Research on Overflow Accident Early Warning Model Based on Improved DBSCAN Clustering for Export Instantaneous Flow

Haibo Liang [a], Haibing Huang [b], Rui Quan [c], Zhenghua Xu [d]

College of Mechanical and Electrical Engineering, Southwest Petroleum University, ChengDu 610500, China.

[a] secondbo@126.com, [b] 983046868@qq.com, [c] 849474525@qq.com,

[d] 781386897@qq.com

Abstract

Oil and gas are still a necessity for social production activities nowadays, but their exploration and exploitation are extremely complicated and dangerous. The overflow accident is one of the most threatening accidents for safe drilling operations in the process of oil and gas exploitation. If early warning can be made timely and accurately to reduce or avoid the occurrence of the overflow accidents in the early stage of the overflow, it is very important for safe drilling. Based on the export instantaneous flow monitoring method and its interference factors, the monitoring parameters of the system are determined, the model of eliminating interference factors is established, and the real-time dynamic correction of the instantaneous flow data is realized. Combined with the pattern recognition technology, an early warning recognition model of overflow is established by using dynamic clustering method. And the early warning model is simulated and analyzed by using the data of overflow accident section in drilling site, which verifies the accuracy of the early warning model monitoring results and provides engineering basis for the early overflow accident monitoring and diagnosis.

Keywords

Early overflow; monitoring and warning; outlet instantaneous flow; dynamic clustering method.

1. Introduction

At present, experts and scholars at home and abroad have done a lot of research on early flood warning technology and achieved certain results. Early foreign overflow warning technology mainly relies on improving monitoring equipment or real-time monitoring of bottom hole parameters to determine early overflow [23]. The monitoring effect is more accurate and intuitive than the traditional methods, but most of the underground measuring devices are comprehensive and difficult to implement, and the monitoring cost is high. The requirements for on-site professional and technical personnel involved in system debugging, proofreading, maintenance, etc., are also very high. Especially in ultra-deep well drilling, the instrument will be affected by the conditions of high temperature and high pressure at bottom hole, limited data transmission capacity and extremely high monitoring cost. The downhole measurement technology in early overflow warning is greatly restricted in industrialization and application [24]. The early overflow warning technology in China mainly relies on real-time monitoring of ground drilling engineering parameters, mud circulation parameters, and changes in gas measurement parameters to identify overflow accidents. However, the judgment method mainly relies on field personnel to complete with personal experience, lacking scientific and systematic theoretical basis. Moreover, the effect of this artificial judgment will vary according to the level of the staff, the level of experience and knowledge, the strong and weak sense of responsibility, and the mental state of work. Therefore, it is difficult to guarantee the timeliness, accuracy and reliability of early overflow warning results, and there are many shortcomings and limitations in monitoring sensitivity and universality. Although some domestic and early artificial overflow warning methods have been gradually tried in the near future, they are almost only in the laboratory research stage. Therefore, it is very important to find a method that is suitable for ground application and more timely and accurate to detect overflow and alarm than existing ground measurement methods. Therefore, this paper will not rely on underground measuring instruments but the ground measurement method, and change the principle of the existing overflow ground monitoring method, and propose a set of suitable for ground applications, which is earlier, faster and more accurate than the existing ground measurement method. A set of early warning system of overflow accident is finally formed, which can identify and alarm the overflow before it is moved to the wellhead, and provide technical support for the realization of safe drilling.

2. The establishment of an early warning model of overflow accident based on instantaneous flow of export

In order to establish a complete and reasonable early overflow warning model based on outlet instantaneous flow parameters during normal drilling, it is necessary to eliminate the interference factors of the instantaneous flow change caused by the non-overflow and combine it with the pattern recognition technology to judge the overflow accident. The factors that cause the instantaneous flow change of the outlet are the bottom hole overflow, the thermal expansion effect of drilling fluid, the change of mud pump displacement and the variation of cuttings concentration in annulus. The specific early warning model is shown in Figure 1:

1. According to the instantaneous flow value of the outlet collected from the drilling site, combined with the thermal expansion effect of the fluid, the change of the pump displacement and the flow variation caused by the annulus debris concentration, the corrected instantaneous flow rate value is obtained by excluding the three types of interference factors. All the changes in the instantaneous flow rate of the outlet are attributed to the bottom hole overflow factor;
2. Combined with pattern recognition technology, the modified instantaneous flow value is synthetically analyzed by using the overflow identification model to judge whether the overflow accident occurs or not
3. According to the diagnostic results of the identification model, if overflow occurs, an early warning signal of overflow is issued.
4. If there is no overflow, return to the first step and continue the loop judgment.

Based on this model, the early warning of overflow can be achieved quickly and accurately.

2.1 Study on the model of thermal expansion effect of drilling fluid

According to the thermal expansion effect of the drilling fluid, the calculation model of the volume change of the well fluid caused by the thermal expansion of the drilling fluid in the vertical section is derived and simplified as follows:

$$\Delta V = \alpha V \frac{t_2 - t_0}{1 + \frac{cm \ln(r_2 / r_1)}{2\lambda \pi L}}$$

The calculation model of fluid volume change in horizontal section caused by thermal expansion of drilling fluid is as follows:

$$\Delta V = \alpha V_{sp} G h_{cz}$$

It can be seen that the mud volume change can be obtained by combining the above static calculation data with the known static data and real-time data (Fig.-2). The static data includes mud thermal expansion coefficient, formation thermal conductivity, mud specific heat capacity, geothermal gradient, and well structure data; real-time data includes mud inlet density and well depth data. Based on this model, the effect of thermal expansion on the instantaneous flow at outlet can be weakened or even eliminated.

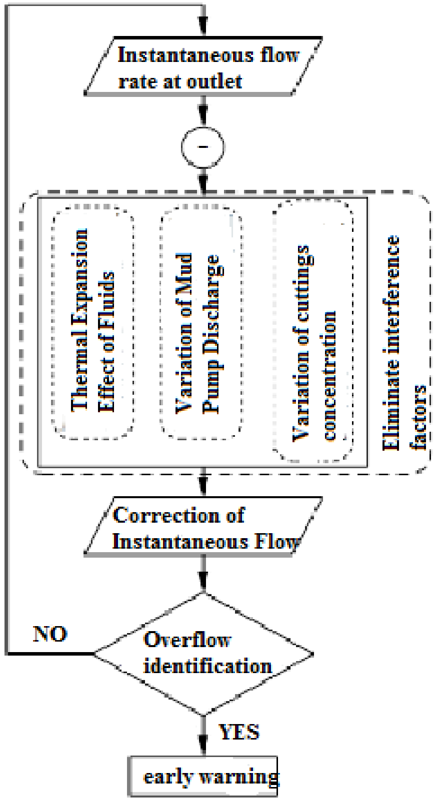

Figure 1 Overflow accident warning model

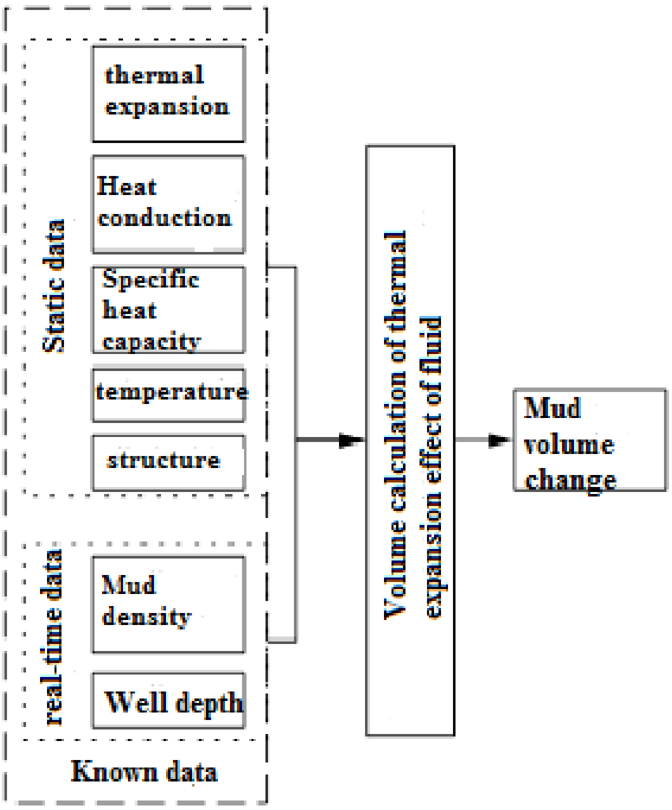

Figure 2 Mud volume change calculation process

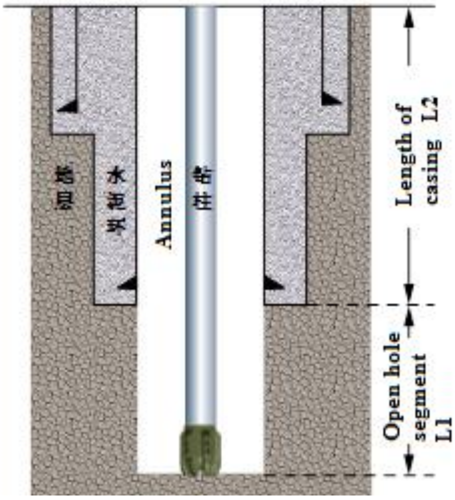

Figure 3 Well structure segmentation diagram

In order to simplify the calculation process and optimize the calculation results, the well body is divided into the casing section and the open hole section as shown in Fig. 3 to calculate the fluid thermal expansion effect model. In the above model, L is divided into L1 and L2 and the corresponding thermal conductivity of different strata is used to calculate the volume variation of mud.

2.2 Study on the Elimination Model of Mud Pump Displacement Factors

According to the variation factor of mud pump displacement, the calculation model of mud pump displacement is derived and simplified as follows:

Single acting pump displacement is

$$Q = \frac{1}{4}\pi\beta mnLD^2$$

Double acting pump displacement is

$$Q = \frac{1}{4}\pi\beta mnL\left(2D^2 - d^2\right)$$

Mud pump displacement can be obtained from the above calculated model based on known pump parameters and real-time data. The pump parameters include pump displacement coefficient, piston rod diameter, cylinder liner diameter, piston stroke, pump cylinder number; real-time data take pump strokes. Based on this model, the interference of slurry pump displacement on the instantaneous flow of the outlet can be reduced or even eliminated.

2.3 Study on the method for eliminating the variation of annulus cuttings concentration

In order to eliminate the influence of annulus cuttings concentration change on the instantaneous flow rate at outlet, the factors of annulus cuttings concentration variation can be considered. There is no direct relationship between the change of the inclination angle of the well and the change of the mud outlet flow, which can be neglected. The effect of pump displacement on the concentration of cuttings in annulus is much smaller than that on instantaneous flow at outlet, and the effect of pump displacement on instantaneous flow at outlet is more direct and obvious. The displacement factor of pump can also be neglected. The remaining factors of mechanical drilling rate and wall rock caving are all the causes of the increase of cuttings caused by the formation rock entering into the well. Therefore, the annulus cuttings concentration is used to exclude changes in outlet flow caused by formation rock entering the wellbore.

According to the actual experience of drilling construction site, the variation of cuttings concentration in annulus is the smallest in comparison with the effect of fluid thermal expansion and pump displacement on the instantaneous flow rate at the outlet, and its value is directly given by the field. At the scene, the concentration of the annulus cuttings is usually obtained by weighing method, that is, the wellhead returns the cuttings at intervals and the average value of the wells is taken, and then the annulus debris concentration is calculated based on the annulus volume. In this paper, the ratio of the annulus cuttings volume Vs to the wellbore annulus volume Vs is used

to represent the annulus debris concentration $C_a = \frac{V_s}{V_a}$. Therefore, the instantaneous flow rate at the outlet caused by the variation of cuttings concentration in the annulus is as follows

$$\Delta V = \Delta V_s = \Delta C_a \bullet V_a .$$

According to the annulus cuttings concentration value given at the site, the mud flow variation caused by the increase of annulus cuttings can be deduced, and the instantaneous flow rate is corrected accordingly.

3. Research on overflow recognition model using dynamic clustering method

According to the interference factor elimination model of the instantaneous flow rate of the exit, the corrected instantaneous flow value can be known. Combined with the principle of pattern recognition, a fast early warning recognition model of early overflow accident based on dynamic clustering method is established. By using this model, the modified instantaneous flow value of exit can be synthetically analyzed, and the real time early warning results can be obtained.

3.1 Fundamental theoretical research on cluster analysis

The traditional Bayesian classification and probability statistics are all supervised classifications by teachers, and they all have their own recognition advantages. First, the sample of the known category is used as the training set, or the corresponding discriminant function is designed by using the class conditional probability density function known or obtained from the training sample set, and then the unknown sample is classified by the function.

However, when the training sample set is not available, there is no way for the teacher-supervised taxonomy to classify unknown samples. At this time, we must seek a classification method without teachers and unsupervised and cluster analysis is one of them. This method is the embodiment of the idea that “things are gathered together and people are divided into groups”. When the discriminant function is designed by the cluster analysis method, the category of the discriminant object is unknown, and the clustering criterion is reasonably selected and automatically classified according to the degree of similarity between samples. According to the clustering process of determining the final attribution category of the object, the cluster analysis method can be divided into the hierarchical clustering method and the dynamic clustering method.

According to the characteristics of each clustering algorithm, combined with the characteristics of the instantaneous flow data of export which the system needs to deal with, there will be a small range fluctuation in a certain range of categories. Based on the advantages of DBSCAN clustering algorithm which is insensitive to noise and high speed, DBSCAN clustering algorithm based on density is chosen as the core algorithm of the early warning model of overflow, in order to realize overflow early warning.

3.2 Application of dynamic clustering method to identify the technical route design of overflow accident

The DBSCAN clustering algorithm is applied to realize the pressure-controlled drilling overflow warning. The specific circuit diagram is as follows:

1.The corrected instantaneous flow values of the outlets obtained by excluding the interference factors are chronologically distributed in the same coordinate system;

2.Set the neighborhood parameters $(Minpts, Eps)$ of the DBSCAN clustering algorithm and the sensitivity of the overflow accident identification S ;

3.Entering the DBSCAN clustering algorithm to obtain clusters of export instantaneous traffic clusters;

4.Using the least squares method to obtain the flow change trend value T_i corresponding to the cluster fitting cluster;

4.Contrast the flow change trend value T_i and the overflow accident identification sensitivity S . When $T_i \geq S$, it is determined that an overflow accident has occurred, and when $T_i < S$, it is determined that no overflow accident has occurred.

3.3 Research on Overflow Accident Identification Model Algorithm Based on Dynamic Clustering Method

The clustering algorithm is the best choice under the prior information of the known overflow accident and the law graph of the instantaneous flow at the exit. Clustering analysis is based on the similarity of samples to cluster the data to find the distribution characteristics of object space. Among them, the DBSCAN clustering algorithm uses a center-based density definition method. The algorithm divides the regions with high density into one class according to the density of the sample data, and therefore, it is possible to find clusters of arbitrary shape in the spatial samples with "noise". The clustering method has the advantages of less input parameters, the user can adjust freely according to the need, and the number of clusters is not known in advance, so it is suitable for classifying the sample data of unknown content. The algorithm requires a given data set c , $Minpts$ and Eps , as defined below.

Definition 1 User-specified parameters $Eps > 0$ are used to specify the radius of the field for each object.

Definition 2 If an object $Eps-$ contains $Minpts$ object in its domain, the object is judged as the core object.

Definition 3 When p is satisfied in the $Eps-$ field of q , and q is the core object, the object a to the object c are directly reachable in density.

Definition 4 If there is an object chain $p_1, p_2, \cdots, p_n$, such that $p_1 = q, p_n = p$, and p_{i+1} is directly reachable from p_i ($0 < i < n$), then the object p is reachable from the object q .

Definition 5 If there is an object q such that objects p_1 and p_2 are both reachable from q with respect to Eps and $Minpts$, then objects p_1 and p_2 are connected to the density of Eps and $Minpts$.

3.4 Improvement of DBSCAN clustering

The traditional DBSCAN clustering has the advantages of fast clustering speed and high efficiency, but it needs to be improved as follows in the early warning of overflow accidents:

1.The traditional DBSCAN clustering algorithm uses the Euclidean distance calculation as the sample metric. For the early warning method, the trend of the instantaneous flow of the exit is studied. The slope value of the data point is used as the clustering principle. There are disturbances in the process of underground and ground factors. The instantaneous flow data of the outlet collected by the sensor may fluctuate within a small range. In order to improve the clustering accuracy, when the pressure data is clustered, the pressure slope and the slope change continue. Time is taken as the slope value of the point, and its formula is expressed as Equation 4-1.

$$y'(t) = \sum_{i=1}^{2} b_i (y(t+i) - y(t-i))$$

$$b_i = \frac{12i}{(N^2-1)N}, N = 2K+1, i = 1, 2, \cdots, K$$

1.The traditional DBSCAN clustering algorithm determines the threshold parameters Eps and $Minpts$ of the density is globally unique

2.The traditional DBSCAN clustering algorithm scans the whole world. For the regularity and timing of the distribution of the instantaneous flow curve of the exit, the scanning is performed in chronological order, thereby reducing the time consumption of the clustering process.

After clustering the exit instantaneous flow curve by improved DBSCAN clustering method, the slope of the straight line segment is fitted by least square method to represent the trend of the exit instantaneous flow curve change Ti,. The Ti is compared with the set sensitivity S to judge whether the overflow accident occurs or not.

3.5 Dynamic Clustering Algorithm Simulation Analysis

In order to verify the accuracy and effectiveness of the selected dynamic clustering algorithm in identifying early overflow accidents, this paper intercepts the historical data of an overflow accident in a well at the drilling site and does a lot of simulation analysis on the identification model. The specific work is as follows.

4. Analysis of the Influence of Time Factor on Clustering Results

According to the previous study, the corrected instantaneous flow values after eliminating the interference factors are distributed in time series in the same coordinate system. Therefore, when clustering the flow data, it is necessary to consider not only the relationship between the flow values but also the flow rates. The time corresponding to the value is analyzed.

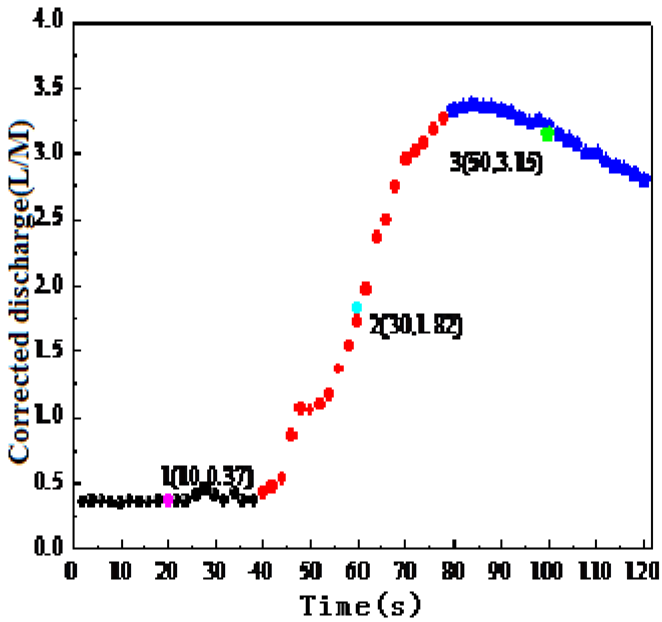

Figure 4 Data clustering results in 120S during overflow accidents

As shown in the figure, the flow rate is about 45s as the inflection point of the overflow accident. When the flow intrusion process ends at about 90s, the flow value begins to fall. After fully considering the time factor, the flow data is clustered, and the clustering effect is in line with the requirements. It is ideal, especially at the inflection point of 45s and 90s, the clustering effect is quite good.

5. Comparative analysis of the effect of clustering method and adjacent point method on overflow judgment

The corrected instantaneous flow data combined with the clustering method is used to estimate the flow change trend, and the overflow accident is judged by this. For the flow change process, the overflow value caused by the sudden change of the flow value or the inflection point of the flow change is misjudged due to various reasons. Has a good inhibitory effect. Figure 5 is a comparison chart of flow rate trends obtained by the neighboring point method and clustering method when the flow rate is relatively stable under normal drilling conditions. Assumption:

The neighboring point method (blue) refers to the slope value of the flow change obtained from two adjacent flow points;

The core method (yellow) refers to using the cluster centers to obtain the slope value of the flow change;

The segmentation method (red) refers to the use of each cluster member segment to obtain the flow rate change slope value;

Overflow accident identification sensitivity (green) S=0.02 L/s2, which can be adjusted in real time according to the actual situation on site.

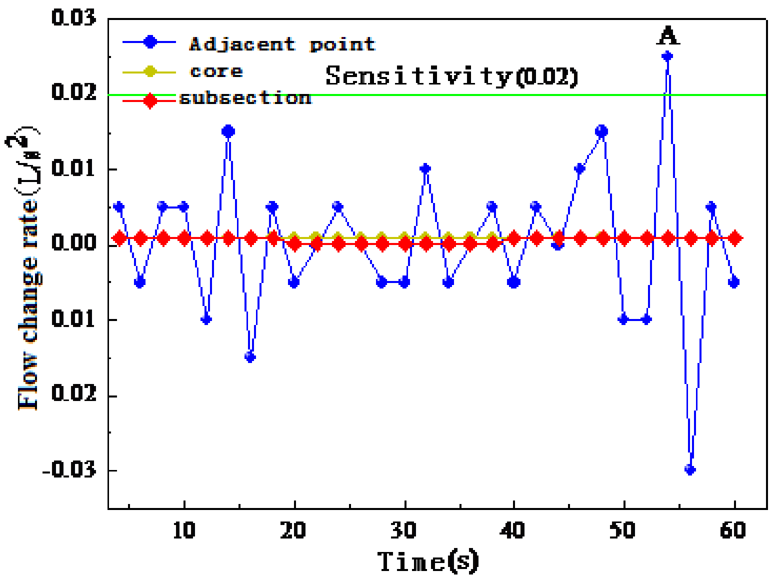

Figure 5 Comparison of stable data segment recognition effects

Table 1 Comparison of different methods to obtain the slope value of flow change

Method	subsection		core	Adjacent points	Sensitivity
Flow Change rate (L/s^2)	1	0.001	0.001	-0.03～0.025	0.02
	2	0.0001			
	3	0.001			

In the normal state, at the sudden change point and the inflection point of the flow value, the slope value of the flow change is obtained by the neighboring point method (blue), and the fluctuation value is even exceeded, and the false alarm is triggered even more than the overflow recognition sensitivity value (point A in the figure); The method (yellow) and the segmentation method (red) find that the flow rate change slope value is consistently below the overflow accident identification sensitivity value (green), that is, no overflow accident occurs, and the recognition effect is correct and effective.

Fig. 6 is a variation slope value obtained by using different methods in the flow data of FIG. 4 when the overflow accident occurs. The flow value is in a steady state during 0 to 20 s, and is continuously increasing in the range of 20 to 60 s. Under normal circumstances, the flow change slope value should be less than the overflow accident identification sensitivity value (green) during the period of 0 to 20 s, and should be greater than the overflow accident identification sensitivity value within 20 to 60 s. Using the neighboring point method (blue) to obtain the flow rate change slope value constantly fluctuates, causing false alarms at points A and B, and triggering leakage alarms at points C and D, which is the early warning result of engineering avoidance and failure; The method (yellow) finds that the slope value of the flow change is greater than the sensitivity value, so it is misjudged during the period of 0 to 20 s; however, the segmentation method (red) finds that the slope value of the flow change is lower than the sensitivity value during the period of 0 to 20 s, 20~ The sensitivity value is higher than 60s, which means that the overflow accident occurs at this time, and the recognition effect is correct and effective.

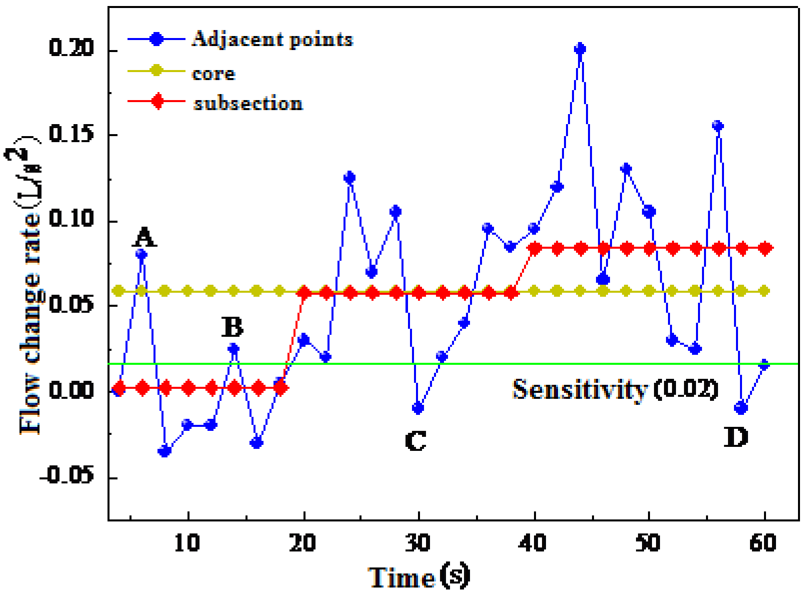

Figure 6 Comparison of the recognition effect of the overflow accident segment

Table 2. Comparison of different methods to obtain the slope value of flow change

Method	subsection		core	Adjacent points	Sensitivity
Flow Change rate (L/s^2)	1	0.002	0.059	-0.035～0.2	0.02
	2	0.058			
	3	0.084			

6. Conclusion

Using segmentation method to identify overflow accidents is more stable and reliable than adjacent point slope discrimination method and cluster core fitting method. The recognition accuracy is higher, which can effectively reduce or even avoid false positives and false negatives. Therefore, in the design of the overflow accident identification technology route, the segmentation method is used to fit the flow change slope to judge the overflow or not.

References

[1] ZhiWei Li,HuiXin Liu,Early micro-overflow monitoring technology in deep wells [J]. Natural Gas Technology and Economy,2011,03:29-31+78.

[2] XiaoYuan Lian,Research on Fault Detection and Diagnosis in Drilling Process [D]. Dalian University of Technology,2013.

[3] YongSheng Cheng,Formation pressure theory and evaluation [M].BeiJin: Petroleum Industry Press,1990.

[4] Huo Huang,Blowout Accident: A Review of Sudden Disasters [J]. Look at Newsweek, 2004,01:22-25.

[5] Xin Cheng, Causes and Reflections on the Extra Large Blowout Event and Its Serious Consequences in Kaixian County [J]. Modern preventive medicine,2007,12:2229-2231.
[6] Haibo Liang, Tang Yongqiang, Li Xiang and Luo Yangyang.RESEARCH ON DRILLING KICK AND LOSS MONITORING METHOD BASED ON BAYESIAN CLASSIFICATION [A]. Pak.J.Statist.2014 Vol.30(6)： 1251-1266.
[7] Tanu Garg,Swetha Gokavarapu,SPE Kuwait International Petroleum Conference and Exhibition. Lessons Learnt From Root Cause Analysis of Gulf of Mexico Oil Spill[C].SPE 163276, Kuwait,2012.
[8] TaiMu Bao,Research Progress on Dealing with Deepwater Horizon Oil Spill Accident in Gulf of Mexico [J]. Journal of Ocean University of China (Natural Science Edition), 2015,01:55-62.
[9] Mendelssohn,Irving A,et al.Oil Impacts on Coastal Wetlands:Implications for the Mississippi River Delta Ecosystem after the Deepwater Horizon Oil Spill[A].BIOSCIENCE.2012 Vol.62(6)：562-574.
[10] WeiYa Zhu,Case Analysis of Safety Production Accidents [M]. Meteorological Publishing House,2012.
[11] Calcar H V， Marsh G L. Gas kick detector：US， 5006845[P].1991-04-09.
[12] En M P， Li X F， Shi F Q， et a1. Research of seabed rescue method of uncontrolled blowout in offshore drilling[J]. Advanced Materials Research， 2013， 6I6：837-843.
[13] Trevor Burgess,A.ALLEN Starkey,David White.Improvements for Kick Detection[J].Oilfield Review.1990 Vol.2(1):43-51.
[14] Choe J， Schubert J， Juvkam-Wold H. Analyses and procedures for kick detection in subsea mudlift drilling[J]. SPE Drilling 8L Completion， 2007， 22(4)； 296-303.
[15] Dhameliya J， Jain S， Gupta V. Liquid lift dual gradient drilling in deep water： early kick detection and control[R]. SPE I65372,2013.
[16] YuTing Kou,An Overflow Detection Method Based on Wellbore Liquid Level Measurement. China,CN2009l0048589.1[P].2009-03-31.
[17] Qiang Zhou,Early Overflow Monitoring Method [J]. Oil and gas field surface engineering, 2012,03:72-73.
[18] YangYang GE,Research and Design of Ground Monitoring and Diagnosis System for Early Overflow and Leakage [D]. Southwest Petroleum University,2014.
[19] XiaoDong Zhang,Progress and Prospect of Deep Water Drilling Technology [J]. Natural Gas Industry,2010,09:46-48+54+123-124.
[20] Guang Ying,Research on the Technology of Ultrasound Flow Measurement [D]. Xi’an Shiyou University,2012.
[21] Trevor Burgess,A.ALLEN Starkey,David White Improvements for Kick Detection[J].Oilfield Review.1990 Vol.2(1):43-51.
[22] Peter Ablard,Chris Bell,David Cook,Ivan Fornasier,Jean-Pierre Poyet,Kevin Fielding,Laura Lawton,George Haines,Mark A.Herkommer,Kevin Mc Carthy,Maja Radakovic,Lawrence Umar. Mud logging technology Advances (1) [J]. World Well Logging Technology,2014,02:67-74.
[23] HongYin Zheng,Analysis and Application of Data Mining Clustering Algorithms [D], Chongqing University, 2002.
[24] Ester M, Kriegel H P, Sander J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]//Kdd. 1996, 96(34): 226-231.