ORIGINAL RESEARCH **Open access**

Algorithmic Confidence as a Control Signal in Materials Research

Daniel Brooks[1*], Amelia Carter[2], Ethan Moore[1]

Abstract

Materials research increasingly relies on machine learning to accelerate property prediction and discovery, yet the trustworthiness of these models remains constrained by their inability to express epistemic limitations. Algorithmic confidence—embodied in principled uncertainty quantification—provides a quantitative measure of model reliability that can extend beyond diagnostic assessment to serve as an active control signal within the research process. This conceptual manuscript synthesizes recent developments in uncertainty-aware machine learning, Bayesian approaches, and adaptive sampling strategies to argue that confidence estimates hold untapped potential as dynamic regulators of investigative workflows. Rather than treating uncertainty solely as a performance metric or sampling criterion, we conceptualize it as a central control variable that modulates decision pathways, balances exploration and exploitation, and informs the transition from computational prediction to empirical validation. A novel framework is proposed wherein algorithmic confidence governs iterative cycles in materials inquiry, enabling self-regulating mechanisms that align model assertions with epistemic boundaries. This perspective reframes uncertainty not as a limitation but as a strategic operator capable of guiding resource-efficient, robust materials exploration in a purely conceptual sense. By elevating confidence to a control role, the approach seeks to foster more deliberate and principled integration of computational intelligence into materials science paradigms.

Keywords Materials discovery, Uncertainty quantification, Algorithmic confidence, Machine learning, Active learning, Bayesian optimization

*Correspondence:
Daniel Brooks
daniel.brooks@outlook.com
[1] Department of Materials Informatics, Faculty of Engineering, University of Manchester, Manchester, United Kingdom
[2] Department of Artificial Intelligence Systems, Faculty of Computer Science, University of Birmingham, Birmingham, United Kingdom

Introduction

Materials science confronts enduring structural challenges as it pursues compounds and systems capable of meeting escalating technological, environmental, and societal demands. The search for materials with targeted functionalities—ranging from high-capacity energy storage media and quantum-efficient semiconductors to corrosion-resistant alloys and carbon-neutral catalysts—requires navigation across vast combinatorial chemical and structural spaces. These spaces are defined not only by compositional permutations but also by crystallographic arrangements, defect landscapes, microstructural heterogeneity, and processing histories. Traditional experimental methodologies, while foundational to scientific progress, remain constrained by throughput limitations, cost intensiveness, and temporal inertia. Iterative synthesis–characterization cycles often depend on heuristic guidance, domain intuition, or serendipitous discovery pathways, rendering systematic exploration of high-dimensional materials landscapes impractical at scale. Consequently, the pace of materials innovation has historically lagged behind application-driven demand, particularly in domains linked to sustainability transitions and next-generation electronics.

The integration of machine learning into materials research has initiated a paradigmatic shift in how these search spaces are navigated. By learning statistical mappings between descriptors—such as stoichiometry, crystal graphs, or electronic structure features—and target properties, machine learning models enable accelerated virtual screening of candidate materials [1-3]. Predictive architectures, including graph neural networks, kernel methods, and deep representation learners, have demonstrated success across a spectrum of property domains. Formation energies, band gaps, elastic moduli, diffusion barriers, and thermodynamic stability metrics can now be estimated at scales previously unattainable through purely ab initio or experimental routes [4, 5]. These advances have repositioned computation from a supportive analytical tool to a generative engine of hypothesis formation within materials design pipelines.

Yet, this acceleration introduces epistemic tensions. Machine learning systems in materials science frequently operate as opaque inferential engines, producing deterministic point predictions without articulating the reliability contours surrounding those estimates. Such black-box characteristics complicate interpretive judgment, particularly when models are deployed in extrapolative regimes—regions of chemical or structural space that are insufficiently represented in the training data. Overconfidence in these predictions poses nontrivial risks: experimental resources may be allocated to spurious candidates, theoretical insights may be misdirected, and discovery pathways may become anchored to algorithmic artifacts rather than to physically grounded plausibility [6, 7]. The high dimensionality and sparsity of materials datasets amplify these concerns, rendering conventional validation heuristics insufficient for gauging predictive trustworthiness.

Uncertainty quantification has emerged as a methodological response to these limitations, introducing probabilistic framing into materials-machine-learning outputs. Within this paradigm, uncertainty is typically decomposed into epistemic components—reflecting model ignorance due to limited knowledge—and aleatoric components—capturing intrinsic variability or noise in data-generating processes [8, 9]. A diverse methodological ecosystem has evolved to operationalize these constructs. Bayesian neural networks embed posterior distributions over model parameters, ensemble approaches approximate predictive variance through model plurality, and Monte Carlo dropout techniques emulate Bayesian inference within deterministic deep learning architectures [10-12]. Calibration frameworks further refine confidence estimates, aligning predicted probabilities with empirical error distributions.

Beyond passive reliability assessment, uncertainty has been mobilized as an active driver of learning efficiency. Active learning frameworks exploit epistemic uncertainty to prioritize data acquisition, selecting candidate materials whose evaluation promises maximal information gain [13, 14]. In parallel, Bayesian optimization leverages uncertainty-aware surrogate models to balance exploitation of high-performing regions with exploration of under-sampled domains in materials discovery landscapes [15, 16]. Through acquisition functions such as expected improvement or upper confidence bounds, these systems formalize decision-making under uncertainty, transforming predictive variance into a navigational compass for experimental and computational campaigns.

Despite these advances, prevailing implementations tend to localize uncertainty within discrete methodological modules —query selection stages, surrogate modeling layers, or calibration routines—rather than embedding it as an organizing principle across the full materials research lifecycle. Confidence estimates inform where to sample or which prediction to trust, but they seldom regulate how entire workflows transition between computational inference, human interpretation, experimental validation, and iterative refinement. This compartmentalization limits the systemic potential of uncertainty as a coordinating signal within discovery ecosystems.

This manuscript advances the proposition that algorithmic confidence can be reconceptualized as a holistic control signal governing materials research processes. Rather than functioning solely as an evaluative afterthought, confidence is framed here as a dynamic modulator that can steer epistemic workflows. Within this perspective, gradients and thresholds of predictive confidence regulate transitions between analytical states: high-confidence regimes may justify autonomous computational prioritization, whereas low-confidence zones may trigger human oversight, targeted experimentation, or representational augmentation. Such regulation does not prescribe specific algorithmic

implementations; instead, it establishes a conceptual architecture through which uncertainty becomes infrastructural to decision orchestration [17-19].

The conceptual implications of this reframing are multifold. First, it decouples the epistemic utility of machine learning from absolute predictive accuracy, foregrounding relational reliability within evolving knowledge conditions. Models need not be universally precise to be scientifically valuable; rather, their confidence landscapes determine where they can responsibly inform action. Second, this approach enhances systemic efficiency by directing investigative resources toward regions where confidence differentials are epistemically meaningful—mitigating both over-exploration of well-understood zones and over-exploitation of uncertain predictions. Third, it aligns materials informatics with broader trajectories in trustworthy artificial intelligence, where transparency, controllability, and calibrated decision support are prioritized alongside performance metrics [20-22].

By elevating algorithmic confidence from a localized statistical descriptor to a cross-workflow regulatory logic, this work contributes a unifying conceptual lens for uncertainty-aware materials discovery. The sections that follow synthesize theoretical foundations spanning uncertainty modeling, active learning dynamics, and epistemic governance in AI-enabled science. Building on this synthesis, a structured conceptual framework is introduced to articulate how confidence gradients can orchestrate transitions among prediction, exploration, validation, and refinement phases. Through this integrative perspective, the manuscript aims to expand interpretive and design vocabularies for next-generation materials AI systems, positioning confidence not merely as a measure of doubt, but as an operational infrastructure for scientific navigation. **Table 1** synthesizes the functional roles of algorithmic confidence as a control signal across investigative regimes, delineating associated actions, epistemic interpretations, and systems-level implications.

Table 1. Algorithmic confidence as a control signal across materials research workflows

Confidence regime/signal state	Control function in workflow	Typical computational and experimental actions	Epistemic interpretation	Systems-level implications for materials discovery
High confidence	Authorizes exploitation and downstream utilization of predictions	Hypothesis advancement; design recommendation; candidate prioritization; resource allocation toward synthesis or deployment	Predictive maturity with strong evidential grounding; model–materials alignment considered reliable within current knowledge bounds	Accelerates translational movement from computation to application; reduces redundant validation; supports efficiency in screening pipelines
Intermediate confidence	Triggers a hybrid regulatory response balancing utilization and verification	Targeted experimental validation; sensitivity analysis; cross-model comparison; localized uncertainty reduction	Partial evidential support; knowledge suggestive but incomplete; warrants cautious interpretive engagement	Prevents premature commitment; sustains investigational flexibility; refines predictive reliability without overextending resources
Low confidence	Activates exploratory control directives	Data acquisition; experimental sampling; descriptor expansion; compositional space exploration; model retraining	Epistemic insufficiency; knowledge gaps or representational sparsity dominate the predictive state	Directs discovery toward underexplored regions; transforms uncertainty into a generative search signal; expands knowledge boundaries
Confidence threshold	Governs discrete workflow state	Shifts between exploration ↔	Nonlinear change in evidential status; signals	Enables adaptive research staging;

crossings	transitions	exploitation modes; reallocation of computational or laboratory effort	regime shift in model reliability	prevents static pipeline behavior
Confidence gradients (temporal)	Modulates iterative learning trajectories	Monitoring convergence across training cycles; guiding active learning iterations	Indicates epistemic stabilization or divergence over time	Supports anticipatory steering; detects knowledge saturation or instability
Confidence gradients (spatial)	Maps reliability across materials design space	Identification of high-certainty clusters vs volatile prediction zones	Reveals heterogeneity in model generalization across chemical/structural domains	Guides region-specific exploration strategies; prevents blind extrapolation
Confidence feedback signals	Recalibrates predictive and control systems	Model updating; retraining; recalibration; descriptor refinement	Knowledge evolves through recursive evidence integration	Establishes closed-loop epistemic governance; sustains adaptive learning ecosystems
Confidence as epistemic currency	Conditions for decision legitimacy	Aligns action proportionality with evidential strength	Reliability prioritized over nominal predictive accuracy	Reorients AI utility toward trustworthiness and interpretive rigor
Confidence as a self-regulatory mechanism	Introduces closed-loop control into discovery pipelines	Continuous adjustment of sampling, prediction use, and validation intensity	Computational inquiry becomes self-monitoring	Reduces premature convergence and wasteful exploration
Confidence hierarchy of action	Orders investigative priorities	High → utilization; Intermediate → refinement; Low → expansion	Structured response to varying knowledge states	Optimizes allocation of experimental and computational resources

Theoretical Background and Literature Synthesis

Machine learning applications in materials property prediction: Machine learning has permeated materials science through supervised models that predict properties from descriptors or atomic configurations [23, 24]. Interatomic potentials learned via neural networks or Gaussian processes have enabled large-scale simulations previously inaccessible [25, 26]. Graph-based architectures and equivariant networks have further improved accuracy for periodic systems [27, 28]. These developments underscore the capacity of data-driven methods to approximate complex quantum-mechanical or empirical relationships [29, 30]. Yet, predictive success often masks underlying variability in generalization, particularly in sparse or noisy regimes characteristic of materials datasets [31, 32].

Uncertainty quantification in materials-relevant machine learning: Uncertainty quantification has emerged as essential for assessing model reliability in materials contexts [1, 8, 18]. Bayesian neural networks offer principled posterior distributions over weights, yielding predictive uncertainty [10, 11]. Ensemble techniques approximate Bayesian inference through model diversity, producing variance estimates that correlate with error [4, 12]. Monte Carlo dropout provides computationally efficient uncertainty via stochastic inference [22, 23]. In materials-specific adaptations, these methods have been applied to force fields, property predictors, and generative models [5, 6, 13]. Calibration of

uncertainty—ensuring that reported confidence intervals align with observed error rates—has received attention, highlighting discrepancies in overconfident regimes [7, 9, 14].

Active learning and adaptive sampling strategies: Active learning exploits uncertainty to guide data acquisition, reducing the need for exhaustive labeling in expensive materials experiments [2, 15, 16]. Query-by-committee, expected model change, and uncertainty sampling variants prioritize informative points [17, 19, 20]. In materials discovery, these strategies have informed closed-loop workflows in which uncertainty guides candidate selection for synthesis or characterization [3, 21, 25]. Bayesian active learning integrates surrogate uncertainty into acquisition functions, enabling efficient exploration [26, 27]. Entropy-based or variance-driven approaches have demonstrated conceptual advantages in balancing informativeness and representativeness [28, 29].

Bayesian optimization in materials search spaces: Bayesian optimization treats materials design as global optimization under uncertainty, using Gaussian processes as surrogates to model objective functions [15, 30]. Acquisition functions such as expected improvement or upper confidence bound explicitly incorporate predictive uncertainty to trade exploration and exploitation [16, 31]. Extensions to multi-objective or constrained settings have broadened the applicability of these methods to complex materials problems [1, 32]. The reliance on uncertainty as a core driver underscores its potential beyond optimization to broader control functions [2, 4].

Challenges and limitations in current uncertainty integration: Despite progress, uncertainty remains compartmentalized: it informs specific algorithmic choices but rarely governs overarching research logic [5, 8, 18]. Overreliance on heuristic acquisition can lead to myopic focus, while miscalibrated uncertainty undermines trust [6, 9, 12]. Few works conceptualize uncertainty as a persistent signal that can modulate multiple workflow stages simultaneously [10, 13, 22]. Bridging this gap requires elevating confidence to a systemic control variable, aligning it with scientific reasoning principles [11, 14, 17].

Proposed conceptual framework: The proposed framework reconceptualizes algorithmic confidence as the central regulatory signal governing materials research workflows. Rather than situating uncertainty as a peripheral diagnostic appended to predictive outputs, the framework positions confidence as an infrastructural control variable embedded within the epistemic mechanics of discovery. In this formulation, confidence estimates—whether probabilistic, interval-based, or categorical—operate as dynamic regulators that condition how knowledge claims propagate through investigative stages. The objective is not to privilege any specific modeling paradigm, but to elevate confidence into a systems-level steering logic capable of coordinating prediction, validation, exploration, and refinement processes.

At the core of the framework lies a closed epistemic loop linking materials representation, machine learning inference, and confidence-mediated control. Materials inputs—spanning compositional descriptors, crystal structures, microstructural encodings, or multimodal characterization data—are processed through predictive architectures that generate dual outputs: (i) point property estimates and (ii) associated confidence measures. These confidence signals may derive from posterior predictive variance, ensemble dispersion, calibration residuals, or other uncertainty proxies, yet the framework remains agnostic to their computational origin. What is critical is their interpretive mobilization.

Both outputs converge upon a central evaluative module conceptualized as a confidence control node. This node operationalizes regulatory assessment by comparing incoming confidence magnitudes against predefined, adaptive, or context-sensitive thresholds. These thresholds need not be static; they may evolve in response to shifting research priorities, resource constraints, or evidential accumulation. Thus, control is not binary but gradient-responsive, enabling proportional alignment between epistemic certainty and downstream action.

Within this architecture, three primary operational regimes emerge.

High-confidence regimes signal predictive maturity. Here, algorithmic outputs exhibit sufficient evidential grounding to justify exploitation-oriented actions. Such actions may include hypothesis advancement, design recommendation,

experimental prioritization, or translational deployment into engineering contexts. Importantly, exploitation is not framed as epistemic closure but as conditional utilization—confidence legitimizes action without foreclosing future reassessment.

Low-confidence regimes, by contrast, denote epistemic insufficiency. Rather than suppressing model output, the framework interprets low confidence as an actionable discovery signal. Control directives in this regime privilege exploratory interventions, such as targeted data acquisition, expansion of compositional search domains, representational enrichment, or architectural retraining. Uncertainty thus becomes generative rather than obstructive, directing attention toward under-characterized regions of materials space.

Intermediate-confidence regimes constitute a dialectical zone between exploitation and exploration. Here, evidence is suggestive yet incomplete, warranting hybridized responses. Actions may include partial experimental validation, sensitivity analysis, cross-model comparison, or localized uncertainty reduction strategies. This regime prevents premature commitment while avoiding exploratory overextension, preserving investigational balance.

Beyond static categorization, the framework incorporates confidence gradients as temporal and spatial modulators of control. Temporal gradients track how uncertainty evolves across iterative training cycles or active learning rounds, indicating epistemic convergence or divergence. Spatial gradients map confidence heterogeneity across materials design spaces, revealing clusters of predictive stability versus volatility. These gradient dynamics enable anticipatory steering—detecting knowledge stabilization, emergent blind spots, or model drift before predictive failure manifests.

Crucially, the framework decouples confidence from specific tasks or algorithmic substrates. Confidence operates orthogonally across predictive inference, sampling logic, and decision orchestration layers. This transversality allows the control signal to coordinate heterogeneous workflows—spanning computational screening, autonomous laboratories, and human-in-the-loop experimentation—without imposing hierarchical rigidity. Instead, the architecture privileges recursive feedback.

Feedback loops constitute the adaptive backbone of the system. Outcomes generated within exploitative, exploratory, or hybrid pathways are re-entered into the modeling environment as updated training signals, validation benchmarks, or representational corrections. These updates recalibrate confidence distributions, which in turn reshape subsequent control directives. The loop thereby sustains epistemic plasticity: action informs knowledge, and knowledge reconditions action.

Conceptually, this recursive structure institutionalizes epistemic humility within AI-mediated materials discovery. By tethering workflow progression to evidential confidence rather than predictive assertion alone, the system ensures proportionality between belief and intervention. Overconfident extrapolation is dampened; underexplored uncertainty is illuminated. Confidence thus evolves from a statistical descriptor into an ethical and operational governor of scientific conduct.

The originality of this framework lies in its unification of previously fragmented uses of uncertainty. Whereas prior paradigms embed uncertainty within isolated modules—such as query strategies, surrogate optimization, or calibration analytics—this model integrates confidence as a continuous regulatory infrastructure spanning the full research lifecycle. Through this synthesis, materials AI systems gain not only predictive capability but also epistemic self-awareness: the ability to recognize the limits of their knowledge and modulate their behavior accordingly. **Figure 1**. Shows the conceptual schematic of the confidence-driven control architecture in materials AI. Machine learning predictions generate both property estimates and associated confidence measures, which converge within a regulatory control node. Confidence thresholds direct workflow progression into exploitative, exploratory, or hybrid refinement pathways. Recursive feedback loops recalibrate model states and confidence distributions, forming a closed epistemic governance system for adaptive materials discovery.

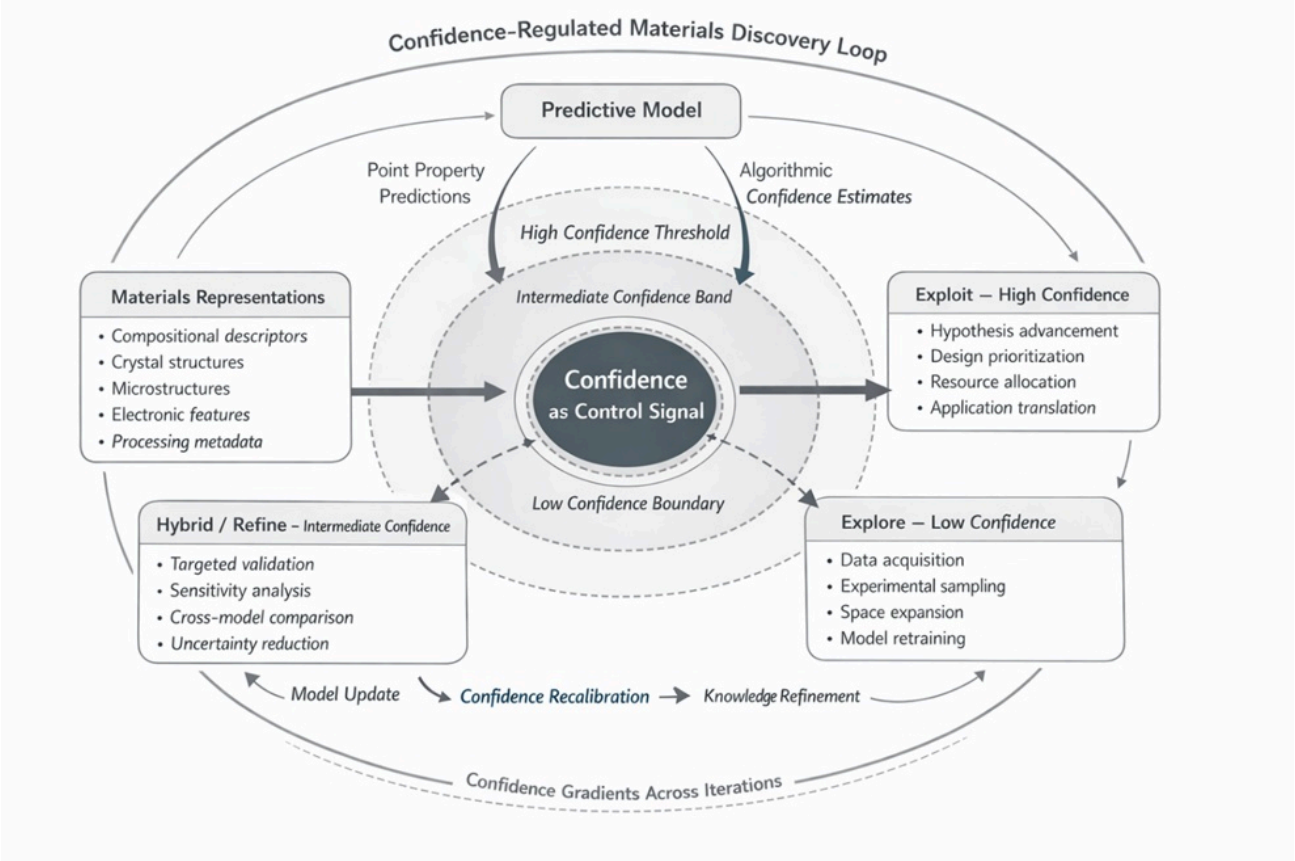

Figure 1. Confidence- driven control architecture in materials AI: a closed-loop epistemic regulation framework

Propositions

The conceptual framework advanced here rests on several interrelated propositions that articulate the role of algorithmic confidence as a control signal in materials research. These propositions are interpretive in nature, deriving from the synthesis of uncertainty quantification practices while extending them into a unified regulatory paradigm.

First, algorithmic confidence serves as a primary epistemic currency in computational materials workflows. Unlike raw predictive outputs, which represent localized approximations, confidence encapsulates the degree of evidential support underpinning those approximations. It therefore serves as a more fundamental indicator of model-world alignment than point estimates alone, enabling decisions to be conditioned on relational reliability rather than nominal accuracy [1, 8, 18].

Second, confidence exhibits control properties analogous to feedback signals in dynamical systems. Threshold crossings in confidence levels can trigger discrete state transitions—such as from exploratory to exploitative modes—while continuous confidence gradients provide proportional modulation of workflow intensity or direction. This controllability arises intrinsically from the probabilistic foundations of modern uncertainty quantification techniques, rendering confidence a naturally regulable variable [10, 15, 30].

Third, elevating confidence to a central control position decouples scientific progress from the requirement of uniformly high predictive fidelity. Materials inquiry proceeds robustly even in regimes of partial knowledge, provided confidence informs adaptive redirection. This proposition reframes uncertainty not as an obstacle to overcome through ever-larger datasets but as an inherent governor that preserves rigor across knowledge states [9, 14, 22].

Fourth, confidence-driven control introduces self-regulation into otherwise open-loop computational pipelines. Iterative feedback from confidence assessments to model refinement, sampling strategy, and decision criteria creates closed conceptual loops that align computational effort with epistemic boundaries. Such a self-regulation conceptually reduces the risk of premature convergence or wasteful exploration [2, 13, 16].

Fifth, the framework implies a hierarchical ordering of investigative actions ordered by confidence magnitude. High-confidence regions justify immediate utilization in downstream reasoning or design; intermediate-confidence zones warrant refinement or hybrid strategies; low-confidence domains necessitate fundamental expansion of the knowledge base. This ordering provides a principled basis for prioritizing computational versus empirical resources without presupposing specific implementations [3, 19, 25].

Sixth, temporal dynamics of confidence—manifest as convergence rates, divergence signals, or oscillatory patterns—offer supplementary control information. Rapid increases in confidence signal productive knowledge accretion, whereas stagnation or decline prompt reevaluation of assumptions, descriptors, or model architectures. These dynamics thus extend static confidence thresholds into trajectory-sensitive governance [11, 20, 27].

Collectively, these propositions delineate a coherent conceptual shift wherein algorithmic confidence transitions from a diagnostic or selective tool to an overarching regulatory mechanism in materials research.

Results and Discussion

The conceptualization of algorithmic confidence as a control signal invites reflection on several broader implications for materials science and computational discovery paradigms.

By positioning confidence at the center of workflow governance, the framework addresses a persistent tension between the speed promised by machine learning and the caution demanded by scientific validity. Current practices often alternate between aggressive exploration via uncertainty sampling and cautious exploitation via high-confidence predictions, yet these phases remain loosely coordinated [4, 17, 21]. The proposed approach integrates them under a single regulatory variable, conceptually enabling smoother, more contextually appropriate transitions. This integration does not eliminate human judgment but augments it by providing a quantitative substrate for deliberative choices.

A further implication concerns resource allocation in materials discovery campaigns. Experimental validation remains the ultimate arbiter of material viability, yet its high cost necessitates judicious selection of candidates. Confidence-based control offers a mechanism to defer expensive measurements until predictive support reaches sufficient strength, while simultaneously identifying regions where computation alone cannot resolve ambiguity [5, 15, 26]. In this sense, the framework promotes epistemic efficiency without sacrificing thoroughness.

The perspective also resonates with emerging discussions on trustworthy artificial intelligence in scientific contexts. Where trust traditionally accrues from predictive performance on held-out data, confidence-centric control shifts the emphasis to transparency of the epistemic state. Researchers interact not merely with predictions but with quantified assertions about those predictions, fostering more critical engagement with computational outputs [7, 12, 23]. This shift aligns with calls for interpretable and controllable machine learning systems in high-stakes domains.

The conceptualization's limitations must be acknowledged. The framework remains abstract and agnostic to specific uncertainty quantification methods, acquisition functions, or model classes. Practical realization would require careful calibration of confidence estimates to ensure meaningful thresholds and gradients [6, 9, 14]. Miscalibration—common in deep learning settings—could distort control behavior, leading to inappropriate exploration or overexploitation. Moreover, the approach assumes that uncertainty estimates capture relevant epistemic dimensions; domain-specific biases or unmodeled sources of variability may undermine this assumption [8, 18, 29].

By comparison, existing Bayesian optimization workflows already incorporate uncertainty in acquisition, yet they confine its influence to candidate selection [16, 30, 32]. The present framework generalizes this influence across prediction utilization, model evolution, and decision staging, suggesting a more pervasive regulatory role. Similarly, active learning paradigms use uncertainty for informativeness but rarely to modulate non-sampling decisions [2, 13, 19]. By synthesizing these strands, the conceptualization offers a more holistic vision.

Ultimately, the framework encourages a reevaluation of how computational tools participate in materials inquiry—not as oracles but as adaptive partners whose contributions are modulated by self-assessed reliability. This perspective may inform future design of integrated discovery platforms where confidence signals orchestrate human–machine collaboration at multiple scales.

Conclusion

Materials research stands at the threshold of profound augmentation enabled by machine learning, yet integrating such tools demands mechanisms that preserve scientific integrity amid accelerating computation. This manuscript has argued that algorithmic confidence—quantified uncertainty in model predictions—can fulfill this role when reconceptualized as a central control signal.

By synthesizing recent advances in uncertainty quantification, active learning, and Bayesian optimization, a novel framework has been delineated in which confidence governs workflow progression, balances exploration and exploitation, and enforces epistemic proportionality in decision-making. The associated propositions articulate confidence as epistemic currency, dynamical feedback, self-regulatory agent, hierarchical action selector, and trajectory monitor.

While remaining purely conceptual, this perspective illuminates pathways toward more deliberate, reliable, and efficient computational assistance in materials science. By treating confidence not as a byproduct of prediction but as its principal regulatory companion, the framework seeks to align machine intelligence more closely with the iterative, evidence-bound nature of scientific discovery. Future elaborations may explore concrete instantiations, yet the core insight—that uncertainty can govern as well as diagnose—offers a foundational reorientation for trustworthy materials computation.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 24 Jul 2022 Revised: 12 Sep 2022 Accepted: 06 Nov 2022
Published online: 18 January 2023

References

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;11:5966.
https://doi.org/10.1038/s41467-020-19597-w.

Tian Y, Li T, Pang J, Zhou Y, Xue D, Ding X, et al. Materials design with target-oriented bayesian optimization. npj Comput Mater. 2025;11:209.
https://doi.org/10.1038/s41524-025-01704-4.

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;9:49.
https://doi.org/10.1038/s41524-023-01006-7.

Tavazza F, DeCost B, Choudhary K. Uncertainty prediction for machine learning models of material properties. ACS Omega. 2021;6(48):32431-40.
https://doi.org/10.1021/acsomega.1c03752.

Vandermause J, Torrisi SB, Batzner S, Smidt T, Uhler C, Kozinsky B. Uncertainty-driven dynamics for active learning of interatomic potentials. Nat Comput Sci. 2023;3(3):249-58.
https://doi.org/10.1038/s43588-023-00406-5.

Chen C, Ong SP. Uncertainty-aware molecular dynamics from bayesian active learning. npj Comput Mater. 2023;9:88.
https://doi.org/10.1038/s41524-023-00988-8.

Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Materials property prediction with uncertainty quantification: A review. Appl Phys Rev. 2023;10(2):021409.
https://doi.org/10.1063/5.0145015.

Zhang H, Wang Y, Lookman T. Performance of uncertainty-based active learning for efficient materials discovery. Sci Rep. 2024;14:76800.
https://doi.org/10.1038/s41598-024-76800-4.

DeCost B, Hattrick-Simpers J, Mehta A. Targeted materials discovery using bayesian algorithm execution. npj Comput Mater. 2024;10:132.
https://doi.org/10.1038/s41524-024-01326-2.

Im J, Lee S, Lee J. A dynamic bayesian optimized active recommender system for efficient materials discovery. npj Comput Mater. 2024;10:191.
https://doi.org/10.1038/s41524-023-01191-5.

Tran A, Wildey T. Uncertainty quantification for bayesian active learning in rupture life prediction of ferritic steels. Sci Rep. 2022;12:6051.
https://doi.org/10.1038/s41598-022-06051-8.

Merchant A, Batzner S, Schoenholz SS, Aykol M, Cheon G, Cubuk ED. Scaling deep learning for materials discovery. Nature. 2023;624(7990):80-5.
https://doi.org/10.1038/s41586-023-06735-9.

Jablonka KM, Ongari D, Smit B. Uncertainty quantification in machine learning for materials. npj Comput Mater. 2022;8:1-10.
https://doi.org/10.1038/s41524-022-00759-4.

Chen C, Zuo Y, Ye W, Li X, Ong SP. Active learning in materials science with emphasis on adaptive sampling using uncertainties. npj Comput Mater. 2020;6:1-11.

Wang AY, Kauwe SK, Murdock R, Thompson A, Sparks TD. Uncertainty quantification in multivariable regression for material property prediction. Sci Rep. 2024;14:61189.
https://doi.org/10.1038/s41598-024-61189-x.

Li Z, Wang Y, Zhang Y. Uncertainty-aware mixed-variable machine learning for materials design. Sci Rep. 2022;12:23431.
https://doi.org/10.1038/s41598-022-23431-2.

Zhang L, Chen C, Ong SP. Uncertainty quantification for neural network potential foundation models. npj Comput Mater. 2025;11:157.
https://doi.org/10.1038/s41524-025-01572-y.

Peterson AA, Khorshidi A. Addressing uncertainty in atomistic machine learning. Phys Rev Mater. 2021;5:024801.
https://doi.org/10.1103/PhysRevMaterials.5.024801.

Bartók AP, Kermode JR, Bernstein N, Csányi G. Machine learning a general-purpose interatomic potential for silicon. Phys Rev X. 2020;10:041044.
https://doi.org/10.1103/PhysRevX.10.041044.

Pan S, Zhang Y, Wang H. Bayesian force fields from active learning for simulation of interatomic interactions. npj Comput Mater. 2021;7:50.
https://doi.org/10.1038/s41524-021-00510-y.

Lin S, Liu Y, Wang H. On-the-fly active learning of interpretable bayesian force fields for materials simulation. npj Comput Mater. 2020;6:28.
https://doi.org/10.1038/s41524-020-0283-z.

Li X, Zhang Y, Chen C. Uncertainty quantification in molecular simulations with dropout neural networks. npj Comput Mater. 2020;6:90.

Smith JS, Nebgen BT, Zubatyuk R, Lubbers N, Devereux C, Barros K, et al. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. Nat Commun. 2019;10(1):2903.
https://doi.org/10.1038/s41467-019-10827-4.

Zhang Y, Wang H, Chen C. Active oversight and quality control in standard bayesian optimization for materials design. npj Comput Mater. 2025;11:148.
https://doi.org/10.1038/s41524-024-01485-2.

Lookman T, Xue D, Balachandran PV. Bayesian optimization with experimental failure for high-throughput materials discovery. npj Comput Mater. 2022;8:59.
https://doi.org/10.1038/s41524-022-00859-8.

Wang Y, Zhang L, Ong SP. Unlocking the black box beyond bayesian global optimization for materials discovery. npj Comput Mater. 2025;11:163.
https://doi.org/10.1038/s41524-025-01639-w.

Zuo Y, Chen C, Li X, Ong SP. Uncertainty quantification in scientific machine learning for materials. J Comput Phys. 2023;478:111965.
https://doi.org/10.1016/j.jcp.2022.111965.

DeCost B, Mehta A. Methods for comparing uncertainty quantifications for material property predictions. Mach Learn Sci Technol. 2020;1:025005.

Tran A, Wildey T. Uncertainty-aware electronic density-functional distributions. Phys Rev Mater. 2025;9:083801.
https://doi.org/10.1103/PhysRevMaterials.9.083801.

Merchant A, Batzner S. Machine learning materials properties with accurate predictions, uncertainty estimates, domain guidance, and persistent online accessibility. Mach Learn Sci Technol. 2025;6:015001.

Jablonka KM, Smit B. A rigorous uncertainty-aware quantification framework is essential for reproducible and replicable machine learning workflows. Digit Discov. 2023;2:1234-45.
https://doi.org/10.1039/D3DD00094J.

Butler KT, Walsh A. Uncertainty in the era of machine learning for atomistic modeling. Digit Discov. 2025;4:567-89.
https://doi.org/10.1039/D5DD00102A.