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

Hazards Australasia 2026

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

Causality-Informed Feature Selection for Cross-Scale Generalisation in Hybrid Bioprocess Models

Kaleb F. Ferede [a,b,c], Ellen Otte [c,d], Greg Martin [a,b,c] , Sally L. Gras [a,b,c]

a Department of Chemical Engineering, The University of Melbourne, Parkville, VIC 3010, Australia.
b The Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Parkville, Vic 3010, Australia.
c ARC Digital Bioprocess Development Hub, The University of Melbourne, Parkville, VIC 3010, Australia.
d CSL Innovation, 655 Elizabeth Street, Melbourne, VIC 3000, Australia

kaleb.ferede@student.unimelb.edu.au

ABSTRACT

Hybrid models are increasingly used in bioprocessing, because they combine mechanistic understanding with the flexibility of machine learning. Their generalisation across different scales remains limited, however, particularly in mammalian cell culture processes, where changes in operating conditions can alter process behaviour. One factor that heavily influences performance is the choice of input features. In practice, feature selection is often based on correlation or optimising for model accuracy on the training data. This can favour dataset-specific relationships that fail when applied to unseen conditions. In this study, we first demonstrate through a systematic ablation study that the choice of input features has a stronger influence on cross-scale generalisation performance than model architecture. Motivated by this, we introduce Rate-Targeted Causal Subset Selection (RT-CSS). RT-CSS is a statistically controlled Granger causality-based feature selection framework that identifies process variables with predictive, temporal relevance to key culture outcomes, such as viable cell density (VCD) and titre concentration.

To assess its effectiveness, the RT-CSS method was benchmarked against established feature selection methods in several bioprocessing scenarios in which hybrid models trained on one scale were applied to scale-up or scale-down experiments. The results show that RT-CSS outperforms baseline approaches, with the largest gains in performance observed in the more challenging generalisation tasks. The framework also produced more targeted and interpretable feature subsets, reducing reliance on broad and less selective input feature sets. These findings demonstrate that causality-informed feature selection improves the robustness and generalisation of hybrid bioprocess models, with direct implications for more reliable scale-up predictions in biopharmaceutical manufacturing.

KEY WORDS

Hybrid modelling, Machine Learning, Feature Selection, Granger Causality, Mammalian Cell Culture

BIOGRAPHY

Kaleb Ferede is a second-year PhD candidate at the University of Melbourne. After earning a double degree in Chemical Engineering (1st Class Honours)/ Biotechnology from RMIT University, he joined the University of Melbourne to pursue his PhD. His research focuses on developing machine learning and hybrid modelling approaches to improve bioprocess development and manufacturing. As part of the ARC Digital Bioprocess Development Hub, a broader collaborative research community, his work contributes to the digitisation of the biopharmaceutical industry.

CONFERENCE PROGRAM

Please indicate which conference program your abstract relates to:

☐ Hazards Australasia

☒ Chemeca