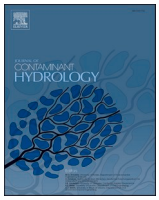
Journal of Contaminant Hydrology 263 (2024) 104337

Contents lists available at ScienceDirect

Journal of Contaminant Hydrology

journal homepage: www.elsevier.com/locate/jconhyd



Statistical and machine learning analysis for the application of microbially induced carbonate precipitation as a physical barrier to control seawater intrusion

Charalampos Konstantinou [a,*], Yuze Wang [b,c,**]

[a] Department of Civil and Environmental Engineering, University of Cyprus, Nicosia, Cyprus
[b] Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen, China
[c] Southern Marine Science and Engineering Guangdong Laboratory, Guangzhou, China

ARTICLE INFO

Keywords:
Seawater intrusion
Biochemical cut-off wall
Microbially induced carbonate precipitation
Machine learning
Statistical analysis
Hydraulic conductivity

ABSTRACT

Seawater intrusion in coastal aquifers is a significant problem that can be addressed through the construction of subsurface dams or physical cut-off barriers. An alternative method is the use of microbially induced carbonate precipitation (MICP) to reduce the hydraulic conductivity of the porous medium and create a physical barrier. However, the effectiveness of this method depends on various factors, and the scientific literature presents conflicting results, making it challenging to generalise the findings. To overcome this challenge, a statistical and machine learning (ML) approach is employed to infer the causes for the reduction in hydraulic conductivity and identify the optimum MICP parameters for preventing seawater intrusion. The study involves data curation, exploratory analysis, and the development of various models to fit the input data (k-Nearest Neighbours – kNN, Support Vector Regression – SVR, Random Forests – RF, Gradient Boosting – XgBoost, Linear model with interaction terms, Ensemble learning algorithms with weighted averages – EnL-WA and stacking – EnL-Stack). The models performed reasonably well in the region where permeability reduction is sensitive to carbonate increase capturing the permeability reduction profile with respect to cementation level while demonstrating that they can be used in initial assessments of the specific conditions (e.g., soil properties). The best performing algorithms were the EnL-Stack and RF followed by XgBoost and SVR. The MICP method is effective in reducing hydraulic conductivity provided that the various biochemical parameters are optimised. Critical biochemical parameters for successful MICP formulations are the bacterial optical density, the urease activity, calcium chloride concentration and flow rate as well as the interaction terms across the properties of the porous media and the biochemical parameters. The models were used to identify the optimum MICP formulation for various porous media properties and the maximum permeability reduction profiles across cementation levels have been derived.

1. Introduction

In recent years, the transition to a climate change adaptation era has necessitated the use of modern technologies and techniques to manage water resources. One such strategy is the use of microbially induced carbonate precipitation (MICP), a technique that involves introducing bacteria into a medium to produce calcium carbonate, which solidifies soil and granular networks. This approach was applied for various purposes in water resources management, including erosion control in seawater environments, drought mitigation, soil stabilization, and reduction of seepage (Hu et al., 2021; Kim et al., 2020; Lambert and Randall, 2019; Lin et al., 2023; Liu et al., 2021; Wang et al., 2022; Yu and Rong, 2022). It can also mitigate groundwater contamination by immobilizing contaminants in the subsurface through the precipitation of calcium carbonate or to control the fracturing behaviour (Konstantinou et al., 2023b; Rajasekar et al., 2021b; Wang et al., 2023a). Moreover, MICP is also being utilised in controlled laboratory-scale studies to investigate groundwater flow and transport by creating artificial specimens with customized properties and characteristics and then conducting fluid flow experiments (Gago et al., 2020; Kirk et al., 2012;

* Corresponding author at: Department of Civil and Environmental Engineering, University of Cyprus, Nicosia, Cyprus.
** Corresponding author at: Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen, China.
E-mail addresses: ckonst06@ucy.ac.cy, ck494@cantab.ac.uk (C. Konstantinou), wangyz@sustech.edu.cn (Y. Wang).

https://doi.org/10.1016/j.jconhyd.2024.104337
Received 16 December 2023; Received in revised form 15 March 2024; Accepted 17 March 2024
Available online 20 March 2024

Konstantinou et al., 2023a; Konstantinou et al., 2022; Konstantinou et al., 2021a; Konstantinou and Biscontin, 2022; Singh et al., 2015).

Apart from the above applications, MICP could potentially be used as a tool to generate a physical barrier for seawater intrusion (SWI) control during excessive exploitation of freshwater in aquifers close to the coast. Seawater intrusion leads to deterioration of freshwater quality and damages the groundwater supply, posing a significant threat to the environment and public health. Various methods and curative measures have been developed to tackle this issue, such as the reduction of the pumping rates (Nasiri et al., 2021), the generation of hydraulic barriers by injecting freshwater into the aquifer (Saad et al., 2022; Zhang et al., 2022), the construction of subsurface dams and cut-off walls (Chang et al., 2019; Fang et al., 2021; Kaleris and Ziogas, 2013; Luyun et al., 2009; Shen et al., 2020; Yang et al., 2021), the use of surface water recharge canals (Motallebian et al., 2019) and, recently, the generation of physical barriers/cut-off walls (Abdoulhalik et al., 2017; Laabidi and Bouhlila, 2021; Sun et al., 2021; Xie et al., 2023). Laabidi and Bouhlila (2021) proposed the generation of a calcite cut-off wall by mixing naturally occurring chemicals, namely aqueous solutions of sodium carbonate (Na_2CO_3) and calcium chloride ($CaCl_2$). Along the same lines, the precipitation of carbonate via MICP could be a potential solution that falls within the same category. Although the potential of the MICP application to prevent seawater intrusion in groundwater has been discussed in the literature, no experiments have been conducted specifically for this purpose (Dawoud, 2020; Mortensen et al., 2011; Phillips et al., 2013).

Since the primary MICP mechanism for this specific application is the reduction in the permeability of the porous medium, there are four critical groups of factors to consider when designing an effective MICP-based seawater intrusion barrier (see Fig. 1). These include the biochemical factors affecting MICP (bacterial strains, urease activity, flow rates, and retention times), the MICP methods (e.g. gravity-based injection, flow rate or pressured controlled and inclusions of substrates in the injecting fluids), the properties of the porous medium (grain size, width of particle size distribution, grain shape, and relative density), and the environmental conditions (temperature, seawater or freshwater, and pH). Although some of these factors have been studied in previous research (Al Qabany and Soga, 2013; Lin et al., 2020; Wang and Nackenhorst, 2020a), they were not explicitly investigated for the purpose of seawater intrusion control. At the same time, the findings of the previous studies cannot be unified to generalise the conclusions in terms of interpreting the underlying factors and identifying an optimum MICP recipe because of the different formulations of each study.

In addition, the hydraulic performance of MICP-treated porous media is influenced by various other factors, including the grain characteristics of the media. According to Konstantinou et al. (2023c), grain shape, angularity, and size distribution have a significant effect on the hydraulic conductivity and strength of MICP-treated sand. The study concluded that understanding the influence of grain characteristics on MICP-treated porous media is crucial for the successful implementation of MICP for both strengthening soils and controlling permeability.

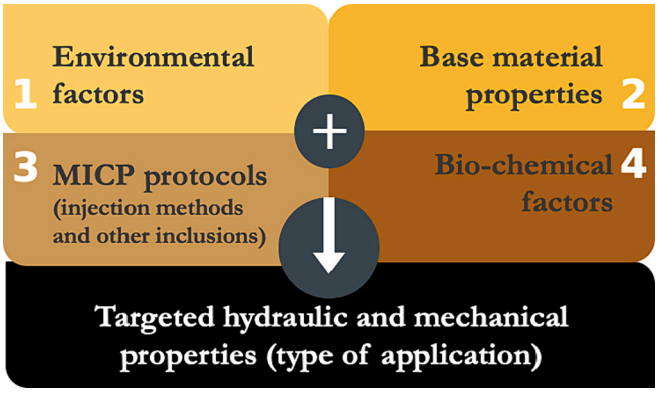


Fig. 1. The critical groups of factors to consider when designing an effective MICP-based formulation.

As stated previously, the findings of these studies cannot be extrapolated to draw conclusions with respect to seawater intrusion control since the MICP formulation, grain characteristics and other environmental factors are inconsistent across the literature. This study adopts a statistical and machine learning approach to overcome this issue and build models to predict the reduction in hydraulic conductivity, to assess the application of MICP and fine-tune its parameters for designing a physical calcite barrier to prevent seawater intrusion into groundwater. First, a comprehensive literature review is conducted for data curation, in which the various input variables are identified and documented. The output variable is the hydraulic conductivity, while the input variables are environmental, biochemical or variables related to the properties of the porous media. This is followed by exploratory analysis where possible patterns are identified. Then, statistical analysis and machine learning (ML) regressors are used to develop models that describe the input data. The models used are the Support Vector Regressors (SVR), k-Nearest Neighbours (kNN), Random Forests (RF), Gradient Boosting (XgBoost), linear regression model with and without interaction terms and ensemble models to combine the previous cases (ensemble learning with weighted averages and with stacking).

Next, inference of the models is conducted with the objectives to (i) assess the influence of the various input parameters and their interactions, (ii) identify the optimum MICP parameters for designing the barrier based on various porous media properties which could exist in the field, (iii) derive standard permeability reduction profiles with respect to cementation levels, and (iv) compare the predictions against empirical models (i.e., Panda-Lake models) and experimental data across a range of cementation levels. These comparisons also show the range of the output variable for which the models are more accurate, while also highlighting the importance of using ML models when the investigated problem becomes more complex.

2. Methodology

The main parameter that defines the ability of a porous material or soil to transmit water through it under a hydraulic gradient is hydraulic conductivity which has units of velocity (m/s). It represents the rate at which water can flow through a soil or other porous media under a given hydraulic gradient or pressure difference. Throughout this study, the hydraulic conductivity is reported using the standard unit of m/s. The overall methodology is presented in the flow chart in Fig. 2. First, the database was developed followed by exploratory analysis. Then, the statistical and ML models were developed. A causality analysis was conducted to infer the parameters that affect the most the output variable and assess their interactions. Then the optimum bio-chemical conditions for various porous media were identified, and the predictions of permeability for a range of cementation levels were compared with the experimental values and with empirical models for permeability.

2.1. Database development

MICP studies were identified that provided measurements of hydraulic conductivity and the various parameters/ variables were stored in a database. In total, the resulting database contains around 1400 entries, 25 input variables were identified, which include environmental factors, grain properties, and MICP parameters. The data was extracted from 62 studies (2 of these studies provide data relating to other studies); these studies are shown in Table 1.

2.2. Data Curation, homogenisation and gap filling

Various variables were extracted from the datasets that provided

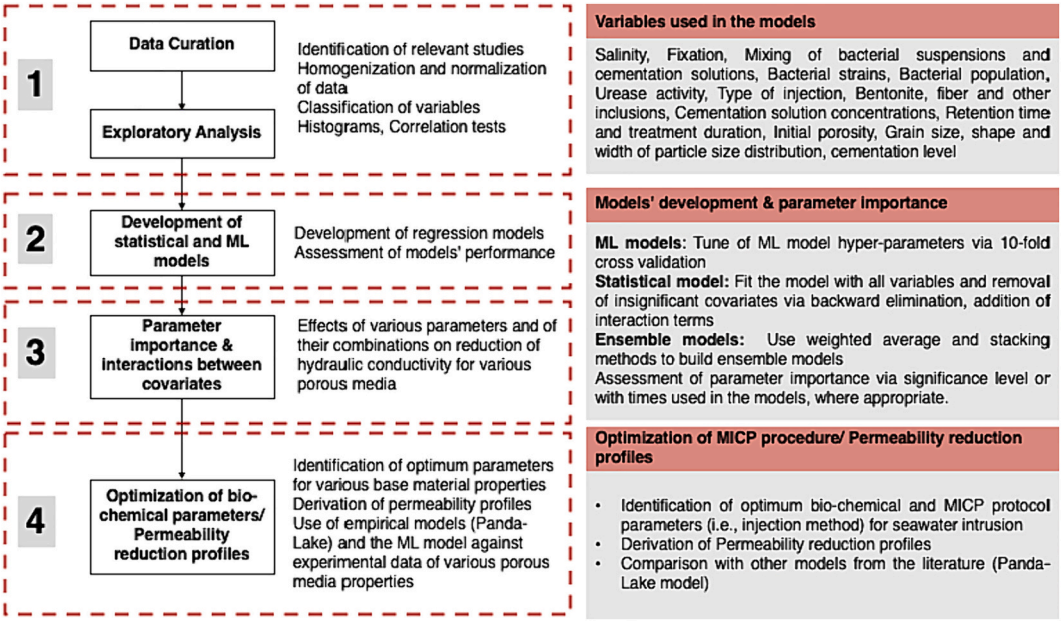


Fig. 2. The methodology of this study.

measurements for hydraulic conductivity (output variable). Hydraulic conductivity was reported in m/s in this study. The entries that represent the un-treated sands (0% cementation) were removed from the dataset resulting to 1348 entries. The variables were normalised before being used in the models.

As mentioned in the introduction, four groups of parameters are of great importance: the MICP bio-chemical parameters, the environmental factors, the injection methods, and the properties of the porous media. A summary of the input variables used in the models is given in Table 2. The variables are divided into three categories: binary, numerical, and categorical variables.

The biochemical parameters include the delivery method of solutions to the medium, the chemical concentrations, the bacterial strain populations, etc. Many studies use bacterial fixation which refers to the immobilization of the bacterial cells onto the soil particles or aggregates because it allows the bacteria to remain near the calcium source and provides a stable environment for bacterial growth and calcium carbonate precipitation. In most of the studies the stabilization solution consists of 50 mM $CaCl_2$. In other studies, a solution containing bacterial cells and a solution containing calcium ions (the cementation solution) are mixed to initiate the precipitation of calcium carbonate at the same time while injecting the solutions into the porous medium. Because there is no standard formulation in the available literature due to various proportions of bacterial suspensions and cementation solutions at various concentrations being used, this variable is binary.

Any bacterial strain could be used in MICP procedures if the strains are capable of hydrolysing urea. In the studies that measure hydraulic conductivity, *S. pasteurii*, *B. megaterium*, *B. sphaericus,* artificially extracted plant enzymes*, B. licheniformis,* biocatalysts, hybrid bacteria*, P. nitroreducens,* seed culture*, V. arenosi* and other ureolytic microorganisms were used. These were categorised in four groups and represented as binary variables in the ML models. The optical density of bacterial suspensions is a common measurement of the bacterial population which is usually given as a range. A numerical variable representing the average value was used. In cases where the bacterial densities were reported in colony forming units (cfu/mL), these were converted into OD_{600} values based on the general form of the standard curve that describes the relationship between these two units. In very few cases the OD was not reported and was taken from relevant research authors which performed the experiments under the same conditions. In the absence of such information an estimation was made based on the densities of the specific bacterial strains and the average value. The urease activity is the rate at which urea breaks down into ammonia and carbon dioxide and in this case, it indicates the rate of calcium carbonate production (urea breakdown rate). Several units have been used in the relevant publications and these were converted into millimoles (mM) of urea hydrolysed per minute (mM/min) so that they could be directly compared. In studies where the urease activity was not reported it was derived using a standard growth curve based on the average values from the literature (8.5 mM/min for OD_{600} values from 1.5 and above and linear interpolation below this value) (Onal Okyay and Frigi Rodrigues, 2013; Whiffin, 2004). The calcium source is also one of the biochemical parameters and an important constituent of the MICP formulation. It is reported as mol/L (M). Various forms of calcium source have been used (calcium chloride, anhydrous calcium chloride, calcium chloride dihydrate and calcium acetate) which are known to have different effects on the calcium carbonate precipitation patterns and the crystal characteristics. This information was not captured by the introduced variable. The ratio of urea to calcium chloride concentration in the cementation solution (CS) was also used as a variable to train the model and was either reported as a ratio, or by stating the individual concentrations.

The time between two subsequent injections (retention time) in the cases of staged injections is another parameter that is of interest since it is important to allow sufficient time for reactions to take place especially when urease activity is low. The variable is reported in hours. For continuous injections, the total treatment duration is reported in days. Studies have also been using various degrees of saturation for the MICP procedure to control the resulting strength and hydraulic conductivity and these are represented by a numerical variable which is a percentage. The method of delivery of bacterial suspensions and chemicals is also of great importance when designing MICP programs. Flow rate-controlled injections, injections via gravity, pressure-controlled injections and immersion of porous media into the solutions have been used in the studies relevant to hydraulic conductivity. The gravity and flow-rate controlled methods which are the main methods for delivering bacterial suspensions and chemical solutions to the granular network were expressed as a binary variable (1 for gravity-based flow). Any modifications of gravity or flow-rated controlled injections or other injection methods (e.g., spraying, mix soil with reactants, wet curing, immerging) were considered using a separate binary variable (OthFlow).

Table 1
The studies used to develop the database classified according to the critical groups of factors to consider when designing an effective MICP-based formulation.

References	Description
1. **Environmental factors**	
(Cheng et al., 2016; Cheng et al., 2014; Cheng et al., 2013; Dekuyer et al., 2012; Jawad and Zheng, 2016; Lin et al., 2023; Liu et al., 2023b; Yu et al., 2022)	MICP under various saturation levels, pH and other environmental conditions and/or in marine/saline environments
(Gomez et al., 2014; Rajasekar et al., 2021a; Sharma et al., 2021; Stabnikov et al., 2013)	Use of in-situ bacterial strains/ hybrid bacteria for MICP
2. **Base material properties**	
(Chen et al., 2023; Duo et al., 2018; Guo et al., 2024a; Li and Chen, 2022; Liu et al., 2023b)	MICP for aeolian sand, calcareous sand, river, and sea sand
(Chen et al., 2023; Li and Chen, 2022; Baek et al., 2024; Konstantinou, 2021; Konstantinou et al., 2023c; Li and Chen, 2022; Song et al., 2021; Song et al., 2020; Zamani et al., 2019; Zamani et al., 2017)	Use of various porous media – various grain characteristics
(Montoya et al., 2019; Phang et al., 2022; Safavizadeh et al., 2018; Sidik et al., 2014; Song et al., 2022; Wang et al., 2023)	Use of various porous media – peat, fly ash, organic soil, granite, rocks etc.
3. **MICP protocols (injection methods & other content inclusion)**	
(Gong et al., 2019; Martinez et al., 2013; Montoya et al., 2018; Niu et al., 2018; Tian et al., 2020, Tian et al., 2018; Wang and Nackenhorst, 2020b; Whiffin et al., 2007; Yang et al., 2022; Yang et al., 2019; Yu et al., 2022; Yu and Yang, 2023)	Use of various MICP procedures (protocols)
(Choi et al., 2020; Choi et al., 2019; Choi et al., 2017; Choi et al., 2016a; Fang et al., 2020; Liu et al., 2023a; Ma et al., 2021; Zhao et al., 2021; Zhao et al., 2020)	Use of fibres, biopolymers, bentonite and other sustainable methods in MICP
4. **Bio-chemical factors**	
(Akoğuz et al., 2019; Choi et al., 2016b; Kadhim and Zheng, 2017)	Effects of different calcium sources in soil improvement
(Al Qabany and Soga, 2013; Dawoud et al., 2014; Soon et al., 2014; Yasuhara et al., 2012, Yasuhara et al., 2011)	Effects of chemical treatment on MICP engineering properties
(Eryürük, 2022; Rowshanbakht et al., 2016; Sharma et al., 2021; Soon et al., 2014; Stabnikov et al., 2013; Yang et al., 2019; Yasuhara et al., 2012)	Bacterial population, strain or enzymatic effects and other related characteristics on MICP

The injection rate in the case of controlled flow rate is expressed as mL/min. This variable is generally reported in these units; however, in few cases it was reported as mM of $CaCl_2$/min or given as a velocity (m/s). The variable was converted to mL/min based on either the concentration of the chemicals and the pore volume (PV) or based on the geometry of the specimen being bio-treated.

The environmental factor relevant to the case of cut-off barriers for seawater intrusion control beyond the classic ones identified in the literature is salinity. Some studies have compared seawater with freshwater and since this is an environment which would be potentially relevant to the case of seawater intrusion, the variable was included in the model. Due to its great heterogeneity (various seawater formulations), this variable is binary and takes the value of 0 when there is no saline environment or 1 when the MICP process is conducted in saline environments.

In terms of the properties of the porous media, the grain diameter, grain shape and the width of the particle size distribution are the main microstructure properties that on the one hand define the initial hydraulic conductivity and on the other hand significantly affect the MICP procedure and the microstructure (Konstantinou and Wang, 2023; Wu

Table 2
The input variables utilised in the models.

	Name	Variables	Description	Units
	Binary variables			
1	Sal	Salinity	Whether the environment is in seawater or freshwater	N/A
2	Fix	Fixation	Whether calcium chloride solution is flushed before bacterial solution to enhance bacterial attachment	N/A
3	BsA	Bacterial Strain A	Whether the strain is *S. pasteurii*	N/A
4	BsB	Bacterial Strain B	Whether the strain is *B. megaterium*	N/A
5	BsC	Bacterial Strain C	Whether the strain is *B. sphaericus*	N/A
6	BgD	Bacterial Group D	Any other bacterial strain or ureolytic method	N/A
7	Grav	Infiltration/ Flowrate	Whether the injection method is based on gravity	N/A
8	OthFlow	Other flow rate	Whether the injection method is other than infiltration and flow rate or a modified version.	N/A
9	OthCon	Other Content	Whether the granular medium contains additives other than fiber or bentonite (silt, etc.)	N/A
10	Mix	Mixing of bacterial solution with cementation solution	Whether mixing of the two solutions during injection is observed	N/A
	Numerical Variables			
11	OD	Optical density (OD_{600})	Bacterial optical density at a wavelength of 600 nm	N/A
12	UreAct	Urease Activity	Bacterial urease activity	mM/min
13	CalChlor	Calcium Chloride	Calcium chloride concentration in the cementation solution	M (mol/L)
14	Ratio	Urea to Calcium Chloride Ratio	The ratio of urea over calcium chloride concentration in the cementation solution	N/A
15	ReTime	Retention time	The time between two subsequent injections	hrs
16	TreatDur	Total Treatment Duration	The total treatment duration for continuous injection	days
17	FlowRat	Injection rate	The injection rate in the case of controlled flow rate	mL/min
18	Por	Initial Porosity	The initial porosity of the porous medium (0–1)	N/A
19	Sat	Saturation	The level of saturation of the porous medium during the MICP treatment (0–1)	N/A
20	FibCon	Fiber content	The fiber content in the porous medium	Percentage by weight
21	BenCon	Bentonite content	The bentonite content in the porous medium	Percentage by weight
22	GrainDia	Grain size	The average grain size of the porous medium	μm
23	Cu	Cu	The uniformity coefficient of the porous medium (=D_{60}/D_{10})	N/A

(*continued on next page*)

Table 2 (*continued*)

	Name	Variables	Description	Units
24	CemLevel	Cementation level	The final cementation level of the bio-treated porous medium (weight of cementation over the total weight)	Percentage by weight
	Categorical variables			
25	GrainShape	Grain Shape	Spherical particles −1 Subrounded particles −2 Angular particles - 3	N/A

et al., 2023). These are parameters that cannot be controlled in the field and should therefore be considered when designing procedures. These characteristics were directly reported or obtained from the particle size distribution curves (either readily available or derived from the proportions of clay, silt, sand, gravel) or from literature data given that most of these sands are 'standard' (e.g., ISO, british standards, ASTM, Ottawa or Toyoura sand). Similarly, initial porosity should also be known, which includes factors such as the density, and degree of compaction of the material. In cases where this was not reported, it was calculated indirectly either utilising the Kozeny-Carman equation based on the grain characteristics, or via the grain density, grain characteristics and relative density.

Some of the studies included fibres, bentonite, silt, or other additives for various reasons other than for reducing hydraulic conductivity, e.g. to increase the tensile strength of the resulting specimens. These were also included in the analysis to assess the effects on permeability. Fiber and bentonite content were expressed as numerical variables whilst any other type of inclusion were represented by a binary variable. The variables did not include effects such as diameter or fiber length, or properties of bentonite which were examined individually in the investigated studies.

Finally, the cementation level (expressed as a percentage) was linked to the permeability based on the provided figures or tables in each study. In some cases, permeability was not directly linked to the cementation level. For example, where hydraulic conductivity was plotted against strength parameters (i.e., shear wave velocity, UCS, etc.), permeability was mapped to the strength parameter which in turn was plotted against the cementation level.

2.3. Exploratory analysis

Prior to constructing the statistical and ML models, an exploratory analysis was conducted to detect any potential patterns emerging from the input data. Histograms were developed for each input variable and a correlation test was conducted to identify highly correlated variables. In cases where the predictor variables were found to be highly correlated, one variable from each pair was removed to address the issue of multicollinearity and ensure the integrity of the models.

2.4. Statistical and machine learning algorithm development

A causality analysis was conducted to identify the parameters that contribute the most to the control of permeability. Various machine learning (ML) algorithms and multiple regression models were employed, based on recommendations outlined in Konstantinou and Stoianov (2020):

1. Random Forests (RF): the algorithm utilizes an ensemble learning method to create multiple decision trees. For regression tasks, each decision tree predicts a continuous value, and the final prediction is determined by averaging the predictions from all the trees (Breiman, 2001; Hastie et al., 2009).
2. eXtreme Gradient Boosting (XgBoost): similar to RF, the algorithm utilizes decision tree ensembles within a gradient boosting framework, optimizing predictive accuracy through the iterative combination of weak learners (Friedman, 2001).
3. Support Vector Regression (SVR): the algorithm minimizes the distance between predicted and actual values while remaining within a specified margin. The margin is determined by a regularization parameter, and the distance is measured by utilising a kernel function (Kecman, 2005).
4. k-Nearest Neighbours (kNN): this algorithm makes predictions based on the majority class or average of k nearest data points in the feature space, where proximity is measured using Euclidean distance.
5. Linear regression with and without interaction terms: A linear multiple regression model was also developed since statistical analysis offers more insights for causal inference. In this model, interaction terms have been introduced (the product of two variables).

The output variable in all cases was the logarithm, with a base of 10, of hydraulic conductivity (m/s). The logarithm was used since there is up to five-six orders of magnitude difference between the maximum and minimum value of hydraulic conductivity in the developed database. Also, this is a way of avoiding the prediction of negative values which is not realistic for this application. The hyperparameters of each ML algorithm were tuned on a training set that consisted of 70% of the whole dataset. The linear regression model was also developed with the use of this dataset utilising a stepwise backward elimination approach: the model was initially constructed using all input variables and then the insignificant variables were removed one-by-one based on the *p*-values (a variable with a p-value above 0.05 is considered to be insignificant). A hold-out sample or a test set which accounted for the other 30% of the whole dataset was tested against the trained models. The data were split into training and testing sets after being shuffled to eliminate any inherent ordering or biases that may be present and which may have disrupted any patterns or sequences in the data.

Ensemble learning algorithms were used to develop models by merging the predictions of the five different models (RF, XgBoost, SVR, kNN, linear model with interaction terms) utilising two methods:

1. Weighted average model (EnL-WA): The five models were assigned various weights defining the contribution of each model. The optimum weights were selected based on the lower error value on the training set.
2. Stacking technique (EnL-Stack): This method uses the predictions of the five models to build a new model. In this study, the new model was selected to be another Random Forest (RF).

The performance of the models was assessed via the root mean square error (RMSE), the mean absolute error (MAE) and the coefficient of determination (R-squared).

RMSE shows how far predictions fall from measured true values using Euclidean distance. RMSE is defined as:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\widehat{y_i} - y_t)^2}{n}} \quad (1)$$

MAE also measures the distance between predicted and actual values as shown in Eq. (2).

$$MAE = \sum_{i=1}^{n} \frac{|\widehat{y_i} - y_t|}{n} \quad (2)$$

R-squared measures how well the model's predictions explain the variation in the actual values and is defined as:

$$R^2 = 1 - \frac{RSS}{TSS} \quad (3)$$

where $\hat{y}_i$ are the predicted values, y_i, the actual values of the output variable, n is the number of observations, RSS is the sum of squares of residuals and TSS is the total sum of squares.

2.5. Inference of models and causality analysis

The parameter importance was used to identify which variables affect the outcome while the interactions between covariates have been also used to examine the synergistic/combined effects of pairs of variables on the output variable which is the logarithm of hydraulic conductivity. The effects of the input parameters (environmental and biochemical factor effects) on the output variable have been assessed via partial dependence plots (PDPs). PDPs provide the marginal effect of an input variable on the predicted outcome while keeping the other input parameters constant. As discussed in Fig. 2, the optimum bio-chemical parameters were used to derive permeability reduction versus cementation level profiles for various grain properties (i.e., grain diameters, widths of particle size distributions, grain shapes). Finally, the PandaLake model for permeability reduction was compared to the ML models against experimental data.

3. Results

3.1. Exploratory analysis

Fig. 3 displays histograms of the input variables. Most experiments were conducted using salt-free water, with approximately one-third utilising bacterial fixation, and a smaller number involving bacterial suspensions mixed with cementation solution. *S. pasteurii* was the most frequently used bacterial strain, with a population OD_{600} between 1 and 2. Urease activity ranged from 0.03 to 50 mM/min, with values between 0.03 and 10 mM/min being the most common. Calcium chloride concentration was mostly 0.25- 2 M, and the ratio of urea to calcium chloride concentration was usually 1.0. Retention time between injections varied considerably, with treatment durations lasting a few days for continuous injections. Over half of the studies used gravity injections, while controlled flow rate injections typically had an injection rate of 10 mL/min. Other injection methods were used in around 200–300 entries, such as immersing the granular media in solutions or pressure-controlled injections. Most studies also utilised full saturation. The initial porosity ranged from 0.2 to 0.5, and few studies included fibres, bentonite, silt, or other inclusions to increase specimen strength, but this would also affect hydraulic conductivity. Grain diameter, grain shape, and coefficient of uniformity (C_u) covered a wide range of soil properties typically found in shallow ground where groundwater is present. Finally, the cementation level indirectly signifies the number of injections, or the total volume of chemicals required to achieve lower hydraulic conductivity.

A correlation test was conducted to detect and eliminate any issues related to multicollinearity. Some pairs of variables exhibit moderate to high correlations (r-values around 0.7 and above or − 0.7 and under), including BsA and BgD, BsA and BsC, Sat and BsC. As most experiments utilised the bacterial strain *S. pasteurii* represented by variable BsA, other variables representing different strains were eliminated due to their high correlations. Therefore, the only variable which characterised the bacterial strain and was incorporated in the models was BsA.

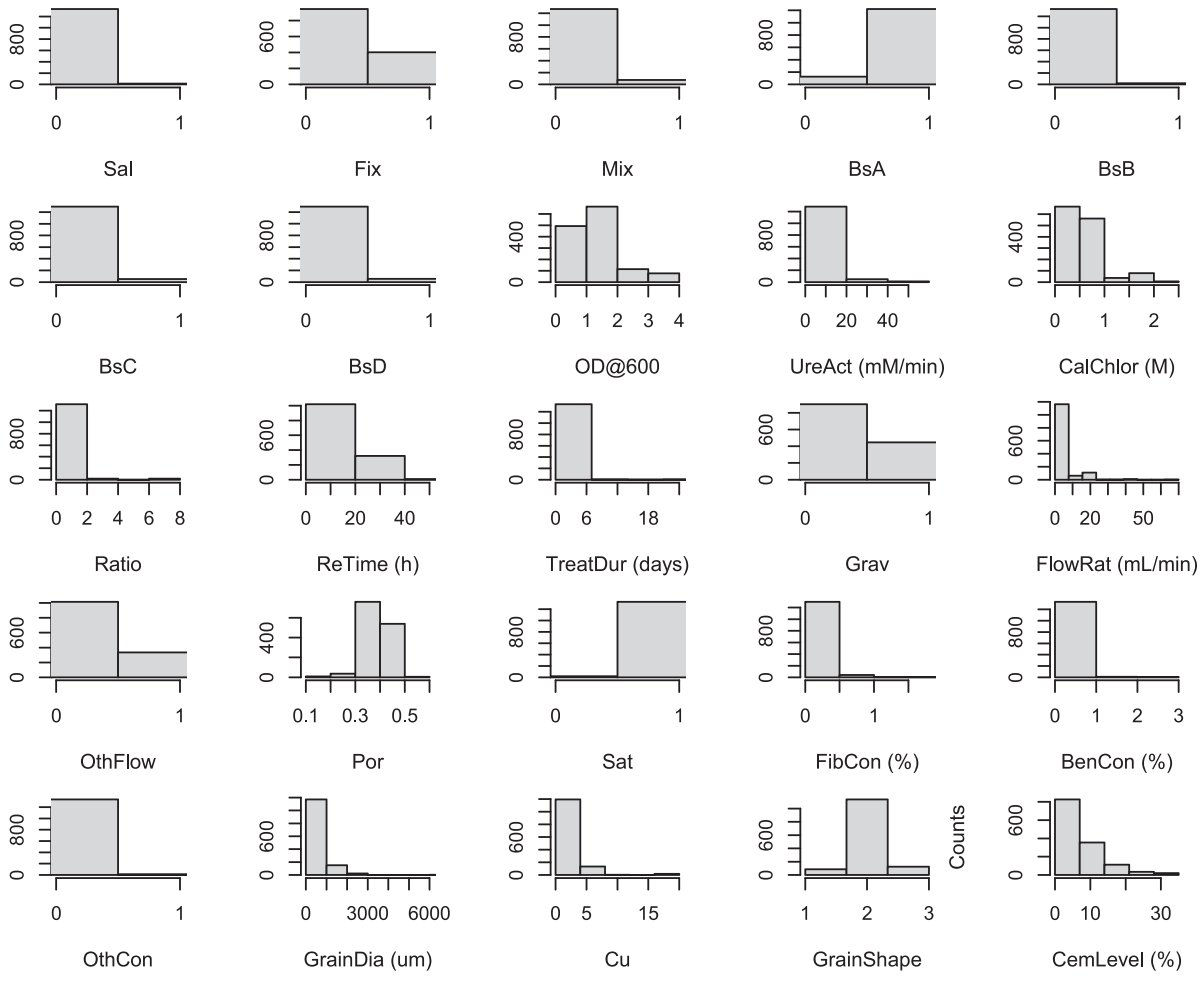


Fig. 3. The histograms of the input variables. The y-axis represents the counts of the entries.

3.2. The performance of the statistical and ML models

The performance of the models was assessed via the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) metrics. The RMSE and MAE values provide a good indication of the performance of a regression model and the smaller they are, the better the fit is. The RF and XgBoost algorithms perform well on the testing sets as shown in Fig. 4 with RMSE values of 0.305 and 0.324, respectively. The SVR and kNN algorithms follow with reasonable fits as the RMSE values are 0.341 (MAE = 0.218) and 0.426 (MAE = 0.274), respectively. The former underpredicts some hydraulic conductivity values, while the latter overpredicts some values. Finally, the linear multiple regression model underperforms giving an RMSE of 0.671. Generally, the different algorithms perform well and seem to capture the complexities of the problem. The problem is highly non-linear as the multiple regression algorithm underperforms mainly by overestimating the permeability values at the lowest range. This suggests that the addition of interaction terms in the linear model could enhance its performance, which is discussed in Section 3.2.

The five models (RF, XgBoost, SVR, kNN, linear model with interaction terms) were used to develop the two ensemble learning algorithms (see Fig. 4 (g-h)). The weighted average model (EnL-WA) provides a similar performance to the original RF model with an RMSE value of 0.296 and a MAE of 0.196. The optimum weights are found to

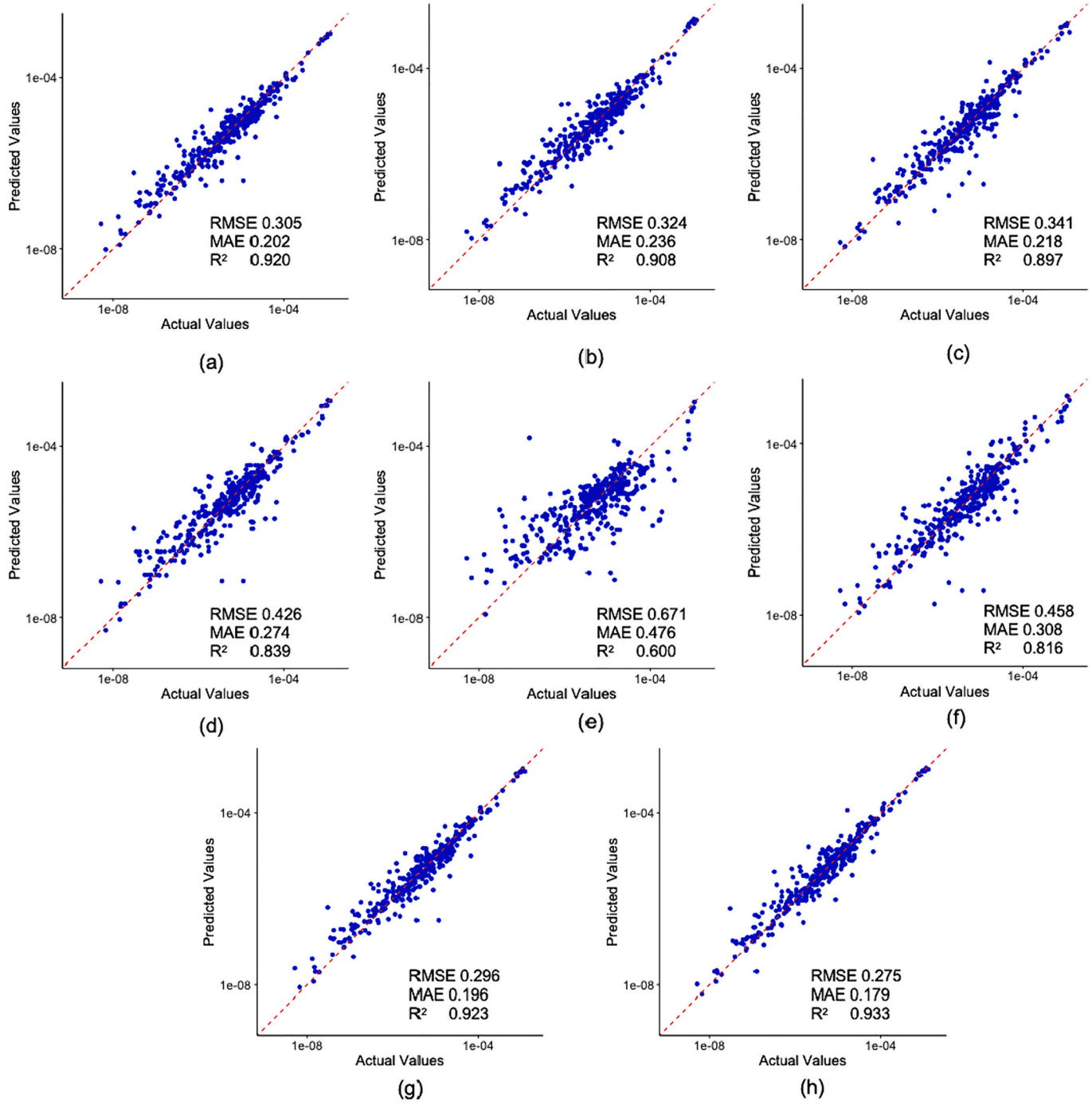


Fig. 4. Predicted Vs. Actual permeability values (testing set) for (a) the RF, (b) the XgBoost, (c) the SVR, (d) the kNN models, (e) the linear model, (f) the linear model with interaction terms, (g) the EnL-WA and (h) the EnL-Stack ensemble models.

be 0.69 for the RF model and 0.31 for the SVR without any contribution from the rest of the models. The ensemble algorithm with the stacking technique (EnL-Stack) provides a better performance compared to the individual RF model with an RMSE of 0.275 and a MAE of 0.179. However, the individual RF and XgBoost models are adequate in terms of their performance as the differences are marginal when compared to EnL-Stack.

3.3. Parameter importance and variable interactions

The importance of input variables is used to identify the contribution of each variable to the alteration of hydraulic conductivity. Fig. 5 shows the parameter importance for each of the four ML models. The most influential parameters for the RF and XgBoost algorithms (which perform better) are the uniformity coefficient, grain mean diameter, initial porosity, cementation level, flow rate, bacterial optical density, calcium chloride concentration, and urease activity. The other two algorithms provide similar results, although in a different order. The SVR algorithm shows that bentonite and fiber contents contribute to some extent on the value of the output variable while kNN ranks grain shape as an important factor. Furthermore, it appears that in addition to the biochemical parameters, which are clearly significant in all models, one variable that depends on time also impacts permeability (in some models FlowRat while in others this is the OthFlow). Despite the fact that the linear model provides poorer fits compared to the ML algorithms, it also gives similar results with respect to variable significance. The most significant variables (with a p-value of <0.0001) are Cu, GrainDia, CemLevel, GrainShape, Por, FlowRat, OD, CalChlor, Ratio, Mix.

The difference in the relative importance of the input variables observed between the various models can be explained by the fact the algorithms have different underlying principles. RF and XgBoost algorithms are complex models capable of fitting intricate patterns in the data resulting in a broader distribution of importance across features. By contrast, SVR and kNN aim to find a simpler, high-dimensional hyperplane or cluster, which may prioritize a subset of features more heavily. The choice of various ML models, results in building confidence about the interpretation of the problem under examination.

The variables' interaction was further examined for the two best performing algorithms (RF and XgBoost algorithms). For the RF model, the Mean Minimal Depth is used to identify the role of the variable within the structure and predictive capacity of a random forest. It corresponds to the depth of the node that splits on that variable and is closest to the root of the tree. A low Mean Minimal Depth indicates that a substantial number of observations are being divided into groups based on a particular variable. The strength of interactions between covariates are also reflected through the mean minimal depth, which represents the mean conditional minimal depth: a variable is taken as a root node or root variable, and the Mean Minimal Depth is calculated for the other variable. For the XgBoost algorithm, interactions between variables exist on pairs of variables where the lower (child) variable exhibits a higher gain than the upper (parent) variable. This assessment is grounded in the 'sumGain' metric, which represents the cumulative gain value across all nodes where a particular variable occurs.

As seen in Fig. 6, for both ML algorithms the properties of the porous media and the final cementation level demonstrate high levels of interactions since the hydraulic conductivity is first defined by the pore size distribution within the porous network. Then, interactions appear between the bio-chemical parameters and the porous media properties. For the RF model these are Cu:OD, Cu:CalChlor, Cu:UreAct, Cu:ReTime, Cu:FlowRat, GrainDia:OD, GrainDia:UreAct. For the XgBoost algorithm the interactions appear mainly with respect to the cementation level. The cementation level shows the amount of carbonate crystals within the porous network, however, the size, type and distribution of these carbonate crystals are defined by the bio-chemical parameters such as bacterial optical density and urease activity, chemicals concentration and time between subsequent injections (Wang et al., 2023b). These are CemLevel:UreAct, CemLevel:FlowRat, FlowRat:GrainDia, BsA:CemLevel, OD:CemLevel, CalChlor:FlowRat, CemLevel:Ratio, Por:CemLevel, CemLevel:GrainDia.

Interaction terms were also added to the linear multiple regression model which are defined as the product between two variables. The modified model provides good fits on the testing dataset as shown in Fig. 4 (f), with an RMSE value of 0.45 which is much lower compared to the initial linear model with an RMSE value of 0.67, demonstrating the highly non-linear nature of the investigated application. It also shows the importance of the interactions across the different variables especially those relating to the bio-chemical parameters and the porous media properties. Some of the most significant interactions occurring in the linear model that includes interaction terms are: CalChlor:Ratio, CalChlor:FlowRat, CalChlor:Por, FlowRat:GrainShape, Por:GrainDia, Por:Cu, GrainDia:CemLevel, GrainDia:GrainShape, OD:Por, Mix:Por.

4. Discussion

The developed models provide reasonable fits to the data and therefore a causality analysis is conducted to examine how the input variables affect permeability. The effects of the environmental and biochemical parameters on the effectiveness of MICP on reducing permeability are quantified with the use of the best performing algorithm (RF). An optimised injection schedule is identified based on the soil's properties while also the permeability reduction profiles with respect to

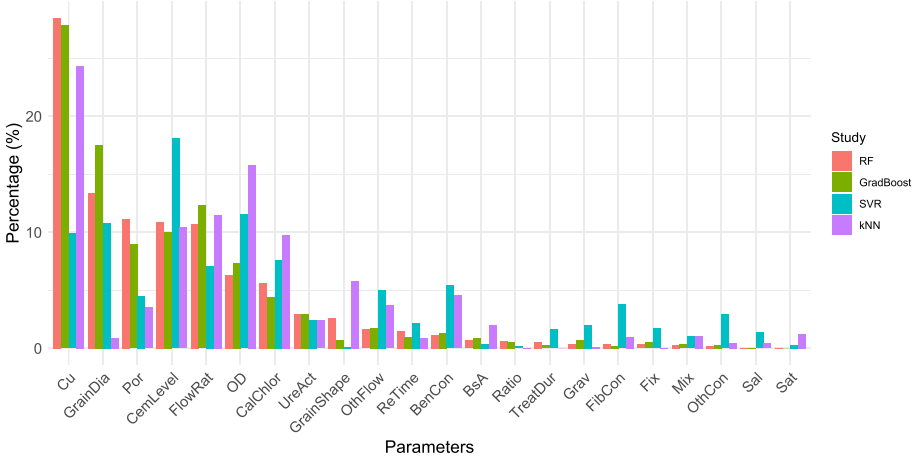
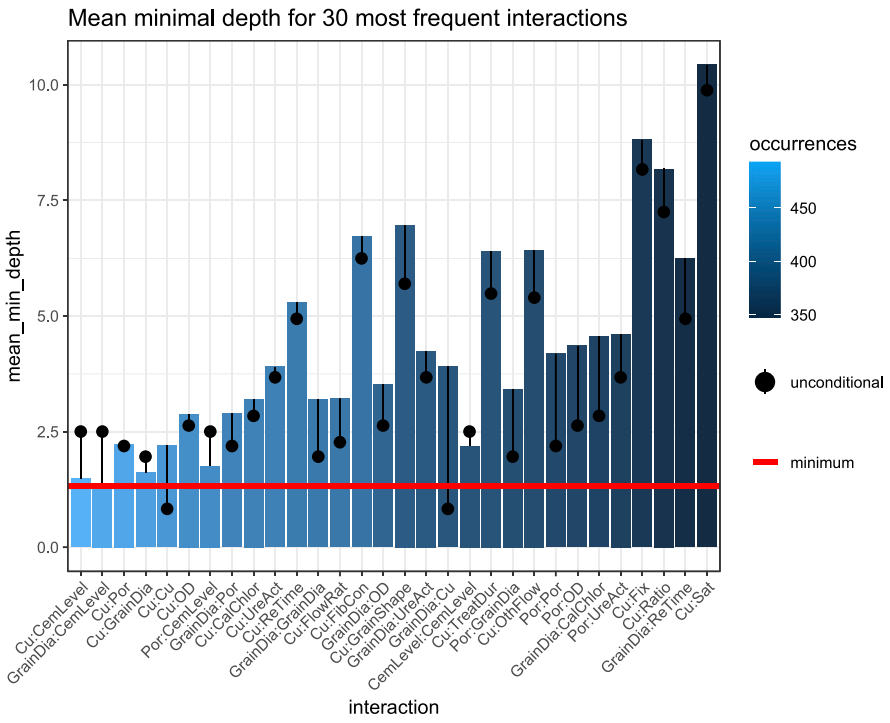
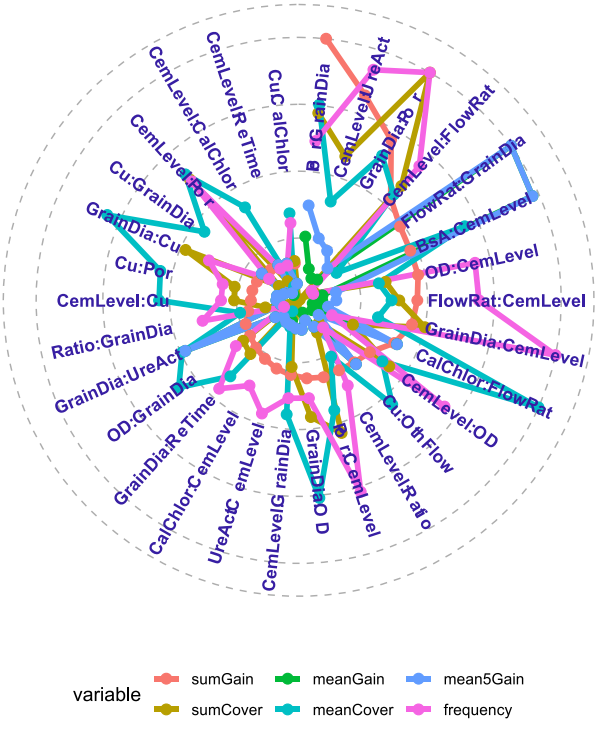


Fig. 5. The parameter importance plots for the different ML algorithms.



(a)



(b)

Fig. 6. (a) Parameter interactions for the RF: the interactions are organized on the plot based on frequency, with the most frequent ones depicted on the left in lighter blue and the least frequent ones on the right in darker blue. The horizontal red line indicates the minimum mean_min_depth, while the uncond_mean_min_depth is represented by the black lollipop. (b) Parameter interactions for the XgBoost algorithm: the plot displays positions arranged in descending order. The variable exhibiting the greatest sumGain value is situated to the right of 12 o'clock. Subsequently, the 'sumGain' values decrease in a clockwise fashion. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

cementation level for various potential porous media properties are produced. Finally, the ML models are compared to the empirical model Panda-Lake which is a modification of the Kozeny-Carman equation for permeability prediction.

4.1. Effects of bio-chemical and environmental parameters

Partial dependence plots (PDPs) are utilised to investigate the relationship between the various parameters and the output variable for the RF model. The PDPs shown in Fig. 7 contain the corresponding input on the x-axis, whilst the y-axis displays power of 10 value of the permeability (for a number 'x' the corresponding value of permeability is 10^{x} m/s). As the cementation level, the bacterial density, the calcium chloride, the ratio of $CaCl_2$ and urea and retention time increase a reduction in hydraulic conductivity is observed. As the flow rate increases the resulting permeability of the bio-cemented soil increases. Marginal effects on permeability take place for a cementation level above 15%, a urease activity of 10 mM/min, a $CaCl_2$ concentration above 1.5 M, a ratio above 4, flow rate above 10 mL/min and a retention time above 24 h. Mixing of bacterial with cementation solution and injection via gravity contribute to the reduction of hydraulic conductivity.

The addition of any content within the granular network contributes to the reduction of permeability. The MICP method can be applied in saline or sterile conditions, saturated or unsaturated conditions since the effects of the effectiveness of the method are marginal.

4.2. Optimization of bio-chemical parameters

To explore the potential of MICP in reducing hydraulic conductivity under various pore space distributions and identify the optimal combination of parameters, various scenarios involving grain characteristics were analysed. The impacts of the mean grain diameter, the width of particle size distribution (uniformity coefficient) and the particle shape were examined separately by taking into consideration various realistic porous media of different properties. The RF model was used to identify the combinations of bio-chemical parameters (OD_{600}, TreatDur, ReTime, FlowRat, OthFlow, Grav, CalChlor, Ratio, Mix, BsA) that provide the highest reduction of hydraulic conductivity. The rest of the values are fixed (Sal is 0, Fix is 0, Sat is 1, FibCon is 0, BenCon is 0, CemLevel is 20%).

The findings indicate that for subrounded grains of medium to high grain sizes, a high optical density (above 2) and urease activity of 10 mM/min are required while the injection mode is via gravity (infiltration) which is the highest possible flow rate without dislocating (moving) the particles or generating conductive channels. Such high flow rates result to uniform reactions (those that occur on all surfaces of the pore structure) resulting in a more uniform reduction of porosity and permeability spatially (Hao and Xu, 2023; Menke et al., 2016). The bacterial mixture should be injected together with cementation solution to increase the probability of carbonate crystals attaching to the pores. Additionally, a high concentration of calcium chloride (about 3 M) with a ratio of urea to calcium chloride of about 2–3 is suggested based on the RF model predictions with a retention time (time between subsequent injections) of 24 h. Longer retention times have been shown to result in greater decrease of hydraulic conductivity in other non-bacterial induced precipitation studies (Guo et al., 2024b). In the case of very small grain sizes (≈20 μm) the RF model suggests a calcium chloride concentration is 0.5 M with an equimolar ratio of calcium chloride and urea, and retention times between 1 and 3 h. Such small pore sizes are susceptible to clogging and this formulation would inhibit these processes around the injection point (Jaho et al., 2016). The mode of injection changes to pressurised injection (OthFlow), which might be due to the smaller pore network. Lower concentrations of calcium chloride result in the precipitation of smaller carbonate crystals which are needed for such small pore spaces (Wang et al., 2022). More frequent injections are allowed for such small concentrations because less time is needed for reactions to take place (Konstantinou et al., 2021b). For both spherical and angular grains, the recipe remains the same suggesting no effects.

As the width of the particle size distribution increases (uniformity coefficient), the retention time between injections decreases substantially, which may help to avoid pore clogging that could be caused by the

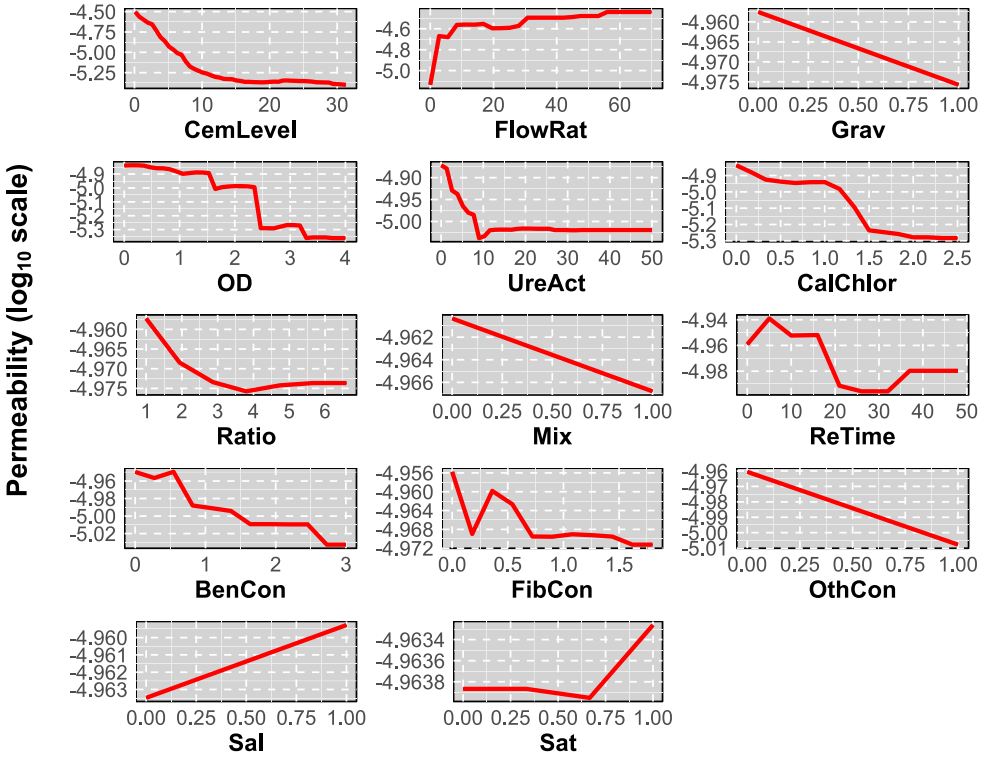


Fig. 7. The partial dependence plots.

injection in a smaller pore network especially close to the injection point. An equimolar recipe of calcium chloride and urea is also suggested and the mode of injection changes to pressurised injection (OthFlow), which might be due to the smaller pore network, similar to the case where grain sizes are smaller and hence the porous network is smaller.

4.3. Permeability curve with respect to cementation level

A general profile of hydraulic conductivity reduction with respect to cementation level is proposed by Song et al. (2020), identifying three phases. Phase I corresponds to low cementations where permeability decreases at a slower rate. This is followed by the rapid decline in permeability at moderate cementation levels (phase II). Finally, in phase III, which corresponds to high cementations, it remains almost constant. The findings of this study are compared to this general profile: the optimum bio-chemical parameters identified in the previous section are applied for the various porous media to derive the permeability curve with respect to cementation level as shown in Fig. 8.

The three phases are identified for the case of soils with a lower degree of sorting (narrower width of PSD) across all grain sizes (Fig. 8 (a)). For very small grain sizes, phase II shows a smaller decline in permeability. For very large grain sizes, permeability profile remains the same across different grain diameters, while for intermediate grain sizes the profile shifts either downwards or upwards. As the uniformity coefficient (Cu) increases for a fixed grain diameter, the profile of hydraulic conductivity changes significantly. At intermediate uniformity coefficients (see Fig. 8 (b) – Cu of 6), permeability still exhibits a large reduction even in phase III. At higher uniformity coefficients (Cu of 15) the profile shifts downwards. This is attributed to the higher pore scale heterogeneity observed in sands of wider PSDs, which exhibit larger reductions in permeability (Masoudi et al., 2024; Noiriel et al., 2016). The observed response is linked to the microscale and is further discussed in Section 4.4. Finally, the grain shape does not have significant effects on the permeability curve as shown in Fig. 8 (c) indicating that an MICP strategy for controlling SWI is not affected by the grain shape.

4.4. Comparison of predictions with theoretical models for calculating permeability

The predictions provided by the RF model are compared against empirical models to predict permeability and examine the performance of the ML algorithms based on a range of cementation levels. The aim is to provide an alternative way to estimate the reduction of hydraulic conductivity using a different set of input data through the ML approach. Porous media with more complicated properties (i.e., coarse grains, various grain shapes and widths of particle size distributions - PSDs) were selected from the work of the same authors of this study and the data points are presented in Fig. 9 (Konstantinou et al., 2023c) The presented SEM images are used explicitly in this study and are not presented in previous works.

There are many permeability-porosity relationships in the literature accounting for dissolution and precipitation mechanisms. The most commonly used are the Kozeny-Carman equations and their modifications (such as the Civan, Marshall, Taylor, Panda-Lake models) and the power law relationships with their modifications (Beckingham, 2017; Guo et al., 2024b; Hommel et al., 2018; Nogues et al., 2013; Sabo and Beckingham, 2021). The relationship between porosity and permeability is influenced not only by variations in pore system characteristics but also by the underlying mechanisms governing porosity alterations (Verma and Pruess, 1988). In MICP, where carbonate precipitates, either the pore throat geometry or the pore body geometry may be heavily affected by the process. Whether the former or the latter occurs, depends on the initial pore structure, as well as the size, shape and frequency of the precipitating crystals which, in turn, depend on the MICP bio-chemical parameters and injection methods. Specifically, in MICP, the pore throat geometry is affected when carbonate crystals land at the contacts between particles, while the pore body geometry is altered when cement lands in the pore network. There is also a case in-between which involves surface coating (cementation appears around a grain). Among the different models, the Panda-Lake model, a modification of the Kozeny-Carman equation, is appropriate for estimating the permeability of cemented soils by considering three different cementation distributions (pore filling, bridging and lining) which aligns with the carbonate crystals resulting from MICP. The Panda-Lake model (Panda and Lake, 1995; Panda and Lake, 1994) is expressed as:

$$k = \frac{\varphi_o^3}{SF\tau(1-\varphi_o)^2 a_v^2}\beta_\varphi\beta_\tau\beta_{a_v} \tag{4}$$

$$\beta_\varphi = \frac{\varphi^3(1-\varphi_o)^2}{\varphi_o^3(1-\varphi)^2}, \beta_\tau = \left(1+\frac{2S}{(1-S)\varphi_o^{0.33}}\right)^{-2}, \beta_{a_v} = \left(\frac{1-\varphi_o}{1-\varphi}+\frac{\alpha_{vc}}{a_v}Pc\right)^{-2} \tag{5}$$

where k is permeability, φ_o is the initial porosity, SF is the shape factor (taken as 3), τ is tortuosity, a_v is the specific surface area, φ is the porosity after cementation is applied, P_c is the proportion of $CaCO_3$ volume to the total volume of solids, a_{vc} is the specific surface area of the $CaCO_3$ crystals and S the cement saturation of the pore space. Tortuosity and specific surface area are given by Eqs. (6 and 7), respectively.

$$\tau = 2.5\left(1+\left(\frac{\sigma}{D_p}\right)^2\right) \tag{6}$$

$$a_v = \frac{6\left(\sigma^2+D_p^2\right)}{\gamma\sigma^3+3D_p\sigma^2+D_p^3} \tag{7}$$

where D_p, σ and γ are the statistical mean of the particle diameter, the

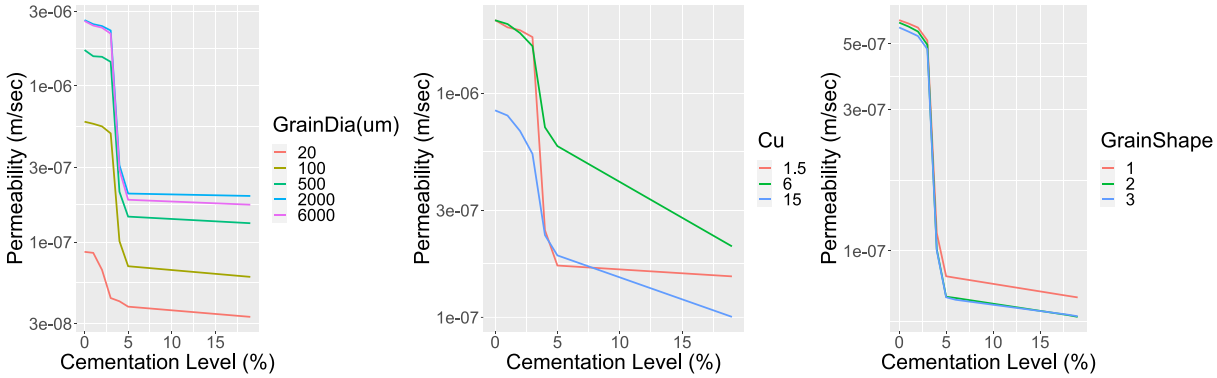
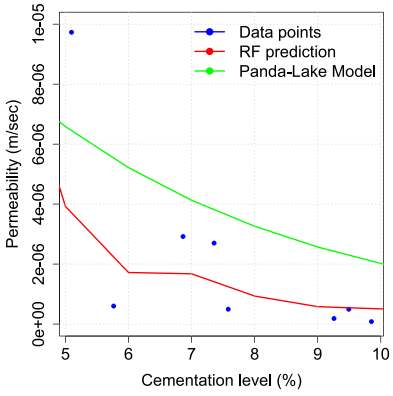
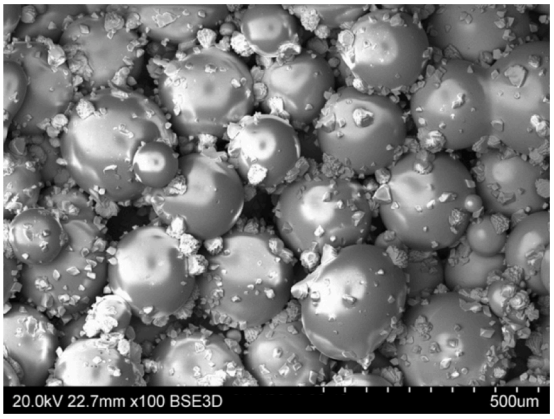
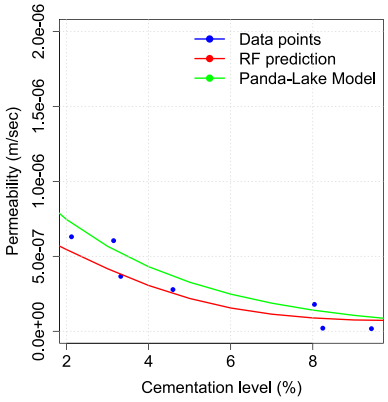


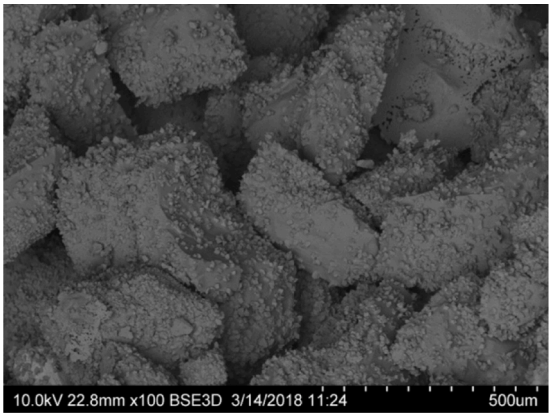
Fig. 8. Permeability profiles with respect to cementation levels for (a) various grain diameters, (b) various uniformity coefficients and (c) the three grain sizes.



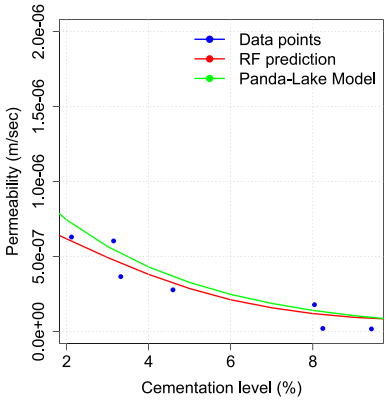


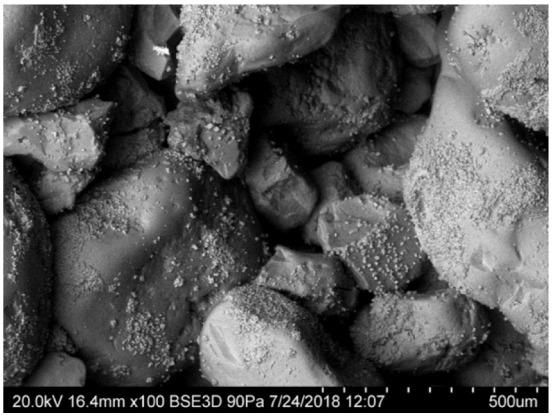
(a)





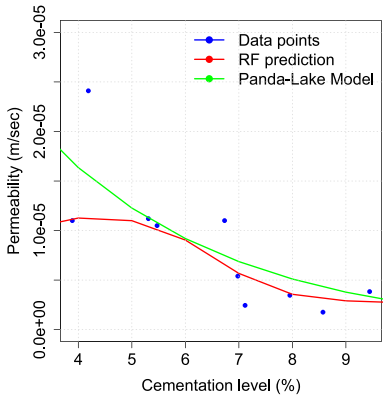
(b)



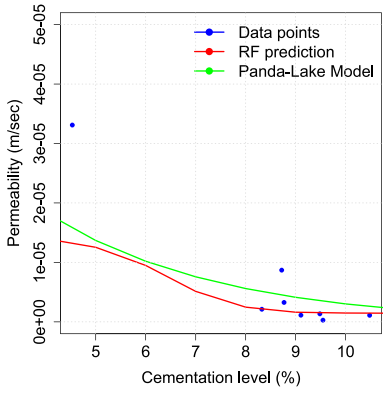


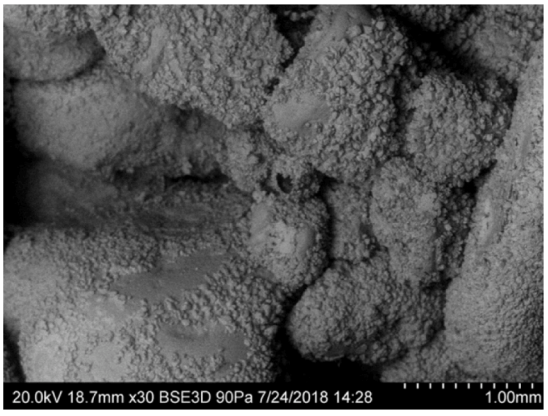
(c)

Fig. 9. The experimental data and predictions of the Panda-Lake model and the RF algorithm along with SEM images for (a) glass beads (D_{50} = 240 μm), (b) subangular sands (D_{50} = 372 μm) (c-e) various widths of particle size distribution (D_{50} = 371 μm and C_u = 2.22, D_{50} = 1180 μm and C_u = 5.6, D_{50} = 2454 μm and C_u = 4.62).



(d)





(e)

Fig. 9. (*continued*).

standard deviation of the particle size distribution (PSD) and the statistical skewness, respectively. The unit weight of water was taken as 9.8 kN/m^3 and the viscosity as 0.0013 Pa*sec.

The Panda-Lake model is a good fit for the experimental data points in all selected porous media, except in the case of glass beads in which the permeability is smaller compared to the predictions of the model (Fig. 9 (a)). This difference is attributed to the different cementation characteristics resulting by the MICP protocol in combination with the initial pore network characteristics. The shape of the grains affects the properties of the crystals that form. In glass beads, the crystals are fewer in number but larger in size, whereas for angular grains, the crystals tend to be smaller. The Panda-Lake model requires information on the distribution of cementation within the porous medium, including crystal size and precipitation. These characteristics are reflected in the specific surface area of the carbonate crystals in Eq. (4) (a_{vc}). Since in glass beads the carbonate crystal size tends to be larger, this term becomes smaller, resulting in a higher permeability value. Consequently, glass beads are not well represented by the Panda Lake model. The Panda-Lake model also failed to capture the high permeability value at low cementations of the porous media with the largest grain sizes and the largest width of PSD (Fig. 9 (f)). Generally, correlations such as the Kozeny-Carman and power laws along with their modifications (e.g., Panda-Lake, Fair-Hatch, Verma-Pruess) are inadequate in capturing the changes in permeability observed in pore size-controlled reactions. This dependency varies depending on the pore and pore-throat size distribution of the sample. Samples characterised by an abundance of small pores may undergo more pronounced changes in both porosity and permeability as shown in the pore-scale study by Beckingham (2017). Conversely, when reactions are confined to large pores and pore throats, substantial shifts in porosity may occur with little impact on permeability. Neither of these phenomena is adequately captured by the examined porosity-permeability relationships, necessitating the development of new correlations that account for the distinct porosity-permeability regions.

On the other hand, the RF algorithm is good at capturing the data points in the range of the moderately cemented specimens even in the areas where the Panda-Lake model failed to adequately predict the permeability values. However, it poorly predicts the permeability at low cementation levels in the cases where the soils had wider PSDs. As seen in the SEM images (Fig. 9 (c-e)), the porous network is composed of a high number of pore throats due to the higher particle interlocking, and thus even a small addition of cementation in these locations results in a greatest reduction in permeability. A similar result was observed in the study by Noiriel et al. (2016), in which the pores exhibit the development of pore-scale heterogeneity as a result of precipitation which in turn leads to a higher reduction in permeability compared to those that experienced uniform precipitation. This region is characterised as the region where permeability decreases at a slower rate, and it indicates that the data in the lower range needs to be populated and more information needs to be obtained to improve the performance of the model.

In uniform particles (Fig. 9 (a-b)), the carbonate crystals have equal probabilities to land on either the surface of the particles and on pore-throats resulting to a more steady reduction of permeability. The Panda-Lake model also revealed a great reduction in permeability even at very low cementation levels which is in contrast to the typical association between permeability reduction and cementation level.

As mentioned earlier, the Panda-Lake model relies on data regarding the distribution of cementation within the porous medium, including factors such as crystal size and precipitating patterns (surface coating or precipitation on particle grains). Obtaining this specific information can be challenging. On the other hand, machine learning (ML) models do require additional data, but it is generally easier to acquire, especially if geochemical parameters and properties of the porous media can be obtained through standard laboratory measurements. Thus, the ML models offer an alternative approach for delivering information relating to permeability using data of a different nature.

4.5. Limitations and significance of the developed models

When developing statistical and machine learning models, there are uncertainties associated with various steps. The highest source of uncertainty is the data collection, gap filling and homogenisation. In this study, the variable that was under-reported was mainly the urease activity. Even though the average behaviour of urease activity is well-documented in the literature, the accuracy of the models would increase dramatically if the findings from the different studies were reported. The optical density is another variable that introduces uncertainties. In cases where it was reported using the units of cfu/mL, standard curves from the literature were used, or estimations were made based on the specific size of the bacterial strain. However, the growth liquid used was not always the same across the studies, thus potentially making the value less accurate. Other challenges associated with the homogenisation of data could be the loss of information. For example, the chemical composition of seawater and the type of calcium source (calcium chloride, dihydrate, anhydrous), which are expected to affect the properties of the precipitated crystals, are not accounted for in this investigation.

Another issue is the fact that only a few studies investigated the addition of fibres, bentonite, kaoline, fines or other materials in their MICP formulation when reporting hydraulic conductivity measurements. Although some of those effects are reflected to some extent in the initial porosity variable, the additives significantly affect the pore network distribution and hence the resulting mechanical and hydraulic properties (Ma et al., 2022). Further work needs to be conducted to properly assess the effects of such additions, as the trained models have not been extensively exposed to such conditions. These advancements in the field need to go hand in hand with experimental investigation.

The bacteria used for the MICP technique should be carefully selected based on their ability to precipitate calcium carbonate and their compatibility with the site conditions. As demonstrated previously, the type of bacteria plays a minor role in the performance of MICP, as long as they produce urease. The bacteria should be able to grow in the presence of seawater and high salinity conditions, which is proven by few studies. However, local urease-producing strains from the site could be isolated and used for this application. Instead of injecting the calcium source externally, the site could have enough calcium ions to support the precipitation of calcium carbonate and the right environmental conditions to support the growth of the bacteria used for the technique. All these factors (i.e., salinity, bacterial strains, calcium source) need to be further studied since the developed ML models have not 'seen' such data based on the current studies. Therefore, such interpretations cannot be done until investigated experimentally.

One solution could be the use of synthetic data (e.g. synthetic minority oversampling technique for regression - SMOTER) which could be generated in such cases where the available dataset is unrepresentative of certain scenarios. However, in this study there are no regions of the target variable space where data is sparse and therefore, it would not benefit to a great extend the developed models. As discussed, the underlying issue is not class imbalance, rather it is the lack of information in other areas outside the investigated ones (e.g., larger bentonite or fiber content). This limitation arises from the inability of machine learning models to extrapolate and forecast outcomes beyond the range of the data they were trained on.

Despite these limitations, the ML and statistical models predict the permeability profiles across various cementation levels performing better than the available empirical models in the literature (i.e., Panda-Lake models) in many cases, which is achieved using less information without the need for specific variables which are not always known. Therefore, the models described in this study could be used as an initial screening tool, which could help to narrow down the available options and these could then be further investigated experimentally. At the same time, the models could be further used to identify patterns and link the hydraulic conductivity response to the microscale characteristics. This is a complex problem which is highly non-linear, with the analysis performed revealing correlations between the properties of the porous media and MICP parameters. This information could be further used to interpret various findings not only with respect to permeability but also with respect to strength enhancement.

4.6. Technological issues associated with the application of MICP for seawater intrusion

The advantage of the MICP technique for preventing seawater intrusion compared to other chemical techniques is that it is more environmentally friendly. However, there are technical issues which might affect the MICP the process when applied in the field which are discussed below.

This study shows that it is important to consider the injection method used to introduce the bacteria and the calcium solution into the aquifer. The injection method should ensure that the bacteria and calcium solution are distributed uniformly throughout the aquifer to form a continuous barrier. Monitoring the progress of the barrier formation is critical to ensure that the barrier is being formed as intended. Parameters such as pH, calcium concentration, and bacterial growth should be monitored regularly. These parameters also act as indicators for any change or deviation from the designed conditions. An example is the fact that the metabolic activity of bacterial strains might be affected by variations in environmental conditions such as salinity, temperature, and nutrient availability, which can affect the overall performance of MICP.

Maintenance of the barrier may require periodic injections of calcium and bacterial solutions to sustain its integrity. Designing subsea barriers that effectively prevent saltwater intrusion while allowing controlled replenishment of freshwater can be very complex in reality. Factors such as tidal fluctuations, wave action, and sediment transport dynamics need to be considered. The environmental impact of the MICP technique should also be carefully evaluated to ensure that it does not negatively affect the environment or groundwater quality (one problem could be the alteration of the natural flow patterns of seawater and freshwater which could potentially disrupt the balance of ecosystems). By restricting the exchange of water between the aquifer and the sea, the natural replenishment and flushing processes may be disrupted. This can lead to changes in water quality, including the accumulation of pollutants or reduced availability of freshwater resources.

All of the above need to be monitored in the long-run to ensure that the effectiveness of subsea barriers continue to function as intended. Regular monitoring programs should be in place to assess the performance of the barriers and make any necessary adjustments or maintenance.

5. Conclusion

The growing demand for freshwater and hence groundwater over-exploitation, coupled with climate change and rising sea-levels, cause saltwater from coastal aquifers to invade groundwater aquifers. Seawater intrusion can lead to significant damage to the aquifer system, including degradation of water quality and loss of natural habitats. Therefore, it is crucial to develop effective techniques to control seawater intrusion.

The application of MICP for generating a barrier to prevent the invasion of seawater into groundwater has great potential since it reduces the hydraulic conductivity. In this study, machine learning and statistical analyses were implemented to build models to identify patterns between the biochemical factors affecting MICP and the properties of the porous medium (together with environmental factors) based on the literature. The ML models (k-Nearest Neighbour, Support Vector Regression, Random Forests, Gradient Boosting) and the modified linear regression model (accounting for interactions) provide reasonable fits. The RF algorithm is the best performing algorithm with an RMSE of 0.305 followed by XgBoost and SVR. The ensemble learning model EnL-Stack which merges the five previous models provides better fits compared to the individual RF algorithm. These models can be used as tools to predict the reduction in hydraulic conductivity under specific conditions.

According to the parameter importance plots, the MICP formulation depends on the characteristics of the porous medium and the most significant biochemical parameters seem to be the bacterial optical density, the urease activity, the calcium chloride concentration, and the flow rate while there are several interaction terms between the porous medium properties and the biochemical parameters that need to be considered.

Since the ML algorithms provide good fits for the data, they can be used in MICP procedures designed for these applications as a guidance for further analysis at the lab scale to fine-tune the formulation. In this study, an optimised injection schedule was identified based on various soil properties while also the permeability reduction profiles with respect to cementation level for various potential porous media properties were produced. The findings suggest specific conditions for effective injection of cementation solutions into porous media. For medium to high grain sizes with subrounded grains, a high bacterial optical density (>2), urease activity of 10 mM/min and a high cementation solution concentration are necessary, along with gravity-based injection of bacterial and cementation solution mixture and a retention time of 24 h between injections. For smaller grain sizes ($\sim$20 μm), a lower calcium chloride concentration with equimolar ratios is suggested, along with shorter retention times ($1-3$h) and pressurised injection. As the particle size distribution widens, shorter retention times are advised to prevent pore clogging, with similar adjustments in calcium chloride concentration and injection mode. The recommendations remain consistent for both spherical and angular grains.

The ML models were also compared to the empirical model Panda-Lake which is a modification of the Kozeny-Carman equation for permeability prediction showing similar performance and, in some cases, better performance. Since the Panda-Lake model requires specific information on the carbonate crystal characteristics which is often difficult to obtain, the ML models could be used as a screening tool to assess MICP applicability for seawater intrusion.

CRediT authorship contribution statement

Charalampos Konstantinou: Writing – original draft, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Yuze Wang:** Writing – review & editing, Methodology, Investigation, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

Charalampos Konstantinou reports financial support was provided by Cyprus Research Promotion Foundation. Yuze Wang reports financial support was provided by Science and Technology Innovation Committee of Shenzhen. Yuze Wang reports financial support was provided by Natural Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

C.K. acknowledges funding from the European Regional Development Fund and the Republic of Cyprus through the Research Promotion Foundation (RPF) (Cyprus RPF, RESTART 2016–2020 PROGRAMMES, Excellence Hubs, Project FLINUGEE, EXCELLENCE/0421/0456). Y.W. acknowledges the support of the Science and Technology Innovation Committee of Shenzhen (Grant No. JCYJ20210324103812033) and Natural Science Foundation of China (Grant Nos. 52171262).

References

Abdoulhalik, A., Ahmed, A., Hamill, G.A., 2017. A new physical barrier system for seawater intrusion control. J. Hydrol. 549, 416–427. https://doi.org/10.1016/j.jhydrol.2017.04.005.

Akoğuz, H., Çelik, S., Bariş, Ö., 2019. The effects of different sources of calcium in improvement of soils by microbially induced calcite precipitation (MICP). Sigma J. Eng. Nat. Sci. 37, 953–965.

Al Qabany, A., Soga, K., 2013. Effect of chemical treatment used in MICP on engineering properties of cemented soils. Geotechnique 63, 331–339. https://doi.org/10.1680/geot.SIP13.P.022.

Baek, S.-H., Kwon, T.-H., DeJong, J.T., 2024. Reductions in hydraulic conductivity of sands caused by Microbially induced calcium carbonate precipitation. J. Geotech. Geoenviron. Eng. 150 https://doi.org/10.1061/JGGEFK.GTENG-11570.

Beckingham, L.E., 2017. Evaluation of macroscopic porosity-permeability relationships in heterogeneous mineral dissolution and precipitation scenarios. Water Resour. Res. 53, 10217–10230. https://doi.org/10.1002/2017WR021306.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Chang, Q., Zheng, T., Zheng, X., Zhang, B., Sun, Q., Walther, M., 2019. Effect of subsurface dams on saltwater intrusion and fresh groundwater discharge. J. Hydrol. 576, 508–519. https://doi.org/10.1016/j.jhydrol.2019.06.060.

Chen, Y., Han, Y., Zhang, X., Sarajpoor, S., Zhang, S., Yao, X., 2023. Experimental study on permeability and strength characteristics of MICP-treated calcareous sand. Biogeotechnics 1, 100034. https://doi.org/10.1016/j.bgtech.2023.100034.

Cheng, L., Cord-Ruwisch, R., Shahin, M.A., 2013. Cementation of sand soil by microbially induced calcite precipitation at various degrees of saturation. Can. Geotech. J. 50, 81–90. https://doi.org/10.1139/cgj-2012-0023.

Cheng, L., Shahin, M.A., Cord-Ruwisch, R., 2014. Bio-cementation of sandy soil using microbially induced carbonate precipitation for marine environments. Geotechnique. https://doi.org/10.1680/geot.14.T.025.

Cheng, L., Shahin, M.A., Mujah, D., 2016. Influence of key environmental conditions on Microbially induced cementation for soil stabilization. J. Geotech. Geoenviron. Eng. 143, 04016083. https://doi.org/10.1061/(ASCE).

Choi, S.G., Wang, K., Chu, J., 2016a. Properties of biocemented, fiber reinforced sand. Constr. Build. Mater. 120, 623–629. https://doi.org/10.1016/j.conbuildmat.2016.05.124.

Choi, S.-G., Wu, S., Chu, J., 2016b. Biocementation for sand using an eggshell as calcium source. J. Geotech. Geoenviron. Eng. 142 https://doi.org/10.1061/(asce)gt.1943-5606.0001534.

Choi, S.G., Chu, J., Brown, R.C., Wang, K., Wen, Z., 2017. Sustainable biocement production via Microbially induced calcium carbonate precipitation: use of limestone and acetic acid derived from pyrolysis of lignocellulosic biomass. ACS Sustain. Chem. Eng. 5, 5183–5190. https://doi.org/10.1021/acssuschemeng.7b00521.

Choi, S.G., Hoang, T., Park, S.S., 2019. Undrained behavior of microbially induced calcite precipitated sand with polyvinyl alcohol fiber. Appl. Sci. 9 https://doi.org/10.3390/app9061214.

Choi, S.G., Chang, I., Lee, M., Lee, J.H., Han, J.T., Kwon, T.H., 2020. Review on geotechnical engineering properties of sands treated by microbially induced calcium

carbonate precipitation (MICP) and biopolymers. Constr. Build. Mater. 246, 1–14. https://doi.org/10.1016/j.conbuildmat.2020.118415.
Dawoud, O., 2020. Modification of hydraulic conductivity of Sandy soil using seawater and alkaline solutions. In: IOP Conference Series: Materials Science and Engineering. Institute of Physics Publishing. https://doi.org/10.1088/1757-899X/800/1/012011.
Dawoud, O., Chen, C.Y., Soga, K., 2014. Microbial Induced Calcite Precipitation for Geotechnical and Environmental Applications.
Dekuyer, A., Cheng, L., Shahin, M.A., Cord-Ruwisch, R., 2012. Calcium Carbonate Induced Precipitation for Soil Improvement by Urea Hydrolysing Bacteria. In: Advances in Civil, Environmental, and Materials Research (ACEM' 12)..
Duo, L., Kan-liang, T., Hui-li, Z., Yu-yao, W., Kang-yi, N., Shi-can, Z., 2018. Experimental investigation of solidifying desert aeolian sand using microbially induced calcite precipitation. Constr. Build. Mater. 172, 251–262. https://doi.org/10.1016/j.conbuildmat.2018.03.255.
Eryürük, K., 2022. Effect of cell density on decrease in hydraulic conductivity by microbial calcite precipitation. AMB Express 12. https://doi.org/10.1186/s13568-022-01448-0.
Fang, X., Yang, Y., Chen, Z., Liu, H., Xiao, Y., Shen, C., 2020. Influence of fiber content and length on engineering properties of MICP-treated coral sand. Geomicrobiol J. https://doi.org/10.1080/01490451.2020.1743392.
Fang, Y., Zheng, T., Wang, H., Guan, R., Zheng, X., Walther, M., 2021. Experimental and numerical evidence on the influence of tidal activity on the effectiveness of subsurface dams. J. Hydrol. 603 https://doi.org/10.1016/j.jhydrol.2021.127149.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 https://doi.org/10.1214/aos/1013203451.
Gago, P.A., Konstantinou, C., Biscontin, G., King, P., 2020. Stress inhomogeneity effect on fluid-induced fracture behavior into weakly consolidated granular systems. Phys. Rev. E 102, 1–5. https://doi.org/10.1103/physreve.102.040901.
Gomez, M.G., Anderson, C.M., Dejong, J.T., Nelson, D.C., Lau, X.H., 2014. Stimulating in-situ soil Bacteria for bio-cementation of sands. In: Geo-Congress, pp. 1674–1682.
Gong, X., Niu, J., Liang, S., Feng, D., Luo, Q., 2019. Environmental effect of grouting batches on microbial-induced calcite precipitation. Ekoloji 28 (107), 929–936.
Guo, L., Wang, B., Guo, J., Guo, H., Jiang, Y., Zhang, M., Dai, Q., 2024a. Experimental study on improving hydraulic characteristics of sand via microbially induced calcium carbonate precipitation. Geomech. Energy Environ. 37 https://doi.org/10.1016/j.gete.2023.100519.
Guo, C., Wu, J., Zhao, Q., Li, H., Yang, H., Wu, Z., Qin, S., 2024b. A model to quantify permeability in solute mixing precipitation porous media. J. Hydrol. 629, 130470 https://doi.org/10.1016/j.jhydrol.2023.130470.
Hao, H., Xu, Z.G., 2023. Pore-scale investigation on porous media morphology evolution considering dissolution and precipitation. Int. J. Multiphase Flow 168. https://doi.org/10.1016/j.ijmultiphaseflow.2023.104569.
Hastie, T., Tibshirani, R., Friedman, J., 2009. Elements of Statistical Learning, 2nd ed. Springer-Verlag. https://doi.org/10.1007/978-0-387-84858-7.
Hommel, J., Coltman, E., Class, H., 2018. Porosity–permeability relations for evolving pore space: a review with a focus on (bio-)geochemically altered porous media. Transp. Porous Media. https://doi.org/10.1007/s11242-018-1086-2.
Hu, L., Wang, H., Xu, P., Zhang, Y., 2021. Biomineralization of hypersaline produced water using microbially induced calcite precipitation. Water Res. 190 https://doi.org/10.1016/j.watres.2020.116753.
Jaho, S., Athanasakou, G.D., Sygouni, V., Lioliou, M.G., Koutsoukos, P.G., Paraskeva, C. A., 2016. Experimental investigation of calcium carbonate precipitation and crystal growth in one- and two-dimensional porous media. Cryst. Growth Des. 16, 359–370. https://doi.org/10.1021/acs.cgd.5b01321.
Jawad, F., Zheng, J.-J., 2016. Improving poorly graded fine sand with microbial induced calcite precipitation. Br. J. Appl. Sci. Technol. 17, 1–9. https://doi.org/10.9734/bjast/2016/28338.
Kadhim, F.J., Zheng, J.-J., 2017. Influences of calcium sources and type of sand on microbial induced carbonate precipitation. Int. J. Adv. Eng. Technol. 10, 20–29.
Kaleris, V.K., Ziogas, A.I., 2013. The effect of cutoff walls on saltwater intrusion and groundwater extraction in coastal aquifers. J. Hydrol. 476, 370–383. https://doi.org/10.1016/j.jhydrol.2012.11.007.
Kecman, V., 2005. Support vector machines – An introduction. In: Wang, L. (Ed.), Support Vector Machines: Theory and Applications. Studies in Fuzziness and Soft Computing. Springer, Berlin Heidelberg. https://doi.org/10.1007/10984697_1.
Kim, D.H., Mahabadi, N., Jang, J., van Paassen, L.A., 2020. Assessing the kinetics and pore-scale characteristics of biological calcium carbonate precipitation in porous media using a microfluidic chip experiment. Water Resour. Res. 56 https://doi.org/10.1029/2019WR025420.
Kirk, M.F., Santillan, E.F.U., McGrath, L.K., Altman, S.J., 2012. Variation in hydraulic conductivity with decreasing pH in a biologically-clogged porous medium. Int. J. Greenhouse Gas Control 11, 133–140. https://doi.org/10.1016/j.ijggc.2012.08.003.
Konstantinou, C., 2021. Hydraulic Fracturing of Artificially Generated Soft Sandstones. University of Cambridge. https://doi.org/10.17863/CAM.64233.
Konstantinou, C., Biscontin, G., 2022. Experimental investigation of the effects of porosity, hydraulic conductivity, strength, and flow rate on fluid flow in weakly cemented bio-treated sands. Hydrology 9. https://doi.org/10.3390/hydrology9110190.
Konstantinou, C., Stoianov, I., 2020. A comparative study of statistical and machine learning methods to infer causes of pipe breaks in water supply networks. Urban Water J. 17, 534–548. https://doi.org/10.1080/1573062X.2020.1800758.
Konstantinou, C., Wang, Y., 2023. Unlocking the potential of microbially induced calcium carbonate precipitation (MICP) for hydrological applications: a review of opportunities, challenges, and environmental considerations. Hydrology 10, 178. https://doi.org/10.3390/hydrology10090178.
Konstantinou, C., Biscontin, G., Jiang, N.-J., Soga, K., 2021a. Application of microbially induced carbonate precipitation (MICP) to form bio-cemented artificial sandstone. J. Rock Mech. Geotech. Eng. 13, 579–592. https://doi.org/10.1016/j.jrmge.2021.01.010.
Konstantinou, C., Wang, Y., Biscontin, G., Soga, K., 2021b. The role of bacterial urease activity on the uniformity of carbonate precipitation profiles of bio-treated coarse sand specimens. Sci. Rep. 11, 1–17. https://doi.org/10.1038/s41598-021-85712-6.
Konstantinou, C., Biscontin, G., Papanastasiou, P., 2022. Interpretation of fluid injection experiments in poorly consolidated sands. In: 56th U.S. Rock Mechanics/Geomechanics Symposium. OnePetro, Santa Fe, New Mexico, USA. https://doi.org/10.56952/ARMA-2022-0632.
Konstantinou, C., Kandasami, R.K., Biscontin, G., Papanastasiou, P., 2023a. Fluid injection through artificially reconstituted bio-cemented sands. Geomech. Energy Environ., 100466 https://doi.org/10.1016/j.gete.2023.100466.
Konstantinou, C., Martínez-Pañeda, E., Biscontin, G., Fleck, N.A., 2023b. Fracture of bio-cemented sands. Extreme Mech. Lett. 64 https://doi.org/10.1016/j.eml.2023.102086.
Konstantinou, C., Wang, Y., Biscontin, G., 2023c. A systematic study on the influence of grain characteristics on hydraulic and mechanical performance of MICP-treated porous media. Transp. Porous Media. https://doi.org/10.1007/s11242-023-01909-5.
Laabidi, E., Bouhlila, R., 2021. A new technique of seawater intrusion control: development of geochemical cutoff wall. Environ. Sci. Pollut. Res. 28, 41794–41806. https://doi.org/10.1007/s11356-021-13677-0.
Lambert, S.E., Randall, D.G., 2019. Manufacturing bio-bricks using microbial induced calcium carbonate precipitation and human urine. Water Res. 160, 158–166. https://doi.org/10.1016/j.watres.2019.05.069.
Li, Y., Chen, J., 2022. Experimental study on the permeability of microbial-solidified calcareous sand based on MICP. Appl. Sci. 12 https://doi.org/10.3390/app122211447.
Lin, H., Suleiman, M.T., Brown, D.G., 2020. Investigation of pore-scale CaCO3 distributions and their effects on stiffness and permeability of sands treated by microbially induced carbonate precipitation (MICP). Soils Found. 60, 944–961. https://doi.org/10.1016/j.sandf.2020.07.003.
Lin, Wenbin, Gao, Y., Lin, Wei, Zhuo, Z., Wu, W., Cheng, X., 2023. Seawater-based bio-cementation of natural sea sand via microbially induced carbonate precipitation. Environ. Technol. Innov. 29 https://doi.org/10.1016/j.eti.2023.103010.
Liu, B., Tang, C.S., Pan, X.H., Zhu, C., Cheng, Y.J., Xu, J.J., Shi, B., 2021. Potential drought mitigation through microbial induced calcite precipitation-MICP. Water Resour. Res. 57 https://doi.org/10.1029/2020WR029434.
Liu, J., Li, X., Li, G., Zhang, J., 2023a. Experimental study on the mechanical behaviors of Aeolian sand treated by Microbially induced calcite precipitation (MICP) and basalt Fiber reinforcement (BFR). Materials 16. https://doi.org/10.3390/ma16051949.
Liu, J., Li, X., Liu, X., Dong, W., Li, G., 2023b. Mechanical properties of Eolian sand solidified by Microbially induced calcium carbonate precipitation (MICP). Geomicrobiol J. 40, 688–698. https://doi.org/10.1080/01490451.2023.2235345.
Luyun, R., Momii, K., Nakagawa, K., 2009. Laboratory-scale saltwater behavior due to subsurface cutoff wall. J. Hydrol. 377, 227–236. https://doi.org/10.1016/j.jhydrol.2009.08.019.
Ma, G., He, X., Jiang, X., Liu, H., Chu, J., Xiao, Y., 2021. Strength and permeability of bentonite-assisted biocemented coarse sand. Can. Geotech. J. 58, 969–981.
Ma, G., Xiao, Y., He, X., Li, J., Chu, J., Liu, H., 2022. Kaolin-nucleation-based biotreated calcareous sand through unsaturated percolation method. Acta Geotech. 17, 3181–3193. https://doi.org/10.1007/s11440-022-01459-y.
Martinez, B.C., DeJong, J.T., Ginn, T.R., Montoya, B.M., Barkouki, T.H., Hunt, C., Tanyu, B., Major, D., 2013. Experimental optimization of microbial-induced carbonate precipitation for soil improvement. J. Geotech. Geoenviron. Eng. 139, 587–598. https://doi.org/10.1061/(asce)gt.1943-5606.0000787.
Masoudi, M., Nooraiepour, M., Deng, H., Hellevang, H., 2024. Mineral Precipitation and Geometry Alteration in Porous Structures: How to Upscale Variations in Permeability-Porosity Relationship? Mineral Precipitation and Geometry Alteration in Porous Structures: How to Upscale Variations in Permeability-Porosity Relationship? (EarthArXiv).
Menke, H.P., Andrew, M.G., Blunt, M.J., Bijeljic, B., 2016. Reservoir condition imaging of reactive transport in heterogeneous carbonates using fast synchrotron tomography - effect of initial pore structure and flow conditions. Chem. Geol. 428, 15–26. https://doi.org/10.1016/j.chemgeo.2016.02.030.
Montoya, B.M., Do, J., Gabr, M.M., 2018. Erodibility of microbial induced carbonate precipitation-stabilized sand under submerged impinging jet. In: IFCEE, pp. 19–28.
Montoya, B.M., Safavizadeh, S., Gabr, M.A., 2019. Enhancement of coal ash compressibility parameters using microbial-induced carbonate precipitation. J. Geotech. Geoenviron. Eng. 145, 04019018 https://doi.org/10.1061/(ASCE)GT.1943.
Mortensen, B.M., Haber, M.J., Dejong, J.T., Caslake, L.F., Nelson, D.C., 2011. Effects of environmental factors on microbial induced calcium carbonate precipitation. J. Appl. Microbiol. 111, 338–349. https://doi.org/10.1111/j.1365-2672.2011.05065.x.
Motallebian, M., Ahmadi, H., Raoof, A., Cartwright, N., 2019. An alternative approach to control saltwater intrusion in coastal aquifers using a freshwater surface recharge canal. J. Contam. Hydrol. 222, 56–64. https://doi.org/10.1016/j.jconhyd.2019.02.007.
Nasiri, M., Moghaddam, H.K., Hamidi, M., 2021. Development of multi-criteria decision making methods for reduction of seawater intrusion in coastal aquifers using SEAWAT code. J. Contam. Hydrol. 242 https://doi.org/10.1016/j.jconhyd.2021.103848.

Niu, J.G., Liang, S.H., Gong, X., Feng, D.L., Luo, Q.Z., Dai, J., 2018. Experimental study on the effect of grouting interval on microbial induced calcium carbonate precipitation. In: IOP Conference Series: Earth and Environmental Science. Institute of Physics Publishing. https://doi.org/10.1088/1755-1315/186/3/012071.

Nogues, J.P., Fitts, J.P., Celia, M.A., Peters, C.A., 2013. Permeability evolution due to dissolution and precipitation of carbonates using reactive transport modeling in pore networks. Water Resour. Res. 49, 6006–6021. https://doi.org/10.1002/wrcr.20486.

Noiriel, C., Steefel, C.I., Yang, L., Bernard, D., 2016. Effects of pore-scale precipitation on permeability and flow. Adv. Water Resour. 95, 125–137. https://doi.org/10.1016/j.advwatres.2015.11.013.

Onal Okyay, T., Frigi Rodrigues, D., 2013. High throughput colorimetric assay for rapid urease activity quantification. J. Microbiol. Methods 95, 324–326. https://doi.org/10.1016/j.mimet.2013.09.018.

Panda, M.N., Lake, L.W., 1994. Estimation of single-phase permeability from parameters of particle- size distribution. Am. Assoc. Pet. Geol. Bull. https://doi.org/10.1306/a25fe423-171b-11d7-8645000102c1865d.

Panda, M.N., Lake, L.W., 1995. A physical model of cementation and its effects on single-phase permeability. Am. Assoc. Pet. Geol. Bull. 79 https://doi.org/10.1306/8d2b1552-171e-11d7-8645000102c1865d.

Phang, I.R.K., Wong, K.S., Chan, Y.S., Lau, S.Y., 2022. Effect of microbial-induced calcite precipitation towards strength and permeability of peat. Bull. Eng. Geol. Environ. 81 https://doi.org/10.1007/s10064-022-02790-0.

Phillips, A.J., Gerlach, R., Lauchnor, E., Mitchell, A.C., Cunningham, A.B., Spangler, L., 2013. Engineered applications of ureolytic biomineralization: a review. Biofouling. https://doi.org/10.1080/08927014.2013.796550.

Rajasekar, A., Moy, C.K.S., Wilkinson, S., Sekar, R., 2021a. Microbially induced calcite precipitation performance of multiple landfill indigenous bacteria compared to a commercially available bacteria in porous media. PLoS One 16. https://doi.org/10.1371/journal.pone.0254676.

Rajasekar, A., Wilkinson, S., Moy, C.K.S., 2021b. MICP as a potential sustainable technique to treat or entrap contaminants in the natural environment: a review. Environ. Sci. Ecotechnol. https://doi.org/10.1016/j.ese.2021.100096.

Rowshanbakht, K., Khamehchiyan, M., Sajedi, R.H., Nikudel, M.R., 2016. Effect of injected bacterial suspension volume and relative density on carbonate precipitation resulting from microbial treatment. Ecol. Eng. 89, 49–55. https://doi.org/10.1016/j.ecoleng.2016.01.010.

Saad, S., Javadi, A.A., Chugh, T., Farmani, R., 2022. Optimal management of mixed hydraulic barriers in coastal aquifers using multi-objective Bayesian optimization. J. Hydrol. 612 https://doi.org/10.1016/j.jhydrol.2022.128021.

Sabo, M.S., Beckingham, L.E., 2021. Porosity-permeability evolution during simultaneous mineral dissolution and precipitation. Water Resour. Res. 57 https://doi.org/10.1029/2020WR029072.

Safavizadeh, S., Montoya, B.M., Gabr, M.A., 2018. Effect of microbial induced calcium carbonate precipitation on compressibility and hydraulic conductivity of Fly ash. In: IFCEE, pp. 69–79.

Sharma, M., Satyam, N., Reddy, K.R., 2021. Hybrid bacteria mediated cemented sand: microcharacterization, permeability, strength, shear wave velocity, stress-strain, and durability. Int. J. Damage Mech. 30, 618–645. https://doi.org/10.1177/1056789521991196.

Shen, Y., Xin, P., Yu, X., 2020. Combined effect of cutoff wall and tides on groundwater flow and salinity distribution in coastal unconfined aquifers. J. Hydrol. 581 https://doi.org/10.1016/j.jhydrol.2019.124444.

Sidik, W.S., Canakci, H., Kilic, I.H., Celik, F., 2014. Applicability of biocementation for organic soil and its effect on permeability. Geomech. Eng. 7, 649–663. https://doi.org/10.12989/gae.2014.7.6.649.

Singh, R., Yoon, H., Sanford, R.A., Katz, L., Fouke, B.W., Werth, C.J., 2015. Metabolism-induced CaCO3 biomineralization during reactive transport in a micromodel: implications for porosity alteration. Environ. Sci. Technol. 49, 12094–12104. https://doi.org/10.1021/acs.est.5b00152.

Song, C., Song, C., Chen, Y., Wang, J., 2020. Plugging high-permeability zones of oil reservoirs by Microbially mediated calcium carbonate precipitation. ACS Omega 5, 14376–14383. https://doi.org/10.1021/acsomega.0c00902.

Song, C., Elsworth, D., Zhi, S., Wang, C., 2021. The influence of particle morphology on microbially induced CaCO3 clogging in granular media. Mar. Georesour. Geotechnol. 39, 74–81. https://doi.org/10.1080/1064119X.2019.1677828.

Song, C., Elsworth, D., Jia, Y., Lin, J., 2022. Permeable rock matrix sealed with microbially-induced calcium carbonate precipitation: evolutions of mechanical behaviors and associated microstructure. Eng. Geol. 304 https://doi.org/10.1016/j.enggeo.2022.106697.

Soon, N.W., Lee, L.M., Khun, T.C., Ling, H.S., 2014. Factors affecting improvement in engineering properties of residual soil through microbial-induced calcite precipitation. J. Geotech. Geoenviron. Eng. 140 https://doi.org/10.1061/(asce)gt.1943-5606.0001089.

Stabnikov, V., Jian, C., Ivanov, V., Li, Y., 2013. Halotolerant, alkaliphilic urease-producing bacteria from different climate zones and their application for biocementation of sand. World J. Microbiol. Biotechnol. 29, 1453–1460. https://doi.org/10.1007/s11274-013-1309-1.

Sun, Q., Zheng, T., Zheng, X., Walther, M., 2021. Effectiveness and comparison of physical barriers on seawater intrusion and nitrate accumulation in upstream aquifers. J. Contam. Hydrol. 243 https://doi.org/10.1016/j.jconhyd.2021.103913.

Tian, K., Wu, Y., Zhang, H., Li, D., Nie, K., Zhang, S., 2018. Increasing wind erosion resistance of aeolian sandy soil by microbially induced calcium carbonate precipitation. Land Degrad. Dev. 29, 4271–4281. https://doi.org/10.1002/ldr.3176.

Tian, K., Wang, X., Zhang, S., Zhang, H., Zhang, F., Yang, A., 2020. Effect of reactant injection rate on solidifying aeolian sand via microbially induced calcite precipitation. J. Mater. Civ. Eng. 32 https://doi.org/10.1061/(asce)mt.1943-5533.0003391.

Verma, A., Pruess, K., 1988. Thermohydrological conditions and silica redistribution near high-level nuclear wastes emplaced in saturated geological formations. J. Geophys. Res. 93, 1159–1173. https://doi.org/10.1029/JB093iB02p01159.

Wang, X., Nackenhorst, U., 2020a. A coupled bio-chemo-hydraulic model to predict porosity and permeability reduction during microbially induced calcite precipitation. Adv. Water Resour. 139 https://doi.org/10.1016/j.advwatres.2020.103563.

Wang, X., Nackenhorst, U., 2020b. A coupled bio-chemo-hydraulic model to predict porosity and permeability reduction during microbially induced calcite precipitation. Adv. Water Resour. 139 https://doi.org/10.1016/j.advwatres.2020.103563.

Wang, Y., Konstantinou, C., Soga, K., Biscontin, G., Kabla, A.J., 2022. Use of microfluidic experiments to optimize MICP treatment protocols for effective strength enhancement of MICP-treated sandy soils. Acta Geotech. 17, 3817–3838. https://doi.org/10.1007/s11440-022-01478-9.

Wang, Y.N., Li, S.K., Li, Z.Y., Garg, A., 2023. Exploring the application of the MICP technique for the suppression of erosion in granite residual soil in Shantou using a rainfall erosion simulator. Acta Geotech. 18, 3273–3285. https://doi.org/10.1007/s11440-022-01791-3.

Wang, Y., Konstantinou, C., Tang, S., Chen, H., 2023a. Applications of microbial-induced carbonate precipitation: a state-of-the-art review. Biogeotechnics, 100008. https://doi.org/10.1016/j.bgtech.2023.100008.

Wang, Y., Wang, Yong, Konstantinou, C., 2023b. Strength behavior of temperature-dependent MICP-treated soil. J. Geotech. Geoenviron. Eng. 149 https://doi.org/10.1061/jggefk.gteng-11526.

Whiffin, V.S., 2004. Microbial CaCO3 Precipitation for the Production of Biocement. Murdor University Repository. Murdor University.

Whiffin, V.S., van Paassen, L.A., Harkes, M.P., 2007. Microbial carbonate precipitation as a soil improvement technique. Geomicrobiol J. 24, 417–423. https://doi.org/10.1080/01490450701436505.

Wu, H., Wu, W., Liang, W., Dai, F., Liu, H., Xiao, Y., 2023. 3D DEM modeling of biocemented sand with fines as cementing agents. Int. J. Numer. Anal. Methods Geomech. 47, 212–240. https://doi.org/10.1002/nag.3466.

Xie, H., Shi, Y., Yan, H., Bouazza, A., Zhu, X., Wang, A., 2023. Analytical model for organic contaminant transport in a cut-off wall and aquifer dual-domain system considering barrier arrangements. J. Contam. Hydrol. 259 https://doi.org/10.1016/j.jconhyd.2023.104259.

Yang, Y., Chu, J., Xiao, Y., Liu, H., Cheng, L., 2019. Seepage control in sand using bioslurry. Constr. Build. Mater. 212, 342–349. https://doi.org/10.1016/j.conbuildmat.2019.03.313.

Yang, J., Graf, T., Luo, J., Lu, C., 2021. Effect of cut-off wall on freshwater storage in small islands considering ocean surge inundation. J. Hydrol. 603 https://doi.org/10.1016/j.jhydrol.2021.127143.

Yang, B., Li, Hui, Li, Haozhen, Ge, N., Ma, G., Zhang, H., Zhang, X., Zhuang, L., 2022. Experimental investigation on the mechanical and hydraulic properties of urease stabilized fine sand for fully permeable pavement. Int. J. Transp. Sci. Technol. 11, 60–71. https://doi.org/10.1016/j.ijtst.2020.12.002.

Yasuhara, H., Hayashi, K., Okamura, M., 2011. Evolution in Mechanical and Hydraulic Properties of Calcite-Cemented Sand Mediated by Biocatalyst.

Yasuhara, H., Neupane, D., Hayashi, K., Okamura, M., 2012. Experiments and predictions of physical properties of sand cemented by enzymatically-induced carbonate precipitation. Soils Found. 52, 539–549. https://doi.org/10.1016/j.sandf.2012.05.011.

Yu, X., Rong, H., 2022. Seawater based MICP cements two/one-phase cemented sand blocks. Appl. Ocean Res. 118 https://doi.org/10.1016/j.apor.2021.102972.

Yu, X., Yang, H., 2023. One-phase MICP and two-phase MISP composite cementation. Constr. Build. Mater. 409 https://doi.org/10.1016/j.conbuildmat.2023.133724.

Yu, X., He, Z., Li, X., 2022. Bio-cement-modified construction materials and their performances. Environ. Sci. Pollut. Res. 29, 11219–11231. https://doi.org/10.1007/s11356-021-16401-0.

Zamani, A., Asce, S.M., Montoya, B.M., Asce, M., 2017. Shearing and hydraulic behavior of MICP treated silty sand. In: Geotechnical Frontiers 281. ASCE, Orlando, Florida, pp. 290–299. https://doi.org/10.1061/9780784480489.029.

Zamani, A., Montoya, B., Gabr, M.A., 2019. Investigating the challenges of in situ delivery of MICP in fine grain sands and silty sand. Can. Geotech. J. 56, 1889–1900. https://doi.org/10.1139/cgj-2018-0551.

Zhang, J., Lu, C., Wu, H., Li, L., 2022. Analytical solutions for seawater intrusion and the maximum pumping rate of a well in unconfined coastal aquifers bounded by L-shaped coastlines. J. Hydrol. 614 https://doi.org/10.1016/j.jhydrol.2022.128589.

Zhao, Y., Xiao, Z., Fan, C., Shen, W., Wang, Q., Liu, P., 2020. Comparative mechanical behaviors of four fiber-reinforced sand cemented by microbially induced carbonate precipitation. Bull. Eng. Geol. Environ. 79, 3075–3086. https://doi.org/10.1007/s10064-020-01756-4.

Zhao, Y., Zhang, P., Fang, H., Guo, C., Zhang, B., Wang, F., 2021. Bentonite-assisted microbial-induced carbonate precipitation for coarse soil improvement. Bull. Eng. Geol. Environ. 80, 5623–5632. https://doi.org/10.1007/s10064-021-02302-6.