Analysis of the Influence Factors of Casing Damage Based on Data Mining

Shujuan Zhang[1(✉)], Xueqing Zhang[2], Wenchang Fu[1], Jin Zhang[3], Liqiu Zhang[1], Qinghong Liu[1], and Lihong Zhu[1]

[1] The Exploration and Development Research Institute of Daqing Oilfield, Daqing City, Heilongjiang Province, China
Zhangshujuan8401@petrochina.com.cn

[2] The Planning and Design Institute, The Second Production Plant of Daqing Oilfield, Daqing City, Heilongjiang Province, China

[3] Natural Gas Company of Daqing Oilfield Limited Company, Daqing City, Heilongjiang Province, China

Abstract. The prolonged period of waterflooding development on the oilfield has considerably increased the risk of casing damage, which would affect the regular production of wells, decrease the oil production rate as well as increase the costing of productive maintenance. Considering the in-depth analysis of casing damage and the selection of main effecting factors, a targeted countermeasure is needed in order to prevent casing damage. The causing factors of well casing damage include three types: geological (i.e., fault, formation dip, and mudstone interlayer), developmental (i.e., water injection pressure, injection-producing intensity, and volume replacement ratio), and engineering (i.e., cementing quality, material of the casing pipe). Data of influencing factors as described various in type and are big in volume, especially the data of developing dynamic history. The article focuses on the optimization of main-controlling developing factors and risk prediction of casing damage based on big data analysis technology. To derive the pattern in which exploiting factors affect on casing damage from mass data using big data analysis technology, the first step is to establish the relations between individual influencing factors with their corresponding casing damage rate, therefore identify the main controlling factor; In the second step, casing damage forewarning parameters system is established while the hysteresis phenomenon of the casing damage time is eliminated both by the computation of the correlation coefficient between parameters and casing damage rate. The last step is to establish the casing damage risk prediction model achieved by the application of principal component analysis and support vector machine, which make prediction of casing damage risk in advance feasible and provide the technical support for the prevention and control of casing damage.

Keywords: Casing damage · Risk prediction · Big data · Support vector machine (SVM)

J. Lin (ed.), *Proceedings of the International Field Exploration and Development Conference 2018*, Springer Series in Geomechanics and Geoengineering, https://doi.org/10.1007/978-981-13-7127-1_21

1 Introduction

Ever since the year of 2009, 1000 wells casing damage are discovered every year in Daqing oilfield. In some blocks, concentrated casing damage had appeared in the marker bed of member II of the Nenjiang Fm, causing the ratio of start-up well number to producing well number in these blocks to fall below 70%. There is a clearly out-of-balance for the injection-production system. The results lead to oil production rate dropped sharply [1]. Because of the casing failure, the cost for monitor, maintain, and update has greatly increased. Currently, a new technique is desperately needed to forecast casing damage risk quickly and efficiently according to production abnormalities, so that casing damage risk can be discovered as quick as possible and being dealt with timely, therefore casing damage well number can be reduced.

All around the world, casing damage is one of the biggest issues in impeding oilfield development. Experts around the world have done tremendous investigations on the main reason for the frequent occurrence of casing damage. They believe there are many complex factors that contribute to the cause of casing damage, include: geographical structure, the balance between injection and production, the overloading pressure on the oil formation, the overall well cement job quality, and the overall quality of operate execution [2, 3]. Many researchers have developed investigations using geomechanics models as a way to figure out the mechanical theory behind the casing damage therefore being able to predict potential risks of future casing damage [4–6]. But this method is time-consuming, which means that it will have low efficiency when it comes to the monitoring of a large area therefore causing the inability to quickly discover and prevent the risks of casing damage.

Currently, big data analysis has been used in a variety of fields around the world, due to the fact that predictions and prevention of casing damage refer to many professional areas, such as geology, development, perforating technology, monitoring technique. So it has very complex data including the structured data of everyday production and unstructured data of injection profile and other graphics. The information amount is huge and the process involved is also super complex. Making big data analysis useful in terms of predicting the trend of casing damage. Different from the normal traditional data analysis, big data analysis is different in three ways: the first is “Maximize the importance of the whole, minimize the importance of singulars”, the second is “Maximize the importance of correlation, minimize the importance of cause and effect”, the third is “Maximize the importance of rate and efficiency, minimize the importance of accuracy”. Big data analysis successfully deals with the problems regarding the difficulty to obtain and analyze qualitative data from casing damage and makes it easier to discover hidden pattern among complex data with high efficiency therefore establishing the foundation for predictions and prevention of casing damage.

This paper provides a detailed description of all the process involved in the big data analysis. The first step is to discover the main influencing factors that cause casing damage and design parameter calculation expression. The next step is to find out the relationship between parameter and the casing damage rate and select the ones with the highest degree of correlation and eliminate the hysteresis phenomenon of casing damage time and the correlations between parameters one by one. The last step is to establish a forewarning model by using these parameters as a way to analyze the development factors data of the oilfield and predict the potential risks therefore being able to quickly establish prevention methods.

2 Main Influence Factors of Casing Damage Analysis and Parameters Optimization

Big data analysis can be summarized in four steps: data collection, standardization and pretreatment, statistics analysis, and excavation [7]. First, data related to casing damage has to be collected. According to the discoveries of researchers in the past, the process of oilfield development has close relationship with the casing damage [8, 9]. These factors can be roughly divided into three groups: geological factor, development factor, and engineering factor. Geological factors are the basic conditions, some geological factors that can potentially increase the risk of casing damage, include dip angle, fault, reservoir development situation, special lithological interface, etc. On the other hand, engineering factors are the casing tube condition, include: the type of casing tube, cement job quality, perforation condition, and operation construction influence. Development factor is the external cause of casing damage and is the most important factor related to casing damage. As a controlled factor, development factor is the most important focus of the investigation. Development factors include: the injection-production relationship, formation pressure, water injection pressure, volume replacement ratio, etc. This article introduces the method and the process of development factor analysis as an example.

According to analysis, formation pressure, water injection pressure, level and variation of injection-production parameters are the main factors influencing casing damage.

2.1 Formation Pressure Is the Main Development Factor that Causes Casing Damage

Here using differential pressure to represent the difference of current formation pressure and the original formation pressure. According to the curve between annual differential pressure and annual casing damage well number of typical blocks, there is a distinct correlation exists between total pressure difference and casing damage well number.

When there are drastic changes of differential pressure or when there is a huge amount of debt, the amount of casing damage well increases. From the time of the significant drop of differential pressure to the significant increase of casing damage wells, there is a time hysteresis.

Based on the graph shown below (Fig. 1), casing damage well number increases when the differential pressure is below −0.5 MPa.

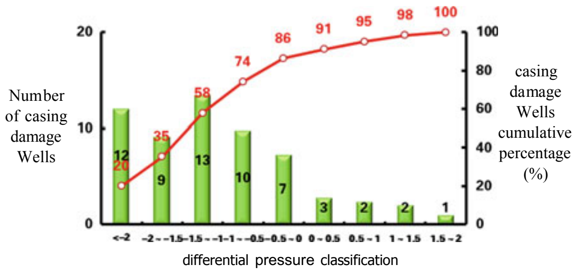

Fig. 1. Histogram of casing damage wells numbers with differential pressure

2.2 High Water Injection Pressure Is a Main Development Factor that Causes Casing Damage

Based on the collected data, back in the 1980s when the water injection pressure is higher, there are more number of casing damages [10]. And then an upper limit is being set on water injection pressure based on the pressure required to break the formation.

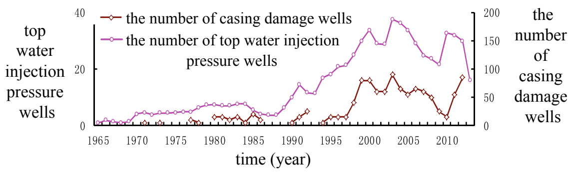

Fig. 2. Relation curve between the number of casing damage wells and top water injection pressure wells

The presence of top water injection pressure wells is also another factor that causes the increasing amount of casing damage wells.

To have a more thorough understanding regarding the relationship between high water injection pressure and the amount of casing damage wells, a graph is designed regarding the relationship between top water injection pressure wells and the number of casing damage wells (Fig. 2).

The graph shows that a distinct correlation exists between water injection pressure and the number of casing damage wells.

Also, keeping high-pressure water injection in a period of time and too much pressure fluctuation will also increase the number of casing damage wells.

2.3 The Injection-Production Parameter Is also a Main Development Factor that Affects Casing Damage

Formation pressure is the main development factor that causes casing damage, but not every well is monitored the formation pressure. So the injection-production parameter, the factor that causes formation pressure changed being selected as representative. Based on the data collected, it can be seen that a positive correlation exists between "Annual Water Injection Change", "Annual Production Quantity Change", and the "Rate of Annual Casing Damage" [11] (Fig. 3).

3 Casing Damage Forewarning Parameter Optimization

3.1 The Design of Casing Damage Forewarning Parameter

Due to the complexity of the factors of casing damage, a variety of changes of development factors may cause casing damage. The intensity, the magnitude of change, the peripheral difference, and the cumulative change can all cause casing damage. That is why factors regarding the action to deal especially with the effect of intensity, magnitude, the cumulation, the structure, and the hysteresis.

Based on the characteristics of the development factors above, 5 category (Formation Pressure, Injected Pressure, Water Injection, Overall Production, Volume Replacement Ratio) and 39 parameters are designed (Tables 1 and 2).

3.2 Correlation Analysis of Development Parameters and Casing Damage Rate

In order to select the more accurate and the most relatable parameters relating to casing damage, the calculation using correlation coefficient to determine the relationship between the parameter and the casing damage rate. For every parameter(X-Axis) and the casing damage rate(Y-Axis). The formula is shown below.

$$\rho\, x, y = \frac{\operatorname{cov}(X, Y)}{\sqrt{D(X)\sqrt{D(Y)}}} \tag{1}$$

In the equation:

$\rho\ x, y$ represents the correlation coefficient.

Covariance: $\operatorname{cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$

Variance: $D(X) = E\{[X - E(X)]2\}$

$D(X)$ is the standard deviation

E is a mathematical expectation of random variables, reflecting the average value of random variables.

Correlation coefficient greater than or equal to 0.4 is a medium correlation. Greater than or equal to 0.7 is a strong correlation. Hysteresis had eliminated by calculating the correlation coefficient of different lag years of the two column data, the corresponding

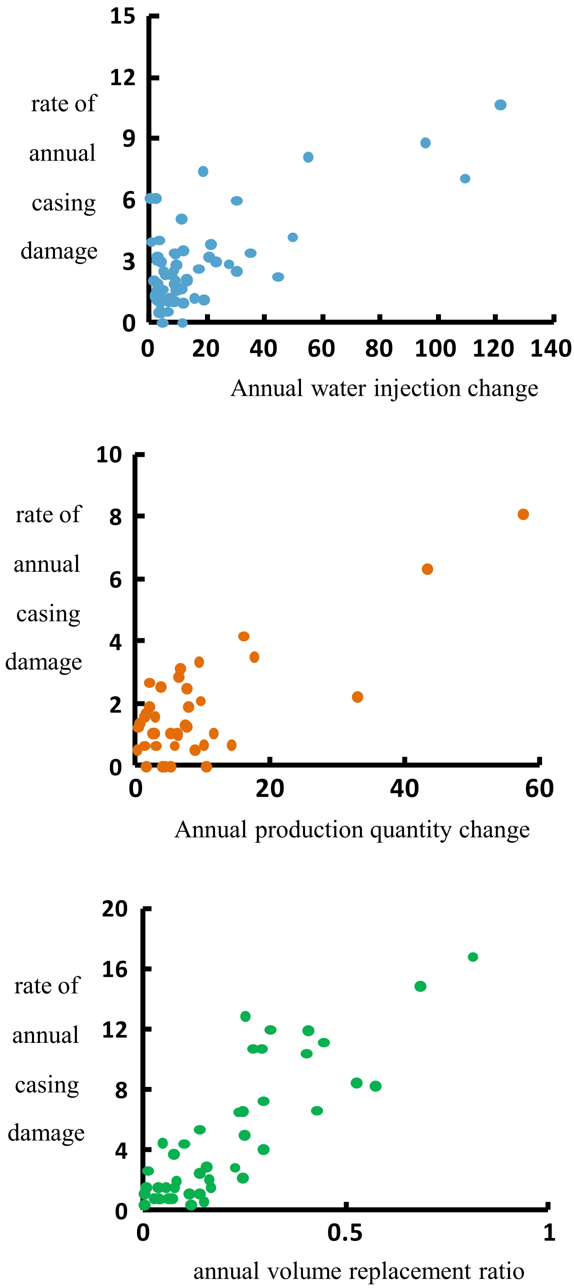

Fig. 3. Annual water injection change, annual production quantity change, annual volume replacement ratio relation with the rate of annual casing damage

time of the maximum correlation coefficient can be determined to be the lag time. Then

Table 1. Development parameters based on pressure system

Development factor	Parameter	Unit
Formation pressure	Average differential pressure	MPa
	Percent of wells with absolute value of differential pressure above 1 MPa	%
	Yearly change of differential pressure	MPa
	Change of differential pressure in the past three years	MPa
	Differential pressure between water flooding development wells and EOR wells	MPa
	Average differential pressure between adjacent blocks	MPa
Water injection pressure	Average water injection pressure	MPa
	Percent of wells with pressure within 0.3 MPa under upper limit	%
	Pressure fluctuation	MPa
	Percent of wells with abnormal high pressure	%
	Percent of wells with abnormal high fluctuations in a year	%
	Magnitude of change of water injection pressure in a year	%

move the parameter column data backward the lag time to obtain new parameter data.

Based on the above 1/2 blocks' parameter has medium or strong correlation with the rate of annual casing damage, 16 Parameters were selected.

3.3 Dealing with Correlation Between Parameter

Due to the fact that relativity does exist between the development parameters, the relativity has to be eliminated using the correlation coefficient before the warning model can be established. Only one among all the parameters with high relativity will be kept. At last, nine parameters be selected which are: average differential pressure, yearly change of differential pressure, change of differential pressure in the past three years, differential pressure between waterflooding development wells and EOR wells, average differential pressure between adjacent blocks, percent of wells with pressure within 0.3 MPa under upper limit, magnitude of yearly change of water injection, magnitude of yearly change in liquid production, volume replacement ratio of waterflooding development wells.

It will be difficult to establish a model with too many parameters being present, that is why the number of parameters needs to be reduced with the method of principal component analysis. In this kind of analysis, the parameter groups will be broken into new parameters with no relativity in attempt to have a more general reflection on the collected information. Then nine new parameters will be created with the first three data being kept as the foundation of the casing damage forewarning model.

Table 2. Development parameters based on injection-production parameter

Development factor	Parameter	Unit
Water injection rate	Injection intensity	m^3/d.m
	Standard deviation of injection intensity	m^3/d.m
	Magnitude of yearly change of water injection	%
	Magnitude of change of water injection in the past three years	%
	Percent of wells with annual magnitude of change above 10%	%
	Percent of wells with annual magnitude of change above 20%	%
	Magnitude of change of water injection in the past year	%
	Percent of wells with magnitude of change above 20% in the past year	%
	Percent of wells with magnitude of change above 30% in the past year	%
Liquid production rate	Production intensity	t/d.m
	Standard deviation of production intensity	t/d.m
	Magnitude of yearly change in fluid production	%
	Change in liquid production in the past three years	%
	Percent of wells with annual production fluctuation above 10%	%
	Percent of wells with annual production fluctuation above 20%	%
	Magnitude of change of annual production in the past year	%
	Percent of wells with production fluctuation above 20% in the past year	%
	Percent of wells with production change fluctuation 30% in the past year	%
Volume replacement ratio	Absolute difference between the parameter and the desired value	
	Annual change of volume replacement ratio	
	Volume replacement ratio of water flooding development wells	
	Volume replacement ratio of EOR wells	
	Total parameter fluctuation in the past year	
	Total parameter fluctuation in the past year (water drive)	
	Absolute difference between the parameter of water flooding development wells and the EOR wells	
	Absolute difference between the fluctuation of water flooding development wells and the fluctuation of EOR wells in the past year	
	Total parameter fluctuation in the past year (EOR)	

4 Comprehensive Evaluation on the Casing Damage Forewarning Parameters

After the parameters are being selected, a comprehensive evaluation is required for all of the parameters. Data mining method will be selected to establish mathematical model, while the historical data of parameters will be used as samples of evaluation in terms of the prediction of the risk of casing damage in the present year [12].

Due to the fact that casing damage is a complicated problem involving many factors, there is a time delay between the factors and the time when the damage happens, and there is complex internal relationship between factors. Based on the patterns discovered among the casing damage data flow, the method of support vector machine (SVM) is being used in the overall research of the casing damage risk forecast. SVM is a small sample statistical learning theory based on the principle of structural risk minimization and the concept of VC dimension. The aim is to obtain the global optimal solution under the existing information, not only the optimal value when the sample tends to infinity, but also ingeniously solves the dimension question. The complexity of the algorithm is independent of the sample size and is suitable for solving complex problems of multiple influencing factors.

Table 3. Matching rate of high casing damage rate of the casing damage forewarning model

Geology level	1st	2nd	3rd	4th	5th	Total
Total number of samples	23	25	31	13	2	94
Matching samples	16	17	25	8	2	68
Matching rate	69.6	68	80.6	61.5	100	72.3

Based on the annual production data of the blocks' development forewarning parameters, different forewarning model is established based on different degree of geological factors of casing damage potential risk. The main focus of this system is to deal with the group of blocks in the category of "High Risk(Casing rate above 3%)". The accuracy of the model is approximately 72.3% (Table 3).

5 Conclusions

In this paper, the big data analysis method is used to predict the casing damage trend, and the basic data of Daqing Changyuan oilfield is used to establish the prediction model and used in verification and interpretation.

1. There are many factors that cause casing damage. The development factor is the primary focus of the research as it is controllable. While evaluating the primary parameters, the hysteresis has to be eliminated.
2. The method of correlation coefficient and principal component analysis is used to establish the forewarning parameters system of casing damage.

3. Support vector machine is being used to finalize the casing damage warning model. The model realizes the prediction of block's casing damage risk which enables high casing damage risk blocks to be detect earlier therefore adjust and control high-risk parameters in advance and prevent casing damage.

Acknowledgements. This project was subordinate to China Petroleum and Natural gas co., LTD., a major science and technology projects "Daqing oil and gas sustainable development effective key technology research and application"(2016E-0205), and this work was support by exploration and development research institute of Daqing oilfield, we thank Yanming Pang, Qinglong Du, Qing Li, for their helpful comments, we are grateful to Chunyu Zhao, Xueyan Jiang for their significant technical contributions to this project.

References

1. Fredrich T, Arguello JG. Geomechanical modeling of reservoir compaction, surface subsidence, and casing damage at the Belridge Diatomite Field. SPE.
2. Guo Y, Blanford M, Candella JD. Evaluating the risk of casing failure caused by high-density perforation: a 3D finite-element-method study of compaction-induced casing deformation in a deepwater reservoir, Gulf of Mexico, SPE.
3. Bozhong H, Zhiliang X. Casing failure mechanisms and protection measures of oil and water wells in Daqing Oilfield. Oil Drill Prod Technol. 1998; 05.
4. Lian Z, Luo Z, Yu H, Bu H, Caixiong, Changping LI, Finite element analysis and control measures on casing failure in sand shale interbeds. Oil Drill Prod Technol. 2016; 38: 887–892.
5. Wu E, T Yan. The fossil layers in Nen2 marker bed of Daqing oilfield cause casing damage mechanism analysis. J Northeast Petrol Univ. 2007;31(2):38–41.
6. Liu H, Liu J, Zhou S, Jin Y. Geological controlling factors of the piece of casing damage in the Nen2 marker bed of Daqing oilfield. Acta Petrol Sin. 2006;27(5):135–8.
7. Du B, Shi Z, Zhu W, Feng C. Integration prevention technique of tubing damage in ShengLi oilfield. Oil Drill Prod Technol. 2004;26(06165):66.
8. Li Z, Song J, Zhou G. Casing damage causes and microscopic mechanism in Sandstone reservoirs. Energy Conserv Petrol Petrochem Ind. 2005;21(8):40–2.
9. Li S, Sun X, Zhao L. Analyse of casing failure and its preventive measure of Gaotaizi fault in Daqing Oilfield. Oil Drill Prod Technol. 2001;23(62–65):84.
10. Liu C, Liu Y, Mingsheng et al. Method and application of regional warninWeng for geohazards in China. Beijing, Geolojical Publishing, 2009; 12.
11. Liu C. Discussion on the causes of casing damage and prevention measures in Pubei 1 fault block. China Petrol Chem Eng Stand Qual. 2012; 32(7): 166–166.
12. Gu R. Research on large data processing technology and system, study for a period. Nanjing University, Internet Publishing; 2017.
13. Yang L, Wang X, Cui Z, Sun S, Chen W. The casing destruction Mechani sm and prevention measures in Jidong Oilfield. Petrol Eng Constr. 2005;31(z1):72–6.
14. Gang Lu. Casing failure mechanism study and control in Gudao Oilfield. Special Oil Gas Reservoir. 2004;11(6):71–3.