Strategies for robust and efficient scientific machine learning in reaction engineering

Kian Hajireza[1,2*], David Eklund[1], Ronnie Andersson[2]

1 RISE Research Institutes of Sweden, Industrial Systems, SE-16440, Stockholm, Sweden;

2 Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-41296 Gothenburg, Sweden

**Corresponding author: kian.hajireza@ri.se*

Highlights

- Methodology for efficient development of hybrid chemical reactor models.
- Pre-training remains efficient and robust in the presence of noise.
- Computational time reduced by approximately 70%.
- The generalization and extrapolation capabilities were elucidated.

1. Introduction

Scientific machine learning and hybrid modelling are used to bridge the gap between machine learning and mechanistic modelling [1]. It has been shown that hybrid modelling methods such as Neural Ordinary Differential Equations (neural-ODEs) [2] and Universal Differential Equations (UDEs) [3] can be used as an alternative to a mechanistic model when data is available and information regarding the systems are limited. Recently, this has found applications within chemical engineering [4]. However, a limitation with the commonly used hybrid modelling methodologies is that they are computationally costly in that they involve solving ODEs and partial differential equations (PDEs) numerous times during the training. It has been shown that there is a way of circumventing the need for an ODE solver by using a collocation-based training method [5] for a significantly improved computational time. Furthermore, the application of collocation-based training has been demonstrated as a form of pre-training [6]. However, both studies [5,6] only used one single set of initial conditions and utilized a generous amount of observation data for training their model. This work explores a novel pre-training strategy for robust and efficient training of hybrid models for chemical reactors, using a large set of initial conditions and limited observation data as would be the case if scientific machine learning is aiming at modelling reaction kinetics. Additionally, this study investigates methods to improve generalization and extrapolation capabilities of hybrid models.

2. Methods

This work aims at exploring pre-training for hybrid models such as neural-ODEs for chemical reactors. A hybrid model was developed for a non-isothermal batch reactor with multiple reactions including both parallel and consecutive chemical reactions. The reactions studied in this work were of higher order. The data for this work was generated synthetically from a reactor model. A design of experiments approach was used to train the neural network, splitting the initial conditions into a training and validation set. For a more realistic evaluation, the pre-training approach was also evaluated when Gaussian noise was present in the data, with a standard deviation up to 0.1. When noise was present, the numerical estimates of derivatives based on the data are significantly misleading, therefore spline interpolation was used to approximate the derivatives of the data and the spline approximation was used to pre-train the neural networks when noise was present.

The pre-training was performed by first training the neural network against the derivatives of the data, that were numerically estimated from linear interpolation and finite difference method. With this approach, the need for an ODE solver is initially circumvented. Once the network has been pre-trained for a sufficient number of iterations; the hybrid model is solved and trained against the actual data instead of the numerical derivatives of the data. Thus, the model gets a significant head start by using pre-training thus, only utilizing an ODE solver for a limited number of iterations. Note that the same network architecture that was used in the pre-training is used when performing the regular training.

The loss function measures the degree of error in a prediction model and is minimized in machine learning. The loss function that was used for pre-training was also used during the actual training as a form of regularization, forming an alternative loss function. The reason for this is to investigate the degree of generalization with this approach. The approach taken in this study was benchmarked against other hybrid modelling methods such as using a neural-ODE without any pre-training.

3. Results and discussion

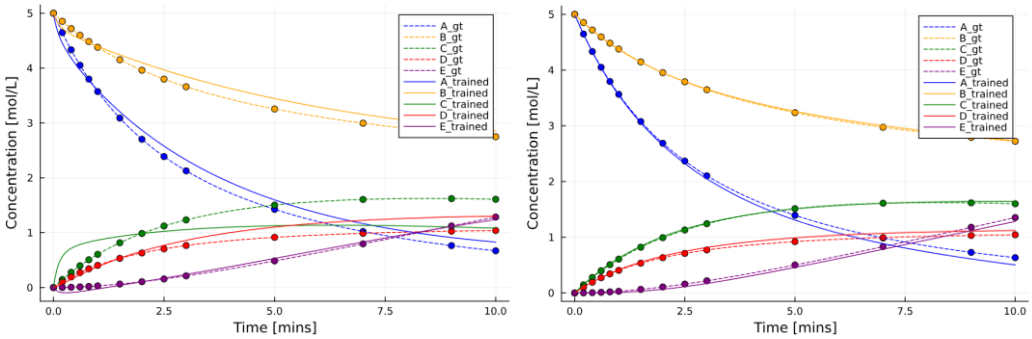

Figure 1. Concentration profiles predicted after a.) 300 regular training iterations without pre-training (benchmark), b.) 200 pre-training iterations and 100 regular training iterations. Dashed lines show the ground truth data and solid lines show the hybrid model prediction.

Figure 1 illustrates the effect of pre-training. Figure 1b shows that the predictions of the concentration profiles are significantly closer to the data compared to Figure 1a, despite being trained for the same number of iterations. After full convergence, the average training time for the benchmark was 55 mins and the total average training time for the scenario with pre-training was 18 mins. Demonstrating that total training time is cut by approximately 70% for this specific case. Furthermore, by using the alternative loss function for the regular training, in average, an improved generalization was obtained compared to the benchmark.

4. Conclusions

As the results demonstrate, the method is capable of achieving a faster computational time and in average, a better model generalization. This work shows that pre-training has the capability to speed up hybrid modelling and scientific machine learning problems significantly with a wide application area that includes chemical reaction engineering.

References

[1] M.L. Thompson, M.A. Kramer, Modeling chemical processes using prior knowledge and neural networks, AIChE Journal. 40. 1328 - 1340, 1994

[2] R.T.Q. Chen, Y. Rubanova, J. Bettencourt, D.K. Duvenaud, Neural Ordinary Differential Equations, Advances in neural information processing systems, 2018

[3] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan, Universal Differential Equations for Scientific Machine Learning, arXiv preprint arXiv:2001.04385, 2020

[4] M. Bangi, K. Kao, J. Kwon, Physics-informed neural networks for hybrid modeling of lab-scale batch fermentation for β-carotene production using saccharomyces cerevisiae, Chemical Engineering Research and Design, 179, 2022

[5] E. Roesch, C.Rackauckas, M.Stumpf, Collocation based training of neural ordinary differential equations, Statistical Applications in Genetics and Molecular Biology, 20, 2021

[6] S. Gong, Q. Meng, Y. Wang, L. Wu, W. Chen, Z. Ma, T.Y. Liu, Incorporating NODE with pre-trained neural differential operator for learning dynamics, Neurocomputing, 528, 48-58, 2023

Keywords

Reactor modelling; Scientific Machine Learning; Hybrid modelling; Pre-training