engineering proceedings

Proceeding Paper

Fault Detection of Multi-Rate Two Phase Reactor Condenser System with Recycle Using Multiple Probabilistic Principal Component Analysis †

Dhrumil Gandhi and Meka Srinivasarao *

Chemical Engineering Department, Dharmsinh Desai University, Nadiad-387001, Gujarat, India; dhrumilgandhi.ch@ddu.ac.in

* Correspondence: msrao@ddu.ac.in

† Presented at the 2nd International Electronic Conference on Processes: Process Engineering—Current State and Future Trends (ECP 2023), 17–31 May 2023; Available online: https://ecp2023.sciforum.net/.

Abstract: Fault detection in multi-rate process systems is a challenging task. Common techniques used for fault detection include threshold-based detectors, statistical detectors, and machine learning-based detectors. One such statistical detector technique is multiple probabilistic principal component analysis (MPPCA). MPPCA uses probabilistic PCA to detect fault signals from multiple sensors without down-sampling or up-sampling. This paper uses MPPCA to detect faults in a two-phase reactor–condenser system with recycle (TPRCR) with three measurement classes. These measurement data are used to build the MPPCA model using expectation maximization (EM). Based on this, T^2 and *SPE* statistics are generated for fault detection in TPRCR systems, and the MPPCA approach's effectiveness for fault detection is satisfactory.

Keywords: fault detection; multi-rate process; MPPCA

check for updates

Citation: Gandhi, D.; Srinivasarao, M. Fault Detection of Multi-Rate Two Phase Reactor Condenser System with Recycle Using Multiple Probabilistic Principal Component Analysis. *Eng. Proc.* **2023**, *37*, 91. https://doi.org/10.3390/ECP2023-14669

Academic Editor: Gade Pandu Rangaiah

Published: 17 May 2023

1. Introduction

Modern chemical industries focus on detecting and diagnosing faults as early as possible to increase production yield [1]. Effective fault-detection techniques available in the literature require the regular availability of measurements [1]. However, some variables in chemical processes are measured online, while other qualitative variables are measured offline. Measurement of these offline quality variables requires human involvement, which makes the system an irregularly sampled multi-rate system [2]. Fault-detection techniques for multi-rate systems include state space estimation techniques and data-based modelling methods. State space estimation techniques require accurate system models, which are difficult to model for complex chemical engineering systems. Compared to the above methods, another data-driven approach uses measurement data to model the system's behaviour. These data-driven methods for multi-rate systems require down-sampling, up-sampling and re-sampling. While the down-sampling approaches will lose essential information during modelling, the up-sampling methods heavily rely on the correctness of the predictions [3]. In most chemical processes, the variation in sample rates is too significant, resulting in unmanageable complexity in the re-sampling models. The MPPCA method does not require down-sampling, up-sampling and re-sampling of multi-rate data. It uses multi-rate data to build an inferential model that can handle multiple measurement classes. The MPPCA method is an extension of probabilistic principal component analysis (PPCA) which uses the EM algorithm for parameter tuning.

In this study, the effectiveness of the MPPCA method in detecting various faults for the multi-rate nonlinear chemical process TPRCR is studied, and fault detection is carried out by using T^2 and *SPE* statistics.

The remainder of this paper is organised as follows. Section 2 gives details about the MPPCA method and model parameter estimation. Then, Section 3 details the TPRCR model. Section 4 implements a fault-detection technique on TPRCR. Finally, conclusions are made in the last section.

2. MPPCA Method

The MPPCA model combines several-rate data into a single model without down- or up-sampling. In our article, we have considered the MPPCA model with three different classes of measurements, and it is given by the following equations:

$$x_1 = \varnothing_1 t + \varepsilon_1 \quad (1)$$

$$x_2 = \varnothing_2 t + \varepsilon_2 \quad (2)$$

$$x_3 = \varnothing_3 t + \varepsilon_3 \quad (3)$$

In Equations (1)–(3), $x_1 \in \mathrm{R}^{\mathrm{K1}\times\mathrm{M1}}$, $x_2 \in \mathrm{R}^{\mathrm{K2}\times\mathrm{M2}}$, and $x_3 \in \mathrm{R}^{\mathrm{K3}\times\mathrm{M3}}$ are three different rate measurements classes in which x_3 is the slowest and x_1 is the fastest measurement. $\varnothing_1 \in \mathrm{R}^{\mathrm{M1}\times\mathrm{D}}$, $\varnothing_2 \in \mathrm{R}^{\mathrm{M2}\times\mathrm{D}}$ and $\varnothing_3 \in \mathrm{R}^{\mathrm{M3}\times\mathrm{D}}$ are loading matrices with three different sampling rates. $t \in \mathrm{R}^{\mathrm{D}}$ is a latent variable which extracts a restricted link between data with varied sampling rates and helps develop one single model. The latent variable is assumed to have a Gaussian distribution with a zero mean and unit variance. $\varepsilon_1 \in \mathrm{R}^{\mathrm{M1}}$, $\varepsilon_2 \in \mathrm{R}^{\mathrm{M2}}$ and $\varepsilon_3 \in \mathrm{R}^{\mathrm{M3}}$ are used to model the corresponding isotropic Gaussian noises.

The sequence of the measurements can be altered for easier notation and visualisation on the premise that all sample variables are independent. The whole observation (V) comprises three divisions of the observed data. The first sample contains all observations with dimensions M1 + M2 + M3 (V_3), the following sample variables have dimensions M1 + M2 (V_2), and the last one contains only M1 (V_1) variables. As a result, the entire observation set is expressed as a union of all three.

$$V = V_3 \cup V_2 \cup V_1 \quad (4)$$

The EM technique is used to estimate model parameters for the MPPCA model. The method repeats the expectation step (E-step) and the maximisation step (M-step) until convergence. In the E-step, the current model parameters are utilised to estimate the posterior distributions of the latent variables. The model parameters are then adjusted in the M-step by maximising log likelihood. One reference contains a detailed explanation of the step of the EM algorithm for MPPCA training [4].

SPE statistics can be used to detect abnormal behaviour in measurements. There are three different classes of measurements, so three different *SPE* statistics are used to detect any anomaly in measurement.

$$SPE_1 = x_1 - \varnothing_1 t \quad (5)$$

$$SPE_2 = x_2 - \varnothing_2 t \quad (6)$$

$$SPE_3 = x_3 - \varnothing_3 t \quad (7)$$

Since each *SPE* statistic is compiled based on the prediction errors of different classes of measurements, it clearly shows that a given fault is caused by a certain class of measurements. The confidence bound of *SPE* statistics can be predicted by χ^2-distributed approximation: $SPE \sim g.\chi_h^2$, in which g and h are the parameters of the χ^2 distribution, and they are given by [5].

$$gh = mean(SPE) \quad (8)$$

$$2g^2h = var(SPE) \tag{9}$$

3. Two-Phase Reactor–Condenser System with Recycle

The process depicted in Figure 1 includes a two-phase reactor and condenser [6]. Reactants A and B are introduced into the reactor at molar flow rates F_A and F_B and temperatures T_A and T_B, respectively, in the vapour and liquid phases. Reactant A diffuses into the liquid phase at rate N_{A1}, where an exothermic reaction occurs, which is given by Equation (10).

$$A + B \rightarrow 2C \tag{10}$$

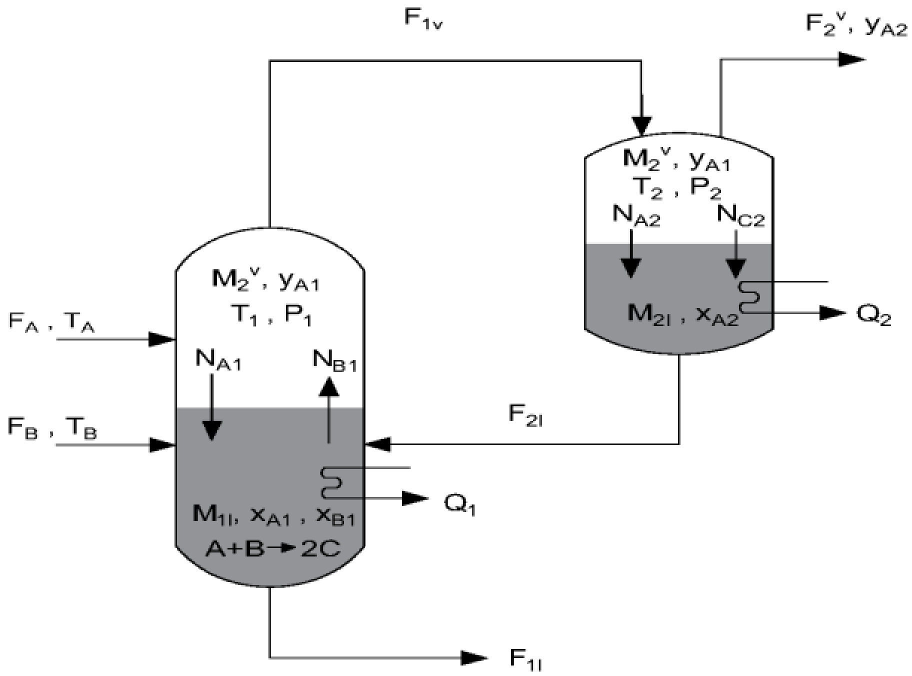

Figure 1. Schematic Diagram of TPRCR system.

Product C diffuses into the vapour phase at a rate $Nc1$, whereas reactant B is non-volatile. The interphase mass transfer resistance is assumed to be minimal, and the Arrhenius equation provides the reaction rate in the bulk liquid phase, which is given by Equation (11).

$$r_A = k_{10} \exp\left(\frac{-E_a}{RT_1}\right) M_1^l \rho x_{A1} x_{B1} \tag{11}$$

where r_A is the rate at which reactant A is consumed at temperature T_1. The preexponential factor and activation energy are denoted by k_{10} and E_a, respectively. M_1^l is the liquid molar holdup in the reactor, and ρ is the liquid density. x_{A1} and x_{B1} are A and B mole fractions in the liquid phase. For the sake of simplicity, heat capacity, density, and molar heat of vaporisation are considered to be constant and equal for all species. The liquid and vapour phases are suitable combinations. The liquid stream from the reactor is withdrawn at a constant flow rate F_{1l}, while the vapour stream enters the condenser at a flow rate F_{1v}. The vapour in the condenser is cooled to T_2 to improve product purity by eliminating reactant A from the liquid.

The reactant A-rich liquid phase in the condenser is returned to the reactor at a flow rate of F_{2l}, while the product vapour phase departs the condenser at a flow rate of F_{2v} and a composition of y_{A2}.

Equations (12)–(29) give a detailed differential algebraic equation (DAE) model used to train the MPPCA model for data generation.

$$\dot{M}_1^l = F_B - F_{1l} + F_{2l} + N_{A1} - N_{C1} \tag{12}$$

$$\dot{x}_{A1} = \left(\frac{1}{M_1^l}\right)[-F_B x_{A1} + F_{2l}(x_{A2} - x_{A1}) + N_{A1}(1 - x_{A1}) + N_{C1}x_{A1} - r_A] \tag{13}$$

$$\dot{x}_{B1} = \left(\frac{1}{M_{1l}}\right)[F_B(1 - x_{B1}) - F_{2l}x_{B1} - N_{A1}x_{B1} + N_{C1}x_{B1} - r_A \tag{14}$$

$$\dot{M}_1^v = F_A - F_{1v} - N_{A1} + N_{C1} \tag{15}$$

$$\dot{y}_{A1} = \left(\frac{1}{M_1^v}\right)[F_A(1 - y_{A1}) - N_{A1}(1 - y_{A1}) - N_{C1}y_{A1}] \tag{16}$$

$$\begin{aligned}\dot{T}_1 = \left(\tfrac{1}{M_1^l + M_1^v}\right)&[F_A(T_A - T_1) + F_B(T_B - T_1) + F_{2l}(T_2 - T_1)\\ &+(N_{A1} - N_{C1})\tfrac{\Delta H^v}{C_P} - \tfrac{Q_1}{C_P} + r_A\left(\tfrac{-\Delta H_r}{C_P}\right)]\end{aligned} \tag{17}$$

$$\dot{M}_2^l = N_{A2} + N_{C2} - F_{2l} \tag{18}$$

$$\dot{x}_{A2} = \left(\frac{1}{M_2^l}\right)[N_{A2}(1 - x_{A2}) - N_{C2}x_{A2}] \tag{19}$$

$$\dot{M}_2^v = F_{1v} - F_{2v} - N_{A2} - N_{C2} \tag{20}$$

$$\dot{y}_{A2} = \left(\frac{1}{M_2^v}\right)[F_{1v}(y_{A1} - y_{A2}) - N_{A2}(1 - y_{A2}) + N_{C2}y_{A2}] \tag{21}$$

$$\dot{T}_2 = \left(\frac{1}{M_2^l + M_2^v}\right)[F_{1v}(T_1 - T_2) + (N_{A2} + N_{C2})\frac{\Delta H^v}{C_p} - \frac{Q_2}{C_P}] \tag{22}$$

$$0 = N_{A1} - k_A a(y_{A1} - y_{A1}^*)\frac{M_1^l}{\rho} \tag{23}$$

$$0 = N_{C1} - k_C a(y_{c1}^* - (1 - y_{A1}))\frac{M_1^l}{\rho} \tag{24}$$

$$0 = N_{A2} - k_A a(y_{A2} - y_{A2}^*)\frac{M_2^l}{\rho} \tag{25}$$

$$0 = N_{C2} - k_C a * (1 - y_{A2} - y_{C2}^*)\frac{M_2^l}{\rho} \tag{26}$$

$$0 = P_1\left(V_{1T} - \frac{M_1^l}{\rho}\right) - M_1^v R T_1 \tag{27}$$

$$0 = P_2\left(V_{2T} - \frac{M_2^l}{\rho}\right) - M_2^v R T_2 \tag{28}$$

$$0 = P_1 - P_2 - \frac{1}{0.09}(F_{1v})^{\frac{7}{4}} \tag{29}$$

The system parameter values are given in Table 1.

Table 1. TPRCR system parameter.

Parameter	Description	Value	Unit
a	Interfacial mass transfer area/unit liquid holdup	1000	m^2/m^3
C_p	Molar heat capacity	80	J/mol K
E_a	Activation energy	110	kJ/mol
K	Proportional gain of pressure controller	−8	mol/s atm
K_{10}	Preexponential factor	2.88×10^{11}	m^3/mol s
k_a	Overall mass transfer coefficient for A	0.2	mol/m^2 s
k_c	Overall mass transfer coefficient for C	0.8	mol/m^2 s
${M_1}^l$	Liquid molar holdup in reactor	14.52	kmol
${M_2}^l$	Liquid molar holdup in condenser	15	kmol
${M_1}^v$	Vapour molar holdup in reactor	3.75	kmol
${M_2}^v$	Vapour molar holdup in condenser	3.90	Kmol
P_1	Pressure in reactor	50	atm
P_2	Pressure in condenser	48.69	atm
P^*	Set point for reactor pressure	50	atm
T_A	Temperature of feed A	315	K
T_B	Temperature of feed B	300	K
T_1	Temperature in reactor	330	K
T_2	Temperature in condenser	304.16	K
V_{1T}	Volume of reactor	3	m^3
V_{2T}	Volume of condenser	3	m^3
ρ	Liquid molar density	15,000	mol/m^3
ΔH_r	Heat of reaction	−50	kJ/mol
ΔH^v	Heat of vaporization	10	kJ/mol

4. Fault Detection Using MPPCA for the TPRCR System

Three types of measurements are used to train the MPPCA model. Fast-rate measurements include temperature, pressure, and flow rates available every second (x_1). Medium-rate measurements include molar holdups available every fifteen seconds (x_2), and slow-rate measurements include mole fractions available every sixty seconds (x_3).

The MPPCA model is trained with 7200 samples of fast-rate measurements, 480 samples of medium-rate measurements, and 120 samples of slow-rate observations. The fault identification capability of the MPPCA approach is assessed using the six categories of faults indicated in Table 2.

Table 2. Fault description in the TRPCR system.

Fault No.	Fault Type	Fault Introduced (s)
1	Step jump in flow rate of A (F_A)	2400
2	Step jump in flow rate of B (F_B)	2400
3	Step jump in temperature of A (T_A)	2400
4	Step jump in temperature of B ($\mathrm{T_B}$)	2400
5	Ramp jump in flow rate of A ($0.0004 \times t$)	2400
6	Ramp jump in temperature of A ($0.003 \times t$)	2400

For a fair comparison, all detection models in this work have a level of significance of 0.99 for the SPE and T^2 statistics. Table 3 shows the false alarm rates for normal data and the missing detection rates for faults, where Fault 0 represents normal test data and that the monitoring results are false alarm rates. The false alarm rate is the fraction of normal data that is interpreted as problem data. Similarly, the missing detection rate is the fraction of the defect data that are treated as normal data. Table 3 shows the monitoring results of all faults using T^2 and different SPE statistics for the MPPCA model.

Table 3. Fault monitoring results using T^2 and *SPE* statistics.

Fault No	T^2	SPE_1	SPE_2	SPE_3
0	0.021	0.001	0.004	0.0001
1	0.035	0.023	0.09	0.008
2	0.067	0.065	0.011	0.034
3	0.001	0.001	0.001	0.001
4	0.854	0.673	0.765	0.231
5	0.313	0.452	0.023	0.045
6	0.201	0.121	0.111	0.201

Three different *SPE* statistics are used to see which fault will have an effect on which SPE statistic. Figure 2 shows *SPE* statistics for the fault in flow rate of A (F_A), which suggests that this fault affects all three *SPE* statistics.

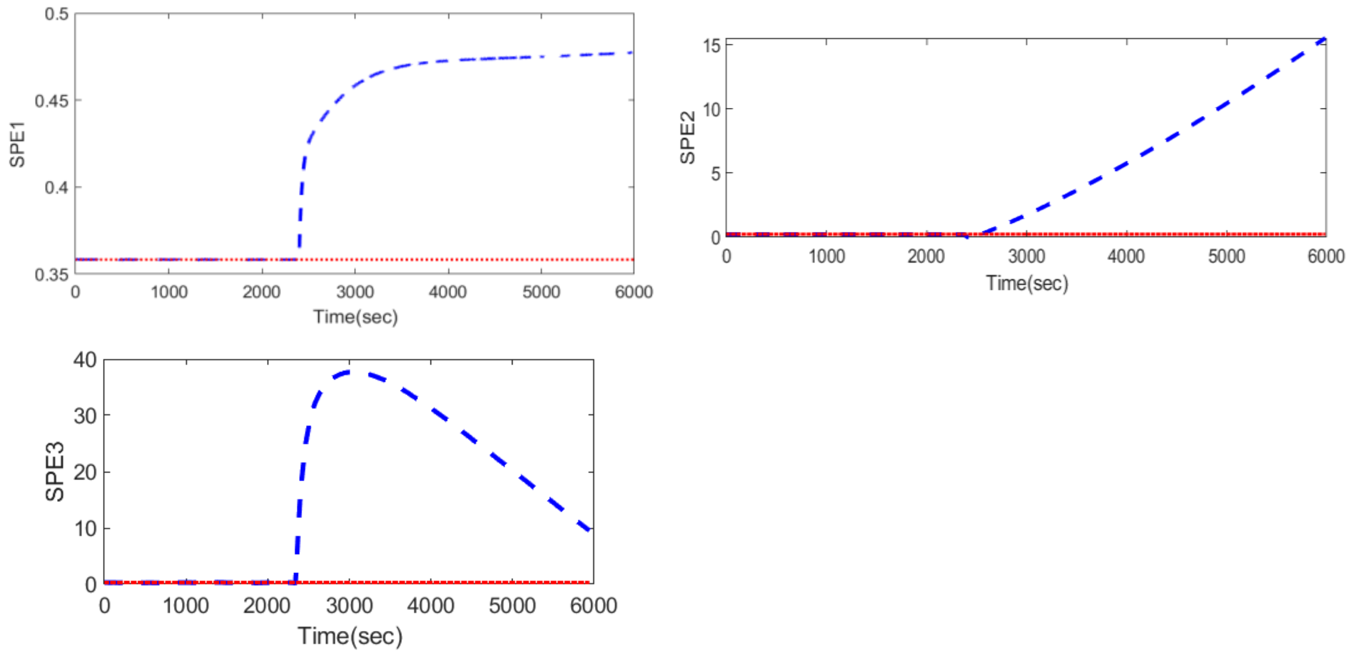

Figure 2. Monitoring results of Fault 1.

5. Conclusions

In this paper, the TPRCR system is modelled as a multi-rate system due to the involvement of qualitative variables, including three different classes of measurements. These measurements are used to develop the MPPCA model using the EM algorithm. This developed MPPCA model is used to detect faults by developing T^2 and three different *SPE* statistics for each measurement class. Six different types of faults are used to check the effectiveness of the developed MPPCA model, and, from the monitoring results, we can clearly say that the MPPCA model can detect faults with a high detection rate.

Author Contributions: Conceptualization, D.G. and M.S.; methodology, D.G. and M.S.; investigation, D.G. and M.S.; writing—original draft preparation, D.G.; writing—review and editing, M.S.; supervision, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Venkatasubramanian, V.; Rengaswamy, R.; Yin, K.; Kavuri, S.N. A review of fault detection and diagnosis. Part III: Process history based methods. *Comput. Chem. Eng.* **2003**, *27*, 327–346. [CrossRef]
2. Srinivasarao, M.; Patwardhan, S.C.; Gudi, R.D. Nonlinear predictive control of irregularly sampled data systems using identified observers. *Lect. Notes Control Inf. Sci.* **2007**, *358*, 141–149. [CrossRef]
3. Zhu, J.; Ge, Z.; Song, Z. Robust semi-supervised mixture probabilistic principal component regression model development and application to soft sensors. *J. Process Control* **2015**, *32*, 25–37. [CrossRef]
4. Zhou, L.; Chen, J.; Jie, J.; Song, Z. Multiple probability principal component analysis for process monitoring with multi-rate measurements. *J. Taiwan Inst. Chem. Eng.* **2019**, *96*, 18–28. [CrossRef]
5. Box, G.E.P. Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification. *Ann. Math. Stat.* **1954**, *25*, 290–302. [CrossRef]
6. Kumar, A.; Daoutidis, P. Feedback Regularization and Control of Nonlinear Differential-Algebraic-Equation Systems. *AIChE J.* **1996**, *42*, 2175–2198. [CrossRef]