, realistic constraints, and multi-objective tasks. Provide open-source implementations of random search, evolutionary algorithms, and expert-rule heuristics so that improvements can be measured against strong baselines. Public leaderboards should rank methods by regret or hypervolume improvement rather than test-set accuracy.

For journal editors and reviewers: require optimization-centric metrics for any manuscript that claims to advance materials design or discovery. Manuscripts that report only prediction accuracy on random splits while asserting design impact should be redirected or revised. Encourage reviewers to ask explicitly: “Does this work demonstrate faster discovery of better materials under constraints?”

For funding agencies: prioritize proposals that integrate optimization-centric methods and closed-loop infrastructure. Fund the development of self-driving laboratory platforms and the creation of shared optimization benchmarks [5, 24, 25]. Support interdisciplinary teams that combine machine learning, robotics, and domain expertise, recognizing that discovery is a systems-level challenge rather than an isolated modeling task.

Collective adoption of these recommendations will accelerate the transition from prediction to design and ensure that resources are directed toward research that measurably advances materials discovery.

## Toward Closed-Loop Materials Discovery

The ultimate expression of optimization-centric workflows is the self-driving laboratory [26]. In this vision, machine-learning models suggest candidates, robotic platforms execute experiments or high-throughput simulations, and the resulting data stream back to update the surrogate model in real time. The loop—predict, suggest, measure, update—runs autonomously, enabling discovery at speeds unattainable by human-led campaigns.

Recent demonstrations of closed-loop optimization have already validated the concept [2, 5, 10, 17]. Platforms for nanoparticle synthesis, organic emitter discovery, and automated materials experimentation show that the integration of Bayesian optimization or active learning with robotics can reduce the number of experiments needed by orders of magnitude [10, 17, 27].

A practical roadmap can be outlined. In the short term (1–2 years), the community should standardize optimization-centric evaluation protocols and benchmark Bayesian optimization across common materials problems [1, 3]. In the medium term (2–5 years), integrate multi-objective and constraint-handling capabilities while connecting models to existing high-throughput experimental facilities. In the long term (5–10 years), deploy fully autonomous self-driving laboratories capable of end-to-end discovery without human intervention [28, 29].

Success will be measured by concrete outcomes: the discovery of a new material meeting a target property specification using fewer than 100 experiments, with full documentation of the optimization trajectory. When this threshold is routinely achieved, materials machine learning will have fulfilled its original promise.

## Conclusion

Materials machine learning must shift from prediction-centric to optimization-centric workflows. Prediction-centric metrics such as MAE on random splits do not predict design success. Optimization requires explicit goal specification, uncertainty quantification, constraint handling, sequential adaptive sampling, closed-loop integration, extrapolation awareness, and multi-objective capability. The proposed evaluation metrics—best property found, regret, iterations to target, Pareto hypervolume, sample efficiency, extrapolation distance, and constraint satisfaction rate—provide the necessary yardsticks.

The community now faces a clear choice. It can continue to optimize for interpolation accuracy on static benchmarks, or it can reorient toward the discovery of new materials that solve real problems under real constraints. The position outlined here is not a dismissal of past progress but a call to align methods, metrics, and priorities with the field’s ultimate purpose.

Adopting optimization-centric workflows will demand changes in how we train models, design benchmarks, review papers, and allocate funding. The reward is a materials science that delivers on its long-standing promise: faster,

cheaper, and more reliable discovery of the materials our society needs. The time to make this shift is now.

## Acknowledgements

None

## Conflict of interest

None

## Financial support

None

## Ethics statement

None

Received: 02 Jun 2024

Revised: 12 Sep 2024

Accepted: 05 Dec 2024

Published online: 18 January 2025

## Rights and permissions

**Open Access** The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

## References

Zuo Y, Qin M, Chen C, Ye W, Li X, Luo J, et al. Accelerating materials discovery with Bayesian optimization and graph deep learning. *Mater Today*. 2021;51:126-35.  
<https://doi.org/10.1016/j.mattod.2021.08.012>.

Kusne AG, Yu H, Wu C, Zhang H, Hattrick-Simpers J, DeCost B, et al. On-the-fly closed-loop materials discovery via Bayesian active learning. *Nat Commun*. 2020;11(1):5966.  
<https://doi.org/10.1038/s41467-020-19597-w>.

Cinquin T, Lo S, Strieth-Kalthoff F, Aspuru-Guzik A, Pleiss G, Bamler R, et al. What actually matters for materials discovery: Pitfalls and recommendations in Bayesian optimization. In: *AI for Accelerated Materials Design - ICLR 2025*; 2025.

Liu Y, Cui Y, Zhou H, Lei S, Yuan H, Shen T, et al. Machine learning-based methods for materials inverse design: A review. *Comput Mater Contin*. 2025;82(2):1463-92.  
<https://doi.org/10.32604/cmc.2025.060109>.

Abolhasani M, Kumacheva E. The rise of self-driving labs in chemical and materials sciences. *Nat Synth.* 2023;2(6):483-92.  
<https://doi.org/10.1038/s44160-022-00231-0>.

Lookman T, Balachandran PV, Xue D, Yuan R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *NPJ Comput Mater.* 2019;5(1):21.  
<https://doi.org/10.1038/s41524-019-0153-8>.

Diwale S, Eisner MK, Carpenter C, Sun W, Rutledge GC, Braatz RD. Bayesian optimization for material discovery processes with noise. *Mol Syst Des Eng.* 2022;7(6):622-36.  
<https://doi.org/10.1039/D1ME00154J>.

Priyadarshini MS, Romiluyi O, Wang Y, Miskin K, Ganley C, Clancy P. PAL 2.0: A physics-driven Bayesian optimization framework for material discovery. *Mater Horiz.* 2024;11(3):781-91.  
<https://doi.org/10.1039/D3MH01474F>.

Chen X, Zhou H, Li Y. Effective design space exploration of gradient nanostructured materials using active learning based surrogate models. *Mater Des.* 2019;183:108085.  
<https://doi.org/10.1016/j.matdes.2019.108085>.

Strieth-Kalthoff F, Hao H, Rathore V, Derasp J, Gaudin T, Angello NH, et al. Delocalized, asynchronous, closed-loop discovery of organic laser emitters. *Science.* 2024;384(6697):eadk9227.  
<https://doi.org/10.1126/science.adk9227>.

Angello NH, Friday DM, Hwang C, Yi S, Cheng AH, Torres-Flores TC, et al. Closed-loop transfer enables artificial intelligence to yield chemical knowledge. *Nature.* 2024;633(8029):351-8.  
<https://doi.org/10.1038/s41586-024-07892-1>.

Lee J, Park D, Lee M, Lee H, Park K, Lee I, et al. Machine learning-based inverse design methods considering data characteristics and design space size in materials design and manufacturing: A review. *Mater Horiz.* 2023;10(12):5436-56.  
<https://doi.org/10.1039/D3MH00039G>.

Ha CS, Yao D, Xu Z, Liu C, Liu H, Elkins D, et al. Rapid inverse design of metamaterials based on prescribed mechanical behavior through machine learning. *Nat Commun.* 2023;14(1):5765.  
<https://doi.org/10.1038/s41467-023-40854-1>.

Herbol HC, Poloczek M, Clancy P. Cost-effective materials discovery: Bayesian optimization across multiple information sources. *Mater Horiz.* 2020;7(8):2113-23.  
<https://doi.org/10.1039/D0MH00062K>.

Khatamsaz D, Vela B, Singh P, Johnson DD, Allaire D, Arróyave R. Bayesian optimization with active learning of design constraints using an entropy-based approach. *NPJ Comput Mater.* 2023;9(1):49.  
<https://doi.org/10.1038/s41524-023-01006-7>.

Wang A, Liang H, McDannald A, Takeuchi I, Kusne AG. Benchmarking active learning strategies for materials optimization and discovery. *Oxf Open Mater Sci.* 2022;2(1):itac006.  
<https://doi.org/10.1093/oxfmat/itac006>.

Park J, Kim YM, Hong S, Han B, Nam KT, Jung Y. Closed-loop optimization of nanoparticle synthesis enabled by robotics and machine learning. *Matter.* 2023;6(3):677-90.  
<https://doi.org/10.1016/j.matt.2023.01.018>.

Arróyave R, Khatamsaz D, Vela B, Couperthwaite R, Molkeri A, Singh P, et al. A perspective on Bayesian methods applied to materials discovery and design. *MRS Commun.* 2022;12(6):1037-49.  
<https://doi.org/10.1557/s43579-022-00288-0>.

Wang J, Wang Y, Chen Y. Inverse design of materials by machine learning. *Materials (Basel)*. 2022;15(5):1811.  
<https://doi.org/10.3390/ma15051811>.

Bassman Oftelie L, Rajak P, Kalia RK, Nakano A, Sha F, Sun J, et al. Active learning for accelerated design of layered materials. *NPJ Comput Mater*. 2018;4(1):74.  
<https://doi.org/10.1038/s41524-018-0129-0>.

Abu-Mualla M, Huang J. Inverse design of 3D cellular materials with physics-guided machine learning. *Mater Des*. 2023;232:112103.  
<https://doi.org/10.1016/j.matdes.2023.112103>.

Liu Z, Zhu D, Raju L, Cai W. Tackling photonic inverse design with machine learning. *Adv Sci (Weinh)*. 2021;8(5):2002923.  
<https://doi.org/10.1002/advs.202002923>.

Chen CT, Gu GX. Generative deep neural networks for inverse materials design using backpropagation and active learning. *Adv Sci (Weinh)*. 2020;7(5):1902607.  
<https://doi.org/10.1002/advs.201902607>.

Seifrid M, Pollice R, Aguilar-Granda A, Chan ZM, Hotta K, Ser CT, et al. Autonomous chemical experiments: Challenges and perspectives on establishing a self-driving lab. *Acc Chem Res*. 2022;55(17):2454-66.  
<https://doi.org/10.1021/acs.accounts.2c00220>.

Baird SG, Sparks TD. What is a minimal working example for a self-driving laboratory? *Matter*. 2022;5(12):4170-8.  
<https://doi.org/10.1016/j.matt.2022.11.007>.

Soldatov MA, Butova VV, Pashkov D, Butakova MA, Medvedev PV, Chernov AV, et al. Self-driving laboratories for development of new functional materials and optimizing known reactions. *Nanomaterials (Basel)*. 2021;11(3):619.  
<https://doi.org/10.3390/nano11030619>.

Zaki M, Prinz C, Ruehle B. A self-driving lab for nano- and advanced materials synthesis. *ACS Nano*. 2025;19(9):9029-41.  
<https://doi.org/10.1021/acsnano.4c17504>.

Tamura R, Tsuda K, Matsuda S. NIMS-OS: An automation software to implement a closed loop between artificial intelligence and robotic experiments in materials science. *Sci Technol Adv Mater Methods*. 2023;3(1):2232297.  
<https://doi.org/10.1080/27660400.2023.2232297>.

Lo S, Baird SG, Schrier J, Blaiszik B, Carson N, Foster I, et al. Review of low-cost self-driving laboratories in chemistry and materials science: The “frugal twin” concept. *Digit Discov*. 2024;3(5):842-68.  
<https://doi.org/10.1039/D3DD00223C>.