A post-stack seismic sparse dictionary learning inversion method based on logging data - A case study of the Tarim Basin

Long Chen
PetroChina Tarim Oilfield Company, Korla, China

Xin Zhang (corresponding author)
College of Petroleum Engineering, China University of Petroleum Beijing, China

Li Zhao, Bao Zhou, Yutian Han
PetroChina Tarim Oilfield Company, Korla, China

Yan Ding
CNPC Engineering Technology R&D Co., Ltd., Beijing, China

Yunhu Lu
College of Petroleum Engineering, China University of Petroleum Beijing, China

ABSTRACT: Highly accurate seismic inversion results help refine geomechanical modeling. In this paper, a post-stack seismic sparse dictionary learning inversion method based on logging data is proposed. First, feature functions are extracted from the logging data. Then, a dictionary is learned adaptively from known observations. This dictionary is composed of a series of feature functions, using which the parametric model can be effectively characterized. This method effectively avoids the problem of single mathematical model assumptions. Finally, the post-stack wave impedance inversion data are solved, and the separation of wave impedance data is performed using the correlation between the velocity data and the density data. This method can effectively improve the resolution of seismic inversion results by extracting a priori information from logging data. It is found that the root-mean-square error of the sparse dictionary learning method is reduced by 9.075% compared to the Tikhonov method.

Keywords: sparse dictionary learning, seismic inversion, highly accurate, geomechanical parameters.

1 INTRODUCTION

Due to the lack of drilled well data in the undeveloped area of the field, the only available information is seismic data. The seismic data effectively reflects the rich stratigraphic characteristics such as lithology, tectonics, and physical properties in the study area. However, the seismic data need to be analyzed and converted through a series of processing to obtain velocity data and density data of the study area which can be used as input data for geomechanical models. Currently, there are several methods of using seismic data to obtain the basic input data for geomechanical models.

In conventional studies, the Dix formula is generally used to calculate the layer velocity as a basic geomechanical data (Dix, 1955; Gholami et al, 2019). The method is mainly based on the stacked velocity obtained from the raw seismic data. After dip and phase correction, the stacked velocity is converted to layer velocity. This method is simple and applicable. However, the accuracy depends on the stratigraphic position and the accuracy of the stacked velocity. In particular, the accuracy of this method is not high when the dip angle of the stratum and the lateral variation of the velocity are large.

In addition, the inversion method has made great progress in obtaining geomechanical parameters. Seismic inversion is the process of imaging the physical properties of subsurface rock formations using surface observed seismic data, constrained by known geological laws and drilling and logging data. Seismic inversion mainly includes pre-stack inversion and post-stack inversion according to the seismic data used in the seismic inversion technique (Hampson et al, 2005; Maurya et al, 2020). The pre-stack inversion mainly includes travel-time based laminar imaging technique and amplitude based AVO analysis technique. The post-stack inversion mainly includes travel-time based tectonic analysis technique and amplitude-based wave impedance inversion technique. The wave impedance inversion technique further separates the predicted wave impedance to obtain important basic geomechanical parameters such as velocity and density.

In recent years, many researchers have explored the potential relationship between seismic and rock properties, and proposed many prediction models for geomechanical parameters based on artificial intelligence algorithms such as neural networks (Sun et al. 2020; Zhang et al. 2021; Chen et al. 2021). A nonlinear mapping relationship between seismic data and geomechanical parameters was established. Schultz proposed a method for predicting logging attributes using multiple seismic attributes (Schultz et al. 1994). They used logging and seismic data from 15 wells to train the relationship between seismic and logging attributes. Then they used the trained model to achieve prediction of stratigraphic attributes in undeveloped areas. Himmer and Link used a neural network approach to predict the porosity of the formation using properties such as seismic amplitude and transient amplitude (Himmer and Link et al. 1997). In addition, Schuelke and Hampson proposed a multi-attribute linear regression method to predict logging attributes using seismic data (Schuelke et al. 1998; Hampson et al. 2001). The method is based on linear regression to select the optimal single seismic attribute for logging parameter prediction. Then, the two optimal attributes containing that single best attribute are filtered. This process is continued until the prediction error of the cross-validation data is no longer reduced.

The data sources for the above methods to obtain the required geomechanical base parameters are all seismic data in the work area. Although seismic data contain rich medium-frequency information, they lack low and high-frequency information. The result is that the geomechanical data obtained by conversion of seismic data have a low-resolution problem in the vertical direction (Veeken et al. 2004). The prediction results obtained using these data suffer from low resolution and large scale. The prediction results cannot meet the refined engineering design in drilling and development. Therefore, improving the vertical resolution of the predicted geomechanical base parameters is a problem that needs to be solved in the current study.

2 METHODOLOGY

2.1 Sparse dictionary-constrained wave impedance inversion

The post-stack seismic record s_t is the fold product of the stratigraphic reflection coefficient sequence r_t and the seismic subwave sequence w_t, which can be expressed by the equation

$$s_t = w_t * r_t \tag{1}$$

The discrete form of this folded product model is

$$s_t = \sum_{k=0}^{p} r_k w_{t-k} \tag{2}$$

where $s_t = \{s_0, s_1, \dots, s_n\}; w_t = \{w_o, w_1, \dots, w_m\}; r_t = \{r_0, r_1, \dots, r_p\}, n > m, n > p.$
The discrete form of the folded product model can be expressed as a matrix form

$$\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{n-1} \end{bmatrix} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ \vdots & w_1 & \ddots & \vdots \\ w_k & \vdots & \ddots & 0 \\ 0 & w_k & \vdots & w_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & w_k \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_{n-1} \end{bmatrix} \tag{3}$$

The following relationship is derived from Russell's approximation formula

$$r(t) = \frac{z_{i+1} - z_i}{z_{i+1} + z_i} \cong \frac{\Delta z_i}{2z_i} = \frac{\Delta \ln(z_i)}{2} \tag{4}$$

where z_i is wave impedance. Write equation 4 in matrix form as

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_{n-1} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} lnZ_1 \\ lnZ_2 \\ \vdots \\ lnr_n \end{bmatrix} \tag{5}$$

Substitute equation 4 into equation 3,

$$\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{n-1} \end{bmatrix} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ \vdots & w_1 & \ddots & \vdots \\ w_k & \vdots & \ddots & 0 \\ 0 & w_k & \vdots & w_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & w_k \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} lnZ_1 \\ lnZ_2 \\ \vdots \\ lnr_n \end{bmatrix} \tag{6}$$

Write equation 6 as

$$s = Gx \tag{7}$$

where G is the orthogonal matrix and x is the logarithm of the stratigraphic wave impedance. Then, the objective function of the seismic inversion can be obtained

$$x = \operatorname{argmin}\{\|s - Gx\|_2^2\} \tag{8}$$

The result of inversion is to solve the above objective function to obtain x.

However, seismic inversion suffers from multi-solvability and low resolution, which are the main problems faced by traditional inversion methods. The most direct way to solve the inversion resolution problem is to add high and low frequency information for constrained inversion by means of regularization. If the Tikhonov regularization constraint is added, the objective function (8) is rewritten as the following equation

$$x = \underset{x}{\operatorname{argmin}}\{\|s - Gx\|_2^2 + \lambda\|Dx\|_2^2\} \tag{9}$$

where λ is the regularization parameter and the matrix D is the smoothing operator. Such regularization constraint methods assume that the strata satisfy a single distribution law and are no longer applicable to the problem of inversion of complex geological features.

The introduction of high-frequency information from logging data can significantly improve the resolution of seismic inversion results. However, the method of introducing information from logging data based on regularization terms has the problem of insufficient mining of valid information. Sparse characterization methods based on learning dictionaries can solve this problem. The main component of dictionary learning is to learn a dictionary adaptively from logging data of drilled wells, which consists of a series of feature functions. These feature functions can effectively characterize the parametric model. Dictionary learning learns formation characteristics directly from logging data without making any assumptions about the structure of the formation model or the distribution of formation parameters. The method effectively avoids the problem of single model assumptions that exist in traditional inversion constraints.

The workflow for dictionary learning using drilled well log data is shown in Figure 1. First, the sliding window size and sliding step size are selected. Second, obtaining the training sample set of the dictionary on the logging data. Third, feature training using dictionary learning algorithm. Finally, the wave impedance is sparsely characterized using the learned dictionary.

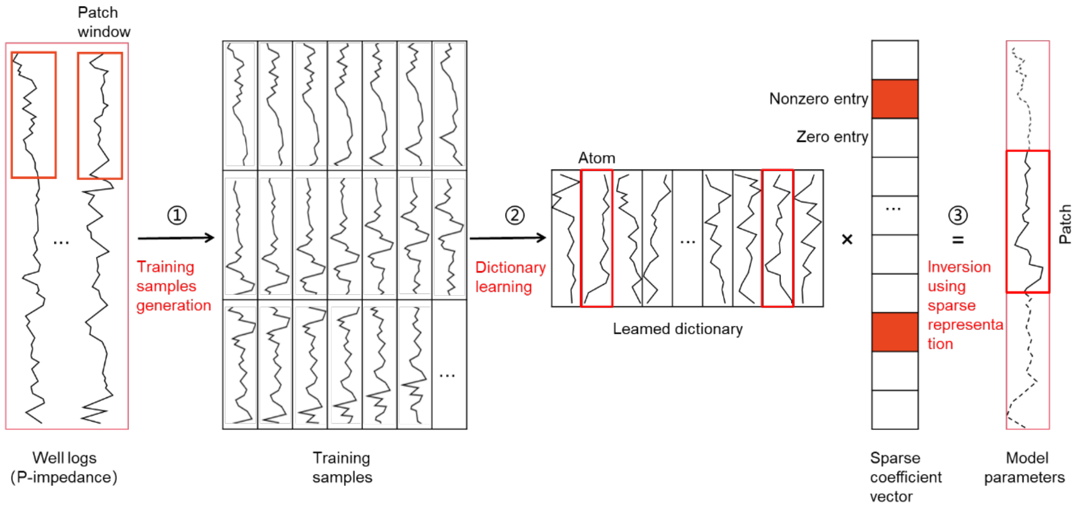

Figure 1. Sparse dictionary learning based on Logging Data.

Using a dictionary-based sparse representation constraint instead of the traditional regularization constraint, equation (9) is rewritten in the following form

$$x = \arg\min_{x,\, v_i} \left\{ \|s - Gx\|_2^2 + \beta\|y - x\|_2^2 + \lambda \sum_i^n \|R_i y - D v_i\|_2^2 \right\} \tag{10}$$

$$\forall i\colon \|v_i\|_0 \le \varepsilon \tag{11}$$

Where λ is the regularization factor; n represents the number of blocks obtained from each column of data; R is the block operator, R_i is the ith block extracted by the sliding window along the direction of increasing depth of wave impedance; D is the sparse dictionary obtained in the previous step; v_i is the sparse coefficient of the ith block, and ε is the maximum allowed by the sparsity.

The inverse problem is then subdivided into the following three subproblems

$$x^{k+1} = \underset{x}{\operatorname{argmin}} \{\|s - Gx\|_2^2 + \beta\|y - x\|_2^2\} \tag{12}$$

$$v_i^{k+1} = \arg\min_{v_i} \left\{ \sum_{i=1}^n \left\|R_i y^k - D v_i\right\|_2^2 \right\} \tag{13}$$

$$y^{k+1} = \underset{y}{\operatorname{argmin}} \left\{ \lambda\left\|y - x^{k+1}\right\|_2^2 + \sum_{i=1}^n \left\|R_i y - D v_i^{k+1}\right\|_2^2 \right\} \tag{14}$$

The inverse result x is obtained by iteratively updating the three parameters in a continuous loop.

2.2 Velocity and density separation

The defining equation of wave impedance is

$$I = \rho v \tag{15}$$

Where ρ is density, v is velocity. There is a significant correlation between the layer velocity and the stratigraphic density. For example, the well-known Gardner's empirical formula is

$$\rho = a \cdot v^b \tag{16}$$

Where a and b are constant parameters. Bringing the above equation into the expression for wave impedance, we get

$$I = a \cdot v^{b+1} \tag{17}$$

Then, the wave impedance data can be separated into density data and velocity data.

3 CASE STUDY

An oil field in the Tyuritag tectonic zone of the Tarim Basin was selected as the object of study. The initial model is constructed using log data from three drilled wells, and the other well is used as a verification well, which is not involved in the construction of the initial model. Sparse dictionary constrained inversion of the established initial model to obtain depth domain impedance data. Figure 2 show the wave impedance inversion results for the Tikhonov regularization constraint and the sparse dictionary constraint, respectively. The figures show a continuous well profile consisting of modeling and verification wells with a total of 549 CDP seismic traces. Comparing the inversion results, we can see that the sparse dictionary constrained inversion method has higher resolution. The method can better characterize the details of the stratum.

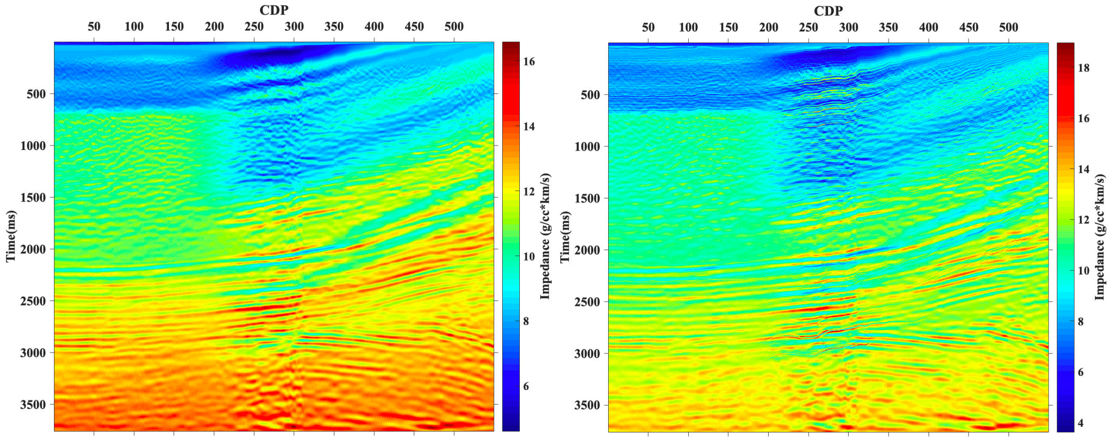

Figure 2. Inversion results based on Tikhonov regularization constraints (left) and sparse dictionary learning constraints (right).

In order to quantify the accuracy of the inversion results, we extracted the inversion results corresponding to one drilled well. Compared with the logging data, the root mean square error of the conventional method is 1.157, while the error of the sparse dictionary learning method is smaller at 1.052. The sparse dictionary learning method improves 9.075% on this performance metric.

Separate the high-resolution 3D wave impedance data obtained from the sparse dictionary-constrained inversion of this area into velocity and density data. Figure 3 shows the 3D velocity and density data. Velocity and density data are the basic data of geomechanics and can be directly used for geomechanical modeling.

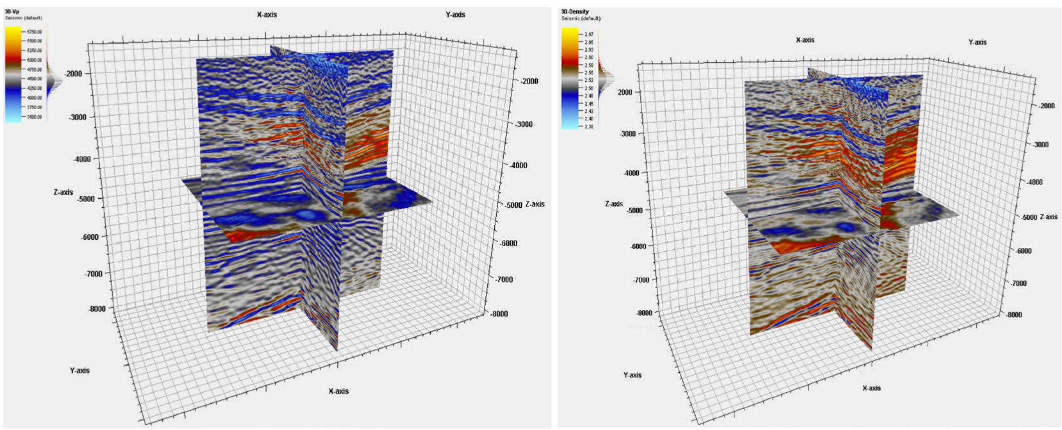

Figure 3. 3-D depth domain p-wave velocity data (left) and density data (right).

CONCLUSIONS

1. a post-stack seismic sparse dictionary learning inversion method based on logging data is proposed. The new method can effectively improve the resolution of seismic inversion results by extracting a priori information from existing logging data.

2. The root-mean-square error of the sparse dictionary learning method is reduced by 9.075% compared to the Tikhonov method.

ACKNOWLEDGEMENTS

This study was supported by the National Key Research and Development Technology Project, No. 2019YFA0708303-05.

REFERENCES

Chen, Y., & Saygin, E., 2021, Seismic inversion by hybrid machine learning. Journal of Geophysical Research: Solid Earth, 126(9), e2020JB021589.

Dix, C. H., 1955, Seismic velocities from surface measurements: Geophysics, 20, 68–86.

Gholami, A., & Zabihi Naeini, E., 2019, 3D Dix inversion using bound-constrained total variation regularization. Geophysics, 84(3), R311-R320.

Hampson, D. P., Russell, B. H., & Bankhead, B., 2005, Simultaneous inversion of pre-stack seismic data. In SEG Technical Program Expanded Abstracts 2005 (pp. 1633-1637). Society of Exploration Geophysicists.

Hampson, D. P., Schuelke, J. S., and Quirein, J. A., 2001, Use of multi-attribute transforms to predict log properties from seismic data: Geophysics, 66, 220–236.

Himmer, P., and Link, C. A., 1997, Reservoir porosity prediction from 3-D seismic data using neural networks: 67th Annual International Meeting, Society of exploration Geophysicists, 842–845.

Maurya, S. P., Singh, N. P., Singh, K. H., Maurya, S. P., Singh, N. P., & Singh, K. H., 2020, Post-stack seismic inversion. Seismic inversion methods: A Practical approach, 39-80.

Schuelke, J. S., and Quirein, J. A., 1998, Validation: A technique for selecting seismic attributes and verifying results: 68th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 936–939.

Schultz, P. S., Ronen, S., Hattori, M. and et al, 1994, Seismic guided estimation of log properties, parts 1, 2, and 3: The Leading Edge, 13, 305–310,674–678,770–776.

Sun, J., Niu, Z., Innanen, K. A., Li, J., & Trad, D. O., 2020, A theory-guided deep-learning formulation and optimization of seismic waveform inversionTheory-guided DL and seismic inversion. Geophysics, 85(2), R87-R99.

Veeken, P. C. H., & Da Silva, A. M., 2004, Seismic inversion methods and some of their constraints. First break, 22(6).

Zhang, J., Li, J., Chen, X., Li, Y., Huang, G., & Chen, Y., 2021, Robust deep learning seismic inversion with a priori initial model constraint. Geophysical Journal International, 225(3), 2001-2019.