

A machine-learning approach for classifying fluvial sandbody types from vertical facies sequences using geological analogues

Luca Colombera^{a,*}, Soma Budai^a, Nigel P. Mountney^b

^a Dipartimento di Scienze della Terra e dell'Ambiente, Università di Pavia, Italy

^b School of Earth and Environment, University of Leeds, UK

ARTICLE INFO

Editor: Dr. Catherine Chagué

Keywords:

Artificial intelligence
Facies models
Channel
Overbank
Braided
Meandering

ABSTRACT

Machine-learning tools exist for classifying sedimentary units according to objective and verifiable labels (e.g., lithological types), but the role of geologists in interpretive facies analyses is considered irreplaceable. This study presents a machine-learning method for automating interpretations of fluvial sandbodies following an approach that mirrors the application of classic one-dimensional facies models presented as idealized vertical facies sequences. Ensemble decision-tree models were trained on descriptors of facies sequences (sandbody thickness, average facies thickness, facies proportion, and vertical thickness trend), documented in many geological analogues studied by different research groups, reflecting global sedimentological thinking. The data were synthesized in two alternative datasets of contrasting dimensionality (50 vs 26 features). Up to 1222 facies sequences were used for model training and testing, with data splits operated in different ways: (i) using a partition algorithm, with options to exclude data from two-dimensional outcrop panels and to prevent the same sedimentary bodies from occurring in both splits; (ii) by arbitrarily picking sandbodies that are thought to be easily misinterpreted. Four ensemble machine-learning models were applied to perform binary classifications of (i) fluvial sandbody types according to their general channel or overbank origin, and (ii) channel sandbody types according to the interpreted planform style of their formative rivers ('low-sinuosity or braided' vs 'meandering'). Across all training approaches, the models for general sandbody classification yield accuracy values ranging between 0.76 and 0.87: on average only 16% of sandbodies are misclassified. Comparable predictive power (accuracy: 0.75–0.91) is seen for models classifying formative river patterns from channel-body facies sequences, a result that challenges the much-cited notion that interpreting river planforms from vertical profiles of channel deposits is futile. A benchmark comparison against interpretations by eight sedimentologists demonstrates that cases of model misclassification are in line with errors by human geologists. The outcomes support the possibility of automating sedimentological interpretations of borehole observations (e.g., image logs) using models trained on geological analogues.

1. Introduction

A key task in the characterization of clastic reservoir and aquifer successions is the classification of sandbodies according to their likely origin, which is commonly done on the basis of sedimentological interpretations of one-dimensional borehole stratigraphies based on cored intervals or interpreted wireline logs (Barwis et al., 1990; Slatt, 2013). Typically, this process leads to a categorization of sandbodies in terms of likely sub-environments of deposition, which is the basis for attempting predictions of the likely geometrical and petrophysical properties of reservoir and aquifer building blocks. In fluvial successions, a broad

classification of sandbodies according to their inferred deposition in overbank and channel settings is commonly operated, since this is an important distinction for subsurface modelling workflows in consideration of how these sandbodies contrast in terms of shape, internal heterogeneity and impact on sandstone connectivity (cf. Miall, 1988; Larue and Hovadik, 2006; Yeste et al., 2020; Colombera and Mountney, 2021). A more detailed classification of the facies associations observed as forming channel and/or overbank sandbodies is also generally desired, since it allows geologists to be more predictive in subsurface studies. For example, overbank sandbodies may be differentiated depending on whether they are deposited following levee overtopping or crevassing,

* Corresponding author.

E-mail address: luca.colombera@unipv.it (L. Colombera).

<https://doi.org/10.1016/j.sedgeo.2026.107047>

Received 25 November 2025; Received in revised form 30 January 2026; Accepted 6 February 2026

Available online 9 February 2026

0037-0738/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

since overbank and crevasse splays exhibit contrasting architectures and topological relationships with related channel deposits (cf. Lewin et al., 2017; Burns et al., 2019). Similarly, a more detailed classification of channel bodies based on the interpreted channel pattern of their formative rivers – whether of meandering or low-sinuosity, including braided, planforms – can support predictions of the distribution of sand-prone volumes, styles of sandstone compartmentalization and possible presence of permeability contrasts (cf. Lunt et al., 2004; Donselaar and Overeem, 2008; Colombera et al., 2017; Willis and Sech, 2019).

Interpretations of sandbody types made on core and wireline-log datasets are often based on facies models presented as idealized one-dimensional (1D) facies sequences ('facies successions'), which overall capture the proportion, thickness and vertical trends of the constituent lithofacies of each sandbody type (Walker, 1992; Dalrymple, 2010). In an alternative approach, interpretations may be guided by reference to sedimentary logs from specific geological analogues consisting of ancient successions observable in outcrop (Fig. 1), or from the shallow subsurface of modern depositional environments for which the depositional context is known (Jordan and Pryor, 1992; Lang et al., 2002). Facies sequences presented in 1D facies models and sedimentary logs are especially applicable to image logs that permit to determine sediment textures and sedimentary structures in absence of core (Fig. 1) and which typically provide a comprehensive coverage of reservoir sandbodies penetrated by wells (Donselaar and Schmidt, 2005; Pöppelreiter et al., 2010). However, it is often claimed that 1D facies models or selected representative sedimentary logs are of limited value for guiding the interpretations of facies associations of subsurface fluvial successions: because of the large variability in the facies organization of fluvial sandbodies and the degree of overlap in the facies make-ups of different sandbody types, it has been argued that traditional 1D facies models fail to capture natural variability and tend to overemphasize differences between types (Miall, 1978; Dott Jr and Bourgeois, 1983; Anderton, 1985). As a result, any interpretation based on such models may remain highly uncertain.

The idea of relying on selected geological analogues or facies models can be overcome by finding ways to leverage on the large knowledge base of sedimentological studies conducted over several decades on fluvial successions: by considering a breadth of case studies, it is possible to capture the natural variability of facies sequences observed in sandbody types, especially if their sedimentological properties are quantified (Colombera and Mountney, 2019; Cosgrove et al., 2021; Budai et al., 2025). This raises the question as to whether we could make use of quantitative facies data from many successions to train machine-learning tools for automating interpretations of sandbody types. Machine-learning classification and regression methods are now widely applied in several tasks of subsurface prediction. Several tools have been proposed for the classification of subsurface stratigraphic units and lithologies, but these are mostly applicable in situations in which the classes of interest are defined objectively (e.g., lithologies; cf. Hall and Hall, 2017; Alaudah et al., 2019; Halotel et al., 2020; Bressan et al., 2020; Li et al., 2021, 2023). The classification of sandbodies according to their origin, instead, is inherently interpretational: the subjective nature of sandbody labelling makes the development of machine-learning tools that can be applied to this task especially challenging. Here we address the research question as to whether machine-learning tools can assist interpretations of 1D facies sequences that are commonly performed on the basis of qualitative facies analysis.

The aim of this work is to develop and test a novel way to use data from sedimentary successions to train machine-learning methods for automating the classification of fluvial sandbodies in subsurface characterization workflows. Specific objectives are as follows: (i) to define quantitative datasets of facies properties of fluvial sandbodies that can be used to train machine-learning models for sandbody classification; (ii) to develop alternative workflows for training and optimizing the machine-learning models depending on analogue data types (outcrop panels vs sedimentary logs); (iii) to assess the performance of the

machine-learning models in tasks of prediction of fluvial sandbody types classified according to environment of deposition (channel vs overbank) and formative-river planform type (meandering vs low-sinuosity or braided).

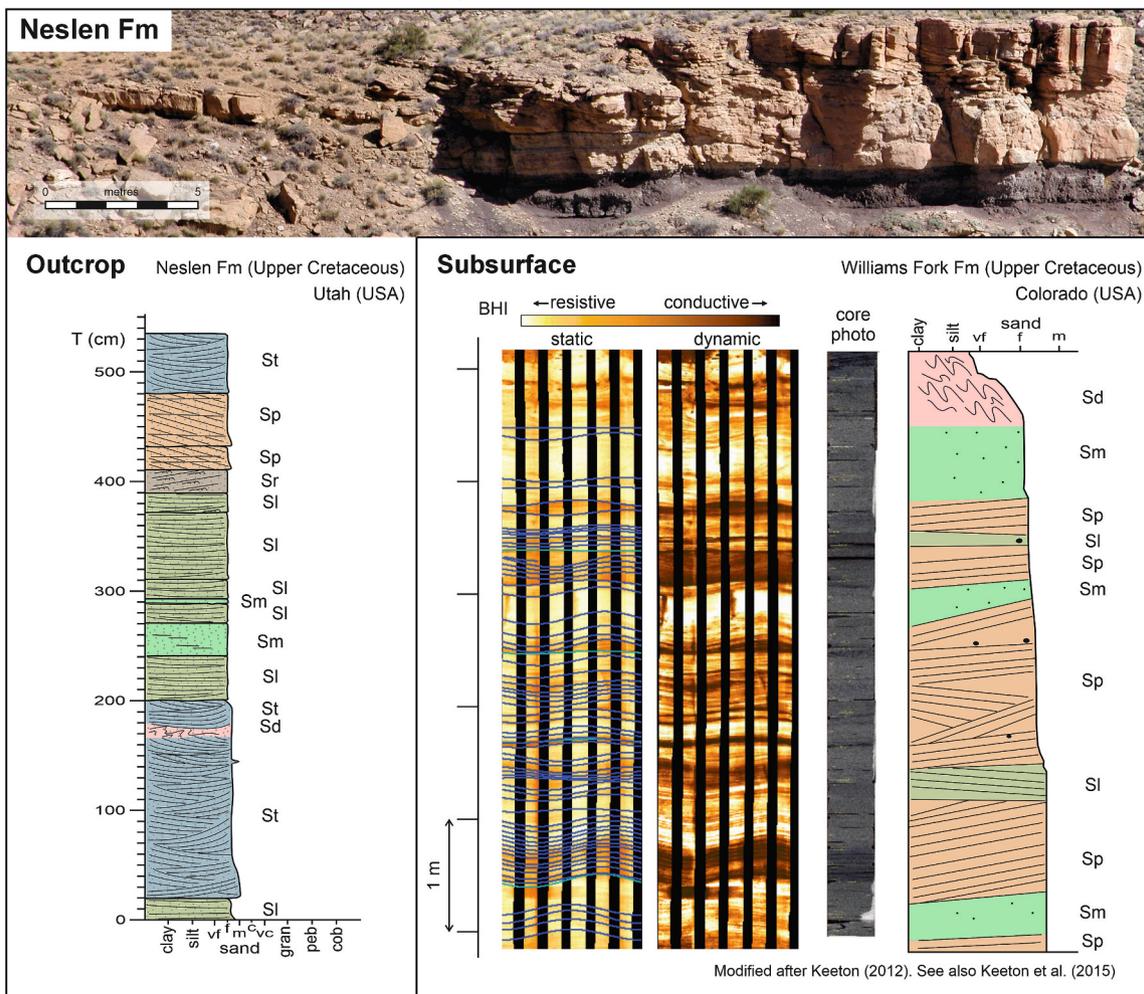
2. Data and methods

2.1. Database

The dataset used in this study has been collated in a large sedimentological database, the Fluvial Architecture Knowledge Transfer System, FAKTS (Colombera et al., 2012, 2013). FAKTS is a SQL (Structured Query Language) relational database that contains data on different types of sedimentary units observed in ancient fluvial successions and in the recent deposits of modern rivers. FAKTS contains data obtained from published sources and unpublished field studies of sedimentary successions of any age and from a broad range of environmental settings. The majority of the sedimentological data included in FAKTS are derived from peer-reviewed literature studies by many different authors; thus, all data are harmonized following a common standard to enable the production of meaningful data syntheses (Colombera et al., 2012, 2013). In the database, the depositional context and geological setting of the case studies are characterized through a number of attributes, which, along with metadata, can be used to select geological analogues to subsurface successions (Colombera et al., 2012, 2013). In FAKTS, the architecture of channel and overbank deposits is categorized into three hierarchies of sedimentary units: higher-order depositional elements, lower-order architectural elements, and the elementary facies units constituting them. Depositional elements are large-scale sedimentary bodies classified as 'channel body' or 'floodplain' elements on the basis of the interpreted origin of their deposits; a channel body is a discrete channelized unit made of channel deposits with no particular genetic or palaeo-geomorphic significance: it may correspond to a preserved channel belt, to an isolated channel fill, to an alluvial valley fill, or to a compound amalgamated multi-storey body. Architectural elements, which may be nested in depositional elements, represent sedimentary bodies with characteristic facies associations and architectural properties interpretable in terms of types of sub-environment of deposition, and classified into types related to channel or overbank sedimentation (e.g., downstream-accreting barforms, crevasse-splay elements; cf. Miall, 1996, Colombera et al., 2013). In FAKTS, facies units are basic lithological units recognized at sub-bed scale resolution, delimited by surfaces marking a change in facies or facies association, a major palaeocurrent change, or an erosional contact. Facies units are classified according to their sediment texture and sedimentary structures; the facies classification scheme adopted in FAKTS is an extension of the scheme by Miall (1996; Table 1). When the sedimentary structure of each facies unit is unknown (15% of all facies units by cumulative thickness), the lithologies are generally classified into generic gravelly (G-), sandy (S-), or fine-grained (F-) categories. All sedimentary facies classes employed by the original authors of the considered literature studies are translated into this facies scheme to ensure consistency. The three-dimensional spatial relationships of depositional elements, architectural elements and facies units are captured in FAKTS in the form of transitions between units (Colombera et al., 2012); for the scope of this work, this enables the users to select facies occurring within individual sandbodies and to quantify vertical trends in facies properties.

The sandbodies considered in this study comprise of sedimentary units consisting of isolated or vertically and/or laterally amalgamated depositional or architectural elements recognized in outcrop or in the shallow subsurface of modern river systems; for example, a single channel sandbody may be formed by vertically stacked channel-fill and barform architectural elements that may not be captured in the database as part of a classified depositional element. Out of all the 1D and 2D datasets of channel and overbank sedimentary units derived from FAKTS literature and field studies ($N = 1267$; Supplementary Material 1), this

A Channel sandbody facies sequences



B Overbank sandbody facies sequences

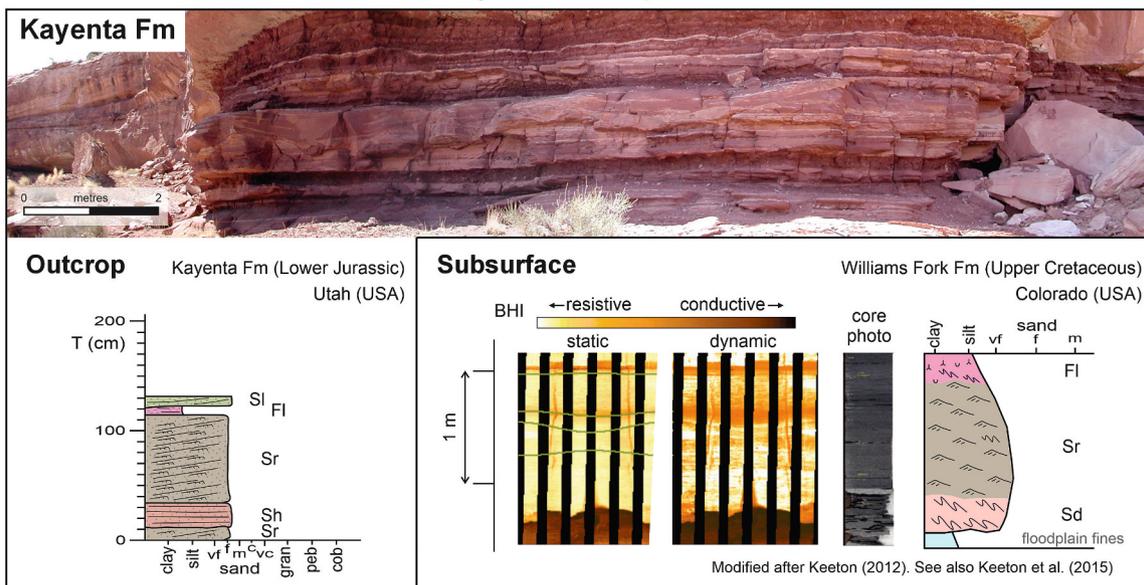


Fig. 1. Examples of channel and overbank sandbodies interpreted in outcrop and borehole data (image logs with core control), exemplifying the types of data that would be used for training the models and to which the models would be applied, respectively. (A) Comparison of the facies sequences of channel sandbodies observed in outcrop (Neslen Fm.; Colombera et al., 2016) and in subsurface data (Williams Fork Fm.; Keeton, 2012; Keeton et al., 2015). (B) Comparison of the facies sequences of overbank sandbodies observed in outcrop (Kayenta Fm.; Colombera et al., 2012) and in subsurface data (Williams Fork Fm.; Keeton, 2012; Keeton et al., 2015).

Table 1

Classes of lithofacies adopted in this study, expanding those by Miall (1996); Colombera et al. (2013).

Code	Characteristics
G-	Gravel-dominated deposit; structure and additional textural characteristics not specified.
Gmm	Matrix-supported, massive or crudely bedded gravel.
Gmg	Matrix-supported, graded gravel.
Gcm	Clast-supported, massive gravel.
Gci	Clast-supported, inversely graded gravel.
Gh	Clast-supported, horizontally or crudely bedded gravel; possibly imbricated.
Gt	Trough cross-stratified gravel.
Gp	Planar cross-stratified gravel.
S-	Sand-dominated deposit; structure and additional textural characteristics not specified.
St	Trough cross-stratified sand.
Sp	Planar cross-stratified sand.
Sr	Current ripple cross-laminated sand.
Sh	Horizontally bedded sand.
Sl	Low-angle (<15°) cross-bedded sand.
Ss	Faintly laminated, cross-bedded, massive or graded sandy fill of a shallow scour.
Sm	Massive sand; possibly locally graded or faintly laminated.
Sw	Symmetrical ripple cross-laminated sand.
Sd	Soft-sediment deformed sand.
F-	Deposit dominated by fine-grained sediment (silt, clay); structure and additional textural characteristics not specified.
Fl	Interlaminated very-fine sand, silt and clay; might include thin isolated cross-laminated sandy lenses.
Fsm	Laminated to massive silt and clay.
Fm	Massive clay.
Fr	Fine-grained root bed.
C	Coal or highly carbonaceous mud.

study only considered those in which the sandbodies are observed as having a fraction of mud-prone lithofacies that is less than 50% ($N = 1222$ instances of 1D sections and 2D outcrop panels). In total, the data extracted from FAKTS are associated with (i) 511 discrete classified fluvial channel sandbodies, corresponding to 441 unique depositional elements and 84 unique architectural elements; (ii) 351 overbank sandbodies described as architectural elements. Given that some elements are intersected by multiple sedimentary logs, a total of 1222 datasets consisting of 1115 unique vertical logged sections and 107 architectural panels are obtained for the sandbodies (Supplementary Material 2); these correspond to 748 sedimentological datasets of channel sandbodies and 474 sedimentological datasets of overbank sandbodies. The facies units contained in these sandbodies return a total measured thickness of 3572.2 m. A total of 352 channel sandbodies could be classified on the observed or inferred channel pattern of their formative river system according to these classes: (i) meandering, (ii) braided, (iii) single-thread low sinuosity; this dataset is associated with 1835.5 m of cumulative facies-unit thickness from 163 distinct sandbodies (Supplementary Material 3).

2.2. Software tools

This sedimentological dataset was used in a machine-learning workflow to attempt binary sandbody classifications (channel vs overbank origin; meandering vs low-sinuosity or braided river planforms) based on facies properties that could be observed in 1D subsurface data, such as image logs. Four alternative supervised machine-learning algorithms for classification based on non-parametric ensemble learning methods were trialled in this study: bagged classification tree, random forest classifier, adaptive boosting classifier, and extreme gradient boosting classifier. Ensemble machine-learning models combine many decision trees to make more accurate and reliable predictions. Compared to simple decision trees, ensemble learners tend to reduce overfitting, to better generalize to unseen data and to capture complex patterns by combining multiple models. A summary of these algorithms is presented in Table 2. Model training and application were performed

Table 2

Machine-learning models used in this study and corresponding methods in the *caret* package (Kuhn, 2008).

Algorithm	Description	<i>caret</i> method
Bagged classification tree (<i>treebag</i>)	Classification algorithm based on the development of an ensemble of decision trees through bootstrap and aggregation (bagging). Multiple decision trees are created by using alternative training datasets created by bootstrap resampling of the original data. The final classification is the modal outcome of all decision trees. See Bühlmann (2012) and references therein.	<i>treebag</i>
Random forest (<i>rf</i>)	Classification algorithm that is also based on an ensemble of decision trees through bagging. Differently from bagged trees, subsets of features selected randomly are used for bootstrap resampling, resulting in decorrelated decision trees. See Parmar et al. (2019) and references therein.	<i>rf</i>
Adaptive boosting (<i>adaboost</i>)	Classification algorithm that is based on an ensemble of base learners, such as decision trees, obtained by developing new learners based on the results of previous learners (boosting). Training data are assigned weights reflecting classification errors. Base learners are created iteratively, and predictions are made by majority vote of the base learners' predictions weighted by their accuracy. See Ferreira and Figueiredo (2012) and references therein.	<i>AdaBoost</i> , <i>MI</i>
Extreme gradient boosting with DART (<i>xgb</i>)	Classification algorithm that is also based on an ensemble of learners (in this case, decision trees) through boosting. Differently from adaptive boosting, this method is based on a gradient descent algorithm to minimize a loss function accounting for the sum of the errors of each tree in the ensemble. The DART algorithm (Dropouts meet Multiple Additive Regression Trees; Rashmi and Gilad-Bachrach, 2015) is employed to drop trees added earlier during boosting iterations to balance their contributions. See Chen and Guestrin (2016) and references therein.	<i>xgbDART</i>

in R 4.2.3 (R Core Team, 2023) in the *tidyverse* framework (Wickham et al., 2019), using the *caret* package (Kuhn, 2008). The *groupdata2* package (Olsen, 2024) has been used for data partitioning. Shapley Additive Explanation (SHAP) values have been derived for selected models to quantify the contribution of the chosen features to the model predictions (Lundberg and Lee, 2017). Model SHAP values have been computed following the Kernel SHAP method using the *kernelshap* library (Mayer and Watson, 2025).

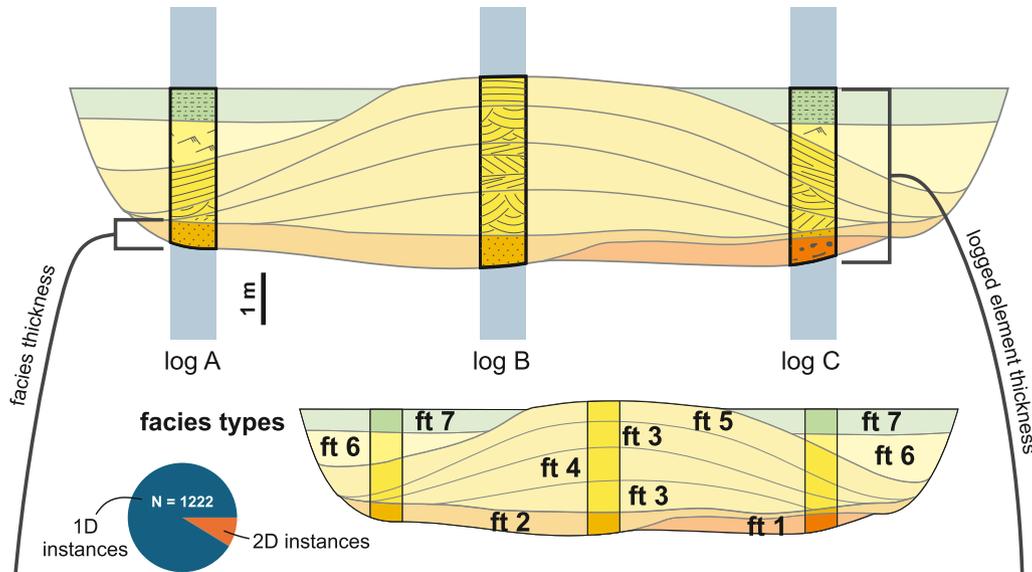
3. Model training

The models were trained to classify sandbodies based on their thickness and internal facies organization recognized in 1D sections.

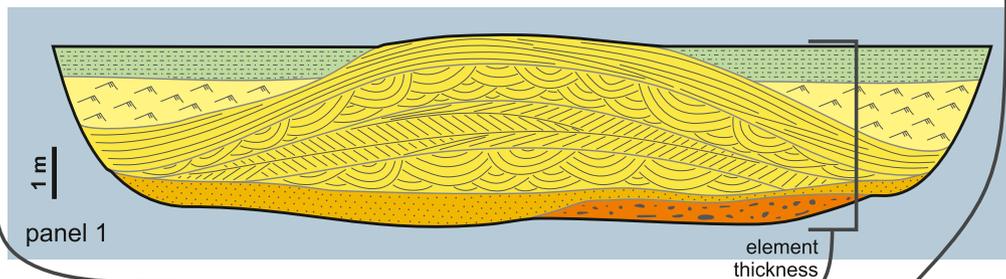
3.1. Training-dataset definition

Sedimentological datasets that could be used for model training and testing were assembled by considering each vertical sedimentary log or 2D outcrop panel of a sandbody as an instance, codified as a row in a table of sandbody data (Fig. 2). Channel sandbodies consist of isolated depositional or architectural elements, or of vertically stacked elements through which there are uninterrupted vertical facies sequences of

A 1D instances - vertical sedimentological sections

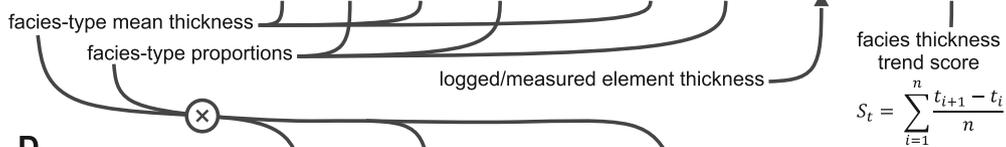


B 2D instances - architectural panels



C Dummy training dataset

inst_id	inst_type	ft1_t̄	ft1_p	ft2_t̄	ft2_p	[...]	ft7_t̄	ft7_p	element_t	t_trend_score
log A	1D	(0)	0.00	0.49	0.16	[...]	0.63	0.20	3.10	0.088
log B	1D	(0)	0.00	0.61	0.17	[...]	(0)	0.00	3.68	-0.028
log C	1D	0.50	0.15	0.14	0.04	[...]	0.65	0.19	3.40	0.035
panel 1	2D	0.50	0.04	0.63	0.11	[...]	0.68	0.12	3.69	0.046



D

inst_id	inst_type	ft1_score	ft2_score	[...]	ft7_score	element_t	t_trend_score
log A	1D	(0)	0.078	[...]	0.126	3.10	0.088
log B	1D	(0)	0.104	[...]	(0)	3.68	-0.028
log C	1D	0.075	0.016	[...]	0.124	3.40	0.035
panel 1	2D	0.020	0.069	[...]	0.082	3.69	0.046

Dummy training dataset – reduced dimensionality

(caption on next page)

Fig. 2. Idealized diagrams that illustrate the compilation of the machine-learning training dataset used in this study based on one-dimensional (1D) vertical sections (A) and two-dimensional (2D) or pseudo-three-dimensional (pseudo-3D) outcrop panels (B). This idealized example shows how facies data related to the same sedimentary unit may be captured from 1D and 2D datasets (C); in reality, only 1D or 2D data on any specific unit are included in the training dataset to avoid redundancy. The lower part of the figure (D) illustrates the way in which the values of proportion and average thickness of each facies type have been combined into a single variable to obtain an alternative training dataset with reduced dimensionality (26 instead of 50 features). ‘inst_id’ denotes a unique identifier of the specific instance of 1D sedimentary section or 2D outcrop panel. ‘inst_type’ indicates the instance type (1D vs 2D). ‘ft’ indicates a specific facies type. ‘t[−]’ indicates the average facies thickness (e.g., ‘ft1_t[−]’ indicates the average thickness of facies type 1). ‘p’ indicates the facies proportion (e.g., ‘ft2_p’ indicates the fraction of facies type 2 in the instance). ‘element_t’ indicate the apparent thickness of the studied sandbody in the 1D section or the maximum sandbody thickness observed in outcrop panels. ‘t_trend_score’ denotes the facies-scale thickness trend score, as defined in the text.

channel deposits. Overbank sandbodies consist of sand-prone floodplain deposits that are variably classified as levee, crevasse-splay, or overbank sandsheet architectural elements (cf. Colombera et al., 2013). In this study, architectural elements classified as crevasse channel fills are not included in either of the two categories of sandbodies.

Two alternative sedimentological datasets were compiled, in which the class label associated with each instance consists of (i) the sandbody type (‘channel’ vs ‘overbank’ sandbody origin; 1222 instances; Supplementary Material 2), (ii) channel-body type based on planform channel pattern (‘meandering’ vs ‘low-sinuosity or braided’ river; 352 instances; Supplementary Material 3). Hence, the two datasets were developed for tasks of fluvial sandbody and fluvial-channel sandbody classification, respectively.

Each of the two datasets contains a total of 50 features, i.e., attributes on which the machine-learning models are trained and the predictions are made. These features are chosen because they represent properties that can generally be determined using image logs. The considered features (Fig. 2) include: (i) the thickness of the sandbody, as measured along the 1D vertical section or where the sandbody is thickest in the 2D outcrop panel; (ii) the proportion of each of the 24 facies types determined in the 1D or 2D instance; (iii) the average thickness of each of the 24 facies types in the 1D or 2D instance; (iv) a facies thickness trend score that quantifies the degree of linear decrease or increase in thickness across vertically stacked facies in the sandbody. The average thickness is replaced with a value of zero in cases where the facies is not present.

For each sandbody instance, the facies thickness trend score is computed as the arithmetic average of all the values of a quantity defined by:

$$S_t = \sum_{i=1}^n \frac{t_{i+1} - t_i}{n}$$

calculated for values of n comprised between 1 (thickness difference across a facies couplet) and 4 (thickness differences across successive facies units in facies sequences consisting of up to five facies), where i denotes the vertical position of the facies units in ascending order through the stratigraphy. The value of n for which S_t is computed is limited to the total number of facies units in a vertical sequence; values of S_t are then averaged based on the number of defined facies windows. Positive values of the facies thickness trend score indicate thickening-upward trends across facies units, whereas negative values indicate thinning-upward trends. In both cases, the magnitude of the stratal trend is quantified by the absolute value of the score. A value equal to zero is assigned by default to sandbody instances consisting of a single facies unit. This particular approach to the quantification of vertical trends in facies thickness adapts to the variable number of facies units forming a facies succession through a particular sandbody, and enables the inclusion of facies sequences consisting of two facies units only.

Values of the attributes considered for compiling the sedimentological datasets are summarized in Fig. 3, for both the generic sandbody types (channel vs overbank origin) and the channel sandbody types (‘meandering’ vs ‘low-sinuosity or braided’ river planforms). Statistical distributions were inspected, but steps of outlier removal or replacement were not implemented as part of the machine-learning workflow, since statistical outliers are not error-related; rather, they represent unusual

facies properties, which may be indicative of certain sandbody types.

The compiled datasets exhibit systematic correlations between the values of mean thickness and proportion of each of the facies types (see Supplementary Material 4), partly because facies proportions are computed based on cumulative thickness, but primarily because the thickness values differ from zero only where a facies is present, i.e., when its proportion differs from zero. Thus, to reduce the dimensionality of the dataset and address the redundancy in information contained in some variables, an alternative sedimentological dataset has been compiled for the instances of general sandbody types, by combining values of proportions and average thickness of each facies type into a new variable defined as their product (or zero if that facies type is absent in the section). This alternative dataset consists of 26 effectively uncorrelated variables.

In both sedimentological datasets, the binary labels included in the dataset are slightly imbalanced (748 channel vs 474 overbank sandbodies; 157 braided vs 195 low-sinuosity channel bodies). As part of the developed workflow, a SMOTE (Synthetic Minority Oversampling Technique; Chawla et al., 2002) was tentatively applied to rebalance the proportions of the minority classes; however, doing so resulted in an increase in model specificity with decrease in sensitivity, with no significant accuracy gain. The results presented below are based on models trained on the original non-resampled sedimentological datasets.

3.2. Training vs testing data partition

Each of the different sedimentological datasets was partitioned into (i) a training dataset, used by the machine-learning algorithms to train the models, and (ii) a validation dataset, to be used to assess the performance of the trained models. Three approaches to data partition were taken. The first two approaches were designed to assess the general predictive power of the models: each of the two datasets were divided so that approximately 80% of the data could be used for model training and ca. 20% of the data could be used for testing. The partition was operated in two alternative ways: (i) by simply ensuring that the proportion of each class of sedimentary unit is equally represented across the two splits; or (ii) by separating 1D facies sequences so that different vertical sections intersecting the same architectural element fall in the same data split, whilst also ensuring equal representation of both classes of sedimentary units (labels). The ratio between the number of sedimentological datasets and the number of distinct sandbodies, which acts as a measure of the degree to which different vertical sections insist on the same sedimentary unit, is equal to 1.42 for sandbodies overall and 2.16 for classified channel bodies. A partition algorithm was used to ensure that the proportions of categorical variables (labels) present in testing and training datasets were comparable, and to optionally force different 1D vertical sections of the same sedimentary unit to occur in the same split to prevent the same unit to be present in both training and testing data. The third approach to data partition was intended to enable a test of the models against datasets that may be difficult to interpret: a division between training and test data was made by arbitrarily selecting six sedimentary units (sandbodies and channel sandbodies, in each respective dataset) whose vertical facies sequences exhibit unexpected characteristics (profile thickness, dominant facies types, and facies thickness trends) for the given unit type; data on these units were chosen for testing the models, whereas the remainder were all used for model

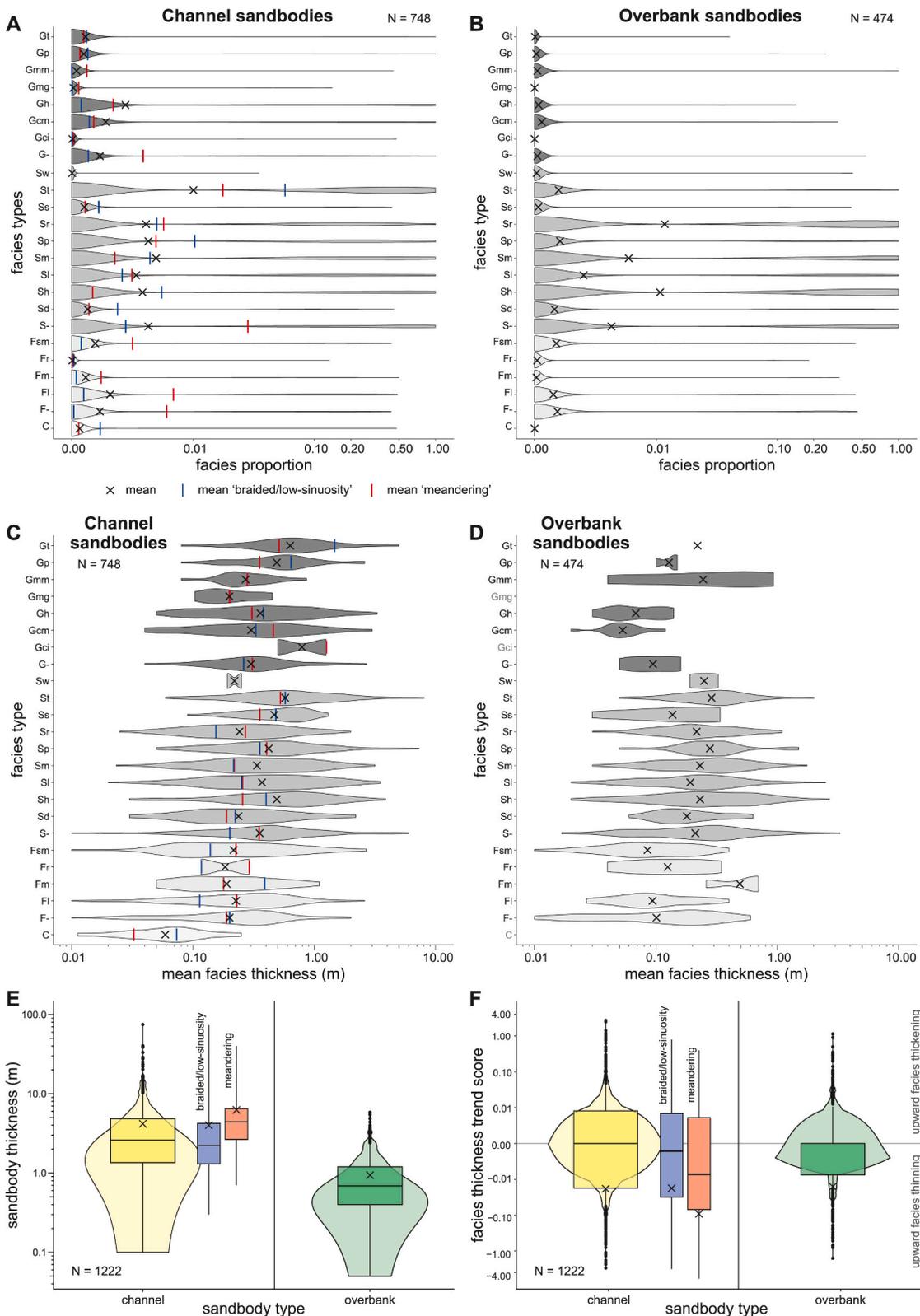


Fig. 3. Summary of the sedimentological datasets used for machine-learning model training and testing. (A-B) Violin plots of the distributions of proportions of facies in sandbody instances, separately presented for channel (A) and overbank (B) deposits; average values for channel bodies categorized on river planforms are also reported; note the pseudo-logarithmic scale. (C-D) Violin plots of the distributions of average facies thicknesses in sandbody instances, separately presented for channel (C) and overbank (D) deposits; average values for channel bodies categorized on river planforms are also reported; note the logarithmic scale. (E) Violin and box plots of the distributions of sandbody thicknesses for channel and overbank sandbodies, and for channel bodies categorized on river planforms; note the logarithmic scale. (F) Violin and box plots of the distributions of facies thickness trend scores for channel and overbank sandbodies, and for channel bodies categorized on river planforms.

training. The six selected examples were picked deliberately as instances for which misinterpretations by humans or models were expected to be likely. For the specific task of general sandbody categorization (channel vs overbank interpretation), the first approach to data partitioning was applied to create two alternative data splits: (i) a split into a training dataset that includes both 1D vertical sections and 2D outcrop panels and a testing dataset that only includes 1D data, since this is the type of data to which the models would be applied in practice; (ii) a split into training and testing datasets that comprise exclusively of 1D data. This

was done to determine the impact of using 2D data for model training on model performance, given that 2D data capture lateral variability and might bias the models towards more general facies arrangements. The training datasets used for channel sandbody classification into 'meandering' vs 'low-sinuosity or braided' types included both 1D and 2D sedimentological data.

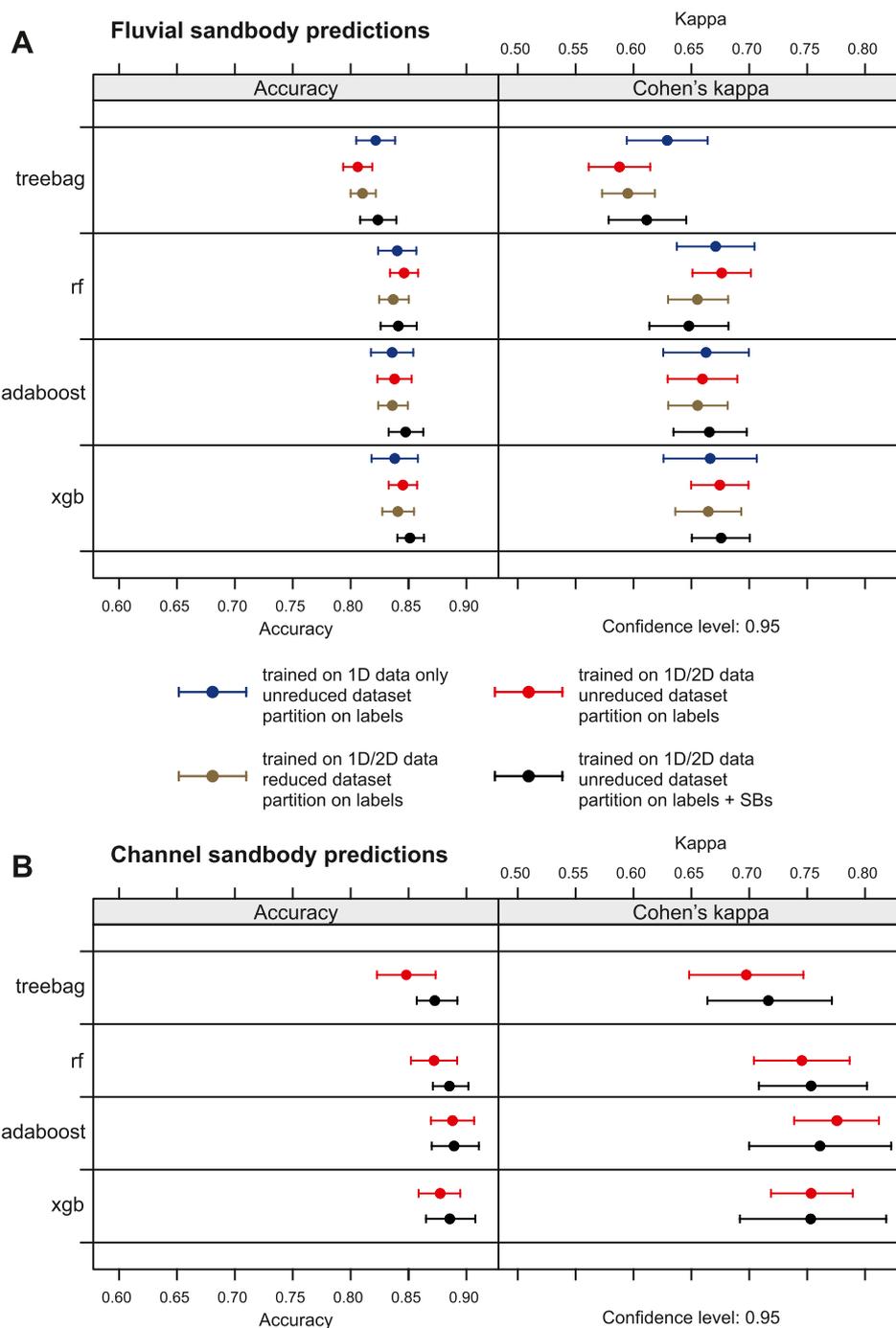


Fig. 4. Dotplots of accuracy and Cohen's kappa values obtained via cross-validation in model training, for models developed for general sandbody (channel vs overbank) interpretations (A), and channel-body (meandering planform vs low-sinuosity or braided planform) interpretations (B). Spots indicate mean values; bars indicate 95% confidence intervals. In A, results for models trained on 1D data only and for models trained on combined 1D and 2D data are presented separately, and results relating to models trained on the dataset with reduced dimensionality are also included. Results are separately shown for models trained using data partitioned according to two approaches: (i) balancing label proportions only (in blue, red, and brown); (ii) keeping different 1D facies sequences of the same element ('SB') in the same split (in black). See Table 2 for model abbreviations.

3.3. Model tuning and cross-validation

Using these different training datasets, the model parameters were tuned by running the algorithms with three repeats of ten-fold cross-validation, with the same seed number. Since the binary labels are only moderately imbalanced, accuracy – i.e., the fraction of correctly classified sandbodies – was considered as a suitable metric for model optimization in the training process. However, sandbody labels are inherently interpretational, and subjectivity is particularly significant for interpretations of formative river planform types; as such, values of accuracy merely quantify a match between the predictions made by a model and the interpretations made by a sedimentologist. Correctness in interpretation remains unknown.

The model training time ranged from few minutes (e.g., for *treemap*, Table 2) to approximately three hours (*xgb*, Table 2), without parallel processing, on a 64-bit version of R running on a Windows machine with i9–12,900 CPU and 64 GB RAM.

The results of model cross-validation on the training datasets are summarized for both binary-classification tasks (sandbody and channel-body classifications) as mean values of model accuracy, Cohen's kappa and associated 95% confidence intervals, as shown in Fig. 4. Accuracy is the proportion of classifications in agreement with the dataset. Cohen's kappa is a metric that quantifies the degree to which the agreement in classification exceeds what would be expected by chance alone, and is more robust than accuracy under class imbalance. These results provide a general measure of model performance (Hothorn et al., 2005; Eugster et al., 2008; Emmert-Streib et al., 2019).

In the task of general sandbody classification (channel vs overbank interpretation), models trained on the unreduced datasets partitioned on the basis of label proportions only return mean cross-validation accuracy ranging between 0.82 (*treemap*) and 0.84 (*rf*) when training is performed on 1D data only, and between 0.81 (*treemap*) and 0.85 (*rf*) when 1D and 2D sedimentological data are both used for training (Fig. 4). Similarly, mean kappa values range between 0.63 (*treemap*) and 0.67 (*rf*) for models trained on 1D data only, and between 0.59 (*treemap*) and 0.68 (*rf*) for models trained on both 1D and 2D (Fig. 4). With the exception of the bagged classification tree (*treemap*), the inclusion of data from 2D outcrop panels in the training dataset does not cause a decrease in the predictive power of the models. For models trained on 1D data only, the area under the receiver operating characteristic (ROC) curve, a metric that takes into account both sensitivity and specificity (i.e., measures of model performance at detecting channel and overbank sandbodies respectively), ranges between 0.89 (*treemap*) and 0.92 (*adaboost* and *xgb*). For models trained using the dataset having a reduced number of features, in which facies proportions and average thickness values are synthesized in a single variable for each facies type, the values of mean accuracy range between 0.81 and 0.84, whereas mean kappa values range between 0.60 and 0.66 (Fig. 4). Therefore, the predictive power of models trained on the reduced dataset is comparable to that of models trained on the full set of facies variables. Cross-validation of the models developed for classification of channel sandbodies (into 'meandering' vs 'low-sinuosity or braided' classes) yields mean accuracy values ranging between 0.85 (*treemap*) and 0.89 (*adaboost*) and mean kappa values between 0.70 (*treemap*) and 0.77 (*adaboost*) (Fig. 4). The models trained on datasets partitioned in a way that balances label proportions whilst keeping sedimentary logs from the same sedimentary unit together exhibit values of mean accuracy that are typically higher than those of models trained on datasets that are only partitioned balancing label proportions (Fig. 4), both for tasks of sandbody classification (0.82–0.85) and channel-body classification (0.87–0.88).

4. Model testing

Overall, the predictive power of the models for general sandbody classification (channel vs overbank interpretation) seen in the cross-validation process is broadly confirmed by the application of the

models to the unseen testing data (Figs. 5 and 6). A key aspect of the testing data is that only 1D sedimentological data were considered, even for testing models that were trained in part using facies data from 2D outcrop panels. The inclusion of 2D data makes the training dataset larger but less representative of 1D facies sequences, yet the performance of models developed on this basis does not differ substantially from that of models trained on 1D data only (Fig. 5). This supports the use of sedimentological data from outcrop panels to increase the volume of data used for model training. Across the four models and the two approaches to data splitting based on dataset type (i.e., with training data including 1D data only vs both 1D and 2D data), the accuracy of the predictions ranges between 0.83 (*treemap* trained on 1D data) and 0.87 (*xgb* trained on 1D data), and Cohen's kappa values range between 0.65 (*treemap* trained on 1D data) and 0.73 (*xgb* trained on 1D data). On average, less than 15% of the sandbodies are misclassified, and 74% of the sandbodies are labelled correctly by all four models: this suggests the viability of using ensemble machine-learning models for supporting automated interpretations of fluvial sandbody types. Lower model performance is seen in application to testing data that only include vertical 1D sections from sandbodies that are not simultaneously represented in the training dataset, which is expected in view of the higher statistical independence of training and testing data. In this case, model accuracy ranges between 0.76 (*treemap*) and 0.80 (*xgb*), whereas values of kappa range between 0.52 (*treemap*) and 0.60 (*xgb*). Across this set of models, ca. 28% of channel sandbodies are misclassified as overbank sandbodies; misclassification of overbank sandbodies is less common, with 13% to 21% of mislabelled examples (Fig. 6).

With regards to the misclassified sandbodies, it is likely that the predictive value of the models is affected by the fact that channel and overbank sandbodies exhibit significant overlaps in the considered sandbody properties: sandbody thickness, facies proportions, average facies thicknesses, and vertical facies-unit thickness trend (Fig. 3). However, it is likely that the models tend to fail especially in cases when these sandbody properties are unusual for the particular sandbody type. For instance, out of the 222 instances extracted for testing from the dataset including 1D data only, 13 have been misclassified by all the models: these consist of unusually thin channel sandbodies (1.1 m thick on average) and thick overbank sandbodies (2.6 m). By comparison, the 164 sandbodies that were labelled correctly by all four models exhibit average thickness value of 5.1 m and 0.7 m for channel and overbank units, respectively. This is consistent with the importance of sandbody thickness as the dominant feature in all models for general sandbody interpretation, as indicated by the results of the SHAP analysis (Fig. 7).

The second approach to model training and testing, whereby a data split is operated manually by extracting six chosen instances as test data, enables a test of the models against sandbodies whose interpretation is considered challenging. This approach was also taken to perform a comparison of the model outcomes with interpretations made by eight experienced geologists trained at doctoral level in clastic sedimentology and with special expertise in the facies analysis of fluvial successions, who were asked to attempt a classification of all six sandbodies only on the basis of their facies sequences. The participants were only provided with the graphical sedimentary logs and relative facies legend, as shown in Fig. 8, with no other sedimentological data or ancillary information on the considered successions. For each selected test sandbody, the predictions made by the four models can be compared with the interpretations made by the eight sedimentologists (Fig. 8). Because of the particular nature of the chosen examples, the outcomes of this exercise should not be considered as an assessment of the general performance of the models compared to the geologists; it is merely an evaluation of whether the models tend to fail for sandbodies exhibiting atypical facies successions, and whether such behaviour would be comparable to the way humans misclassify vertical sections that are at odd with common facies models. The developed machine-learning models all misclassified (i) two cases of channel bodies characterized by vertical profiles that are remarkably thin (less than 80 cm) and that consist of a single bed of

model evaluation against test dataset: fluvial sandbody types

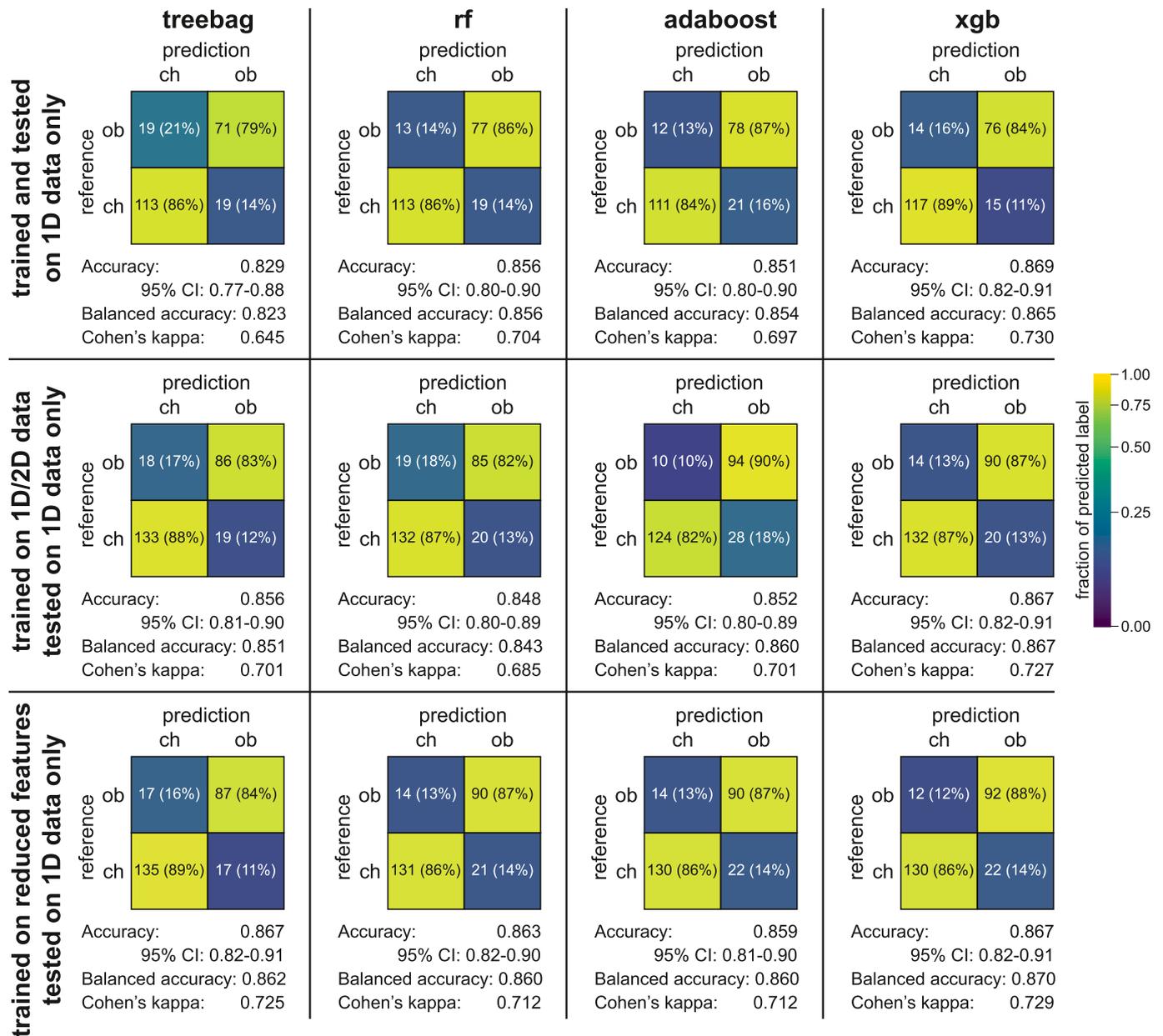


Fig. 5. Confusion matrices summarizing the performance of the models for general sandbody interpretation in application to the test data. Label 'ch' indicates channel sandbodies; 'ob' indicates overbank sandbodies. Results for models trained on 1D data only and for models trained on combined 1D and 2D data are presented separately; results relating to models trained on the dataset with reduced dimensionality obtained by combining facies proportions and average thicknesses are also shown. See Table 2 for model algorithms. Balanced accuracy values account for the slightly imbalanced nature of the dataset (152 channel sandbodies vs 104 overbank sandbodies), whereas Cohen's kappa accounts for agreement by chance.

massive sandstone (Fig. 8A), as well as (ii) one case of crevasse-splay architectural element dominated by trough cross-stratified sandstone with local soft-sediment deformation (Fig. 8B). A specific model (*adaboost*) even failed to classify correctly any of the six examples. These results confirm the idea that model failure is associated with uncommon styles of facies sequences, for which we expect human interpretations to be equally error prone. Any quantification of model performance has limited value in this exercise, given the very small sample size; however, it is notable that the interpretations made by the sedimentologists return: (i) an average accuracy value (0.54) that is comparable with random guessing (i.e., 0.5 accuracy); (ii) values of Fleiss's Kappa and Krippendorff's Alpha – employed as measures of consistency in classification – equal to 0.21 and 0.23, respectively. For the six selected

examples, the considered facies data are not suitable for making reliable inferences on sandbody type, and the degree of agreement across the geologists is limited. Yet, the models (average accuracy of 0.21) perform considerably worse than random picking when applied to these unusual facies sequences.

The predictive power of the models for classification of channel sandbodies according to the channel pattern of the formative rivers is also confirmed by applying the models to a testing dataset (Fig. 9) that, unlike the training dataset, only includes instances of 1D facies sequences. For models trained and tested on data partitioned on the basis of label proportions only, accuracy values range between 0.88 (*rf* and *xgb*) and 0.91 (*adaboost*), whereas Cohen's kappa ranges between 0.75 (*rf* and *xgb*) and 0.81 (*adaboost*) (Fig. 9). On average, only 11% of the

model evaluation against test dataset: fluvial sandbody types

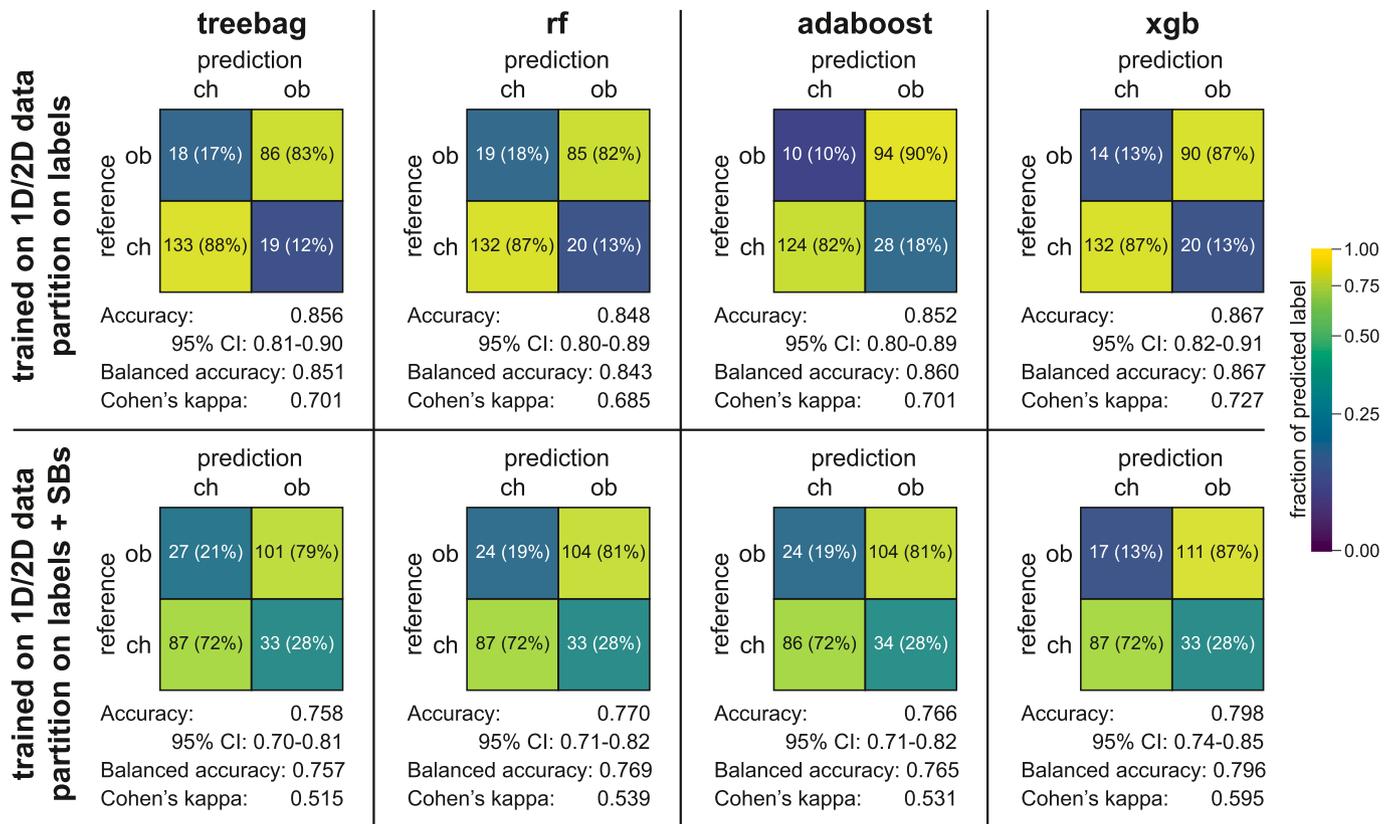


Fig. 6. Confusion matrices summarizing the test performance of models for general sandbody interpretation trained and applied using sets of 1D and 2D data obtained via two alternative partition approaches: (i) by splitting the dataset to maintain label proportions only; (ii) by operating a separation that keeps 1D facies sequences from the same sandbody ('SB') in the same split. Label 'ch' indicates channel sandbodies; 'ob' indicates overbank sandbodies. See Table 2 for model algorithms. Balanced accuracy values account for the slightly imbalanced nature of the dataset, whereas Cohen's kappa accounts for agreement by chance.

channel bodies are misclassified across the applications of the four models. For models trained and tested using data partitioned on the basis of both label proportions and architectural-element identifiers, such that sedimentary logs of the same element are not present in both training and testing data, accuracy values range between 0.75 (*treebag*) and 0.81 (*adaboost* and *xgb*), whereas Cohen's kappa ranges between 0.49 (*treebag*) and 0.62 (*adaboost* and *xgb*) (Fig. 9). This finding suggests that, by training ensemble machine-learning models on geological analogues for which river channel patterns are observed (modern rivers) or interpreted with some relative confidence (in outcrop), it becomes feasible to attempt automatic interpretations of the planforms of the formative rivers of channel deposits based on 1D sedimentological data. This task is commonly considered futile and subject to significant uncertainty even when attempted by expert sedimentologists (e.g., Allen, 1983; Friend, 1983; Bridge, 1985; Miall, 1992; Bristow, 1996; Fielding et al., 2025). The SHAP feature importance analysis of models trained on data partitioned by balancing label proportions only reveals that the relative contribution of the considered attributes is more variable compared to models developed for general sandbody classification, in terms of both range of features that dominate the decisions and magnitude of their influence on model outcomes (Fig. 10).

Additionally, the models were developed and tested by using the manual data split according to which six instances of channel bodies were selected as test data, to assess the proposed method against channel bodies whose interpretation is considered difficult. These examples include, for instance, deposits of meandering river channels whose sand-prone facies are exclusively made of planar or trough cross-stratified sands, and interpreted braid-bar elements displaying drapes of fine-grained deposits and dominance by ripple cross-laminated sandstones

(Fig. 11); these are cases in which the channel deposits of a meandering river exhibit characteristics that are typically represented in 1D facies models for braided rivers, and vice versa. The model predictions were again compared with interpretations of the facies sequences by the same eight sedimentologists who completed the earlier task (Fig. 11). The machine-learning models all misclassified one particular case of channel facies sequence: a vertical section through point-bar deposits of the Wabash River at the Helm bend (Fig. 11A), which was specifically chosen because it had originally been recognized as displaying characters that are not typical of the channel deposits of meandering rivers (Jackson, 1976). The corresponding log was also misclassified by seven out of the eight sedimentologists, whose classifications have been operated inconsistently across the set of six examples (returning values of Fleiss's Kappa and Krippendorff's Alpha equal to 0.50 and 0.24, respectively). Overall, for this task, the models have marginally outperformed the human geologists, with a performance that is merely slightly better than random labelling (average accuracy values of 0.58 vs 0.40; Fig. 11C).

5. Discussion

Many machine-learning methods have recently been proposed for completing routine tasks of subsurface studies, including those performed by sedimentologists in the form of observations and interpretations (e.g., Di Martino et al., 2023; Simon and Worden, 2025). Several tools exist for the classification of sedimentary deposits based on 1D borehole data (e.g., Bestagini et al., 2017; Hall and Hall, 2017; Martin et al., 2021). Some machine-learning approaches (cf. Valentín et al., 2019; Zheng et al., 2021; Tian et al., 2023; Gharieb et al., 2024;

SHAP analysis — prediction of fluvial sandbody types

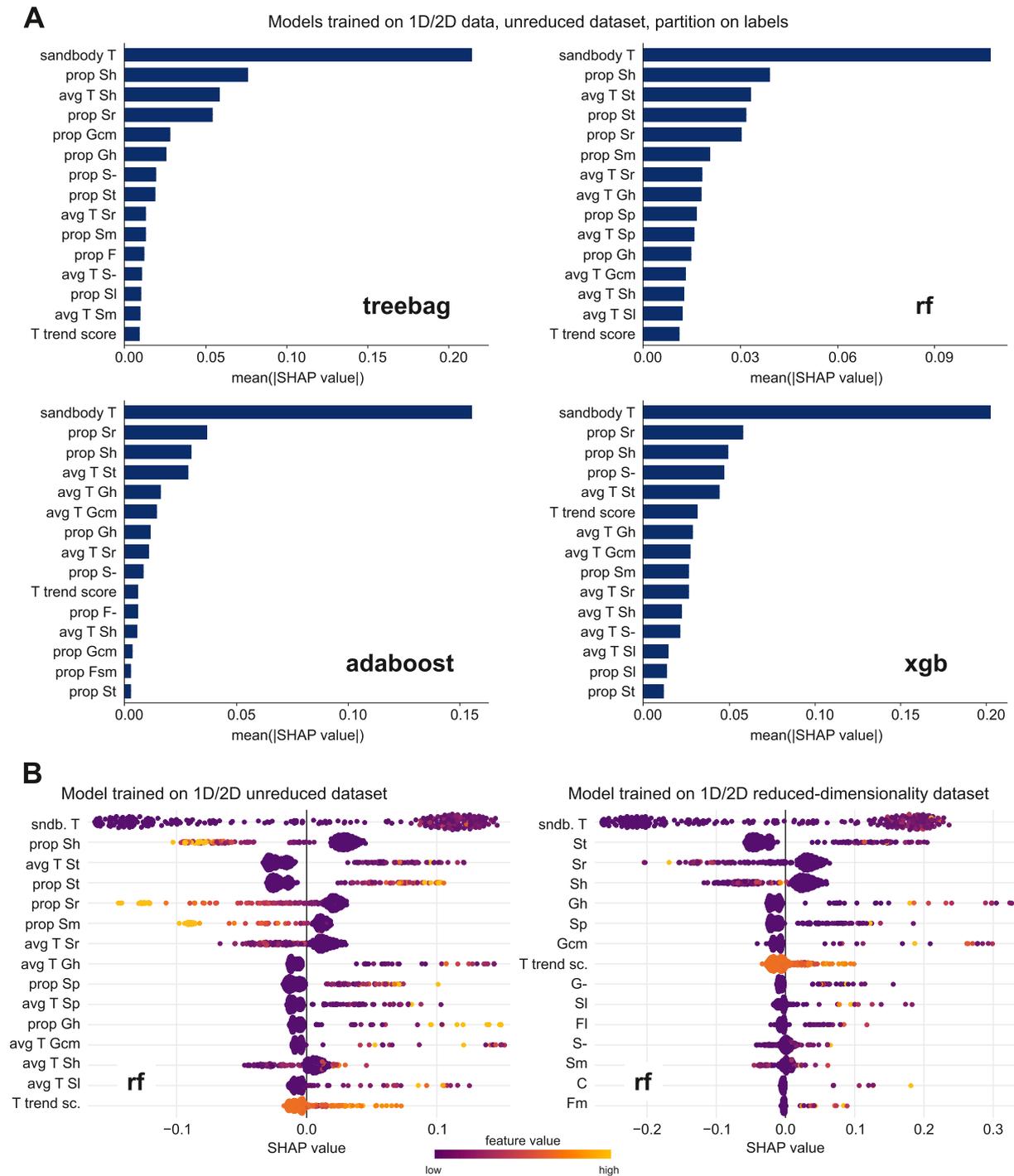
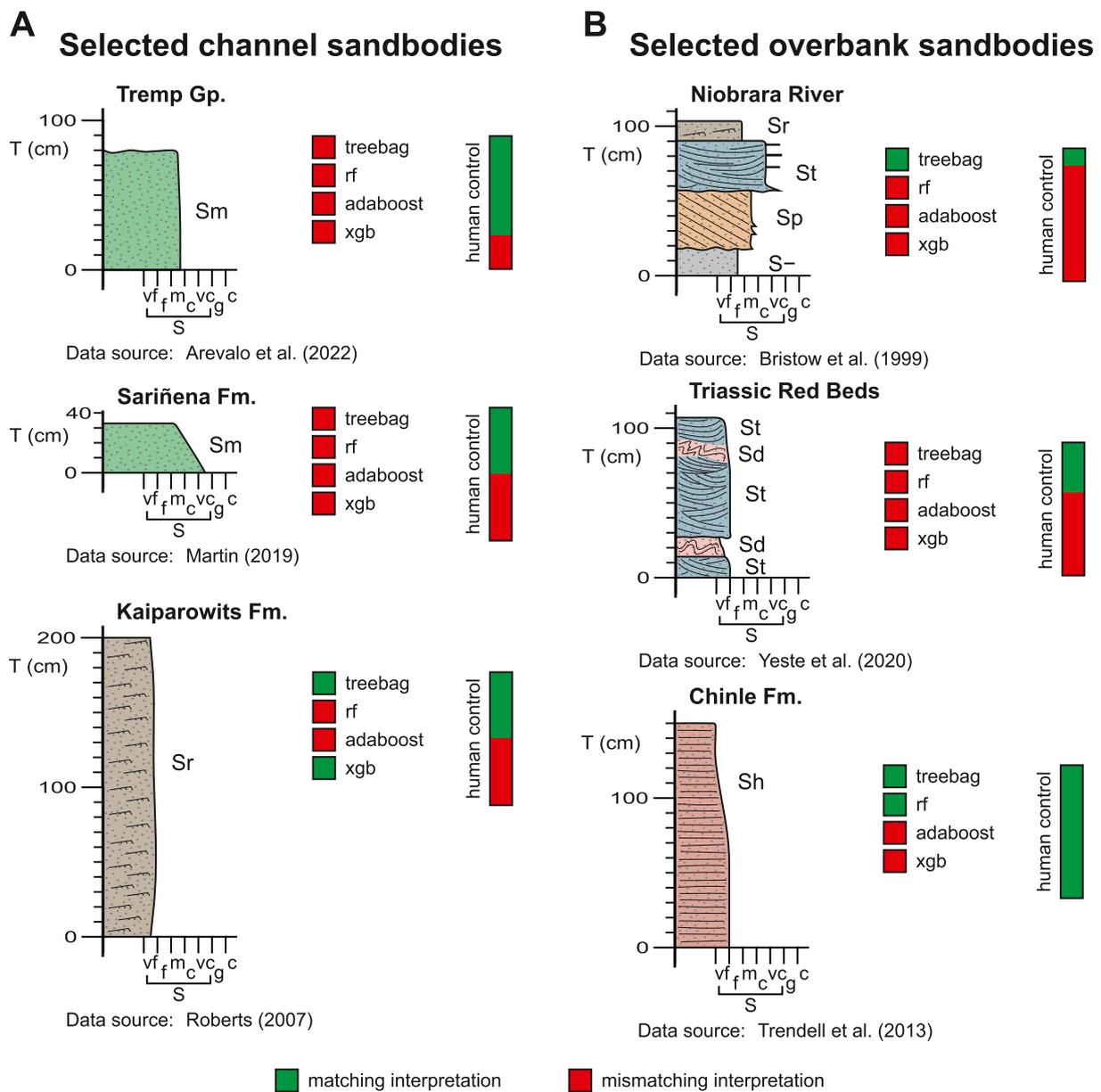


Fig. 7. Results of SHAP analyses for models predicting sandbody types, expressed as SHAP values representing the contribution of each feature to a classifier's prediction compared to a baseline prediction. (A) bar plots ranking the 15 most influential features based on absolute SHAP value for each of the four models trained using 1D and 2D unreduced data partitioned to maintain label proportions only. (B) beeswarm plots showing the direction and magnitude of the influence of each of the top 15 features on predictions by random forest models trained on 1D and 2D data; results for models trained on unreduced and reduced dimensionality datasets are separately presented. 'prop' indicates proportions; 'avg T' indicates average thickness values; 'sndb. T' indicates sandbody thickness; 'T trend sc.' indicates the thickness trend score.

Hou et al., 2024) can in principle be applied for classifying sedimentary facies types using image logs that demonstrate textural characteristics and sedimentary structures at mm- to cm-scale vertical resolutions. In parallel, some methods have been proposed for the automatic picking of surfaces that bound stratigraphic packages of certain types using exclusively 1D borehole data (e.g., flooding surfaces that bound

parasequences; cf. Kusumah et al., 2022; Colombera and Budai, 2026). These recent advancements provide the opportunity to develop workflows that can ultimately enable partially automatic recognition of facies associations in well data, which could incorporate the approach proposed in this study. Differently from approaches for the classification of facies associations through the direct application of machine-learning



C Model performance summary

test against manually selected examples

model	x-val. accuracy	test accuracy	test kappa
treebag	0.828	0.500	0.000
rf	0.860	0.167	-0.667
adaboost	0.854	0.000	-1.000
xgb	0.858	0.167	-0.667

1D instances only – train N = 1216, test N = 6

Accuracy of human control interpretations (N = 8): 0.333–0.833 (avg: 0.542)

Fig. 8. Vertical profiles of the sandbodies chosen for the sandbody classification exercise proposed to the models, trained as explained in the text, and to eight volunteering sedimentologists, shown separately for channel (A) and overbank (B) deposits. Cases of correct and incorrect classifications by the four types of models and the geologists ('human control') are reported. (C) summary of the performance of humans and models in the sandbody classification challenge. See Table 2 for model abbreviations.

models to well logs (e.g., Tognoli et al., 2024), such workflows would reflect sedimentological thinking, in the sense that they would build on the recognition of sedimentary facies sequences and physical stratigraphic divisions. However, whereas machine-learning algorithms for facies classification typically handle objective labels (lithological types) for purposes of both model training and testing, the recognition of facies

associations is inherently interpretational. It is therefore important that, for training the machine-learning models, the approach proposed in this study relies on the use of geological analogues in which the depositional context is either known, for present-day rivers, or interpretable reliably, thanks to outcrop observations. In this perspective, it is convenient that the inclusion of data from 2D outcrop panels for the purpose of training

model evaluation against test dataset: channel-body types

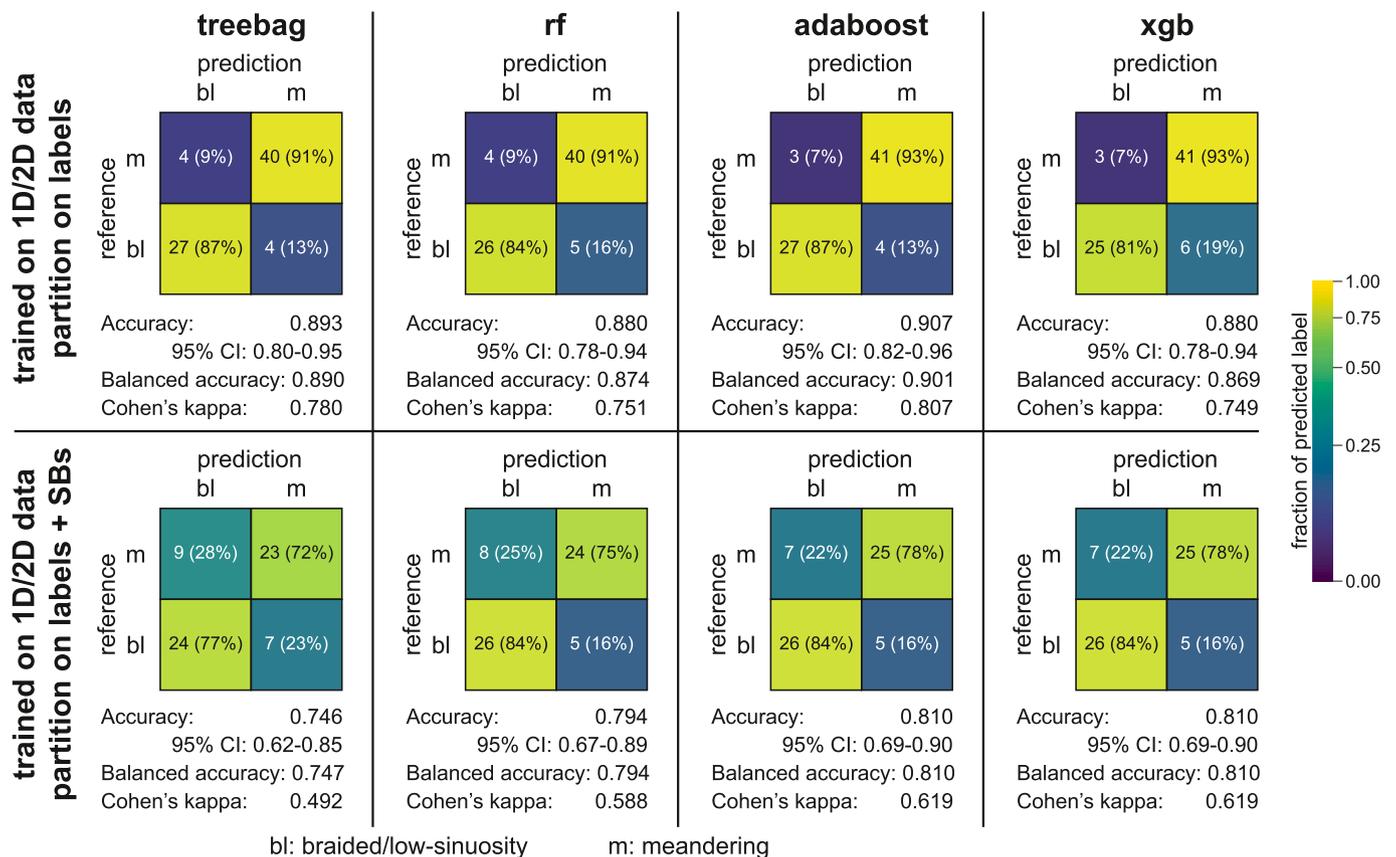


Fig. 9. Confusion matrices summarizing the performance of the models for channel-body river-planform interpretation in application to the test data. Results are shown separately for models trained and applied using alternative sets of 1D and 2D data obtained via two different partition approaches: (i) by splitting the dataset to maintain label proportions only; (ii) by operating a separation that keeps 1D facies sequences from the same architectural element ('SB') in the same split. Label 'm' indicates channel bodies of meandering river systems; 'bl' indicates channel sandbodies of braided or single-thread low-sinuosity river systems. See [Table 2](#) for model algorithms. Balanced accuracy values account for the slightly imbalanced nature of the dataset (44 'meandering' channel sandbodies vs 31 'braided or low-sinuosity' channel sandbodies), whereas Cohen's kappa accounts for agreement by chance.

models that are meant to operate on 1D data did not cause a significant loss in predictive power.

The principal finding of this study lies in the feasibility of an approach for automating subsurface interpretations of fluvial sandbodies that mimics the application of 1D facies models by geologists. The ability to distinguish between channel and overbank sandbodies using facies attributes that may be extracted from full-bore formation microimaging (FMI) data can facilitate the processing of large subsurface datasets, for which the distinction between these two sandbody types is important in consideration of their contrasting architectures and heterogeneities. In subsurface workflows, labelling these types of sandbodies in boreholes is a key step both for constraining the inputs of stochastic reservoir models – especially for assigning realistic input proportions of reservoir-unit types – as well as for conditioning the same models using well observations ([Deutsch and Tran, 2002](#); [Keogh et al., 2007](#); [Pyrz et al., 2009](#); [Tye, 2013](#)). The performance of the models in the classification of the six selected examples of 'challenging' fluvial sandbodies was worse than guesswork, but the error rate of the human sedimentologists involved in the benchmark comparison was also significant, even though humans could inspect graphic logs that present more data (detailed grain size and vertical facies trends) than what is fed to the models. As such, the use of machine-learning tools based on the proposed approach in industry workflows appears to be viable, even if only as a way for obtaining alternative interpretations to be used as measures of uncertainty in subsurface reconstructions.

Some unexpected results were obtained in the attempt at training

and using the models for the classification of channel bodies according to the planform style of their formative rivers, since this task has hitherto been considered by many experts to be nearly hopeless. It is widely acknowledged that classic 1D facies models categorized on channel pattern (e.g. [Visher, 1965](#); [Jackson, 1978](#); [Miall, 1978](#); [Cant, 1982](#)) are flawed, since facies sequences are not by themselves diagnostic of the planform style of the channels that led to their accumulation ([Allen, 1983](#); [Friend, 1983](#); [Bridge, 1985](#); [Miall, 1992](#); [Bristow, 1996](#); [Fielding et al., 2025](#)). It was therefore surprising that, when applied to their respective test datasets, the models for predicting channel planform demonstrated accuracy values that are overall comparable to those for general sandbody interpretation, for pairs of models trained using the same data partition approach ([Figs. 5 and 7](#)). This is remarkable given that the two channel-body types do not appear to exhibit more significant differences in their facies-sequence characteristics, compared to the two general fluvial sandbody classes. Facies-scale thinning-upward trends are more marked across channel body types than for fluvial sandbody types, but differences in facies properties and sandbody thickness are generally more modest ([Fig. 3](#)). Accordingly, the importance of model features on sandbody classification is more variable across models trained to make channel-body interpretations ([Figs. 7 and 10](#)). The proposed approach for the recognition of channel patterns in 1D borehole data is based on statistical discrimination, rather than mechanistic process interpretation. Its application is likely to produce highly uncertain results, but outcomes of this study suggest that attempting these interpretations using machine-learning tools may be

SHAP analysis — prediction of channel sandbody types

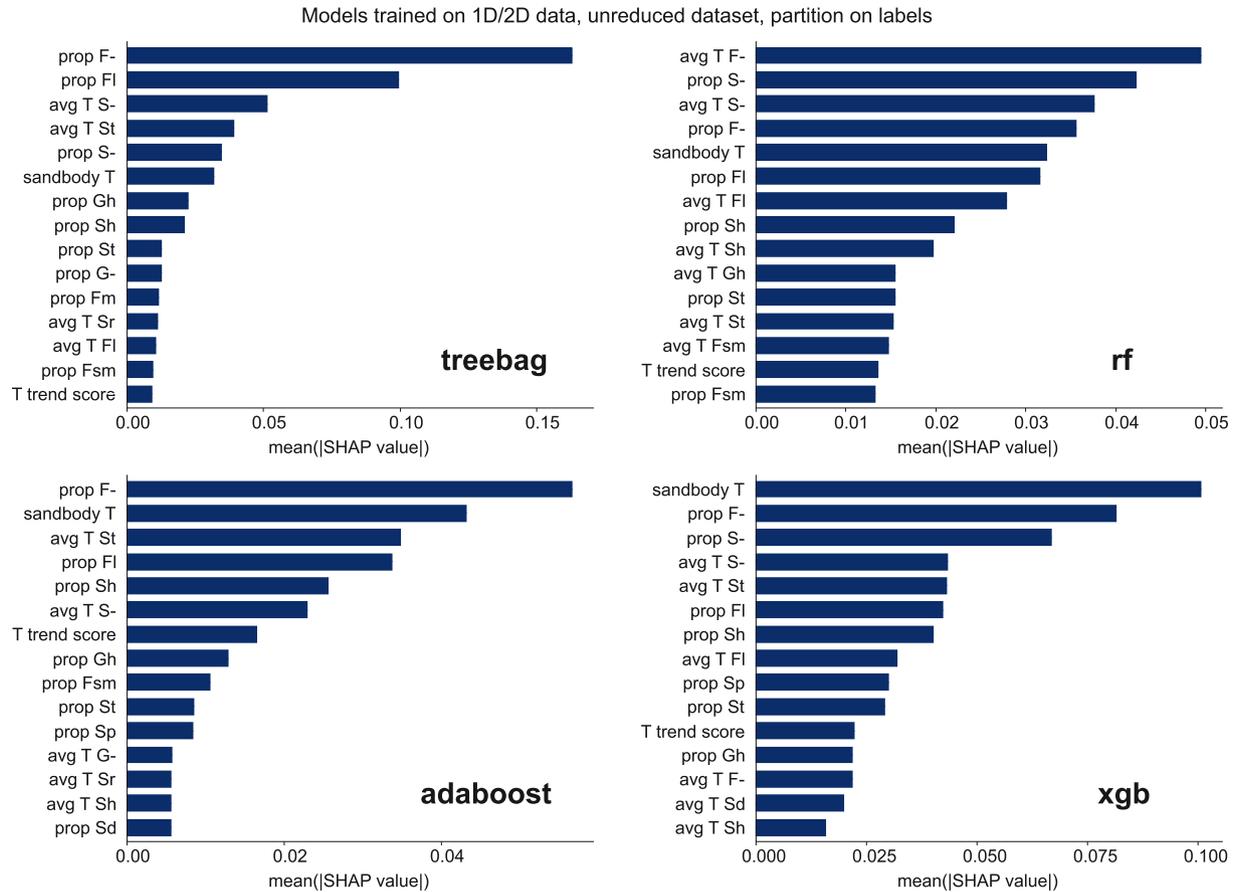


Fig. 10. Results of the SHAP analysis for models predicting channel-body types classified on formative river planforms, shown as bar plots ranking the 15 most influential features based on absolute SHAP value for each of the four models trained using 1D and 2D unreduced data partitioned to maintain label proportions only. 'prop' indicates proportions; 'avg T' indicates average thickness values. SHAP values quantify the contribution of each feature to a classifier's prediction compared to a baseline prediction.

valuable, especially given the links between formative river planforms, channel sandbody architecture and sedimentary heterogeneity (cf. Lunt et al., 2004; Donselaar and Overeem, 2008; Colombera et al., 2017; Willis and Sech, 2019).

5.1. Limitations and applicability

The proposed approach to the training and application of machine-learning models for sandbody classification can be refined and expanded in several ways. For example, the collated sedimentological dataset ignored vertical trends in the occurrence of the facies types, and therefore fails to capture systematic vertical changes in grain size and sedimentary structures that can be interpreted in a process-oriented perspective. This was done to limit the number of features used for model development; however, such trends could be readily incorporated in the sedimentological dataset in the form of vertical transition statistics (cf. Colombera et al., 2013). The adopted facies codes are widely applied in literature studies of fluvial successions, which makes it easy to bring together many field datasets in a consistent way, but this comes at the expense of sedimentological details that may be captured to describe facies characteristics of the sandbodies that can be linked to depositional processes (e.g., presence of systematic grain-size patterns in unit-bar cross-stratified sands; cf. Reesink and Bridge, 2011; types of low-angle cross-stratification indicative of upper flow regime; cf. Fielding, 2006). The use of a standard facies scheme is an inevitable source of bias. The method also ignored the potential presence and nature of surfaces (e.g.,

erosional vs accretionary basal surface) picked up by FMI logs. Additionally, FMI logs can be processed to extract data on the geometry (orientation, dip angle) of surfaces that describe palaeoflow and/or accretion style, which may therefore reflect the morphodynamics of fluvial sub-environments. Attributes of sandbody bounding surfaces and bedding surfaces could be used as additional training-data features. Furthermore, river planform interpretations could leverage on sedimentological data that can be used to establish the consistency of the labelling with inferences of likely planform based on palaeohydraulics (cf. Lyster et al., 2022); however, the use of these data and their systematic processing by a machine is likely to be challenging in practice. A limitation that affects both the proposed machine-learning approach and its benchmark test against expert geologists lies in having disregarded contextual information that is normally available in addition to facies sequences seen on a vertical profile (e.g., information on the general depositional setting, interpretations of neighbouring facies associations); this information can be readily used by a human, but was not provided in the sandbody classification challenge exercise. In principle, however, machine-learning models of the type presented here could also be expanded to incorporate soft knowledge of this type, alongside other predictors (e.g., geochemical logs). Additionally, the predictive power of models for labelling sandbody types based on facies characteristics can plausibly be improved by filtering the sedimentological knowledge base on attributes describing depositional context and geological boundary conditions, so as to exclusively extract data from analogue case studies that are comparable to the subsurface

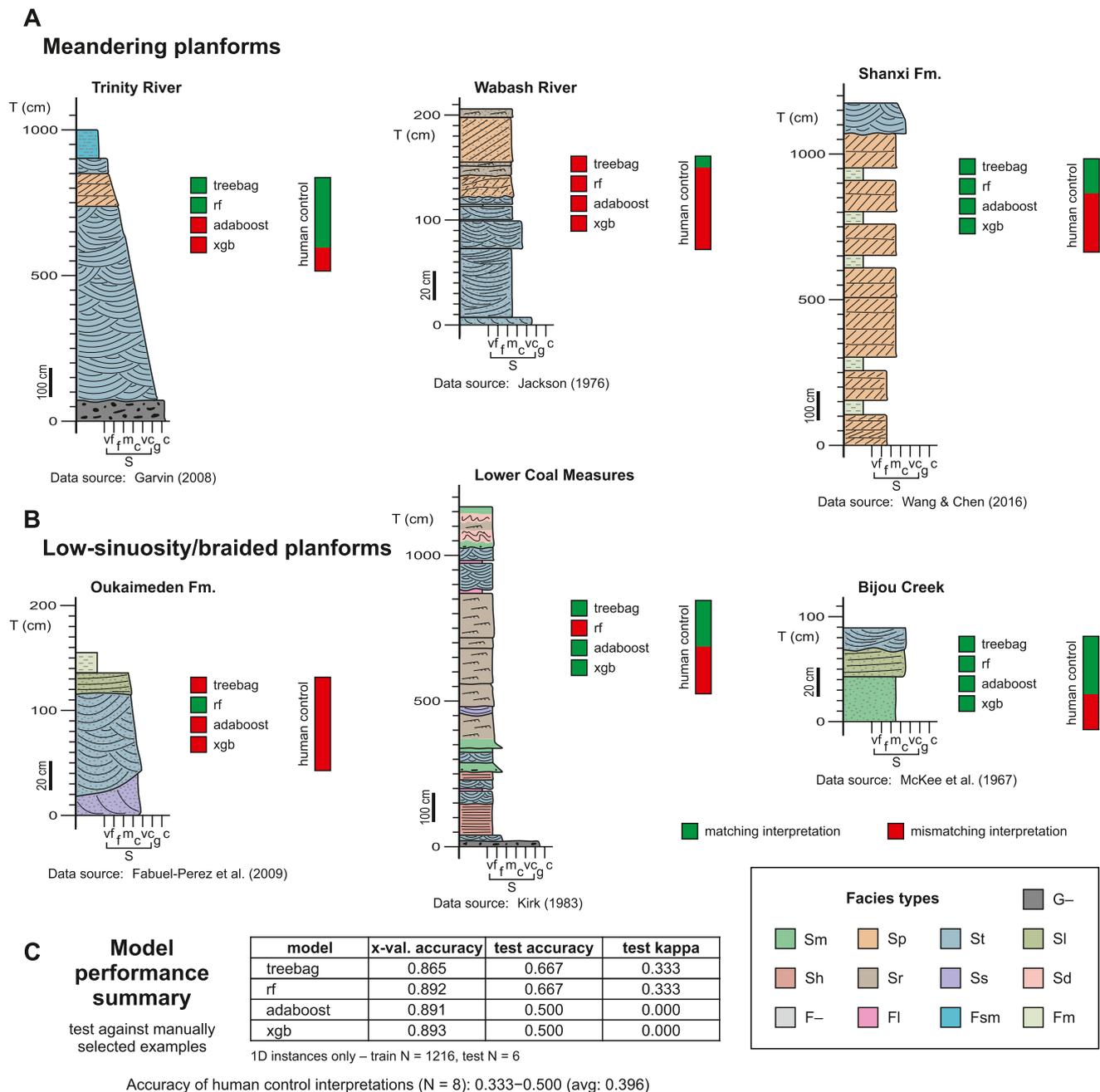


Fig. 11. Vertical profiles of the channel bodies chosen for the river-planform interpretation exercise proposed to the models, trained as explained in the text, and to eight volunteering sedimentologists, shown separately for channel deposits of meandering (A) and braided or low-sinuosity (B) river systems. Cases of correct and incorrect classifications by the four types of models and the geologists ('human control') are reported. (C) summary of the performance of humans and models in the channel-body classification challenge. See Table 2 for model abbreviations.

succession of interest (cf. Colombera et al., 2013; Colombera and Moutney, 2019). Given the mass of data required for training machine-learning models, producing tailored database outputs for specific analogue types requires continued growth of the global sedimentological knowledge base.

In spite of these limitations, this study demonstrates how the use of outcrop and modern analogues for training machine-learning tools widens the range of tasks that can be partially automated in subsurface workflows, provided that a sufficiently large analogue knowledge base is available, either for direct use (cf. Colombera and Budai, 2026) or for supporting the creation of synthetic datasets of appropriate size for model training (e.g., artificial sedimentary logs; cf. Budai et al., 2025). The approach can be generalized to other types of sedimentary

successions for which analogue datasets of appropriate size and type exist (cf. Colombera et al., 2016; Cullis et al., 2019; Cosgrove et al., 2021), and may be especially useful in cases where 1D facies sequences are more meaningful indicators of depositional sub-environments (e.g., in shallow-marine depositional systems). This addresses the need to fill the gap left by the vocation crisis affecting geoscience university degrees in many countries (Rogers et al., 2024, and references therein), and the difficulty in recruiting subsurface geologists and retaining sedimentological expertise in subsurface industries.

6. Conclusions

This study has presented a new method for the use of machine

learning to make binary interpretations of fluvial sandbodies, according to (i) their general channel or floodplain origin, and to (ii) the interpreted river planform style ('low-sinuosity or braided' vs 'meandering' channel bodies). Key aspects of the method are as follows: (i) it mirrors the way in which 1D facies models are used as a reference for classification, especially by geologists who lack specific sedimentological knowledge for this type of successions; (ii) it is based on the training of machine-learning models based on a large knowledge base of global sedimentological studies of ancient outcropping successions and of the recent sedimentary record of modern rivers. Four ensemble machine-learning methods (bagged classification tree, random forest classifier, adaptive boosting classifier, and extreme gradient boosting classifier) were trained and tested using data from the geological analogues. Some differences in performance are seen across the four models, but their application to the test datasets is generally robust, with balanced accuracy values that are typically around 0.8 for tasks of both general sandbody classification and river planform interpretation, for models applied to sedimentary logs from sandbodies that are excluded from the training dataset. A benchmark test against selected sandbodies that were deemed as being difficult to interpret and that were independently labelled by eight expert sedimentologists highlights that the way the models fail to perform is comparable to the way a geologist may be misguided by relying excessively on 1D facies models. Such uncertainty is inherent in facies analysis: it is unlikely to ever be resolved, but this study proposes a way to quantify it through the simultaneous application of a set of models that can integrate geological reasoning. The approach can be refined in many ways, but the outcomes of this study reveal that there is scope for trying to automate the sedimentological interpretation of large 1D datasets (e.g. FMI logs) documenting the facies organization of subsurface sedimentary bodies.

CRediT authorship contribution statement

Luca Colombera: Writing – review & editing, Writing – original draft, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Soma Budai:** Writing – review & editing, Methodology, Investigation. **Nigel P. Mountney:** Writing – review & editing, Resources, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the sponsors of the Fluvial, Eolian & Shallow-Marine Research Group for financial support to this research, and in particular the companies sponsoring phase six of the consortium (AkerBP, ENI, Equinor, Murphy Oil, Oxy). We are grateful to the geologists who took part in the sandbody classification challenge for their contribution. We thank Richard Worden and two anonymous reviewers for their constructive comments to earlier versions of this paper, which have improved the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.sedgeo.2026.107047>.

Data availability

Data available in the supplementary material

References

- Alaudah, Y., Michalowicz, P., Alfarraj, M., AlRegib, G., 2019. A machine-learning benchmark for facies classification. *Interpretation* 7 (3), SE175–SE187. <https://doi.org/10.48550/arXiv.1901.07659>.
- Allen, J.R.L., 1983. *Studies in fluvial sedimentation: bars, bar-complexes and sandstone sheets (low-sinuosity braided streams) in the Brownstones (L. Devonian), Welsh Borders*. *Sedimentary Geology* 33 (4), 237–293.
- Anderton, R., 1985. Clastic facies models and facies analysis. In Brenchly, P.J., Williams, P.J. (Eds.), *Sedimentology: recent developments and applied aspects*, Geological Society, London, Special Publications, 18, 31–47.
- Barwis, J.H., McPherson, J.G., Studlick, J.R., 1990. *Sandstone petroleum reservoirs*. Springer Science & Business Media, p. 583.
- Bestagini, P., Lipari, V., Tubaro, S., 2017. A machine learning approach to facies classification using well logs. SEG technical program expanded abstracts 2017, 2137–2142. (Society of Exploration Geophysicists).
- Bressan, T.S., de Souza, M.K., Girelli, T.J., Junior, F.C., 2020. Evaluation of machine learning methods for lithology classification using geophysical data. *Computational Geosciences* 139, 104475. <https://doi.org/10.1016/j.cageo.2020.104475>.
- Bridge, J.S., 1985. Paleochannel patterns inferred from alluvial deposits: a critical evaluation. *Journal of Sedimentary Research* 55 (4), 579–589.
- Bristow, C.S., 1996. Reconstructing fluvial channel morphology from sedimentary sequences. In: Carling, P.A., Dawson, M.R. (Eds.), *Advances in Fluvial Dynamics and Stratigraphy*. Wiley, pp. 351–371.
- Budai, S., Colombera, L., McArthur, A., Patacci, M., 2025. Quantitative bed-type classification for a global comparison of deep-water sedimentary systems. *Computational Geosciences* 199, 105917. <https://doi.org/10.1016/j.cageo.2025.105917>.
- Bühlmann, P., 2012. Bagging, boosting and ensemble methods. In: *Handbook of Computational Statistics: Concepts and Methods*. Springer, pp. 985–1022.
- Burns, C.E., Mountney, N.P., Hodgson, D.M., Colombera, L., 2019. Stratigraphic architecture and hierarchy of fluvial overbank splay deposits. *Journal of the Geological Society* 176 (4), 629–649.
- Cant, D.J., 1982. Fluvial facies models and their application. In: Scholle, P.A., Spearing, D. (Eds.), *Sandstone Depositional Environments*, AAPG Memoir vol. 31, pp. 115–137.
- Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* 16, 321–357.
- Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794.
- Colombera, L., Budai, S., 2026. Identification of shallow-water clastic parasequences in borehole data using machine learning trained on geologic analogs. In: Feldman, H., Ainsworth, R.B., Colombera, L., Caldwell, R.L. (Eds.), *Are Parasequences Still Relevant?* SEPM Special Publication, 115. <https://doi.org/10.2110/sepm.115.01> in press.
- Colombera, L., Felletti, F., Mountney, N.P., McCaffrey, W.D., 2012. A database approach for constraining stochastic simulations of the sedimentary heterogeneity of fluvial reservoirs. *AAPG Bulletin* 96 (11), 2143–2166.
- Colombera, L., Mountney, N.P., 2019. The lithofacies organization of fluvial channel deposits: a meta-analysis of modern rivers. *Sedimentary Geology* 383, 16–40.
- Colombera, L., Mountney, N.P., 2021. Influence of fluvial crevasse-splay deposits on sandbody connectivity: Lessons from geological analogues and stochastic modelling. *Marine and Petroleum Geology* 128, 105060. <https://doi.org/10.1016/j.marpetgeo.2021.105060>.
- Colombera, L., Mountney, N.P., McCaffrey, W.D., 2013. A quantitative approach to fluvial facies models: Methods and example results. *Sedimentology* 60 (6), 1526–1558.
- Colombera, L., Mountney, N.P., Russell, C.E., Shiers, M.N., McCaffrey, W.D., 2017. Geometry and compartmentalization of fluvial meander-belt reservoirs at the bar-form scale: Quantitative insight from outcrop, modern and subsurface analogues. *Marine and Petroleum Geology* 82, 35–55.
- Colombera, L., Shiers, M.N., Mountney, N.P., 2016. Assessment of backwater controls on the architecture of distributary-channel fills in a tide-influenced coastal-plain succession: Campanian Neslen Formation, USA. *Journal of Sedimentary Research* 86 (5), 476–497.
- Cosgrove, G.I.E., Colombera, L., Mountney, N.P., 2021. A Database of Aeolian Sedimentary Architecture for the characterization of modern and ancient sedimentary systems. *Marine and Petroleum Geology* 127, 104983. <https://doi.org/10.1016/j.marpetgeo.2021.104983>.
- Cullis, S., Patacci, M., Colombera, L., Bührig, L., McCaffrey, W.D., 2019. A database solution for the quantitative characterisation and comparison of deep-marine siliciclastic depositional systems. *Marine and Petroleum Geology* 102, 321–339.
- Dalrymple, R.W., 2010. Interpreting sedimentary successions: Facies, facies analysis and facies models. In: James, N.P., Dalrymple, R.W. (Eds.), *Facies Models 4*. Geological Association of Canada, pp. 3–18.
- Deutsch, C.A., Tran, T.T., 2002. FLUVSIM: a program for object-based stochastic modeling of fluvial depositional systems. *Computational Geosciences* 28 (4), 525–535.
- Di Martino, A., Carlini, G., Castellani, G., Remondini, D., Amorosi, A., 2023. Sediment core analysis using artificial intelligence. *Scientific Reports* 13 (1), 20409. <https://doi.org/10.1038/s41598-023-47546-2>.
- Donselaar, M.E., Overeem, I., 2008. Connectivity of fluvial point-bar deposits: an example from the Miocene Huesca fluvial fan, Ebro basin, Spain. *AAPG Bulletin* 92 (9), 1109–1129.

- Donselaar, M.E., Schmidt, J.M., 2005. Integration of outcrop and borehole image logs for high-resolution facies interpretation: example from a fluvial fan in the Ebro Basin, Spain. *Sedimentology* 52 (5), 1021–1042.
- Dott Jr., R.H., Bourgeois, J., 1983. Hummocky stratification: significance of its variable bedding sequences: discussion and reply: reply. *Geological Society of America Bulletin* 94 (10), 1249–1251.
- Emmert-Streib, F., Moutari, S., Dehmer, M., 2019. A comprehensive survey of error measures for evaluating binary decision making in data science. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9 (5), e1303. <https://doi.org/10.1002/widm.1303>.
- Eugster, M.J., Hothorn, T., Leisch, F., 2008. Exploratory and inferential analysis of benchmark experiments. Technical Report 30, 28 Pp. Institut für Statistik, Ludwig-Maximilians-Universität München, Munich, Germany.
- Ferreira, A.J., Figueiredo, M.A., 2012. Boosting algorithms: A review of methods, theory, and applications. In: Zhang, C., Ma, Y. (Eds.), *Ensemble Machine Learning: Methods and Applications*. Springer, pp. 35–85.
- Fielding, C.R., 2006. Upper flow regime sheets, lenses and scour fills: extending the range of architectural elements for fluvial sediment bodies. *Sedimentary Geology* 190 (1–4), 227–240.
- Fielding, C.R., Alexander, J., Argiro, A., 2025. Towards a new generation of fluvial facies models for the interpretation of ancient deposits, based on inter-annual peak discharge variance of modern rivers. *Sedimentology* 72 (1), 275–329.
- Friend, P.F., 1983. Towards the field classification of alluvial architecture or sequence. In: Collinson, J.D., Lewin, J. (Eds.), *Modern and Ancient Fluvial Systems*, IAS Special Publication vol. 6, 345–354.
- Gharieb, A., Ibrahim, A.F., Gabry, M.A., Elsayy, M., Algarhy, A., Darraj, N., 2024. Automated borehole image interpretation using computer vision and deep learning. *SPE Journal*, 1–16.
- Hall, M., Hall, B., 2017. Distributed collaborative prediction: results of the machine learning contest. *The Leading Edge* 36 (3), 267–269.
- Halotel, J., Demyanov, V., Gardiner, A., 2020. Value of geologically derived features in machine learning facies classification. *Mathematical Geoscience* 52 (1), 5–29.
- Hothorn, T., Leisch, F., Zeileis, A., Hornik, K., 2005. The design and analysis of benchmark experiments. *Journal of Computational and Graphical Statistics* 14 (3), 675–699.
- Hou, L., Ma, C., Tang, W., Zhou, Y., Ye, S., Chen, X., Zhang, X., Yu, C., Chen, A., Zheng, D., Cao, Z., Zhang, Y., Hou, M., 2024. DDViT: advancing lithology identification on FMI image logs through a dual modal transformer model with less information drop. *Geoenvironment Science and Engineering* 234, 212662. <https://doi.org/10.1016/j.geoen.2024.212662>.
- Jackson, R.G. II, 1976. Depositional model of point bars in the lower Wabash River. *Journal of Sedimentary Research*, 46(3), 579–594.
- Jackson, R.G. II, 1978. Preliminary evaluation of lithofacies models for meandering alluvial streams. In: Miall, a.D. (Ed.), *Fluvial Sedimentology*, Canadian Society of Petroleum Geologists Memoir, 5, 543–576.
- Jordan, D.W., Pryor, W.A., 1992. Hierarchical levels of heterogeneity in a Mississippi River meander belt and application to reservoir systems. *AAPG Bulletin* 76 (10), 1601–1624.
- Keeton, G.I., 2012. Sedimentological and Stratigraphic Characteristics of Fluvial Sandstones Based on Outcrop Spectral-Gamma-Ray Data and Borehole Images, Williams Fork Formation, Piceance Basin, Colorado. Master's Thesis, University of Colorado at Boulder, p. 153.
- Keeton, G.I., Pranter, M.J., Cole, R.D., Gustason, E.R., 2015. Stratigraphic architecture of fluvial deposits from borehole images, spectral-gamma-ray response, and outcrop analogs, Piceance Basin, Colorado. *AAPG Bulletin* 99 (10), 1929–1956.
- Keogh, K.J., Martinius, A.W., Osland, R., 2007. The development of fluvial stochastic modelling in the Norwegian oil industry: a historical review, subsurface implementation and future directions. *Sedimentary Geology* 202 (1–2), 249–268.
- Kuhn, M., 2008. Building predictive models in R using the caret package. *Journal of Statistical Software* 28, 1–26.
- Kusumah, E.P., Riadi, R.S., Setiawan, T.S., 2022. Linear Regression for Flooding Surface Identification in well Log, and Outcrop image. *Bulletin of Geology* 6 (2), 978–993.
- Lang, S.C., Kassan, J., Benson, J., Grasso, C., Hicks, T., Hall, N., Avenell, C., 2002. Reservoir Characterisation of Fluvial, Lacustrine and Deltaic Successions—Applications of Modern and Ancient Geological Analogues. *Indonesian Petroleum Association 28th Annual Convention Proceedings* vol. 1, 557–580.
- Larue, D.K., Hovadik, J., 2006. Connectivity of channelized reservoirs: a modelling approach. *Petroleum Geoscience* 12 (4), 291–308.
- Lewin, J., Ashworth, P.J., Strick, R.J., 2017. Spillage sedimentation on large river floodplains. *Earth Surface Processes and Landforms* 42 (2), 290–305.
- Li, W., Yue, D., Colombera, L., Du, Y., Zhang, S., Liu, R., Wang, W., 2021. Quantitative prediction of fluvial sandbodies by combining seismic attributes of neighboring zones. *Journal of Petroleum Science and Engineering* 196, 107749. <https://doi.org/10.1016/j.petrol.2020.107749>.
- Li, W., Yue, D., Colombera, L., Duan, D., Long, T., Wu, S., Liu, Y., 2023. A novel method for seismic-attribute optimization driven by forward modeling and machine learning in prediction of fluvial reservoirs. *Geoenvironment Science and Engineering* 227, 211952. <https://doi.org/10.1016/j.geoen.2023.211952>.
- Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* 30, 4765–4774. Presented at the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, California.
- Lunt, I.A., Bridge, J.S., Tye, R.S., 2004. A quantitative, three-dimensional depositional model of gravelly braided rivers. *Sedimentology* 51 (3), 377–414.
- Lyster, S.J., Whittaker, A.C., Hajek, E.A., 2022. The problem of paleo-planforms. *Geology* 50 (7), 822–826.
- Martin, T., Meyer, R., Jobe, Z., 2021. Centimeter-scale lithology and facies prediction in cored wells using machine learning. *Frontiers in Earth Science* 9, 659611. <https://doi.org/10.3389/feart.2021.659611>.
- Mayer, M., Watson, D., 2025. kernelshap: Kernel SHAP. R package version 0.9.1.
- Miall, A.D., 1978. Lithofacies types and vertical profile models in braided river deposits: a summary. In: Miall, a.D. (Ed.), *Fluvial Sedimentology*, Canadian Society of Petroleum Geologists Memoir, 5, pp. 597–604.
- Miall, A.D., 1988. Reservoir heterogeneities in fluvial sandstones: lessons from outcrop studies. *AAPG Bulletin* 72 (6), 682–697.
- Miall, A.D., 1992. Alluvial deposits. In: Walker, R.G., James, N.P. (Eds.), *Facies Models: Response to Sea Level Change*. Geological Association of Canada, pp. 119–142.
- Miall, A.D., 1996. *The Geology of Fluvial Deposits*. Springer Verlag, p. 582.
- Olsen, L., 2024. groupdata2: Creating groups from Data. R package version 2.0.5. <https://github.com/ludvigolsen/groupdata2>.
- Parmar, A., Katariya, R., Patel, V., 2019. A review on random forest: An ensemble classifier. In: Hemanth, J., Fernando, X., Lafata, P., Baig, Z. (Eds.), *International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018: Lecture Notes on Data Engineering and Communications Technologies*, 26, pp. 758–763.
- Pöppelreiter, M., García-Carballido, C., & Kraaijveld, M., 2010. Dipmeter and Borehole Image Log Technology. *AAPG Memoir* 92. American Association of Petroleum Geologists, Tulsa OK, USA. 357 pp.
- Pyrzc, M.J., Boisvert, J.B., Deutsch, C.V., 2009. ALLUVSIM: a program for event-based stochastic modeling of fluvial depositional systems. *Computational Geosciences* 35 (8), 1671–1685.
- R Core Team, 2023. R: A Language and Environment for Statistical Computing. R 642 Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rashmi, V.K., Gilad-Bachrach, R., 2015. Dart: Dropouts meet multiple additive regression trees. In: *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics, JMLR: W&CP Volume 38*, pp. 489–497.
- Reesink, A.J., Bridge, J.S., 2011. Evidence of bedform superimposition and flow unsteadiness in unit-bar deposits, South Saskatchewan River, Canada. *Journal of Sedimentary Research* 81 (11), 814–840.
- Rogers, S.L., Giles, S., Doney, N., Greene, S.E., Bhatia, R., Van Landeghem, K., King, C., 2024. “You just look at rocks, and have beards” Perceptions of Geology from the United Kingdom: a Qualitative Analysis from an Online Survey. *Earth Science, Systems and Society* 4 (1), 10078. <https://doi.org/10.3389/esss.2024.10078>.
- Simon, N., Worden, R.H., 2025. A digital framework for estuarine stratigraphy: an example of a machine learning approach to paleo-environmental classification and coastal evolution. *Sedimentary Geology* 492, 107008. <https://doi.org/10.1016/j.sedgeo.2025.107008>.
- Slatt, R.M., 2013. *Stratigraphic Reservoir Characterization for Petroleum Geologists, Geophysicists, and Engineers*. Developments in Petroleum Science. Elsevier, Amsterdam, The Netherlands. 671 pp.
- Tian, M., Tan, M., Wang, M., 2023. Identification of Shale Lithofacies from FMI Images and ECS Logs using Machine Learning with GLCM Features. *Processes* 11 (10), 2982. <https://doi.org/10.3390/pr11102982>.
- Tognoli, F.M.W., Spaniol, A.F., de Mello, M.E., de Souza, L.V., 2024. A machine-learning based approach to predict facies associations and improve local and regional stratigraphic correlations. *Marine and Petroleum Geology* 160, 106636.
- Tye, R.S., 2013. Quantitatively modeling alluvial strata for reservoir development with examples from Krasnoleninskoye field, Russia. *Journal of Coastal Research* 69, 129–152.
- Valentim, M.B., Bom, C.R., Coelho, J.M., Correia, M.D., De Albuquerque, M.P., de Albuquerque, M.P., Faria, E.L., 2019. A deep residual convolutional neural network for automatic lithological facies identification in Brazilian pre-salt oilfield wellbore image logs. *Journal of Petroleum Science and Engineering* 179, 474–503.
- Visher, G.S., 1965. Fluvial processes as interpreted from ancient and recent fluvial deposits. In: Middleton, G.V. (Ed.), *Primary Sedimentary Structures and their Hydrodynamic Interpretation*. SEPM Special Publication vol. 12, 116–132.
- Walker, R.G., 1992. *Facies, facies models and modern stratigraphic concepts*. In: Walker, R.G., James, N.P. (Eds.), *Facies Models: Response to Sea Level Change*. Geological Association of Canada, Stittsville, ON, Canada, pp. 1–14.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D.A., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. *Journal of Open Source Software* 4 (43), 1686. <https://doi.org/10.21105/joss.01686>.
- Willis, B.J., Sech, R.P., 2019. Quantifying impacts of fluvial intra-channel-belt heterogeneity on reservoir behaviour. In: Ghinassi, M., Mountney, N.P., Colombera, L., Reesink, A.J. (Eds.), *Fluvial Meanders and their Sedimentary Products in the Rock Record*, IAS Special Publication vol. 48, 543–572.
- Yeste, L.M., Varela, A.N., Viseras, C., McDougall, N.D., García-García, F., 2020. Reservoir architecture and heterogeneity distribution in floodplain sandstones: key features in outcrop, core and wireline logs. *Sedimentology* 67 (7), 3355–3388.
- Zheng, W., Tian, F., Di, Q., Xin, W., Cheng, F., Shan, X., 2021. Electrofacies classification of deeply buried carbonate strata using machine learning methods: a case study on ordovician paleokarst reservoirs in Tarim Basin. *Marine and Petroleum Geology* 123, 104720. <https://doi.org/10.1016/j.marpetgeo.2020.104720>.