Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Building trust in automated experimentation: uncertainty quantification in the era of high-throughput biolabs

Wolfgang Wiechert[1,2,*], Laura M Helleckes[3,4] and Katharina Nöh[1]

Uncertainty quantification (UQ) is central to data analytics, particularly in the life sciences, where experiments are often affected by significant measurement noise. In emerging automated high-throughput biolabs, such as biofoundries, parallel cultivation systems, and smart analytics platforms, UQ should be a built-in feature rather than an optional add-on. These environments pose a unique challenge: robotic liquid handling must be combined with miniaturized biochemical analytics (including omics), process monitoring, online data analytics, and digital control. Although traditional UQ methods from classical and computational statistics remain valid and applicable, integrating them into highly parallelized experimental and digital workflows presents new challenges. These include data preprocessing, model-based data integration, decision-making, and experimental control. In this review, we examine the emerging demands on UQ in automated experimentation and survey recent frameworks, strategies, and computational tools designed to address them.

Addresses
[1] Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Jülich, Germany
[2] RWTH Aachen University, Aachen, Germany
[3] The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London, UK
[4] I-X Centre for AI in Science, Imperial College London, London, UK

Corresponding author: Wiechert, Wolfgang (w.wiechert@fz-juelich.de)

Current Opinion in Biotechnology 2025, **96**:103382

This review comes from a themed issue on **Systems Biology**

Edited by **Christoph Wittmann** and **Pablo Iván Nikel**

Available online xxxx

https://doi.org/10.1016/j.copbio.2025.103382

Introduction

Generating extensive quantitative data sets is essential to all systemic biological disciplines, including systems biology, systems medicine, metabolic engineering, and process development. These data sets can contain raw measurements, such as mass spectrometry data; processed data, such as peak areas or image analysis results; model-based parameter estimates, such as growth parameters; or data-based model predictions. In biotechnology, key performance indicators (KPIs), such as growth rates and product or space-time yields, are important for characterizing production organisms and processes. Signal-to-noise ratios are much worse in biology than in other disciplines, such as physics or engineering (Figure 1). This underlines the importance of UQ for the life sciences, which is the only way to ensure trust in data.

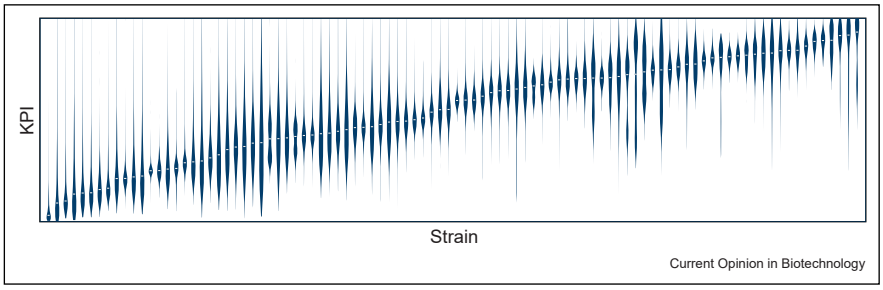
Automation speeds up data generation tremendously and eliminates human lab-worker variability [1–3]. However, miniaturization also introduces even higher noise levels caused by small-scale liquid handling and other operations [4]. In addition, miniaturized instruments often replace standard measurement devices with faster, but noisier alternatives that can be better integrated into robotic platforms [5]. For instance, optical density (OD) measurements in lab-scale cultivations are replaced by backscatter data in parallelized micro cultures [4], or an High-performance liquid chromatography device is replaced by enzymatic assays [4].

Figure 1 illustrates UQ with a strain selection experiment performed in the Jülich Biofoundry. In this proof-of-concept study, a library of 96 microbial production strains was screened in six batches performed in parallelized micro-cultivations with 48 strains each. The maximum product formation rate of a strain achieved during the cultivation period (max dP/dt) was chosen as KPI. Computing this KPI required a Bayesian process model based on noisy and scarce product data. Replicates were performed to reduce uncertainty. This study clearly demonstrates some key challenges of UQ:

1. Thorough UQ is necessary to judge data quality. Otherwise, the quality of the data is strongly overestimated; that is, reducing the information to the medians shown in Figure 1 is completely misleading. The partially even bimodal KPI distributions give much more information.
2. UQ is crucial for automatic experimental design. By repeating experiments, the uncertainty of KPIs for high-performing strains can be reduced through further exploitation, whereas low performers can be

* Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungs-zent-rum Jülich GmbH, Wilhelm-Johnen-Straße, 52428 Jülich, Germany.

Figure 1



Importance of UQ for biological data processing illustrated with microbial strain selection: The violin plot with medians visualizes the uncertainty within a difficult-to-determine KPI for a library of 96 microbial production strains. A total of 288 single experiments were performed, including replicates. Details are given in the main text.

excluded from being revisited, thus enabling more efficient exploration of the strain library. This explains why the low-performer KPIs have higher uncertainties than those of the high-performers.

3. Making an informed decision for strain selection needs to account for uncertainty. After 6×48 experiments, there is no doubtlessly best-performing strain; rather, a group of high-performers can be selected for further investigation.

Generally, automation entails significant changes in how data is evaluated because manually inspecting large volumes of heterogeneous data becomes impractical. Thus, to reach their full potential, automated systems require a versatile digital infrastructure [2,3,6,7] to manage concurrent experimental workflows, collect data, autonomously make decisions, or plan new experiments. Since UQ must be an integral part of such digital infrastructures, new data-scientific challenges arise. Automated experimentation is a dynamically developing field, and no unique approach or even a standard for workflow-integrated UQ has thus emerged yet. The major obstacles and some partial solutions will be discussed below.

Uncertainty quantification in a nutshell

Although alternative approaches, such as interval calculus or fuzzy set theory, have been proposed, mainstream UQ is today based on a probabilistic framework. However, because UQ is a broad field [8–11], it comes in several flavors. In the context of computational workflows, uncertainty propagation (UP) is the most important aspect discussed below:

Probabilistic approach: Probability distributions provide the most informative way to describe uncertainty because they support probabilistic reasoning and the derivation of confidence intervals. While the theoretical foundations for propagating distributions through (nonlinear) functions are well established, exact solutions are in most cases practically infeasible, especially in high-dimensional and multivariate settings.

Moment-based approach: Variance, and respectively standard deviation, reduce a probability distribution to characteristic measures. In the multivariate case, this concept generalizes to the covariance matrix. UP for these quantities by linear approximation (aka Gaussian error propagation) [12–14] is well established. However, this linearized approach to UP can be risky in the presence of strong nonlinearities, poor signal-to-noise ratios, and violated assumptions on distribution normality and independence [15].

Monte Carlo approach: Replacing a probability distribution with a large sample and propagating this data through computational workflows is conceptually straightforward [16,17] and broadly applicable to any kind of distribution and nonlinearity. However, the required large sample sizes entail a computationally expensive statistical evaluation, which can become prohibitive in high-dimensional and time-critical experimental settings [18].

The Bayesian approach: Bayesian statistics is grounded in classical probability theory, but differs from frequentist approaches in its interpretation. Probabilities represent degrees of belief or states of knowledge, rather than long-run frequencies [19]. Prior knowledge is explicitly encoded in the form of a prior distribution, which is updated via Bayes' theorem when new data arrives. Like other approaches, Bayesian methods rely on statistical measures, or special Monte Carlo (MC) algorithms, but the results can reflect data, model, and parameter uncertainty [4,20–22].

Approximation approaches: In this review, the term refers to a diverse family of methods that replace distribution functions or models with surrogates that are easier to handle. Examples are polynomial chaos [23,24],

spectral expansion [25], surrogate modeling [26], or conformal prediction [18]. Such approximations can be combined with any of the aforementioned approaches.

UQ and UP methodologies are well developed in theory [8,27,28], and supported by numerous open-source implementations [12,24,29–33]. Some methodological comparisons are found in Refs. [17,22,27]. Also, the necessary computing power is available today, even for compute-intensive MC methods. What, then, are the data science-related challenges in the context of automated experimental workflows?

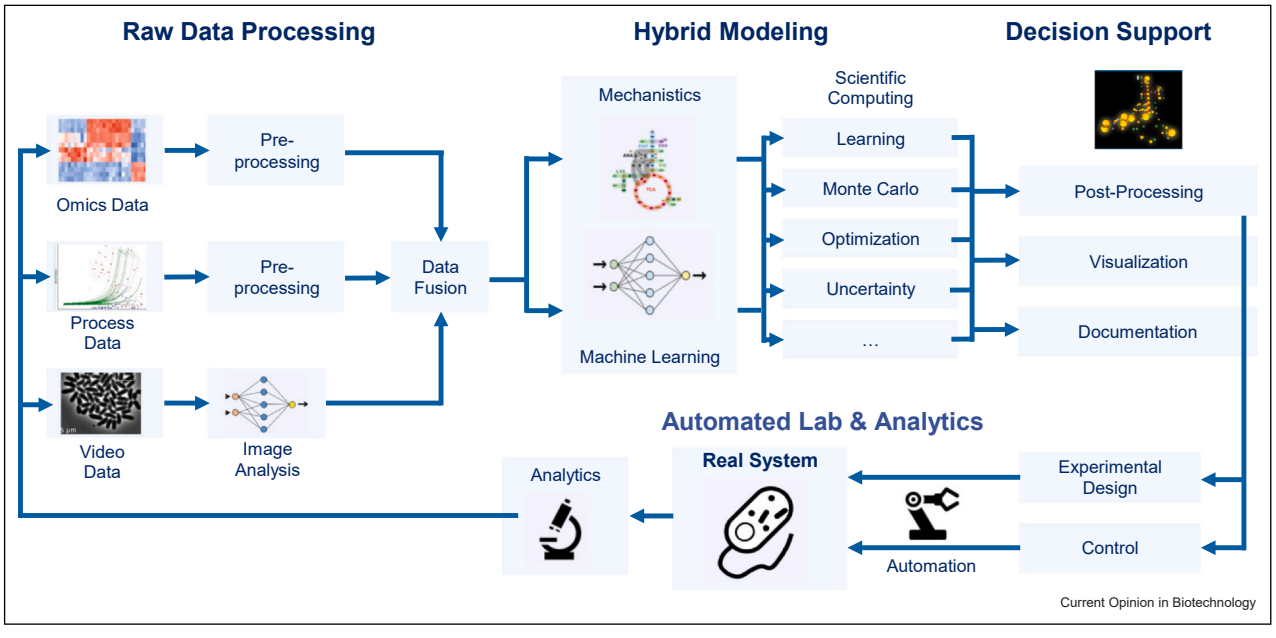
Preprocessing measurement data with uncertainty quantification

UQ starts with the preprocessing of raw measurement data (Figure 2). For historical reasons, metrology and the analytical sciences have developed distinct vocabularies and even industry standards for UQ [34,35]. Importantly, analytical terms such as *accuracy*, *precision*, *trueness*, *limit of detection*, and *calibration* can be mapped onto corresponding statistical terminology, such as *location*, *spread*, *bias*, etc. [28]. In the context of robotic workflows, this terminological confusion becomes critical because a sharp separation between analytical steps and other experimental operations within robotic systems no longer exists.

Raw data produced by analytical instruments, such as High-performance liquid chromatographys, mass spectrometers, laser backscatter instruments, or microscopes, do not directly represent the quantities of interest. Moreover, miniaturized automated experimentation often requires operating outside the linear calibration regime. For instance, dilution series for measuring concentrations are often avoided. Measurement noise can have a non-normal distribution and may strongly depend on the measured quantity (heteroscedasticity). In this situation, which exceeds analytical textbook knowledge, advanced automatic calibration methods have recently become available [4]. It is important to note that a calibration function maps the wanted quantity to the corresponding measurement (e.g. biomass concentration to OD), but is actually used in the inverse direction (e.g. from OD to biomass). This poses additional challenges for correct UP [4,28,36].

In the life sciences, experimental protocols heavily rely on liquid handling operations [3,6,37]. All pipetting steps influence the experimental error, and these errors accumulate. Usually, the uncertainty of elementary operations in an experimental protocol (e.g. pipetting, weighing, detection, and counting) is well characterized. Thus, bottom-up physical modeling of the experimental workflow (mostly mixing of liquids) provides an excellent tool to estimate the cumulative error arising from

Figure 2



General scheme of the physical and digital workflows in an automated Biolab. Measurement data from the real system is pre-processed by different computational procedures to calculate the use data. The use data are then fused and processed based on mechanistic or ML models, which may be hybrid. Highly condensed data can be used to maintain human involvement for potential interventions or decision-making. Ideally, low-level control of the process, high-level decisions, and even smart design of new experiments will be performed in closed-loop operation.

a sequence of elementary operations, each contributing some noise [12–14,38]. Such error models — often based on cause-and-effect diagrams [14] — have the added advantage of identifying and improving major factors influencing the overall noise level [13,14,17]. For complex operations like chromatography [39], cell counting [40,41], or Polymerase Chain Reaction procedures [42], error modeling based on deep process understanding is possible, but demanding.

In high-tech analytical instruments, such as mass spectrometers, the calculation of desired data (e.g. peak areas) from raw data (e.g. spectra) is often performed using built-in vendor software. The code is usually proprietary, and intrinsic UQ is often unavailable or poorly documented. Consequently, poor quantifications (e.g. for overlapping or irregular peaks) must be manually corrected [43,44]. This gap can be bridged by making raw spectra data evaluation available as open source software [44–46]. Nevertheless, in automated workflows, it is critically important to have a quality indicator that expresses the regularity of a signal [47,48]. It is recommended to define quality standards and skip data sets exposing irregular patterns in fully automated procedures.

Single-cell cytometry [41] and live-cell imaging [49] are increasingly important in automated analytics. Image and video processing pipelines involve preprocessing, cell segmentation, cell tracking, and feature quantification, each of which presents unique challenges — especially in time-resolved (video) data, where real-time or near-real-time processing is required [50]. While deep learning methods, particularly convolutional neural networks, have advanced many of these tasks, robust uncertainty quantification remains an underdeveloped area, despite its critical role in assessing model robustness and guiding biological interpretation [50,51].

Identifying, eliminating, or at least correcting often very subtle systematic errors [52] is a common theme for generating reliable data on automated platforms. Data correction always requires additional information given by empirically determined correction factors or by using quantities derived from other measured data. Consequently, data correction involves additional uncertainties [13]. Here, the well-known bias-variance tradeoff must be considered: the bias of corrected data may decrease, but the variance increases at the same time. Thus, data correction is not always beneficial. As a special case, batch effects [4,34,53] are common in parallelized experiments (Figure 3), which can often be attenuated by normalization operations [54,55].

Uncertainty quantification for advanced data processing algorithms

Figure 2 distinguishes between data preprocessing (cf. Section *Preprocessing measurement data with uncertainty quantification*) and subsequent model-based data integration. The latter is concerned with merging heterogeneous data sources, model-based evaluation of data, making predictions, as well as computing quantities and visualizations for decision support.

Mechanistic mathematical models, such as cell models on different levels of detail [11], are one approach for integrating heterogeneous data, such as concentration and process data, and for making predictions based on this data. Nonlinear regression (including its Bayesian counterparts) is the most common method for deriving parameters and decision-relevant data, such as KPIs, from pre-processed data sets. Although UP for nonlinear regression models is well established [11,20,21,27], it notoriously suffers from multiple optima and approximation errors due to model linearization. However, low-dimensional models with well-understood nonlinearity and sensitivity to noise influences (e.g. growth models) can often be adjusted to run robustly in an automated framework [4]. Here, Bayesian regression approaches [20–22] have specific advantages at a higher computational cost. Novel methods like conformal prediction [18] aim to reduce this cost while providing valid uncertainty estimates without the need for Bayesian approaches.

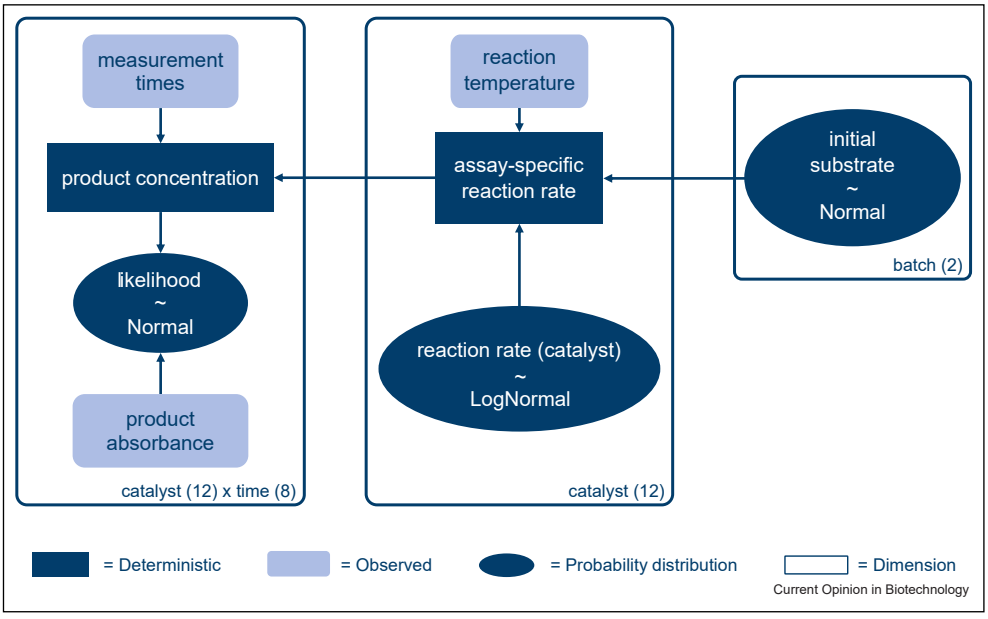
In recent times, data-based machine learning (ML) has matured into a powerful alternative to mechanistic model-based classical regression [56]. However, deep learning models require extensive training with high-quality annotated data, which is often unavailable even in the context of high-throughput experimentation. For this reason, hybrids between mechanistic models and ML approaches are becoming popular [11,23,57,58], which require less data for model training. Different hybrid model architectures have been proposed, such as bio/physics-informed neural networks [59,60] or physical balance equations with integrated ML modules for learning kinetic laws [57]. Importantly, UQ for ML is still in its infancy [51,61].

Model uncertainty remains a critical but underexplored aspect in quantitative biology, particularly in complex experimental systems. Single-model approaches often yield overconfident estimates — both in predictions and parameter inferences. In contrast, multimodel strategies, such as Bayesian model selection [62] and averaging [63], offer principled statistical ways to reflect this uncertainty and improve robustness [63]. Looking ahead, such cutting-edge methods can play an increasingly important role in automated experimentation because they enhance the robustness and trustworthiness of the results.

Integrating uncertainty quantification into multistep concurrent workflows

Automatic experimentation couples an experimental with a concurrent digital workflow [3,10,37,64]. Generally, a UQ framework for automated experimentation

Figure 3



Exemplary probabilistic model for enzyme kinetic measurements. Product absorbance and reaction temperature of a reaction are measured in the laboratory at eight time points (light blue). The aim is to model product concentration and compare it with the measured absorbance using a likelihood function to quantify uncertainty. The reaction of each of the 12 biocatalysts (distribution at bottom center) is influenced by reaction temperature as well as a batch effect. Since experiments are conducted in two batches, pipetting errors cause slight variation in the initial substrate concentration. The model accounts for this batch effect as a distribution (right), multiplying this fluctuation with the individual reaction rate to model the specific reaction rate in the assay.

must enable UP from measurement preprocessing, model-based data analysis, and decision-making all the way to control actions [19,31,65] (Figure 2). However, end-to-end UQ in multistep workflows still suffers from a lack of standardization and interoperability between UQ concepts and software representations of uncertainty (cf. Section *Uncertainty quantification in a nutshell*). Simplifying assumptions, such as treating intermediate results as independent normal distributions, to bridge between different workflow components, can produce grossly incorrect results, as demonstrated in Ref. [15] for regression model predictions. One of the most overlooked UP problems is neglecting the often strong correlations between intermediate computational results, such as model parameter estimates [66]. This leads to inaccurate uncertainty representations, which then propagate through all subsequent steps of the workflow.

In a rigorous approach, UP should be based on the same UQ concept (distributions, moments, MC, etc.), applied throughout the whole workflow. Such a unifying framework requires more effort in terms of method and software development, but this is worthwhile because typical statistical flaws are avoided. Three cases demonstrate that practical solutions are emerging to support a rigorous UQ approach:

- MC simulation provides a universally applicable method for UP, but its computational cost can be prohibitive for high-dimensional or computing-intensive models.
- Automatic code differentiation allows efficient and exact Gaussian UP through extensive computational workflows by leveraging algorithmic differentiation to compute moments of output distributions [67].
- Probabilistic programming frameworks (Figure 3) enable full Bayesian workflows by combining model specification and inference in a unified language, supporting UQ across hierarchical and data-driven models [29].

Some of the discussed methods (image analysis, regression, ML, and MC) require substantial computing resources when UQ is embedded, especially when evaluating many experiments in parallel. Therefore, high-performance compute clusters should be an integral part of automation facilities. To facilitate parallelization and orchestration of many interdependent tasks, it is popular to represent workflows as directed acyclic graphs (DAG) [3,6], which is well supported by widely used workflow management systems such as Airflow, Nextflow, or Snakemake. However, because DAGs inherently disallow feedback loops or iterations, circular

logic and adaptive workflows must be handled through higher-level orchestration layers.

In this context, the close connection to the concept of digital twins (DT) [10,68] must be pointed out. DTs provide modular architectures capable of representing real processes, evaluating data, integrating heterogeneous data sources into unified models, making decisions, guiding experiment design, and feeding results back into the physical system. The experimental and digital closed-loop workflow, shown in Figure 2, can also be seen as a DT architecture. UQ for DTs is an emerging field [68], and automatic experimentation in the life sciences can learn from these developments.

Autonomy versus human intervention

Once a robotic high-throughput system is running, increasing the degree of autonomy is a big challenge. Simple, low-level actions can easily be triggered by measurable events. Higher-level decision-making, which involves goal-oriented steering of screening processes under time and resource constraints, is on the horizon. In this context, smart design of experiment strategies and active learning concepts [11] are emerging to make biofoundry workflows more efficient and information-rich [2,53]. Rather than relying on exhaustive or trial-and-error experimentation, these approaches aim to select the most informative experiments with minimal experimental resources [69]. In particular, active learning adapts experiment selection on-the-fly, based on how much new information each candidate experiment is expected to yield about a new strain variant or process parameters.

However, because irregular and unforeseen behavior is notoriously encountered with biological systems, it makes sense to keep the human in the loop. The big question is, on which level of information condensation this should be done. Doubtlessly, if low-level decisions need to be made during the runtime of high-throughput experiments, having humans in the loop can be a showstopper. To support high-level decision-making, powerful information displays, visualization tools, and intervention mechanisms are needed, which facilitate surveying a complex situation in a short time. UQ must become an essential part of all these frameworks.

Recommendations for biofoundry newcomers and the community

As emphasized by Figure 1, quantitative biological data should always be accompanied by UQ to build trust in the data, as well as to understand what conclusions can or cannot be drawn. It is crucial that researchers understand how to assess data, inference, and prediction quality. As shown, the complexity of this task increases due to the intricate workflows in automated systems and the inability to include low-level manual data evaluation steps. Some recommendations for newcomers:

1. Analytical calibration requires careful analysis of uncertainties in measurement systems and should go beyond linearity or normal distribution assumptions. If in doubt, compare different calibration functions.
2. Calibration must be combined with UQ within an integrative framework because both parts cannot be treated separately in automated systems.
3. Question or benchmark automatic analysis in vendor software.
4. Use existing UQ frameworks like probabilistic programming languages and toolboxes [30–33] and apply them rigorously throughout all steps of the data analysis pipeline.

This will involve a steep learning curve and ease the specification and execution of probabilistic models (Figure 3). Additionally, two biofoundry community activities are suggested:

1. A survey of UQ/UP approaches and frameworks would be highly beneficial. Based on the results, it might be possible to derive best practice examples and even minimal community standards. Benchmarking studies would further help to compare and evaluate the performance, robustness, and applicability of different UQ approaches.
2. The strongest obstacle to making UQ available as a tool is the variety of different digital control systems for automated labs developed around the world. A common abstraction layer is needed to exchange data processing workflows between different locations.

CRediT authorship contribution statement

Wolfgang Wiechert conceptualized the review paper, surveyed the literature, and finalized the manuscript. Laura Helleckes and Katharina Nöh significantly contributed to the collection and evaluation of literature, wrote some paragraphs, streamlined the text, and did major revisions.

Data Availability

No data were used for the research described in the article.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used DeepL Write in order to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgements

The authors thank their colleagues in the Helmholtz-School for Data Science in Life, Earth and Energy (HDS-LEE) for many fruitful discussions. The authors thank Eric von Lieres for his advice on process modeling.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Tellechea-Luzardo J, Otero-Muras I, Goñi-Moreno A, Carbonell P: **Fast biofoundries: coping with the challenges of biomanufacturing**. *Trends Biotechnol* 2022, **40**:831-842.

2. •• HamediRad M, Chao R, Weisberg S, Lian J, Sinha S, Zhao H: **Towards a fully automated algorithm driven platform for biosystems design**. *Nat Commun* 2019, **10**:5150.

This is one of the first papers to demonstrate the power of autonomous robotic systems in optimizing biological systems. As a proof of concept, it shows how ML algorithms can autonomously optimize a biosynthetic pathway. To this end, the BioAutomata automation platform is introduced.

3. Kaspersetz L, Schröder-Kleeberg F, Mione FM, Martinez EC, Neubauer P, Cruz-Bournazou MN: **Automation of experimental workflows for high throughput robotic cultivations**. *Comput Aided Chem Eng* 2024, **53**:2971-2976.

4. •• Helleckes LM, Osthege M, Wiechert W, Von Lieres E, Oldiges M: **Bayesian calibration, process modeling and uncertainty quantification in biotechnology**. *PLoS Comput Biol* 2022, **18**:e1009223.

This is an in-depth discussion of the calibration problem for heteroscedastic noise from statistical and Bayesian perspectives. State-of-the-art interactive Python packages for fitting calibration curves in the context of automated labs are supplied. Applications for assay evaluation, backscatter data, and growth curves obtained by a robotic system are presented.

5. Reiter A, Herbst L, Wiechert W, Oldiges M: **Need for speed: evaluation of dilute and shoot-mass spectrometry for accelerated metabolic phenotyping in bioprocess development**. *Anal Bioanal Chem* 2021, **413**:3253-3268.

6. • Casas A, Bultelle M, Kitney R: **An engineering biology approach to automated workflow and biodesign**. *Synth Biol* 2024, **9**:ysae009.

This paper describes a software architecture for workflow orchestration in a biofoundry. The solution uses directed graphs to represent workflows, orchestrators to execute them, and existing standards. A simple example is an assay based on a workflow involving dilution, measurement, and data analysis.

7. Porr M, Lange F, Marquard D, Niemeyer L, Lindner P, Scheper T, *et al.*: **Implementing a digital infrastructure for the lab using a central laboratory server and the SiLA2 communication standard**. *Eng Life Sci* 2021, **21**:208-219.

8. Smith RC: **Uncertainty Quantification: Theory, Implementation, And Applications**. Second ed., SIAM; 2024.

9. In *Uncertainty in Biology: A Computational Modeling Approach*. Edited by Geris L, Gomez-Cabrero D. Springer; 2016.

10. Kimpton LM, Paun LM, Colebank MJ, Volodina V: **Challenges and opportunities in uncertainty quantification for healthcare and biological systems**. *Philos Trans R Soc A* 2025, **383**:20240232.

11. Balsa-Canto E, Campo-Manzanares N, Moimenta AR, Roudaut G, Troitiño-Jordedo D: **Quantifying and managing uncertainty in systems biology: mechanistic and data-driven models**. *Curr Opin Syst Biol* 2025, **42**:100557.

12. Klu JK, Puch-Solis R, Mudie R, Marland V, Nic Daeid N: **Open-access measurement uncertainty valculator MUCalc and its application to the quantification of etizolam in tablets using high-performance liquid chromatography (HPLC)**. *Sci Justice* 2025, **65**:70-81.

13. • Tillack J, Paczia N, Nöh K, Wiechert W, Noack S: **Error propagation analysis for quantitative intracellular metabolomics**. *Metabolites* 2012, **2**:1012-1030.

This paper analyzes UP in complex experimental protocols for quantitative metabolomics using Gaussian error propagation and Monte Carlo methods. This enables a thorough examination of how each error source contributes to the final result.

14. • Beasley-Green A, Heckert NA: **Estimation of measurement uncertainty for the quantification of protein by ID-LC–MS/MS**. *Anal Bioanal Chem* 2023, **415**:3265-3274.

This is a nice example of the application of cause-effect diagrams, bottom-up error modeling, and Gaussian error propagation to a protein quantification protocol.

15. •• Van Mourik S, Ter Braak C, Stigter H, Molenaar J: **Prediction uncertainty assessment of a systems biology model requires a sample of the full probability distribution of its parameters**. *PeerJ* 2014, **2**:e433.

This study uses six different systems biology models to show that parameter uncertainty does not automatically lead to uncertain model predictions. The study emphasizes that separating parameter estimation and prediction without losing information is only possible if the full distribution (or respective sample) is propagated.

16. Osberghaus A, Baumann P, Hepbildikler S, Nath S, Haindl M, von Lieres E, *et al.*: **Detection, quantification, and propagation of uncertainty in high-throughput experimentation by Monte Carlo methods**. *Chem Eng Technol* 2012, **35**:1456-1464.

17. Kim SY, Shin DW, Hyun J, Kwon NH, Cheong JC, Paeng KJ, *et al.*: **Uncertainty evaluation for the quantification of urinary amphetamine and 4-hydroxyamphetamine using liquid chromatography–tandem mass spectrometry: comparison of the guide to the expression of uncertainty in measurement approach and the Monte Carlo method with R**. *Molecules* 2023, **28**:6803.

18. • Garmulewicz A, Sabin N, Reed-Tsochas F: **Conformal prediction for uncertainty quantification in dynamic biological systems**. *PLoS One* 2025, **20**:e0321963.

Conformal prediction is a state-of-the art approach to overcome the computational complexity of MC methods while maintaining their power for UQ. The paper develops novel conformal algorithms specifically designed for dynamic biological models, providing rigorous, distribution-free uncertainty quantification with non-asymptotic guarantees across predictive tasks. These methods offer scalable, flexible alternatives to conventional Bayesian methods, especially suited for complex, real-time systems where rapid and robust UQ is critical.

19. • Gelman A, Vehtari A, Simpson D, Margossian CC, Carpenter B, Yao Y, *et al.*: **Bayesian workflow**. *arXiv* 2020, arXiv:2011.01808 [stat.ME].

An in depth introduction to the workflows of Bayesian statistics written by one of the leading scientists in the field. A textbook on this theme is in preparation.

20. Linden NJ, Kramer B, Rangamani P: **Bayesian parameter estimation for dynamical models in systems biology**. *PLoS Comput Biol* 2022, **18**:e1010651.

21. Groves T, Cowie NL, Nielsen LK: **Bayesian regression facilitates quantitative modeling of cell metabolism**. *ACS Synth Biol* 2024, **13**:1205-1214.

22. Theorell A, Leweke S, Wiechert W, Nöh K: **To be certain about the uncertainty: Bayesian statistics for ^{13}C metabolic flux analysis**. *Biotechnol Bioeng* 2017, **114**:2668-2684.

23. Makrygiorgos G, Berliner AJ, Shi F, Clark DS, Arkin AP, Mesbah A: **Data-driven flow-map models for data-efficient discovery of dynamics and fast uncertainty quantification of biological and biochemical systems**. *Biotechnol Bioeng* 2023, **120**:803-818.

24. Narayan A, Liu Z, Bergquist JA, Charlebois C, Rampersad S, Rupp L, *et al.*: **UncertainSCI: uncertainty quantification for computational models in biomedicine and bioengineering**. *Comput Biol Med* 2023, **152**:106407.

25. Deneer A, Molenaar J, Fleck C: **Spectral expansion methods for prediction uncertainty quantification in systems biology**. *Front Syst Biol* 2024, **4**:1419809.

26. Alden K, Cosgrove J, Coles M, Timmis J: **Using emulation to engineer and understand simulations of biological systems**. *IEEE/ACM Trans Comput Biol Bioinform* 2020, **17**:302-315.

27. • Villaverde AF, Raimúndez E, Hasenauer J, Banga JR: **Assessment of prediction uncertainty quantification methods in systems biology**. *IEEE/ACM Trans Comput Biol Bioinform* 2023, **20**:1725-1736.
This is a well-written comparison of four different UQ techniques (Gaussian, Bayesian, likelihood, and ensembles) applied to four different biological regression models of various dimensions. It includes a thorough discussion of the pros and cons of each method.

28. Miller JN, Miller JC, Miller RD: **Statistics and Chemometrics for Analytical Chemistry**. Seventh ed., Pearson; 2018.

29. Štrumbelj E, Bouchard-Côté A, Corander J, Gelman A, Rue H, Murray L, *et al.*: **Past, present and future of software for Bayesian inference**. *Stat Sci* 2024, **39**:46-61.

30. Fjelde TE, Xu K, Widmann D, Tarek M, Pfiffer C, Trapp M, *et al.*: **Turing.jl: a general-purpose probabilistic programming language**. *ACM Trans Probabilistic Mach Learn* 2025, **1**:14, https://doi.org/10.1145/371189

31. Cabezas A, Corenflos A, Lao J, Louf R, Carnec A, Chaudhari K, *et al.*: **BlackJAX: composable Bayesian inference in JAX**. *arXiv* 2024, arXiv:2402.10797 [cs.MS].

32. Bingham E, Chen JP, Jankowiak M, Obermeyer F, Pradhan N, Karaletsos T, *et al.*: **Pyro: deep universal probabilistic programming**. *J Mach Learn Res* 2025, **20**:1-6.

33. Abril-Pla O, Andreani V, Carroll C, Dong L, Fonnesbeck CJ, Kochurov M, *et al.*: **PyMC: a modern, and comprehensive probabilistic programming framework in python**. *PeerJ Comput Sci* 2023, **9**:e1516.

34. •• Feinberg M, Rudaz S: **Quantification, Validation and Uncertainty in Analytical Sciences: An Analyst's Companion**. Wiley; 2024.
This is a timely ' must read ' book on the principles of analytical science, with a strong focus on UQ in modern analytical workflows. Hopefully, this will prevent the unnecessary reinvention of methods when UQ is applied to robotic labs.

35. Rozet E, Marini RD, Ziemons E, Boulanger B, Hubert P: **Advances in validation, risk and uncertainty assessment of bioanalytical methods**. *J Pharm Biomed Anal* 2011, **55**:848-858.

36. Visconti G, Boccard J, Feinberg M, Rudaz S: **From fundamentals in calibration to modern methodologies: a tutorial for small molecules quantification in liquid chromatography–mass spectrometry bioanalysis**. *Anal Chim Acta* 2023, **1240**:340711.

37. Ko SC, Cho M, Lee HJ, Woo HM: **Biofoundry palette: planning-assistant software for liquid handler-based experimentation and operation in the biofoundry workflow**. *ACS Synth Biol* 2022, **11**:3538-3543.

38. Lee JH, Choi JH, Youn JS, Cha YJ, Song W, Park AJ: **Comparison between bottom-up and top-down approaches in the estimation of measurement uncertainty**. *Clin Chem Lab Med CCLM* 2014, **53**:1025-1032.

39. Heymann W, Glaser J, Schlegel F, Johnson W, Rolandi P, Von Lieres E: **Advanced error modeling and Bayesian uncertainty quantification in mechanistic liquid chromatography modeling**. *J Chromatogr A* 2023, **1708**:464329.

40. Schmidt PJ, Emelko MB, Reilly PM: **Quantification of analytical recovery in particle and microorganism enumeration methods**. *Environ Sci Technol* 2010, **44**:1705-1712.

41. Patrone PN, DiSalvo M, Kearsley AJ, McFadden GB, Cooksey GA: **Reproducibility in cytometry: signals analysis and its connection to uncertainty quantification**. *PLoS One* 2023, **18**:e0295502.

42. DeJaco RF, Roberts MJ, Romsos EL, Vallone PM, Kearsley AJ: **Reducing bias and quantifying uncertainty in fluorescence produced by PCR**. *Bull Math Biol* 2023, **85**:83.

43. Gika HG, Theodoridis GA, Plumb RS, Wilson ID: **Current practice of liquid chromatography–mass spectrometry in metabolomics and metabonomics**. *J Pharm Biomed Anal* 2014, **87**:12-25.

44. Nießer J, Osthege M, Von Lieres E, Wiechert W, Noack S: **PeakPerformance – a tool for Bayesian inference-based fitting of LC-MS/MS peaks**. *J Open Source Softw* 2024, **9**:7313.

45. Röst HL, Sachsenberg T, Aiche S, Bielow C, Weisser H, Aicheler F, *et al.*: **OpenMS: a flexible open-source software platform for mass spectrometry data analysis**. *Nat Methods* 2016, **13**:741-748.

46. Tyanova S, Temu T, Cox J: **The MaxQuant computational platform for mass spectrometry-based shotgun proteomics**. *Nat Protoc* 2016, **11**:2301-2319.

47. Yang C, Hsiao YC, Lee CC, Yu JS: **Systematic evaluation of chromatographic peak quality for targeted mass spectrometry via variational autoencoder**. *Anal Chem* 2024, **96**:2849-2856.

48. • Zhang Z, Yu H, Wong-Ma E, Dokouhaki P, Mostafa A, Shavadia JS, *et al.*: **Reducing quantitative uncertainty caused by data processing in untargeted metabolomics**. *Anal Chem* 2024, **96**:3727-3732.
The quantification of raw data (here, LC-MS peaks) is often incorrect if the underlying ideal peak models do not match signals with irregularities. In an automated setting, there is no time for manual correction. This study focuses on automatically detecting poorly behaved peaks using support vector machine-based ML.

49. Chai B, Efstathiou C, Yue H, Draviam VM: **Opportunities and challenges for deep learning in cell dynamics research**. *Trends Cell Biol* 2024, **34**:955-967.

50. Paul RD, Seiffarth J, Rügamer D, Scharr H, Nöh K: **How to make your cell tracker say "I dunno!"**. *arXiv* 2025, arXiv:2503.09244 [cs.CV].

51. Kendall A, Gal Y: **What uncertainties do we need in Bayesian deep learning for computer vision?** *arXiv* 2017, arXiv:1703.04977 [cs.CV].

52. Kune C, Tielens S, Baiwir D, Fléron M, Vandormael D, Eppe G, *et al.*: **Significant impact of consumable material and buffer composition for low cell number proteomic sample preparation**. *Anal Chem* 2025, **97**:3836-3845.

53. Helleckes LM, Müller C, Griesbach T, Waffenschmidt V, Moch M, Osthege M, *et al.*: **Explore or exploit? A model-based screening strategy for PETase secretion by *Corynebacterium glutamicum***. *Biotechnol Bioeng* 2023, **120**:139-153.

54. Wulff JE, Mitchell MW: **A comparison of various normalization methods for LC/MS metabolomics data**. *Adv Biosci Biotechnol* 2018, **9**:339-351.

55. Wu Y, Li L: **Sample normalization methods in quantitative metabolomics**. *J Chromatogr A* 2016, **1430**:80-95.

56. Lawson CE, Martí JM, Radivojevic T, Jonnalagadda SVR, Gentz R, Hillson NJ, *et al.*: **Machine learning for metabolic engineering: a review**. *Metab Eng* 2021, **63**:34-60.

57. Espinel-Ríos S, Montaño López J, Avalos JL: **Omics-driven hybrid dynamic modeling of bioprocesses with uncertainty estimation**. *Biochem Eng J* 2025, **216**:109637.

58. Narayanan H, Von Stosch M, Feidl F, Sokolov M, Morbidelli M, Butté A: **Hybrid modeling for biopharmaceutical processes: advantages, opportunities, and implementation**. *Front Chem Eng* 2023, **5**:1157889.

59. Lagergren JH, Nardini JT, Baker RE, Simpson MJ, Flores KB: **Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. Lavrik I, Herausgeber**. *PLoS Comput Biol* 2020, **16**:e1008462.

60. Hu H, Cheng Q, Guo S, Wen H, Zhang J, Song Y, *et al.*: **A framework for parameter estimation and uncertainty quantification in systems biology using quantile regression and physics-informed neural networks**. *Bull Math Biol* 2025, **87**:60.

61. • Gawlikowski J, Tassi CRN, Ali M, Lee J, Humt M, Feng J, *et al.*: **A survey of uncertainty in deep neural networks**. *Artif Intell Rev* 2023, **56**:1513-1589.
Though written from an earth system research perspective, this is a comprehensive description of sources of uncertainty in deep neural networks, recent methodological developments, and the practical limitations of uncertainty quantification for mission- and safety-critical real-world applications.

62. Rybiński M, Möller S, Sunnåker M, Lormeau C, Möller J: **TopoFilter: a MATLAB package for mechanistic model identification in systems biology**. *BMC Bioinform* 2020, **21**:34.

63. Theorell A, Jadebeck JF, Wiechert W, McFadden J, Nöh K: **Rethinking ^{13}C-metabolic flux analysis – the Bayesian way of flux inference**. *Metab Eng* 2024, **83**:137-149.

64. Bartley B, Beal J, Rogers M, Bryce D, Goldman RP, Keller B, *et al.*: **Building an open representation for biological protocols**. *ACM J Emerg Technol Comput Syst* 2023, **19**:28.

65. Ye D, Veen L, Nikishova A, Lakhlili J, Edeling W, Luk OO, *et al.*: **Uncertainty quantification patterns for multiscale models**. *Philos Trans R Soc A* 2021, **379**:20200072.

66. Gutenkunst R, Waterfall J, Casey F, Brown K, Myers C, Sethna J: **Universally sloppy parameter sensitivities in systems biology models**. *PLoS Comp Biol* 2007, **3**:e189.

67. Margossian CC: **A review of automatic differentiation and its efficient implementation**. *WIREs Data Min Knowl Discov* 2019, **9**:e1305.

68. •• Thelen A, Zhang X, Fink O, Lu Y, Ghosh S, Youn BD, *et al.*: **A comprehensive review of digital twin – Part 2: roles of uncertainty quantification and optimization, a battery digital twin, and perspectives**. *arXiv* 2022, arXiv:2208.12904 [cs.LG].
Together with Part 1, this paper provides an exhaustive overview of DT methodology, particularly w.r.t. UQ. Although it is written from the perspectives of mechanical, industrial, and systems engineering, where DTs are already well-established, most concepts can be transferred to DTs for automated experimentation.

69. Radivojević T, Costello Z, Workman K, Garcia Martin H: **A machine learning automated recommendation tool for synthetic biology**. *Nat Commun* 2020, **11**:4879.