Geomechanics and Engineering, *Vol. 15, No. 6 (2018) 1207-1217*
DOI: https://doi.org/10.12989/gae.2018.15.6.1207

Coupling relevance vector machine and response surface for geomechanical parameters identification

Hongbo Zhao*[1], Zhongliang Ru[1] and Shaojun Li[2]

[1]School of Civil Engineering, Henan Polytechnic University, Jiaozuo, Henan 454003, PR China
[2]State Key Laboratory of Geomechanics and Geotechnical Engineering, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan, Hubei 430071, PR China

(Received October 14, 2016, Revised April 2, 2018, Accepted April 25, 2018)

Abstract. Geomechanics parameters are critical to numerical simulation, stability analysis, design and construction of geotechnical engineering. Due to the limitations of laboratory and in situ experiments, back analysis is widely used in geomechancis and geotechnical engineering. In this study, a hybrid back analysis method, that coupling numerical simulation, response surface (RS) and relevance vector machine (RVM), was proposed and applied to identify geomechanics parameters from hydraulic fracturing. RVM was adapted to approximate complex functional relationships between geomechanics parameters and borehole pressure through coupling with response surface method and numerical method. Artificial bee colony (ABC) algorithm was used to search the geomechanics parameters as optimal method in back analysis. The proposed method was verified by a numerical example. Based on the geomechanics parameters identified by hybrid back analysis, the computed borehole pressure agreed closely with the monitored borehole pressure. It showed that RVM presented well the relationship between geomechanics parameters and borehole pressure, and the proposed method can characterized the geomechanics parameters reasonably. Further, the parameters of hybrid back analysis were analyzed and discussed. It showed that the hybrid back analysis is feasible, effective, robust and has a good global searching performance. The proposed method provides a significant way to identify geomechanics parameters from hydraulic fracturing.

Keywords: back analysis; response surface; hydraulic fracturing; geomechanics parameters; relevance vector machine

1. Introduction

Geomechanics parameters are critical to numerical simulation, stability analysis, design and construction of geotechnical engineering. But geomaterials are inherently anisotropic, inhomogeneous, and have discontinuities, and these variations in characteristics inevitably lead to complexity, nonlinear and uncertainty of geomechanics. Geomechanics parameters identification still remains one of the most challenging tasks. Laboratory experiment, in-situ test and back analysis are the three main ways to identify geomechanics parameters. Due to the limitations of laboratory experiments and the complexity of geomaterials, in-situ tests have been advocated to identify geomechanics parameters (Agarwal and Triggs 2004). However, it is difficult to conduct in-situ tests for most of geotechnical engineering problems, constrained by costs and time. So back analysis is widely employed to identify geomechanics parameters through coupling the numerical simulation, measure information of fields and geomechancis theory and various back analysis models were developed (Gioda and Maier 1980, Gioda and Jurina 1981, Sakurai and Takeuchi 1983, Deng and Lee 2001, Pichler *et al*. 2003, Feng *et al*. 2004, Gomes and Awruch 2004, Oreste 2005, Yu *et al*. 2007, Ghorbani and Sharifzadeh 2009, Vardakos *et al*. 2012, Zhang and Yin 2013, Ferrero *et al*. 2013, Zhou *et al*. 2016).

The basic procedure of back analysis is to adjust some parameters of the geomaterials so that the model predictions agree as closely as possible with the measurements of fields such as displacement, pressure, stress etc. Numerical simulation and optimal algorithm are the two key elements in back analysis. Due to the nonlinear, complexity relationship between geomechanics parameters and measure information of fields, numerical simulation is time consuming to the large scale project. To overcome this problem, the response surface method (RSM) is a good way to replace of numerical analysis in back analysis. By using RS, the complexity relationships between geomechanics parameters and measure information of fields can be presented by selecting a set of numerical simulations with the aid of the design of experiment technique (Khuri and Cornell 1996). The major benefits of RS are the significant reduction of the number required of numerical evaluations and improve the efficiency of back analysis. In the most commonly used polynomial-based RSM, the number of samples required increases in tandem with the order of polynomial used. This is time-consuming for practical engineering problems when a high-order polynomial is needed, because of the large number of input variables. Some researchers combined response surface with artificial neural networks (ANN) or support vector machines (SVM) overcome this problem (Deng and Lee 2001, Feng *et al*.

∗Corresponding author, Professor
E-mail: bxhbzhao@hotmail.com

ISSN: 2005-307X (Print), 2092-6219 (Online)

2004, Yu *et al*. 2007, Zhang and Yin 2013, Zhao and Yin 2016). ANN and SVM based response surface models have the advantage of providing high-order approximations with fewer samples than polynomial functions of comparable order (Gomes and Awruch 2004, Deng *et al*. 2005, Zhao *et al*. 2014, Zeinab *et al.* 2015); however, they have some inherent drawbacks, such as their slow convergence, a less generalized performance, arriving at a local minimum, over-fitting problems and selection of a suitable kernel function. Relevance vector machines (RVM) are somewhat similar to SVM, but appear to have more merit (Zhao *et al*. 2012). Tipping (2001) proposed an RVM that does not suffer from the above disadvantages. RVM is a general Bayesian learning framework for obtaining sparse solutions to regression and classification tasks utilizing models with linear parameters, and has been shown to perform a comparable generalization with a more sparse solution than SVM. State-of-the-art predictions have been reported in many applications where RVMs have been used (Chen *et al*. 2001, Agarwal and Triggs 2004, Farkhondeh *et al*. 2016).

Optimal method is another key element in back analysis (Oreste 2005). The optimal method searches trial values of the unknown geomechanics parameters until the discrepancy between measured value and predicted value associated with the trial geomechanics parameters is minimized. Many optimal method such as genetic algorithm and particle swarm optimization algorithm have been successfully used in back analysis (William 1981, Cividini 1988, Okabe 1998, Pichler *et al*. 2003, Rechea *et al*. 2008, Zhao and Yin 2009, Vardakos *et al*. 2012). The ABC algorithm has the advantage of better robustness and global optimization performance over the genetic algorithm (GA) and the particle swarm optimization (PSO) algorithm (Karaboga and Basturk 2008, Karaboga and Ozturk 2011). Artificial bee colony (ABC) algorithm, as an alternative optimization algorithm, is employed in this paper to search for the optimal geomechanics parameters.

Nowadays, hydraulic fracturing is the most common technique to perform well stimulation and in-situ stress characterization of hydrocarbon reservoirs, especially for unconventional reservoirs such as shale gas, tight gas and coal bed methane. Hydraulic fracturing tests are considered the most effective method for determination of the in situ stress and mechanical parameters of rock mass (Haimson and Fairhurst 1969, Haimson 1978, 1993, White *et al*. 2002, Fang and Khaksar 2011, Seyed *et al.* 2014, Zhu *et al.* 2014, Roman *et al.* 2015, Xu *et al*. 2016, Puller *et al*. 2016). But this method also needs to determine some poroelastic coefficients which it is difficult to determine in practice. In this paper, a hybrid back analysis was proposed to identify geomechanics parameters from hydraulic fracturing by coupling relevance vector machine (RVM), response surface and numerical simulation, and artificial bee colony (ABC) was adopted in the proposed method as optimal method.

Details will be presented in the rest of the paper and is structured as follows. Section 2 describes the classic breakdown formula of hydraulic fracturing in detail. The algorithm of RS, RVM and the basic principle of ABC are described briefly in Section 3. In Section 4, the procedure of the proposed method is presented in detail including basic conceptions of RVM and back analysis, generating samples, building RVM model and fitness function, etc. An example about hydraulic fracturing and discussion is introduced in Section 5. And finally, some conclusions are made in Section 6.

2. Hydraulic fracturing model

2.1 Breakdown formula of hydraulic fracturing

Hydraulic fracturing is a widely accepted technology for the determination of in-situ stress magnitudes and directions. One principal stress σ_v has a magnitude equal to the overburden pressure in vertical directions. The least horizontal principal stress $\sigma_{h\min}$ is usually determined directly in the experiment from the shut-in pressure. The greatest horizontal principle stress $\sigma_{H\max}$ must be calculated using a breakdown formula derived from an appropriate hydraulic fracturing model. Hubbert and Willis (1957) proposed a classic breakdown formula (Eq. (1)) to calculate $\sigma_{H\max}$ for nonporous impermeable rocks with hydraulic fracturing. However, the pore pressure term has been ignored.

$$P_b = 3\sigma_{hmin} - \sigma_{Hmax} + T \tag{1}$$

To consider the pore pressure, the breakdown formulas of porous impermeable rocks and porous permeable rocks were built (Eq. (2) and Eq. (3))

$$P_b = 3\sigma_{hmin} - \sigma_{Hmax} + T - P_p \tag{2}$$

$$P_b = \frac{3\sigma_{hmin} - \sigma_{Hmax} + T - \alpha\dfrac{1-2\upsilon}{1-\upsilon}P_p}{2 - \alpha\dfrac{1-2\upsilon}{1-\upsilon}} \tag{3}$$

Where P_b is the breakdown pressure, $\sigma_{H\max}$ and $\sigma_{h\min}$ are the greatest and least horizontal principal stress, T is the rock tensile strength and P_p is the pore pressure, α is the Biot poroelastic parameters and υ is the Posson's ratio. Although the Eq. (3) may best conform to the conditions under which hydraulic fracturing is conducted from open borehole, Eq. (3) is used in practice due to the difficulty of determining α and υ.

2.2 Numerical model of hydraulic fracturing

Hydraulic fracturing involves a strong coupling between fracture propagation, rock deformation, fluid flow, and even heat transfer at great depth. Numerical model is proposed to build the hydraulic fracturing model. For a single phase fluid flow in the formation, the constitutive relationship in incremental form is expressed (Lewis and Schrefler 1998) as

$$\mathrm{d}\boldsymbol{\sigma}' = \mathrm{D}^e\,\mathrm{d}\varepsilon^e + \mathbf{m}\alpha\mathrm{d}p \tag{4}$$

where $\mathrm{d}\sigma'$ is the effective stress increment; $\mathrm{d}\boldsymbol{\varepsilon}^e$ is the elastic strain increment; $\mathbf{m}=[1, 1, 0]^T$; α is Biot's coefficient; $\mathrm{d}p$ is the pore pressure increment. Moreover, $\mathbf{D}^e$ is elastic stress-strain matrix, which can be written as

$$\mathbf{D}^{e} = \frac{E}{(1+\nu)(1-2\nu)}\begin{bmatrix} 1-\nu & \nu & 0 \\ \nu & 1-\nu & 0 \\ 0 & 0 & \frac{1-2\nu}{2} \end{bmatrix} \quad (5)$$

where E is Young's modulus; ν is Poisson's ratio.

The flow rate in a single fracture of length, l, subject to a pressure difference of dp, is given by the following equation based on the cubic law of flow in fracture (Witherspoon *et al*. 1980, Zhang *et al*. 1999)

$$q = -\frac{a^3}{12\mu}\frac{\mathrm{d}p}{l} \quad (6)$$

where μ is dynamic viscosity, the contact hydraulic aperture, a is given by the following relationship

$$a = a_0 + u \quad (7)$$

where a_o is the fracture aperture at zero normal stress, u is the fracture normal displacement, which is related to rock properties and normal stress. The stress-displacement relation at the contact is assumed to be linear and governed by the normal stiffness k_n and the shear stiffness k_s as

$$k_n = \frac{d\sigma_n'}{du_n}, \quad k_s = \frac{d\sigma_s'}{du_s} \quad (8)$$

where σ'_n and σ'_s are effective normal stress and effective shear stress, respectively, u_n and u_s are normal displacement and shear displacement, respectively.

3. Methodologies

3.1 Response surface method

The response surface method (RSM) is an important technique that avoids lengthy computations in the analyses of complex physical systems. The basic response surface procedure is to approximate the response by an nth-order polynomial with undetermined coefficients, and thus to generate a polynomial equation using regression analysis and an approximately complex functional relationship between the dependent output y and the input variables

$$y = f(x_1, x_2, x_3, \ldots) + e\,. \quad (9)$$

Where x_i is the variables which have a significant influence on the response of geotechnical structure such as Young's modulus, in-situ stress, etc. y is the response of geotechnical structure such as displacement, stress, and pressure, etc.

The conventional RSM uses polynomial functions to fit the actual response of physics system based on sample points selected according to some experimental design. A second-order regression model without interaction terms containing the two input variables x_1, x_2 is given by

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_1^2 + a_4x_2^2 + e\,, \quad (10)$$

where a_0, a_1, …, a_4 are regression coefficients, and e is the error generated by neglecting other uncertainty sources.

3.2 Relevance vector machine (RVM)

RVM is simply a specialization of a sparse Bayesian model that utilizes the same data-dependent kernel basis (Tipping 2001). Supposing the system is identified as a multiple-input/single-output model with a sampled data set of N input vectors $\{X_n\}_{n=1}^N$ and N corresponding scalar-valued output $\{y_n\}_{n=1}^N$, and further assuming that the outputs are independent, identically distributed observations, then from the engineering viewpoint, some observations could be assumed to contain mean-zero Gaussian noise with variance $\sigma^2 : p(\varepsilon^n \mid \sigma^2) = N(0,\sigma^2)$. Then

$$y_n = f(X_n;W) + \varepsilon_n \quad (11)$$

That is

$$p(y_n \mid x_n, W, \sigma^2) = N(f(X_n;W), \sigma^2) \quad (12)$$

where $W = [\omega_1, \omega_2, \ldots, \omega_N]^T$ is the weight vector. After the Bayesian learning process, the regression estimate $\hat{y}$ at a value X is given by

$$\hat{y} = f(X;W) = \sum_{i=0}^{N}\omega_i K(X, X_i) = \varphi W \quad (13)$$

where $K(X, X_i)$ is a kernel function. In this study, we make the common choice of utilizing the general polynomial kernel function and the radial basis function kernel (RBF) function. φ is an $N \times (N+1)$ design matrix, with $\phi(X_N) = [1, K(X_N, X_1), K(X_N, X_2), \ldots, K(X_N, X_N)]^T$, in which $\phi = [\phi(X_1), \phi(X_2), \ldots, \phi(X_N)]^T$.

The classic approach to estimating $\hat{y}$ is to maximize likelihood or to minimize the least-squares of the measured training dataset to estimate W and σ^2; however, this may lead to overfitting.

$$p(y|W,\sigma^2) = (2\pi\sigma^2)^{-N/2}\exp\left(-\frac{1}{2\sigma^2}\|y-\phi W\|^2\right) \quad (14)$$

To control the complexity of the model and to avoid overfitting, a zero-mean Gaussian prior probability distribution is defined over every ω_i with variance σ_i^{-1}, and the likelihood of W is written as

$$p(W|\alpha) = \prod_{i=0}^{N} N(\omega_i|0, \alpha_i^{-1}) \quad (15)$$

where the hyperparameter vector $\alpha = [\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_N]^T$ controls the amount each weight is allowed to deviate from zero. For completion of the hierarchical prior, hyperpriors over α:$p(\alpha)$ and noise variance σ^2:$p(\sigma^2)$ are specified as gamma distributions.

Consequently, using Bayesian posterior inference, the posterior over W is given by

$$p(W|y,\alpha,\sigma^2) = \frac{p(y|W,\sigma^2)p(W|\alpha)}{p(y|\alpha,\sigma^2)} \quad (16)$$

where $p(y|\alpha,\sigma^2)$ is the normalizing factor, and $p(y|W,\boldsymbol{\sigma}^2)$ and

$p(W|\boldsymbol{\alpha})$ are both Gaussian priors. Thus the posterior is also conveniently Gaussian: $p(W|y,\alpha,\sigma^2) \sim N(\mu,\Sigma)$. Here the posterior mean μ and covariance Σ are given by

$$\begin{aligned} \mu &= \sigma^{-2}\sum\varphi^T y \\ \sum &= (\sigma^{-2}\varphi^T\varphi + A)^{-1} \end{aligned} \tag{17}$$

where $A = diag(\alpha_0, \alpha_1, \alpha_2, ..., \alpha_N)$. Note that W in Eq. (13) can be set to the fixed value of μ for the purpose of point prediction.

For unseen data, X^*, predictions are made for the corresponding output y^* in terms of the predictive distribution

$$\begin{aligned} p(y^*|y) &= \int p(y^*|W,\sigma^2)p(W,\alpha,\sigma^2|y)dWd\alpha d\sigma^2 \\ &= \int p(y^*|W,\sigma^2)p(W,\alpha,\sigma^2|y)p(\alpha,\sigma^2|y)dWd\alpha d\sigma^2 \\ &\approx \int p(y^*|W,\sigma^2)p(W,\alpha,\sigma^2|y)\delta(\alpha_{MP},\delta^2_{MP}|y)dWd\alpha d\sigma^2 \\ &\approx \int p(y^*|W,\sigma^2_{MP})p(W|y,\alpha_{MP},\delta^2_{MP})dW \end{aligned} \tag{18}$$

Good approximation of Eq. (18) needs to search hyperparameters posterior mode - that is, maximization

$$p(\alpha,\sigma^2|y) \propto p(y|\alpha,\sigma^2) = \int p(y|W,\sigma^2)p(W|\alpha)dW \cdot \tag{19}$$

The values of $\alpha_{MP}, \sigma^2_{MP}$ in Eq. (18) is learned using a *type-II maximum likelihood* method, and the iterative re-estimation is formulated

$$\begin{aligned} \alpha_i^{new} &= \frac{1-\alpha_i\sum_{ii}}{\mu_i^2} \\ (\sigma^2)^{new} &= \frac{\|y-\varphi\mu\|^2}{N-\sum_{i=1}(1-\alpha_i\sum_{ii})} \end{aligned} \tag{20}$$

Now, the computed result of Eq. (18) is $p(y^*|y) \sim N(\mu_*, \delta^2_*)$, where

$$\begin{aligned} \mu_* &= y(X^*;\mu) \\ \sigma^2_* &= \sigma^2_{MP} + \varphi(X^*)^T\sum\varphi(X^*) \end{aligned} \tag{21}$$

The mean μ_* is the predictor of the model output with unseen data X^* and posterior mean weight μ. The predictive variance σ^2_* is the sum of the variances associated both with noise processing and the uncertainty of the weight estimates. In this optimization process, the vector from the training set associated to the remaining nonzero weights is called the relevance vector (RV).

3.3 Artificial bee colony algorithms

The artificial bee colony (ABC) algorithm was originally developed by Karabogain (2005). In the ABC algorithm, the colony of artificial bees consists of three groups of bees: employed bees, onlookers and scouts. The position of a food source represents a possible solution for the problem under consideration and the nectar amount of a food source represents the quality of the solution represented by the fitness value (Karaboga and Basturk 2008, Karaboga and Ozturk 2011). ABC algorithm requires cycle of four phases: initialization phase, employed bees phase, onlooker bees phase and scout bee phase.

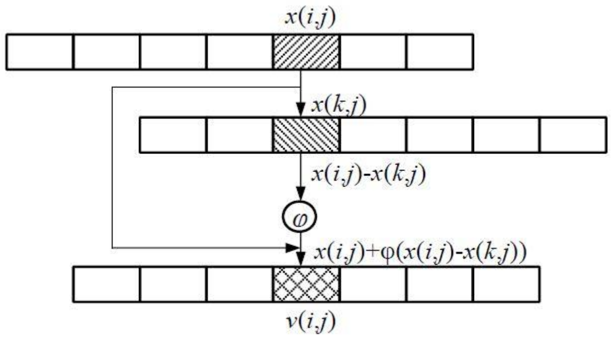

Fig. 1 Illustrating a simple position update equation execution

3.3.1 Initialization phase

In the algorithm, the number of the employed bees or the onlooker bees is equal to the number of solutions in the population. At the first step, the ABC generates a randomly distributed initial population of *SN* solutions and calculates the fitness of each solution.

$$x(i,j) = x^j_{\min} + rand(0,1)(x^j_{\max} - x^j_{\min}) \tag{22}$$

where $x(i,j)$ is the candidate solution of problem; i=1, 2,..., $SN/2$ and $SN/2$ denotes the size of population; j=1,2,…,D and D is the dimension number of each solution; $rand(0,1)$ is a random number between [0, 1]; $x^i_{\min}$ and $x^i_{\max}$ are the upper and lower bound of each solution.

3.3.2 Employed bees phase

Once initialization is completed, employed bees search for specific food sources (solution) and calculate the amount of nectars (fitness value). A candidate food position can be produced by the memory of bees (seen in Fig. 1), which is defined as

$$v(i,j) = x(i,j) + \varphi_{ij}(x(i,j) - x(k,j)) \tag{23}$$

where k used to be different from i is randomly chosen indexes from {1, 2,...,$SN/2$}, j is also randomly chosen indexes from {1, 2,...,D}, φ_{ij} is a random number in [-1,1] and controls the generation of neighbor food sources around $x(i,j)$ and represents the comparison of two food positions seen by a bee.

3.3.3 Onlooker bees phase

Onlooker bees choose a food source based on the nectars (probability of food source) which shared by employed bees and determine the source to be abandoned and allocate its employed bee as scout bees. The probability of being selected for each fitness value, p_i can be expressed as

$$p_i = \frac{fitness_i}{\sum_{n=1}^{SN} fitness_n} \tag{24}$$

where $fitness_i$ is the fitness value of the solution.

3.3.4 Scout bees phase

Scout bees randomly search for a new food source. In ABC algorithm, a food source that its position cannot be improved further through a predetermined number of cycles is assumed to be abandoned by onlookers. $x(i,j)$ used to represent the abandoned source is replaced by $x^{'}(i,j)$ that is a new food source the scout bees find, which is conducted by Eq. (22). Each candidate source position $v(i,j)$ produced by $x(i,j)$ can be evaluated using the comparison between $x(i,j)$ and its old source position. The old food source will be replaced by the new food source when it is equal or better than the old food source. Otherwise, the old food source is retained in the memory.

4. Hybrid back analysis through coupling RS and RVM

A hybrid back analysis approach is proposed to identify the geomechanics parameters from hydraulic fracturing through coupling RS and RVM in this section. The RVM is used to build the response surface which approximates the nonlinear relationship between the geomechanics parameters and the borehole pressure. The ABC algorithm is adopted to search for the geomechanics parameters based on the fitness value that is established by comparing the monitored pressure and the RVM-predicted pressure.

4.1 Building response surface

Response surface is adopted to approximate the nonlinear relationship between the pressure and geomechanics parameters. The RVM has more powerful regression capabilities than polynomial-based response surfaces. It is able to reflect nonlinear relationship between the pressure and geomechanics parameters. Response surface can be build using Relevance vector machine RVM($\boldsymbol{X}$) as

$$RVM(\boldsymbol{X}):\ R^{N} \rightarrow R \tag{25}$$

$$\boldsymbol{y}=RVM(\boldsymbol{X}) \tag{26}$$

where $\boldsymbol{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N)$ is vector of geomechanics parameters, for example, in-situ stress, Young's modulus and Poisson's ratio. $\boldsymbol{y}$ is the borehole pressure.

In order to obtain RVM($\boldsymbol{X}$), a training process based on the known data set is required. To train the RVM, it is needed to create the necessary training samples and to determine the training parameters of RVM. The former is performed by using numerical simulation of hydraulic fracturing for the given set of tentative determined parameters to obtain the corresponding borehole pressure. The other is to determine the parameter of the kernel function in RVM algorithm.

4.2 Objective function

To use the ABC algorithm to identify geomechanics parameters, as in any conventional approach to the back analysis, it is necessary to build the objective function for ABC, which is defined as follows

$$fitness = \sqrt{\left(RVM(X) - y\right)^2} \tag{27}$$

where $RVM(X)$ is the predicted pressure using RVM model, $\boldsymbol{y}$ is the monitored pressure.

4.3 Procedure of hybrid back analysis coupling RVM and RS

If a RVM model that represents the non-linear relationship between the borehole pressures and geomechanics parameters, the response surface of RVM can be used to predict the borehole pressures. Then, the ABC algorithm is used to search for the optimal parameters through the error minimization between the predicted pressures by the RVM model and the monitored pressures. The flowchart of hybrid back analysis is shown in Fig. 2. This back analysis algorithm can be described as follows:

- Step 1: Determine the general information and data such as the unknown (need to determine by back analysis) and known parameters of numerical model, experiment design method, parameters of RVM and ABC algorithm and the range of parameters to be determined.
- Step 2: Generate the set of tentative parameters based on the range of parameters to be identify using experiment design method.
- Step 3: Based on the above tentative samples, calculate the borehole pressure with each set using numerical model, and then build the training samples for RVM.
- Step 4: Based on the training samples of above step, built the RVM model by RVM algorithm in section 3.2.
- Step 5: Build RVM based response surface through coupling RS and RVM which build in the above step.
- Step 6: Active the ABC algorithm and search the geomechanics parameters based on the monitored pressure by combining with back analysis method.

5. Numerical example

In this paper, a numerical example is adopted to verify the above proposed method. UDEC is used to model Hydraulic fracturing modeling. Model size is a 24.0 m×24.0 m, and located 1000 m below the ground surface, with a wellbore of 0.2 m in diameter. Fig. 3 showed the numerical model of hydraulic fracturing. To simulate the hydraulic fracturing propagating from wellbore, an existing fracture in the rock is assumed to be parallel to the maximum horizontal in situ stress passing through the center of the borehole. The in-situ stresses of fracturing formation are composed of a vertical stress σ_v, and the maximum and minimum horizontal in situ stresses $\sigma_{H\max}$ and σ_{hmin}. Here σ_y is perpendicular to the two-dimensional model in vertical borehole so that it is ignored in the plane strain modeling. The vertical in situ stress is set equal in this work to 25.0 MPa, which is calculated from weight of the overburden.

The goal of this study is to determine the horizontal in situ stress and geomechanics parameters of rock based on the borehole pressures from hydraulic fracturing tests. Inputs of training samples are the maximum and minimum horizontal in-situ stresses, Young's modulus, and Poisson's

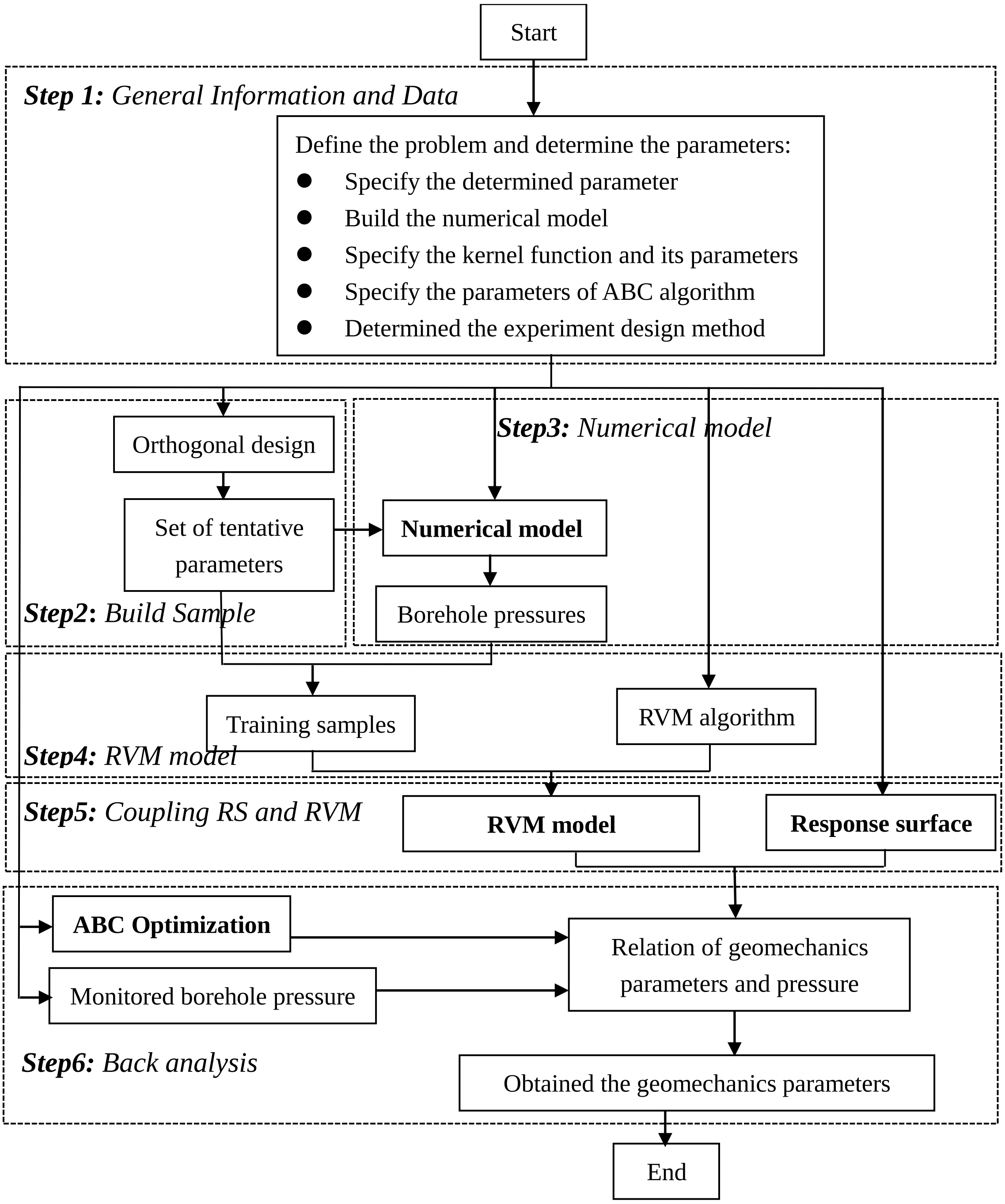

Fig. 2 The flowchart of the proposed method

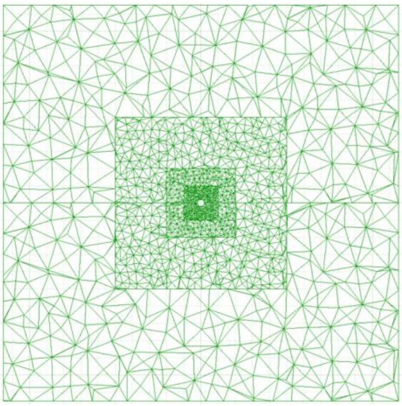

Fig. 3 The numerical model of hydraulic fracturing

ratio. Outputs are specified for formation breakdown pressure P_1 (FBP), fracture propagation pressure P_2 (FPP), and instantaneous shut-in pressure P_3 (ISIP). Fluid injection at a constant flow rate of 0.004 m^3/s is specified for the wellbore, and the injection point is located at the center of the model. The fracture tensile strength σ_t=8.0 MPa is used to enforce the zero toughness condition. The compressible flow algorithm is selected and the fluid bulk modulus K_f =100.0 MPa is specified to simulate an "incompressible" fluid. Horizontal in-situ stresses ($\sigma_{H\max}$, $\sigma_{h\min}$), Young's modulus E and Poisson's ration v are used to unknown, and others parameters are determined. Based on experiment design theory, 50 training samples are built for RVM (Fig. 4).

	A	B	C	D	E	F	G	H	I	J	K
1											
2											
3		Relevance Vector Machine									
4											
5		**Training samples**			**Parameter of RVM**						
6		Number of input	Number of samples	Number of output	σ						
7		4	50	3	2.5						
8											
9		**Training samples**							**Models of RVM**		
10		Input(xi)				Output(yi)			υ_i		
11		E	υ	σ_{Hmax}	σ_{hmin}	p1	p2	p3	7.932	37.6	6.982214
12	1	20.00	0.14	5.00	2.00	9.33	9.5	3.2	4E-05	-19.7	-5.52013
13	2	20.00	0.18	12.50	6.00	9.39	14.73	7.07	-0.02	-7.8	-0.56592
14	3	20.00	0.22	20.00	10.00	11.39	21.96	11.07	0.878	-12.1	0.985863
15	4	20.00	0.26	27.50	14.00	16.49	35.89	14.95	9.831	8.197	1.300181
16	5	20.00	0.30	35.00	18.00	9.98	39.97	19.62	-3.61	11.48	10.08788
17	6	30.00	0.14	12.50	10.00	14.68	27.42	11.31	4.479	-2.82	-0.03154
18	7	30.00	0.18	20.00	14.00	8.427	33.32	15.14	-16.8	-10.1	-0.34879
19	8	30.00	0.22	27.50	18.00	13.57	44.89	20.01	-1.57	9.069	5.614032
20	9	30.00	0.26	35.00	2.00	2.12	3.22	2.88	-9.27	-22.7	-5.89746
21	10	30.00	0.30	5.00	6.00	11.3	14.98	7.385	-1.32	-17.9	-0.70698
22	11	40.00	0.14	20.00	18.00	16.25	52.42	19.65	0.134	19.61	3.880626
23	12	40.00	0.18	27.50	2.00	2.497	3.656	2.931	-9.98	-16.4	-3.91352
24	13	40.00	0.22	35.00	6.00	9.158	14.15	7.635	4.108	-0.3	-0.02297
25	14	40.00	0.26	5.00	10.00	7.364	29.24	11.43	-11.1	-2.04	1.532986
26	15	40.00	0.30	12.50	14.00	13.12	38.06	15.55	-1.99	-0.76	0.530199
27	16	50.00	0.14	27.50	6.00	9.774	24.58	7.789	5.034	28.31	0.901239
28	17	50.00	0.18	35.00	10.00	8.72	14.04	11.89	-3.73	-16.3	1.768801
29	18	50.00	0.22	5.00	14.00	15.44	47.95	15.39	8.456	19.93	1.450332
30	19	50.00	0.26	12.50	18.00	8.66	54.8	21.14	-18.4	11.06	6.919286
31	20	50.00	0.30	20.00	2.00	2.398	7.389	3.167	-11.6	-16.5	-3.57681
32	21	60.00	0.14	35.00	14.00	9.534	28.2	15.74	-3.74	5.902	6.539832
33	22	60.00	0.18	5.00	18.00	11.48	58.68	20.07	-8.73	27.92	9.661363
34	23	60.00	0.22	12.50	2.00	8.124	15.46	3.294	-5.17	-12	-5.79543
35	24	60.00	0.26	20.00	6.00	11.69	29.32	7.395	2.618	21.29	-0.53009
36	25	60.00	0.30	27.50	10.00	10.03	14.85	11.59	-4.46	-24.1	0.440706
37	26	20.00	0.14	5.00	14.00	16.54	45.87	15	0.166	9.456	3.032878
38	27	20.00	0.18	12.50	18.00	20.05	57.75	19.88	8.396	26.07	8.483815
39	28	20.00	0.22	20.00	2.00	0.5753	3.122	2.815	-13.9	-28.2	-5.18918
40	29	20.00	0.26	27.50	6.00	9.26	24.94	6.993	1.444	19.03	-1.85747
41	30	20.00	0.30	35.00	10.00	9.84	14.03	10.89	-0.02	-30.4	-1.79325
42	31	30.00	0.14	12.50	2.00	7.946	12.19	2.823	-1.26	-2.88	-4.05479
43	32	30.00	0.18	20.00	6.00	10.16	25.93	7.147	1.046	22.86	-3.8E-10
44	33	30.00	0.22	27.50	10.00	17.27	22.79	11.11	15.55	-7.78	-1.1657
45	34	30.00	0.26	35.00	14.00	11.26	35.48	15.27	-3.08	10.78	1.235731
46	35	30.00	0.30	5.00	18.00	25.34	59.93	19.75	18.32	26.75	8.166722
47	36	40.00	0.14	20.00	10.00	19.77	22.15	11.5	17.02	-12.3	-0.45762
48	37	40.00	0.18	27.50	14.00	14.84	33.91	15.56	-4.6	-1.19	-0.82791
49	38	40.00	0.22	35.00	18.00	18.71	41.53	19.53	9.793	9.877	6.996741
50	39	40.00	0.26	5.00	2.00	9.705	17.63	3.505	3.678	1.243	-4.76728
51	40	40.00	0.30	12.50	6.00	9.791	14.91	7.492	-2.5	-8.79	-0.22636
52	41	50.00	0.14	27.50	18.00	21.13	41.53	19.83	7.763	3.344	4.683534
53	42	50.00	0.18	35.00	2.00	2.41	3.866	3.09	-6.04	-18.1	-5.68438
54	43	50.00	0.22	5.00	6.00	10.7	15.1	7.542	-2.22	-20.4	-2.18861
55	44	50.00	0.26	12.50	10.00	14.97	26.86	11.81	4.226	-9.09	-0.98743
56	45	50.00	0.30	20.00	14.00	20.15	41.37	15.99	7.181	5.158	-1.09086
57	46	60.00	0.14	35.00	6.00	8.583	15.11	7.672	5.932	1E-11	-1.66774
58	47	60.00	0.18	5.00	10.00	17.75	23.87	11.79	8.349	-26.2	-0.75688
59	48	60.00	0.22	12.50	14.00	22.83	45.11	15.78	11.86	5.011	-0.0521
60	49	60.00	0.26	20.00	18.00	26.24	51.48	20.34	16.31	15.75	7.708756
61	50	60.00	0.30	27.50	2.00	1.787	4.247	3.307	-9.73	-24.6	-4.94321

Fig. 4 The worksheet of RVM

Based on the algorithm of RVM, the code of RVM is written in Excel and VBA. Parameters of RVM, the value of some υ_i and samples are shown in Fig. 4. The performance of the RVM model is shown in Fig. 5. It can be seen that the predicted pressure is well agreement with the monitored value. Therefore, RVM model can be used to calculate the borehole pressures as the response surface. It also shows the RVM model can represent well the nonlinear relationship between the pressures and geomechanics parameters.

Once the RVM model is finished, it can be used to combine with RS for back analysis. Based on the hybrid back analysis method, ABC algorithm is adopted to search for geomechanics parameters. The parameters of ABC, the values of geomechanics parameters and its searching range

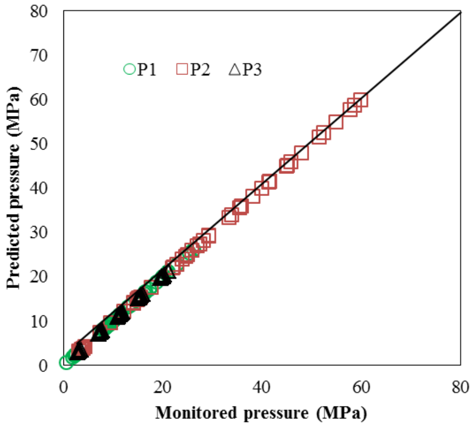

Fig. 5 The comparison between predicted pressure by RVM and monitored

Table 1 Results and its comparisons

	E (Gpa)	υ	$\sigma_{H\max}$ (MPa)	$\sigma_{h\min}$ (MPa)
Real parameters	40.0000	0.22	20.0000	10.0000
RVM-RS	40.1021	0.14	20.0214	10.5102
Eq. (1)	-	-	31.5305	-
Eq. (2)	-	-	21.5305	-

Table 2 The different searching range

Range	E (GPa)	υ	$\sigma_{H\max}$ (MPa)	$\sigma_{h\min}$ (MPa)
Range-1	[30.0-50.0]	[0.18-0.26]	[10.0-25.0]	[5.0-15.0]
Range-2	[20.0-60.0]	[0.14-0.30]	[5.0-35.0]	[2.0-18.0]
Range-3	[10.0-70.0]	[0.12-0.35]	[3.0-40.0]	[1.0-20.0]
Range-4	[5.0-80.0]	[0.10-0.40]	[1.0-50.0]	[0.2-30]

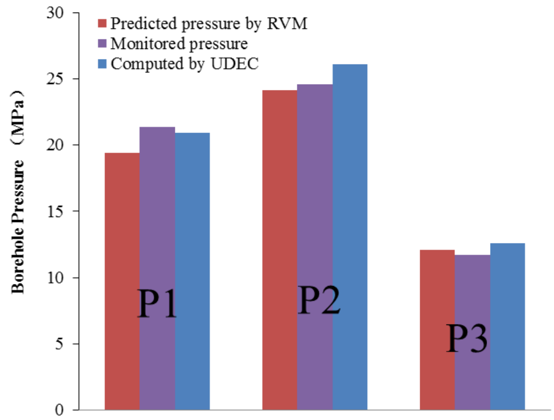

Fig. 7 The comparison pressure based on different method

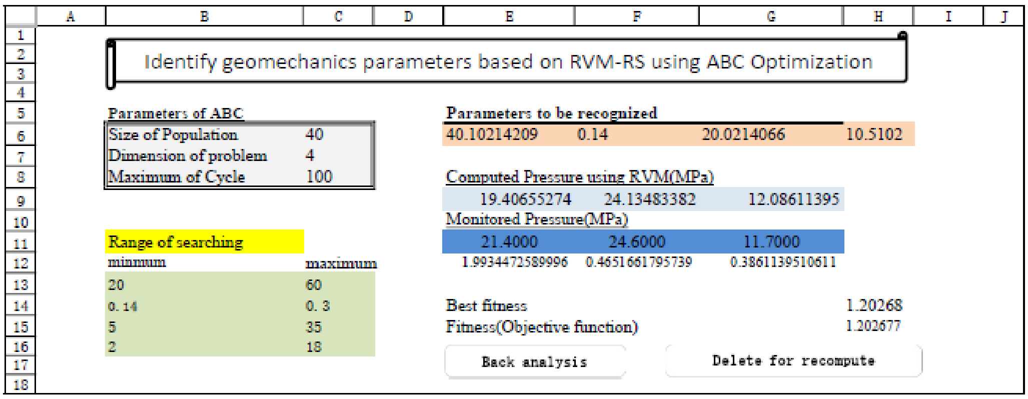

Fig. 6 The worksheet of hybrid back analysis using ABC

are shown in Fig. 6. Fig. 7 shows the borehole pressure comparison of predicted by RVM, calculated by UDEC based on identified geomechanics parameters and monitored value. It shows coupling RVM and RS is feasible to back analysis and the RVM model has a good performance for predicting the borehole pressure. Table 1 lists the geomechanics parameters in different method. It can be seen the identified maximum horizontal in situ stress is well agreement with the values base on Eq. (2). It shows the hybrid back analysis can be used to determine the maximum horizontal in situ stress without the poroelastic coefficient from the hydraulic fracturing. It is very useful in practice.

5.1 The effect of population size in ABC algorithm

Population size is important for the hybrid back analysis based on ABC algorithm. The bigger population size needs more time to search and the smaller population size maybe not find the global solution. Fig. 8 shows the convergence of hybrid back analysis using different population size. Fig.

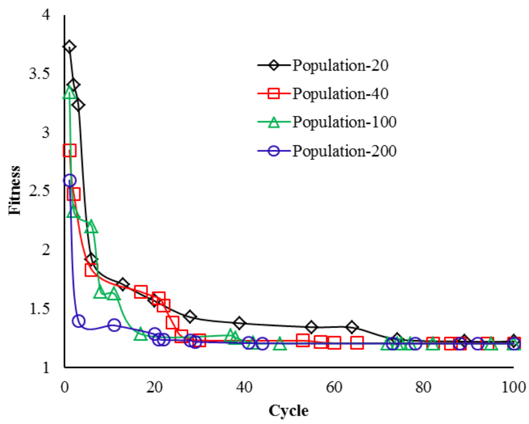

Fig. 8 The variation process of fitness in different population size

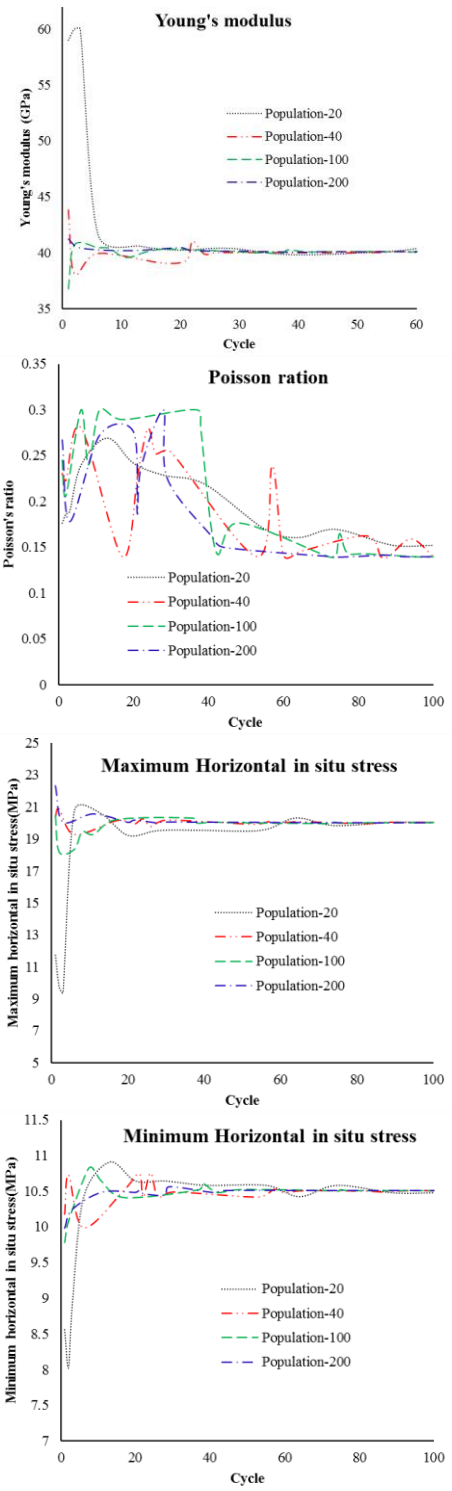

Fig. 9 The relationship between identified parameters and population size

9 shows the relation between identified geomechanics parameters and population size. It can be seen the hybrid back analysis can find the solution in cycle 40 when population size is 40. But it can find the solution with more cycle (about 80) when population size is 20. With the increasing of population size, the solution can be found with more little cycle (about cycle 20 with population 100 and 200). Need to say is the Poisson's ratio is not steady in

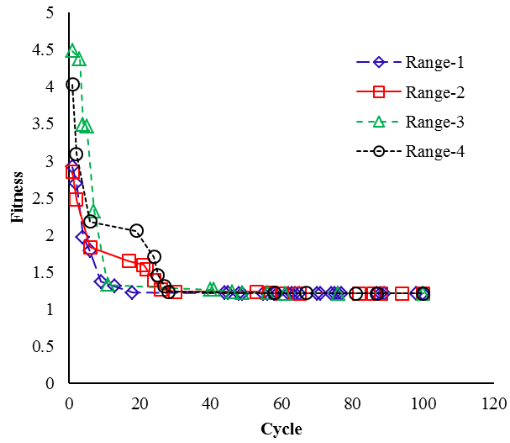

Fig. 10 The variation process of fitness in different searching range

convergence. The value of Poisson's ration also shows bigger error. It shows the Poisson's ration is not sensitive to borehole pressure.

5.2 The effect of searching range in ABC algorithm

The convergence and efficiency of back analysis are important to back analysis, special to large scale project problem. Fig. 10 shows relation between the fitness and the searching range (Table 2). It can be seen the small range has a quick convergence. While the hybrid back analysis can identify the goemechancis parameters to the larger searching range. Fig. 11 shows the relation between identified geomechanics parameters and searching range. It is evident that hybrid back analysis shows good performance in identifying the global optimum in a larger search range. For a complex, nonlinear response surface, there are some local minimum solutions. The global search performance is important in back analysis. The results show the hybrid back analysis is a robust and can find the global solution to the complex problem of hydraulic fracturing in a bigger searching range.

6. Conclusions

In this paper, a hybrid analysis procedure was proposed through coupling hydraulic fracturing model, response surface and RVM. And it has been successfully applied to a numerical example of geomechanics parameters from hydraulic fracturing. RVM was adopted to present the complex, nonlinear relationship between geomechanics parameters and borehole pressure. And then RVM model was coupled with response surface method to instead the numerical method in back analysis. ABC algorithm was used to search the geomechanics parameters. The proposed method was applied to a numerical example. The results matched the practice of hydraulic fracturing more reasonably. The effect of population size and searching range were discussed and its showed the proposed method is robust and has a good performance in global searching. In the classic breakdown formula, it is difficult to determine

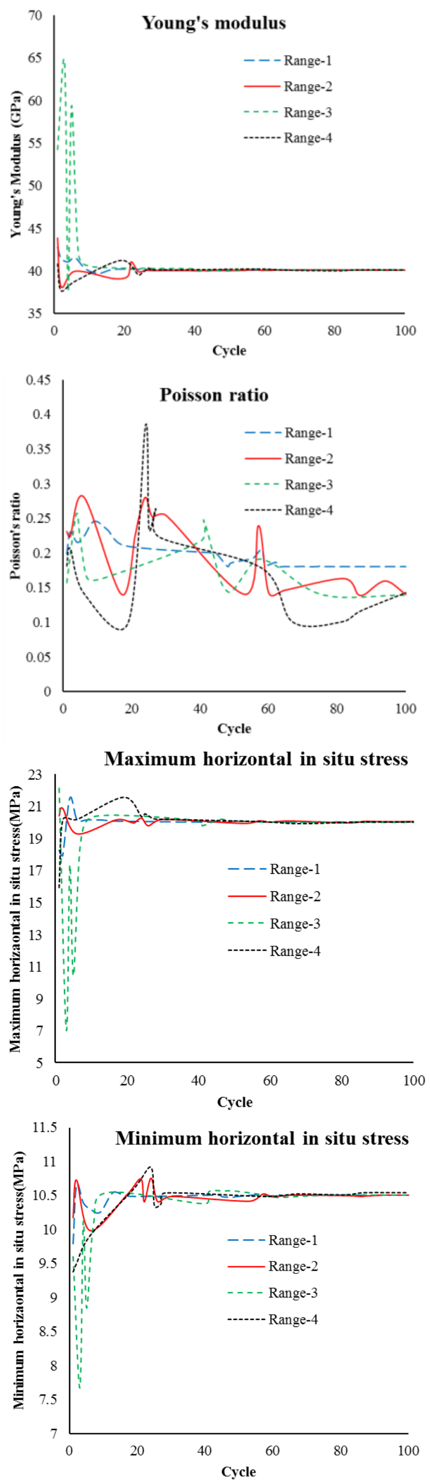

Fig. 11 The relationship between identified parameters and searching range

the maximum horizontal in-situ stress in practice while considering the poroelastic coefficient. The proposed method can predict the maximum horizontal in-situ stress based on the borehole pressures. This is important to determine the in-situ stress from hydraulic fracturing in practice. The proposed method is practical and accurate, and makes it convenient to be applied to identify geomechanics parameters from hydraulic fracturing.

Acknowledgements

Supports by the Program for Innovative Research Team (in Science and Technology) in University of Henan Province (No. 15IRTSTHN029) and the National Natural Science Foundation of China (No. 41072224) are gratefully acknowledged.

Reference

Agarwal, A. and Triggs, B. (2004), “3D human pose from silhouettes by relevance vector regression”, *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, 882-888.

Chen, S., Gunn, S. and Harris, C. (2001), “The relevance vector machine technique for channel equalization application”, *IEEE Tran. Neur. Network.*, **12**(6), 1529-1532.

Cividini, A. (1988), “Parameter estimation of a static geotechnical model using a Bayes approach”, *Int. J. Rock Mech. Min. Sci.*, **20**(5), 215-226.

Deng, J., Gu, D., Li, X. and Yue, Z.Q. (2005), “Structural reliability analysis for implicit performance functions using artificial neural network”, *Struct. Saf.*, **27**(1), 25-48.

Deng, J.H. and Lee, C.F. (2001), “Displacement back analysis for a steep slope at the Three Gorges Project site”, *Int. J. Rock Mech. Min. Sci.*, **38**, 259-268.

Fang, Z. and Khaksar, A. (2011), “Complexity of minifrac tests and implications for in-situ horizontal stresses in coalbed methane reservoirs’, IPTC14630, 1-13.

Farkhondeh, K., Hamid, S. and Mahabadi, E.S. (2016), “Relevance Vector Machine for Survival Analysis”, *IEEE Tran. Neur. Network. Learn. Syst.*, **27**(3), 648-660.

Feng, X.T., Zhao, H. and Li, S. (2004), “A new displacement back analysis to identify mechanical geo-material parameters based on hybrid intelligent methodology”, *Int. J. Numer. Anal. Meth. Geomech.*, **28**, 1141-1165.

Ferrero, A.M., Migliazza, M., Segalini, A. and Gullì, D. (2013), “In-situ stress measurements interpretations in large underground marble quarry by 3D modeling”, *Int. J. Rock Mech. Min. Sci.*, **60**, 103-113.

Ghorbani, M. and Sharifzadeh, M. (2009), “Long term stability assessment of Siah Bisheh powerhouse cavern based on displacement back analysis method”, *Tunnel. Underg. Space Technol.*, **24**(5), 574-583.

Gioda, G. and Jurina, L. (1981), “Numerical identification of soil structureinteraction pressures”, *Int. J. Numer. Anal. Meth. Geomech.*, **5**, 33-56.

Gioda, G. and Maier, G. (1980), “Direct search solution of an inverse problem in elastoplasticity: identification of cohesion, friction angle and in situ stress by pressure tunnel tests”, *Int. Numer. Meth. Eng.*, **15**, 1823-48.

Gomes, H.M. and Awruch, A.M. (2004), “Comparison of response surface and neural network with other methods for structural reliability analysis”, *Struct. Saf.*, **26**, 49-67.

Haimson, B.C. (1978), “The hydrofracturing stress measuring method and recent field results”, *Int. J. Rock Mech. Min. Sci. Geomech. Abstr*, **15**, 167-178.

Haimson, B.C. (1993), “The hydraulic fracturing method of stress measurement: theory and practice”, Ed. Hudson, J.A., Comprehensive Rock Engineering, Pergamon Press, Oxford, 395-412.

Haimson, B.C. and Fairhurst, C. (1969), “In-situ stress determination at great depth by means of hydraulic fracturing”, *The 11th U.S. Symposiumon RockMechanics (USRMS)*, Berkeley, CA.

Hubbert, M.K. and Willis, D.G. (1957), “Mechanics of hydraulic

fracturing", *Tran. AIME*, **210**, 153-163.
Karaboga, D. (2005), "An idea based on honey bee swarm for numerical optimization", Technical Report-TR06, Erciyes University, Engineering Faculty, Computer Engineering Department.
Karaboga, D. and Basturk, B. (2008), "On the performance of artificial bee colony (ABC) algorithm", *Appl. Soft Comput.*, **8**, 687-697.
Karaboga, D. and Ozturk, C. (2011), "A novel clustering approach: Artificial Bee Conoly (ABC) algorithm", *Appl. Soft Comput.*, **11**, 652-657.
Khuri, A. and Cornell, J.A. (1996), *Response Surfaces. Designs and Analyses*, Marcel Dekker Inc., New York.
Lewis, R.W. and Schrefler, B.A. (1998), *The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media*, 2nd Edition, Wiley, London.
Mirzaei, Z., Akbarpour, A., Khatibinia, M. and Siuki, A.K. (2015), "Optimal design of homogeneous earth dams by particle swarm optimization incorporating support vector machine approach", *Geomech. Eng.*, **9**(6), 709-727.
Okabe, T. (1998), "Inverse of drilling-induced tensile fracture data obtained from a single inclined borehole", *Int. J. Rock Mech. Min. Sci.*, **35**(6), 747-758.
Oreste, P. (2005), "Backanalysis techniques for the improvement of the understanding of rock in underground constructions", *Tunnel. Underg. Space Technol.*, **201**, 7-21.
Pichler, B., Lackner, R. and Mang, H.A. (2003), "Back analysis of model parameters in geotechnical engineering by means of soft computing", *Int. J. Numer. Meth. Eng.*, **57**, 1943-1978.
Puller, J.W., Mills, K.W., Jeffrey, R.G. and Walker, R.J. (2016), "In-situ stress measurements and stress change monitoring to monitor overburden caving behaviour and hydraulic fracture pre-conditioning", *Int. J. Min. Sci. Technol.*, **26**(1), 103-110.
Rechea, C., Levasseur, S. and Finno, R. (2008), "Inverse analysis techniques for parameter identification in simulation of excavation support systems", *Comput. Geotech.*, **35**, 331-345.
Rodriguez, R.F., Nicieza, C.G., Gayarre, F.L. and Lopez, F.L.R. (2015), "Application of hydraulic cylinder testing to determine the geotechnical properties of earth-filled dams", *Geomech. Eng.*, **9**(4), 483-498.
Sakurai, S. and Takeuchi, K. (1983), "Back analysis of measured displacements of tunnels", *Rock Mech. Rock Eng.*, **16**, 173-180.
Seyed, E.S., Reza, K. and Kaveh, A. (2014), "A new geomechanical approach to investigate the role of in-situ stresses and pore pressure on hydraulic fracture pressure profile in vertical and horizontal oil wells", *Geomech. Eng.*, **7**(3), 233-246.
Tipping, M. (2001), "Sparse Bayesian learning and the relevance vector machine", *J. Mach Learn. Res.*, **1**, 211-244.
Vardakos, S., Gutierrez, M. and Xia, C. (2012), "Parameter identification in numerical modeling of tunneling using the Differential Evolution Genetic Algorithm (DEGA)", *Tunnel. Underg. Space Technol.*, **28**, 109-123.
White, A.J. (2002), "The use of leak-off tests as means of predicting minimum in-situ stress", *Pet. Geosci*, **8**, 189-193.
William, W.G.Y. (1981), "Aquifer parameter identification with optimum dimension in parameterization", *Water Resour. Res.*, **17**(3), 664-672.
Witherspoon, P.A., Wang, J.S., Iwai, K. and Gale, J.E. (1980), "Validity of cubic law for fluid flow in a deformable rock fracture", *Water Resour. Res.*, **6**, 1016-1024.
Xu, H., Sang, S., Yang, J., Jin, J., Hu, Y., Liu, H., ... and Gao, W. (2016), "In-situ stress measurements by hydraulic fracturing and its implication on coalbed methane development in Western Guizhou, SW China", *J. Unconvent. Oil Gas Resour.*, **15**, 1-10.
Yu, Y.Z., Zhang, B.Y. and Yuan, H.N. (2007), "An intelligent displacement back-analysis method for earth-rockfill dams", *Comput. Geotech.*, **34**(6), 423-434.
Zhang, S. and Yin, S. (2013), "Reservoir geomechanical parameters identification based on ground surface movements", *Acta Geotech.*, **8**, 279-292.
Zhang, X., Last, N., Powrie, W. and Harkness, R. (1999), "Numerical modeling of wellbore behavior in fractured rock masses", *J. Petrol. Sci. Eng.*, **23**(2), 95-115.
Zhao, H. and Yin, S. (2016), "Inverse analysis of geomechanical parameters by artificial bee colony algorithm and multi-output support vector machine", *Inver. Prob. Sci. Eng.*, **24**(7), 1266-1281.
Zhao, H., Ru, Z. and Yin, S. (2012), "Relevance vector machine applied to slope stability analysis", *Int. J. Numer. Anal. Meth. Geomech.*, **36**(5), 643-652.
Zhao, H., Ru, Z., Chang, X., Yin, S. and Li, S. (2014), "Reliability analysis of tunnel using least square support vector machine", *Tunnel. Underg. Space Technol.*, **41**, 14-23.
Zhao, H.B. and Yin, S.D. (2009), "Geomechanical parameters identification by particle swarm optimization and support vector machine", *Appl. Math. Model.*, **33**, 3997-4012.
Zhou, W., Li, S.L., Ma, G., Chang, X.L., Cheng, Y.G. and Ma, X. (2016), "Assessment of the crest cracks of the Pubugou rockfill dam based on parameters back analysis", *Geomech. Eng.*, **11**(4), 571-585.
Zhu, H.Y., Guo, J.C., Zhao, X., Lu, Q., Luo, B. and Feng, Y.C. (2014), "Hydraulic fracture initiation pressure of anisotropic shale gas reservoirs", *Geomech. Eng.*, **7**(4), 403-430.

CC