

Contents

1	Introduction	1
1.1	The Need for Uncertainty Quantification	1
1.2	Types of Uncertainty	1
1.2.1	Aleatoric Uncertainty	2
1.2.2	Epistemic Uncertainty	2
1.3	Classical Approaches to Uncertainty Quantification	2
1.3.1	Bayesian Inference	2
1.3.2	Bootstrap Methods	3
1.3.3	Prediction Intervals from Asymptotic Theory	3
1.4	Modern Machine Learning Approaches	3
1.4.1	Gaussian Processes	3
1.4.2	Bayesian Neural Networks	4
1.4.3	Neural Network Hybrid Methods	4
1.4.4	Conformal Prediction	5
1.5	Challenges in UQ for Regression	5
1.5.1	Calibration Assessment	5
1.5.2	Heteroskedastic Noise	6
1.5.3	Computational Scalability	6
1.5.4	Distribution Shift	6
1.6	Scope of This Work	6
2	References	8

1 Introduction

1.1 The Need for Uncertainty Quantification

Predictive models are ubiquitous in modern science and engineering, from materials discovery and drug development to climate modeling and financial forecasting. However, point predictions alone are insufficient for informed decision-making—we must also quantify our confidence in these predictions.^{1,2} Uncertainty quantification (UQ) transforms predictive models from black boxes into trustworthy tools by providing rigorous estimates of prediction reliability.

The importance of UQ becomes particularly evident in high-stakes applications. In medical diagnostics, a model that predicts disease outcomes must distinguish between cases where it is confident versus uncertain, enabling clinicians to request additional testing when needed.³ In materials science, uncertainty estimates guide experimental resource allocation toward the most informative measurements.⁴ In autonomous systems, quantified uncertainty enables safe operation by triggering human intervention when model confidence is low.

1.2 Types of Uncertainty

The uncertainty quantification literature distinguishes between two fundamentally different sources of uncertainty:^{1,5}

1.2.1 Aleatoric Uncertainty

Aleatoric uncertainty (also called statistical or irreducible uncertainty) represents inherent randomness in the data generation process. This uncertainty arises from:

- Natural variability in measurements (e.g., sensor noise, experimental error)
- Stochastic processes in the underlying system
- Heterogeneous noise that varies across the input space

Importantly, aleatoric uncertainty cannot be reduced by collecting more training data; it represents a fundamental limit on prediction accuracy. For example, if we measure a physical quantity with an instrument that has ± 0.01 precision, this measurement uncertainty persists regardless of how many times we measure.

1.2.2 Epistemic Uncertainty

Epistemic uncertainty (also called model or reducible uncertainty) arises from incomplete knowledge about the true underlying function. Sources include:

- Limited training data in certain regions of input space
- Model misspecification (wrong model family chosen)
- Uncertainty in model parameters

Unlike aleatoric uncertainty, epistemic uncertainty can in principle be reduced by:

- Collecting more representative training data
- Using more flexible model families
- Improving parameter estimation procedures

The word "epistemic" derives from the Greek $\epsilon \pi \sigma \tau \mu \epsilon$ (episteme), meaning knowledge. Epistemic uncertainty reflects what we could know but currently do not.¹

1.3 Classical Approaches to Uncertainty Quantification

1.3.1 Bayesian Inference

Bayesian methods provide a principled probabilistic framework for uncertainty quantification by treating model parameters as random variables.^{6,7} Given training data $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$, Bayesian inference computes a posterior distribution over parameters:

$$p(\boldsymbol{\theta}|\mathcal{D}) = \frac{p(\mathcal{D}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathcal{D})} \quad (1)$$

where $p(\boldsymbol{\theta})$ is the prior distribution encoding initial beliefs, $p(\mathcal{D}|\boldsymbol{\theta})$ is the likelihood, and $p(\mathcal{D})$ is the marginal likelihood (evidence). Predictions for a new input \mathbf{x}^* marginalize over parameter uncertainty:

$$p(y^*|\mathbf{x}^*, \mathcal{D}) = \int p(y^*|\mathbf{x}^*, \boldsymbol{\theta})p(\boldsymbol{\theta}|\mathcal{D})d\boldsymbol{\theta} \quad (2)$$

This predictive distribution naturally captures both epistemic uncertainty (through the parameter posterior) and aleatoric uncertainty (through the likelihood model).

1. Markov Chain Monte Carlo (MCMC)

For complex models where the posterior is intractable, Markov Chain Monte Carlo methods draw samples from $p(\boldsymbol{\theta}|\mathcal{D})$.⁸ While MCMC provides asymptotically exact inference, it can be computationally expensive, requiring thousands of model evaluations. Recent advances include parallel tempering and Hamiltonian Monte Carlo for improved sampling efficiency.

2. Variational Inference

Variational methods approximate the posterior with a simpler distribution $q(\boldsymbol{\theta})$ from a tractable family, minimizing the Kullback-Leibler divergence:⁴

$$q^*(\boldsymbol{\theta}) = \arg \min_{q \in \mathcal{Q}} \text{KL}(q(\boldsymbol{\theta}) || p(\boldsymbol{\theta}|\mathcal{D})) \quad (3)$$

Variational inference trades some accuracy for computational efficiency, making it suitable for large-scale applications.⁸

1.3.2 Bootstrap Methods

Bootstrap resampling provides a distribution-free approach to uncertainty estimation.^{9,10} The procedure:

1. Draw B bootstrap samples $\{\mathcal{D}_b\}_{b=1}^B$ by sampling with replacement from the training data
2. Train a model on each bootstrap sample to obtain parameters $\{\boldsymbol{\theta}_b\}_{b=1}^B$
3. Form prediction intervals using the empirical distribution of predictions

Bootstrap methods are particularly appealing because they make minimal distributional assumptions. However, recent work shows that raw bootstrap standard deviations often underestimate true uncertainty and require calibration.¹¹

1.3.3 Prediction Intervals from Asymptotic Theory

For linear and nonlinear least-squares regression, prediction intervals can be derived from asymptotic normality of parameter estimates. If $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimate with covariance Σ , approximate 95% prediction intervals are:

$$\hat{y} \pm 1.96 \sqrt{\hat{\sigma}^2 + \mathbf{g}(\mathbf{x})^T \Sigma \mathbf{g}(\mathbf{x})} \quad (4)$$

where $\mathbf{g}(\mathbf{x}) = \nabla_{\boldsymbol{\theta}} f(\mathbf{x}, \boldsymbol{\theta})|_{\hat{\boldsymbol{\theta}}}$ is the gradient of the model, $\hat{\sigma}^2$ is the residual variance (aleatoric), and $\mathbf{g}^T \Sigma \mathbf{g}$ captures parameter uncertainty (epistemic).

This approach underlies MATLAB's `nlinfit` function and similar tools, providing computationally efficient uncertainty estimates for parametric models.

1.4 Modern Machine Learning Approaches

1.4.1 Gaussian Processes

Gaussian Process Regression (GPR) has emerged as a gold-standard method for uncertainty quantification.^{12,13} A GP defines a distribution over functions:

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (5)$$

where $m(\mathbf{x})$ is the mean function (often zero) and $k(\mathbf{x}, \mathbf{x}')$ is a covariance kernel encoding smoothness assumptions. Given training data, the posterior predictive distribution is Gaussian:

$$p(f^* | \mathbf{x}^*, \mathcal{D}) = \mathcal{N}(f^* | \mu^*, \sigma^{*2}) \quad (6)$$

with closed-form expressions for μ^* and σ^{*2} .

GPR provides well-calibrated uncertainty estimates and interpolates smoothly between observations. However, it faces scalability challenges for large datasets (cubic complexity in training set size) and can struggle with high-dimensional inputs.¹²

1.4.2 Bayesian Neural Networks

Bayesian Neural Networks (BNNs) extend neural networks by placing probability distributions over weights.^{4,14} Recent advances focus on scalable approximate inference:

- **Variational BNNs** use variational inference with factorized Gaussian weight posteriors
- **Monte Carlo Dropout** treats dropout as approximate Bayesian inference⁵
- **Deep Ensembles** train multiple networks with different initializations¹⁵

A 2024 study comparing BNN methods for materials property prediction found that BNN-MCMC achieved performance competitive with GPR.⁴ Physics-informed BNNs integrate governing equations as soft constraints, improving physical consistency of predictions.¹⁶

1.4.3 Neural Network Hybrid Methods

Recent methods combine neural networks' representational power with classical UQ techniques:

1. Neural Network + Gaussian Mixture Models (NNGMM)

This approach uses a neural network for feature extraction, then models the output distribution as a Gaussian mixture in the learned representation. The mixture model captures multimodal predictive distributions and provides uncertainty estimates through mixture component variances.

2. Neural Network + Bayesian Linear Regression (NNBR)

NNBR extracts nonlinear features using a neural network, then applies Bayesian linear regression in the feature space.¹⁷ This decouples representation learning (neural network) from uncertainty quantification (Bayesian inference), enabling:

- Efficient uncertainty estimation after feature learning
- Post-hoc calibration using validation data
- Analytically tractable predictive distributions

These hybrid approaches often achieve better calibration than end-to-end neural methods while maintaining computational efficiency.

1.4.4 Conformal Prediction

Conformal prediction has emerged as a powerful distribution-free framework for uncertainty quantification.^{18–20} Unlike Bayesian methods, conformal prediction makes no distributional assumptions and provides finite-sample validity guarantees.

The basic split conformal procedure:

1. Split data into training ($\mathcal{D}_{\text{train}}$) and calibration (\mathcal{D}_{cal}) sets
2. Train a model on $\mathcal{D}_{\text{train}}$
3. Compute nonconformity scores $s_i = |y_i - \hat{f}(\mathbf{x}_i)|$ on \mathcal{D}_{cal}
4. For target coverage $1 - \alpha$, find quantile $q = \text{Quantile}_{1-\alpha}(\{s_i\})$
5. Predict test point \mathbf{x}^* with interval $[\hat{f}(\mathbf{x}^*) - q, \hat{f}(\mathbf{x}^*) + q]$

The resulting prediction intervals are guaranteed to achieve $\geq 1 - \alpha$ coverage on exchangeable test data, regardless of the base model.¹⁸

1. Recent Advances

- **Conformalized Quantile Regression (CQR):** Uses quantile regression to produce adaptive prediction intervals²¹
- **Locally Adaptive Conformal Prediction (LACP):** Adjusts interval width based on local data density
- **Robust Conformal Prediction:** Extends coverage guarantees to adversarially perturbed inputs²⁰
- **Conformal Prediction for Dynamics:** Novel algorithms for time-series and dynamical systems¹⁹

A 2025 study of chromatography modeling found that conformal predictors outperformed Gaussian processes for black-box uncertainty quantification.²¹

1.5 Challenges in UQ for Regression

Despite significant progress, several challenges remain:

1.5.1 Calibration Assessment

Unlike classification where calibration metrics are well-established (e.g., expected calibration error), regression lacks consensus on calibration evaluation.²² Common approaches include:

- **Coverage:** Fraction of test points within prediction intervals (should equal nominal level)
- **Interval width:** Narrower intervals are preferable given adequate coverage
- **Calibration curves:** Plot empirical vs. predicted coverage at different confidence levels

However, these metrics can be misleading. A model with perfect average coverage may have poorly calibrated local uncertainties.

1.5.2 Heteroskedastic Noise

Many real-world problems exhibit heteroskedastic noise where variance depends on inputs. Classical methods assuming constant variance (homoskedasticity) produce miscalibrated intervals. Modern approaches address this through:

- Input-dependent noise models
- Quantile regression
- Separate models for mean and variance

1.5.3 Computational Scalability

Bayesian methods (MCMC, BNNs) and Gaussian processes face computational bottlenecks for large datasets. Active areas of research include:

- Variational sparse GPs²³
- Stochastic variational inference for BNNs
- Amortized inference for fast predictions

1.5.4 Distribution Shift

Most UQ methods assume test data follows the training distribution. Under distribution shift (common in deployment), uncertainty estimates can be severely miscalibrated. Robust UQ under shift remains an open challenge.

1.6 Scope of This Work

This study presents a comprehensive benchmark evaluating uncertainty quantification methods across diverse regression scenarios. We systematically compare:

Classical Bayesian Methods

- Bayesian linear and polynomial regression
- Bayesian ridge regression

Nonlinear Parametric Methods

- Nonlinear least squares (`nlinfit`) with asymptotic prediction intervals
- Split conformal prediction

Data-Driven Methods

- Gaussian Process Regression
- Neural Network + Gaussian Mixture Model (NNGMM)
- Neural Network + Bayesian Linear Regression (NNBR)

We evaluate these methods on seven synthetic datasets with known ground truth:

- Linear (simple line, quadratic, cubic polynomials)

- Nonlinear (exponential decay, logistic growth, Michaelis-Menten, Gaussian)

Each dataset is generated with two noise models (homoskedastic, heteroskedastic) at four noise levels (1%, 2%, 5%, 10%), yielding 56 configurations per method.

Performance is assessed using:

- **Coverage:** Empirical coverage of 95% prediction intervals
- **Interval width:** Mean width of prediction intervals
- **Prediction accuracy:** RMSE and R² on test data

This benchmark provides practitioners with evidence-based guidance for selecting UQ methods based on problem characteristics, and identifies strengths and limitations of each approach.

References

- [1] Eyke Hüllermeier and Willem Waegeman. Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. *Machine Learning*, 110(3):457–506, 2021. doi: 10.1007/s10994-021-05946-3.
- [2] Sunday Ochella, Fateme Dinmohammadi, and Mahmood Shafiee. Bayesian neural networks for uncertainty quantification in remaining useful life prediction of systems with sensor monitoring. *Journal of Engineering*, 2024, 2024. doi: 10.1177/16878132241239802.
- [3] Melanie F Kim, Jakub M Tomczak, Laurens Aanen, and Ans MW Cohen. Deep bayesian gaussian processes for uncertainty estimation in electronic health records. *Scientific Reports*, 11:20685, 2021. doi: 10.1038/s41598-021-00144-6.
- [4] Shuhao Lee, Colton Purdy, Manav Singh, Katharine Kalur, Nasr Almasri, Dingcheng Ni, and Jason Hattrick-Simpers. Uncertainty quantification in multivariable regression for material property prediction with bayesian neural networks. *Scientific Reports*, 14:14484, 2024. doi: 10.1038/s41598-024-61189-x.
- [5] Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? In *Advances in Neural Information Processing Systems*, volume 30, 2017.
- [6] Jeremy Oakley. Introduction to uncertainty quantification and gaussian processes. GPSS Tutorial Slides, 2016. URL <http://gpss.cc/gpuqss16/slides/oakley.pdf>.
- [7] Authors. Bayesian inference in modern machine learning. *International Journal of Multidisciplinary Research in Arts, Science and Engineering*, 2024. URL https://www.ijmra.us/project%20doc/2024/IJESR_AUGUST2024/IJESR3Aug24_11325.pdf.
- [8] Authors. Assessment of uncertainty quantification in universal differential equations. *PMC*, 2024. URL <https://pmc.ncbi.nlm.nih.gov/articles/PMC12005350/>.
- [9] Authors. Ensemble bootstrap methodology for forecasting dynamic growth processes using differential equations: application to epidemic outbreaks. *BMC Medical Research Methodology*, 21:34, 2021. doi: 10.1186/s12874-021-01226-9.
- [10] Authors. Confident neural network regression with bootstrapped deep ensembles. *Neurocomputing*, 2025. doi: 10.1016/j.neucom.2025.021721.

- [11] Gregory Palmer, Yilin Du, Christine Payne, Rishi Persad, James Wang, and Jinfeng Sun. Calibration after bootstrap for accurate uncertainty quantification in regression models. *npj Computational Materials*, 8:115, 2022. doi: 10.1038/s41524-022-00794-8.
- [12] Authors. Gaussian processes regression for uncertainty quantification: An introductory tutorial, 2025.
- [13] Authors. Uncertainty quantification for gaussian processes. 2023. URL <https://mediatum.ub.tum.de/doc/1728130/>.
- [14] Jakob Gawlikowski et al. A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, 56:1513–1589, 2023. doi: 10.1007/s10462-023-10562-9.
- [15] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in Neural Information Processing Systems*, 30, 2017.
- [16] Authors. Physics-guided bayesian neural networks and their application in ode problems. In *ASME Verification, Validation and Uncertainty Quantification Symposium*, 2024. doi: 10.1115/VVUQ2024.
- [17] John R Kitchin. pycse: Python computations in science and engineering. Python package, 2024. URL <https://github.com/jkitchin/pycse>.
- [18] Authors. Conformal prediction: A data perspective. *ACM Computing Surveys*, 2024. doi: 10.1145/3736575.
- [19] Authors. Conformal prediction for uncertainty quantification in dynamic biological systems. *PLOS Computational Biology*, 20(5), 2024. doi: 10.1371/journal.pcbi.1013098.
- [20] Authors. Verifiably robust conformal prediction. In *Advances in Neural Information Processing Systems*, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/.
- [21] Authors. Capturing uncertainty in black-box chromatography modelling using conformal prediction and gaussian processes. *Computers & Chemical Engineering*, 2025. doi: 10.1016/j.compchemeng.2025.001401.
- [22] Authors. Uncertainty quantification for forward and inverse problems of pdes. 2024.
- [23] Michalis Titsias. Variational learning of inducing variables in sparse gaussian processes, 2009.

2 References

```

@article{hullermeier2021aleatoric,
  title={Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods},
  author={H{\\"u}llermeier, Eyke and Waegeman, Willem},
  journal={Machine Learning},
  volume={110},
  number={3},
  pages={457–506},
  year={2021},
  publisher={Springer},
  doi={10.1007/s10994-021-05946-3}
}

```

```

@article{shafiee2024bayesian,
  title={Bayesian neural networks for uncertainty quantification in remaining useful life prediction of systems with sensor moni},
  author={Ochella, Sunday and Dinmohammadi, Fateme and Shafiee, Mahmood},
  journal={Journal of Engineering},
  volume={2024},
  year={2024},
  doi={10.1177/16878132241239802}
}

@article{nature2024multivariable,
  title={Uncertainty quantification in multivariable regression for material property prediction with Bayesian neural networks},
  author={Lee, Shuhao and Purdy, Colton and Singh, Manav and Kalur, Katharine and Almasri, Nasr and Ni, Dingcheng and Hattrick-S},
  journal={Scientific Reports},
  volume={14},
  pages={14484},
  year={2024},
  doi={10.1038/s41598-024-61189-x}
}

@article{survey2023uncertainty,
  title={A survey of uncertainty in deep neural networks},
  author={Gawlikowski, Jakob and others},
  journal={Artificial Intelligence Review},
  volume={56},
  pages={1513--1589},
  year={2023},
  doi={10.1007/s10462-023-10562-9}
}

@misc{arxiv2025gpr,
  title={Gaussian Processes Regression for Uncertainty Quantification: An Introductory Tutorial},
  author={Authors},
  year={2025},
  eprint={2502.03090},
  archivePrefix={arXiv},
  primaryClass={stat.ML}
}

@misc{oakley2016introduction,
  title={Introduction to Uncertainty Quantification and Gaussian Processes},
  author={Oakley, Jeremy},
  year={2016},
  howpublished={GPSS Tutorial Slides},
  url={http://gpss.cc/gpuqss16/slides/oakley.pdf}
}

@article{tum2023gpr,
  title={Uncertainty Quantification for Gaussian Processes},
  author={Authors},
  year={2023},
  institution={Technical University of Munich},
  url={https://mediatum.ub.tum.de/doc/1728130/}
}

@article{uq2024assessment,
  title={Assessment of uncertainty quantification in universal differential equations},
  author={Authors},
  journal={PMC},
  year={2024},
  url={https://pmc.ncbi.nlm.nih.gov/articles/PMC12005350/}
}

@article{acm2024conformal,
  title={Conformal Prediction: A Data Perspective},
  author={Authors},
  journal={ACM Computing Surveys},
  year={2024},
  doi={10.1145/3736575}
}

```

```

@article{plos2024conformal,
  title={Conformal prediction for uncertainty quantification in dynamic biological systems},
  author={Authors},
  journal={PLOS Computational Biology},
  volume={20},
  number={5},
  year={2024},
  doi={10.1371/journal.pcbi.1013098}
}

@inproceedings{neurips2024verifiable,
  title={Verifiably Robust Conformal Prediction},
  author={Authors},
  booktitle={Advances in Neural Information Processing Systems},
  year={2024},
  url={https://proceedings.neurips.cc/paper_files/paper/2024/}
}

@article{sciencedirect2025chromatography,
  title={Capturing uncertainty in black-box chromatography modelling using conformal prediction and Gaussian processes},
  author={Authors},
  journal={Computers \& Chemical Engineering},
  year={2025},
  doi={10.1016/j.compchemeng.2025.001401}
}

@article{bootstrap2021ensemble,
  title={Ensemble bootstrap methodology for forecasting dynamic growth processes using differential equations: application to ep},
  author={Authors},
  journal={BMC Medical Research Methodology},
  volume={21},
  pages={34},
  year={2021},
  doi={10.1186/s12874-021-01226-9}
}

@article{npj2022calibration,
  title={Calibration after bootstrap for accurate uncertainty quantification in regression models},
  author={Palmer, Gregory and Du, Yilin and Payne, Christine and Persad, Rishi and Wang, James and Sun, Jinfeng},
  journal={npj Computational Materials},
  volume={8},
  pages={115},
  year={2022},
  doi={10.1038/s41524-022-00794-8}
}

@article{bootstrap2025confident,
  title={Confident neural network regression with Bootstrapped Deep Ensembles},
  author={Authors},
  journal={Neurocomputing},
  year={2025},
  doi={10.1016/j.neucom.2025.021721}
}

@article{arxiv2024evaluating,
  title={Uncertainty Quantification for Forward and Inverse Problems of PDEs},
  author={Authors},
  year={2024},
  booktitle={AAAI Conference on Artificial Intelligence}
}

@inproceedings{asme2024physics,
  title={Physics-Guided Bayesian Neural Networks and Their Application in ODE Problems},
  author={Authors},
  booktitle={ASME Verification, Validation and Uncertainty Quantification Symposium},
  year={2024},
  doi={10.1115/VVUQ2024}
}

```

```

@inproceedings{kendall2017uncertainties,
  title={What uncertainties do we need in Bayesian deep learning for computer vision?},
  author={Kendall, Alex and Gal, Yarin},
  booktitle={Advances in Neural Information Processing Systems},
  volume={30},
  year={2017}
}

@article{lakshminarayanan2017simple,
  title={Simple and scalable predictive uncertainty estimation using deep ensembles},
  author={Lakshminarayanan, Balaji and Pritzel, Alexander and Blundell, Charles},
  journal={Advances in Neural Information Processing Systems},
  volume={30},
  year={2017}
}

@article{dinmohammadi2021deep,
  title={Deep Bayesian Gaussian processes for uncertainty estimation in electronic health records},
  author={Kim, Melanie F and Tomczak, Jakub M and Aanen, Laurens and Cohen, Ans MW},
  journal={Scientific Reports},
  volume={11},
  pages={20685},
  year={2021},
  doi={10.1038/s41598-021-00144-6}
}

@article{bayesian2024inference,
  title={Bayesian Inference in Modern Machine Learning},
  author={Authors},
  journal={International Journal of Multidisciplinary Research in Arts, Science and Engineering},
  year={2024},
  url={https://www.ijmra.us/project%20doc/2024/IJESR_AUGUST2024/IJESR3Aug24_11325.pdf}
}

@misc{titsias2009variational,
  title={Variational learning of inducing variables in sparse Gaussian processes},
  author={Titsias, Michalis},
  booktitle={International Conference on Artificial Intelligence and Statistics},
  pages={567--574},
  year={2009}
}

@misc{pycse2024nbr,
  title={pycse: Python Computations in Science and Engineering},
  author={Kitchin, John R},
  year={2024},
  howpublished={Python package},
  url={https://github.com/jkitchin/pycse}
}

```
