

# Optimized feature selection assists lithofacies machine learning with sparse well log data combined with calculated attributes in a gradational fluvial sequence



David A. Wood

DWA Energy Limited, Lincoln, United Kingdom

## ARTICLE INFO

### Keywords:

Derivative/volatility log attributes  
Sparse well-log datasets  
Multi-k-fold analysis  
Optimizer comparisons  
Lithofacies imbalance.

## ABSTRACT

Machine learning (ML) to predict lithofacies from sparse suites of well-log data is difficult in laterally and vertically heterogeneous reservoir formations in oil and gas fields. Meandering, braided fluvial depositional environments tend to form clastic sequences with laterally discontinuous layers due to the continuous shifting of relatively narrow sandstone channels. Three cored wells drilled through such a reservoir in a large oil field, with just four recorded well logs available, are used to classify four lithofacies using ML models. To augment the well-log data, six derivative and volatility attributes were calculated from the recorded gamma ray and density logs, providing sixteen log features for the ML models to select from. A novel, multiple-optimizer feature selection technique was developed to identify high-performing feature combinations with which seven ML models were used to predict lithofacies assisted by multi-k-fold cross validation. Feature combinations with just seven to nine selected log features achieved overall ML lithofacies accuracy of 0.87 for two wells used for training and validation. When the trained ML models were applied to a third well for testing, lithofacies ML prediction accuracy declined to 0.65 for the best performing extreme gradient boosting model with seven features. However, an accuracy of ~0.76 was achieved by that model in predicting the presence of the pay bearing sandstone and siltstone lithofacies in the test well. A model using only the four recorded well logs was only able to predict the pay-bearing lithofacies with ~0.6 accuracy. Annotated confusion matrices and feature importance analysis provide additional insight to ML model performance and identify the log attributes that are most influential in enhancing lithofacies prediction.

## 1. Introduction

The classification of lithofacies in wellbores is a fundamental component of reservoir characterization (Dubois et al., 2007). Extensive use is made of recorded well-log dataset to conduct that classification in both clastic (Rider, 1990) and carbonate (Stowe and Hock, 1988) reservoirs. This is very effective, particularly where the lithofacies are easily distinguished in core and well-log terms, laterally extensive across the reservoir, reasonably homogeneous and a sufficient diversity of well-log features is recorded in multiple wellbores. If some of these requirements are missing it is typically necessary to complement the available core/well-log data with additional geological information (Reverdy et al., 1983).

A common issue to contend with is lack of recovered core material from a reservoir. Coring a reservoir adds substantially to drilling costs and in some formations core recovery is poor. Consequently in the

majority of fields cores are only collected from a small fraction of the development wells drilled. A similar data limitation issue also often exists with the well-log data collected. In exploration and appraisal wellbores it is standard practice to collect a broad suite of well logs that is sufficient to characterize a reservoir on a field-wide scale. However, when it comes to development drilling, to save costs and time, only a sparse suite of well logs is normally collected. For reservoir-wide lithofacies classification purposes these dataset limitations introduce substantial uncertainty, particularly in heterogeneous reservoir sequences (Ma, 2019). Recent studies have demonstrated that sparse well-log datasets can be augmented by adding calculated derivative and volatility attributes of selected well logs to improve lithofacies classification reliability (Wood, 2021, 2022). These studies represent the first time such well-log attributes were applied to assist lithofacies classification.

The selection of which features to use for lithofacies classification, geological (Halotel et al., 2020) and/or well-log (Wood, 2021) has a

E-mail address: dw@dwasolutions.com.

substantial influence on its effectiveness and accuracy. There are many approaches that can be adopted to conduct feature selection in lithofacies analysis. At the simplest level a trial-and-error approach can be used, if there are relatively few features to select from. However, as the number of available features increases the possible combinations for those features runs into many thousands, making trial-and-error both inefficient and unreliable. Indeed, feature selection is an NP-hard combinatorial challenge (Cortes and Vapnik, 1995); as the number of possible feature combinations grows exponentially as the number of available features increases. A more effective approach is to apply optimizers to conduct more comprehensive feature selection (Abiodun et al., 2021).

In recent decades, machine learning (ML) methods have substantially improved upon statistical and regression methods for lithofacies predictions based on well logs (Goncalves et al., 1995). However, combining clustering and principal-component analysis to substantially reduce the number of features can improve the efficiency of some ML models (Ma, 2011). Distinct types of ML algorithm are now used for well-log-based lithofacies prediction. This commenced three decades ago with neural networks (Rogers et al., 1992; Agrawal et al., 2022) and has diversified to include support vector classification (Sarkar and Majundar, 2020), K-nearest neighbor (Merembayev et al., 2021), optimized data matching (Wood, 2019) and various tree-ensemble methods (Farzi and Bolandi, 2016; Al-Mudhafar et al. (2022)).

This study applies a novel multi-optimizer feature selection method to a three-well dataset sampling a vertically and laterally heterogeneous clastic reservoir deposited in a meandering braided stream environment. The well data consists of a core-derived lithofacies classification and just four recorded well logs augmented by twelve calculated derivative and volatility attributes. The multi-optimizer feature selection technique proposed and developed involving well-log attributes and coupled with multi-k-fold cross validation applied to seven ML models achieves improved lithofacies classification. It also makes the identification of the more influential well-log attributes more efficient.

## 2. Materials & methods

### 2.1. Well-log derivative and volatility attributes considered

Calculated attributes derived from recorded gamma-ray (GR) logs have recently been proposed and exploited to enhance lithofacies classification with ML models (Wood, 2022). This study expands that approach to include attributes of the density (PB) log. It evaluates the impact of GR and PB log attributes on lithofacies classification of a cored three-well clastic sequence displaying lateral and vertical gradations across an oil reservoir. Only four recorded well logs exist in the three wells studied, and these are derived from a spectral gamma-ray log. The recorded logs are total GR, PB (gamma-gamma system) and the spectral gamma rays for thorium (TH) and potassium (K). These recorded-log values are referred to as GRO, PBO, Th0 and KO.

Six attributes (equations (1)–(6)) are calculated from each of the GRO and PBO values for Wells V, W and X. The term “AL” refers generically to the recorded well log for which an attribute is calculated.

#### 2.1.1. First-derivative attribute (AL1)

$$AL1_d = (AL0_d - AL0_{d-1}) / Abs(d - (d - 1)) \quad (1)$$

where  $AL0_d = AL0$  value at depth  $d$ , and  $AL0_{d-1} = AL0$  at depth  $d - 1$ . The depth interval between each sampled well-log value is approximately 15 cm. Attributes GR1 and PB1 are calculated with equation (1).

#### 2.1.2. Moving average of first-derivative attribute (AL2)

$$AL2_{da} = \left( \sum_{i=1}^{i=a} AL1_{d-i} \right) / \alpha \quad (2)$$

where  $\alpha =$  a specified number of sampled intervals above depth  $d$ , and  $\alpha$  is assigned a value of 10 for this dataset, as determined by trial and error. Attributes GR2 and PB2 are calculated with equation (2).

#### 2.1.3. Second-derivative attribute (AL3)

$$AL3_{dp} = (AL1_d - AL1_{d-\beta}) / Abs(d - (d - \beta)) \quad (3)$$

where  $\beta =$  a specified number of sampled intervals above depth  $d$ , and  $\beta$  is assigned a value of 10 for this dataset, as determined by trial and error. Attributes GR3 and PB3 are calculated with equation (3).

#### 2.1.4. Natural logarithm of variation between adjacent recorded well-log values (AL4)

$$AL4_{i(d)} = \ln(AL0_d / AL0_{d-1}) \quad (4)$$

where  $i(d) =$  the interval between recorded log values at depths  $d-1$  and  $d$ . Attributes GR4 and PB4 are calculated with equation (4).

#### 2.1.5. Standard deviation of AL4 for specified overlying interval (AL5 ; "volatility")

$$AL5_{i(\gamma)} = \sqrt{\frac{\sum_{j=0}^{\gamma} (AL4_{i(d-j)} - AL4_{i(\gamma)} \text{mean})^2}{\gamma - 1}} \quad (5)$$

where  $i(\gamma) =$  the interval between recorded log values at depths  $d-\gamma$  and  $d$ , and  $\gamma$  is assigned a value of 10 for this dataset, as determined by trial and error. Attributes GR5 and PB5 are calculated with equation (5).

#### 2.1.6. Moving average of volatility attribute (AL6)

$$AL6_{i(\delta)} = \left( \sum_{i=0}^{i=\delta} AL5_{i(d-i)} \right) / \delta \quad (6)$$

where  $\delta =$  a specified number of sampled intervals above depth  $d$ , and  $\delta$  is assigned a value of 10 for this dataset, as determined by trial and error. Attributes GR6 and PB6 are calculated with equation (6).

Taking into account the six calculated GR attributes (GR1 to GR6), the six calculated PB attributes (PB1 to PB6) together with the four recorded well logs (GRO, PBO, TH0 and KO), there are sixteen influential variables for consideration by the ML lithofacies classification algorithms.

There are no specific geological factors used to define or control any of the log attributes selected. However, the fluctuating values of each attribute do tend to respond differently to different formation lithologies, textures, and degrees of lamination. The attributes selected are generic and could be applied to any log curve that shows fluctuations in its absolute values versus depth that relate to the physical rock properties. Hence, attributes calculated for gamma-ray, bulk density, and acoustic well logs are more likely to give consistent responses for formations across several wells than resistivity and neutron logs. The reason for this is that the resistivity and neutron logs are strongly affected by the fluid type present in the porous formations present. So, for lithofacies analysis, attributes applied to gamma-ray, bulk density, and acoustic logs commonly recorded in many wells are the most obvious logs for which attributes can provide useful complementary information to the raw log values.

The log attributes selected are designed to provide complementary information to the raw log values relating to their instantaneous rates of change (first and second derivative), fluctuations in absolute values (volatility), and the moving averages of those values over specified depth intervals. The depth intervals specified for the volatility and moving average attributes can be usefully varied to suit the conditions of specific rock sequences being studied. For instance, in thinly bedded sequences or those with rapidly changing facies and lithology, it is

appropriate to calculate the volatility and moving averages over quite narrow depth intervals. On the other hand, in sequences with more massive formations and with facies and lithology changing at relatively low frequencies it is appropriate to calculate the volatility and moving averages over wider depth intervals. The adjustment of depth intervals considered in this way provides flexibility and enables some of the attributes to be tuned to suit geological conditions.

## 2.2. Three well reservoir dataset

Three cored wellbores, V, W and X sample a fluvial clastic reservoir up to 95m thick with each well containing multiple oil-saturated sandstones and siltstones. Core gamma-ray and bulk-density data were used to ensure that core and log depths were appropriately calibrated. The producing field extends over a closed oil-pay area of approximately 100 m<sup>2</sup>. The wells are located between 5 km and 10 km apart. Core analysis has revealed that the clastic reservoir sequence consists of stacked meandering braided river-channels deposited in a regionally extensive flood plain (Fig. 1). Each channel, dominated by sandstone becoming more silty towards the channel margins, is extensive in the sinuous flow direction but of limited lateral extent (up to about 1 km). The braided channels do not coalesce to form laterally continuous sandstones but do regionally cut into one another forming an extensive reservoir system with pore fluids in pressure communication between many of the channels. The lateral discontinuity of the channels poses difficulties in correlating specific channels from one wellbore to another and in lithofacies classification. Channel depositional conditions vary substantially from one channel to another in a specific wellbore. Depositional conditions depend upon the flow regime and energy (fast-flowing or slow-flowing) prevailing in the channel at each location governed by its position in the meandering system.

The channels are encased in regionally extensive shales, deposited originally as clay-rich mudstones deposited extensively over time across the flood plain. Some of those mudstones were deposited in lacustrine conditions, associated with extensive isolated ox-bow lakes that existed from time to time and shifted their positions periodically across the flood plain. Isolated crevasse splays also occurred from time to time as overbank spills from the active river channels in times of flood (Fig. 1). This depositional setting led to the accumulation of four distinct lithofacies

currently preserved at depths of about 3000m in the reservoir section. These are: clay-rich shale (Sh) forming laterally extensive layers; siltstone (Slt) mainly towards the margins and tops of the preserved channels; sandstone (Sd) dominating the bulk of the preserved channels; and, sandy shale (SSh) primarily associated with the crevasse splays that also contain thin siltstone and sandstone bands in places.

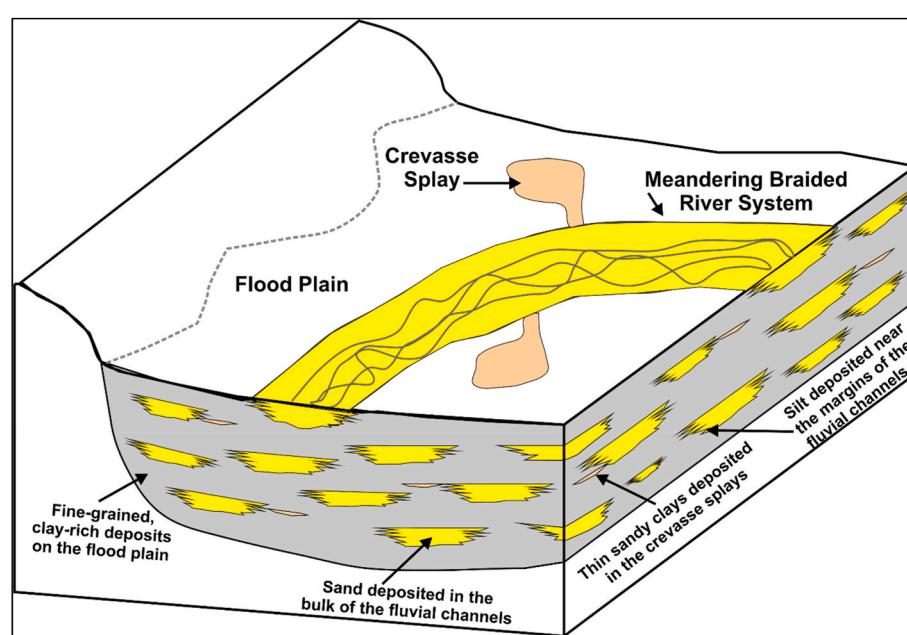
**Table 1** details the fractional distributions of the Sh, Slt, Sd and SSh lithofacies in the three cored wells, which are displayed in Fig. 2. It reveals that more than 50% of the reservoir is formed of Sd and a little less than 30% formed of Sh. On the other hand, Slt and SSh are present in minor amounts distributed through the reservoir. The Slt component varies substantially from well to well (6%–27%) averaging about 13%. SSh is the least abundant lithofacies, varying from about 3% to 6%, and averaging about 4%. These lithofacies distributions constitute a class imbalance problem for the ML classification models making it more difficult for them to predict the least abundant lithofacies.

Fig. 3 displays the vertical distributions of the cored lithofacies in wells V, W and X and Table 2 statistically summarises the recorded well logs and each of the calculated GR and PB attributes in the cored sections of those three wells.

Wells V, W and X are the only boreholes drilled through the reservoir in this oil field in which cores were recovered. The core analysis and interpretation has made it possible to establish a detailed depositional model for those well locations. However, there are multiple other non-cored production wells drilled in the oil field for which spectral gamma ray logs were recorded. In those non-cored wells substantial

**Table 1**  
Fractional distribution of the four lithofacies identified from core data for the 3-well meandering braided fluvial clastic dataset.

Cored Lithofacies	Well V	Well W	Well X	3 Wells Combined
Number of data records	623	362	251	1236
Clay-rich shale (Sh)	0.3178	0.2127	0.2789	0.2791
Siltstone (Slt)	0.0562	0.2652	0.1155	0.1294
Sandstone (Sd)	0.5923	0.4751	0.5418	0.5477
Sandy shale (SSh)	0.0337	0.0470	0.0637	0.0437
Total of fractions:	1.0000	1.0000	1.0000	1.0000



**Fig. 1.** Schematic representation of the meandering braided fluvial depositional environment associated with the clastic reservoir sequence characterized by gradations between four lithofacies in vertical and lateral directions.

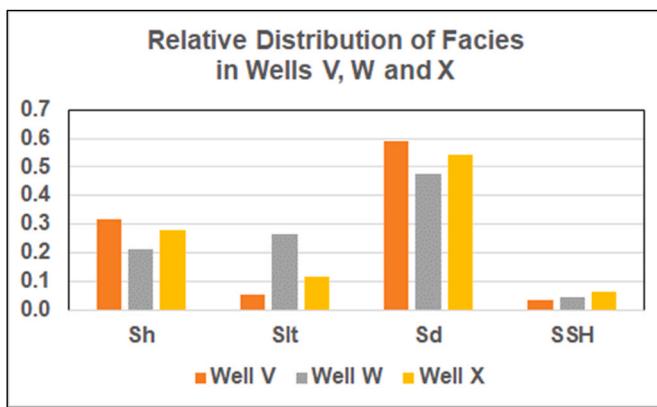


Fig. 2. Facies distribution compared in wells V, W and X.

uncertainty remains regarding lithofacies distribution and relationships between the productive Sd and Slt sequences. Establishing a reliable calibration between the cored lithofacies interpretation and the recorded well logs is highly desirable as it offers the opportunity to reliably extend the lithofacies interpretation to the non-cored wells. However, to

do so it is important to ensure that the well logs used for this purpose are appropriately corrected for environmental conditions in the boreholes to remove obvious anomalies caused by drilling conditions. The well log responses also need to be carefully scrutinized in conjunction with the facies identified in the available cores to ensure that there are no systematic shifts in the log values between the logging suites recorded in each well. The dataset used satisfied these requirements.

It is apparent from Fig. 3 that Sd forms the thickest reservoir zones in each well, up to 10–15 m in some cases. Sh constitutes some reservoir zones up to 6 m thickness but with some thinner zones of about 1 m thick. However, Slt and SSH exist as several thin layers of less than 2 m thick, except in Well W which includes an Slt zone almost 10 m thick.

Comparison of the mean and P50 values in Table 3 reveal that the PBO value distribution is approximately symmetrical (mean = P50). On the other hand, the other recorded logs the mean is greater than the P50 value indicating they form positively skewed distributions. Similar comparisons for the calculated GR and PB log-attribute distributions indicate that most of them are approximately symmetrical, although GR1 and GR2 are slightly negatively skewed (P50 > mean) and GR3 slightly positively skewed. The mean values for the well log and attribute variables in the three wells are similar in most cases. However, for GRO, TH0 and K0 there is a substantial difference. For GRO and K0 the

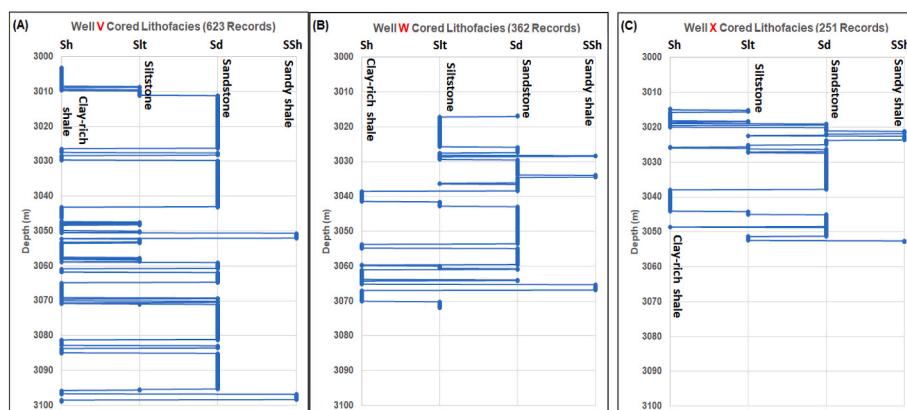


Fig. 3. Distribution of cored lithofacies versus depth in the three-well clastic sequence dataset: (A) Well V; (B) Well W; and, (C) Well X. Each well consists of an interbedded sequence of more massive sandstones (Sd), clay-rich shales (Sh) interbedded with thinner, more sporadic layers of siltstone (Slt) and sandy shale (SSH).

Table 2

Value distribution statistics for variables associated with the meandering braided fluvial 3-well clastic dataset.

	Statistical summary of well-log features including GR and PB attributes for 1236-record cored dataset from wells V,W and X												
	Min	Max	Range	Mean	P10	P50	P90	SD	SE	CoV	Mean Well V	Mean Well W	Mean Well X
Depth	3003.2	3098.9	95.7	3045.6	3017.8	3042.9	3080.1	23.074	0.65631	0.008	3051.1	3044.4	3033.8
GRO	13.167	172.704	159.536	80.749	39.406	74.743	127.334	33.453	0.95155	0.414	74.141	83.994	92.468
GR1	-20.314	25.316	45.630	-0.031	-5.223	-0.173	5.163	4.768	0.13562	-151.640	-0.030	-0.047	-0.012
GR2	-5.484	8.907	14.391	-0.026	-2.326	-0.104	2.131	2.001	0.05691	-75.931	-0.028	-0.020	-0.031
GR3	-2.126	2.048	4.174	-0.002	-0.611	0.003	0.554	0.512	0.01456	-264.046	-0.002	-0.003	0.001
Gr4	-0.446	0.442	0.888	0.000	-0.107	-0.003	0.113	0.101	0.00287	-217.454	0.000	-0.001	0.000
Gr5	0.014	0.238	0.223	0.083	0.030	0.075	0.148	0.046	0.00131	0.553	0.082	0.087	0.079
GR6	0.018	0.202	0.184	0.083	0.037	0.076	0.141	0.040	0.00115	0.486	0.082	0.087	0.079
PBO	2.229	2.842	0.613	2.530	2.396	2.532	2.662	0.102	0.00289	0.040	2.512	2.581	2.503
PB1	-0.059	0.070	0.129	0.000	-0.019	0.000	0.020	0.016	0.00046	114.565	0.000	0.000	0.000
PB2	-0.018	0.022	0.040	0.000	-0.007	0.000	0.008	0.006	0.00017	-220.686	0.000	0.000	0.000
PB3	-0.006	0.007	0.013	0.000	-0.002	0.000	0.002	0.002	0.00005	121.785	0.000	0.000	0.000
PB4	-0.035	0.044	0.080	0.000	-0.011	0.000	0.012	0.010	0.00028	107.232	0.000	0.000	0.000
PB5	0.002	0.029	0.028	0.009	0.004	0.008	0.015	0.004	0.00013	0.512	0.009	0.008	0.008
PB6	0.002	0.024	0.021	0.009	0.004	0.008	0.014	0.004	0.00011	0.458	0.009	0.008	0.008
TH0	0.503	16.960	16.457	6.584	2.842	5.894	11.856	3.507	0.09976	0.533	6.441	7.014	6.320
K0	0.036	5.843	5.807	2.585	0.657	2.462	4.739	1.325	0.03770	0.513	2.049	2.952	3.389
Facies #	0	3	3	1.356	0	2	2				1.342	1.356	1.390

Notes to Table 2: P10 = 10th percentile; P50 = 50th percentile; P90 = 90th percentile; SD = standard deviation; SE = standard error of the mean; CoV = coefficient of variation; and Facies# referring to lithofacies expressed as numbers (Sh = 0, Slt = 1, Sd = 2 and SSH = 3) to each data record.

**Table 3**

Optimizer algorithms applied in feature selection routine.

Thirteen Optimizers Applied to Entire three-well Dataset with 16 Features						
1263 data records		Execution Times in seconds		Example Applications for Each Optimizer		
Optimizer Algorithms	Acronym	N = 10	N = 100	Source		
Bat Flight	BF	14	113	Wood (2016a)		
Cuckoo Search	CS	26	226	Wood (2016b)		
Differential Evolution	DE	14	90	Zhang et al. (2020)		
Firefly	FF	15	787	Wood (2018)		
Flower Pollination	FP	19	115	Alyasseri et al. (2018)		
Genetic Algorithm	GA	22	147	Mansouri et al. (2015)		
Grey Wolf	GW	14	80	Mirjalili et al. (2014)		
Harris Hawk	HH	22	224	Hussien et al. (2022)		
JAYA	JY	15	62	Rao (2016)		
Particle Swarm	PS	25	133	Atashnezhad et al. (2014)		
Sine Cosine	SC	8	62	Abualigah & Diabat (2021)		
SALP	SP	15	126	Faris et al. (2020)		
Whale Optimizer	WH	15	114	Ning & Cao (2021)		

Note: N refers to the population size used.

mean values are ordered such that Well X > Well W > Well V. For TH0 the mean values are ordered such that Well W > Well X > Well V. This is due to the variable relative distributions of the lithofacies in the three wells (Table 1).

The substantial variations in distribution ranges displayed by the recorded well logs and calculated attributes makes it essential to normalize their values for ML analysis. For this study all variables were normalized to a scale of -1 to +1 using equation (7).

$$Nx_i^m = 2 * \left( \frac{x_i^m - xmin^m}{xmax^m - xmin^m} \right) - 1 \quad (7)$$

Where  $Nx_i^m$  is a log's value adjusted to a scale of -1 to +1 based on the minimum and maximum values of  $m$ th well log's value distribution.  $x_i^m$  refers to the recorded or calculated value of the  $i$ th data point of the  $m$ th variable distribution,  $xmin^m$  is the minimum value of  $m$ th variable distribution, and  $xmax^m$  is the maximum value of  $m$ th variable distribution.

Fig. 4 displays the Pearson's ( $R$ ) and Spearman's ( $p$ ) correlation coefficients between the variables and the lithofacies class. To calculate these correlations, each lithofacies is assigned a numerical value (facies#: Sh = 0, Slt = 1, Sd = 2 and SSh = 3). The  $R$  calculation makes parametric assumptions, whereas the  $p$  calculation is non-parametric being based on ranked values. For symmetrical variable distributions, approximating normal distributions,  $R$  and  $p$  values are normally quite similar. On the other hand, for highly skewed distributions, indicative of non-parametric conditions  $R$  and  $p$  values are normally quite distinct.

For most of the sixteen well-log variables considered  $R$  and  $p$  values with facies# are relatively similar, indicative of approximately parametric conditions. However, for variables GR1, GR2, PB2, PB5 and PB6,  $R$  and  $p$  values with facies# do show greater differences, indicating that those calculated attribute variables do involve non-parametric relationships. Variables GR0 and TH0 display high negative correlations ( $<-0.6$ ) with facies#, whereas PBO and K0 display moderate negative correlations ( $<-0.4$ ) with facies# (Fig. 4). GR0 has high positive  $R$  correlations with TH0 and K0 ( $>+0.78$ ) and a moderate positive  $R$  correlations with PBO (+0.42). From a machine-learning perspective, the negative correlations of all the recorded well logs with facies# and, the relative high correlations among the recorded well logs, is not a benefit. ML models are better able to discriminate facies in cases where there are positive correlations with some log variable and negative correlations with others.

The calculated well log attributes display poorer  $R$  and  $p$  correlations (between -0.33 and +0.17) with facies# than the recorded well logs (Fig. 4). Most of those correlations are negative, but those for GR5, GR6, PB5 and PB6 are positive. Attributes GR2 and PB2 possess the most

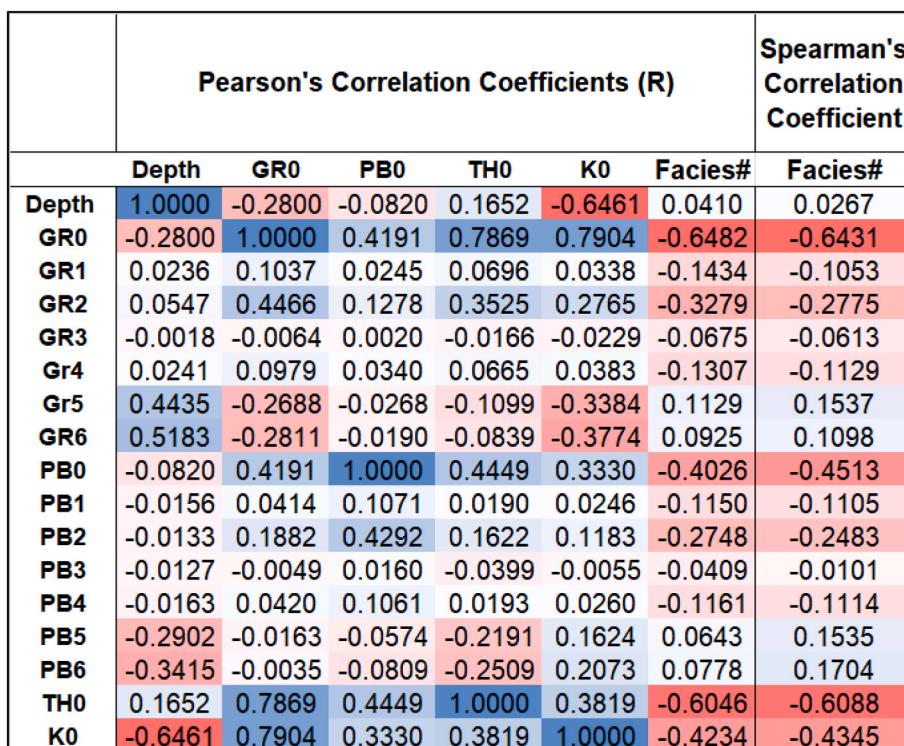


Fig. 4. Heat map displaying selected correlation coefficients of the recorded well logs and well-log attributes with observed facies expressed as numbers (Facies#: Sh = 0; Slt = 1, Sd = 2, SSh = 3) for the 1236-record, cored dataset.

negative  $R$  and  $p$  correlations with facies# of the attributes considered. Based on these correlations, it should be expected that attributes GR2, GR5, GR6, PB2, PB5 and PB6 should be most useful, to the ML algorithms in their attempts to classify lithofacies from the available variables, and usefully complement the recorded well log variables.

Other correlation relationship that are likely to influence facies classification are the high negative  $p$  correlation of K0 with depth ( $<-0.64$ ) and the moderately negative  $p$  correlation of GRO with depth ( $<-0.28$ ). The negative K0 correlation with depth is a consequence of the two lowermost thick sandstones in Well V (Fig. 3) displaying substantially lower K0 values than the sandstones present in the middle and upper parts of the reservoir in all three wells. From core analysis this is a consequence of the lowermost sandstones being poorer in feldspathic minerals than the other sandstones indicating a change in provenance of the sediment during the deposition of the complete reservoir section. The negative GRO correlation with depth reflects the higher proportion of high GR value shales in the upper part of the reservoir section, particularly in the uppermost shale zones. These slight variations in sandstone and shale compositions with depth through the drilled reservoir sections make it more difficult for ML models to accurately characterize the Sh and Sd lithofacies using the available recorded well logs.

### 2.3. Optimized feature selection

With as many as sixteen potentially influential well-log variables to choose from it is appropriate to conduct feature selection to establish whether some of the least influential of those variables can be disregarded, thereby helping to make the ML models more efficient and effective. There are multiple possible variable combinations that could be considered. Two combination that are used as benchmarks are Case 0 using only the four recorded well logs (GRO, PB0, TH0, K0) and Case 1 using all sixteen available variables.

Optimization is an effective method to assess the potential performance of a large number of possible feature combinations and to identify the best performing feature combinations. There are many optimizers available to conduct that analysis, all using slightly different algorithms involving either gradient descent or evolutionary methods. Becoming potentially trapped at local minima is a problem for most optimizers and justifies the value of running several, and multiple runs of each, to establish a set of potentially “optimum” feature selections.

Thirteen evolutionary optimizer algorithms, previously applied to a range of engineering and geoscience datasets were run for this study. These are listed in Table 3, together with citations for studies in which they have been applied, which include detailed descriptions of the methodologies and functions involved for each algorithm. The optimizers are coded for this study in Python, using their standard functions, and are applied to the entire three-well dataset (1236 data records) with the K-nearest neighbor (KNN) ML algorithm. KNN is used because it executes rapidly and requires relatively few control parameters to work effectively. Each optimizer is run multiple times with different population sizes (N values varying from 10 to 100) for 100 iterations, applying the same KNN configuration to the three-well dataset. As the population size of the optimizer (the number of solutions evaluated in each iteration) increases from N = 10 the execution time also increases. Execution times in seconds for the three-well data set modelled with KNN and each optimizer with N = 10 and N = 100 are listed in Table 3. The optimizers execute rapidly with a small population (8–26 s for N = 10) but execution times are higher for high populations (62–787 s for N = 100).

The output of the optimizer models is the feature combination that generates the highest overall accuracy ( $\sum$ Accuracy: the fraction of correct lithofacies classifications). Together with the  $\sum$ Accuracy value, the number of features the optimum solution includes and a fitness score (FS). Rather than simply try to optimize  $\sum$ Accuracy the optimizers are configured to minimize FS calculated by equation (8).

$$FS = \sigma * \varepsilon + \mu \left( \frac{Z}{\max Z} \right) \quad (8)$$

where  $\sigma$  = a constant slightly less than 1,  $\varepsilon$  =  $(1 - \sum \text{Accuracy})$  representing the overall error fraction,  $\mu = 1 - \sigma$ ,  $Z$  = the number of features involved in a solution, and  $\max Z$  = the maximum number of features available. The purpose of the FS, as configured, is to penalize solutions according to the number of features they involve. The small penalty applied increases as the value of  $Z$  increases. Hence the first part of the FS calculation,  $\sigma * \varepsilon$ , the slightly adjusted absolute error dominates the FS value, but the second part of the calculation,  $\mu \left( \frac{Z}{\max Z} \right)$ , adds a penalty to the FS that increases as the number of features in the solution increases. This is an effective way of encouraging the optimizers to seek good solutions with fewer features.

### 2.4. Machine-learning algorithms applied to predict lithofacies

Seven supervised ML algorithms are employed to separately classify the lithofacies of the three-well dataset based on the sixteen defined variables (four recorded well logs and twelve calculated well-log attributes). These algorithms were because they are based on distinct mathematical principles, can be evaluated relatively rapidly, with standard Python code available in public function libraries (SciKit Learn, 2022a).

These ML models are: adaptive boosting (ADA), a tree-ensemble method introduced by Freund & Schapire (1997); decision tree (DT) introduced by Quinlan (1986); K-nearest neighbor (KNN) introduced by (Fix and Hodges, 1951); logistic regression (LR) introduced by Cox (1958); random forest (RF) introduced by Ho, (1998); support vector regression (SVC) introduced by Cortes & Vapnik (1995); and extreme gradient boosting (XGB) introduced by Chen and Guestrin (2016). These models have been widely applied to lithofacies classification studies and their mathematical formulations are well described in the literature. Several recent studies have evaluated and fully described these algorithms. Al-Mudhafar et al. (2022) applied ADA and XGB to classify a heterogeneous carbonate reservoir using well log data. Sarkar and Majundar (2020) compared the performance of DT, RF and SVC in lithofacies classification based on wireline data. Merembayev et al. (2021) compared the performance of DT, KNN, RF and XGB in predicting lithofacies in oil and gas fields located in Norway and Kazakhstan. Masapanta (2021) compared the performances of multiple ML models, including LR, KNN, SVC, DT and RF, to classify lithofacies in North Sea field reservoirs.

Each ML model has hyperparameters that determine its performance with respect to specific datasets. These parameters (Table 4) have been

**Table 4**

Set up and hyperparameters values of ML algorithms used to classify lithofacies from well-log variables.

ML Models Applied	Hyperparameter Values Applied
Adaptive Boosting (ADA)	Number of estimator = 750; learning rate = 0.1; splitting criterion = Gini; base estimator is DT with depth = 15; splitter = best
Decision Tree (DT)	Maximum depth = no limit; splitter = best; splitting criterion = Gini
K Nearest Neighbor (KNN)	Number of nearest neighbours assessed K = 3; distance metric = Minkowski with p = 1 (Manhattan); neighbor selection algorithm = auto
Logistical Regression (LR)	Penalty = elasticnet (applying both L1 and L2 penalty terms); L1 ratio = 0.5; tolerance = 0.0001
Random Forest (RF)	Number of estimators = 750; maximum depth = 20; splitting criterion = Gini
Support Vector Classifier (SVC)	Kernel = radial basis function; C = 750; gamma = 0.2; tolerance = 0.001
Extreme Gradient Boosting (XGB)	Number of estimators = 500; eta = 0.3; maximum depth = 15; splitting criterion = Gini subsample = 0.4; columns sampled/tree = 0.9

selected by combining grid-search analysis (SciKit Learn, 2022b), Bayesian-optimizer prioritization (SciKit Learn, 2022c) and trial-and-error test cases applied to the three-well dataset.

Once the hyperparameters are established, it is necessary to establish the appropriate splits of data records to use for the training and validation dataset. A multi-k-fold cross validation technique is conducted to identify the appropriate splits to use. Four different k values are evaluated: 3-fold, 4-fold, 5-fold and 10-fold. Each fold is computed for multiple runs with the combined data records from Wells V and X (the training and validation subset) to provide mean and standard deviations for mean absolute error (MAE) and root mean squared error (RMSE) of the facies#. A comparison of the mean and standard deviation MAE and RMSE values of the four different k-folds considered indicates the appropriate splits to use. Typically, the split with the lowest mean plus standard deviation is the best split to use. Also, k-folds that generate high mean and/or standard deviation values typically identify the least efficient splits to apply. The K-fold-cross-validation function available in Python for use with the ML packages of SciKit Learn (2022d) is customized to execute the multi-k-fold technique.

### 2.5. Classification performance metrics

Two groups of prediction error metrics are determined to assess the relative lithofacies classification performance of the ML algorithms with the three-well dataset studied. These are: 1) error metrics calculated using the core-interpreted (actual) and predicted facies# comprising MAE, RMSE and correlation coefficient squared ( $R^2$ ); and, 2) classification accuracy measures taking into account all classes ( $\sum$ Accuracy; the fraction of correct classifications achieved), and those considering specific lithofacies categories that measure accuracy (A), precision (P), recall (R) and F1-score (F<sub>1</sub>). The equations used for determining these widely used metrics are provided in Appendix A. Although it is usual to focus on accuracy-related metrics in classification problems, by using a combination of error and accuracy metrics complementary information can often be extracted regarding the prediction performance of ML models.

**Table 5**

Feature selection with the lowest fitness scores found by the optimizers evaluated with the KNN model. The optimization algorithm abbreviations are those defined in Table 3.

Feature Selections of Best Performing Optimizer Runs								
Optimizer	SC	DE	PS	JY	CS	GW	WH	SP
Rank	1 Case 2	2 Case 3	3 Case 4	4	5	6	7	8
GR0	X	X	X	X	X	X	X	X
GR1				X				X
GR2	X		X		X		X	X
GR3	X							
GR4	X							
GR5		X			X	X	X	
GR6	X	X	X	X	X	X	X	
PB0			X	X	X	X		X
PB1					X			X
PB2	X	X	X	X	X	X	X	X
PB3				X			X	X
PB4			X		X			
PB5	X			X		X	X	X
PB6	X	X	X	X	X	X	X	
TH0		X	X		X	X	X	X
K0	X	X	X	X				X
Number of Features Selected	9	7	9	9	10	8	9	10
Accuracy	0.9395	0.9315	0.9315	0.9315	0.9274	0.9234	0.9194	0.9194
Population Size	20	100	50	50	10	100	15	10
Iterations	100	100	100	100	100	100	100	100
Fitness Score (FS)	0.0655	0.0722	0.0735	0.0735	0.0781	0.0808	0.0855	0.0861
Time (sec)	15.6	86.6	69.85	61.6	25.8	79.6	19.3	14.6

Notes: "X" marks features selected by optimizer solution. Optimizer FF found a similar solution to CS but took much longer to do so.

**Table 6**

Multi-K-fold results expressed in terms of mean absolute error (MAE) and root mean squared error for five lithofacies classification feature selections.

Multi-fold Cross Validation Results for KNN Model Applied to Five Datasets Each Predicting Four Lithofacies Classes				
	MAE Mean	MAE Standard Deviation	RMSE Mean	RMSE Standard Deviation
<b>Case 0 (Four Variables: recorded well logs only)</b>				
3-fold validation	0.2731	0.0391	0.7287	0.0673
4-fold validation	0.2807	0.0592	0.7379	0.1070
5-fold validation	0.2700	0.0506	0.7209	0.0943
10-fold validation	0.2707	0.0773	0.7228	0.1320
<b>Case 1 (Sixteen Variables: all attributes included)</b>				
3-fold validation	0.1686	0.0150	0.5523	0.0394
4-fold validation	0.1682	0.0255	0.5545	0.0571
5-fold validation	0.1674	0.0294	0.5452	0.0718
10-fold validation	0.1643	0.0503	0.5377	0.1081
<b>Case 2 (Nine Variables: Sine Cosine optimizer best solution)</b>				
3-fold validation	0.1858	0.0311	0.5757	0.0618
4-fold validation	0.1838	0.0367	0.5740	0.0734
5-fold validation	0.1751	0.0241	0.5557	0.0593
10-fold validation	0.1655	0.0507	0.5316	0.1055
<b>Case 3 (Seven Variables: Differential Evolution optimizer best solution)</b>				
3-fold validation	0.1968	0.0274	0.5898	0.0612
4-fold validation	0.1952	0.0442	0.5875	0.0848
5-fold validation	0.1857	0.0338	0.5746	0.0767
10-fold validation	0.1815	0.0494	0.5574	0.0943
<b>Case 4 (Nine Variables: Particle Swarm optimizer best solution)</b>				
3-fold validation	0.1838	0.0173	0.5724	0.0431
4-fold validation	0.1948	0.0454	0.5895	0.0864
5-fold validation	0.1823	0.0337	0.5690	0.0761
10-fold validation	0.1735	0.0501	0.5483	0.0966

with all the K-folds evaluated. However, the 5-fold results are considered the most attractive for each of the five cases as it generates a relatively low mean values associated with relatively low standard deviations. For this reason, training and validation cases are split 80%: 20% in this study. Fig. 5 displays RMSE versus MAE values for the best-performing ML models for each of the five cases evaluated (Case 0 to Case 4) evaluated by 5-fold cross validation.

For Case 0 (four recorded well-logs only) the RF model provides the lowest MAE and RMSE values (Fig. 5). However, that case is substantially out-performed by the four cases including well-log attributes. Whereas, the SVC model generated the lowest RMSE and MAE values for Case 1, for the other three cases (Cases 2, 3 and 4) the KNN model generated the best 5-fold cross validation results. These 5-fold cross validation comparisons are encouraging regarding the positive contributions the calculated well log attributes can potentially make in terms of improving the lithofacies predictions.

### 3.3. Lithofacies classification results for training/validation subset

Table 7 displays the accuracy and statistical error measures achieved for random validation-subset selections. Results are included for seven ML models trained (80%) and validated (20%) using data from wells V and X. Each model is applied separately to the five cases with distinctive feature selections.

The KNN model achieves the highest  $\sum$ accuracy in lithofacies prediction for the validation subset reported in Table 7. The  $\sum$ accuracy versus MAE results for the KNN models applied to Cases 0 to 4 are illustrated in Fig. 6. The relative positions of the case results in Fig. 6 are consistent with the 5-fold cross validation results (Fig. 4). Case 0 lithofacies predictions for randomly selected validation subsets are poorer than those achieved by Cases 1 to 4, which achieve higher  $\sum$ accuracy ( $\sim 0.87$ – $0.88$ ) and lower MAE ( $\sim 0.17$ – $0.19$ ) and RMSE ( $\sim 0.55$ – $0.57$ ).

For Case 0, the accuracy of the KNN model applied to the randomly generated validation subset is matched by the RF model (Table 7), which also displays lower MAE and RMSE values than the KNN model. That is to be expected based on the RF model's performance in the 5-fold cross validation analysis for Case 0 (Fig. 5). For Case 1, the SVC model achieves slightly lower MAE and RMSE values than the KNN model, but the

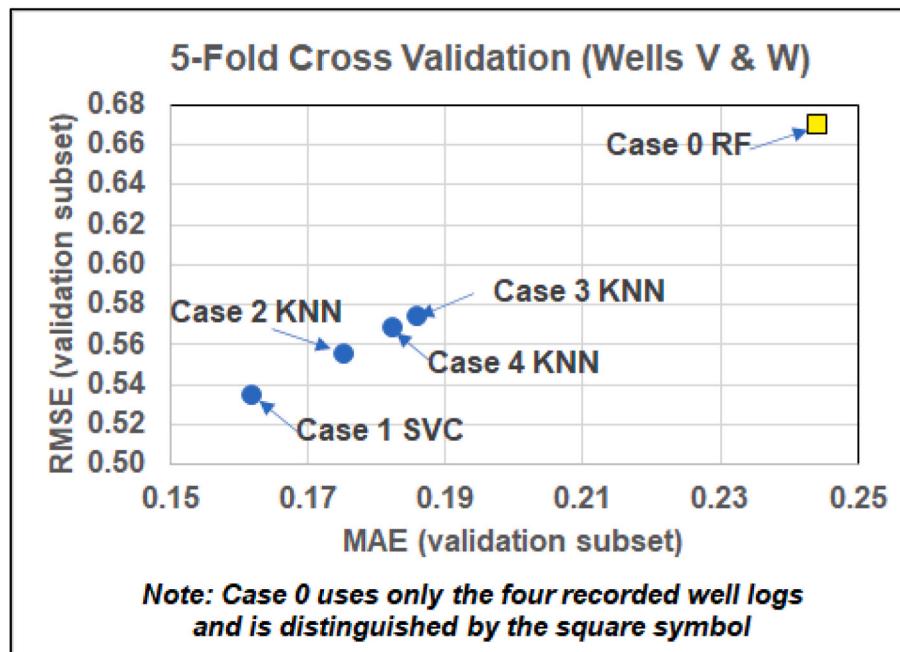
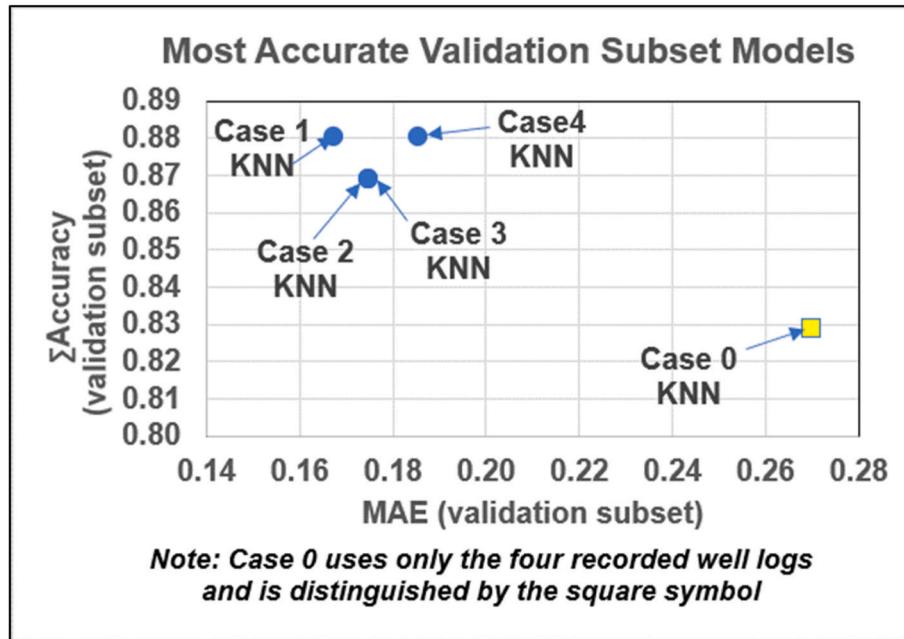


Fig. 5. RMSE versus MAE results achieved for 5-fold cross validation displaying the most accurate model for each of the five cases evaluated. Case 0, using only the four recorded well logs, is distinguished from the cases incorporating well log attributes.

**Table 7**

Validation subset lithofacies classification results for seven ML models applied to five feature-selection cases.

Example Validation Subset Results (20% of Wells V and X)							
	ADA	DT	KNN	LR	RF	SVC	XGB
<b>Case 0 (Four Variables: recorded well logs only)</b>							
MAE	0.3009	0.3081	0.2700	0.2986	0.2441	0.2513	0.2647
RMSE	0.7555	0.7542	0.7209	0.7516	0.6702	0.6791	0.6979
R <sup>2</sup>	0.5644	0.4422	0.5765	0.4192	0.6128	0.5341	0.5644
ΣAccuracy	0.7829	0.7727	0.8286	0.7943	0.8286	0.8229	0.8000
<b>Case 1 (Sixteen Variables: all attributes included)</b>							
MAE	0.2906	0.2956	0.1674	0.2967	0.1880	0.1617	0.1850
RMSE	0.7095	0.7119	0.5452	0.7486	0.5724	0.5349	0.5732
R <sup>2</sup>	0.6128	0.6368	0.6854	0.5160	0.6793	0.7096	0.6914
ΣAccuracy	0.8114	0.8182	0.8800	0.8057	0.8457	0.8686	0.8571
<b>Case 2 (Nine Variables: Sine Cosine optimizer best solution)</b>							
MAE	0.2921	0.2937	0.1751	0.3043	0.1995	0.2258	0.2094
RMSE	0.7444	0.7569	0.5557	0.7457	0.5926	0.6424	0.6020
R <sup>2</sup>	0.5042	0.5200	0.7580	0.4676	0.7398	0.7096	0.5825
ΣAccuracy	0.7714	0.7841	0.8686	0.7943	0.8686	0.8571	0.8171
<b>Case 3 (Seven Variables: Differential Evolution optimizer best solution)</b>							
MAE	0.2807	0.2860	0.1857	0.3039	0.2113	0.2383	0.2143
RMSE	0.7070	0.7103	0.5746	0.7636	0.6224	0.6571	0.6263
R <sup>2</sup>	0.5160	0.4682	0.7519	0.4192	0.6430	0.5220	0.5825
ΣAccuracy	0.7886	0.7386	0.8686	0.8057	0.8457	0.8114	0.8343
<b>Case 4 (Nine Variables: Particle Swarm optimizer best solution)</b>							
MAE	0.2742	0.2814	0.1823	0.2990	0.1934	0.1873	0.1918
RMSE	0.7136	0.7047	0.5690	0.7582	0.5807	0.5774	0.5854
R <sup>2</sup>	0.6067	0.4811	0.8185	0.4434	0.7035	0.6975	0.6491
ΣAccuracy	0.8114	0.8068	0.8800	0.8000	0.8514	0.8457	0.8457

**Fig. 6.** ΣAccuracy versus MAE results achieved with a random validation-subset for the five cases evaluated. Case 0, using only the four recorded well logs, is distinguished from the cases incorporating well log attributes.

KNN model achieves higher accuracy than the SVC model. Slight differences between the error and accuracy performances among the models are to be expected. These are explained in terms of the extent of the misclassification of data records that are incorrectly predicted by a model. The cases where prediction misclassifications involve numerically adjacent facies (e.g. Sd = 2 misclassified as SSh = 3) will have a smaller impact on MAE, RMSE and R<sup>2</sup> than for misclassification that involve more numerically separated facies (e.g. Sh = 0 misclassified as Sh = 3). On the other hand, the accuracy measure only considers whether the prediction for each data record is correct or incorrect.

#### 3.4. Lithofacies classification results for independent testing subset

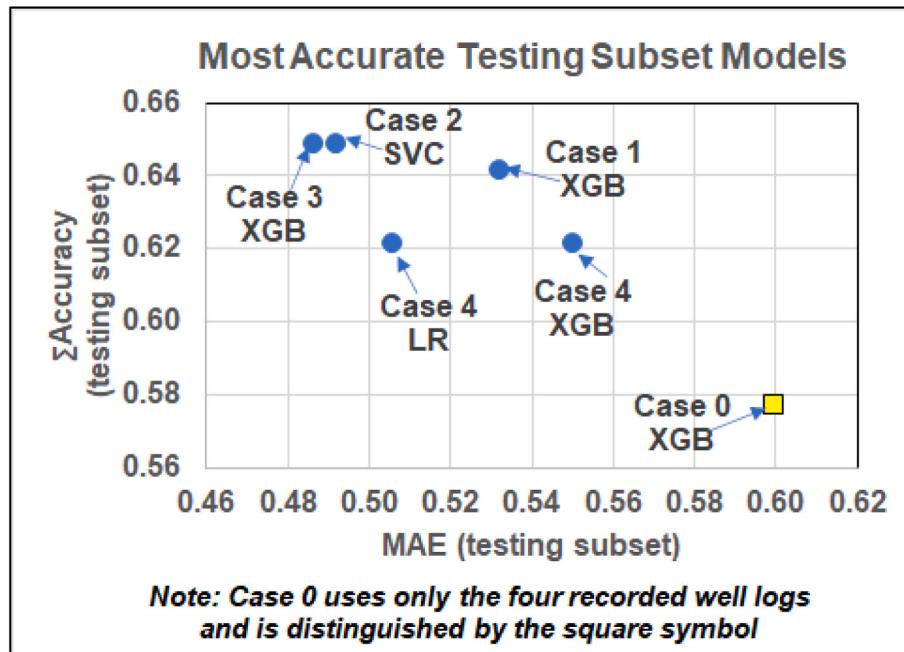
The decisive test for lithofacies classification assesses how each of the models trained and validated with data from one set of wells (i.e., Wells V and X) performs when applied to data from an independent dataset (Well W). The results for the trained/validated ML models applied to all 362 data records of Well W are displayed in Table 8.

The accuracy versus MAE results for the best-performing models applied to Cases 0 to 4 are illustrated in Fig. 7. The prediction errors generated are substantially higher and the accuracy achieved is substantially lower when the ML models are applied to the testing subset

**Table 8**

Testing subset (Well W) lithofacies prediction results for seven ML models (trained and validated with data from Wells V and X) applied to five feature-selection cases.

Example Testing Subset Results (100% of Well W)							
	ADA	DT	KNN	LR	RF	SVC	XGB
<b>Case 0 (Four Variables: recorded well logs only)</b>							
MAE	0.6436	0.6160	0.6003	0.6105	0.6160	0.6519	0.5994
RMSE	1.0629	0.9762	1.0232	0.9239	1.0313	1.0485	1.0041
R <sup>2</sup>	0.1491	0.1045	0.1844	0.2457	0.2281	0.2248	0.2338
ΣAccuracy	0.5691	0.5359	0.5754	0.5683	0.5718	0.5414	0.5773
<b>Case 1 (Sixteen Variables: all attributes included)</b>							
MAE	0.6198	0.6639	0.5372	0.5179	0.5344	0.5702	0.5317
RMSE	1.0245	0.9972	0.9315	0.9271	0.9477	1.0028	0.9889
R <sup>2</sup>	0.1505	0.0858	0.2285	0.2909	0.2738	0.2035	0.1776
ΣAccuracy	0.5730	0.4876	0.6061	0.6171	0.6143	0.6006	0.6419
<b>Case 2 (Nine Variables: Sine Cosine optimizer best solution)</b>							
MAE	0.5691	0.6354	0.5221	0.4972	0.4972	0.4917	0.5193
RMSE	0.9577	1.0668	0.9058	0.8826	0.9043	0.9073	0.9461
R <sup>2</sup>	0.1976	0.1105	0.2536	0.3133	0.2945	0.2944	0.2093
ΣAccuracy	0.5829	0.5691	0.6105	0.6188	0.6381	0.6492	0.6354
<b>Case 3 (Seven Variables: Differential Evolution optimizer best solution)</b>							
MAE	0.5276	0.5801	0.5414	0.4890	0.4696	0.5055	0.4862
RMSE	0.9328	0.9490	0.9431	0.8873	0.8475	0.9149	0.8950
R <sup>2</sup>	0.2319	0.2028	0.2318	0.3154	0.3568	0.2779	0.2857
ΣAccuracy	0.6160	0.5663	0.6133	0.6298	0.6326	0.6298	0.6492
<b>Case 4 (Nine Variables: Particle Swarm optimizer best solution)</b>							
MAE	0.7293	0.8232	0.5470	0.5055	0.5497	0.5856	0.5497
RMSE	1.1658	1.1634	0.9663	0.9058	0.9562	1.0273	0.9847
R <sup>2</sup>	0.0501	0.0085	0.2117	0.3158	0.2734	0.1951	0.1964
Accuracy	0.5331	0.4282	0.6133	0.6215	0.6050	0.5967	0.6215



**Fig. 7.** ΣAccuracy versus MAE results achieved for the independent testing subset evaluated by the five cases evaluated. Case 0, using only the four recorded well logs, is distinguished from the cases incorporating well log attributes.

compared to validation subset (Fig. 6). Nevertheless, the models achieve better results for the feature selections involving including calculated well-log attributes (Cases 1 to 4) than the case based on just the four recorded well logs (Case 0). Of the cases including attributes, Cases 2 and 3 generate more accurate results than Cases 1 and 4 with lower MAE and RMSE values. In general, the XGB model performs better with the testing subset than the other models, providing the best prediction performance for Cases 0, 1 and 3 (Table 8; Fig. 7). For Case 4, the XGB model matches the LR model in terms of accuracy. However, the XGB model applied to the Case 4 testing subset generates higher MAE and RMSE values than the LR, KNN and RF models.

### 3.5. Confusion matrixes and classification scores for individual lithofacies

Confusion matrices and a more detailed consideration of the misclassification metrics (A, P, R, F1 score) provide valuable insight as to the ways in which the models are classifying and misclassifying each of the lithofacies classes. Combining such analysis in annotated confusion matrixes (Wood, 2021) is the most effective way to present such analysis.

Fig. 8 displays an annotated confusion matrix for the XGB model's validation subset for the Case-3 feature selection. As to be expected the lithofacies are sampled in an imbalanced distribution by this subset

XGB Case 3		20% of Wells V and X (Validation Subset)			
175 Total Records					
Records	Class	Sh	Slt	Sd	Ssh
58	Sh	46	4	8	0
17	Slt	8	5	4	0
93	Sd	0	1	92	0
7	Ssh	2	0	2	3
29	Errors	10	5	14	0
<b>0.8343</b> Successful Prediction Fraction ( $\Sigma$ Accuracy)					
Accuracy	0.8690	0.8957	0.9068	0.9733	
Precision	0.8214	0.5000	0.8679	1.0000	
Recall	0.7931	0.2941	0.9892	0.4286	
<b>F1 Score</b>	<b>0.8070</b>	<b>0.3704</b>	<b>0.9246</b>	<b>0.6000</b>	

**Fig. 8.** Annotated confusion matrix for validation subset achieved by the best performing XGB Case 3 model.

reflecting the relative abundance of the lithofacies in Wells V and X (Table 1). The model is much more successful at predicting the more abundant lithofacies (Sh and Sd) than the relatively sparse lithofacies (Slt and SSh). Indeed, even though the overall accuracy is relatively high (0.8343), the recall values for facies Slt ( $R = 0.2941$ ; 5 correct predictions from 17 available records) and SSh ( $R = 0.4286$ ; 3 correct predictions from 7 available data records) are low. This leads to high F1 scores for facies Sh and Sd and low F1 scores for facies Slt and SSh.

It is also apparent from Fig. 8 that whereas facies Sd has a high recall value ( $R = 0.9892$ ; 92 correct predictions from 93 available records), the recall value for facies Sh is  $< 0.8$  ( $R = 0.7931$ ; 46 correct predictions from 58 available records). Moreover, the misclassifications associated with facies Sh, Slt and SSh confuse each of those facies with two other facies. The confusion of facies Slt and SSh with both facies Sh and Sd is to be expected due to the gradational nature of the lithofacies, vertically and laterally, in this sedimentary environment (Fig. 1). In terms of physical and chemical properties facies Slt and SSh typically fit

compositionally and texturally in between shale (Sh) and sandstone (Sd). Of more concern is that 8 out of 58 data records identified from core analysis to be facies Sh are misclassified as facies Sd in the model validation subset.

Fig. 9 displays an annotated confusion matrix for the XGB model's testing subset for the Case-0 feature selection (recorded well logs only). The overall prediction accuracy is  $< 0.6$  and, apart from facies Sd achieving an F1 score of 0.8324, the other three facies achieve F1 scores of  $< 0.5$ , ranging from 0 for SSh to 0.4571 for Sh.

The misclassification is more profound for the Case-0 testing subset in that data records from each of the facies are confused with all three of the other facies. Whereas facies Sh is most likely to be confused with facies Sd, the other three facies are most likely to be confused with facies Sh. Indeed most samples for facies Slt (79 out of 96) are misclassified as facies Sh (Fig. 9). None of the data records identified from core analysis as facies SSh are identified correctly. These results suggest that the models trained and validated with just the four recorded well logs cannot be generalized to predict with a reliable degree of confidence the facies from those well logs in other wells, except for facies Sd.

Fig. 10 displays an annotated confusion matrix for the XGB model's testing subset for the Case-3 feature selection (7 features). This model and case generated the highest overall accuracy with the testing subset (Fig. 7). The classification performance is substantially improved compared to Case 0 (Fig. 9) with the spread of misclassifications much reduced. Facies Sh and Sd achieve recall values of 0.7922 and 0.9244, respectively. Facies Sh is only confused with facies Sd, and facies Sd and SSh are only confused with two other facies. However, the classification performance for facies Slt remains poor, being confused with all three other facies. Slt achieves a recall value of 0.1488 (only 14 data records out of 96 correctly classified). Additionally, only one of the 17 data records of facies SSh is correctly classified.

Fig. 10 suggests that the better performing trained models, involving feature-selected well-log attributes, can be generalized for application to predict with reliable confidence lithofacies Sh and Sd in other wells (not involved in the training/validation process). However, the predictions made by these models of the minor lithofacies (Slt and SSh) remain unreliable.

In such circumstances, another way to apply these models is use

XGB Case 0:		100% of Well W (Testing Subset)			
362 Total Records					
Records	Class	Sh	Slt	Sd	Ssh
77	Sh	56	3	17	1
96	Slt	79	9	6	2
172	Sd	24	2	144	2
17	Ssh	9	1	7	0
153	Errors	112	6	30	5
<b>0.5773</b> Successful Prediction Fraction ( $\Sigma$ Accuracy)					
Accuracy	0.6111	0.6921	0.7828	0.9048	
Precision	0.3333	0.6000	0.8276	0.0000	
Recall	0.7273	0.0938	0.8372	0.0000	
<b>F1 Score</b>	<b>0.4571</b>	<b>0.1622</b>	<b>0.8324</b>	<b>0.0000</b>	

**Fig. 9.** Annotated confusion matrix for testing subset achieved by the best performing XGB Case 0 model (four recorded well-log features only).

XGB Case 3:		100% of Well W (Testing Subset)			
362 Total Records					
Records	Class	Sh	Slt	Sd	Ssh
77	Sh	61	0	16	0
96	Slt	47	14	30	5
172	Sd	12	0	159	1
17	Ssh	8	0	8	1
127	Errors	67	0	54	6
<b>0.6492 Successful Prediction Fraction (<math>\Sigma</math>Accuracy)</b>					
Accuracy    0.7390    0.7413    0.7781    0.9144					
Precision    0.4766    1.0000    0.7465    0.1429					
Recall    0.7922    0.1458    0.9244    0.0588					
<b>F1 Score</b> <b>0.5951</b> <b>0.2545</b> <b>0.8260</b> <b>0.0833</b>					

Fig. 10. Annotated confusion matrix for testing subset achieved by the best performing XGB Case 3 model (seven well-log and attribute features).

them to distinguish between two groups of facies: one group representing the potential oil pay zones, facies Slt and Sd; the other group representing the non-pay zones, facies Sh and SSh. Taking the results presented in Fig. 10 (model XGB Case 3), it is apparent that the combination of facies Slt and Sd can be predicted with an accuracy of 0.7575, and the combination of facies Sh and SSh can be predicted with an accuracy of 0.7447. On the other hand, Taking the results presented in Fig. 9 (model XGB Case 0), it is apparent that the combination of facies Slt and Sd can only be predicted with an accuracy of 0.6007, whereas the combination of facies Sh and SSh can be predicted with an accuracy of 0.7021. Being able to use the XGB Case-3 model to predict pay zones from well-log and attribute data in the independent testing well with >0.75 accuracy makes it a valuable tool.

#### 4. Discussion

##### 4.1. Relative influence of well-log features on case solutions

Monitoring the Gini indices of each feature involved in the XGB model solutions makes it possible to establish the relative contributions (importance weights) the XGB model assigns to each feature involved in

its specific trained model solutions. Fig. 11 displays the XGB feature contributions applied to its Case 0 model established using the training/validation subset. The ranking of those contributions is GR0>TH0>PBO>KO, and that ranking order is the same as the relative magnitudes of the Spearman correlation coefficients between the features and facies numbers (Fig. 4). The dominance of GR0 and TH0 in determining these trained XGB solutions can therefore be substantially explained in terms of their high Spearman's correlation coefficients with facies number: -0.6431 and -0.6088, respectively.

It is also apparent from Fig. 4, that high Pearson's correlation coefficients exist between GR0 versus TH0 and KO: 0.7869 and 0.7904, respectively. Moreover, the Pearson's correlation coefficient between GR0 and PBO is 0.4191, which is still moderately high but less so than for GR0 versus TH0 and KO. The relatively high correlations between the four recorded well logs available may act to limit the ability of the four recorded well logs to accurately discriminate between the four lithofacies. Fig. 12 displays the XGB feature contributions applied to the Case-1 to Case-4 models established using the training/validation subset.

For models XGB Case 1 (Fig. 12A) and Case 4 (Fig. 12D) the GR0, TH0 and PBO well logs values, in that order are the most influential features. This is consistent with the influences established for XGB Case 0 (Fig. 11). For Case 1 relatively small importance weights (up to ~0.05) are assigned to the other 13 features included. Similarly, for XGB Case 4 relatively small importance weights (up to ~0.08) are assigned to the other 6 features included. XGB Case 2 is distinct from the other high-performing feature selections identified with the aid of the optimizers in that it excludes TH0. Case 2 also excludes PBO and assigns the highest influence to GR0 (~0.3) of all the models displayed in Fig. 12. Case 2 assigns relatively small importance weights (from ~0.6 to ~0.11) to the other 8 features included. XGB Case 3 also excludes PBO but assigns its highest importance weights to GR0 and TH0, in that order. Case 3 assigns similar relatively small importance weights (~0.1) to the other 5 features included.

Of the attributes included in XGB Cases 2, and 4, features GR2, PB2, and PB6 display slightly higher Pearson's correlation coefficients with facies number than the other attributes (Fig. 4). The attributes GR2, GR6, PB2 and PB6 are also assigned slightly more influence in XGB Case 1 (Fig. 12A) than the other attributes. In XGB Case 3 (Fig. 12C) the

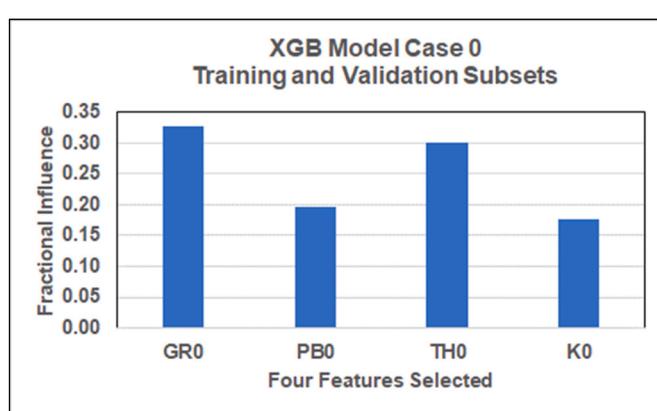
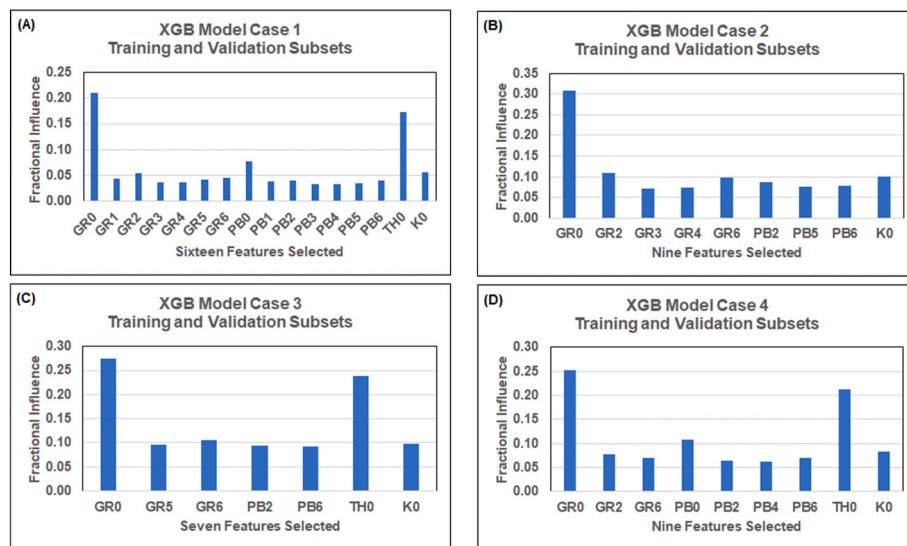


Fig. 11. Feature influences on the trained and validated XGB Case 0 model (four recorded well-log features only).



**Fig. 12.** Feature influences on the trained and validated XGB models: (A) Case 1; (B) Case 2; (C) Case 3; and (D) Case 4.

attribute GR6 is assigned the third highest importance weight ahead of K0. The feature influences displayed in Fig. 12 highlight the value of derivative and volatility attributes of the GR ad PB logs to all four high performing feature selections in improving their lithofacies classification compared to that achieved by Case 0.

#### 4.2. Approaches to improve lithofacies predictions in gradational formations

Generating reliable lithofacies classification from a sparse suite of recorded well logs in gradational and spatially heterogeneous reservoir formations, such as those generated by braided fluvial environments poses a substantial challenge. In the three-well dataset evaluated, using only the four recorded well logs to develop ML prediction models for the lithofacies determined by core analysis leads to substantial inaccuracy. This is particularly so when applying trained models to well locations not involved in the training/validation of the models. (e.g., XGB Case 0 with  $<0.58 \sum$  accuracy, and  $\sim 0.6$  accuracy in predicting the testing subset lithofacies associated with oil pay) Calculating well-log attributes for some of the recorded logs and using optimizers to select features from the recorded well logs and the calculated attributes generates models that substantially improve the lithofacies classification performance (e.g., XGB Case 3 with  $\sim 0.65 \sum$  accuracy and  $\sim 0.75$  accuracy in predicting lithofacies associated with oil pay). However, scope remains to further improve lithofacies prediction accuracy achieved by the attributed-enhanced models developed in this study. Three possible ways to do this are:

1. Core additional reservoir section in development wells yet to be drilled to improve the core dataset to be used for calibration across the field. This is a very expensive option and is the reason why cores are typically only recovered from very few wells in most producing oil fields.
2. Expand the suite of well logs recorded across the reservoir sections in both cored and non-cored wellbores. This is not easily achieved retrospectively in wells already drilled and completed in a producing reservoir. Additional well logs that record physical properties of the reservoir sections, such as compressional-wave and shear wave-sonic logs, would add an additional dimensions and diversity to the existing four-log suite. One of the problems with the existing four log suite is that the GR0, Th0 and K0 log distributions are highly correlated resulting in a lack of diversity and contrast within the

well-log data base. However, such an approach is also costly and would take substantial time and cost to achieve.

3. Evaluate alternative well-log attributes calculated from the existing suite of well logs to determine whether they could further improve upon the lithofacies classification achieved by the derivative and volatility attributes calculated using the recorded GR0 and PB0 curves. Combined with the feature-selection optimization method applied in this study, such an approach offers the quickest and lowest cost approach to apply. Future studies are planned to evaluate other mathematical attributes of the sparse recorded well-log suite available for the studied field, to determine if, together with recorded well logs and derivative and volatility attributes, they can further improve the lithofacies predictions.

#### 5. Conclusions

Meandering, braided fluvial depositional environments generate clastic reservoir formations that tend to lack laterally continuity of their sandstone and siltstone pay zones. This makes it difficult for lithofacies classification machine-learning (ML) models, supervised with available core and well-log data in a few wells, to reliably predict lithofacies in non-cored wellbores distributed across the entire field area. That problem is heightened when only a small suite of well logs is recorded, and those recorded well logs are highly correlated with each other. The dataset evaluated consists of three cored wells drilled into such a reservoir system with only a spectral gamma ray log recorded in the field development wells. The available recorded log suite is restricted to the total gamma ray (GR), density (PB), Thorium (TH) and Potassium (K) spectral signals. To supplement the recorded well-log data six derivative and volatility attributes were calculated for each cored depth interval covered by the GR and PB curves. This generated a dataset of 1236 data records with sixteen log-derived features and four lithofacies classes verified by core analysis.

A novel feature-selection technique coupling multiple evolutionary optimization algorithms with a K-nearest neighbor (KNN) model, applied to the complete dataset, was able to identify several high-performing feature combinations. The best feature combinations utilized just seven to ten of the available features for lithofacies classification. Five feature-combination cases were evaluated in detail with seven ML algorithms. Case 0 involved just the four recorded well logs, Cases 2 and 4 both included nine different features, and Case 3 included just seven features. Two of the three wells (V and X) were used for ML model training and

validation, and the other well (W) was used for independently testing the trained models.

Comprehensive multi-k-fold cross validation applied to the training/validation subset established that a data split of 80% training: 20% validation was most effective. It also demonstrated convincingly that Cases 1 to 4 achieved lithofacies classification of the validation subset with substantially fewer errors and higher overall accuracy ( $\sum$ accuracy) than Case 0. The KNN model achieved  $\sum$ accuracy  $\sim 0.87$  for Cases 1 to 3 versus  $\sim 0.83$  for Case 0 with respect to randomly selected validation subsets. However, when the trained models were applied to the testing subset (Well W) substantially reduced  $\sum$ accuracy was achieved  $\sim 0.65$  for Cases 2 and 3 versus  $\sim 0.58$  for Case 0. Generalizability of the lithofacies classification model to other reservoir wells in this heterogeneous reservoir poses a clear challenge. The extreme gradient boosting (XGB) model generated the most reliable results with the testing subset in this regard. Confusion matrices revealed that the classification performance for the individual facies was more encouraging for the models enhanced with well-log attributes. The trained and validated XGB model for Case 3 applied to the testing subset achieved an accuracy of  $\sim 0.76$  for the combined pay zone facies (sandstone plus siltstone) compared to only  $\sim 0.60$  accuracy for that combination for XGB Case 0.

Consideration of the Gini indices for the XGB models made it possible to establish the relative contributions of the selected features to the

model solutions. For Case 0, the feature importance ranking of the recorded logs was revealed as GR0>TH0>PB0>K0. For the best-performing model with attributes, XGB Case 3, the feature ranking was GR0>TH0>GR6>K0>GR5>PB2>PB6, highlighting the importance both volatility and derivative attributes to that solution. The limited success in generalizing the trained and validated ML models with attributes to well location away from the training/validation wells indicates that scope remains to improve the lithofacies classification using the available well-logs across this heterogeneous reservoir. The lithofacies classification improvements generated by including GR and PB log attributes, justify future work to search for additional well-log attributes that could complement the set of attributes considered in this study.

## Funding

No funding was received for this study.

## Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Statistical measures used to assess lithofacies predictions

Figure A defines prediction performance assessment metrics referred to in this study.

### Statistical Measures of Prediction Performance Assessed

**Mean Squared Error (MSE):**

$$MSE = \frac{1}{n} \sum_{i=1}^n ((X_i) - (Y_i))^2$$

**Root Mean Squared Error (RMSE):**

$$RMSE = \sqrt{MSE}$$

**Mean Absolute Error (MAE):**

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i - Y_i|$$

**Correlation Coefficient (R):**

$$R = \frac{\sum_{i=1}^n (X_i - Xmean)(Y_i - Ymean)}{\sqrt{\sum_{i=1}^n (X_i - Xmean)^2 \sum_{i=1}^n (Y_i - Ymean)^2}}$$

R is expressed on a scale of -1 to +1

$R^2$  is expressed on a scale of 0 to 1

**Where:**

$X_i$  is the measured value and  $Y_i$  is the predicted value for data record  $i$

$n$  is the number of data records in the set or subset being evaluated

$D_i$  is  $(X_i - Y_i)$  for data record  $i$

$$Dmean = \frac{1}{n} \sum_{i=1}^n (X_i - Y_i)$$

### Measures specifically focused on classification and misclassification

**Fraction of correct classifications considering all classes ( $\Sigma$  Accuracy)**

$$\text{Accuracy (A) of a specific class: } A = (TP+TN)/(TP+TN+FP+FN)$$

$$\text{Precision (P) of a specific class: } P = TP/(TP+FP)$$

$$\text{Recall (R) of a specific class: } R = TP/(TP+FN)$$

$$\text{Balanced F-score (F1) of a specific class: } F1 = (2 \times (R \times P))/(R+P)$$

**Where:**

$TP$  = true number of positive predictions,  $TN$  = true number of negative predictions,

$FP$  = number of false positive predictions and  $FN$  = number of false negative predictions.

These classification prediction error assessments are most usefully displayed and interpreted in conjunction with an annotated confusion matrix.

**Fig. A.** Prediction performance measures used to assess lithofacies classification.

### References

- Abiodun, E.O., Alabdulatif, A., Abiodun, O.I., Alawida, M., Abdullah Alabdulatif, A., Alkhawaldeh, R.S., 2021. A systematic review of emerging feature selection optimization methods for optimal text classification: the present state and prospective opportunities. *Neural Comput. Appl.* 33, 15091–15118. <https://doi.org/10.1007/s00521-021-06406-8>.
- Abualigah, L., Diabat, A., 2021. Advances in sine cosine algorithm: a comprehensive survey. *Artif. Intell. Rev.* 54, 2567. <https://doi.org/10.1007/s10462-020-09909-3>, 260.
- Agrawal, R., Malik, A., Samuel, R., Saxena, A., 2022. Real-time prediction of litho-facies from drilling data using an artificial neural network: a comparative field data study with optimizing algorithms. *J. Energy Resour. Technol.* 144, 043003 <https://doi.org/10.1115/1.4051573>, 12.
- Al-Mudhafer, W.J., Abbas, M.A., Wood, D.A., 2022. Performance evaluation of boosting machine learning algorithms for lithofacies classification in heterogeneous carbonate reservoirs. *Mar. Petrol. Geol.* 145, 105886 <https://doi.org/10.1016/j.marpgeo.2022.105886>.
- Alyasser, Z.A.A., Khader, A.T., Al-Betar, M.A., Awadallah, M.A., Yang, X.S., 2018. Variants of the flower pollination algorithm: a review. In: Yang, X.S. (Ed.), *Nature-Inspired Algorithms and Applied Optimization*, Studies in Computational Intelligence, vol. 744. Springer, Cham, 10.1007/978-3-030-12127-3\_11.
- Farzani, R., Bolandi, V., 2016. Estimation of organic facies using ensemble methods in comparison with conventional intelligent approaches: a case study of the South Pars Gas Field, Persian Gulf, Iran. *Model. Earth Syst Environ.* 2, 105. <https://doi.org/10.1007/s40808-016-0165-z>.
- Fix, E., Hodges Jr., J.L., 1951. *Discriminatory Analysis, Nonparametric Discrimination: Consistency Properties*. USAF School of Aviation Medicine, Technical Report.
- Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* 55, 119–139. <https://doi.org/10.1007/jcss.1997.1504>.
- Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Krishnapuram, Balaji, Shah, Mohak, Smola, Alexander J., Aggarwal, Charu C., Shen, Dou, Rastogi, Rajeev (Eds.), *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 2016. ACM, San Francisco, CA, USA, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>, August 13–17.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20 (3), 273–297. <https://doi.org/10.1007/BF00994018>.
- Cox, D.R., 1958. The regression analysis of binary sequences. *J. Roy. Stat. Soc. B* 20 (2), 215–242. <http://www.jstor.org/stable/2983890>.
- Dubois, M.K., Bohling, G.C., Chakrabarti, S., 2007. Comparison of four approaches to a rock facies classification problem. *Comput. Geosci.* 33, 599–617.
- Faris, H., Mirjalili, S., Aljarah, I., Mafarja, M., Heidari, A.A., 2020. Salp swarm algorithm: theory, literature review, and application in extreme learning machines. In: Mirjalili, S., Song Dong, J., Lewis, A. (Eds.), *Nature-inspired Optimizers*, Studies in Computational Intelligence, vol. 811. Springer, Cham. [https://doi.org/10.1007/978-3-030-12127-3\\_11](https://doi.org/10.1007/978-3-030-12127-3_11).
- Farzi, R., Bolandi, V., 2016. Estimation of organic facies using ensemble methods in comparison with conventional intelligent approaches: a case study of the South Pars Gas Field, Persian Gulf, Iran. *Model. Earth Syst Environ.* 2, 105. <https://doi.org/10.1007/s40808-016-0165-z>.
- Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Krishnapuram, Balaji, Shah, Mohak, Smola, Alexander J., Aggarwal, Charu C., Shen, Dou, Rastogi, Rajeev (Eds.), *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 2016. ACM, San Francisco, CA, USA, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>, August 13–17.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20 (3), 273–297. <https://doi.org/10.1007/BF00994018>.
- Cox, D.R., 1958. The regression analysis of binary sequences. *J. Roy. Stat. Soc. B* 20 (2), 215–242. <http://www.jstor.org/stable/2983890>.
- Dubois, M.K., Bohling, G.C., Chakrabarti, S., 2007. Comparison of four approaches to a rock facies classification problem. *Comput. Geosci.* 33, 599–617.
- Faris, H., Mirjalili, S., Aljarah, I., Mafarja, M., Heidari, A.A., 2020. Salp swarm algorithm: theory, literature review, and application in extreme learning machines. In: Mirjalili, S., Song Dong, J., Lewis, A. (Eds.), *Nature-inspired Optimizers*, Studies in Computational Intelligence, vol. 811. Springer, Cham. [https://doi.org/10.1007/978-3-030-12127-3\\_11](https://doi.org/10.1007/978-3-030-12127-3_11).
- Farzi, R., Bolandi, V., 2016. Estimation of organic facies using ensemble methods in comparison with conventional intelligent approaches: a case study of the South Pars Gas Field, Persian Gulf, Iran. *Model. Earth Syst Environ.* 2, 105. <https://doi.org/10.1007/s40808-016-0165-z>.
- Fix, E., Hodges Jr., J.L., 1951. *Discriminatory Analysis, Nonparametric Discrimination: Consistency Properties*. USAF School of Aviation Medicine, Technical Report.
- Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* 55, 119–139. <https://doi.org/10.1007/jcss.1997.1504>.

- Goncalves, C.A., Harvey, P.K., Lovell, M.A., 1995. Application of a multilayer neural network and statistical techniques in formation characterization. In: SPWLA 36th Annual Logging Symposium, Paris, 26–29 Jun, Society of Petrophysicists and Well Log Analysts, Houston (U.S.A.), p. 12.
- Halotel, J., Demyanov, V., Gardiner, A., 2020. Value of geologically derived features in machine learning facies classification. *Math. Geosci.* 52, 5–29. <https://doi.org/10.1007/s11004-019-09838-0>.
- Ho, T.K., 1998. The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (8), 832–844. <https://doi.org/10.1109/34.709601>.
- Hussien, A.G., Abualigah, L., Abu Zitar, R., Hashim, F.A., Amin, M., Saber, A., Almotairi, K.H., Gandomi, A.H., 2022. Recent advances in Harris Hawks optimization: a comparative study and applications. *Electronics* 11, 1919. <https://doi.org/10.3390/electronics11121919>.
- Ma, Y.Z., 2011. Lithofacies clustering using principal component analysis and neural network: applications to wireline logs. *Math. Geosci.* 43, 401–419. <https://doi.org/10.1007/s11004-011-9335-8>, 2011.
- Ma, Y.Z., 2019. Facies and lithofacies classifications from well logs. In: Quantitative Geosciences: Data Analytics, Geostatistics, Reservoir Characterization and Modeling. Springer, Cham. [https://doi.org/10.1007/978-3-030-17860-4\\_10](https://doi.org/10.1007/978-3-030-17860-4_10).
- Mansouri, V., Khosravaniyan, R., Wood, D.A., Aadnoy, B.S., 2015. 3-D well path design using a multi-objective genetic algorithm. *J. Nat. Gas Sci. Eng.* 27 (1), 219–235. <https://doi.org/10.1016/j.jngse.2015.08.051>.
- Masapanta, J., 2021. Machine and Deep Learning for Lithofacies Classification from Well Logs in the North Sea. Thesis. University of Stavanger, Norway, p. 178. <https://uis-brage.unit.no/uis-xmlui/handle/11250/2786292>.
- Merembayev, T., Kurmangaliyev, D., Bekbauov, B., Amanbek, Y.A., 2021. Comparison of machine learning algorithms in predicting lithofacies: case studies from Norway and Kazakhstan. *Energies* 1896. <https://doi.org/10.3390/en14071896>.
- Mirjalili, S., Mirjalili, S.M., Lewis, A., 2014. Grey wolf optimizer. *Adv. Eng. Software* 69, 46–61. <https://doi.org/10.1016/j.advengsoft.2013.12.007>.
- Ning, G.-Y., Cao, D.-Q., 2021. Improved whale optimization algorithm for solving constrained optimization problems. *Discrete Dynam. Nat. Soc.* 2021, 8832251, 13. <https://doi.org/10.1155/2021/8832251>, 13.
- Quinlan, J.R., 1986. Induction of decision trees. *Mach. Learn.* 1, 81–106. <https://doi.org/10.1007/BF00116251>.
- Rao, R.V., 2016. Jaya: a simple and new optimization algorithm for solving constrained and unconstrained optimization problems. *Int. J. Ind. Eng. Comput.* 7, 19–34. <https://doi.org/10.5267/j.ijiec.2015.8.004>.
- Reverdy, X., Argaud, M., Walgenwitz, F., 1983. Mineralogical analysis required for log interpretation in complex lithologies. In: Paper H, Transactions of the SPWLA 8th European Symposium.
- Rider, M.H., 1990. Gamma-ray log shape used as a facies indicator: critical analysis of an oversimplified methodology. Geological Society, London, Special Publications 48, 27–37. <https://doi.org/10.1144/GSL.SP.1990.048.01.04>.
- Rogers, S.J., Fang, J., Karr, C., Stanley, D., 1992. Determination of lithology from well logs using a neural network (1). *AAPG (Am. Assoc. Pet. Geol.) Bull.* 76, 731–739.
- Sarkar, S., Majundar, C., 2020. A comparative analysis of supervised classification algorithms for lithofacies characterization. In: EAGE Digitalization Conference and Exhibition, Nov 2020. European Association of Geoscientists & Engineers, pp. 1–5. <https://doi.org/10.3997/2214-4609.202032090>.
- SciKit, Learn, 2022a. Supervised and Unsupervised Machine Learning Models in Python, 2022a. Accessed 24th September 2022. <https://scikit-learn.org/stable/>.
- SciKit, Learn, 2022b. GridSearchCV: Exhaustive Search over Specified Parameter Values for an Estimator in Python, 24th September 2022. [https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.GridSearchCV.html](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html).
- SciKit, Learn, 2022c. Bayesian Optimization of Hyperparameters in Python. Accessed 24th September 2022. <https://scikit-optimize.github.io/stable/modules/generated/skopt.BayesSearchCV.html>.
- SciKit, Learn, 2022d. Cross-validation: Evaluating Estimator Performance. Accessed 24th September 2022. [https://scikit-learn.org/stable/modules/cross\\_validation.html](https://scikit-learn.org/stable/modules/cross_validation.html).
- Stowe, L., Hock, M., 1988. Facies analysis and diagenesis from well logs in the Zechstein carbonates of Northern Germany. In: Conference Paper SPWLA 29th Annual Logging Symposium, San Antonio, Texas, June 1988 (SPWLA-1988-HH).
- Wood, D.A., 2016a. Hybrid bat flight optimization algorithm applied to complex wellbore trajectories highlights the relative contributions of metaheuristic components. *J. Nat. Gas Sci. Eng.* 32, 211–221. <https://doi.org/10.1016/j.jngse.2016.04.024>.
- Wood, D.A., 2016b. Hybrid cuckoo search optimization algorithms applied to complex wellbore trajectories aided by dynamic, chaos-enhanced, fat-tailed distribution sampling and metaheuristic profiling. *J. Nat. Gas Sci. Eng.* 34, 236–252. <https://doi.org/10.1016/j.jngse.2016.06.060>.
- Wood, D.A., 2018. Thermal maturity and burial history modelling of shale is enhanced by use of Arrhenius time-temperature index and memetic optimizer. *Petroleum* 4, 25–42. <https://doi.org/10.1016/j.petlm.2017.10.004>.
- Wood, D.A., 2019. Lithofacies and stratigraphy prediction methodology exploiting an optimized nearest-neighbour algorithm to mine well-log data. *Mar. Petrol. Geol.* 110, 347–367. <https://doi.org/10.1016/j.marpgeo.2019.07.026>.
- Wood, D.A., 2021. Enhancing lithofacies machine learning predictions with gamma-ray attributes for boreholes with limited diversity of recorded well logs. *Artificial Intell. Geosci* 2, 148–164. <https://doi.org/10.1016/j.aiig.2022.02.007>.
- Wood, D.A., 2022. Gamma-ray log derivative and volatility attributes assist facies characterization in clastic sedimentary sequences for formulaic and machine learning analysis. *Adv. Geo-Energy Res* 6 (1), 69–85. <https://doi.org/10.46690/ager.2022.01.06>.
- Zhang, Y., Li, Y., Guo, W., Li, Y., Dang, H., 2020. Differential evolution and the influencing factors of low-maturity terrestrial shale with different types of kerogen: a case study of a Jurassic shale from the northern margin of Qaidam Basin, China. *Int. J. Coal Geol.* 230, 103591 <https://doi.org/10.1016/j.coal.2020.103591>.