Cognitive Seismic Data Modelling based Successive Differential Evolution Algorithm for Effective Exploration of Oil-Gas Reservoirs

Jing Zhao[1], Jinchang Ren[2,3*], Jaime Zabalza[3], Jinghuai Gao[4], Xinying Xu[3*], Gang Xie[5]

1 School of Earth Science and Engineering, Xi'an Shiyou University, Xi'an, China.
2 School of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan, China
3 Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK
4 Institute of Wave & Information, School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China.
5 Taiyuan University of Science and Technology, Taiyuan, China

Abstract— **A cognitive modelling based new inversion method, the successive differential evolution (DE-S) algorithm, is proposed to estimate the Q factor and velocity from the zero-offset vertical seismic profile (VSP) record for oil-gas reservoir exploration. The DE algorithm seeks optimal solutions by simulating the natural species evolution processes and makes the individuals become optimal. This algorithm is suitable for the high-dimensional nonseparable model space where the inversion leads to recognition and prediction of hydrocarbon reservoirs. The viscoelastic medium is split into layers whose thicknesses equal to the space between two successive VSP geophones, and the estimated parameters of each layer span the related subspace. All estimated parameters span to a high dimensional nonseparable model space. We develop bottom-up workflow, in which the Q factor and the velocity are estimated using the DE algorithm layer by layer. In order to improve the inversion precision, the crossover strategy is discarded and we derive the weighted mutation strategy. Additionally, two kinds of stopping criteria for effective iteration are proposed to speed up the computation. The new method has fast speed, good convergence and is no longer dependent on the initial values of model parameters. Experimental results on both synthetic and real zero-offset VSP data indicate that this method is noise robust and has great potential to derive reliable seismic attenuation and velocity, which is an important diagnostic tool for reservoir characterization.**

***Index Terms*—Successive differential evolution algorithm, VSP data, high dimensional data, velocity and *Q* inversion**

I. INTRODUCTION

The exploration targets are turning from conventional to unconventional reservoirs with the development of oil-gas exploration technology [1]. How to finely describe the medium structure, lithology and saturation of fluids is a critical

problem in exploration geophysics. Seismic waves propagating through the earth suffer attenuation and dispersion due to the viscosity of the media. The inherent attenuation of the medium is usually quantified by the quality factor Q which is a diagnostic tool for hydrocarbon detection and reservoir characterization. Meanwhile, velocity is one of the most important earth parameters that determine the accuracy of seismic imaging in exploration geophysics. Therefore, Q and velocity estimation are of great importance.

Direct estimation methods and inversion methods are usually two different approaches to estimate the Q factor and velocity. On one hand, the direct estimation approaches can be divided into time domain, frequency domain, and time-frequency domain methods. On the other hand, there are also many inversion estimation methods such as waveform inversion and tomography. This paper focuses on the waveform inversion methods.

Researchers have already developed various waveform inversion approaches such as local optimization waveform inversion and global optimization. In local optimization, the minimum of the objective function can be determined by gradient, conjugate gradient, Newton, Gauss-Newton, and quasi-Newton methods. Among these, the gradient-based waveform inversion methods have got some success, but they have the limitation of nonlinearity of the inversion and dependence on the initial model. In fact, Virieux [2] pointed out that some challenges of waveform inversion are related to building exact initial models, defining new minimization criterion and improving multi-parameter inversion capability.

On the other hand, there are many global optimization methods such as evolutionary algorithm (EA) [3], simulated annealing algorithm [4] and the Monte Carlo method [5]. Moreover, EA can be divided into differential evolution (DE), genetic algorithm (GA), ant algorithm, and so on. Compared with the gradient-based waveform inversion method, the global optimization methodology is less dependent on the initial values and requires no gradient. However, it needs a large amount of computation. Despite this fact, it is widely used in recent years with the development of high performance computing.

On a different note, cognitive computing can handle human problems. It has the capability in studying, reasoning, and solving specific problems. Cognitive systems can understand the problems adequately and model the thought process as computing models. Their ability of studying abstract features is similar to the study processing of the brain. Hence the cognitive computing can enhance the human intellectual and decision-making capacity.

The EAs are a kind of cognitive computing, they simulate the natural species evolution processes by computers, and make-the individual become optimal. However, the conventional EAs often lose their effectiveness when applied

to high-dimensional problems. Z. Pan [6] proposed CRsADE method which use individual crossover rate and subcomponent crossover rate to adaptively improve the efficiency of crossover operation. C. Wang proposed DE with cooperative coevolution selection (DE-CCS) [7] and DE with cooperative coevolution mutation (DE-CCM) [8]. These methods divided the high-dimension problem to several subproblems and judged the subproblems by local fitness functions. R. Chandra [9] proposed coevolutionary multi-task learning used for multi-step chaotic time series prediction. This paper presents a network architecture which is capable predict multi-step. Z. Gao [10] proposed a highly efficient DE algorithm (HEDE) which used a new population evolution strategy to decrease the population size. A new multimutation scheme was later proposed to converge quickly. [11]. X. Cui [12] improved particle swarm optimization used for poststack impedance inversion. This method combined the swarm intelligence and probabilistic theory for global optimization. S. Mahdavip [13] reviewed metaheuristics in large-scale global continues optimization, including large-scale global optimization (LSGO), evolutionary algorithm (EAs), cooperative coevolution (CC). Although lots of DEs such as cooperative coevolution (CC) [14], ND-CC [15] and Co-evolutionary multi-task learning were proposed to solve high-dimensional optimization problems, they only work well with the separable problem, i.e., the existing CCs often lose their advantages when applied to high-dimensional nonseparable model space. Meanwhile, most existing methods have the problems of "dimensional bottleneck". The optimizing ability of these methods declined sharply with the increasing of the dimensions. To make the fine subsurface medium, we need to divide the medium into small pieces which let the parameters of the model to be huge. To inverse thousands of parameters can result in the existing CCs becoming invalid or making a large number of iterations.

In this paper, we propose a new DE to derive Q and velocity from zero-offset vertical seismic profiling (VSP) data. Compared with reflection seismic data, VSP data has high signal-to-noise ratio (SNR), high resolution, and can provide more information about medium layers and signal frequency. As the upgoing primary reflection of the VSP record is much stronger than the related multiples, our research will focus on the direct downgoing wave and upgoing primary reflection.

By considering the property of viscoelastic media and the examples of DE-CCM, we estimate the Q factor and velocity using the DE algorithm along successive medium layers from bottom to top. Another innovation is that a weighted mutation strategy and a stopping criteria are introduced to speed up the computation. We name the new DE as successive differential evolution (DE-S), where the validity of our proposed method is tested using both synthetic model and real record.

II. Our Proposed DE-S Method

A. The conventional DE method and its disadvantages

The Q factor and velocity of the viscoelastic medium that need to be estimated, span the related nonseparable model parameter space. Fig. 1(a) shows the relationship between the downgoing and the upgoing wave at a given interface and at the top and the bottom of a layer. $D_k^{'}(\omega)$ and $D_k(\omega)$ are the spectrum of the downgoing wave at the bottom and the top of layer k, respectively. Similarly, $U_k^{'}(\omega)$ and $U_k(\omega)$ correspond to the spectrum of the upgoing wave. Fig. 1(b) is the zero-offset VSP model diagram. The space between two successive geophones is defined as a layer. The parameters of each layer span the subspace and the parameters are defined as subcomponents.

VSP is a kind of seismic observation method. As shown in Fig. 1(b), the source of the VSP excites the seismic waves at some points at the surface, but receives the seismic signals by geophones which are vertically settled at various depths along the well. The received seismic waves include downgoing waves propagating from up to down and upgoing waves propagating from down to up. The record received by a geophone is called a trace of record. Thousands of traces construct a VSP data. For a trace of VSP data, the amplitude, phase and frequency of the direct downgoing wave changes because of the attenuation when seismic wave propagates from the source to the geophones in the viscoelastic medium.

Thus the direct downgoing wave is a function of the parameters of the upper layers over the geophone. The downgoing wave is reflected on each reflecting interface and then it arrives to the geophones to form the upgoing wave. Thus, the upgoing wave is a function of the parameters of all layers. Finally, the trace of record is a function of all the estimated parameters. Therefore, the model space spanned by the estimated parameters is nonseparable along both time and depth directions, so it is not possible to estimate the subcomponents (parameters) independently and simultaneously.

Storn [16] first proposed the DE algorithm for global optimization, which has three operations. First, it creates the mutant individual by adding a weighted difference vector between two individuals to a third individual. Then, it creates the trial individual by crossing the mutated individual and the original individual. Finally, it selects the next generation according to the fitness values of the parent individual and the trial individual. DE-CCM incorporates the decomposition strategy of cooperative coevolution into DE to decompose the large-size individual into subcomponents in the mutation step.

We can estimate the Q factor by using DE-CCM to show the disadvantages of the existing methods. Take the VSP direct downgoing wave as an example. We get the forward simulation from the method proposed by Ganley:

$$D_{k+1} = T_k e^{-\alpha \Delta z_k - i\omega \Delta z_k / v_k} D_k \tag{1}$$

where T_k is the transmission coefficients, ω is the frequency, v_k is the layered velocity, Δz_k is the layered distance, $\alpha = 1/2v_k Q_k$ is the attenuation factor, and D_k is the spectrum of the downgoing wave of the kth layer.

The 5-layer depth model is shown in Fig. 2(a), and Fig. 2(b) is the synthetic record based on the model. There are 100 geophones, and the record received by a geophone is called a trace of data. So, the number of horizontal ordinate 'Traces' equals to the number of geophones. The estimated Q factors span a 100-dimension model space. All subcomponents are updated simultaneously during each iteration. Fig. 3(a) and (b) show the curves of the estimated Q and velocity after 300 iterations, while Fig. 4(a) and (b) provide the curves after 1100 iterations. From Fig. 3(a-b), we can see that the estimated Q of the first two layers are accurate, and the parameters of other layers are not accurate though they are updated at each iteration. With the increasing time of iteration, Fig. 4(a-b) shows that the Q-factor of the first four layers are accurate, but the Q of the last deeper layer differ greatly from the true values. This means that the parameters of the deeper layers only can be estimated accurately if the parameters of the upper layers have been estimated accurately too. The convergence curve of the error energy is shown in Fig. 5.

This example indicates that the DE-CCM algorithm is not efficient when applied to nonseparable model spaces, and this inspires us to adopt a kind of layer-by-layer inversion method to improve the estimation precision and the computational efficiency. Thus a new global optimization method called DE-S is proposed.

B. Principle of DE-S

The DE algorithm seeks optimal solutions by simulating the processes in natural species evolution. Cognitive science can be utilized to mimic the evolution process. First, we provide an initial population, and then mutate some individuals of it. We use the algorithm to learn autonomously, to think independently and judge whether the mutation is the optimal direction of evolution. Finally, all the individuals become optimum by iteration. And based on the result, people can find the optimum parameters and make quicker and better decisions. The proposed DE-S algorithm starts with a population $\mathbf{X}$ which has M initial individuals:

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M] \tag{2}$$

The kth individual can be expressed as:

$$\mathbf{x}_k = [x_{k,1}, x_{k,2}, \mathrm{L}\ , x_{k,N}], k = 1, \mathrm{L}\ , M \tag{3}$$

where N is the number of the total traces as well as the dimension of each estimated parameter. There are two parameters, Q and velocity, to be estimated, so the dimension of the estimated parameters is $2N$. The local fitness function is used to control the mutation direction defined by the error energy between the observed and the calculated data. For the pure direct downgoing wave, the local fitness function of the kth individual is defined as:

$$f_{k,j} = \sum_{n=T-T_1}^{T+T_2} \left| \left[d_{o,j}(n) - d_{k,j}(n) \right] \cdot g(n) \right| \tag{4}$$

where $d_{o,j}(\cdot)$ is the observed downgoing and primary upgoing wave of the jth trace, $d_{k,j}(\cdot)$ is the forward simulated data in which the kth individual is used as parameters. T is the location of the envelope-peak of instantaneous amplitude (IA) of the direct downgoing wave, T_1 is the interval of the first trough above T and T_2 is the interval of the first trough below T. To make the maximum use of the waveform information, we introduce a Gaussian window to cut the downgoing and primary upgoing wave. g(n) is the Gaussian window function:

$$g(n) = \exp\left(-\frac{1}{2} \cdot \frac{(n-P)^2}{\sigma^2} \right) \tag{5}$$

where n is the variable and the standard deviation σ controls the width of the "bell". $P = (T_1 + T_M)/2$ is the middle point of the window. Local fitness value is an $M \times N$ matrix, $\mathbf{f} = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,N} \\ \cdots & & & \\ f_{M,1} & f_{M,2} & \cdots & f_{M,N} \end{bmatrix}$.

For the direct downgoing wave and upgoing primary reflection, the local fitness function of the kth individual is defined as:

$$f_{k,j} = \sum_{n=T_1}^{T_M} \left| \left[d_{o,j}(n) - d_{k,j}(n) \right] \cdot g(n) \right| \tag{6}$$

Fig.6 is the diagram of the peak locations. P_1 is the location of the first envelope-peak of the instantaneous amplitude (IA) of the observed data, P_M is the location of the last IA peak over the given threshold, T_1 is the location of the first trough above P_1, and T_M is the location of the first trough below P_M.

C. Cognitive Modelling based Improved DE-S

A good mutation strategy can modify the genes (namely, parameters) of the individuals effectively, and the optimal solution can be approached after several iterations. To clarify, here we divide the field stratum as thin layers according to geophones, so in this paper the interval between two neighboring geophones is defined as a layer. We assume there are N geophones to acquire data, so we have N layers. The signal of a geophone construct a record trace, hence there are N traces of data. Q and velocity of each layer are to be estimated, so there are $2N$ parameters estimated. There are two populations, one is used to estimate Q and another is for velocity. A population includes M individuals. M can be

any number, if *M* is too small, we do not always obtain the global optimal parameters; if *M* is too big, it takes too long to run. In the experiments we choose M as 100. An individual includes the whole estimated parameters. The genes of an individual mean the parameters *Q* or velocity. If there are *N* estimated parameters, the number of the genes is *N*. In the synthetic data examples *N* is 100, and in the real data examples, *N* is 553 because there are 553 traces of data. The *M* initial individuals of each population are mutated and updated, and finally we will choose an optimum individual as the result.

For the direct downgoing VSP record, a data trace is the function of all the parameters of the upper layers. When an individual cycles, we fix the parameters of other layers, and mutate that of the current layer. To search for the correct mutation direction, we sort all the variables of the *j*th row of the population **X** in an ascending order when inverting the parameters of the *j*th layer:

$$[x_{1j}^s \;\; x_{2j}^s \;\; \cdots \;\; x_{Nj}^s] = sort[x_{1j} \;\; x_{2j} \;\; \cdots \;\; x_{Nj}] \tag{7}$$

The corresponding local fitness values of the ordered variables are；

$$\overset{\vee}{f}_j^s = \left[f_j\left(x_{1j}^s\right) \;\; f_j\left(x_{2j}^s\right) \;\; \mathrm{L} \;\; f_j\left(x_{Nj}^s\right) \right] \tag{8}$$

Pick the minimal value of $\overset{\vee}{f}_j^s$ and the *NP* values on the two sides of the minimum, and the analytic expression of the ($2\times NP$+1) values is acquired by polynomial fitting using a concave function which has a minimum:

$$y = Ae^{-mx} + Be^{nx} \tag{9}$$

where, A, B, *m*, *n* are the four parameters of the compound function which are calculated using the inversion method. The location of the minimum of the concave function might be a better variable. The objective function of polynomial fitting is defined as:

$$E = \sum_{i=1}^{NP+1} \left| f_j^s(x_i^L) - y(x_i^L) \right|^2 \tag{10}$$

Where, $f_j^s(x_i^L)$ is the picked local fitness values, $y(x_i^L)$ is the polynomial values to be fitted. There are only four parameters to be estimated, the inverse of the Hessian matrix is easy to calculate. We hence use the Gaussian-Newton method to inverse.

The global optimum method needs lots of iterations which take a long time to calculate. Polynomial fitting is kind of inversion method which will slow down the calculation. To speed up the calculation, we take the following simple method to mutate the variable in practice.

Picking the variable $x_{r,j}$ corresponding to the minimal value of the sorted local fitness function, and choosing other two variables $x_{r1,j}$ and $x_{r2,j}$ randomly, the mutation strategy of the *j*th component of the *k*th individual is derived as:

$$x_{k,j}^{m} = x_{r,j} + F \cdot \left| x_{r1,j} - x_{r2,j} \right| \tag{11}$$

where F is the control factor whose value is between 0 and 1. The mutated gene disturbs nearby the optimum value.

We circulate all the individuals to update the jth column genes of the population, and then do several times of individual cyclic operation in updating the jth column genes to improve the inversion precision. During the inner cycle, we found the following phenomenon: the local fitness value is still large, but the difference between the genes is small. This is because all components of the jth column approach but do not reach the optimal point, and now the difference between components is small. This leads to small mutation. To increase the amplitude of mutation, we propose a weighted mutation method to amplify the mutation perturbation when the error energy is less than the given threshold:

$$x_{k,j}^{m} = x_{r,j} + F \cdot \left| x_{r1,j} - x_{r2,j} \right| \cdot p_1 \cdot \lambda \tag{12}$$

where λ is another control factor that is chosen according to the estimation precision. λ is fixed, but F is a random number at each iteration. Weighting coefficient p_1 is the normalized local fitness value:

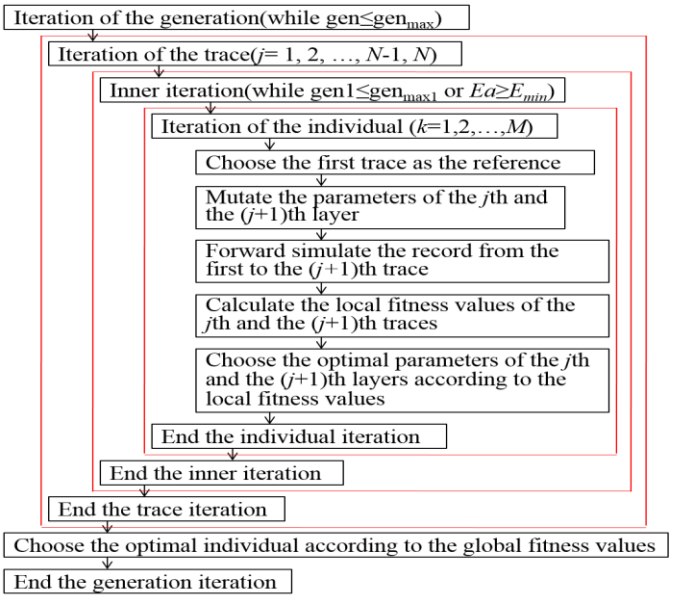
$$p_1 = \frac{1/f_{r,j}}{p_2}, p_2 = \frac{1}{f_{r,j}} + \frac{1}{f_{r1,j}} + \frac{1}{f_{r2,j}} \tag{13}$$

where $f_{r,j}, f_{r1,j}, f_{r2,j}$ are the local fitness values corresponding to $x_{r,j}, x_{r1,j}, x_{r2,j}$.

C1. The application in the pure direct downgoing wave

The table of the DE-S method for the pure direct downgoing wave is summarized as follow:

ALGORITHM 1 – THE UP-DOWN METHOD WORKFLOW



where gen_{max} and gen_{max1} define a given hard threshold. We set two kinds of iteration stopping criteria. One is hard threshold criteria, where if the iteration times reach the given threshold, then the program exits from the loop. The other one is soft threshold criteria:

$$Ea_j < E_{\min} \tag{14}$$

where $Ea_j = 1/M \sum_{k=1}^{M} f_{k,j}$ is the average error energy and $E_{\min}$ is the given threshold of the minimal precision. If the above inequality in Eq. 14 is satisfied, then the program is forced to exit from the iteration.

We can see from the flow that the parameters of the upper layers remain unchanged during the iteration of the *j*th trace, and then the parameters of the *j*th and (*j*+1)th layers are mutated and selected. The first (*j*+1) traces are forward simulated after mutation, and the local fitness values of the *j*th and the (*j*+1)th layers are updated. We update local fitness value of the (*j*+1)th layer to eliminate the effect of the updated parameters to the value.

The mutation strategy of the DE-S method differs from that of the DE-CCM method, given that all the subcomponents in DE-CCM method are mutated and selected simultaneously and independently under the guidance of the local fitness function, being the new individuals and the original ones crossed randomly. However, in the DE-S method, only the variables of the current layer are mutated and selected; then the subcomponents of the individual are updated layer-by-layer, and the mutation is implemented on single variable, so there is no need to cross the new individual and the original one.

C2. The application in the direct downgoing wave and upgoing primary reflection

For the VSP data including the direct downgoing wave and upgoing primary reflection, we adopt a layer-by-layer, bottom-up strategy. A data trace is the function of all the estimated variables: the downgoing wave contains the parameters information of the medium from the source to the geophone, while the upgoing wave contains the information of the deeper medium behind the geophone, thus the individual is nonseparable. If we window and remain the direct downgoing wave and inverse the parameters using the top-down strategy, the upgoing wave will not be used and the overlapping of the reflected wave at the interfaces will reduce the precision.

If the upgoing primary reflections are not reflected downward, there are no multiples. Based on this, we get the improved forward simulation from the method proposed by Ganley [17]:

$$\begin{cases} D_k = \dfrac{1}{T_k} e^{\alpha \Delta z_k + i\omega \Delta z_k / v_k} D_{k+1} \\ U_k = \dfrac{1}{T_k} R_j e^{-\alpha \Delta z_k - i\omega \Delta z_k / v_k} D_{k+1} + \dfrac{1}{T_k} (1\text{-}R_k^2) e^{-\alpha \Delta z_k - i\omega \Delta z_k / v_k} U_{k+1} \end{cases} \tag{15}$$

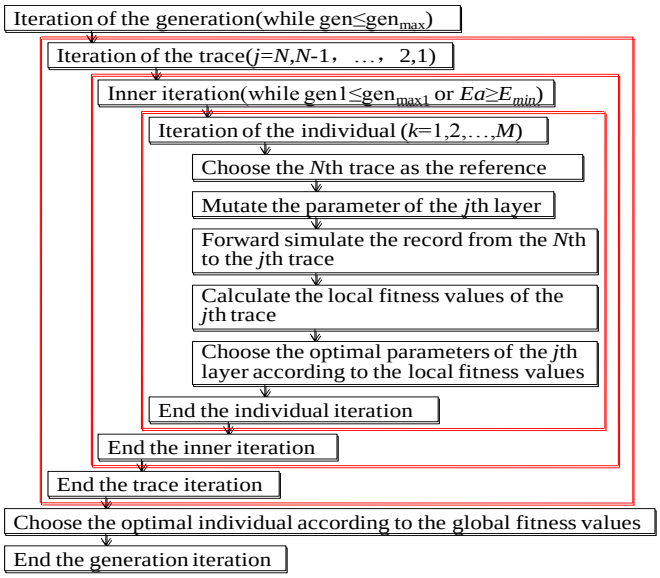
where $R_k = \frac{\rho_k v_k - \rho_{k+1} v_{k+1}}{\rho_k v_k + \rho_{k+1} v_{k+1}}, T_k = \frac{2\rho_k v_k}{\rho_k v_k + \rho_{k+1} v_{k+1}}$ are respectively the reflection coefficients and transmission coefficients. They both are the functions of the density ρ and the velocity v of the k^{th} layer and the $(k+1)^{th}$ layer. For ω, v_k and Δz_k, they denote respectively the frequency, the layered velocity, and the layered distance. $\alpha = 1/2 v_k Q_k$ is the attenuation factor; D and U are the spectrum of the downgoing and upgoing waves, respectively. Eq. 15 indicates that the seismogram of the kth layer is derived from the record of the $(k+1)$th layer through the transfer matrix. For a $(N+1)$ layers medium, $U_N = 0$. The matrix form of formula (15) is given as following:

$$\begin{bmatrix} D_k \\ U_k \end{bmatrix} = A_k \begin{bmatrix} D_{k+1} \\ U_{k+1} \end{bmatrix} \tag{16}$$

where $A_k = \frac{1}{T_k} \begin{bmatrix} e^{\alpha \Delta z_k + i\omega \Delta z_k / v_k} & 0 \\ R_j e^{-\alpha \Delta z_k - i\omega \Delta z_k / v_k} & (1-R_k^2) e^{-\alpha \Delta z_k - i\omega \Delta z_k / v_k} \end{bmatrix}$ is the transfer matrix. The upgoing wave of the kth layer is determined according to the downgoing and upgoing waves of the $(k+1)$th layer. As this forms a function with all the relevant parameters, the parameters span a non-separable model space. This inspires us to adopt a layer-by-layer, bottom-up inversion method to improve the estimation precision and the computational efficiency. In this way, we need to calculate and save all the downgoing waves, followed by calculating the upgoing waves layer by layer from bottom to top.

There are 4 iterations in the DE-S method implementation. The traces are iterated from the last to the first trace. For the jth trace, the whole M individuals are updated and we finally choose an optimal individual according to the fitness values. Algorithm 2 summarizes the DE-S method.

ALGORITHM 2 – THE DOWN-UP METHOD WORKFLOW



We can see from the flow chart that the parameters are updated from the Nth layer to the first layer. The parameters of the deeper layers remain unchanged during the iteration of the jth trace, and then the parameters of the jth layer are mutated and selected. The last (N-j+1) traces are forward simulated after mutation, and the local fitness value of the jth layer is updated. According to Eq. 15 and the workflow, we can see that the DE-S method is suitable for the zero-offset VSP record without multiples.

III. EXPERIMENTS AND RESULTS

A. Synthetic zero-offset VSP data including direct downgoing wave

We test the validity of the proposed method using a synthetic VSP record including the direct downgoing wave. Fig. 7 is the direct downgoing wave cut from Fig. 2(b) using a slide window. The source signature is a 40Hz Ricker wavelet located at the surface with zero-offset. We invert the velocity and Q using the top-down method. Fig. 8 is the estimated Q and velocity using DE-S method. The dash line is the theory values, and the solid line is the estimated values. We can see that the estimated velocity is accurate except at the interfaces, and the estimated Q have more error. This is because in the geology field, the velocities of real seismic waves are in the range of 1000-6500 m/s, and the Q factors are in the interval 5~200. The waveform is more sensitive to the variety of the velocity. The serious overlapping of the reflected waves at the interfaces affects the precision. The accumulative effect of error validate that the parameters are nonseparable, but the trend of estimated Q is correct. The difference profile between the observed and calculated record in Fig. 9 indicates that the error is at 10^{-7} order magnitude which is acceptable. Notice that the

bottom-up method can also be used for pure downgoing wave. In this case, we need to remove the upgoing wave calculation module when forward simulating.

In the DE-S method, not the whole forward record but the partial traces of the record are synthesized during each trace iteration. This is different to the DE-CCM method, where the whole synthetic record is received in each iteration. Take the direct downgoing wave as an example. Assuming the time taken for the forward of one layer is T seconds, then the whole time for the DE-CCM method is $T\times N\times M\times N$ after the trace iteration, where N is the number of layers, which equals to the number of traces, and M is the number of individuals. But for our DE-S method, the forward data are obtained layer-by-layer, and the running time is $(T+2T+3T+\cdots+N\times T)\times M=(1+N)\times N\times T\times M/2$, nearly half time of the DE-CCM method. Thus DE-S can double the computational efficiency when iteration times are the same.

The results in Fig. 8 and Fig.9 also indicate that the residual upgoing waves will affect the estimated values. We hence use the bottom-up strategy to estimate the parameters of the VSP data including direct downgoing wave and upgoing primary reflection.

B. Synthetic zero-offset VSP data including direct downgoing wave and upgoing primary reflection

The proposed method is tested using a synthetic zero-offset VSP record. The source signature is a 40Hz Ricker wavelet located at the surface with zero-offset. There are 100 geophones with 10m apart. The theoretical velocity and Q factor change every 100m as shown in Fig. 10(a). The snapshots of the synthetic record are shown in Fig. 10(b)-10(d).

There are 100 initial individuals in the population, and the values of variables of each individual are generated randomly. In order to optimize the individuals quickly and efficiently, the search space is defined by the lower and upper bounds. The initial Q factor is calculated using the WEPIF (wavelet's envelope peak instantaneous frequency) method [18-20] and the initial velocity is calculated using travel-time. The low frequency trends of the initial velocity and Q factor are used to define the upper and lower bounds by superimposing ±0.7km/s and ±50 on them respectively (as shown in Fig. 11). We choose these values according to the average value in the field and experience. According to the range of the velocity and Q, choosing ±0.7km/s and ±50 as the boundary of the search space is able to fulfill our requirements. DE-S uses only the low frequency trends of the initial values but not the initial values itself, hence DE-S is not dependent on the initial values.

The estimated results are shown in Fig. 12. The solid line represents the estimated values, and the dash line represents the theoretical values. We can see that the method works well in most areas where, in general, the errors are below 6%. The peak error of *Q* at several points is about 20% which is acceptable. The estimated velocity is accurate, but the estimated *Q* factor has small oscillations. This is because the velocity is several times larger than *Q*. As a result, the waveform is more sensitive to the variations in velocity than in the *Q* factor.

Overall, the error in Fig. 12 indicates that a generational cycle is good enough to obtain accurate parameters, where the times of inner iteration is 10, so there are 1000 inner iterations after 100 traces iterations.

B.1 Comparison with DE-CCM

In this subsection, we compare the performance of the proposed DE-S method with the DE-CCM method. Fig.13 (a) is the time of the iteration of each layer taken by the DE-S method. Fig.13 (b) is the total time after the iteration of the generation. We can see from the results that the running time of each layer is increased gradually. This is because the layers of forward modeling is increased successfully. The total time is a curve. Fig.14 (a) is the time of each iteration of the DECCM method which is in the range of 3.2~3.56. Fig.14 (b) is the total time after 1000 iterations which is a straight line. This is consistent with the fact that the DECCM method forward simulates the complete synthetic record. Fig.15 is the convergence curve of the two methods. The red line is obtained by DE-S method while the black line represents for the convergence speed of the DECCM method. The horizon axis represents for time. The figure indicates that the DE-S method runs two times as fast as the DECCM method.

B.2 Analysis of Anti-noise Capability

To tackle the noise within the real seismic records, it is necessary to analyze the anti-noise capability of the proposed DE-S method. We illustrate in Fig.16 noisy signals with 10 layers, where the Gaussian white noise is used with a SNR (signal noise ratio) of 20. Fig.17 (a) is the synthetic seismic record after filtering the noise, Fig.17 (b) is the residual error of the seismic record before and after noise filtering. As can be seen from Fig.17 (b), there are some residual downgoing signals at the near offset area, but it is almost all random noise in the residual section. Fig.18 (a) and (b) are the *Q* and velocity curves obtained by the proposed DE-S method. The inverted velocity from the noisy data is very consistent with the real values, where the error rate of less than 10% seems acceptable. We also show the inverted results with different SNRs in Fig. 19 and Fig. 20, where the SNR is 10 and 30 respectively. Fig.19 indicated that the estimated velocity and Q factor are quite different from the true values. Fig.20 shows that the estimated values are much more accurate when the SNR becomes larger. In addition, we show in Fig.21 the average error energy of *Q*

when the SNR varies from 2~50 with an interval of 2. As can be seen, the error becomes stable when the SNR is larger than 20. These experiment results have demonstrated that the robustness of the proposed DE-S method to Gaussian noise when the SNR is not too low.

C. Real zero-offset VSP data

The proposed method is also tested using real zero-offset VSP data shown in Fig. 22, in which the depth ranges from 110m to 6100m. The distance between two adjacent geophones is 10m in the depth of 110-5200m and 20m in the depth of 5200-6100m. There are 553 data traces. As mentioned above, we define the interval of two geophones as a layer, and the parameters of each layer are to be estimated. The estimated Q and velocity are 553 respectively. There are two generations. Each generation has 100 individuals and each individual has 553 genes. We use the low frequency trend of the initial Q and velocity, which is calculated by WEPIF method and the travel time, to define the boundary of the search ranges. Fig. 23 is the obtained Q factor and velocity. Green lines represent the estimated values, and black lines are the smoothed values using an interpolation method. The smoothed component can help us analyze the accuracy and the trend of the results.

As expected, the Q factor is lower in gas-bearing sandstone, which means that the absorption is more intensive in this area than in others. There are several strong absorption zone, but according to the interpretation by geology, the beneficial target areas estimated by our method is in the depth ranges of 1600-2000m, 3500-4100m, 4200~5000m, and 5200~6000m as marked by ovals. The results draws a distinction between the gas-bearing areas and the non-gas-bearing areas. This is one of the aid tools to predict gas-oil reservoir. The attenuation is effected by lots of factors such as temperature, pressure, strain amplitude, lithology, frequency, saturation, and fluid property. The velocity usually has the same trend with the Q-factor, but it is not absolute. In the target areas, Fig. 23(b) shows that the velocity has smaller values in the depth range of 4200~6000m which is corresponding to the small Q-factor. The purpose of this article is to estimate strong attenuation zone and to predict favorable reservoirs.

For comparative evaluations, Fig. 24(a) shows the estimated Q values by the automatic time-domain waveform inversion method (ATWI) [21], which is another popular method used for Q estimation. The DE-S method divided the layers according to the interval of the geophones which is 10m or 20m. And the initial parameters of each layer are different. The ATWI method inverses the parameters according to the waveform in time domain. The waveforms in an actual stratum are similar while the waveforms are obviously different in different stratums. The stratum thickness is from a few centimeters to tens of meters. Hence the estimated values are not as noisy as those from the

proposed method. Fig. 24(b) is the smoothed component obtained by a multi-scale analysis method which can decompose the signal to different levels, which is the 2-level decomposition. The higher the level is, the smoother the results are, which means the results lose more high frequency components and keep more low frequency. We take this method because we can choose the most suitable level to decompose. The lower areas of *Q* values are represented by A, B, C, and D. It is worth to mention that the inversion result shown here is consistent with our DE-S method, which helps to confirm the validity of the proposed method further.

IV. Conclusion

In this paper, we proposed a successive differential evolution (DE-S) algorithm that is suitable for the high-dimensional nonseparable model space. Being based on cognitive computing, a weighting factor was introduced into the mutation strategy to improve the inversion precision, where two kinds of-stopping criteria for flexible and effective iteration were proposed to improve the computational efficiency. We also provided an up-bottom and a bottom-up workflow for zero-offset VSP record. Unlike the local optimization waveform inversion method and other CCs, the new method is not dependent on the initial values and has good feasibility for nonseparable model space, fast computation and good convergence. Test results on both synthetic and real zero-offset VSP data indicate that this method has great potential to derive reliable seismic attenuation and velocity, even under certain degree of noise, which may have a large impact for reservoir characterization. Our approach will be further improved along with the development of high performance computing.

Acknowledgment

We thank National Natural Science Foundation of China (NSFC) (41604113, E070101) and National Nature Science Foundation Project of International Cooperation (41711530128) for their support and the joint project funded by NSFC and Royal Society of Edinburgh. We also thank Changqing oilfield for their field data.

References

[1] C. Zhou, R. Zhu, S. et al. "Types, Characteristics, Genesis and Prospect of Conventional and Unconventional Hydrocarbon Accumulations: Taking Tight Oil and Tight Gas in China as an Instance," *Acta Petrolei Sinica*, vol. 33, 2012.

[2] J. Virieux and S. Operto, "An overview of full-waveform inversion in exploration geophysics," *Geophysics*, vol. 74, pp. WCC1-WCC26, 2009.

[3] K. V. Price, "Differential evolution: a fast and simple numerical optimizer," in *Fuzzy Information Processing Society, 1996. NAFIPS., 1996 Biennial Conference of the North American*, 1996, pp. 524-527.

[4] A. Corana, M. Marchesi, et al, "Minimizing multimodal functions of continuous variables with the "simulated annealing" algorithm Corrigenda for this article is available here," *ACM Transactions on Mathematical Software (TOMS)*, vol. 13, pp. 262-280, 1987.

[5] F. Press, "Earth models obtained by Monte Carlo inversion," *Journal of Geophysical Research*, vol. 73, pp. 5223-5234, 1968.

[6] Z. Pan, J. Wu, Z. Gao, J. Gao，"Adaptive Differential Evolution by Adjusting Subcomponent Crossover Rate for High-Dimensional Waveform Inversion, " *IEEE Geoscience and Remote Sensing Letters*, 2015, pp:1327-1331.

[7] C. Wang and J. Gao, "A new differential evolution algorithm with cooperative coevolutionary selection operator for waveform inversion," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2010, pp. 688-690.
[8] C. Wang and J. Gao, "High-Dimensional Waveform Inversion With Cooperative Coevolutionary Differential Evolution Algorithm," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, pp. 297-301, 2012.
[9] R. Chandra, YS. Ong, CK. Goh, "Co-evolutionary multi-task learning with predictive recurrence for multi-step chaotic time series prediction," *Neurocomputing*, 2017, vol. 243, Pages: 21-34.
[10] Z. Gao, Z Pan, J. Gao, "A New Highly Efficient Differential Evolution Scheme and Its Application to Waveform Inversion," *IEEE Geoscience and Remote Sensing Society*, 2014, pp:1702 – 1706.
[11] Z. Gao, Z. Pan, J. Gao, "Multimutation Differential Evolution Algorithm and Its Application to Seismic Inversion," *IEEE Transactions on Geoscience and Remote Sensing*, 2016, vol. 54, Issue 6.
[12] XF. Cui, JH. Gao, B. Zhang and Z. Wang, "Poststack impedance inversion using improved particle swarm optimization," *SEG Technical Program Expanded Abstracts*, 2016, pp: 3809-3813.
[13] S. Mahdavi, M.E. Shiri, S. Rahnamayan, "Metaheuristics in large-scale global continues optimization: A survey," *Information Sciences*, 2015, vol. 295, pp:407-428.
[14] R. Govindan, R. Kumar, S. Basu, and A. Sarkar, "Altimeter-derived ocean wave period using genetic algorithm," *IEEE Geoscience and Remote Sensing Letters*, 2011, vol. 8, pp. 354-358.
[15] J.Gomes, P. Mariano and A. Lyhne, "Novelty-Driven Cooperative Coevolution," *Evolutionary Computation*, 2017, pp. 275-307.
[16] R. Storn and K. Price, "Minimizing the real functions of the ICEC'96 contest by differential evolution," *Evolutionary Computation, Proceedings of IEEE International Conference on*, 1996, pp. 842-844.
[17] D. Ganley, "A method for calculating synthetic seismograms which include the effects of absorption and dispersion," *Geophysics*, vol. 46, 1981, pp. 1100-1107.
[18] J. Gao, S. Yang, D. Wang, and R. Wu, "Estimation of quality factor Q from the instantaneous frequency at the envelope peak of a seismic signal," *Journal of Computational Acoustics*, vol. 19, 2011, pp. 155-179.
[19] G. J. Huai, Y. S. Lin, and W. D. Xing, "Quality factor extraction using instantaneous frequency at envelope peak of direct waves of VSP data," *Chinese Journal of Geophysics*, vol. 51, 2008, pp. 853-861.
[20] S. Yang and J. Gao, "Seismic attenuation estimation from instantaneous frequency," *Geoscience and Remote Sensing Letters, IEEE*, vol. 7, 2010, pp. 113-117.
[21] Gao J H, Wang C, Zhao W, "On the method of adaptive waveform inversion with zero-offset VSP data," Chinese J. Geophys.,52(12) , 2009, pp. 3091~3100.

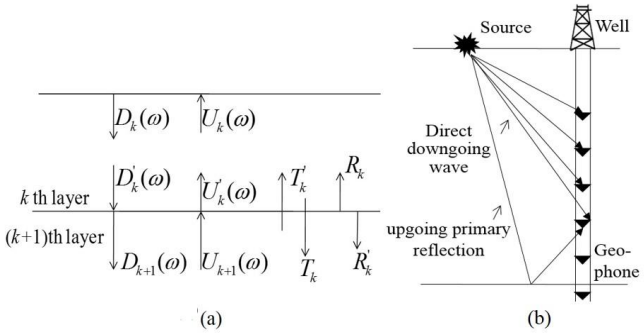


Fig.1. (a) The relationship between the downgoing and the upgoing wave. (b) The diagram of the zero-offset VSP record including the direct downgoing wave and upgoing primary reflection.

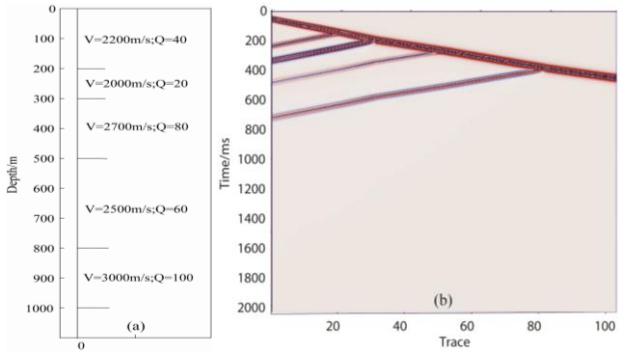


Fig.2. (a) A 5-layer depth model. (b) A synthetic record based on the model.

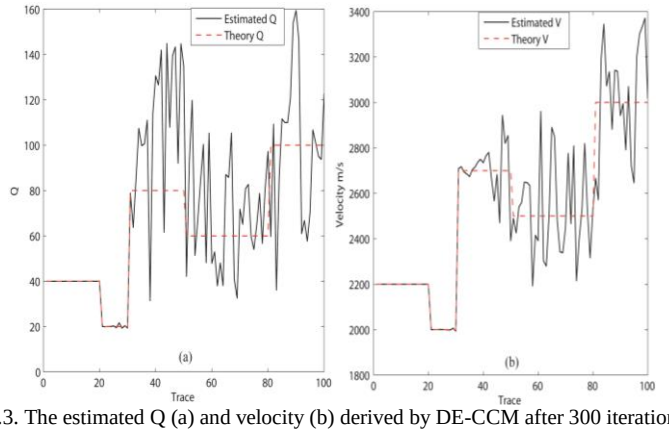


Fig.3. The estimated Q (a) and velocity (b) derived by DE-CCM after 300 iterations.

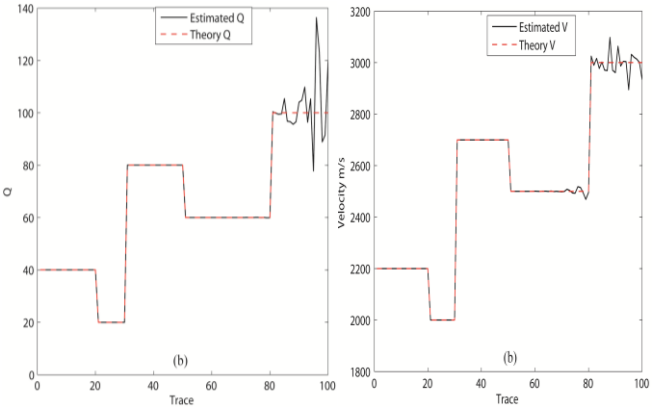


Fig.4. The estimated Q (a) and velocity (b) derived by DE-CCM after 1100 iterations.

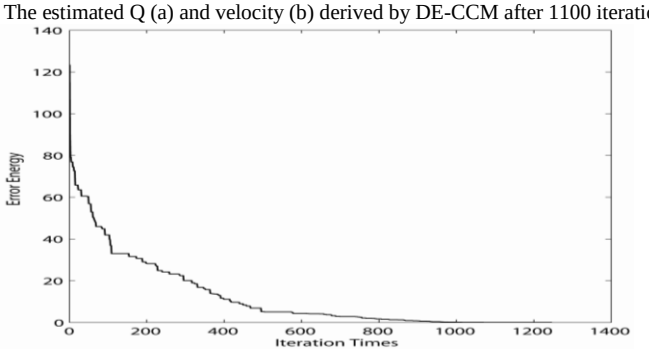
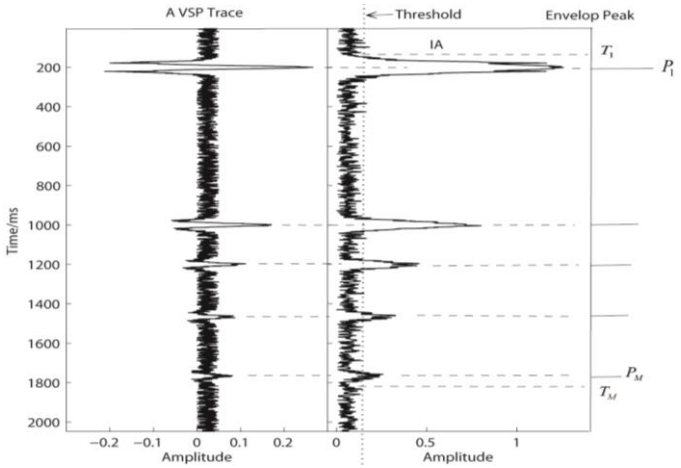


Fig.5. Best global fitness values (error energy) recorded in the evolution process. The horizontal axis is the number of iterations.



(a) (b)

Fig. 6. The diagram of the peak locations. (a) A trace of the VSP data. (b) The instantaneous amplitude (IA) of (a).

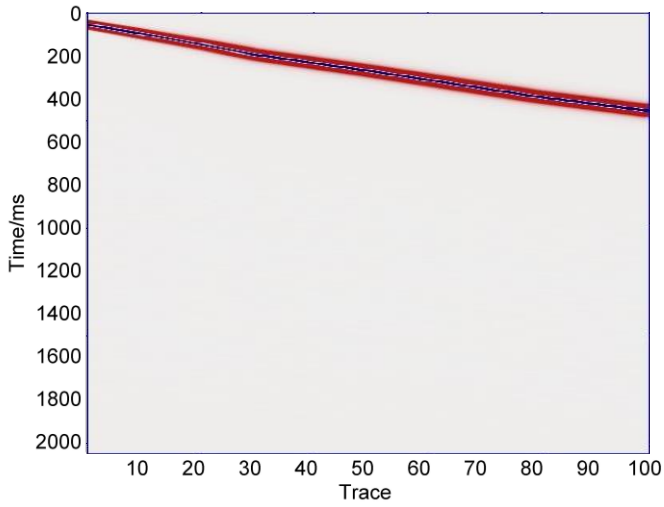


Fig. 7 The direct downgoing wave cut from Fig. 2(b).

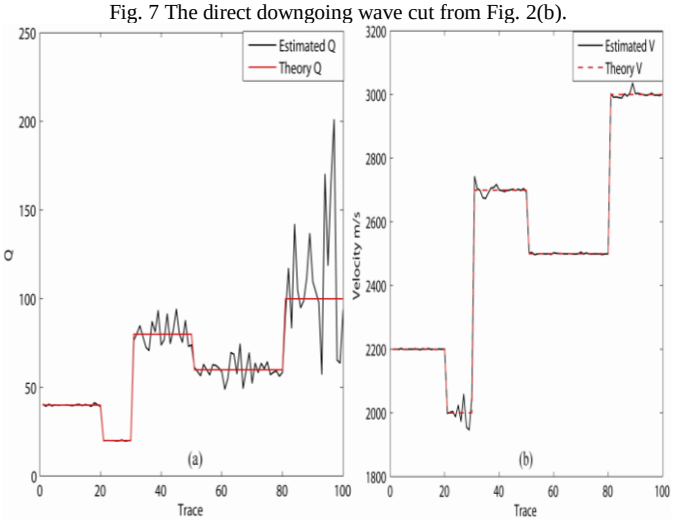


Fig. 8 The estimated (a)*Q* and (b)velocity using DE-S top-to-down method.

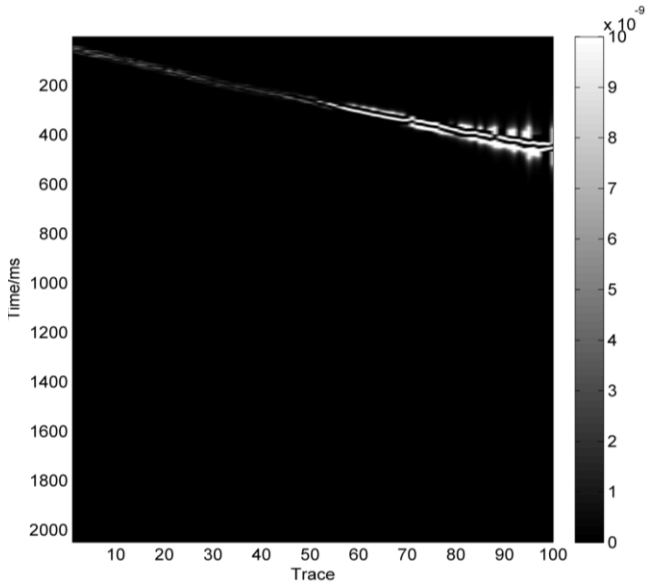


Fig. 9. The difference between observed and resulted record.

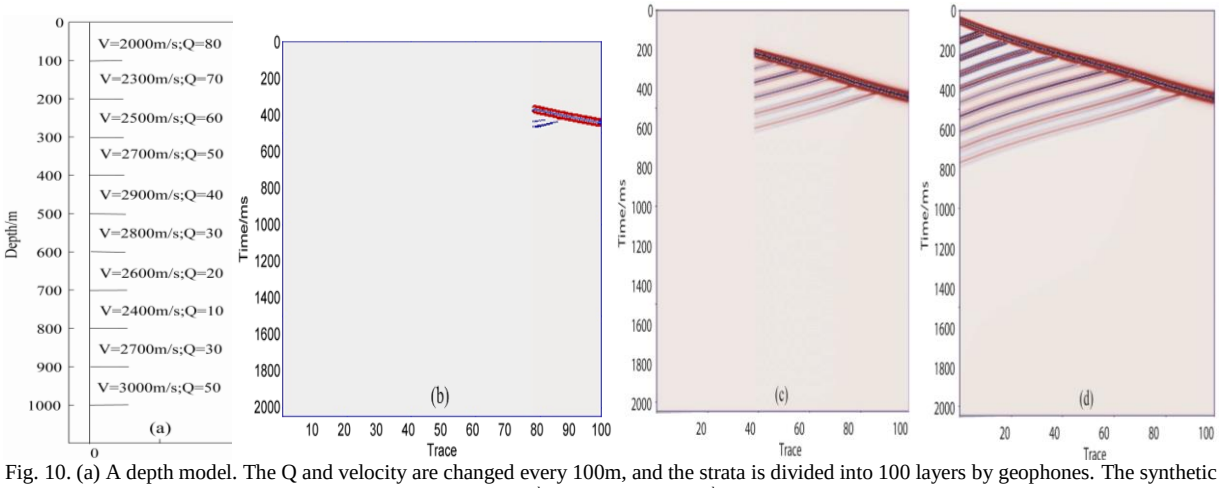
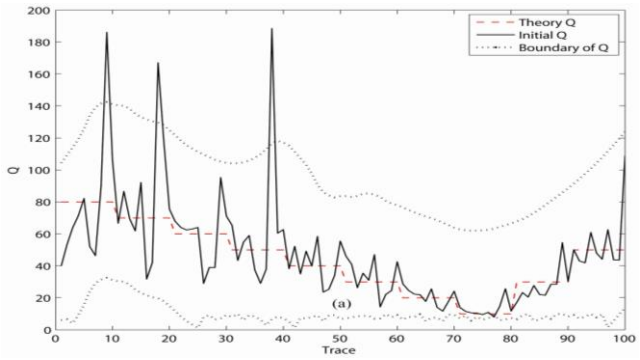


Fig. 10. (a) A depth model. The Q and velocity are changed every 100m, and the strata is divided into 100 layers by geophones. The synthetic records derived from the last to the 80th layer (b), to the 40th layer (c), and to the first layer (d).



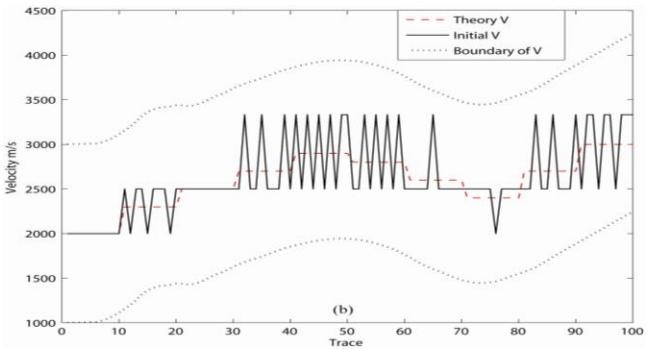


Fig. 11. Red dashed lines indicate theoretical Q-factors (a) and velocity (b). Black solid lines indicate estimated initial Q-factors (a) and velocity (b). Black dotted lines indicate the search space defined by the upper and lower bounds.

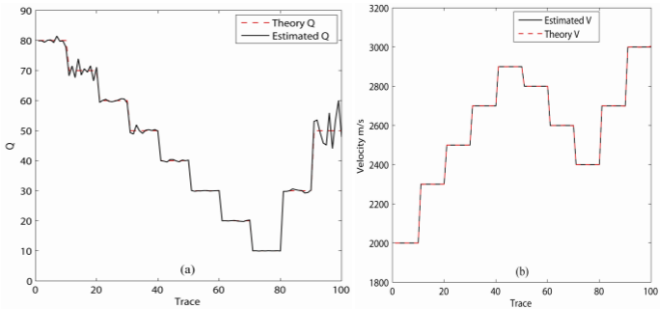


Fig. 12. The estimated (a) *Q* and (b) velocity by using DE-S bottom-to-up method.

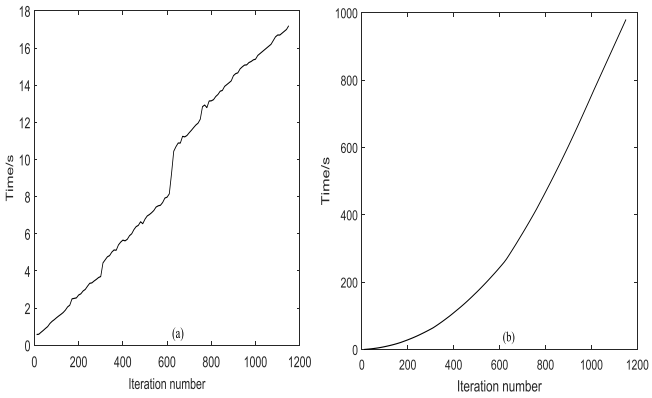


Fig.13 (a) The time of the iteration of each layer taken by the DE-S method. (b) The total time of the layer iteration taken by the DE-S method.

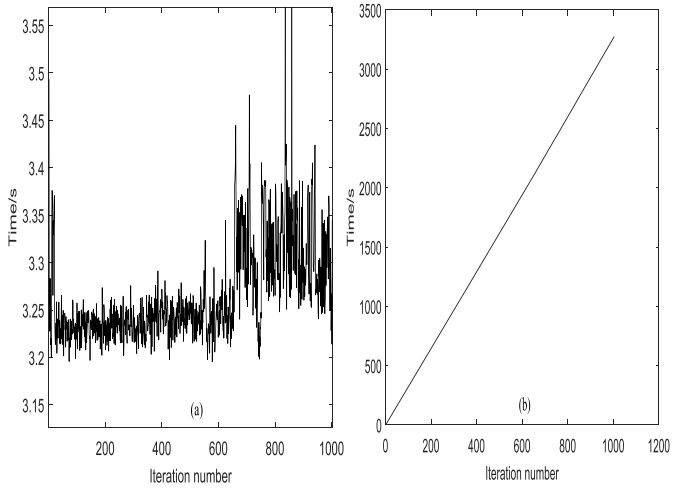


Fig.14 (a) The time of the iteration taken by the DECCM method. (b) The total time along with the iteration taken by the DECCM method.

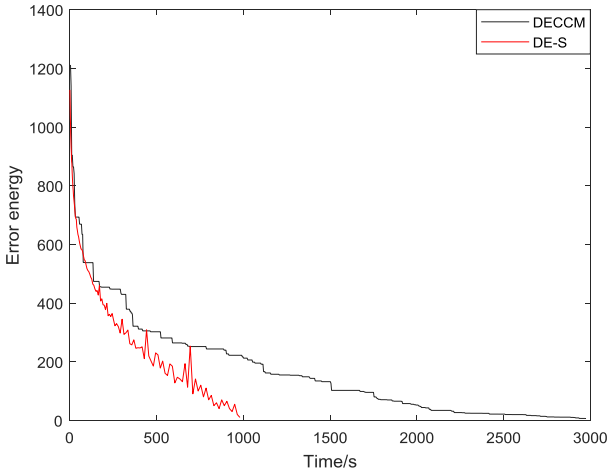


Fig.15 The convergence curve of the error energy. The horizontal axis is the running time. The red curve is obtained by DE-S method. The black curve is obtained by DECCM method.

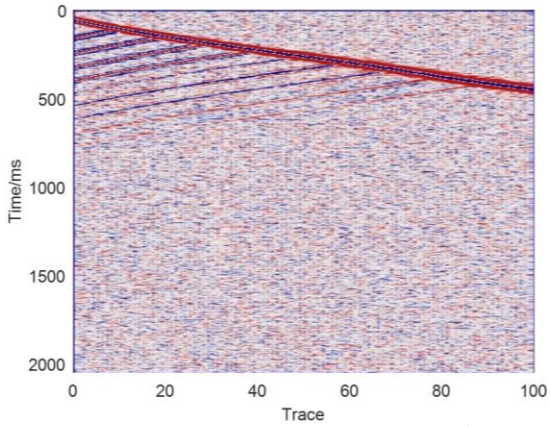


Fig.16: The synthetic noisy signal with a SNR of 20.

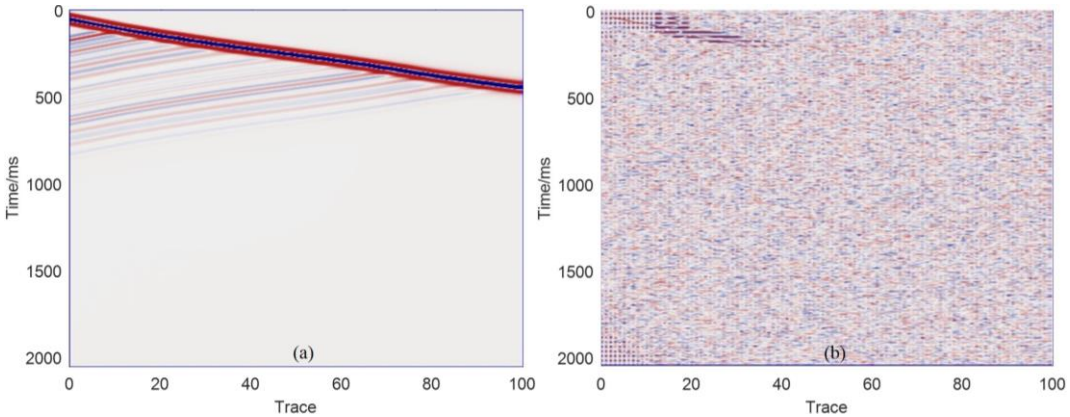


Fig.17 (a) The synthetic seismic record after filtering. (b) The residual error of the seismic record.

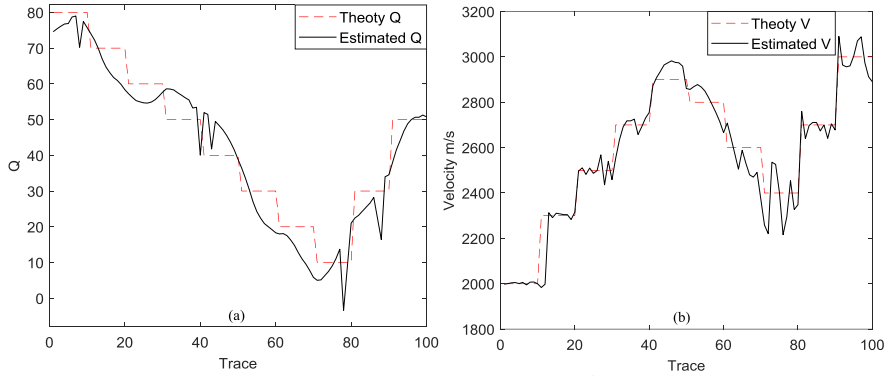


Fig.18. The estimated (a) *Q* and (b) velocity by our proposed DE-S method from the noisy record when the SNR is 20.

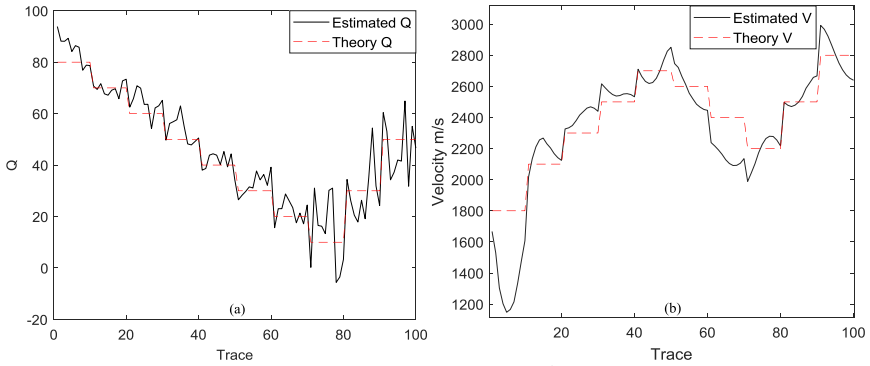


Fig.19. The estimated (a) *Q* and (b) velocity by using DE-S method from the noisy record when the SNR is 10.

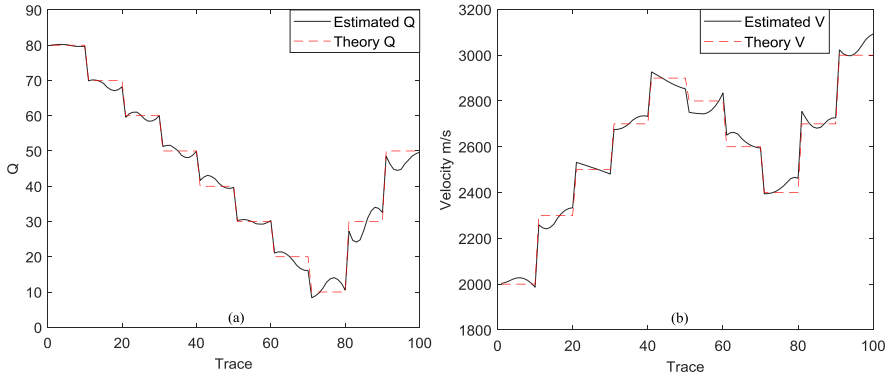


Fig.20. The estimated (a) *Q* and (b) velocity by using DE-S method from the noisy record when the SNR is 30.

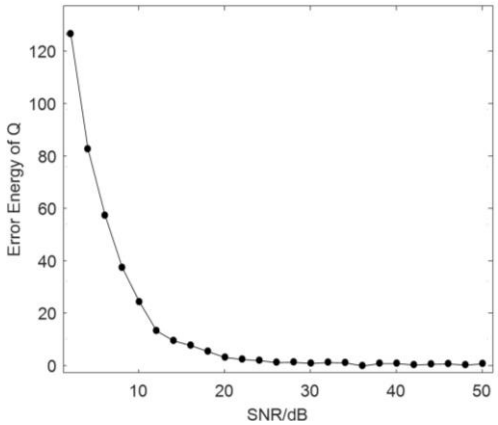


Fig.21. The average error energy of *Q* when the SNR is from 2~50 in the interval of 2.

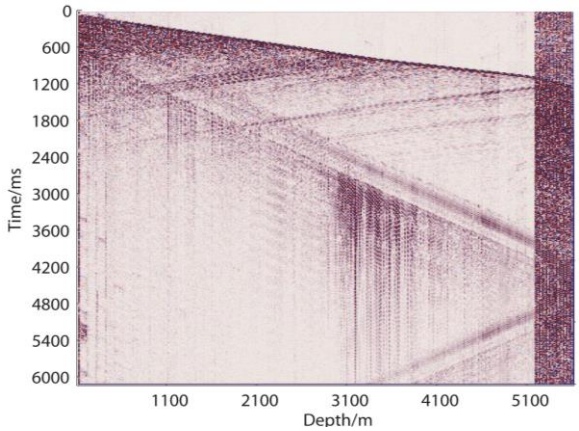


Fig.22. A real zero-offset VSP record.

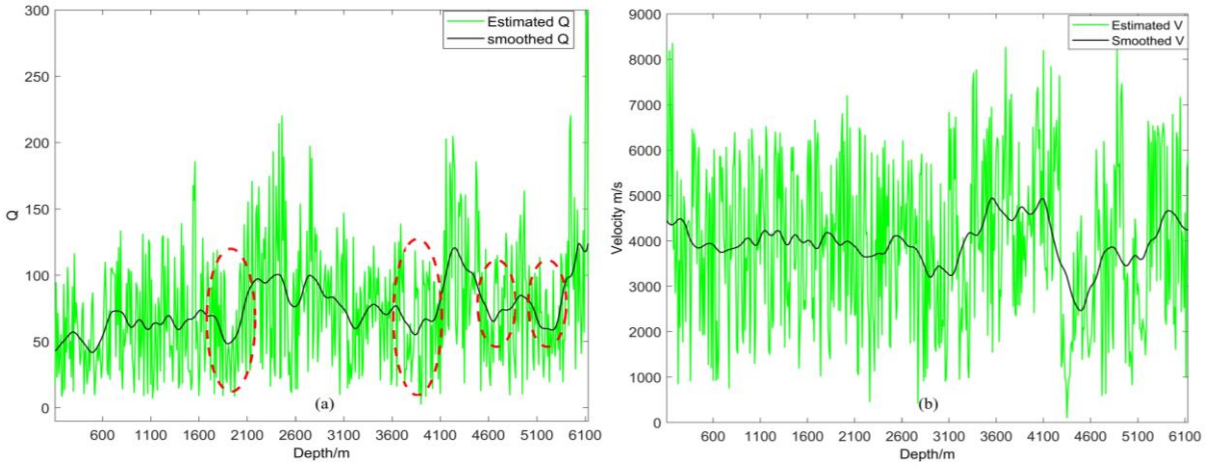


Fig. 23. The estimated (a) *Q* and (b) velocity by using our DE-S bottom-to-up method from real data.

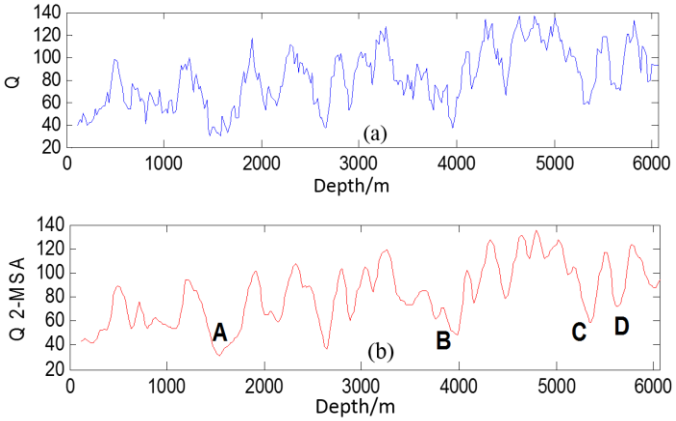


Fig.24. The estimated *Q* (a) and the smoothed *Q* (b) by using the ATWI method from real data.