Review

URF PUBLISHERS
connect with research world

Journal of Petroleum & Chemical Engineering

https://urfpublishers.com/journal/petrochemical-engineering

Vol: 3 & Iss: 1

Shale Characterization Approaches with XRD, Cation Exchange Capacity (CEC) and Linear Swell Meter (LSM): Experimental Results and Statistical Assessment

Ekrem Alagoz[1]*, Artur Davletshin[2] and Muhammed Kemal Ozel[2]

[1]Turkish Petroleum Corporation (TPAO), Ankara, Turkey

[2]The University of Texas at Austin, Texas, USA

Citation: Alagoz E, Davletshin A, Ozel MK. Shale Characterization Approaches with XRD, Cation Exchange Capacity (CEC) and Linear Swell Meter (LSM): Experimental Results and Statistical Assessment. *J Petro Chem Eng* 2025;3(1): 71-79.

Received: 30 January, 2025; **Accepted:** 20 February, 2025; **Published:** 24 February, 2025

***Corresponding author:** Ekrem Alagoz, Turkish Petroleum Corporation (TPAO), Ankara, Turkey, E-mail: ealagoz@tpao.gov.tr

ABSTRACT

Shale characterization plays a critical role in assessing its viability as a hydrocarbon reservoir and optimizing hydraulic fracturing strategies. This study investigates the effectiveness of three key techniques—X-ray diffraction (XRD), cation exchange capacity (CEC) and linear swell meter (LSM)—in evaluating shale properties. A set of shale samples from a specific reservoir was analyzed using XRD to determine mineralogical composition, CEC to quantify ion exchange capacity and LSM to assess swelling behavior. The results were subjected to statistical analysis to examine correlations among these characterization methods. Findings indicate that the combined use of clay stabilizers and KCl salt significantly enhances shale stability, while the concentrations of various additives exert both positive and negative influences on swelling behavior. Moreover, CEC values for each formation can be reliably estimated using a statistical approach based on XRD data. This study introduces a robust, data-driven framework for predicting clay swelling, demonstrating the effectiveness of integrating XRD, CEC and LSM for a comprehensive shale characterization methodology.

Keywords: Hydraulic fracturing; Shale rock properties; Unconventional shales; X-ray diffraction (XRD), Cation exchange capacity (CEC); Linear swell meter (LSM)

Introduction

Fracturing fluids, also known as fracking fluids, are liquids used in the process of hydraulic fracturing. They are injected into rock formations at high pressure to create fractures, which allow for the extraction of natural gas and oil. Fracturing fluids typically consist of water, sand and chemicals. The water helps to create and propagate the fractures, while the sand acts as a proppant to keep the fractures open. The chemicals are added for various purposes, such as to reduce friction, prevent bacterial growth and control the pH of the fluid. The composition of the fracturing fluid can vary depending on the specific characteristics of the rock formation and the type of hydrocarbon being extracted. Rock-fluid and fluid-fluid interactions have been studied by many researchers before[1-4]. The fracturing fluid additives used in this study are biocide, surfactant, friction reducer and clay stabilizer. The abbreviations of the chemicals used are given in **(Table 1)**. The general properties of these additives are given below, respectively.

Biocides are added to hydraulic fracturing fluids to kill or inhibit the growth of microorganisms that can be present in the rock formations. These microorganisms can consume the nutrients in the hydraulic fracturing fluid, which can cause the fluid to become less effective. They can also produce gases that can increase the pressure in the fractures, which can cause them

to close. Additionally, they can cause corrosion in the pipelines and other equipment used in the hydraulic fracturing process. Biocides are added to the fluid to kill or inhibit the growth of these microorganisms, which helps to maintain the effectiveness of the fluid and prevent damage to the equipment.

Table 1: Chemicals used to prepare frac fluids.

Abbreviations	Fluid Type	Concentration
DI	DI Water	98-99%
TAP	Tap Water	98-99%
S2	Surfactant	0.1-0.5%
FR3	Friction Reducer	0.1-1%
CSTAB2	Clay Stabilizer	0.1-1%
CIDE	Biocide	0.01-0.1%

Surfactants are used in hydraulic fracturing fluids as they reduce the surface tension between the fluid and the rock, allowing the fluid to more easily flow into the fractures in the rock. They also help to prevent the fractures from closing after the fluid is injected, which helps to keep the fractures open and increases the effectiveness of the hydraulic fracturing process.

Friction reducers, also known as viscosity reducers, are added to hydraulic fracturing fluids to decrease the fluid's resistance to flow, allowing it to move more easily through the fractures in the rock. This in turn allows the fluid to be pumped into the fractures at a higher rate, increasing the efficiency and effectiveness of the hydraulic fracturing process. Additionally, friction reducers can also help to reduce the amount of energy required to pump the fluid into the fractures, which can lower costs and reduce the environmental impact of the process.

Clay stabilizers, also known as clay inhibitors, are added to hydraulic fracturing fluids to prevent the clays in the rock from swelling and closing the fractures created during the hydraulic fracturing process. These clays can absorb water and increase in volume, which can cause the fractures to close, reducing the effectiveness of the hydraulic fracturing process. Clay stabilizers work by altering the properties of the clay minerals, inhibiting their ability to absorb water and maintaining the fractures open. By preventing the clays from swelling, clay stabilizers help to keep the fractures open and increase the effectiveness of the hydraulic fracturing process.

Experimental Methods

As the experimental setup, XRD, CEC and LSM tests were carried out, respectively. Each step was carried out in the procedure described below and the experiments were completed[5-6].

X-Ray diffraction (XRD) for mineralogy

X-ray diffraction (XRD) is a technique used to determine the crystal structure of a material. It is based on the interaction of X-rays with the atoms in a crystal, which causes the X-rays to be scattered in many directions. The pattern of these scattered X-rays can be used to determine the arrangement of atoms in the crystal. XRD is used in a variety of fields, including materials science, chemistry and mineralogy, to identify and study the crystal structures of materials, such as minerals, metals and polymers. It is also used to determine the purity, crystallinity and defects of a material.

Cutting samples taken from the wells, which are in the hydraulic fracturing program for XRD (X-ray Diffraction) bulk powder mineral and clay mineral analyses. The samples ground with Retsch RS-200 vibratory disc mill to have bulk powder and then representatively selected and plated for the XRD bulk powder analyses. After the bulk XRD analyses, two glass slides were prepared from each powdered samples for clay fraction analysis by using the "smear mount method" described in Moore and Reynolds[7]. After the completion of analyses of air-dried slides, same slides left in the 60°C ethylene glycol vapor bath for 2.5 hours and then analyzed. The other slides heated up to 540°C for 2.5 hours and then analyzed. The semi-quantitative XRD bulk powder and clay mineral analyses performed under the conditions given below:

- Generator : Rigaku D/Max-2200 Ultima
- X-ray Tube : Cu
- Voltage : 40 kV
- Current : 20 mA
- Wavelength : (CuKα1) 1.54059 Angstrom
- Scan Speed : 1°/min.
- Software : MDI Jade 7.0

The X-ray diffractograms interpreted based on the Inorganic Crystal Structure Database (ICSD) of International Center for Diffraction Data (ICCD) by using MDI's "Jade 7.0" software. The outputs of the XRD analysis evaluated according to profile-based matching of the software and reference intensity ratios (RIR) by using "Easy Quant" patch of the software. The relative abundances of the bulk and clay minerals in the samples were determined as weight percentages are given on **(Table 2)**.

Table 2: XRD Analysis Results of Cutting Samples.

	XRD Bulk Powder Mineralogy (wt%)								XRD Clay Mineralogy (wt%)						Total %	CEC, (meq/100g)
	Quartz	Feldspar	Calcite	Dolomite	Pyrite	Barite	Siderite	Gypsum	Total Clay+Mica	Illite-Mica	Smectite	Kaolin	Chlorite	Mixed-Layer Illite-Smectite		
sample-1	19	2	0	8	2	6	0	0	63	28	1	20	10	4	100	8
sample-2	23	2	1	2	2	0	4	0	66	16	1	37	12	0	100	6
sample-3	33	7	6	6	3	2	0	0	43	19	1	16	6	1	100	5
sample-4	26	4	9	3	4	4	0	0	50	30	2	10	5	3	100	4
sample-5	39	5	8	7	4	0	0	0	37	29	1	3	1	3	100	3
sample-6	46	4	11	0	3	6	0	0	30	18	1	6	4	1	100	1,5
sample-7	34	4	10	2	0	0	0	3	47	15	0	19	4	10	100	6,2
sample-8	40	5	8	2	0	0	1	4	40	23	0	8	7	2	100	2,5
sample-9	33	5	0	21	0	0	1	5	35	21	0	8	4	2	100	2
sample-10	36	0	23	3	7	0	0	0	31	12	0	12	6	1	100	3,5
sample-11	21	0	35	14	4	0	1	0	25	7	0	9	6	2	100	3
sample-12	22	0	33	6	6	0	0	0	33	19	1	10	2	1	100	3,5
sample-13	16	0	48	10	4	0	4	0	18	3	0	7	3	5	100	5
sample-14	34	0	13	6	6	0	1	0	40	25	0	9	4	2	100	3
sample-15	21	0	2	41	3	0	0	0	33	17	1	9	2	5	100	3
sample-16	31	13	12	4	4	0	1	0	35	25	0	7	2	1	100	2
sample-17	25	0	15	16	5	0	3	0	36	17	0	8	8	3	100	4
sample-18	23	4	17	34	1	0	0	0	21	12	0	4	3	2	100	2,6
sample-19	35	41	1	4	3	0	0	0	16	8	0	4	4	0	100	3,5
sample-20	30	8	3	7	10	0	2	0	40	22	1	10	7	1	100	3

Based on the XRD results of the samples, quartz and clay minerals are dominant mineral types in all samples. Calcite, dolomite and feldspar are common and close to the original abundances of the XRD results of cutting samples of these wells that have presented in previous studies. In terms of clay mineral content, Illite and smectite, which has the highest swelling potential, are the most abundant clay types with minor amounts of kaolin, chlorite and mixed-layer Illite-smectite (I/S) minerals.

Cation Exchange Capacity (CEC) for Activity of Shales

Cation exchange capacity (CEC) is the total capacity of a rock to hold exchangeable cations. The higher CEC value, the more the shale sample tends to exchange cations, which increases shale swelling. CEC is measured in milli-equivalents (meq) of methylene blue dye adsorbed per 100 g dry clay. (Methylene blue solution: 1 mL = 0.01 milli-equivalents, (meq) containing 3.20 g USP grade methylene blue (Cl6Hl8N3SCI) per liter).

According to API RP 13B-1[8] standard, methylene blue capacity of a water-based fluid is an indication of the amount of reactive clays (added bentonite and/or drill solids) present as determined by the methylene blue test. The methylene blue capacity provides an estimate of the total cation exchange capacity of the drilling cuttings. The test for the Methylene blue capacity of cuttings performed according to API RP 13I[9] procedure, which shows the methodology below. Organic materials, if present in the sample, destroyed by oxidation with hydrogen peroxide. The sample is titrated with standard methylene blue solution until the adsorptive capacity is satisfied.

Methodology

- Drill cuttings are grinded.
- Grinded cuttings dried for CEC tests.
- 25 cc %2 Tetra Sodium Pyrophosphate solution added to erlenmeyer.
- 1 gram of cutting sample added to erlenmeyer.
- Erlenmeyer is magnetically stirred and boiled.
- After 10 minutes boiling, 15 cc %3 Hydrogen Peroxide and 1 cc 5N Sulfuric Acid added to solution and boiled 10 minutes more.
- Erlenmeyer kept cool.
- Methylene Blue (MB) solution started adding to erlenmeyer.
- Clay mineral with absorbed MB dropped to Whatman No.1 filtrate paper with a pipette.
- MB solution continued adding up to observing the full-saturated blue circle and turquoise ring near the blue circle.
- Test repeated again in ten minutes later. If turquoise ring occurs again, test is completed.

Cation Exchange Capacity (CEC) results for all samples illustrated in **(Table 3)**. CEC values of the samples vary from 1.5 meq/100 gr to 8 meq/100 gr. Based on these results, reactivity of cutting samples is about 1/9 or 1/10 less than Sodium Montmorillonite Bentonite clay which has CEC with 70-130 meq/100 gr according to literature[10] in **(Table 4)**.

Test Procedure of Lineer Swell Meter (LSM) Test for Shale Swelling

The Linear Swell Meter (LSM) device used for determining shale hydration or swelling by measuring the expansion just in vertical axis of radially confined sample plugs, which exposed to HVFR fracturing fluids. Testing procedure of LSM test is not exist in API Specifications or Recommended Practice documents as a standard procedure. Therefore, performed experimental test results used for comparisons between candidate fluids. Due to the nature of the experimental application and the multiple components availability, measurement inconsistencies may occur in LSM tests. Maximum effort and consideration have made to minimize the measurement inconsistencies and to provide high accuracy, high precision and standardization in all LSM tests.

Table 3: Cation Exchange Capacity Results.

Sample Name	CEC, (meq/100g)
Shale Sample-1	8
Shale Sample-2	3
Shale Sample-3	6
Shale Sample-4	5
Shale Sample-5	4
Shale Sample-6	6,2
Shale Sample-7	1,5
Shale Sample-8	2,5
Shale Sample-9	2
Shale Sample-10	2
Shale Sample-11	3
Shale Sample-12	3,5
Shale Sample-13	3,5
Shale Sample-14	3
Shale Sample-15	3,5
Shale Sample-16	5
Shale Sample-17	4
Shale Sample-18	3
Shale Sample-19	3
Shale Sample-20	2,6

Table 4: CEC Values of Clay Minerals[10].

Clay Mineral	CEC, (meq/100 gr)
Montmorillonite	70-130
Vermiculite	100-200
Illite	10-40
Chlorite	10-40
Attapulgite-Sepiolite	10-35
Kaolinite	3-15

Grace Instrument M4600 HPHT LSM equipment used to perform the tests, which is an automated, dual core, high pressure and high temperature linear swell meter (up to 2,000-psi and 500 °F) as seen in **(Figure 1)**. Dual core/plug compactor was also used to create cylindrical plugs from ground samples by applying 6,000-psi pressure for 3-hours shown in **(Figure 2)**. M4600 includes a windows-based software for data acquisition. Real-time data displayed along with customized charts during a test. Test data exported to Microsoft Excel for reporting after a test is completed.

All LSM tests conducted at ambient temperature. Results recorded as plots of swelling percent versus time in minutes. These tests provide a graphical comparison of multiple inhibitive fluids simultaneously. The M4600 is composed of two independent pressure cells: Cell A and Cell B **(Figure 3)**. Each

cell has its own controls and one can get them worked either individually or simultaneously.

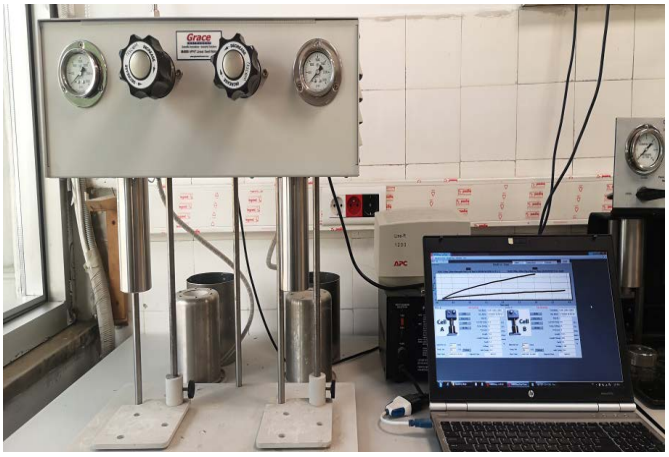


Figure 1: M4600 LSM Equipment.

Figure 2: Dual Core Plug Compactor and Pump.

Figure 3: Pressure Cell Components[11].

Each pressure cell assembly is comprised of the following components.

A. Sample Cup A

B. Sample Cup B

C. Wafer Holder A

D. Wafer Holder B

E. LVDT sensor tip A

F. LVDT sensor tip B

G. Bottom Plate A

H. Bottom Plate B

I. Steel Rod with Top Plate A

J. Steel Rod with Top Plate B

M4600 apparatus uses a LVDT (Linear Variable Differential Transformer) sensor, which detects mechanical linear movement as displacement occurs and converts it to electrical signals. Contact displacement sensors based on this method read changes in the shape of the target by converting it into electrical signals. These requests and guidance include which reservoir sample to test and which brine composition to use. Testing procedure of LSM Test is below.

Methodology

- Drill cuttings are grinded.
- Grinded cuttings dried for Linear Swell Meter tests.
- 10.5-gr sample poured into the steel cylinder to create the plug. The steel cell is hitted manually on a hard surface, allowing the air between the particles to escape and the first compaction.
- Top and bottom acrylic spacers used to compress the cutting particles.
- Plugs created by applying 6,000-psi for 3-hours with compactor pistons.
- Prior to every test, Grace Instrument M4600 HPHT LSM equipment calibration is done by using steel spacers which have exact length of 0.40-in, 0.65-in and 0.90-in for both Cell A and Cell B.
- The height of the plugs measured with a caliper manually.
- Plugs placed in the LVDT sensors and bottom plate tightened.
- LVDT sensor tightened to LSM device.
- The length of the plug, the tested solution information for Cell A and Cell B and the material information for the plug defined to the LSM device.
- 80 ml of solutions poured to pressure Cell A and Cell B. Test started before introducing liquid sample with the plug.
- As soon as test started, pressure cell screwed into the LSM device. Pressure Cell already covers the LVDT sensor in this position.
- Swelling and hydration between the solution and plug starts immediately. Swelling percentage vs time displayed in real time graph on computer.

This method always followed based on the advices of manufacturer manual and our experiences for R&D research studies. Total 65 LSM tests carried out by using mentioned methodology. Test results of Lineer Swell Meter demonstrated in **(Figures 4,5,6,7,8)**.

Results And Discussion

CEC vs LSM

Effect of fracturing fluids on shale swelling: The use of frac fluid in hydraulic fracturing can have a significant impact on the swelling of shale rock. When the fluid is injected into the rock formation at high pressure, it can cause an increase in the volume of the shale rock, known as “swelling.” This swelling occurs because the fluid can interact with the minerals in the rock and cause a chemical reaction that leads to an expansion of the rock. The swelling of the shale rock can have a negative impact on the hydraulic fracturing process. When the rock swells, it can cause the fractures created by the high-pressure injection to close, which reduces the flow of gas or oil from the

well. This can make it difficult to extract the desired resources and it can also lead to damage to the well. The pressure created by the swelling can cause the well bore to become misaligned, which can make it difficult to repair.

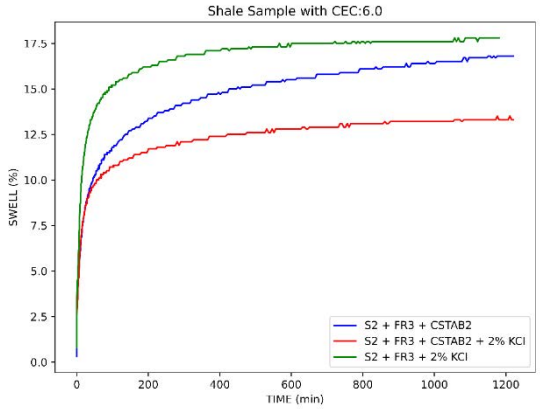


Figure 4: Clay stabilizer with KCI worked better together.

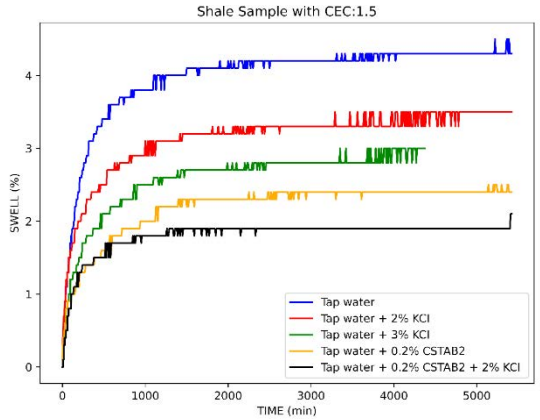


Figure 5: Clay stabilizer with KCI worked better together.

The studies[1-4] on shale-fluid interaction have investigated the impact of different factors on the rate of imbibition, which is the process by which fluids are absorbed into a porous material. One of the key findings from these studies is that the use of clay stabilizers in conjunction with KCI (potassium chloride) can effectively decrease the rate of imbibition. This means that by using these materials together, the fluid is absorbed at a slower rate into the porous material.

These findings were determined through the use of two specific test methods: proppant embedment and NMR (Nuclear Magnetic Resonance). Proppant embedment is a technique that measures the amount of fluid that is absorbed into a porous material, while NMR is a non-destructive testing method that is used to study the properties of fluids. Both of these methods provided detailed information on the rate of imbibition and how it was affected by the presence of clay stabilizers and KCI.

The test setup used in this study also revealed that clay stabilizer and KCI work well together. This means that when these materials are used in combination, they have a synergistic effect, resulting in an even greater decrease in the rate of imbibition than when used individually. Overall, these research studies provide valuable insights into how shale-fluid interactions can be influenced by different materials and can be used to optimize the extraction of resources from shale formations.

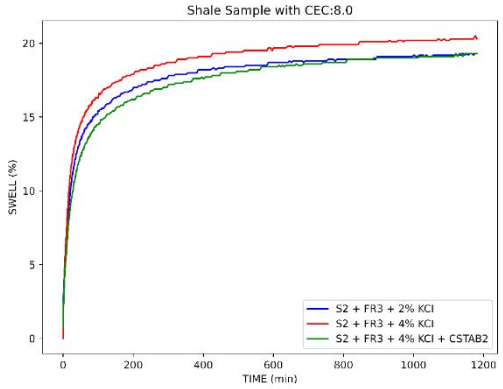


Figure 6: When CEC= 8, the Fracturing fluid effect is not seen much.

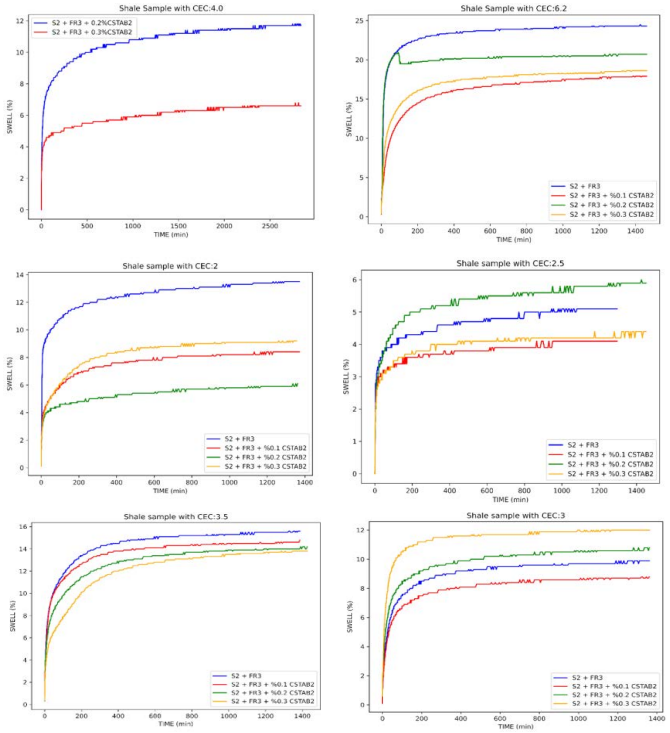


Figure 7: Optimization of Clay stabilizer with different CEC values.

XRD vs CEC

Statistical model: The collected data from experiments and published papers was prepared for performing statistical analysis and generating a comprehensive statistical model for CEC determination with mineralogical composition of shale samples. Multi-linear regression was selected as modelling method as a result of large number of predictor variables. All of the calculations, statistical tests and generation of ANOVA table were performed using Minitab statistical analysis program along with all required plots and graphs. Important steps of the model generation process were explained in details below.

The confidence interval is chosen as % 95 for the model generation. The hypothesis tests performed by using P values, i.e. the P values used as significance criterion of the predictor's parameters included in the model. So, the parameter which have P value greater than 0.05 ($> \alpha = 0.05$) considered insignificant for the model. All predictor variables were included in the modelling phase to investigate the significance level of each predictor variables and generating the best model. The generated preliminary model and ANOVA table were given in Figure 9.

As seen in the ANOVA table most of the predictors did not satisfy the significance condition and most of them have to be eliminated from the model. The multi-collinearity among the predictors were also investigated for the preliminary model with VIF (variance inflation factor) values. As seen in **(Figure 9)** none of the predictor parameters' VIF value greater than 10 which is the threshold value for existence of multi-collinearity. Hence, the model did not suffer from multi-collinearity effect and any treatment (like centralizing the data) for dataset is not required.

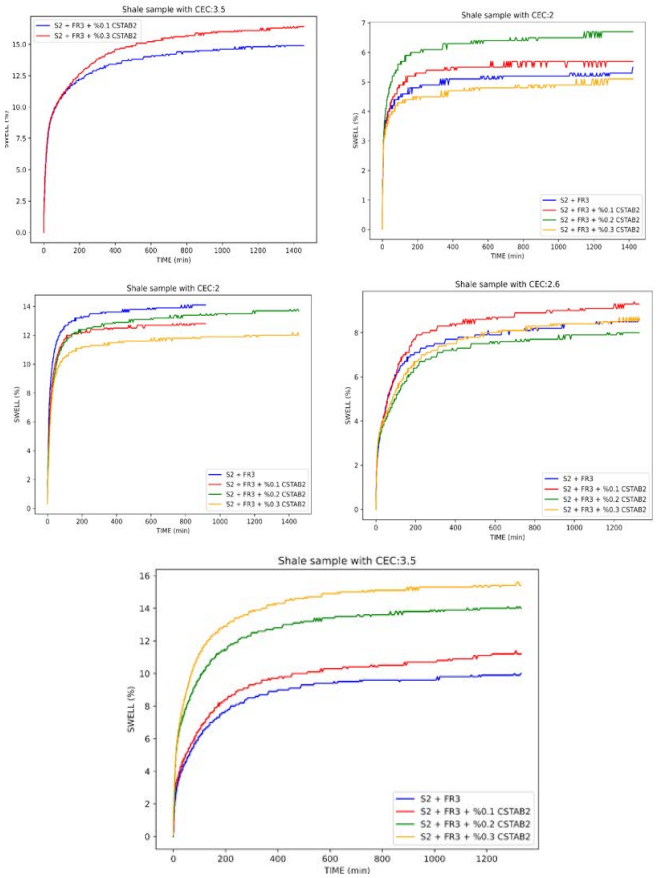


Figure 8: Optimization of Clay stabilizer with different CEC values cont'd.

The resultant normal probability, residual vs. fitted values, residual histogram and residual vs. observation order graphs (four-in-one graph) for the preliminary model were given in **(Figure 10)**. When the normal probability plot is investigated, some point can visually be detected as possible outliers. For differentiation and elimination influential outliers Cook's distance and DFFITS values of each data row was calculated. The DFFITS threshold value was taken as "1" as suggested in the literature. A helpful plot was generated with residual and Cook's distance values to determine threshold value for elimination **(Figure 11)**. The threshold for Cook's distance values visually determined as 0.15, as shown in the **(Figure 11)**. The points fall above these threshold limits considered as influential outlier and deleted from the dataset.

In preliminary model a significant curvature effect, which is an indicator of violence of normality and may be considered as an indicator for presence of non-linearity in the generated model, was also observed in normal probability plot **(Figure 10)**. Hence, square power of each predictor variable was also added to the dataset as predictor variable to eliminate the observed curvature effect. Continuing models was performed with the updated dataset.

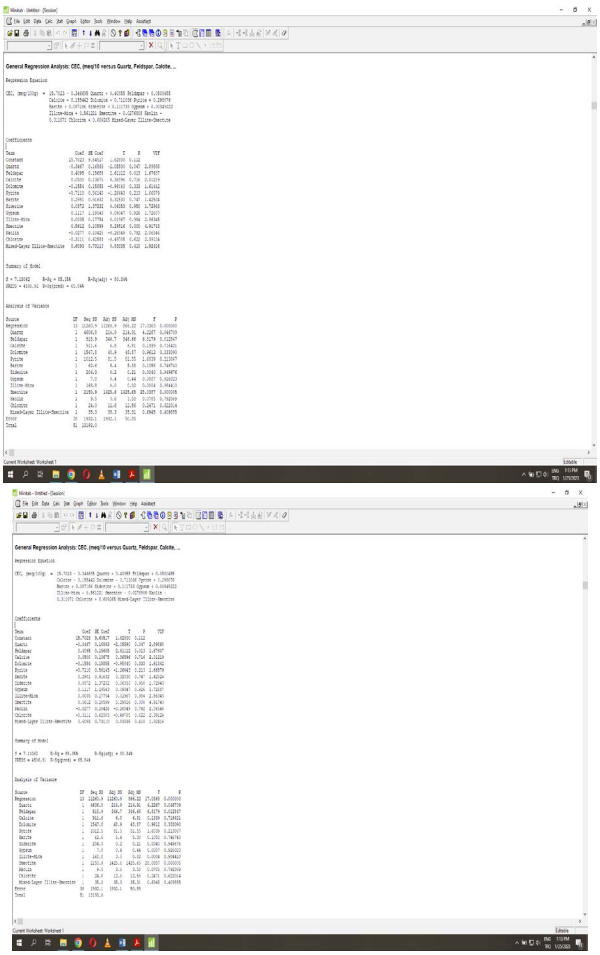


Figure 9: Generated preliminary model and ANOVA table.

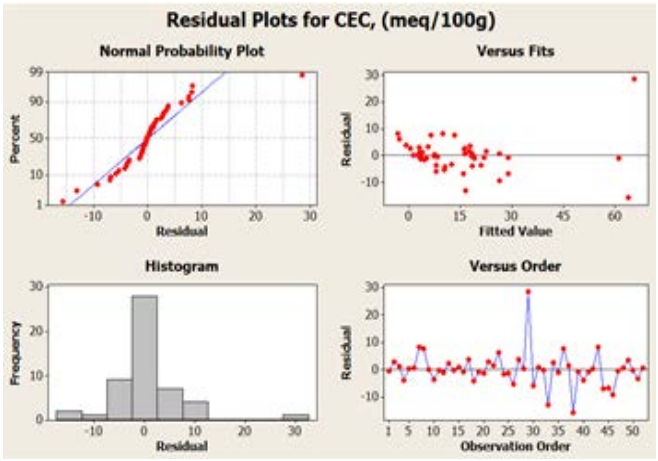


Figure 10: Resultant four-in-one graph for the preliminary model.

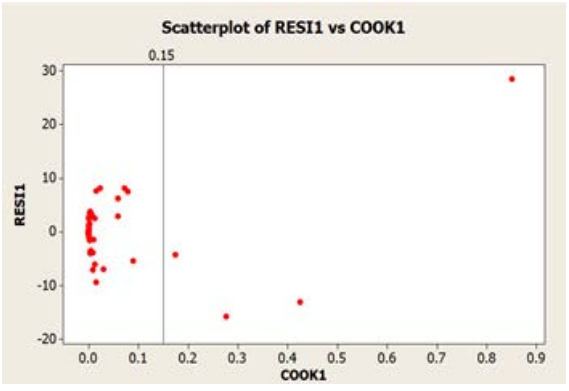


Figure 11: Resultant four-in-one graph for the preliminary model.

Generating the best model with large number of the predictor variables is significantly time consuming when trying the trial-error approach. Hence best subsets algorithm was used for selecting the possible best models. In this algorithm the program takes one of the predictor parameters as free parameter and creates all possible models with remaining predictors. The best model than can be selected using the selection parameters (R2-adjusted, Mallows cp and standard deviation (S)) and existing knowledge and field experience in the target industry. The models resulting from the best subsets algorithm were summarized in **(Figure 12)**. The best options among the generated models can be identified with the lowest difference between "P" and "Mallow's Cp" values, i.e. the closer value indicates the better model. The low standard deviation (S) value is also can be considered a preference factor when selecting the best model. When the generated models investigated, a group of models (13, 14, 15 and 16) were come front as proper models with acceptable "P-Mallow's Cp" difference. The best model among this group is Model-14 with seven predictor parameters (quartz, feldspar, smectite, quartz2, feldspar2, smectite2 and kaolin2). Although Model-14 stands as best option, some modification has to be performed before continuing the analysis with this model. At first glance kaolin has to be implemented to the model since the square power of this parameter included in the model. Although they did not be founded significant and did not include in Model-14 the parameters calcite and chlorite considered as influencer parameters for determining CEC in numerous publications. Hence, these parameters have to be included in the model. The square power of both parameters also added to the model for same reason mentioned above.

After performing the mentioned modifications, Model -14 regenerated and same model remedial operations were performed (testing the predictor parameters influence level and eliminating influential outliers). Same procedure was performed till eliminating all insignificant parameters and influential outliers from the model and the analysis was finalized. The final model includes nine predictor parameters (quartz, feldspar, calcite, smectite, kaolin, chlorite, calcite2, smectite2 and kaolin2) with the R2 and R2-adjusted values are founded as 87.71 and 84.55, respectively. The final model is given in Equation 1;

$$CEC, \frac{meg}{100g} = 11.67 - 0.277 * quarts + 0.247 * feldspar + 0.544 * calcite + 0.981 * smectite - 0.784 * kaolin + 0.963 * chlorite - 0.016 * calcite^2 - 0.014 * smectite^2 + 0.012 * kaolin^2 \quad (1)$$

Model	P	Vars	R-Sq	R-Sq(adj)	Mallows-Cp	S	Quartz	Feldspar	Calcite	Dolomite	Pyrite	Barite	Siderite	Gypsum	Illite-Mica	Smectite	Kaolin	Chlorite	Mixed-Layer	Smectite	Quartz2	Feldspar2	Calcite2	Dolomite2	Pyrite2	Barite2	Siderite2	Gypsum2	Illite-Mica2	Smectite2	Kaolin2	Chlorite2	Mixed-Layer	Smectite2
1	2	1	75.9	75.4	84	7.977										X																	
2	2	1	69	68.3	121.9	9.0504																								X				
3	3	2	81.1	80.4	57.3	7.1287		X								X																		
4	3	2	78.8	78	69.8	7.5472									X	X																		
5	4	3	84.4	83.5	41.2	6.5402		X								X						X												
6	4	3	83.6	82.6	45.7	6.7121	X	X								X																		
7	5	4	87.6	86.5	25.9	5.9006	X	X								X						X												
8	5	4	86.2	85.1	33.3	6.2133		X								X					X	X												
9	6	5	90.6	89.6	11.5	5.1966	X	X								X					X	X												
10	6	5	87.9	86.6	26.1	5.8856	X	X		X						X						X												
11	7	6	91.3	90.2	9.4	5.0383	X	X								X					X	X								X				
12	7	6	91.1	89.9	10.9	5.1169	X	X								X	X				X	X												
13	8	7	92.1	90.9	7.1	4.8577	X	X								X	X				X	X								X				
14	8	7	92	90.7	7.8	4.8954	X	X								X					X	X								X	X			
15	9	8	92.4	91	7.4	4.8156	X	X								X	X				X	X	X							X				
16	9	8	92.4	91	7.5	4.82	X	X	X							X	X				X	X								X				
17	10	9	93.3	91.8	4.8	4.5961	X	X						X		X	X				X	X						X		X				
18	10	9	93.1	91.6	6	4.6712	X	X						X		X					X	X						X		X	X			
19	11	10	93.6	92.1	4.9	4.5297	X	X						X		X	X				X	X						X		X	X			
20	11	10	93.6	92	5	4.5378	X	X						X		X	X				X	X						X		X				X
21	12	11	93.9	92.2	5.5	4.4897	X	X						X		X	X				X	X	X					X		X	X			
22	12	11	93.9	92.2	5.6	4.5024	X	X	X					X		X	X				X	X						X		X	X			
23	13	12	94.2	92.4	5.6	4.4214	X	X						X		X	X				X	X	X	X				X		X	X			
24	13	12	94.1	92.3	6.1	4.4548	X	X	X					X		X	X				X	X		X				X		X	X			
25	14	13	94.5	92.6	6.4	4.3877	X	X						X	X	X	X				X	X	X	X				X		X	X			
26	14	13	94.4	92.5	6.4	4.3938	X	X						X		X	X				X	X	X	X					X	X	X			
27	15	14	94.6	92.5	7.7	4.3955	X	X						X	X	X	X				X	X	X	X		X		X		X	X			
28	15	14	94.6	92.5	7.7	4.3989	X	X						X		X	X				X	X	X	X		X		X	X	X	X			
29	16	15	94.7	92.5	9	4.4091	X	X		X				X	X	X	X				X	X	X	X		X		X		X	X			
30	16	15	94.7	92.5	9	4.4091	X	X		X				X		X	X				X	X	X	X		X		X	X	X	X			
31	17	16	94.9	92.5	10.2	4.4045	X	X						X	X	X	X			X	X	X	X	X		X		X		X	X			X
32	17	16	94.8	92.5	10.2	4.4076	X	X						X		X	X			X	X	X	X	X		X		X	X	X	X			X
33	18	17	95	92.5	11.4	4.4085	X	X					X	X	X	X	X			X	X	X	X	X		X		X		X	X			X
34	18	17	95	92.5	11.5	4.4135	X	X					X	X		X	X			X	X	X	X	X		X		X	X	X	X			X
35	19	18	95.1	92.4	13	4.4382	X	X						X		X	X	X		X	X	X	X	X		X		X	X	X	X	X		X
36	19	18	95.1	92.4	13	4.4412	X	X					X	X	X	X	X			X	X	X	X	X		X	X	X		X	X			X
37	20	19	95.2	92.3	14.5	4.4642	X	X		X			X	X	X	X	X			X	X	X	X	X		X	X	X		X	X			X
38	20	19	95.2	92.3	14.5	4.4645	X	X		X			X	X		X	X			X	X	X	X	X		X	X	X	X	X	X			X
39	21	20	95.3	92.2	15.9	4.4839	X	X					X	X		X	X	X		X	X	X	X	X		X	X	X	X	X	X	X		X
40	21	20	95.2	92.2	16.1	4.5021	X	X					X	X	X	X	X	X		X	X	X	X	X		X	X	X		X	X	X		X
41	22	21	95.4	92.2	17.1	4.4931	X	X		X			X	X		X	X	X		X	X	X	X	X		X	X	X	X	X	X	X		X
42	22	21	95.4	92.2	17.3	4.505	X	X		X		X	X	X		X	X	X		X	X	X	X	X			X	X	X	X	X	X		X
43	23	22	95.4	91.9	19.1	4.568	X	X		X			X	X	X	X	X	X		X	X	X	X	X		X	X	X	X	X	X	X		X
44	23	22	95.4	91.9	19.1	4.5694	X	X		X	X		X	X		X	X	X		X	X	X	X	X		X	X	X	X	X	X	X		X
45	24	23	95.4	91.7	21	4.641	X	X		X	X		X	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		X
46	24	23	95.4	91.7	21.1	4.6455	X	X		X	X	X	X	X		X	X	X		X	X	X	X	X	X		X	X	X	X	X	X		X
47	25	24	95.4	91.4	23	4.7248	X	X		X	X		X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		X
48	25	24	95.4	91.4	23	4.7249	X	X		X	X	X	X	X		X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		X
49	26	25	95.4	91	25	4.8138	X	X		X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		X
50	26	25	95.4	91	25	4.8148	X	X	X	X	X		X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		X
51	27	26	95.4	90.7	27	4.909	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		X

Figure 12: Generated models using best subsets algorithm.

The statistical analysis for the final model, including a regression analysis and an analysis of variance (ANOVA) table, were presented in **(Figure 13)**. As shown in the figure, the P values for all predictor variables included in the model were less than the threshold value, indicating that all predictor variables met the criteria for statistical significance. This means that the relationship between each predictor variable and the outcome variable is likely not due to chance. The regression analysis and ANOVA table were used to examine the relationship between various predictor variables and the outcome variable and the results suggest that these predictor variables have a statistically significant effect on the outcome.

General Regression Analysis: CEC, (meq/10 versus Quartz, Feldspar, Calcite, ...

Regression Equation

CEC, (meq/100g) = 11.6702 - 0.276945 Quartz + 0.247063 Feldspar + 0.544461 Calcite + 0.981397 Smectite - 0.784201 Kaolin + 0.962551 Chlorite - 0.0161412 Calcite2 - 0.0136772 Smectite2 + 0.0122765 Kaolin2

Coefficients

Term	Coef	SE Coef	T	P
Constant	11.6702	3.64483	3.20187	0.003
Quartz	-0.2769	0.08691	-3.18640	0.003
Feldspar	0.2471	0.10915	2.26362	0.030
Calcite	0.5445	0.23214	2.34537	0.025
Smectite	0.9814	0.17947	5.46820	0.000
Kaolin	-0.7842	0.25419	-3.08506	0.004
Chlorite	0.9626	0.32702	2.94339	0.006
Calcite2	-0.0161	0.00720	-2.24039	0.032
Smectite2	-0.0137	0.00483	-2.83142	0.008
Kaolin2	0.0123	0.00500	2.45543	0.019

Summary of Model

S = 3.16629 R-Sq = 87.71% R-Sq(adj) = 84.55%
PRESS = 633.234 R-Sq(pred) = 77.82%

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	2503.71	2503.71	278.190	27.7486	0.0000000
Quartz	1	223.64	101.79	101.789	10.1531	0.0030269
Feldspar	1	511.20	51.37	51.370	5.1240	0.0299030
Calcite	1	48.56	55.15	55.147	5.5008	0.0248047
Smectite	1	1498.18	299.77	299.771	29.9012	0.0000039
Kaolin	1	34.94	95.42	95.418	9.5176	0.0039606
Chlorite	1	6.38	86.86	86.855	8.6636	0.0057308
Calcite2	1	21.97	50.32	50.321	5.0194	0.0315136
Smectite2	1	98.40	80.37	80.373	8.0169	0.0076315
Kaolin2	1	60.44	60.44	60.444	6.0291	0.0191819
Error	35	350.89	350.89	10.025		
Total	44	2854.60				

Figure 13: Generated final model and ANOVA table.

The resultant normal probability, residual vs. fitted values, residual histogram and residual vs. observation order graphs for the final model were given in **(Figure 14)**. It can be seen that the curvature on the normal probability plot line significantly reduced after eliminating insignificant predictor parameters and elimination of influencer outliers. Although there still exist a slight curvature, the normal probability plot satisfyingly good for the constant variance assumption. The remaining curvature may be eliminated via including the interaction between the predictor variables to the model or performing a parameter transformation on response parameter. When the residual vs. fitted plot was inspected no pattern was observed. Hence, it can be concluded that the constant variance assumption is still satisfied after elimination of outliers.

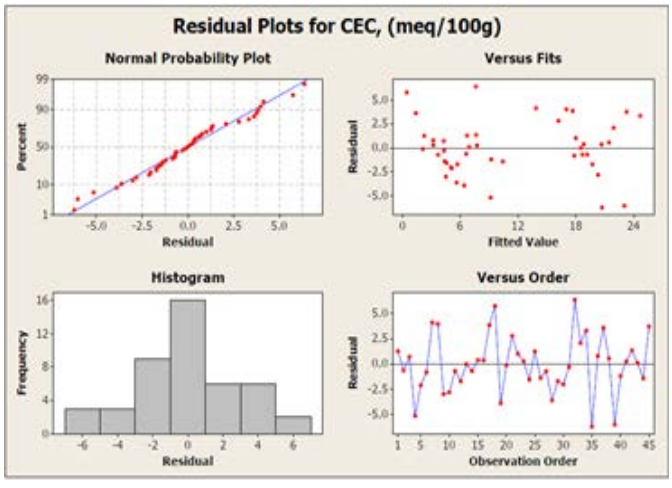


Figure 14: Resultant four-in-one graph for the final model.

Data driven model for shale swelling prediction: Data based model is one of the tools that can be used for petrophysical and technological properties prediction. Linear regression method was used to develop a swelling prediction pipeline enables convenient and robust clay swelling prediction. Mineralogical composition, fracturing fluid type, cation exchange capacity was used to develop a liner regression model. **(Figure 15)** shows the correlation matrix between the all-available features representing the available data. Authors acknowledge the existence of multicollinearity between smectite and cation exchange capacity values and since the model is used for swelling prediction it was decided to keep the multicollinearity.

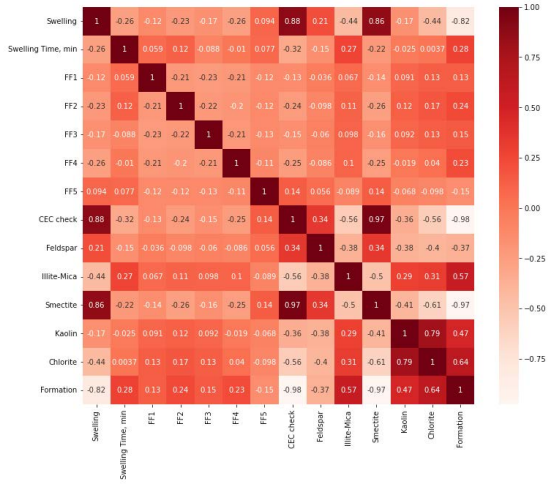
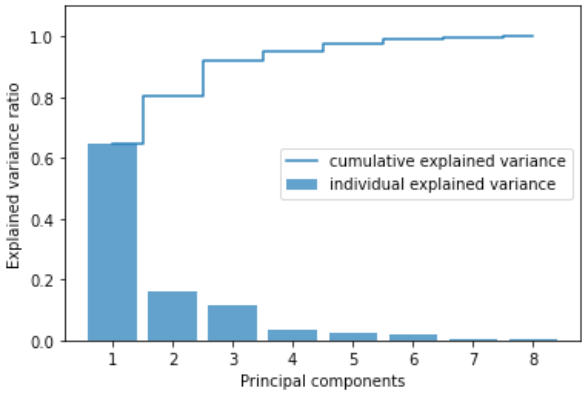


Figure 15: Correlation matrix.

To explore the feature importance and the variance that is carries by the features principal component analysis was performed. **(Figure 16)** shows the principal components ranked by its variance and explained cumulative variance explained. Typically, principal components that explain the 90-95 % of the variance are used to develop a prediction model. Here we use four main components that explain 97 % of the variance to develop our multi linear regression model.



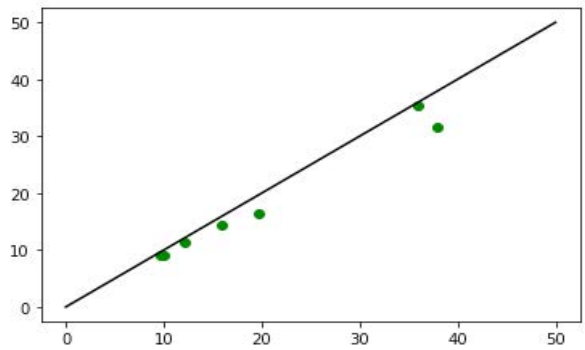


Figure 18: PCA components and explained variance **Figure 19:** Predicted values from

To avoid the state selection bias and since the limited data was available for training the multilinear regression model the 1000 different state were modeled how to split data set to training and test sub datasets. Train to test split was 80 to 20 of the datasets. **(Figure 19)** shows how swelling values were predicted for one single state. The average R2 value for all states equals 0.80 and mean square error value equal 5.79. Thus, swelling can be predicted with high accuracy only using limited available data with multilinear regressing models.

Conclusion

- By looking at the XRD results, the effect of clays on swelling observed independently of the fracturing fluids.
- An increase in CEC values observed according to the number of active clays and the degree of activity.
- Clay stabilizers inhibit more effectively with KCI.
- Presented methods can be used to optimize fracturing fluids before the treatments.
- The equation obtained as a result of statistical analyzes can be used to estimate CEC from XRD results on a regional basis in Dadas Shale formations.
- The most active mineral in the equation created as a result of the statistical model is Smectite. The reason for its high activity is due to its weak bonding and high repulsive potentials on its surface. This unique characteristic allows water to enter between the layers, which in turn causes an increase in the c-spacing. This expanding lattice greatly increases the colloidal activity of Smectite by making all the layer surfaces available for hydration and cation exchange. This results in a significant increase in specific surface, as observed in reference 8.
- Swelling behavior is predicted with high confidence using multilinear regression model.

Acknowledgement

We gratefully acknowledge Turkish Petroleum Corporation (TPAO) for the permission to publish this paper. We would also like to thank Engin Ozgur Ozmen, who helped us with our lab studies.

References

1. Alagoz E, Wang H, Russell RT and Sharma MM. New Experimental Methods to Study Proppant Embedment in Shales. Paper ARMA 2020-1933, 54th US Rock Mechanics/ Geomechanics Symposium held in Golden, Colorado, USA 2020.
2. Alagoz E and Sharma MM. Investigating Shale-Fluid Interactions and Its Effect on Proppant Embedment Using NMR techniques. Paper ARMA 2021-1129, 55th US Rock Mechanics/ Geomechanics Symposium held in Houston, Texas, USA 2021.
3. Alagoz E, Wang H, Russell RT, et al. New Experimental Methods to Study Proppant Embedment in Shales. Rock Mech Rock Eng 2022;55:2571-2580.
4. Alagoz E. Interaction of Fracturing Fluids with Shales: Proppant Embedment Mechanisms. MS Thesis, The University of Texas at Austin, Austin/Texas 2020.
5. Alagoz E, Mengen AE, Yaradilmis Y. Evaluation of XRD, CEC and LSM Methods for Fracturing Fluid Optimization: Experimental Findings. Paper no: 185, 21th International Petroleum and Natural Gas Congress and Exhibition of Turkiye. IPETGAS held in Ankara, Turkiye 2023.
6. Alagoz E, Mengen AE. Shale Characterization Methods Using XRD, CEC and LSM: Experimental Findings. Petro and Petrochem Eng J 2024:8(1).
7. Moore DM and Reynolds RC Jr. "X-Ray Diffraction and the Identification and Analysis of Clay Minerals, Oxford University Press", New York 1997:215-225,
8. API Recommended Practice 13B-1, Third Edition 2003
9. ANSI/API Recommended Practice 13I, Eighth Edition 2009
10. From Grim RE. Clay Mineralogy. McGraw Hill Book Co., New York and Weaver, C.E. and Pollard, L.D., 1973. The Chemistry of Clay Minerals. Elsevier Scientific Publ Co 1953.
11. M4600 HPHT Liner Swell Meter, Grace Instrument Operational Manual 2012.