Journal of Water Resource and Protection, 2026, 18(11), 812-835
https://www.scirp.org/journal/jwarp
ISSN Online: 1945-3108
ISSN Print: 1945-3094

Predicting the Outcome of a Borehole during Drilling from Regolith and Fissured-Layer Thickness: A 2432-Borehole Random-Forest Study with Spatial Cross-Validation (Burkina Faso, West Africa)

Relwinde Abdoul-Karim Nassa[1*], Hoboubé Nicole Abotsi[2], Julien Nikiema[2]

[1]Félix Houphouët-Boigny National Polytechnic Institute, Yamoussoukro, Côte d'Ivoire
[2]Geosciences and Environment Laboratory, Joseph KI-ZERBO University, Ouagadougou, Burkina Faso
Email: *nassa.ing.inphb@gmail.com

How to cite this paper: Nassa, R.A.-K., Abotsi, H.N. and Nikiema, J. (2026) Predicting the Outcome of a Borehole during Drilling from Regolith and Fissured-Layer Thickness: A 2432-Borehole Random-Forest Study with Spatial Cross-Validation (Burkina Faso, West Africa). *Journal of Water Resource and Protection*, **18**, 812-835.
https://doi.org/10.4236/jwarp.2026.1811040

Received: September 21, 2026
Accepted: November 16, 2026
Published: November 19, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

Siting and drilling water boreholes in hard-rock aquifers remains costly and uncertain. This study uses 2432 boreholes from a 108 × 111 km area of the Precambrian crystalline basement within the Ouagadougou map sheet of Burkina Faso (12.0˚ - 13.0˚N, 1.0˚ - 2.0˚W), West Africa, to establish the rule that governs the depth drilled and to predict the outcome of a borehole from the thicknesses of the regolith and of the fissured layer. Recorded depth is the exact sum of the two recorded thicknesses for all 2432 boreholes, so the depth of a borehole is governed by a calibrated stopping rule at the base of the fissured layer; a random forest reproduces it with $R^2 = 0.994$ and MAE = 0.42 m on a hold-out set and multiple linear regression reproduces it exactly, which converts the estimation of depth into the estimation of two thicknesses, with an error that transfers one for one and sets the accuracy that pre-drilling geophysics must reach. For the outcome, a random-forest classifier of four use-based yield classes (negative <0.7 m^3/h, low, medium, high ≥5 m^3/h), using the position, the regolith thickness and the thickness of fissured rock already crossed, reaches an accuracy of 48.3%, a balanced accuracy of 40.5% and a macro-AUC of 0.70, with a recall of 71% and an AUC of 0.82 for the negative class, against 35.0% and 25% and 0.50 for the majority class and chance. Ten-fold spatial block cross-validation with a 2 km buffer, whose minimum train-test separation is 2002 m, gives 38.9% and 0.69, so the model transfers to areas not represented in training; extra trees, gradient boosting and logistic regression converge to the same ceiling on

identical folds, which places the limit in the information content of the predictors rather than in the algorithm. The prediction is strongest where it is most useful: among the boreholes in which more than 40 m of fissured rock was crossed, 66% ended negative, against 12% among those in which less than 20 m was crossed, which supports a quantified stop-or-continue decision while the rig is on site. Success also rises with regolith thickness from 56% to 77%, a targeting signal at programme scale, whereas the regolith × fissured index used locally as a "hydro-potential" adds no information once both thicknesses are known.

Keywords

Hard-Rock Aquifer, Borehole Drilling, Random Forest, Spatial Cross-Validation, Drilling Decision Support, Weathered Layer, Burkina Faso

1. Introduction

Groundwater is the main source of drinking water for rural populations of the West African crystalline basement, where surface water is seasonal and aquifers are discontinuous [1]. Boreholes equipped with hand pumps or small piped schemes are the standard supply solution, but the operation remains expensive and uncertain: failure rates of 20% - 40% are common, and the depth actually drilled, hence the cost, is decided in the field, often without a quantitative forecast. Two questions therefore govern any siting decision: how deep will the borehole have to be, and what yield can be expected?

Hard-rock aquifers are today described by a stratiform conceptual model [2] [3] in which a capacitive, low-permeability regolith (saprolite) overlies a fissured layer whose dense sub-horizontal fracturing, generated by weathering processes rather than by tectonics, provides the transmissive function. The borehole must cross the regolith and capture the upper part of the fissured layer, where hydraulic conductivity is highest and decreases downward [4]. This model gives a physical basis for expecting the thickness of the two layers to control both the depth to be drilled and the probability of success. Large operational databases confirm this expectation. Courtois *et al.* [5] mapped the thicknesses of the regolith and fissured layer over Burkina Faso from 16,116 boreholes; Adeotan *et al.* [6], on 2347 Burkinabe boreholes, found the thickness of the weathered layers to be the best predictor of drilling success whereas lineaments, long used as the primary siting criterion, showed no significant correlation with productivity; Iglesias-Martínez *et al.* [7] reported, near Bobo-Dioulasso, yields three times higher under weathered covers thicker than 40 m than under covers thinner than 20 m. In Côte d'Ivoire, specific-capacity analyses [8] [9] identified lithology, weathered-zone thickness and the type of electrical-sounding anomaly as the main productivity controls, and depth-optimisation strategies were developed from drilling parameters [10].

Machine learning has entered this field mainly through groundwater potential mapping, where tree-based ensembles, random forest [11], boosted regression trees and extra trees, consistently outperform single trees and linear models [12] [13]. Gómez-Escalonilla *et al.* [14] showed, on 483 boreholes of southern Mali, that borehole yield is tractable when reformulated as a multiclass problem, reaching a balanced accuracy of about 0.80 with lithological, structural and geomorphological predictors. Two aspects remain little explored. First, most studies validate by random splits, although boreholes are spatially clustered and random folds let neighbouring boreholes fall on both sides of the split, which inflates apparent transferability [15]. Second, operational databases record the depth at which drilling stopped and the thickness actually crossed, not the geological thickness of the layers; the consequences of this censoring for what a model can learn have rarely been examined.

This study addresses both aspects with a database of 2432 boreholes from the crystalline basement within the Ouagadougou map sheet of Burkina Faso, several times larger than those of the Ivorian and Malian studies cited above. Its objectives are 1) to establish the rule that links the depth drilled to the thicknesses of the two layers, and the accuracy that an estimate of those thicknesses must reach; 2) to build and validate a prediction of the outcome of a borehole, as four use-based yield classes, from the information available while the rig is on site; 3) to compare random and buffered spatial block cross-validation and to benchmark random forest against extra trees, gradient boosting and linear models on identical folds; 4) to assess the contribution of a derived index, the product of regolith and fissured thicknesses, presented locally as a "hydro-potential"; and 5) to interpret the results in the light of the conceptual model of hard-rock aquifers and of the censoring inherent in operational databases. The analysis is delivered with an open, browser-based tool (HydroPrédict) that trains, validates, predicts and maps from a spreadsheet without installation.

2. Study Area and Data

2.1. Study Area

The study area is a 108 km (east-west) by 111 km (north-south) block within the Ouagadougou map sheet of Burkina Faso, between 12.0˚ and 13.0˚N and 1.0˚ and 2.0˚W (UTM zone 30N, eastings 609 - 717 km, northings 1327 - 1437 km), which includes Ziniaré and Loumbila. The Ouagadougou dataset was chosen in order to have a larger amount of data for a more accurate model (**Figure 1**). The bedrock belongs to the Palaeoproterozoic Birimian domain of the West African craton, granitoids, greenstone belts and metasediments, overlain by a lateritic-to-clayey regolith developed under a Sudano-Sahelian climate (annual rainfall 700 - 900 mm concentrated in June-September) (**Figure 2**). Groundwater is abstracted almost exclusively from the fissured layer of the basement through boreholes 40 - 80 m deep equipped with hand pumps, and, for larger villages, from higher-yielding boreholes feeding small piped schemes.

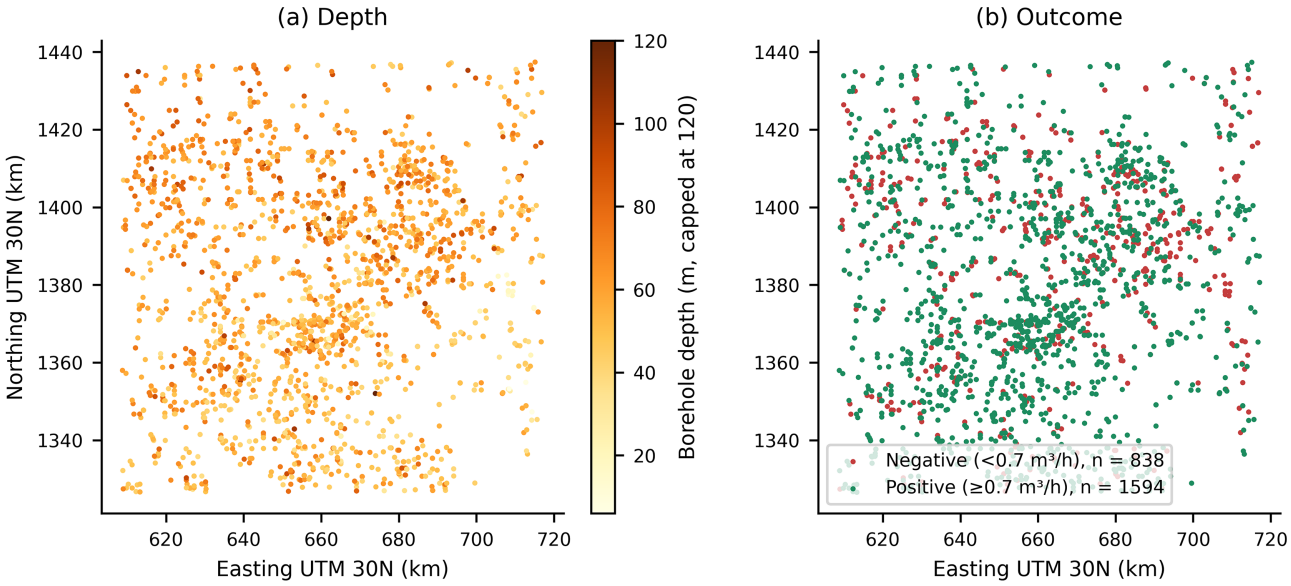


Figure 1. Study area: location of the 2432 boreholes in UTM 30N coordinates, coloured by (a) Final depth and (b) Outcome at the 0.7 m³/h threshold.

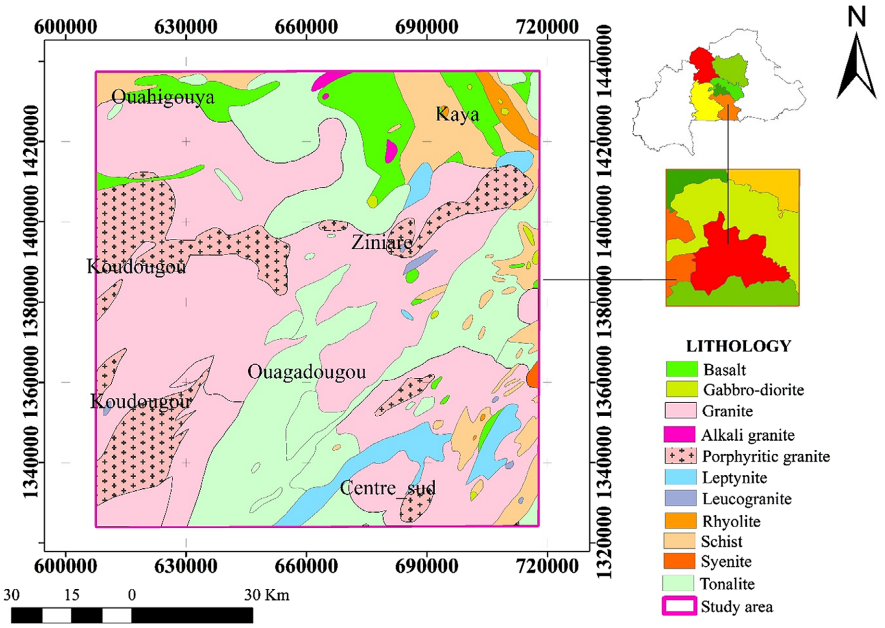


Figure 2. Geology of the study area (modified from Castaing *et al.* [16]).

2.2. Borehole Database

The database compiles 2432 water boreholes drilled for rural water supply and documented by the drilling contractors. The compilation was assembled from the drilling reports held by the Direction Générale des Ressources en Eau et d'Hydraulique (DGREH), the Bureau des Mines et de la Géologie du Burkina (BUMIGEB), the Ouvrages de Mobilisation des Eaux Souterraines (OMES) programme and the Bureau de Recherches Géologiques et Minières (BRGM); the properties of the file supplied to the analysis are reported in the paragraph below. Each record contains the UTM

coordinates, the thickness of the regolith (weathered layer, in m), the thickness of the fissured layer crossed (m), the end-of-development yield (air-lift or pumping test, m^3/h) and the final depth (m). Records with missing values were discarded before analysis. Descriptive statistics are given in Table 1 and distributions in Figure 3. Median depth is 55 m (mean 57.4 m); 2 boreholes deeper than 150 m (455 and 672 m) are exceptional exploration holes and were retained because they test the extrapolation behaviour of the models. Yield is strongly right-skewed: median 1.20 m^3/h, mean 3.11 m^3/h, 22.7% of boreholes with a recorded yield of zero and 65.5% of positive boreholes at the 0.7 m^3/h threshold used in Burkina Faso to equip a hand pump.

Table 1. Descriptive statistics of the borehole database (n = 2432).

Variable	Min	Q1	Median	Mean	Q3	Max	SD
Northing (m, UTM 30N)	1,326,648	1,360,852	1,377,880	1,379,301	1,401,146	1,437,380	27,565
Easting (m, UTM 30N)	608,751	639,622	661,897	660,225	679,702	717,063	26,017
Regolith thickness (m)	0.0	16.0	25.0	25.3	33.0	85.0	12.1
Fissured-layer thickness (m)	0.0	19.0	30.0	32.1	43.0	669.0	22.4
Regolith × fissured index (m^2)	0	444	627	698	900	13,926	458
Yield (m^3/h)	0.00	0.20	1.20	3.11	3.14	108.0	6.19
Depth (m)	6.0	47.0	55.0	57.4	66.0	672.0	20.5

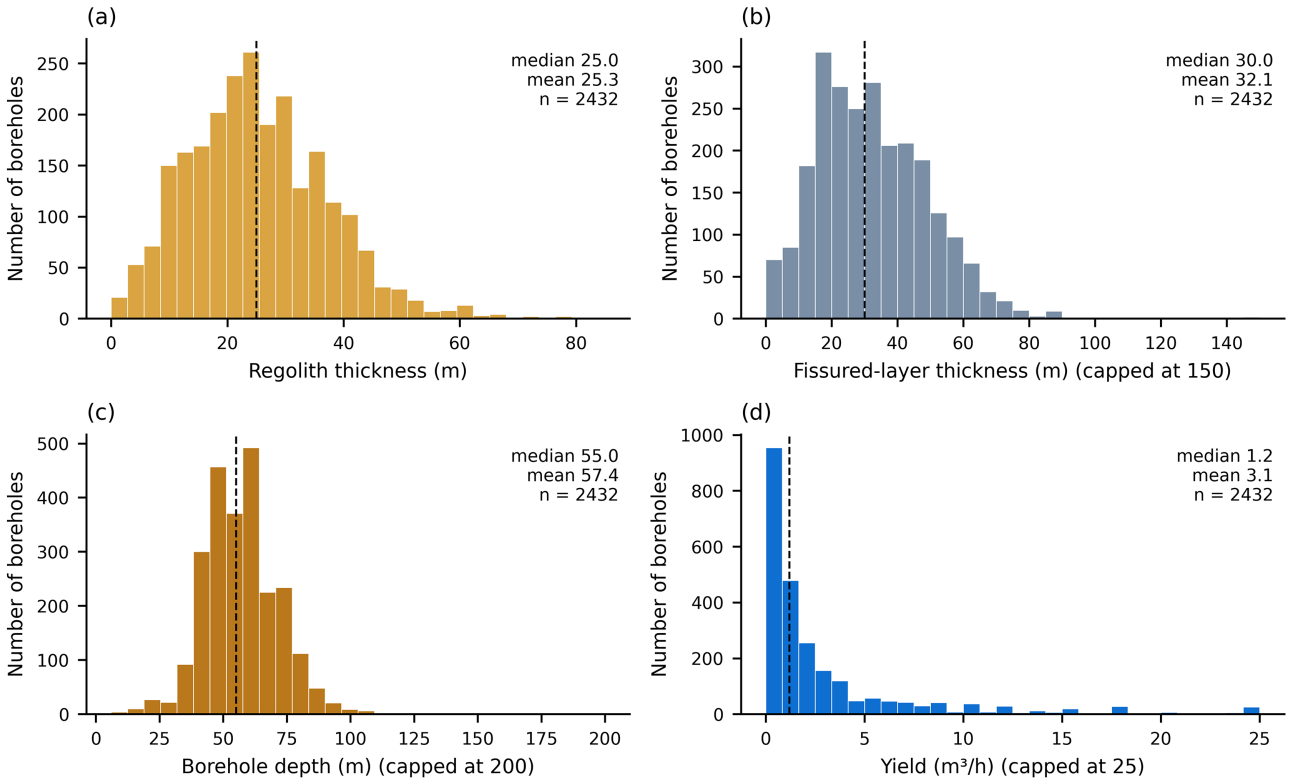


Figure 3. Distributions of (a) Regolith thickness, (b) Fissured-layer thickness, (c) Borehole depth and (d) Yield; dashed lines mark medians.

Two properties of the database matter for the interpretation. First, the recorded depth is, for every borehole, exactly the sum of the recorded regolith and fissured thicknesses (Pearson r = 1.000): the fissured thickness is the thickness crossed before drilling stopped, not the thickness of the geological layer. Second, the database is operational: drilling stops when a sufficient yield is obtained and continues, within budget, when it is not, so that recorded thicknesses are censored by the outcome. Both properties are examined in Section 5.

Three properties of the file analysed here are reported for transparency, following the data-provenance requirement. First, the file contains 2432 records and no missing value in any of the seven columns, so the complete-case selection was applied upstream of the file supplied to the analysis; the number of records before that selection and the missingness by variable are held by the compiling institutions and were not available to us. Second, 8 records are strictly identical on all variables and 574 records share their coordinates with at least one other record, which is expected where several boreholes were drilled at one village but has a direct consequence for validation design (Section 3.3). Third, the end-of-development yield is reported as a single figure per borehole, without the test type (airlift during development or pumping test) or its duration; where both procedures coexist in such compilations the yields are not strictly comparable, and the test procedure applied to each record is not documented in the file supplied to the analysis. The four yield classes used here are wide (0.7, 2.5 and 5 m^3/h), so a difference of method would have to displace a borehole across one of these thresholds to change its class.

2.3. Predictors, Derived Index and Targets

Four measured predictors were used: northing and easting (m), regolith thickness (m) and fissured-layer thickness (m). These predictors do not all become available at the same moment. The coordinates and the regolith thickness can be estimated at a candidate site before drilling, from geophysics or from neighbouring boreholes; the thickness of fissured rock is the thickness actually crossed, and it becomes available, metre by metre, while the borehole is being drilled. The models of this paper are therefore specified at the drilling stage, where all three are known, and we refer to this as the operational predictor set. A prospective set restricted to the coordinates and the regolith thickness is also reported, in Sections 4.3 and 5.2, to quantify what is gained by waiting for the fissured thickness and to justify the stage at which the prediction is positioned. A fifth, derived predictor was added: the product of regolith and fissured thicknesses (m^2), which the original spreadsheet stored under the name "hydro-potential" as a spreadsheet formula. Because it is entirely determined by two measured predictors, it is treated here as a derived index and its contribution is tested explicitly (Section 4.5). Two targets were modelled: the final depth (m), by regression, and the yield class, by classification into four use-based classes defined by the thresholds applied in rural water supply: negative (<0.7 m^3/h, 838 boreholes, 34.5%), low (0.7 - 2.5 m^3/h, hand pump, 850,

35.0%), medium (2.5 - 5 m^3/h, small piped scheme, 323, 13.3%) and high (≥5 m^3/h, multi-village scheme, 421, 17.3%). Continuous yield was also tested as a regression target and is reported for completeness.

3. Methods

3.1. Modelling Framework

The workflow comprised 1) descriptive analysis and correlation of the predictors and targets; 2) training of a random-forest regressor for depth and of a random-forest classifier for yield class; 3) validation by three complementary protocols; 4) benchmarking of three alternative algorithms on identical folds; 5) ablation of predictor groups; and 6) analysis of variable importance and of the empirical relationships between layer thicknesses and drilling success. All computations were performed with scikit-learn 1.x [17] in Python 3, with fixed random seeds; the same pipeline is implemented in JavaScript in the companion tool HydroPrédict, which reproduces the analysis in a web browser and generates this report's tables and figures from any database with the same columns.

3.2. Random Forest

Random forest [11] averages the predictions of decorrelated regression or classification trees, each grown on a bootstrap sample of the training set with a random subset of predictors considered at every split. It handles non-linear and interacting effects without distributional assumptions, is robust to outliers in predictors and provides an impurity-based measure of variable importance. Forests of 100 trees were used, with unlimited depth, a minimum of one observation per leaf, all predictors candidate at each split (the scikit-learn default for regression) and a fixed seed (**Table 2**). For the classifier, class weights inversely proportional to class frequencies were applied to counter the imbalance between the medium class (13.3%) and the two dominant classes; the classifier outputs a probability for each class, the predicted class being the most probable. The probability of a positive borehole is the complement of the probability of the negative class.

Table 2. Hyper-parameters of the models (scikit-learn implementation).

Model	Setting	Value
Random forest (regression and classification)	Number of trees	100
	Minimum samples per leaf	1
	Maximum depth	Unlimited
	Predictors per split	All (regression); √p (classification)
	Class weights (classifier)	Balanced (1/class frequency)
	Random seed	42 (+ fold index in cross-validation)

Continued

Extra trees	Number of trees/split rule	100/random thresholds, no bootstrap
Gradient boosting	Trees/depth/learning rate/subsample	300 (regr.), 200 (class.)/3/0.1/0.8
Linear/logistic regression	Standardisation/regularisation	z-scores/L2 (logistic), balanced class weights
Validation	Hold-out	80%/20%, stratified by random shuffling
	Random k-fold	10 folds, shuffled
	Spatial block k-fold	6 × 6 grid of ~18 × 18 km blocks, 10 folds of 3 - 4 non-contiguous blocks

3.3. Validation Design

Three protocols were applied. 1) A hold-out split: 80% of the boreholes (1945) for training and 20% (487) for testing, after random shuffling. 2) A 10-fold random cross-validation: the database was shuffled and divided into ten subsets of ~243 boreholes; each subset served once for testing while the other nine served for training. 3) A 10-fold spatial block cross-validation [15]: the study area was divided into a 6 × 6 grid of ~18 × 18 km blocks and the 36 blocks were assigned to ten folds so that each fold groups three or four non-contiguous blocks. Because block boundaries cut through the borehole network, grouping non-contiguous blocks does not by itself guarantee that a test borehole has no close neighbour in the training set, and the separation actually achieved was therefore measured for every protocol (**Figure 4**). Under random folds, 828 of the 2432 test boreholes have a training borehole at the same coordinates, 893 within 100 m and 1263 within 500 m, so the random protocol is optimistic by construction. Under spatial blocks the minimum separation is 164 m and only 10 test boreholes lie within 500 m of a training borehole. To remove this residual proximity, a buffer was added: training boreholes within a given distance of any test borehole of the fold are discarded. A 2 km buffer raises the minimum train-test separation to 2002 m and the fifth percentile to 2327 m, at the cost of 2% of the training data (2142 boreholes per fold on average, against 2188 without buffer); buffers of 1 and 5 km were also tested. All spatial results reported below are given with and without the 2 km buffer. For cross-validation, metrics were computed both on the pooled out-of-fold predictions of the whole database (n = 2432) and as the mean of the per-fold values; the two differ when a fold contains extreme values, which is the case here for the two boreholes deeper than 150 m.

3.4. Performance Metrics

Regression performance was assessed with the mean absolute error (MAE), the root-mean-square error (RMSE), the coefficient of determination $R^2 = 1 - SSE/SST$, the Pearson correlation coefficient R, and the indicators used in hydrological model evaluation: the Nash-Sutcliffe efficiency NSE [18], which is algebraically identical to R^2 as defined above; the Kling-Gupta efficiency KGE = 1 − [(r −

$1)^2 + (\alpha - 1)^2 + (\beta - 1)^2]^{0.5}$, where α and β are the ratios of the standard deviations and of the means of predicted and observed values [19]; Willmott's index of agreement d [20]; the percent bias PBIAS; and the RMSE-to-observations standard-deviation ratio RSR. Performance classes follow Moriasi *et al.* [21]: NSE > 0.75, RSR ≤ 0.50 and |PBIAS| < 10% for a "very good" model. Classification performance was assessed with overall accuracy, balanced accuracy (mean of per-class recalls, 0.25 at chance for four classes), per-class recall, precision and one-vs-rest area under the ROC curve (AUC, 0.50 at chance), and the macro-averaged AUC; confusion matrices are reported on pooled out-of-fold predictions.

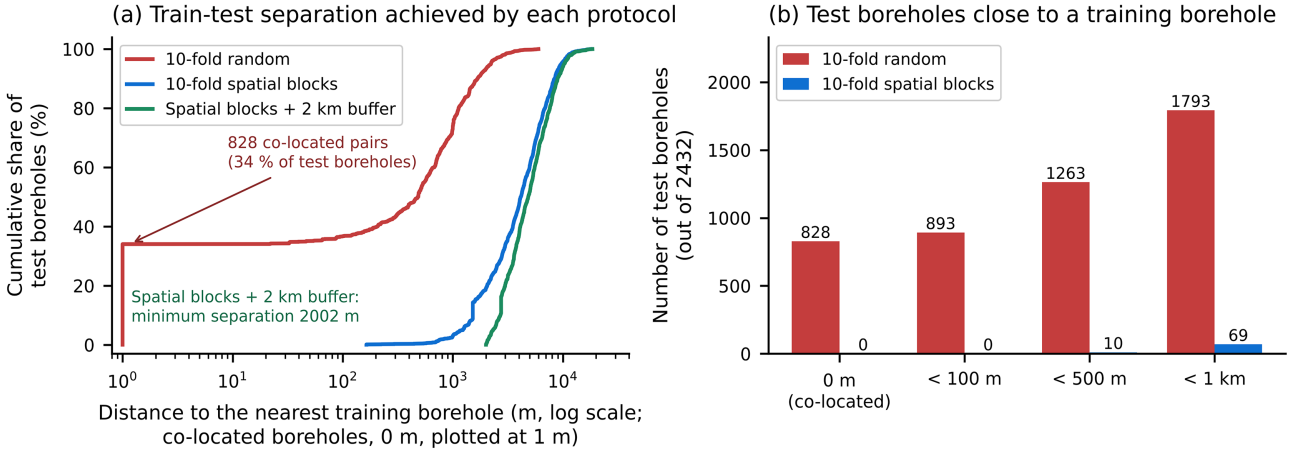


Figure 4. Distance from each test borehole to the nearest training borehole under the three validation protocols: (a) Cumulative distribution (logarithmic scale); (b) Number of test boreholes located at the same coordinates as, or within 100 m, 500 m and 1 km of, a training borehole. Random folds leave 828 co-located pairs; spatial blocks leave none.

3.5. Benchmark Algorithms and Ablation

Three alternatives were trained on exactly the same folds as the random forest: extra trees [22], which draw split thresholds at random and use no bootstrap; gradient boosting [23], which fits shallow trees sequentially to the residuals of the previous ones; and, as linear references, multiple linear regression for depth and one-vs-rest logistic regression for yield classes, both on standardised predictors. Settings are listed in **Table 2**. An ablation then re-trained the random forest on subsets of predictors, all five; the four measured predictors without the index; the two thicknesses only; the two coordinates only; the index only; and the prospective set (position and regolith thickness), which is the only subset available before drilling, to isolate the information carried by each group.

3.6. Variable Importance and Thickness-Success Relationships

Variable importance was computed as the normalised mean decrease in impurity over all trees of the forests trained on the whole database. Because the index is a deterministic function of the two thicknesses, importance is shared between correlated predictors and is interpreted jointly with the ablation. Finally, the empirical relationship between each layer thickness and drilling success was examined

by computing the success rate (yield ≥ 0.7 m^3/h) and the median yield in three thickness classes (<20 m, 20 - 40 m, ≥40 m), the classes used by Iglesias-Martínez *et al.* [7].

4. Results

4.1. Correlation Structure

Depth is correlated with fissured-layer thickness (r = 0.85), with the derived index (r = 0.73) and only weakly with regolith thickness (r = 0.13) and with the coordinates ($|r| \leq 0.17$); its correlation with the sum of the two thicknesses is 1.000. Yield, in contrast, is correlated with none of the predictors: r = 0.10 with regolith thickness, −0.17 with fissured thickness, −0.11 with the index and $|r| \leq 0.11$ with the coordinates. Regolith represents on average 46% of the final depth.

4.2. Prediction of Borehole Depth

On the hold-out test set (n = 487), the random forest reproduced depth with R^2 = 0.994, MAE = 0.42 m and RMSE = 1.04 m; NSE = 0.994, KGE = 0.980, d = 0.999, PBIAS = −0.20% and RSR = 0.077 all fall in the "very good" class of Moriasi *et al.* [21] (**Table 3**, **Figure 5(a)**). In 10-fold random cross-validation the fold-mean R^2 was 0.945 and the fold-mean MAE 0.59 m; nine folds exceeded R^2 = 0.98 and the fold containing the 672 m borehole fell to 0.490 (**Figure 6(a)**). On pooled out-of-fold predictions the R^2 dropped to 0.785 and the RMSE rose to 9.50 m, entirely because of the two boreholes deeper than 150 m, which no tree ensemble can extrapolate: excluding them, pooled R^2 = 0.995 and RMSE = 0.99 m (**Figure 5(b)**). Spatial block cross-validation gave almost identical results, fold-mean R^2 0.934, fold-mean MAE 0.84 m, pooled R^2 0.756 (**Figure 5(c)**, **Figure 6(b)**), showing that the fit transfers to areas not represented in training, and the 2 km buffer leaves it unchanged (pooled R^2 0.778, fold-mean 0.939, MAE 0.65 m).

Table 3. Performance of the random-forest depth model under the three validation protocols. Pooled: metrics on all out-of-fold predictions; fold-mean: average of per-fold values. Verdicts follow Moriasi *et al.* [21]. The last four rows give the prospective predictor set (position and regolith thickness only) and the buffered spatial protocol, both added in revision.

Metric	Hold-out (20%)	10-fold random, pooled	10-fold random, fold-mean	10-fold spatial, pooled	10-fold spatial, fold-mean
n	487	2432	~243	2432	~243
MAE (m)	0.42	0.59	0.59	0.64	0.84
RMSE (m)	1.04	9.50	—	10.13	—
R^2 = NSE	0.994	0.785	0.945	0.756	0.934
R (Pearson)	0.997	0.890	—	0.873	—
KGE	0.980	0.780	—	0.754	—
Index of agreement d	0.999	0.931	—	0.919	—

Continued

PBIAS (%)	−0.20	−0.53	—	−0.55	—
RSR	0.077	0.463	—	0.494	—
Verdict (NSE)	Very good	Very good	Very good	Very good	Very good
Prospective set, R^2 = NSE	—	−0.113	−0.216	−0.277	−0.479
Prospective set, MAE (m)	—	11.34	—	12.08	—
Operational set, spatial + 2 km buffer, R^2	—	—	—	0.778	0.939
Operational set, spatial + 2 km buffer, MAE (m)	—	—	—	0.65	—

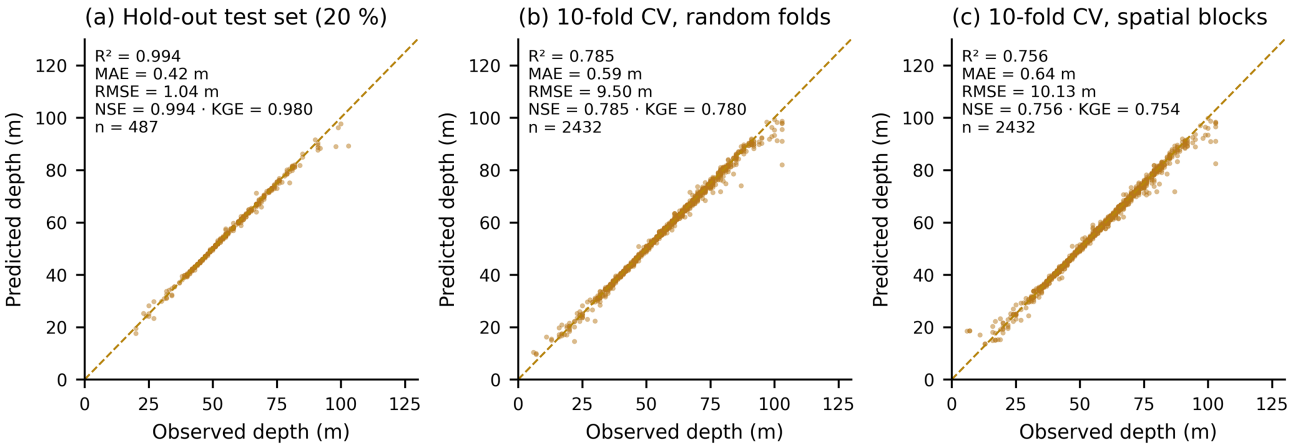


Figure 5. Observed versus predicted borehole depth for the random forest: (a) Hold-out test set; (b) Pooled out-of-fold predictions, 10-fold random cross-validation; (c) Pooled out-of-fold predictions, 10-fold spatial block cross-validation. Axes are truncated at 130 m; the two boreholes deeper than 150 m lie outside the frame.

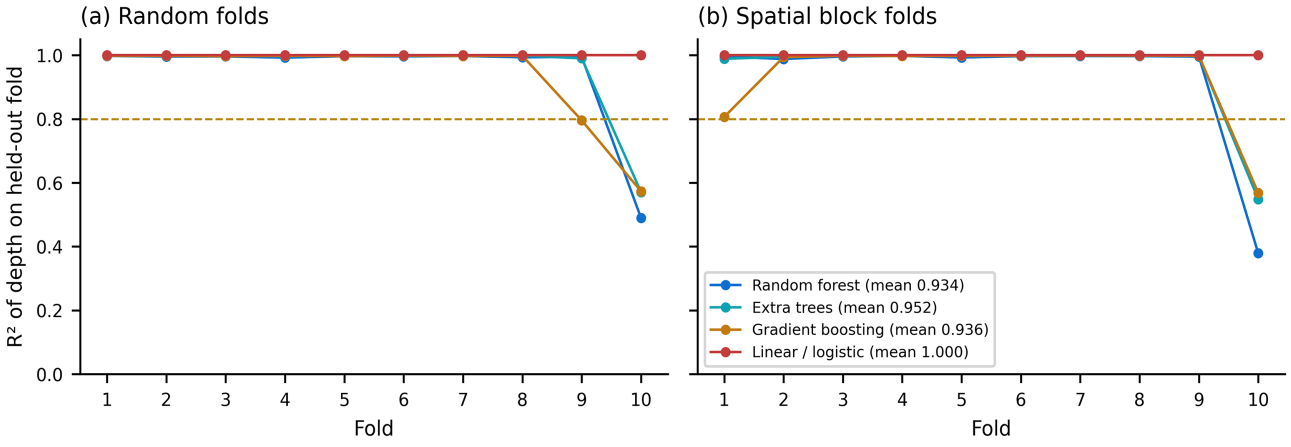


Figure 6. R^2 of depth on each held-out fold for the four algorithms under (a) random and (b) spatial block cross-validation. The dashed line marks R^2 = 0.80.

The relation behind these figures is exact rather than statistical. Recorded depth equals the sum of the recorded regolith and fissured thicknesses for every one of

the 2432 boreholes, with a maximum absolute deviation of 0.0 m, so the model is a calibration of the stopping rule applied in the field: the borehole is stopped at the base of the fissured layer crossed, a few metres below the last water strike. Two practical consequences follow. The error on a predicted depth equals the error on the sum of the two estimated thicknesses, so a depth estimate to within 5 m requires thicknesses estimated to within 5 m, a target that electrical sounding or resistivity profiling can be tested against. And the fissured thickness is a necessary input: with the prospective set alone the same model returns a negative R^2 (Table 3), which is why the depth relation is used at the drilling stage, where the thickness crossed is logged as the rig advances, and as a specification for pre-drilling geophysics rather than as a stand-alone forecast.

4.3. Prediction of the Yield Class at the Drilling Stage

As a continuous variable, yield was not predictable: in 10-fold random cross-validation the random forest reached a pooled R^2 of 0.042 (fold-mean 0.029) with an MAE of 3.03 m^3/h for a median yield of 1.20 m^3/h, *i.e.* no better than predicting the mean, consistent with the absence of correlation reported in Section 4.1. As classes, yield was partly predictable (Table 4, Figure 7). On pooled random-fold predictions the random forest reached an accuracy of 48.3%, against 35.0% for always predicting the majority class, and a balanced accuracy of 40.5% (chance 25%). Discrimination was strongest for the negative class (recall 71%, AUC 0.82) and the high class (AUC 0.73), and weakest for the intermediate classes (recall 15% and AUC 0.64 for medium; AUC 0.60 for low), which are mostly confused with their neighbours (Figure 7(a)). The macro-AUC was 0.70. Hold-out results were similar (balanced accuracy 41.8%, macro-AUC 0.69) and spatial block validation lowered them only marginally (balanced accuracy 39.5%, macro-AUC 0.69), indicating that the modest skill of the classifier is at least geographically stable.

Table 4. Yield-class performance of the random-forest classifier (pooled out-of-fold predictions). Recall and one-vs-rest AUC are given per class: Neg. = negative (<0.7 m^3/h), Low (0.7 - 2.5), Med. (2.5 - 5), High (≥5). The operational set, used at the drilling stage, comprises the position, the regolith thickness, the thickness of fissured rock crossed and the derived index; the prospective set, restricted to position and regolith thickness, is reported to quantify the information that drilling itself provides.

Predictor set and protocol	Accuracy	Balanced acc.	Recall Neg.	Recall Low	Recall Med.	Recall High	AUC Neg.	AUC Low	AUC Med.	AUC High	Macro-AUC
Operational set											
Hold-out (20%)	49.1%	41.8%	73%	49%	12%	34%	0.80	0.58	0.65	0.73	0.69
10-fold random	48.3%	40.5%	71%	49%	15%	27%	0.82	0.60	0.64	0.73	0.70
10-fold spatial	48.4%	39.5%	69%	54%	8%	27%	0.81	0.59	0.67	0.70	0.69
10-fold spatial + 2 km buffer	47.7%	38.9%	69%	53%	11%	23%	0.81	0.59	0.67	0.71	0.69
Prospective set											

Continued

10-fold random	36.9%	31.3%	50%	41%	13%	22%	0.63	0.52	0.56	0.64	0.59
10-fold spatial	35.1%	29.2%	47%	40%	8%	22%	0.62	0.50	0.56	0.61	0.57
10-fold spatial + 2 km buffer	35.3%	29.0%	45%	45%	9%	18%	0.62	0.49	0.57	0.60	0.57
Chance/majority	35.0%	25%	—	—	—	—	0.50	0.50	0.50	0.50	0.50

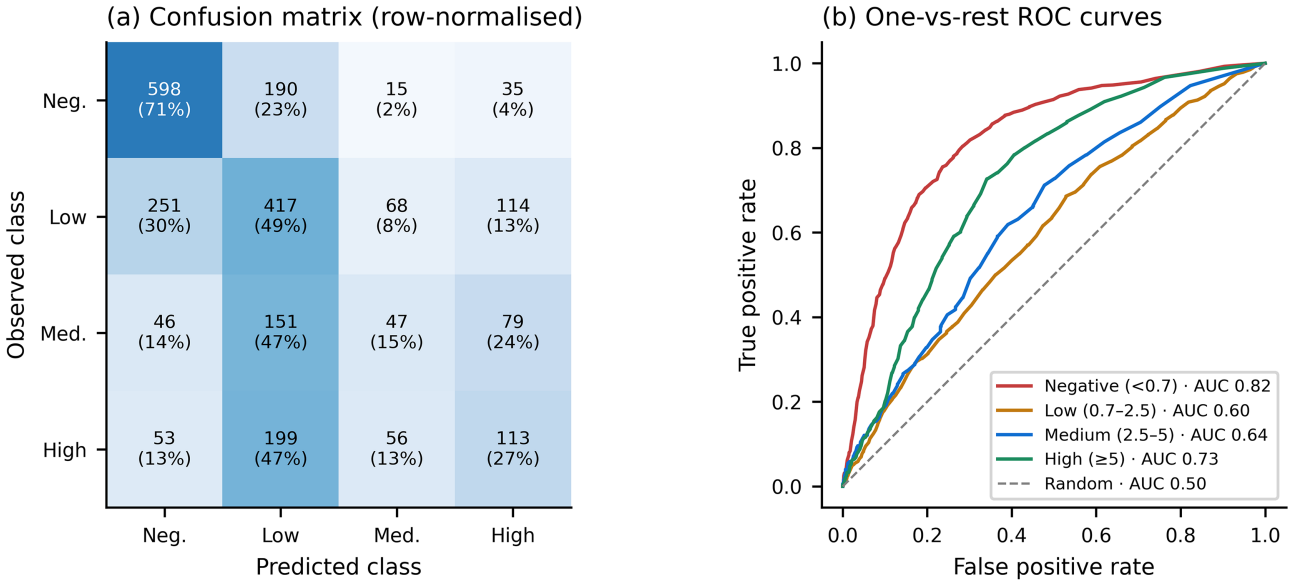


Figure 7. Yield-class classifier, random forest, pooled out-of-fold predictions of the 10-fold random cross-validation: (a) Confusion matrix with row percentages (recall on the diagonal); (b) One-vs-rest ROC curves and AUC per class.

The choice of the drilling stage is what makes this performance available. Restricted to the prospective set, that is to position and regolith thickness alone, the same classifier returns a balanced accuracy of 31.3% and a macro-AUC of 0.59 under random folds, and 29.0% and 0.57 under spatial blocks with the 2 km buffer, the AUC of the negative class falling from 0.81 to 0.62 (Table 4, Figure 8). The ten points of balanced accuracy and the 0.19 of AUC that separate the two sets are carried by the thickness of fissured rock crossed, and they are gained as soon as drilling begins. The prediction is therefore specified where the information is: the siting decision is supported by the regional gradient of success with regolith thickness (Section 4.6), and the borehole-level prediction is made and updated while the rig is on site.

4.4. Benchmark of Algorithms

Table 5 and Figure 9 compare the four algorithms on identical folds. For depth, multiple linear regression reproduced the target exactly (R^2 = 1.000, MAE = 0.00 m under both protocols), because recorded depth is the sum of two predictors; among tree ensembles, extra trees performed best (fold-mean R^2 0.955, MAE 0.37 m), followed by random forest and gradient boosting, all three being limited only

by their inability to extrapolate beyond the training range (Figure 6). For yield classes, the three ensembles were equivalent within a few tenths of a percent, balanced accuracy 39.8% - 41.4%, macro-AUC 0.69 - 0.71, and logistic regression was only slightly behind (balanced accuracy 40.6%, macro-AUC 0.68). The convergence of methods of very different flexibility toward the same ceiling indicates that the limit is set by the information content of the predictors, not by the form of the model.

Figure 10 compares the observed and simulated depths of the four algorithms on the pooled out-of-fold predictions. The three tree ensembles cluster tightly along the 1:1 line up to about 100 m and saturate beyond the deepest values seen in training, extra trees showing the narrowest scatter (MAE 0.37 m) and gradient boosting the widest (0.68 m); the linear model follows the diagonal exactly over the whole range.

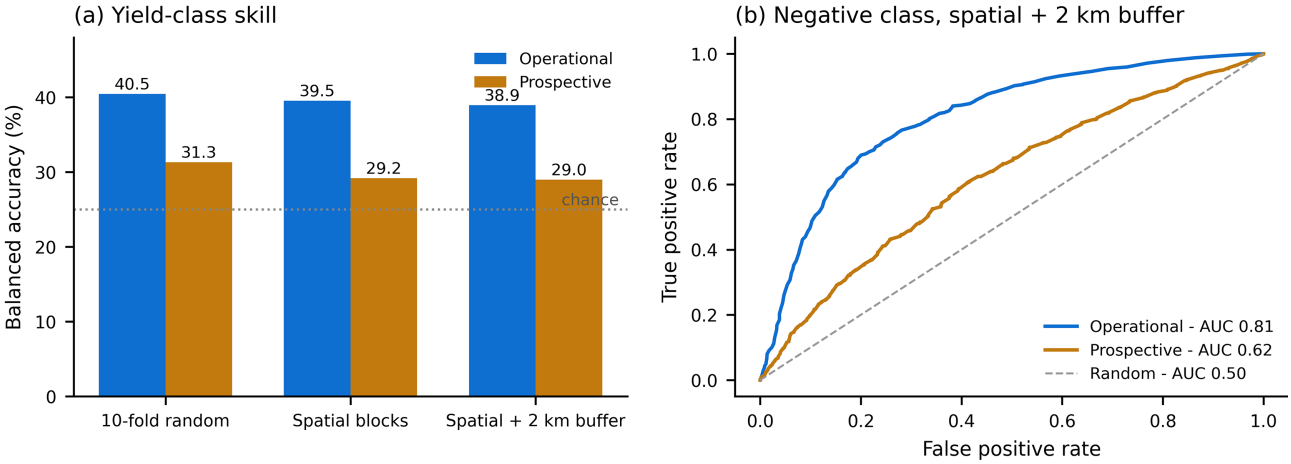


Figure 8. Information gained at the drilling stage: (a) Balanced accuracy of the four yield classes for the operational and prospective predictor sets under the three validation protocols (dotted line: chance level, 25%); (b) One-vs-rest ROC curves of the negative class under spatial block validation with a 2 km buffer.

Table 5. Benchmark of the four algorithms on identical folds (pooled out-of-fold predictions; fold-mean R^2 in parentheses).

Algorithm	Protocol	Depth R^2 (fold-mean)	Depth MAE (m)	Depth KGE	Class accuracy	Balanced accuracy	Macro-AUC
Random forest	Random	0.785 (0.945)	0.59	0.780	48.3%	40.5%	0.70
	Spatial	0.756 (0.934)	0.64	0.754	48.4%	39.5%	0.69
Extra trees	Random	0.819 (0.955)	0.37	0.774	47.8%	39.8%	0.69
	Spatial	0.822 (0.952)	0.39	0.775	47.1%	38.6%	0.69
Gradient boosting	Random	0.781 (0.935)	0.68	0.864	50.0%	41.4%	0.71
	Spatial	0.795 (0.936)	0.68	0.866	46.1%	37.5%	0.68
Multiple linear/logistic regression	Random	1.000 (1.000)	0.00	1.000	43.5%	40.6%	0.68
	Spatial	1.000 (1.000)	0.00	1.000	44.0%	41.1%	0.67

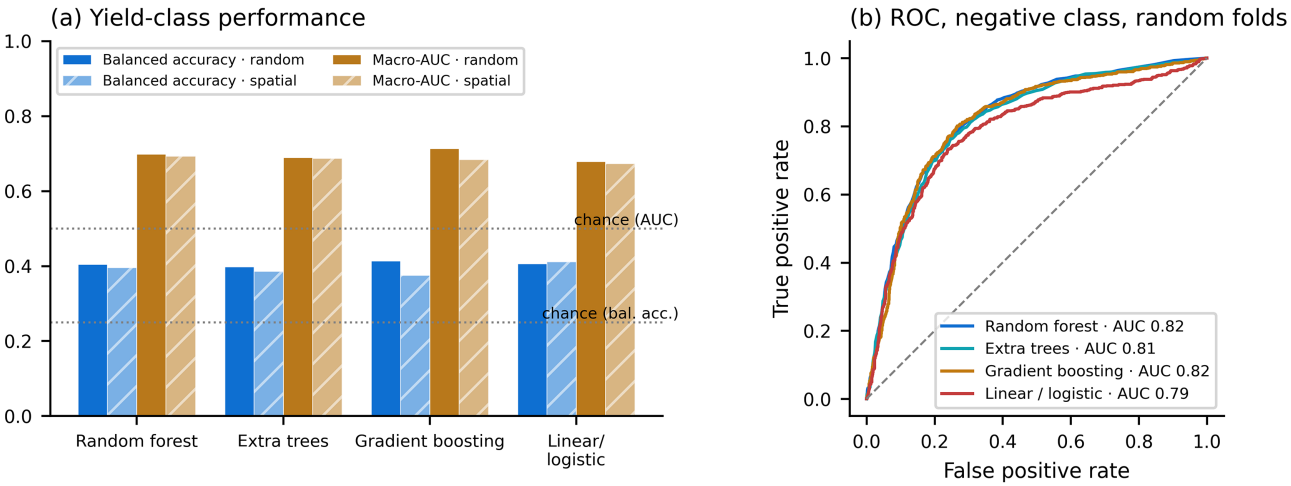


Figure 9. Benchmark of algorithms for yield classes: (a) Balanced accuracy and macro-AUC under random and spatial block cross-validation (dotted lines: chance levels); (b) One-vs-rest ROC curves of the negative class, random folds.

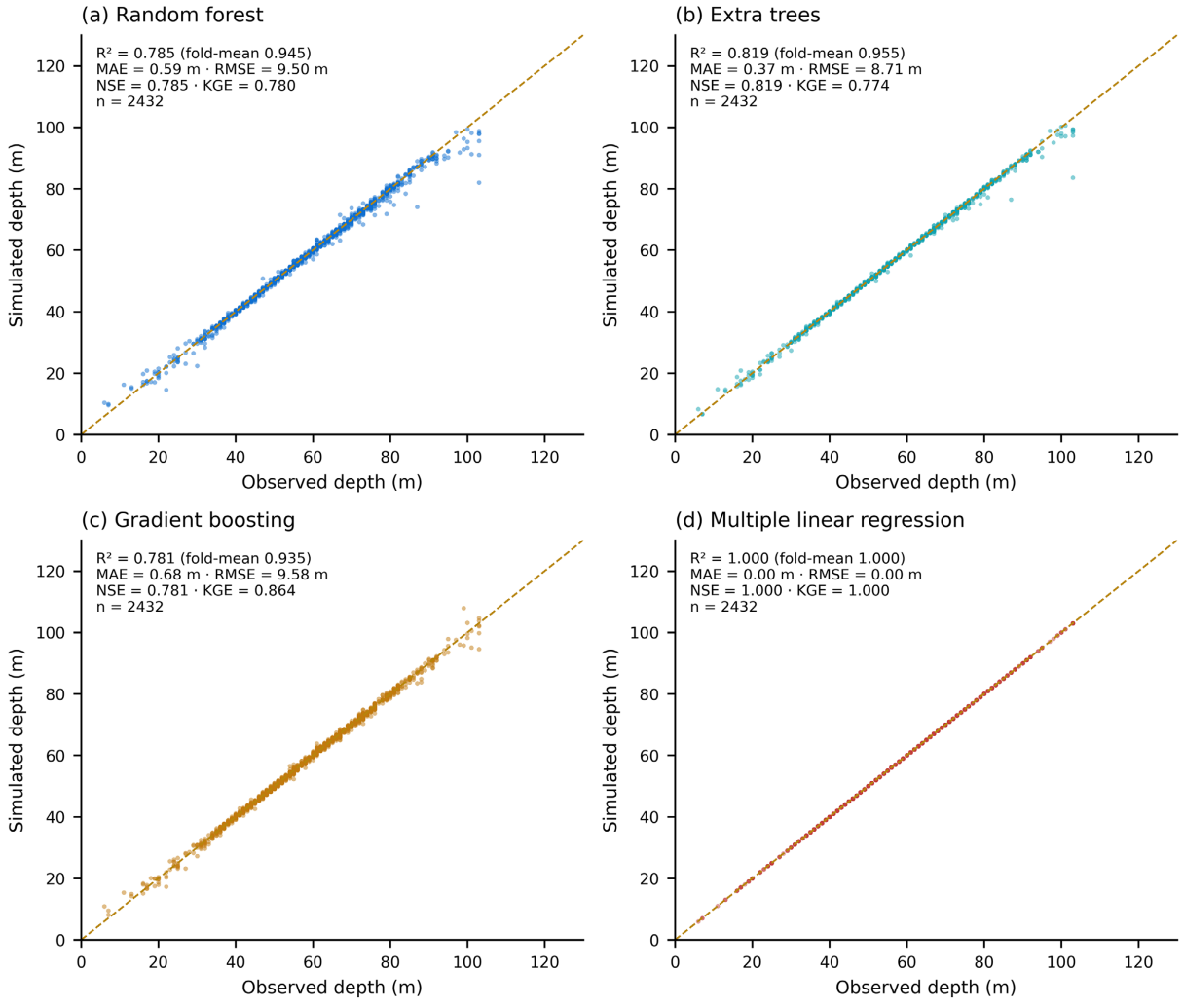


Figure 10. Observed versus simulated borehole depth for the four benchmarked algorithms, pooled out-of-fold predictions of the 10-fold random cross-validation: (a) Random forest, (b) Extra trees, (c) Gradient boosting, (d) Multiple linear regression. Axes truncated at 130 m.

4.5. Variable Importance and Ablation

In the depth regressor, importance was shared almost equally between the fissured thickness (47.3%) and the derived index (47.3%), regolith thickness accounting for 4.7% and the coordinates for less than 1% (Figure 11(a)). In the yield-class classifier, importance was spread over all predictors, with the coordinates together weighing 44% against 56% for the thicknesses and the index (Figure 11(b)), *i.e.* a zone effect not captured by the thickness variables. The ablation (Table 6) resolves the apparent weight of the index. Removing it did not degrade depth prediction (pooled R^2 0.863 against 0.785 with it; the pooled value even improved because the index, with values up to 13,926 m^2, amplifies the extrapolation error on the deepest boreholes) nor yield classification (balanced accuracy 41.3% against 40.5%). The two thicknesses alone carried the whole depth information (R^2 0.864) and most of the class information (balanced accuracy 35.7%); the coordinates alone carried no depth information (R^2 −0.16) but a small class information (balanced accuracy 32.7%, macro-AUC 0.59) that adds to that of the thicknesses; the index alone explained 40% of depth variance and almost nothing of the classes.

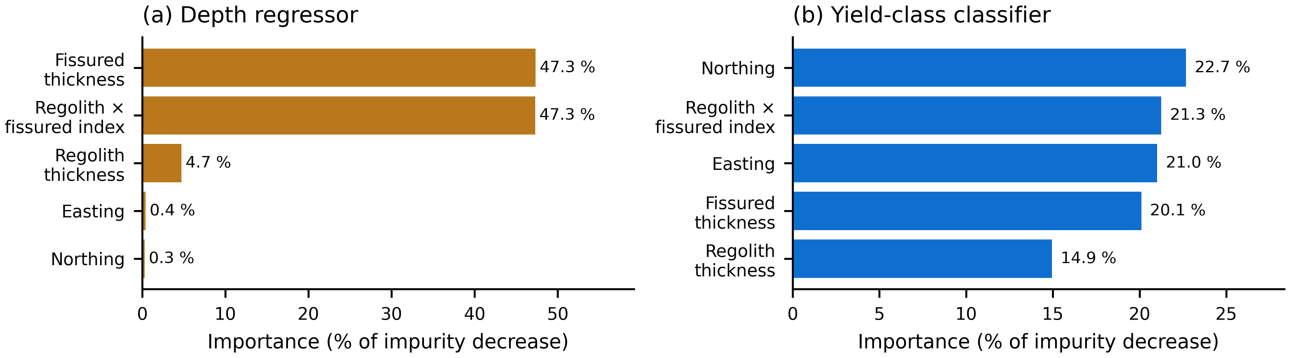


Figure 11. Normalised impurity-based variable importance of (a) The depth regressor and (b) The yield-class classifier, random forests trained on the whole database.

Table 6. Ablation of predictor groups, random forest, 10-fold random cross-validation (pooled out-of-fold predictions).

Predictors	Depth R^2	Depth MAE (m)	Balanced accuracy	Macro-AUC
All five (coordinates, thicknesses, index)	0.785	0.59	40.5%	0.70
Four measured (without index)	0.863	0.65	41.3%	0.70
Two thicknesses only	0.864	0.50	35.7%	0.63
Coordinates only	−0.16	11.73	32.7%	0.59
Index only	0.398	5.15	31.4%	0.58
Position + regolith thickness (prospective)	−0.113	11.34	31.3%	0.59

4.6. Layer Thickness and Drilling Success

The success rate rose monotonically with regolith thickness, from 56% for covers thinner than 20 m (n = 829, median yield 0.87 m^3/h) to 69% for 20 - 40 m (n =

1313, 1.40 m³/h) and 77% beyond 40 m (n = 290, 1.70 m³/h) (Figure 12(a)). The relationship with recorded fissured thickness was the reverse and much stronger: 88% success (median 2.60 m³/h) below 20 m, 75% (1.40 m³/h) between 20 and 40 m and 34% (0.20 m³/h) beyond 40 m (Figure 12(b)). This inverse relationship is the single most discriminant pattern in the database and is what the classifier exploits to recognise negative boreholes.

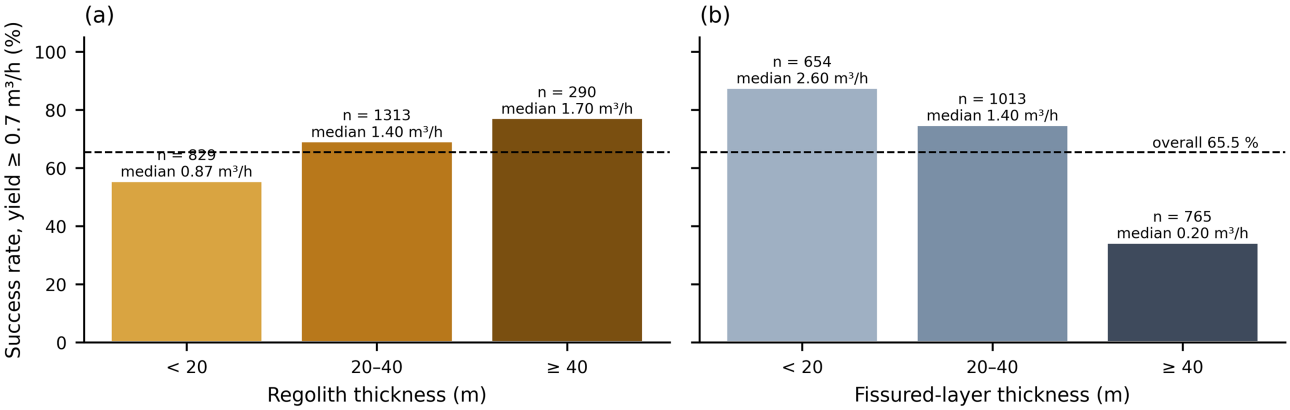


Figure 12. Success rate (yield ≥ 0.7 m³/h) and median yield by class of (a) regolith thickness and (b) recorded fissured-layer thickness. The dashed line is the overall success rate.

5. Discussion

5.1. Depth Prediction: A Drilling Stopping Rule Consistent with the Conceptual Model

The depth model is an exact relation calibrated on 2432 boreholes rather than a statistical fit: recorded depth is the sum of the two recorded thicknesses, so what the model stores is the rule applied in the field, namely that the borehole is stopped at the base of the fissured layer crossed, a few metres below the last water strike. This is not an artefact but the numerical expression of the stratiform conceptual model of hard-rock aquifers [2] [3], in which the fissured layer is the conductive compartment the borehole must fully penetrate, and of the field observation that its transmissivity concentrates in the upper 20 - 40 m below the regolith [4]. The share of regolith in total depth in our database (46% on average) is close to the 44% reported by Kouamé *et al.* [4] for 148 boreholes of northern Côte d'Ivoire, and our median fissured thickness of 30 m lies in the productive interval they identified. The rule is usable in two ways. During drilling, it converts the thickness already crossed into the depth at which the borehole will be stopped if the fissured layer is fully penetrated. Before drilling, it transfers the uncertainty of the estimated thicknesses one for one onto the estimated depth, which makes the accuracy of geophysical thickness estimation, rather than the choice of algorithm, the quantity to improve; a validation of that route on seven electrical resistivity tomography profiles of the same area with known borehole outcomes is in progress and will be reported separately.

Two consequences follow. First, tree ensembles are not the best tool for this

target: multiple linear regression reproduces the identity exactly and extrapolates, whereas forests saturate at the deepest training values (**Figure 5(b)**, **Figure 5(c)**; **Figure 6**; **Figure 10**). Second, the division of labour between the two targets becomes clear: depth is governed by an exact relation and needs measurement, not learning, whereas yield needs learning but is limited by the predictors available, which is where the modelling effort belongs.

Earlier Ivorian work pursued the same operational goal with different predictors. Kouadio *et al.* [10], on 214 boreholes of the Denguélé region, derived from principal-component analysis and multiple regression an optimisation equation of the stopping depth based on drilling parameters, advancement rate, depth of the first water strike and number of strikes, achieving 90% success and about 8% saving on drilled depth; statistical models of the depth limit of drilling in Ivorian crystalline rocks were later proposed by Kouassi *et al.* [24]. Those models share with ours the use of quantities observed during drilling, and the same caution applies to their interpretation; the difference is that they were explicitly designed as during-drilling decision rules, which is the status we now give to our operational model. Combined with the regional thickness maps of Courtois *et al.* [5], an independent, pre-drilling estimate of the two thicknesses could be assembled and the depth relation tested prospectively, which is the natural continuation of this work.

5.2. Outcome Prediction at the Drilling Stage

Yield in hard-rock aquifers is a heavy-tailed variable controlled by local fracture connectivity, lithology and landscape position [8] [9], which two thickness measurements and a position cannot resolve as a continuous quantity; the correlations of $|r| \leq 0.17$ observed here are of the same order as those commonly reported. The reformulation into use-based classes, following Gómez-Escalonilla *et al.* [14], makes the problem tractable, and specifying it at the drilling stage makes it useful: the classifier separates negative boreholes with an AUC of 0.81 and a recall of 69% under buffered spatial validation, which is the decision that matters most once the rig is on site. The two predictor sets correspond to two moments of the same operation rather than to a valid and an invalid model, and the ten points of balanced accuracy between them measure the value of the information that drilling itself produces. The gap with the Malian results (balanced accuracy ≈ 0.80, AUC up to 0.87) measures the information contained in the lithological, structural and geomorphological predictors used there and absent here; the convergence of random forest, extra trees, gradient boosting and logistic regression toward the same ceiling (**Table 5**) shows that this gap cannot be closed by changing the algorithm. The stability of the scores under buffered spatial validation (38.9% against 39.5% without buffer) shows that the skill is not an artefact of spatial autocorrelation [15] and that the model transfers to sectors not represented in training; the random protocol, by contrast, is optimistic, since 828 of its test boreholes share their coordinates with a training borehole (**Figure 4**).

The role of regolith thickness in success is consistent with recent West African

evidence. Adeotan *et al.* [6], on 2347 Burkinabe boreholes, found the thickness of the weathered layers to be the best predictor of success and lineaments to be uninformative; Iglesias-Martínez *et al.* [7] reported yields three times higher under covers thicker than 40 m. Our database shows the same gradient (56% to 77% success, median yield 0.87 to 1.70 m^3/h), which supports the capacitive role of the regolith feeding the fissured layer [2]. The effect is graded, however, and cannot classify an individual borehole, which explains the low recalls of the intermediate classes.

5.3. Censoring in Operational Databases

The inverse relationship between recorded fissured thickness and success (**Figure 12(b)**) deserves emphasis because it is easily misread as a hydrogeological property. It is, in our interpretation, the signature of censoring: when water is not found, the driller continues into the fissured layer within the budgeted depth, so that negative boreholes accumulate the largest recorded fissured thicknesses, whereas successful boreholes are stopped early. The recorded thickness is therefore an outcome-dependent measurement, not the geological thickness of the layer, and the classifier's ability to recognise negative boreholes partly reflects this: it reads in the fissured thickness the trace of an unsuccessful search for water. A recent analysis of weathered-basement borehole databases [25] reached a similar conclusion, showing that such databases underestimate aquifer productivity because drilling stops as soon as a sufficient yield is obtained and failures are unevenly reported. Two methodological consequences follow: depth models trained on such data reproduce a practice, and yield models are trained on a censored predictor. The size of the second effect is measured in Section 4.3: admitting the drilled fissured thickness raises the balanced accuracy of the yield-class model from 29.0% to 38.9% and the AUC of the negative class from 0.62 to 0.81. Because the predictor encodes the outcome, that increment cannot be transferred to a site that has not been drilled, and we now report the two models separately rather than a single figure. Distinguishing the geological thickness of the fissured layer, from geophysics or from the thickness above the last water strike, from the thickness drilled is a prerequisite for a genuinely predictive model, and separating regolith from saprock, as recommended by Adeotan *et al.* [6], would refine the capacitive term.

5.4. The Regolith × Fissured Index

The product of regolith and fissured thicknesses, stored in the source spreadsheet as a "hydro-potential", combines the capacitive and conductive compartments of the conceptual model and has an evident physical appeal. The results nevertheless show that it carries no independent information: the ablation gives identical or better scores without it (**Table 6**), its apparent importance in the depth regressor is a sharing of splits with the fissured thickness from which it is computed, and it even degrades the pooled depth R^2 by amplifying extrapolation error on extreme

boreholes. Such indices should be presented as derived variables and evaluated by ablation before being credited with a predictive role; the same test applies to the many composite "potential" indices of groundwater potential mapping.

5.5. Implications for Borehole Siting and Drilling

For practitioners, the results translate into a decision supported at two moments, each with its measured accuracy. Before drilling, the regional gradient established here is the robust signal: success rises from 56% to 77% as the regolith thickens from under 20 m to over 40 m, which supports the targeting of thick-cover sectors at the scale of a programme, and the expected thicknesses at a candidate site give a depth and cost estimate whose uncertainty is that of the thickness estimate itself. Once drilling has begun, the thickness of fissured rock already crossed is entered as it is logged and the model returns an updated probability for each yield class: among the boreholes of this database in which more than 40 m of fissured rock was crossed, 66% ended negative, against 12% among those in which less than 20 m was crossed, which is a quantified basis for a stop-or-continue decision and the strongest predictive use the database supports. The companion tool implements both steps and reports the probability with its inter-tree interval rather than a categorical answer. The mapped probabilities also make a zone effect explicit: the coordinates carry class information that the thickness variables do not, pointing to lithological or structural domains that a geological map or a lineament density layer would capture directly.

5.6. Limitations and Further Work

Five limitations bound the scope of this study. The predictors are limited to position and two thicknesses; adding lithology, distance to and density of lineaments, static water level, depth of the first water strike and terrain descriptors is the condition for a yield classifier comparable to that of Gómez-Escalonilla *et al.* [14]. The depth target is the final depth; the depth of the first water strike, the decision variable of the driller [10], would be a more defensible target. The fissured thickness is a drilled, not a geological, thickness, and the regolith is not split into saprolite and saprock. Validation, although spatially blocked and buffered, remains internal to one database; an external test on boreholes drilled after its compilation is the next step. Finally, the provenance of the compilation is documented only in part: the number of records available before the complete-case selection, the missingness by variable and the yield-test procedure applied to each record are held by the compiling institutions and were not available to us, and any residual heterogeneity of the yield measurement would act on the class assignment near the 0.7, 2.5 and 5 m^3/h thresholds. The companion tool was designed to accept these additional variables without modification.

6. Conclusion

A random forest trained on 2432 boreholes of the crystalline basement within the

Ouagadougou map sheet of Burkina Faso reproduces the final depth of a borehole from its position and the thicknesses of the regolith and fissured layers with R^2 = 0.994 on a hold-out set and a mean absolute error below 0.7 m. Recorded depth is the exact sum of the two recorded thicknesses, so this relation is a calibrated stopping rule, in agreement with the stratiform conceptual model of hard-rock aquifers: it converts the prediction of depth into the estimation of two thicknesses, with an error that transfers one for one, and it sets the accuracy that pre-drilling geophysics must reach. The outcome of a borehole is predicted at the drilling stage, where the thickness of fissured rock crossed is known: in four use-based classes the classifier reaches an accuracy of 48.3%, a balanced accuracy of 40.5% and a macro-AUC of 0.70, with an AUC of 0.82 and a recall of 71% for the negative class, and it keeps 38.9% and 0.69 under spatial block cross-validation with a 2 km buffer, which establishes its transferability to sectors not represented in training. Its most useful expression is the stop-or-continue reading: 66% of the boreholes that crossed more than 40 m of fissured rock ended negative, against 12% of those that crossed less than 20 m. The regolith × fissured index adds no information beyond its two components. The increase in success with regolith thickness confirms recent regional findings, while its decrease with recorded fissured thickness reveals the censoring inherent in operational databases, which must be accounted for before such data are used to infer aquifer properties. Extending the prediction to the siting stage with comparable accuracy requires lithological and structural predictors and the geological, rather than drilled, thickness of the fissured layer; the database, the validation design and the companion tool assembled here provide the baseline against which such a model would have to be judged.

Acknowledgements

The authors thank the Direction Générale des Ressources en Eau et d'Hydraulique (DGREH), the Bureau des Mines et de la Géologie du Burkina (BUMIGEB), the Ouvrages de Mobilisation des Eaux Souterraines (OMES) programme and the Bureau de Recherches Géologiques et Minières (BRGM) for access to the borehole data used in this study.

Author Contributions

Conceptualization, methodology, software, formal analysis and writing of the original draft, R.A.-K.N.; data curation, validation, writing, review and editing, H.N.A.; supervision, writing, review and editing, J.N. All authors have read and agreed to the published version of the manuscript.

Data Availability

The borehole database analysed in this study was compiled from the drilling reports held by the Direction Générale des Ressources en Eau et d'Hydraulique (DGREH) of Burkina Faso, the Bureau des Mines et de la Géologie du Burkina (BUMIGEB), the Ouvrages de Mobilisation des Eaux Souterraines (OMES) pro-

gramme and the Bureau de Recherches Géologiques et Minières (BRGM). The compilation is not publicly available. The subset of 2432 complete records analysed here, restricted to the seven variables used in this study (UTM coordinates, regolith thickness, fissured-layer thickness, end-of-development yield and final depth), is available from the corresponding author on reasonable request and with the permission of these institutions. The companion tool HydroPrédict, which reproduces the whole analysis in a web browser from any spreadsheet with the same columns, is provided as Supplementary Material.

Supplementary Materials

The supplementary material associated with this article includes the per-fold metrics of all algorithms, the confusion matrices of the benchmark models, the continuous-yield regression results and the companion tool, together with the train-test separation statistics of the three validation protocols.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] MacDonald, A.M., Bonsor, H.C., Dochartaigh, B.É.Ó. and Taylor, R.G. (2012) Quantitative Maps of Groundwater Resources in Africa. *Environmental Research Letters*, **7**, Article 024009. https://doi.org/10.1088/1748-9326/7/2/024009

[2] Dewandel, B., Lachassagne, P., Wyns, R., Maréchal, J.C. and Krishnamurthy, N.S. (2006) A Generalized 3-D Geological and Hydrogeological Conceptual Model of Granite Aquifers Controlled by Single or Multiphase Weathering. *Journal of Hydrology*, **330**, 260-284. https://doi.org/10.1016/j.jhydrol.2006.03.026

[3] Lachassagne, P., Wyns, R. and Dewandel, B. (2011) The Fracture Permeability of Hard Rock Aquifers Is Due Neither to Tectonics, Nor to Unloading, but to Weathering Processes. *Terra Nova*, **23**, 145-161. https://doi.org/10.1111/j.1365-3121.2011.00998.x

[4] Kouamé, I.K., Douagui, A.G., Bouatrin, D.K., Kouadio, S.K.A. and Savané, I. (2021) Assessing the Hydrodynamic Properties of the Fissured Layer of Granitoid Aquifers in the Tchologo Region (Northern Côte d'Ivoire). *Heliyon*, **7**, e07620. https://doi.org/10.1016/j.heliyon.2021.e07620

[5] Courtois, N., Lachassagne, P., Wyns, R., Blanchin, R., Bougaïré, F.D., Somé, S., *et al.* (2010) Large-Scale Mapping of Hard-Rock Aquifer Properties Applied to Burkina Faso. *Groundwater*, **48**, 269-283. https://doi.org/10.1111/j.1745-6584.2009.00620.x

[6] Adeotan, Y.S.O., Vouillamoz, J.-M., Lawson, F.M.A., Koïta, M., Zorom, M., Traoré, Y.S., *et al.* (2025) Beyond Lineaments: Exploring the Influence of Weathered Layers on Borehole Productivity in Basement Rocks of Burkina Faso, West Africa. *Hydrogeology Journal*, **33**, 411-424. https://doi.org/10.1007/s10040-025-02876-5

[7] Iglesias-Martínez, M., Alcalá, F.J. and Rull, M. (2024) Predicting Well Productivity in Cratonic Regions Using Remote Sensing Lineaments and Weathered Cover Thickness: A Case Study from Bobo Dioulasso, Burkina Faso. *Journal of African Earth Sciences*, **220**, Article 105461. https://doi.org/10.1016/j.jafrearsci.2024.105461

[8] Douagui, A.G., Kouadio, S.K.A., Mangoua, J.O.M., Kouassi, A.K., Kouamé, B.K. and

Savané, I. (2019) Using Specific Capacity for Assessing of the Factors Controlling Borehole Productivity in Crystalline Bedrock Aquifers of N'zi, Iffou and Moronou Regions in the Eastern Area of Côte d'Ivoire. *Groundwater for Sustainable Development*, **9**, Article 100235. https://doi.org/10.1016/j.gsd.2019.100235

[9] Kouamé, B.K., Douagui, A.G., Kouamé, I.K., Yeo, E.W. and Savané, I. (2019) Borehole Productivity Controlling Factors in Crystalline Bedrock Aquifer of Gbêkê Region, Center of Côte d'Ivoire. *Journal of Water Resource and Protection*, **11**, 728-739. https://doi.org/10.4236/jwarp.2019.116043

[10] Kouadio, E.K., Nagnin, S. and Savané, I. (2010) Stratégie d'optimisation de la profondeur des forages en contexte de socle: Application à la région du Denguélé, Nord-Ouest de la Côte d'Ivoire. *Revue des sciences de l'eau*, **23**, 1-15. https://doi.org/10.7202/038921ar

[11] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. https://doi.org/10.1023/A:1010933404324

[12] Naghibi, S.A., Pourghasemi, H.R. and Dixon, B. (2016) GIS-Based Groundwater Potential Mapping Using Boosted Regression Tree, Classification and Regression Tree, and Random Forest Machine Learning Models in Iran. *Environmental Monitoring and Assessment*, **188**, Article No. 44. https://doi.org/10.1007/s10661-015-5049-6

[13] Golkarian, A., Naghibi, S.A., Kalantar, B. and Pradhan, B. (2018) Groundwater Potential Mapping Using C5.0, Random Forest, and Multivariate Adaptive Regression Spline Models in GIS. *Environmental Monitoring and Assessment*, **190**, Article No. 149. https://doi.org/10.1007/s10661-018-6507-8

[14] Gómez-Escalonilla, V., Diancoumba, O., Traoré, D.Y., Montero, E., Martín-Loeches, M. and Martínez-Santos, P. (2022) Multiclass Spatial Predictions of Borehole Yield in Southern Mali by Means of Machine Learning Classifiers. *Journal of Hydrology: Regional Studies*, **44**, Article 101245. https://doi.org/10.1016/j.ejrh.2022.101245

[15] Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., *et al*. (2017) Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure. *Ecography*, **40**, 913-929. https://doi.org/10.1111/ecog.02881

[16] Castaing, C., Le Métour, J., Billa, M., Chèvremont, P., Donzeau, M., Zida, B., Ouédraogo, I., Koté, S., Ouédraogo, O.F. and Egal, E. (2003) Notice explicative de la carte géologique du Burkina Faso à 1/200 000. Bureau de Recherches Géologiques et Minières.

[17] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., *et al*. (2011) Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*, **12**, 2825-2830. https://www.jmlr.org/papers/v12/pedregosa11a.html

[18] Nash, J.E. and Sutcliffe, J.V. (1970) River Flow Forecasting through Conceptual Models Part I—A Discussion of Principles. *Journal of Hydrology*, **10**, 282-290. https://doi.org/10.1016/0022-1694(70)90255-6

[19] Gupta, H.V., Kling, H., Yilmaz, K.K. and Martinez, G.F. (2009) Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling. *Journal of Hydrology*, **377**, 80-91. https://doi.org/10.1016/j.jhydrol.2009.08.003

[20] Willmott, C.J. (1981) On the Validation of Models. *Physical Geography*, **2**, 184-194. https://doi.org/10.1080/02723646.1981.10642213

[21] Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D. and Veith, T.L. (2007) Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, **50**, 885-900. https://doi.org/10.13031/2013.23153

[22] Geurts, P., Ernst, D. and Wehenkel, L. (2006) Extremely Randomized Trees. *Machine Learning*, **63**, 3-42. https://doi.org/10.1007/s10994-006-6226-1

[23] Friedman, J.H. (2001) Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, **29**, 1189-1232. https://doi.org/10.1214/aos/1013203451

[24] Kouassi, A.M., Kouakou, K.E., Kadio, K.K.R. and Biemi, J. (2014) Design of Statistical Models Using Hydrogeological Variables for Predicting the Depth Limit Stop of Water Drilling in Crystalline and Crystallophyllian Rocks in Ivory Coast. *International Journal of Innovation and Scientific Research*, **11**, 181-197.

[25] Kinglo, A.A., Hector, B., Galiez, C., Koïta, M., Lachassagne, P., Lawson, F. and Vouillamoz, J.-M. (2026) Are There Biases in Borehole Databases of Weathered Basement Aquifers Affecting Their Reliability to Estimate Aquifer Productivity? *Hydrogeology Journal*, **34**, 497-519. https://doi.org/10.1007/s10040-025-02991-3