Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Fault detection for geological drilling processes using multivariate generalized Gaussian distribution and Kullback Leibler divergence ⋆

Yupeng Li *,** **Weihua Cao** *,** **Wenkai Hu** *,** **Chao Gan** *,**
Min Wu *,**

* *School of Automation, China University of Geosciences, Wuhan 430074, China.*
** *Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, China.*

e-mail: {yupengli, weihuacao, wenkaihu, ganchao, wumin}@cug.edu.cn

Abstract: The presence of downhole faults compromises the safety and also leads to increased maintenance costs in complex geological drilling processes. In order to achieve timely and accurate detection of downhole faults, a systematic fault detection method is proposed based on the Multivariate Generalized Gaussian Distribution (MGGD) and the Kullback Leibler Divergence (KLD). Uncorrelated components are obtained from the original drilling process signals using the principle component analysis; then, the distribution of components is estimated using the MGGD; afterwards, the KLD is calculated based on a deduced analytic formula; last, the downhole faut is detected by comparing the calculated KLD with the alarm threshold obtained from normal data. The effectiveness and practicality of the proposed method are demonstrated by application to a real drilling process.

Keywords: Geological drilling processes, fault detection, Kullback Leibler Divergence, multivariate generalized Gaussian distribution.

1. INTRODUCTION

Geological drilling is a complex process for geological and energy exploration. However, conditions, such as the small hole diameter, high pressure, and alternation of the hard and soft formations, bring challenges to the drilling processes, e.g., causing serious safety issues and increasing drilling costs (Gan et al., 2019a). For instance, a blowout or kick incident would lead to catastrophic consequences in the offshore drilling (Sule et al., 2019; Nayeem et al., 2016). Moreover, according to (Godhavn, 2010; Willersrud et al., 2015b), non-productive time caused by downhole incidents was more than 20% of the total project time. Therefore, early downhole fault detection in drilling processes is in great demand, so as to ensure drilling safety and reduce costs. In practice, the drilling process monitoring relies on proficient experiences and knowledge of drilling workers; however, such a monitoring form is time and resource intensive, and also lead to increased workload. In order to reduce the cost and the labor intensity, the fault detection for drilling processes has received plentiful attentions.

Existing methods on downhole fault detection are divided into two categories, namely the mechanism model-based and data-driven methods. In the first category, a hydraulic differential equation model was established to kick mitigation in a circulation system (Hauge et al., 2013). Willersrud et al. (2015a) presented a downhole fault detection and isolation scheme by employing analytical redundancy relations. However, the establishment of mancinism models is difficult since the mechanism knowledge and measured variables are usually limited. By contrast, the data driven methods do not require any mechanism knowledge, making it more feasible. Zhang et al. (2018) proposed a real-time fault diagnosis method that extracted qualitative trends first and then diagnosed faults using a multi-class SVM. Li et al. (2020) presented a fault diagnosis method based on multi-time scale features and probabilistic neural networks. Further, Tang et al. (2019) defined two indicators related to the kick incident and conducted online calculation of the probability of the kick. These results were demonstrated to be effective by case studies. But how to further improve the reliability and accuracy of fault detection is an ongoing hard problem, especially in view of the fact that the abnormal data is usually very limited.

In the area of fault detection, the Kullback Leibler Divergence (KLD) has received increasing attentions owing to

⋆ This work was supported by the National Natural Science Foundation of China under Grants 61733016 and 61903345, the Hubei Provincial Technical Innovation Major Project under Grant 2018AAA035, the National Key R&D Program of China under Grant 2018YFC0603405, the Hubei Provincial Natural Science Foundation under Grant 2019CFB251, the 111 project under Grant B17040, and the Fundamental Research Funds for the Central Universities under Grant CUGCJ1812.
† Corresponding author: Weihua Cao (weihuacao@cug.edu.cn)

its capability of detecting incipient faults Aggoune et al. (2016). The analytical form of the KLD between two univariate Gaussian Distributions (GDs) was deduced in (Xie et al., 2015). The KLD was proven to follow the Chi-square distribution, and then the threshold corresponding to the normal operating region was calculated (Zeng et al., 2014). Harmouche et al. (2014); Chen et al. (2018) extended the KLD from the univariate GD to the Multivariate Gaussian Distribution (MGD), where the Principal Component Analysis (PCA) was exploited to obtain main components of process variables; the fault detection problem was solved by calculating the KLD between these components. However, not all process signals can be accurately described by the GD. Therefore, the KLD for univariate zero mean Generalized Gaussian Distribution (GGD) was proposed in (Do and Vetterli, 2002). Xiong et al. (2019) presented a detailed derivation of the analytic form of the KLD between two univariate non-zero mean GGDs; the method was demonstrated to be more effective compared to GD-KLD based method.

The objective of this work is to achieve effective and timely detection of downhole faults in geological drilling processes. Motivated by the above discussions, a systematic downhole fault detection is proposed based the Multivariate Generalized Gaussian Distribution (MGGD) and the Kullback Leibler Divergence. The method consists of four major steps:

(1) Uncorrelated components are extracted from original drilling process signals;
(2) MGGD is used to calculate the distribution of components, and parameters of the MGGD are estimated;
(3) The analytic form of KLD between two MGGDs is obtained;
(4) Downhole fault detection is achieved by comparing the KLD with a threshold determined from normal data.

An industrial case study with historical data from real drilling processes is presented to demonstrate the effectiveness of the proposed method.

The remainder of the paper is organized as follows: Section 2 introduces the geological drilling process and the problem. The method of downhole fault detection is presented in Section 3, including the extraction of uncorrelated components, estimation of the MGGD, calculation of the KLD, and implementation procedures. The industrial case study is given in Section 4, followed by conclusions in the final section.

2. PRELIMINARIES

This section introduces the geological drilling process, discusses the process variables involving the drilling system, and points out the fault detection problem.

2.1 Description of drilling systems

The schematic of a typical geological drilling process is shown in Fig. 1, where the drill rig applies pressure on the rotating drill bit, to break the downhole formation, while the mud pump pumps the drilling fluid into the bottom of the well to carry the broken rocks up to the ground. The Weight On Bit (WOB), HooK Load (HKL), Mud Flow In (MFI) and Stand Pipe Pressure (SPP) are four key variables in geological drilling (Gan et al., 2019b); the WOB and HKL determine the pressure on the drill bit, and the MFI and SPP relate to the drilling fluid flow and pressure. A real geological drilling system is shown in Fig. 2. The ZKLG1 well is a geological exploration well located in Shandong Province, China. The design depth of the well is 3000 meters. The operating rig shown in Fig. 2-(a) is the XY-9 spindle-type core drill. The mud pit shown in Fig. 2-(b) is used to store drilling fluid. Faults were found to occur in this drilling system. For example, a lost circulation appeared when drilling at round the depth of 1900 meters and seriously reduced the drilling efficiency.

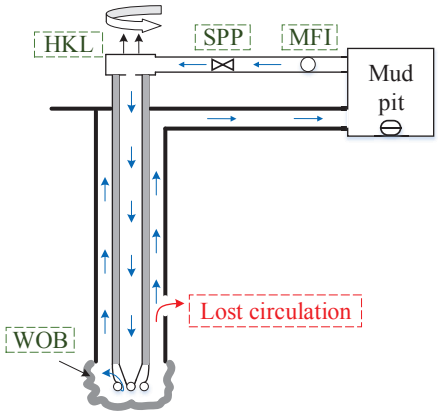

Fig. 1. Schematic of a geological drilling process. The blue arrows represent flow directions of the drilling fluid. The red arrows point to the lost circulation. Dashed green rectangles highlight key process variables.

(a) Drilling rig (b) Mud pit

Fig. 2. A real drilling system for a geological exploration well in Shandong Province, China.

In geological drilling processes, downhole incidents are usually reflected by the changes of process signals, mainly associated with variables in the rotational and circulation subsystems. Due to the uncertainty in downhole formations and the nonlinearity of rock breaking processes, signals in the rotational subsystem are involved with large volatilities. By contrast, signals in the circulation subsystem are relatively stationary during normal drilling. In addition, variables in the same subsystem are correlated, thus the information redundancy exists in process measurement signals. These characteristics pose a challenge to the signal analysis and influence the performance of fault detection.

2.2 Problem description

The fault detection problem is that the condition (normal or faulty) of the drilling process is detected given measured signals of key drilling process variables. A KLD based fault detection method is proposed to solve this problem. Two major phases are involved: In the training phase, uncorrelated components are extracted from historical data at first; then, the MGGD under normal condition is obtained and the threshold of KLD is determined from sufficient normal data. In the test phase, uncorrelated components are extracted online from process signals in a sliding window and the distribution is estimated using the MGGD; then, the KLD between estimated MGGD and the expected MGGD is used to measure the change. If the KLD exceeds the threshold, an alarm is generated.

3. DOWNHOLE FAULT DETECTION METHOD

This section proposes a systematic downhole fault detection method based on the MGGD and KLD. Detailed calculations in each step are presented in the following subsections.

3.1 Extraction of uncorrelated components via PCA

In a drilling system, downhole fault detection requires a comprehensive consideration of multiple process variables. Since there are correlations between process variables of drilling system, the PCA is adopted to obtain uncorrelated principal components at first. The historical data under the normal condition is prepared as a matrix $X_0 = [\mathbf{x}_{0,1}, \mathbf{x}_{0,2}, \cdots, \mathbf{x}_{0,n}] \in \mathbb{R}^{m\times n}$ with n process variables and m samples. In view of that different variables have different volatilities, the matrix X_0 is normalized to $X_1 \in \mathbb{R}^{m\times n}$ with zero means. For each column in X_0, the normalized variable $x_{1,ij}$ is calculated as

$$x_{1,ij} = \frac{(x_{0,ij} - x_{min,i})}{(x_{max,i} - x_{min,i})} - \frac{1}{m}\sum_{j=1}^{m}\frac{(x_{0,ij} - x_{min,i})}{(x_{max,i} - x_{min,i})}, \quad (1)$$

where $i \in \{1, ..., n\}$ and $j \in \{1, ..., m\}$. Variables $x_{max,i}$ and $x_{min,i}$ denote the maximum and minimum of $\mathbf{x}_{0,i}$, respectively.

Then, the covariance matrix S is calculated as (Bakdi et al., 2017)

$$S = \frac{1}{m-1}X_1^T X_1. \quad (2)$$

Using the singular value decomposition, S is rewritten as

$$S = P\Lambda P^T, \quad (3)$$

where Λ represents a diagonal matrix composed by eigenvalues of S ordered in a magnitude decreasing manner, and each columns of $P \in \mathbb{R}^{n\times n}$ corresponds to an eigenvector of an eigenvalue. Since the PCA is used only to multivariate decoupling, the number of retained principal components is set to m. The data matrix X_1 is transformed as

$$X = X_1 P P^T, \quad (4)$$

where $X \in \mathbb{R}^{m\times n}$ is a drilling process components matrix containing uncorrelated columns.

3.2 Parameter estimation for MGGD

In order to describe the characteristics of drilling process variables, the probability distributions of process signals are needed under different conditions. The MGGD provides a suitable tool for signal processing and data analysis. The Probability Density Function (PDF) of the MGGD is expressed as (Verdoolaege and Scheunders, 2012)

$$\mathrm{MGGD}(\mathbf{m}, d, Q, \beta) = \frac{\Gamma(\frac{d}{2})}{\pi^{\frac{d}{2}}\Gamma(\frac{d}{2\beta})2^{(\frac{d}{2\beta})}}\frac{\beta}{|Q|^{\frac{1}{2}}}\cdot \exp\left\{-\frac{[(\mathbf{x}-\mathbf{m})'\mathrm{Q}^{-1}(\mathbf{x}-\mathbf{m})]^{\beta}}{2}\right\}, \quad (5)$$

where d denotes the dimension of the probability space, and $\Gamma(.)$ represents the Gamma function. The other three parameters β, $\mathbf{m}$, and Q denote the shape parameter, the mean vector, and the symmetric positive definite matrix of the MGGD, respectively. Note that $\beta = 1/2$ results in a Laplacian distribution and $\beta = 1$ yields to the MGD. In this study, the mean value of each uncorrelated component is zero, so only the zero-mean condition is considered and therefore $\mathbf{m}$ in (5) can be omitted.

The parameters of MGGD can be obtained using the moment based estimation or the maximum likelihood estimation (Boubchir and Fadili, 2005; Verdoolaege and Scheunders, 2012). In (5), there are two parameters Q and β to be estimated. The input matrix $X = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n]$ is assumed to follow the zero mean MGGD. The symmetric positive definite matrix Q is (Mardia et al., 1982)

$$Q = \frac{d\Gamma(\frac{d}{2\beta})}{2^{1/\beta}\Gamma(\frac{d+2}{2\beta})}\Sigma, \quad (6)$$

where Σ is the covariance matrix. Since each column in X is uncorrelated, Σ is a diagonal matrix comprised by variances of each column as

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \sigma_d^2 \end{bmatrix}, \quad (7)$$

where $d = n$. The variance σ_i is estimated as

$$\hat{\sigma}_i = \frac{1}{m}\sum_{i=1}^{m}(\mathbf{x}_i^T\mathbf{x}_i). \quad (8)$$

Another critical parameter of MGGD is the shape parameter β, which is estimated by the maximum likelihood estimation. The likelihood function is given as follow

$$a(\beta) = L(\mathbf{x}|Q, \beta) = \ln\prod_{j=1}^{m} MGGD(\mathbf{x}_j|Q, \beta). \quad (9)$$

Then, $\hat{\beta}$ is obtained by solving

$$a(\beta)|_{\beta=\hat{\beta}} = \frac{dn}{2\sum_{i=1}^{n}u_i^{\beta}}\sum_{i=1}^{n}\left[u_i^{\beta}In(u_i)\right] - \frac{dn}{2\beta}\left[\Psi(\frac{d}{2\beta}) + In2\right] - n - \frac{dn}{2\beta}In\left(\frac{\beta}{dn}\sum_{i=1}^{n}u_i^{\beta}\right) = 0, \quad (10)$$

where $u_i = \mathbf{x}_i^T Q^{-1}\mathbf{x}_i$, and $\Psi(.)$ denotes the Diagamma function. The (6) and (10) can be solved numerically

by using a recursive solution of the maximum likelihood equations (Boubchir and Fadili, 2005; Verdoolaege and Scheunders, 2012).

3.3 Calculation of the KLD between MGGDs

The KLD is used to measure the distance between two PDFs (Xie et al., 2015), and is adopted here to capture statistical changes on distributions of uncorrelated components in $\tilde{X}$. The KLD between the MGGD(Q,β) obtained from historical data X under normal condition and the MGGD$(\hat{Q},\hat{\beta})$ of online data $\tilde{X}$ is given by

$$\begin{aligned} &D_{kl}\left(\text{MGGD}(\hat{Q},\hat{\beta}),\text{MGGD}\left(Q,\beta\right)\right) \\ &\triangleq \int_{-\infty}^{+\infty}\text{MGGD}(\hat{Q},\hat{\beta})\ln\frac{\text{MGGD}(\hat{Q},\hat{\beta})}{\text{MGGD}\left(Q,\beta\right)}\mathrm{d}x. \end{aligned} \tag{11}$$

Based on the MGGD(Q,β) estimated from the normal samples, the fault detection problem is to test if the MGGD$(\hat{Q},\hat{\beta})$ estimated from $\tilde{X}$ is normal, i.e.,

$$\begin{cases} D_{kl} < D_{th} : \tilde{X} \text{ is under normal condition,} \\ D_{kl} \geq D_{th} : \tilde{X} \text{ is not under normal condition,} \end{cases} \tag{12}$$

where D_{th} is the alarm threshold to be determined.

According to (Xiong et al., 2019), the KLD for univariate GGD is shown in Lemma 1.

Lemma 1. The KLD $D_{kl,\beta}$ between GGD(σ^2,β) and GGD$(\hat{\sigma}^2,\hat{\beta})$ under the assumption that $\beta=\hat{\beta}$ has a simplified form as

$$\begin{aligned} D_{kl,\beta} =& \ln\frac{\hat{\beta}\sigma\Gamma(1/\beta)\sqrt{\Gamma(1/\beta)\Gamma(3/\hat{\beta})}}{\beta\hat{\sigma}\Gamma(1/\hat{\beta})\sqrt{\Gamma(1/\hat{\beta})\Gamma(3/\beta)}}-\frac{1}{\hat{\beta}} \\ &\left(\frac{\hat{\sigma}\sqrt{\Gamma(1/\hat{\beta})\Gamma(3/\beta)}}{\sigma\sqrt{\Gamma(1/\beta)\Gamma(3/\hat{\beta})}}\right)^{\beta}\frac{\Gamma(\beta/\hat{\beta}+1/\hat{\beta})}{\Gamma(1/\hat{\beta})}. \end{aligned} \tag{13}$$

Two methods were proposed in (Xiong et al., 2019) to determine the thresholds, including a constant threshold and an adaptive threshold. The main difference between the two methods is whether to assume the shape parameters of the two distributions are the same or not. In this work, β is assumed to be not deterministic, i.e., $\hat{\beta}=\beta$. Generalizing to the multivariate case, Lemma 2 is proposed as follows:

Lemma 2. Given d uncorrelated variables that follow the MGGD, the KLD under the assumption that $\beta=\hat{\beta}$ is given by

$$D_{kl,\beta}=\sum_{i=1}^{d}[-\frac{1}{2}In(\frac{\hat{\sigma_i}}{\sigma_i})+\frac{1}{\beta}(\frac{\hat{\sigma_i}}{\sigma_i})^{\beta}-\frac{1}{\beta}], \tag{14}$$

where $\sigma_i\in\{\sigma_1,...,\sigma_d\}$. Then, the KLD in (14) is to be used as the test statistic, and the key step is to find the PDF of $D_{kl,\beta}$ to determine the D_{th}.

3.4 Determination of the alarm threshold using KDE

There are two ways to calculate the PDF of the KLD, including the analytic calculation and the non-parametric estimation. Kernel Density Estimation (KDE) is a non-parametric method to calculate the PDF without any predefined distribution (Gonzalez et al., 2015). Since it is hard to calculate the analytical form of the PDF of D_{kl} in (11), the KDE is used to estimate it under the normal condition. Let $y_1,y_2,...,y_N$ denote N i.i.d samples collected from the real density function $p(y)$ of $D_{kl,\beta}$ in (14). The estimated function $\hat{p}(y)$ is calculated by summing kernels for y, i.e.,

$$\hat{p}(y)=\frac{1}{n}\sum_{i=1}^{N}\frac{1}{h}K(\frac{y-y_i}{h}), \tag{15}$$

where $K(\cdot)$ represents the kernel function, N stands for the number of KLD samples, y_i denotes the ith point in Y, and h indicates the bandwidth parameter. In this study, the Gaussian kernel is adopted to achieve smooth estimation of the PDF. A commonly used formula to determine the bandwidth h of the Gaussian kernel is given by (Gonzalez et al., 2015)

$$h=\left(\frac{4\sigma_{kl}^5}{3N}\right)^{\frac{1}{5}}, \tag{16}$$

where σ_{kl} is the variance of the KLD in normal condition and N is the number of KLD samples. Based on the PDF in (15) under the normal condition, a test under a confidence level α is used to determine D_{th} in (12).

3.5 Performance assessment

The performance of the fault detection can be evaluated by metrics, such as the detection delay, False Alarm Rate (FAR), and Missed Alarm Rate (MAR). A detection delay represents the time interval from when a fault actually appears to when it is detected. FAR and MAR are exploited to calculate the accuracy of fault detection. The sample based calculations of FAR and MAR are given by

$$\text{FAR}=\frac{n_{an}}{n_{tn}}\times 100\%, \tag{17}$$

$$\text{MAR}=(1-\frac{n_{af}}{n_{tf}})\times 100\%, \tag{18}$$

where n_{an} is the number of alarm samples under the normal condition, n_{tn} gives the number of total samples under the normal condition, n_{af} denotes the number of alarm samples under faulty conditions, and n_{tf} indicates the number of total samples under faulty conditions.

4. INDUSTRIAL CASE STUDY

To illustrate the effectiveness and practicality of the MGGD-KLD based fault detection method in geological drilling processes, the drilling data from the ZKLG1 well was used in this case study. The sampling period of the real drilling process was 1 s. The 4 involved drilling process signals are shown in Fig. 3. A lost circulation fault occurred at t=2000 s after the drilling worker had changed the WOB. Then, significant variations were shown in signals of WOB and HKL, and slight changes were shown in the signal of SPP. The MGGD-KLD based fault detection method was applied. In the foremost step, uncorrelated components were extracted using PCA from the four drilling signals.

For comparison, the distribution of the collected data after PCA was approximated using the MGGD and MGD,

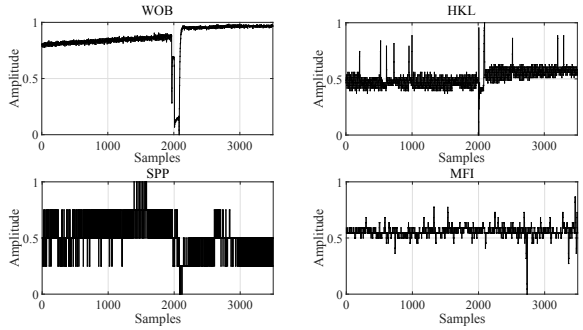

Fig. 3. Signals of WOB, HKL, SPP, and MFI with the sampling interval 1 s. The lost circulation fault occurred at t=2000 s.

respectively. An example for a single variable under the GD and GGD is presented in Fig. 4, where the histogram was obtained from the normal training data. Because the MGGD can adjust shapes of the top and the tail by changing β, it can be seen that the solid red curve with MGGD is more suitable to describe the histogram compared to the dashed blue curve with MGD.

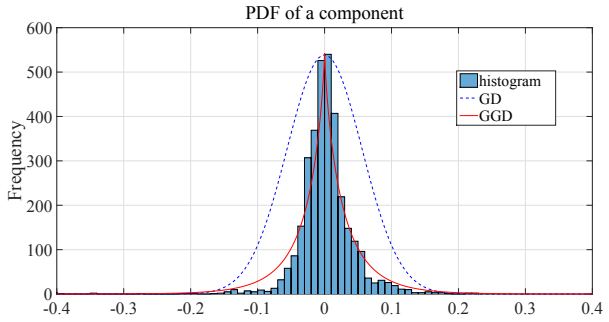

Fig. 4. An exmaple of the estimated PDFs and the histogram. The dashed blue line represents the estimated GD and the solid red line denotes the estimated GGD.

In this part, 2500 samples under the normal condition were used as the training set to obtain the reference MGGD(Q, β) and the alarm threshold D_{th}. The data was divided into two parts, i.e., one part was used as a test set while the rest served as a template. Then, the PDF of the KLD under the normal condition was estimated by the KDE in Section 3.4 with the kernel length $h = 0.023$. Accordingly, the alarm thresholds were calculated for MGGD and MGD (shown as the dashed horizontal lines in Figs. 5 and 6, respectively). Furthermore, the sliding window with the length of 100 samples was adopted to analyze online samples. The fault detection results by MGGD-KLD and MGD-KLD are shown as the black curves in Figs. 5 and 6, respectively. Samples on the right of the blue vertical line indicate the faulty condition. The dashed red horizontal line represents the alarm threshold at the significance level of 0.05. The thresholds for MGGD-KLD and MGD-KLD are 1.65 and 4.50, respectively. In addition, the T^2 statistic and the SPE statistic (shown in Fig. 7) in the classical PCA fault detection (Ding, 2014) were calculated for comparison. The alarm thresholds at the significance level of 0.05 were thresholds $T^2_{th} = 7.85$ and $Q_{th} = 5.79$ for the T^2 and SPE statistics, respectively.

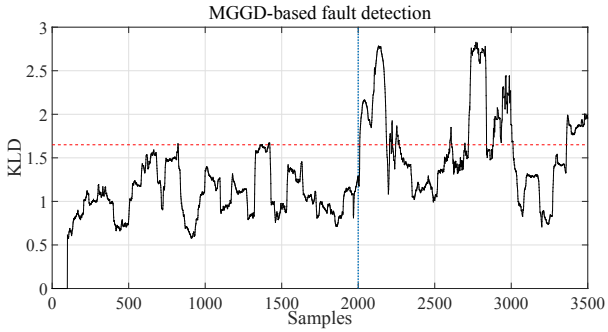

Fig. 5. Fault detection result using the MGGD-KLD. The horizontal red dashed line represents the alarm threshold and the vertical dotted blue line stands for the fault occurrence time instant.

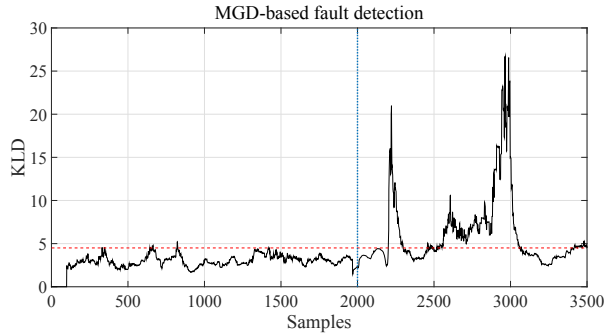

Fig. 6. Fault detection result using the MGD-KLD.

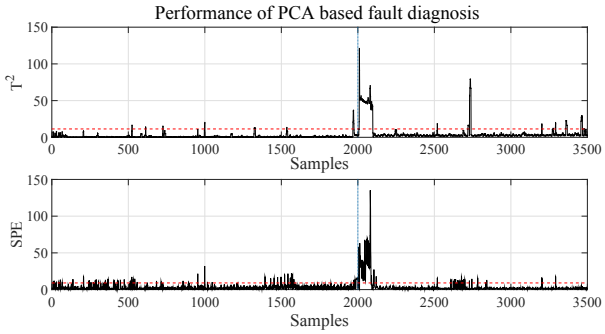

Fig. 7. Fault detection results based on the T^2 statistic and the SPE statistic.

To better understand the fault detection results, the FAR, MAR, and detection delay were calculated to assess the performance of the above 3 methods. Since the influence of the fault on WOB and HKL signals was short, the 1400 samples before and 140 samples after the fault were exploited in the calculation. As summarized in Table 1, the MGGD-KLD has the lowest FAR and MAR compared to all the other three methods. There is only 7 s detection delay using the MGGD-KLD. Even though not the smallest, it is close to the detection delay using the PCA-based method. It can also be found that the fault was not detected over the studied period using the MGD-KLD. This is probably because that the estimation of the data distribution based on MGD has a large fitting error shown in Fig. 4, causing the erroneous conclusion. In summary, the fault was detected in time using either SPE or T^2 statistic, but FAR and MAR are relatively large. By contrast, the MGGD-KLD based method detected downhole

fault with low FAR and MAR, as well as a small detection delay. Based on the above 3 indices, it can be found that the proposed method outperformed other methods with a superior performance.

Table 1. Comparison of the results using different fault detection methods.

Method	FAR(%)	MAR(%)	Detection delay(s)
SPE	8.56	11.43	6
T^2	2.05	13.33	6
MGD-KLD	1.10	100.00	-
MGGD-KLD	0.89	5.41	7

5. CONCLUSION

This work proposes a downhole fault detection method for geological drilling processes based on the MGGD and KLD. The MGGD is adapted to estimate the PDF of the uncorrelated components obtained from original drilling signals after PCA. The KLD between the MGGD estimated from training data under the normal condition and the MGGD of online data is calcuated to capture changes of process signals. The PDF of the KLD under the normal condition is estimated using the KDE, such that the alarm threshold of the KLD is determined at a certain significant level. According to the results from the industrial case study, successful detection was achieved with low FAR and MAR by the proposed method, which has significant better performance compared to the conventional PCA based and MGD-KLD based methods.

REFERENCES

L. Aggoune, Y. Chetouani, and T. Raissi. Fault detection in the distillation column process using Kullback Leibler divergence. *ISA Transactions*, 63:394–400, 2016.

A. Bakdi, A. Kouadri, and A. Bensmail. Fault detection and diagnosis in a cement rotary kiln using PCA with EWMA-based adaptive threshold monitoring scheme. *Control Engineering Practice*, 66:64–75, 2017.

L. Boubchir and J. M. Fadili. Multivariate statistical modeling of images with the curvelet transform. In *Proceedings of the Eighth International Symposium on Signal Processing and Its Applications.*, volume 2, pages 747–750, Aug 2005.

H. Chen, B. Jiang, and N. Lu. An improved incipient fault detection method based on Kullback-Leibler divergence. *ISA Transactions*, 79:127–136, 2018.

S. X. Ding. *Data-driven design of fault diagnosis and fault-tolerant control systems*. Springer, 2014.

M. N. Do and M. Vetterli. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE Transactions on Image Processing*, 11 (2):146–158, 2002.

C. Gan, W. Cao, and M. Wu et. al. Two-level intelligent modeling method for the rate of penetration in complex geological drilling process. *Applied Soft Computing*, 80: 592–602, 2019a.

C. Gan, W. Cao, and M. Wu et. al. Prediction of drilling rate of penetration (ROP) using hybrid support vector regression: A case study on the shennongjia area, central china. *Journal of Petroleum Science and Engineering*, 181:106200, 2019b.

J.-M. Godhavn. Control requirements for automatic managed pressure drilling system. *SPE Drilling & Completion*, 25(3):336–345, 2010.

R. Gonzalez, B. Huang, and E. Lau. Process monitoring using kernel density estimation and Bayesian networking with an industrial case study. *ISA Transactions*, 58:330–347, 2015.

J. Harmouche, C. Delpha, and D. Diallo. Incipient fault detection and diagnosis based on Kullback-Leibler divergence using principal component analysis: Part I. *Signal Processing*, 94:278–287, 2014.

E. Hauge, O.M. Aamo, J. M. Godhavn, and G. Nygaard. A novel model-based scheme for kick and loss mitigation during drilling. *Journal of Process Control*, 23(4):463–472, 2013.

Y. Li, W. Cao, W. Hu, and M. Wu. Diagnosis of downhole incidents for geological drilling processes using multi-time scale feature extraction and probabilistic neural networks. *Process Safety and Environmental Protection*, 137:106–115, 2020.

K. V. Mardia, J. T. Kent, and J.M. Bibby. *Multivariate analysis*. London: Academic Press, 1982.

A. A. Nayeem, R. Venkatesan, and F. Khan. Monitoring of down-hole parameters for early kick detection. *Journal of Loss Prevention in the Process Industries*, 40:43 – 54, 2016.

I. Sule, S. Imtiaz, F. Khan, and S. Butt. Risk analysis of well blowout scenarios during managed pressure drilling operation. *Journal of Petroleum Science and Engineering*, 182:106296, 2019.

H. Tang, S. Zhang, F. Zhang, and S. Venugopal. Time series data analysis for automatic flow influx detection during drilling. *Journal of Petroleum Science and Engineering*, 172:1103–1111, 2019.

G. Verdoolaege and P. Scheunders. On the geometry of multivariate generalized Gaussian models. *Journal of Mathematical Imaging and Vision*, 43(3):180–193, 2012.

A. Willersrud, M. Blanke, and L. Imsland. Incident detection and isolation in drilling using analytical redundancy relations. *Control Engineering Practice*, 41:1–12, 2015a.

A. Willersrud, M. Blanke, L. Imsland, and A. Pavlov. Fault diagnosis of downhole drilling incidents using adaptive observers and statistical change detection. *Journal of Process Control*, 30:90–103, 2015b.

L. Xie, J. Zeng, U. Kruger, X. Wang, and J. Geluk. Fault detection in dynamic systems using the Kullback-Leibler divergence. *Control Engineering Practice*, 43: 39–48, 2015.

Y. Xiong, Y. Jing, and T. Chen. Abnormality detection based on the Kullback-Leibler divergence for generalized Gaussian data. *Control Engineering Practice*, 85:257–270, 2019.

J. Zeng, U. Kruger, J. Geluk, X. Wang, and L. Xie. Detecting abnormal situations using the Kullback-Leibler divergence. *Automatica*, 50(11):2777–2786, 2014.

X. Zhang, L. Zhang, and J. Hu. Real-time diagnosis and alarm of down-hole incidents in the shale-gas well fracturing process. *Process Safety and Environmental Protection*, 116:243–253, 2018.