Internal Report of the manuscript

Kernel Latent State Space Model for Nonlinear Dynamic Data Modeling and Monitoring

Jiaxin Yu, *Student Member, IEEE*, Yining Dong, and S. Joe Qin, *Fellow, IEEE*

The main purpose of this internal report is to facilitate the peer review process and only for internal reference. The contents of this report are organized as follows:

- Section-A presents the online process monitoring strategy based on the proposed nonlinear dynamical system framework.
- In Section-B, the detailed analysis of the fault detection results in the TEP case is illustrated.
- In Section-C, the TEP fault identification results are analyzed.
- In Section-D, the fault detection results from the multiphase flow facility (MPFF) are visualized and analyzed.
- Section-E describes the consecutive fault identification and process understanding of the MPFF case.

A. Online Process Monitoring Strategy based on KLaSS

To perform online monitoring of the real-time data, we need to calculate the indices defined in Section IV in the manuscript, i.e., $Q_{e,k}$ in the outer layer, $T^2_{v,k}$ with $Q_{v,k}$ in the inner layer, with their respective control limits. Note that the equation numbers are consistent with those in Section IV in the manuscript. Then the online process monitoring procedures based on KLaSS could be given as follows:

1. Scale the real-time samples Each real-time sample is first standardized using the mean and standard deviation of the training set to obtain $\boldsymbol{y}_k$.

2. Transformation, projection, and transition At the outer layer, the standardized sample $\boldsymbol{y}_k$ is transformed into the feature space via the kernel function. Subsequently, the mapped features are projected onto the latent space to estimate the DLV sample, $\boldsymbol{v}_k$ as in (14). This step leverages the kernel trick to enhance computational efficiency. With the DLVs, we resort to (6) to transition the latent states $\boldsymbol{x}_k$ at the inner layer.

3. Monitor dynamic variations at inner layer The DLV prediction residual, $\boldsymbol{e}^v_k$, can be computed as in (28). By conducting PCA on the DLV prediction residual $\boldsymbol{e}^v_k$, we can compute the indices T^2_v in (29) and Q_v in (30) to monitor the dynamic variations at the middle layer.

4. Monitor static variations at outer layer The reconstruction residuals of the measurements, $\boldsymbol{e}^y_k$, can be nominally calculated as in (26) and further monitored by inspecting the Q_e index given in (27). This index allows for the detection of static variations at the outermost layer, which is comparable to the Q_e indices in KPCA, KDPCA, and KCVA, complementing a comprehensive scheme of fault detection.

5. Evaluating monitoring performance The final step involves comparing the calculated indices against their respective control limits. Any testing index exceeding its control limit with the specified consecutive samples indicates a fault or variation. Then fault detection delays can be counted to assess the responsiveness of the monitoring models. Otherwise, the process is under NOC.

As mentioned in the paper, the above integration of residuals from static and dynamic perspectives presents a comprehensive assessment for complex process monitoring. Results from the case studies not only verified the improved detection capabilities of the proposed monitoring strategy, but also deepen the understanding of variations. All of these advantages notably contributed to the system reliability and the afterwards fault identification.

B. TEP Fault Detection Results

For the TEP case, the process monitoring performance of the proposed algorithm and the alternatives is summarized in Table I. The FAR, FDR, and detection delays metrics are crucial for evaluating the models' sensitivity and timeliness in detecting process deviations or adversarial attacks [1]. The results in the IDV(0) case indicate the FAR of these models. It is evident in Table I that the proposed KLaSS model consistently exhibits high FDRs with shorter detection delays in nearly all of the IDV cases, signifying its capability in identifying various fault types.

It can be seen from Table I that the proposed KLaSS model achieved the FDR averages of over 88% for the indices T^2_v in the inner DLVs layers, and the average of 86.2% in Q_e in the outer measurement layer, respectively. However, in the alternative methods, the largest FDR values are the 77.5% of T^2_y of the PCs index in CVA and 72.6% of Q_e of the residual index in DPCA. Take IDV(10) as a specific instance, the FDRs of T^2_y and Q_e are 80.6% and 82.7% in DPCA, 83.0% and 75.9% in CVA, 75.4% and 53.4% in KPCA, 74.8% and 74.8% in DKCPA, 71.4% and 71.7% in KCVA, whereas 88.1%, 84.1%, and 86.6% with T^2_v, Q_v, and Q_e for KLaSS, respectively. The advantages of the detection performance can be attributed to the designed model architecture for outer non-linearity handling and inner dynamics capturing in the proposed KLaSS model.

C. TEP Fault Identification Results

Following the detected variations of the IDVs, identifying the most relevant process variables related to the anomaly is crucial for root cause diagnosis and targeted interventions. Since all the methods are based on the LVs, these models can be used to facilitate the fault identification through FR-based approach linking the latent variables with fault labels. Here, we select several representative fault types including the step variations (IDV(4)-IDV(7)), random variations (IDV(11)), and unknown variations (IDV(17), IDV(19)). Their corresponding fault identification results are listed in Table II, where only the

TABLE I. Process Monitoring Results (%) of DPCA, CVA, KPCA, DKPCA, KCVA, and KLaSS models in the TEP case.

	DPCA			CVA			KPCA			DKPCA			KCVA			KLaSS			
IDV	T_y^2	Q_e	delay	T_y^2	Q_e	delay	T_y^2	Q_e	delay	T_y^2	Q_e	delay	T_y^2	Q_e	delay	T_v^2	Q_v	Q_e	delay
0	0.1	0.6	-	0	2.1	-	1.5	5.5	-	0	6.5	-	0.1	0.6	-	7.0	4.5	4.5	-
1	98.9	99.0	10	99.4	98.0	6	98.5	98.5	14	94.3	90.4	12	98.2	98.5	12	99.3	99.1	99.3	7
2	90.2	85.5	84	94.1	38.3	59	93.0	57.1	51	94.3	78.5	52	89.7	90.6	66	96.3	92.6	96.6	19
3	17.8	25.1	597	51.0	3.6	217	35.1	11.9	221	40.0	25.7	57	17.8	21.5	224	85.0	52.4	80.4	6
4	100	100	0	100	100	0	99.9	86.4	1	88.6	80.2	3	100	100	0	99.9	99.9	99.9	1
5	100	100	0	100	100	0	99.9	24.7	1	99.9	99.9	1	100	100	0	99.9	99.9	99.9	1
6	100	100	0	100	100	0	99.8	94.8	1	99.8	98.5	1	100	100	0	99.8	99.8	99.8	1
7	100	100	0	100	100	0	99.9	99.9	1	99.6	99.9	1	99.9	100	0	99.9	99.9	99.9	1
8	89.2	88.6	87	89.6	81.9	86	88.6	74.9	89	88.1	89.6	83	88.1	88.3	92	91.4	89.1	90.9	82
9	41.1	46.9	110	82.9	5.7	96	39.0	10.6	99	38.8	34.3	100	21.1	25.3	99	87.4	60.6	88.0	90
10	80.6	82.7	140	83.0	75.9	136	75.4	53.4	146	74.8	74.8	146	71.4	71.7	147	88.1	84.1	86.6	133
11	95.5	97.6	17	97.6	79.0	20	89.0	72.5	21	68.3	66.2	25	81.5	87.1	20	97.9	95.4	97.8	17
12	98.0	98.3	14	98.2	97.7	15	97.9	61.2	17	97.5	97.0	18	98.0	98.1	15	98.0	98.1	98.1	16
13	81.7	82.2	150	84.3	53.4	132	82.7	64.8	142	82.2	69.2	143	82.3	82.5	141	88.4	83.9	87.9	133
14	99.8	100	0	100	42.7	0	90.6	66.9	4	37.3	17.7	405	62.2	99.0	2	99.9	98.1	99.9	1
15	1.1	3.3	-	8.8	2.3	-	6.0	6.5	-	6.1	10.1	164	0.9	2.1	-	59.6	32.8	46.6	160
16	3.1	8.7	788	18.0	2.0	75	5.5	7.4	786	6.2	16.6	334	2.1	3.1	-	63.8	34.3	53.7	70
17	78.6	78.7	173	79.9	57.2	169	73.8	29.0	196	75.7	40.1	145	69.1	71.0	195	87.3	82.9	84.5	170
18	49.0	49.1	420	56.2	40.2	419	56.9	39.2	167	56.8	54.6	145	47.7	50.3	186	80.4	67.8	75.9	163
19	97.4	97.0	21	97.5	96.2	22	97.1	90.9	23	96.9	96.3	24	96.9	97.4	21	98.9	97.5	97.9	22
20	77.9	78.3	173	79.3	77.1	173	72.4	43.1	167	76.4	74.7	145	61.1	72.6	166	87.6	81.5	84.4	169
21	1.6	4.1	-	7.0	2.6	-	3.1	7.5	-	2.6	10.0	215	0.8	1.1	-	55.9	29.7	42.1	203
Avg	71.5	72.6	147	77.5	59.7	86	71.6	52.4	113	67.8	63.1	106	66.1	69.5	77	88.8	80.0	86.2	70

TABLE II. Fault identification results with FR (%) using DPCA, CVA, KPCA, DKPCA, KCVA, and KLaSS models for selected IDVs in the TEP case.

	DPCA		CVA		KPCA	
IDV	Var1: FR (%)	Var2: FR(%)	Var1: FR (%)	Var2: FR(%)	Var1: FR (%)	Var2: FR(%)
4	XMV10: 66.9	XMV3: 9.9	XMV3: 65.0	XMV6: 31.0	XMV10: 81.1	XMV2: 2.2
5	XMEAS22: 93.5	XMEAS18: 4.6	XMV6: 67.6	XMEAS22: 27.5	XMEAS22: 93.8	XMEAS18: 4.4
6	XMV3: 90.2	XMEAS1: 3.3	XMV3: 99.8	XMV6: 0.01	XMV3: 85.4	XMEAS16: 4.6
7	XMV4: 99.5	XMV3: 0.2	XMV3: 58.7	XMV4: 31.6	XMV4: 99.6	XMEAS7: 0.1
11	XMV10: 61.8	XMV1: 7.6	XMV6: 58.2	XMV3: 36.8	XMV10: 69.0	XMEAS21: 14.5
17	XMEAS21: 92.9	XMV10: 6.1	XMV3: 82.6	XMV6: 12.3	XMEAS21: 81.7	XMV10: 12.6
19	XMEAS22: 39.6	XMV7: 32.4	XMV6: 97.9	XMEAS22: 1.6	XMEAS22: 42.8	XMV7: 33.1

	DKPCA		KCVA		KLaSS	
IDV	Var1: FR (%)	Var2: FR(%)	Var1: FR (%)	Var2: FR(%)	Var1: FR (%)	Var2: FR(%)
4	XMV10: 62.9	XMV1: 11.3	XMV6: 44.6	XMV10: 44.1	XMV10: 85.9	XMEAS21: 2.0
5	XMEAS22: 89.4	XMEAS18: 6.4	XMEAS22: 93.5	XMEAS18: 3.5	XMEAS22: 99.2	XMEAS18: 0.7
6	XMV3: 40.9	XMEAS16: 30.1	XMV3: 65.6	XMEAS11: 9.5	XMV3: 72.7	XMEAS7: 11.3
7	XMV4: 99.2	XMV1: 0.3	XMV4: 96.4	XMV6: 1.6	XMV4: 99.7	XMEAS11: 0.05
11	XMV10: 56.6	XMV1: 8.5	XMV10: 35.8	XMV6: 29.1	XMV10: 35.3	XMEAS9: 18.7
17	XMEAS21: 52.8	XMV10: 25.8	XMEAS21: 82.3	XMV10: 10.8	XMEAS21: 94.9	XMV1: 1.6
19	XMEAS22: 37.8	XMV7: 32.0	XMEAS22: 36.9	XMV8: 24.0	XMEAS18: 50.1	XMEAS22: 10.5

measured variables with the largest two FRs are displayed due to the page limitation.

It can be proven that almost all methods have identified the correct fault variables to some extent, which verifies the effectiveness of the fault identification method based on FR. Take IDV(5) for example, the fault type is the step deviation in cooling water inlet temperature of the condenser unit [2], and all the methods identified XMEAS22 (condenser cooling water outlet temperature) as the abnormal variable. And for IDV(6), a step variation of loss in A feed in stream 1 occurred, and all the methods successfully identified XMV3 (valve position feed component A (stream 1)). Besides, a step anomaly in C header pressure loss (stream 4) was introduced in IDV(7), then the XMV4 (valve position feed component A & C in stream 4) was identified by the proposed model and the alternatives.

However, we also see that in the comparison methods, erroneous variables in the fault scenario could be identified, as in IDV(4), IDV(11), IDV(17), and IDV(19). Through analysis, in IDV(4), the step deviation happened in cooling water inlet temperature of reactor, but the CVA model erroneously found XMV3 (valve position feed component A) and XMV6 (valve position purge); the KCVA model obtained vague FR results with 44.6% for XMV6 and 44.1% for XMV10. The other methods identified the correct XMV10 (valve position cooling water outlet of reactor), which can be evidenced by the FR results and the measured variables in Fig. 1. This fault identification result is not surprising since the CVA and KCVA models adopt larger lags of measured variables than the other methods for dynamic modeling, which suffer from the mingled effects while conducting reconstruction procedure.

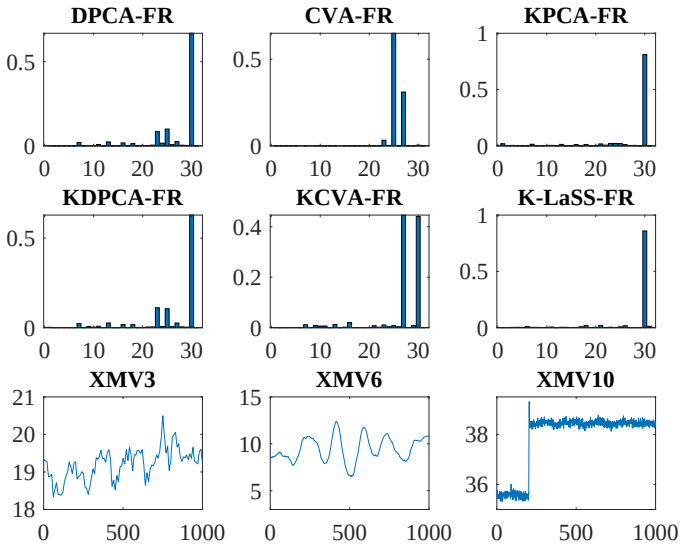

Fig. 1. The fault identification results and process variables of IDV(4) in the TEP case, where Var25 (XMV3), Var27 (XMV6), and Var30 (XMV10) are identified by these methods.

Similar results can be found in IDV11 (random fault in cooling water inlet temperature of reactor): the CVA model failed to identify the correct XMV10. In addition, deviations of heat transfer within reactor happened in IDV(17), all the method identified XMEAS21 (reactor cooling water outlet temperature), except the CVA model.

D. MPFF Fault Detection Results

As mentioned, the MPFF is nonlinear dynamic process with pressurized system managing the flows of air, water, and oil [3]. The fault description information of the test datasets is summarized in Table III. In addition, the flow diagram of the MPFF process is depicted in Fig. 2.

TABLE III. Fault description of the testsets in MPFF process.

Dataset	Fault type	Fault starts	Fault ends
F-1.2	air line blockage	657	3777
F-3.2	top separator input blockage	333	5871
F-4.2	open direct bypass	851	3851

The fault detection indices of F-3.2 and F-4.2 are plotted in Fig. 3 and Fig. 4, respectively. From these indices, we observe that the false alarms of KLaSS model are less frequent than the alternative methods. In the traditional methods, there are always one index that exceeds the control limits to great extent, although we have tested distinct hyper-parameter settings. As mentioned in the paper, the proposed KLaSS model consistently show much more stable FAR, better FDR, and quicker detection across all tested scenarios, which is attributed to the tiered dynamical system design. However, the traditional methods DPCA and CVA may have failed to deal with the highly nonlinear, non-stationary data from the MPFF; the KPCA model didn't consider the dynamics contained in the data; the KDPCA and KCVA methods suffered from the mentioned *mingled* effects of the extracted LVs.

E. MPFF Fault Identification Results

After the faults are detected, we continue to utilize the detection results to facilitate the proposed fault identification as mentioned. The identification results are listed in Table IV. In F-1.2, the manual valve of the air line was gradually closing, simulating the fault of *air line blockage*. But the valve angle was not included in the measurement list for monitoring [3]. Hence, the most related abnormal measurement variable in this scenario will be PT312, the air delivery pressure in Fig. 2. According to Table IV, most of the models identified PT-312 in this dataset especially KLaSS with the largest FR 95.8%, However, CVA and DKPCA failed to find the correct variables.

As for F-3.2, the fault mode is *top separator input blockage*. It can be found in Table IV that most of the models successfully identified critical variables that correlate directly with this fault type: PT401 (Pressure at the bottom of the riser), the pressure would be directly affected by any blockage in the flow path leading to the top separator; and also PT312 (Air delivery pressure), this variable again indicates disruptions in the air supply path to the separator. As mentioned, these two measurements corresponds to the two input paths of the top separator, which complies with the flow diagram as in 2.

As mentioned, the fault F-4.2: *open direct bypass* means a leakage at the bottom of the riser is, impacting the flow through the primary processing path. As depicted in Fig. 2, the bypass provides an alternative route for the multi-phase mixture, bypassing the riser section and leading directly to the

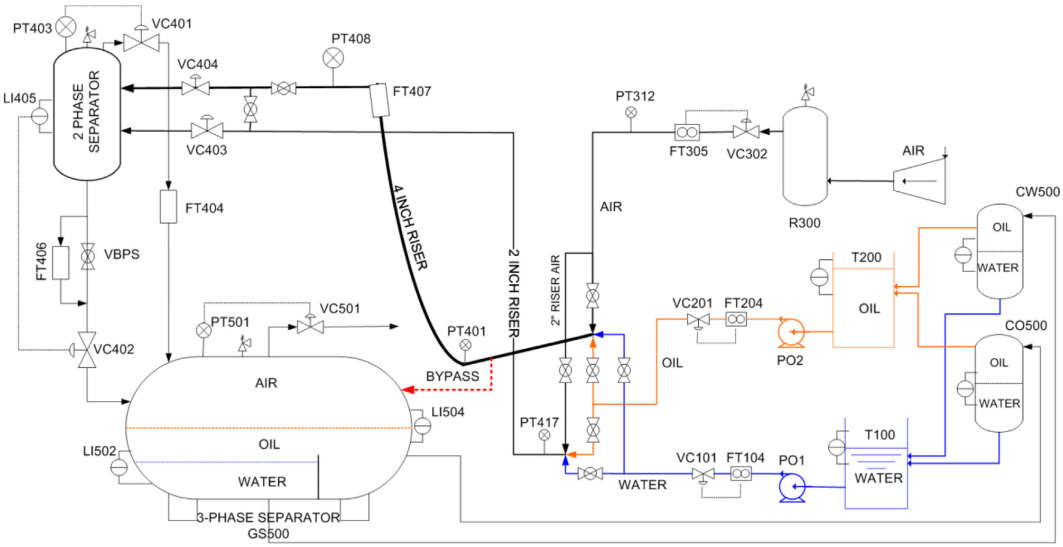

Fig. 2. The flow diagram of the MPFF process.

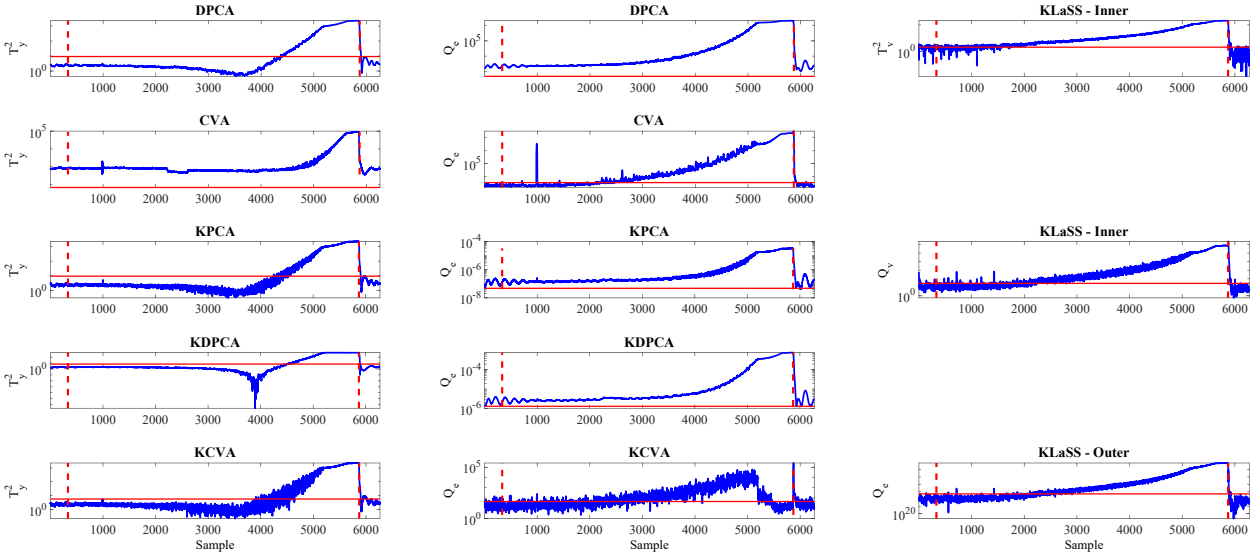

Fig. 3. The monitoring results of Fault3.Set2 in the MPFF process.

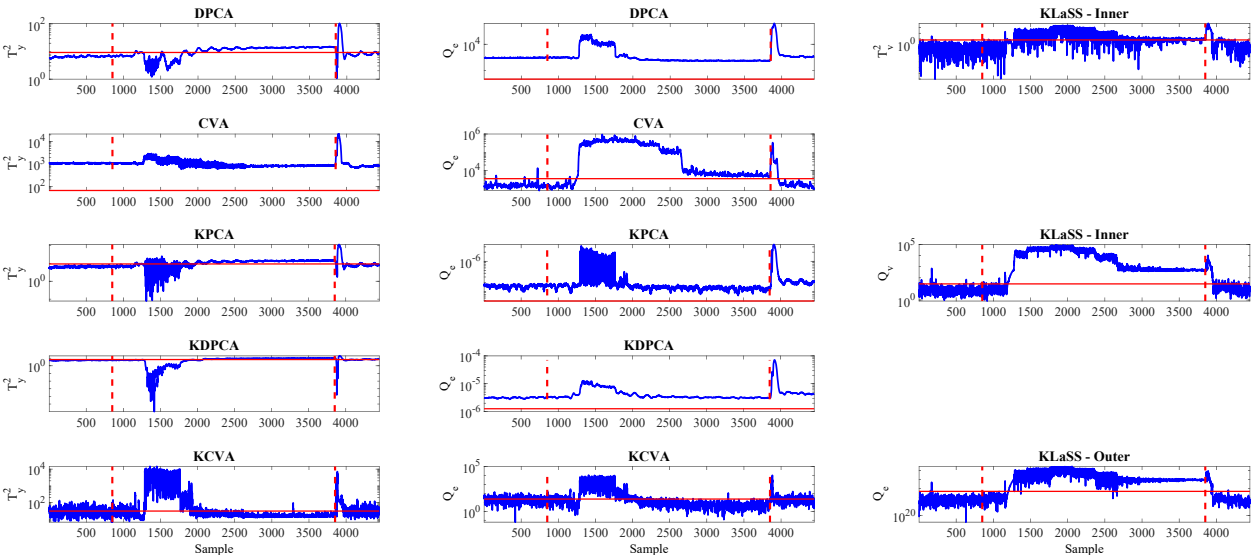

Fig. 4. The monitoring results of Fault4.Set2 in the MPFF process.

TABLE IV. Fault identification results (FR in %) in the MPFF process.

	DPCA		CVA		KPCA	
F-	Var1-FR	Var2-FR	Var1-FR	Var2-FR	Var1-FR	Var2-FR
1.2	PT312-66.9	PT408-6.5	FT406-85.5	VC501-11.5	PT312-67.0	FT407-9.1
3.2	PT401-47.5	PT312-44.1	FT406-96.2	PT401-1.5	PT403-78.0	PT408-20.3
4.2	FT407-21.4	PT408-15.6	FT406-98.4	VC501-1.2	FT104-18.5	FT406-15.6
	DKPCA		KCVA		KLaSS	
F-	Var1-FR	Var2-FR	Var1-FR	Var2-FR	Var1-FR	Var2-FR
1.2	PT408-41.0	PT312-27.4	PT312-27.7	VC501-16.1	PT312-95.8	PT401-2.2
3.2	PT401-57.2	PT312-33.0	PT401-47.6	PT312-47.1	PT401-62.0	PT312-20.9
4.2	PT408-26.8	FT407-20.1	PT403-31.4	PT408-23.0	PT401-35.1	PT312-32.5

3-phase separator. The fault identification results of the KLaSS model are particularly effective for this fault scenario, since it pinpoints pressure variables that are logically connected to the introduced fault. The identified variables include: PT401 with FR 35.1%, PT312 with 32.5% in FR and PT408 (Differential pressure [PT401-PT408]) with FR 31.7% (this variable is not listed in Table IV due to space constraint). Changes in the differential pressure between PT401 and PT408 can effectively capture the resultant pressure disparities introduced by the bypass. The sensitivity of PT408 to changes in pressure dynamics due to the bypass makes it a key indicator of the fault in this scenario. However, the other methods failed to identify all these three most relevant measurements.

By systematically evaluating the fault detection metrics and the fault identification accuracy with process flow diagrams, these case studies demonstrated the superiority of process monitoring performance of the proposed KLaSS algorithm in complex industries. The results could provide valuable algorithm insights and process interpretability for real-world applications and the potential to enhance fault diagnosis in industrial electronic systems.

REFERENCES

[1] V. Pozdnyakov, A. Kovalenko, I. Makarov, M. Drobyshevskiy, and K. Lukyanov, "Adversarial attacks and defenses in fault detection and diagnosis: A comprehensive benchmark on the tennessee eastman process," *IEEE Open Journal of the Industrial Electronics Society*, 2024.

[2] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Computers & Chemical Engineering*, vol. 17, no. 3, pp. 245–255, 1993.

[3] C. Ruiz-Cárcel, Y. Cao, D. Mba, L. Lao, and R. Samuel, "Statistical process monitoring of a multiphase flow facility," *Control Engineering Practice*, vol. 42, pp. 74–88, 2015.