12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

Data-driven nonlinear system identification and SIR particle filtering for chemical process monitoring and prediction

Sarmilan Santhakumaran * **Yuri A.W. Shardt** **

* *Covestro Deutschland AG, Leverkusen, North-Rhine Westphalia, Germany, 51365*
** *Technical University of Ilmenau, Ilmenau, Thuringia, Germany, 98694 (e-mail: sarmilan.santhakumaran@covestro.com , yuri.shardt@tu-ilmenau.de)*

Abstract: Chemical process monitoring is essential for product quality, plant efficiency, and safety. Conventional methods often prove inaccurate, particularly when dealing with nonlinear process behaviour. This paper presents a new approach that combines data-driven nonlinear system identification using smoothed L_1 regularisation and a state prediction method using a sequential importance resampling (SIR) particle filter to provide a basis for process monitoring. The results obtained from the polycondensation reaction in an operator training simulator (OTS) with real process conditions validate the effectiveness of the method in detecting anomalies, addressing challenges in nonlinear process modeling, and reliable state prediction for chemical process monitoring.

Keywords: **Process monitoring, nonlinear system identification, closed-loop, state prediction, data-driven modelling**

1. INTRODUCTION

As industrial chemical processes grow increasingly complex, the demand for reliable monitoring and diagnostic methods has risen. This necessitates the implementation of reliable system identification and prediction methods. A significant trend is the shift from model-based to data-driven methods, enhanced by modern sensors providing large datasets. Deep learning, while effective at identifying complex patterns and correlations in data, has difficulty in understanding process dynamics due to hidden models and their static nature (Guidotti et al. (2018)). However, data-driven nonlinear process modelling remains challenging due to unknown functional structures and lack of prior knowledge (Santhakumaran et al. (2023)). In the area of state estimation, well-established methods are developed in Bayesian frameworks using state-space models, but not often used in state prediction (Patwardhan et al. (2012)).

This paper introduces a new data-driven method for nonlinear system identification and prediction. Therefore, smoothed L_1 regularisation is used to simultaneously perform structure identification and parameter estimation, while initial regression is performed to preselect nonlinear candidates. The identified model allows a multi-step-ahead state prediction using a sequential importance resampling (SIR) particle filter.

2. DATA-DRIVEN NONLINEAR SYSTEM IDENTIFICATION

Data-driven system identification to be presented is based on Koopman theory and extended dynamic mode decomposition (eDMD). Koopman theory maps nonlinear process dynamics to an augmented state space using the Koopman operator and augmented state variables, called a *set of observables*. eDMD complements this by addressing infinite-dimensional challenges in Koopman theory using eigenmode decomposition and a library of nonlinear functions. Detailed derivations of both methods are provided in Korda et al. (2020) and Williams et al. (2015). As eDMD relies on using a set of observables, that may not always be given, Brunton et al. (2016) demonstrated the feasibility of eDMD using state variables. Based on these investigations, consider a general nonlinear process that can be expressed using a time-invariant, state-space model as

$$x_{k+1} = f(x_k) + g(u_k), \tag{1}$$

where $x_k \in \mathrm{R}^n$ are the state variables and $u_k \in \mathrm{R}^m$ are the controller outputs that exist in the state space Ω, and $f, g : \Omega \to \Omega$ are the process and controller dynamics, which are Lipschitz continuous. Using the results from eDMD with measured data, the nonlinear process dynamic can be described as

$$\begin{aligned} x_{k+1} &= \sum_{k=1}^{N_k} a_k \omega_k \Psi_k, \\ &= \Psi(x_k)\alpha, \end{aligned} \tag{2}$$

where the function library $\Psi(x_k)$ is given as $\Psi(x_k) = [\psi_1(x_k) \dots \psi_{N_k}(x_k)] \in \mathbb{R}^{nxN_k}$ and N_k refers to the number of function candidates $\psi_{1:N_k}(x_k)$ in a Hilbert space L^2 with a general index set U. The weighting vector for each function candidate $\psi_{1:N_k}(x_k)$ is represented by $a \in \mathbb{R}^{N_k}$, while the Koopman eigenmodes, which include the Koopman dynamics of evolution, are represented by $\omega \in \mathbb{R}^{N_k}$. The product of the Koopman eigenmodes ω and the weighting vector a is represented by $\alpha \in \mathbb{R}^{N_k}$. In

order to provide the state derivatives x_{k+1}, the Savitzky-Golay filter is used to perform numerical state estimation. The feasibility and accuracy of using state derivatives has been discussed and demonstrated in Santhakumaran et al. (2022).

2.1 Smoothed L_1 regularisation

As proportional-integral (PI) controllers are predominantly used in process industries, a typical closed-loop nonlinear process using Equation 1 can be written as

$$x_{k+1} = f(x_k) + g(u_k) + K_p z_k + K_I z_{k+1}, \tag{3}$$
$$z_k = r_k - y_k, \tag{4}$$
$$y_k = d(x_k). \tag{5}$$

where K_p is the controller gain, K_I is the controller reset time, and z_k is the control deviation and an auxiliary state to consider the closed-loop dynamics. Note that this reformulation is valid for a PI controller with a parallel structure of the controller parameters. The process dynamics of Equations 3-5 can be reformulated using the results from Equation 2 as

$$\begin{aligned} x_{k+1} &= \Psi_1(x_k)\alpha_1 + \Psi_2(u_k)\alpha_2 + K_p z_k + K_I z_{k+1}, \\ z_k &= r_k - y_k, \\ y_k &= \Psi_3(x_k)\alpha_3. \end{aligned} \tag{6}$$

Reformulating Equation 6 in matrix notation gives

$$\begin{bmatrix} x_{k+1} \\ y_k \\ z_k \end{bmatrix} = \begin{bmatrix} \Psi_1(x_k) & \Psi_2(u_k) & 0 \\ 0 & 0 & \Psi_3(x_k) \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} K_p & K_I \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z_k \\ z_{k+1} \end{bmatrix},$$

$$\mathbf{X} = \mathbf{\Psi}\boldsymbol{\alpha} + \mathbf{c}, \tag{7}$$

where $\mathbf{X} = \begin{bmatrix} x_{k+1} \\ y_k \\ z_k \end{bmatrix}$, $\mathbf{\Psi} = \begin{bmatrix} \Psi_1(x_k) & \Psi_2(u_k) & 0 \\ 0 & 0 & \Psi_3(x_k) \\ 0 & 0 & 0 \end{bmatrix}$, $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$ and $\mathbf{c} = \begin{bmatrix} K_p & K_I \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z_k \\ z_{k+1} \end{bmatrix}$.

Equation 7 shows that performing eDMD to a closed-loop process results in nonlinear system identification, combining functional structure identification and parameter estimation. However, using all function candidates in $\mathbf{\Psi}$ may introduce inaccuracies due to potential inclusion of irrelevant functions in the linear combination. Additionally, since $\mathbf{\Psi}$ is singular, inversion for regression can lead to numerical issues. Therefore, sparse regression methods are able to deal with these challenges by choosing suitable functions from $\mathbf{\Psi}$ and decreasing its dimensionality.

Sparse regression combines an optimisation problem with a regularisation term to shift the optimal solution, achieving sparsity in the parameter estimates. Therefore, a regularisation problem for Equation 7 can be expressed as

$$\min_{\alpha}\{\|\mathbf{X} - (\mathbf{\Psi}\boldsymbol{\alpha} + \mathbf{c})\|_2^2 + \lambda \left|\left|\boldsymbol{\alpha}\right|\right|_p\}, \tag{8}$$

where the regularisation term includes the hyperparameter λ and $||\cdot||_p$, which represents the L_p norm. For nonlinear system identification, L_1 regularisation method is suitable due to its ability to select suitable function candidates in $\mathbf{\Psi}$ by assigning zero to several model coefficients $\boldsymbol{\alpha}$ according to the properties of the L_1-norm. As a result, it efficiently eliminates unsuitable function candidates from $\mathbf{\Psi}$ and performs feature selection. However, optimising L_1 regularisation is challenging due to its nondifferentiability, whereas L_2 regularisation demands less effort but cannot ensure a generally sparse solution.

Therefore, an approximation of the L_1-norm based on the L_2-norm is developed for system identification to combine both properties. In this context, Fan et al. (2001) performed a comprehensive analysis using the Huber-Loss regularisation, which uses a switching approach between the L_1 and L_2 regularisations, to express an analytical solution that combines the properties of both methods. As a result, the following relationship was derived:

$$||\boldsymbol{\alpha}^{[1]}|| = ||\boldsymbol{\alpha}^{[0]}|| + \frac{1}{2}||\boldsymbol{\alpha}^{[0]}||^{-1}(\boldsymbol{\alpha}^2 - (\boldsymbol{\alpha}^{[0]})^2), \tag{9}$$

where $\boldsymbol{\alpha}^{[0]}$ is an initial vector, which needs to be set for regularisation and $\boldsymbol{\alpha}^{[1]}$ is the next iteration. For a general regression problem $\|\mathbf{y} - \mathbf{Ab}\|_2^2 + \lambda \left|\left|\boldsymbol{\alpha}\right|\right|_1$ that needs to be solved using L_1 regularisation, van Wieringen (2015) demonstrated that using the quadratic approximation to the absolute value function in the elements of $\boldsymbol{\alpha}$ during the k^{th} update and using the properties from Equation 9 leads to an iterative L_2 regularisation, that is

$$\boldsymbol{\alpha}^{[k+1]} = [\mathbf{A}^T\mathbf{A} + \frac{1}{2}\lambda||\boldsymbol{\alpha}^{[k]}||^{-1}]^{-1}\mathbf{y}^T\mathbf{A}. \tag{10}$$

Using the result from Equation 10, an analytical solution for Equation 7 can be obtained by solving the following optimisation problem:

$$\begin{aligned} \min_{\boldsymbol{\alpha}^{[k+1]}}\{||(\,\mathbf{X} - (\mathbf{\Psi}\boldsymbol{\alpha}^{[k+1]} + \mathbf{c})\,||_2^2 + \frac{1}{2}\lambda||\boldsymbol{\alpha}^{[k]}||^{-1}(\boldsymbol{\alpha}^{([k+1])})^2\} \\ -2\mathbf{\Psi}^T\mathbf{X} + 2\mathbf{\Psi}^T\mathbf{\Psi}\boldsymbol{\alpha}^{[k+1]} + \lambda||\boldsymbol{\alpha}^{[k]}||^{-1}(\boldsymbol{\alpha}^{[k+1]}) = 0 \\ \boldsymbol{\alpha}^{[k+1]} = [\mathbf{\Psi}^T\mathbf{\Psi} + \frac{1}{2}\lambda||\boldsymbol{\alpha}^{[k]}||^{-1}]^{-1}\mathbf{X}^T\mathbf{\Psi} \end{aligned} \tag{11}$$

The result of Equation 11 provides a suitable estimate that has the properties of the L_1 regularisation by achieving convergence of $\boldsymbol{\alpha}$. This result is defined as *smoothed L_1 regularisation*. Thus, smoothed L_1 regularisation promotes a sparse solution while structure identification and parameter estimation are simultaneously performed.

The application of smoothed L_1 regularisation would allow an unlimited number N_k in the function library $\mathbf{\Psi}$, that can lead to a considerable computational effort. To improve the performance, an initial regression is proposed that starts with a polynomial regression using $\mathbf{X}$. For validation, R^2 scores are calculated, defined as

$$R^2 = \frac{SSR}{TSS} \tag{12}$$

where SSR is the sum of squares due to regression that is the difference between the predicted values and the mean value, while TSS is the total sum of squares that measures the total variance in the dataset. Polynomial regression is performed by increasing the polynomial order until a sufficiently high R^2 is achieved. Based on this result, the sum series of possible nonlinear functions $\psi_{1:N_k}(x_k)$ are then compared and the suitable candidates are selected using forward selection. The linear combination of the nonlinear function candidates that exceed a R^2 score of 0.95 are added to the function library $\mathbf{\Psi}$ (Burnham et al. (2004)). Thus, M_k function candidates with $M_k \ll N_k$ are preselected for the subsequent steps.

3. STATE PREDICTION USING A SIR PARTICLE FILTER

A sequential importance sampling (SIS) particle filter, which is derived from a recursive Bayesian Filter, is used to iteratively estimate the probability density function (PDF) of process states from measurements. The filter prioritises real-time performance (Patwardhan et al. (2012), Setoodeh et al. (2022)). In practical applications, the sequential importance resampling (SIR) method is often preferred due to its lower computational requirements for calculating particle weights compared to SIS and its simpler particle generation (Haßkerl et al. (2016)). Therefore, SIR particle filtering will be used for state prediction.

Therefore, consider the nonlinear process model provided in Equation 6 and identified using the result from Equation 11. The particle filter's state estimation relies on the fundamental framework of the Bayesian filter, which is,

$$p(x_{1:k}|y_{1:k}) = p(x_{1:k-1}|y_{1:k-1})p(x_k|x_{k-1})p(y_k|x_k), \quad (13)$$

where $p(\cdot)$ is the PDF and describes the probability of the first variable, based on the observation of the second variable, $x_{1:k}$ are the state variables and $y_{1:k}$ is the process output obtained from time step 1 to k. Equation 13 outlines a method for estimating the state variable at time k using the process outputs $y_{1:k}$ and past state variables $x_{1:k-1}$ based on sampling the related PDF. If the PDF can be parametrically described with a known function, it simplifies the generation of random variables. However, this assumption does not hold when dealing with arbitrary PDFs. In such cases, the SIS method is used to perform Monte Carlo integration using a proposed density function, known as the importance density (Haßkerl et al. (2016)), that is

$$E\{f(x)\} = \int_{-\infty}^{+\infty} \frac{p(x)}{q(x)} f(x)q(x)dx. \quad (14)$$

where $E\{\cdot\}$ is the expectation operator, x are the samples drawn, $f(\cdot)$ is the process dynamics, $p(\cdot)$ the related PDF, $q(x)$ is the proposed distribution, and the ratio of $\frac{p(x)}{q(x)}$ provides the degree of similarity between the samples drawn from $q(x)$ and $p(x)$. In the context of SIS, the samples drawn are called particles. The weight $\omega(x)$ assigned to each particle is represented by

$$\omega(x) = \frac{p(x)}{q(x)}. \quad (15)$$

Using Equation 14, Equation 15 can be expressed as

$$E\{f\} = \int_{-\infty}^{+\infty} \omega(x)f(x)q(x)dx \approx \frac{1}{N}\sum_{i=1}^{N} \omega(x^{(i)})f(x^{(i)}), \quad (16)$$

Using Equation 16, Bayesian-filter state estimation from Equation 13 can be performed by providing a proposed density function q and thus providing particles. However, the weights of the particle ω must be iteratively obtained at each step resulting in increased computational effort. Therefore, Equation 15 can be rewritten using Bayes' theorem to provide a recursive expression, that is,

$$\omega_k = \frac{p(x^{(i)}_{0:k}|y_{0:k})}{q(x^{(i)}_{0:k}|y_{0:k})} = \frac{p(y_{0:k}|x^{(i)}_{0:k})p(x^{(i)}_{0:k})}{q(x^{(i)}_k, x^{(i)}_{0:k-1}|y_{0:k-1})}. \quad (17)$$

Expanding Equation 17 using Bayes' theorem and the Markov assumption that each state depends only on the current state and not on previous measurements, gives

$$\begin{aligned}\omega_k &= \frac{p(y_k|x^{(i)}_k)p(y_{0:k-1}|x^{(i)}_{0:k-1})p(x^{(i)}_k|x^{(i)}_{k-1})p(x^{(i)}_{0:k-1})}{q(x^{(i)}_k|x^{(i)}_{0:k-1}, y_{0:k-1})q(x^{(i)}_{0:k-1}|y_{0:k-1})} \\ &= \frac{p(y_k|x^{(i)}_k)p(x^{(i)}_k|x^{(i)}_{k-1})}{q(x^{(i)}_k|x^{(i)}_{0:k-1}, y_{0:k-1})} \frac{p(y_{0:k-1}|x^{(i)}_{0:k-1})p(x^{(i)}_{0:k-1})}{q(x^{(i)}_{0:k-1}|y_{0:k-1})} \\ &= \frac{p(y_k|x^{(i)}_k)p(x^{(i)}_k|x^{(i)}_{k-1})}{q(x^{(i)}_k|x^{(i)}_{0:k-1}, y_{0:k-1})} \frac{p(x^{(i)}_{0:k-1}|y_{0:k-1})}{q(x^{(i)}_{0:k-1}|y_{0:k-1})}. \end{aligned} \quad (18)$$

Comparing Equation 18 with Equation 15, it can be seen that a decomposition for the particle weights has been performed. The first fraction represents an expression based on the current measurements, while the second fraction represents the weights from the previous iteration step. Thus, a recursive expression for the particle weights can be described as

$$\omega_k = \frac{p(y_k|x^{(i)}_k)p(x^{(i)}_k|x^{(i)}_{k-1})}{q(x^{(i)}_k|x^{(i)}_{0:k-1}, y_{0:k-1})}\omega_{k-1}. \quad (19)$$

Initially, particles representing potential hypotheses about the process state are generated using the importance density. State estimation is performed using the process model identified from Equation 11, and particle weights are updated based on their correlation with measurements using Equation 19. To address *degeneracy issues*, resampling replaces low-weight particles with duplicates of high-weight ones. This SIS and resampling (SIR) method improves particle representation and approximates the state probability distribution. The state estimate is derived by averaging particles, neglecting the correction step in favor of multi-step-ahead state prediction.

4. CASE STUDY

This case study examines chemical process monitoring of a polycondensation reaction using the proposed smoothed L_1 regularisation and SIR particle filter. Figure 1 presents an overview of the polycondensation reaction process.

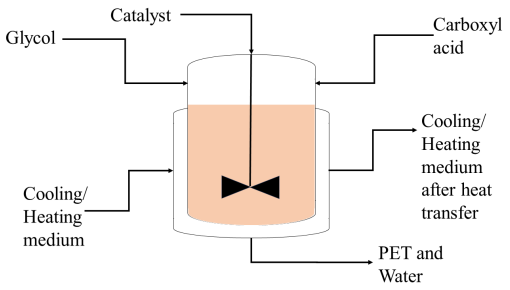

Fig. 1. Overview of PET reaction unit

The focus is on polyethylene terephthalate (PET) production, resulting from a reaction between carboxyl acid and glycol. A by-product of this reaction is water that is later distilled downstream. The reaction requires a catalyst for initiation, accompanied by specific temperature control using heating and cooling media. Efficient monitoring of the reaction rate is crucial to ensure product quality and plant safety. The simulation study is based on the use of

an operator training simulator (OTS) that incorporates plant dynamics and real-plant-related data, as well as an emulated distributed control system (DCS) and emergency shutdown system (ESD) from the plant setup. To evaluate the effectiveness of state prediction and its contribution to process monitoring, the OTS randomly activates a fouling scenario within the prediction step.

4.1 Nonlinear system identification of a polycondensation reaction

In order to provide a process model for the PET reaction behaviour, data-driven nonlinear system identification is performed using smoothed L_1 regularisation. The process is run in closed loop using four PI controllers to control the reactant flows, the catalyst flow, reactor heating, and reactor cooling. Table 1 describes the process parameters used and Figures 2-4 show the process data used for system identification.

Table 1. Process parameters used for the polycondensation reaction

Process parameters	Description	units
f_{r1}	Flow of glycol	t/h
f_{r2}	Flow of carboxyl acid	t/h
f_c	Flow of catalyst	t/h
T	Reactor temperature	°C
r_r	Reaction rate	%
C_1	Flow controller of glycol	-
C_2	Flow controller of carboxyl acid	-
C_3	Flow controller of catalyst	-
C_4	Reactor temperature controller	-
z_1	Auxiliary variable for C_1	-
z_2	Auxiliary variable for C_2	-
z_3	Auxiliary variable for C_3	-
z_4	Auxiliary variable for C_4	-

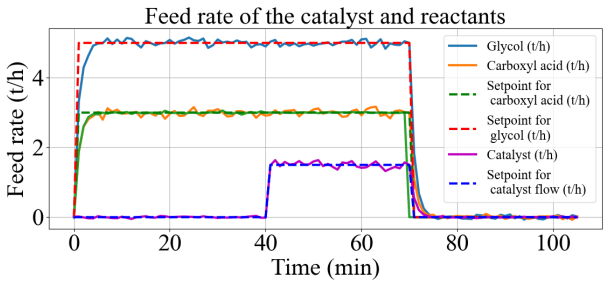

Fig. 2. Reactants and catalyst flow for the PET reaction

Figure 2 shows the flows of reactants and catalyst during the process. At the beginning, glycol and carboxyl acid are controlled using PI flow controllers with setpoints of 3 and 5 $\frac{t}{h}$ over 70 minutes. The catalyst is then added with a setpoint of 1.5 $\frac{t}{h}$ after 40 minutes for 25 minutes to start the reaction. Figure 3 shows the reactor-temperature profile. It can be seen that after 20 minutes, the reactor is preheated to 40°C. In order to prevent a safety shutdown, the reactor temperature is controlled at 220°C during the reaction. The reaction rate of PET is shown in Figure 4. The reaction rate refers to the amount of PET formed during the process. After the catalyst is added to the process, 20% PET is produced in the first 3 minutes. Thereafter, the reaction rate increases directly to 91%

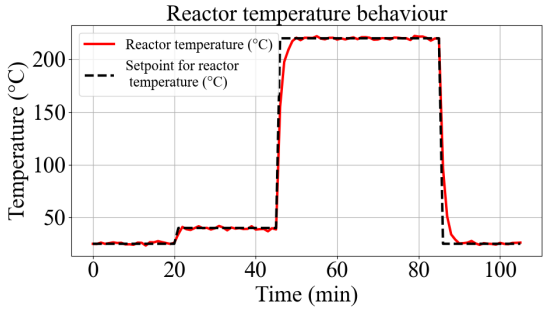

Fig. 3. Temperature behaviour of PET reaction

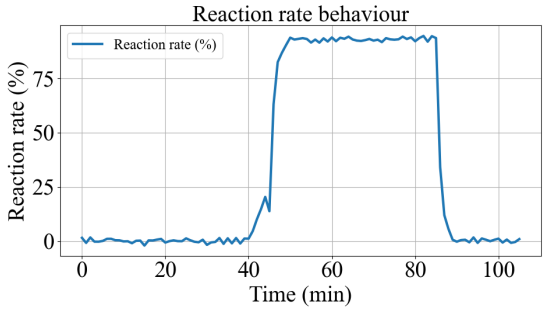

Fig. 4. Reaction rate of PET

PET and remains almost constant with uncertainties due to the reactor temperature control and measurement. The supply of reactants and catalyst is completed after 70 minutes, while the reaction is completed after 85 minutes. The product is then discharged to the downstream process for further processing and the process continues with the next batch.

Based on the dataset obtained, smoothed L_1 regularisation is performed to provide a process model. First, an initial regression based on polynomial regression and forward selection is performed in order to obtain a set of nonlinear-function candidates. The preselected nonlinear function candidates are shown in Table 2.

Table 2. Nonlinear-function candidates obtained from initial regression

Nonlinear function candidates	Regression fit (R^2 score)
exp	0.68
$(.) + (.)^2$	0.15
sin	0.07
cos	0.06
Total fit	0.96

Table 2 shows that the initial regression proposes a superposition of an exponential (exp), sinusoidal (sin and cos) and a second-order polynomial ($(.) + (.)^2$) candidates with a R^2 score of 0.96 for the polynomial regression. Thus, these candidates are considered for the next step.

Next, data-driven system identification is performed using the preselected nonlinear-function candidates from Table 2. Identification is performed using a 1000-fold cross-validation, where the process data obtained is distributed into 1000 uniform data packages of 630 data points each. In this process, 950 data packages are used as the training dataset for system identification and the remaining 50 data packages are used as the test dataset for validation. This procedure is performed 1000 times, with the training and

test datasets randomly generated. The result of the cross-validation is shown in Figure 5.

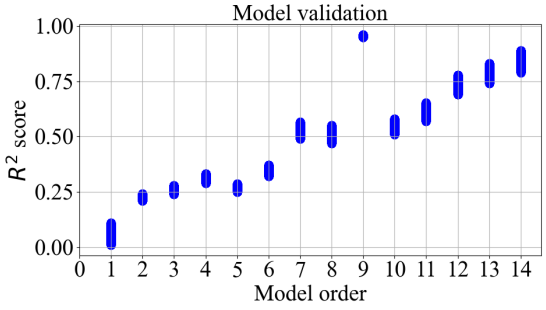

Fig. 5. Validation of smoothed L_1 regularisation using 1000-fold cross-validation

The cross-validation shown in Figure 5 evaluates the identified process models using the model order, which evaluates the number of nonlinear function candidates considered in relation to the R^2 score. The results show that the process model with 9 nonlinear function candidates has the best fit with an R^2 score that ranges between 0.91 and 0.98. Thus, the process model with 9 candidate functions shows that it was able to accurately capture the process dynamics present. As there are variations in the R^2 score range, and thus variations in the parameter estimates, the identified process parameters are selected with the averaged R^2 score of 0.955. The identified process model for the R^2 score of 0.955 can be described as

$$
\begin{aligned}
\frac{dr_r}{dt} &= -5.71 \cdot 10^4 r_r^2 \exp(-0.68T) - 1.8r_r + 4 \cdot \cos(z_4) + 79 \\
\frac{dT}{dt} &= 8.21 \cdot 10^5 r_r^2 \exp(-0.93T) + 3.2r_r^2 - 14.3r_r + C_4 \\
\frac{df_{r1}}{dt} &= 37.64 \exp(0.58 f_{r1}) + 0.45 + C_1 \\
\frac{df_{r2}}{dt} &= 37.64 \exp(0.43 f_{r2}) + 0.31 + C_2 \\
\frac{df_c}{dt} &= 37.64 \exp(0.47 f_c) + 0.24 + C_3 \\
\frac{dz_1}{dt} &= 120 + f_{r1} \\
\frac{dz_2}{dt} &= 120 + f_{r2} \\
\frac{dz_3}{dt} &= 120 + f_c \\
\frac{dz_4}{dt} &= 380 - T \\
C_1 &= (120 + f_{r1}) + 130 z_1 \\
C_2 &= (120 + f_{r2}) + 130 z_2 \\
C_3 &= (120 + f_c) + 130 z_3 \\
C_4 &= -(380 - T) + 420 z_4
\end{aligned}
\tag{20}
$$

The result of smoothed L_1 regularisation shows a nonlinear relationship between the reaction rate and temperature. As well, the flow behaviour is influenced by the controller and the previous state of itself. In addition, it can be shown that this approach allows the identification of the controller parameters (K_p: 120 and -380/ K_I: 130 and 420), if a PI controller structure is considered. This property is useful in case the controller parameters cannot be accessed. The behaviour of the model is shown by the reaction rate in Figure 6.

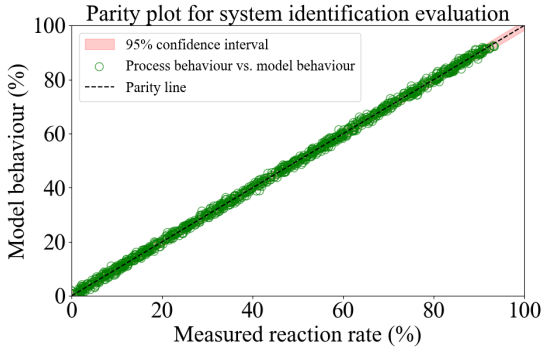

Fig. 6. Process model behaviour resulting from smoothed L_1 regularisation

Figure 6 shows a parity plot, which visually assesses the regression performance by evaluating the closeness of the measurements to the identified process model along the parity line. The close clustering of points around this line indicates the accuracy of the process model in capturing the inherent process dynamics. The comparison between the process behaviour and the model behaviour shows a reliable fit. Validation is further supported by the 95% confidence interval, which reinforces confidence in the accuracy of the model and confirms the efficency of the smoothed L_1 regularisation for data-driven nonlinear system identification.

4.2 State prediction using SIR particle filter

In the first step of the SIR particle filter, the time horizon is divided into two phases: training phase and test phase. In the training phase, the particle weights are estimated based on the process model identified from Equation 20 and updated using measurements. In the test phase, the correction step is neglected, and thus, state prediction is performed using the process model identified. In order to initialise the particle set, a univariate importance density covering the measurement range of the reaction rate is assumed and 1000 particles are randomly drawn. The performance of state estimation is shown in Figure 7.

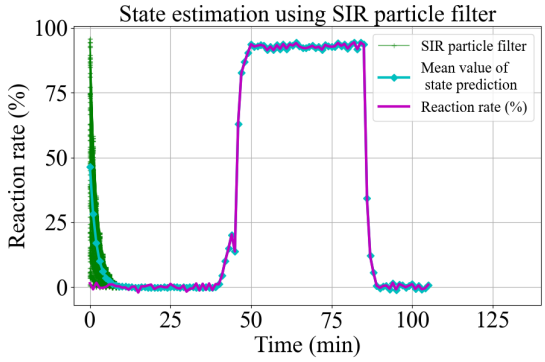

Fig. 7. State estimation using the SIR particle filter

The training phase, lasting one batch, begins with the distribution of particles over the measurement range of the reaction rate. Using the process model and measurement

obtained, the particles gradually converge towards the true process state. Particle weighting is performed every minute using Equation 19, with higher weights given to states that are more probable. Lower-weighted particles are resampled to maintain the same number of particles within the range of higher weights. After 12 minutes, precise tracking of the particle set can be ensured in the batch, which demonstrates the efficiency of this method for state estimation. Based on the training, state prediction is performed using the process model identified without a correction step. The performance of the state prediction as well as the accuracy of this approach are shown in Figure 8.

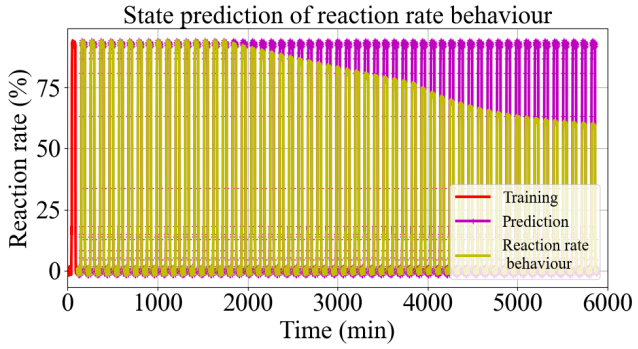

Fig. 8. State prediction using SIR particle filter

Figure 8 shows the prediction of the reaction rate over a period of 5800 minutes. The red curve represents the training phase, that is shown and explained in Figure 7. The yellow curve shows the measured reaction-rate behaviour, while the violet curve shows the predicted reaction rate using the SIR particle filter. It can be noted that the prediction closely aligns with the reaction rate up to 1800 minutes, followed by a quasilinear decrease from 1800 to 4000 minutes, and an exponential decrease from 4000 to 5800 minutes. The prediction follows the process dynamics using from the process model and remains similar to the behaviour in the training phase. The pattern of the reaction rate corresponds to the gradual fouling process randomly activated in the OTS, that involves the accumulation of fouling on reactor surfaces and pipes, reduces the heat transfer and increases the energy consumption and pressure. Consequently, these factors contribute to the observed decrease in reaction rate. As a result, this fouling can lead to overheating and reactor damage, resulting in costly repairs and a decrease in product quality. Therefore, by setting a threshold specific to the process and performing smoothed L_1 regularisation within the SIR particle filter, accurate state prediction can be performed, anomalies can be detected, and a maintenance plan can be scheduled. Thus, the proposed method can help to monitor chemical processes.

5. CONCLUSION

This paper has introduced and evaluated a data-driven nonlinear system identification and state-prediction method using smoothed L_1 regularisation and SIR particle filter for monitoring a polycondensation reaction process. The study started with data-driven system identification to capture the relevant process dynamics within a closed-loop process. Subsequently, state prediction of the reaction rate was performed based on the process model identified. The system identification was successfully performed, yielding an averaged R^2 score of 0.955, and the process model was effectively used for state prediction. Notably, the state prediction successfully detected a deviation related to fouling within the reactor, highlighting its ability to detect anomalies.

The results highlight the significant role that data-driven nonlinear system identification and state prediction contribute to process monitoring. However, it is crucial to incorporate appropriate threshold values and examine model-based fault diagnosis and isolation approaches to enhance the proposed method. Additionally, the applicability of this framework across different chemical processes and conditions should be investigated to establish and develop further the scope.

REFERENCES

S.L. Brunton, J.L. Proctor, and J.N. Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15), pp. 3932–3937, 2016.

K.P. Burnham and D.R. Anderson. Model selection and multimodel inference: A practical information-theoretic approach. *Sociological Methods & Research*, 33(2): pp. 261–304, 2004.

J. Fan and R. Li. Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties. *Journal of the American Statistical Association*, 96(456): pp. 1348–1360, 2001.

R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti and D. Pedreschi. A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*, 2018.

D. Haßkerl, M. Arshad, R. Hashemi, S. Subramanian, S. Engell. Simulation Study of the Particle Filter and the EKF for State Estimation of a Large-scale DAE-system with Multi-rate Sampling *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 490-495, 2016.

M. Korda, M. Putinar, I. Mezić. Data-driven spectral analysis of the Koopman operator. *Applied and Computational Harmonic Analysis*, Volume 48, Issue 2, pp. 599-629. ISSN 1063-5203, 2020.

S. C. Patwardhan, S. Narasimhan, P. Jagadeesan, B.Gopaluni, and S. L. Shah. Nonlinear Bayesian State Estimation: A Review of Recent Developments. *Control Engineering Practice*, vol. 20, no. 10, pp. 933-953, 2012.

S. Santhakumaran, Y.A.W. Shardt, J. Rejek and C. Maul. Data-driven nonlinear system identification of a closed-loop CSTR. *ARGESIM Report 17*, pp. 29-30, 2022.

S. Santhakumaran and Y.A.W. Shardt. Data-driven nonlinear system identification of blood glucose behaviour in Type I diabetics. *Control Engineering Practice*, 132:105405, 2023.

P. Setoodeh, S. Habibi and S. Haykin. Nonlinear Filters: Theory and Applications. *John Wiley & Sons*, 2022.

W.N. van Wieringen. Lecture notes on ridge regression. *arXiv preprint*, stat.ME, 2015.

M.O. Williams, I.G Kevrekidis and C.W. Rowley. A Data–Driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition. *J Nonlinear Sci* 25, pp. 1307–1346, 2015.