Journal of Petroleum Science and Engineering 182 (2019) 106280

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: www.elsevier.com/locate/petrol

Meta-analysis of propped fracture conductivity data—fixed effects regression modeling

Rashnur M. Rahman, Obadare O. Awoleke*, Scott D. Goddard

University of Alaska, Fairbanks, USA

ARTICLE INFO

Keywords:
Fracture conductivity
Statistical analysis of pooled data
Regression modeling

ABSTRACT

Previous empirical models of propped fracture conductivity are based either on data sourced from single investigations or on data not in the public domain. In this work, statistically rigorous models of propped fracture conductivity are developed using a database of fracture conductivity experiments reported in technical literature over the last 40 years.

The database contains about 2700 data points. Propped fracture conductivity is the dependent variable and proppant type, mesh size, proppant concentration, formation hardness, closure stress, formation temperature, and polymer concentration are the independent variables. The full, original database was pared down to various subsets, each having complete information in relation to some subset of the independent variables. The number of independent variables included in each of the resulting databases varied from three to six. Seventy percent of the data was used to develop the models while thirty percent of the data was used to validate them.

Fixed effect models were selected using all subsets regression analysis in three, four, and five experimental factors for each of two types of proppant. The five factor model appeared to be the most useful model for both proppant types from a predictive point of view. Five-factor model predictions also compared favorably with an existing propped fracture conductivity model and different case histories published in literature.

This project provides engineers with access to a propped fracture conductivity database based on experiments reported over the past 40 years in technical literature. The models developed based on this database can be used to generate predictions of propped fracture conductivity for a variety of proppant characteristics and formation conditions. Also, the models presented here are based on data from experimental investigations in different laboratories, potentially reducing biases that may be present in single laboratory investigations.

1. Introduction

Propped fracture conductivity is one of the important variables required for the accurate estimation of productivity from hydraulically fractured completions. Therefore, it is important that engineers can predict it in the laboratory by simply varying some highly accessible proxies, such as formation and proppant characteristics, and by relating these proxies to fracture conductivity. These laboratory fracture conductivity values can thereafter be used as benchmark estimates for field fracture conductivity values.

Although a large number of laboratory conductivity tests have been conducted over the last 45 years to understand the effects of various factors on fracture conductivity and to improve laboratory predictions, to date no research has pooled this publicly available data and analyzed it in aggregate. This is in spite of clear incentives that exist to do so, which include greater precision in estimating linear regression coefficients, greater power to detect significant factors, greater ability to correctly discern insignificant factors, and potential attenuation of the biases inherent in individual studies. The predictive power of the resulting models may presumably be greater for the aggregate data than for those of any single study. Moreover, an aggregated analysis of the experimental data produces results and inferences to which users can uniformly refer, rather than a patchwork of differing results from disparate studies. In short, the aggregation of these data sets for complete analysis may provide new or clearer knowledge that can improve engineering practice, spur further research, and drive the development of new theory.

This study pools the data from reported fracture conductivity experiments in the technical literature. Where the raw data in certain cases were not publicly available, graphical summaries of the data that appeared in the published articles were digitized to produce data with which to work. Studies were selected for inclusion in this meta-analysis

* Corresponding author.
E-mail address: ooawoleke@alaska.edu (O.O. Awoleke).

https://doi.org/10.1016/j.petrol.2019.106280
Received 10 September 2018; Received in revised form 15 May 2019; Accepted 15 July 2019
Available online 16 July 2019

if they were published in the journals or conference proceedings of the SPE (with one exception) by 2015 and passed an examination for errors, unreasonable conductivity values, and questionable methodology. The resulting set of selected studies, we believe, is of sufficient quality to be relied upon in meta-analysis. The following discussion reviews these studies selected for inclusion.

Cooke (1973) conducted laboratory tests to understand the effect of temperature and closure stress on fracture conductivity. The apparatus was used to measure the permeability of vertical fractures packed with several layers of brittle proppants at different stress levels. Two types of fracturing fluid were used in his experiments: water and oil based. He stated that permeability of some fractures is negatively affected by stress in the presence of high-temperature brine. At 250 °F, there was a significant difference in proppant pack permeability in the presence of water and oil.

Cooke (1975) developed theoretical models to evaluate the effect of fracturing fluid on conductivity. The major finding was the adverse effect of gel residue on conductivity. Polymer concentration, breaker type, and concentration are the important factors controlling the amount of residue. Vlis et al. (1975) conducted experiments using sands, glass beads, and walnut hulls as proppants at different concentration and stress levels. They developed a set of empirical relations predicting fracture conductivity by using Brinell Hardness Number (BHN) as a rock strength parameter. Volk et al. (1981) developed empirical equations to assess proppant embedment in sand and shale by using sintered bauxite as proppant. Cutler et al. (1985) measured fracture conductivity at different stress levels by using wide varieties of resin-coated and ceramic proppants. At higher stress levels, they determined bauxite and zirconia A provided the highest conductivities. Cutler et al. (1985) recommended these two proppants for deep wells and resin-coated proppant to prevent proppant flow back. It was demonstrated by Parker and McDaniel (1987) that filter cakes formed by the fracturing fluid on fractures can have harmful effects on conductivity. 20/40 Ottawa sand was used as proppant in their studies. They also reported that using an intermediate strength proppant is required even at low closure stress. And that regardless of proppant type, they concluded that it is necessary to pump proppants at high concentration to get the desired conductivity in the fractures.

McDaniel (1986) studied proppant conductivity using five different proppant types. The objectives of the study include a comparison between short-term and long-term conductivity under different test conditions, characterizing the effect a change from ambient to elevated temperatures can have on fracture conductivity, and how brine flow at high temperature and stress affects fracture conductivity. He found a significant reduction in long-term conductivity at elevated temperature and stress. Penny (1987) demonstrated effects of closure stress, temperature, proppant embedment, fracturing fluid interactions, and fluid loss additives on short-term as well as long-term conductivity. The importance of considering time as a factor to evaluate effects of stress on conductivity was reported by Much and Penny (1987).

Fredd et al. (2001) conducted a series of laboratory experiments using cores obtained from Texas Cotton Valley formation. Experiments were conducted for both smooth and rough fracture surfaces. They stated that displacement of rough fracture walls can provide increased conductivity in the absence of proppant. Their results also indicate that formation properties such as the degree of fracture displacement, the size and distribution of asperities, and rock mechanical properties can significantly affect the success of a water fracturing treatment. In addition, they recommended using high strength proppants of low concentrations but at least of 1 lb-m/ft^2 to overcome any uncertainties associated with these formation properties.

To simulate field conditions Marpaung et al. (2008) developed a dynamic fracture conductivity apparatus. They concluded that both increasing gas flux and using a breaker enhance cleanup efficiency. On the other hand, high polymer concentration in the fracturing fluid has a deleterious effect on cleanup.

Awoleke et al. (2012) conducted experimental investigations on tight gas reservoir using factorial design. The factors considered for the experiments were arranged in order of decreasing influence: closure stress, temperature, nitrogen rate, polymer loading, proppant concentration, and the presence of a breaker. The work is also an application of advanced statistical analysis in the field of petroleum engineering. Rivers et al. (2012) studied coated and uncoated proppant pack conductivity using consolidated sandstone cores. Resin from coated proppants reduces the pore throat diameter as a result of proppant pack consolidation, so uncoated proppants demonstrated higher conductivity performance than coated proppant. To consider geochemical reactions of formation and proppant diagenesis, a new dynamic system was set up by Aven et al. (2013). The dynamic flow setup was run for six months to simulate real field diagenesis. The results emphasized the importance of considering proppant compatibility with geochemical reactions before designing fracturing treatments.

Barree et al. (2016) developed a generic conductivity model for proppants. The model is applicable for almost all types of proppants including sand, resin-coated, and ceramic proppants. The main objective of this work was to compare the performance of different proppants, and to present correlations for predicting the performance of unknown or non-standard proppant materials. The factors considered in this model include closure stress, proppant concentration, and proppant size/size distribution.

Kainer et al. (2017) collected four shale samples from Barnett, Fayetteville, Marcellus, and Eagle Ford to analyze their fracture conductivity behaviors. The parameters studied for the different shale formation include closure stress, Young's modulus, Poisson's ratio, Brinell hardness number, mineralogy, proppant size, proppant concentration, and fracture surface. They employed multiple linear regression analysis to develop models of fracture conductivity for each test sample. Influential parameters for each sample were ranked by their corresponding p-value. A more general conductivity model was proposed which considers all four types of shale samples. They concluded that closure stress and proppant concentration hold have more influence on fracture conductivity than rock mechanical properties and fracture surface characteristics. Young's Modulus is also found to be one of the most important parameters and it has positive correlation to fracture conductivity at each stress level.

Few statistical approaches to model fracture conductivity are found in the literature so far. However, Awoleke et al. (2016) developed one empirical and one semi-empirical model of fracture conductivity using statistical tools. They used fractional factorial design to define the number of experiments and later employed regression analysis to model fracture conductivity data. For the semi-empirical model, they performed dimensional analysis to derive two dimensionless groups. They developed the relationship between the two dimensionless groups by using Buckingham Pi Theorem (1914). Afterwards, non-linear regression analysis was employed to develop an alternate fracture conductivity model.

Lastly, to validate the models developed in this work, we will need the following case studies. Lindley (1983) investigated proppant placement in a hydraulic fracturing treatment using “tail-in” technique at McAllen Ranch Field, Texas. Pressure build-up test data from two wells are used here for comparison: McAllen 16 and McAllen 17. For well 16, pressure build up tests were done on 02/10/1978 and on 04/26/1983. Pressure build up tests were also carried out twice for well 17 on 02/19/1980 and on 05/17/1982. 20/40 bauxite is used as proppant for the hydraulic fracturing treatment.

Palisch et al. (2007) reported realistic hydraulic fracture conductivity using PredictK (Stim Lab Proppant and Fluid Consortia, 1997–2006) for Wamsutter Gas Field, Wyoming. The gas field includes around 1300 wells, and over 150 wells were fractured from 2002 to 2004. PredictK was used to report baseline conductivity for 20/40 mesh size white sand, resin-coated sand/proppant (RCS/P), economic light-weight ceramic (ELWC) proppant, and intermediate strength proppant (ISP).

Table 1
Fracture conductivity data sources for our models.

Paper/Thesis	Objective of study
Cooke (1975)	To evaluate the effect of fracturing fluid on fracture conductivity
Vlis et al. (1975)	To determine criteria for proppant admittance and placement and to develop empirical correlations to estimate fracture conductivity
Cutler et al. (1985)	To study the performance of different proppant materials at intermediate and high closure stress
McDaniel (1986)	To investigate the effects of time, closure stress and temperature on fracture conductivity for various proppants
McDaniel (1987)	To evaluate realistic fracture conductivity of different proppants by varying temperature
Much and Penny (1987)	To study long term fracture conductivity using realistic test conditions
Parker and McDaniel (1987)	To study and minimize the effects of gel filter cakes on fracture conductivity by achieving high proppant concentration
Penny (1987)	To investigate the effects of environment and fracturing fluids on long term fracture conductivity
Fredd et al. (2001)	To study the effects of different fracture properties on fracture conductivity for water fracturing and conventional fracturing treatments
Marpaung et al. (2008)	To investigate dynamic fracture conductivity and measure cleanup efficiency
Awoleke et al. (2012)	To investigate the effects of different factors affecting propped fracture conductivity using factorial design
Rivers et al. (2012)	To investigate fracture width and fracture conductivity with high proppant loading at high pressure and temperature
Aven et al. (2013)	To evaluate proppant damage with long term dynamic flow testing apparatus
Zhang (2014)	To determine unpropped and propped fracture conductivity, fracture conductivity impairment by water
Li et al. (2015)	To develop mathematical models to calculate proppant embedment, proppant deformation, the change in fracture aperture, and fracture conductivity
McGinley et al. (2015)	To evaluate fracture conductivity with samples fractured in both horizontal and vertical orientation

Handren and Palisch (2007) illustrated a case history of hybrid slickwater-fracture designs in the Cotton Valley Taylor completions of East Texas. The authors simulated the actual well conditions and determined proppant performance at temperature and stress conditions. The actual conditions were maintained for 45 days to investigate the loss in fracture conductivities. The authors conducted tests using different proppants. Results from 40/70 sand, 20/40 ELWC, and 30/50 ELWC proppants are used here to carry out the model comparison.

In summary, by pooling the raw observations from the included studies, we obtain approximately 2700 data points from fifteen research papers and one thesis dissertation on hydraulic fracturing. The list and the main objectives of these included studies are given below in Table 1. Our objective is to use the aggregated experimental dataset as the basis for the development of empirical models of propped fracture conductivity, whereupon we might ascertain significant factors, assess uncertainty of model coefficient estimates, and produce propped fracture conductivity predictions.

Accordingly, we present a sequence of statistical models of fracture conductivity based on a compilation of fracture conductivity data from the technical literature. All of the models presented in this work not only account for the effects of closure stress, proppant type/concentration/mesh size (like Barree et al.'s (2016) model); we also account for additional variables like reservoir temperature, polymer loading and Brinell hardness that were not accounted for in previous models. Elevated reservoir temperature can have either beneficial effect on fracture conductivity because it helps the gel cleanup process. However, at above 250 °F, it can help accelerate proppant diagenesis which, in turn, lowers proppant conductivity. Polymer loading directly affects fracture conductivity because the likelihood of gel damage is increased at elevated values of polymer loading. Proppant embedment is much more of an issue in softer formations. Accordingly, the higher the rock Brinell hardness number, the potential of fracture conductivity impairment due to embedment is reduced. An exhaustive discussion of impairment mechanisms can be seen in Palisch et al., (2007).

The models presented here are based on data from experimental investigations in different laboratories. The pooling of such data into a single database requires that we assume that the physical principles at work in determining fracture conductivity operate in nearly the same way across the various investigations. As discussed in the final section, this assumption is a potential limitation of this study, as it is in every published meta-analysis. Notwithstanding, the safeguards inherent in both statistical model diagnostic checking and in the cross-validations used in model-building offer protection and resilience from input data which possesses either material bias or disproportionate error (or both). The upshot is that a meta-analysis like this can dilute biases present in the results of single-laboratory investigations.

The balance of this article is organized as follows: first, we outline the study methodology followed by a section summarizing the model results. Another section compares the model results to those of a previously published model, and then we apply the fitted models to make predictions on published case histories for validation purposes. Finally, some discussion and conclusions are provided.

2. Methodology

2.1. Experimental factors

Potentially significant factors that can affect resulting fracture conductivity of a propped well include proppant particle size, type, concentration and embedment, closure stress, reservoir temperature, polymer loading and residue, formation type and hardness, flow back rate and the presence or absence of breaker. In this study, we focus on developing empirical models based on the subsets of these listed in the third column of Table 2 (proppant mesh size and concentration, closure stress, reservoir temperature, polymer concentration, and formation hardness). These factors were chosen from the above list based on the preponderance of data available for each one of them in the scientific literature.

Additionally, proppant type is used as an independent variable which takes three levels: sand, resin-coated sand, and ceramics. For notational convenience, these levels will be referred to as types 1, 2, and 3, respectively. In defining models, the variables fracture conductivity, mesh size, closure stress, proppant concentration, temperature, polymer concentration, proppant type, and formation hardness will be denoted by Fc, P_{ms}, σ_c, C_p, T, C_{pl}, t_P, and BHN respectively.

The key difficulty inherent in pooling multivariate data from multiple studies lies in these studies' inconsistent controlling-for and reporting-of experimental factors. For example, certain studies may have reported mesh size and temperature, but not polymer concentration. Other studies may have reported mesh size and polymer concentration, omitting temperature. Still other studies may have left out mesh size but reported temperature and polymer concentration. Since standard regression models make no allowance for observations being incomplete in one or more variables, statistical software packages will crop the data set in this example to those observations (if any) which are complete in all three experimental factors only. Rather than accept the loss of information that results from fitting a model only to data which are complete in all factors, we adopt a strategy of fitting a model to data which are complete in only some of the factors. We may do this for several different subsets of the factors, resulting in partially distinct sets of observations being used in each model. Each row in Table 2 represents a different subset of factors. Under proppant type 1 there are

Table 2
Fixed Regression models developed in this study with the independent variables used in each model.

Type of Proppant	Developed Models in this Study	Independent Variables	Sample size
Proppant Type 1 (sand)	Three Factor Model	Mesh Size (Proppant Particle Size), Closure Stress, and Temperature	786
	Four Factor Model	Mesh Size (Proppant Particle Size), Closure Stress, Proppant Concentration, and Temperature	655
		Closure Stress, Proppant Concentration, Temperature and Formation Hardness (BHN)	425
	Five Factor Model	Mesh Size (Proppant Particle Size), Closure Stress, Proppant Concentration, Temperature, and Polymer Loading	179
Proppant Type 2 (resin-coated)	Three Factor Model	Mesh Size (Proppant Particle Size), Closure Stress, Temperature	198
Proppant Type 3 (ceramics)	Three Factor Model	Mesh Size (Proppant Particle Size), Closure Stress, Temperature	718
	Four Factor Model	Mesh Size (Proppant Particle Size), Closure Stress, Proppant Concentration, and Temperature	343
	Five Factor Model	Mesh Size (Proppant Particle Size), Closure Stress, Proppant Concentration, Temperature, and Polymer Loading	133
All Proppant Type	Six Factor Model	Mesh Size (Proppant Particle Size), Closure Stress, Proppant Concentration, Temperature, Polymer Loading, and Type of Proppant	353

four different subsets of factors used for fitting a regression model. For proppant type 2, only the three-factors model is developed and for proppant type 3, three subsets of factors are used. The six-factors model is developed for the all proppant type dataset.

Fitting a model to proppant type 2 with only three factors is necessary due to the small amount of data available for resin-coated sand. Similarly, two four-factor models are possible in proppant type 1 due to the large amount of data available for sand. In all cases, the models developed for different subsets of factors can be evaluated using the same diagnostics and predictive performance metrics described later on. A side-benefit of fitting models to so many different combinations of factors is that an engineer with data in fewer than all six factors may still find a model developed herein for their particular combination of factors.

With several models being fit for proppant types 1 and 3, we would like a way to compare the various models' performance. The predictive power of the fitted regression models is easily measured using cross-validation procedures. Such procedures make random partitions of the data into a model training set and a validation set. The training set is used to build the model and make predictions for the data points contained in the validation set. Overall prediction strength of a model can be quantified through mean squared prediction error (MSPE) or through cross validation R-squared. Low values of the former and high values of the latter indicate a model that predicts well.

In this research we use cross validation in two different ways to achieve two different types of comparisons. First, a global cross validation is performed wherein the original data set within each proppant type is randomly partitioned from the outset. The validation data (representing 30% of the original data) is reserved until after all models are finalized (using the other 70%) in order to make out-of-sample prediction comparisons with those of the model from Barree et al. (2016). The results of these comparisons are discussed in *Comparisons to Another Model section* of this work. Second, a K-fold cross validation approach is used for developing the models themselves. This approach randomly partitions the data set into k sets so that each set can be used for prediction validation of a model fit with the other $k-1$ sets used as training data. In the end a prediction is made for every observation in the data and model strength is assessed by these predictions. The results of these K-fold validations, where we have taken k to be 10, are summarized in the *Results and Discussion section* of this work.

A further measure of predictive performance uses new, independent data from published case histories, as described in the *Model Validation with Case Histories* section of this work.

2.2. Transformations

In order to standardize regression coefficients, a normalizing transformation is applied to all experimental factors which take quantitative values. For each factor, this is accomplished by subtracting out its mean and dividing out its standard deviation. Appendix A contains the means, standard deviations used to transform the datasets. Qualitative factors are left untransformed and enter regression models through dummy variables.

In addition, in order to improve regression model diagnostics (see below), the response variable fracture conductivity is transformed via a natural logarithm.

2.3. Model building

In this study, multiple linear regression analysis is performed using the *lm()* function in R (R Core Team, 2017). In building each model, we consider including all main effect terms and all possible two-way interaction effect terms as fixed effects since they are all continuous variables. When the relationship between the response variable and one of the quantitative factor variables is judged unlikely to be linear, we also consider including a quadratic effect term for the variable.

Developing a regression model with multiple independent variables, their interactions, and polynomial terms (collectively referred to as *regressors*) can result in a potentially large number of predictors. It is normally not necessary to include all these predictors in a model. It is reasonable to choose those predictors that contribute significantly to improve the model's performance in diagnostic checks, prediction

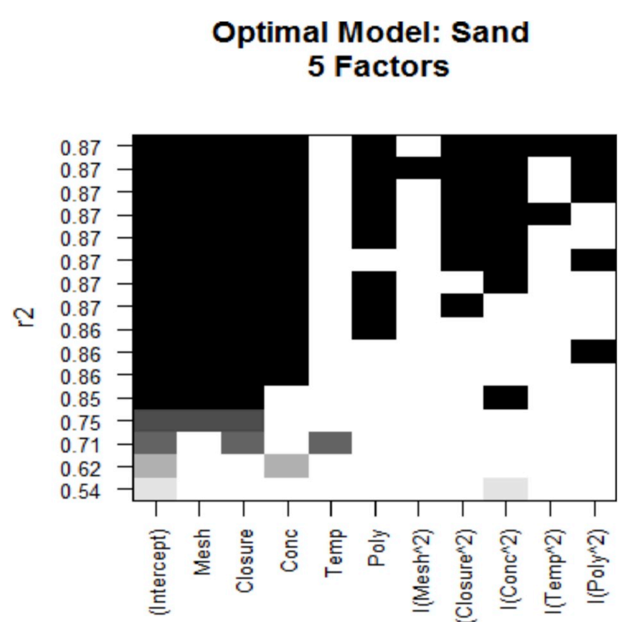

Fig. 1. Optimal model from all subsets regression for five factor model of proppant type 1.

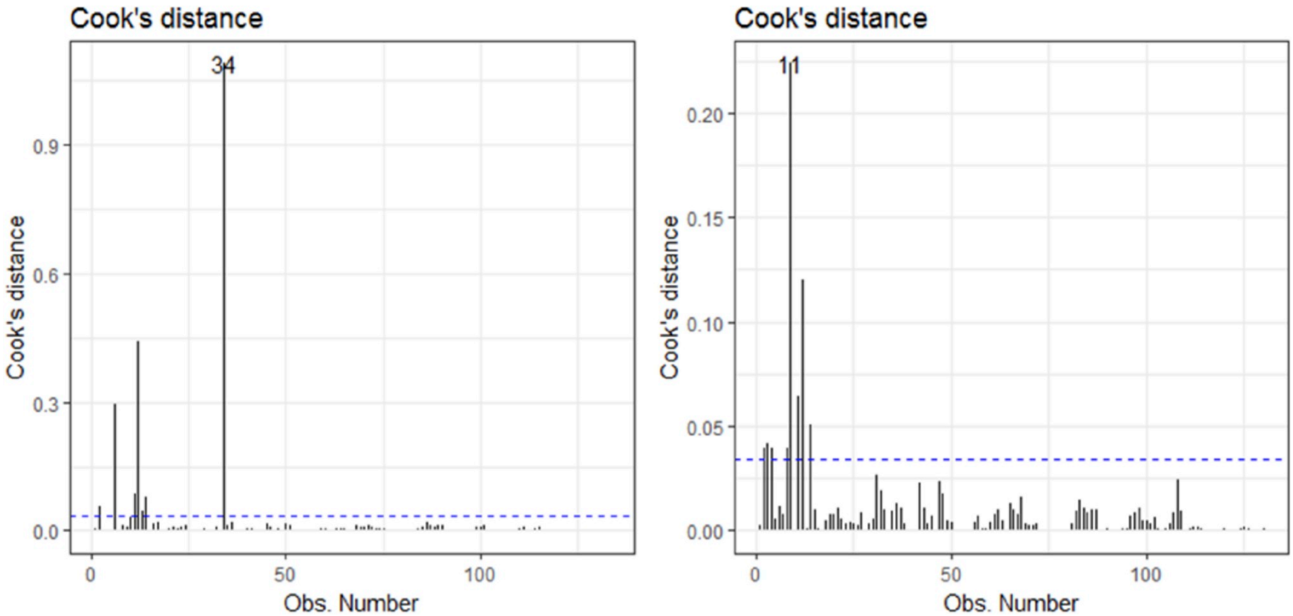

Fig. 2. Cook's distance plot for the five factor model of proppant type 1 (Left-before deletion, Right-after deletion).

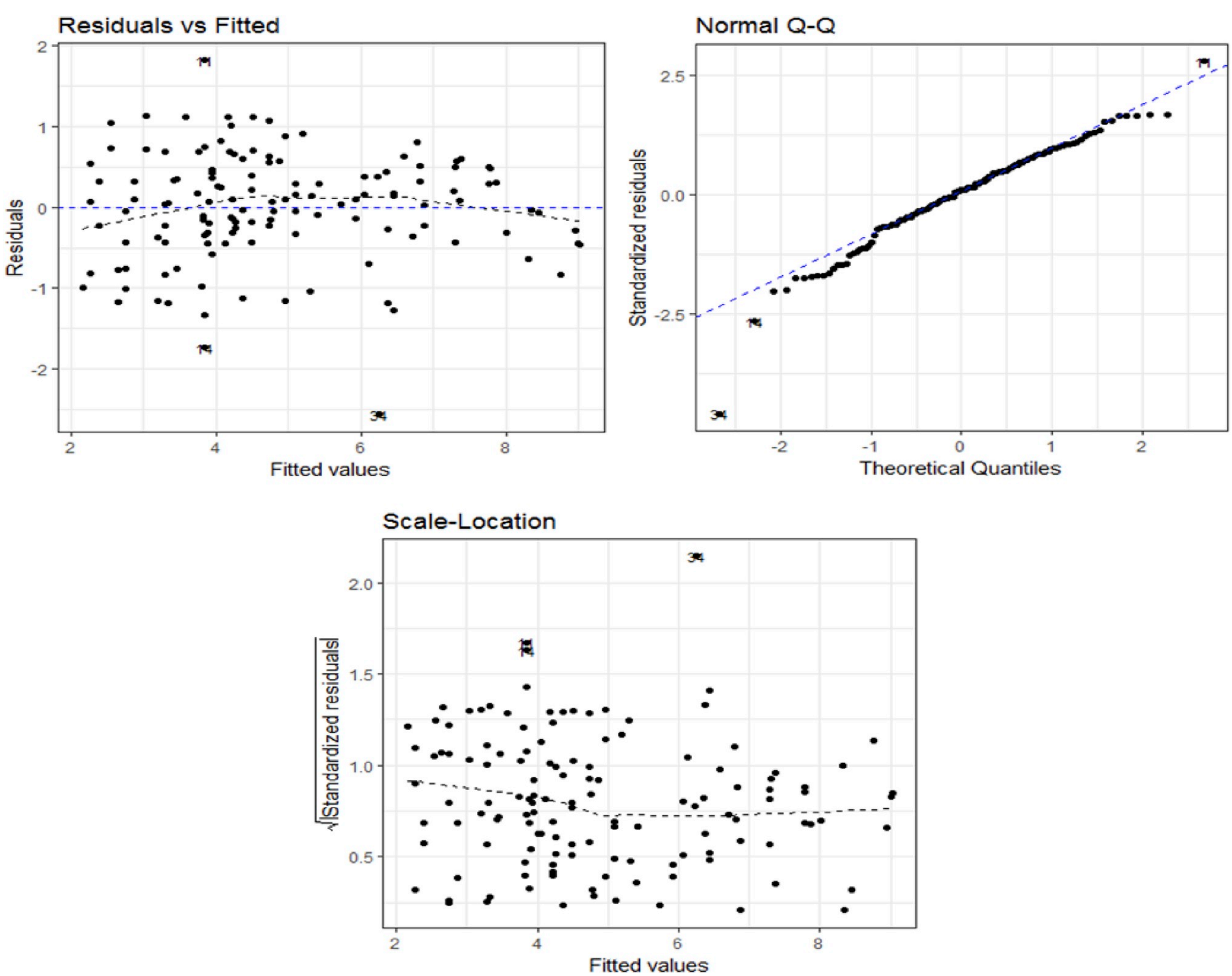

Fig. 3. Regression diagnostics before deleting influential observations as the validation of model assumptions for the five factor model of proppant type 1, (Upper left-Linearity, Upper right-Normality, Bottom-Homoscedasticity).

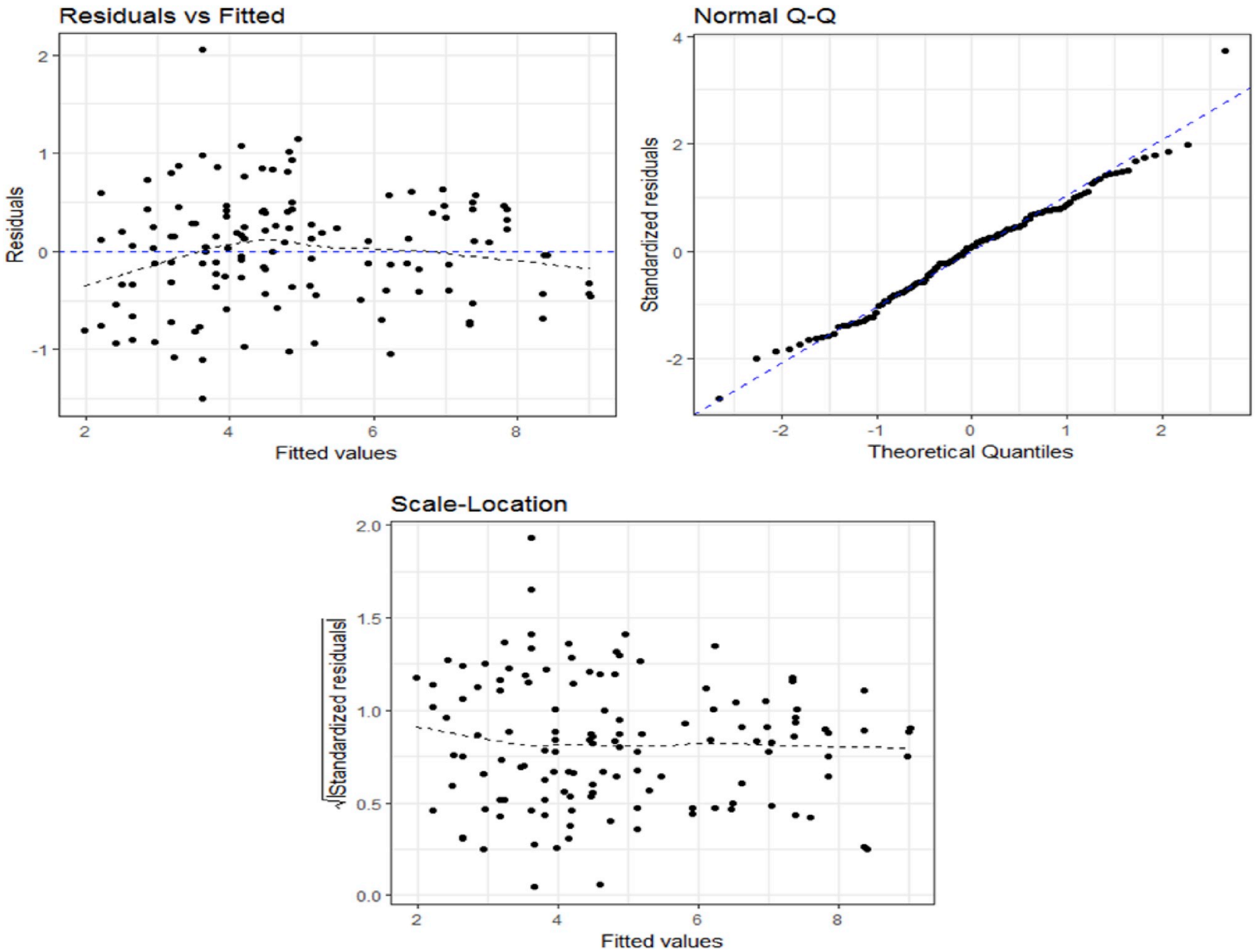

Fig. 4. Regression diagnostics after deleting influential observations as the validation of model assumptions for the five factor model of proppant type 1, (Upper left-Linearity, Upper right-Normality, Bottom-Homoscedasticity).

validation, and goodness-of-fit. To choose the best model all possible models (consisting of the models formed from every subset of the set of all possible regressors) are fit. This method is called all subsets regression. It is done with the *regsubsets()* function of the leaps package in R (Lumley, 2017). It compares all possible models using different subsets of predictors. The function's output shows a small number of the best fitting models according to some user-selected criterion, such as R-squared or adjusted R-squared. When a selected model contains an interaction effect term or a polynomial term as a significant regressor we then force the main effect(s) of that particular term into the best model regardless of its (or their) reported significance in order to aid models interpretability.

2.4. Regression diagnostics

Regression diagnostics are checks performed to validate the assumptions underlying the developed models. Although regression software will automatically provide model results including estimated regression coefficients, R-squared, adjusted R-squared, and residual standard error, it is not appropriate to immediately assume that the fitted models are useful for future prediction. This is because characteristics of the sample data (and hence the underlying population's data) might not be consistent with assumptions made by the model. In such a case, results could be misleading, resulting in incorrect inferences about the relationship between the response variable and the independent variables.

To validate a model, one customarily checks that the data satisfy the following major statistical assumptions: linearity between the regressors and the response variable, normality in the model errors, and homoscedasticity in the model errors.

This first assumption states that the response variable has a linear relationship with the independent variables. If the response variable is linearly related to the independent variables then there should not be any evident association between the model's residuals and its fitted values. This assumption is validated from a scatterplot graph of the residuals versus predicted log-fracture conductivities. The points on the graph should be randomly distributed around a horizontal line at zero.

The second assumption states that the model errors are normally distributed. If this is true then the residuals will also approximately be normally distributed with a mean of zero. This assumption is checked with a normal probability plot of standardized residuals. It is also called a normal Quantile-Quantile (Q-Q) plot. The points on the graph should fall roughly on a straight line.

The third assumption states that the model errors have a constant variance. It can be evaluated with a graph of the square root of the absolute value of the standardized residuals versus the predicted log-fracture conductivity values. This is also called a scale location plot in statistics. If the constant variance assumption is met then the points on the graph should fall randomly in a band centered on a horizontal line.

In addition to the model assumptions listed above, it is important to check for influential observations that disproportionately impact the model's fit. Removing such observations, when justified by non-statistical considerations, can result in improvement in the model performance. Cook (1977) introduced a method to measure each

Table 3
Summary of fitted regression models.

Type of Proppant	[a] Developed Models in this Study	Fitted Regression Models
Proppant Type 1 (sand)	Three Factors Model	$\ln Fc = 7.21611 + (0.75318{*}P_{ms}) - (0.35250{*}\sigma_c) - (0.02203{*}T) - (1.04327{*}P_{ms}{*}\sigma_c) - (0.64078{*}\sigma_c{*}T)$
	Four Factors Model	$\ln Fc = 7.20960 + (0.59221{*}P_{ms}) - (0.59399{*}\sigma_c) + (1.63724{*}C_p)$) $- (0.38014{*}T) - (0.78710{*}P_{ms}{*}\sigma_c) - (0.60512{*}P_{ms}{*}C_p) - (0.69822{*}\sigma_c{*}T) + (0.81040{*}C_p{*}T$
	Four Factors Model	$\ln Fc = 8.15643 - (0.82407{*}\sigma_c) + (86302{*}C_p) - (0.23893{*}T) - (0.20613{*}BHN) + (0.24599{*}\sigma_c{*}BHN) + (0.36996{*}C_p{*}BHN)$
	Five Factors Model	$\ln Fc = 6.93165 + (1.26923{*}P_{ms}) - (0.90890{*}\sigma_c) + (3.44982{*}C_p) - (0.10642{*}T) - (0.08941{*}C_{pl}) + (0.16575{*}\sigma_c^2) - (1.90421{*}C_p^2) + (0.05304{*}T^2) - (0.40768{*}C_{pl}^2)$
Proppant Type 2 (resin-coated proppant)	Three Factors Model	$\ln Fc = 8.09765 - (0.19166{*}P_{ms}) - (0.34484{*}\sigma_c) - (0.37392{*}T) - (0.31527{*}P_{ms}{*}\sigma_c) + (0.10521{*}M{*}T) - (0.47903{*}\sigma_c{*}T)$
Proppant Type 3 (ceramic proppant)	Three Factors Model	$\ln Fc = 8.64340 + (0.41235{*}P_{ms}) - (0.28032{*}\sigma_c) - (0.31361{*}T) - (0.21124{*}P_{ms}^2) - (0.07990{*}T^2)$
	Four Factors Model	$\ln Fc = 8.57212 + (0.06329{*}P_{ms}) - (0.53543{*}\sigma_c) + (1.53781{*}C_p) - (0.09027{*}T) - (0.37050{*}P_{ms}^2) - (0.44576{*}C_p^2)$
	Five Factors Model	$\ln Fc = 8.22287 + (0.78182{*}P_{ms}) - (0.86664{*}\sigma_c) + (1.70340{*}C_p) - (0.05282{*}T) - (0.03940{*}C_{pl} - (0.87666{*}P_{ms}^2) - (0.61180{*}C_p^2) + (0.33036{*}T^2) + (0.07949{*}C_{pl}^2)$
All Proppant Types	Six Factors Model	$\ln Fc = 5.29372 + (1.50163{*}P_{ms}) - (0.92900{*}\sigma_c) + (1.47756{*}C_p) - (0.49934{*}T) - (0.57253{*}C_{pl} + (0.26464{*}t_P) + (0.22944{*}\sigma_c^2) - (0.69801{*}C_p^2) + (0.67987{*}T^2)$

[a] To use the models in Table 3 directly, the engineer should standardize the actual values of the variables using the mean and standard deviation data specific to the model the engineer wants to use (see Tables A-1 to A-9 of Appendix A).

observation's influence, which is known as Cook's distance. It is both a function of the observation's residual and its leverage, which is a measure of how far the observation's independent variable values lie from those of the other observations. A large residual or a large leverage will contribute to large values of Cook's distance.

One proposed threshold is to regard observations as being unduly influential when their Cook's distance is greater than $4/(n - p - 1)$, where n is the sample size and p is the number of model regressors (Kabacoff, 2015). For these models we adopt a policy of deletion of influential points and outlying points one-at-a-time until no such points remain. The automatic deletion of such points is not, in general, recommended in the statistical literature because it can result in biased estimates and incorrect conclusions. However, for the purposes of this study we justify such an approach with the following rationale. First, this work is based on data from sixteen studies spanning four decades, and it is possible that fracture conductivity was underestimated or overestimated in some cases. Second, the primary researchers of these data sources are, in many cases, not available to verify whether they encountered any issues during any particular measurement.

3. Results and discussion

We will illustrate the methodology associated with fixed effects (or traditional) regression modeling here with the five factors model for proppant type 1. Thereafter, we will summarize the results for all the models as enumerated in Table 2.

3.1. Five factors model—proppant type 1

The very first linear model developed with the dataset has an R-squared of 0.9126 and an adjusted R-squared of 0.903. However, the main effects of mesh size and temperature are reported as negative, which violates general understanding of the nature of these variables' relationship with fracture conductivity. We therefore employ polynomial terms in these variables to account for potential non-linearity between the response and independent variables. The resulting polynomial model has an R-squared and an adjusted R-squared of 0.8713 and 0.8608 respectively. Fig. 1 depicts the optimal model according to R-squared when only models with nine or fewer regressors are considered. All the possible main and polynomial effect terms are plotted on the x-axis (we omit interaction terms here). The shaded blocks in each row indicate which variables are included in that row's model, and the y-axis gives the model's corresponding R-squared. Starting from the bottom, the first row represents a model with intercept and the polynomial term of proppant concentration, which has an R-squared of 0.54. In the top two rows, the black blocks indicate regressors present in the best two models, which have R-squared of 0.87. Starting from the top, the first row does not include the main effect of temperature and polynomial effect of mesh size. Afterwards, in the second row, the main and quadratic effects of temperature are not included in the best model as predictors. The predictors from the first row (starting from the top) are chosen to be the optimal model or best model because this study is more interested in developing effective models than fitting models which contain all possible independent variables. Since the polynomial effect of temperature is found in the optimal model, we force the main effect of temperature into the final optimal model as a regressor.

Fig. 2 shows the Cook's distance plot before deletion (left) and after deletion (right). The Cook's distance plot before deletion (left) indicates that very few data points are above the blue line which might have significant influence on the regression model. Only three data points have to be removed to reach the point where no influential points remain. The final R-square and adjusted R-square for this model are 0.9056 and 0.8985 respectively. Deleting these three data points decreases the mean squared error from 0.44 to 0.33.

The regression diagnostic plot before deletion is displayed in Fig. 3. In the residuals versus fitted values plot (upper left), 11th, 14th, and

Table 4
Comparison of model performance between three, four and five factors models of proppant type 1.

Performance Assessing Measure		Three Factors Model	Four Factors Model	Five Factors Model
K-fold cross validation	Original R-squared	0.5206	0.7442	0.9056
	Cross validated R-squared	0.3219	0.5106	0.8826
MSE		1.83	1.18	0.33
MSPE		3.42	2.91	0.48
Suggested Power Transformation		1.895	2.09	1.07
Satisfaction Levels of Model Assumptions	Linearity	Moderate	Low	High
	Normality	High	Moderate	High
	Homoscedasticity	Moderate	Low	High
Overall Model Adequacy		Fair	Poor	Good

Table 5
Comparison of model performances between three, four and five factors models of proppant type 3.

Performance Assessing Measure		Three Factors Model	Four Factors Model	Five Factors Model
K-fold cross validation	Original R-squared	0.3436	0.8032	0.6395
	Cross validated R-squared	0.3271	0.7850	0.5096
MSE		0.64	0.18	0.79
MSPE		1.6	0.65	0.80
Suggested Power Transformation		4.86	5.52	1.81
Satisfaction Levels of Model Assumptions	Linearity	Low	Low	High
	Normality	Low	Low	High
	Homoscedasticity	Low	Low	High
Overall Model Adequacy		Poor	Poor	Good

34th number observations are certainly outliers, but overall there is no distinctive pattern created by the data points. All the data points are almost equally spread around a horizontal line, which indicates that the relationship we developed to relate the dependent variable to the independent variables is adequate. In the normal Q-Q plot (upper right), most of the data points fall on the line, although there are still a few data points at the top and the bottom including the outliers, which do not fall on that line. The distribution in the normal Q-Q plot appears to be normal. In the scale location plot (bottom), there is a random band around the horizontal line, which suggests that the homoscedasticity assumption is met.

The regression diagnostic plots after deletion are displayed in Fig. 4. There are no major changes in Fig. 4 from Fig. 3, indicating that the model plausibly meets all the model assumptions.

In summary, the regression diagnostics suggest that the model assumptions are satisfied. The final five-factors model for proppant type 1 is given below in Equation (1).

$$\ln Fc = 6.93165 + (1.26923{*}P_{ms}) - (0.90890{*}\sigma_c) + (3.44982{*}C_p) - (0.10642{*}T) - (0.08941{*}C_{pl}) + (0.16575{*}\sigma_c^2) - (1.90421{*}C_p^2) + (0.05304{*}T^2) - \left(0.40768{*}C_{pl}^2\right) \tag{1}$$

Statistically significant predictors are identified by testing that each predictor's coefficient is significantly different from 0 in a standard *t*-test. We use an α-level of 0.05 for declaring predictors significant. For instance, in Equation (1), temperature, polymer concentration, and the polynomial term of temperature have p-values higher than 0.05. Regardless of the significance, all the predictors reported from all subsets regression are reported in the final models in Table 3.

3.2. Other regression models

The above procedure was repeated for all the models mentioned in Table 2. Table 3 presents a summary of all the fixed regressions models

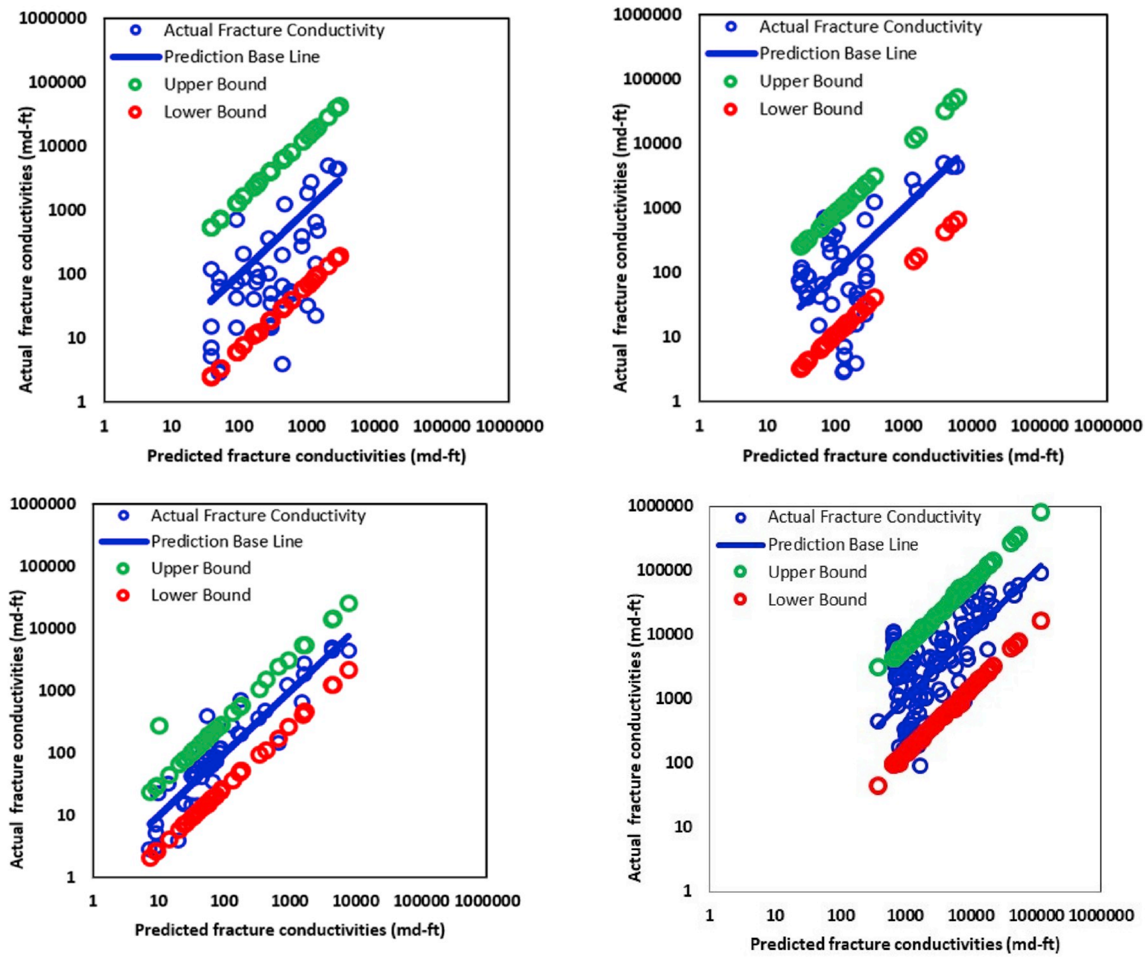

Fig. 5. Graphical representation of predictions for the models of proppant type 1 on 30% validation data with 95% prediction interval (Upper left-three Factors, Upper right- 1st four factors model, Bottom left—five factors, Bottom right—2nd four factors model).

developed during this study. Details about the procedure for the rest of the models shown in Table 3 are fully described in Rahman (2017). The mean and standard deviation of the independent variables for each fixed regression model is provided in Appendix A. Appendix A also contains the ranges of the independent variables.

3.3. Model comparisons

The performance of all three models of proppant type 1 is summarized in Table 4. The percent decreases in cross validated R-squared from original R-squared are listed as 38%, 31%, and 2.5% for three, four, and five factor models respectively. We see that the five factors model shows better predictions with newer data than the other two models because of its lower percent decrease in R-squared. Moreover, the mean squared error (MSE) and mean squared prediction error are the lowest for the five factors model.

The performances of all the models of proppant type 3 is summarized in Table 5. The percent decrease from original R-squared to cross validated R-squared is the highest for the five factors model and the lowest for the four factors model. The mean squared error and mean squared prediction error are also lowest for the four factors model and highest for the three factors model. Thus, the five factors model for proppant type 3 shows intermediate performance.

The performance of the models for proppant type 2 is not given for conciseness and since fewer data points are available for this proppant type.

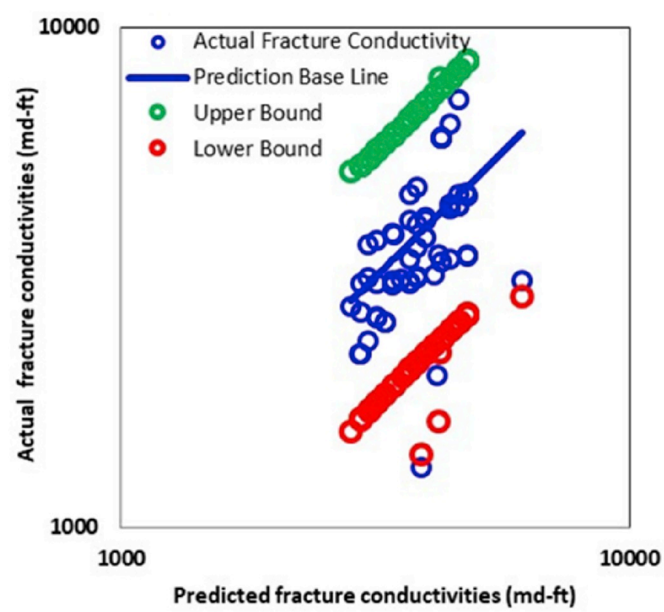

Fig. 6. Graphical representation of predictions for the three factors model of proppant type 2 on 30% validation data with 95% prediction interval.

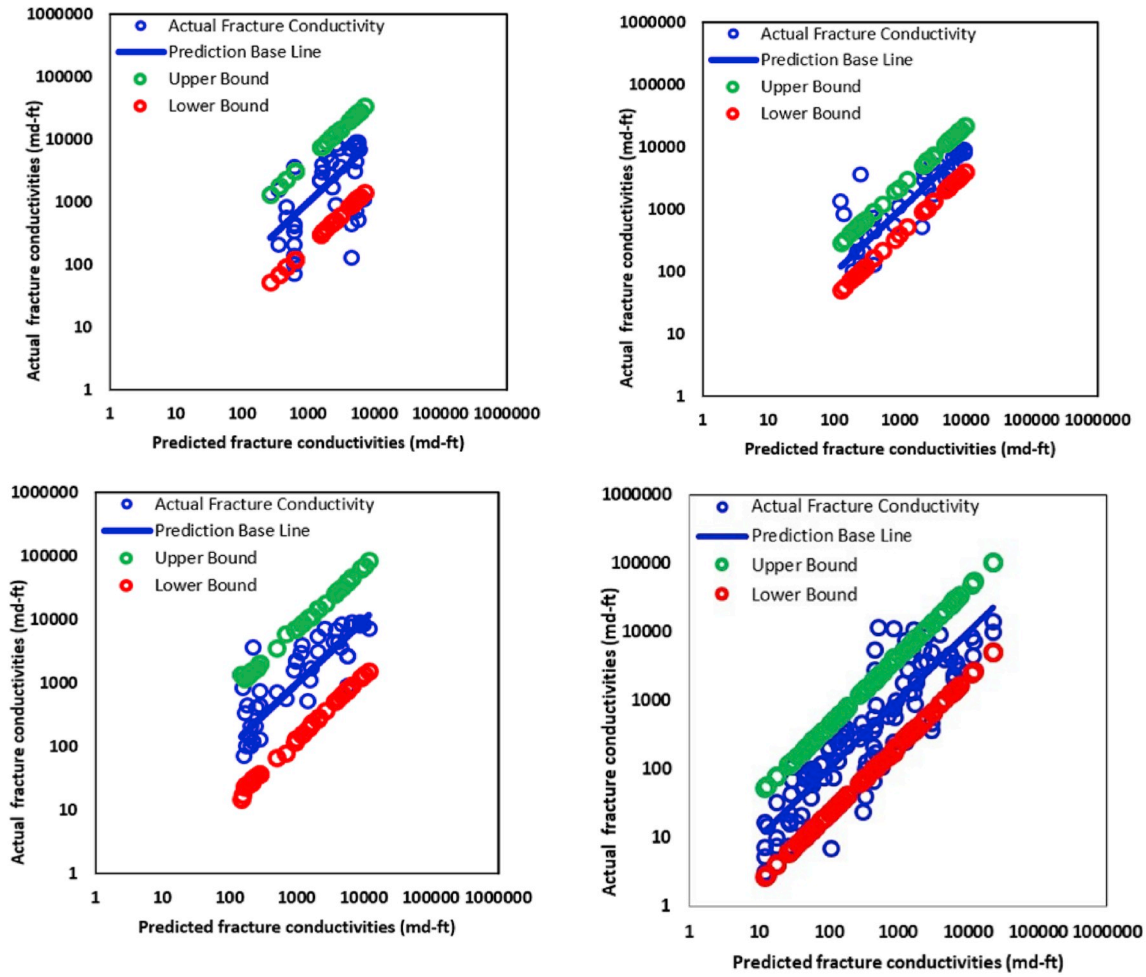

Fig. 7. Graphical representation of predictions for the models of proppant type 3 on 30% validation data with 95% prediction interval (Upper left-three factors, Upper right—four factors, Bottom left-five factors, Bottom right-six factors model, all proppant type).

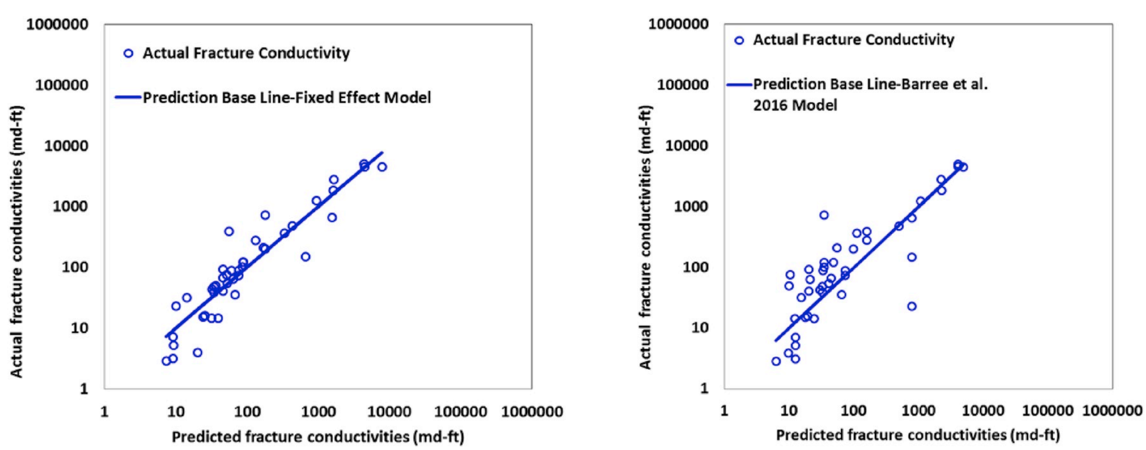

Fig. 8. Graphical measures of model predictive performance for models of proppant type 1 used in this research, as well that of Barree et al. model Fracture Conductivity Model (left-our model, right-Barree et al. (2016) model).

3.4. Model validations

All the models developed for proppant type 1 are validated with the 30% validation dataset (43 data points) that was separated at the beginning of this research. Fig. 5 is the graphical representation of predictions using the validation dataset with a 95% prediction interval for these three models. Equation (2) is used to estimate the prediction interval where $\hat{Y}_i$ is the predicted natural log of fracture conductivity, t_{crit} is the critical value of t distribution, MSE is the mean squared error of residuals, X is the $(p+1) \times n$ matrix of independent variables, and X_0 is the $(p+1) \times 1$ column vector of values of the independent variables for the observations being predicted. Here, p represents the number of independent variables.

$$\textit{Prediction Interval} = \hat{Y}_i \pm t_{crit}\sqrt{MSE\left(1 + X_0^T(X^TX)^{-1}X_0\right)} \tag{2}$$

For all the models, most of the actual conductivity data fall inside

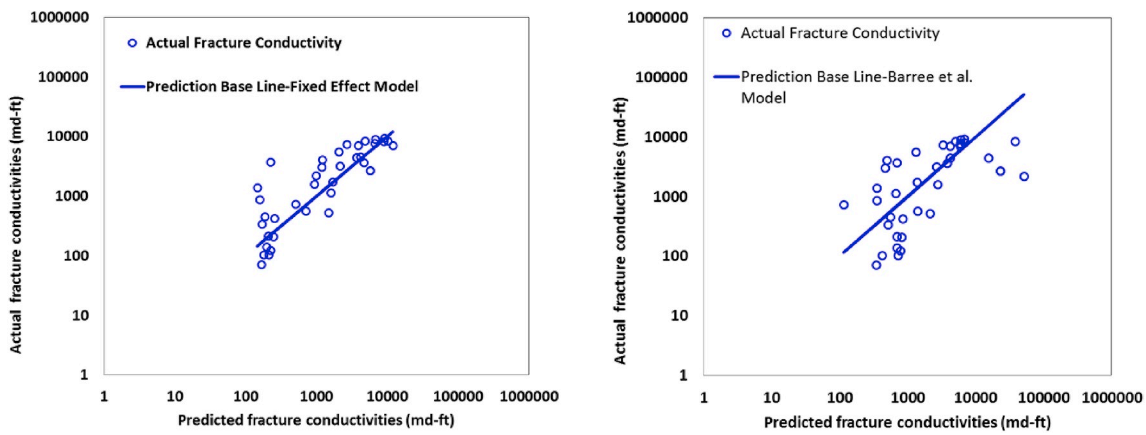

Fig. 9. Graphical measures of model predictive performance for models of proppant type 3 used in this research, as well that of Barree et al. model Fracture Conductivity Model (left-our model, right-Barree et al. (2016) model).

the prediction interval. In the bottom-left plot of Fig. 5 (five-factors model), the actual conductivities are closer to the prediction base line and only one or two data points are outside the prediction interval. On the other hand, the actual conductivity is more scattered from the prediction baseline in the other two plots at the top of Fig. 5. The mean squared prediction errors for three, four, and five factor models are 3.42, 2.91, and 0.48 respectively. Therefore, the five factor model provides more accurate predictions than the other two models. The graphical representation of prediction for the other four factor model is shown in the bottom right corner of Fig. 5. Actual conductivities are mostly inside the prediction envelope. The mean squared prediction error for this model is 1.44.

As expected, under the same conditions, the predicted fracture conductivity from the models in Table 3 are highest for ceramic proppant (proppant type 3) and least for sand (proppant type 1). Resin coated sand conductivity values generally lie between those for sand and ceramic proppants. For example, if the standardized values of all the variables in Table 3 is set at zero, we see that for proppant type 1, the natural log of fracture conductivity (ln *Fc*) is between 6.93 and 8.15 (actual conductivity between 1022 and 3463 md-ft), depending on the factors model used (average ln *Fc* is 7.38, that is 1604 md-ft). For proppant type 2, ln *Fc* is approximately 8.1 (that is 3294 md-ft) and for proppant type 3, ln *Fc* is between 8.22 and 8.64 (between 3714 and 5653 md-ft) depending on the factors model used (average ln *Fc* is 8.48, that is 4817 md-ft).

In the six factors model, proppant type is treated as a categorical variable with two levels, and proppant type 1 is categorized as “1” and proppant type 3 as “2” in the database. Proppant type 2 was not included in the six factors model because of data paucity.

The graphical representation of predictions using the three factors model of proppant type 2 is displayed in Fig. 6. 55 data points are used to validate the model. The mean squared prediction error of predictions is 0.08. The regression diagnostics performed for this model suggest that it is inadequate in terms of model assumptions. In addition, the model is developed using only three factors as no data was available for any additional factors. Altogether, we recommend that the model not be relied upon until more data becomes available.

43 data points are used to validate the rest of the models developed for proppant type 3. Fig. 7 is the graphical representation of predictions using the validation data with a 95% prediction interval for these three models. The upper left and upper right plots of Fig. 7 show predictions for the three and four factors model respectively. Three or more data points are outside the prediction interval in these two plots. In the upper left plot, actual conductivities are widely scattered around the prediction baseline. On the other hand, the plot at the bottom-left of Fig. 7 shows more accuracy in prediction. Only a single data point is outside the prediction envelope in this plot. Actual conductivities are well around the prediction baseline. The mean squared prediction errors for three, four, and five factor models are 1.61, 0.65, and 0.80 respectively. Evidently, the four and five factors models provide better predictions than the three factors model.

All proppant types includes both sand, resin-coated sand, and ceramic proppants. The bottom right corner of Fig. 7 represents the predictions on validation data for this model. Around 115 data points are used to validate the model. More than 90% of actual conductivities are inside the 95% prediction envelope and they are scattered closely around the prediction baseline. The mean squared prediction error is 1.10 for this case.

3.5. *Comparisons to Another Model*

Barree et al. (2016) provided a general model of fracture conductivity based on work and experiments done at the Stim-Lab Proppant Conductivity Consortium. The model developed in their study used datasets containing a wide range of proppant types and sizes. The main aim was to estimate correlations between fracture conductivity and other variables to assess newer proppant types or sizes. This model correlated fracture conductivity to closure stress and proppant properties and concentration.

We used the 30% validation dataset used to validate all the models of this research to predict fracture conductivity using the correlations provided by Barree et al. (2016). Barree et al.'s model specified different regression coefficients for the different proppants under an umbrella proppant type. To use their correlation with our validation data, these regression coefficients are averaged for each umbrella proppant type. Mesh size is transformed from US sieve to microns according to the API specifications and later averaged for each umbrella proppant type. We thereafter compared the five factor models for both proppant type 1 and proppant type 3 developed above with the Barree et al. (2016) model. These comparisons are shown in Figs. 8 and 9 respectively. The mean squared prediction error for proppant type 1 and proppant type 3 data using the Barree et al. (2016) model are 1.19 and 3.69 respectively.

This implies that the five-factor models presented in this paper for both proppant types 1 and 3 (MSPE gave better predictions (MSPE = 0.48 and 0.80 respectively—see Tables 4 and 5)) compared to the Barree et al. model. This makes sense because the Barree at al. model only accounts for proppant type, concentration, and stress with the expectation that the engineer would correct the conductivity value for the effects of gel damage. However, in addition to proppant type, concentration and stress, the models presented here explicitly account for both polymer loading and temperature.

3.6. Model Validation with case histories

Although the global cross-validation method used in the previous section employs out-of-sample prediction as a way to validate the models developed, the data held back for validation was still originally part of the same data sets used for developing models. As such, it may possess more similarity to the training data than data would which was not from the same sources. Here we validate the models using case history data which was not derived from the original data sources, providing a robust evaluation of the models' predictive performance.

All the case histories discussed in this section did not report any values for polymer concentration. Therefore, a sensitivity analysis on the case histories is done using a range between 10 lb/1000 gal to 60 lb/1000 gal of polymer concentration.

Our assumption is that the fracture conductivity prediction will be closest to the 'true or actual' fracture conductivity at the actual polymer concentration used in the case histories. The most important consideration here is that the 95% prediction interval from our model actually brackets or includes the 'true or actual' fracture conductivity value. 'True or actual' conductivity value here refers to fracture conductivity estimated from other independent sources by other researchers.

More recent data that explicitly provide values for all the independent variables investigated in this research are scarce. Therefore, the most recent data included here are from 2007 (Handren and Palisch, 2007). However, the case histories presented here illustrate the applicability of these models. Data from three previous case histories are collated from the literature to validate the models (five factor fixed models for proppant type 1 and proppant type 3) developed in this study. These case histories are (1) Case History I: Lindley (1983), (2) Case History II: Palisch et al. (2007), and (3) Case History III: Handren and Palisch (2007) as described in the *Introduction* section of this paper.

Table 6 shows the conditions under which Figs. 10–18 were developed. It also shows the methods that were used to calculate the fracture conductivity values referenced in the case histories and the actual fracture conductivity values from the sources.

Figs. 10–18 show the comparisons between the predictive performances of the five factor regression models compared with realistic fracture conductivities extracted from the case histories. Sensitivity analysis was carried out for polymer concentration since the references did not report polymer concentration. The dashed green line in Figs. 10–18 represent the true fracture conductivity values from the sources while the red squares represent the prediction from the five factors model. The vertical line with short horizontal bars at each end in the figures is the 95% prediction interval for all the predictions.

More specifically, Figs. 10 and 11 compare actual and predicted proppant type 1 conductivities for Case Histories II and III respectively. In Fig. 10, the prediction interval does not encompass the actual conductivity values at lower polymer concentrations. It does, however, encompass it when the polymer concentration is 60 lb/1000 gal. This supports the hypothesis that the polymer loading used in this well is in the 40–60 lb/1000 gal range. In Fig. 11, it is most likely that polymer was not used during the conductivity measurement procedure. Accordingly, the predictions at the low polymer concentration ranges are a better match when compared the true fracture conductivity value.

Figs. 12–15 provide the comparison between predicted and actual proppant type 3 conductivity values for Case History I. We suspect that for Figs. 12–15 the polymer loading would be in the 30–60lb/1000gal range (since the actual value was not reported in the paper). This is based on the proppant concentration and the time the fracture was pumped (1978–1980). The true conductivity values in Figs. 12 and 13 are as reported from build-up tests from the same well conducted approximately five years apart (see Table 6). In both figures, the prediction intervals from our model encompass the true value, especially when polymer concentration is in the 30–60 lb/1000 gal range.

The true conductivity values referred to in Figs. 14 and 15 are from build-up tests from the same well conducted approximately two years apart. In the earlier build-up (1.5 months after fracture treatment), the fracture conductivity was determined to be 876 md-ft at an effective closure stress of 2750 psi (green line in Fig. 14). In the later build-up (conducted approximately 2 years after the initial build-up), the

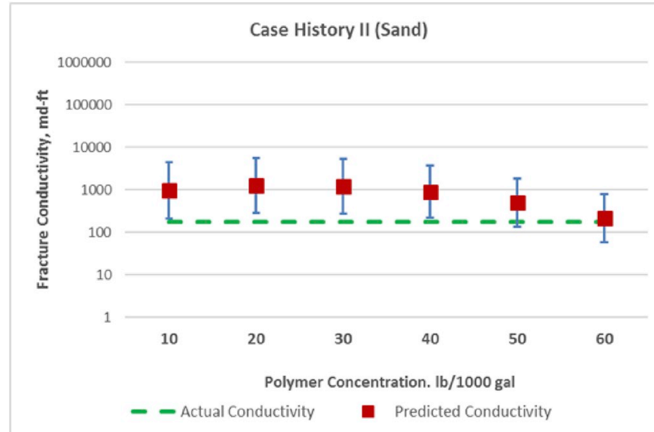

Fig. 10. Comparison of prediction interval and actual conductivity value for proppant type 1 (Case History II) for a variety of polymer concentrations.

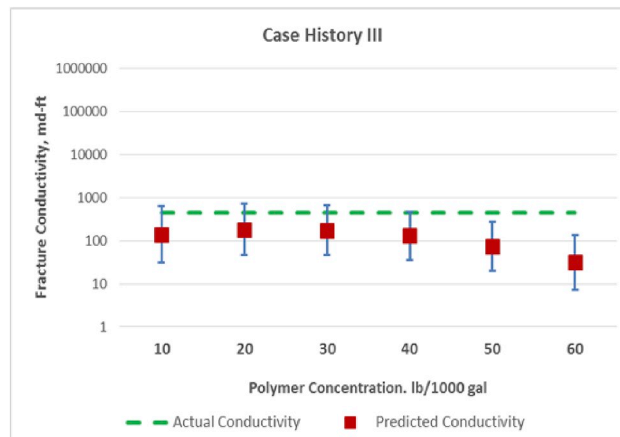

Fig. 11. Comparison of prediction interval and actual conductivity value for proppant type 1 (Case History III) for a variety of polymer concentrations.

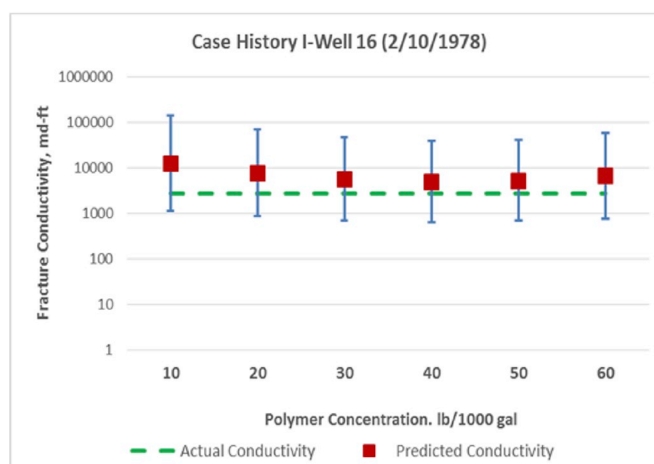

Fig. 12. Comparison of prediction interval and actual conductivity value for proppant type 3 (Well 16, Case History I) for a variety of polymer concentrations.

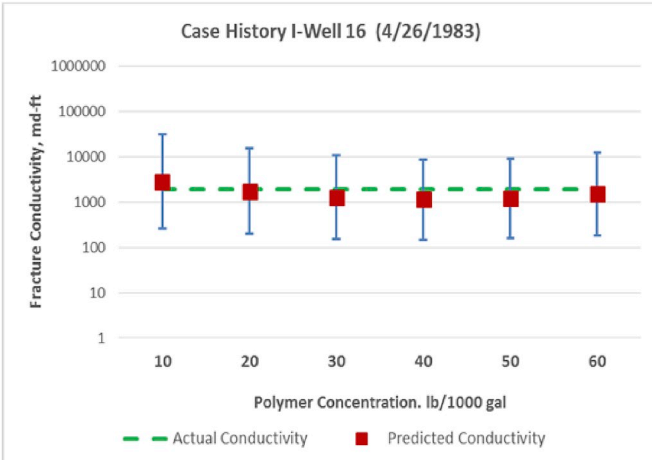

Fig. 13. Comparison of prediction interval and actual conductivity value for proppant type 3 (Well 16, Case History I) for a variety of polymer concentrations.

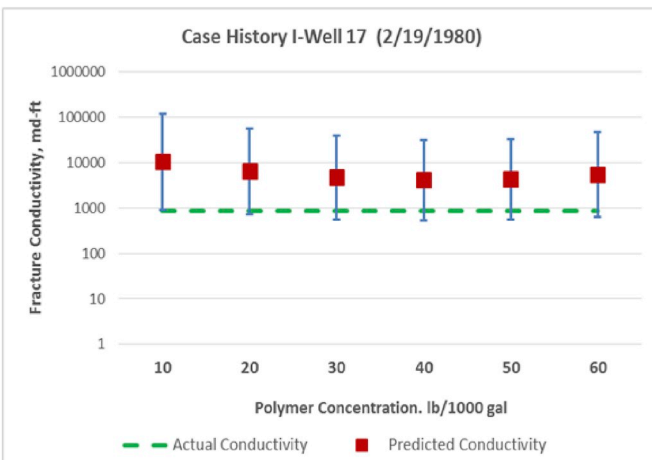

Fig. 14. Comparison of prediction interval and actual conductivity value for proppant type 3 (Well 17, Case History I) for a variety of polymer concentrations.

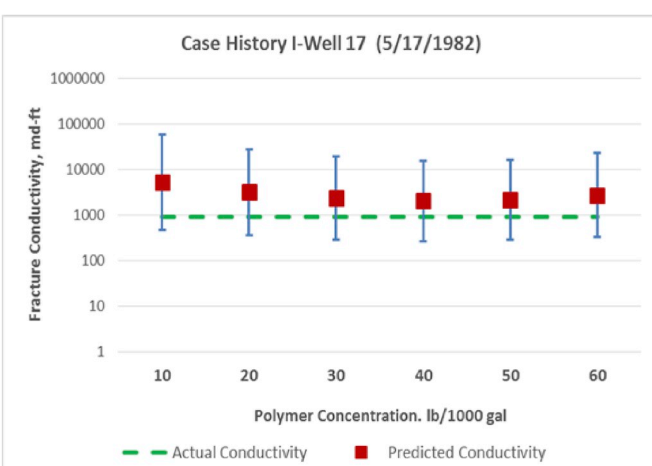

Fig. 15. Comparison of prediction interval and actual conductivity value for proppant type 3 (Well 17, Case History I) for a variety of polymer concentrations.

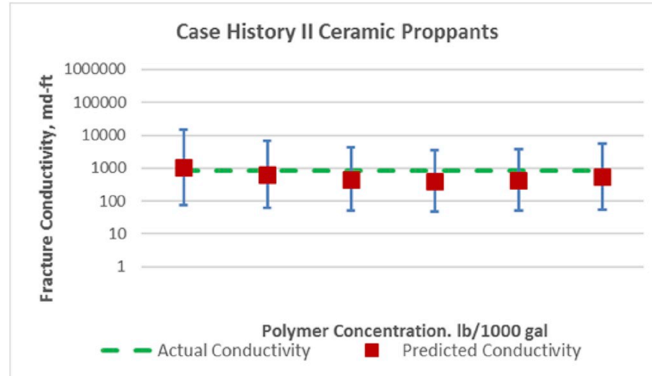

Fig. 16. Comparison of prediction interval and actual conductivity value for proppant type 3 (Case History II) for a variety of polymer concentrations.

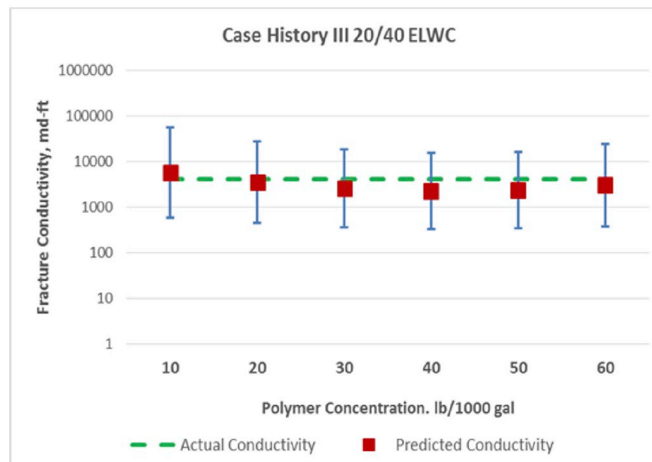

Fig. 17. Comparison of prediction interval and actual conductivity value for proppant type 3 (Case History III) for a variety of polymer concentrations.

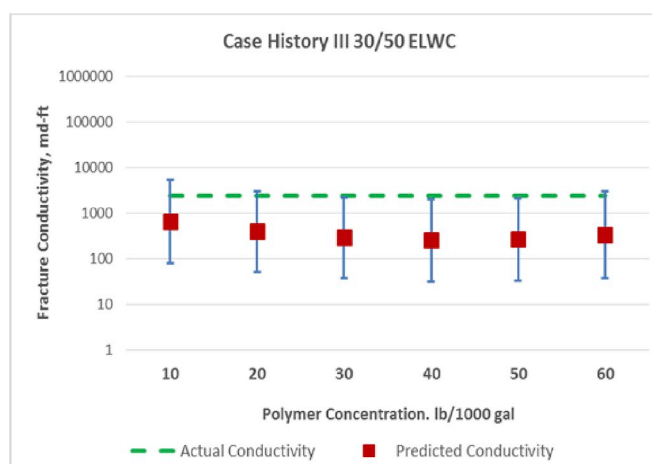

Fig. 18. Comparison of prediction interval and actual conductivity value for proppant type 3 (Case History III) for a variety of polymer concentrations.

fracture conductivity was determined to be 932 md-ft (green line in Fig. 14) at an effective closure stress of 5100 psi. One of these tests is likely to be wrong because it does not make physical sense to have higher conductivity values at a higher closure stress. We hypothesize that the true value in Fig. 14 is inaccurate as can be seen by the mismatch between the prediction intervals from our model and the true value. We believe that the true conductivity value in Fig. 15 is a more

Table 6
Summary table for case histories.

Figure number	Proppant type (1 or 3)	Proppant mesh size	Proppant concentration (Lb/ft2)	Polymer loading (Lb/gal)	Closure stress (psi)	Static reservoir temperature (degF)	'True' Fracture Conductivity Source	'True' Fracture Conductivity Value from source or reference (md-ft)
10	1	20/40	0.5	Varied from 10 to 60 lb/1000gal for all figures because polymer concentration values were not reported.	6500	225	Case History II Palisch et al., 2007 (Predict K)	177
11	1	40/70	2		6000	250	Case History III Handren and Palisch, 2007 (Lab work and Predict K)	446
12	3	20/40	1.08		2000	300	Case History I Lindley, 1983 (Pressure Build up Test)	2783
13	3	20/40	1.08		7200	300	Case History I Lindley, 1983 (Pressure Build up Test)	1900
14	3	20/40	1.083		2750	300	Case History I Lindley, 1983 (Pressure Build up Test)	876
15	3	20/40	1.083		5100	300	Case History I Lindley, 1983 (Pressure Build up Test)	932
16	3	20/40	0.5		6500	225	Case History II Palisch et al., 2007 (Predict K)	864
17	3	20/40	2		6000	250	Case History III Handren and Palisch, 2007 (Lab work and Predict K)	4263
18	3	30/50	2		6000	250	Case History III Handren and Palisch, 2007 (Lab work and Predict K)	2392

accurate representation of fracture flow capacity. This assertion is supported by the fact that the conductivity predictions from our model encompass this true value especially at higher polymer loadings. Essentially, this particular case study shows that models like the ones developed in this work can be used in some scenarios as a quality assurance/control measure for results from diagnostic tests like pressure build-ups.

Fig. 16 makes the comparison for Case History II and Figs. 17 and 18 make the comparison for Case History III, all for proppant type 3. For most of the cases, the predictions are satisfactory. The empirical models are likely to perform least accurately when the values for the new cases are at or lie beyond the variable ranges used in this study (Appendix A). For example, in Fig. 18, the 30/50 mesh size is near the lower limit for the mesh size range for the five factor model and the actual proppant used to estimate the true value in Fig. 18 is the economic light-weight ceramic proppant as opposed to the ceramic proppants data used to develop the five factors model for proppant type 3.

In summary, the case histories provide a separate validation that the models presented here are useful and are anchored in reality. In reference to a famous remark by the statistician George Box, we do not make the claim that these models are right because all models are wrong. We only make the claim that these models are useful in the hands of an engineer that is cognizant of their inherent limitations.

3.7. How to use the models

As an illustration of how these models can be used, we suppose that a well was fractured years ago and for some reason, the case file is incomplete. However, it is still necessary to determine the expected productivity of the well because that would inform the engineer if the well was optimally fractured. One of the important variables an engineer working the problem requires is an estimate of the fracture conductivity.

To illustrate how the engineer can estimate the fracture conductivity to be used in a reservoir simulator, let us suppose that the engineer has the following data: mesh size or median proppant diameter (P_{ms}) of proppant type 1 = 0.01 inches, closure stress (σ_c) = 3,000 psi and reservoir temperature (T) = 200 $degF$ [column (2) of Table 7]. The mean and standard deviation used to calculate the standardized values [column (3) of Table 7] are from the appropriate table in Appendix A. In this case, Table A-1. For this case, the engineer does not have an idea of the proppant concentration anywhere in the fracture.

The standardized values for the mesh size, closure stress and temperature are substituted in the appropriate model from Table 3 (in this case, the three factors model for proppant type 1). We have the following:

$$\ln Fc = 7.21611 + (0.75318{*}\boldsymbol{P_{ms}}) - (0.35250{*}\boldsymbol{\sigma_c}) - (0.02203{*}\boldsymbol{T}) - (1.04327{*}\boldsymbol{P_{ms}}{*}\boldsymbol{\sigma_c}) - (0.64078{*}\boldsymbol{\sigma_c}{*}\boldsymbol{T})$$

$$\ln Fc = 7.21611 + (0.75318{*}{-}1.375) - (0.35250{*}{-}0.527) - (0.02203{*}1.229) - (1.04327{*}(-1.375{*}{-}0.527)) - (0.64078{*}(-0.527{*}1.229))$$

$\boldsymbol{Fc} \approx 403\ \boldsymbol{md-ft}$

Using equation (2), the 95% prediction interval (PI) for the above prediction is 28 md-ft ≤ 403 md-ft ≤ 5779 md-ft. The 95% PI is large because there are many other factors not considered in the three-factor model that can contribute to fracture conductivity.

Table 7
Example calculation I.

(1)	(2)	(3)
Variable	Engineer's dataset (Real values)	Standardized value (variable-mean)/standard deviation
P_{ms}, Mesh size (inches)	0.01	−1.375
σ_c, Closure stress (psi)	3000	−0.527
T, Temperature (degF)	200	1.229

Table 8
Example calculation II.

(1)	(2)	(3)
Variable	Engineer's dataset (Real values)	Standardized value (variable-mean)/standard deviation
P_{ms}, Mesh size (inches)	0.01	−0.605
σ_c, Closure stress (psi)	3000	0.006
T, Temperature (degF)	200	0.961
C_{pl}, Polymer concentration (lbs/gal)	0.01	−0.941
C_p, Proppant concentration (lbs/ft2)	0.5	−0.051

Let us now say that the engineer continues to look for more information and find that a 10lb gel system was used to place the proppant. Let us also say that the engineer has access to either a coupled hydraulic fracture-reservoir simulator or can input the fracture dimensions/properties in a reservoir simulator. This means the engineer has an idea of the proppant concentration as a function of space in the fracture/reservoir simulator and it was determined that the average proppant concentration in a part of the fracture was 0.5 lbs/ft2. The engineer now has enough information to use another empirical model that takes into consideration more of the available data, that is, the five factors model for proppant type 1. Table 8 shows the input data for the five factors model. Again, the mean and standard deviation used to calculate the standardized values [column (3) of Table 8] are from the appropriate table in Appendix A. In this case, the relevant table is Table A-4.

The standardized values for the mesh size, closure stress, temperature, polymer concentration and proppant concentration are substituted in the appropriate model from Table 3 (in this case, the five factors model for proppant type 1). We have the following:

$$\ln Fc = 6.93165 + (1.26923{*}\boldsymbol{P_{ms}}) - (0.90890{*}\boldsymbol{\sigma_c}) + (3.44982{*}\boldsymbol{C_p}) - (0.10642{*}\boldsymbol{T}) - (0.08941{*}\boldsymbol{C_{pl}}) + (0.16575{*}\boldsymbol{\sigma_c^2}) - (1.90421{*}\boldsymbol{C_p^2}) + (0.05304{*}\boldsymbol{T}^2) - \left(0.40768{*}\boldsymbol{C_{pl}^2}\right)$$

$\boldsymbol{Fc} \approx 283\ \boldsymbol{md-ft}$

The five factors model essentially corrects the three factors model for the effect of proppant concentration and polymer loading. The 95% PI associated with this prediction is 73 md-ft ≤ 283 md-ft ≤ 1099 md-ft. The 95% PI for the five factors model is narrower than the 95% PI for the three factors model, reflecting additional information the engineer has included in the computation of fracture conductivity.

The fact that the five factors model produced a predicted fracture conductivity quite distinct from the three factors model is an expected result of the bias that can develop when important regressors are omitted from a predictive model (see, for example, Weisberg, 2014). Although the various models with different subsets of factors presented in Table 3 may provide somewhat inconsistent results for some data, they each represent the best model available for that particular subset of factors. In cases where a user only has access to data in fewer than five factors, the models presented would be expected to outperform any other model with those same factors.

4. Other considerations

4.1. Applicability of empirical models to subsurface conditions

A long-standing criticism of laboratory based fracture conductivity empirical model development is that the results are not correlated with actual (field) or sub-surface fracture conductivity values. One of the authors of this article made a variant of this argument in Awoleke et al. (2016). However, we have shown in this paper, especially the preceding section, that the probabilistic estimates of fracture conductivity from our models do encompass 'true or actual' values of fracture conductivity. We maintain that it is important that researchers continue to attempt to develop new methods for estimating fracture conductivity in the laboratory, but we nonetheless claim that our use of existing data to develop statistically rigorous empirical models of fracture conductivity has successfully achieved the objectives we outlined in the introductory section.

4.2. Dependence of conductivity on time-changing proppant manufacturing practices

It is also important to point out that experimental fracture conductivity estimates might depend on the time epoch in which the experiments were carried out. This is because proppant manufacturing practices might have changed over time. Whether the effect of such shifts over time is material to the results of the models presented here will be addressed in a future article, wherein we endeavor to tease out the effect of data provenance on the resulting fracture conductivity models using mixed effects regression. But for now we observe that this simpler approach performs adequately for many applications.

Nomenclature

BHN	Formation hardness/Brinnel hardness number, M/L^2, kg/mm^2
C_p	Proppant concentration, ML^{-2}, lb/ft^2
C_{pl}	Polymer concentration, M/L^3, lb/1000 gal
ELWC	Economic light-weight ceramic, type of proppant
Fc	Fracture conductivity, L^3, md-ft
ISP	Intermediate strength proppant, type of proppant
k	Number of data set partitions made in K-fold cross validation
MSE	Mean squared error
n	Sample size
p	Number of regressors (not including an intercept) in a model
P_{ms}	Proppant particle size/mesh size, L, inch
RCS/P	Resin coated sand/proppant, type of proppant
t_{crit}	Critical value of t-distribution
t_P	Types of proppant
T	Temperature, T, °F
X	A matrix of $(p + 1) \times n$ values of regressors
X_0	$(p + 1) \times 1$ column vector of new regressors
$\hat{Y}_i$	Predicted natural log of fracture conductivity, L^3, md-ft
σ_c	Closure stress, $ML^{-1}t^{-1}$, psi

5. Conclusions

Three, four and five factors models were developed for sand (proppant type 1) and ceramic (proppant type 3) proppants using an aggregated dataset pooled from the technical literature. A three factors model was developed for resin-coated proppant (proppant type 2). The five factors model proved to be the most useful for predicting conductivity from both sand and ceramic proppants. The three factors model for resin coated sand presented here is to be used with caution because of challenges with model adequacy. The five factor model compared favorably with the Barree at al. model for both sand and ceramic proppants and was also successfully validated with some case histories.

Since the aggregated data discussed in this work come from disparate studies, their response values are generated from possibly distinct experimental setups. Such variability in experimental conditions and parameters may introduce additional study-level variability into the aggregated data which might possibly be accounted for with a more sophisticated regression model than those employed here. In particular, the study-level variability may be amenable to modeling using random effect terms in a mixed modeling framework. As mentioned earlier, we will describe such models in a future article.

Appendix A

Table A1
Mean, Standard Deviation and Range of independent variables for three factor fixed regression model of proppant type 1.

	Mesh Size (inch)	Closure Stress (psi)	Temperature (^{0}F)
Mean	0.032	4574	114
Standard Deviation	0.016	2985	70
Range	0.0059–0.0559	200–1400	70–325

Table A2
Mean, Standard Deviation and Range of independent variables for four factor fixed regression model of proppant type 1.

	Mesh Size (inch)	Closure Stress (psi)	Proppant Concentration (lb/ft^2)	Temperature (^{0}F)
Mean	0.0329	4222	1.47	116
Standard Deviation	0.0171	2553	1.98	78
Range	0.0059–0.0559	200–10000	0.013–10	70–325

Table A3
Mean, Standard Deviation and Range of independent variables for 2nd four factor fixed regression model of proppant type 1.

	Formation BHN (kg/mm2)	Closure Stress (psi)	Proppant Concentration (lb/ft^2)	Temperature (^{0}F)
Mean	40.19	4607	1.84	74
Standard Deviation	27.48	2720	2.47	23
Range	20–100	200–8000	0.02–10.33	70–275

Table A4
Mean, Standard Deviation and Range of independent variables for five factor fixed regression model of proppant type 1.

	Mesh Size (inch)	Closure Stress (psi)	Proppant Concentration (lb/ft^2)	Temperature (^{0}F)	Polymer Concentration (lb/gal)
Mean	0.0149	2988	0.54	126	0.026
Standard Deviation	0.0081	1870	0.78	77	0.017
Range	0.0059–0.0248	500–8000	0.03–2	70–275	0–0.09

Table A5
Mean, Standard Deviation and Range of independent variables for three factor fixed regression model of proppant type 2.

	Mesh Size (inch)	Closure Stress (psi)	Temperature (^{0}F)
Mean	0.0251	7110	132
Standard Deviation	0.0009	3770	46.5
Range	0.0248–0.02815	1000–14000	70–325

Table A6
Mean, Standard Deviation and Range of independent variables for three factor fixed regression model of proppant type 3.

	Mesh Size (inch)	Closure Stress (psi)	Temperature (^{0}F)
Mean	0.0257	7078	168
Standard Deviation	0.00306	3807	70
Range	0.01745–0.03505	900–15000	70–325

Table A7
Mean, Standard Deviation and Range of independent variables for four factor fixed regression model of proppant type 3.

	Mesh Size (inch)	Closure Stress (psi)	Proppant Concentration (lb/ft^2)	Temperature (^{0}F)
Mean	0.0266	6787	2.15	217
Standard Deviation	0.0042	3669	1.71	74
Range	0.01745–0.03505	900–15000	0.10–8.00	70–325

Table A8
Mean, Standard Deviation and Range of independent variables for five factor fixed regression model of proppant type 3.

	Mesh Size (inch)	Closure Stress (psi)	Proppant Concentration (lb/ft2)	Temperature (^{0}F)	Polymer Concentration (lb/gal)
Mean	0.0252	5268	2.44	200	0.040
Standard Deviation	0.0064	3006	2.34	70	0.009
Range	0.01745–0.03505	900–12000	0.25–8.00	70–300	0.01–0.06

Table A9
Mean, Standard Deviation and Range of independent variables for six factor fixed regression model of all proppant types.

	Mesh Size (inch)	Closure Stress (psi)	Proppant Concentration (lb/ft2)	Temperature (^{0}F)	Polymer Concentration (lb/gal)
Mean	0.0196	3980	1.52	155	0.081
Standard Deviation	0.0092	2766	2.16	83	0.050
Range	0.0059–0.03505	500–12000	0.03–8.00	70–300	0.01–0.144

References

Aven, K.N., Weaver, J.D., Tang, T., 2013. Long-term dynamic flow testing of proppants. In: Presented at SPE International Symposium on Oilfield Chemistry, the Woodlands, Texas. Society of Petroleum Engineers, pp. 8–10. April. SPE-164082-MS. https://doi.org/10.2118/164082-MS.

Awoleke, O.O., Romero, J.D., Zhu, D., et al., 2012. Experimental investigation of propped fracture conductivity in tight gas reservoirs using factorial design. In: Presented at SPE Hydraulic Fracturing Technology Conference, the Woodlands, Texas, 6-8 February, SPE-151963-MS. https://doi.org/10.2118/151963-MS.

Awoleke, O.O., Zhu, D., Hill, A.D., 2016. New propped-fracture-conductivity models for tight gas sands. SPE J. 21 (05), 1508–1517. SPE-179743-PA. https://doi.org/10.2118/179743-PA.

Barree, R.D., Miskimins, J.L., Conway, M.W., et al., 2016. Generic correlations for proppant pack conductivity. In: Presented at SPE Hydraulic Fracturing Technology Conference, the Woodlands, Texas, 9-11 February, SPE-179135-MS. https://doi.org/10.2118/179135-MS.

Buckingham, E., 1914. On physically similar systems; illustrations of the use of dimensional equations. Phys. Rev. IV (4), 345–376. https://doi.org/10.1103/PhysRev.4.345.

Cook, R.D., 1977. Detection of influential observations in linear regression. Technometrics Am. Stat. Assoc. 19, 15–18. https://doi.org/10.2307/1268249.

Cooke, C.E., 1973. Conductivity of fracture proppants in multiple layers. SPE J. Petrol. Technol. 25 (09), 1101–1107. SPE-4117-PA. https://doi.org/10.2118/4117-PA.

Cooke, C.E., 1975. Effect of fracturing fluids on fracture conductivity. SPE J. Petrol. Technol. 27 (10), 1273–1282. SPE-5114-PA. https://doi.org/10.2118/5114-PA.

Cutler, R.A., Enniss, D.O., Jones, A.H., et al., 1985. Fracture conductivity comparison of ceramic proppants. SPE J. 25 (02), 157–170. SPE-11634-PA. https://doi.org/10.2118/11634-PA.

Fredd, C.N., McConnell, S.B., Boney, C.L., et al., 2001. Experimental study of fracture conductivity for water-fracturing and conventional fracturing applications. SPE J. 6 (03), 288–298. SPE-74138-PA. https://doi.org/10.2118/74138-PA.

Handren, P.J., Palisch, T.T., 2007. Successful hybrid slickwater fracture design evolution–an East Texas Cotton Valley Taylor case history. In: Presented at SPE Annual Technical Conference and Exhibition, Anaheim, California, 11-14 November, SPE-110451-MS. https://doi.org/10.2118/110451-MS.

Kabacoff, R.I., 2015. R in Action, Data Analysis and Graphics with R, second ed. Manning Publications Co, Shelter Island, NY.

Kainer, C., Guerra, D., Zhu, D., et al., 2017. A comparative analysis of rock properties and fracture conductivity in shale plays. In: Presented at SPE Hydraulic Fracturing Technology Conference and Exhibition, the Woodlands, Texas, 24–26 January, SPE-184877-MS. https://doi.org/10.2118/184877-MS.

Li, K., Gao, Y., Lyu, Y., et al., 2015. New mathematical models for calculating proppant embedment and fracture conductivity. SPE J. 20 (03), 496–507. SPE-155954-PA. https://doi.org/10.2118/155954-PA.

Lindley, B.W., 1983. An investigation of a high-strength proppant tail-in at McAllen Ranch field. In: Presented at SPE Annual Technical Conference and Exhibition, San Francisco, California, pp. 5–8. October. SPE-11935-MS. https://doi.org/10.2118/11935-MS.

Lumley, T., 2017. Based on Fortran Code by Miller, A. Leaps: Regression Subset Selection. R package. https://CRAN.R-project.org/package=leaps, version 3.0.

McGinley, M., Zhu, D., Hill, A.D., 2015. The Effects of Fracture Orientation and Elastic Property Anisotropy on Hydraulic Fracture Conductivity in the Marcellus Shale. In: Presented at SPE Annual Technical Conference and Exhibition, Houston, Texas, pp. 28–30. September. SPE-174870-MS. https://doi.org/10.2118/174870-MS.

Marpaung, F., Chen, F., Pongthunya, P., et al., 2008. Measurement of gel cleanup in a propped fracture with dynamic fracture conductivity experiments. In: Presented at SPE Annual Technical Conference and Exhibition, Denver, Colorado, 21-24 September, SPE-115653-MS. https://doi.org/10.2118/115653-MS.

McDaniel, B.W., 1986. Conductivity testing of proppants at high temperature and stress. In: Presented at SPE California Regional Meeting, Oakland, California, pp. 2–4. April. SPE-15067-MS. https://doi.org/10.2118/15067-MS.

McDaniel, B.W., 1987. Realistic fracture conductivities of proppants as a function of reservoir temperature. In: Presented at Low Permeability Reservoirs Symposium, Denver, Colorado, pp. 18–19. May. SPE-16453-MS. https://doi.org/10.2118/16453-MS.

Much, M.G., Penny, G.S., 1987. Long-term performance of proppants under simulated reservoir conditions. In: Presented at Low Permeability Reservoirs Symposium, Denver, Colorado, pp. 18–19. May. SPE-16415-MS. https://doi.org/10.2118/16415-MS.

Palisch, T.T., Duenckel, R.J., Bazan, L.W., et al., 2007. Determining realistic fracture conductivity and understanding its impact on well performance - theory and field examples. In: Presented at SPE Hydraulic Fracturing Technology Conference, College Station, Texas, 29-31 January, SPE-106301-MS. https://doi.org/10.2118/106301-MS.

Parker, M.A., McDaniel, B.W., 1987. Fracturing treatment design improved by conductivity measurements under in-situ conditions. In: Presented at SPE Annual Technical Conference and Exhibition, Dallas, Texas, pp. 27–30. September.SPE-16901-MS. https://doi.org/10.2118/16901-MS.

Penny, G.S., 1987. An evaluation of the effects of environmental conditions and fracturing fluids upon the long-term conductivity of proppants. In: SPE Annual Technical Conference and Exhibition, Dallas, Texas, pp. 27–30. September. SPE-16900-MS. https://doi.org/10.2118/16900-MS.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rahman, M.R., 2017. Meta-Analysis of Hydraulic Fracturing Data. University of Alaska Fairbanks, Fairbanks, AK (December, 2017).

Rivers, M., Zhu, D., Hill, A.D., 2012. Proppant fracture conductivity with high proppant loading and high closure stress. In: Presented at SPE Hydraulic Fracturing Technology Conference, the Woodlands, Texas, 6-8 February, SPE-151972. https://doi.org/10.2118/151972-MS.

Vlis, A.C., Haafkens, R., Schipper, B.A., et al., 1975. Criteria for proppant placement and fracture conductivity. In: Presented at Fall Meeting of the Society of Petroleum Engineers of AIME, Dallas, Texas, 28 September-1 October, SPE-5637-MS. https://doi.org/10.2118/5637-MS.

Volk, L.J., Raible, C.J., Carroll, H.B., et al., 1981. Embedment of high strength proppant into low-permeability reservoir rock. In: Presented at SPE/DOE Low Permeability Gas Reservoirs Symposium, Denver, Colorado, pp. 27–29. May. SPE-9867-MS. https://doi.org/10.2118/9867-MS.

Weisberg, S., 2014. Applied Linear Regression, fourth ed. John Wiley & Sons, Inc, Hoboken, NJ.

Zhang, J., 2014. Creation and Impairment of Hydraulic Fracture Conductivity in Shale Formations. Texas A & M University, College Station, TX (August, 2014). http://hdl.handle.net/1969.1/153365.

Stim Lab Proppant and Fluid Consortia.