Comparison of machine learning based hybrid modelling methodologies for dynamic simulation of chemical reaction networks

Harry Kay[1], Fernando Vega-Ramon[1], Dongda Zhang[1]*

1 Department of Chemical Engineering, University of Manchester, Oxford Road, M1 3AL, Manchester, UK,

*Corresponding author: dongda.zhang@manchester.ac.uk

Highlights
- Developed a comparative study for different hybrid modelling strategies.
- Estimating dynamics of key parameters offers high quality, interpretable solutions.
- The discrepancy model provided no improvement over the PSSH benchmark model.
- The two-step approach learnt soft constraints imposed through parameter estimation.

1. Introduction

Historically, mathematical modelling has dominated the field of chemical reaction kinetics, where insight into the complex relationships held between predictor and response matrices are expressed such that any physical phenomena of interest can be described. This is largely achieved through developing systems of ordinary differential equations (ODE) representing the system under study. This allows for rapid predictions of current and future states, control, optimisation, and accelerated system design (reduced experimental burden). It is common however, within industrial processes, the nature of the phenomena is too complex to adequately model via first principles and inductive bias (assumptions) must be introduced to make simplifications. To circumvent this issue, the paradigm of machine learning can be exploited to account for the non-linearities within the process and mimic the dynamic system. This is accompanied by further problems such as lack of interpretability and data quality issues.

It is therefore intuitive to combine the aspects of data-driven and physical modelling to account for the flaws present within each methodology. This technique is named hybrid modelling and exists such that the kinetic component retains the underlying structure of the reaction network (ODEs), and the data-driven component accommodates for any missing information. Within the literature, two major forms of hybrid models exist: (1) estimating dynamic changes within key parameters of a kinetic model, using machine learning to capture the non-linear properties of the parameters (increased interpretability as no change to the structure is made), and (2) modifying a pre-defined kinetic structure with an unknown data-driven parameter that accounts for missing information. This allows for incomplete or incorrect model structures to be rectified and is referred to as hybrid discrepancy modelling.

2. Methods

This study focuses on a C16 hydroisomerisation reaction network (Figure 1) in which a rigorous microkinetic model was developed and reduced by (Vega-Ramon et al., 2023) and used to generate 8 training datasets with varying Pd catalyst concentrations, C and initial mass fractions of $C16$ (x_{c16}), mono branch products (x_{mo}), multi branch products (x_{mul}) and cracking byproducts (x_{cn}). A further 3 testing datasets were generated to simulate real experimental conditions of $x_{c16}, x_{mo}, x_{mul}, x_{cn} = 0$ and $C = \{0.1, 0.3, 0.5\}$ wt.%. The pseudo steady state hypothesis (PSSH) was applied to simplify the model, and the result used as a benchmark to represent more traditional modelling efforts.

$$\frac{dx_{c16}}{dt} = \frac{-(\overline{k_1} + \overline{k_2} + \overline{k_3}) \cdot x_{c16}}{1 + \overline{k_7} \cdot x_{c16} + \overline{k_8} \cdot x_{mo} + \overline{k_9} \cdot x_{cn}}$$

$$\frac{dx_{mo}}{dt} = \frac{\overline{k_1} \cdot x_{c16} - (\overline{k_4} + \overline{k_5}) \cdot x_{mo}}{1 + \overline{k_7} \cdot x_{c16} + \overline{k_8} \cdot x_{mo} + \overline{k_9} \cdot x_{cn}}$$

$$\frac{dx_{mul}}{dt} = \frac{\overline{k_2} \cdot x_{c16} + \overline{k_4} \cdot x_{mo} + \overline{k_6} \cdot x_{cn}}{1 + \overline{k_7} \cdot x_{c16} + \overline{k_8} \cdot x_{mo} + \overline{k_9} \cdot x_{cn}}$$

$$\frac{dx_{cn}}{dt} = \frac{\overline{k_3} \cdot x_{c16} + \overline{k_5} \cdot x_{mo} - \overline{k_6} \cdot x_{cn}}{1 + \overline{k_7} \cdot x_{c16} + \overline{k_8} \cdot x_{mo} + \overline{k_9} \cdot x_{cn}}$$
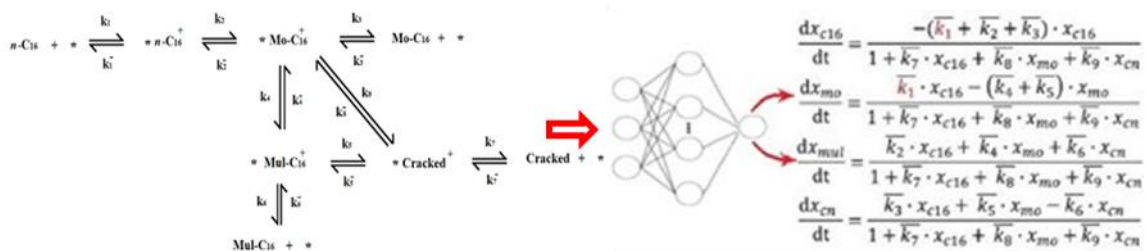
Figure 1: C16 hydroisomerisation reaction network and a schematic of the hybrid model strategy.

We can define two strategies for solving hybrid models being the one and two-step approaches. The two-step approach is defined by initially discretising the set of parameters, which are required to be modelled by machine learning, as constants over set time intervals. Once the optimal solution for the parameters was found, a neural network (ANN) was employed to learn the correlations between the input data (state parameters and catalyst concentration) and the time varying parameters that define the system until the subsequent timestep. The one-step approach differs in that we simultaneously estimate the network parameters and kinetic constants rather than separating it into two tasks. In this work, automatic differentiation is used in conjunction with gradient descent to estimate all parameters.

3. Results and discussion

A multistep ahead method is used to develop concentration profiles from a set of initial conditions. The predictive capabilities of the PSSH, two-step discrepancy, one-step and two-step time varying models are presented in Table 1.

Table 1. MAPE and MAE testing results averaged across all testing batches for different models.

Model	Mean average percentage error (MAPE) / %				Mean absolute error (MAE)			
	x_{c16}	x_{mo}	x_{mul}	x_{cn}	x_{c16}	x_{mo}	x_{mul}	x_{cn}
PSSH (kinetic)	10.3	6.0	71.1	18.9	0.030	0.023	0.030	0.022
Discrepancy (hybrid)	15.8	12.4	36.8	14.0	0.048	0.049	0.008	0.018
One-step (hybrid)	4.6	4.5	65.0	10.3	0.017	0.016	0.010	0.009
Two-step (hybrid)	7.9	7.2	16.7	10.0	0.023	0.030	0.006	0.010

With regards to the hybrid configurations, the discrepancy model performed the worst with MAE values often 80% larger than the rest with the exception of multi branch predictions. This can be explained by the structure of the ODE system not missing information and the uncertainty and error being held within the kinetic parameters themselves. The benchmark PSSH model was seemed competent in its predictions and predicted the mono branch profiles with high accuracy, however it struggled to represent the cracking byproduct and multi branch profiles. Both of the time varying parameter hybrid models simulated all mass fraction profiles with a high degree of accuracy and show similar performance when compared to one another. The large percentage error associated with the one-step approach is an artifact of values close to zero giving rise to large errors. The absolute MAPE/MAE improvement over the benchmark PSSH model for the one and two-step approach is 5.5 %/0.013 and 11.7 %/0.009. The two-step approach is restricted by the parameter estimation solution, however learns the soft constraints imposed through parameter estimation. The one-step approach can theoretically achieve lower errors, however, does not learn any physical constraints and generates more rapid changes in gradients.

4. Conclusions

In conclusion, three hybrid model configurations were explored and compared to a traditional modelling technique (PSSH) for the simulation of a C16 hydroisomerisation reaction network. The time varying strategy outperforms the benchmark model, offering improvements in MAE over 34%. The discrepancy model failed to improve upon the benchmark model and provided a lower quality representation of the system. Through the use of parameter estimation, automatic differentiation, neural networks, and kinetic modelling, high quality solutions that incur low computational expense can be attained which offer increased interpretability as opposed to black box modelling techniques. This study outlines the potential for hybrid modelling within the chemical reaction kinetics and on a wider scale.

References

[1]   Vega-Ramon, F., Wang, W., Wu, W., & Zhang, D. (2023). Developing a rigorous chemical reaction network reduction strategy for n-hexadecane hydroisomerisation, 52, 157-162