LLNL-JRNL-628332

Surrogate-based optimization of hydraulic fracturing in pre-existing fracture networks

M. Chen, Y. Sun, P. Fu, C. R. Carrigan, Z. Lu, C. H. Tong

March 20, 2013

Computers & Geosciences

Disclaimer

Surrogate-based optimization of hydraulic fracturing in pre-existing fracture networks

Mingjie Chen[a]*, Yunwei Sun[a], Pengcheng Fu[a], Charles R. Carrigan[a], Zhiming Lu[b], and Charles H. Tong[c]

[a]Atmospheric, Earth and Energy Division, Lawrence Livermore National Laboratory, Livermore, CA, USA
[b]Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, USA
[c]Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, CA, USA

*Corresponding author:

Mingjie Chen, PhD
Lawrence Livermore National Laboratory
P.O. Box 808, L-223
Livermore, California 94551
1-925-423-5004 (office)
1-925-423-0153 (fax)
chen70@llnl.gov

Abstract

Hydraulic fracturing has been used widely to stimulate production of oil, natural gas, and geothermal energy in formations with low natural permeability. Numerical optimization of fracture stimulation often requires a large number of evaluations of forward hydraulic fracturing models, which are computationally expensive and even prohibitive in some situations. Moreover, there are a variety of uncertainties associated with the pre-existing fracture distributions and rock mechanical properties, which require evaluation of their impact on the optimal treatment. In this study, a surrogate-based approach is developed for efficient optimization of hydraulic fracturing well design in the presence of natural-system uncertainty. The fractal dimension is derived from the simulated fracturing network as the objective for maximizing energy recovery sweep efficiency. The surrogate model mapping the input parameters to the fractal dimension is constructed using training data generated by high-fidelity fracturing models, and it provides fast approximation of the objective functions. A suite of surrogate models constructed using different fitting methods is evaluated and validated for fast predictions. Global sensitivity analysis is conducted to gain insights into the impact of the input variables on the output of interest, and further used for parameter screening. The high efficiency of the surrogate-based approach is demonstrated for three optimization scenarios with different, uncertain ambient conditions. Our results suggest the critical importance of considering uncertain pre-existing fracture networks in optimization studies of hydraulic fracturing.

Keywords: Hydraulic fracturing; Fractal dimension; Surrogate model; Optimization; Global sensitivity

1. Introduction

Hydraulic communication is a key factor determining hydrocarbon or thermal energy recovery sweep efficiency in an underground reservoir. Sweep efficiency is a measure of the effectiveness of heat, gas or oil recovery process that depends on the volume of the reservoir contacted by an injected fluid. In the petroleum industry, hydraulic fracturing techniques have been used for over 60 years to increase hydraulic communication and stimulate oil and gas production (Britt, 2012). Artificial (stimulated) hydraulic fractures are usually initiated by injecting fluids into the borehole to increase the pressure to the point where the minimal principal stress in the rock becomes tensile. Continued pumping at an elevated pressure causes tensile failure in the rock, forcing it to split and generate a fracture that grows in the direction normal to the least principal stress in the formation. Hydraulic fracturing activities often involve injection of a fracturing fluid with proppants in order to better propagate fractures and to keep them open (Britt, 2012). The design of fracturing treatment should involve the optimization of operational parameters, such as the viscosity of the fracturing fluid, injection rate and duration, proppant concentration, etc., so as to create a fracture geometry that favors increased sweep efficiency. The net present value (NPV) introduced by Ralph and Veatch (1986) as the economic criteria, is usually used as the objective for optimal fracturing treatment design. Some studies have been reported to use a sensitivity-based optimization procedure coupled with a fracture propagation model and an economic model to find the optimal design parameters leading to maximum NPV (Balen et al., 1988; Hareland et al., 1993; Aggour and Economides, 1998). Nevertheless, this procedure, requiring brute-force parameter-sensitivity analysis, is tedious and incapable of exploring parameter space globally, which could potentially lead to the problem of converging to a local minimum of the objective function.

Rueda et al. (1994) optimized fracturing variables, including the injected fluid volume, pumping rate, fluid and proppant type, by applying a mixed integer linear programming (MILP) approach, which also lacks a global optimization capability. Mohaghegh et al. (1999) proposed a surrogate-based optimization approach by using a genetic algorithm to fit the dataset generated from a fracturing simulator that models both fracture propagation and proppant transport. Surrogate-based optimization refers to the idea of speeding optimization processes by using fast surrogate models. Queipo et al. (2002) applied a neural network algorithm to construct a “surrogate” of the NPV for an optimal design of hydraulic fracturing treatments. The objective function (NPV) was trained as a function of inputs by a synthetic dataset produced from a high-fidelity physics model, which integrated a fracturing simulator, a proppant transport and sedimentation model, a post-fracturing production model, and an economic model. This surrogate-based procedure is computationally less expensive for obtaining global minima without executing physics-model simulations, which are computationally prohibitive in some optimizations. However, none of these studies has considered optimizing the hydraulic fracturing of a pre-existing fracture network, which is a very common feature of rocks (Odling, 1992). Moreover, uncertainties of geomechanical properties and of the pre-existing fracture networks, resulting from the geologic architecture and fracture properties, such as fracture density, length, and orientation, etc. (Reeves et al., 2008), have not been rigorously studied for optimization of hydraulic fracturing treatment.

It has been demonstrated from field studies that fluid flow in fractured rock is primarily controlled by the fracture geometry and the interconnectivity between fractures (Long and Witherspoon, 1985; Cacas et al, 1990). A fractal is a self-similar geometric set (Mandelbrot, 1983) with Hausdorff-Besicovitch dimension exceeding the topological or Euclidian dimension, which is referred as fractal dimension. It is well recognized that natural fracture networks are fractal over a wide scale range (Barton, 1995; Bonnet et al., 2001), and fractal dimensions have been demonstrated to be efficient metrics for natural fracture patterns (e.g. LaPointe, 1988; Barton, 1995; Berkowitz and Hadad, 1997).

In this work, a surrogate-based optimization approach is proposed for optimizing hydraulic fracturing design in the presence of uncertainties in a pre-existing natural fracture network and its geomechanical properties. A state-of-the-art 2-D hydraulic fracturing model, GEOS-2D (Fu, et al., 2012), is used to simulate dynamic fracture propagation within a pre-existing facture network. Instead of integrating physical models and economic models to maximize NPV as the objective function, we focus on physical criteria, that is, the optimal hydraulic fracture propagation under uncertain natural conditions. The fractal dimension of the connected fractures can be derived from the post-fracturing network simulated by GEOS-2D to represent the network density and connectivity. More importantly, the scale-invariant feature of fractals allows observations from the core scale to be applied in another scale (e.g. reservoir scale). Therefore, the fractal dimension is chosen as the objective function to optimize the hydraulic fracturing well design. While a line, square, and cubic have the integer dimensions of 1, 2, and 3 respectively, the fractals in this study, which are applied to linear fractures in a 2-D plane, have a non-integer fractional dimension between 1 and 2.

In this paper, both non-parametric and parametric algorithms are used to construct surrogate models. Both types of surrogate models are quantitatively evaluated for prediction performance by cross-validations, and the best quality model is then selected for optimization. BOBYQA (Powell, 2009), a powerful and efficient derivative-free nonlinear optimization algorithm, is applied to drive a global search on the surrogate-modeled response surface. Compared to previous studies, our optimization methodology includes advances in (1) incorporating uncertain pre-existing natural fracture networks, (2) constructing both non-parametric and parametric surrogate models and conducting rigorous quality evaluations, (3) applying the high-efficient state-of-the-art optimizing algorithm, and (4) deriving the scale-invariant fractal dimension as the objective function.

2. Surrogate-based optimization approach

The proposed surrogate-based approach includes the following key steps (Fig. 1).

1. Construct samples of the parameter space.
2. Conduct mathematical simulations using the input samples from the previous step.
3. Derive the objective function from the simulated results.
4. Construct and validate surrogate models using the data from the previous steps for predication.
5. Perform optimization using selected surrogate model.

2.1. Sampling in parameter space

As shown in Fig. 1 and Table 1, an 11-dimensional parameter space is constrained by the ranges of the 11 input parameters. Latin Hypercube Sampling (LHS) procedure is used to draw N samples in the designed space following probability distribution functions (PDF) for each parameter. LHS is an effective stratified sampling approach ensuring that all portions of a given partition are sampled (McKay et al., 1979). Each point in the parameter space represents a deterministic vector for the 11 input variables. Fig. 1 shows an example of a 3-D input space, in which $N = 800$ points are sampled from the uniform distribution within the specified parameter range.

2.2. Hydraulic fracturing simulations

In this step, the computationally expensive physical models are executed N times with each input configuration sampled in the previous step. Each model includes generating an initial fracture network followed by simulation of hydraulic fracturing. The initial discrete fracture network is generated with fracture lengths controlled by the Pareto distribution (Odling, 1997)

$$P(L > l) = C \cdot l^{-a} \quad (1)$$

where P is the probability of a fracture of length larger than l, C is a constant that depends on the minimum fracture length in the system, which is assumed to be 5% of the domain size (100 m) in this study, and a is the power law exponent varying between 1 and 3 for natural fracture networks (Davy, 1993; Renshaw, 1999; Reeves et al., 2008). Typically, the mean fracture length of the fracture network increases as a decreases. Natural fracture networks usually consist of two fracture sets with most fractures in a set oriented in the same direction (LaPointe and Hudson, 1985; Ehlen, 2000). In this study, the fracture orientation refers to the angle between the fracture and the maximum principal stress direction (east). The orientation of the first fracture set ranges between 0^{o} and 135^{o}, while that of the second set is always 45^{o} more than the first one. For example, the orientation of the first fracture sets in the pre-existing fracture network shown in Fig. 1 is 25° from the input sample, hence that of the second set is 70°, with 45° from the first set.

Hydraulic fracturing under injected fluid pressure is simulated using an explicitly coupled hydro-geomechanical code, GEOS-2D, developed at Lawrence Livermore National Laboratory (Fu et al., 2012). This code couples a solid solver, a flow solver, a joint model, and a remeshing module, and is capable of dynamically simulating fracture propagation in a pre-existing fracture network. Fig. 1 presents the simulated fracture distribution after hydraulic fracturing with an injection well located at (0, 0), at sample point 1 with parameter values provided in Table 1.

2.3. Fractal dimension calculation

The fractal dimension of fractures opened by pressurized fluids can be reasonably representative of the density and connectivity of the network. Owing to self-similarity of fractals, the fractal dimension calculated from borehole samples can be extrapolated to reservoir-scale fracture networks. Due to these attractive features, the fractal dimension calculated from the simulated post-fracturing distribution is used as objective function of surrogate models for optimization. In the widely used box-counting method, measuring the fractal dimension of the fracture network (Barton and Larsen, 1985; Chil□s, 1988; Walsh and Watterson, 1993), the number of boxes of side length r needed to cover the fractal shape, $N(r)$, is approximated as a power law relation

$$N(r) = k \cdot (1/r)^{D}, \qquad (2)$$

where k is a constant and D is the fractal dimension. By log-transforming the both sides, we obtain

$$Log(N(r)) = D \cdot Log(1/r) + Log(k). \quad (3)$$

Thus, the fractal dimension D can be derived as the slope of the line linearly regressed from a series of size r and the corresponding $N(r)$. Fig. 1 shows that the fractal dimension of the simulated network is 1.725 from the well-fitted regression line with an R^2 value of 0.9963.

2.4. Surrogate-based optimization

Since surrogate models can be quickly constructed once the expensive training dataset is generated, we build alternatives from which the best one is selected according to the model validation results. The selected surrogate model is then used for evaluating objective functions for optimization or for other analyses.

2.4.1. Surrogate model construction

The calculated fractal dimensions, paired with the corresponding sample inputs, constitute the training data set for construction of the non-linear relations between them. For n paired observations, the model is given by

$$Y_i = f(\mathbf{x}_i) + \varepsilon_i, i = 1 \; to \; N. \quad (4)$$

Here, $\mathbf{x}_i$ is the input variable vector of sample i, Y_i is the response observation (calculated fractal dimension), $f(\mathbf{x}_i)$ is the mean response, ε_i is the error, and N is the sample number. Generally speaking, there are two kinds of fitting methods, namely, parametric and non-parametric regression. The parametric approaches, such as Gaussian Process

(GSP) and Polynomial Regression (PRG), presume a uniform global function form between input variables and the response variable, and require the estimation of a finite number of coefficients (Williams and Rasmussen, 1996; Draper and Smith, 1998), while non-parametric approaches, such as Multivariate Adaptive Regression Splines (MARS), use different types of local models in different regions of the data to construct the overall model (Friedman, 1991). In our approach, we build MARS, GSP, and PRG models and determine which one performs the best by follow-up validation. Various PRG models are also built with different order and different number of input variables that are the most sensitive ones ranked by global sensitivity analysis to be discussed next section. The first, second, and third order PRG including N_v input variables can be expressed as

$$f1(\mathbf{x}) = \beta_0 + \sum_{i=1}^{N_v} \beta_i x_i,$$

$$f2(\mathbf{x}) = f1(\mathbf{x}) + \sum_{i=1}^{N_v} \sum_{j=i}^{N_v} \beta_{ij} x_i x_j, \tag{5}$$

$$f3(\mathbf{x}) = f2(\mathbf{x}) + \sum_{i=1}^{N_v} \sum_{j=i}^{N_v} \sum_{k=j}^{N_v} \beta_{ijk} x_i x_j x_k,$$

where $\beta_0, \beta_{ij}, \beta_{ijk}$ are coefficients to be estimated. Higher order PRG can be formulated by adding higher-order terms. With more input variables included in higher order PRG, the fitting is better, but the number of coefficients increases, which must be less than the number (N) of observations (training dataset). Because of the limited training data, there is a trade-off between the order of PRG and the number of included variables for the best fit.

2.4.2. Global sensitivity analysis

Sensitivity is a measure of the contribution of an independent variable to the total variances of the dependent variable. Sensitivity analysis of a model system can be used as the following purposes.

1. Parameter screening: fix one or more of the input variables with negligible influence on the output variability.
2. Variable prioritization: rank input variables according to their sensitivity indices.
3. Variable selection for reducing uncertainty: invest money to measure those sensitive variables that can reduce output uncertainty to maximum extent.

There are numerous methods for sensitivity analysis (Frey and Patil, 2002), among which the Sobol' method (Sobol', 1993) is used to drive global sensitivity analysis of input variables for the response variable, i.e., the fractal dimension. Using the training dataset, the relative importance of input variables is quantified by Sobol' total sensitivity indices. In this study, the sensitivity analysis for the preliminary experiment screens out the non-sensitive parameters to reduce the parameter dimension for the 2nd stage experiment of optimization. The selection of input variables in the PRG models is also based on parameter ranking by Sobol' indices.

2.4.3. Model validation and selection

A well-fitted surrogate model does not necessarily mean that it is good for prediction. It is easy to over-fit data by including too many degrees of freedom. One way to measure the predictive ability of a surrogate model is to test it using a test dataset, which is split from the sample data and not used in training. Nevertheless, it will limit the data available for constructing the surrogate models. Alternatively, the popular leave-one-out cross-validation (LOOCV) method can make use of the available sample data much more efficiently (Picard and Cook, 1984). Given *N* input samples, a surrogate model is constructed *N* times, each time leaving out one of the input sample from training, and using the omitted sample to test the model. The generalization error of the LOOCV can be estimated using the root mean square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - f_i^{(-i)})^2}{N}} \qquad (6)$$

where Y_i represents the i^{th} response observation (calculated fractal dimension), and $f_i^{(-i)}$ denotes the prediction (interpolated fractal dimension) tested by sample *i* using the surrogate model fitted by all the other *N*-1 samples. The surrogate model with a minimum RMSE is selected for optimization.

2.4.4. Optimizer

Bound Optimization BY Quadratic Approximation (BOBYQA) algorithm is applied to search the minimal objective function (negative fractal dimension) of the surrogate model $f(\mathbf{x}), \mathbf{x} \in \mathcal{R}^N$, where $\mathcal{R}^N$ is the *N*-dimensional parameter space constrained by the range of each input variable. BOBYQA is a powerful numerical optimization solver for derivative-free nonlinear problems, subject to simple bound constraints (Powell, 2009). In the case studies, optimal hydraulic fracturing design parameters and natural field properties corresponding to the minimal objective function are found on the response surface using BOBYQA optimizer.

2.5. Implementation

The proposed approach was implemented in a Python code that couples the hydraulic fracturing simulator GEOS-2D (Fu et al., 2012) with the uncertainty quantification tools contained within the PSUADE code (Tong, 2009). PSUADE (Problem Solving environment for Uncertainty quantification And Design Exploration) is a suite of uncertainty quantification modules capable of addressing high-dimensional sampling, parameter screening, global sensitivity analysis, response surface analysis, uncertainty assessment, numerical calibration, and optimization (Hsieh, 2007; Wemhoff and Hsieh, 2007; Sun et al., 2012). The computationally expensive hydraulic fracturing simulations for generation of the synthetic training dataset (GEOS-2D) are executed using the high performance computing facilities at Lawrence Livermore National Laboratory. The box-counting method for deriving fractal dimension of connected fractures from the post-fracturing distribution is implemented in a Fortran code.

3. Case study: hydraulic fracturing well design optimization

In this section, the developed surrogate-based approach is applied to optimizing the hydraulic fracturing well design (location and length) in a 2-D domain under uncertain natural-system conditions. To reduce the dimensionality of the input parameter space, preliminary simulations are performed to generate a training dataset used to conduct global sensitivity analysis for parameter screening. The input parameter sampling and numerical simulations are presented in Fig. 1 and Table 1. Based on $N = 800$ observation pairs, Sobol' total sensitivity indices are derived and parameter importance is ranked (Table 1). Of the $Nv = 11$ input parameters, two operational ones, working fluid viscosity and injection pressure, are found to be the most important for effective fracturing. The four least sensitive parameters with Sobol' indices less than 0.01 are screened out. The remaining variables – two parameters related to pre-existing network, fracture orientation and number, and three parameters related to rock mechanical properties, Young's modulus, minimum principal stress, and stress anisotropy, are included for the optimization experiment described below.

3.1. Experimental design

As illustrated in Fig. 2, a horizontal injection well is placed in an experimental 2-D physical domain along its left-most boundary (along the y axis). The pertinent design parameters of interest here include the length of the open (perforated) injection interval (anywhere from 0 to 40 m) and its center lying between y = -20 and 20 m. The design parameters, plus the five most important natural-system parameters determined above, are treated as uncertain parameters. A total of 529 input samples are drawn from the seven-dimensional parameter space using the LHS sampling method. Two of the seven parameters, fracture orientation and the number of fractures in the pre-existing network, are fed into the pre-existing fracture model and the remaining five are applied to the hydraulic fracturing model. Instead of injection pressure, injection rate is used as the source term of the fracturing model. The total injection rate is fixed at 0.25 m^3/s, and is averaged over the perforated well length, which is subdivided into 2-m long injection nodes. As a result, the injection rate applied on each injection node decreases linearly with increasing horizontal-well length.

3.2. Synthetic dataset analysis

For each of 529 input samples, pre-existing network are generated and GEOS-2D models are executed, and nine snapshots of post-fracturing network distributions are exported in nine sequential time steps from which the fractal dimensions are derived.

Mean values of the 529 fractal dimensions increase with the injection time or fluid volume (Fig. 3a), suggesting that fracture networks keep growing with the continuous injection of fluid. The time series of the mean fractal dimensions also indicate that their growth rates are very high initially, and gradually decreases to nearly zero from 11.4 to 51.1 seconds, suggesting the economic benefit of hydraulic fracturing declines with time. The probability distribution of the 529 fractal dimension results in the last snapshot at 51.1 seconds shows that most of them are between 1.5 and 1.7, and the value with highest possibility (10%) is around 1.65 (Fig. 3b). The corresponding cumulative probability indicates that about 25% of 529 fractal dimensions are less than 1.5, 50% less than 1.6, and 75% less than 1.65. Only 10% of these fracture dimension values are above 1.7 and the maximum value is 1.79. The nine sets of 529 observation pairs consisting of the seven input variables and the corresponding fractal dimension are served as the training and testing dataset for surrogate models.

3.3. Global sensitivity analysis

All seven input variables are normalized between zero and one, based on their upper and lower bounds. For each input sample, fracture distributions at nine sequential injection time steps were generated, from which the corresponding fractal dimensions are derived. Fig. 4 shows the global sensitivity of nine sets of fractal dimensions to the seven input variables sorted by the last set. For all the nine time steps, the variability of fractal dimensions is largely influenced by the initial fracture number and well length (Sobol' indices > 0.5), and moderately by the other 5 input variables, indicating that the initial fracture number is the key uncertain parameter influencing post-fracturing conditions. Injection lengths (and the corresponding averaged injection rate) are the key contributors to the variability of fractal dimensions at the earlier injection stages, while initial fracture number becomes the key contributor at the later stages. Well center location strongly affects the fractal dimension (Sobol' indices = 0.4), while becoming marginally important (Sobol' indices = 0.1) as injection proceeds. Overall, two stress parameters, minimum principal stress, stress anisotropy, and fracture set orientation, influence the objective somewhat more than does Young's modulus. The sensitivity information inferred above is used to rank variable prioritization to be included in PRG models below.

3.4. Surrogate models evaluation

Non-parametric MARS, parametric GSP and 11 PRG models with various parameters and orders are constructed for the nine snapshots, each using 529 observation pairs (input parameters vs. fractal dimension). Table 2 shows the comparison of MARS, GSP and 11 PRG models constructed for the post-fracturing distribution (i.e., the last snapshot). The natural-system parameters included in PRG models are determined according to the importance ranking by Sobol' indices (Fig. 3). For examples, minimum principal stress is dropped off for the 6-parameter PRG model, and Young's modulus is further excluded from the 5-parameter PRG model, because the two parameters are ranked as least important for fracturing at the final timestep. In terms of fitting error, the more coefficients that are included, the higher the accuracy of PRG models becomes. In fact, when the number of coefficients is greater than 125, PRG models fit the training dataset better than the MARS model does. Nevertheless, the predictive ability, tested against a new dataset, will usually get worse as more terms are included, due to over-fitting. As shown in Table 2, the RMSE of cross-validation for each surrogate model confirms that the best fitted PRG model with 461 coefficients turns out to be the worst in prediction performance, and the quadratic PRG, with seven variables and just 35 estimated coefficients, had the best prediction performance among 11 PRG models. Finally, the MARS model is selected for optimization due to its better prediction performance than both GSP and the best PRG model.

To illustrate the surrogate model quality regarding fitting and validation, the scatter plots of fractal dimension simulated by surrogate models versus GEOS-2D from 529 sample inputs are compared between MARS model and the best-fitted, but worst-validated PRG model (5-order 6-parameter) (Fig. 5). The closer the points are to the diagonal line, the better the surrogate model matches the physical model. It is seen that the points are clustered closely along the diagonal line for the PRG model fitting (RMSE=0.00805), but are significantly scattered for cross-validation (RMSE=0.401). Conversely, points in both the MARS fitting and cross-validation scatter plots are moderately spread with 0.0257 and 0.0410 of RMSE, respectively.

3.5. Horizontal well design optimization

The problem of interest is to find the favorable fracture-stimulation well design variables, namely, well center y location and the perforation length, in the presence of natural-system uncertainty. To investigate how natural-system uncertainty affects optimal well design, three optimization cases with sequentially decreasing natural-system uncertainty are performed for the last snapshot at an injection time of 51.1 seconds. Case A searches the minimum objective function (maximum fractal dimension) in a 7-D parameter space, with two design variables and with five natural-system variables treated as uncertain. Case B is adapted from case A, with the uncertainty reduced by fixing the fracture orientation and number, which are two parameters describing the pre-existing fracture network. In case C, only well location and length are allowed to vary within the specified ranges during the optimization process, by further fixing the three geomechanical variables affecting fracture propagation, minimal principal stress, stress anisotropy, and Young's modulus. The objective function to be minimized is the negative fractal dimension. All the three optimization cases are efficiently conducted using surrogate models without rerunning the expensive physics-based GEOS-2D, due to the flexibility of our surrogate-based approach. The BOBYQA optimizer, coupled with the selected MARS models, is executed for the three inverse problems.

Fig. 6 depicts the optimization processes, which involves searching the minimal objective function for each of the three cases. It is seen that the number of evaluations of the surrogate model required to satisfy the convergence criteria (10^{-6}) is 337, 269 and 994, respectively. Each of the optimizations requires hundreds of model evaluations and can be completed in less than a minute, while a single realization conducted with the GEOS-2D code costs tens of hours. Moreover, a physics-based model is usually not as smooth as its surrogate, implying that a greater number of model evaluations are required for convergence than required by surrogate-based optimization. As a result, the high-efficient surrogate-based optimization approach can make the otherwise computationally prohibitive procedure practically achievable. An example of an expensive procedure is Bayesian stochastic joint inversion modeling using hard (borehole core) and soft data (geophysical survey), which usually entails expensive Markov Chain Monte Carlo sampling. Another advantage of the surrogate-based approach is its high degree of flexibility. Once the training data is generated from the expensive physics-model simulations, numerous surrogate models can be constructed and validated for optimization within a very short time.

The optimal values of the parameter sets corresponding to the minimum objectives are listed in Table 3. Case A represents a scenario in which the hydraulic fracturing treatment is designed with minimal knowledge of the targeted field; thus, a wide range of the natural-system properties must be accounted for. The optimal location of the well center is found to be 4.31 m on the y axis, and the optimal well length is 0.08 m. This indicates that, to obtain a maximum fractal dimension, the fluid should be injected in just one injection node at y = 4 m, and at the rate of 0.25 m^3/s, if fracturing is to be optimized for this level of natural-system uncertainty. With the entire injection rate concentrated at one node, the maximum possible hydraulic pressure is achieved, which confirms our intuition about what will maximize the growth of the fracture network.

Case B assumes both the fracture orientation and fracture number of the pre-existing network are already determined to be 1^{o} and 250, respectively, on the basis of borehole core data or other geophysical measurements. The optimal well design parameters (position and length) are found to be 5.09 m and 21.3 m, which corresponds to a hydraulic fracturing scheme where fluid is injected into 11 nodes, centered at y = 5 m, with each injected at a rate of 0.25/11 = 0.0227 m^3/s. Unlike case A, where all of the fluid injection (and pressurization) is concentrated in one node, pressurization in case B is distributed along 11 nodes, suggesting that both the distribution and magnitude of pressure are important for creating a favorable fracturing network, and must be traded off given the limited total injection volume. The maximum fractal dimension is 1.622, which has been significantly reduced from 1.872 in case A, demonstrating the importance of considering uncertainty of the pre-existing fracture network for optimizing the hydraulic fracturing treatment. Sensitivity analysis has shown that the fractal dimension is highly sensitive to the initial fracture number (Fig. 3), so it is reasonable to conclude that the large decrease of fractal dimension from case A to B results from the large reduction of the initial fracture number from 486 to 250. It is also seen that fracture orientation and Young's modulus differ a lot from case A to B, but since they were found not to strongly affect fractal dimension, they are not likely to be the main contributors to its decrement.

Case C is designed to investigate the optimal well injection scheme given full knowledge of the natural system, with all five natural-system properties fixed as listed in Table 3. The optimized well injection design parameters turn out to be similar to those in case B, suggesting that uncertainty of the three rock mechanical parameters has a small influence on the optimization results. On the other hand, the comparison with case A shows that the uncertainties of the two input variables for pre-existing fracture network can lead to a big difference in the optimization results. These findings demonstrate the importance of addressing uncertainty of the pre-existing fracture network, rather than addressing that of the rock geomechanical properties in optimizing hydraulic-fracturing treatments, which was lacking in previous studies. The moderate decrement of maximal fractal dimension from case B to C is believed primarily caused by the increment of stress anisotropy from 1.0 to 1.2, based on the fact of its relatively small sensitivity to the other varied rock properties (Fig. 3). The 2-D response surface for case C is shown in Fig. 7. Apparently, multiple local minimal objective functions exist, with the global minimum being found using the BOBYQA optimizer.

Fig. 8 plots the three post-fracturing distributions simulated using the corresponding optimal input parameter sets. It is apparent that the network connected by fluid injection for case A sweeps a larger area than the other 2 cases, demonstrating that the fractal dimension of opened fracture network is an appropriate indicator of the potential energy sweep efficiency in the target field. The fractures in case C propagate mainly along x axis (maximum principal stress direction) since the stress field is moderately anisotropic while the stresses in case A and B are almost isotropic.

4. Summary and conclusion

A surrogate-based optimization approach involving high-dimensional parameter space sampling, numerical physics-model simulations, objective-function derivation, surrogate-model construction and validation, with the coupled execution of the optimizer and surrogate models, is proposed and implemented for optimizing hydraulic-fracturing treatments. For a strongly non-linear process, such as hydraulic fracturing considered in this study, the surrogate model constructed by the non-parametric MARS method is demonstrated to have the best prediction performance according to the cross-validation, and hence was selected for optimizing the hydraulic fracturing treatment. The 3 optimization cases, each requiring hundreds of surrogate model evaluations to meet convergence tolerance, are completed in less than one minute, demonstrating the high efficiency of the approach. A comparison study of 3 optimization cases is conducted by varying the dimensionality of the parameter space without rerunning expensive physics-model simulations. Moreover, additional optimizations using surrogate models can be performed quickly and easily for particular purposes if necessary, for example, reducing the uncertainty of an input variable by narrowing its range.

The comparison study shows the optimization results depend on the degree of uncertainty of the pre-existing fracture networks. This indicates the importance of incorporating information about pre-existing fracture networks into the process of optimizing hydraulic fracturing treatment, which has been largely overlooked by previous optimization studies in the literature. In contrast, the influence of uncertainty in rock geomechanical properties on the optimal injection scheme is found to be less important. These findings suggest that the pre-existing fracture network, rather than the geomechanical properties, should be the top priority to be characterized before designing a hydraulic fracturing treatment.

The statistical analysis of the training data and fracture networks for the three optimized hydraulic-fracturing cases indicates that fractal dimension is a useful metric for quantifying the density and connectivity of a fracture network. Furthermore, the scale-invariant nature of the fractal makes it a universal indicator for the fracture network across wide range of spatial scales, from core through outcrop to aerial image scale. The successful incorporation of fractal dimension into the efficient surrogate-based approach in this study provides a useful solution for other inverse problems that suffer from the heavy computational burden and multi-scale measurements, such as the stochastic joint inversion problem.

The decreasing growth rate of the mean fractal dimension with injection time implies the diminishing value of continuing the hydraulic fracturing operation. Therefore, there exists a cost-efficient time to stop the fracturing operation, that is, the injection time and the rate need to be optimized for economic objective. Although this paper is focused on incorporating uncertainty of the natural system into optimization and hence only considers the physical criterion as the objective function, the presented surrogate-based optimization approach, can be modified to find optimal injection rate and time by integrating an energy production model and economic model to derive both physical and economic criteria as the objective function.

Acknowledgements

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. We would like to thank LDRD-SI program of Lawrence Livermore National Laboratory for supporting this project. We greatly appreciate the review comments from Tom Buscheck and Andy Tompson that improved the paper.

References

Aggour, T.M., Economides, M.J., 1998. Optimization of the performance of high-permeability fractured wells. In: SPE International Symposium on Formation Damage Control, Lafayette, LA, SPE Paper 39474.

Balen, M.R., Meng, H.Z., Economides, M.J.,1988. Application of net present value (NPV) in the optimization of hydraulic fractures. In: SPE Eastern Regional Meeting, Charleston, South Carolina, USA, SPE Paper 18541, pp. 181-191.

Barton, C.C., Larsen, E., 1985. Fractal geometry of two-dimensional fracture networks at Yucca Mountain, southwestern Nevada. In: Stephannson O. (Ed.), Proceedings of the International Symposium on Fundamentals of Rock Joints, Gentek, Lulea, Sweden, pp. 77-84.

Barton, C.C., 1995. Fractal analysis of scaling and spatial clustering of fractures. In: Barton C.C., LaPointe, P.R. (Eds.), Fractals in the Earth Sciences, Plenum, New York, pp. 141-178.

Berkowitz, B., Hadad, A., 1997. Fractal and multifractal measures of natural and synthetic fracture networks. Journal of Geophysical Research 103, 12205-12218.

Bonnet, E.,O. Bour, N.E. Odling, P. Davy, Main,I., Cowie P., Berkowitz B., 2001. Scaling of fracture systems in geological media. Reviews of Geophysics 39(3), 347–383. doi:10.1029/1999RG000074.

Britt, L., 2012. Fracture stimulation fundamentals. Journal of Natural Gas Science and Engineering 8, 34-51. http://dx.doi.org/10.1016/j.jngse.2012.06.006.

Cacas, M.C., Ledoux, E., De Marsily, G., Tilie, B., Barbreau, A., Durand, E, Feuga, B., Peaudecerf, P., 1990. Modeling fracture flow with a stochastic discrete fracture network: calibration and validation: 1. The flow model. Water Resources Research 26, 469-489. doi:10.1029/WR026i003p00479.

Chil□s, J.P., 1988. Fractal and geostatistical methods for modeling of a fracture network. Mathematical Geology 20, 631-654.

Davy, P., 1993. On the frequency-length distribution of the San Andreas fault system. Journal of Geophysical Research 98(B7), 12141-12151.

Draper N.R., Smith H., 1998. Applied Regression Analysis, 3rd edn., Wiley Interscience, New York, NY, 736pp.

Ehlen, J., 2000. Fractal analysis of joint patterns in granite. International Journal of Rock Mechanics and Mining Sciences 37, 902-922.

Frey, H. Patil, S., 2002. Identification and review of sensitivity analysis methods. Risk Analysis 22, 553-578.

Friedman, J.H., 1991. Multivariate adaptive regression splines. The Annals of Statistics 19, 1-67.

Fu, P., Johnson, S.M., Carrigan, C.R., 2012. An explicitly coupled hydro-geomechanical model for simulating hydraulic fracturing in arbitrary discrete fracture networks. International Journal for Numerical and Analytical Methods in Geomechanics. doi: 10.1002/nag.2135.

Hareland G.l., Rampersad, P., Dharaphop, J., Sasnanand S., 1993. Hydraulic fracturing design optimization. In: SPE Eastern Regional Conference and Exhibition, Pittsburgh, PA, USA, SPE Paper 26950, pp. 493-500.

Hsieh, H., 2007. Application of the PSUADE tool for sensitivity analysis of an engineering simulation. Lawrence Livermore National Laboratory, UCRL-TR-237205.

Larry, B., 2012. Fracture stimulation fundamentals. Journal of Natural Gas Science and Engineering 8, 34-51. doi:10.1016/j.jngse.2012.06.006.

LaPointe, P.R., Hudson, J.A., 1985. Characterization and interpretation of rock mass joint patterns. Geological Society of America Special Paper, 37pp.

LaPointe, P.R., 1988. A method to characterize fracture density and connectivity through fractal geometry. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts 25, 421-429.

Long, J.C.S., Witherspoon, P.A., 1985. The relationship of the degree of interconnection to permeability of fracture networks. Journal of Geophysical Research: Solid Earth 90(B4), 3087-3098.

Mandelbrot, B.B., 1982. The Fractal Geometry of Nature, W. H. Freeman, New York, NY, 468pp.

McKay, M., Beckman, R., Conover, W., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2), 239-245.

Mohaghegh, S., Balanb, B., Platon, V., Ameri, S., 1999. Hydraulic fracture design an optimization of gas storage wells. Journal of Petroleum Science and Engineering 23, 161-171.

Odling, N.E., 1992. Network properties of a two-dimensional fracture pattern. Pure and Applied Geophysics 138, 95-114.

Odling, N.E., 1997. Scaling and connectivity of joint systems in sandstones from western Norway. Journal of Structural Geology 19, 1257-1271.

Picard, R.R., Cook, R.D., 1984. Cross-validation of regression models. Journal of the American Statistical Association 79, 575-583.

Powell, M.J.D., 2009. The BOBYQA algorithm for bound constrained optimization without derivatives. Report DAMTP 2009/NA06, Centre for Mathematical Sciences, University of Cambridge, UK .

Queipo, N.V., Verde, A.J., Canelon, J., Pintos, S., 2002. Efficient global optimization for hydraulic fracturing treatment design. Journal of Petroleum Science and Engineering 35, 151-166.

Ralph, W., Veatch Jr., W., 1986. Economics of fracturing: some methods, examples and case studies. In: SPE 61th Annual Technical Conference and Exhibition, New Orleans, Atlanta, USA, SPE Paper 15509, pp. 1-16.

Renshaw, C.E., 1999. Connectivity of joint networks with power law length distributions. Water Resources Research 35(9), 2661-2670.

Reeves, D.M., Benson D.A., Meerschaert, M.M., 2008. Transport of conservative solutes in simulated fracture networks: 1. Synthetic data generation. Water Resources Research 44. doi: 10.1029/2007WR006069.

Rueda, J.I., Rahim, Z., Holditch, S.A., 1994. Using a mixed integer linear programing technique to optimize a fracture treatment design. In: SPE Eastern Regional Meeting, Charleston, South Carolina, USA, SPE Paper 29184, pp.233-244.

Shook, G.M., 2003. A simple, fast method of estimating fractured reservoir geometry from tracer tests. Geothermal Resources Council Transactions 27, 407-411.

Sobol' I., 1993. Sensitivity estimates for non-linear mathematical models. Mathematical Modeling and Computational Experiment 4, 407-414.

Sun, Y., Tong, C.H., Duan, Q., Buscheck, T.A., Blink, J.A., 2012. Combining simulation and emulation for calibrating sequentially reactive transport systems. Transport in Porous Media. doi 10.1007/s11242-011-99170-4.

Tong, C.H., 2009. PSUADE User's Manual (Version 1.2.0). Lawrence Livermore National Laboratory, LLNL-SM-407882.

Walsh, J., Watterson, J.J., 1993. Fractal analysis of fracture pattern using the standard box-counting technique: Valid and invalid methodologies. Journal of Structural Geology 15, 1509-1512.

Wemhoff, A.P., Hsieh, H., 2007. TNT Prout-Tompkins kinetics calibration with PSUADE. Lawrence Livermore National Laboratory, UCRL-TR-230194.

Williams, C.K.I., Rasmussen C.E., 1996. Gaussian process for regression. In: Touretzky D.S., Mozer M.C., Hasselmo M.E. (eds.), Advances in Neural Information Processing Systems 8. MIT Press, Cambridge, pp. 514-520.

Table list:

Table 1. Preliminary experiment: parameter importance ranking for the fractal dimensions of opened fractures in post-fracking networks according to Sobol's total sensitivity indices.

Table 2. Evaluation of surrogate models for fracture network at final time.

Table 3. Optimization of well center location and length for fracture network at final time.

Figure list:

Fig. 1. Surrogate-based modeling approach for simulated hydraulic fracturing.

Fig. 2. An example of a horizontal well (center at y = 0 m and length=40 m) placed in a pre-existing network (orientation =0^{o} and number =250). The red solid line is horizontal injection well with uncertain location and length along left y axis. The pre-existing fracture orientation and number of natural network are also uncertain. The maximum and minimum principal stress are assumed x and y direction respectively.

Fig. 3. Global sensitivity of fractal dimension to the 7 input parameters for 9 sequential injection time steps. The 9 parameter sequences are ordered according to the last one.

Fig. 4. Statistics of 529 derived fractal dimension: (a) mean for 9 injection time or volumes, (b) PDF and CDF at finial time(51.1 seconds or 12.8 m^3 injected fluid volume).

Fig. 5. Scatter plots of fractal dimension simulated using surrogate model data versus physical model from 529 input samples: the comparison of fitting (red dots) and cross-validation (blue dots) between MARS and 5-order 6-parameter PRG surrogate models. The tighter the points clustered along diagonal, the closer the surrogate model data match the physical model data.

Fig. 6. Minimal objective searching curve for optimization case with (a) 7 uncertain parameters, (b) 5 uncertain parameters, and (c) 2 uncertain parameters. The optimal parameter values for the 3 cases are proved in Table 3.

Fig. 7. The visualized response surface with 2 uncertain design parameters, i.e., horizontal well center at y axis and well length.

Fig. 8. The post-fracking network corresponding to the optimal parameter set (values provided in Table 3) with (a) 7 uncertain parameters, (b) 5 uncertain parameters, and (c) 2 uncertain parameters. The optimal parameters for the 3 cases are provided in Table 3. The color of the fractures is based on the hydro-pressure. Red indicates the maximum pressure, while blue denotes zero pressure, meaning closed fractures which are not included in fractal dimension calculation.

Table 1. Preliminary experiment: parameter importance ranking for the fractal dimensions of opened fractures in post-fracking networks according to Sobol' total sensitivity indices

Parameter name	PDF [1]	Min	Max	Sample#1	Indices	Rank
11.Fluid viscosity (Pa-s)	Log-U	0.0001	0.001	0.00025	0.51	1
6.Injection pressure / σ_h	U	1	2	1.7	0.43	2
1.Fracture orientation (degree)	U	0	135	25	0.050	3
2.Initial fracture numbers	U	50	500	250	0.031	4
7.Young's modulus (GPa)	U	5	50	31	0.026	5
4.Minimum principal stress σ_h (MPa)	U	10	15	10.1	0.022	6
5.Stress anisotropy (σ_H/σ_h)	U	1	2	1.3	0.014	7
9.Poisson's ratio	U	0.1	0.5	0.2	0.0024	8
3.Fracture power law exponent	U	1	3	1.8	0.001	9
8.Joint friction coefficient	U	0.5	1.2	0.7	0.0	10
10.Fracture toughness (MPa-m$^{0.5}$)	U	0.2	2.0	1.0	0.0	11

[1] U and Log-U denote uniform and log-uniform distribution.

Table 2. Evaluation of surrogate models for fracture network at final time.

Construction method	Estimated Coefficients	RMSE	
		Fitting	Validation
MARS	-	0.0257	0.0410
GSP	-	0.0278	0.0428
1-order 7-parameter PRG	7	0.0473	0.0483
2-order 6-parameter PRG	27	0.0390	0.0458
2-order 7-parameter PRG	35	0.0378	0.0436
3-order 5-parameter PRG	55	0.0326	0.0452
3-order 6-parameter PRG	83	0.0300	0.0458
3-order 7-parameter PRG	119	0.0283	0.0462
4-order 5-parameter PRG	125	0.0258	0.0506
4-order 6-parameter PRG	209	0.0221	0.0589
5-order 5-parameter PRG	251	0.0184	0.0568
4-order 7-parameter PRG	329	0.0169	0.0865
5-order 6-parameter PRG	461	0.00805	0.401

Table 3. Optimization of well center location and length for fracture network at final time.

Input sample space	Case A: 7-D		Case B: 5-D		Case C: 2-D	
	Range	Opt.	Range	Opt.	Range	Opt.
Fracture orientation	0 - 135	121	1	-	1	-
Initial fracture number	50 - 500	486	250	-	250	-
Minimum principal stress σ_h (MPa)	10 - 15	13.02	10 - 15	13.27	12.5	-
Stress anisotropy (σ_H/σ_h)	1.0 - 1.5	1.07	1.0 - 1.5	1.00	1.2	-
Young's modulus (GPa)	5 - 50	38.33	5 - 50	48.65	25	-
Injection well center on y axis (m)	-20 - 20	4.31	-20 - 20	5.09	-20 - 20	5.09
Injection well length (m)	0 - 40	0.08	0 - 40	21.3	0 - 40	21.9
Maximum fractal dimension		1.872		1.622		1.547

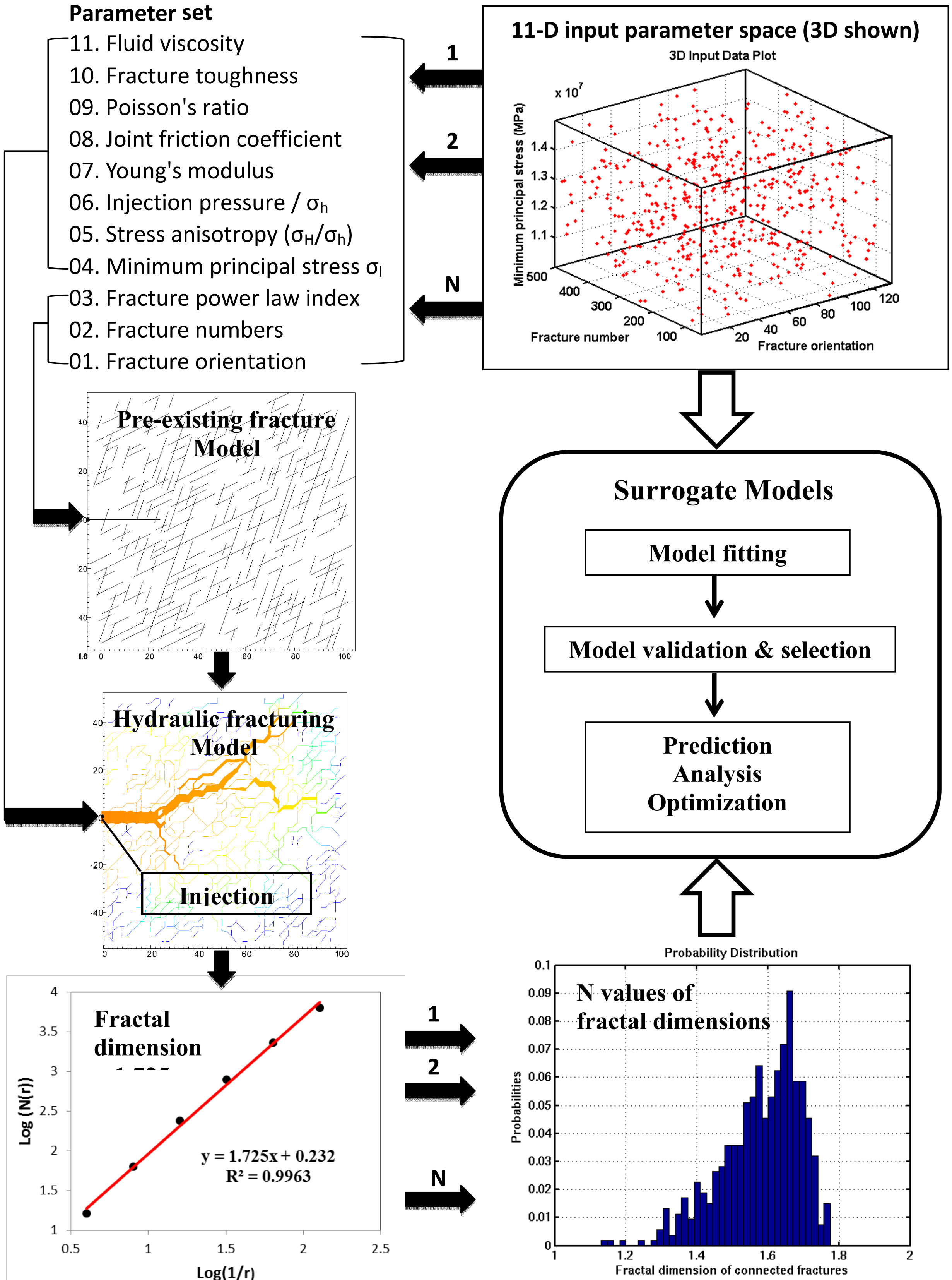

Fig. 1. Surrogate-based modeling approach for simulated hydraulic fracturing

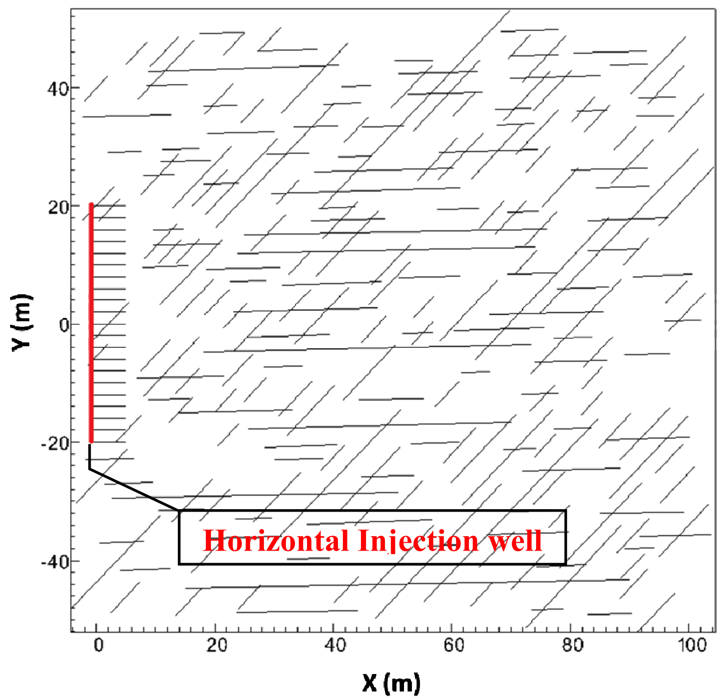

Fig. 2. An example of a horizontal well (center at y = 0 m and length=40 m) placed in a pre-existing network (orientation $=0^{o}$ and number =250). The red solid line is horizontal injection well with uncertain location and length along left y axis. The pre-existing fracture orientation and number of natural network are also uncertain. The maximum and minimum principal stress are assumed x and y direction respectively.

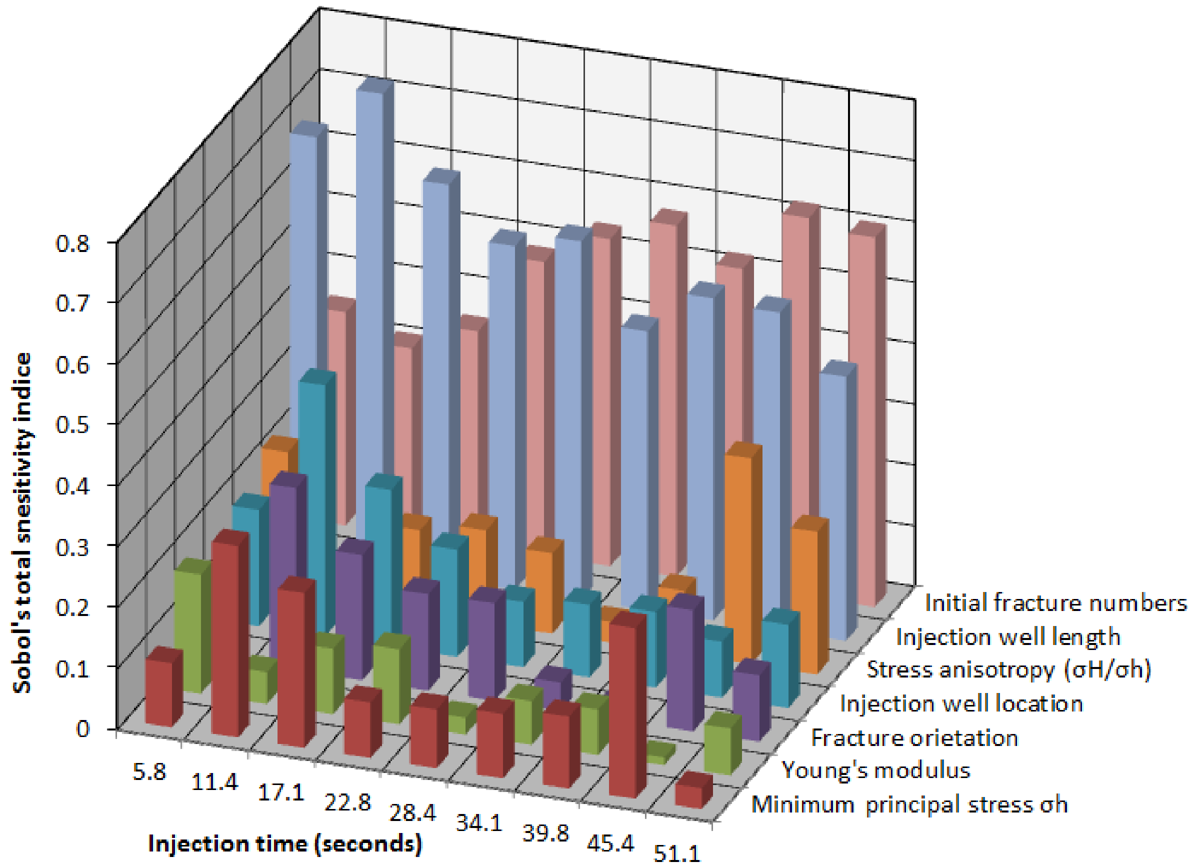

Fig. 3. Global sensitivity of fractal dimension to the 7 input parameters for 9 sequential injection time steps. The 9 parameter sequences are ordered according to the last one.

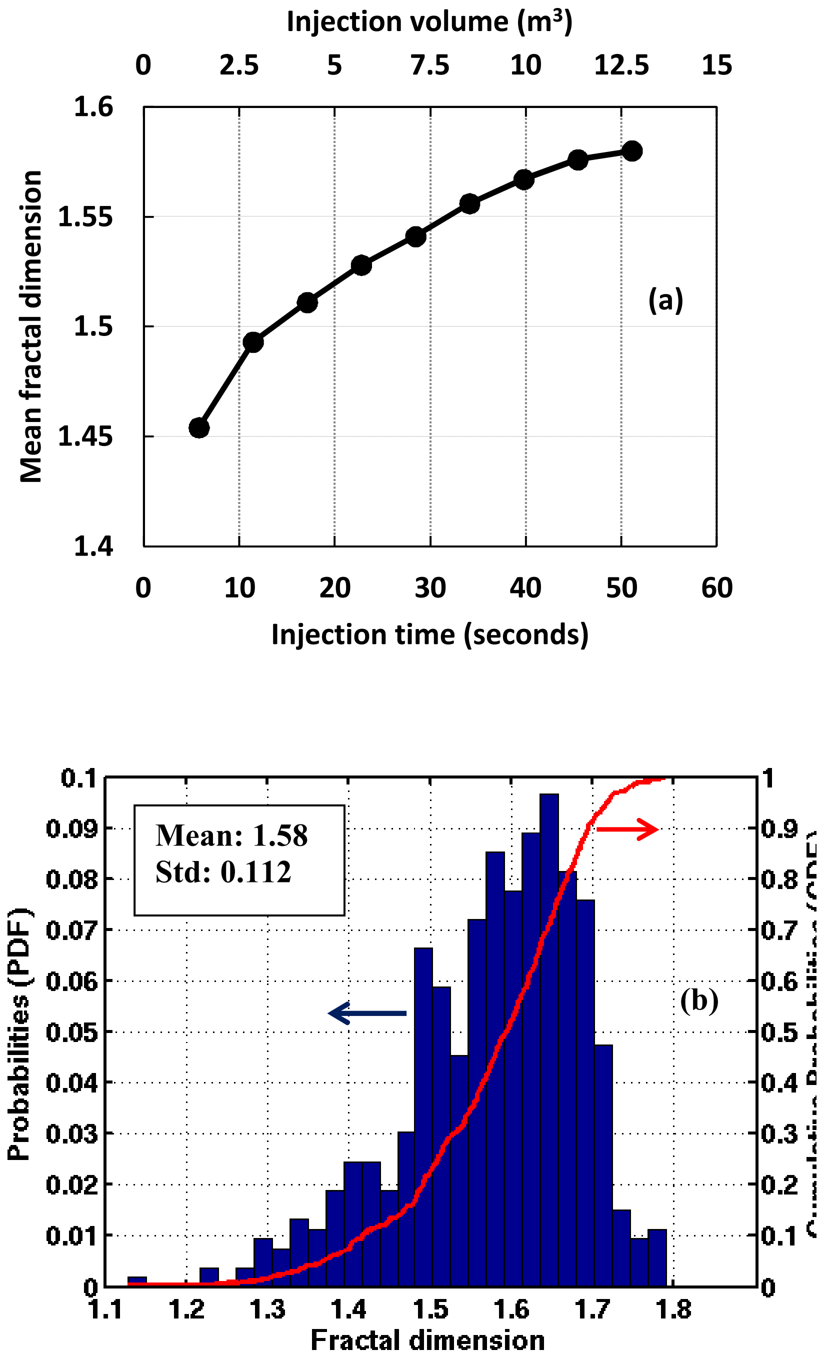

Fig. 4. Statistics of 529 derived fractal dimensions: (a) mean for 9 injection time or volumes, (b) PDF and CDF at finial time(51.1 seconds or 12.8 m^3 injected fluid volume).

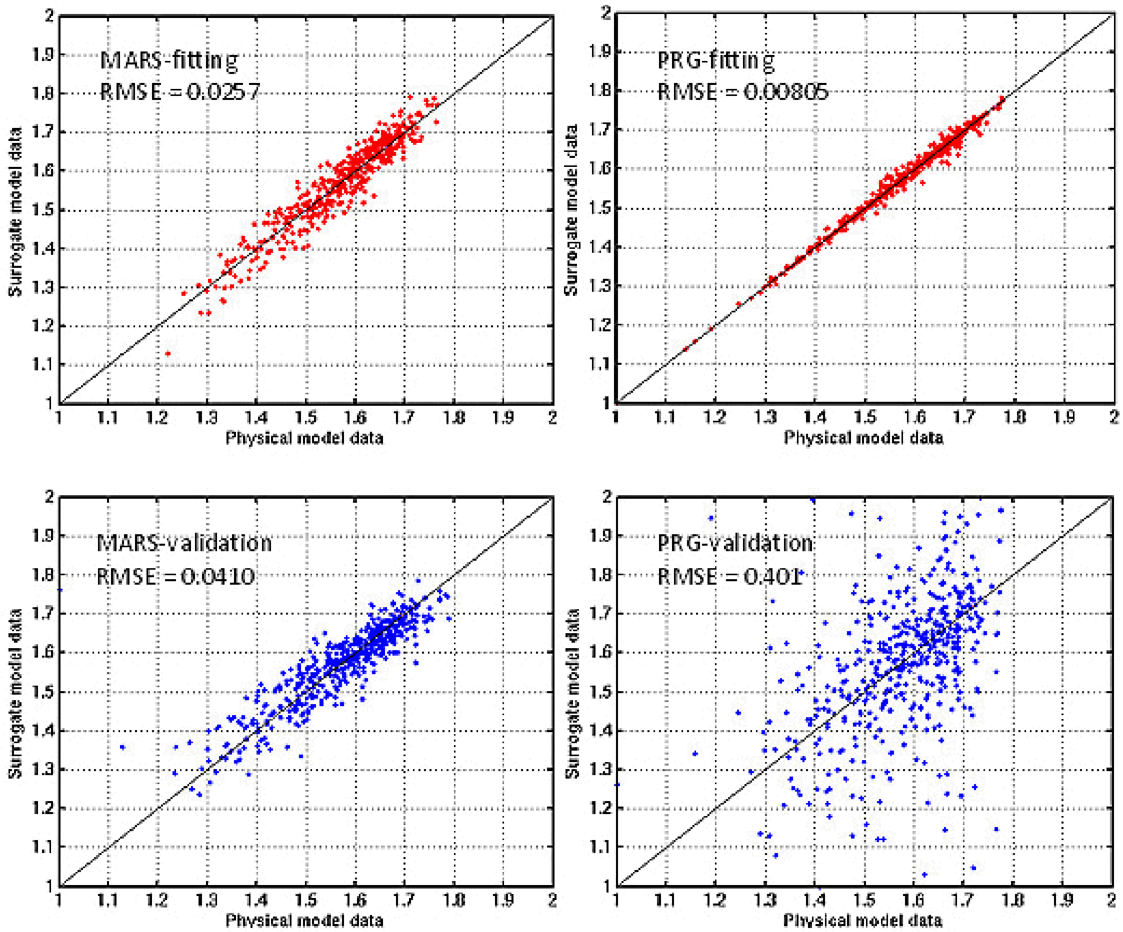

Fig. 5. Scatter plots of fractal dimension simulated using surrogate model data versus physical model from 529 input samples: the comparison of fitting (red dots) and cross-validation (blue dots) between MARS and 5-order 6-parameter PRG surrogate models. The tighter the points clustered along diagonal, the closer the surrogate model data match the physical model data.

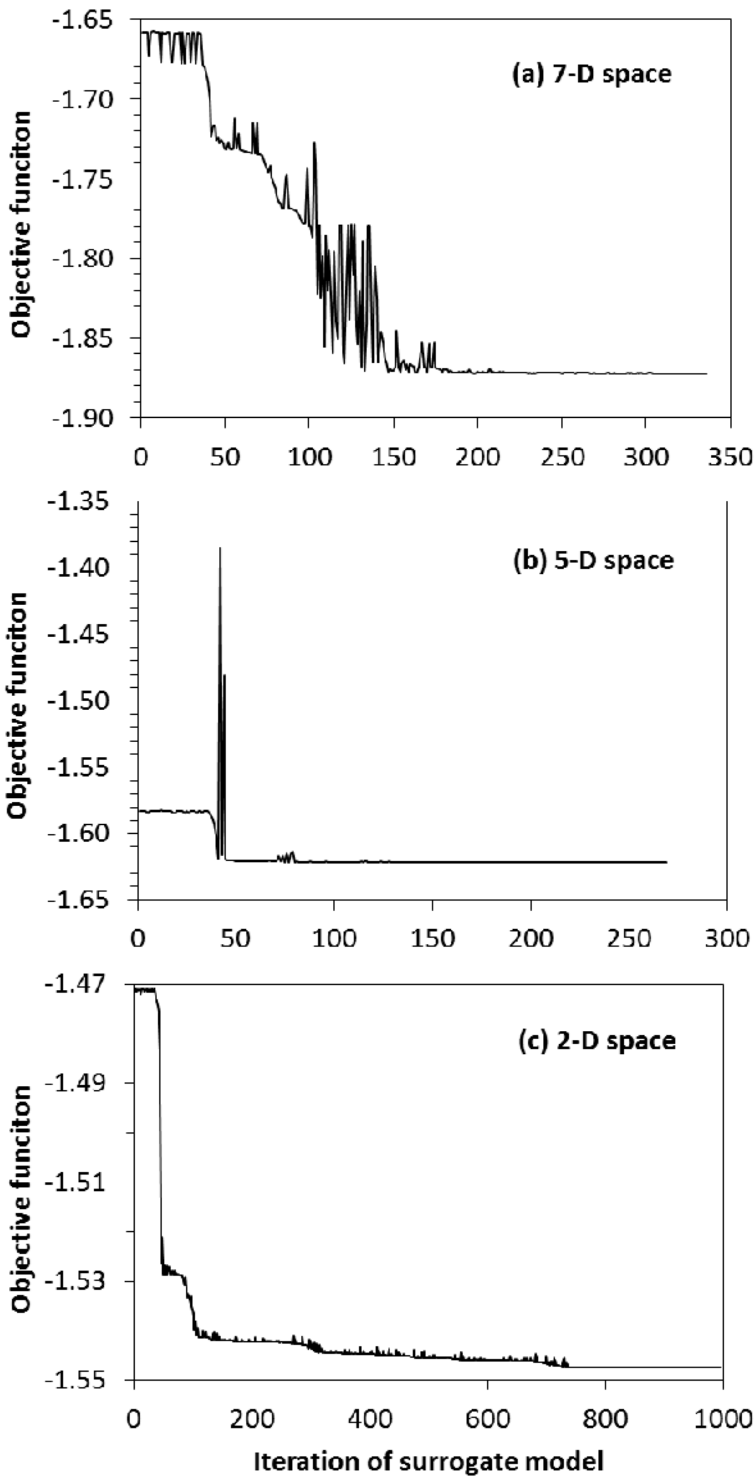

Fig. 6. Minimal objective searching curve for optimization case with (a) 7 uncertain parameters, (b) 5 uncertain parameters, and (c) 2 uncertain parameters. The optimal parameter values for the 3 cases are proved in Table 3.

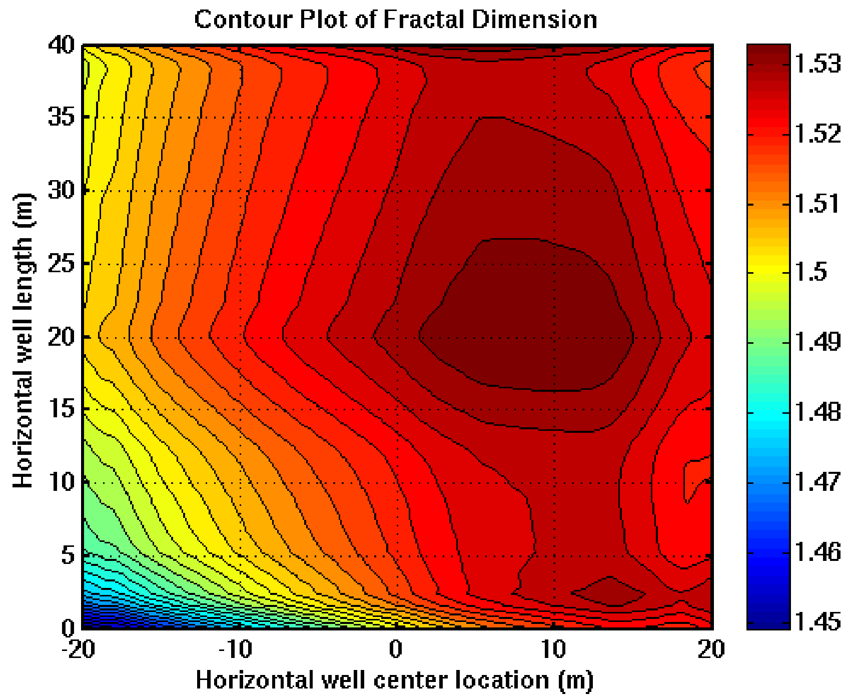

Fig. 7. The visualized response surface with 2 uncertain design parameters, i.e., horizontal well center at y axis and well length.

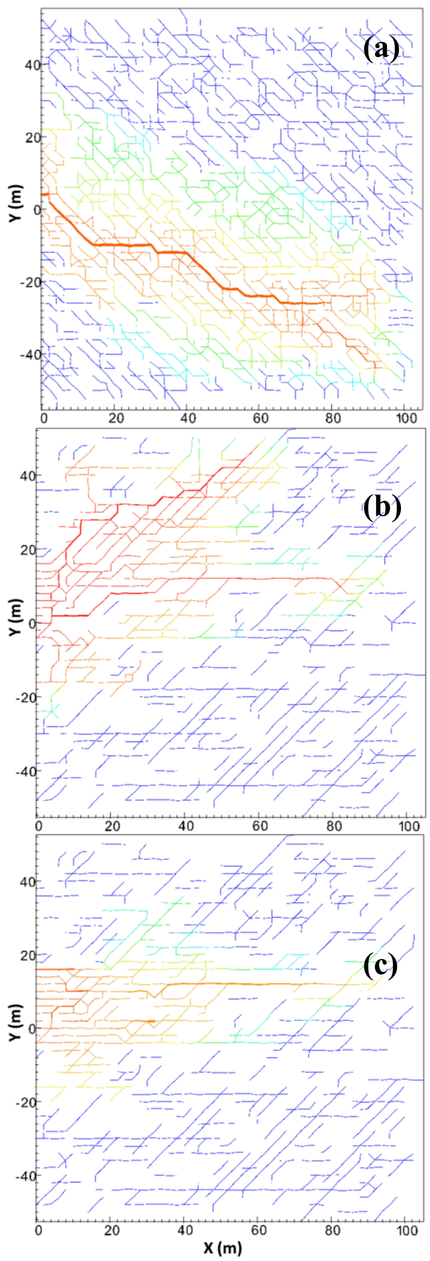

Fig. 8. The post-fracking network corresponding to the optimal parameter set (values provided in Table 3) with (a) 7 uncertain parameters, (b) 5 uncertain parameters, and (c) 2 uncertain parameters. The optimal parameters for the 3 cases are provided in Table 3. The color of the fractures is based on the hydro-pressure. Red indicates the maximum pressure, while blue denotes zero pressure, meaning closed fractures which are not included in fractal dimension calculation.