Process Fault Diagnosis with Model- and Knowledge-based Approaches: Advances and Opportunities

Weijun Li, Hui Li, Sai Gu, Tao Chen*

Department of Chemical and Process Engineering, University of Surrey, Guildford GU2 7XH, United Kingdom

**Corresponding author. Tel.: +44 1483 686593. E-mail: t.chen@surrey.ac.uk (T.Chen).*

Abstract

Fault diagnosis plays a vital role in ensuring safe and efficient operation of modern process plants. Despite the encouraging progress in its research, developing a reliable and interpretable diagnostic system remains a challenge. There is a consensus among many researchers that an appropriate modelling, representation and use of fundamental process knowledge might be the key to addressing this problem. Over the past four decades, different techniques have been proposed for this purpose. They use process knowledge from different sources, in different forms and on different details, and are also named model-based methods in some literature. This paper first briefly introduces the problem of fault detection and diagnosis, its research status and challenges. It then gives a review of widely used model- and knowledge-based diagnostic methods, including their general ideas, properties, and important developments. Afterwards, it summarises studies that evaluate their performance in real processes in process industry, including the process types, scales, considered faults, and performance. Finally, perspectives on challenges and potential opportunities are highlighted for future work.

Key words: Fault diagnosis; Fault detection; Process monitoring; Process knowledge; Process safety

List of abbreviations

BN	Bayesian network
CR	Compensatory response
CSTR	Continuous stirred tank reactor
DAE	Differential algebraic equation
DBN	Dynamic Bayesian network
DCS	Distributed control system
DE	Differential equation
EKF	Extended Kalman filter
ESDG	Extended signed directed graph
FDD	Fault detection and diagnosis
FT	Fault tree
FTA	Fault tree analysis
HAZOP	Hazard and operability analysis
HMM	Hidden Markov model
ICA	Independent component analysis
IR	Inverse response
KF	Kalman filter
NOC	Normal operating condition
OOBN	Objective oriented Bayesian network
P&ID	Piping and instrumentation diagrams
PCA	Principle component analysis
PF	Particle filter
PLS	Partial least square
PR	Parity relation
PS	Parity space
QS	Qualitative simulation
QTA	Qualitative trend analysis
SDG	Signed directed graph
TNBN	Temporal nodes Bayesian network
T-S	Takagi-Sugeno
UIO	Unknown input observer
UKF	Unscented Kalman filter
XML	Extensible Mark-up Language

1. Introduction

Modern process systems are large in scale and complex in process, resulting in increasing demands on their safety and reliability. There is a growing interest in developing techniques to address abnormalities that occur frequently in process industry hence to ensure safe and efficient productions. Fault detection and diagnosis (FDD) is the key technology for addressing this problem. It specifically addresses the unacceptable deviation between the observed and calculated variables. Ideally, FDD provides timely, reliable and interpretable decision supports to the operators about the process states.

Three tasks are usually associated with this practice: 1) fault detection, 2) fault diagnosis, 3) fault prognosis. Fault detection is the most basic task with the aim to detect potential abnormality and determine its occurrence time. It is accomplished by checking the deviations between observed and calculated variables or departures from an acceptable operational region. Fault detection is relatively easy compared to the other two tasks and has achieved remarkable success (Qin, 2012; Yin et al., 2014). Fault diagnosis is the “downstream” process of fault detection which further requires to identify the type, location and magnitude of the fault (Isermann & Ballé, 1997). It still remains a challenge due to several reasons such as complex process dynamics (Dash & Venkatasubramanian, 2000), fault propagations, various possible faults and scarcity of fault data. Fault prognosis is to forecast future abnormality so that predictive maintenance and shutdown can be scheduled. It requires accurate knowledge in how abnormality evolves and propagates and is considered as a future technology with limited reported solutions (Reis & Gins, 2017).

Various approaches to fault diagnosis have been developed and they can be divided into three

main categories: data-driven approaches, deep-knowledge-based approaches and analytical-model-based approaches. It is worth noting that all methods require operational data to estimate some parameters and data-driven methods herein specifically refer to methods that rely purely on operational data without using pre-exist process knowledge. Data-driven approaches learn from history and place no requirements on first-principle models or expert knowledge. Because of this merit and the large amount of operational data coming from the wide use of distributed control systems (DCSs), they have been very popular in recent years. The extensive investigations have generated many exciting outcomes, which have been reviewed in some dedicated papers (Ge et al., 2013; Qin, 2012; Severson et al., 2016; Yin et al., 2014) and will not be discussed in detail in this paper.

In contrast, analytical-model- and deep-knowledge-based methods rely heavily on fundamental understanding of the underlying physics and chemistry (and sometimes biology) of the process. They are distinguished herein by the form (analytical or non-analytical) of knowledge used. Despite the need for more expert efforts in the deployment and maintenance, using such knowledge has many benefits. For instance, it might provide causal effect relationships which are essential for fault diagnosis, prognosis and risk assessment. Another desirable characteristic is the interpretability or explicability of decisions they make. Operators would continue to participate in the diagnostic task in near future, and they may want to understand decisions and judge their reliability from a mechanistic perspective. In addition, much mechanistic knowledge is available even for a new process, including flowsheets, main chemical reactions, causal effect relationships, and increasingly often, sophisticated mathematical models. The fault diagnosis exercise may benefit from modelling, representing and use in depth of such knowledge. The

exploitation of model- and knowledge-based methods is essential and of long-term interest.

In past decades, analytical-model- and deep-knowledge-based methods have been extensively studied. Those reported in the literature use different levels of knowledge from different aspects, whether alone or in combination with data-driven methods, providing effective and flexible FDD solutions. However, to the authors' knowledge, apart from some early comprehensive reviews (Venkatasubramanian, Rengaswamy, Yin, & Kavuri, 2003; Venkatasubramanian, Rengaswamy, & Kavuri, 2003), there has been limited effort in systematically reviewing this topic, especially in process industry.

Against this background, this paper provides a systematic review on the state-of-the-art fault diagnosis methods in this category. Different types of strategies are introduced and the method properties are discussed. Fig. 1 summarises the main techniques considered in this paper. As most developments are tested on simulations, their performance on real processes may be questioned. Therefore, in Section 4, we summarise in detail the studies that evaluate FDDs on real processes in process industry. Some perspectives on challenges and promising research directions are highlighted at the end of the paper.

Fig. 1. Overview of model- and knowledge-based FDD methods

2. Analytical-model-based methods

Analytical-model-based FDD methods have been extensively studied in last century and are relatively mature. The basic idea is to run process models (either in state space form or input-output representations) in parallel to the real process and apply certain algorithms to reconstruct process behaviour online. When a fault occurs, model relations are no longer satisfied, thereby generating differences between the real and estimated behaviour or the so-called residuals, which are considered as fault indicator.

A general scheme of analytical-model-based FDD is presented in Fig. 2. As shown, faults can occur either in actuators, sensors, or process components. The faults can be further classified into additive faults and multiplicative faults and the former faults affect processes as unmeasured inputs whilst the latter faults can change process dynamics. Based on the measured process inputs and process model, the FDD methods generate features including residuals $\mathbf{r}$, state estimates $\hat{\mathbf{x}}$ and parameter estimates $\hat{\boldsymbol{\theta}}$. They are further evaluated to determine the occurrence of a fault and its time, location, type and so on.

Fig. 2. General scheme of analytical-model-based methods

2.1. State observer

State observer is to use process inputs, measurements and models to reconstruct some observable states of the process. The errors between the model and process outputs are fed back to recalibrate the reconstructed states. The corresponding feedback gain matrix is designed such that the reconstructed states converge fast to the plant states regardless of the process inputs and initial conditions. In observer-based FDD, the residuals are generated from the reconstructed states and compared with fixed or adaptive thresholds to determine the occurrence of a fault.

Such fault detection scheme works satisfactorily given perfect process models. However, in practice the residuals can be affected by unknown inputs such as modelling errors, process and measurement noise and disturbances, thereby resulting in high false alarm. To address this shortcoming, efforts have been made to design residuals that are robust to those unknown inputs but sensitive to fault signals. Unknown input observer (UIO) is an effective method for this purpose (Frank & Wunnenberg, 1989). The main concern is to take some design freedoms to decouple the effects of unknown inputs from that of faults. By treating specific faults as unknown inputs and designing a bank of UIOs, fault diagnosis can be achieved (Termehchy & Afshar, 2014; Du et al., 2013). In doing this, each observer is designed to be sensitive to a specific fault whilst invariant to other faults and unknown inputs (Gertler, 1988), or to be invariant to a particular fault and unknown inputs but sensitive to other faults (Frank, 1990).

Introduced above are linear-model-based observers which are a relatively mature field of study. Many recent efforts have been devoted to design of nonlinear observers. The design of FDD system for a general type of nonlinear processes remains a challenge and most methods were

developed for certain kinds of nonlinear systems, such as Lipschitz nonlinear systems (Zhang et al., 2010) and affine nonlinear systems (Yang et al., 2015). Keliris et al. (2015) discussed a distributed fault detection observer with adaptive threshold, for input–output interconnected Lipschitz nonlinear systems with modelling uncertainties and measurement noise. To deal with general nonlinear systems, one could use linear models to approximate the system around the operating points. Takagi-Sugeno (T-S) fuzzy model is a representative technique for this purpose (Takagi & Sugeno, 1985). It develops a set of linear sub-models to represent local dynamics and further combines them by nonlinear membership functions to describe global system behaviour. Li, Ding, Qiu, et al., (2016) proposed a weighted piecewise-fuzzy observer that built local fuzzy Lyapunov functions and weighted their residual signals. Furthermore, its design in presence of norm bounded external disturbances was discussed in (Li, Ding, Yang, et al., 2016).

Another category of study focuses on design of advanced observers, such as adaptive observer (Zhang et al., 2008), proportional and integral observer (Gao et al., 2008) and sliding mode observer (Alwi & Edwards, 2014). Adaptive observers treat faults as unknown parameters and perform a recursive joint estimation of states and parameters. The estimated parameters are further examined for fault identification. Proportional integral observers can be considered as a higher-order observer for an extended state vector composed of both faults and system states. Such that it reconstructs not only the states but also the faults. Sliding mode observers, on the other hand, introduce variable structure control inputs into the system to drive its trajectory to a predefined sliding surface in finite time and force it to remain on that surface thereafter (Edwards, Spurgeon, & Patton, 2000). Usually the output error or state error equalling to zero

is chosen as such surface, so that the observer can track the true process states whilst being robust and insensitive to unknown inputs such as disturbances and unmodelled dynamics.

2.2. Parameter estimation

The basic assumption behind this approach is that a fault can cause changes in some unmeasurable physical parameters of the system, and these changes will be further reflected in variations of some model parameters. Two steps are usually required in performing parameter estimation based FDD: 1) determining the relationships between physical and model parameters by expert knowledge, $\mathbf{p} = \mathbf{f}(\boldsymbol{\theta})$; 2) estimating variations in model parameters and the corresponding changes in physical parameters, as illustrated in Fig. 3.

This straightforward approach, initiated in (Baskioti et al., 1979; Isermann, 1985), has been enhanced with various techniques subsequently. Reppa & Tzes (2011) used hyper ellipsoid to represent the acceptable parameter set (nominal parameter values with bounded uncertainties). It detected a fault when the calculated parameter set exceeds the hyper ellipsoid. By retaining the new nominal parameter set, it identified the fault type and magnitude. Recently, Díaz et al. (2016) achieved fault diagnosis by solving an inverse problem. Although effective and being able to diagnose multiple faults, it did not consider model errors, so its applications in large processes may be limited.

Fig. 3. Parameter identification based fault diagnosis

Clearly, the parameter estimation requires an accurate parametric model which is not an easy task. In addition, it requires solving nonlinear optimisation problem, hence can be computationally intensive for online application. Some efforts for simplifying online calculation can be found in (Doraiswami et al., 2010; Vachhani et al., 2001). The latter paper reported a two-stage diagnosis strategy. In the offline phase, it used simulated data to train a neural network to distinguish the fault patterns caused by the deviation of each parameter. In the online stage, it identified the faulty parameter by the neural network and then estimated only that parameter, thereby resulting in faster and more accurate estimation.

2.3. Parity space

Parity space (PS) is another common method for fault detection and diagnosis. It transforms the process model to parity relations (PRs) which describe the inputs and outputs relationships independent on the states of the system. Such relations are used to generate residuals to check the consistency between the model and process outputs. In presence of a fault, they are no longer satisfied, thereby generating nonzero residuals as fault indication. PR can be represented

by algebraic equations or difference equations, which provide direct redundancy and temporal redundancy, respectively. Direct redundancy relates instantaneous outputs of different sensors and is usually restricted to sensor fault detection. Whilst temporal redundancy relates histories of sensor outputs and actuator inputs and is useful for detecting both sensor and actuator faults (Chow & Willsky, 1984).

This technique was pioneered by (Desai & Ray, 1981; Desai & Ray, 1984) and further explored by (Gertler & Monajemy, 1995; Gertler & Yin, 1996; Wang & Wu, 1993). Patton & Chen (1991) provided a dedicated review. The residuals can be affected by model uncertainties, process and measurement noise, and unknown disturbances. Therefore, significant efforts have been made on improving the robustness of PS based residual generators. Chow & Willsky (1984) achieved this by minimising the maximum expected residual values under normal operating conditions (NOCs). Alternatively, Ding et al., (1999) minimised the ratio between the sensitivity of the residuals to unknown disturbances and to faults, which was a widely adopted performance index.

By designing the residuals to be sensitive to one fault but robust to other faults, fault diagnosis can be achieved. Odendaal & Jones (2014) presented a robust diagnosis solution: 1) decrease the residual sensitivities to unknown disturbances; 2) when a specific fault occurs, increase the sensitivity to this fault but decrease the sensitivities to other faults. It is worth noting that this strategy needs to be repeated for each of the concerned faults. Another diagnosis strategy was developed by (Gertler & Singer, 1990) in which PRs were designed to be orthogonal to concerned additive faults and arranged into specific structures to provide isolability. Gustafsson

(2002) explicitly computed the misdiagnosis probabilities and discussed the measures to improve diagnosis performance and robustness to sensor noise and design parameters.

Another interesting study (Izadi et al., 2008) considered the case of non-uniform sampling where the key idea was not to fix the length of time window but to fix the number of samples.

It is worth noting that most studies focused on linear systems. The PS based FDD methods can be generalised to nonlinear systems by using the T-S fuzzy model (Nguang et al., 2007).

2.4. Kalman filter

The residual generation methods given above were developed for deterministic systems.

However, in practice many systems have random fluctuations in their state values and measurements. Kalman filter (KF) is an efficient approach for state estimation in such systems.

It is a recursive algorithm that takes two steps to approximate the optimal state estimate: 1) predicting forward the current state and error covariance to obtain a priori estimate; 2) feeding back the error between the model and process outputs by a gain matrix (Kalman gain) to recalibrate the state estimate to obtain an improved posteriori estimate. Here the error gain matrix is determined by minimising the posteriori state estimate; readers are referred to (Brown & Hwang, 1992) for the detailed introduction.

Fault detection with KF can be formulated as a hypothesis test problem with the NOC considered as null hypothesis (Mehra & Peschon, 1971). FDD can be achieved by designing a bank of KFs based on all known fault models and performing multiple hypothesis tests (Bøgh, 1995).

KF is originally defined for linear system. Extended Kalman filter (EKF) and unscented Kalman filter (UKF) are representative extensions for nonlinear systems (Romanenko & Castro, 2004). EKF propagates previous state estimate through nonlinear state dynamic and observation models to obtain priori estimate and output prediction. It further linearises the state dynamic model (around the estimate at the previous time step) and the observation model (around the priori estimate), and uses the resulting first-order partial derivative matrix (or so-called Jacobian matrix) for calculating the priori and posteriori error covariance of the estimate. EKF has been widely investigated for FDD (Li & Olson, 1991; Chetouani, 2004). EKF has two important tuning parameters, i.e. covariances of process and measurement noise, which are usually empirically set as constants. In order to reduce dependence on prior knowledge and improve estimation performance, adaptive EKF performing online estimation of the covariances has been developed (Salahshoor et al., 2008). To speed up computation and robustness (in case of sensor failure), they further investigated parallel EKFs. Another drawback of EKF is the significant errors introduced in first-order linearisation. To address this problem, Huang et al., (2003) introduced pseudomeasurements which were obtained by interpolating new and the previous measurements during each sampling period. Such that linearisation can be performed at these new additional measurements.

Conventional EKF cannot handle equality or inequality constraints. Vachhani et al. (2005) addressed this issue by integrating EKF and moving-horizon optimisation technique which used EKF to provide a priori estimates and solved the constrained optimisation problem in the update step to obtain updated estimates using the new measurement. Prakash et al. (2014) proposed an alternative which obtained priori estimates based on samples generated from truncated

multivariate Gaussian distributions that satisfied state constraints. It is worth noting that EKFs require calculation of first-order partial derivatives (or so-called Jacobian matrix) and only work on differentiable process models. In addition, EKFs only adopt the first-order approximation, which may be inadequate for highly nonlinear processes.

Unscented Kalman filter (Wan & Menve, 2000) is another nonlinear variant. It selects a minimal set of sample points (called sigma points) associated with corresponding weights to capture the same mean and covariance of the previous state estimate. These sigma points are propagated through the nonlinear state dynamic model and the resulting transformed points together with their weights are used to estimate the priori mean and covariance of the state. This process is the so-called unscented transformation (UT) and approximates the distribution accurately to the third order for Gaussian inputs for any nonlinearity (Wan & Menve, 2000). By further propagating these points through the nonlinear observation model, the mean and covariance of the output as well as the cross covariance between output and state estimate can be computed. These statistics will be used in the update step for computing the posteriori state estimate and covariance.

The absence of linear approximation step and, consequently, errors and the calculation of Jacobian matrix introduced by it, enables the application of UKF in a wider range of situations where the process is highly nonlinear or the model is indifferentiable. It is worth noting that because UKF uses only a small number of samples, its computational complexity is the same order as EKF (Wan & Menve, 2000). Its superior performance over EKF has been reported in some comparative studies (Romanenko & Castro, 2004; Qu & Hahn, 2009; Kandepu et al., 2008). In a recent development, Mosalanejad & Arefi (2018) modified UKF for joint state and

parameter estimation in which multi-sensor fusion and asynchronous measurements issues were considered and addressed by using some rules.

2.5. Particle filter

Particle filter (PF) is an alternative filtering technology that does not require linear models, model linearisation or Gaussian posterior distribution hence can be applied to a wider range of systems. The basic idea is to generate a large number of samples (called particles) to approximate the posterior density of the states. Due to the difficulty in sampling directly from target density (Chen et al., 2005), importance sampling technique is applied which samples from an alternative density function (called importance density) and assigns weight to each sample based on the ratio of its likelihood in the target density to the likelihood in the importance density. The target posterior state density is approximated by these weighted samples.

In the sequential state estimation problem, when a new measurement is received, new samples are drawn from the importance density and their weights are recursively updated (based on sample weights at the previous time). Notice that depending on the choice of importance density, new samples may be conditioned on the previous sample distribution or measurements, resulting in different filtering performance. Readers are referred to dedicated tutorials (Tulsyan et al., 2016; Arulampalam et al., 2007) for more information.

The implementation of PF may face one problem that the weights will become increasingly concentrated over time and therefore most sampled particles become meaningless (Doucet et al., 2000). This limitation has been addressed by breaking big particles into smaller ones with

equal weights. PF is capable to approximate any nonlinear and non-Gaussian systems and improve the accuracy by simply increasing the “number of particles”. Although it is computationally intensive, the advances in computation technologies will allow its extensive use.

PF for fault diagnosis was pioneered in (Kadirkamanathan et al., 2000; Li & Kadirkamanathan, 2001). In the former study, fault diagnosis was achieved by developing an augmented state model in which changes in process states and parameters were simultaneously estimated and tracked. The latter study combined PF and the likelihood ratio test, and estimated the magnitude and time of the fault simultaneously. PF based FDD has been illustrated on simulations of batch processes (Chen et al., 2005), pipeline leak (Liu et al., 2005) and CSTR (Bahmanpour et al., 2007; Chen et al., 2008).

Recently, Alrowaie et al., (2012) developed a bank of PFs to represent models of all known faults and NOC. Hidden states were estimated by such filters and then used in likelihood ratio test to determine the fault type. Yin & Zhu (2015) proposed an intelligent PF that used the genetic operators, such as mutation and crossover, to modify the small-weight particles into large-weights ones. This strategy can improve the particle diversity and thus mitigate the impoverishment problem.

The choice of the importance density is critical to the performance of PF. State transition probability is a popular choice, which however does not take the current measurement into account and therefore can be inefficient in presence of significant modelling errors. To address this problem, (Jayaprasanth & Kanthalakshmi, 2016; Shenoy et al., 2013) combined UKF and

PF and generated the importance density by unscented transformation to incorporate the knowledge of measurement, thereby retaining the ability to estimate non-Gaussian distributions whilst providing robustness against modelling errors.

3. Deep-knowledge-based methods

The analytical methods viewed in Section 2 express the fundamental understanding of the process in terms of mathematical functional relationships between inputs and outputs of the system. In many situations, it is difficult to obtain accurate knowledge to build such models. Quite often we have process knowledge in other forms, such as qualitative behaviour of the system, if-then rules, probabilistic conditional relationships, causal-effect relations and process connectivity. The flexibility in knowledge representation allows effective use of available knowledge. Given such knowledge, different techniques have been developed for fault diagnosis, and typical techniques will be reviewed below.

3.1. Qualitative simulation

Qualitative simulation (QS) provides an abstract description of the behaviour of a system and ignores much of its quantitative details. The simplified qualitative model can provide a reliable prediction of system behaviour under normal and various faulty conditions, thereby providing crucial diagnostic knowledge. QS allows to draw partial conclusions from uncertain and incomplete process knowledge. Kler & Brown (1984) and Kuipers (1986) pioneered this technique and proposed two representative QS methods, termed IQA and QSIM, respectively. IQA uses a set of confluence equations (which can be derived from mathematical models) to capture the qualitative relationships between physical variables. Each variable takes $[+]$, $[-]$, $[0]$, or $[?]$ to represent its change direction. Sign arithmetic is defined so as to “solve” the confluence

equations to get all qualitative behaviour of the system given a certain deviation.

Herbert & Williams (1987) made early attempts on IQA-based FDD. A set of confluences was used to monitor the process, and once it was violated, a fault was detected. Fault diagnosis was performed by suspending each confluence until all remaining confluences can be satisfied. Waters & Ponton (1989) discussed its use in fault propagation prediction and indicated that it can have serious efficiency problem due to combinatorial explosion. As qualitative models can produce many spurious solutions, to improve diagnosis resolution, Zhang et al. (1990) included information on magnitude order relations between different variables. The same authors discussed a self-learning diagnostic system with enhanced robustness (Zhang et al., 1991). Although easy to use, IQA is fundamentally limited to algebraic equations and hence to describing steady-state behaviour.

In contrast, QSIM abstracts dynamic behaviour of a system by using qualitative differential equations and can therefore reason about dynamic behaviour of the system. QSIM represents the qualitative state of a variable by $\langle qval, qdir \rangle$. *qval* uses an ordered set of landmark values to loosely describe the value of the variable, whilst *qdir* stands for the change direction. A set of qualitative constraints has been developed to describe the relations between these variables, including their arithmetic, differential and monotonic relations, referring to (Kuipers, 1986) for the details.

The original QSIM suffers from two limitations: 1) the inherent ambiguity; 2) the lack of temporal information as the time domain is discretised. To this end, Shen & Leitch (1993) proposed a fuzzy QSIM algorithm that allowed more quantitative information to be included, such as relative strengths of qualitative relations and the rate-of-change of variables. Such that

the spurious solutions can be greatly reduced and the temporal durations and possible transitions of each state can be computed. The diagnosis problem was formulated as a parameter identification problem in ref (Steele & Leitch, 1997). By using fuzzy qualitative models, only a smaller and finite search space needed to be explored. A recent study in combining QSIM and qualitative trend analysis can be found in (Lu et al., 2016).

3.2. Expert system

An expert system is an organised knowledge system that mimics a human expert to solve problems in a specific domain (Liao, 2005). Plant operators accumulate theoretical and practical expertise over years, based on which they can reason out the cause of a potential fault and suggest corresponding corrective actions. The expert system is designed to automate this process. A typical expert system includes four components: knowledge base, inference machine, knowledge management and user interface (Qian et al., 2003), as shown in Fig. 4.

Fig. 4. Schematic diagram of expert systems

Proper selection of the knowledge base is a critical factor for the online performance of an expert system. The domain knowledge can be obtained from various sources and in flexible forms, such as “if-then” rules, mathematical models and signed directed graphs. The inference

machine reasons according to the knowledge base and makes judgement on fault patterns and appropriate actions. Inferring time and accuracy can relate to the inference mechanism, quality and structure of the knowledge base. Knowledge management is to make necessary modifications of knowledge base to address issues like knowledge conflict, redundancy, insufficiency. User interface facilitates the communication between operators and the expert system. The expert system has advantages of transparent reasoning and flexibility in knowledge representation (Dhaliwal & Benbasat, 1996). The main difficulty in applying the diagnostic expert system is the generation, management and hierarchisation of the process knowledge. Particularly, the accurate and adequate knowledge representation is quite challenging.

Chester et al. (1984) and Rich et al. (1989) were the first studies to develop the expert system for fault diagnosis in process industry. Addanki & Sethuraman (1992) developed a rule-based expert system for the water-steam cycle system of a 210MW industrial thermal power station. Puñal et al. (2002) discussed an expert system for a pilot-scale wastewater treatment plant. Qian et al. (2003) applied expert system to fluid catalytic cracking system in refineries. It used a knowledge base containing both general static rules and time-variant rules and a bidirectional reasoning strategy for computational efficiency. In order to examine inconsistencies in heuristic rules including contradiction, redundancy, circulation, repetition and subsumption, integrality verification modulars and algorithms have been developed in (Qian et al., 2005). Rule-based reasoning usually requires some boundaries to distinguish certain process states. In order to handle uncertainties in these boundaries, fuzzy logic has been introduced in expert system (Feng et al., 1998; Zahedi et al., 2011).

The above expert systems used rule-based diagnostic knowledge, which are easy to develop without resorting to complex quantitative domain knowledge. However, it may have limitations such as low resolution, combinatorial explosion, and inability to capture accurately time-varying and spatially varying process knowledge. Model-based expert system can be a solution to this problem. It provides a more general and compact knowledge base and has advantage in handling novel situations.

Surgenor & Jofriet (1992) discussed an expert system that utilised quantitative governing equations to generate a set of residuals and associated their deviations with certain fault scenarios. Kordon & Dhurjati (1995) discussed an expert system that obtained rules from both expert experience and process simulation. Terpstra et al. (1992) represented a process in hierarchical object-oriented description. These objects can be modelled in different formats (quantitative, qualitative, rules, inequalities, and boolean equations) and diagnosed by different techniques. Its main advantages include: 1) ease in maintaining and reusing the knowledge base; 2) allowing to use different techniques simultaneously.

In addition to quantitative models, expert systems have been combined with many other techniques. Nan et al. (2008) combined the expert system with the qualitative trend analysis (QTA). It extracted the trend patterns (temporal behaviour) of the variable and matched them with the patterns related to each fault scenario in the knowledge base. Chan (2005) discussed a knowledge base construction strategy based on ontology modelling (discussed in later section), which can facilitate knowledge sharing, reusing and improve knowledge base construction efficiency. An integration of neural network and fuzzy expert systems can be found in (Özyurt & Kandel, 1996).

It is worth noting that the difficulty in deployment of a diagnostic expert system increases dramatically with system size and complexity. This can be mitigated by decomposing the system into manageable units and applying multitiered hierarchical diagnostic procedures (Prasad et al., 1998). Finch & Kramer (1988) decomposed a system into a set of subsystems or units according to their functions. By checking their functional states and dependencies, some possible fault sources can be quickly located for further more detailed diagnosis, thereby avoiding unnecessary details in early stage of diagnosis. Zhang & Roberts (1991) discussed a structural decomposition strategy that decomposed a system into subsystems according to the process topology and then used Boolean matrices to explicitly represent the relations and causalities between and inside subsystems. Structural decomposition is straightforward to implement and explain, and can effectively narrow the diagnostic focus.

3.3. Signed directed graphs

Signed directed graphs (SDGs) are among the most extensively studied qualitative methods for process fault diagnosis. They use nodes and directed arcs to represent the events or variables of the system, and the causal relationships among them, respectively. Sign is attached to each node to qualitatively represent its status. The directed arcs among the nodes are drawn from the “cause” nodes to the “effect” nodes. Such that a large amount of process knowledge is explicitly represented in a compact form. SDGs can be obtained via two ways:

- Mathematical models such as differential equations (DEs) or differential algebraic equations (DAEs) (Iri et al., 1979; Oyeleye & Kramer, 1988)

- Expert knowledge on process connectivity and causality. Note that building SDGs by experts can be error-prone and tedious for large processes (He et al., 2014).

A fundamental assumption of the SDG techniques is that causality must connect the fault origins to the observed symptoms. SDGs achieve diagnosis by searching the consistent paths which potentially explain the local fault propagation pathways, and through reasoning locate all possible fault origins. SDGs have advantages in completeness, which means that the origin of the unusual event is often included in the multiple interpretations. However, as qualitative models, SDGs using alone can generate spurious solutions. It is worth noting that signs of the directed arcs are dependent on the status of the process. However, in most cases they are derived from the causal relationships at the initial or steady state response of the system and are therefore valid only in early stage of the faulty operation (He et al., 2014). As a result, only the early faulty samples can be used for the diagnosis purpose.

The concept of SDGs for fault diagnosis was first proposed by (Iri et al., 1979). Umeda et al. (1980) comprehensively discussed the generation of SDGs from DEs of the process. In following developments, considerable efforts have been made to improve the diagnostic resolution. Kokawa et al. (1983) suggested to use priori knowledge such as fault propagation probabilities and fault propagation rates to rank the fault origin candidates generated by standard SDGs. Shiozaki et al. (1985) discussed an approach that extended the node state from original three states ("+", "-", "0") to five states ("+", "+?", "0", "-?", "-") and thus allowed to describe transition states and handle uncertainties on the node states. Another popular technique is to introduce fuzzy set theory (Han et al., 1994; Wang et al., 1995; Tarifa & Scenna, 1997; Chen & Chang, 2009). The key idea is to use membership functions to represent the seriousness

of the node deviation, thus enable to handle uncertainties in node states and incorporate more quantitative information.

Another approach for improving resolution is to introduce additional constraint information such as analytically redundant equations derived from steady state process equations (Oyeleye & Kramer, 1988). Above methods use only static symptom, which also limits the diagnostic resolution. Against this background, Shiozaki et al. (1989) proposed a new method that further utilised the temporal information, specifically the abnormality revealing order of each node, resulting in higher diagnostic accuracy.

Another important contribution was made by (Kramer & Jr., 1987; Tarifa & Scenna, 1997) where the SDG was used as a qualitative simulator so that the diagnostic rules can be derived as response to the deviation of each node. These rules can be cast into an expert system framework and integrated with other rules on plant operations. Tarifa & Scenna (2003) developed a fuzzy expert system where the knowledge based was automatically generated from SDGs. However, due to completeness of SDG models, this approach may lead to rule explosion. To address this issue, Lee et al. (1997) proposed a knowledge compression technique which used a cluster of variables instead of a single variable to represent a node. The combination of variable states in a node constitutes the qualitative statues that the node can take. In doing this, the SDG model size and diagnostic rules can be greatly reduced.

Other problems that SDGs need to address include the compensatory response (CR) and inverse response (IR) which occur due to the presence of negative feedback loops (e.g. controller) or multiple feed forward paths with conflicting effects (Venkatasubramanian, Rengaswamy, &

Kavuri, 2003). CR brings the node state back to normal, and IR brings it to an opposite value.

This can break the associated consistent paths during the reasoning. Oyeleye & Kramer (1988) proposed a solution which drew additional arcs across the associated variables to capture the IRs and CRs, termed extended SDG (ESDG). Maurya et al. (2003a, 2003b) systematically studied the generation and analysis of SDGs for fault diagnosis of chemical processes, strategies to deal with CR, IR and spurious fault origins.

Most of the above studies assume that single fault is responsible for the process abnormality. However, multiple faults do exist in process industry. To handle such cases, Vedam & Venkatasubramanian (1997) made an important extension by using the concept of minimal cut sets. It aimed to find the minimal number of fault origins able to explain the process failure. Another interesting study can be found in (Zhang et al., 2005) which diagnosed multiple faults based an inverse inference strategy instead of the commonly used forward inference strategy.

More recently, probabilistic SDGs have been proposed to account for the conditional relationships among the states of the nodes (Lü et al., 2011; Peng et al., 2014; Liu et al., 2016). Han et al. (2018) and He et al. (2014) integrated principle component analysis (PCA, a dimension reduction technique that reserves major variations in the dataset) and SDGs where PCA detected faulty variables by finding the combination of variables that contributes most to the deviation of the monitoring statistics from a predefined threshold. Such variables were considered as possible root causes and sent to SDGs for further confirmation. This hybrid data-model scheme can improve the diagnosis accuracy and resolution. In addition, using this approach the implementation of SDGs does not require manual choice of the threshold for each node variable. Similarly, Lee et al. (2006) and Ahn et al. (2008) combined SDGs with partial

least square (PLS, a method for constructing regression models). In these two studies, the qualitative statuses of variables in the SDG were determined by the residuals between their measured values and the values predicted by the PLS model. When a fault occurs, the abnormal node status will activate the consistency check to identify the fault origin. Another study attempted to construct the regression model by combining kernel PLS (to deal with data nonlinearity) and support vector regression (Lü & Wang, 2008). SDGs have also been integrated with QTA (Maurya et al., 2007; Gao et al., 2010). In such studies, SDGs first provided a set of possible fault candidates, from which the true fault origin was identified by using the temporal evolution of the process variables.

3.4. Fault tree

Fault tree (FT) is a widely used technique for reliability analysis and fault diagnosis. It is a graphical and hierarchical model that propagates primal (fault origins) and intermediate events to the top event (a hazard). The primal event can include a variety of concerns such as equipment malfunction, component contamination, human factors or protection barriers. The intermediate event, for example, can be either an abnormal symptom or an induced fault. The propagation is performed by using different logic gates (e.g. AND, OR and XOR), thereby offering a more flexible representation of causal knowledge compared to SDGs which predominantly use the OR logic gate. The root cause of an undesired event can be deduced by a top down analysis, whilst the consequences of a basic event can be inferred in a down-to-up manner.

A FT is generally constructed by hypothetical analysis of all possible events that can bring out the predefined top event and the sub-events as well. The construction terminates when the

further analysis is not possible or not necessary. It is worth noting that improper FTs can be produced by incorrect logic or omission of certain events. As a result, the main concern with fault tree analysis (FTA) is that FT development is time-consuming and error-prone. Furthermore, it is difficult to evaluate the quality of the constructed FT.

FTA is usually performed in a qualitative manner to identify possible paths to the occurrence of the top event, which potentially explains the fault propagation pathways. A common approach is to convert the fault tree into the minimal cut sets and rank their consistency. The minimal cut set is a set of primal events necessary for the top event to occur. If the past statistical data is available, the FTA can be extended to a quantitative version. In such case, given real-time data the probability of the top event as well as the contribution from each event can be estimated (Ulerich & Powers, 1988).

Due to the efforts in FT construction, early research focused on the computer-aided FT synthesis. Kavčič & Juričić (1997) developed a software for rapid prototyping of FT in a CAD environment. It offered process components library, mini-trees associated with each component, algorithms for FT synthesis and generating fault symptoms and diagnostic rules, and flexibilities in modifying the automatically generated structures and rules. Another research interest is to synthesise FTs from SDGs. Lapp & Powers (1977) and Ulerich & Powers (1988) used matrixes to represent SDGs, and explicitly included some common faults that can cause node deviations, e.g. control valve stiction and sensor fault. The FT synthesis algorithm assumed the deviation of each node (top event) and searched its causes through the matrixes. Chang et al. (2002) further explored this approach by considering the order of fault symptoms and the partially developed fault symptoms as well. These symptoms formed diagnostic rules,

which were incorporated into a fuzzy inference system. However, this study has not considered the CRs and IRs of control loop variables. Chang & Chang (2003) and Chen & Chang (2006) addressed this issue by introducing transient state to describe these variables.

FT has also been widely used for reliability assessment and identifying hazards. This task requires the probabilities or frequencies of the concerned primal events, or expert experience on this. Recent industrial case studies include reliability assessment of a 50 tons chlorine storage and filling facility (Renjith et al., 2010), deethanizer failure in petrochemical plant (Lavasani et al., 2015), kick control in managed pressure drilling process (Sule et al., 2018). In order to facilitate probabilistic FTA of chemical processes, Khan & Abbasi (1999) and Khan & Abbasi (2000) developed a so-called PROFAT software. Several advanced techniques such as Boolean algebra, cut sets generation, Monte-Carlo simulation and fuzzy probability set have been included in this package for efficient analysis. FTA has also been combined with Hazard and Operability Analysis (HAZOP) studies (Guo & Kang, 2015; Melani et al., 2018). The motivation is that HAZOP study generally cannot provide quantitative assessment results, or graphical representation of fault propagation path. FTAs can serve as a complement to it.

On the other hand, it is important to quantify the overall uncertainties of the fuzzy FTA. For this objective, Purba et al. (2015) used fuzzy probability based FTA and developed rules to propagate epistemic uncertainties from primal events to the top event.

3.5. Bayesian network

Bayesian network (BN) is another commonly used technique in knowledge representation and probabilistic reasoning under uncertainty. It is a directed acyclic graphical model that consists

of qualitative and quantitative parts. The qualitative part has nodes and their links to represent a set of random variables (either continuous or discrete) and their causal relationships. The link leads from a parent node to a child node. A node without any child node is called “leaf node” whilst a parentless node is called “root node”. The marginal probabilities are assigned to the root nodes, and the conditional probabilities are assigned to the others as quantitative information of the network. Observations of some nodes, once available, can be propagated throughout the network to update the knowledge and estimate posterior probabilities on the unobservable nodes according to the Bayes theorem

$$P(X_E|X_O) = \frac{P(X_1, \dots, X_n)}{\sum_O P(X_1, \dots, X_n)} \quad (1)$$

Here X_O and X_E are the observed nodes and the nodes to be estimated, respectively. Such practice is also called inference.

This technique possesses several desired properties. Firstly, it can handle various uncertainties and provide confidence in the diagnosis. Secondly, it can provide robust diagnosis in the presence of missing and unreliable data (Qi et al., 2010). Furthermore, BNs allow to take into account the prior process knowledge, in both qualitative form (e.g. causal dependencies among nodes) and quantitative form (e.g. marginal and conditional probabilities). Therefore, it is increasingly used for fault detection and diagnosis in process industry.

Fault diagnosis with BNs generally includes three main steps: BN structure modelling, BN parameter modelling and BN inference. The structure modelling is to determine the nodes and directed arcs among them, which can be accomplished by process experts, mapping from other qualitative knowledge or models, or learning from historical data. Parameter modelling of BNs

is to assign prior probabilities to the root nodes and conditional probabilities to the other nodes, which are obtained from expert judgement and statistical results of historical or simulated data. Note that some BNs-based diagnosis methods are purely data-driven and will not be discussed in this paper. A more detailed and comprehensive BN introduction can be found in (Cai et al., 2017).

Fault diagnosis with BN was pioneered by (Rojas-Guzman & Kramer, 1993) and further exploited by (Kang & Golay, 1999). In these early studies, BNs perform diagnosis at certain time point without consideration of temporal information, which are also called static BNs. Such temporal dependency is often critical in describing how abnormality evolves and propagates throughout the entire system. Considerable efforts have been made to develop BN methods that can effectively model the temporal information.

Dynamic BN (DBN) is a popular technique for this purpose. It extends the BN by considering multiple copies of the network nodes, connecting them by interslice arcs and assigning additional transition probabilities to each node. The hidden Markov model (HMM) is a simple example of DBN. Weber et al. (2006) developed a DBN that modelled the time degradation of sensors with HMM. DBN for control loop diagnosis was discussed in (Qi & Huang, 2011). They further suggested to model only the temporal dependencies of highly autocorrelated nodes in the DBN, such as to reduce the evidence transition space (thereby reducing the required historical data). Zhang & Dong (2014) proposed a three-time slice DBN that was able to deal with incomplete measurements and non-Gaussian process data. Lerner et al. (2000) discussed the diagnosis of complex systems with both continuous and discrete variables. This allows the modelling of a much wider range of problems.

On the other hand, Arroyo-Figueroa et al. (1998) and Arroyo-Figueroa & Sucar (2005) proposed the so-called temporal nodes Bayesian network (TNBN). It allocated a certain number of time intervals to each node, and determined the conditional distributions of the nodes connected with the potential fault nodes according to their fault propagation speed. Additional novelty of this TNBN is that the nodes represent the state change events instead of the state values in other DBNs, resulting a simplified dynamic system modelling. In order to facilitate the use of TNBN, Arroyo-Figueroa et al. (2000) developed a software package based on it, in which many modules had been included including knowledge base, inference engine, plant operation database and operator interface.

It is worth noting that most above techniques require conditional probability distributions of node states under concerned fault conditions, which is not easy to be satisfied. This limitation was addressed in two DBN based studies (Yu & Rashid, 2013; Amin et al., 2019). They developed novel indexes to quantify the abnormality likelihood at each node and assumed the nodes in true fault propagation pathways should present higher abnormality likelihood. Thus, the fault propagation pathways can be identified by searching for the nodes with high abnormality likelihood throughout the network.

An important technique in BN research is the object-oriented BN (OOBN) (Weidl et al., 2005; Cai et al., 2016). A node in OOBN can be a variable node as in standard BN or instance node which abstracts a network fragment into a single unit. Modelling BNs in this way can improve model reusability and building efficiency, and reduce the complexity of the overall BN for fault diagnosis, and therefore is particularly suitable for modelling large-scale, complex and hierarchical systems.

Another way of rapid BN construction is to convert FTs into BNs (Chiremsel et al., 2016), which provides several benefits: 1) more concise knowledge representation as multiple faults can be incorporated in one BN (Bobbio et al., 2001); 2) ability to handle unreliable and missing readings (Khakzad et al., 2011). A tool has been developed in (Montani et al., 2006) for automatically converting dynamic fault trees into DBNs.

Recently, BNs have been integrated with a couple of data-driven methods. Yu et al. (2015) discussed a hybrid method that combines independent component analysis (ICA) and BN. The ICA extracted independent components from NOC data (by minimising the difference between their joint distribution and marginal distributions), and established monitoring statistics based on them for fault detection. In abnormal situations, the statistics exceeded a predefined threshold, and the variables (referred as faulty variables) that contributed the most to this were identified and used as evidence to update the BN. Similarly, Gharahbagheri et al. (2017) used Kernel PCA to identify the faulty variables. In addition, to avoid subjective expert judgment in BN structure modelling, this study used Granger causality and transfer entropy to discover the causal relationships in the network. Granger causality detects causality by testing whether the past values of one variable (denoted as X) can help predict future values of another variable (denoted as Y). Transfer entropy achieves this by measuring the amount of information transferred from X to Y , which is determined by the reduced uncertainty in prediction of Y by knowing the past values of X . It is worth noting that transfer entropy can be applied to non-linear data, whilst the original Granger causality cannot.

3.6. Topology and ontology models

When a process is designed, the topology, describing the physical or signal connections between process units, is usually clearly known. It can be derived from readily available technical documents like piping and instrumentation diagrams (P&IDs) or process flow diagrams. Topology alone cannot perform diagnostics, and it is often used in combination with data-driven techniques to improve the reliability and resolution of their diagnosis.

Thambirajah et al. (2007) and Thambirajah et al. (2009) made some important contributions in use of topology: 1) representing plant schematic through Extensible Mark-up Language (XML) according to the Computer Aided Engineering Exchange (CAEX) schema; 2) parsing XML file to derive connectivity matrix. The extracted connectivity was used to examine the candidates of fault propagation pathways generated from transfer entropy. Yim et al. (2006) developed a corresponding prototype software which included parsers for XML files interpretation, inference engine and graphical user interface. Similarly, Landman et al. (2014) integrated topology with Granger causality to identify oscillations propagation paths in control loops.

One limitation of the above methods is that they cannot distinguish between direct and indirect causality. For example, knowing that there is causality and an intermediate variable Z between X and Y , the above methods cannot determine whether the causality is directly from X to Y or from X through Z to Y . Landman & Jämsä-Jounela (2016) overcame this limitation by introducing a direct transfer entropy developed in (Duan et al., 2013). The key is to calculate the amount of information that can be transferred from X to Y after losing the information transferred from Z to Y . If it is greater than 0, it means that there is a direct causal relationship between X and Y . A prototype tool to integrate, visualise and analyse different connectivity, including electrical, logical and process connectivity, was presented in (Dorantes et al., 2015).

Topology has been integrated with process models in (Gil et al., 2011). In this study, the process models of basic units were stored in a library, and when units were identified in the topology, corresponding models were extracted from the library and connected according to the topology model to generate an enhanced causal model.

Ontology modelling is another effective method for knowledge representation. It represents process knowledge in different details by using general abstract concepts and their instantiations and specialisations. It specifies attributes and characteristics to each object, relationships between them and some axioms such as mass and energy conservation (Yang et al., 2009; Yang et al., 2008). As an explicit organisation of deep process knowledge, ontology can facilitate knowledge sharing, reuse, and query, as well as integration and consolidation of heterogeneous knowledge.

Some multi-purpose and reusable ontologies have been developed for process industry, e.g. OntoCAPE (Morbach et al., 2007), and they have also been extended to specific applications like process design and supervision. For example, on basis of OntoCAPE, Natarajan et al. (2012) developed a specific ontology named OntoSafe. It additionally introduced concepts, classes and relations to represent states and conditions of a plant, making it possible to explicitly represent the dependencies between process descriptors (e.g. control system, instrumentation, flowsheet) and various elements of a supervision system. A comprehensive judgment can be made on the state and conditions of the process. Natarajan & Srinivasan (2014) further extended it to an architecture consisting of multiple supervision methods which can be applied according to the identified process changes.

Elhaddad et al. (2013) discussed a framework for supervision of the shutdown process of petroleum plants. Some work has used ontology for risk assessment such as automatic HAZOP (Zhao et al., 2005; Zhao et al., 2009). Modelling of fault-relevant knowledge is the key to ontology application in process supervision.

4. Real process applications

Most of the studies presented above were illustrated on some well-established simulation benchmarks and satisfactory results have been reported. There is a strong motivation to understand their performance in real processes. However, there are very limited practical applications of model- and knowledge-based FDD methods in process industry and they are summarised in Tables 1 and 2 below. The tables are comprehensive and self-explanatory in nature, covering studied processes, their scales, methods used, faults, and reported performance, and therefore do not need any further elucidation.

They successfully performed multiple FDD tasks including fault detection, diagnosis, and, in some studies, fault magnitude estimation, for different process scales. Sensor faults, actuator faults and process faults have been covered in these studies. In the research on laboratory- and pilot- scale processes, faults are usually physically simulated. In the research on industrial-scale processes, the FDD methods have either been implemented online in the practical process or tested on historical fault data of the process. Analytical-model-based approaches have less applications in industrial scale processes, compared with knowledge-based methods.

Table 1 Real process applications of analytical-model-based methods

Process	Method	Scale	Fault	Performance	Ref
Three tank system	A bank of UIOs	Laboratory	Tank leak and clogging	Fast detection and diagnosis	(Zolghadri et al., 1996)
	Observer	Laboratory	Regulating valves bias	Fast detection, diagnosis and fault signal reconstruction	(Orani et al., 2010)
	Nonlinear observer	Laboratory	Tank leak; level sensor bias	Fast detection, diagnosis, and magnitude estimation	(Fang et al., 2007)
	Observer	Laboratory	Tank leak; level sensor bias	Fast detection	(Ding et al., 2019)
	A bank of UIOs	Laboratory	Level sensor bias, failure; pump degrading and failure	Detection (delay within 16s), correct diagnosis and magnitude estimation	(Theilliol et al., 2002)
Semi-batch reaction process	Nonlinear observer	Laboratory	Flow valve fault; heating power change	Fast detection and diagnosis	(Hammouri et al., 2002)
	EKF	Laboratory	Step changes in flow rate and in cooling temperature	70s and 170s delay in detection	(Chetouani, 2004)
	A bank of EFKs	Laboratory	Step changes in cooling temperature	320s delay in detection, success in magnitude estimation	(Chetouani, 2008)
Level tank	PF	Laboratory	Valve faults	Detection rate higher than 90%	(Morales-Menéndez et al., 2003)
Continuous stirred tank heater	Observer, T-S fuzzy model	Laboratory	Temperature sensor offset; valve close	Fast detection	(Li et al., 2016)
Air heater	EKF	Laboratory	Faults of heater and fan actuators; their combination; unknown fault	Correct detection rate: 84.2%; effective diagnosis	(Castillo et al., 2014)
Three tank system	A bank of nonlinear observers	Pilot	Flowrate sensor offset	Detection and diagnosis within 35s	(Aguirre & Pereira, 1998)
			Flow valve failure; pipe leak, blockage; pump failure	Detection with delays ranging from 25s to 284s	
Heat exchange process	A bank of PRs	Pilot	Sensor bias and drift	Fast detection and diagnosis	(Simonart & Kinnaert, 1994)

	PR	Pilot	9 faults including: valve stick, offset, and characteristic deviation; sensor offset and gain deviation	High detection rate and no false alarm for most faults	(Höfling et al., 1995)
	PR	Pilot	Temperature sensor offsets	Fast detection, diagnosis and magnitude estimation	(Peng et al., 1997)
Waster water reactor	A bank of PRs	Industrial	Sensor bias, drift, failure and precise degradation	Detection within 20s, success in diagnosis and magnitude estimation	(Qin & Li, 2001)
Imperial smelting furnace	PR	Industrial	Two fault scenarios	Success in detection	(Jiang et al., 2008)
Heating furnace; debutaniser and reactive distillation	A bank of UIOs	Industrial	Sensor bias; drop in feed flow; condenser fault	Detection and diagnosis within 15 min	(Schubert et al., 2011)
Gas turbine	PR	Industrial	Vibration fault	Fast detect and diagnosis	(Hafaifa et al., 2015)
Drying section of an industrial board machine	PR	Industrial	Valve blockage; steam flow rate sensor	Success in detection and diagnosis	(Zakharov et al., 2013)
Activated sludge and clarification processes of a wastewater treatment plants	PE	Industrial	Insufficient oxygen in different aeration tanks	Detection within 10 min	(Fuente et al., 1996)
120MW power plant	A bank of KFs	Industrial	Bias and drift of input and output sensors	Identified the minimum detectable fault magnitudes	(Simani, 1999)

Table 2 Real process applications of deep-knowledge-based methods

Process	Method	Scale	Fault	Performance	Ref
Two tank mixing process	ES; structural decomposition	Laboratory	Control valve fault; temperature and level sensors faults; hand valves blocked; their combinations	100% diagnosis rate for single fault; in all multiple faults scenarios, only one component fault can be identified	(Zhang & Roberts, 1991)
	QS	Laboratory	Control valve fault; temperature and level sensor faults; hand valves blocked	85% diagnosis rate in average	(Zhang et al., 1990)
	QS; threshold self-learning	Laboratory	Temperature and level sensor faults; flow rate control vale faults	100% diagnosis rate	(Zhang et al., 1991)

Tank-pipeline system	SDG; fault revealing order	Laboratory	Pipeline blockage; tank leak; valve misoperation	Fast diagnosis (less than 1s) with high resolution	(Shiozaki et al., 1989)
	SDG; fault revealing order	Laboratory	Pipeline blockage, tank leak	Blockage diagnosis gave unique correct cause; leak diagnosis gave two cause candidates including the correct one.	(Tsuge et al., 2000)
Wastewater treatment plant	ES; if-then rules	Pilot	Hydraulic overload; organic overload	Success in process states estimation, root cause diagnosis, future trend prediction and action recommendation	(Puñal et al., 2002)
Lactic acid fermentation process	Fuzzy ES; if-then rules	Pilot	Faults in pre-culture quantity and time	Satisfactory performance based on online experiments	(Nakajima et al., 1994)
Distillation column	BN; ES	Pilot	Cooling water failure; full open stuck reflux valve	Successful diagnosis with high confidence	(Leung & Romagnoli, 2000)
Water-steam cycle of thermal power plants	ES; if-then rules	Industrial	Not available	Enabled operators to take quick decisions and corrective actions in one year's implementation	(Addanki & Sethuraman, 1992)
Subsea production system in oil field	OOBN	Industrial	Production flow loop leak; production wing valve failure; hydraulic power unit control valve and fill pump failure; annulus flow loop leak; annulus wing valve failure; electrical power unit line coupler failure	Successful diagnosis of multiple simultaneous faults with high confidence	(Cai et al., 2016)
Steam boiler process	SDG; ES	Industrial	Leak in the secondary superheater	Fast and correct diagnosis	(Lee et al., 1997)
Shaft-furnace roasting process	ES; if-then rules	Industrial	Working situation faults: flames escaping from the combustion chamber; flames reach out of the top of the furnace; iron ore sticks inside the furnace; underdeoxidization; overdeoxidization	Fault frequency reduced by more than 50%; production rate increased by 0.7 t/h; magnetic tube recovery rate increased by 2%; equipment operation rate increased by 2.98%	(Chai et al., 2007)
Lubricating oil refining process	ES; 632 pieces of if-then rules	Industrial	Not available	Successful and rapid diagnosis and suggestion of corrective actions in one year's implementation; reduced process instability and economic loss	(Qian et al., 2008)
Leaching process	ES; if-then rules;	Industrial	Not available	Diagnosis rate higher than 90%; much	(Wu et al., 2002)

of a nonferrous metals smeltery	fault probabilities			reduced frequency of serious faults	
Hot strip mill process	BN; PLS	Industrial	Malfunction of gap control loop; fault of cooling valve between rolling stands	Effective detection; successful diagnosis and propagation path identification	(Ma et al., 2018)
Gas fraction unit	Probabilistic SDG	Industrial	Feed Pump Shut Down; reflux tank liquid level sensor bias	Correct diagnosis with high resolution and correct propagation pathway identification	(Lü et al., 2011)
Fluid catalytic cracking unit	ES; if-then rules	Industrial	Not available	Successful and rapid diagnosis and suggestion of corrective actions in two years' implementation	(Qian et al., 2003; Qian et al., 2005)
Ethylene cracker	SDG; PCA	Industrial	Naphtha feed ratios deviation; non-uniform fuel gas feed in nozzles	True positive rate over 98%, false alarm rate 1.56% and correct root cause diagnosis with test on three months' industrial data	(Han et al., 2018)
Drying section of paperboard machine	Topology	Industrial	Valve stiction in a control loop	Successful diagnosis and oscillation propagation pathway identification	(Landman et al., 2014; Landman & Jämsä-Jounela, 2016)
Anaerobic digester	Modular fuzzy ES; if-then rules	Industrial	Toxicity; acidogenesis; underload; organic overload; hydraulic overload	Successful fault diagnosis in 220 days' test despite of missing alarms for 40 days	(Lardon et al., 2005)

5. Challenges and opportunities

As reviewed above, tremendous efforts and many insightful discussions have been made on developing model- and knowledge-based FDD methods. Despite the long-term widespread attention and remarkable achievements, satisfactory industrial applications are not many (Shu et al., 2016). Many factors make this exercise a challenge, including: 1) various possible fault types, 2) fault data scarcity, 3) insufficient understanding of complex process dynamics (especially in faulty conditions), 4) imperfect measurement, and 5) variable operating policies. FDD represents a complex task that requires a better understanding and modelling of process mechanism and causalities (especially under faulty conditions) and efficient use of online and historical data. Rapid advances in data analytics should be leveraged. Assumptions behind current methods need to be relaxed and their reliability and effectiveness need to be improved. On the other hand, given the fact that most methods were demonstrated on simulation benchmarks, their reliability in industrial applications may be questioned. It would be beneficial to conduct more collaborative research between industry and academia, which could provide more specific and practical cases for study.

In FDD practice, there is a long way to go and are a lot of content deserving further studying. Some issues or directions that we consider interesting and may worth more efforts in future are listed below.

5.1. Fault propagation path identification

Faults can occur locally somewhere in a plant and then quickly have their effect propagated across the entire plant. Identifying fault propagation paths plays an important role in performing reliable diagnosis. A diagnosis method providing propagation paths would also be more

acceptable to operators as they may want to understand how the fault evolved from a basic event and its possible consequences before taking any corrective actions.

Modelling and discovering causal information are essential for this task. SDGs and BNs (Yu & Rashid, 2013; Amin et al., 2019) are effective methods for modelling priori causal knowledge. On the other hand, there has also been some techniques for discovering causality directly from data, such as granger causality (Landman et al., 2014) and transfer entropy (Landman & Jämsä-Jounela, 2016), which could be leveraged to discover unknown causalities and complement SDGs and BNs. Such integrations of causal relationship modelling and causal relationship discovery techniques may deserve more investigations.

In addition, sound signals, vibration signals, images, and spectroscopy are increasingly available in process industry (Severson et al., 2016). Incorporating them into causal models may provide new opportunities in identifying fault propagation paths.

Notice that many fault propagation studies are reported on small scale problems. However, modern industrial processes are usually large systems with many monitored variables, which can result in low resolution and efficiency in the diagnosis. A possible solution is hierarchical modelling and system decomposition strategies (Zhang & Roberts, 1991). For example, Finch & Kramer (1988) proposed a structural decomposition strategy which can quickly narrow the focus of diagnosis, thereby avoiding unnecessary details in the early stages of diagnosis. These strategies also deserve further development.

5.2. Hybrid methods

Each method has its strength and weakness. A comparative study of various methods can be found in (Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003). Existing studies show that the desired characteristics of a diagnostic system can hardly be satisfied by a single method or single source of knowledge. The main advantage of a hybrid model is that it not only captures the strengths of each method, but also leverages different knowledge sources, resulting in improved efficiencies and accuracies. For example, reconstruction-based PCA can be used to quickly isolate faulty variables to reduce the focus, and then informative and expensive methods such as SDGs can be used in the focus region (He et al., 2012). ICA based monitoring statistics can be used to identify faulty variables as evidence for updating BN (Yu et al., 2015). An well-known hybrid diagnostic system is Dkit developed by (Mylaraswamy & Venkatasubramanian, 1997). It included SDGs, state observers, QTA, expert rules and neural networks and they were developed for common units in process industry including CSTR, distillation column and pulp. Tidriri et al. (2016) gave a recent review on approaches combining data analytics and first-principle knowledge.

We believe that hybrid systems still remain an important research direction and there are more potentials to explore. Since each method relies on its own assumptions, reliable integration of various methods is not straightforward (Tidriri et al., 2016). It is important to find specific applications and scenarios (in terms of the available knowledge and concerned process characteristics), in which each hybrid strategy has advantageous over other alternatives. On the other hand, current research usually uses a single hybrid strategy for the entire system. Deployment of different individual hybrid methods in different subsystems may provide better performance, which however has not been extensively explored. In addition, the decision fusion

strategy may require further research to reliably combine solutions from individual methods and resolve the contradictions between them (Tidriri et al., 2016).

5.3. Enrichment of fault samples

Generally, rich fault samples are critical to developing a reliable FDD model (Shu et al., 2016). However, process fault rarely happens and the research suffers from a lack of fault samples. In addition, even though fault data are available, most of them are unlabelled. Fault data scarcity is an important obstacle to research on fault diagnosis and needs to be addressed in future research.

The recent success of AlphaGo Zero taught some lessons. It was successfully trained using only computer simulated data, without utilisation of any historical data. This naturally arises an interesting question: can we utilise synthetic fault data to help train a diagnostic model? It is key to remember that process monitoring is different from the Go game: the Go game has clear rules, perfect model and measurements, which unfortunately is not the case in our domain, as discussed by (Qin & Chiang, 2019) in detail. How to effectively use computer-simulated fault data and handle the uncertainties arose from mismatches between simulation and real processes can be an interesting problem.

Data scarcity can also be addressed by collecting and adapting fault samples from different plant sites. The idea presented in (Shu et al., 2016) that various process plants share their process data (especially fault data) in a cloud computing environment is constructive. However, the processes where the data comes from can be operated under different conditions, and even the processes themselves can be different due to retrofit designs. How to effectively utilise data

from different plant sites and processes in presence of domain discrepancies can be an important research direction. Wu & Zhao (2020) made an early attempt by using a transfer learning strategy.

5.4. Rapid model development

For complex systems (typical case in process industry), enormous amounts of time, efforts and money can be involved in developing knowledge-based models, e.g. SDGs, BNs, FTs. The complexity and integration degree of process systems continue to increase, posing challenges for manual development of such models. For example, the construction of SDGs, typically by experts with specialised knowledge, will require the identification of key variables, their connectivity and causal relationships, which can be time consuming and error prone. On the other hand, process plants seldom keep invariant and the models can change with operating strategies (under NOC and different abnormal conditions) and retrofit designs. Hence, there exists a strong motivation for developing tools for fast model development.

Rapid modelling can be achieved by reusing models from similar processes. Note that many principles and knowledge models (e.g. SDGs, FTs, BNs) of systems or subsystems can often remain the same under different operating conditions or processes. Based on this idea, some recent studies (Yuan et al., 2019; Li et al., 2019) attempted to adapt BNs from similar process. Other notable contributions include OOBNs (Weidl et al., 2005; Cai et al., 2016) and computer-aided FT synthesis prototype (Kavčič & Juričić, 1997) which aimed to improve model reusability and building efficiency. More efforts may be needed along this direction, and development of toolboxes or software can be useful for promoting their practical applications.

6. Conclusion

This review paper first provides an overview of the process fault diagnosis problem and discusses the benefits of using fundamental process knowledge. Then, many existing approaches based on process models and knowledge are reviewed in terms of their basic ideas, weaknesses, strengths and recent progresses. Although the current achievements are remarkable, most studies are evaluated in simulation benchmarks. In order to understand their performance in real processes, the paper also summarises studies tested on real processes, including their processes, scales, concerned faults and performance. Finally, some perspectives on important problems and future directions are discussed and highlighted for future research at the end of the paper.

Acknowledgments

This work has been partially supported by the UK EPSRC (EP/R001588/1), BBSRC (BB/S020896/1), and the Unilever-IPE-Surrey collaborative doctoral training programme.

Reference

- Addanki, N. V. L., & Sethuraman, R. (1992). On Line Diagnosis of Water Chemistry in Thermal Power Plant. *IFAC Proceedings Volumes*, 25(4), 299–302. [https://doi.org/10.1016/s1474-6670\(17\)50258-x](https://doi.org/10.1016/s1474-6670(17)50258-x)
- Aguirre, L. A., & Pereira, M. F. S. (1998). A modified observer scheme for fault detection and isolation applied to a poorly observed process with integration. *Journal of Process Control*, 8(1), 47–56. [https://doi.org/10.1016/S0959-1524\(97\)00026-7](https://doi.org/10.1016/S0959-1524(97)00026-7)
- Ahn, S. J., Lee, C. J., Jung, Y., Han, C., Yoon, E. S., & Lee, G. (2008). Fault diagnosis of the multi-stage flash desalination process based on signed digraph and dynamic partial least square. *Desalination*, 228(1–3), 68–83. <https://doi.org/10.1016/j.desal.2007.08.008>
- Alrowaie, F., Gopaluni, R. B., & Kwok, K. E. (2012). Fault detection and isolation in stochastic non-linear state-space models using particle filters. *Control Engineering Practice*, 20(10), 1016–1032. <https://doi.org/10.1016/j.conengprac.2012.05.008>
- Alwi, H., & Edwards, C. (2014). Robust fault reconstruction for linear parameter varying systems using sliding mode observers. *International Journal of Robust and Nonlinear Control*, 24, 1947–1968. <https://doi.org/10.1002/rnc>

- Amin, M. T., Khan, F., & Imtiaz, S. (2019). Fault detection and pathway analysis using a dynamic Bayesian network. *Chemical Engineering Science*, 195, 777–790. <https://doi.org/10.1016/j.ces.2018.10.024>
- Arroyo-Figueroa, G., Alvarez, Y., & Sucar, L. E. (2000). SEDRET-an intelligent system for the diagnosis and prediction of events in power plants. *Expert Systems with Applications*, 18(2), 75–86. [https://doi.org/10.1016/S0957-4174\(99\)00054-8](https://doi.org/10.1016/S0957-4174(99)00054-8)
- Arroyo-Figueroa, Gustavo, & Sucar, L. E. (2005). Temporal Bayesian Network of Events for Diagnosis and Prediction in Dynamic Domains. *Applied Intelligence*, 23(2), 77–86. <https://doi.org/10.1007/s10489-005-3413-x>
- Arroyo-Figueroa, Gustavo, Sucar, L. E., & Villavicencio, A. (1998). Probabilistic temporal reasoning and its application to fossil power plant operation. *Expert Systems with Applications*, 15(3–4), 317–324. [https://doi.org/10.1016/S0957-4174\(98\)00038-4](https://doi.org/10.1016/S0957-4174(98)00038-4)
- Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/nongaussian bayesian tracking. *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, 50(2), 174–188. <https://doi.org/10.1109/9780470544198.ch73>
- Bahmanpour, S., Bashooki, M., & Refan, M. H. (2007). State Estimation and Fault Diagnosis of Industrial Process by Using of Particle Filters. *The 6th WSEAS International Conference on Signal Processing, Robotics and Automation*, 208–213.
- Baskioti, C., Raymond, J., & Rault, A. (1979). Parameter identification and discriminant analysis for jet engine mechanical state diagnosis. *The 18th IEEE Conference on Decision and Control/Decision and Control*, 648–650.
- Bobbio, A., Portinale, L., Minichino, M., & Ciancamerla, E. (2001). Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. *Reliability Engineering and System Safety*, 71, 249–260.
- Bøgh, S. (1995). Multiple hypothesis-testing approach to FDI for the industrial actuator benchmark. *Control Engineering Practice*, 3(12), 1763–1768.
- Brown, R. G., & Hwang, P. Y. C. (1992). *Introduction to Random signals and Applied Kalman Filtering* (Third Edit). <https://doi.org/10.1017/CBO9781107415324.004>
- Cai, B., Huang, L., & Xie, M. (2017). Bayesian Networks in Fault Diagnosis. *IEEE Transactions on Industrial Informatics*, 13(5), 2227–2240. <https://doi.org/10.1109/TII.2017.2695583>
- Cai, B., Liu, H., & Xie, M. (2016). A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks. *Mechanical Systems and Signal Processing*, 80, 31–44. <https://doi.org/10.1016/j.ymssp.2016.04.019>
- Castillo, I., Edgar, T. F., & Dunia, R. (2014). Nonlinear detection and isolation of multiple faults using residuals modeling. *Industrial and Engineering Chemistry Research*, 53(13), 5217–5233. <https://doi.org/10.1021/ie4016655>
- Chai, T., Wu, F., Ding, J., & Su, C. Y. (2007). Intelligent work-situation fault diagnosis and fault-tolerant system for the shaft-furnace roasting process. *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering*, 221(6), 843–855. <https://doi.org/10.1243/09596518JSCE364>
- Chan, C. W. (2005). An expert decision support system for monitoring and diagnosis of petroleum production and separation processes. *Expert Systems with Applications*, 29(1),

- 131–143. <https://doi.org/10.1016/j.eswa.2005.01.009>
- Chang, S.-Y., & Chang, C.-T. (2003). A fuzzy-logic based fault diagnosis strategy for process control loops. *Chemical Engineering Science*, 58(15), 3395–3411. [https://doi.org/10.1016/S0009-2509\(03\)00218-5](https://doi.org/10.1016/S0009-2509(03)00218-5)
- Chang, S.-Y., Lin, C.-R., & Chang, C.-T. (2002). A fuzzy diagnosis approach using dynamic fault trees. *Chemical Engineering Science*, 57(15), 2971–2985. [https://doi.org/10.1016/S0009-2509\(02\)00178-1](https://doi.org/10.1016/S0009-2509(02)00178-1)
- Chen, J. Y., & Chang, C.-T. (2006). Fuzzy diagnosis method for control systems with coupled feed forward and feedback loops. *Chemical Engineering Science*, 61(10), 3105–3128. <https://doi.org/10.1016/j.ces.2005.11.062>
- Chen, J. Y., & Chang, C.-T. (2009). Development of fault diagnosis strategies based on qualitative predictions of symptom evolution behaviors. *Journal of Process Control*, 19(5), 842–858. <https://doi.org/10.1016/j.jprocont.2008.11.006>
- Chen, T., Morris, J., & Martin, E. (2005). Particle filters for state and parameter estimation in batch processes. *Journal of Process Control*, 15, 665–673. <https://doi.org/10.1016/j.jprocont.2005.01.001>
- Chen, T., Morris, J., & Martin, E. (2008). Dynamic data rectification using particle filters. *Computers and Chemical Engineering*, 32(3), 451–462. <https://doi.org/10.1016/j.compchemeng.2007.03.012>
- Chester, D., Lamb, D., Dhurjati, P. (1984). Rule-based computer alarm analysis in chemical process plants. *The 7th Annual Conference on Computer Technology*, 22–29.
- Chetouani, Y. (2008). Design of a multi-model observer-based estimator for fault detection and isolation (FDI) strategy: Application to a chemical reactor. *Brazilian Journal of Chemical Engineering*, 25(4), 777–788. <https://doi.org/10.1590/S0104-66322008000400015>
- Chetouani, Yahya. (2004). Fault detection by using the innovation signal: Application to an exothermic reaction. *Chemical Engineering and Processing-Process Intensification*, 43(12), 1579–1585. <https://doi.org/10.1016/j.cep.2004.02.002>
- Chiremsel, Z., Said, R. N., & Chiremsel, R. (2016). Probabilistic Fault Diagnosis of Safety Instrumented Systems based on Fault Tree Analysis and Bayesian Network. *Journal of Failure Analysis and Prevention*, 16(5), 747–760. <https://doi.org/10.1007/s11668-016-0140-z>
- Chow, E. Y., & Willsky, A. S. (1984). Analytical Redundancy and the Design of Robust Failure Detection Systems. *IEEE Transactions on Automatic Control*, 29(7), 603–614. Retrieved from <http://dspace.mit.edu/handle/1721.1/2828>
- Dash, S., & Venkatasubramanian, V. (2000). Challenges in the industrial applications of fault diagnostic systems. *Computers and Chemical Engineering*, 24(2–7), 785–791. [https://doi.org/10.1016/S0098-1354\(00\)00374-4](https://doi.org/10.1016/S0098-1354(00)00374-4)
- Desai, M., & Ray, A. (1981). A fault detection and isolation methodology. *IEEE Conference on Decision and Control Including the Symposium on Adaptive Processes*, 1363–1369.
- Desai, M., & Ray, A. (1984). A Fault Detection and Isolation Methodology Theory and Application. *American Control Conference*, 262–270.
- Dhaliwal, J. S., & Benbasat, I. (1996). The Use and Effects of Knowledge-based System Explanations: Theoretical Foundations and a Framework for Empirical Evaluation. *Information Systems Research*, 7(3), 342–362. <https://doi.org/10.1287/isre.7.3.342>

- Díaz, C. A., Echevarría, L. C., Prieto-Moreno, A., Neto, A. J. S., & Llanes-Santiago, O. (2016). A model-based fault diagnosis in a nonlinear bioreactor using an inverse problem approach and evolutionary algorithms. *Chemical Engineering Research and Design*, 114, 18–29. <https://doi.org/10.1016/j.cherd.2016.08.005>
- Ding, S.X., Ding, E. L., & Jeinsch, T. (1999). An Approach to Analysis and Design of Observer and Parity Relation Based FDI Systems. *IFAC Proceedings Volumes*, 32(2), 7718–7723. [https://doi.org/10.1016/s1474-6670\(17\)57317-6](https://doi.org/10.1016/s1474-6670(17)57317-6)
- Ding, Steven X., Li, L., & Krüger, M. (2019). Application of randomized algorithms to assessment and design of observer-based fault detection systems. *Automatica*, 107, 175–182. <https://doi.org/10.1016/j.automatica.2019.05.037>
- Doraiswami, R., Diduch, C. P., & Tang, J. (2010). A New Diagnostic Model for Identifying Parametric Faults. *IEEE Transactions on Control Systems Technology*, 18(3), 533–544. <https://doi.org/10.3182/20080706-5-KR-1001.01246>
- Dorantes, D., Graven, T., & Thornhill, N. F. (2015). Linking process, electrical and logical connectivity for supported fault diagnosis. *Computer Aided Chemical Engineering*, 37, 965–970. <https://doi.org/10.1016/B978-0-444-63577-8.50006-1>
- Doucet, A., Godsill, S., & Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3), 197–208. <https://doi.org/10.1023/A:1008935410038>
- Du, M., Scott, J., & Mhaskar, P. (2013). Actuator and sensor fault isolation of nonlinear process systems. *Chemical Engineering Science*, 104, 294–303. <https://doi.org/10.1016/j.ces.2013.08.009>
- Duan, P., Yang, F., Chen, T., & Shah, S. L. (2013). Direct Causality Detection via the Transfer Entropy Approach. *IEEE Transactions on Control Systems Technology*, 21(6), 2052–2066. <https://doi.org/10.1109/TCST.2012.2233476>
- Edwards, C., Spurgeon, S. K., & Patton, R. J. (2000). Sliding mode observers for fault detection and isolation. *Automatica*, 36(4), 541–553. [https://doi.org/10.1016/S0005-1098\(99\)00177-6](https://doi.org/10.1016/S0005-1098(99)00177-6)
- Elhdad, R., Chilamkurti, N., & Torabi, T. (2013). An ontology-based framework for process monitoring and maintenance in petroleum plant. *Journal of Loss Prevention in the Process Industries*, 26(1), 104–116. <https://doi.org/10.1016/j.jlp.2012.10.001>
- Fang, M., Tian, Y., & Guo, L. (2007). Fault diagnosis of nonlinear system based on generalized observer. *Applied Mathematics and Computation*, 185(2), 1131–1137. <https://doi.org/10.1016/j.amc.2006.07.034>
- Feng, E., Yang, H., & Rao, M. (1998). Fuzzy expert system for real-time process condition monitoring and incident prevention. *Expert Systems with Applications*, 15(3–4), 383–390. [https://doi.org/10.1016/s0957-4174\(98\)00053-0](https://doi.org/10.1016/s0957-4174(98)00053-0)
- Finch, F. E., & Kramer, M. A. (1988). Narrowing diagnostic focus using functional decomposition. *AIChE Journal*, 34(1), 25–36. <https://doi.org/10.1002/aic.690340105>
- Frank, P. (1990). Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge-based Redundancy A Survey and Some New Results. *Automatica*, 26(3), 459–474. <https://doi.org/10.1109/ROBIO.2011.6181275>
- Frank, P., & Wunnenberg, J. (1989). Robust fault diagnosis using unknown input observer schemes. *Fault Diagnosis in Dynamic Systems*.

- Fuente, M. J., Vega, P., Zarrop, M., & Poch, M. (1996). Fault detection in a real wastewater plant using parameter-estimation techniques. *Control Engineering Practice*, 4(8), 1089–1098.
- Gao, D., Wu, C., Zhang, B., & Ma, X. (2010). Signed Directed Graph and Qualitative Trend Analysis Based Fault Diagnosis in Chemical Industry. *Chinese Journal of Chemical Engineering*, 18(2), 265–276. [https://doi.org/10.1016/S1004-9541\(08\)60352-3](https://doi.org/10.1016/S1004-9541(08)60352-3)
- Gao, Z., Breikin, T., & Wang, H. (2008). Discrete-time proportional and integral observer and observer-based controller for systems with both unknown input and output disturbances. *Optimal Control Applications and Methods*, 29, 171–189. <https://doi.org/10.1002/oca>
- Ge, Z., Song, Z., & Gao, F. (2013). Review of recent research on data-based process monitoring. *Industrial and Engineering Chemistry Research*, 52(10), 3543–3562. <https://doi.org/10.1021/ie302069q>
- Gertler, J. J. (1988). Survey of Model-Based Failure Detection and Isolation in Complex Plants. *IEEE Control Systems Magazine*, 8(6), 3–11. <https://doi.org/10.1109/37.9163>
- Gertler, J. J., & Monajemy, R. (1995). Generating Directional Residuals with Dynamic Parity Relations. *Automatica*, 31(4), 627–635. [https://doi.org/10.1016/0005-1098\(95\)98494-Q](https://doi.org/10.1016/0005-1098(95)98494-Q)
- Gertler, J., & Singer, D. (1990). A New Structural Framework for Parity Equation-based Failure Detection and Isolation. *Automatica*, 26(2), 381–388. [https://doi.org/10.1016/0005-1098\(90\)90133-3](https://doi.org/10.1016/0005-1098(90)90133-3)
- Gertler, J., & Yin, K. (1996). Statistical Decision Making for Dynamic Parity Relations. *IFAC Proceedings Volumes*, 29(1), 6337–6342. [https://doi.org/10.1016/s1474-6670\(17\)58697-8](https://doi.org/10.1016/s1474-6670(17)58697-8)
- Gharahbagheri, H., Imtiaz, S. A., & Khan, F. (2017). Root Cause Diagnosis of Process Fault Using KPCA and Bayesian Network. *Industrial and Engineering Chemistry Research*, 56(8), 2054–2070. <https://doi.org/10.1021/acs.iecr.6b01916>
- Gil, G. J. D. G., Alabi, D. B., Iyun, O. E., & Thornhill, N. F. (2011). Merging Process Models and Plant Topology. *The 4th International Symposium on Advanced Control of Industrial Processes*, 15–21. IEEE.
- Guo, L., & Kang, J. (2015). An extended HAZOP analysis approach with dynamic fault tree. *Journal of Loss Prevention in the Process Industries*, 38, 224–232. <https://doi.org/10.1016/j.jlp.2015.10.003>
- Gustafsson, F. (2002). Stochastic Fault Diagnosability in Parity Spaces. *IFAC Proceedings Volumes*, 35(1), 41–46. <https://doi.org/10.3182/20020721-6-ES-1901.00738>
- Hafaifa, A., Guemana, M., & Daoudi, A. (2015). Vibrations supervision in gas turbine based on parity space approach to increasing efficiency. *Journal of Vibration and Control*, 21(8), 1622–1632. <https://doi.org/10.1177/1077546313499927>
- Hammouri, H., Kabore, P., Othman, S., & Biston, J. (2002). Failure diagnosis and nonlinear observer. Application to a hydraulic process. *Journal of the Franklin Institute*, 339(4–5), 455–478. [https://doi.org/10.1016/S0016-0032\(02\)00027-3](https://doi.org/10.1016/S0016-0032(02)00027-3)
- Han, C., Shih, R., & Lee, L. (1994). Quantifying Signed Directed Graphs with the Fuzzy Set for Fault Diagnosis Resolution Improvement. *Industrial and Engineering Chemistry Research*, 33(8), 1943–1954. <https://doi.org/10.1021/ie00032a008>
- Han, X., Tian, S., Romagnoli, J. A., Li, H., & Sun, W. (2018). PCA-SDG Based Process Monitoring and Fault Diagnosis: Application to an Industrial Pyrolysis Furnace. *IFAC-*

- PapersOnLine*, 51(18), 482–487. <https://doi.org/10.1016/j.ifacol.2018.09.378>
- He, B., Chen, T., & Yang, X. (2014). Root cause analysis in multivariate statistical process monitoring: Integrating reconstruction-based multivariate contribution analysis with fuzzy-signed directed graphs. *Computers and Chemical Engineering*, 64, 167–177. <https://doi.org/10.1016/j.compchemeng.2014.02.014>
- He, B., Yang, X., Chen, T., & Zhang, J. (2012). Reconstruction-based multivariate contribution analysis for fault isolation: A branch and bound approach. *Journal of Process Control*, 22, 1228–1236. <https://doi.org/10.1016/j.jprocont.2012.05.010>
- Herbert, M. R., & Williams, G. H. (1987). An initial evaluation of the detection and diagnosis of power plant faults using a deep knowledge representation of physical behaviour. *Expert Systems*, 4(2), 90–99.
- Höfling, T., Deibert, R., & Hecker, O. (1995). Fault Detection of Flowrate and Temperature Control Loops using Estimated Parity Equations. *IFAC Proceedings Volumes*, 28(12), 193–198. [https://doi.org/10.1016/s1474-6670\(17\)45421-8](https://doi.org/10.1016/s1474-6670(17)45421-8)
- Huang, Y., Reklaitis, G. V., & Venkatasubramanian, V. (2003). A Heuristic Extended Kalman Filter Based Estimator for Fault Identification in a Fluid Catalytic Cracking Unit. *Industrial & Engineering Chemistry Research*, 42(14), 3361–3371. <https://doi.org/10.1021/ie010659t>
- Iri, M., Aoki, K., Shima, E. O., & Matsuyama, H. (1979). An algorithm for diagnosis of system failures in the chemical process. *Computers and Chemical Engineering*, 3(1–4), 489–493. [https://doi.org/10.1016/0098-1354\(85\)80006-5](https://doi.org/10.1016/0098-1354(85)80006-5)
- Isermann, R. (1985). Process Fault Diagnosis with Parameter Estimation Methods. *IFAC Proceedings Volumes*, 18(11), 51–60. [https://doi.org/10.1016/s1474-6670\(17\)60110-1](https://doi.org/10.1016/s1474-6670(17)60110-1)
- Isermann, R., & Ballé, P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. *Control Engineering Practice*, 5(5), 709–719. [https://doi.org/10.1016/S0967-0661\(97\)00053-1](https://doi.org/10.1016/S0967-0661(97)00053-1)
- Izadi, I., Shah, S. L., & Chen, T. (2008). Parity space fault detection based on irregularly sampled data. *Proceedings of the American Control Conference*, 2798–2803. <https://doi.org/10.1109/ACC.2008.4586917>
- Jayaprasanth, D., & Kanthalakshmi, S. (2016). Improved State Estimation for Stochastic Nonlinear Chemical Reactor using Particle Filter based on Unscented Transformation. *Control Engineering and Applied Informatics*, 18(4), 36–44.
- Jiang, S., Gui, W., Ding, S. X., & Xie, Y. (2008). Direct identification of model-based fault detection system and its application to the process of lead-zinc Smelting Furnace. 453–457.
- Kadirkamanathan, V., Li, P., Jawaxd, M. H., & Fabri, S. . (2000). A sequential Monte Carlo filtering approach to fault detection and isolation in nonlinear systems. *The 39th IEEE Conference on Decision and Control*, (1), 4341–4346. Retrieved from <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=914708>
- Kandepu, R., Foss, B., & Imsland, L. (2008). Applying the unscented Kalman filter for nonlinear state estimation. *Journal of Process Control*, 18(7–8), 753–768. <https://doi.org/10.1016/j.jprocont.2007.11.004>
- Kang, C. W., & Golay, M. W. (1999). A Bayesian belief network-based advisory system for operational availability focused diagnosis of complex nuclear power systems. *Expert*

- Systems with Applications*, 17(1), 21–32. [https://doi.org/10.1016/s0957-4174\(99\)00018-4](https://doi.org/10.1016/s0957-4174(99)00018-4)
- Kavčič, M., & Juričić, Đ. (1997). Process Diagnosis Using Fault-Tree Propagation Structures: The CAD Approach. *IFAC Proceedings Volumes*, 30(18), 531–536. [https://doi.org/10.1016/s1474-6670\(17\)42456-6](https://doi.org/10.1016/s1474-6670(17)42456-6)
- Keliris, C., Polycarpou, M. M., & Parisini, T. (2015). A robust nonlinear observer-based approach for distributed fault detection of input-output interconnected systems. *Automatica*, 53, 408–415. <https://doi.org/10.1016/j.automatica.2015.01.042>
- Khakzad, N., Khan, F., & Amyotte, P. (2011). Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliability Engineering and System Safety*, 96(8), 925–932. <https://doi.org/10.1016/j.ress.2011.03.012>
- Khan, F. I., & Abbasi, S. A. (1999). PROFAT: A User Friendly System for Probabilistic Fault Tree Analysis. *Process Safety Progress*, 18(1), 42–49.
- Khan, F. I., & Abbasi, S. A. (2000). Analytical simulation and PROFAT II: A new methodology and a computer automated tool for fault tree analysis in chemical process industries. *Journal of Hazardous Materials*, 75(1), 1–27. [https://doi.org/10.1016/S0304-3894\(00\)00169-2](https://doi.org/10.1016/S0304-3894(00)00169-2)
- Kleer, J. De, & Brown, J. S. (1984). A Qualitative Physics Based on Confluences. *Artificial Intelligence*, 24, 7–83. <https://doi.org/10.1016/B978-1-4832-1447-4.50013-4>
- Kokawa, M., Miyazaki, S., & Shingai, S. (1983). Fault Location Using Digraph and Inverse Direction Search with Application. *Automatica*, 19(6), 729–735. [https://doi.org/10.1016/0005-1098\(83\)90039-0](https://doi.org/10.1016/0005-1098(83)90039-0)
- Kordon, A., & Dhurjati, P. (1995). An Expert System for Crude Unit Process Supervision. *IFAC Proceedings Volumes*, 28(12), 143–147. [https://doi.org/10.1016/s1474-6670\(17\)45413-9](https://doi.org/10.1016/s1474-6670(17)45413-9)
- Kramer, M. A., & Jr., B. L. P. (1987). A Rule-Based Approach to Fault Diagnosis Using the Signed Directed Graph. *AIChE Journal*, 33(7), 1067–1078. <https://doi.org/10.1002/aic.690330703>
- Kuipers, B. (1986). Qualitative simulation. *Artificial Intelligence*, 29(3), 289–338. [https://doi.org/10.1016/0004-3702\(86\)90073-1](https://doi.org/10.1016/0004-3702(86)90073-1)
- Landman, R., & Jämsä-Jounela, S.-L. (2016). Hybrid approach to casual analysis on a complex industrial system based on transfer entropy in conjunction with process connectivity information. *Control Engineering Practice*, 53, 14–23. <https://doi.org/10.1016/j.conengprac.2016.04.010>
- Landman, R., Kortela, J., Sun, Q., & Jämsä-Jounela, S. L. (2014). Fault propagation analysis of oscillations in control loops using data-driven causality and plant connectivity. *Computers and Chemical Engineering*, 71, 446–456. <https://doi.org/10.1016/j.compchemeng.2014.09.017>
- Lapp, S. A., & Powers, G. J. (1977). Computer-aided Synthesis of Fault-trees. *IEEE Transactions on Reliability*, R-26(1), 2–13.
- Lardon, L., Puñal, A., Martinez, J. A., & Steyer, J. P. (2005). Modular expert system for the diagnosis of operating conditions of industrial anaerobic digestion plants. *Water Science and Technology*, 52(1–2), 427–433. <https://doi.org/10.2166/wst.2005.0549>
- Lavasani, S. M., Zendegani, A., & Celik, M. (2015). An extension to Fuzzy Fault Tree Analysis

- (FFTA) application in petrochemical process industry. *Process Safety and Environmental Protection*, 93, 75–88. <https://doi.org/10.1016/j.psep.2014.05.001>
- Lee, G., Mo, K. J., & Yoon, E. S. (1997). The methodology for knowledge base compression and robust diagnosis: Application to a steam boiler plant. *Expert Systems with Applications*, 12(2), 263–274. [https://doi.org/10.1016/S0957-4174\(96\)00099-1](https://doi.org/10.1016/S0957-4174(96)00099-1)
- Lee, G., Tosukhowong, T., Lee, J. H., & Han, C. (2006). Fault Diagnosis Using the Hybrid Method of Signed Digraph and Partial Least Squares with Time Delay: The Pulp Mill Process. *Industrial and Engineering Chemistry Research*, 45(26), 9061–9074. <https://doi.org/10.1021/ie060793j>
- Lerner, U., Parr, R., Koller, D., & Biswas, G. (2000). Bayesian Fault Detection and Diagnosis in Dynamic Systems. *The Seventeenth National Conference on Artificial Intelligence*, 531–537.
- Leung, D., & Romagnoli, J. (2000). Dynamic probabilistic model-based expert system for fault diagnosis. *Computers and Chemical Engineering*, 24(11), 2473–2492. [https://doi.org/10.1016/S0098-1354\(00\)00610-4](https://doi.org/10.1016/S0098-1354(00)00610-4)
- Li, H., Wang, F., Li, H., & Wang, Q. (2019). Safety control modeling method based on Bayesian network transfer learning for the thickening process of gold hydrometallurgy. *Knowledge-Based Systems*. <https://doi.org/10.1016/j.knosys.2019.105297>
- Li, L., Ding, S. X., Qiu, J., Yang, Y., & Zhang, Y. (2016). Weighted fuzzy observer-based fault detection approach for discrete-time nonlinear systems via piecewise-fuzzy lyapunov functions. *IEEE Transactions on Fuzzy Systems*, 24(6), 1320–1333. <https://doi.org/10.1109/TFUZZ.2016.2514371>
- Li, L., Ding, S. X., Yang, Y., & Zhang, Y. (2016). Robust fuzzy observer-based fault detection for nonlinear systems with disturbances. *Neurocomputing*, 174(Part B), 767–772. <https://doi.org/10.1016/j.neucom.2015.09.102>
- Li, P., & Kadirkamanathan, V. (2001). Particle Filtering Based Likelihood Ratio Approach to Fault Diagnosis in Nonlinear Stochastic Systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 31(3), 337–343. <https://doi.org/10.1109/5326.971661>
- Li, R., & Olson, J. H. (1991). Fault Detection and Diagnosis in a Closed-Loop Nonlinear Distillation Process: Application of Extended Kalman Filters. *Industrial and Engineering Chemistry Research*, 30(5), 898–908. <https://doi.org/10.1021/ie00053a012>
- Liao, S. H. (2005). Expert system methodologies and applications-a decade review from 1995 to 2004. *Expert Systems with Applications*, 28(1), 93–103. <https://doi.org/10.1016/j.eswa.2004.08.003>
- Liu, M., Zang, S., & Zhou, D. (2005). Fast leak detection and location of gas pipelines based on an adaptive particle filter. *International Journal of Applied Mathematics and Computer Science*, 15(4), 541–550.
- Liu, Y., Meng, Q., Zeng, M., & Ma, S. (2016). Fault Diagnosis Method Based on Probability Extended SDG and Fault Index. *The 12th World Congress on Intelligent Control and Automation*, 2868–2873. <https://doi.org/10.1109/WCICA.2016.7578605>
- Lü, N., & Wang, X. (2008). Fault Diagnosis Based on Signed Digraph Combined with Dynamic Kernel PLS and SVR. *Industrial and Engineering Chemistry Research*, 47(23), 9447–9456. <https://doi.org/10.1021/ie8009457>

- Lü, N., Xiong, Z., Wang, X., & Ren, C. (2011). Integrated Framework of Probabilistic Signed Digraph Based Fault Diagnosis Approach to a Gas Fractionation Unit. *Industrial and Engineering Chemistry Research*, 50(17), 10062–10073. <https://doi.org/10.1021/ie200016t>
- Lu, Y., Wang, F., Jia, M., & Qi, Y. (2016). Centrifugal compressor fault diagnosis based on qualitative simulation and thermal parameters. *Mechanical Systems and Signal Processing*, 81, 259–273. <https://doi.org/10.1016/j.ymssp.2016.03.018>
- Ma, L., Dong, J., & Peng, K. (2018). A practical propagation path identification scheme for quality-related faults based on nonlinear dynamic latent variable model and partitioned Bayesian network. *Journal of the Franklin Institute*, 355(15), 7570–7594. <https://doi.org/10.1016/j.jfranklin.2018.07.035>
- Maurya, M.R., Rengaswamy, R., & Venkatasubramanian, V. (2007). A signed directed graph and qualitative trend analysis-based framework for incipient fault diagnosis. *Chemical Engineering Research and Design*, 85(A10), 1407–1422. [https://doi.org/10.1016/S0263-8762\(07\)73181-7](https://doi.org/10.1016/S0263-8762(07)73181-7)
- Maurya, Mano Ram, Rengaswamy, R., & Venkatasubramanian, V. (2003a). A Systematic Framework for the Development and Analysis of Signed Digraphs for Chemical Processes. 1. Algorithms and Analysis. *Industrial and Engineering Chemistry Research*, 42(20), 4789–4810. <https://doi.org/10.1021/ie020644a>
- Maurya, Mano Ram, Rengaswamy, R., & Venkatasubramanian, V. (2003b). A Systematic Framework for the Development and Analysis of Signed Digraphs for Chemical Processes. 2. Control Loops and Flowsheet Analysis. *Industrial and Engineering Chemistry Research*, 42(20), 4811–4827. <https://doi.org/10.1021/ie0206453>
- Mehra, R. K., & Peschon, J. (1971). An Innovations Approach to Fault Detection and Diagnosis in Dynamic Systems. *Automatica*, 7(5), 637–640. [https://doi.org/10.1016/0005-1098\(71\)90028-8](https://doi.org/10.1016/0005-1098(71)90028-8)
- Melani, A. H. A., Murad, C. A., Netto, A. C., Souza, G. F. M. de, & Nabeta, S. I. (2018). Criticality-based maintenance of a coal-fired power plant. *Energy*, 147, 767–781. <https://doi.org/10.1016/j.energy.2018.01.048>
- Montani, S., Portinale, L., Bobbio, A., Varesio, M., & Codetta-Raiteri, D. (2006). A tool for automatically translating dynamic fault trees into dynamic Bayesian networks. *Annual Reliability and Maintainability Symposium*, 434–441. <https://doi.org/10.1109/RAMS.2006.1677413>
- Morales-Menéndez, R., Freitas, N. de, & Poole, D. (2003). Real-Time Monitoring of Complex Industrial Processes with Particle Filters. *Advances in Neural Information Processing Systems* 15.
- Morbach, J., Yang, A., & Marquardt, W. (2007). OntoCAPE-A large-scale ontology for chemical process engineering. *Engineering Applications of Artificial Intelligence*, 20(2), 147–161. <https://doi.org/10.1016/j.engappai.2006.06.010>
- Mosalanejad, M., & Arefi, M. M. (2018). UKF-based soft sensor design for joint estimation of chemical processes with multi-sensor information fusion and infrequent measurements. *IET Science, Measurement & Technology*, 12(6), 755–763. <https://doi.org/10.1049/iet-smt.2017.0340>
- Mylaraswamy, Dinkar Venkatasubramanian, V. (1997). A Hybrid Framework for Large Scale

- Process Fault Diagnosis. *Computers and Chemical Engineering*, 21(Supplement), S935–S940. [https://doi.org/10.1016/S0098-1354\(97\)87622-3](https://doi.org/10.1016/S0098-1354(97)87622-3)
- Nakajima, M., von Numershs, C., Yada, H., Siimes, T., Pokkinen, M., Endo, I., & Linko, P. (1994). An on-line advisory control system for the lactic acid fermentation process. *Applied Microbiology and Biotechnology*, 42(2–3), 204–211. <https://doi.org/10.1007/BF00902718>
- Nan, C., Khan, F., & Iqbal, M. T. (2008). Real-time fault diagnosis using knowledge-based expert system. *Process Safety and Environmental Protection*, 86(1), 55–71. <https://doi.org/10.1016/j.psep.2007.10.014>
- Natarajan, S., Ghosh, K., & Srinivasan, R. (2012). An ontology for distributed process supervision of large-scale chemical plants. *Computers and Chemical Engineering*, 46, 124–140. <https://doi.org/10.1016/j.compchemeng.2012.06.009>
- Natarajan, S., & Srinivasan, R. (2014). Implementation of multi agents based system for process supervision in large-scale chemical plants. *Computers and Chemical Engineering*, 60, 182–196. <https://doi.org/10.1016/j.compchemeng.2013.08.012>
- Nguang, S. K., Zhang, P., & Ding, S. X. (2007). Parity Relation Based Fault Estimation for Nonlinear Systems: An LMI Approach. *International Journal of Automation and Computing*, 4(2), 164–168. <https://doi.org/10.1016/B978-008044485-7/50062-2>
- Odendaal, H. M., & Jones, T. (2014). Actuator fault detection and isolation: An optimised parity space approach. *Control Engineering Practice*, 26(1), 222–232. <https://doi.org/10.1016/j.conengprac.2014.01.013>
- Orani, N., Pisano, A., & Usai, E. (2010). Fault diagnosis for the vertical three-tank system via high-order sliding-mode observation. *Journal of the Franklin Institute*, 347(6), 923–939. <https://doi.org/10.1016/j.jfranklin.2009.11.010>
- Oyeleye, O. O., & Kramer, M. A. (1988). Qualitative simulation of chemical process systems: Steady-state analysis. *AIChE Journal*, 34(9), 1441–1454. <https://doi.org/10.1002/aic.690340906>
- Özyurt, B., & Kandel, A. (1996). A hybrid hierarchical neural network-fuzzy expert system approach to chemical process fault diagnosis. *Fuzzy Sets and Systems*, 83(1), 11–25.
- Patton, R. J., & Chen, J. (1991). A Review of Parity Space Approaches to Fault Diagnosis. *IFAC Proceedings Volumes*, 24(6), 65–81. [https://doi.org/10.1016/s1474-6670\(17\)51124-6](https://doi.org/10.1016/s1474-6670(17)51124-6)
- Peng, D., Geng, Z., & Zhu, Q. (2014). A Multilogic Probabilistic Signed Directed Graph Fault Diagnosis Approach Based on Bayesian Inference. *Industrial and Engineering Chemistry Research*, 53(23), 9792–9804. <https://doi.org/10.1021/ie403608a>
- Peng, Y., Youssouf, A., Arte, P., & Kinnaert, M. (1997). A complete procedure for residual generation and evaluation with application to a heat exchanger. *IEEE Transactions on Control Systems Technology*, 5(6), 542–555. <https://doi.org/10.1109/87.641400>
- Prakash, J., Huang, B., & Shah, S. L. (2014). Recursive constrained state estimation using modified extended Kalman filter. *Computers and Chemical Engineering*, 65, 9–17. <https://doi.org/10.1016/j.compchemeng.2014.02.013>
- Prasad, P. R., Davis, J. F., Jirapinyo, Y., Josephson, J. R., & Bhalodia, M. (1998). Structuring diagnostic knowledge for large-scale process systems. *Computers and Chemical Engineering*, 22(12), 1897–1905. [https://doi.org/10.1016/S0098-1354\(98\)00227-0](https://doi.org/10.1016/S0098-1354(98)00227-0)

- Puñal, A., Roca, E., & Lema, J. M. (2002). An expert system for monitoring and diagnosis of anaerobic wastewater treatment plants. *Water Research*, 36(10), 2656–2666.
- Purba, J. H., Tjahyani, D. T. S., Ekariansyah, A. S., & Tjahjono, H. (2015). Fuzzy probability based fault tree analysis to propagate and quantify epistemic uncertainty. *Annals of Nuclear Energy*, 85, 1189–1199. <https://doi.org/10.1016/j.anucene.2015.08.002>
- Qi, F., & Huang, B. (2011). Bayesian methods for control loop diagnosis in the presence of temporal dependent evidences. *Automatica*, 47(7), 1349–1356. <https://doi.org/10.1016/j.automatica.2011.02.015>
- Qi, F., Huang, B., & Tamayo, E. C. (2010). A Bayesian Approach for Control Loop Diagnosis with Missing Data. *AIChE Journal*, 56(1), 179–195. <https://doi.org/10.1002/aic>
- Qian, Y., Li, X., Jiang, Y., & Wen, Y. (2003). An expert system for real-time fault diagnosis of complex chemical processes. *Expert Systems with Applications*, 24(4), 425–432. [https://doi.org/10.1016/S0957-4174\(02\)00190-2](https://doi.org/10.1016/S0957-4174(02)00190-2)
- Qian, Y., Xu, L., Li, X., Lin, L., & Kraslawski, A. (2008). LUBRES: An expert system development and implementation for real-time fault diagnosis of a lubricating oil refining process. *Expert Systems with Applications*, 35(3), 1252–1266. <https://doi.org/10.1016/j.eswa.2007.07.061>
- Qian, Y., Zheng, M., Li, X., & Lin, L. (2005). Implementation of knowledge maintenance modules in an expert system for fault diagnosis of chemical process operation. *Expert Systems with Applications*, 28(2), 249–257. <https://doi.org/10.1016/j.eswa.2004.10.005>
- Qin, S. J. (2012). Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, 36(2), 220–234. <https://doi.org/10.1016/j.arcontrol.2012.09.004>
- Qin, S. J., & Chiang, L. H. (2019). Advances and opportunities in machine learning for process data analytics. *Computers and Chemical Engineering*, 126, 465–473. <https://doi.org/10.1016/j.compchemeng.2019.04.003>
- Qin, S. J., & Li, W. (2001). Detection and identification of faulty sensors in dynamic processes. *AIChE Journal*, 47(7), 1581–1593. <https://doi.org/10.1002/aic.690470711>
- Qu, C. C., & Hahn, J. (2009). Process monitoring and parameter estimation via unscented Kalman filtering. *Journal of Loss Prevention in the Process Industries*, 22(6), 703–709. <https://doi.org/10.1016/j.jlp.2008.07.012>
- Reis, M. S., & Gins, G. (2017). Industrial Process Monitoring in the Big Data/Industry 4.0 Era: from Detection, to Diagnosis, to Prognosis. *Processes*, 5(4), 35. <https://doi.org/10.3390/pr5030035>
- Renjith, V. R., Madhu, G., Nayagam, V. L. G., & Bhasi, A. B. (2010). Two-dimensional fuzzy fault tree analysis for chlorine release from a chlor-alkali industry using expert elicitation. *Journal of Hazardous Materials*, 183(1–3), 103–110. <https://doi.org/10.1016/j.jhazmat.2010.06.116>
- Reppa, V., & Tzes, A. (2011). Fault detection and diagnosis based on parameter set estimation. *IET Control Theory & Applications*, 5(1), 69–83. <https://doi.org/10.1049/iet-cta.2009.0202>
- Rich, S. H., Venkatasubramanian, V., Nasrallah, M., & Matteo, C. (1989). Development of a diagnostic expert system for a whipped toppings process. *Journal of Loss Prevention in the Process Industries*, 2(3), 145–154.
- Rojas-Guzman, C., & Kramer, M. A. (1993). Comparison of Belief Networks and Rule-based

- Expert Systems for Fault Diagnosis of Chemical Processes. *Engineering Applications of Artificial Intelligence*, 6(3), 191–202. [https://doi.org/10.1016/0952-1976\(93\)90062-3](https://doi.org/10.1016/0952-1976(93)90062-3)
- Romanenko, A., & Castro, J. A. A. M. (2004). The unscented filter as an alternative to the EKF for nonlinear state estimation: A simulation case study. *Computers and Chemical Engineering*, 28(3), 347–355. [https://doi.org/10.1016/S0098-1354\(03\)00193-5](https://doi.org/10.1016/S0098-1354(03)00193-5)
- Salahshoor, K., Mosallaei, M., & Bayat, M. (2008). Centralized and decentralized process and sensor fault monitoring using data fusion based on adaptive extended Kalman filter algorithm. *Measurement*, 41(10), 1059–1076. <https://doi.org/10.1016/j.measurement.2008.02.009>
- Schubert, U., Kruger, U., Arellano-Garcia, H., de Sá Feital, T., & Wozny, G. (2011). Unified model-based fault diagnosis for three industrial application studies. *Control Engineering Practice*, 19(5), 479–490. <https://doi.org/10.1016/j.conengprac.2011.01.009>
- Severson, K., Chaiwatanodom, P., & Braatz, R. D. (2016). Perspectives on process monitoring of industrial systems. *Annual Reviews in Control*, 42, 190–200. <https://doi.org/10.1016/j.arcontrol.2016.09.001>
- Shen, Q., & Leitch, R. (1993). Fuzzy Qualitative Simulation. *IEEE Transactions on Systems, Man and Cybernetics*, 23(4), 1038–1061. <https://doi.org/10.1109/21.247887>
- Shenoy, A. V., Prakash, J., Prasad, V., Shah, S. L., & McAuley, K. B. (2013). Practical issues in state estimation using particle filters: Case studies with polymer reactors. *Journal of Process Control*, 23(2), 120–131. <https://doi.org/10.1016/j.jprocont.2012.09.003>
- Shiozaki, J., Matsuyama, H., Tano, K., & O'Shima, E. (1985). Fault Diagnosis of Chemical Processes by Use of Signed Directed Graphs: Extension to Five-Range Patterns of Abnormality. *International Chemical Engineering*, 4(25), 651–659.
- Shiozaki, Junichi, Shibata, B., Matsuyama, H., & O'Shima, E. (1989). Fault Diagnosis of Chemical Processes Utilizing Signed Directed Graphs-Improvement by Using Temporal Information. *IEEE Transactions on Industrial Electronics*, 36(4), 469–474. <https://doi.org/10.1109/41.43004>
- Shu, Y., Ming, L., Cheng, F., Zhang, Z., & Zhao, J. (2016). Abnormal situation management: Challenges and opportunities in the big data era. *Computers and Chemical Engineering*, 91, 104–113. <https://doi.org/10.1016/j.compchemeng.2016.04.011>
- Simani, S. (1999). Sensor fault diagnosis of a power plant: An approach based on state estimation techniques. *Recent Advances in Signal Processing and Communications*, 274–281.
- Simonart, X., & Kinnaert, M. (1994). Sensor Fault Detection and Isolation on a Pilot Heat Exchanger. *IFAC Proceedings Volumes*, 27(5), 423–428. [https://doi.org/10.1016/s1474-6670\(17\)48064-5](https://doi.org/10.1016/s1474-6670(17)48064-5)
- Steele, A. D., & Leitch, R. (1997). Time-constrained qualitative model-based diagnosis. *Engineering Applications of Artificial Intelligence*, 10(5), 417–427. [https://doi.org/10.1016/s0952-1976\(97\)00031-6](https://doi.org/10.1016/s0952-1976(97)00031-6)
- Sule, I., Khan, F., Butt, S., & Yang, M. (2018). Kick control reliability analysis of managed pressure drilling operation. *Journal of Loss Prevention in the Process Industries*, 52, 7–20. <https://doi.org/10.1016/j.jlp.2018.01.007>
- Surgenor, B. W., & Jofriet, P. J. (1992). Thermal Fault Analysis and the Diagnostic Model Processor. *IFAC Proceedings Volumes*, 25(4), 37–42. <https://doi.org/10.1016/s1474->

- Takagi, T., & Sugeno, M. (1985). Fuzzy Identification of Systems and Its Applications to Modeling and Control. *IEEE Transactions on Systems, Man and Cybernetics, SMC-15*(1), 116–132. <https://doi.org/10.1109/TSMC.1985.6313399>
- Tarifa, E. E., & Scenna, N. J. (1997). Fault diagnosis, direct graphs, and fuzzy logic. *Computers and Chemical Engineering, 21*(Supplement), S649–S654. [https://doi.org/10.1016/S0098-1354\(97\)87576-X](https://doi.org/10.1016/S0098-1354(97)87576-X)
- Tarifa, E. E., & Scenna, N. J. (2003). Fault diagnosis for a MSF using a SDG and fuzzy logic. *Desalination, 152*(1–3), 207–214. <https://doi.org/10.1016/j.desal.2004.06.063>
- Termehchy, A., & Afshar, A. (2014). A novel design of unknown input observer for fault diagnosis in non-minimum phase systems. *The 19th IFAC World Congress, 19*(3), 8552–8557. <https://doi.org/10.3182/20140824-6-za-1003.02104>
- Terpstra, V. J., Verbruggen, H. B., Hoogland, M. W., & Ficke, R. A. E. (1992). A Real-Time Fuzzy, Deep-knowledge Based Fault-diagnosis System for a CSTR. *IFAC Proceedings Volumes, 25*(4), 73–78. [https://doi.org/10.1016/s1474-6670\(17\)50219-0](https://doi.org/10.1016/s1474-6670(17)50219-0)
- Thambirajah, J., Benabbas, L., Bauer, M., & Thornhill, N. F. (2007). Cause and Effect Analysis in Chemical Processes Utilizing Plant Connectivity Information. *IEEE Advanced Process Control Applications for Industry Workshop*, 1–6.
- Thambirajah, J., Benabbas, L., Bauer, M., & Thornhill, N. F. (2009). Cause-and-effect analysis in chemical processes utilizing XML, plant connectivity and quantitative process history. *Computers and Chemical Engineering, 33*(2), 503–512. <https://doi.org/10.1016/j.compchemeng.2008.10.002>
- Theilliol, D., Noura, H., & Ponsart, J. C. (2002). Fault diagnosis and accommodation of a three-tank system based on analytical redundancy. *ISA Transactions, 41*(3), 365–382. [https://doi.org/10.1016/s0019-0578\(07\)60094-9](https://doi.org/10.1016/s0019-0578(07)60094-9)
- Tidriri, K., Chatti, N., Verron, S., & Tiplica, T. (2016). Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges. *Annual Reviews in Control, 42*, 63–81. <https://doi.org/10.1016/j.arcontrol.2016.09.008>
- Tsuge, Y., Hiratsuka, K., Takeda, K., & Matsuyama, H. (2000). A fault detection and diagnosis for the continuous process with load-fluctuations using orthogonal wavelets. *Computers and Chemical Engineering, 24*(2–7), 761–767. [https://doi.org/10.1016/S0098-1354\(00\)00368-9](https://doi.org/10.1016/S0098-1354(00)00368-9)
- Tulsyan, A., Bhushan Gopaluni, R., & Khare, S. R. (2016). Particle filtering without tears: A primer for beginners. *Computers and Chemical Engineering, 95*, 130–145. <https://doi.org/10.1016/j.compchemeng.2016.08.015>
- Ulerich, N. H., & Powers, G. J. (1988). On-Line Hazard Aversion and Fault Diagnosis in Chemical Processes: The Digraph + Fault-Tree Method. *IEEE Transactions on Reliability, 37*(2), 171–177. <https://doi.org/10.1109/24.3738>
- Umeda, T., Kuriyama, T., O'Shima, E., & Matsuyama, H. (1980). A graphical approach to cause and effect analysis of chemical processing systems. *Chemical Engineering Science, 35*(12), 2379–2388.
- Vachhani, P., Rengaswamy, R., Gangwal, V., & Narasimhan, S. (2005). Recursive estimation in constrained nonlinear dynamical systems. *AIChE Journal, 51*(3), 946–959.

<https://doi.org/10.1002/aic.10355>

- Vachhani, P., Rengaswamy, R., & Venkatasubramanian, V. (2001). A framework for integrating diagnostic knowledge with nonlinear optimization for data reconciliation and parameter estimation in dynamic systems. *Chemical Engineering Science*, 56(6), 2133–2148. [https://doi.org/10.1016/S0009-2509\(00\)00488-7](https://doi.org/10.1016/S0009-2509(00)00488-7)
- Vedam, H., & Venkatasubramanian, V. (1997). Signed Digraph Based Multiple Fault Diagnosis. *Computers and Chemical Engineering*, 21(Supplement), S655–S660.
- Venkatasubramanian, V., Rengaswamy, R., & Kavuri, S. N. (2003). A review of process fault detection and diagnosis part II: Qualitative models and search strategies. *Computers and Chemical Engineering*, 27(3), 313–326. [https://doi.org/10.1016/S0098-1354\(02\)00161-8](https://doi.org/10.1016/S0098-1354(02)00161-8)
- Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., & Yin, K. (2003). A review of fault detection and diagnosis. Part III: Process history based methods. *Computers and Chemical Engineering*, 27, 327–346.
- Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2003). A review of process fault detection and diagnosis part I: Quantitative model-based methods. *Computers and Chemical Engineering*, 27(3), 293–311. [https://doi.org/10.1016/S0098-1354\(02\)00160-6](https://doi.org/10.1016/S0098-1354(02)00160-6)
- Wan, E. A., & Menve, R. van der. (2000). The Unscented Kalman Filter for Nonlinear Estimation. *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, 153–158.
- Wang, X. Z., Yang, S. A., Veloso, E., Lu, M. L., & Mcgreavy, C. (1995). Qualitative process modelling-a fuzzy signed directed graph method. *Computers and Chemical Engineering*, 19(Supplement), 735–740.
- Wang, Y. Y., & Wu, N. E. (1993). An approach to configuration of robust control systems for robust failure detection. *IEEE Conference on Decision and Control*, 1704–1709.
- Waters, A., & Ponton, J. W. (1989). Qualitative simulation and fault propagation in process plants. *Chemical Engineering Research and Design*, 67(4), 407–422. [https://doi.org/10.1016/S0076-6879\(08\)03802-0](https://doi.org/10.1016/S0076-6879(08)03802-0)
- Weber, P., Theilliol, D., Aubrun, C., & Evsukoff, A. (2006). Increasing effectiveness of model-based fault diagnosis: A dynamic bayesian network design for decision making. *IFAC Proceedings Volumes*, 39(13), 90–95. <https://doi.org/10.1016/B978-008044485-7/50016-6>
- Weidl, G., Madsen, A. L., & Israelson, S. (2005). Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes. *Computers and Chemical Engineering*, 29(9), 1996–2009. <https://doi.org/10.1016/j.compchemeng.2005.05.005>
- Wu, H., & Zhao, J. (2020). Fault detection and diagnosis based on transfer learning for multimode chemical processes. *Computers & Chemical Engineering*, 135, 106731. <https://doi.org/https://doi.org/10.1016/j.compchemeng.2020.106731>
- Wu, M., She, J.-H., Nakano, M., & Gui, W. (2002). Expert control and fault diagnosis of the leaching process in a zinc hydrometallurgy plant. *Control Engineering Practice*, 10(4), 433–442.
- Yang, D., Dong, M., & Miao, R. (2008). Development of a product configuration system with an ontology-based approach. *Computer-Aided Design*, 40(8), 863–878.

- <https://doi.org/10.1016/j.cad.2008.05.004>
- Yang, D., Miao, R., Wu, H., & Zhou, Y. (2009). Product configuration knowledge modeling using ontology web language. *Expert Systems with Applications*, 36(3), 4399–4411. <https://doi.org/10.1016/j.eswa.2008.05.026>
- Yang, Y., Ding, S. X., & Li, L. (2015). On observer-based fault detection for nonlinear systems. *Systems and Control Letters*, 82, 18–25. <https://doi.org/10.1016/j.sysconle.2015.05.004>
- Yim, S. Y., Ananthakumar, H. G., Benabbas, L., Horch, A., Drath, R., & Thornhill, N. F. (2006). Using process topology in plant-wide control loop performance assessment. *Computers and Chemical Engineering*, 31(2), 86–99. <https://doi.org/10.1016/j.compchemeng.2006.05.004>
- Yin, S., Ding, S. X., Xie, X., & Luo, H. (2014). A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, 61(11), 6414–6428. <https://doi.org/10.1109/TIE.2014.2301773>
- Yin, S., & Zhu, X. (2015). Intelligent Particle Filter and Its Application to Fault Detection of Nonlinear System. *IEEE Transactions on Industrial Electronics*, 62(6), 3852–3861. <https://doi.org/10.1109/TIE.2015.2399396>
- Yu, H., Khan, F., & Garaniya, V. (2015). Modified Independent Component Analysis and Bayesian Network-Based Two-Stage Fault Diagnosis of Process Operations. *Industrial and Engineering Chemistry Research*, 54(10), 2724–2742. <https://doi.org/10.1021/ie503530v>
- Yu, J., & Rashid, M. M. (2013). 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. <https://doi.org/10.1002/aic>
- Yuan, P., Sun, Y., Li, H., Wang, F., & Li, H. (2019). Abnormal Condition Identification Modeling Method Based on Bayesian Network Parameters Transfer Learning for the Electro-Fused Magnesia Smelting Process. *IEEE Access*, 7, 149764–149775. <https://doi.org/10.1109/access.2019.2947499>
- Zahedi, G., Saba, S., al-Otaibi, M., & Mohd-Yusof, K. (2011). Troubleshooting of crude oil desalination plant using fuzzy expert system. *Desalination*, 266(1–3), 162–170. <https://doi.org/10.1016/j.desal.2010.08.020>
- Zakharov, A., Tikkala, V. M., & Jämsä-Jounela, S. L. (2013). Fault detection and diagnosis approach based on nonlinear parity equations and its application to leakages and blockages in the drying section of a board machine. *Journal of Process Control*, 23(9), 1380–1393. <https://doi.org/10.1016/j.jprocont.2013.03.006>
- Zhang, J., & Roberts, P. D. (1991). Process fault diagnosis with diagnostic rules based on structural decomposition. *Journal of Process Control*, 1(5), 259–269. [https://doi.org/10.1016/0959-1524\(91\)85017-D](https://doi.org/10.1016/0959-1524(91)85017-D)
- Zhang, J., Roberts, P. D., & Ellis, J. E. (1990). Fault diagnosis of a mixing process using deep qualitative knowledge representation of physical behaviour. *Journal of Intelligent and Robotic Systems*, 3(2), 103–115. <https://doi.org/10.1007/BF00242159>
- Zhang, J., Roberts, P. D., & Ellis, J. E. (1991). A self-learning fault-diagnosis system. *Transactions of the Institute of Measurement and Control*, 13(1), 29–35.
- Zhang, K., Jiang, B., & Cocquempot, V. (2008). Adaptive Observer-based Fast Fault Estimation. *International Journal of Control, Automation and Systems*, 6(3), 320–326.

- Zhang, X., Polycarpou, M. M., & Parisini, T. (2010). Fault diagnosis of a class of nonlinear uncertain systems with Lipschitz nonlinearities using adaptive estimation. *Automatica*, 46(2), 290–299. <https://doi.org/10.1016/j.automatica.2009.11.014>
- Zhang, Zhaoqian, Wu, C., Zhang, B., Xia, T., & Li, A. (2005). SDG multiple fault diagnosis by real-time inverse inference. *Reliability Engineering and System Safety*, 87(2), 173–189. <https://doi.org/10.1016/j.ress.2004.04.008>
- Zhang, Zhengdao, & Dong, F. (2014). Fault detection and diagnosis for missing data systems with a three time-slice dynamic Bayesian network approach. *Chemometrics and Intelligent Laboratory Systems*, 138, 30–40. <https://doi.org/10.1016/j.chemolab.2014.07.009>
- Zhao, C., Bhushan, M., & Venkatasubramanian, V. (2005). PHASuite: An Automated HAZOP Analysis Tool for Chemical Processes Part: II: Implementation and Case Study. *Process Safety and Environmental Protection*, 83(6), 533–548. <https://doi.org/10.1205/psep.04056>
- Zhao, J., Cui, L., Zhao, L., Qiu, T., & Chen, B. (2009). Learning HAZOP expert system by case-based reasoning and ontology. *Computers and Chemical Engineering*, 33(1), 371–378. <https://doi.org/10.1016/j.compchemeng.2008.10.006>
- Zolghadri, A., Henry, D., & Monsion, M. (1996). Design of nonlinear observers for fault diagnosis: A case study. *Control Engineering Practice*, 4(11), 1535–1544. [https://doi.org/10.1016/0967-0661\(96\)00167-0](https://doi.org/10.1016/0967-0661(96)00167-0)