Proceedings of CONF-CIAP 2027 Symposium: Applications of Complex Analysis, Optimization and Integrability to Physical Theories
DOI: 10.54254/2753-8818/2027.TA38934

# *An Overview and Applications of Uncertainty Quantification*

**Wenjia Zhang**

*Qingdao No.2 High School of Shandong Province, Qingdao, China*
*15192497327@163.com*

***Abstract.*** The analysis and decision-making of scientific data and models have become increasingly important with the progress of data science, artificial intelligence, and computing. Although such analyses draw on available data, limited samples, measurement errors, missing data, and model assumptions all affect the results. In addition, many complex systems are inherently random. Therefore, in addition to prediction of a subject under study by a model, the uncertainty in the prediction must be studied. Uncertainty quantification (UQ) uses probability theory, statistics, and computational methods to study the sources of uncertainty and their magnitudes, and to study the effects of uncertainty in model results. This paper discusses sources of uncertainty in data, models, and randomness; uses confidence intervals, probability distributions, Bayesian methods, and Monte Carlo methods for UQ; and presents applications of these methods in healthcare, climate science, and artificial intelligence. UQ does not aim to make unknowns vanish; it makes model results more credible by describing the uncertainties in the results.

***Keywords:*** uncertainty quantification, statistics, probability distribution, Bayesian methods, Monte Carlo simulation

## 1. Introduction

Modern scientific research is increasingly based on data. Researchers in the natural sciences, medicine, climate research, and AI-related fields first analyze existing data and then use statistical and numerical models to draw conclusions about the phenomena they study. However, model outputs—whether numerical or graphical—are rarely completely correct. They are approximations of real systems under specific, often limiting assumptions.

There are several basic problems in statistics. First, one wants to use information from samples to make inferences about populations. Samples exhibit natural random variation, data contain errors, and estimates of model parameters vary. For complex problems, different models may yield completely different results. Natural and social systems themselves have characteristics that make their changes difficult to fully predict, which contributes to uncertainty in data analysis. Uncertainty quantification (UQ) not only focuses on the results themselves, but also pays attention to the degree and source of variation behind these results, as well as whether this variation affects the final judgment. Its goal is to use mathematical and statistical methods to describe uncertain factors rather than to reach absolute conclusions [1]. This article explores the main sources of uncertainty,

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2027.TA38934

common quantification methods, and practical application examples to clarify its role in data analysis and scientific decision-making.

## 2. Main sources of uncertainty

### 2.1. Data uncertainty

Data uncertainty primarily arises during the data collection and recording processes, and its sources include limited sample size, measurement errors, missing observations, and sampling bias [1]. Since most studies cannot observe the entire population, analyses typically rely on subsamples of that population. Due to the presence of random factors, differences exist between samples, meaning that the calculated statistics are often biased from the true population parameters.

The measurement process itself may also introduce uncertainty. Measuring instruments have inherent accuracy limitations, manual operations may lead to recording errors, and survey responses may be influenced by subjective judgment. Missing data reduce the completeness of available information; in particular, when data are non-randomly missing, simply excluding incomplete observations may introduce systematic bias [2].

Although increasing the amount of data can reduce some of the uncertainty caused by sampling, the size of a dataset does not determine its quality. Even with a large sample size, if bias is present, the data may still fail to accurately reflect the characteristics of the population as a whole. Therefore, assessing data uncertainty requires careful consideration of sample size, data sources, measurement methods, and the sampling process.

### 2.2. Model uncertainty

Model uncertainty arises from the simplification of real-world systems. Mathematical models use variables, parameters, and assumptions to describe a problem, but no single model can encompass every factor affecting the subject of study. Therefore, making different decisions regarding variable selection, model structure, and parameter values may lead to differing results.

This uncertainty exists at both the parameter level and the structural level. Parameters are typically estimated based on limited data, introducing uncertainty into their values. Model structures involve representing real-world mathematical relationships; even when using the same dataset, models constructed based on different assumptions may produce different predictive outcomes [2].

A more complex model is not necessarily more reliable. Although complex models may fit existing data better, they are also prone to overfitting, resulting in poor performance on new observational data. Therefore, evaluating model uncertainty requires a comprehensive assessment of assumptions, parameter stability, model applicability, and predictive performance.

### 2.3. Random uncertainty

Random uncertainty concerns variation inherent in the phenomenon being studied. Unlike missing data or model simplification, it does not arise entirely from a lack of information on the researcher's part; the system itself may have random features [1].

Random variables and probability distributions are central statistical tools for describing such phenomena. Even if the probability law governing a random process is known, the outcome of any particular event usually cannot be predicted exactly. More data can help researchers estimate that law more accurately, but they cannot eliminate random variation altogether.

DOI: 10.54254/2753-8818/2027.TA38934

It is therefore useful to distinguish uncertainty caused by limited information from randomness inherent in a system. The former may decrease as data and knowledge improve, whereas the latter usually requires a probability model that describes the range of possible outcomes. This distinction helps guide the choice of UQ method.

## 3. Common methods of uncertainty quantification

### 3.1. Confidence intervals

A confidence interval is a basic way to express uncertainty in parameter estimation. A point estimate provides a single estimated value for a population parameter, whereas a confidence interval gives a range at a specified confidence level and thus indicates the precision of the estimate [1].

Its width depends on sample size, variability in the data, and the chosen confidence level. Other things being equal, a larger sample generally reduces sampling error and narrows the interval, whereas greater variability in the data generally increases uncertainty. Raising the confidence level usually widens the interval. Reporting an interval therefore conveys more about statistical error than reporting a point estimate alone.

### 3.2. Probability distributions

A probability distribution describes the possible values of a random variable and their probabilities. It is a foundation of UQ because it represents possible outcomes and their likelihoods together, turning abstract uncertainty into something that can be calculated and analyzed [3].

Different random phenomena call for different distributions. Discrete and continuous random variables, for example, generally require different types of models. Distributions can also be used to calculate means, variances, and probability intervals, thereby describing both the center and the spread of the data.

A distribution is usually chosen in light of theoretical assumptions and observed data. If the assumed distribution differs substantially from the actual data, subsequent uncertainty estimates may be affected. Selecting an appropriate probability model is therefore a key step in UQ.

### 3.3. Bayesian methods

Bayesian methods offer a statistical framework for updating judgments about uncertainty as new data become available. Prior information is represented by a prior distribution. As new observations become available, Bayes' theorem is applied to update the prior distribution and obtain a posterior distribution [4].

The posterior distribution describes possible parameter values and their probability characteristics. Bayesian methods can thus provide an estimate while also expressing the relative plausibility of different results given the available evidence [5]. They are especially useful when data arrive over time or an analysis needs to incorporate prior knowledge.

These methods also have limitations. The choice of prior may affect the result, and a prior without a sound basis may introduce further uncertainty. Researchers should therefore explain the origin of their prior assumptions and assess how alternative choices affect their conclusions [4].

DOI: 10.54254/2753-8818/2027.TA38934

### 3.4. Monte Carlo simulation

Monte Carlo simulation uses repeated random sampling to investigate uncertainty in complex systems. When a model contains several random variables and an analytical solution is difficult to obtain, researchers can assign probability distributions to its inputs and run the model repeatedly [5].

Each run generates one set of possible inputs and a corresponding output. After many runs, the outputs form an empirical distribution from which the mean, variance, quantiles, or probability of a particular event can be calculated. This shows how uncertainty in the inputs propagates to the final result.

Monte Carlo simulation can handle several uncertain factors at once and is useful for complex models. Its results still depend on the input distributions and the model itself. If the underlying assumptions are substantially biased, more runs will improve computational stability but will not correct a misspecified model.

## 4. Applications

### 4.1. Healthcare

Healthcare involves considerable uncertainty in both data and models. Clinical data usually come from limited patient samples, while patients differ in age, genetic traits, lifestyle, and the course of their illnesses [3]. A treatment effect or disease risk estimated from a sample therefore cannot be assumed to apply with certainty to every individual.

In clinical trials, confidence intervals can indicate how precisely a treatment effect has been estimated. In disease-risk prediction, probability models can express the likelihood of different outcomes. As machine learning is used in medical imaging and decision support, UQ can also help assess whether a model's prediction for a particular case is sufficiently reliable [6].

Medical decisions often have direct consequences for patients, making highly uncertain results important to identify. When a model lacks sufficient information, communicating its uncertainty can help justify further testing or clinician review.

### 4.2. Climate science

The climate system comprises interacting components such as the atmosphere, oceans, land, and human activities. Climate models must mathematically represent these complex processes. Differences in model structure, parameter settings, initial conditions, and future socioeconomic scenarios can all influence long-term projection results [7].

Therefore, climate projections should report both expected trends and their associated uncertainties. Comparing results across different models and scenarios helps identify changes characterized by high consistency versus those subject to significant uncertainty [8, 9]. Probabilistic analysis and simulation also facilitate the assessment of potential risks under various climate scenarios, including extreme events.

Long-term policy decisions often address not a single, certain future but a range of possible scenarios. Uncertainty quantification (UQ) methods allow for a more systematic elucidation of these differences, enabling risk management strategies that leverage the information provided by projections while also accounting for their limitations [10].

DOI: 10.54254/2753-8818/2027.TA38934

### 4.3. Artificial intelligence

Advances in artificial intelligence, particularly machine learning, have created new applications for UQ. Machine-learning models learn patterns from training data, but those data cannot cover every real-world situation. A model may become less reliable when it encounters inputs substantially different from those on which it was trained [6].

Traditional evaluations of machine learning often focus on overall measures such as accuracy and error. Such measures do not fully convey how trustworthy an individual prediction is. In high-risk areas, including medical AI and autonomous driving, providing a prediction alone is insufficient; the model's uncertainty about that prediction must also be assessed [4].

Uncertainty estimates can help systems recognize the limits of their capabilities. An unfamiliar input or a highly uncertain prediction may call for human review or additional information to reduce risk. As AI takes on more complex tasks, enabling models to express uncertainty appropriately will remain essential to their trustworthiness [6].

## 5. Conclusion

In data- and model-driven scientific research, uncertainty is ubiquitous. This text explores the three main sources of uncertainty—data, models, and randomness—and four common methods: confidence intervals, probability distributions, Bayesian methods, and Monte Carlo simulation. Although these methods address different problems and employ fundamentally different computational approaches, they share a common goal: to transform information that lacks absolute certainty into a form that can be described, analyzed, and compared.

Applications in fields such as healthcare, climate science, and artificial intelligence demonstrate that uncertainty quantification (UQ) is vital for both practical decision-making and theoretical research. In complex systems, a single forecast is often insufficient to fully reflect reality; the range of possible outcomes, the confidence of the forecast, and the conditions under which the model may fail are also of equal importance [1].

From a statistical perspective, rigorous analysis does not require absolute certainty. Clearly explaining the degree of uncertainty involved in a conclusion is often more rigorous than reporting a figure that appears precise but lacks a solid foundation. Quantifying uncertainty is not meant to mask the unknown, but to make analysis possible by assigning a concrete scope and structure to unknown factors.

As datasets grow larger and models become increasingly complex, data analysis must strike a balance between pursuing predictive accuracy and ensuring reliability and interpretability. Identifying the sources of uncertainty, quantifying its magnitude, and appropriately utilizing this information in decision-making will remain key topics for future research.

## References

[1] Zhang, J., et al. (2020). Basic framework and main methods of uncertainty quantification. *Mathematical Problems in Engineering, 2020*, 6068203. https://doi.org/10.1155/2020/6068203

[2] Abdar, M., Pourpanah, F., Hussain, S., et al. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion, 76*, 243–297. https://doi.org/10.1016/j.inffus.2021.05.008

[3] Huang, L., Ruan, S., Xing, Y., & Feng, M. (2024). A review of uncertainty quantification in medical image analysis: Probabilistic and non-probabilistic methods. *Medical Image Analysis, 97*, 103223. https://doi.org/10.1016/j.media.2024.103223

DOI: 10.54254/2753-8818/2027.TA38934

[4] Seoni, S., Jahmunah, V., Salvi, M., et al. (2023). Application of uncertainty quantification to artificial intelligence in healthcare: A review of last decade (2013-2023). *Computers in Biology and Medicine, 165*, 107441. https://doi.org/10.1016/j.compbiomed.2023.107441

[5] Vasquez, V. R. (2005). Accounting for both random errors and systematic errors in uncertainty propagation analysis of computer models involving experimental measurements with Monte Carlo methods. *Risk Analysis, 25*(6), 1669–1681. https://doi.org/10.1111/j.1539-6924.2005.00704.x

[6] Lambert, B., Forbes, F., Doyle, S., Dehaene, H., & Dojat, M. (2024). Trustworthy clinical AI solutions: A unified review of uncertainty quantification in deep learning models for medical image analysis. *Artificial Intelligence in Medicine, 150*, 102830. https://doi.org/10.1016/j.artmed.2024.102830

[7] Majda, A. J., & Gershgorin, B. (2010). Quantifying uncertainty in climate change science through empirical information theory. *Proceedings of the National Academy of Sciences, 107*(34), 14958–14963. https://doi.org/10.1073/pnas.1007009107

[8] Wu, Y., Miao, C., Fan, X., et al. (2022). Quantifying the uncertainty sources of future climate projections and narrowing uncertainties with bias correction techniques. *Earth's Future, 10*, e2022EF002963. https://doi.org/10.1029/2022EF002963

[9] Pathak, R., Dasari, H. P., El Mohtar, S., et al. (2021). Uncertainty quantification and Bayesian inference of cloud parameterization in the NCAR Single Column Community Atmosphere Model (SCAM6). *Frontiers in Climate, 3*, 670740. https://doi.org/10.3389/fclim.2021.670740

[10] Oppenheimer, M., Little, C. M., & Cooke, R. M. (2016). Expert judgement and uncertainty quantification for climate change. *Nature Climate Change, 6*, 445–451. https://doi.org/10.1038/nclimate2959