

Prédiction de la résistance à la compression du matériau terre crue à l'aide de l'intelligence artificielle

Zidane Rigasse TCHOUPOU SEGNING^{1,2}, Philippe POULLAIN¹, Beudelair SAHA TCHINDA², Nabil ISSAADI¹, Aurelle TCHAGNA KOUANOU³, Stéphanie BONNET¹, Daniel TCHIOTSOP²

¹ GeM, UMR 6183 CNRS, Nantes Université, Centrale Nantes, F-44600 Saint-Nazaire, France.

² Université de Dschang, IUT Fotso Victor de Bandjoun, UR-AIA, Rue Dja'a BP134 Bandjoun, Cameroun

³ College of Technology, University of Buea, BP63 Buea, Cameroun

RESUME Dans un contexte de transition vers des matériaux de construction plus durables, la terre crue suscite un regain d'intérêt en raison de ses qualités environnementales. Cependant, sa faible résistance en compression demeure un obstacle majeur à son utilisation structurelle. Cette étude explore l'usage de l'intelligence artificielle pour prédire cette propriété à partir des paramètres géotechniques du sol mesurés en laboratoire, des paramètres de formulation et la contrainte de compactage. Un modèle de Gradient Boosting a été entraîné sur une base de données comprenant 135 échantillons de briques de terre crue, stabilisées ou non. Après une optimisation des hyperparamètres, le modèle prédit avec une racine carrée de l'erreur quadratique moyenne de 0,38 MPa, une erreur absolue moyenne de 0,25 MPa et un coefficient de détermination de 0,984 sur les données de test. L'analyse de l'importance des variables révèle que la résistance à la compression est principalement influencée par le pourcentage massique de liant, la contrainte de compactage et le pourcentage massique d'eau du mélange. Ces résultats montrent l'efficacité du Gradient Boosting dans la caractérisation mécanique des matériaux terreux à faible impact environnemental.

Mots-clés Terre crue, résistance à la compression, intelligence artificielle, Gradient Boosting, optimisation des hyperparamètres

I. INTRODUCTION

Dans un contexte mondial marqué par l'épuisement progressif des ressources naturelles, l'urbanisation croissante et les émissions massives de gaz à effet de serre, les matériaux de construction conventionnels comme le béton et l'acier posent aujourd'hui de sérieux défis environnementaux (Morel et al., 2001). Leur fabrication est énergivore, fortement émettrice de CO₂ et dépendante de ressources non renouvelables, ce qui remet en question leur durabilité à long terme. En particulier, la production du ciment représente environ 8 % des émissions anthropiques mondiales de CO₂ (Habert et al., 2020; Morel et al., 2001). Face à ces limitations, un intérêt croissant se manifeste pour des alternatives plus responsables, notamment les matériaux issus de ressources locales (Morel et al., 2001). Parmi eux, la terre crue, matériau ancestral utilisé depuis des millénaires, suscite un regain d'intérêt (Hamard et al., 2016). Utilisée sous forme de briques de terre comprimée (BTC), pisé, adobe ou torchis, la terre présente de nombreux avantages écologiques : faible énergie grise, recyclabilité, disponibilité locale, faible émission de CO₂ et bon confort thermique (Medjelekh et al., 2016; Zami and Lee, 2010). Cependant, malgré ses nombreux atouts, l'usage de la terre reste limité par des performances mécaniques modestes,

notamment en compression. Des études rapportent que les briques de terre non stabilisée affichent des valeurs généralement comprises entre 0,5 et 5 MPa (Miccoli et al., 2014; Thompson et al., 2025).

Pour améliorer ces propriétés, diverses techniques de stabilisation sont couramment mises en œuvre. La stabilisation chimique, par adjonction de liants tels que le ciment ou la chaux, permet la formation de produits de réaction comme les silicates ou aluminates de calcium hydratés, qui renforcent la structure interne du matériau, mais au prix d'un impact environnemental accru (Taallah et al., 2014). La stabilisation physique ou mécanique, telle que le compactage sous haute pression, augmente la densité et réduit la porosité (Taallah et al., 2014). L'incorporation de fibres naturelles (paille, chanvre, sisal, etc.) a également été explorée afin d'améliorer la résistance à la rupture, la ductilité et de limiter la fissuration (Poullain et al., 2019). Cette technique s'inscrit toutefois dans une pratique ancienne, déjà mise en œuvre depuis longtemps dans de nombreuses cultures de construction en terre (Danso et al., 2015). L'efficacité de cette approche dépend néanmoins fortement des propriétés des fibres (nature, longueur, dosage, résistance). Par ailleurs, les propriétés mécaniques dépendent aussi de la composition du sol et des conditions de mise en œuvre. Les caractéristiques géotechniques, telles que la granulométrie, les teneurs en argile, limon, sable et gravier, ainsi que les limites d'Atterberg, jouent un rôle central dans la cohésion du matériau (Houben and Guillaud, 1994). Un excès de plasticité peut entraîner un retrait excessif au séchage, tandis qu'un sol trop sableux présente généralement une cohésion insuffisante. La teneur en eau au moment du moulage, proche de l'optimum Proctor, conditionne également la densification et, par conséquent, les performances finales. La grande variabilité des sols et des techniques de formulation rend difficile une évaluation systématique de la résistance à la compression de la terre crue. Traditionnellement, celle-ci est déterminée par des essais expérimentaux nécessitant des équipements spécifiques et des protocoles longs et coûteux. Pour pallier ces contraintes, plusieurs approches ont été développées : modèles analytiques (Burroughs, 2009), équations constitutives (Yan et al., 2022), méthodes statistiques basées sur les fonctions de densité de probabilité (Poullain et al., 2023; Tchiotso et al., 2022), ainsi que des techniques d'intelligence artificielle (Houcine et al., 2023; Mustafa et al., 2025; Ngo et al., 2025; Turco et al., 2021). En ce qui concerne ces dernières approches, plusieurs algorithmes ont été utilisés pour estimer la résistance des briques à partir de données expérimentales. Ces modèles incluent les réseaux de neurones artificiels (ANN) (Houcine et al., 2023; Turco et al., 2021), les arbres de décision, les machines à vecteurs de support (SVM) (Mustafa et al., 2025), ainsi que des modèles d'ensemble tels que le Random Forest, l'Adaboost et le Gradient Boosting (Ngo et al., 2025). Bien que ces méthodes aient montré leur efficacité, de nombreuses études se limitent typiquement à 5 à 9 à un nombre restreint de variables d'entrée, ou à un seul type de stabilisant (ciment, chaux, etc.), négligeant parfois des facteurs importants tels que la contrainte de compactage, le type de fibre. Ces limites peuvent nuire à la capacité des modèles à généraliser correctement.

Dans ce contexte, l'objectif de cette étude est de développer et d'évaluer un modèle d'apprentissage automatique, le Gradient Boosting, pour prédire la résistance en compression des briques de terre crue, à partir de données issues de la littérature, en tenant compte des caractéristiques géotechniques du sol, des paramètres de formulation et de la contrainte de compactage.

II. METHODOLOGIE

II.1. Présentation et analyse exploratoire de la base de données

Les modèles de machine learning nécessitent des données fiables et représentatives afin de prédire avec précision la résistance à la compression des briques de terre. Une base de données initiale comprenant

494 échantillons issus de 21 études a été constituée à partir de la littérature. Pour cette étude, un sous-ensemble de 135 échantillons provenant de six publications (Danso et al., 2015; Malonga et al., 2024a, 2024b; Poullain et al., 2019; Sadouri et al., 2024; Taallah et al., 2014) a été retenu, sélectionnés en raison de la richesse des paramètres mesurés. La base de données finale comprend 18 variables, dont une variable cible (la résistance à la compression, RC) et 17 variables d'entrée. Ces dernières couvrent les caractéristiques géotechniques du sol (teneurs en gravier, sable, limon, argile : G, S, Lm, A ; limites d'Atterberg : LP, LL), les paramètres de formulation du matériau (teneurs en eau, liant, fibres, sable ajouté : W, L, F, Sa, toutes en pourcentage massique du mélange sec), ainsi que les conditions de compactage (CC). Trois variables sont de nature catégorielle : le type de liant (T_L), le type de fibre (T_f) et le type de malaxage (T_m), comprenant respectivement 7, 7 et 2 modalités, c'est-à-dire des valeurs qualitatives exprimées sous forme de catégories. Une analyse statistique descriptive a été menée sur les variables numériques afin d'évaluer leur distribution et leur variabilité. Il est à noter que les valeurs nulles associées aux propriétés des fibres correspondent à l'absence de fibres dans les échantillons concernés. Le tableau 1 présente les statistiques clés (moyenne, écart-type, minimum, maximum) de ces variables numériques, mettant en évidence la diversité des formulations étudiées. La figure 1 illustre la distribution des variables catégorielles ainsi que celle de la teneur en liant et de la résistance à la compression, la teneur en liant étant présentée en raison de son rôle clé dans la formulation.

TABLEAU 1. Statistiques descriptives des variables

Variables (Unité)	Notation	Moyenne	Std	Min	Max
Teneur en gravier (%)	G	6,16	8,72	0,00	24,80
Teneur en sable (%)	S	52,71	12,63	27,70	68,50
Teneur en limon (%)	Lm	24,45	6,88	13,00	38,4
Teneur en argile (%)	A	16,69	5,36	9,10	30,00
Limite de plasticité (%)	LP	20,98	4,36	14,00	27,30
Limite de liquidité (%)	LL	33,84	8,32	18,60	51,20
Longueur moyenne de la fibre (mm)	L_f	24,12	23,07	0,00	100,00
Diamètre moyen de la fibre (mm)	D_f	0,54	0,84	0,00	3,00
Résistance à la traction de la fibre (MPa)	Rt_f	172,95	245,15	0,00	960,00
% massique d'eau du mélange (%)	W	13,65	3,19	10,00	19,02
% massique de sable ajouté au mélange (%)	Sa	8,23	12,73	0,00	28,5
% massique de liant (%)	L	4,98	4,51	0,00	15,00
% massique de fibre (%)	F	0,32	0,42	0,00	1,50
Contrainte de compactage (MPa)	CC	8,24	2,43	1,50	10,00
Résistance à la compression (MPa)	RC	4,99	2,93	0,83	12,7

Par ailleurs, la matrice de corrélation de Pearson présentée à la figure 2 permet d'identifier les relations linéaires entre les différentes variables. Il en ressort que la résistance à la compression est positivement corrélée à la teneur en liant (L, 0,69) et à la teneur en sable ajouté (Sa, 0,53). En revanche, des corrélations négatives sont observées avec la teneur en eau (W, -0,59) et la teneur en fibres (F, -0,50). Ces quatre paramètres pourraient avoir une forte influence sur la prédiction. D'autres variables, telles que la contrainte de compactage (CC), les teneurs en sable et en argile (A), ou encore les limites d'Atterberg (LP, LL), présentent des corrélations plus faibles, indiquant une influence indirecte ou combinée. On observe également une forte corrélation négative entre la teneur en sable et celle en gravier (G, -0,93),

traduisant leur interdépendance dans la composition granulométrique du sol, puisque les fractions sable, gravier, argile et limon doivent totaliser 100%, ainsi qu'une corrélation positive notable entre la teneur en sable et le diamètre moyen des fibres (D_f , 0,70) qui pourrait résulter d'un biais lié au nombre restreint d'études ayant alimenté la base de données. Ces fortes corrélations peuvent introduire des effets de colinéarité susceptibles de dégrader les performances prédictives des modèles. L'utilisation d'algorithmes robustes face à la colinéarité, tels que les modèles d'ensemble, s'avère donc essentielle dans ce contexte (Mustafa et al., 2025).

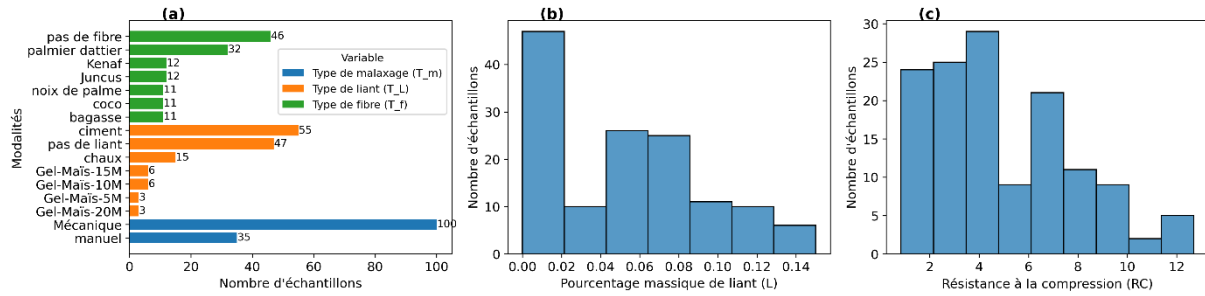


FIGURE 1. Distribution des variables : (a) variables catégorielles, (b) pourcentage massique de liant (L), (c) résistance à la compression (RC). Les notations 5M, 10M, 15M et 20M indiquent les % d'amidon dans le gel

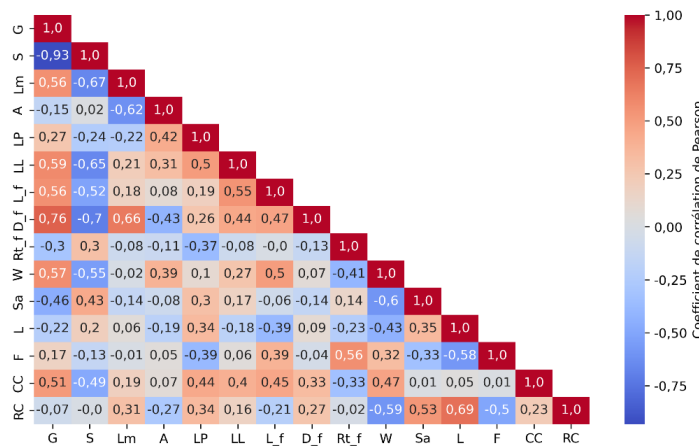


FIGURE 2. Matrice de corrélation de Pearson entre les différentes variables de la base de données

II.2. Division et prétraitement des données

Le jeu de données a été mélangé aléatoirement, puis divisé en deux ensembles : un ensemble de test représentant 20 % des données, soit 27 échantillons, utilisé pour l'évaluation finale des modèles, et un ensemble d'entraînement constitué des 80 % restants. Afin d'évaluer la robustesse du modèle et de limiter les biais liés à une unique partition des données, une validation croisée stratifiée à k plis, avec $k = 6$, a été appliquée sur l'ensemble d'entraînement. La moyenne des performances obtenues sur les six plis a été retenue comme score de validation final du modèle. La stratégie de validation croisée à 6 plis repose sur un découpage de l'ensemble d'entraînement en 6 sous-ensembles. À chaque itération, un sous-ensemble est utilisé pour la validation, tandis que les autres servent à l'entraînement. Les performances moyennes sur l'ensemble des plis permettent une évaluation plus robuste et moins biaisée du modèle (Pedregosa et al., 2011).

Pour utiliser efficacement les données en entrée des modèles, certains traitements préalables ont été nécessaires, notamment l'encodage des variables catégorielles et la normalisation des variables numériques. Dans cette étude, les variables catégorielles ont été encodées à l'aide de la méthode One

Hot Encoding de la bibliothèque scikit-learn (Pedregosa et al., 2011), qui consiste à transformer chaque variable en autant de colonnes binaires que de modalités, chaque colonne valant 1 si la modalité est présente et 0 sinon. La normalisation, quant à elle, a consisté à mettre à l'échelle les variables numériques selon l'équation (1). Les méthodes StandardScaler et MinMaxScaler de la bibliothèque scikit-learn, ont été testées et ont conduit à des performances similaires. Toutefois puisqu'elles ne peuvent être appliquées simultanément, seule l'équation du StandardScaler est présentée ici.

$$x_i^{(scaler)} = \frac{x_i - \mu}{\sigma} \quad (1)$$

où x_i est la valeur de la variable x pour l'observation i , μ est la valeur moyenne de la variable x , σ est l'écart-type de la variable x et $x_i^{(scaler)}$ est la valeur transformée.

II.3. Présentation du modèle Gradient Boosting

Le Gradient Boosting est un algorithme d'ensemble puissant, applicable aussi bien aux tâches de classification qu'à celles de régression. Il cherche à approximer une fonction de prédiction $F(x)$ qui relie les variables d'entrées x à la sortie y en construisant de manière séquentielle un modèle $F_m(x)$. Le processus débute par l'entraînement d'un modèle initial $F_0(x)$ directement sur la cible y . Ensuite, à chaque itération, un apprenant faible — généralement un arbre de décision peu profond, qui divise itérativement les données en sous-groupes homogènes selon les variables les plus discriminantes — est ajouté afin de corriger les erreurs des prédictions précédentes. Ce nouvel arbre est entraîné pour minimiser une fonction de perte, telle que l'erreur quadratique moyenne dans le cas de la régression (fonction de perte utilisée dans cette étude), conformément à l'équation (2). Plus précisément, il est ajusté aux gradients négatifs de la fonction de perte par rapport aux prédictions actuelles, définis à l'équation (3). La contribution de chaque nouvel arbre est régulée par un facteur d'apprentissage (learning rate) et le modèle final est obtenu par une somme pondérée des arbres successifs, comme indiqué à l'équation (4) :

$$h_m = \underset{h}{\operatorname{argmin}} \sum_{i=1}^n l(y_i, F_{m-1}(x_i) + h_m(x_i)) \quad (2)$$

$$g_i^m = - \left[\frac{\partial l(y_i, F_m(x_i))}{\partial F_m(x_i)} \right]_{F=F_{m-1}} \quad (3)$$

$$F_m(x) = F_{m-1}(x) + \nu \cdot h_m(x) \quad (4)$$

où F_m est le modèle après m itérations, F_{m-1} est le modèle après $m - 1$ itérations, h_m est le $m - i$ ème apprenant faible, ν est le pas d'apprentissage, g_i^m est le gradient négatif du $i - i$ ème échantillon de la fonction de perte l après $m - 1$ itérations, x est la matrice d'entrée et y_i est la $i - i$ ème cible.

Lors de l'apprentissage, des hyperparamètres structurels comme `max_features`, qui définit le nombre maximal de variables considérées pour chaque division d'un arbre, peuvent également être ajustés. Ce paramètre permet de contrôler la diversité entre les arbres et de réduire le surapprentissage en limitant leur complexité. Ce cadre théorique a été introduit par Friedman (Friedman, 2001).

II.4. Métriques d'évaluation du modèle

L'évaluation des performances du modèle a été réalisée à l'aide de trois métriques couramment utilisées en régression : la racine carrée de l'erreur quadratique moyenne (RMSE), l'erreur Absolue moyenne (MAE) et le coefficient de détermination (R^2). Ces indicateurs permettent de quantifier à la fois la précision globale des prédictions et leur écart moyen par rapport aux valeurs mesurées (Mustafa et al.,

2025; Ngo et al., 2025; Turco et al., 2021). RMSE de même unité que la cible, mesure la racine carrée de la moyenne des carrés des écart entre les valeurs prédites \hat{y}_i et réelles y_i . Elle pénalise les grandes erreurs. L'erreur absolue moyenne quant à elle correspond à l'erreur moyenne en valeur absolue entre les prédictions et les valeurs réelles. Le coefficient de détermination mesure la proportion de la variance des données qui est expliquée par le modèle. Un coefficient de détermination très proche de 1 traduit une forte capacité prédictive, il varie entre 0 et 1. Leur expression mathématique est présentée par les équations (5), (6) et (7)

$$RMSE = \sqrt{MSE} = \frac{1}{n} \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \frac{1}{n} \sum_{i=1}^n y_i)^2} \quad (7)$$

II.5. Optimisation des hyperparamètres du modèle

L'efficacité d'un modèle de Gradient Boosting repose non seulement sur son principe d'apprentissage séquentiel, mais aussi sur le choix approprié de ses hyperparamètres, qui influence directement sa complexité et ses performances. Pour ce faire, une recherche exhaustive par grille (GridSearchCV) a été utilisée, telle qu'implémentée dans la bibliothèque scikit-learn (Pedregosa et al., 2011). Cette méthode consiste à évaluer l'ensemble des combinaisons possibles de paramètres à l'aide d'une validation croisée, et à sélectionner celle minimisant une fonction de perte, ici l'erreur quadratique moyenne (MSE) sur les données de validation. Les plages testées étaient : `n_estimators` de 10 à 350 avec un pas de 10 (valeur initiale =100), `max_depth` de 1 à 5 (valeur initiale =3) et `max_features` de 0,5 à 1,0 avec un pas de 0,1 (valeur initiale =1,0). Les paramètres optimaux, sont présentés à la figure 2.

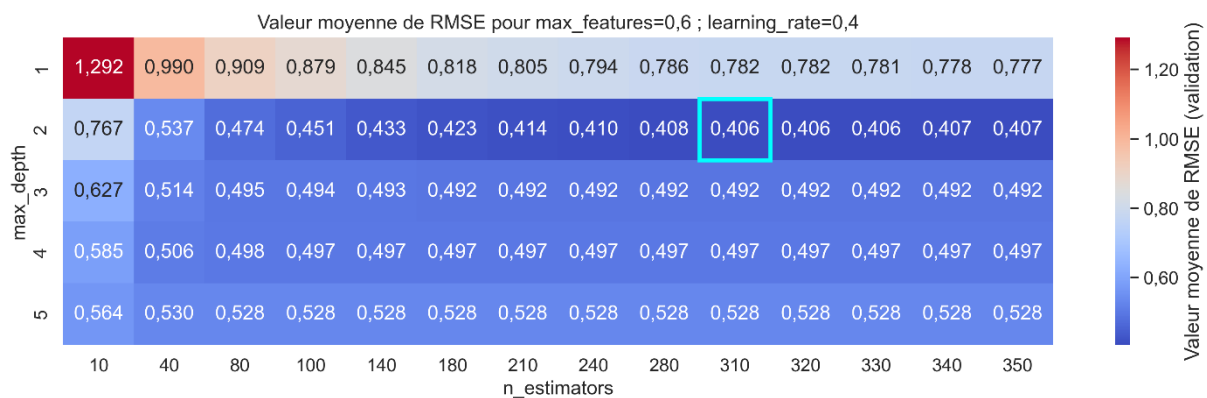


FIGURE 3. Impact des hyperparamètres sur la performance du modèle en validation, la zone encadrée indique la meilleure performance

III. RESULTATS ET DISCUSSION

III.1. Évaluation globale des performances

Le tableau 2 présente les résultats obtenus sur les ensembles d'entraînement, de validation et de test. Le modèle optimisé présente une excellente capacité de prédiction avec un R^2 de 0,999 sur l'entraînement, 0,976 sur la validation et 0,984 sur le test. Ces valeurs indiquent que le modèle explique près de 98% de la variance des données sur l'ensemble de validation et de test et quasiment 100 % sur l'ensemble

d'entraînement. En ce qui concerne les erreurs, le modèle affiche une RMSE de 0,056 sur l'entraînement, 0,406 sur la validation et 0,380 sur le test, ce qui témoigne d'un très bon ajustement, y compris sur des données non vues. L'erreur absolue moyenne est également faible 0,042 en entraînement, 0,283 en validation et 0,25 en test, ce qui indique une bonne précision dans la prédiction de la résistance à la compression. Ces résultats montrent que le modèle généralise bien sans surapprentissage notable. Des erreurs comparables ont été rapportées par Ngo et al (Ngo et al., 2025) dans une étude utilisant également le Gradient Boosting pour la prédiction de la résistance à la compression des matériaux terreux stabilisés, avec un RMSE de 0,259 en entraînement et 0,612 en test. Toutefois une comparaison directe reste délicate en raison de la variabilité des bases de données utilisées.

Le tableau 2 compare également les résultats du modèle optimisé au modèle non optimisé, le modèle optimisé affiche des performances supérieures sur l'ensemble des jeux de données, en particulier sur le test, avec une amélioration de (+0,8 %) sur le R^2 , une réduction de RMSE de 17% et une baisse de MAE de près de 26 %. Ces gains soulignent l'importance de l'optimisation et mettent en évidence une meilleure capacité de généralisation, qui devra néanmoins être confirmée sur de nouvelles études.

TABLEAU 2. Performances globales du modèle avant et après optimisation

Modèles	Entraînement			Validation			Test		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
Optimisé	0,999	0,056	0,042	0,976	0,406	0,283	0,984	0,380	0,250
Non optimisé	0,996	0,188	0,138	0,965	0,513	0,357	0,976	0,456	0,335

III.2. Présentation visuelle des performances

La figure 3 illustre visuellement les performances du modèle sur les données d'entraînement et de test. La figure 3 (a) et la figure 3 (b) montrent les courbes de régression entre les résistances réelles et prédites. L'alignement des points autour de la diagonale témoigne d'une prédiction précise. Sur le test, les prédictions sont bien réparties autour de la droite, sans biais systématique. L'équation de la régression montre une pente très proche de 1, signe d'une très bonne fidélité du modèle.

La figure 3 (c) complète cette analyse en affichant les valeurs réelles, prédites et l'erreur absolue pour chaque échantillon du jeu de test. L'erreur reste globalement faible, avec 66,66% des erreurs inférieures à 0,25 MPa correspondant à RMSE. Les prédictions suivent bien la tendance des données, avec des écarts raisonnables même aux extrêmes. La régularité de l'erreur absolue confirme une bonne stabilité du modèle face aux variations internes du matériau.

III.3. Analyse de l'importance des variables

L'importance d'une variable est déterminée comme la somme normalisée des gains de réduction de la erreur quadratique moyenne (MSE) à travers l'ensemble des arbres du modèle (Louppe et al., 2013).

La figure 5 (a) présente l'importance des variables dans la prédiction. Le type et la teneur en liant sont les plus influente (avec 34% et 27% respectivement), suivie de la contrainte de compactage (21 %). Ces observations sont cohérentes avec plusieurs travaux antérieurs qui ont souligné l'influence majeure du taux de liant et de la pression de compactage sur les propriétés mécaniques des matériaux terreux stabilisés (Sadouri et al., 2024; Taallah et al., 2014). Les caractéristiques géotechniques du sol et les fibres apparaissent secondaires mais ne sont pas à négliger.

En complément, une analyse par élimination récursive avec validation croisée (RFECV) a été réalisée. Cette méthode consiste à retirer progressivement certaines variables et à réévaluer les performances du modèle à chaque étape, afin d'identifier celles qui contribuent réellement à la prédiction. La figure 5 (b) illustre l'évolution du RMSE en fonction du nombre de variables conservées. Le retrait de trois variables — la teneur en sable, la teneur en gravier et le type de malaxage — permet une légère amélioration du modèle, avec un RMSE de 0,390 MPa obtenu pour 260 estimateurs, contre 0,406 MPa avec 310 estimateurs lorsque toutes les variables sont incluses. Avec 310 estimateurs et les trois variables retirées, le RMSE reste similaire (0,391 MPa). En revanche, la suppression de variables supplémentaires entraîne une dégradation des performances. Ces résultats confirment que l'ensemble des variables retenues est globalement nécessaire à une bonne prédiction, tout en indiquant que certaines caractéristiques granulométriques ont un rôle plus limité, probablement en raison de leur corrélation avec d'autres variables explicatives déjà intégrées dans le modèle.

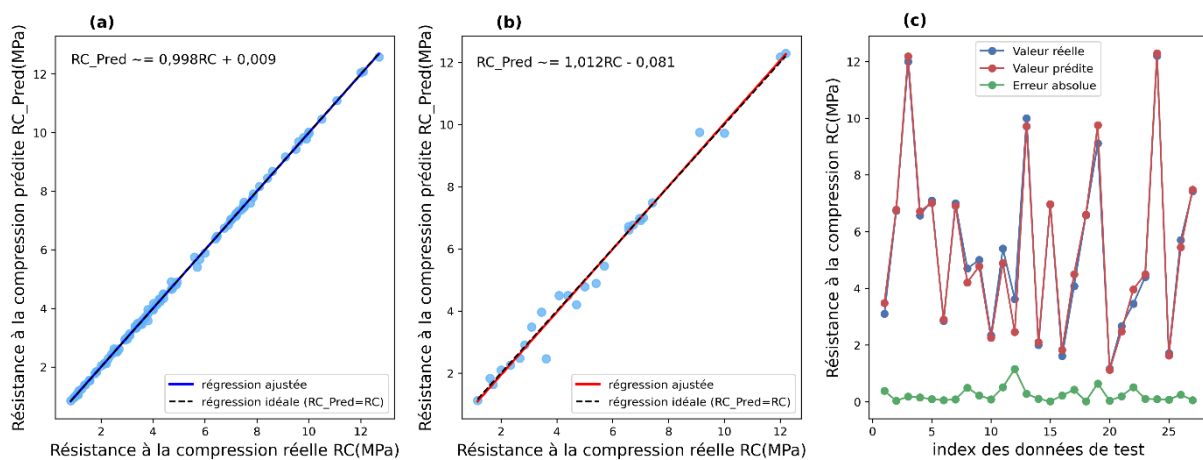


FIGURE 4. Évaluation des performances du modèle : (a) Régression entre les résistances réelles et prédites sur l'entraînement, (b) Régression entre les résistances réelles et prédites le test, (c) comparaison des valeurs réelles et prédites sur les données de test avec erreur absolue.

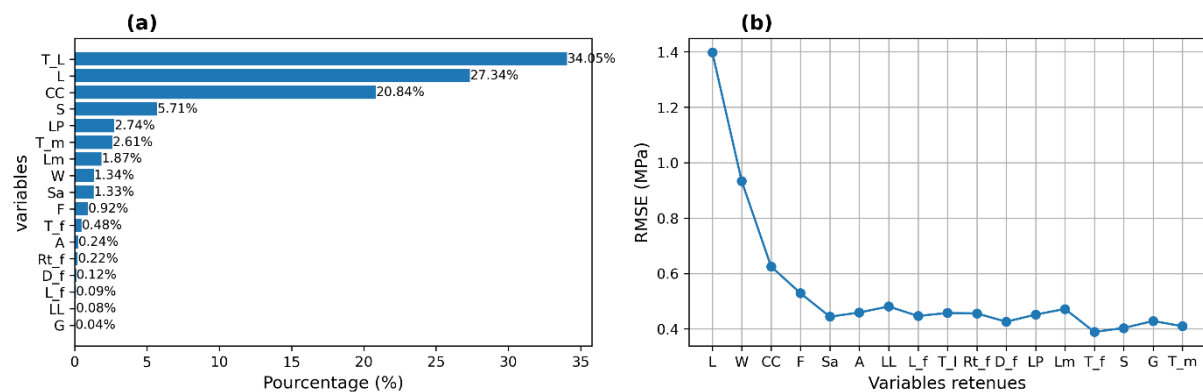


FIGURE 5. (a) : importance relative des variables d'entrée (%) sur la prédiction du modèle. (b) dégradation du RMSE (MPa) lors de l'élimination récursive des variables en validation croisée.

IV. CONCLUSION

Nous avons proposé dans cette étude l'utilisation du Gradient Boosting dans la prédiction de la résistance à la compression des matériaux en terre crue à partir des caractéristiques géotechniques

mesurables, des paramètres de formulation et la contrainte de compactage. Pour cela, une base de données constituée de 135 échantillons, incluant des briques stabilisées et non stabilisées, a été mobilisée. Les résultats montrent que le modèle offre d'excellentes performances prédictives, avec une racine carrée de l'erreur quadratique moyenne de 0,346 MPa, une erreur absolue moyenne de 0,239 MPa et un coefficient de détermination de 0,986 sur les données de test. Le modèle présente une capacité élevée à restituer avec précision la résistance à la compression à partir de données d'entrées tout en mettant en évidence l'influence dominante de variables telles que le pourcentage massique de liant, la contrainte de compactage, et le pourcentage massique d'eau du mélange. Ces résultats confirment le potentiel de l'intelligence artificielle pour prédire la résistance à la compression des briques en terre crue. Cependant, la taille restreinte de la base de données, notamment sur les données de test constitue une limite importante à la généralisation des résultats. Des futures études devront porter sur l'extension de la base de données, en intégrant une plus grande diversité de sols, de procédés de fabrication, de condition et durée de cure, afin de permettre une évaluation plus robuste et généralisable des performances du modèle.

Remerciements

Ce travail a été réalisé au sein de l'Institut de Recherche en Génie civil et Mécanique (GeM – CNRS UMR 6183), à Saint-Nazaire, dans le cadre du projet CNRS-Africa JRP AI4EB, coordonné par BONNET Stéphanie et TCHIOTSOP Daniel en charge de l'encadrement scientifique du projet. Les auteurs remercient toute l'équipe pour son soutien et sa contribution à la réussite de ce travail.

REFERENCES

- Burroughs, S., 2009. Relationships between the density and strength of rammed earth. *Proc. Inst. Civ. Eng.-Constr. Mater.* 162, 113–120. <https://doi.org/10.1680/coma.2009.162.3.113>
- Danso, H., Martinson, D.B., Ali, M., Williams, J.B., 2015. Physical, mechanical and durability properties of soil building blocks reinforced with natural fibres. *Constr. Build. Mater.* 101, 797–809. <https://doi.org/10.1016/j.conbuildmat.2015.10.069>
- Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* 29. <https://doi.org/10.1214/aos/1013203451>
- Habert, G., Miller, S.A., John, V.M., Provis, J.L., Favier, A., Horvath, A., Scrivener, K.L., 2020. Environmental impacts and decarbonization strategies in the cement and concrete industries. *Nat. Rev. Earth Environ.* 1, 559–573. <https://doi.org/10.1038/s43017-020-0093-3>
- Hamard, E., Cazacliu, B., Razakamanantsoa, A., Morel, J.-C., 2016. Cob, a vernacular earth construction process in the context of modern sustainable building. *Build. Environ.* 106, 103–119. <https://doi.org/10.1016/j.buildenv.2016.06.009>
- Houben, H., Guillaud, H., 1994. *Earth Construction: A Comprehensive Guide*, Craterre-Eag, ed, Earth Construction Series; Grenoble, France.
- Houcine, B., Mohamed, R., Samir, K., Sarra, B., 2023. Artificial intelligence for the prediction of the physical and mechanical properties of a compressed earth reinforced by fibers. *J. Eng. Exact Sci.* 9, 15910–01e. <https://doi.org/10.18540/jcecvl9iss4pp15910-01e>
- Loupe, G., Wehenkel, L., Sutera, A., Geurts, P., 2013. Understanding variable importances in forests of randomized trees. *Adv. Neural Inf. Process. Syst.* 26, 431–439.
- Malonga, B., Poullain, P., Bendahmane, F., Bonnet, S., Issaadi, N., Ahouet, L., 2024a. Mechanical and Thermal Characterisation of Compressed Earth Blocks Made of Termite Mound Soil (*Macrotermes* Sp.) Stabilised with Corn Starch Gel, in: Beckett, C., Bras, A., Fabbri, A., Keita, E., Perlot, C., Perrot, A. (Eds.), *Second RILEM*

International Conference on Earthen Construction, RILEM Bookseries. Springer Nature Switzerland, Cham, pp. 122–131. https://doi.org/10.1007/978-3-031-62690-6_13

Malonga, B., Poullain, P., Bonnet, S., Bendahmane, F., Issaadi, N., Ahouet, L., 2024b. STABILISATION OF COMPRESSED EARTH BRICKS BASED ON CORN STARCH GEL.

Medjelekh, D., Ulmet, L., Gouny, F., Fouchal, F., Nait-Ali, B., Maillard, P., Dubois, F., 2016. Characterization of the coupled hygrothermal behavior of unfired clay masonries: Numerical and experimental aspects. *Build. Environ.* 110, 89–103. <https://doi.org/10.1016/j.buildenv.2016.09.037>

Miccoli, L., Müller, U., Fontana, P., 2014. Mechanical behaviour of earthen materials: A comparison between earth block masonry, rammed earth and cob. *Constr. Build. Mater.* 61, 327–339. <https://doi.org/10.1016/j.conbuildmat.2014.03.009>

Morel, J.C., Mesbah, A., Oggero, M., Walker, P., 2001. Building houses with local materials: means to drastically reduce the environmental impact of construction. *Build. Environ.* 36, 1119–1126. [https://doi.org/10.1016/S0360-1323\(00\)00054-8](https://doi.org/10.1016/S0360-1323(00)00054-8)

Mustafa, Y.M.H., Wudil, Y.S., Zami, M.S., Al-Osta, M.A., 2025. Machine Learning Approach for Assessment of Compressive Strength of Soil for Use as Construction Materials. *Eng* 6, 84. <https://doi.org/10.3390/eng6050084>

Ngo, T.-P., Vu, H.-N., Bui, Q.-B., 2025. Application of machine learning models for the optimisation of compressive strength and water resistance of geopolymer stabilised compacted earth. *Case Stud. Constr. Mater.* 22, e04203. <https://doi.org/10.1016/j.cscm.2025.e04203>

Pedregosa, F., Pedregosa, F., Varoquaux, G., Varoquaux, G., Org, N., Gramfort, A., Gramfort, A., Michel, V., Michel, V., Fr, L., Thirion, B., Thirion, B., Grisel, O., Grisel, O., Blondel, M., Prettenhofer, P., Prettenhofer, P., Weiss, R., Dubourg, V., Dubourg, V., Vanderplas, J., Passos, A., Tp, A., Cournapeau, D., 2011. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.

Poullain, P., Barnaure, M., Bonnet, S., 2023. Mechanical properties of cob-earth composites: Variability and focus on the different calculation methods of Young's modulus. *J. Build. Eng.* 72, 106622. <https://doi.org/10.1016/j.jobe.2023.106622>

Poullain, P., Leklou, N., Laïbi, A., Gomina, M., 2019. Properties of Compressed Earth Blocks Made of Traditional Materials from Benin. *Rev. Compos. Matér. Avancés* 29, 233–241. <https://doi.org/10.18280/rcma.290407>

Sadouri, R., Kebir, H., Benyoucef, M., 2024. The Effect of Incorporating Juncus Fibers on the Properties of Compressed Earth Blocks Stabilized with Portland Cement. *Appl. Sci.* 14, 815. <https://doi.org/10.3390/app14020815>

Taallah, B., Guettala, A., Guettala, S., Kriker, A., 2014. Mechanical properties and hygroscopicity behavior of compressed earth block filled by date palm fibers. *Constr. Build. Mater.* 59, 161–168. <https://doi.org/10.1016/j.conbuildmat.2014.02.058>

Tchiotsop, J., Issaadi, N., Poullain, P., Bonnet, S., Belarbi, R., 2022. Assessment of the natural variability of cob buildings hygric and thermal properties at material scale: Influence of plants add-ons. *Constr. Build. Mater.* 342, 127922. <https://doi.org/10.1016/j.conbuildmat.2022.127922>

Thompson, D., Augarde, C., Osorio, J.P., 2025. A review of chemical stabilisation and fibre reinforcement techniques used to enhance the mechanical properties of rammed earth. *Discov. Civ. Eng.* 2, 27. <https://doi.org/10.1007/s44290-025-00184-1>

Turco, C., Funari, M.F., Teixeira, E., Mateus, R., 2021. Artificial Neural Networks to Predict the Mechanical Properties of Natural Fibre-Reinforced Compressed Earth Blocks (CEBs). *Fibers* 9, 78. <https://doi.org/10.3390/fib9120078>

Yan, J., Yuan, K., Zhang, F., Guo, L., 2022. Study on Mechanical Properties and Constitutive Equation of Earth Materials under Uniaxial Compression. *Appl. Sci.* 13, 19. <https://doi.org/10.3390/app13010019>

Zami, M.S., Lee, A., 2010. Economic benefits of contemporary earth construction in low-cost urban housing – State-of-the-art review. *J. Build. Apprais.* 5, 259–271. <https://doi.org/10.1057/jba.2009.32>