Electronic Supplementary Material

An integrated approach for machine-learning-based system identification of dynamical systems under control: application towards the model predictive control of a highly nonlinear reactor system

Ewan Chee[1], Wee Chin Wong[2], Xiaonan Wang (✉)[1]

1 Department of Chemical & Biomolecular Engineering, Faculty of Engineering, National University of Singapore, Singapore 117585, Singapore

2 Chemical Engineering & Food Technology Cluster, Singapore Institute of Technology, Singapore 138683, Singapore

E-mail: chewxia@nus.edu.sg

Appendixes

Appendix A. Values for model parameters.

Table A1 Values for model parameters of CSTR system.

Constant	Value
C_{A0}	0.8
$k_{0,j\in\{1,..,4\}}$	$[1.0, 0.7, 0.1, 0.006]^T$
$\left[\frac{E}{RT_0}\right]_{j\in\{1,...,4\}}$	$[8.33,10.0,50.0,83.3]^T$

Appendix B. Model hyperparameters.

Table B1 Candidate models f_i and their hyperparameters η_i.

f_i	η_i
Polynomial regression	Polynomial degree $\in \{1, \dots, 10\}$
SVR	Polynomial degree $\in \{1, \dots, 10\}$. C $\in [1,3]$. $\epsilon \in [0,0.3]$.

DT regression	Polynomial degree $\in \{1, \dots, 10\}$. Min samples split $\in \{2, \dots, 40\}$. Min samples leaf $\in \{1, \dots, 20\}$. Max features $\in$ {'Square root', 'Log 2', 'None'}.
ET regression	Polynomial degree $\in \{1, \dots, 10\}$. Min samples split $\in \{2, \dots, 15\}$. Min samples leaf $\in \{1, \dots, 10\}$. Number of estimators $\in \{70, \dots, 100\}$.
GB regression	Polynomial degree $\in \{1, \dots, 10\}$. Min samples split $\in \{2, \dots, 40\}$. Min samples leaf $\in \{1, \dots, 20\}$. Number of estimators $\in \{50, \dots, 100\}$. Learning rate $\in [0.01, 0.2]$.

Appendix C. Tuned parameters for NMPC.

Table C1 Tuned parameters for NMPC.

Property	Value
m	1
p	5
Jacobian	2-point
Bounds for q	$[0.6, 0.9]$
Bounds for T	$[0.8, 1.1]$
Bounds for Δq	$[-0.05, 0.05]$
Bounds for ΔT	$[-0.02, 0.02]$
Q	$\mathrm{diag}([100, 300])$
R	$\mathrm{diag}([10, 100])$
ODE solver	RK23
Optimizer	SLSQP

Appendix D. HO results.

Table D1 Candidate models f_i and their tuned hyperparameters $\hat{\eta}_i$.

f_i	$\hat{\eta}_i$	Validation R^2
Polynomial regression	Polynomial degree = 7.	0.559
SVR	Polynomial degree = 4. C = 2.27. ϵ = 0.02.	0.542
DT regression	Polynomial degree = 1. Min samples split = 38. Min samples leaf = 10. Max features = "None".	0.462

ET regression	Polynomial degree = 1. Min samples split = 15. Min samples leaf = 1. Number of estimators = 98.	0.523
GB regression	Polynomial degree = 2. Min samples split = 4. Min samples leaf = 14. Number of estimators = 99. Learning rate = 0.2.	0.513