Unlocked decision making based on causal connections strength

Mohamed Amine Atoui, Achraf Cohen, Vincent Cocquempot

Version of Record: https://www.sciencedirect.com/science/article/pii/S094735802100073X
Manuscript_c1eba4c95e73d99cbb66aac2eba6cb63

Unlocked decision making based on causal connections strength

M. Amine Atoui[1,3], Achraf Cohen[2] and Vincent Cocquempot[3]

***Abstract*— Fault detection and diagnosis are crucial to reducing risks and costs in any process. The identification of the propagation path and the variables responsible for faulty operating conditions is also vital. This paper presents a causal network-based approach to detect, diagnose, and identify root causes in multivariate processes. We discuss aspects such as complexity and rules related to modeling such network approaches. The proposed strategy is established on statistical justifications. The introduced decision rules deal with unknown faults and offer new perspectives to data-driven methods for fault diagnosis. The proposed approach is evaluated and demonstrated using the well-known Tennessee Eastman Process (TEP) benchmark.**

I. Introduction

Data-driven methods for process control are crucial to monitoring complex systems because the model-based and knowledge-based approaches may fail to produce an acceptable monitoring model for such systems. Process monitoring includes fault detection and isolation/identification procedures that aim at determining whether or not a fault has occurred and which variables are responsible, respectively. The statistical process monitoring notion relies on the statistical theory that data variations (e.g. mean and variance) are similar unless a fault happens in the process. A fault is any deviation from the normal operating condition of the process. Faults are often modeled as deviations from the in-control process mean or variance, or both.

The large amount of data collected from processes are often multivariate of nature, say m-variate, where m variables are monitored to decide on the process's state. Therefore, multivariate statistical methods have been extensively utilized to reduce the dimension or extract relevant information and latent knowledge [1], [2], [3].

Some of the techniques that have been widely studied for process monitoring include principal component analysis (PCA) [1], [4], [5], partial least squares (PLS) [6], Bayesian networks (BN) [7], [8], subspace methods [9], wavelets analysis [10]. Fault detection and diagnosis is a classification problem. Therefore several machine learning methods [11] have been applied to solve it, such as support vector machines [12], neural networks [13], decision trees [14]. These methods are powerful but require a considerable amount of reliable data. Probabilistic graphical models are an important tool for modeling decision support systems in uncertain environments [15], [16]. In the last decades, probabilistic graphical models have received great attention in the statistical process monitoring field [17], [18], [19], [20], [21], [22], [23], [24], [25], [23].

Fault detection and diagnosis are important tasks to help the operator to decide about the actual operating conditions. A challenging task is the root cause identification, where the goal is to identify which variables are responsible for the observed fault. Various approaches are used to deal with these tasks. For instance, principal component analysis (PCA) and kernel principal component analysis (KPCA) techniques were used to detect faults and Bayesian Networks to identify root causes in [26]. The authors used transfer entropy and Granger causality to identify the structure of the network. Another fault diagnosis method was proposed by combining dynamic Bayesian anomaly index-based control chart and dynamic Bayesian models [27]. In [28], the authors proposed nonparametric probabilistic indices to detect and identify the root cause variables. [29] proposed another index, the G-index, estimated by adopting a nonparametric estimator.

One particular method of interest uses the statistical tests based on quadratic statistics to detect faults in a multivariate process. When a fault is detected, the next step is to look for the variables explaining faulty data. A couple of methods have been proposed for identification including the Mason, Young, and Tracy (MYT) approach [30], multivariate contribution plots [31], reconstruction analysis [32], [33], [34], and non-negative garrote [35]. Causality analysis techniques were also proposed to identify the variables such as Granger causality test [36], transfer entropy [37], and causal networks [38], [39]. The MYT decomposition holds solid statistical properties. However, it requires a large number of decompositions, which makes it less attractive in practice.

In this work, a new causal monitoring system is proposed. It integrates fault detection and diagnosis as well as an identification stage where the potential variables responsible for the faults and their propagation path are determined. The rest of the paper is organized as follows. Section 2 provides a theoretical basis and definitions. Section 3 presents the decision rules and introduces the proposed method. In Section 4, the performance of the proposed causal network is evaluated on the Tennessee Eastman Process (TEP). Finally, the conclusions are summarized in Section 5.

[1] IEMN, CNRS, University of Lille, amine.atoui@gmail.com
[2] Mathematics and Statistics Department, University of West Florida, acohen@uwf.edu
[3] UMR 9189 - CRIStAL, CNRS, University of Lille, vincent.cocquempot@univ-lille.fr

II. Probabilistic graphical models

A probabilistic graphical model [16] is a Directed Acyclic Graph (DAG) G with a set of random variables $\mathbf{X}_1, \ldots, \mathbf{X}_m$ assigned to its nodes. The nodes are connected by directed arcs. The arcs describe the dependencies between variables. These dependencies are governed by the Markov conditions. Thus, DAG conditional independence relationships allow, applying the chain rule, the decomposition of the joint distribution of $\mathbf{X}$ as a set of independent conditional probabilities such that:

$$p(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_m) = \prod_{i=1}^{m} p(\mathbf{X}_i | \rho(\mathbf{X}_i)) \quad (1)$$

where $p(\mathbf{X}_i|\rho(\mathbf{X}_i))$ is the conditional probability of the $\mathbf{X}_i$ given its parents $\rho(\mathbf{X}_i)$ in G. A parent A in a directed graph to B can be informally interpreted as indicating that A "causes" B. Basically, these dependencies between variables represent a causal model. The structure of the model can be built by experts or learned from data. We are interested in the DAG's adjacency matrix encoding the causal relationships between variables.

One particular form of the probabilistic graphical model is the Conditional Continuous Network (CCN). The nodes of the CCN can represent discrete or continuous (univariate or multivariate) random variables. Continuous nodes $\mathbf{X}_i$ given their continuous parents follow regression models with parameters depending on the values k of their discrete parents, $\mathbf{D}$. Each model is characterized by the response variable mean and variance, μ_k and Σ_k, respectively.

A. Statistical monitoring in Causal Networks

We consider $K+1$ classes, C_k, which represent the system K faults plus the normal operation conditions (NoC). In practice, it is almost impossible to enumerate, simulate, or gather data about all the system possible operating conditions. Indeed, a current challenge with fault diagnosis is how to deal with unknown or new considered operating conditions [19], [40].

Often the statistical decision about whether or not an observation belongs to a class of faults is made by choosing the class with maximum posterior probability. Although this rule is efficient, it does not handle false alarms and unknown operating conditions. It also does not allow explicitly isolate the variables at fault. Based on the posterior probabilities of each state of C_k, of $\mathbf{D}$, in respect to the new observations $X_1, X_2, \ldots, X_m$, $p(C_k|X_1, X_2, \ldots, X_m)$, new rules are proposed to address these shortcomings about decision making under causal network. Based on the properties of the causal network G, from equation (1), $p(C_k|X_1, X_2, \ldots, X_m)$ can be written as a product of Gaussians like follows

$$\begin{aligned} p(C_k|X_1, X_2, \ldots, X_m) &\propto p(X_1, X_2, \ldots, X_m|C_k)p(C_k) \\ &\propto \prod_{i=1}^{m} p(X_i|\rho(\mathbf{X}_i), C_k)p(C_k) \end{aligned} \quad (2)$$

with $p(X_i|\rho(\mathbf{X}_i), C_k) \equiv \mathcal{N}(\mu_{k_i} + W_{k_i}^{\rho_i}\mathbf{X}_{\rho_i}, \sigma^k_{X_i \cdot X_{\rho_i}})$, where $\sigma^k_{X_i \cdot X_{\rho_i}}$ is the variance of variable $\mathbf{X}_i$ conditional on its parents. μ_{k_i} is a constant associated with $\mathbf{X}_i$. Let's denote a connection strength from a variable $\mathbf{X}_i$ to other variables, its direct parents, $\mathbf{X}_{\rho_i}$, in the DAG by $W_{k_i}^{\rho_i}$, then the model can be represented by $\mathbf{X}_i = \mu_{k_i} + W_{k_i}^{\rho_i}\mathbf{X}_{\rho_i} + \mathbf{e}_i$, where $\mathbf{e}_i \sim \mathcal{N}(0, \sigma^k_{X_i \cdot X_{\rho_i}})$. By developing (2) we obtain

$$\frac{p(C_k|X_1, X_2, \ldots, X_m)}{p(C_k)} \propto (2\pi)^{-\frac{m}{2}} \frac{\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{m} \Delta^k_{X_i \cdot X_{\rho_i}}}}{\prod_{i=1}^{m} \sigma^k_{X_i \cdot X_{\rho_i}}} \quad (3)$$

$\Delta^k_{X_i \cdot X_{\rho_i}}$ is noted the Hotelling distance associated to $\mathbf{X}_i$ in respect to $\mathbf{X}_{\rho_i}$ and the value k of $\mathbf{D}$.

B. Bayesian causal-based decomposition

Consider Δ^k the Hotelling quadratic statistic associated to a multivariate variable $\mathbf{X}, \mathbf{X} = [\mathbf{X}_1, \ldots, \mathbf{X}_m]^T$. Δ^k admits $m!$ possible decompositions [41], [42], as a sum of m independent orthogonal quadratic statistics, and two types of decomposition can be distinguished. For a given decomposition of the Δ^k, if it exists a term $\Delta_{i \cdot 1,\ldots,i-1}$ such that the set of variables $\mathbf{X}_1, ..., \mathbf{X}_{i-1}$ includes at least one descendant of $\mathbf{X}_i$, then this decomposition is a Type A. The remaining ones are of Type B. These latter retained as the more accurate as they converge to the same decomposition of Δ^k. It is then enough to focus only on the Type B decomposition (alternative to MYT), which involves the regression of the variable on its causes [41]. This can be deduced from the relations of dependencies between the networks' variables and given by the decomposition $\Delta^k = \sum_{i=1}^{m} \Delta^k_{X_i \cdot X_{\rho_i}}$, which can be deduced from equation (3). Therefore, without loss of generality, we can write it as follows:

$$p(C_k|X_1, X_2, \ldots, X_m) \propto (2\pi)^{-\frac{m}{2}} p(C_k) \frac{\mathrm{e}^{-\frac{1}{2}\Delta^k}}{\prod_{i=1}^{m} \sigma^k_{X_i \cdot X_{\rho_i}}} \quad (4)$$

where Δ^k is a function of μ_k and Σ_k, the mean and covariance of the joint variable $\mathbf{X}$ in respect to the values of $\mathbf{D}$.

III. Decision rules

Statistical rules for fault detection and diagnosis, as well as the identification of the variables responsible for faulty operating conditions are developed. We propose a complete data-driven monitoring approach taking advantage of a unified framework offered by Bayesian models and dealing with known and unknown faults. To our knowledge, this is the first attempt to meet such a challenge.

In the following, the different rules to cope with causal network' decisions for fault detection, diagnosis, and root cause identification are presented.

A. Fault detection

Early detection is necessary to avoid an undesirable situation. Accurate detection is crucial also. Indeed, as mentioned previously, the majority of BN-based fault detection methods makes decisions without respecting the false alarm rate. Here, we propose a detection rule in the context of causal networks that respect a false alarm rate.

Consider the quadratic term Δ^{NoC}, which can be deduced from the distribution of $\prod_{i=1}^{m} p(\mathbf{X}_i|\rho(\mathbf{X}_i), C_k)$. Its associated distribution is given by

$$CL_{\Delta^{NoC}} = \frac{(N+1)(N-1)}{N(N-m-1)} F_\alpha(1, N-m-1) \tag{5}$$

which is considered as its control limit for a given significance level α. A system is considered in normal operating conditions if Δ^{NoC} is less than or equal to its $CL_{\Delta^{NoC}}$. By developing the inequality, in respect of $X_1, \ldots, X_m$, as below

$$\Delta^{NoC} \leq CL_{\Delta^{NoC}} \tag{6}$$

$$\sum_{i=1}^{m} \Delta^{NoC}_{X_i \cdot X_{\rho_i}} \leq CL_{\Delta^{NoC}}$$

$$\frac{1}{(2\pi)^{\frac{m}{2}}|\Sigma^{NoC}|} \mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{m} \Delta^{NoC}_{X_i \cdot X_{\rho_i}}} \geq \frac{\mathrm{e}^{-\frac{1}{2}CL_{\Delta^{NoC}}}}{(2\pi)^{\frac{m}{2}}|\Sigma^{NoC}|}$$

From (3) we can write

$$p(X_1, \ldots, X_m|NoC) \geq p(X^*|NoC) \tag{7}$$

The probabilistic rule to decide about the presence of faulty operating conditions is then deduced and given by, $X_1, \ldots, X_m \in NoC$

$$\text{if } p(NoC|X_1, \ldots, X_m) \geq PL\Delta^{NoC}_{X_1,\ldots,X_m} \tag{8}$$

Where the probabilistic limit $PL_{\Delta^{NoC}_{X_1,\ldots,X_m}}$, in respect to (7), can be deduced as follows

$$PL_{\Delta^{NoC}_{X_1,\ldots,X_m}} = \tau \cdot \frac{1}{1 + \sum_{j=1}^{K-1} \mathbf{e}^{-\frac{1}{2}(\varphi_j - \beta_j)}} \tag{9}$$

with

$$\varphi_j = \Delta^{C_j/\{NoC\}} - \Delta^{NOC}, \tau = \frac{1}{\mathbf{e}^{-\frac{1}{2}(\Delta^{NOC} - CL_{\Delta^{NOC}})}},$$

$$\beta_j = 2\ln(\omega_j \frac{\prod_{i=1}^{m} \sigma^{NoC}_{X_i \cdot X_{\rho_i}}}{\prod_{i=1}^{m} \sigma^{C_j/\{NoC\}}_{X_i \cdot X_{\rho_i}}}), \omega_j = \frac{p(C_j/\{NoC\})}{p(NOC)}$$

B. Fault diagnosis

Data-driven diagnosis is a classification problem. Each operating condition is represented as a class (learned from data or determined by an expert). In practice, it is not evident to describe efficiently the process operating conditions. Also, it is not always possible to identify the exact number of faults that could influence or change the process from its no-fault operating conditions. Then, one can envision an unlocked supervised classification, where a new class representing Unknown operation Conditions (UoC) is added. This class covers unknown faults or unknown normal conditions. This class would hold new observations that are not classified into any of the known classes. We introduce the following rule to take account of UoC after fault detection:

If the Fault k with the highest posterior probability is smaller than its corresponding probabilistic limit (which can be obtained similarly to equation 9) therefore the observation belongs to UoC class, otherwise it belongs to Fault k.

It is possible to deal with multiple simultaneous faults when information about the robustness or sensitivity of $X_1, \ldots, X_m$ for each fault is available. Such information is not naturally or directly available in the data-driven context. However, in the model-based context, faults are isolated through their signature regarding the behaviors of the generated residuals. These residuals are defined in a way they are decoupled, which means that two faults can not have the same signature [43]. We can extend the detection rule to deal with multiple simultaneous faults when such information is available.

C. Root cause identification

A suitable alternative to the MYT approach and their updates is proposed next. The proposed rule is a root-cause variable identification made entirely under a causal network and without any pre-treatment. To identify the root cause variable responsible for a deviation from NoC, the posterior probability of the latest is compared to different probabilistic limits made for each node of G. So, for any variable $\mathbf{X}_i$, if $p(NoC|X_1, X_2, \ldots, X_m) \geq PL_{\Delta^{NoC}_{X_i \cdot X_{\rho_i}}}$ then $\mathbf{X}_i \notin NoC$, where $PL_{\Delta^{NoC}_{X_i \cdot X_{\rho_i}}}$ is the probabilistic limit associated to NoC to monitor $X_i \cdot X_{\rho_i}$ and can be obtained following the development given below.

Consider Γ the set of all combinations $(X_i \cdot X_{\rho_i})$. Let's note γ one of these combinations, $\gamma \in \Gamma$, and $\Gamma/\{\gamma\}$ the remaining ones, $\phi \in \Gamma/\{\gamma\}$. From the equivalence between 3 and 4 we can write

$$\Delta^{NoC} = \sum_{i=1}^{m} \Delta^{NoC}_{X_i \cdot X_{\rho_i}} = \Delta^{NoC}_{\gamma} \sum^{\Gamma/\{\gamma\}} \Delta^{NoC}_{\phi} \tag{10}$$

Based on (10) we expand and develop the rule (6) associated to the statistic Δ^{NoC}_{γ}, as below

$$\Delta_\gamma \leq CL_{\Delta^{NoC}_{\gamma}}$$

$$\mathrm{e}^{-\frac{1}{2}\Delta^{NoC}_{\gamma}} e^{-\frac{1}{2}\sum^{\Gamma/\{\gamma\}} \Delta^{NoC}_{\phi}} \geq \mathrm{e}^{-\frac{1}{2}CL_{\Delta^{NoC}_{\gamma}}} e^{-\frac{1}{2}\sum^{\Gamma/\{\gamma\}} \Delta^{NoC}_{\phi}}$$

$$\mathrm{e}^{-\frac{1}{2}\Delta^{NoC}} \geq \mathrm{e}^{-\frac{1}{2}CL_{\Delta^{NoC}_{\gamma}}} e^{-\frac{1}{2}\sum^{\Gamma/\{\gamma\}} \Delta^{NoC}_{\phi}}$$

$$p(X|NOC) \geq p(X^*_\gamma|NoC) p(X_{\Gamma/\{\gamma\}}|NoC) \tag{11}$$

From (11), $PL\Delta^{NoC}_{X_i \cdot X_{\rho_i}}$ is deduced and given by

$$PL_{\Delta^{NoC}_{X_i \cdot X_{\rho_i}}} = \frac{\mathbf{e}^{\frac{1}{2}(\Delta^{NOC} - CL_{\Delta^{NoC}_{\gamma}} - \sum^{\Gamma/\{\gamma\}} \Delta_\phi)}}{(1 + \sum_{j=1}^{K-1} \mathbf{e}^{-\frac{1}{2}(\varphi_j - \beta_j)})} \tag{12}$$

D. Summary

We propose several probabilistic limits that guarantee statistical boundaries in respect to a significance level α. The step-by-step procedure of the proposed method is summarized below:

- (1) Collect process historical data as a training set that includes all different operating conditions;
- (2) Learn the causal relationships between variables (experts or data, or both);
- (3) Estimate the parameters and prior probability of each operating condition;
- (4) For each monitored observation of $X_1, \ldots, X_m$, at a given instant, compute its posterior probability belonging to every operating condition class;
- (6) Further, deduce the probabilistic limit $PL_{\Delta^{C_k}_{X_1,\ldots,X_m}}$, for each k and i;
- (7) If the probabilistic index value $p(NOC|X_1, \ldots, X_m)$ is less than $PL_{\Delta^{NoC}_{X_1,\ldots,X_m}}$, the corresponding observation is labelled normal operating condition. Otherwise, the sample is potentially faulty and go to next step for further fault diagnosis;
- (8) Discriminate between the faulty known operating conditions and decide to which operating conditions the new observation belongs to (Faulty or UoC);
- (9) similarly to (7), compare $p(NOC|X_1, \ldots, X_m)$ to the probabilistic limit $PL^{NOC}_{\Delta_{X_i \cdot X_{\rho_i}}}$ associated to each node $\mathbf{X}_i$.
- (10) The variables whose probabilistic index value is significantly high are considered responsible for process abnormality.
- (11) investigate and update training data-sets for further use.

IV. Application

TEP is a complex process and consists of 52 continuous variables. The full process flow sheet is given in Figure 1. It consists of five main units: a reactor, a condenser, a compressor, a separator, and a stripper. The entire process consists of four gaseous reactants A, C, D, E, and a small amount of inert B are sent into the reactor to form liquid products F, G, and H. The process includes 12 manipulated variables and 41 measured variables, among which the measured variables include 22 continuous process variables and 19 composition variables. TEP's data are issued from normal and faulty operating conditions.

Tennessee Eastman Process (TEP) is a well-known chemical process simulation [44]. It has been widely used to verify the effectiveness of fault monitoring methods.

Four TEP's faulty data-sets are considered in this study: Faults 1, 4, 10, and 14. This set of faults includes step changes, random variations, and sticking. Also, they represent both stationary and non-stationary faulty processes. Fault 1 is an uncontrollable fault when it occurs

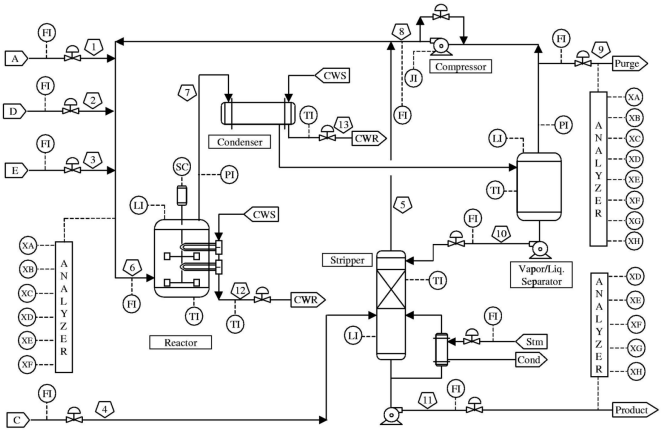

Fig. 1. Tennessee Eastman Process

activates a non-stationary process. It consists of a step-change in the A/C feed ratio in Stream 4. This causes an increase in the C feed and a decrease in the A feed. This fault affects XMEAS (4) and XMEAS (44) and gets propagated to other variables. Faults 4 and 14 influence the reactor. These operating conditions are hard to identify since they overlap. They have in common the root variables XMEAS(9) and XMV(51), which measure the reactor temperature. However, they are different in terms of type and nature. Step-changes for Fault 4 and Fault 14 imply that the valve of reactor cooling water is sticking, which affects the process dynamics, fluctuations in the reactor temperature, and the reactor cooling water flow. Fault 10 corresponds to a random variation in the C feed temperature. It affects the stripper temperature and therefore relates to XMEAS (18).

Causes Var.	Effects var.	Causes Var.	Effects var.
XMEAS (1)	$\{44\}$	XMEAS (27)	$\{33, 41\}$
XMEAS (3)	$\{43\}$	XMEAS (29)	$\{20, 17\}$
XMEAS (4)	$\{38, 45\}$	XMEAS (30)	$\{18\}$
XMEAS (5)	$\{32\}$	XMEAS (31)	$\{34\}$
XMEAS (7)	$\{13\}$	XMEAS (33)	$\{20\}$
XMEAS (9)	$\{2, 51\}$.	XMEAS (34)	$\{30, 33\}$
XMEAS (10)	$\{47\}$	XMEAS (35)	$\{31\}$
XMEAS (11)	$\{20, 22\}$	XMEAS (36)	$\{34, 35\}$
XMEAS (13)	$\{46\}$	XMEAS (37)	$\{4, 40\}$
XMEAS (15)	$\{49\}$	XMEAS (38)	$\{20\}$
XMEAS (16)	$\{7, 46\}$	XMEAS (40)	$\{19\}$
XMEAS (18)	$\{38, 50, 11\}$	XMV (42)	$\{2, 9, 21\}$
XMEAS (19)	$\{29\}$	XMV (43)	$\{32, 35\}$
XMEAS (21)	$\{2, 16\}$	XMV (46)	$\{5, 20\}$
XMEAS (22)	$\{14\}$	XMV (47)	$\{35\}$
XMEAS (23)	$\{20, 29\}$	XMV (48)	$\{12\}$
XMEAS (24)	$\{37\}$	XMV (50)	$\{19\}$
XMEAS (25)	$\{11\}$	XMV (51)	$\{21\}$
XMEAS (26)	$\{36\}$	XMV (52)	$\{22, 17\}$

TABLE I
The adjacency matrix describing the strength relationships between variables

First, the structure of the causal network is learned. The PC algorithm is used. It consists of a series of

statistical significance tests of conditional independence (e.g. the Fisher's z-transform for the estimation of partial correlations). More details about the learning algorithms structures are in [16], [15]. Second, to complete the structure of the classifier, a discrete node **D** is added and linked to all the network's variables. The resulting structure can be easily deduced from the adjacency matrix presented in Table I, where the connection strength relationships between variables are provided. The resulting graphical network is given in 2.

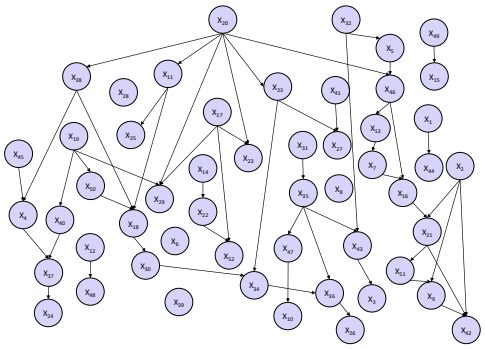

Fig. 2. TEP's causal network

To demonstrate the efficiency and ability of the proposed approach, we assume each faulty operating condition once as unknown operating conditions. That means we did not include such operating conditions in the learning stage of the classifier. We create four different scenarios (Sce.) in Table II. For each evaluation scenario, the causal network' uses 480 training samples of each class to train and estimate its parameters.

Sce.	Fault 1 non-stationary	Fault 4 Step	Fault 10 Variation	Fault 14 Sticking	
I	✓	✓	✓	×	
II	✓	✓	×	✓	NoC
III	✓	×	✓	✓	
IV	×	✓	✓	✓	

TABLE II

The studied scenarios, where ✓and × holds for known (part of the training set) and unknown fault, respectively

For each scenario, 800 observations from each of NoC, known and unknown operating conditions testing sets are presented to the network. First, fault detection is performed. The posterior probability corresponding to NoC of each sample is compared with its corresponding probabilistic limit, we set $\alpha = 1\%$ to control the overall false alarm rate. When a fault is detected the remaining steps of our approach are applied. The results of the fault detection and diagnosis procedures are presented in the confusion matrix presented in Table III.

Overall, the misdetection rate and the misclassification of known and unknown faults are presented. The proposed approach yields an excellent detection rate of different types of fault. The causal network successfully detects known and unknown faults. Regarding fault diagnosis, the approach discriminates well the various faults, an excellent average rate of 91.04% is obtained. This rate is the overall mean results of the discrimination between the different combinations of three known faults picked from the set composed of faults 1, 4, 10, and 14. To each one of this combination, the remaining fault is considered as unknown/undefined. Unknown faults are also well diagnosed by the proposed approach. An average diagnosis rate of 80.9%, which is an interesting result for a causal network that is not build to deal with data it wasn't trained with. The rate is however slightly lower than the one obtained by the known operating conditions. It would have been better if the considered faults do not overlap. The overall rates demonstrate the ability of the causal network to correctly discriminate between known and unknown faults.

	True classes		
Predic. classes ↓	Normal cdt.	Known cdt.	Unknown cdt.
Normal cdt.	**93.6**	5.6	0.78
Known cdt.	4.7	**91.04**	4.2
Unknown cdt.	4.7	14.3	**80.9**

TABLE III

The average classification rates %

Figure 3 shows the results of the identification procedure applied to the detected samples. It gives comparisons of the 52 variables in each scenario, regarding a class among faults 1, 4, 10 and 14. It shows all candidates either in known or unknown operating conditions. These candidates are obtained by the network. A significance level of α/m is considered following the Bonferroni correction [45].

Fault 1 shifts the system to uncontrollable operating conditions due to the conflict generated between the controller increasing the A feed flow in stream 1 in reaction to the decrease in the same feed in the recycle stream 5. This fault implies that XMEAS (4) and XMV (44) are directly affected, which then spreads to other parts. The fault identification results of our approach are shown in Figure 3, where root cause candidates stand out. Noticeably, five variables were appointed to characterize fault 1, which are variables 4, 18, 19, 44 and 50. The strength connections given in Table I would play as a second decision layer to reduce the number of candidates. Variables 4 and 44, are directly validated. it is not the case for variables 18, 19, and 50 as they share causal relationships (see Figure 2). From the adjacency matrix, it can be noticed a change in XMEAS (19) leads to changes in the stripper temperature XMEAS (18) and steam valve XMV (50). Therefore, the results of the proposed approach are perfectly in line with the expected results and the prior knowledge of the process.

The random variation in C feed, named Fault 10, tends to increase the stripper steam flow, see 1. This variation change has been well detected by the causal

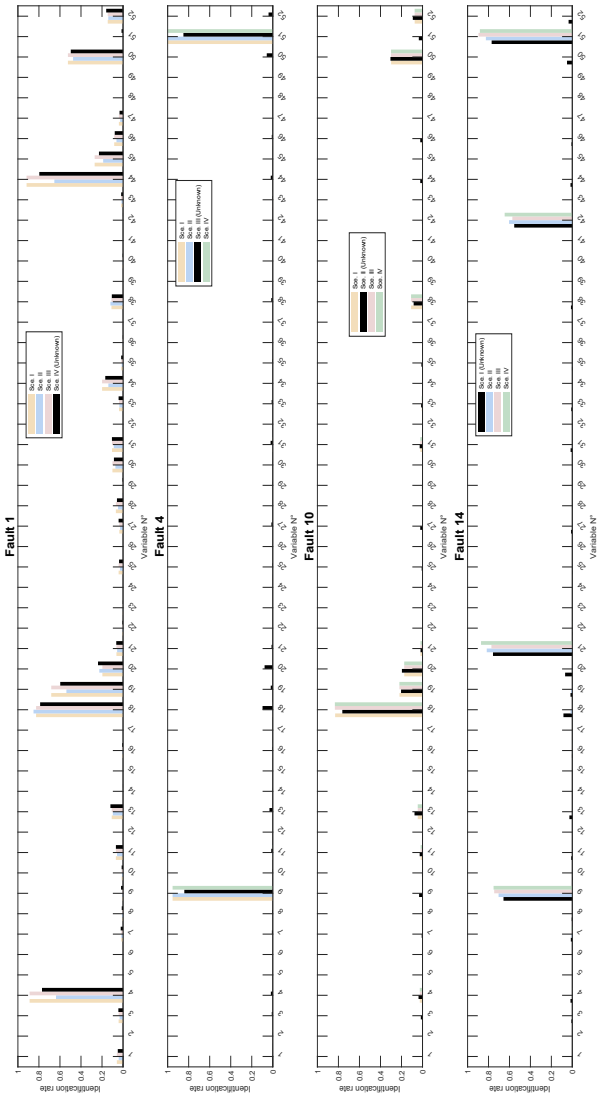

Fig. 3. Identification rates for Faults 1, 4, 10, and 14 in four scenarios

network and their responsible candidates illustrated in Figure 3. XMEAS (18) can be considered among the main references to fault 10.

Regarding Fault 4 and Fault 14, they both influence directly the reactor and they do share some candidates. From figure 3 the root variables 9 and 51 are correctly identified by the network. The variable XMEAS (51) is diagnosed the most frequently, which is consistent with the expectations that came with the descriptions of each fault. The increase in the cooling water flow tends to decrease the effect of the reactor cooling water temperature (towards normal operating conditions), which explains the great impact of XMEAS (51) regarding XMEAS (9). Both faults influence the reactor cooling water inlet temperature and since no sensor is measuring the inlet temperature, the value of the reactor temperature XMEAS (9) and the reactor cooling water flow XMEAS (51) can be used as references for fault 4.

Besides, other variables are also identified as candidates to explain fault 14 (whether simultaneously or not): XMEAS (21) and XMV (42). This actually can be explained by the propagation path deduced from the connection strength relationships between variables (see Figure 2). The variable XMEAS (9) dilutes its effect by spreading it to its descendant. Thus, under abnormal conditions, the deviation of XMEAS (9) from its normal behavior is spread through causality, thereby affecting its child node XMEAS (42). Variable XMEAS (21) plays a major role with regards to fault 14. Indeed, a striking in the reactor cooling water valve may cause as well a change in the reactor cooling water outlet temperature. Furthermore, it leads to a shift in the reactor temperature and its cooling water flow. This can be seen clearly from the causal network. One can deduce XMEAS (9) and XMEAS (21) as a reference for fault 14.

Here, both stationary and non-stationary faulty processes have been used to demonstrate the proposed approach. Faults with different natures, step, variation, and striking, have been introduced to show the feasibility of the proposed scheme. Therefore, this TEP case study fairly demonstrates the potential and the ability of this approach to monitor complex processes dealing with multivariate variables. It also shows how root cause variables tone down their effect on their descendants, but not the other way around. This confirms the interest of using a causal network to understand the propagation of one variable on other variables and also to locate the root-cause faulty variables rather than just identifying the variables contributing to significant indices.

V. CONCLUSIONS

In this paper, we presented a complete monitoring system that integrates fault detection and diagnosis as well as root cause identification for multivariate processes. The proposed monitoring approach (1) detects faults with respect to a false alarm rate; (2) identifies which variables are responsible for the faulty observations; and (3) diagnoses known and unknown faults. Four scenarios were evaluated to assess the performance of the proposed method. The results of detection and diagnosis are encouraging, and the identification of root cause results in variables that are consistent with the mechanism of the TEP.

Multivariate systems are the product of the interaction of their different elements and the behavior of each of them. It is then paramount to use this information to enhance decision making. The new approach follows this direction. Thus, it merits further investigation. It may be useful for the study of this and other related fields. Future directions would aim to investigate means to capture causality, extend the proposal to deal with multi-modal data and enhance decision making by an ensemble of causal models.

References

[1] Shen Yin, Steven X Ding, Xiaochen Xie, and Hao Luo. A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, 61(11):6418–6428, 2014.

[2] Steven X Ding. *Data-driven design of fault diagnosis and fault-tolerant control systems*. Springer, 2014.

[3] William H Woodall. Bridging the gap between theory and practice in basic statistical process monitoring. *Quality Engineering*, 29(1):2–15, 2017.

[4] Mohamed Amine Atoui, Sylvain Verron, and Kobi Abdessamad. Conditional gaussian network as pca for fault detection. *IFAC Proceedings Volumes*, 47(3):1935–1940, 2014.

[5] M Amine Atoui and Achraf Cohen. Fault diagnosis using pca-bayesian network classifier with unknown faults. In *2020 European Control Conference (ECC)*, pages 2039–2044. IEEE, 2020.

[6] Kaixiang Peng, Kai Zhang, Bo You, and Jie Dong. Quality-related prediction and monitoring of multi-mode processes using multiple pls with application to an industrial hot strip mill. *Neurocomputing*, 168:1094–1103, 2015.

[7] Mohamed Amine Atoui, Sylvain Verron, and Abdessamad Kobi. Fault detection with conditional gaussian network. *Engineering Applications of Artificial Intelligence*, 45:473–481, 2015.

[8] Chuyue Lou, Xiangshun Li, M Amine Atoui, and Jin Jiang. Enhanced fault diagnosis method using conditional gaussian network for dynamic processes. *Engineering Applications of Artificial Intelligence*, 93:103704, 2020.

[9] SX Ding, P Zhang, A Naik, EL Ding, and B Huang. Subspace method aided data-driven design of fault detection and isolation systems. *Journal of process control*, 19(9):1496–1510, 2009.

[10] Achraf Cohen and Mohamed Amine Atoui. On wavelet-based statistical process monitoring. *Transactions of the Institute of Measurement and Control*, 2020.

[11] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.

[12] Daniel Jung, Kok Yew Ng, Erik Frisk, and Mattias Krysander. Combining model-based diagnosis and data-driven anomaly classifiers for fault isolation. *Control Engineering Practice*, 80:146–156, 2018.

[13] Seongmin Heo and Jay H Lee. Fault detection and classification using artificial neural networks. *IFAC-PapersOnLine*, 51(18):470–475, 2018.

[14] Shuangyin Liu, Longqin Xu, Qiucheng Li, Xuehua Zhao, and Daoliang Li. Fault diagnosis of water quality monitoring devices based on multiclass support vector machines and rule-based decision trees. *IEEE Access*, 6:22184–22195, 2018.

[15] David Barber. *Bayesian reasoning and machine learning*. Cambridge University Press, 2012.

[16] Ruocheng Guo, Lu Cheng, Jundong Li, P Richard Hahn, and Huan Liu. A survey of learning causality with data: Problems and methods. *ACM Computing Surveys (CSUR)*, 53(4):1–37, 2020.

[17] Chuyue Lou, Xiangshun Li, and M. Amine Atoui. Bayesian network based on an adaptive threshold scheme for fault detection and classification. *Industrial & Engineering Chemistry Research*, 59(34):15155–15164, 2020.

[18] Jinxin Wang, Zhongwei Wang, Viacheslav Stetsyuk, Xiuzhen Ma, Fengshou Gu, and Wenhui Li. Exploiting bayesian networks for fault isolation: A diagnostic case study of diesel fuel injection system. *ISA transactions*, 86:276–286, 2019.

[19] M Amine Atoui, Achraf Cohen, Sylvain Verron, and Abdessamad Kobi. A single bayesian network classifier for monitoring with unknown classes. *Engineering Applications of Artificial Intelligence*, 85:681–690, 2019.

[20] Yalan Wang, Zhiwei Wang, Suowei He, and Zhanwei Wang. A practical chiller fault diagnosis method based on discrete bayesian network. *International Journal of Refrigeration*, 102:159–167, 2019.

[21] Mihiran Galagedarage Don and Faisal Khan. Dynamic process fault detection and diagnosis based on a combined approach of hidden markov and bayesian network model. *Chemical Engineering Science*, 201:82–96, 2019.

[22] Iago Pachêco Gomes and Denis Fernando Wolf. A health monitoring system with hybrid bayesian network for autonomous vehicle. In *2019 19th International Conference on Advanced Robotics (ICAR)*, pages 260–265. IEEE, 2019.

[23] Guohua Wu, Jiejuan Tong, Liguo Zhang, Yunfei Zhao, and Zhiyong Duan. Framework for fault diagnosis with multi-source sensor nodes in nuclear power plants based on a bayesian network. *Annals of Nuclear Energy*, 122:297–308, 2018.

[24] Aisong Qin, Qin Hu, Yunrong Lv, and Qinghua Zhang. Concurrent fault diagnosis based on bayesian discriminating analysis and time series analysis with dimensionless parameters. *IEEE Sensors Journal*, 19(6):2254–2265, 2018.

[25] Javier Herrera-Vega, Felipe Orihuela-Espina, Pablo H Ibargüengoytia, Uriel A García, Eduardo F Morales, Luis Enrique Sucar, et al. A local multiscale probabilistic graphical model for data validation and reconstruction, and its application in industry. *Engineering Applications of Artificial Intelligence*, 70:1–15, 2018.

[26] H. Gharahbagheri, S. A. Imtiaz, and F. Khan. Root cause diagnosis of process fault using kpca and bayesian network. *Industrial & Engineering Chemistry Research*, 56(8):2054–2070, 2017.

[27] Md. Tanjin Amin, Faisal Khan, and Syed Imtiaz. Fault detection and pathway analysis using a dynamic bayesian network. *Chemical Engineering Science*, 2018.

[28] Jie Yu and Mudassir M Rashid. A novel dynamic bayesian network-based networked process monitoring approach for fault detection, propagation identification, and root cause diagnosis. *AIChE Journal*, 59(7):2348–2365, 2013.

[29] Hongyang Yu, Faisal Khan, and Vikram Garaniya. Nonlinear gaussian belief network based fault diagnosis for industrial processes. *Journal of Process Control*, 35:178–200, 2015.

[30] Robert L Mason, Nola D Tracy, and John C Young. Decomposition of t 2 for multivariate control chart interpretation. *Journal of quality technology*, 27(2):99–108, 1995.

[31] Johan A Westerhuis, Stephen P Gurden, and Age K Smilde. Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and intelligent laboratory systems*, 51(1):95–114, 2000.

[32] Zhengbing Yan, Yuan Yao, Tsai-Bang Huang, and Yi-Sern Wong. Reconstruction-based multivariate process fault isolation using bayesian lasso. *Industrial & Engineering Chemistry Research*, 57(30):9779–9787, 2018.

[33] Bo He, Xianhui Yang, Tao Chen, and Jie Zhang. Reconstruction-based multivariate contribution analysis for fault isolation: A branch and bound approach. *Journal of Process Control*, 22(7):1228–1236, 2012.

[34] H Henry Yue and S Joe Qin. Reconstruction-based fault identification using a combined index. *Industrial & engineering chemistry research*, 40(20):4403–4414, 2001.

[35] Jian-Guo Wang, Xue-Zhi Cai, Yuan Yao, Chunhui Zhao, Bang-Hua Yang, Shi-Wei Ma, and Sen Wang. Statistical process fault isolation using robust nonnegative garrote. *Journal of the Taiwan Institute of Chemical Engineers*, 107:24–34, 2020.

[36] Hongchao Cheng, Jing Wu, Yiqi Liu, and Daoping Huang. A novel fault identification and root-causality analysis of incipient faults with applications to wastewater treatment processes. *Chemometrics and Intelligent Laboratory Systems*, 188:24–36, 2019.

[37] Yi Luo, Bhushan Gopaluni, Yuan Xu, Liang Cao, and Qun-Xiong Zhu. A novel approach to alarm causality analysis using active dynamic transfer entropy. *Industrial & Engineering Chemistry Research*, 59(18):8661–8673, 2020.

[38] Md Tanjin Amin, Faisal Khan, and Syed Imtiaz. Dynamic availability assessment of safety critical systems using a dynamic bayesian network. *Reliability Engineering & System Safety*, 178:108–117, 2018.

[39] Jialin Liu, David Shan Hill Wong, and Ding-Sou Chen. Bayesian filtering of the smearing effect: Fault isolation in chemical process monitoring. *Journal of Process Control*, 24(3):1–21, 2014.

[40] Daniel Jung. Residual generation using physically-based grey-

box recurrent neural networks for engine fault diagnosis. *arXiv preprint arXiv:2008.04644*, 2020.

[41] Jing Li, Jionghua Jin, and Jianjun Shi. Causation-based t 2 decomposition for multivariate process monitoring and diagnosis. *Journal of Quality Technology*, 40(1):46–58, 2008.

[42] Robert L Mason, Nola D Tracy, and John C Young. A practical approach for interpreting multivariate t 2 control chart signals. *Journal of Quality Technology*, 29(4):396–406, 1997.

[43] M Amine Atoui and Achraf Cohen. Coupling data-driven and model-based methods to improve fault diagnosis. *Computers in Industry*, 128:103401, 2021.

[44] Andreas Bathelt, N Lawrence Ricker, and Mohieddine Jelali. Revision of the tennessee eastman process model. *IFAC-PapersOnLine*, 48(8):309–314, 2015.

[45] Mohamed Amine Atoui, Sylvain Verron, and Abdessamad Kobi. A bayesian network dealing with measurements and residuals for system monitoring. *Transactions of the Institute of Measurement and Control*, 38(4):373–384, 2016.