

REVIEW

Open access

Uncertainty Quantification in Computational Materials Engineering: Methods and Deployment Contexts

Ahmed Mansour^{1*}, Omar Saeed¹

Abstract

Computational materials engineering has undergone a transformative shift with the integration of data-driven methodologies and artificial intelligence, enabling accelerated discovery and design of novel materials. Uncertainty quantification (UQ) plays a pivotal role in this paradigm, addressing inherent variabilities in simulations, experimental data, and model predictions to ensure reliable decision-making in materials development. This review synthesizes recent advancements in UQ methods within computational and data-driven materials engineering, focusing on probabilistic modeling, sensitivity analysis, and Bayesian inference techniques deployed across multiscale simulations and machine learning frameworks. We examine deployment contexts ranging from molecular dynamics to additive manufacturing, highlighting how UQ enhances robustness in property prediction, process optimization, and autonomous discovery systems. By integrating insights from high-impact studies the review delineates a systems-level perspective on UQ infrastructures, emphasizing their role in bridging computational predictions with experimental validation. Key challenges such as computational efficiency and data scarcity are contextualized, alongside opportunities for multimodal integration. Ultimately, this synthesis positions UQ as an essential infrastructure for advancing materials informatics toward industrial applicability, offering a forward-looking outlook on scalable, uncertainty-aware workflows in materials engineering.

Keywords Uncertainty quantification, Bayesian optimization, Machine learning potentials, Computational materials science, Data-driven engineering, Multiscale modeling

*Correspondence:

Ahmed Mansour
ahmed.mansour@gmail.com

¹ Department of Materials Engineering and Data Modeling, Faculty of Engineering, Cairo University, Cairo, Egypt

Introduction

Computational materials engineering has undergone a profound epistemic and technological transformation over the past several decades, evolving from physics-constrained deterministic modeling toward hybrid computational–data ecosystems driven by artificial intelligence (AI) and machine learning (ML) [1, 2]. This transition reflects not merely an expansion of computational capability but a restructuring of how materials knowledge is generated, validated, and operationalized within discovery pipelines.

Historically, materials design was governed by empiricism—iterative trial-and-error experimentation guided by heuristic thermodynamic and metallurgical principles. While foundational breakthroughs emerged from this paradigm, the approach was inherently resource-intensive, time-bound, and constrained by narrow exploration bandwidths. The integration of atomistic simulation frameworks, particularly density functional theory (DFT) and molecular dynamics (MD), marked a pivotal inflection point in this trajectory [3, 4]. These computational tools enabled predictive interrogation of materials

behavior at electronic and molecular scales, facilitating insights into bonding environments, defect energetics, diffusion kinetics, and phase stability landscapes.

Yet, despite their predictive power, such simulation frameworks are intrinsically approximation-laden. Exchange–correlation functional selections, finite supercell representations, boundary condition assumptions, and numerical convergence thresholds collectively introduce epistemic and computational uncertainties into simulated outputs [5, 6]. These uncertainties propagate through downstream modeling layers, influencing thermodynamic extrapolations, property predictions, and design recommendations. Consequently, uncertainty quantification (UQ) has emerged as a critical methodological scaffold for interpreting simulation reliability and bounding predictive confidence.

The subsequent rise of data-driven materials science has further amplified the centrality of UQ. Large-scale open repositories—including high-throughput computational databases such as the Materials Project and AFLOW—have transformed materials engineering into a data-abundant discipline [7, 8]. Machine learning models trained on these repositories now predict functional properties spanning electronic bandgaps, elastic moduli, ionic conductivity, catalytic activity, and thermal transport coefficients. However, predictive acceleration introduces new epistemic vulnerabilities.

Two principal uncertainty classes dominate ML-enabled materials modeling. Epistemic uncertainty arises from model incompleteness, limited training coverage, and representation bias—particularly acute in underexplored compositional or structural regimes [3, 9]. Aleatoric uncertainty, by contrast, reflects intrinsic stochasticity within experimental measurements, synthesis variability, and environmental fluctuations. Distinguishing and quantifying these uncertainty forms is essential for calibrating predictive trustworthiness.

To address these challenges, probabilistic ML frameworks—including dropout neural networks, Bayesian neural architectures, and Gaussian process regression—have been integrated into materials informatics pipelines [3, 10, 11]. These methods generate predictive distributions rather than point estimates, enabling uncertainty propagation across screening workflows. Such probabilistic inference supports risk-aware decision-making in candidate selection, experimental prioritization, and inverse design targeting.

In manufacturing contexts, particularly additive manufacturing (AM), UQ assumes an operationally critical role. AM processes exhibit pronounced sensitivity to thermal gradients, scan strategies, powder morphology, and layer deposition kinetics. UQ frameworks enable probabilistic mapping between process parameters and resulting microstructural or mechanical properties, supporting robust process optimization [6, 10, 12]. Integrated computational materials engineering (ICME) platforms, augmented by CALPHAD thermodynamic modeling, further embed UQ to reduce compositional and phase evolution uncertainties in alloy fabrication pathways [5, 12].

Parallel advances have emerged in molecular and atomistic modeling domains. Active learning strategies, informed by UQ metrics, iteratively identify high-uncertainty configurational regions for targeted simulation sampling [9, 13]. This adaptive sampling paradigm significantly enhances interatomic potential development, enabling high-accuracy models to be trained with reduced computational expenditure.

Beyond single-scale prediction, UQ is increasingly central to multiscale materials modeling. Hierarchical frameworks propagate uncertainties originating at the quantum mechanical level through mesoscopic microstructure evolution and ultimately into macroscopic performance predictions [14, 15]. For example, in woven fiber composites, variability in fiber orientation distributions, matrix–fiber interfacial properties, and manufacturing defects can be probabilistically modeled, producing structurally reliable performance envelopes rather than deterministic strength estimates [14].

This multiscale uncertainty propagation underpins the emergence of integrated simulation–experimental ecosystems. Self-driving laboratories exemplify this convergence, wherein robotic synthesis platforms, high-throughput characterization tools, and AI inference engines operate within closed-loop discovery architectures [16, 17]. Within such

systems, UQ guides experimental acquisition, model retraining, and hypothesis refinement, enabling iterative convergence toward optimized materials solutions.

Motivated by the escalating complexity of materials systems and the increasing reliance on AI-mediated discovery infrastructures, this review examines the methodological landscape and deployment contexts of UQ in computational and data-driven materials engineering. We synthesize literature, with emphasis on review and perspective contributions that articulate conceptual advances, methodological innovations, and translational deployment strategies [1, 16, 18, 19].

The analytical scope encompasses probabilistic modeling frameworks, Bayesian inference systems, sensitivity analysis methodologies, and uncertainty propagation techniques across property prediction, inverse design, and autonomous experimentation workflows. Empirical benchmarking studies and hardware-specific optimization reports are intentionally excluded to maintain a workflow- and systems-oriented synthesis.

By situating UQ as a foundational infrastructural layer rather than a peripheral statistical tool, this review advances an integrative perspective on uncertainty-aware materials engineering. We argue that the future of computational materials science hinges not solely on predictive acceleration but on the co-evolution of reliability, interpretability, and epistemic governance across data generation, model training, and experimental validation ecosystems.

Landscape of Computational and Data-Driven Materials Science

Core infrastructures for materials information

Within computational and data-driven materials science, foundational information infrastructures constitute the epistemic bedrock upon which predictive and generative discovery systems are constructed. These infrastructures extend far beyond passive repositories; they operate as active knowledge environments that structure, validate, and operationalize heterogeneous materials evidence. In this context, uncertainty quantification (UQ) functions as a systemic reliability layer, safeguarding the interpretive robustness, traceability, and downstream applicability of integrated datasets [1, 2].

Large-scale materials repositories—populated through high-throughput density functional theory calculations, combinatorial experimentation, and historical thermodynamic assessments—inevitably encode multi-source variability. Such variability emerges from instrumental noise, calibration drift, sample heterogeneity, finite-size simulation effects, and exchange–correlation functional approximations within electronic structure modeling [5, 7]. Without formal quantification, these uncertainties propagate through data pipelines, embedding hidden biases within machine learning training regimes and inverse design heuristics.

To address these epistemic vulnerabilities, statistical recalibration techniques and global variance decomposition methods are employed to characterize uncertainty distributions across compositional and thermodynamic spaces [18, 20]. Sensitivity analysis frameworks further identify parameter regimes where predictive stability is most susceptible to perturbation, enabling targeted data refinement strategies.

Thermodynamic modeling provides a canonical illustration. Within CALPHAD (CALculation of PHase Diagrams) methodologies, UQ propagates parametric uncertainties embedded in Gibbs energy descriptions across multicomponent phase equilibria. Rather than deterministic phase boundaries, probabilistic phase diagrams emerge, expressing stability fields as confidence intervals conditioned on underlying data fidelity [5, 12]. Such probabilistic thermodynamic maps transform alloy design from a deterministic extrapolation exercise into a risk-aware optimization process.

From an infrastructural systems perspective, contemporary materials data ecosystems increasingly prioritize multimodal convergence. Spectroscopic signatures, electron microscopy imaging, diffraction patterns, and in situ process diagnostics are fused with ab initio simulations to generate context-dense knowledge graphs [21]. UQ frameworks preserve modality-

specific confidence signatures during fusion, enabling interoperable yet epistemically transparent archives. Consequently, materials databases evolve into adaptive discovery substrates rather than static storage systems.

Designs for encoding structures with uncertainty

Representation—or encoding—architectures form the translational interface between raw materials data and machine learning inference engines. These architectures transform atomic configurations, bonding environments, and microstructural motifs into structured mathematical embeddings suitable for predictive modeling. However, representational compression inevitably introduces ambiguity, making UQ indispensable for diagnosing encoding fidelity and interpretive stability [3, 11].

Neural interatomic potentials augmented with dropout regularization exemplify uncertainty-aware encoding. By stochastically deactivating network weights during inference, dropout approximates Bayesian posterior sampling, generating predictive distributions rather than deterministic outputs. In atomistic simulations, such probabilistic force fields enable confidence-calibrated molecular dynamics trajectories, particularly valuable in extrapolative thermodynamic regimes [3].

Multi-fidelity encoding frameworks extend representational robustness by integrating datasets of heterogeneous precision. Low-fidelity empirical potentials, semi-empirical quantum approximations, and high-accuracy *ab initio* calculations are hierarchically fused. UQ mechanisms calibrate fidelity weighting, ensuring that predictive embeddings reflect both computational efficiency and physical reliability [7, 11]. This hierarchical encoding is critical for scaling discovery workflows across expansive chemical design spaces.

Graph neural networks (GNNs), widely deployed for molecular and crystalline property prediction, further embed UQ within architecture optimization. Automated neural architecture search platforms increasingly incorporate Bayesian hyperparameter tuning, wherein uncertainty estimates guide topology selection, message-passing depth, and feature aggregation strategies [19]. Such integration produces encoding systems that are not only predictive but epistemically self-aware.

Collectively, these encoding designs enable uncertainty-conscious representation learning—an essential capability for managing sparse datasets, noisy experimental inputs, and multimodal descriptor heterogeneity within materials informatics ecosystems [9].

Forecasting properties via AI with reliability checks

AI-driven property prediction has dramatically accelerated materials screening, enabling rapid evaluation of millions of hypothetical compounds across mechanical, electronic, catalytic, and thermal domains. Yet, predictive throughput without calibrated reliability risks epistemic overconfidence. UQ therefore emerges as the interpretive counterweight to algorithmic acceleration [1, 13].

Active learning infrastructures exemplify this synthesis. Within these iterative frameworks, machine learning models identify high-uncertainty regions within chemical or process design spaces and selectively query new simulations or experiments. This targeted sampling strategy minimizes redundant data acquisition while maximizing knowledge gain—an approach particularly effective in developing interatomic potentials and phase stability predictors [9, 13].

In additive and advanced manufacturing environments, UQ-integrated ML models quantify variability in microstructural evolution, porosity formation, and residual stress accumulation. By embedding probabilistic forecasts within process simulations, these systems enable reliability-aware certification of fabricated components [6, 10].

Bayesian optimization frameworks further institutionalize uncertainty within design exploration. By representing objective landscapes—such as catalytic activity or alloy strength—as probabilistic functions, these methods balance exploratory

sampling of uncertain regions with exploitative refinement of known optima [16, 17]. Advanced high-dimensional variants incorporate heteroscedastic noise modeling, constraint-aware acquisition functions, and adaptive kernel learning to enhance optimization fidelity. The most common UQ method families used across simulation- and ML-driven prediction workflows, along with their typical roles and limitations, are summarized in Table 1.

Table 1. Uncertainty Quantification Methods in Computational Materials Engineering: Primary Function, Strengths, and Limitations

UQ method family	Typical implementation examples	Primary role in materials workflows	Key strengths	Common limitations / failure modes
Bayesian inference	Bayesian calibration; Bayesian neural models	Posterior parameter confidence; model-form uncertainty	Principled uncertainty semantics; supports decision-making	Computationally intensive; prior sensitivity
Gaussian processes (GPs)	GP regression; surrogate modeling	Sample-efficient prediction + uncertainty for BO/AL	Strong in low–medium data; well-calibrated intervals	Poor scaling with dataset size; kernel dependence
Deep UQ via dropout / ensembles	MC dropout; deep ensembles	Predictive intervals for ML potentials & predictors	Simple to deploy; strong empirical performance	Can miscalibrate in OOD regimes; ensemble cost
Sampling-based propagation	Monte Carlo; bootstrapping	Uncertainty propagation through pipelines	Flexible; can approximate complex distributions	Prohibitive in high dimensions; slow convergence
Sensitivity analysis	Sobol indices; variance decomposition	Identify dominant uncertainty contributors	Interpretability; guides data collection	Requires sampling; may miss interactions if simplified
Multi-fidelity UQ	Hierarchical surrogates; fidelity-weighted learning	Balance accuracy/cost; calibrate fidelity layers	Efficient scaling; integrates heterogeneous evidence	Cross-fidelity bias; mis-weighting risk
Calibration & reliability assessment	Reliability diagrams; conformal checks	Align predicted confidence with observed error	Improves trust and deployment readiness	Needs validation data; may not fix model bias

Through these infrastructures, AI property prediction transitions from deterministic estimation toward probabilistic decision intelligence, delivering reliability-calibrated forecasts across polymers, ceramics, alloys, and hybrid materials systems [4, 20].

Structures for reverse engineering materials

Inverse design architectures invert conventional forward prediction by targeting structural configurations that realize predefined functional properties. This inversion amplifies epistemic uncertainty because multiple candidate structures may satisfy identical performance constraints. UQ thus becomes central to solution ranking, feasibility filtering, and experimental prioritization [2, 16].

Generative modeling systems—including variational autoencoders, generative adversarial networks, and diffusion models—construct latent design spaces encoding structural variability. UQ metrics embedded within these latent manifolds evaluate generative confidence, novelty, and physical plausibility, distinguishing viable materials from interpolation artifacts [11, 19].

Applications in metal–organic framework discovery illustrate the utility of uncertainty-guided inverse design. Machine learning systems trained to detect adsorption-active motifs incorporate UQ scoring to prioritize candidates exhibiting both high predicted storage capacity and robust confidence margins [21]. This dual-criterion ranking improves experimental translation efficiency.

Adaptive inverse design pipelines further integrate statistical inference and uncertainty sampling, dynamically refining generative trajectories based on posterior updates [9, 16]. Multiscale UQ propagation frameworks extend this capability across modeling hierarchies, transmitting uncertainty signals from quantum simulations to continuum performance predictions [14, 15]. Such propagation ensures that inverse solutions remain robust under scale transitions, enhancing synthesis viability.

Fusion of diverse data modes

Multimodal data fusion represents the integrative frontier of computational materials science, uniting computational predictions, experimental measurements, geometric modeling outputs, and process telemetry within unified discovery environments. UQ functions as the harmonizing mechanism enabling interoperability across disparate data modalities [18, 22].

Parametric geometric representations, used in architected and metamaterial design, employ UQ to quantify structural variability and performance sensitivity across shape manifolds [22]. In stochastic dynamical systems, machine learning models propagate uncertainty across temporal processing trajectories, supporting predictive control of synthesis pathways [23].

Extreme-state materials research provides further illustration. Density response investigations in warm dense matter integrate experimental scattering observations with quantum simulations, with UQ reconciling theoretical and empirical discrepancies [24]. Sequential multi-fidelity Bayesian optimization frameworks similarly quantify cross-modal uncertainties within coupled experimental–computational loops [25].

Parallel methodologies in environmental and structural systems modeling—such as watershed simulations and infrastructure reliability analysis—deploy stochastic UQ to integrate spatially distributed datasets [26, 27]. When translated into materials contexts, these approaches enable robust multiscale fusion of imaging, spectroscopy, and performance testing evidence.

Through this integrative architecture, UQ emerges as the epistemic connective tissue binding multimodal materials ecosystems into coherent, confidence-aware discovery systems.

Self-sustaining and iterative discovery mechanisms

Self-sustaining discovery mechanisms represent the operational apex of computational materials science, integrating artificial intelligence, robotic automation, and high-fidelity simulation into closed-loop learning ecosystems. Within these autonomous infrastructures, UQ operates as the decision-theoretic engine governing iteration, prioritization, and convergence dynamics [16, 17].

Autonomous laboratories operationalize the conceive–synthesize–characterize–analyze cycle through robotic experimentation platforms. UQ-guided acquisition functions rank candidate experiments based on expected knowledge gain, thereby optimizing resource allocation and accelerating discovery trajectories [9].

Active learning algorithms embedded within these platforms leverage probabilistic outputs from Gaussian processes, Bayesian neural networks, or ensemble models to select subsequent experimental conditions. Such strategies have demonstrated effectiveness in layered materials optimization and thin-film process calibration [6, 13].

Simulation–experiment coupling further embeds UQ within translational workflows. Variabilities in molecular dynamics simulations or quantum calculations inform robotic synthesis parameters through uncertainty-aware control interfaces [3, 4]. For instance, in glass transition studies, UQ-calibrated atomistic simulations guide automated temperature ramping and cooling schedules [4, 20].

In additive manufacturing, iterative process refinement loops integrate UQ to dynamically adjust laser power, scan velocity, and thermal gradients—minimizing defect formation in alloy fabrication [10, 12]. These adaptive control systems transform manufacturing environments into learning entities capable of self-optimization.

Simulation–experiment coupling leverages UQ for discrepancy modeling, ensuring alignment between virtual and real-world outcomes [5, 15]. In woven composites, multiscale UQ propagates errors across loops [14]. Bayesian frameworks in these systems formalize the cycle as:

$$\theta_{t+1} = \arg \max_{\theta} [U(\theta) + \alpha \cdot I(\theta)] \quad (1)$$

where $U(\theta)$ is the utility function, $I(\theta)$ the information metric from UQ, and α a balancing parameter [16, 17]. This conceptual formula synthesizes exploration-exploitation in autonomous workflows. This closed-loop structure can be summarized as an uncertainty-aware discovery architecture linking data ecosystems, probabilistic modeling, decision policies, and experimental feedback (Figure 1).

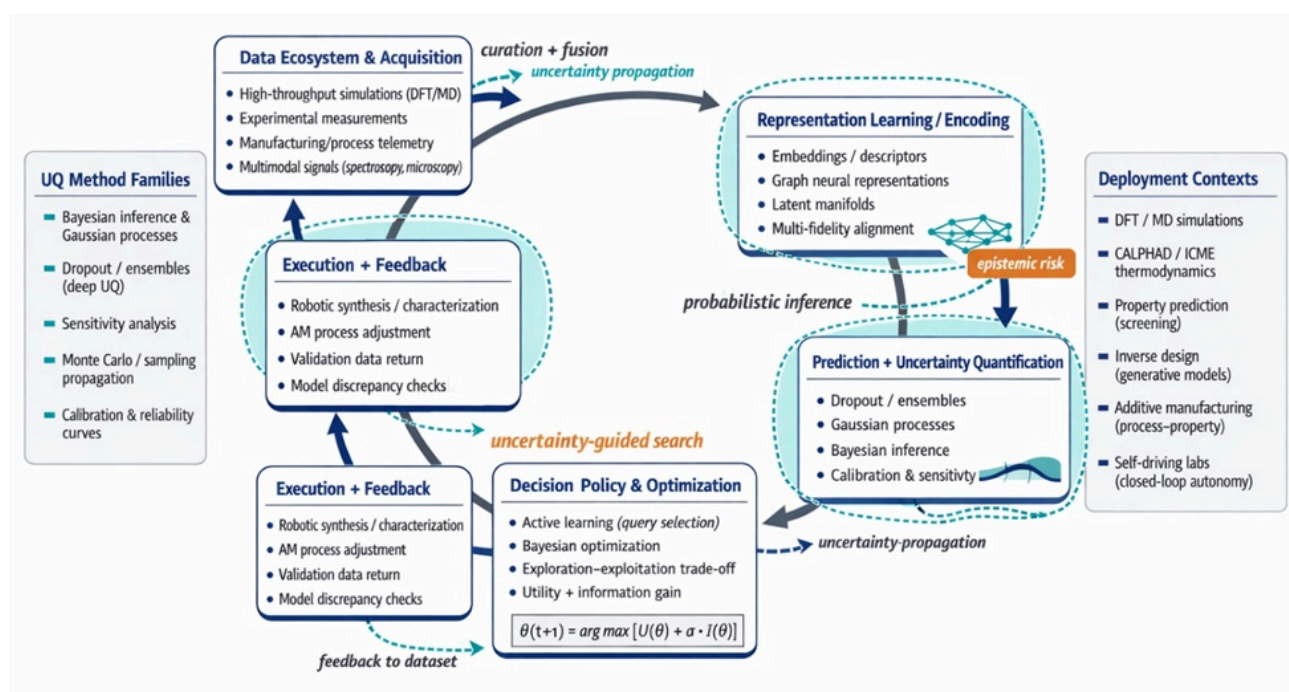


Figure 1. Uncertainty-Aware Closed-Loop Materials Discovery Architecture Across Simulation–AI–Experiment Workflows

These systems exemplify UQ's deployment in dynamic contexts, fostering acceleration in materials innovation [1, 2, 11].

Results and Discussion

The integration of uncertainty quantification (UQ) within computational and data-driven materials engineering has catalyzed a fundamental shift in how predictive reliability, model interpretability, and discovery risk are conceptualized across simulation–AI ecosystems [1, 3, 5, 10]. Rather than functioning as a post hoc statistical correction, UQ now operates as a structural design principle embedded within modeling architectures, data infrastructures, and experimental coupling frameworks. This systemic embedding reveals nuanced interdependencies between data-driven inference engines and physics-grounded simulation pipelines.

To clarify how uncertainty signals translate into actionable decisions across major materials workflows, we summarize key deployment contexts, uncertainty types, and decision outputs in Table 2.

Table 2. Deployment Contexts for UQ Across the Materials Discovery Pipeline: Inputs, Uncertainty Types, and Decision Outputs

Deployment context	Typical inputs	Dominant uncertainty types	What UQ produces	Primary decision enabled
DFT / electronic-structure screening	Structures, compositions, computed energies	Epistemic (functional choice); numerical	Confidence bounds; discrepancy estimates	Candidate ranking; screening thresholds
Molecular dynamics & ML potentials	Configurations, forces, trajectories	Epistemic (coverage); aleatoric (thermal noise)	Predictive variance; OOD flags	Active learning queries; safe extrapolation
CALPHAD / ICME thermodynamics	Thermodynamic datasets, phase equilibria	Parametric + model-form	Probabilistic phase diagrams; stability intervals	Alloy composition optimization; process windows
Property prediction (ML screening)	Descriptors/embeddings, training corpora	Epistemic (data gaps); aleatoric (label noise)	Calibrated predictive intervals	Risk-aware selection; triage for experiments
Inverse design (generative modeling)	Target properties; latent representations	Epistemic (underconstraint); representational ambiguity	Confidence scoring for candidates	Prioritize synthesizable designs
Additive manufacturing (process–property)	Process telemetry; microstructure proxies	Aleatoric (process variability); epistemic (model mismatch)	Uncertainty-aware process maps	Parameter tuning; defect minimization
Self-driving labs (closed-loop autonomy)	Mixed simulation + experiment streams	Time-varying epistemic + aleatoric	Information gain metrics; acquisition functions	Next experiment selection; autonomous convergence
Multimodal fusion (imaging + spectroscopy + sim)	Multimodal signals + simulations	Cross-modal mismatch; spatial/temporal variability	Harmonized confidence layers	Data integration; robust digital twin updates

A central point of discussion concerns the scalability and contextual adaptability of UQ methodologies across deployment domains. In molecular and atomistic simulations, dropout-enabled neural network potentials offer computationally efficient

uncertainty estimation by approximating Bayesian inference through stochastic weight sampling [3, 4]. These approaches are particularly effective for interpolative regimes within well-represented configurational spaces. However, in extrapolative domains—such as high-temperature phase transitions or non-equilibrium defect dynamics—dropout methods may inflate predictive variance, producing overly conservative uncertainty envelopes that complicate decision prioritization.

By contrast, thermodynamic and manufacturing-oriented frameworks—especially CALPHAD-integrated UQ infrastructures—propagate uncertainties across composition–process–performance linkages [5, 6, 12]. In additive manufacturing, this propagation enables probabilistic mapping between alloy chemistry, thermal gradients, microstructural evolution, and mechanical performance metrics. Such system-level uncertainty tracing aligns computational forecasts with real-world fabrication variability, strengthening translational reliability.

This comparative synthesis exposes a fundamental trade-off. Probabilistic modeling enhances interpretive robustness and decision confidence, yet it introduces significant computational overhead. Sampling-intensive inference, posterior estimation, and ensemble modeling increase computational cost, particularly in high-dimensional materials design spaces. Hybrid paradigms—combining multi-fidelity learning, surrogate modeling, and sensitivity analysis—have therefore emerged as pragmatic compromises, preserving uncertainty awareness while maintaining tractable computational demands [7, 11, 15].

A second critical dimension of discussion centers on trust calibration in AI-driven materials workflows. Active learning systems exemplify the operationalization of UQ as a discovery steering mechanism. By identifying high-uncertainty regions within compositional or structural design spaces, these systems guide targeted simulation or experimental acquisition, thereby accelerating materials optimization cycles [9, 13, 16].

However, trust calibration becomes more complex in multimodal and inverse design environments. Sparse datasets, modality misalignment, and representational compression can introduce epistemic distortions that bias generative outputs or inverse predictions [2, 14, 19]. In such contexts, UQ serves a corrective function—quantifying confidence asymmetries and flagging structurally underdetermined design candidates.

Multiscale simulation studies further demonstrate UQ's stabilizing influence. In woven fiber composites and architected geometries, probabilistic propagation of uncertainties—from fiber orientation distributions to mesoscale stress fields—produces performance predictions bounded by realistic confidence intervals [14, 22]. This probabilistic framing transforms structural design from deterministic optimization into risk-aware engineering.

Within autonomous discovery infrastructures, UQ assumes an adaptive calibration role. Self-driving laboratories and robotic experimentation platforms integrate stochastic simulation outputs with experimental feedback loops, enabling dynamic model recalibration under evolving evidence regimes [17, 23]. Through such coupling, uncertainty becomes an active learning signal rather than a passive statistical descriptor.

Collectively, these developments underscore UQ's transformative role in transitioning computational materials science from deterministic prediction paradigms toward probabilistic, evidence-weighted discovery architectures. This transition carries implications for methodological standardization, benchmarking, and governance within materials informatics ecosystems [1, 18, 28].

Challenges

Despite rapid methodological maturation, significant technical and epistemic challenges continue to constrain the deployment of UQ across computational materials engineering workflows.

Computational scalability and cost

Computational efficiency remains a primary barrier. Sampling-intensive UQ methods—including Monte Carlo propagation, Bayesian posterior sampling, and ensemble neural modeling—scale poorly with dimensionality [3, 7, 10]. In high-throughput screening environments, the computational burden of propagating uncertainty across millions of candidate materials can become prohibitive.

Thermodynamic modeling presents analogous constraints. CALPHAD parameter spaces often contain hundreds of coupled variables, and uncertainty propagation across such high-dimensional Gibbs energy landscapes requires advanced optimization and surrogate acceleration strategies to avoid exponential cost escalation [5, 18].

Data scarcity and epistemic instability

Data sparsity amplifies epistemic uncertainty, particularly in emerging materials classes or extreme operating regimes. Machine learning models trained on limited datasets exhibit high variance and poor extrapolative generalization, undermining predictive confidence [1, 9, 13]. This issue is especially acute in inverse design contexts where generative models must operate beyond empirically sampled design spaces.

Cross-scale interoperability

Multiscale integration introduces additional complexity. Propagating uncertainties from quantum mechanical simulations to continuum performance models requires consistent statistical translation across modeling hierarchies [14, 15, 23]. Disparities in boundary conditions, constitutive assumptions, and discretization schemes can amplify uncertainty during scale transitions.

In additive manufacturing, disentangling epistemic uncertainty (model incompleteness) from aleatoric uncertainty (process variability) remains particularly challenging. Thermal fluctuations, powder heterogeneity, and machine noise introduce stochastic variability that is difficult to isolate within predictive models [6, 10, 12].

Experimental validation constraints

Validation of UQ frameworks against empirical observations is complicated by measurement noise and instrumentation limits. Glass transition experiments, density response analyses, and high-temperature phase measurements often contain stochastic fluctuations that obscure ground-truth benchmarking [4, 20, 24]. This complicates calibration of predictive uncertainty envelopes.

Methodological calibration and overconfidence

Ensuring well-calibrated uncertainty estimates remains a methodological frontier. Overconfident models risk catastrophic design errors, whereas excessively conservative estimates impede decision efficiency [11, 17, 19]. In graph neural networks and multi-fidelity architectures, embedding UQ within architecture optimization introduces additional bias risks if uncertainty metrics distort hyperparameter selection [19, 25].

Comparable calibration challenges arise in environmental and structural modeling domains, where spatial uncertainty quantification must reconcile heterogeneous data resolutions and measurement inconsistencies [26, 27].

Addressing these intersecting challenges requires coordinated advances in algorithm design, benchmark standardization, and cross-disciplinary methodological transfer.

Future research directions

Future research trajectories in uncertainty-aware computational materials engineering will likely coalesce around hybridization, autonomy, and infrastructural standardization.

Hybrid physics–AI uncertainty frameworks

Integrating deep learning with physics-informed modeling offers a promising pathway for scalable UQ deployment [1, 2, 11]. Physics-constrained neural networks, surrogate thermodynamic models, and mechanistically regularized generative systems can embed physical priors within probabilistic inference architectures—enhancing both interpretability and uncertainty calibration.

Next-generation active learning and inverse design

Advancing active learning frameworks through richer uncertainty metrics—such as information gain, epistemic entropy, and ensemble disagreement—could dramatically improve data efficiency in inverse design pipelines [9, 16, 19]. Generative models require robust uncertainty bounding mechanisms to ensure physical plausibility of synthesized candidates.

Extreme-condition and quantum-informed UQ

Emerging research domains—including warm dense matter, radiation-tolerant materials, and ultra-high-pressure systems—demand quantum-informed UQ frameworks capable of handling extreme thermodynamic regimes [23, 24]. Stochastic electronic structure modeling and uncertainty-aware quantum simulations represent critical frontiers.

Autonomous and closed-loop discovery

Real-time UQ integration within self-driving laboratories will enable adaptive experiment–simulation coupling at unprecedented temporal resolution [13, 15, 17]. Closed-loop Bayesian optimization platforms capable of updating uncertainty posteriors during live experimentation will accelerate convergence toward optimal materials solutions.

Multi-fidelity and multiscale optimization

Future multi-fidelity Bayesian optimization frameworks must accommodate dynamically evolving uncertainty landscapes across additive manufacturing, composites engineering, and hierarchical materials design [10, 12, 14, 25]. Coupling process telemetry with simulation priors will enhance predictive stability.

Multimodal data fusion and spatial–temporal UQ

Developing multimodal UQ architectures capable of integrating spatial imaging, temporal processing data, and atomistic simulations remains a critical frontier [22, 26, 27]. Such frameworks will underpin digital twins and predictive manufacturing ecosystems.

Standardization and open infrastructure

Establishing standardized UQ protocols, benchmark datasets, and open-source validation platforms will be essential for reproducibility and industrial adoption [5, 18, 28]. Community infrastructures analogous to open materials databases could host uncertainty-annotated datasets and model calibration benchmarks.

Translational and industrial integration

Finally, embedding UQ within high-throughput screening pipelines, autonomous synthesis platforms, and industrial ICME workflows will accelerate deployment into sustainable materials manufacturing, energy systems, and structural engineering applications [6, 7, 21].

Conclusion

Uncertainty quantification stands as a cornerstone in advancing computational and data-driven materials engineering, providing the necessary rigor for reliable predictions and optimizations. This review has synthesized key methods and deployment contexts, from molecular potentials to autonomous systems, illustrating UQ's role in enhancing workflow robustness. By addressing variabilities through probabilistic and active learning approaches, UQ bridges computational

efficiency with experimental fidelity. Despite challenges in scalability and interoperability, future directions promise integrated, uncertainty-aware infrastructures that will propel materials discovery. In essence, UQ not only mitigates risks but empowers transformative applications in materials science.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 09 Sep 2021

Revised: 25 Oct 2021

Accepted: 17 Jan 2022

Published online: 18 March 2022

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ramprasad R, Batra R, Pilania G, Mannodi-Kanakthodi A, Kim C. Machine learning in materials informatics: Recent applications and prospects. *npj Comput Mater*. 2017;3(1):54.
<https://doi.org/10.1038/s41524-017-0056-5>.
- Pilania G, Gubernatis JE, Lookman T. Multi-fidelity machine learning models for accurate bandgap predictions of solids. *Comput Mater Sci*. 2017;129:156-63.
<https://doi.org/10.1016/j.commatsci.2016.12.004>.
- Wen M, Tadmor EB. Uncertainty quantification in molecular simulations with dropout neural network potentials. *npj Comput Mater*. 2020;6(1):124.
<https://doi.org/10.1038/s41524-020-00390-8>.

Vu-Bich H, Nguyen-Trung C, Pham D, Le T, Nguyen H, Ngo N, et al. Uncertainty quantification in molecular dynamics studies of the glass transition temperature. *Acta Mater.* 2017;128:344-52.
<https://doi.org/10.1016/j.actamat.2017.02.016>.

Otis R. Uncertainty reduction and quantification in computational thermodynamics. *Comput Mater Sci.* 2022;212:111590.
<https://doi.org/10.1016/j.commatsci.2022.111590>.

Hu Z, Mahadevan S. Uncertainty quantification in prediction of material properties during additive manufacturing. *Acta Mater.* 2019;176:37-48.
<https://doi.org/10.1016/j.actamat.2019.06.041>.

Bassman L, Hexner D, Owen L, Zok F, Hexemer A, Ushizima D, et al. Active learning for accelerated design of layered materials. *npj Comput Mater.* 2018;4(1):74.
<https://doi.org/10.1038/s41524-018-0129-7>.

Zhang L, Lin DY, Wang H, Car R, Weinan E. Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys Rev Mater.* 2019;3(2):023804.

Xue D, Balachandran PV, Hogden J, Theiler J, Xue D, Lookman T. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput Mater.* 2019;5(1):28.
<https://doi.org/10.1038/s41524-019-0153-8>.

Wang Z, Jiang C, Liu P, Yang W, Zhao Y, Horstemeyer MF, et al. Uncertainty quantification and reduction in metal additive manufacturing. *npj Comput Mater.* 2020;6(1):175.
<https://doi.org/10.1038/s41524-020-00444-x>.

Tran A, Furlan JM, Pagalthivarthi KV, Viswanathan R, Wildey T, Wang Y. eQual: Uncertainty quantification for material property predictions with theory-guided machine learning. *Adv Mater.* 2022;34(21):2200983.
<https://doi.org/10.1002/adma.202200983>.

Wang Z, Jiang C, Liu P, Yang W, Zhao Y, Horstemeyer MF, et al. Uncertainty quantification and composition optimization for alloy additive manufacturing through a CALPHAD-based ICME framework. *npj Comput Mater.* 2020;6(1):188.
<https://doi.org/10.1038/s41524-020-00454-9>.

Zhang L, Lin DY, Wang H, Car R, Weinan E. Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys Rev Mater.* 2019;3(2):023804.
<https://doi.org/10.1103/PhysRevMaterials.3.023804>.

Zheng P, Rongvong W, Dong W, Lopez J, Hu Z. Uncertainty quantification in multiscale simulation of woven fiber composites. *Comput Methods Appl Mech Eng.* 2018;338:506-32.
<https://doi.org/10.1016/j.cma.2018.04.024>.

Honarmandi P, Arróyave R. Uncertainty quantification and propagation in computational materials science and simulation-assisted materials design. *Acta Mater.* 2020;190:1-14.
<https://doi.org/10.1016/j.actamat.2020.03.012>.

Lookman T, Balachandran PV, Xue D, Hogden J, Theiler J. Statistical inference and adaptive design for materials discovery. *Curr Opin Solid State Mater Sci.* 2017;21(3):121-28.
<https://doi.org/10.1016/j.cossms.2016.10.002>.

Ling J, Hutchinson M, Antono E, Paradiso S, Meredig B. High-dimensional materials and process optimization using data-driven experimental design with well-calibrated uncertainty estimates. *Integr Mater Manuf Innov.* 2017;6(3):207-17.
<https://doi.org/10.1007/s40192-017-0098-z>.

Pernot P, Cailliez F. A critical review of statistical calibration modelling in property calculation and QSAR. *Comput Mater Sci.* 2020;173:109473.
<https://doi.org/10.1016/j.commatsci.2019.109473>.

Pernot P. Uncertainty quantification for molecular property predictions with graph neural architecture search. *Digit Discov.* 2022;1(3):266-77.
<https://doi.org/10.1039/D2DD00006A>.

Zhou T, Gui C, Sun L, Hu Y, Lyu H, Wang Z, et al. Energy applications of ionic liquids: Recent developments and future prospects. *Chem Rev.* 2023;123(21):12170-253.

Fernandez M, Boyd PG, Daff TD, Aghaji MZ, Woo TK. Rapid and accurate machine learning recognition of high performing active sites in metal–organic frameworks for H₂ storage. *J Mater Chem A.* 2018;6(5):2432-42.
<https://doi.org/10.1039/C7TA09919A>.

Bostanabad R, Kearney T, Tao S, Apley DW, Chen W. Leveraging the nurbs-based parametric geometry representations for uncertainty quantification of geometric models. *Comput Aided Des.* 2018;104:103-15.
<https://doi.org/10.1016/j.cad.2018.06.002>.

Ouyang Y, Bai L, Tian H, Li X, Yuan F. Recent progress of thermal conductive ploymer composites: Al₂O₃ fillers, properties and applications. *Compos Part A Appl Sci Manuf.* 2022;152:106685.

Dornheim E, Vorberger J, Böhme M, Moldabekov UA, Diessel N, Schlages M, et al. Density response of the warm dense electron gas: models, methods and experiments. *Phys Rep.* 2022;956:1-74.
<https://doi.org/10.1016/j.physrep.2022.01.001>.

Tran A, Wildey T, McCann S. sMF-BO-2CoGP: A sequential multi-fidelity constrained Bayesian optimization framework for design applications. *J Comput Inf Sci Eng.* 2020;20(3):031006.
<https://doi.org/10.1115/1.4045670>.

Foks NL, Starn JJ, Stengel DE, Yager RM. Uncertainty quantification of environmental performance measures and spatial ranking of watersheds using stochastic watershed modeling. *J Hydrol.* 2021;598:126259.
<https://doi.org/10.1016/j.jhydrol.2021.126259>.

Shang Z, Sun Y. Local sensitivity analysis of multiscale structural reliability using separation distance. *Theor Appl Frac Mech.* 2020;105:102417.
<https://doi.org/10.1016/j.tafm.2019.102417>.

Almufarreh A, Arshad M. Promising emerging technologies for teaching and learning: Recent developments and future challenges. *Sustainability.* 2023;15(8):6917.