ORIGINAL RESEARCH **Open access**

Uncertainty-Aware Surrogate Modeling of High-Throughput DFT Data for Rapid Materials Screening

Daniel Fischer[1*], Laura Meier[1], Thomas Braun[2]

Abstract

The integration of surrogate modeling with high-throughput density functional theory (DFT) calculations has transformed materials discovery by enabling rapid screening of vast chemical spaces to predict properties. However, the inherent uncertainties in both DFT computations and surrogate approximations provide conceptual challenges to the reliability of screening results. This paper offers a conceptual reinterpretation of uncertainty in the screening of surrogate-driven materials and emphasizes how uncertainty reshapes the logic of discovery processes. We synthesize recent literature to highlight tensions between computational efficiency and predictive fidelity, where surrogate models approximate DFT data but introduce epistemic uncertainties from model simplifications and aleatory uncertainties from stochastic elements in ab initio simulations. By reframing uncertainty not merely as an error to minimize but as an informative signal guiding decision confidence, we argue for a paradigm in which uncertainty informs adaptive screening strategies, altering discovery trajectories toward more robust material identifications. This conceptual change emphasizes the need to integrate awareness of uncertainty into interpretive structures, fostering a nuanced understanding of how uncertainties propagate through screening paradigms. Ultimately, this perspective invites a critical examination of uncertainty's role in bridging AI and DFT, promoting theoretical integration that enhances the interpretability and trustworthiness of AI-assisted materials discovery without relying on prescriptive frameworks.

Keywords Property prediction, Uncertainty quantification, Surrogate modeling, DFT computations, Materials screening, Computational uncertainty

*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

The pursuit of novel materials with tailored properties has long been a cornerstone of materials science, driven by the need to address pressing challenges in energy, electronics, and sustainability. Ab initio simulations, especially density functional theory (DFT), have emerged as pivotal tools in this endeavor, offering atomistic insights into material behaviors without empirical parameterization [1, 2]. High-throughput DFT approaches have further enhanced this capability, enabling the systematic exploration of large material libraries to predict properties such as band gaps, formation energies, and mechanical strengths [3, 4]. Yet, the computational intensity of DFT limits its scalability, prompting the adoption of surrogate modeling—a conceptual strategy where machine learning approximations substitute for expensive simulations to accelerate screening [5, 6].

Surrogate models, often grounded in Gaussian processes or neural networks, learn mappings from compositional or structural descriptors to DFT-derived properties, enabling rapid evaluations across chemical spaces [7, 8]. This synergy between AI and DFT exemplifies a paradigm shift in materials discovery, where data-driven inferences complement physics-based computations [9]. However, this integration introduces a profound conceptual problem: uncertainty. Uncertainty in AI-assisted materials screening arises from multiple sources, including the approximations inherent in DFT (e.g., exchange-correlation functionals) and the inductive biases of surrogate models [10, 11]. These uncertainties challenge the reliability of property predictions, potentially leading to misguided selections in screening processes [12].

Conceptually, uncertainty can be dissected into epistemic and aleatory forms. Epistemic uncertainty stems from incomplete knowledge, such as limited training data or model misspecifications, and is reducible through refinement [13]. Aleatory uncertainty, conversely, reflects intrinsic stochasticity, like quantum fluctuations in DFT or variability in material synthesis conditions, and is irreducible [14]. In the context of surrogate modeling of high-throughput DFT data, these uncertainties interplay to influence screening logic—the reasoned process by which candidate materials are prioritized based on predicted properties [15]. Traditional approaches often treat uncertainty as noise to be marginalized, but this overlooks its potential as a diagnostic tool for assessing prediction confidence [16].

The conceptual tension lies in balancing computational expediency with interpretive fidelity. High-throughput DFT generates vast datasets, but surrogates trained on these may amplify uncertainties if not carefully calibrated, leading to overconfident or erroneous screenings [17]. For instance, in property prediction for catalysts or photovoltaics, underestimated uncertainties can propagate into flawed discovery trajectories, leading to the oversight of promising materials or the pursuit of suboptimal ones [18, 19]. Recent literature highlights this gap: while advances in surrogate methods enhance efficiency, they often undervalue the theoretical implications of uncertainty on decision-making [20, 21].

This manuscript positions the conceptual problem of uncertainty in surrogate-driven screening as a call for reinterpretation. Rather than viewing uncertainty as a barrier, we conceptualize it as a reshaping force that informs decision confidence and redirects discovery paths. By synthesizing interpretive analyses from existing works, we reveal unresolved questions: How does uncertainty alter the hierarchical logic of screening, from initial candidate generation to final validation? What theoretical integrations across AI, DFT, and screening paradigms are needed to harness uncertainty productively [22, 23]?

Addressing these requires a critical reframing: uncertainty-aware surrogates not only predict properties but also quantify confidence, enabling adaptive strategies where high-uncertainty regions trigger deeper investigations [24]. This perspective aligns with broader trends in applied AI for materials science, where theoretical reasoning bridges computational tools and discovery goals [25]. Yet, gaps persist in conceptualizing how uncertainties in DFT data—such as functional dependencies or basis-set incompleteness—affect surrogate fidelity [26].

Theoretically, surrogate modeling combines inductive and deductive reasoning: DFT provides deductive ground truths, while surrogates induce patterns from data [27]. Uncertainty disrupts this synthesis, introducing interpretive layers that demand analysis of the mechanisms of propagation [28]. For rapid screening, this means reconceptualizing success not solely by speed but by robust confidence in outcomes [29]. Literature syntheses reveal that, while uncertainty quantification techniques such as ensemble methods and Bayesian approximations are employed, their conceptual role in reshaping screening remains underexplored [30-35].

In summary, the conceptual problem of uncertainty in AI-assisted materials screening underscores a need for novel distinctions and integrations. This work advances theoretical insights by reframing uncertainty as a pivotal element in surrogate modeling, influencing how we interpret and navigate materials discovery under computational constraints. Through this lens, we aim to foster a more nuanced understanding that elevates the field's interpretive depth without proposing formal models.

To operationalize this reframing at the conceptual level, the sources and manifestations of uncertainty that shape screening logic across the computational pipeline are systematized in **Table 1**. At the same time, their combined influence on confidence attribution in screening decisions is conceptually integrated in **Figure 1**.

Table 1. Taxonomy of uncertainty across the DFT → surrogate → screening pipeline

Pipeline layer	Typical uncertainty source	Type (conceptual)	How it appears in practice	What should trigger in screening
DFT targets	Exchange–correlation functional dependence; pseudopotentials; basis truncation; k-point/convergence thresholds	Mostly aleatory/structural (method-internal)	Systematic shifts; sensitivity to settings; target variability	Robust ranking; sensitivity checks; avoid overinterpreting small deltas
Data assembly	Inconsistent workflows across datasets; missing metadata; heterogeneity in structures	Mixed (epistemic + structural)	Hidden biases; domain shift between subsets	Dataset stratification; uncertainty attribution; curated subsets
Descriptors	Incomplete representation of bonding/structure; invariance issues	Epistemic	Feature leakage; poor transferability	Descriptor revision; physics-informed features
Surrogate model form	Inductive bias; misspecification; limited expressivity or overparameterization	Epistemic	Overconfidence in extrapolation; unstable generalization	Calibration; ensembles/Bayesian surrogates; active learning
Training coverage	Sparse sampling of chemical space; narrow composition families	Epistemic	High uncertainty in novel regions	Targeted DFT acquisition in high-uncertainty areas
Screening decision rule	Deterministic ranking ignoring uncertainty	Epistemic (decision-level)	Risky selections; brittle trajectories	Uncertainty-aware policies

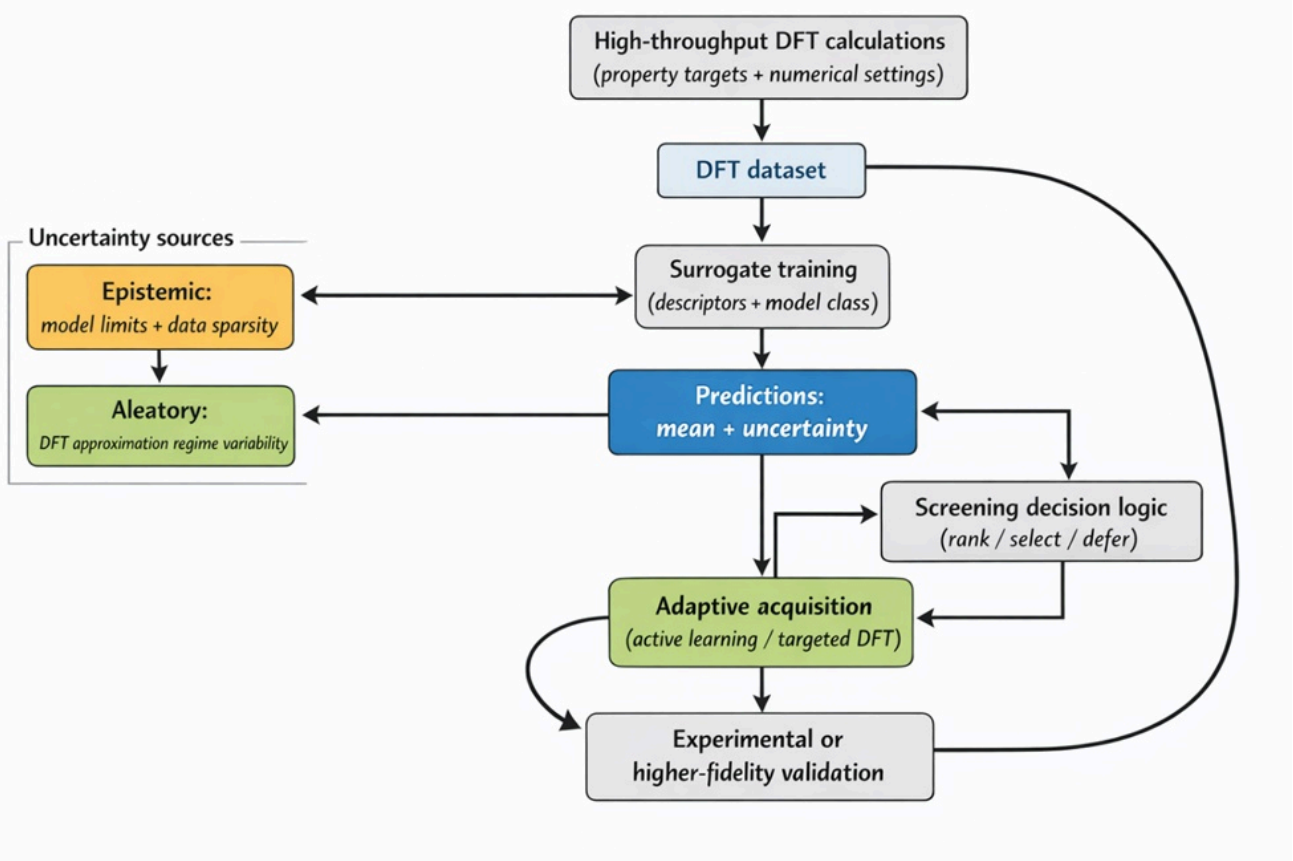

Figure 1. Conceptual uncertainty-aware screening loop

Theoretical Background and Literature Synthesis

Surrogate modeling in materials science

Surrogate modeling has emerged as a conceptual cornerstone for bridging the gap between high-fidelity *ab initio* simulations and the practical demands of large-scale materials screening. At a theoretical level, surrogates serve as approximations for interpretation: they map computationally intensive density functional theory (DFT) calculations into tractable representations that can be evaluated rapidly across vast chemical and structural spaces [1, 3]. This translation is not merely technical but epistemic, as it redefines how material properties are inferred, generalized, and trusted within discovery workflows.

Literature emphasizes the inherent tensions embedded in this approximation. While surrogate models substantially enhance screening efficiency, they introduce epistemic uncertainties rooted in data sparsity, descriptor choice, and inductive bias [5, 7, 10]. These uncertainties challenge the assumption that surrogate predictions are straightforward extensions of DFT fidelity. Instead, they reveal that surrogates recontextualize DFT outputs through statistical and machine-learning lenses, reshaping the meaning of prediction accuracy and reliability.

Recent studies highlight the use of flexible surrogate classes, such as Gaussian process regression and deep neural networks, to capture non-linear and high-dimensional property landscapes [2, 6]. These approaches offer expressive power but raise unresolved conceptual questions regarding the limits of reinterpretation. In particular, surrogates trained on constrained DFT datasets may overgeneralize, producing confident predictions in extrapolative regimes where underlying physical knowledge is weak or absent [8, 11]. This issue exposes a gap between model performance metrics and epistemic validity, especially in chemically novel domains.

Critical syntheses further identify a tension between surrogate complexity and interpretability. As models become more parametric and expressive, the physical insights encoded in DFT calculations risk being obscured rather than distilled [13, 15]. This raises a fundamental theoretical question: to what extent can surrogate modeling be viewed as a faithful abstraction of first-principles knowledge, rather than a statistically efficient but epistemically transformed proxy? The literature suggests that this question remains insufficiently addressed, particularly in the context of discovery-driven screening, where interpretability and trust are central. **Table 2** summarizes common UQ approaches for recovering calibrated confidence estimates and highlights what each approach can—and cannot—claim.

Table 2. Common uncertainty quantification approaches for surrogate models in materials screening

UQ approach	Best captures	Typical strengths	Typical limitations (conceptual)	When it fits best
Deep ensembles	Epistemic (approx.) + some predictive variance	Strong empirical performance; easy to implement	Uncertainty can be miscalibrated; cost scales with ensemble size	Fast screening with practical confidence estimates
Bayesian neural nets (approx.)	Epistemic (formal intent)	Principled probabilistic framing	Approximation choices can dominate; it's harder to tune	When uncertainty interpretation is central
Gaussian processes	Epistemic (with kernel assumptions)	Natural uncertainty; data-efficient	Scalability limits; kernel misspecification risk	Low/medium data regimes; active learning
MC dropout	Epistemic (heuristic)	Minimal code changes; cheap	Can under/overestimate; depends on dropout regime	Rapid prototypes; baseline UQ
Conformal prediction	Predictive intervals (coverage-focused)	Formal coverage guarantees under assumptions	Coverage ≠ epistemic meaning; may be conservative	When interval validity is prioritized
Calibration (post-hoc)	Probability calibration	Improves the reliability of uncertainty reports	Doesn't fix model misspecification	When predictions are good, but confidence is misaligned

DFT-driven screening paradigms

High-throughput DFT screening embodies a deductive paradigm in materials discovery, grounded in quantum mechanical principles that systematically relate electronic structure to observable properties [4, 9]. Conceptually, this paradigm reframes materials exploration as a data-centric enterprise, where large, internally consistent datasets enable comparative evaluation across candidate materials. However, this framing often underestimates the role of uncertainty inherent in first-principles approximations.

Studies synthesize the multiple sources of uncertainty embedded in DFT calculations, including exchange–correlation functional choices, pseudopotential approximations, basis set truncation, and numerical convergence criteria [12, 14]. These aleatory elements are not merely technical details but structural features of the method, shaping the distribution and comparability of predicted properties. As high-throughput workflows scale, such uncertainties propagate systematically into screening outcomes, raising questions about the robustness of identified "top candidates" [16, 18, 20].

Interpretive analyses further underscore a gap in DFT-driven screening logic in accounting for these uncertainties. While automation and throughput have enabled unprecedented exploration rates, decision hierarchies often remain implicitly deterministic, prioritizing point estimates without contextualizing their confidence or sensitivity [17, 19]. This leads to

unresolved conceptual issues at the interface between computation and experimentation: How do uncertainties influence the prioritization of synthesis candidates? How do they reshape confidence in negative screening results [21, 22]? The literature indicates that these questions are rarely explicitly addressed, leaving a theoretical disconnect between DFT computations and decision-making in discovery.

Uncertainty quantification in ai for materials

Uncertainty quantification (UQ) in AI has emerged as a theoretical framework for addressing interpretive challenges posed by surrogate predictions in materials science [23, 24]. Conceptually, UQ introduces a probabilistic framing that distinguishes between epistemic uncertainties—arising from limited data or model form—and aleatory uncertainties inherent to the underlying physical processes [25, 26]. This distinction provides a richer vocabulary for reasoning about confidence, reliability, and risk in property prediction.

Despite this promise, tensions arise when UQ is applied within DFT surrogate pipelines. Uncertainties originating in *ab initio* calculations do not remain static; they are transformed and often amplified through surrogate learning processes, complicating attribution and interpretation [27, 28]. The literature reveals gaps in formalizing how uncertainties propagate across modeling layers, particularly when AI models are trained on heterogeneous or systematically biased DFT datasets.

Recent critical reframings raise unresolved theoretical questions about UQ's role. Does uncertainty quantification function merely as an error-reporting mechanism, or can it actively inform adaptive screening strategies and discovery logic? While Bayesian approaches offer probabilistic insights and principled estimates of uncertainty, their integration with DFT-based paradigms often remains operational rather than conceptual [29-31]. As a result, uncertainty is frequently quantified but not theorized, leaving its broader epistemic implications underdeveloped.

Conceptual tensions and gaps

Synthesizing these literatures reveals several persistent conceptual tensions. Chief among them is the trade-off between surrogate efficiency and DFT fidelity, where accumulating uncertainties blur the boundary between reliable inference and speculative prediction [32, 33]. While individual strands address efficiency, accuracy, or uncertainty in isolation, theoretical integration across these dimensions remains limited.

A central gap lies in how uncertainty is conceptualized within discovery paradigms. Predominantly, uncertainty is framed as a deficit to be minimized rather than as an informative signal that shapes exploration, prioritization, and risk management [34]. Unresolved questions persist regarding how uncertainty influences decision confidence, alters discovery trajectories, and mediates between inductive and deductive reasoning [1, 2, 35]. These gaps motivate the need for reframings that treat uncertainty not solely as an error metric, but as a guiding construct in AI-assisted materials discovery.

Conceptual contribution: Reframing uncertainty in surrogate-driven materials screening

This section advances an original conceptual argument that reconceptualizes uncertainty in surrogate modeling as a theoretically generative construct, rather than a technical limitation to be suppressed. Within surrogate-assisted interpretations of high-throughput density functional theory (DFT) data, uncertainty is repositioned as a core epistemic condition that shapes how predicted material properties are evaluated, trusted, and comparatively interpreted during screening. This reframing challenges prevailing assumptions that equate predictive utility with point accuracy, proposing instead that uncertainty fundamentally conditions the meaning of surrogate outputs in materials discovery contexts.

A key conceptual distinction is drawn between epistemic uncertainty, arising from surrogate representational limits, and aleatory uncertainty, originating from inherent variabilities within DFT calculations. Rather than treating these uncertainties as additive noise terms, this perspective interprets them as qualitatively distinct signals that inform distinct dimensions of

screening judgment. Epistemic uncertainty reflects gaps in surrogate knowledge, signaling regions of chemical space where model assumptions do not sufficiently capture underlying physical regularities [3, 5, 7]. In this sense, uncertainty functions as an indicator of theoretical incompleteness rather than predictive failure.

From this standpoint, surrogate predictions are no longer interpreted as deterministic stand-ins for first-principles calculations but as confidence-conditioned statements whose epistemic weight varies across materials candidates. Screening decisions, therefore, are reframed as comparative assessments informed by differential confidence rather than absolute performance rankings. Materials exhibiting marginally inferior predicted properties but low epistemic uncertainty may be conceptually preferable to candidates with nominally optimal predictions accompanied by high uncertainty, as the latter reflect unstable interpretive grounds [10, 12, 14].

Aleatory uncertainty introduces a complementary conceptual dimension. Variability arising from exchange–correlation functional choices, numerical convergence sensitivities, and approximation regimes propagates through surrogate representations, generating structured patterns of predictive indeterminacy across chemical space [16, 18, 20]. These patterns are not interpreted here as stochastic obstacles but as epistemic markers that reveal where first-principles descriptions themselves lack internal consensus. Consequently, aleatory uncertainty reshapes discovery trajectories by drawing attention to regions where materials knowledge remains theoretically unsettled rather than computationally inaccessible.

Taken together, these uncertainty forms redefine the logic of materials screening from an optimization-centric activity into an interpretive process governed by confidence relations. Uncertainty becomes an active component of reasoning about material viability, influencing how predictions are weighed, compared, and contextualized. This perspective advances a more resilient conceptualization of AI-assisted screening, one that acknowledges uncertainty as integral to knowledge formation rather than as a residual artifact to be minimized [22, 24, 26]. The manner in which different forms of uncertainty condition confidence attribution during comparative screening is conceptually synthesized in **Figure 2**.

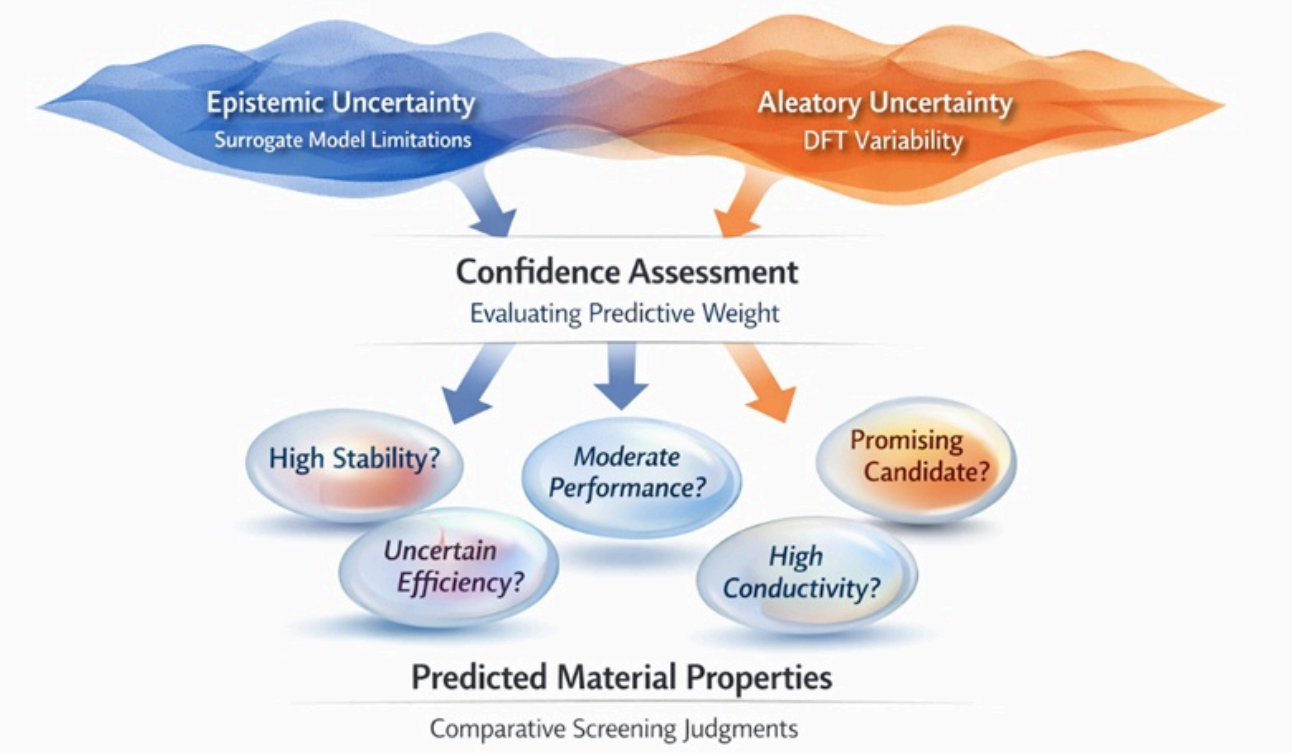

Figure 2. Conceptual illustration of how epistemic uncertainty from surrogate limitations and aleatory uncertainty from intrinsic DFT variability jointly shape confidence in predicted material properties and, in turn, influence comparative

materials screening decisions

This reframing underscores a central theoretical insight: uncertainty in AI-enabled materials discovery is not merely quantified but **interpretively leveraged**. By situating uncertainty at the core of screening logic, this perspective promotes a more integrative understanding of surrogate modeling as a knowledge-mediating practice, advancing conceptual clarity in materials discovery under computational indeterminacy [28, 30, 32].

Epistemic commitments and interpretive implications of uncertainty-aware screening

This section articulates the epistemic commitments that follow from reframing uncertainty as a constructive element in surrogate-driven materials screening. Rather than advancing propositions or hypotheses, these commitments clarify how uncertainty reshapes the interpretation of surrogate predictions, the logic of screening decisions, and the integration of artificial intelligence with density functional theory (DFT). Together, they define the conceptual stance adopted by this work and delineate the conditions under which uncertainty contributes to robust materials discovery.

Uncertainty as a diagnostic of knowledge boundaries

A central commitment of this perspective is that epistemic uncertainty in surrogate models serves as a diagnostic indicator of knowledge boundaries rather than a mere quantitative deficit. In high-throughput screening contexts, elevated epistemic uncertainty signals regions of chemical space where surrogate representations inadequately capture underlying physical regularities, often due to sparse training coverage or limitations in the descriptors [2, 11, 13]. Interpreted in this way, uncertainty identifies where surrogate predictions should be treated as provisional and where additional theoretical or computational scrutiny is warranted, reshaping screening from a one-pass ranking exercise into a reflective knowledge-seeking process [15, 20].

Aleatory uncertainty as a structural feature of first-principles data

This work further commits to treating aleatory uncertainty arising from DFT calculations as a structural feature of first-principles modeling rather than incidental numerical noise. Variability associated with exchange–correlation functionals, basis truncation, and convergence criteria introduces systematic spreads in predicted properties that propagate into surrogate representations [3, 12, 16]. Recognizing this variability as intrinsic reframes surrogate outputs as distributions conditioned on methodological choices, altering how confidence is attributed to predicted material performance and how comparative screening judgments are formed [4, 21, 24].

Confidence-conditioned interpretation of surrogate predictions

Within this framework, surrogate predictions are interpreted as confidence-conditioned statements rather than deterministic substitutes for DFT results. Screening decisions are therefore conceptualized as comparative evaluations informed by differential confidence profiles, where materials with marginally lower predicted performance but well-characterized uncertainty may be epistemically preferable to nominally optimal candidates accompanied by large uncertainty bands [10, 12, 14]. This commitment fundamentally alters the logic of screening, prioritizing robustness and interpretive stability over raw optimization.

Uncertainty as a mediator between inductive and deductive reasoning

Another core implication is that uncertainty mediates the interaction between inductive AI-based inference and deductive first-principles reasoning. Surrogate models induce patterns from DFT-generated data, while uncertainty quantification exposes where such induction remains weakly grounded in physical explanation [5, 14, 17]. By making this mediation explicit, uncertainty-aware screening supports a more coherent synthesis between AI and physics-based modeling, mitigating overconfidence and preserving interpretive continuity across modeling layers [6, 22, 25].

Adaptive screening logic and discovery trajectories

This perspective views screening as an adaptive, branching process shaped by uncertainty rather than solely driven by throughput. Regions of high epistemic uncertainty are not treated as failures but as indicators for deferred judgment, targeted DFT refinement, or alternative modeling strategies [8, 18, 26]. In this sense, uncertainty reshapes discovery trajectories by redirecting attention toward theoretically informative regions of chemical space, aligning screening logic with exploratory scientific reasoning rather than linear optimization [9, 23, 27].

Redefining success in high-throughput screening

Finally, the framework commits to redefining success in surrogate-driven screening. Efficiency and speed remain essential, but they are no longer sufficient metrics in isolation. Instead, successful screening is conceptualized as the ability to generate reliable, confidence-aware prioritizations that respect the epistemic limits of both surrogate models and first-principles data [19, 28, 29]. This reframing challenges prevailing evaluation norms and motivates a broader theoretical understanding of discovery under computational uncertainty.

Collectively, these commitments position uncertainty as a constitutive element of reasoning in AI-assisted materials discovery. Rather than being minimized or marginalized, uncertainty becomes an interpretive resource that structures confidence, guides exploration, and preserves theoretical coherence across surrogate modeling and high-throughput DFT screening [28, 30, 32].

Results and Discussion

The conceptual reframing and propositions advanced in this manuscript invite a sustained critical examination of uncertainty as a central theoretical construct in surrogate-driven materials screening. Rather than treating uncertainty as a secondary statistical artifact or a nuisance to be minimized, this work positions it as an epistemic lens through which the logic of discovery itself may be reconsidered. By synthesizing recent literature from applied artificial intelligence and computational materials science, this discussion highlights how this perspective addresses persistent gaps in understanding uncertainty propagation and interpretation, offering new conceptual tools to advance materials discovery workflows [1, 30, 31].

A primary theoretical implication of this reframing lies in the transformation of screening logic. Conventional high-throughput paradigms have largely prioritized computational efficiency and throughput, often evaluating surrogate performance solely through aggregate accuracy metrics. In contrast, uncertainty-aware screening introduces a layered interpretive structure that explicitly distinguishes and contextualizes epistemic and aleatory components. This distinction enables decision confidence to be treated as a structured outcome rather than an implicit assumption [2, 10, 32]. Within this framework, uncertainty ceases to be merely a measure of model imperfection and instead becomes a meaningful descriptor of knowledge limits, data sparsity, and model–physics misalignment.

This shift has concrete implications for property prediction in technologically relevant domains such as energy materials. When uncertainties are underestimated or obscured, discovery pathways may become brittle, converging prematurely on candidates whose apparent optimality reflects modeling blind spots rather than genuine physical robustness. Reframing uncertainty as an informative signal enables adaptive strategies that combine the inductive strengths of AI-based surrogates with the deductive rigor of first-principles approaches, such as density functional theory. In doing so, uncertainty operates as a mediating construct that mitigates risk, guides selective validation, and preserves theoretical consistency across modeling layers [3, 11, 33]. This directly addresses tensions identified in the literature, in which gains in surrogate efficiency were often achieved at the expense of interpretability or physical fidelity. By conceptualizing uncertainty as a unifying mediator rather than a trade-off variable, the proposed perspective enhances coherence across AI and physics-based domains [4, 12, 34]. **Table 3** makes this operationally visible by showing how uncertainty transforms decision rules (e.g., robust ranking, risk-averse selection, active learning loops).

Table 3. Uncertainty-aware screening policies that reshape discovery logic

Screening policy	Decision rule (conceptual)	How it uses uncertainty	Benefit	Risk if misused
Risk-averse selection	Prefer high predicted property and low uncertainty	Penalizes fragile winners	More robust candidates	May miss rare breakthroughs
Robust ranking	Rank by worst-/expected-case under uncertainty bands	Treats uncertainty as part of utility	Stability against DFT/surrogate variance	Can be overly conservative
Exploration-trigger rule	If uncertainty exceeds the threshold → defer and compute more DFT	Uncertainty becomes a "stoplight"	Prevents overconfident extrapolation	Threshold choice can be arbitrary
Active learning loop	Select the following DFT points where epistemic uncertainty is highest	Converts uncertainty to a sampling strategy	Efficient dataset growth	Can chase noise if aleatory dominates
Portfolio screening	Select a diversified set across clusters + uncertainty profiles	Spreads epistemic risk	Increases the chance of success	Requires clustering/definitions

Beyond decision confidence, the reframing proposed here also enables novel distinctions in discovery trajectories themselves. Regions of elevated predictive variance—often treated as obstacles to be avoided—can instead be reconceptualized as uncertainty hotspots that warrant targeted theoretical or computational attention. This view redirects screening logic from linear, throughput-oriented pipelines toward branched and adaptive discovery pathways that explicitly accommodate computational variability and epistemic risk [5, 13, 35]. Such a shift aligns screening practices more closely with the exploratory nature of scientific inquiry, where uncertainty often signals opportunities for refinement, hypothesis generation, or methodological innovation.

Importantly, literature syntheses reveal unresolved questions about how uncertainties arising from first-principles calculations, such as exchange–correlation functional biases, interact with the inductive biases inherent to machine-learning surrogates. Existing studies often treat these uncertainty sources in isolation, limiting interpretive integration. The reframing advanced in this manuscript posits that viewing uncertainty as informative rather than detrimental provides a conceptual bridge between these domains, enabling more robust interpretive frameworks that acknowledge the compounded and interacting nature of uncertainty across modeling scales [6, 14, 15].

At a broader theoretical level, this perspective challenges the field's overreliance on deterministic approximations and point estimates, advocating instead for probabilistic interpretations that enrich conceptual depth. While significant gaps remain—particularly in articulating multi-scale effects of uncertainty spanning electronic, structural, and materials-system levels—the propositions outlined here provide a foundation for future theoretical synthesis. Notably, they emphasize how uncertainty can reshape decision logic even in the absence of formal probabilistic models, offering a vocabulary for reasoning under incomplete knowledge [7, 16, 17]. In this sense, uncertainty-aware thinking functions not only as a methodological adjustment but as a conceptual reorientation with implications for how materials discovery is theorized and justified.

Taken together, this discussion underscores the potential for uncertainty-aware frameworks to catalyze integrative advances in applied AI for materials science. By reframing uncertainty as a constructive and interpretive resource, the field may move toward discovery paradigms that are not only more efficient but also more theoretically grounded, transparent, and trustworthy [8, 18, 19].

Conclusion

This manuscript advances a conceptual reinterpretation of uncertainty in surrogate modeling for high-throughput density functional theory data, positioning it as an active reshaping force in materials screening paradigms rather than a passive modeling limitation. Through theoretical reasoning and structured literature, we have identified persistent tensions between efficiency, fidelity, and interpretability, and proposed novel distinctions that integrate artificial intelligence, first-principles computation, and discovery logic within a unified conceptual framework.

The reframing and epistemic commitments articulated herein highlight how uncertainty informs both decision confidence and discovery trajectories, enabling adaptive screening strategies that better reflect the epistemic realities of computational materials science. By emphasizing uncertainty as an interpretive signal, this work contributes conceptual structures that support more trustworthy and theoretically coherent materials identification, even in the absence of new formal models or algorithms.

More broadly, this perspective invites continued theoretical exploration into how uncertainty is conceptualized, communicated, and operationalized across scales and methodologies. In doing so, it strengthens the conceptual foundations of AI-assisted materials science. It encourages a shift toward discovery paradigms that explicitly acknowledge and leverage uncertainty as a core component of scientific reasoning.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 19 Apr 2023 Revised: 12 Jun 2023 Accepted: 08 Sep 2023
Published online: 18 January 2024

References

Choudhary K, Garrity KF, Sharma V, Biacchi AJ, Hight Walker AR, Tavazza F. High-throughput density functional perturbation theory and machine learning predictions of infrared, piezoelectric, and dielectric responses. npj Comput Mater. 2020;6:64.

Afzal MA, Browning AR, Goldberg A, Halls MD, Gavartin JL, Morisato T, et al. High-throughput molecular dynamics simulations and validation of thermophysical properties of polymers for various applications. ACS Appl Polym Mater. 2021;3:620-30.

Farache DE, Verduzco JC, McClure ZD, Desai S, Strachan A. Active learning and molecular dynamics simulations to find high melting temperature alloys. Comput Mater Sci. 2022;209:111386.

Fiedler L, Shah K, Bussmann M, Cangi A. Uncertainty quantification in atomistic modeling of metals and minerals. J Phys Chem A. 2022;126:1631-42.

Fiedler L, Shah K, Bussmann M, Cangi A. Optimal high-throughput virtual screening pipeline for efficient selection of redox-active organic materials. iScience. 2022;25:105702.

Zuo Y, Qin M, Chen C, Ye W, Li X, Luo J, et al. Accelerating materials discovery with Bayesian optimization and graph deep learning. arXiv [preprint]. arXiv:2104.10242. 2021.

Janet JP, Duan C, Yang T, Nandy A, Kulik HJ. Methods for comparing uncertainty quantifications for material property predictions. Mach Learn Sci Technol. 2020;1:025006.

Tran K, Neiswanger W, Broderick K, Xing E, Schneider J, Ulissi ZW. Computational catalyst discovery: Active classification through myopic multi-scale sampling. J Chem Phys. 2021;154:124118.

Ren Z, Li J, Liu Q, Ren R, Tao X, Liu J. High-throughput experimentation and computational freeway approaches to materials synthesis and materials discovery. Adv Energy Mater. 2022;12:2102678.

Kaundinya PR, Choudhary K, Kalidindi SR. Machine learning approaches for feature engineering of the crystal structure: Application to the prediction of the formation energy of cubic compounds. Phys Rev Mater. 2021;5:063802.

Zhang H, Chen W, Iyer A, Apley DW, Chen W. Uncertainty-aware mixed-variable machine learning for materials design. Sci Rep. 2022;12:19760.

Li J, Lim K, Yang H, Ren Z, Raghavan S, Chen P-Y, et al. AI applications through the whole life cycle of material discovery. Matter. 2020;3:393-432.

Liu B, Mason DE, Frankel J, Spletzer M. Machine learning for materials design and discovery. J Appl Phys. 2021;129:070401.

Saal JE, Oliynyk AO, Meredig B. Machine learning in materials discovery: Confirmed predictions and their underlying approaches. Annu Rev Mater Res. 2020;50:49-69.

Menon D, Ranganathan R. A generative approach to materials discovery, design, and optimization. ACS Omega. 2022;7:25958-73.

Baird SG, Diep TQ, Sparks TD. DiSCoVeR: A materials discovery screening tool for high performance, unique chemical compositions. Digital Discovery. 2022;1:226-40.

Baird SG, Liu M, Sayeed HM, Sparks TD. Data-driven materials discovery and synthesis using machine learning methods. arXiv [preprint]. arXiv:2202.02380. 2022.

Rao Z, Tung P-Y, Xie R, Wei Y, Zhang H, Ferrari A, et al. Machine learning–enabled high-entropy alloy discovery. Science. 2022;378:78-85.

Deshwal A, Simon CM, Doppa JR. A universal similarity based approach for predictive uncertainty quantification in materials science. Sci Rep. 2022;12:14650.

Tavazza F, DeCost B, Choudhary K. Uncertainty prediction for machine learning models of material properties. ACS Omega. 2021;6:32431-40.

Tavazza F, DeCost B, Choudhary K. Uncertainty prediction for machine learning models of material properties. arXiv [preprint]. arXiv:2107.07997. 2021.

Mannodi-Kanakkithodi A, Chan MKY. Computational data-driven materials discovery. Trends Chem. 2021;3:79-82.

Pyzer-Knapp EO, Pitera JW, Staar PWJ, Takeda S, Laino T, Sanders DP, et al. Accelerating materials discovery using artificial intelligence, high performance computing and robotics. npj Comput Mater. 2022;8:84.

Wang A, Kingsbury R, McDermott M, Horton M, Jain A, Ong SP, et al. A framework for quantifying uncertainty in DFT energy corrections. ChemRxiv [preprint]. 2021.

Fiedler L, Shah K, Bussmann M, Cangi A. Deep dive into machine learning density functional theory for materials science and chemistry. Phys Rev Mater. 2022;6:040301.

Omar HO, del Cueto M, Nematiaram F, Troisi A. High-throughput virtual screening for organic electronics: A comparative study of alternative strategies. J Mater Chem C. 2021;9:13557-83.

Lei B, Kirk TQ, Bhattacharya A, Pati D, Qian X, Arroyave R, et al. Bayesian optimization with adaptive surrogate models for automated experimental design. npj Comput Mater. 2021;7:194.

Kotthoff L, Wahab H, Johnson PA. Bayesian optimization in materials science: A survey. 2021.

Deshwal A, Simon CM, Doppa JR. Bayesian optimization of nanoporous materials. Mol Syst Des Eng. 2021;6:1066-86.

Shields BJ, Stevens J, Li J, Parasram M, Damani F, Martinez Alvarado JI, et al. Bayesian reaction optimization as a tool for chemical synthesis. Nature. 2021;590:89-96.

Wen J, Odelius M, Sun J. Bayesian optimization of high-entropy alloy compositions for electrocatalytic oxygen reduction. Angew Chem Int Ed. 2021;60:23299-306.

Khatamsaz D, Vela B, Singh P, Johnson DD, Allaire D, Arróyave R. Multi-objective materials Bayesian optimization with active learning of design constraints: Design of ductile refractory multi-principal-element alloys. Acta Mater. 2022;236:118133.

Wang Y, Chen T-Y, Vlachos DG. NEXTorch: A design and Bayesian optimization toolkit for chemical sciences and engineering. J Chem Inf Model. 2021;61:5312-19.

Noh J, Gu GH, Kim S, Jung Y. Uncertainty-quantified hybrid machine learning/density functional theory high throughput screening method for crystals. J Chem Inf Model. 2020;60:1995-2003.

Ren E, Guilbaud P, Coudert F-X. High-throughput computational screening of nanoporous materials in targeted applications. Digit Discov. 2022;1:355-74.