IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON URBAN COMPUTING AND INTELLIGENCE

Received March 15, 2019, accepted April 1, 2019, date of publication April 11, 2019, date of current version April 19, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2909647

An Sand Plug of Fracturing Intelligent Early Warning Model Embedded in Remote Monitoring System

HAIBO LIANG[1], JIALING ZOU[1], MUHAMMAD JUNAID KHAN[1], AND HAN JINXUAN[2]
[1]School of Mechatronic Engineering, Southwest Petroleum University, Chengdu 610500, China
[2]Department of Geology, Lomonosov Moscow State University, 119991 Moscow, Russia

Corresponding author: Haibo Liang (secondbo@126.com)

This work was supported in part by the Applied Basic Research Program of Sichuan Province (CN), under Grant 2016JY0049.

ABSTRACT In recent years, data mining technology has been widely used in different fields. For petroleum industry, the application of data mining method not only can analyze the data faster and more accurately but also can reduce the defects such as misjudgment and missed judgment caused by relying on artificial monitoring. Fracturing is one of the key techniques for increasing petroleum production. Due to the contingency of the abnormal variation of the fracturing construction curve, there is also a large lag and contingency by using the construction curve for sand block warning. Frequent false alarms and delayed alarms often occur. Therefore, the paper proposes an early-warning method of the sand plug of fracturing based on data mining. First, an early warning model of double logarithmic curve sand plug of fracturing is established, and the time series analysis algorithm is used to predict the oil pressure and casing pressure in the double logarithmic curve sand plug of fracturing risk warning model, so the early warning accuracy is improved. And then the general regression neural network (GRNN) algorithm is designed to optimize the prediction results of the time domain analysis. The improved affinity propagation (AP) clustering algorithm is used to cluster the monitoring data to help to improve the accuracy of the subsequent slope calculation, so as to improve the coincidence rate of the fracturing and sand plug risk. Finally, the risk warning model is applied and analyzed on site, and the validity and accuracy of the model are verified. The model is embedded in the remote monitoring system to realize online remote intelligent monitoring of risks by urban office workers.

INDEX TERMS AP clustering, data mining, early warning, sand plug of fracturing, GRNN algorithm.

I. INTRODUCTION

As global energy consumption increases and conventional oil resources decrease, fracturing technology is one of the important technologies to improve the exploitation of oil and gas resources. It is great significance for the development of low-permeability reservoirs and the stimulation of low-yield wells [1]–[5]. Due to the complexity of the stratum, various risks will be faced in the process of fracturing construction, especially sand plug, which is the most common, causing economic losses and environmental pollution, destroying seepage in the formation, and scraps the construction well, etc. [6], [7].

At present, the Internet of Things (IoT) has been widely used in different fields, which makes big data analysis full of challenges [8]–[12]. Accurate data analysis is very important to establish reasonable mathematical models [13]–[15]. The oil industry is gradually moving towards intelligence. Different sensors installed at the well site which can collect data constitute an IoT environment. The analysis of the data collected at the well site is aimed at extracting key information by using data mining technology [16], which can identify data trends, and conduct risk prediction. Therefore, the application of data mining technology to the early warning of fracturing construction is of great significance for avoiding the sand plug accident that occurs during the fracturing process.

Big data and IoT technologies are widely used in the oil industry [17]–[21]. Machine learning of big data and multi-dimensional analyses techniques within PALM were used

The associate editor coordinating the review of this manuscript and approving it for publication was Rongbo Zhu.

to analyze all relevant data from oil field, which let the company make the faster and better decision [22]. A new algorithm based on extreme learning machine model was proposed to determine the marine oil spill regions in SAR images [23]. A large number of historical data of oil and water wells were applied to monitor changes of some important parameters of the well site, which were used in the trend prediction and the early warning system [24].

The relevant researchers have done a lot of researches about fracturing, and have achieved some results. But most studies concentrated on the simulation of fracturing [25]–[28]. Song *et al.* proposed a mathematical model for simulating the flow in ultra-low permeability reservoir with hydraulic fracturing. The artificial fracture length, flow conductivity, and angle to development effect were illuminated in the paper [29]. Fully coupled hydro-mechanical numerical simulations were used to research fracture propagation and containment behavior during single- and multi-stage fracturing under reservoir conditions in the paper [30]. Wasantha *et al.* researched the suitability of low-frequency borehole resistivity measurements for detecting and appraising hydraulic fractures [31]. A robust proppant detection and classification workflow using machine learning for subsurface fractured rock samples post hydraulic fracturing operations was proposed [32].

In summary, there are few studies on the early warning technology of fracturing. At present, the warning of sand plug in fracturing construction mainly relies on the experience of on-site technicians, which causes a series of problems such as misjudgments and missed judgments. It is necessary to establish a set of intelligent early warning models in order to solve the low accuracy and poor early warning, which is the research motivation of the paper.

The paper proposes an early warning method for the risk of sand plug based on double logarithmic curve. Firstly, the coupled time domain analysis and GRNN algorithm are used to predict the oil pressure and casing pressure parameters in the double logarithmic curve slope sand plug risk warning. And then the slope change is applied to identify and judge the sand plug, which can realize the early warning of sand plug of fracturing. Finally, in order to improve the accuracy of curve slope calculation, the improved AP clustering algorithm is used to segment the oil pressure and pressure curve followed by curve fitting, at the same time calculate the slope of the fitted curve.

The main contributions of the paper are as follows:

(1) An early warning model for the double logarithmic curve of sand plug of fracturing is established in the paper.

(2) The time series analysis algorithm is proposed which can be predict the oil pressure and casing pressure in the early warning model, and the GRNN algorithm is used to optimize the prediction results in the time domain analysis.

(3) Improved AP clustering algorithm is used to cluster the monitoring data to improve the accuracy of risk warning.

The rest of the paper is organized as follows. Section 2 provides four mathematical models which includes double logarithmic curve model, time series model, GRNN, improved AP clustering. Section 3 describes a predictive model for coupling time series time domain analysis with GRNN. Section 4 illustrates the improved AP clustering early warning model. Section 5 provides engineering application analysis of model. Section 6 provides conclusion and future works.

II. DATA MINING MODEL THEORY

A. METHOD OF PRESSURE-TIME LOGARITHMIC CURVE SLOPE

The relationship between the net bottom pressure and the time change in the fracturing construction is shown in formula (1).

$$P_w(t) \propto t^e \tag{1}$$

When the fracturing is applied, the relationship between the net pressure of the bottom hole and the time is shown in a double logarithmic graph. $logp_w - logt$ is a linear relationship with a slope of e. e reflects the extension of the fractures in the underground formation.

B. TIME SERIES TIME DOMAIN ANALYSIS ALGORITHM

In time series analysis prediction model, the differential autoregressive moving average model (ARIMA) [33], [34] compared to the autoregressive moving average model (ARMA) [35]–[37] can better solve the problem that most time series in actual data are difficult to meet the requirements of stationarity. The method explains as below.

$$\varphi(B)\nabla^d x_t = \theta(B)\varepsilon_t \tag{2}$$

where d represents operational order, $\varphi(B)$ represents autoregressive polynomial, $\theta(B)$ represents moving average coefficient polynomial, ε_t represents random interference term.

Autocorrelation coefficient and the partial autocorrelation coefficient of the model samples help in identification and order of the model. The calculation formula is as follows in respective order.

$$\rho_k = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} + \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} \tag{3}$$

$$\varphi_k = \frac{\rho_k - \sum_{t=1}^{n-k}\varphi_{k-1}\rho_{k-t}}{1 - \sum_{t=1}^{k-1}\varphi_{k-1}\rho_k} \tag{4}$$

where ρ_k represents autocorrelation coefficient, φ_k represents partial autocorrelation coefficient, x_t and x_{t+k} represent time series values at different moments, $\bar{x}$ represents average value. When $|\varphi_k| \leq 1$, k=1, φ_k is equal to ρ_k.

In order to test the adaptability of the model, the Ljung-Box statistic is used to test the time series.

$$LB = n(n+2)\sum_{k=1}^{m}\left(\frac{p_k}{n-k}\right) \tag{5}$$

n represents number of samples, p_k represents residual sequence autocorrelation coefficient. When the probability value p of the Ljung-Box statistic is greater than 0.05, the original hypothesis is established and the model is significantly adaptable.

C. GRNN ALGORITHM

Generalized Regression Neural Network (GRNN) can analyze the mapping relationship between influencing factors and characterization values in data, and is often used for pattern recognition and prediction [38]–[42].

Let $f(x, y)$ be the joint probability density function of random variables x and y, and the conditional mean of the observed value X of y with respect to x:

$$E[y|X] = \frac{\int_{-\infty}^{+\infty} yf(X, y)dy}{\int_{-\infty}^{+\infty} f(X, y)dy} \tag{6}$$

By estimating the sample observations of x and y, a non-parametric estimate of the unknown probability density $f(x, y)$ can be obtained. After a series of calculations such as exchange summation and integration, the network output $Y(X)$ is obtained.

$$Y(X) = \frac{\sum_{i=1}^{n} Y_i exp[-\frac{(x-x_i)^T(x-x_i)}{2\sigma^2}]}{\sum_{i=1}^{n} exp[-\frac{(x-x_i)^T(x-x_i)}{2\sigma^2}]} \tag{7}$$

D. AN IMPROVED AP CLUSTERING ALGORITHM

In the oil field, data mining methods have been widely used [43], [44]. The AP clustering algorithm can eliminate the influence of artificial initial setting on clustering results, and has good clustering effect, and has been widely used in different fields [45]–[50]. However, AP clustering needs to set the reference degree. The algorithm is complex, and it runs for a long time under a large amount of data. Therefore an improved clustering is used by this article.

The iterative clustering process of the AP clustering algorithm is based on two kinds of information exchange updates between data points. The two kinds of exchange information represent two information parameters by using the representative matrix $r(i, k)$ and the suitable matrix $a(i, k)$ respectively. In the update calculation of the above $r(i, k)$ information parameter and $a(i, k)$ information parameter, the initial value $a_0 = 1$, and update calculation formula are defined as follows.

$$r(i, k) = S(i, k) - max_{k=k'} \left\{ a\left(i, k^{'}\right) + s\left(i, k^{'}\right) \right\} \tag{8}$$

$$a(i, k) = min\left\{0, r(k, k) + \sum\nolimits_{i \notin \{i,k\}} max\left\{0, r(i^{'}, k)\right\}\right\} \tag{9}$$

$$a(k, k) = \sum\nolimits_{i^{'} \neq k} max\left\{0, r(i^{'}, k)\right\} \tag{10}$$

The AP algorithm introduces a damping coefficient λ. Update formula is as follows.

$$r_{i+1}(i, k) = \lambda * r_i(i, k) + (1 - \lambda) * r_i(i, k) \tag{11}$$

$$a_{i+1}(i, k) = \lambda * a_i(i, k) + (1 - \lambda) * a_i(i, k) \quad (i \neq k) \tag{12}$$

Compared with the traditional AP algorithm, the adaptive AP clustering algorithm still oscillates. When the damping factor $\lambda \geq 0.85$, then gradually the p value reduces to adaptively exit the oscillation, whereas the p value is the value of the elements on the diagonal of the similarity matrix. And conduct the iterate with the new p. The p value iteration formula is as follow (13).

$$p = p - b * 0.01 * pm/(0.1 * sqrt(K + 50)) \tag{13}$$

where pm represents the row vector composed of the mean of each column of the similarity matrix S, K represents the number of clusters at the ith step iteration.

III. COUPLING PREDICTION MODEL

In the paper, the oil pressure and casing pressure fitting values predicted by time series time domain analysis are used as the input of the neural network. The obtained oil pressure and casing pressure value are used as the output of the network during the fracturing construction, so that the fitted value can approach the true value. Therefore, the network is trained. After completion of network training, the predicted values of oil pressure and casing pressure are used as input. Predicted value with higher precision and improved accuracy have been obtained accurate predicted values of oil pressure and casing pressure.

A. TIME SERIES ANALYSIS ALGORITHM FOR TIME SERIES ANALYSIS

In order to realize the advance warning of the real-time monitoring and early warning system of the whole sand plug of fracturing, the first problem to be solved in the process of establishing the time series analysis and prediction calculation model is the prediction step length problem as the follow-up warning support data. This study analyzes the casing pressure data of Well A in the X-block fracturing construction well. The results are shown in Fig 1.

Eight sets of data are taken for prediction in Fig 1. In the 8 sets of data, the prediction step size appears as 30, and the prediction sample numbers appears as 70, 80, 90, 100, 110, 120, 130, and 140 respectively.

According to the Fig 1 (a) to (f), the error of each group of prediction model is 0.006, 0.012, 0.017, 0.003, 0.002, 0.005, 0.005, and 0.004, respectively. When the prediction step size is constant, and The increase of sample size indicates that the error of the model prediction is the smallest when the sample numbers is 110, and it can well reflect the trend of the forecast data. The model and the actual changes in the data gradually deviate. When the sample data is less than 110, the error fluctuates greatly. Subsequently, the error fluctuations tend to be stable. The gradual increase of the sample does not change the gradual increase of the prediction error. The rate of error gradually slows down and tends to be stable with the increase in the sample. The model also tends to be stable. The increase of the sample must increase the accuracy of the sequence prediction.

In order to ensure the advanced nature of the model prediction, and consider the error size, the prediction step sizes selected are 13 in the paper and the prediction error is within 0.1. Predictive simulation analysis of the time series has been performed on the selected prediction step sizes.

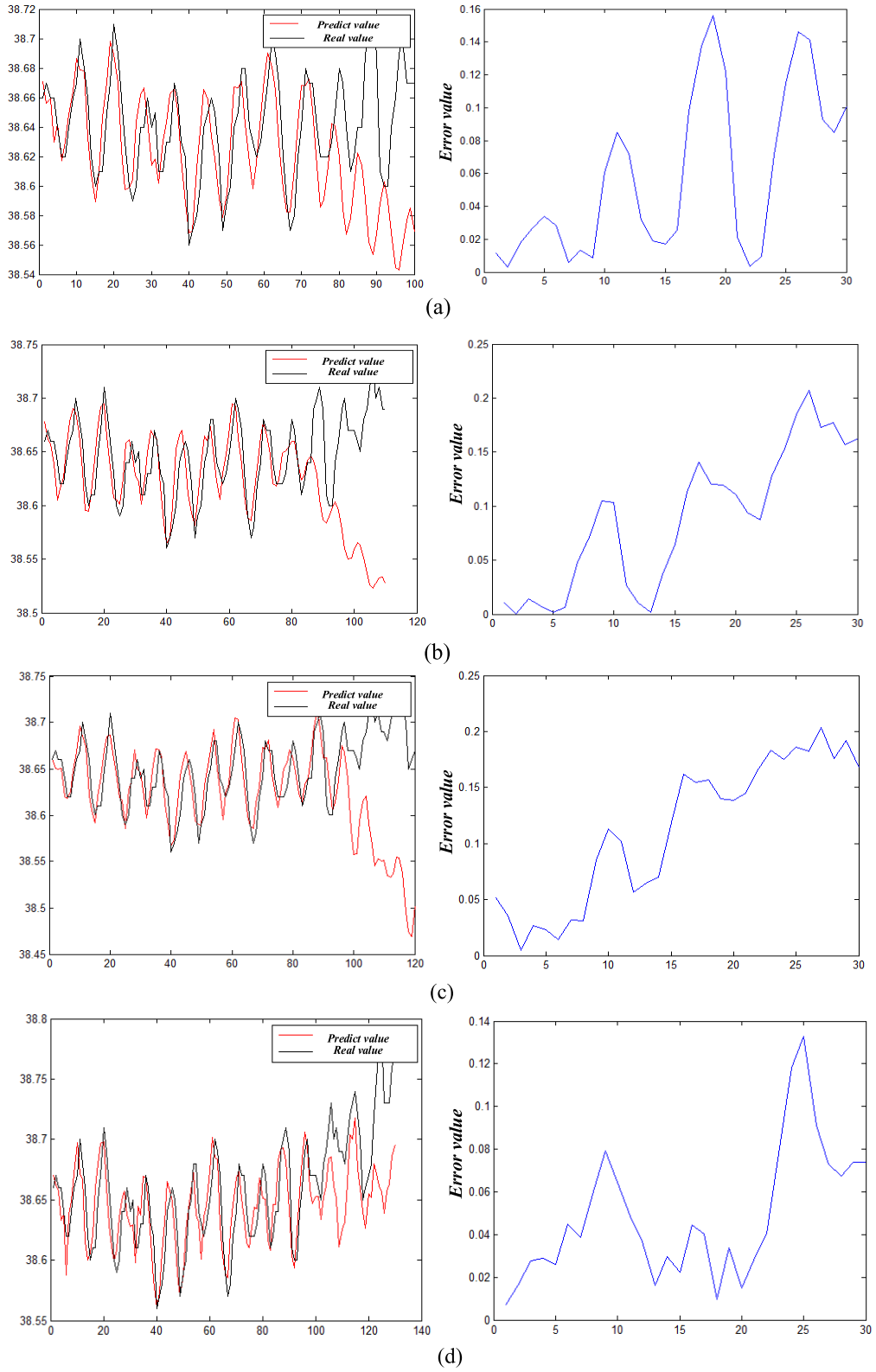


FIGURE 1. Different sample prediction error results. (a) Prediction error variation of 70 groups of samples. (b) Prediction error variation of 80 groups of samples. (c) Prediction error variation of 90 groups of samples. (d) Prediction error variation of 100 groups of samples.

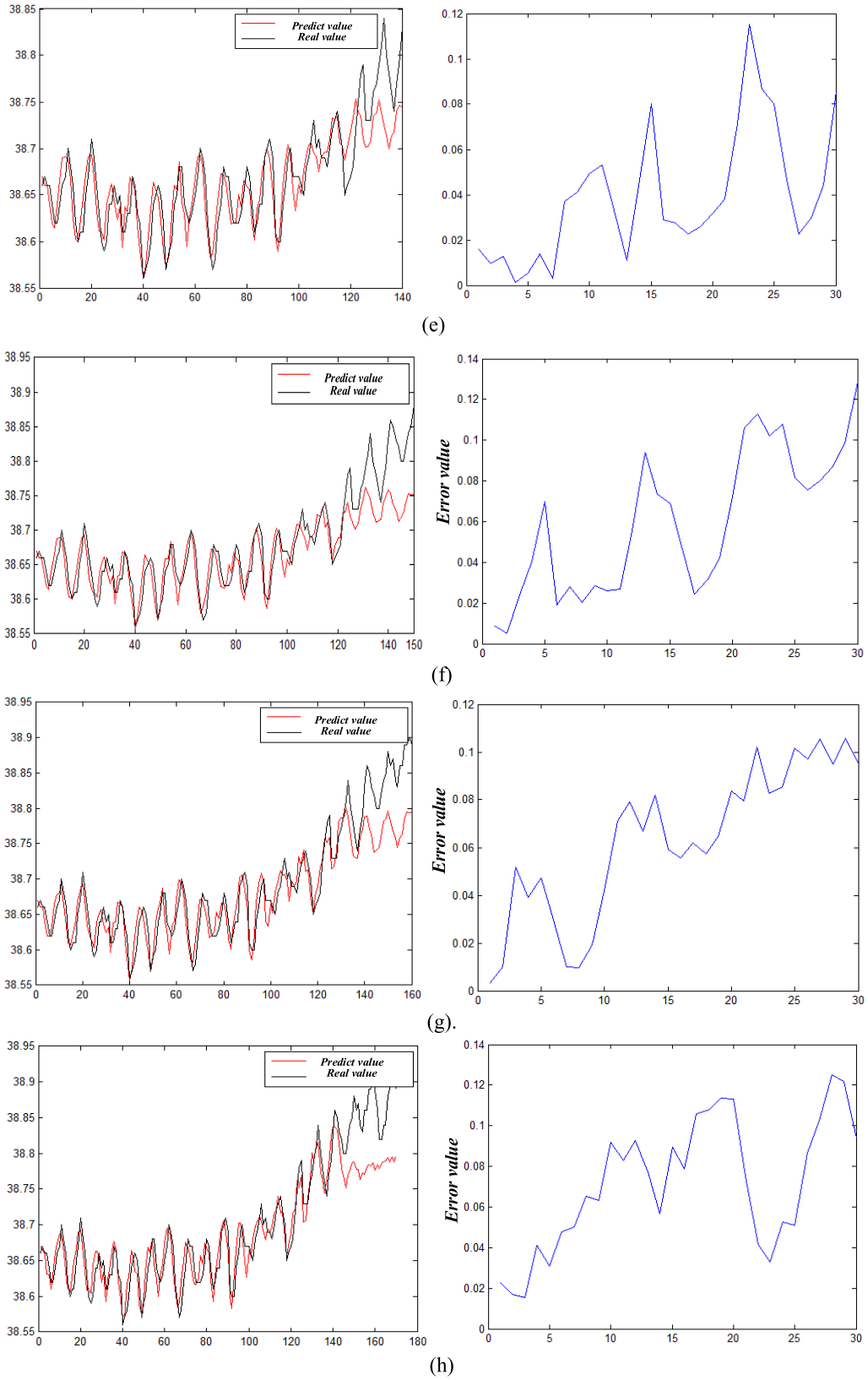
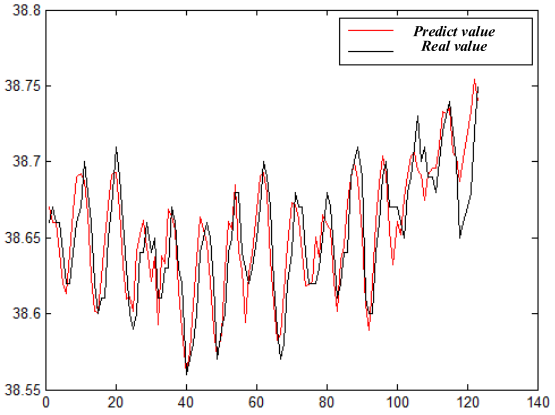


FIGURE 1. *(Continued.)* **Different sample prediction error results. (e) Prediction error variation of 110 groups of samples. (f) Prediction error variation of 120 groups of samples. (g) Prediction error variation of 130 groups of samples. (h) Prediction error variation of 140 groups of samples.**



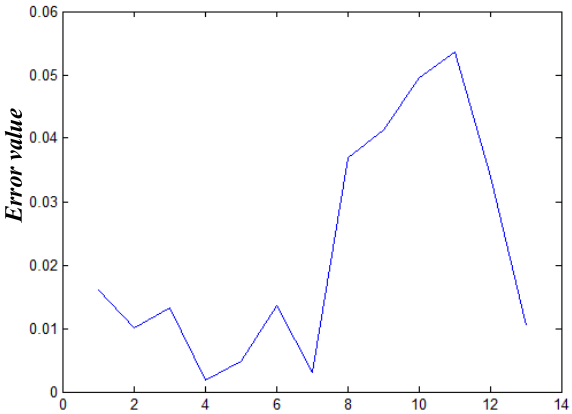


FIGURE 2. Casing pressure prediction and prediction error.

And the casing pressure prediction map and the prediction error map are obtained as follows.

Fig. 2 shows that the maximum prediction error appears as 0.054, the minimum value appears as 0.0018, and the absolute value of the average absolute error is calculated to be 0.1. The time series of the casing pressure changes slowly around 0.01Mpa during the actual fracturing construction, compared with the actual variation of the casing pressure results shows that the prediction error is still quite large. To ensure the accuracy of the subsequent warning, the accuracy of the casing pressure prediction have to be improved.

B. MODEL FOR IMPROVING EARLY WARNING COMPLIANCE RATE BASED ON GRNN

A total of 123 pieces of predicted sample data and predicted data are selected as sample sets for training and prediction. 8 fitting data are randomly select as data to be estimated (0.6450 0.6733 0.6444 0.618 0.6057 0.6559 0.6214 0.6911) to train the network, which determines the smoothing factor σ. σ starts from 0.01 and changes by 0.01 each time until the maximum value settled at 0.2. The paper predicts the fitted data and then calculates the mean square error of the fitted and predicted values. The figure of the mean square error is draw, and the value of the smoothing factor σ is determined (Fig 3).

According to our analysis, when the smoothing factor appears as $\sigma = 0.018$, the MSE value to be estimated as smallest, and the GRNN performance is optimal. In our paper, the trained GRNN has used for data fitting and prediction, and the model output is inversely normalized to obtain the true value and output as follows in Fig 4.

The GRNN is used to predict 13 data, and the predicted values are compared with the actual values and the ARIMA model output values. The results are shown in Table 1 and Fig 5.

According to Table 1 and Figure 5, the MAPE of ARIMA-GRNN is 0.025, and the MAPE of ARIMA is 0.057. The effect of ARIMA-GRNN model is better than ARIMA model.

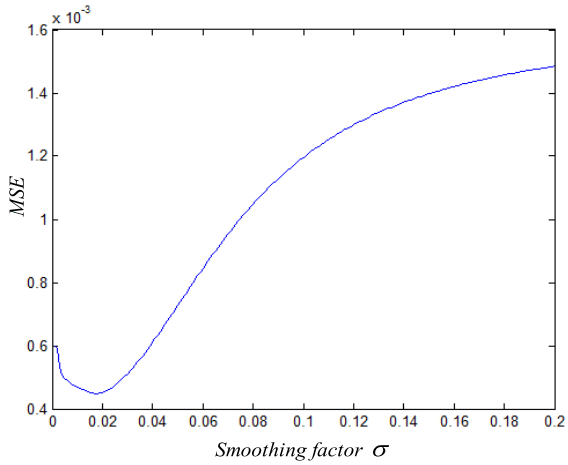


FIGURE 3. MSE values of different smoothing factors.

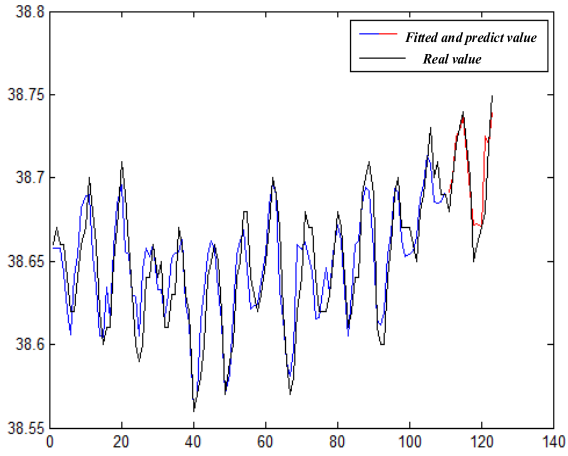


FIGURE 4. GRNN true value and output value.

IV. INTELLIGENT EARLY WARNING MODEL BASED ON AP CLUSTERING

Since the above-mentioned predicted oil pressure, casing pressure and other construction data are a series of discrete data about the time point. The real-time warning consists of

TABLE 1. Comparison of GRNN and ARIMA model results.

True value (MPa)	Predict value of ARIMA (MPa)	Absolute Relative error (%)	Predict value of ARIMA-GRNN (MPa)	Absolute Relative error (%)
38.68	38.696	0.041	38.691	0.028
38.70	38.71	0.026	38.7	0
38.72	38.733	0.034	38.725	0.013
38.73	38.732	0.005	38.729	0.003
38.74	38.735	0.013	38.736	0.01
38.72	38.706	0.035	38.715	0.013
38.70	38.703	0.008	38.692	0.021
38.65	38.687	0.095	38.672	0.057
38.66	38.701	0.107	38.673	0.034
38.67	38.72	0.128	38.67	0
38.68	38.734	0.139	38.725	0.116
38.72	38.754	0.088	38.72	0
38.75	38.74	0.027	38.74	0.026
MAPE		0.057	MAPE	0.025

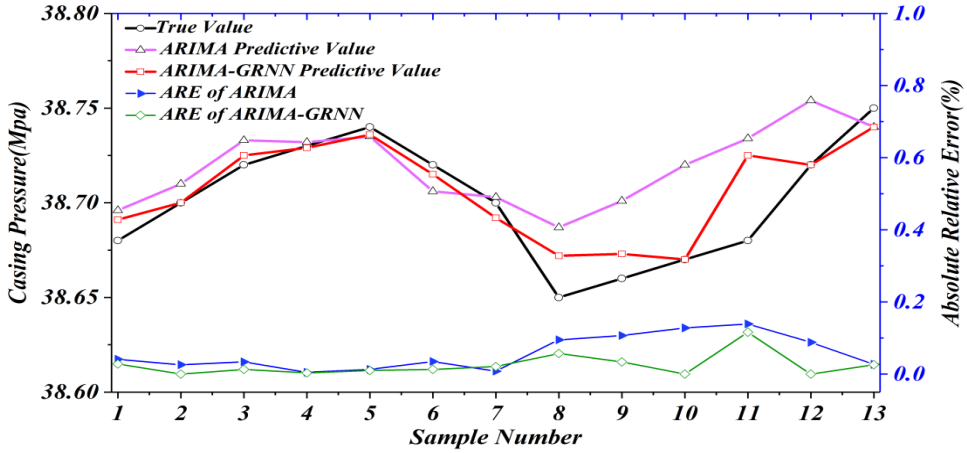


FIGURE 5. Flow chart analysis comparison of GRNN and ARIMA.

two parts, one explains fitting of the oil pressure and the casing pressure curve, and the second explains the real-time calculation of the slope of the double logarithmic curve. Compared with the traditional method, the fitting curve appears more smooth and continuous, which can effectively avoid the oscillation of local data points. The fitting curve passes all the data points and accurately reflect the overall curve in local trend while ensuring fitting accuracy. Therefore, the cubic spline interpolation method has been selected to fit the curve of oil pressure and casing pressure based on the accuracy of data fitting. This section studies the real-time calculation of the slope of the double logarithmic curve. Early warning flow chart of improved AP clustering is shown in Fig 6. In order to measure the pros and cons of the clustering algorithm, the paper introduces the Fowlkes-Mallows index.

$$\mathrm{FM} = \sqrt{\frac{TP}{TP+FP} \cdot \frac{TP}{TP+FN}} \tag{14}$$

where TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives.

This indicator is proportional to the number of TP of the clustering result, the similarity between the two clusters used to determine the index is greater for a higher value

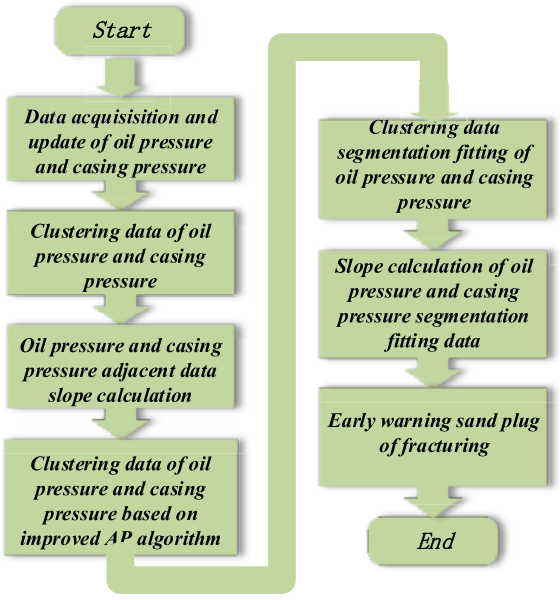


FIGURE 6. Early warning flow chart of improved AP clustering.

of the indicator. FM of the traditional AP clustering and improved AP clustering are 0.825 and 0.965 respectively. The clustering results are shown in Figures 7 and 8 respectively.

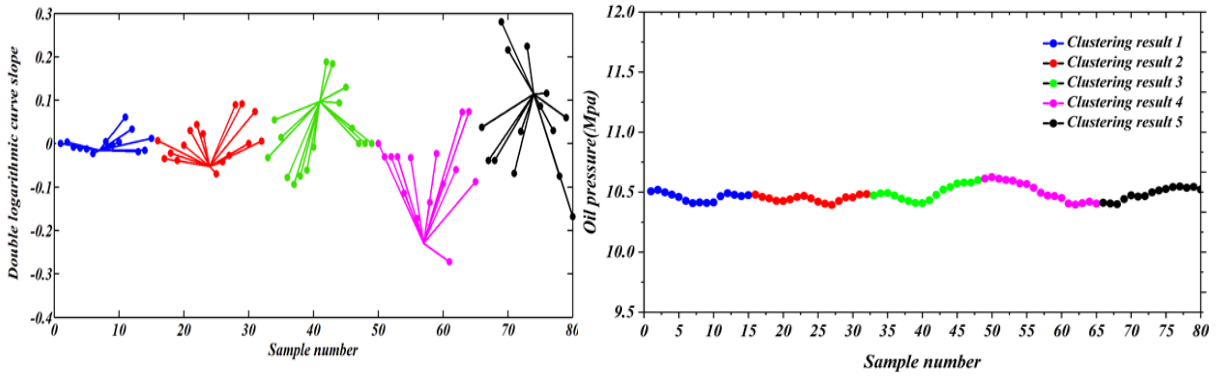


FIGURE 7. Traditional AP clustering results and clustering curves.

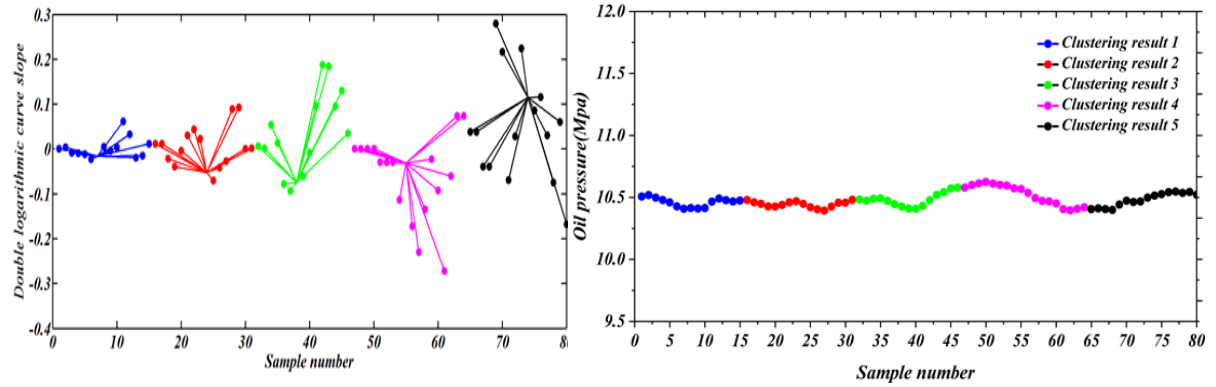


FIGURE 8. Improved AP clustering results and clustering curves.

According to the slope clustering result, the oil pressure data points can be segmented, realizing the data segmentation fitting and oil pressure data point slope achieved by the improved AP algorithm.

V. APPLICATION CASES

The composition of the well site fracturing construction Internet of Things is shown in Figure 9. First, the fracturing construction data is collected by sensors. Then the data is transmitted to the client through the network, and the system processes and analyzes the data. Finally, the office staff can know the construction information in real time through the monitoring system and make management decisions.

This section presents a number of use cases for well sites data analytics. Well A in YY block was fractured by conventional fracturing method. The well was fractured in the well section of 2456.3-2478 m. According to the oil pressure and casing pressure obtained by real-time acquisition, and the predicted oil pressure and casing pressure discrete data points, the curve is drawn. The double log curve slope calculation is then performed and compared to the set warning slope. Finally, whether there is a risk of sand plugging is judged. At the same time, new acquisition data filling and existing prediction data refreshing are performed, thereby realizing refreshing and drawing of the entire curve and realizing continuous warning. Comparison of forecasted data and real-time data have been shown in table 2.

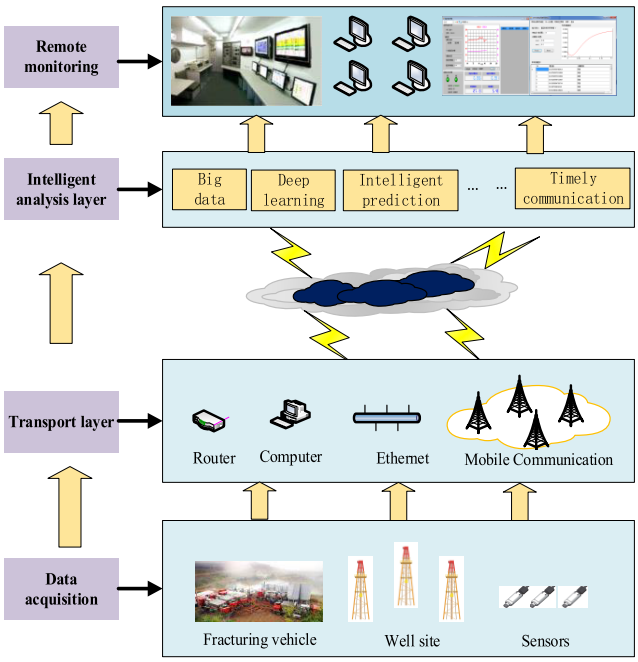


FIGURE 9. Internet of Things system in fracturing.

From the error analysis of the table 2, average relative error between the predicted oil pressure value and the actual oil pressure value at this time point shows as 0.018%, and the error appears within 0.05%. The application requirements are met, the prediction accuracy get higher, and the accuracy can

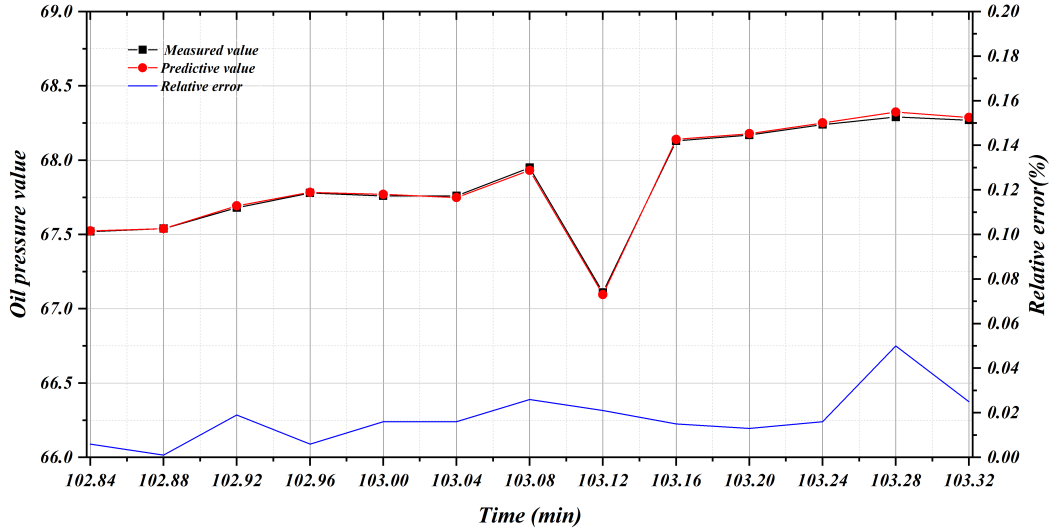


FIGURE 10. Comparative analysis of oil pressure prediction data and real-time data.

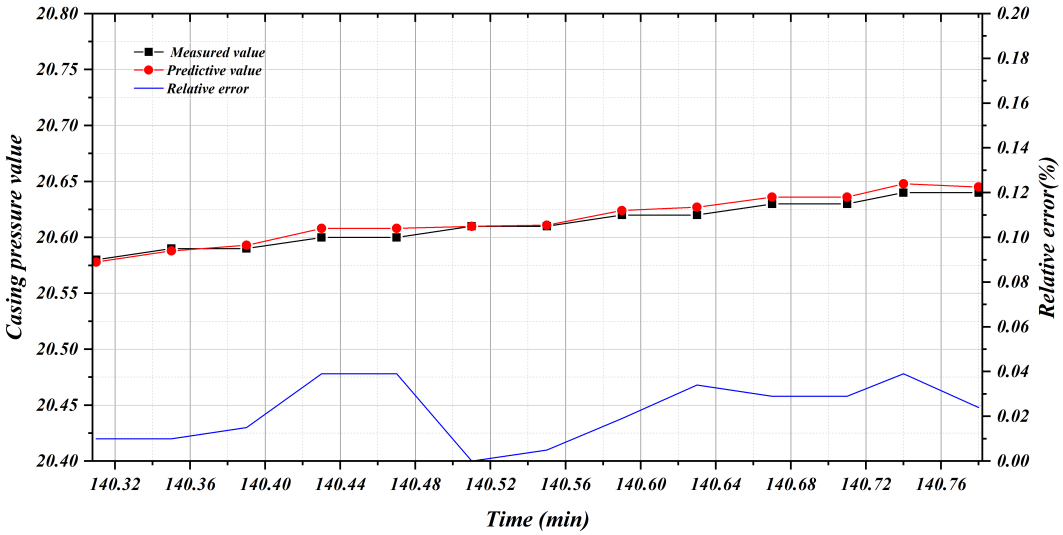


FIGURE 11. Comparative analysis of casing pressure prediction data and real-time data.

TABLE 2. Comparative analysis of oil pressure data.

True value (MPa)	*Predict value (MPa)*	*relative error (%)*
67.52	67.524	0.006
67.54	67.539	0.001
67.68	67.693	0.019
67.78	67.784	0.006
67.76	67.771	0.016
67.76	67.749	0.016
67.95	67.932	0.026
67.11	67.096	0.021
68.13	68.14	0.015
68.17	68.179	0.013
68.24	68.251	0.016
68.29	68.324	0.05
68.27	68.287	0.025
	MAPE	0.018

be more accurate. Reflecting the trend of oil pressure change and during calculation of the slope is shown that the trend has been predicted accurately and realized the intelligent warning of the risk of sanding in fracturing construction.

Well B in YY block was fractured by conventional fracturing method, and the well was fractured in the well section of 2563.5-2584.3 m. The system is accustomed to monitor and predict the sand plugging risk during the construction operation, and the pressure prediction data and real-time. The data comparison has been shown in the table 3.

It can be seen from the error analysis of table 3 that the average relative error between the casing pressure value and the actual casing pressure value predicted at this time point is 0.022%, which can predict the casing pressure change trend more accurately, and finally realize the double logarithmic slope prediction of the curve. It enables early identification of fracturing risk.

It can be seen from the error analysis (Fig 10 and 11) that the average relative error between the casing pressure value and the actual casing pressure value predicted at this

TABLE 3. Comparative analysis of casing pressure data.

True value (MPa)	*Predict value (MPa)*	*Absolute error (%)*
20.58	*20.578*	*0.01*
20.59	*20.588*	*0.01*
20.59	*20.593*	*0.015*
20.6	*20.608*	*0.039*
20.6	*20.608*	*0.039*
20.61	*20.61*	*0*
20.61	*20.611*	*0.005*
20.62	*20.624*	*0.019*
20.62	*20.627*	*0.034*
20.63	*20.636*	*0.029*
20.63	*20.636*	*0.029*
20.64	*20.648*	*0.039*
20.64	*20.645*	*0.024*
MAPE		*0.022*

time point appears as 0.022%, which can predict the casing pressure change trend more accurately, and finally realize the double logarithmic slope prediction of the curve to realize the early identification of the risk of sanding in fracturing construction. In summary, from the field application analysis and prediction error analysis of the wells A and B in the YY block of the system, the risk warning model of fracturing sand plug based on coupled time series time domain analysis algorithm and GRNN algorithm has been identified. The real-time monitoring and early warning system for sand plug of fracturing developed in combination with on-site fracturing equipment appears stable and reliable in the field, and can accurately predict the risk of sand plug of fracturing to ensure safe and efficient fracturing construction.

VI. CONCLUSION

In recent years, the well site has gradually turned to intelligent transformation, and sensor devices are generally placed in the well site to collect a large amount of monitoring data. This study needs to process and analyze the data collected from the wellsite based on the Internet of Things and big data interaction. Firstly, an double logarithmic curve slope fracturing sand risk warning model is established, and it couple time series time domain analysis algorithm and GRNN algorithm. Secondly, the time series analysis method is applied to predict the oil pressure and casing pressure. The time prediction of the fracturing sand block warning is ensured by the advance prediction. The GRNN algorithm is used to improve the time series analysis algorithm for oil pressure and set the coincidence rate of the predicted results. Finally, the improved AP clustering algorithm is used to improve the double logarithmic curve slope fracturing sand risk warning model. From the data precision point of view that the accuracy of fracturing sand plug risk warning is improved. Combined with field application, the improved sand plug risk warning model appears to be more accurate and fast, and has a good industrial application prospect. The early warning model proposed is embedded in the remote system development to enable the city office staff to remotely monitor. In the future research, intellectualization is a hot spot. In addition to intelligent early warning of risks, study of intelligent control after the occurrence of risks worth great importance.

REFERENCES

[1] B. Wang *et al.*, "Quantitative investigation of fracture interaction by evaluating fracture curvature during temporarily plugging staged fracturing," *J. Petroleum Sci. Eng.*, vol. 172, pp. 559–571, Jan. 2019.

[2] D. Shang *et al.*, "Local asymmetric fracturing to construct complex fracture network in tight porous reservoirs during subsurface coal mining: An experimental study," *J. Natural Gas Sci. Eng.*, vol. 59, pp. 343–353, Jan. 2018.

[3] M. He, W. J. Chen, L. Tian, B. Shao, and Y. Lin, "Plant-microbial synergism: An effective approach for the remediation of shale-gas fracturing flowback and produced water," *J. Hazardous Mater.*, vol. 363, no. 5, pp. 170–178, Feb. 2019.

[4] G. Liu, T. Zhang, M. He, M. Li, and J. Li, "Thermal performance analysis of drilling horizontal wells in high temperature formations," *Appl. Therm. Eng.*, vol. 78, pp. 217–227, Mar. 2015.

[5] K. H. S. M. Sampath *et al.*, "Effect of coal maturity on CO_2-based hydraulic fracturing process in coal seam gas reservoirs," *Fuel*, vol. 236, pp. 179–189, Jan. 2019.

[6] B. Figueiredo, C. F. Tsang, J. Rutqvist, and A. Niemi, "Study of hydraulic fracturing processes in shale formations with complex geological settings," *J. Petroleum Sci. Eng.*, vol. 152, pp. 361–374, Apr. 2017.

[7] Z. Zhang, J. Mao, X. Yang, J. Zhao, and G. S. Smith, "Advances in waterless fracturing technologies for unconventional reservoirs," *Energy Sources A, Recovery, Utilization, Environ. Effects*, vol. 41, no. 2, pp. 237–251, 2019.

[8] M. Marjani *et al.*, "Big IoT data analytics: Architecture, opportunities, and open research challenges," *IEEE Access*, vol. 5, pp. 5247–5261, 2017.

[9] F. Ghofrani, Q. He, R. M. P. Goverde, and X. Liu, "Recent applications of big data analytics in railway transportation systems: A survey," *Transp. Res. C, Emerg. Technol.*, vol. 90, pp. 226–246, May 2018.

[10] M. D. Assunção, R. N. Calheiros, S. Bianchi, M. A. S. Netto, and R. Buyya, "Big data computing and clouds: Trends and future directions," *J. Parallel Distrib. Comput.*, vols. 79–80, pp. 3–15, May 2015.

[11] A. M. S. Osman, "A novel big data analytics framework for smart cities," *Future Gener. Comput. Syst.*, vol. 91, pp. 620–633, Feb. 2019.

[12] R. Zhu, X. Zhang, X. Liu, W. Shu, T. Mao, and B. Jalaian, "ERDT: Energy-efficient reliable decision transmission for cooperative spectrum sensing in industrial IoT," *IEEE Access*, vol. 3, pp. 2366–2378, 2015.

[13] X. Liu, R. Zhu, B. Jalaian, and Y. Sun, "Dynamic spectrum access algorithm based on game theory in cognitive radio networks," *Mobile Netw. Appl.*, vol. 20, no. 6, pp. 817–827, Dec. 2015.

[14] B. Wu, L. Zong, X. Yan, and C. G. Soares, "ERDT: Energy-efficient reliable decision transmission for intelligent cooperative spectrum sensing in industrial IoT," *Ocean Eng.*, vol. 164, pp. 590–603, Sep. 2018.

[15] B. Wu, X. Yan, Y. Wang, and C. G. Soares, "An evidential reasoning-based CREAM to human reliability analysis in maritime accident process," *Risk Anal.*, vol. 37, no. 10, pp. 1936–1957, 2017.

[16] S. Shadroo and A. M. Rahmani, "Systematic survey of big data and data mining in Internet of Things," *Comput. Netw.*, vol. 139, pp. 19–47, Jul. 2018.

[17] W. Gharibi, M. Aalsalem, W. Z. Khan, N. Armi, and W. Ghribi, "Monitoring gas and oil fields with reliable wireless sensing and Internet of Things," in *Proc. Int. Conf. Radar, Antenna, Microw., Electron., Telecommun. (ICRAMET)*, Oct. 2017, pp. 188–191.

[18] M. Rahmati, H. Yazdizadeh, and A. Yazdizadeh, "Leakage detection in a gas pipeline using artificial neural networks based on wireless sensor network and Internet of Things," in *Proc. 15th Int. Conf. Ind. Inform. (INDIN)*, Jul. 2017, pp. 659–664.

[19] Y. Wan-Jun, W. Jian, and Z. Huai-Lin, "Research on risk management of gas safety based on big data," in *Proc. Int. Conf. Intell. Inform. Biomed. Sci. (ICIIBMS)*, Oct. 2018, pp. 90–92.

[20] R. M. Alguliyev, R. M. Aliguliyev, and M. S. Hajirahimova, "Big data integration architectural concepts for oil and gas industry," in *Proc. 10th Int. Conf. Appl. Inf. Commun. Technol. (AICT)*, Oct. 2016, pp. 1–5.

[21] H. Liang, T. Yongqiang, and L. Xiang, "Research on drilling kick and loss monitoring method based on Bayesian classification," *Pakistan J. Statist.*, vol. 30, pp. 1251–1266, Dec. 2014.

[22] R. N. Anderson, "'Petroleum analytics learning machine' for optimizing the Internet of Things of today's digital oil field-to-refinery petroleum system," in *Proc. IEEE Int. Conf. (Big Data)*, Dec. 2017, pp. 4542–4545.
[23] X. Lyu, "Oil spill detection based on features and extreme learning machine method in SAR images," in *Proc. 3rd Int. Conf. Mech., Control Comput. Eng. (ICMCCE)*, Sep. 2018, pp. 559–563.
[24] B. Xu, W. Wang, Y. Wu, Y. Shi, and C. Lu, "Internet of Things and big data analytics for smart oil field malfunction diagnosis," in *Proc. IEEE 2nd Int. Conf. Big Data Anal.*, Mar. 2017, pp. 178–181.
[25] F. Cruz, D. Roehl, and E. A. do Vargas, "An XFEM element to model intersections between hydraulic and natural fractures in porous rocks," *Int. J. Rock Mech. Mining Sci.*, vol. 112, pp. 385–397, Dec. 2018.
[26] M. Roostaei, S. Taghipoor, A. Nouri, V. Fattahpour, and D. Chan, "Smeared modeling of hydraulic fracture using partially coupled reservoir and geomechanics simulators," *Int. J. Rock Mech. Mining Sci.*, vol. 113, pp. 99–111, Jan. 2019.
[27] M. Safari, V. Rastegari, and S. Shahmarvand, "Hydraulic fracturing in condensate reservoirs: A simulation study," in *Proc. Int. Conf. Environ. Impacts Oil Gas Industries, Kurdistan Region Iraq Case Study (EIOGI)*, Apr. 2017, pp. 92–96.
[28] N. M. Rayudu, X. Tang, and G. Singh, "Simulating three dimensional hydraulic fracture propagation using displacement correlation method," *Tunnelling Underground Space Technol.*, vol. 85, pp. 84–91, Mar. 2019.
[29] H. Song, X. Huang, and Y. Lou, "Study on the effect of fracture characteristics to ultra-low permeability reservoirs development," in *Proc. 2nd Int. Conf. Mechanic Automat. Control Eng.*, Jul. 2011, pp. 3559–3562.
[30] K. Yang, C. Torres-Verdín, and A. E. Yılmaz, "Detection and quantification of three-dimensional hydraulic fractures with horizontal borehole resistivity measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4605–4615, Aug. 2015.
[31] P. L. P. Wasantha, H. Konietzky, and C. Xu, "Effect of *in-situ* stress contrast on fracture containment during single- and multi-stage hydraulic fracturing," *Eng. Fract. Mech.*, vol. 205, pp. 175–189, Jan. 2019.
[32] D. Maity and J. Ciezobka, "Designing a robust proppant detection and classification workflow using machine learning for subsurface fractured rock samples post hydraulic fracturing operations," *J. Petroleum Sci. Eng.*, vol. 172, pp. 588–606, Jan. 2019.
[33] C. N. Babu and B. E. Reddy, "Temperature using variants of ARIMA models," in *Proc. Int. Conf. Adv. Eng. Sci. Manage.*, Mar. 2012, pp. 256–260.
[34] Q. Wang, X. Song, and R. Li, "A novel hybridization of nonlinear grey model and linear ARIMA residual correction for forecasting U.S. shale oil production," *Energy*, vol. 165, pp. 1320–1331, Dec. 2018.
[35] R. Skopal, "Short-term hourly price forward curve prediction using neural network and hybrid ARIMA-NN model," in *Proc. Int. Conf. Inf. Digit. Technol.*, Jul. 2015, pp. 335–338.
[36] H. Kang, Y. Qi, G. Liu, J. Liu, and C. Li, "Application of the ARMA model in non-stationary vibration signals," in *Proc. Int. Conf. Qual., Rel., Risk, Maintenance, Saf. Eng.*, Jun. 2012, pp. 751–754.
[37] H. Zhu and X. Lu, "The prediction of PM2.5 value based on ARMA and improved bp neural network model," in *Proc. Int. Conf. Intell. Netw. Collaborative Syst. (INCoS)*, Sep. 2016, pp. 515–517.
[38] H. Majumder and K. Maity, "Application of GRNN and multivariate hybrid approach to predict and optimize WEDM responses for Ni-Ti shape memory alloy," *Appl. Soft Comput.*, vol. 70, pp. 665–679, Sep. 2018.
[39] X. Meng, W. Zhang, and S. Cong, "The objective evaluation model on wearing touch and pressure sensation based on GRNN," in *Proc. 3rd Int. Conf. Inf. Comput.*, Jun. 2010, pp. 289–292.
[40] X. Kuang, C. Wu, Y. Huang, and L. Xu, "Traffic flow combination forecasting based on grey model and GRNN," in *Proc. Int. Conf. Intell. Comput. Technol. Automat.*, May 2010, pp. 1072–1075.
[41] Y. Zhengxiang, X. Guimin, and W. Jinwen, "Transport volume forecast based on GRNN network," in *Proc. 2nd Int. Conf. Future Comput. Commun.*, May 2010, pp. V3-629–V3-632.
[42] L. Zhang, G. Lu, and Y. Qi, "Permeability extracting using GRNN method," in *Proc. Asia–Pacific Symp. Electromagn. Compat.*, Nov. 2010, pp. 1657–1659.
[43] H. Liang, X. Q. Huang, and Y. Sun, "A diagnostic model based on support vector machine for the collapse of horizontal well borehole wall," *J. Residuals Sci. Technol.*, vol. 13, pp. 167–175, Jan. 2016.
[44] H. Liang, T. Yongqiang, L. Xiang, and L. Yangyang, "Esearch on drilling kick and loss monitoring method based on Bayesian classification," *Pakistan J. Statist.*, vol. 30 no. 6, pp. 1251–1266, 2014.
[45] Z. Zhang, P. Tang, and T. Corpetti, "Satellite image time series clustering via affinity propagation," in *Proc. Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 2419–2422.
[46] X. Zhu, J. Li, Z. Liu, and F. Yang, "A joint grid segmentation based affinity propagation clustering method for big data," in *Proc. IEEE 18th Int. Conf. High Perform. Comput. Commun.; IEEE 14th Int. Conf. Smart City; IEEE 2nd Int. Conf. Data Sci. Syst. (HPCC/SmartCity/DSS)*, Dec. 2016, pp. 1232–1233.
[47] P. A. Karegar, "Wireless fingerprinting indoor positioning using affinity propagation clustering methods," *Wireless Netw.*, vol. 24, no. 8, pp. 2825–2833, 2018.
[48] Y. Zhu, J. Yu, and C. Jia, "Initializing K-means clustering using affinity propagation," in *Proc. 9th Int. Conf. Hybrid Intell. Syst.*, Aug. 2009, pp. 338–343.
[49] J. Bi, Y. Wang, X. Li, H. Qi, H. Cao, and S. Xu, "An adaptive weighted KNN positioning method based on omnidirectional fingerprint database and twice affinity propagation clustering," *Sensors*, vol. 18, no. 8, p. 2502, 2018.
[50] H. Du, Y. Wang, and L. Duan, "A new method for grayscale image segmentation based on affinity propagation clustering algorithm," in *Proc. 9th Int. Conf. Comput. Intell. Secur.*, Dec. 2013, pp. 170–173.

HAIBO LIANG was born in Weihai, Shandong, China, in 1987. He received the B.E., M.E., and Ph.D. degrees from Southwest Petroleum University, in 2008, where he is currently a Professor. He is mainly engaged in oil drilling related work, such as expert system, data mining, oil equipment development, and related system development. He is engaged in oil and gas geological guidance analysis and information interpretation analysis. He is a Subject Technology Leader. He has published many high-quality papers, patents, and a lot of soft.

JIALING ZOU was born in Chengdu, Sichuan, China, in 1992. She received the M.E. degree from Southwest Petroleum University, in 2018, where she is currently pursuing the Ph.D. degree.

From 2015 to 2018, she has been engaged in petroleum work research, mainly engaged in drilling related data analysis and data mining. At the same time, she is also engaged in oil and gas development related work. She is interested in artificial intelligence and data processing and has been engaged in related research. She has published two papers.

MUHAMMAD JUNAID KHAN was born in Rawalpindi, Pakistan, in 1992. He received the M.E. degree in electronic science and technology from the University of Electronic Science and Technology, Chengdu, China, in 2018. He is currently pursuing the Ph.D. degree with Southwest Petroleum University. His current research interests include data mining, drilling, computer, petroleum, substation automation, and smart grid.

HAN JINXUAN received the bachelor's degree of bioengineering, and the master's and Ph.D. degrees in oil-gas field exploitation from Southwest Petroleum University, China, in 2009, 2012, and 2015, respectively. She is currently pursuing the Ph.D. degree in oil-gas field geology with Lomonosov Moscow State University. Her current research interest includes unconventional oil and gas.