Chemeca2026
Innovate. Integrate. Impact.
28 – 30 September 2026
Melbourne, Australia

Hazards Australasia 2026

Chemeca 2026 and Hazards Australasia
28 – 30 September, Melbourne, Australia

Addressing Small Data Challenges in Biopharmaceutical Development and Manufacturing: A Mini Review of Multi-Fidelity Techniques

Mohammad Golzarijalal[a,c], Uwe Aickelin[a,c], Ellen Otte[b,c]

[a] School of Computing and Information Systems, The University of Melbourne, Parkville, VIC, 3010, Australia

[b] CSL Innovation, 655, Elizabeth Street, Melbourne, VIC 3000, Australia

[c] The ARC Digital Bioprocess Development Hub, The University of Melbourne, Australia

Email address: m.golzarijalal@unimelb.edu.au

ABSTRACT

The growing demand for biopharmaceutical products reflects their effectiveness in medical treatments. However, developing new biopharmaceuticals remains a major bottleneck, often taking up to a decade before market approval. Machine learning (ML) models have the potential to accelerate this process, but their success depends on access to large and diverse datasets for training. Multi-fidelity ML techniques offer a promising solution by integrating abundant, low-cost and less accurate low-fidelity (LF) data with limited, expensive and more accurate high-fidelity (HF) data. In this framework, LF data capture global system trends, while HF data refine and align model predictions with the available ground truth. Such integration can substantially reduce development costs and timelines by minimizing the need to acquire HF data, for example, through extensive experimental campaigns. This presentation reviews developments in surrogate modelling within the biopharmaceutical context, including Gaussian processes, neural networks and physics-informed approaches. It also provides insights on interpretability and explainability of multi-fidelity models, as well as practical recommendations for identifying appropriate LF and HF data. Existing research has primarily focused on upstream processing and drug discovery, highlighting opportunities to extend these methods to other stages like downstream processing. While Gaussian processes and neural networks remain most frequently used models, emerging architectures such as Transformer and diffusion models present promising directions for future research.

KEY WORDS

Multi-fidelity modelling, Biopharmaceutical processes, Gaussian processes, Neural networks,

Graph neural networks, Model interpretability

BIOGRAPHY

Dr. Mohammad Golzarijalal earned his B.Sc. in Chemical Engineering from Amirkabir University of Technology, Iran, where he also completed his master's degree, focusing on mathematical modeling of microalgae flocculation. He later joined Prof. Gras's research group in the Chemical Engineering Department at the University of Melbourne, earning his PhD in 2024. Following this, he joined the Digital Bioprocessing Development Hub in the Computer and Information Systems Department at the same university. There, in collaboration with CSL, one of the world's leading pharmaceutical companies, he addresses industry-relevant challenges using mechanistic and data-driven modeling techniques.

CONFERENCE PROGRAM

Please indicate which conference program your abstract relates to:

☐ Hazards Australasia

☒ Chemeca