SPE 57453

Performance Drivers in Restimulation of Gas Storage Wells

Shahab Mohaghegh, Khalid Mohamad, Andrei Popa, and Sam Ameri, West Virginia University and Dan Wood, CNG Transmission

This paper was prepared for presentation at the 1999 SPE Eastern Regional Meeting held in Charleston, WV, 21–22 October 1999.

ABSTRACT

In order to maintain or enhance deliverability of gas storage wells in the Clinton Sand in Northeast Ohio an annual restimulation program has been in place since the late sixties. The program calls for as many as twenty hydraulic fractures and refractures per year. Several wells have been refractured three to four times while there are still wells that have only been fractured once in the past thirty years.

As the program continues many wells will be stimulated a second, third or fourth time. This paper summarizes an attempt to carefully study the response of the Clinton sand to hydraulic fractures and identify the performance drivers in each series of frac jobs. Do the performance drivers remain the same for the later fractures (second, third and fourth frac jobs) as they were in the first ones? Or do they change? This paper attempts to answer such questions.

Identification of major performance drivers becomes important when new jobs are to be designed. They not only play an important role in enhancing the response of the wells to new stimulation jobs; but also may prove to be an important economic factor in the design of new stimulation procedures. If for instance it is concluded that an increase in proppant volume does not influence the stimulation outcome after the second refracture, then fewer resources can be used for proppant volume and be directed toward parameters that are more influential.

This study employs a combined neural network and fuzzy logic tool to identify the performance drivers.

INTRODUCTION

In many industrial and manufacturing processes it is important to know the role and influence of each component or parameter in the process outcome. Oil field operations are no exception. Such information can contribute significantly to process efficiency and prevent wasteful usage of the resources. If the engineer knows in advance which component is the main driver of the process performance, she/he can concentrate her/his efforts on manipulating that component to achieve the desired outcome. On the other hand lack of such information can result in wasting resources by using more of a component that does not make a significant difference in the process outcome and therefore increase the cost and reduce the efficiency of the process.

The conventional method of approaching this problem is to build an accurate mathematical model of the process and perform parametric sensitivity analysis on the model. Authors believe this still is the best approach and whenever possible all the efforts should be concentrated on developing such models, or models approaching the accuracy of a mathematical model, and perform detailed analysis on the model. We will later discuss the efficiency and accuracy of this approach in comparison of other methodologies.

The reality of many complex processes including some in the petroleum and natural gas industry is that there are no known mathematical models that can accurately describe these processes. The problem being discussed in this paper is one such problem. Restimulation of gas storage wells is a convoluted and complex problem that can not be mathematically modeled for many reasons. The most important reason for the inability of constructing a mathematical model for restimulation of gas storage wells is the lack of detailed reservoir data and the complexity of modeling a stimulated reservoir and its response to restimulation. As it has been shown in the past[1-3], the next best thing to mathematical or numerical modeling of a complex problem is using virtual intelligence techniques (neural networks, fuzzy logic, and genetic algorithms) to approximate the process behavior.

In analyzing petroleum and natural gas engineering problems, when a mathematical model of a complex process can not be constructed, other means have been used to identify the most influential parameters in such processes. These approaches can be as simplistic as statistical methods such as linear regression analysis to more complex analysis such as fuzzy curves[4] and neural networks[5]. In this paper we apply all these methods to the problem of restimulation of gas storage wells and discuss their potential shortcomings. We would also discuss two important issues and provide some ideas that might shed some light on the problem in hand. The first issue concerns the use of fuzzy curves in identifying the most important parameters and provides an extension that might help to further solidify their contribution to problems such as the one being discussed here. The second issue is concerned with the use of neural networks to identify important parameters in a complex problem. Here we apply a method that has never been used in the past to address a petroleum and natural gas engineering problem and show its usefulness and value in providing important information about the process.

STATEMENT OF THE PROBLEM

The restimulation of gas storage wells being discussed here takes place in the Clinton Sand in Northeast Ohio. A hydraulic fracturing and refracturing program has been in place in this field for about three decades. There are storage wells in this field that have not been hydraulically fractured as well as storage wells that have been re-fractured more than four times in the past 25 years. Every year several wells are selected for stimulation in order to maintain deliverability of the field. Identification of the most influential parameters in this process is an important part of designing new fractures and refractures in order to maximize return on investment.

DATA SET

The data set used in this study was collected using the well files that included the design of the hydraulic fractures. The following parameters were extracted from the well files for each hydraulic fracture treatment: year the well was drilled, total number of fractures performed on the well, number of years since the last fracture, fracture fluid, amount of fluid, amount of sand used as proppant, sand concentration, acid volume, nitrogen volume, average pumping rate, and the service company performing the job. The goal of the study is to identify the most important and most influential of the above mentioned parameters when they are correlated with post fracture deliverability. The match up between hydraulic fracture design parameters and the available post fracture deliverability data produces a data set with approximately 560 records.

STATISTICAL ANALYSIS

Using straightforward regression analysis, that can be performed using any widely available spreadsheet software, on the data set shows that some trends may quickly be identified on these fracture jobs. Table 1 shows the ranking of the correlation between each parameter and the post fracture deliverability. The ranking is based on the calculated R squared values.

It is notable that all the R squared values are very low with highest being less than 0.2, yet from a comparative point of view one may argue that since one parameter exhibits a slightly higher correlation thus it may be more influential. Simple regression analysis identifies sand volume, fluid volume and average rate as the top three influential parameters.

Statistical Analysis		
Fracture Parameter	**R Squared**	**Ranking**
Sand Volume	0.1756	1
Fluid Volume	0.1670	2
Average Rate	0.1092	3
Number of Fracs	0.0893	4
Frac Fluid	0.0504	5
Years Before	0.0266	6
Year Drilled	0.0158	7
Acid Volume	0.0049	8
Nitrogen Volume	0.0037	9
Sand Concentration	0.0002	10
Service Company	0.0001	11

Table 1. Most influential parameters based on statistical analysis.

Figures 1, 2 and 3 show the data for the first three parameters to give reader a visual representation of the existing correlation in this data set. Several parameters in the data set have discrete values and it is obvious that they would provide very poor correlation using regression-based approaches. The year the well was drilled, total number of fractures performed on the well, number of years since the last fracture, fracture fluid, and the service company performing the job are such parameters.

FUZZY CURVES

Lin and Coningham[6] introduced use of fuzzy curves to identify important inputs. Fuzzy curves were later applied to petroleum engineering problems[4]. In applying fuzzy curves to a problem to identify the most important parametr, one should plot a fuzzyfied version of all the data of each parameter versus the output (post fracture deliverability in our case). The parameter that has the highest range is identified as the most important one. It is interesting to note that when applied to this problem, the fuzzy curve approach identified the same first three parameters as the linear regression analysis. It also should be noted that this phenomenon has been observed in other cases as well. Table 2 demonstrates the ranking of the parameters resulted from fuzzy curve analysis of the data in our data set.

Fuzzy Curve Analysis		
Fracture Parameter	**Range**	**Ranking**
Sand Volume	2812.40	1
Fluid Volume	2651.87	2
Average Rate	1440.65	3
Sand Concentration	1007.40	4
Nitrogen Volume	853.77	5
Years Before	825.84	6
Number of Fracs	775.93	7
Year Drilled	580.51	8
Frac Fluid	327.36	9
Service Company	210.38	10
Acid Volume	172.38	11

Table 2. Most influential parameters based on fuzzy curve analysis.

When applying any kind of technique to identify parameters' importance to the process outcome the overwhelming tendency is to perform the analysis on one parameter at a time. This is a direct result of unintentional linearization of complex problems in our mind. We have the tendency to look for easy answers. It would be nice to have that single most important parameter that we could control the entire process with. But this is in many cases, and this problem is one of them, an oversimplification of a complex problem. That is the reason that in the previous section we mentioned that building a model of the process and carefully studying the model as a whole is the best way of addressing such issues.

In order to get one step closer to the reality of parameters' role in the hydraulic fracturing of the Clinton sand we attempted to use two parameters at a time and examine the parameters' role in binary combinations. The process was conducted for all 55 possible combinations of two of the parameters. Table 3 shows the top 10 binary combination of parameters that have the most influence on the post frac deliverability in the Clinton sand. Figures 4, 5, and 6 show the plot of these parameters.

Fuzzy Curve Analysis- Binary Combination	
Fracture Parameter	**Ranking**
Sand Volume & Number of Fracs	1
Sand Volume & Fluid Volume	2
Sand Volume & Sand Concentration	3
Sand Volume & Nitrogen Volume	4
Sand Volume & Years Before	5
Sand Volume & Average Rate	6
Sand Volume & Year Drilled	7
Sand Volume & Frac Fluid	8
Fluid Volume & Number of Fracs	9
Fluid Volume & Years Before	10
Years Before & Nitrogen Volume	11

Table 3. Most influential binary combination of parameters based on an extension of fuzzy curve analysis.

Since sand volume was identified as the single most important parameter (Table 2), it makes sense that it dominates the top eight in the binary combination list as well. But it is interesting to note that number of fractures that was ranked seven is paired with sand volume for the number one spot in the list. This simply means that although sand volume and fluid volume are ranked number 1 and 2 in the ranking of single parameters their combination does not necessarily rank as number 1 in binary combination. On the other hand the parameter that was ranked 7 in the single parameter ranking moves up to number 1 in combination with sand volume. This observation underlines the nonlinear nature of this process and adds credibility to the fuzzy curve approach as a method that can identify nonlinear relationship between parameters. It demonstrates that although a parameter may not rank high by itself in the process, its importance to the process may depend on the way it is combined with other parameters. Here we examined binary combinations of parameters. The next step in this study can include combining three, four or more of the parameters and studying their influence on post fracture deliverability.

NEURAL NETWORKS

Among the three techniques being demonstrated in this study (regression, fuzzy curves, and neural nets) neural networks, when used properly, have the best chance of discovering nonlinear relationships between parameters. Linear regression by definition is a linear approach. It seems that fuzzy curves are not utilizing all the potentials of fuzzy logic and can be improved. Fuzzy curves seem capable of detecting some nonlinear behavior when used on binary combinations (and may be other combinations) of the parameters. It must be noted here that fuzzy logic can be used as universal function estimator and have the capability of approximating highly nonlinear functions.

Neural networks are also universal function estimators. The architecture of neural networks makes them capable of discovering highly nonlinear relationships between parameters in a data set. In this study we use a technique that has not been used in the petroleum and natural gas industry before. This technique is called backward elimination. The essence of this technique is to use the capabilities of neural networks to identify most influential parameter or parameters in a data set. This is a long and tedious process and requires a fair amount of familiarity with the problem in hand as well as the essence of neural networks.

The process starts with using all the parameters as input to a network and developing a neural model of the process using a supervised learning algorithm. The learning is stopped at a certain point and the correlation between input and output is calculated using the R squared method. Then one at a time the inputs are removed from the data set and the exact same process is repeated. This means that in case of this study the original neural network that has all the inputs is consists of eleven inputs and one output. The consequent networks have ten inputs and one output. There are a total of twelve neural networks trained.

When one of the input parameters is removed and a new network is trained the R squared of this network is compared with the R squared of the original network. A decrease in the R squared should be directly related to the importance of that input parameter to the process since its removal has inversely affected the overall correlation coefficient of the system. The larger the difference between the two R squared values the higher the importance of the input parameter. Figure 7 shows the R squared values of eleven networks as compared to the original network. The name on the x-axis identifies the name of the parameter that was removed during the construction and training of that network. Table 4 shows the list of the parameters and their ranking based on the neural network analysis.

Unlike linear regression and fuzzy curve, neural network analysis identifies the average rate as the most important parameter. Sand volume that shows up as the most important parameter in the past analysis is ranked number three. Another important finding of the neural network is the influence of the service company performing the hydraulic fracture on the job outcome.

Neural Network Analysis	
Fracture Parameter	**Ranking**
Average Rate	1
Service Company	2
Sand Volume	3
Sand Concentration	4
Number of Fracs	5
Acid Volume	6
Years Before	7
Frac Fluid	8
Nitrogen Volume	9
Fluid Volume	10
Year Drilled	11

Table 4. Most influential parameters based on neural network analysis.

This makes a lot of sense especially on the later jobs (on the same well) where the stimulation procedure becomes more sensitive to engineering design and know-how. The wells that have already been stimulated might be more vulnerable to imperfect execution of stimulation than those that are being stimulated for the first time. In other words there is only so much space for making mistakes.

As was mentioned before some of the wells in the field have been stimulated once while some have been stimulated up to four times. It seems only logical that as the number of stimulation treatments increases, a well's response to the stimulation is altered. Therefore as a whole, wells that are being stimulated for the first time should react different to the stimulation practices when compared to wells that are being stimulated for the second or third time. During the binary combination fuzzy curve analysis, the number of fractures in combination with sand volume has been ranked as the most important parameter. Number of fractures has ranked number 4 in the statistical analysis and number 5 in the neural network analysis ahead of all fluid related parameters. This makes perfect engineering sense.

Therefore in the next step in our neural network analysis we decided to divide the data set into three separate data sets based on the number of the hydraulic fractures. The first data set consists of all the wells that have been fractured at least once and includes the fracture designs. The second data set includes all the wells that have been fractured at least two times and includes only the second fracture designs, and so on. This means that a particular well that has been fractured three times will appear in all three data sets. This allows us to focus the neural model building efforts on the most relevant data. The goal is to see if the influence of parameters changes as the number of hydraulic fractures increases in a well. For example the most important parameter in the first hydraulic fracture may not necessarily be the most important parameter in the second and third fractures. Identification of such changes in the parameter's influence can prove valuable in the design of new fracture jobs.

The neural network backward elimination process was implemented for this part of the analysis and the results are shown Table 5. Figures 8, 9 and 10 are visual representations of these results.
The above analysis indicates that the influence of the sand volume decreases as the number of fractures increase. On the other hand as the number of fractures treatments increases it becomes more important which service company is performing the job. It may be due to the expertise and quality control procedures that are in place for different companies. Fluid volume has ranked eight in all three procedures. Number of years from the last fracture becomes the most important parameter in the second fractures and is the

third most important parameter in the third fractures. This table can assist the engineers in many ways during the design of the new fractures for each well.

A FINAL NOTE

An important issue that needs to be discussed here and was briefly mentioned in the previous sections is the interdependency of parameters to one another. In all three techniques that are presented here as well as in the literature on this topic the objective is to identify one or two parameters that control the entire process. Lotfi A. Zadeh, the father of fuzzy logic, has been quoted to say that "as complexity rises, precise statements lose meaning and meaningful statements lose precision". It seems that trying to achieve such objectives (identification of most important parameters in a process) becomes increasingly imprecise as the process becomes more and more complex.

The restimulation of gas storage wells is a very complex problem. Each parameter's importance is related to the other parameters' role in a complex and nonlinear fashion. Attempts to identify the most important performance driver in such a complex system may be an oversimplification of the problem. It must be noted that by using a data set as the basis of our analysis, and not looking at every case in an independent manner (which is not practical) we are averaging the impacts of all the parameters throughout the time and space. Therefore, we need to qualify our finding by the word average. This is identification of the performance drivers "on the average".

The authors believe that the best way of approaching this problem is to attempt to build an accurate and representative model of the process and then query the model on a case by case basis to identify the important performance drivers for each design. Therefore, the attempt made in the last portion of the previous section to divide the database to smaller subsets and study them separately is a step in the right direction in deconvoluting this complex problem.

CONCLUSIONS

Three different techniques for the identification of performance drivers were presented. These techniques were applied to restimulation of gas storage wells in the Clinton sand. Statistical analysis (regression) identified the sand volume to be the most important parameter. Results provided using the fuzzy curve technique were quite similar to that of linear regression. The fuzzy curve technique was extended to analyze binary combinations of parameters. This approach revealed that the fuzzy curve technique is capable of extracting nonlinear information from the database.

Neural networks were the third technique used in this paper. Backward elimination showed that average rate is the most important parameter and emphasized the role of the service company in the success of the fracture jobs. The neural network backward elimination process was applied to first, second, and third fractures separately. It revealed that as the number of fractures performed on a particular well increases the importance of sand volume decreases and influence of service company increase.

REFERENCES

1. Mohaghegh, S., Arefi, R., and Ameri S.: "Virtual Measurement of Heterogeneous Formation Permeability Using Geophysical

Well Log Responses", *The Log Analyst*, Society of Professional Well Log Analysts, March-April 1996, pp. 32-39.

2. Mohaghegh, S., Balan, B., McVey, D., and Ameri, S.: "A Hybrid Neuro-Genetic Approach to Hydraulic Fracture Treatment Design and Optimization", *SPE 36602, Proceedings*, SPE 71st Annual Technical Conference, October 6-9, 1996, Denver, Colorado.

3. Mohaghegh, S., Richardson, M., and Ameri, S.: "Virtual Magnetic Imaging Logs: Generation of Synthetic MRI Logs from Conventional Well Logs ", *SPE 51075, Proceedings*, SPE Eastern Regional Conference and Exhibition, November 9-11, 1998, Pittsburgh, Pennsylvania.

4. Ali, M., and Chawathe, A.: "Using Artificial Intelligence to Predict Permeability from Petrographic Data", in press.

5. Mohaghegh, S., Mc Vey, D., and Ameri, S.: "Predicting Well Stimulation Results in a Gas Storage Field in the Absence of Reservoir Data, Using Neural Networks", *SPE Reservoir Engineering Journal*, November 1996, pp. 54-57.

6. Lin, Y., and Coningham G.: "A Fuzzy Approach to Input Variable Identification", proceedings of the third IEEE International Conference on Fuzzy Systems, June 26 - July 2, 1994, Lake Buena Vista, Florida.

Neural Network Analysis			
	Fracture Parameter		
Ranking	**First Fracture**	**Second Fracture**	**Third Fracture**
1	Sand Volume	Years Before	Service Company
2	Nitrogen Volume	Sand Concentration	Fracture Fluid
3	Year Drilled	Sand Volume	Years Before
4	Sand Concentration	Service Company	Average Rate
5	Service Company	Average Rate	Sand Volume
6	Fracture Fluid	Nitrogen Volume	Year Drilled
7	Acid Volume	Fracture Fluid	Acid Volume
8	Fluid Volume	Fluid Volume	Fluid Volume
9	Average Rate	Acid Volume	Sand Concentration
10		Year Drilled	Nitrogen Volume

Table 5. Most influential parameters based on neural network analysis.

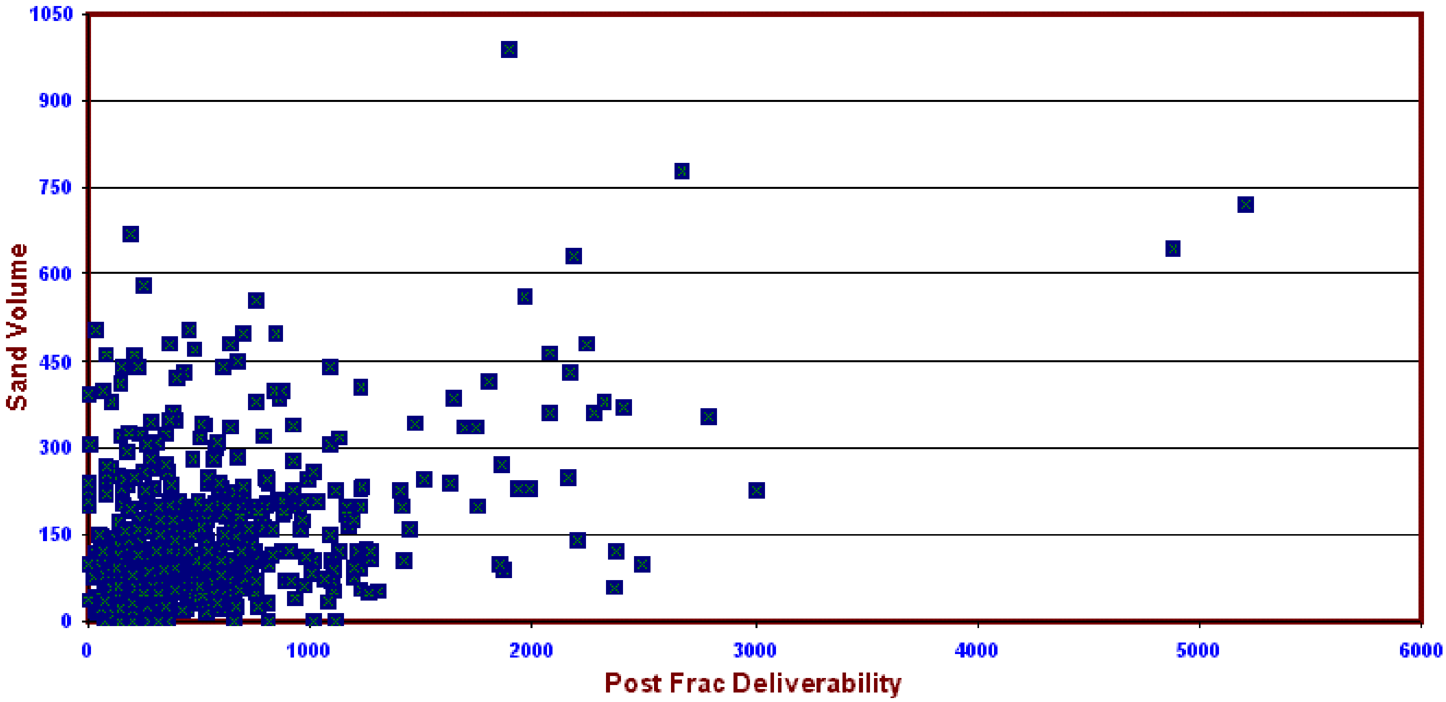

Figure 1. Sand volume versus post fracture deliverability in the data set. Raw data.

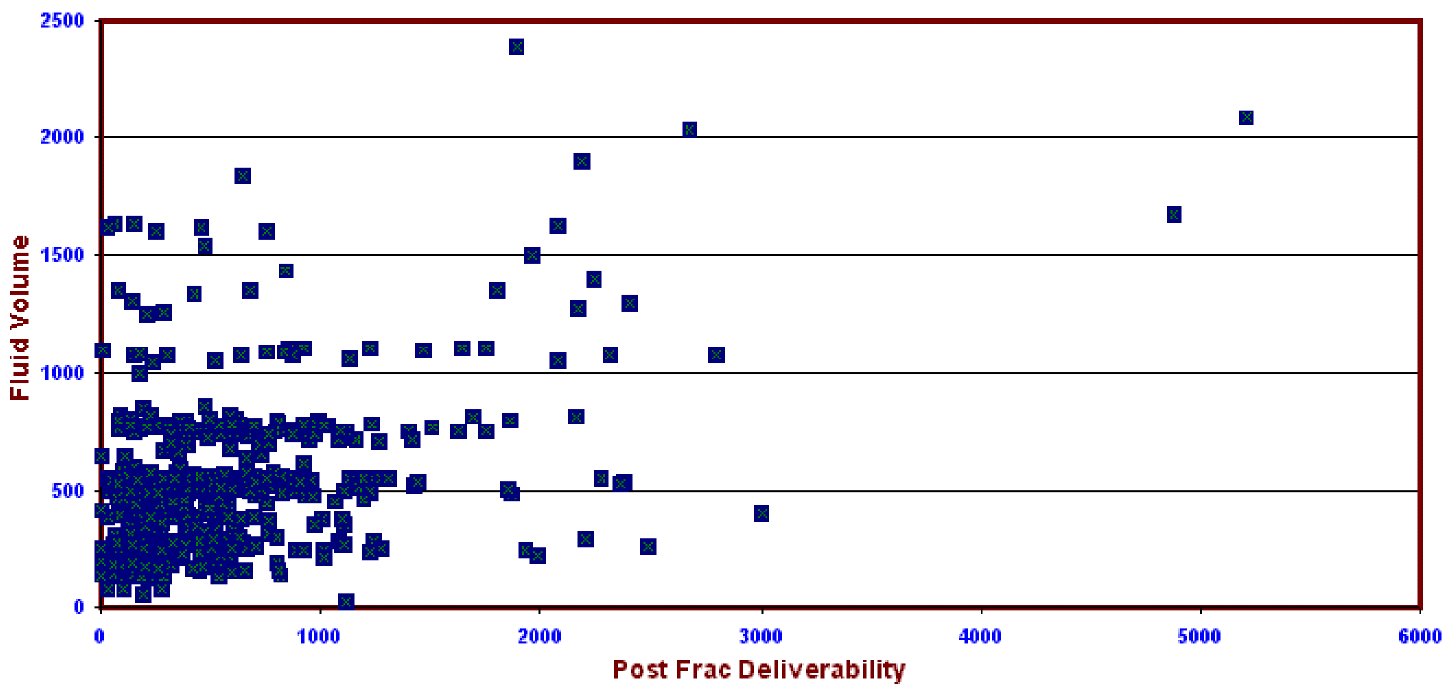

Figure 2. Fluid volume versus post fracture deliverability in the data set. Raw data.

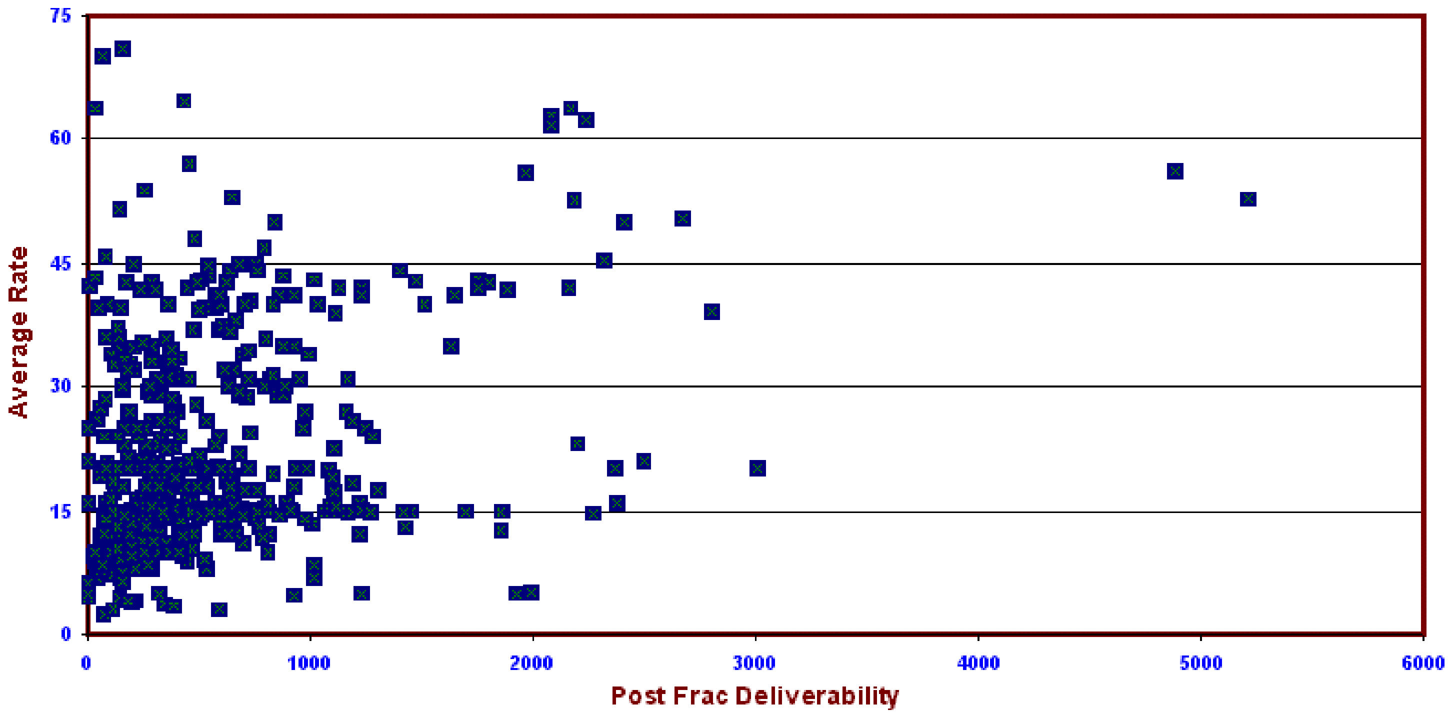

Figure 3. Average rate versus post fracture deliverability in the data set. Raw data.

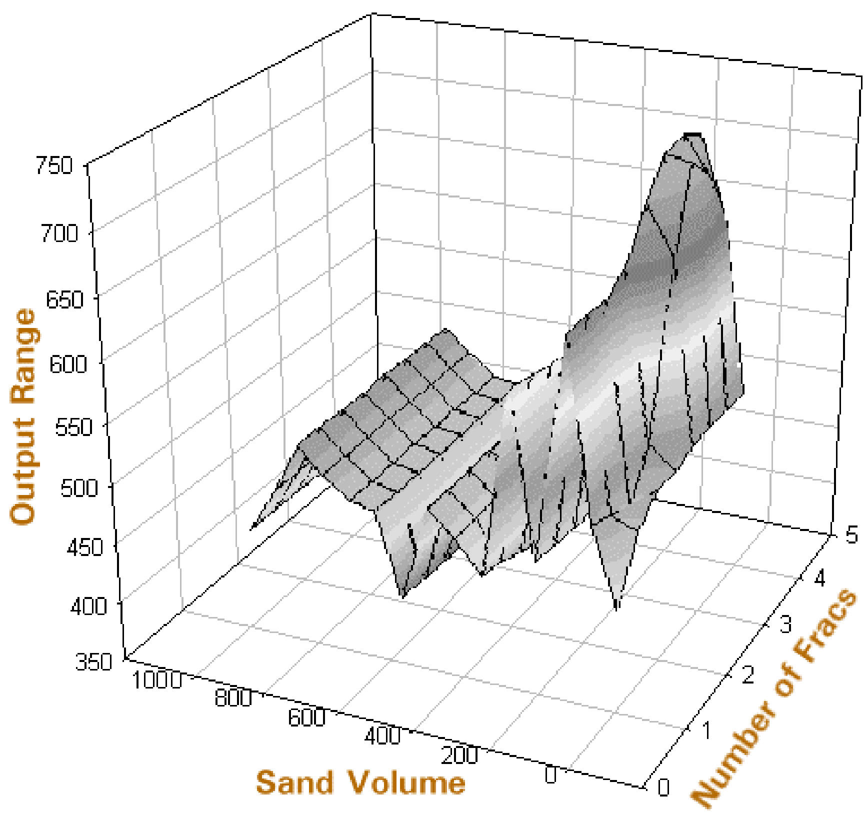

Figure 4. Fuzzy curve for binasy combination of sand volume and number of fractures.

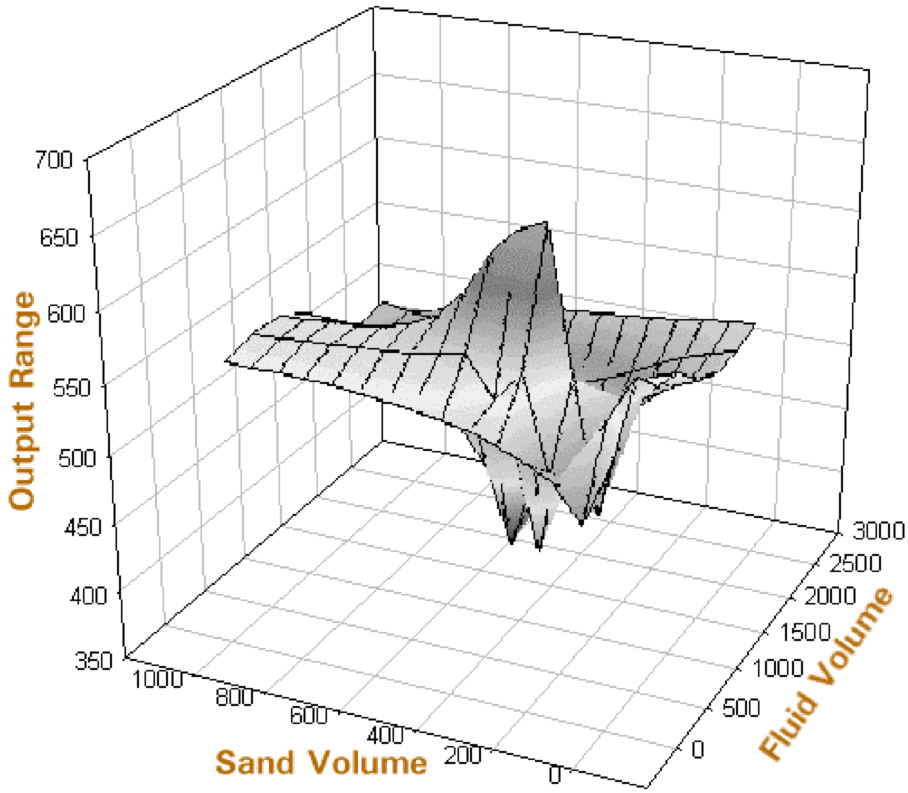

Figure 5. Fuzzy curve for binasy combination of sand volume and fluid volume.

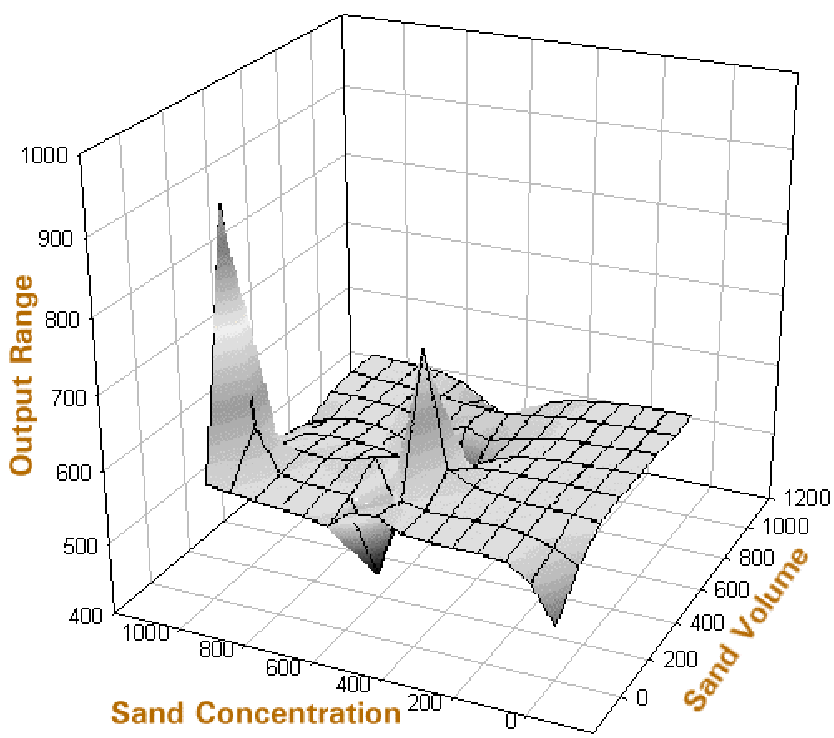

Figure 6. Fuzzy curve for binasy combination of sand volume and sand concentration.

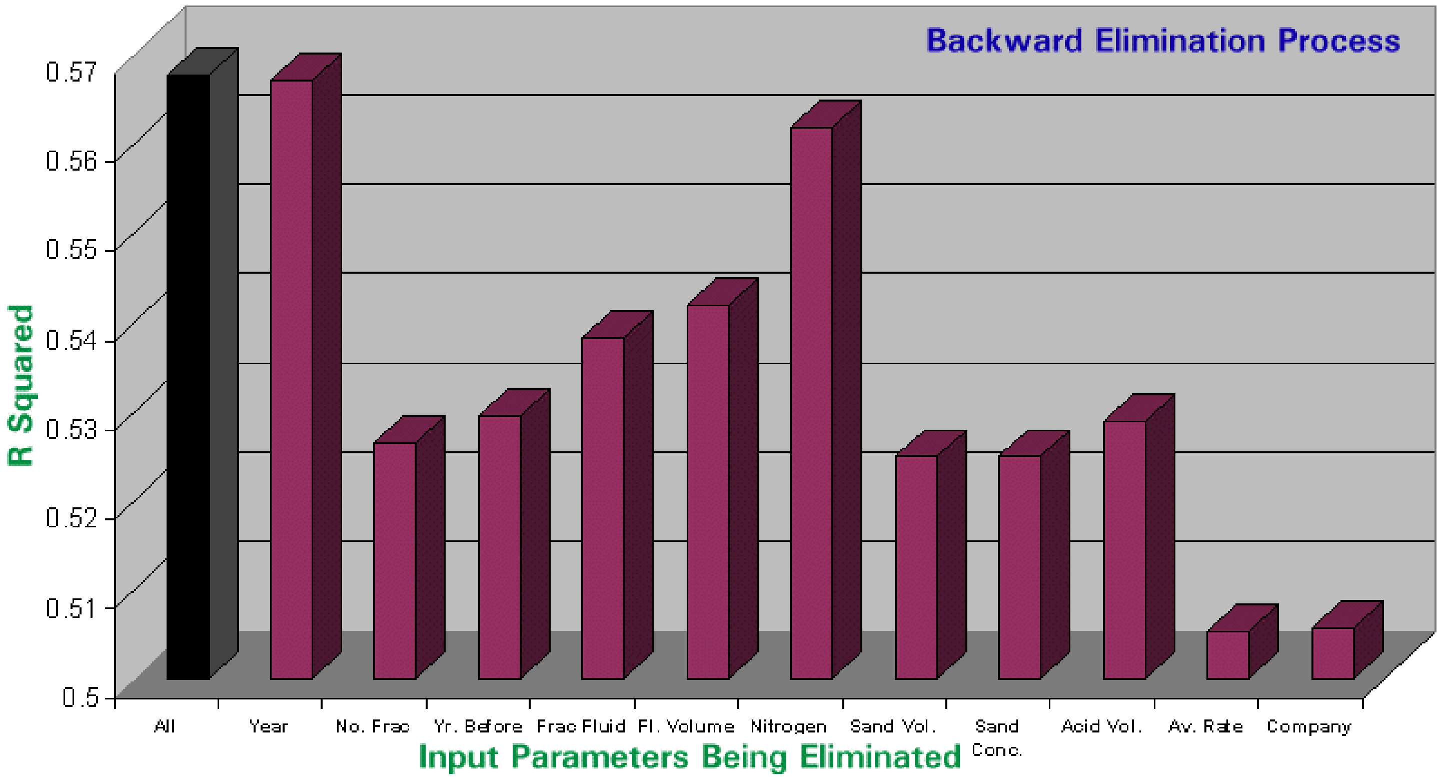

Figure 7. Neural network backward elimination analysis on the entire data set.

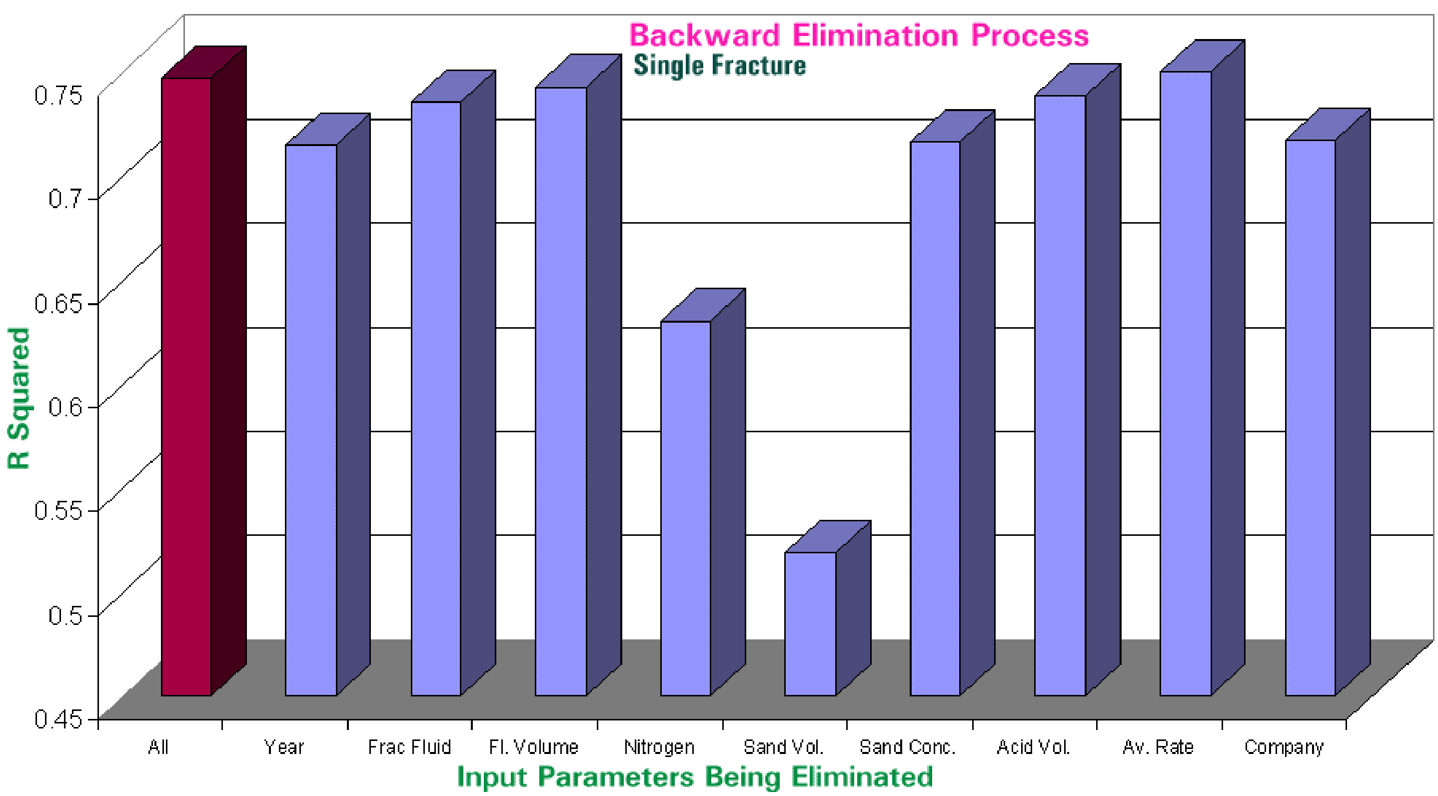

Figure 8. Neural network backward elimination analysis for first fractures.

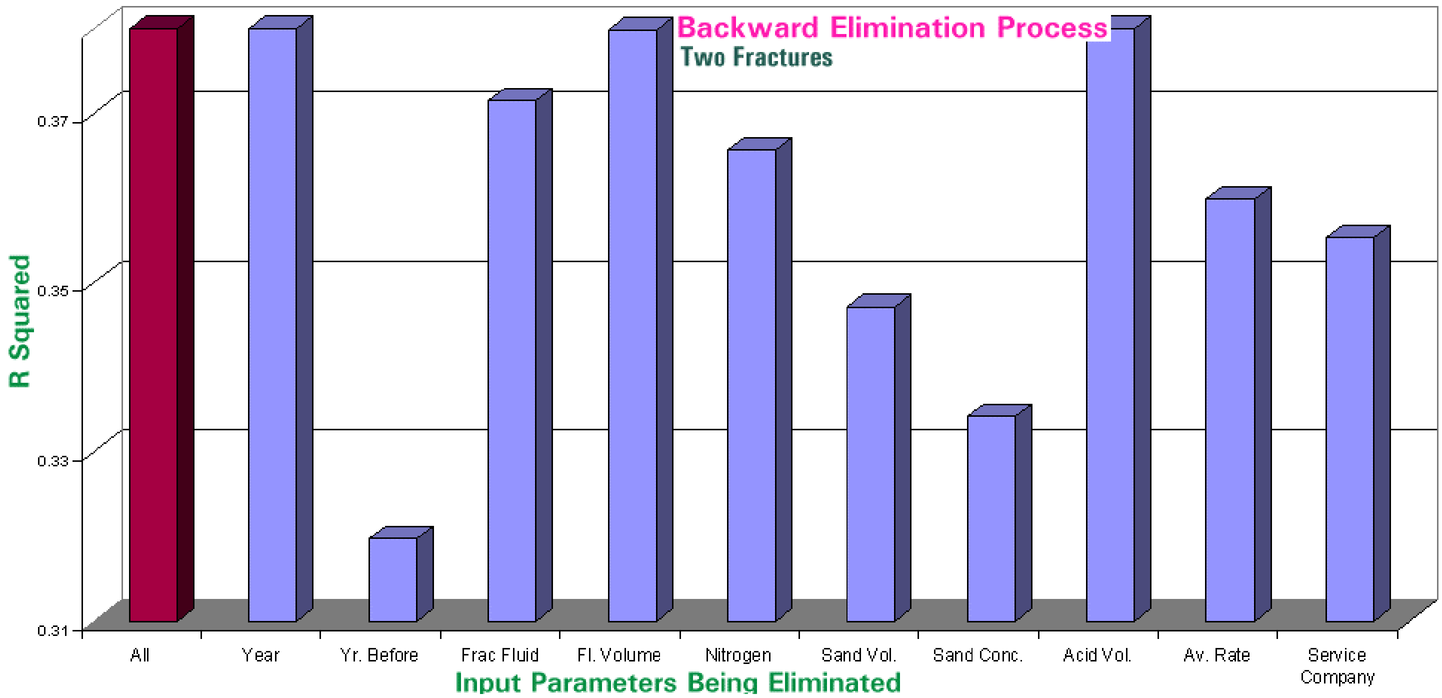

Figure 9. Neural network backward elimination analysis for second fractures.

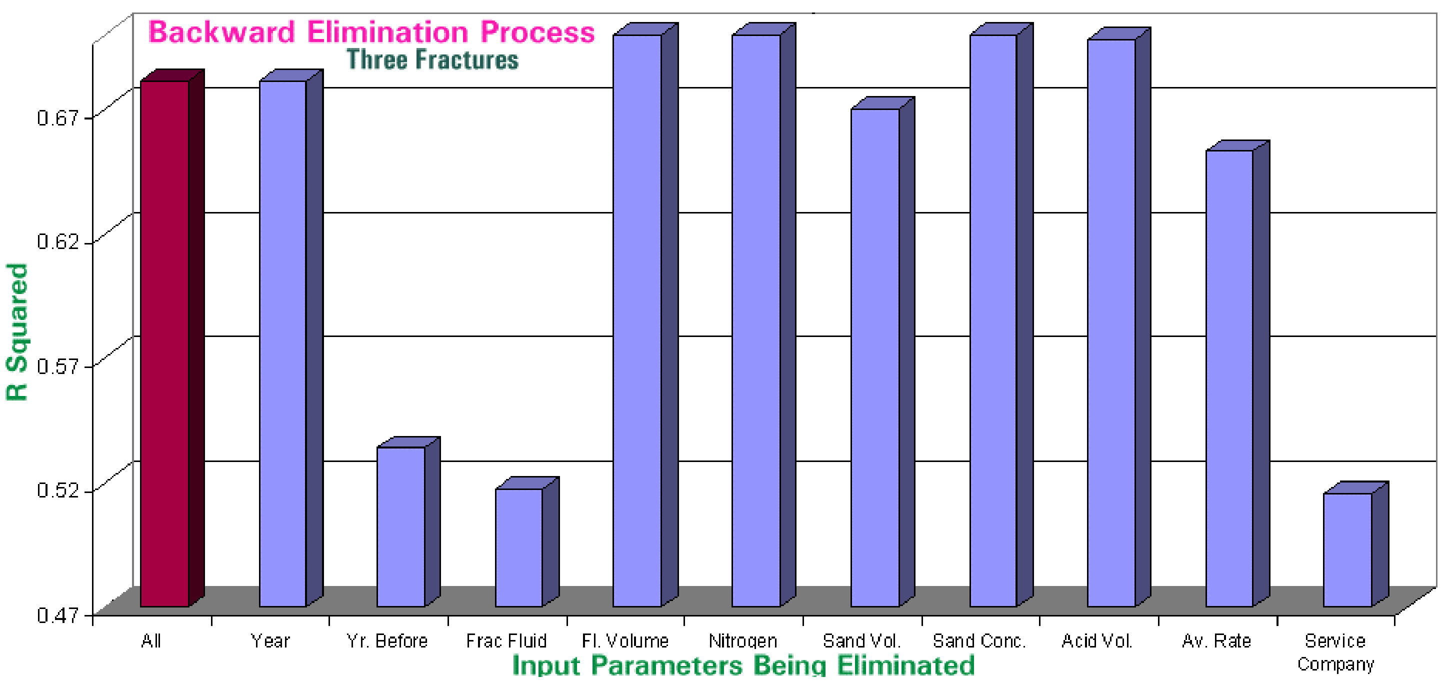

Figure 10. Neural network backward elimination analysis for third fractures.