AADE-22-FTCE-082

Drilling Fluids Recommender System — Artificial Intelligence in Action

Oleh Petryshak, Sergey Makarychev-Mikhailov, Fatma Mahfoudh, Yezid Arevalo, Viktor Antoniuk and Serhiy Seredynskyy, Schlumberger

Abstract

In the current drilling fluid selection process, engineers often decide on a fluid subjectively, based on customer requirements, local practices and regulations, products availability, and personal experience. Such decision does not always rely on the historical data analysis and so might lead to a suboptimal or inappropriate mud selection due to the large number of available options, poor visibility of *relevant* historical data, and complexity of the analysis.

The industry needs an intelligent recommender system to automate the drilling fluid selection process. While the recommender prototype presented in this work is not yet expected to completely replace human engineering decisions, the system can quickly eliminate less favorable options and make recommendations about optimal fluid systems.

The fully automated workflow starts from retrieving nearly 33,000 wells and 90,000 sections from up-to-date historical drilling fluids databases. Vigorous cleaning and processing algorithms are then applied to the data. As an end-user inputs the upcoming (target) interval parameters in a web application, the new data is processed and intervals with similar properties are identified. The offset intervals are grouped by fluid name and reported along with various performance metrics aggregated for each fluid system. The fluids are further ranked by various criteria (prevalence, similarity, average drilling rate of penetration (ROP), fluid complexity, etc.) to assist the engineer with the decision on fluid selection.

The presented drilling fluid recommender enables data-driven technical and business decisions and provides a benchmark for drilling performance under similar conditions. The system becomes a vital part of the higher-level workflow, integrating different drilling technologies, intended to provide holistic optimization of the complex well construction process.

Introduction

Drilling fluid, being the essential part of the well construction process, is a complex system engineered to serve many functions and to satisfy numerous, often contradicting, requirements and specifications. While maintaining wellbore stability and well control are arguably the most important functions of drilling fluids, wellbore cleanup and removing cuttings from the wellbore, sealing permeable formations, transmitting hydraulic energy to downhole tools and the bit, cooling and lubricating the bit, are worth mentioning, among other functions. Of the key requirements, minimum risk to personnel, environment, and drilling equipment are outstanding. However, operational efficiency, low-maintenance, and cost are next in the long list of benefits. Finally, drilling fluid specifications cover dozens of physical and chemical properties as density, rheology, filtration rates, phase ratios, chemical composition, etc.

Drilling fluids are traditionally classified into two big groups, as water-based muds (WBM) and nonaqueous fluids (NAF), within the latter group, oil-based muds (OBM) and synthetic-based muds (SBM) are often distinguished. Historically, NAFs were better performing fluids, however, as they are not as environmentally benign as their WBM counterparts (Friedheim et al. 1991), the industry strives for replacing NAFs with more sustainable water-based systems. Substantial progress has been reported in engineering of WBMs to match or even outperform NAFs (Patel et al. 2007; Young and Friedheim 2013).

The drilling fluid design is often considered to be an art. The artists (product development engineers) formulate complex systems, using various chemical additives carefully selected from the inventory of many hundreds of available items. The fluid must demonstrate required/predictable properties in a wide range of conditions, which are adjustable upon addition of functional chemicals. Yet it must remain stable during the prolonged circulation at temperature and pressure, and in contact with reservoir solids and fluids. The drilling fluids are usually designed with specific requirements related to formation type, logistics, location, etc. It is common for service companies to have many dozens of fluid systems in their portfolios to cover the variety of drilling conditions.

When project or drilling fluid engineers select a fluid system for the next interval (section) of a well, they traditionally rely on customer requirements, local practices and regulations, products availability, and personal experience. These are all important reasons, which can narrow down to available options. However, fluid selection decisions can still be subjective and there is the risk of making a suboptimal or inappropriate choice.

The E&P industry has accumulated a large amount of data over the decades, which is now actively used in data-driven decisions and various optimization workflows. The progress of the industry in digitalization is steady, but it is not equally fast when it comes to different oilfield technologies. Drilling fluids digitalization is perhaps a little behind other well construction technologies. However, with the advent of new generation digital tools, online access to cross-domain databases, extensive data engineering efforts, focusing on database mapping, data cleaning and processing, data visualization and analytics, the drilling fluids selection process rapidly advances into the digital world.

Our recently published results have demonstrated how profiling of drilling fluids by a magnitude of physical properties (Whyte et al. 2022) can help to identify WBMs closely matching properties of NAFs, which is an important step in switching to more environmentally benign systems. We have further utilized the “big data” approach to evaluate drilling performance exhibited by WBMs versus NAFs in many places in around the world and used vigorous statistical analysis to support our conclusions on where WBMs can outperform NAFs (Khvostichenko et al. 2022).

The present work is a continuation of the effort to democratize drilling fluids data and enable data-driven technical and business decisions. Here, our focus is on developing a prototype of a recommender system for drilling fluid selection for a target interval (section), based on user’s input and historical drilling data.

Recommender systems is a popular class of artificial intelligence (AI) tools, implemented in many consumer applications. Music, video, and book recommendation services, product-buying recommendations, social platform content recommenders, and dating match advisers are just a few examples around us.

Recommender systems are tentatively divided into two classes, as collaborative filtering and content-filtering ones. The former systems are based on a database with historical user preferences, which can be used to predict additional topics or products a new user might like (Breese et al. 2013). The second, content-filtering systems, are based on a set of items that are known to be of interest to the user, such as a set of items previously purchased by user (Linden et al. 2001). The recommender system presented in this work belongs to this type. Different modifications of the above classes, hybrid systems, and new recommender systems continue to appear in literature, as the topic remains very popular in data science. For example, the so-called multi-criteria recommender systems (MCRS) concept (Adomavicius and Kwon 2015) is particularly attractive, as it allows the user to formulate not just a preference, e.g., for a movie, but also considers various movie aspects (story, acting, visual effects, etc.). The MCRS concept has been used in the present work.

It must be noted that the developed Drilling Fluids Recommender (DFR) system is still a prototype, which is reported here to share initial (promising) results with the drilling community and initiate discussions with domain experts. Several similar recommender systems have already been developed, e.g., for bits, power sections, bottomhole assemblies (BHA), which are currently in the field-testing phase. Furthermore, our ultimate goal is to merge as many of the well construction technologies as possible to develop a holistic recommender system, which will provide guidance on optimal combinations of the technologies, e.g., bit–BHA–motor–fluid–etc. It is often difficult to digitize and solve many well construction challenges with artificial intelligence tools, however, this is an exciting avenue for digital research, which will undoubtedly change the way we drill wells.

Methodology

The major service company’s global drilling fluid database, called ONE-TRAX Central 2.0 (referred below as “Fluid database”), was used in this study. Fluid database front end allows field engineers to enter data daily through the friendly graphical user interface, and in the back end has a complex schema with data spread across many tables stored in a relational database. Fluid database is used as a primary database for drilling fluid related technical/engineering, inventory and accounting data. As of January 2022, the database contains information on 33,000 wells drilled since 2015 all around the globe with over 90,000 intervals (well sections) and over 850,000 daily reports, the numbers are growing

A custom data engineering workflow was developed to process Fluid database to enable the straightforward data consumption and data analytics. Structured query language (SQL) and Python languages were used to query and join data from different original tables, clean and process data, pivot and/or aggregate values, remove outliers and irrelevant records, etc. The workflow was automated using the Dataiku Data Science Studio ver. 9.0 deployed in the cloud. Many clean tables generated with the data workflow are currently used for various business and technical analytics purposes, including data visualization with Power BI and Tableau dashboards. More details on the workflow and examples of how clean data are used can be found elsewhere (Whyte et al. 2022; Khvostichenko et al. 2022).

The data engineering workflow further generates the clean interval table. It is designed specifically for the DFR and includes the extended summary of well and interval properties (75 attributes). Some of the attributes are inherited from the higher level well table (e.g., well name, location, operator, etc.), and some are aggregated for whole intervals using multiple lower-level daily tables (e.g., numbers of additives used, median values of key drilling fluid properties, sums of time of different activities, etc.). Some attributes are further normalized per interval length (divided by 1,000 ft). The rigorous data cleaning with ruthless rejection of many irrelevant intervals (e.g., those drilled with spud muds), poorly populated ones and those not passing several quality control checks, resulted in significant data attrition. Just over 42,000 intervals remained in the clean interval table whose records, however, are considered to be the most accurate and reliable. The data engineering script is automatically re-run weekly to keep the source table up to date.

The DFR application was fully written in Python ver. 3.9 using external libraries, as pandas (McKinney 2010), Plotly (Plotly Technologies Inc. 2015), scikit-learn (Pedregosa et al.

2011) and a few others. The Streamlit ver. 1.4 library (Streamlit 2022) was used to build the web application with graphical user interface. Streamlit is an open-source Python library that makes it easy to create web applications with literally a few lines of code and is very useful for building fully functional prototypes. The DFR web app is deployed and runs on Microsoft Azure, retrieving the source table from a SQL server.

The end user is required to enter the set of target interval key parameters (features). The list is rather basic at the moment but is expected to expand in the future. The nine key parameters include: country, well water depth, interval top measured depth (MD), interval top true vertical depth (TVD), interval length, interval bottom TVD, bit size, expected maximum inclination and expected maximum bottomhole circulation temperature (BHCT). The interval parameters are immediately processed with the engine, as a) values are converted to oilfield units of measure, if necessary (oilfield units are used as default in the DFR); b) the interval bottom MD is calculated; c) depths and interval length are checked for consistency (with a warning issued if the check fails); d) country latitude and longitude coordinates are retrieved and assigned to the target interval. The resulting 11 interval parameters (features) now all are represented by continuous variables.

In the next step, offset intervals in the historical database are identified, which are intervals most similar to the target one in the multitude of parameters. Different approaches were tested for identification of the offset intervals.

Clustering analysis is an unsupervised machine learning technique focused on finding natural groups in the feature space of input data. A cluster is defined as an area of density in the feature space where samples (i.e. intervals) are more similar to each other that those in other clusters. Clustering was performed on the historical intervals data and multiple combinations of clustering algorithms and numbers of clusters were evaluated. Different clustering efficiency/consistency metrics were screened, for example, the Elbow method was used to find the optimal number of clusters for algorithms that required pre-set numbers. Silhouette was eventually selected as the main metric to compare different clustering algorithms. Out of several algorithms tested (k-means, agglomerative clustering, Gaussian mixture, isolation forest, and a few others), the k-means method was found most promising for the interval clustering.

The general idea was that a preoptimized clustering method can be applied to the source table, so each historical interval is assigned with a cluster label. The clustering was intended to be performed on every (weekly) database update offline, as this is a computationally intensive process. The target interval, once entered by the end user, is quickly classified into a particular cluster, and so all intervals in the target cluster are considered offsets. Additional filtering can still be possible within the cluster to refine the selection of offset intervals for further analysis.

While the clustering approach was implemented in the early prototype, it revealed several shortcomings. First, the approach did not allow any end user control over clustering, as changing the algorithm or the number of clusters was not possible on the fly due to computational time. Therefore, the pre-defined clustering method had to be always used, which was not necessarily optimal for all intervals of interest. Second and the most important, interval clustering resembled a black box, as it was not clear to the end user why and how the target interval was assigned to a particular cluster. In addition, it was possible that target intervals, especially those located near the cluster border, were still closer to historical intervals belonging to a neighboring cluster. However, reclassifying target interval is difficult to implement within the clustering paradigm.

While the clustering approach was found useful for certain drilling fluids analytics applications and it still finds its applications in other recommender systems, clustering was rejected as the DFR offset interval selection method.

The selected approach was based on the multidimensional distances. The idea is that the target interval and all historical intervals can be represented as points in the 11-dimensional feature space. The distance between any two points (intervals) is the measure of similarity, as the smaller the distance the more similar are the intervals. The offset interval similarity (OIS) index was further introduced, which is a derivative of the distance, and has a range between 0 for the most different and 1 for identical intervals. A number of performance and other interval metrics are further calculated by the engine and reported in visual and tabular forms as described in detail in the following section.

Results and Discussion

The Drilling Fluid Recommender web application home page screenshot is shown in **Fig. 1**. It contains a high-level overview of the workflow, release notes (being a prototype, the application is updated monthly with introduction of new features and options) and a quick start note. The left-hand side bar provides fields for the target interval parameters input.

The end user is expected to enter target interval parameters for the upcoming drilling job and submit them to the engine. The OIS index is calculated on the fly for all historical intervals based on the input. The preselected offset intervals with the default OIS index of 0.9 then appear on the screen. The index can be changed by the user to focus the analysis on more (or less) similar offset intervals when moving the slider. Reducing the OIS index value obviously returns more intervals, but with less similarity, and setting the index to zero allows the user to review all intervals available in the database.

Once the OIS index is defined and the user satisfied with the initial number of intervals for analysis, additional filters can be applied to narrow down the offset interval selection. The following filters can be used: drilling fluid (mud) type (WMB or NAF), well offshore or onshore, the service company office, operator, drilling fluid density (mud weight), and the drilling year

After applying additional filters, the user can start analyzing the results. First, a series of pie charts is shown giving the offset interval selection overview **(Fig. 2)** to simplify assessment of interval distributions by key attributes, as drilling fluid families, countries, and operators.

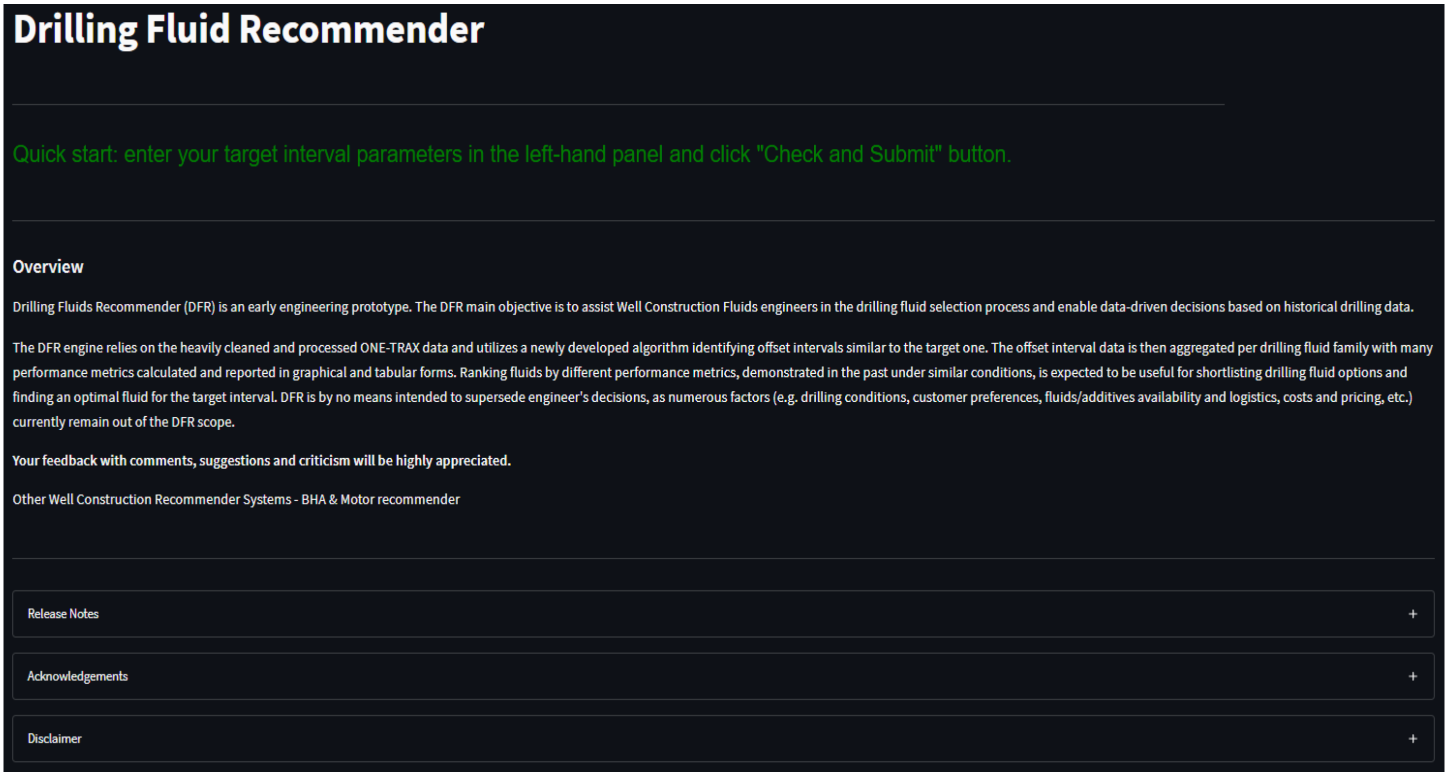


Fig. 1 — Web based Drilling Fluid Recommender (DFR) application home page.

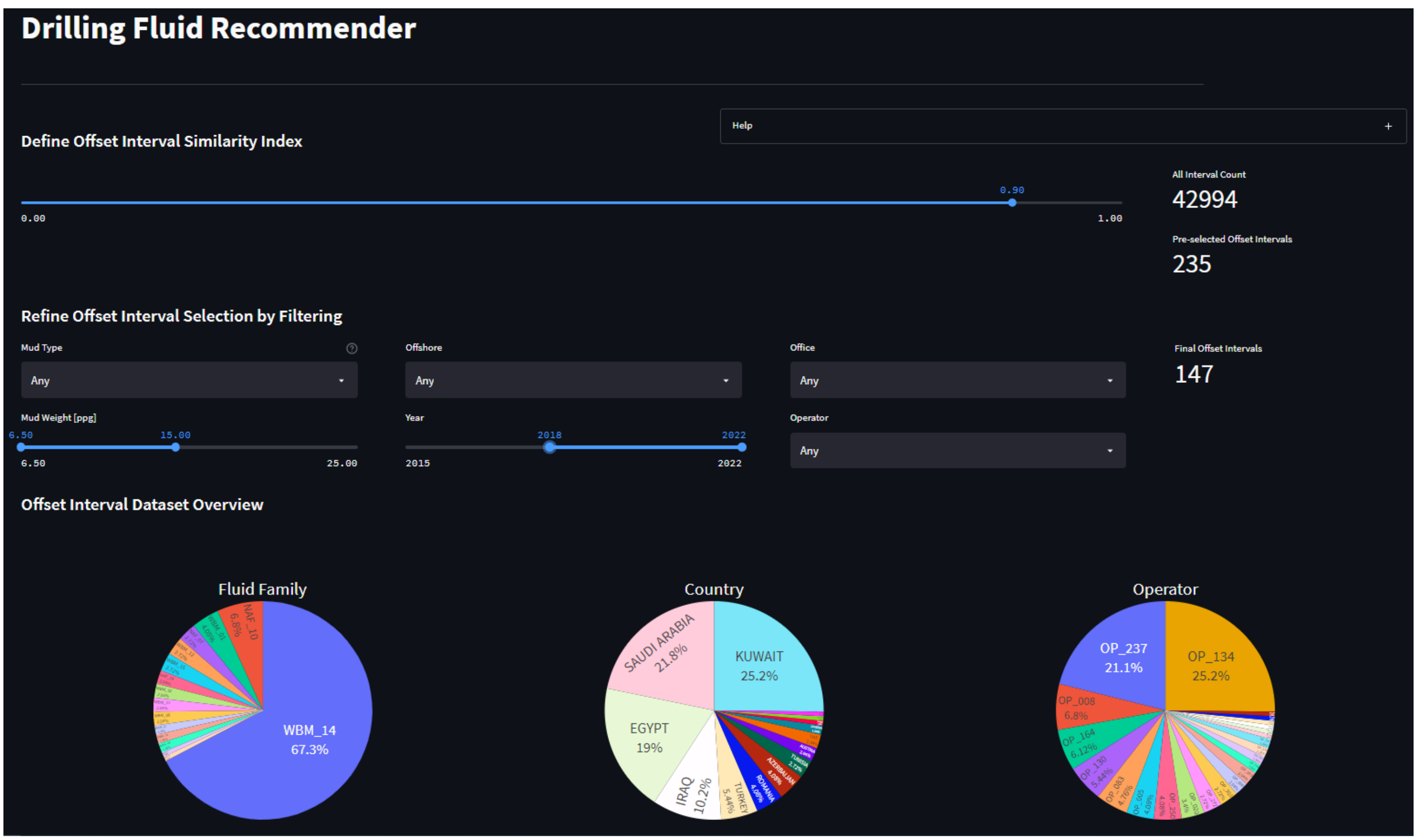


Fig. 2 — Offset interval dataset overview.

The following table is the actual drilling fluid recommendation summary and offers the main outcome of the DFR **(Fig. 3).** Drilling fluid families are ranked according to the key performance factors, calculated based the offset interval historical drilling data, averaged/aggregated per fluid family.

The ranked performance factors are:

- «Prevalence» reflects the relative number of offset intervals drilled with each fluid (another representation of the fluid family pie chart above);
- «Similarity» ranks fluids in terms of how similar they were to the intervals drilled with them and the target interval;
- «Drilling ROP» ranks fluids in terms of drilling rate of penetration (see detailed explanation below);
- «Remedial Time» reflects the relative time spent on various remediation activities while using different fluids;
- «Fluid Treatments» ranks fluids by the number of chemical additions (normalized by drilled distance);
- «Fluid Complexity» ranks fluids by the number of individual chemical products used in fluid treatments (normalized by drilled distance);
- «Additives Cost» ranks the total cost of chemical products used in fluid treatments (normalized by drilled distance);
- «Average Rank» is the arithmetic mean of all of the metrics above.

Data in the recommendation summary table can be sorted by different factors (the lower the rank–the better) individually or using the average rank value. This is a simple implementation of the multi-criteria recommendation, as in MCRS. The table is expected to assist users in selecting optimal drilling fluid for the target interval, depending on their priorities.

It should be noted that performance factors are estimates with variable degrees of uncertainty depending on the original data quality, reporting practices and drilling conditions. "Fluid Treatments", "Fluid Complexity" and "Additives Cost" must be treated with particular care, as under different contracts and scenarios (whole mud, rented mud, customer products, returns, etc.) that account for mud treatments and side-by-side comparisons can be challenging on such a high level. It is, however, perceived that a simplified and unitless ranking is still beneficial, because it could allow quick shortlisting of available drilling fluid options based on historical performance. The in-depth analysis of offset interval is still possible with the DFR provided data, which gives the end user a more complete, yet more complex picture.

In the next module, the DFR further provides detailed insights on drilling performance and activity time distribution **(Fig. 4)** for particular fluid systems used to drill the offset intervals.

Drilling Fluid Recommendation Summary Table

Help +

Ranked performance factors, see warning under Help button.

	Fluid Family	Prevalence	Similarity	Drilling ROP	Remedial Time	Fluid Treatments	Fluid Complexity	Additives Cost	Average Rank
10	WBM_14	1	8	4	14	8	8	7	7
2	NAF_10	2	12	9	4	3	10	12	7
4	WBM_01	3	10	12	11	6	11	11	9
8	WBM_08	4	5	17	17	11	2	4	9
9	WBM_12	5	14	1	8	5	4	3	6
16	WBM_21	6	7	8	13	2	3	10	7
11	WBM_15	6	10	9	9	9	16	14	10
0	NAF_07	8	15	5	10	14	13	5	10
5	WBM_02	8	16	12	6	10	18	9	11
7	WBM_06	11	8	14	16	13	14	13	13
3	NAF_14	11	2	5	15	16	8	8	9
17	WBM_22	11	1	12	5	17	15	15	11
13	WBM_17	13	12	14	7	7	5	2	9
6	WBM_04	14	6	7	2	1	1	1	5
14	WBM_19	14	3	2	12	12	6	17	9
1	NAF_09	17	4	16	2	15	12	6	10
12	WBM_16	17	18	18	18	18	17	16	17
15	WBM_20	17	16	3	2	4	7	18	10

Download Recommendation Summary Table

Fig. 3 — Drilling fluid recommendation summary.

Fig. 4 — Drilling fluid ROP and activity time distributions.

The rate of penetration (ROP) is presented as the main drilling performance metric in two values. Both ROP metrics are aggregated for fluid families (top 20), as median values for the offset intervals.

- «Interval Average Drilling ROP» is calculated for each interval using its length and *drilling time*, as reported in Fluid database daily activities time distribution. This is the most accurate estimate of the ROP;
- «Interval Average ROP» is calculated for the *whole interval* dividing the interval length by the total number of days reported for this interval, no matter what was actually happening (e.g., nonproductive time (NPT), etc.). This is a less accurate, yet meaningful metric, which adds all nondrilling events into consideration.

The Time Distribution plot visualizes reported activities, which are aggregated for individual fluids families (top 20) and normalized to the total of 100%. The activities, not only include drilling and tripping, but also numerous specific remedial ones, such as hole conditioning, stuck pipe, NPT, etc.

In the following module, the fluid treatment metrics are aggregated per fluid family, as median values of the offset intervals as shown on **Fig. 5.** These are characteristics of the fluids themselves.

This is still a relatively high-level overview, which may miss certain details, however, it is important to highlight fluid systems, which have exhibited certain undesired activities in the past, e.g., stuck pipe, or remedial hole cleaning that might be related to the fluid performance.

- «Number of Treatments» is a number of Fluid database reported chemical addition transactions, normalized per 1,000 ft of interval length. This is a basic indicator of the drilling fluid reliability and maintenance demand.
- «Number of Products» utilized in fluid treatments (also normalized per 1,000 ft) is an indicator of the drilling fluid complexity. The products here are chemical additives, branded or nonbranded chemicals. The metric must be treated with care, particularly for nonaqueous fluids, as preformulated muds are often used, which will have a smaller number of reported products. Also, in case of severe subsurface mud losses numerous lost circulation materials (LCM) can show up artificially inflating the number.
- «Total Additive Cost» is the cost of all chemical additions (in US dollars) reported per interval, which value is normalized by 1,000 ft. The numbers are still approximate, as reported costs of individual chemicals may vary depending on location, some additives can be provided by operators, etc. Again, the in-depth analysis is possible, if necessary, to get the full picture on the drilling fluids' cost structures.

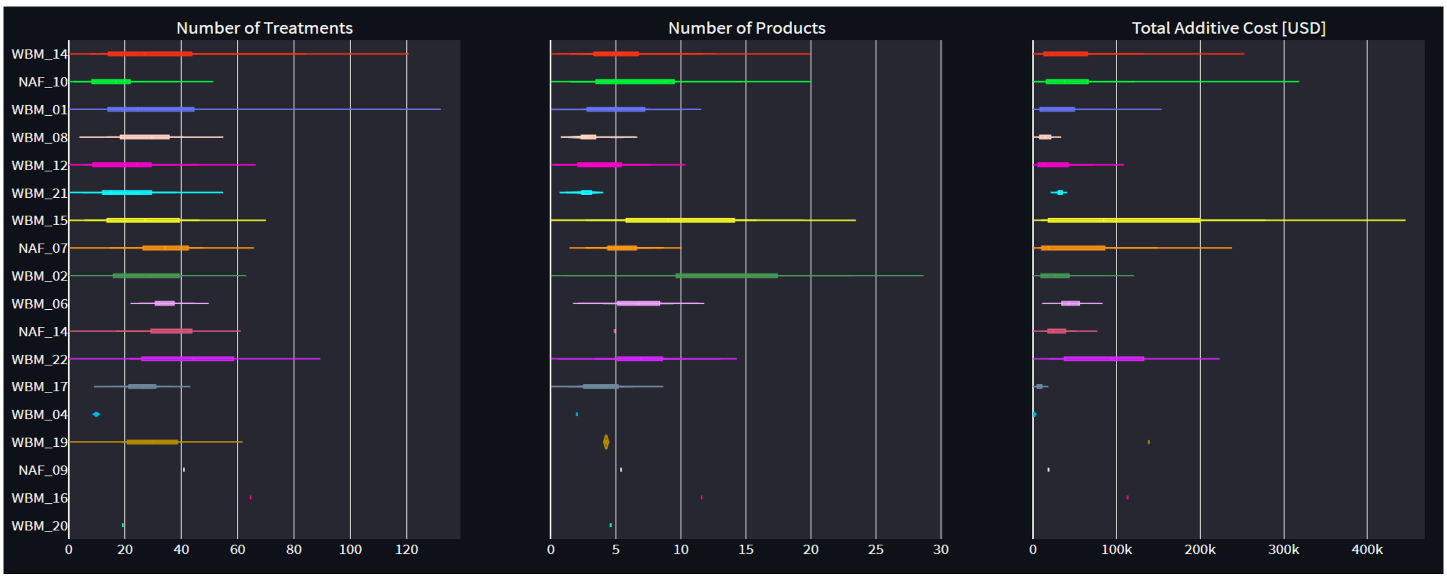


Fig. 5 — Drilling fluid treatment metrics.

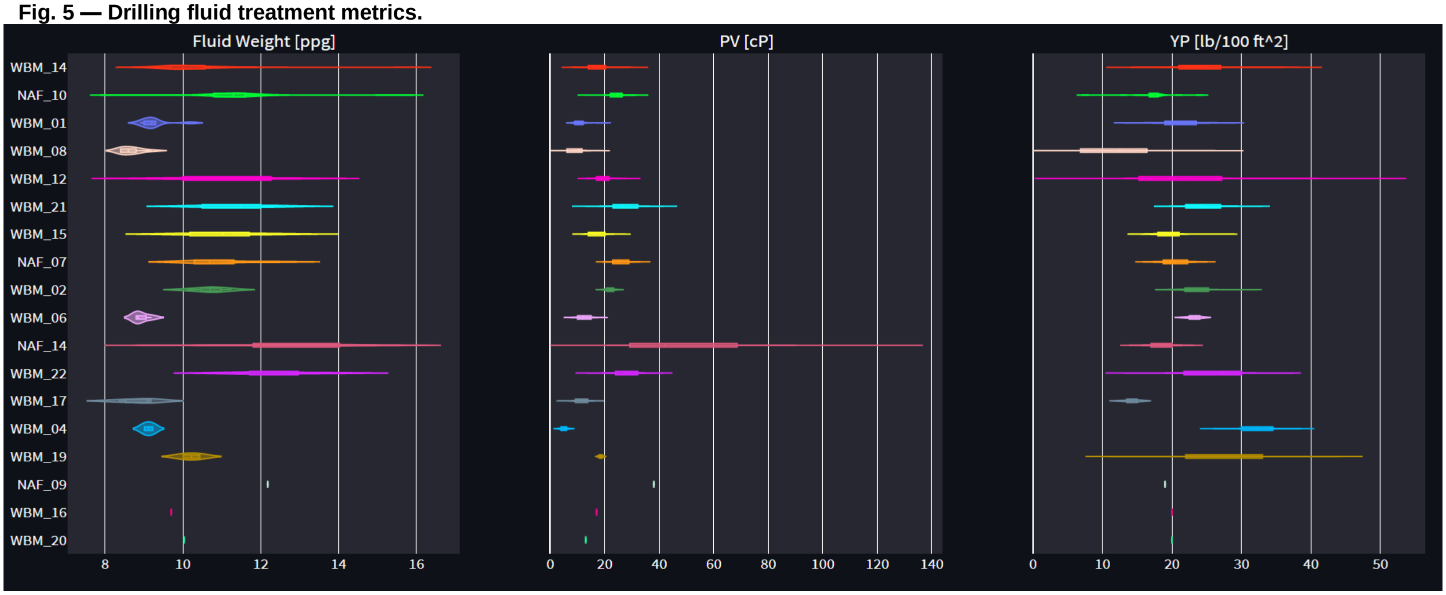


Fig. 6 — Drilling fluid key property distributions.

The last visual DFR module depicts distributions of drilling fluid properties **(Fig. 6).** This module purely aims at describing the historical ranges of key fluid properties for information purposes. While fluid density range can be narrowed down with the filter above, plastic viscosity and yield point violin plot demonstrate the most common ranges of properties across offset intervals in a simplified way, as they are still aggregated based on many daily mud checks and many intervals.

Finally, DFR provides an option to generate, visualize, and download raw tables **(Fig. 7 and Fig. 8)**, where all attribute values plotted above and many more are reported. The first table contains drilling fluid aggregated data, and the second one is the snapshot of the clean interval table with the top 500 offset intervals (with OIS index), also available for download and offline analysis.

Fluid Family	Number of Wells	Number of Intervals	Median Similarity Index	Sum of Interval Length, ft	Sum of Drilling Duration, ...	Median Interval Length, ft	Median Drilling Duration, ...	Median Water Depth, ft	Median Top Depth, ft
WBM_14	148	148	0.911	449779	1915	2978	10	0	4600
NAF_10	17	17	0.907	34410	169	1994	7	0	6039
WBM_01	10	10	0.908	31710	192	3344	18	0	5612
WBM_08	9	9	0.914	24002	161	2864	20	0	4524
WBM_12	7	7	0.905	19147	85	2313	9	0	4209
WBM_21	6	6	0.912	21255	91	3770	15	0	5040
WBM_15	6	6	0.908	12591	55	1922	8	0	5988
NAF_07	5	5	0.904	11673	72	2483	15	156	5230
WBM_02	5	5	0.902	8040	65	1111	10	0	6589
WBM_06	4	4	0.911	12064	89	2948	23	0	5285
NAF_14	4	4	0.922	11020	72	2968	15	52	6140
WBM_22	4	4	0.932	11311	131	3054	36	141	5010
WBM_17	3	3	0.907	8740	52	2874	15	0	3284
WBM_04	2	2	0.913	2901	12	1450	6	0	5674
WBM_19	2	2	0.921	6437	27	3218	14	0	5138
NAF_09	1	1	0.920	2765	13	2765	13	0	4774
WBM_16	1	1	0.900	1337	17	1337	17	0	6800
WBM_20	1	1	0.902	2603	7	2603	7	0	3512

Fig. 7 — Drilling fluid performance metrics table.

WellId	ProjectId	IntervalId	Similarity_Index	IntervalTypeName	IsOffshore	Country_Name	Office_Name	Operator	Well_Name_Clean	Lease_Name_Clean	Well_API_Clean	Rig_Name	Well_Water_Depth
W23300	P185673	4	0.959	Liner	1	TURKEY	Turkey	OP_272	KUZU PINARI 1	KUZU PINARI 1	<NA>	Rowan Norway	282
W25711	P182209	4	0.959	Liner	1	TURKEY	Turkey	OP_272	KUZU PINARI 1	KUZEY ERDEMLI 1	<NA>	Rowan Norway	282
W30266	P213317	5	0.944	Casing	0	IRAQ	Iraq Basra	OP_020	MJ 90	MJ 90	<NA>	HH-030	0
W26866	P167520	4	0.939	Casing	0	EGYPT	Egypt	OP_176	ESAEN 1	ESAEN 1	<NA>	ST-02	0
W06638	P145310	4	0.938	Casing	0	IRAQ	Iraq Basra	OP_083	ZB 305	ZB305	<NA>	SAXON RIG-202	0
W30670	P177753	4	0.938	Casing	0	IRAQ	Iraq Basra	OP_083	ZB 376VERTICAL	ZB 376	<NA>	SAXON RIG-201	0
W18321	P173589	13	0.935	Casing	1	AZERBAIJAN	Azerbaijan	OP_250	SWG 332	GUNESHLI 332	<NA>	GUNESHLI - Guneshli (Pla...	351
W21123	P188050	5	0.935	Casing	1	AZERBAIJAN	Azerbaijan	OP_250	SWG 335	GUNASHLI 335	<NA>	GUNESHLI - Guneshli (Pla...	351
W11677	P156751	4	0.934	Casing	0	KUWAIT	Kuwait	OP_134	MN 239	MN 239	<NA>	BWD-151	0
W10103	P155203	4	0.934	Casing	0	KUWAIT	Kuwait	OP_134	RA 0655	RA 655	<NA>	SINOPEC-174	0
W28658	P209609	4	0.934	Casing	0	KUWAIT	Kuwait	OP_134	SA 0881	SA 881	<NA>	KDC-29	0
W05324	P138548	4	0.932	Casing	0	IRAQ	Iraq Basra	OP_248	ZB 276	ZB 276	<NA>	SAXON RIG-201	0
W22377	P154718	5	0.932	Open Hole	0	TURKEY	Turkey	OP_299	YERKA 6	YERKA 6	<NA>	F200-6	0
W18323	P182109	5	0.932	Casing	1	AZERBAIJAN	Azerbaijan	OP_250	SWG 334	GUNESHLI 334	<NA>	GUNESHLI - Guneshli (Pla...	351
W22464	P172873	5	0.931	Side Track	1	GREECE	Greece	OP_176	EL 1	EL 1	<NA>	GSP Jupiter	125
W28491	P190752	5	0.931	Casing	0	EGYPT	Egypt	OP_005	MEL WEST DEEP 2	MEL WD 2	<NA>	WDI-147	0
W28586	P185525	5	0.931	Casing	0	EGYPT	Egypt	OP_005	BASMA 1X	BASMA	<NA>	EDC-41	0
W28474	P210775	5	0.930	Casing	0	KUWAIT	Kuwait	OP_134	SA 831	SA 831	<NA>	SINOPEC-288	0
W22349	P148829	5	0.930	Casing	0	ITALY	Italy	OP_248	MOIRAGO 1 DIR	MOIRAGO 1 DIR	<NA>	EMSCO C3 Pergemine	0
W28467	P201337	4	0.930	Casing	0	KUWAIT	Kuwait	OP_134	RA 0785	RA 785	<NA>	SINOPEC-285	0
W01057	P128492	7	0.930	Casing	0	IRAQ	Iraq Basra	OP_083	ZB 356	ZB356	<NA>	RMG Rig -102	0
W03129	P132998	4	0.930	Casing	0	KUWAIT	Kuwait	OP_134	RA 594	RA 594	<NA>	KDC-53	0
W02600	P133417	4	0.929	Casing	0	IRAQ	Iraq Basra	OP_083	ZB 359	ZB359	<NA>	RMG-101	0
W29112	P167473	7	0.929	Casing	0	KUWAIT	Kuwait	OP_134	UG 266	UG 266	<NA>	KDC-48	0
W28519	P188413	3	0.928	Casing	0	IRAQ	Iraq Basra	OP_090	WQ1 480	WQ1 480	<NA>	SAXON RIG-203	0
W21271	P191318	3	0.928	Casing	0	ROMANIA	Romania	OP_164	3641 VALCELE	3641 VALCELE VEST	<NA>	Upet 4	0
W05035	P137460	4	0.927	Casing	0	IRAQ	Iraq Basra	OP_083	ZB 251	ZB251	<NA>	SAXON RIG-202	0
W23616	P196241	4	0.927	Casing	1	AZERBAIJAN	Azerbaijan	OP_250	SWG 336	GUNASHLI 336	<NA>	GUNESHLI - Guneshli (Pla...	351
W28469	P193837	4	0.927	Casing	0	KUWAIT	Kuwait	OP_134	RA 0738	RA 738	<NA>	SINOPEC-288	0
W04746	P137459	4	0.927	Casing	0	IRAQ	Iraq Basra	OP_083	ZB 275	ZB275	<NA>	SAXON RIG-201	0

Fig. 8 — Offset interval properties (Data anonymized).

Example – US land

Target interval were chosen in US land, with the plan to drill 2,500 ft interval, previous casing set at 5,500 ft, planned interval bit size set at 12.25 in and maximum BHCT equal to 145 degF. Engine calculated about 194 preselected intervals with OIS index set up as default on 0.90. To widen the range of preliminary chosen offset intervals, OIS index was changed to 0.85, which gave about 1,000 different intervals in US land or close to it.

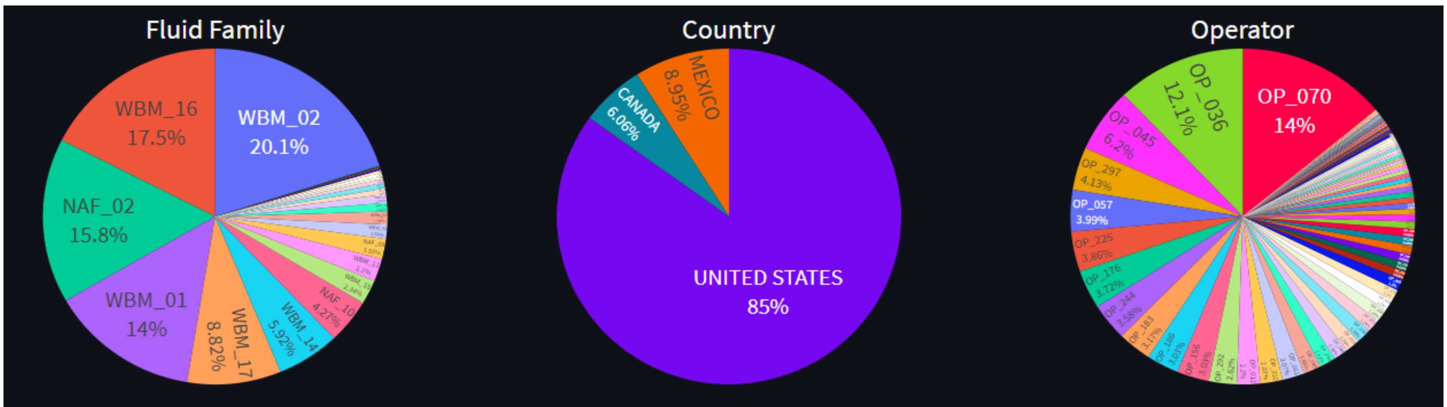
To filter out unnecessary data, mud weight was narrowed to the range of 8.5 to 12.0 lb/galUS, considering planned pore and fracturing pressures.

Currently the prototype does not include information about pore and frac pressures, as well as type of formations. However, chosen mud weight range and location of drilling are an indirect representation of well pressures and formations, which are sufficient for the prototype, until actual data will be implemented (under development at the moment).

Based on the above preparations, 726 final offset intervals are left for further analysis. Statistically, similar intervals were drilled mainly with WBM with some rare intervals drilled with NAF **(Fig. 9).** As can be observed from summary table, WBM_02 shows the highest prevalence over all other types of fluid systems with high similarity to target interval parameters. Drilling ROP using WBM_02 provides excellent performance. At the same time, it results in medium remedial time activities, low quantity of fluid treatments, and low level of fluid complexity. Last, but not least, additives cost for WBM_02 system are relatively small.

Closest NAF system to WBM_02 is NAF_02. Comparing them, it becomes clear that NAF_02 results in much worse Drilling ROP and higher additives cost, being similar to WBM-6 on all other metrics.

Considering target parameters used in this example, it looks like WBM_02 would provide the best solution for this case.



	Fluid Family	Prevalence	Similarity	Drilling ROP	Remedial Time	Fluid Treatments	Fluid Complexity	Additives Cost	Average Rank
8	WBM_02	1	9	5	14	4	4	2	6
17	WBM_16	2	6	7	13	13	17	8	9
1	NAF_02	3	14	10	12	5	2	5	7
7	WBM_01	4	12	4	17	2	4	3	7
18	WBM_17	5	4	6	18	11	12	7	9
15	WBM_14	6	17	11	16	13	13	15	13
5	NAF_10	7	20	2	15	1	2	6	8
16	WBM_15	8	15	16	11	16	16	13	14
13	WBM_11	9	8	19	7	22	22	21	15
4	NAF_09	10	6	17	9	7	6	12	10
12	WBM_10	11	18	14	5	21	19	18	15
11	WBM_09	12	2	1	22	2	7	14	9
14	WBM_13	14	11	13	19	19	14	11	14
21	WBM_21	14	23	3	10	9	3	1	9
22	WBM_22	16	5	19	6	18	9	19	13
2	NAF_05	16	10	23	23	13	9	20	16
19	WBM_19	17	3	9	20	15	18	9	13
10	WBM_06	19	1	7	2	6	14	4	8
6	NAF_11	19	16	12	8	8	9	16	13
20	WBM_20	19	12	17	2	17	11	10	13
9	WBM_03	22	19	14	2	20	23	22	17
3	NAF_07	22	21	22	21	10	20	17	19
0	NAF_01	22	22	21	4	23	21	23	19

Fig. 9—Drilling fluid recommendation summary.

Conclusions

The fully functional Drilling Fluid Recommender web application prototype was developed. Recommendations on the drilling fluid selection for the target interval are provided based on the historical data on similar intervals from the operational database of the world's largest service company.

The historical data is automatically cleaned and processed with algorithms, performing data conversions, transformations, and outlier detection operations. Based on the user's input of the target interval parameters, offset intervals are identified with the newly developed similarity index. Analytic results on the offset intervals are summarized with many drilling performance metrics and drilling fluids characteristics in the form of user-friendly charts and tables. Comparison of different drilling fluid systems in terms of their performance, operational complexity, cost, etc. is streamlined. The recommender is complemented with detailed documentation. and methodology.

Being a prototype, the DFR system can already assist with quick data analysis and enable data-driven business and technical decisions, providing critical information on historical performance of drilling fluids under similar conditions.

The Drilling Fluid Recommender system is one of the first steps of recommender systems journey for well construction. Papers are prepared for publication on other drilling recommender systems. Further research is ongoing on the high-level recommender system, which combines several drilling technologies together and strives to provide holistic recommendations on their optimal combinations.

Acknowledgments

The authors would like to thank Greg Skoff, John Whyte, Mario Bouguetta, Daria Khvostichenko, and Charles Jeong (all from Schlumberger) for their insightful comments and fruitful discussions; and Schlumberger for permission to publish this work.

Nomenclature

Definitions for symbols used in the text:

AI = *Artificial intelligence*
BHA = *Bottomhole assembly*
BHCT = *Bottomhole circulation temperature*
DFR = *Drilling fluid recommender*
LCM = *Lost circulation material*
MCRS = *Multi-criteria recommender system*
MD = *Measured depth*
NAF = *Nonaqueous fluid*
NPT = *Nonproductive time*
OBM = *Oil-based mud*
OIS = *Offset interval similarity*
ROP = *Rate of penetration*
SBM = *Synthetic-based mud*
SQL = *Structured query language*
TVD = *True vertical depth*
WBM = *Water-based mud*

References

1. Adomavicius, G., Kwon, Y. 2015. "Multi-Criteria Recommender Systems." In: Ricci F., Rokach L., Shapira B. (eds) "Recommender Systems Handbook". Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7637-6_25.
2. Breese, J. S., Heckerman, D., Kadie, C. 2013. "Empirical Analysis of Predictive Algorithms for Collaborative Filtering", Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI1998), https://arxiv.org/abs/1301.7363v1
3. Friedheim, J. E., Hans, G. J., Park, A. and Ray, C. R. 1991. "An Environmentally Superior Replacement for Mineral-Oil Drilling Fluids", SPE-23062. SPE Offshore Europe, Aberdeen, United Kingdom, September 1991, https://doi.org/10.2118/23062-MS
4. Khvostichenko, D., Champeau, M., Powell, J., Vesselinov, V., Skoff, G., Bouguetta, M., Arevalo, Y. and Makarychev-Mikhailov, S. 2022. "How Drilling Fluids Affect Drilling Performance: Big Data Analysis," SPE-208688, IADC/SPE International Drilling Conference and Exhibition, Galveston, Texas, USA, 8–10 March 2022.
5. Linden, G. D., Jacobi, J. A., Benson, E. A., 2001. "Collaborative recommendations using item-to-item similarity mappings", US Patent 6,266,649, July 24, 2001.
6. McKinney, W., 2010. "Data Structures for Statistical Computing in Python." Proceedings of the 9th Python in Science Conference, 445, 51.
7. Patel, A., Stamatakis, E, Young, S. and Friedheim, J. 2007. "Advances in Inhibitive Water-Based Drilling Fluids—Can They Replace Oil-Based Muds?" SPE-106476. The International Symposium on Oilfield Chemistry, February 28–March 2, 2007. https://doi.org/10.2118/106476-MS
8. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É. 2011. "Scikit-learn: Machine Learning in Python." Journal of Machine Learning Research, 12 (85), 2825, 2011. https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html
9. Plotly Technologies Inc. Collaborative data science. Montréal, QC, 2015. https://plot.ly
10. Streamlit Inc. 2022. https://Streamlit.io/
11. Whyte, J., Arevalo, Y., Makarychev-Mikhailov, S. 2022. "Drilling Fluids Profiling—Is Your Mud Right for the Job?" SPE-208693. IADC/SPE International Drilling Conference and Exhibition, Galveston, Texas, USA, 8–10 March 2022.
12. Young, S. and Friedheim, J. 2013. "Environmentally Friendly Drilling Fluids for Unconventional Shale." OMC-2013-102. The Offshore Mediterranean Conference and Exhibition, Ravenna, Italy, March 2013.