

# Vibration-based Processing and Classification Method for Oil Well-testing Data from Downhole Pressure Gauges

Feng Xin<sup>1</sup>, Li Shaohui<sup>2\*</sup>, Feng Qiang<sup>3</sup>, Liu Shugui<sup>1\*</sup>

<sup>1</sup>State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China; fengxin182@tju.edu.cn (F.X.); sgliu@tju.edu.cn (L.S.)

<sup>2</sup>Tianjin Research Institute of Water Transport Engineering, Tianjin 300000, China; shlisfw@hotmail.com

<sup>3</sup>CNPC Bohai Drilling Engineering Company Ltd., Tianjin 300457, China; feng\_qiang@cnpc.com.cn

**Abstract.** During petroleum exploration and exploitation, the oil well-testing data collected by pressure gauges are used for monitoring the well condition and recording the reservoir performance. However, due to the large number of the collected data, the classification of this large volume of data requires a previous processing for the removal of noise and outliers. It is impractical to partition and process these data manually. Vibration-based features reflect geological properties and offer a promising option to fulfil such requirements. Based on the 75 on-site measured samples, the time-frequency-domain features are extracted and the classification performance of three classical classifiers are investigated. Then the downhole data processing and classification method is present by analysing the cross interaction of different types of data features and different classification mechanism. Several feature combinations are tested to establish a processing flow that can efficiently remove the noise and preserve the shape of curves, high signal to noise ratio rates, with minimum absolute errors. The results show that optimal multi-feature combination can achieve the highest working stage identification rate of 72%, the parameters optimized support vector machine can achieve the better classification performance than other listed classifiers. This paper provides a theoretical study for the data denoising and processing to enhance the working stage classification accuracy.

## 1 INTRODUCTION

Long-term downhole pressure monitoring is playing an important role in improving reservoir management and workflow optimization. However, with the deepening of oil well exploration and development, the downhole formation structure is becoming more complex, the types of the difficult operating wells is enriching, including the deep and ultra-deep wells, the high temperature and high pressure wells, and the high-angle horizontal well. Such exploration faces the problems with noise, attenuation and nonlinear distortion [1].

Also, to identify the underground production status under different well types and working conditions, the ground analysts identify the working stages from the huge underground measurement data artificially, which brings a big inconvenience and man-made interference to data interpretation. Meanwhile, the data are gathered under unattended situation and usually contain some outliers and noises. Add up the large amount of downhole pressure data, the traditional manual classification of flow process is more time-consuming and laborious. Therefore, the smart well technology is becoming an important methodology to study the downhole well-testing data processing and the data interpretation [2]. Many published papers described the methods of processing and interpreting pressure data.

Elsharif [3] extracted the distribution characteristics of pressure mutation and identified the damaged zones based on the pressure transient analysis when evaluating reservoir matrix chemical treatment in real time. The method used linear transformation to extract single feature of signal, but cannot fully reflect the nonlinear structure of downhole parameter data. Andreas [4] used wavelet analysis to detect the four stages of deterministic sedimentary pattern. However, few attempts have been made to develop a comprehensive model for working stage decisions. The latest researches [5] introduced the multi-feature extraction and the deep learning methods, although obtained a good classification effect, its complex structure is more suitable for offline data processing. Generally, the current difficulties can be summarized as the following three aspects:

1. The stochastic of data due to the variousness and the complexity of the downhole geological conditions.
2. The incompleteness of the well testing record and the unsatisfactory quality of the available datasets.
3. The features used for classification has few types and the superiority of each is not considered.

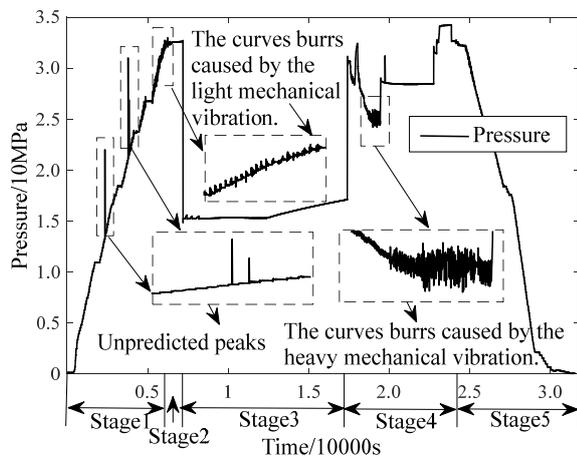
Our objective is to analyse the classification effects of the different feature extraction methods and the different classifiers, and propose an integrated vibration-based processing flow to classify the well-testing segments, determining the working stage they belong to.

\* Corresponding author: sgliu@tju.edu.cn, shlisfw@hotmail.com

## 2 DATA SOURCE AND METHOD

### 2.1 Data Source

The datasets are collected from the well testing platform of Huabei Oilfield, China, which is the same with the datasets we used in our previous study [6]. There are 15 field samples containing 75 stage samples in total, all of which have good quantity. The reservoirs in this filed have the feature of special lithology, structural fractures, and strong edge water. As shown in Fig. 1, a complete well test process includes five stages. The pressure data processed are divided into data segments according to their working stages. The processing objects are the segments that belong to the above five stages.



**Fig. 1.** The actual well-testing curve. Five stages are: Stage1: Lowering the oil string. Stage2: Waiting stage. Stage3: Well opening. Stage4: Well closing. Stage5: Pulling up the oil string.

Fig. 1 shows that the pressure and time are nonlinear relationship, there exists unpredicted peaks from the changeable geological conditions and various intensity of interference from the inevitable mechanical vibration. These peaks and interference are mixed into the pressure data that we used for further working stage classification, which will lead to the inaccurate identification of signal feature points and the poor characteristics representing ability of extracted features. If applying unfiltered data to working stage classification directly, it could lead to false-decision of the system due to the measured data corrupted with noise and outliers.

As for data filtering, the peaks seriously deviate from the average trajectory of data points. The time-domain method can predict the motion trend of the data and quickly find the outliers. The mechanical vibration interference belongs to the local characteristic and it shows the aggregation with different density. The time-frequency-domain method can observe the data from both the time-domain and the frequency-domain aspects and abstractly characterize and denoise the data [6].

### 2.2 Proposed Method

The main task of this paper mainly includes three aspects:

1. Apply the coefficient denoising approach with wavelet transform to filter the data noise.

2. Extract multiple vibration features of well-testing data from the viewpoint of the time domain.

4. Perform feature classification of well-testing working stage by using different classifiers.

3. Analyse the classification effect of the combination features formed by different types of features.

#### 2.2.1 Wavelet Based Denoising

Currently, there is no unified standard for wavelet base selection as its applications are different. Based on our previous research, the Symlets wavelet has good local characteristics in the frequency domain [7]. Therefore, the wavelet denoising in this paper performs the sym6 based 6-scale decomposition, which includes two steps:

1. Decompose a signal in low-frequency and high-frequency coefficients by using the discrete wavelet transform (DWT) based multiresolution analysis.

2. Modify the coefficients according to the thresholding method and remove the noise and outliers.

3. Reconstruct the original data and obtain  $S_d$  by using the  $C_r$  after threshold quantization.

Where, the thresholding method is the Hierarchical based Wavelet method.  $S_d$  is the denoised data after reconstructing the original pressure data.  $C_r$  is the wavelet coefficients that used for data reconstruction.

#### 2.2.2 Vibration Features Extraction

Considering the influence of mechanical vibration on underground geological data, we introduce the common time domain characteristic indexes as the input features. Also, one type of feature is not enough to characterize the information of data fully. As shown in Table 1, here we use the following five indicators as the time domain characteristics for each working stage data. Then the vibration feature  $V_f$  can be expressed as:

$$V_f = [K_v, Cl_f, I_f, C_f, S_f], \quad (1)$$

**Table 1.** Time domain characteristic index.

Quantity	Symbol	Definition
Kurtosis Value	$K_v$	$K_v = \beta / X_{rms}$
Clearance Factor	$Cl_f$	$Cl_f = X_{max} / X_r$
Impulse Factor	$I_f$	$I_f = X_{max} / \text{abs}(X)$
Crest Factor	$C_f$	$C_f = X_{max} / X_{rms}$
Shape Factor	$S_f$	$S_f = X_{rms} / \text{abs}(X)$
Pressure derivatives	$P_d$	$P_d = \Delta P / \Delta t$

Where  $X_{rms}$ ,  $\text{abs}(X)$ ,  $X_{max}$ ,  $X_r$  and  $\beta$  are mean square value, mean absolute value, maximum, root mean square value, and kurtosis. Here, we take another time-domain based feature pressure derivatives  $P_d$  into consideration and use all these features for classification and performance judgement. P is pressure and t is time.

### 2.2.3 Working Stage Classification

This paper uses the majority vote algorithm in the LIBSVM developed by Lin Chih-Jen [8] for the final working stage classification. This algorithm has strong compatibility and needs only two adjustable parameters: penalty parameter  $c$  and kernel function parameter  $g$ .

Then the identified stage  $Stage'$  can be expressed as:

$$Stage' = \operatorname{argmax} \{n_c \mid f(D_n, n_c) = m_c\}, \quad (2)$$

where  $f(D_n, n_c)$  ( $n_c = 1, 2, \dots, \text{num}$ ) denotes the decision function of the  $n_c^{\text{th}}$  classifier,  $\text{num}$  is the combination of  $m_c$  and 2,  $m_c$  is the type number of the well-testing working stage,  $D_n$  is the data segments representing each stage. For a given test sample, define the number of classifiers that classify  $D_n$  as a stage  $m_c$  as  $num_{m_c}$ .

$$num_{m_c} = \{n_c \mid f(D_n, n_c) = m_c\} \quad (3)$$

At last, the stage with the largest  $num_{m_c}$  will be the stage that the sample belongs to. What's more, in order to obtain the best parameters  $c$  and  $g$ , we introduce particle swarm algorithm (PSO) to avoid the occurrence of overlearning and underfitting states.

## 3 Results and Discussion

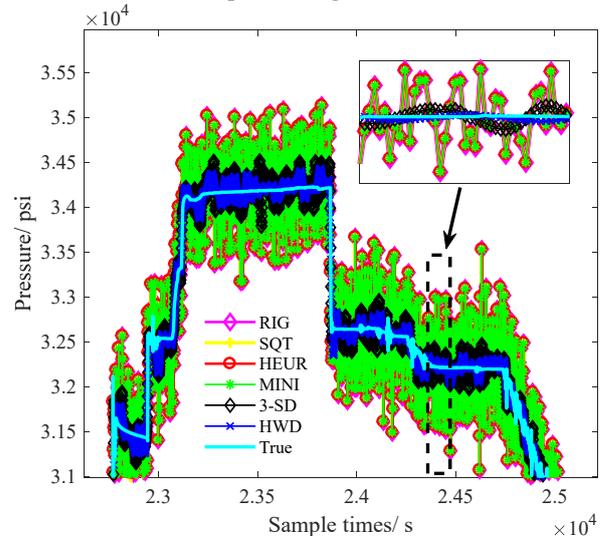
A signal processing method is hard to extract the features with strong data characteristics for uncertain change of fluctuating and trend. To improve the final classification performance, firstly, the original downhole pressure data are processed by wavelet based denoising method. As shown in Table 2, the robust of four global threshold methods and two high-frequency threshold methods are analysed. Hierarchical based threshold method [9] obtains the lowest signal to noise ratio (SNR). The resolution scale used for the data decomposition determines the final data denoising results. The 6 decomposition scale (6-DS) is more suitable for the well-testing data processing than that of less than 6, which can obtain better denoising effect.

**Table 2.** Robust analysis of data denoising based on different threshold methods. 3-SD: 3-times Standard Deviation, HWD: Hierarchical Wavelet Denoising, DS: Decomposition Scale.

Type	Method	SNR
Global Threshold	Rigsure (RIG)	28.864
	Sqtwolog (SQT)	28.948
	Heursure (HEUR)	28.864
	Minimaxi (MINI)	28.921
High-frequency Threshold	3-SD	36.897
	HWD(3-DS)	25.873
	HWD(6-DS)	23.921

All the simulations are conducted using Matlab 2017a. Figure 2 shows that, through output comparison, the HWD can better remove the data noise and enhance the data quality. The recognition rate of edge enhanced

data has a remarkable improvement compared with those of the original data and other processed data. Compared with the global threshold and 3 decomposition scale (3-SD) threshold algorithms, the HWD can better achieve the purpose of pressure data denoising caused by mechanical vibration. It strengthens the pressure change trends apparently and keeps the time domain characteristics of original data highly at the same time. By using the HWD and wavelet reconstruction, the curve burrs and unpredicted peaks in pressure data successions can be easily detected, filtered and eliminated, which is more conducive to the recognition of data features and further classification processing.



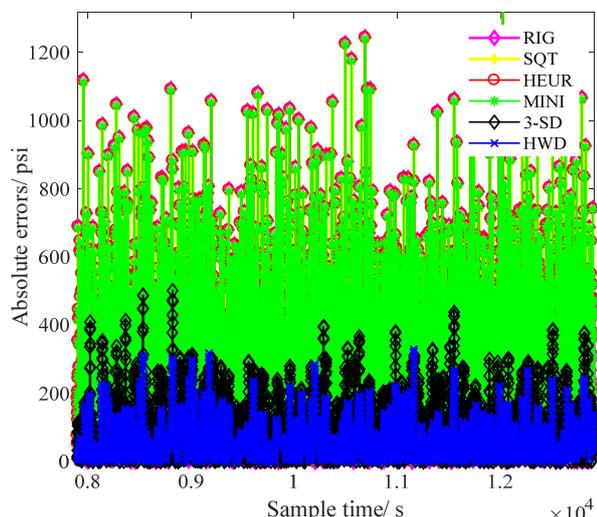
**Fig. 2.** Output comparison under different threshold methods. True means the data from far-surface, where the pressure gauge can capture more precise geological changes.

Secondly, the multiple types of vibration features are extracted and the accuracy comparison of three classical classifiers are made. Three classifiers are Neural Networks (NNs), k-Nearest Neighbors (KNN), and support vector machines (SVM). As show in Table 3, the classification effect of different methods varied with integrating different types of features. The combination of SVM and feature with  $V_f$  and  $P_d$  can achieve the highest classification efficiency, it is 72%, though it is not high, the method is proved efficient and feasible in deleting meaningless noise information and improving greatly feature mining and classification ability.

**Table 3.** Accuracy comparison of different classifiers under different feature combinations. P-SVM: PSO based SVM

Features	Accuracy(%)			
	NNs	KNN	SVM	P-SVM
$V_f$	48.00	49.33	48.00	49.33
$P_d$	42.67	62.67	61.33	68.00
$W_f$	41.33	52.00	54.67	61.33
$V_f + P_d$	49.33	58.67	52.00	72.00
$P_d + W_f$	61.33	62.67	61.33	65.33
$V_f + P_d + W_f$	61.33	60.00	61.33	72.00

This paper also introduces the wavelet features and systematically analysis the classification efficiency under different combinations of the total seven time-domain and time-frequency-domain features. Define the wavelet feature as  $W_f$ . Table 3 and Figure 3 demonstrate through simulations with downhole pressure data that the performances of the proposed method surpass that of previously classical method both visually and in terms of the classification accuracy and the absolute error.



**Fig. 3.** Absolute errors comparison under different threshold methods.

Meanwhile, this paper uses PSO to solve the difficulty of determining the optimal parameters. Table 3 gives the comparisons with NNs and KNN, by simulation of the datasets in our previous research [6], the result shows that PSO based SVM classifier can do better in global optimal solution, which is mainly manifested in its powerful classification ability. The proposed integrated data processing and classification flow for downhole pressure, which based on the wavelet decomposition, the HWD thresholding based coefficients modification, the vibration analysis based domain feature extraction, and the P-SVM classifier, can satisfy the requirement of actual production and provide theoretical foundation for exact downhole working stage identification.

The proposed method also exists some limitations and shortcomings. Although the influences of vibration are analysed, only seven classical indicators in total are used for classification, and more types of influencing factors should be considered. What's more, 75 field stage samples are analysed in this paper and the sample size is small. Future work should collect more downhole well-testing samples including more complete and complex working conditions to make the proposed method more convincing.

## 4 Conclusion

This paper aims at reflecting the real characteristic of oil well reservoirs, making the test data be more close to the true value, and completing the well testing stage classification. The hierarchical based wavelet denoising and the vibration analysis based time-frequency-domain

characteristics are used successfully for this purpose. The experimental results show that the proposed flow improves the signal-to-noise ratio, achieves the high average classification of 72%, which demonstrates that the data processing, feature extraction method, and the working stage classification approach are correct and effective. It can reduce the effort of oil data analysts and provide reliable reference for future academic research.

## Acknowledgement

This research was supported in part by the Research on key measurement standards and Traceability Technology of water transport engineering, National key research and development plan under No. 2018YFF0212200 and the Special fund for basic scientific research business expenses of central public welfare scientific research institutes under No. TKS20200202.

The filed work was conducted in Huabei Oilfield, China. Thanks to all authors and the staffs on-site from CNPC Bohai Drilling Engineering Company Ltd. for their help in field measurements and data analyses.

## References

1. S.M. Godoy, F.C Massatoshi. A novel competitive learning neural network based acoustic transmission system for oil-well monitoring, *IEEE Trans. Ind. Appl.*, **36**, 484-491(2000)
2. G. Carvajal, M. Maucec, S. Cullick, *Smart Wells and Techniques for Reservoir Monitoring*, (Intelligent Digital Oil and Gas Fields, 2018)
3. T.A. Elsherif, A.A. Balto, F. Baez, et al. *The Use of Real-Time Downhole Pressure and Distributed Temperature Surveying in Quantifying the Skin Evolution and Zonal Coverage in Horizontal Well Stimulation*, (Society of Petroleum Engineers 2012)
4. A. Prokoph, F. Barthelmes. Detection of non-stationarities in geological time series: wavelet transform of chaotic and cyclic sequences, *Comput. Geosci.*, **22**, 1097-1108 (1996)
5. X. Feng, Q. Feng, et al. Automatic Deep Vector Learning Model Applied for Oil-Well-Testing Feature Mining, Purification and Classification, *IEEE Access*, **8**,151634-151649 (2020)
6. X. Feng, Q. Feng, et al. Wavelet-Based Kalman Smoothing Method for Uncertain Parameters Processing: Applications in Oil Well-Testing Data Denoising and Prediction, *Sensors*, **20**, 4541 (2020)
7. Wavelets Information. The MathWorks. Available online: <https://ww2.mathworks.cn/help/wavelet/ref/waveinfo.html> (accessed on 8 Sept. 2020).
8. C.W. Hsu, C.J. Lin, A comparison of methods for multiclass support vector machines, *IEEE Transl. Neural Netw.*, **13**, 415-425 (2002)
9. T. Yi, H. Li, X. Zhao, Noise smoothing for structural vibration test signals using an improved wavelet thresholding technique, *Sensors*, **12**, 11205-11220 (2012)