Predicting permeability of compacted clay filtrated with landfill leachate by *k*-Nearest Neighbors modelling method

M. Sukru Ozcoban, M. Erdem Isenkul, S. Güneş-Durak, T. Ormanci-Acar, S. Övez and N. Tüfekci

ABSTRACT

It is important to protect the soil and groundwater from the pollution originating from leachate. Compacted clay soils is a favorable and economic method to protect groundwater and soil against contamination. In this study, compaction tests of leachate was done by using Modified Proctor method. The effects of microbial activity on the permeability of compacted clay soils were analyzed and the obtained data were applied to k-Nearest Neighbors (k-NN) method to predict the permeability of soils in landfill sites. k-NN method, which is a non-parametric distance-based machine learning method and widely used in classification and regression problems was applied to model the relationship between the microorganisms and the permeability. By using k-NN classification method, total heterotrophic bacteria and fungi microorganisms correctly classified the permeability variance as 78.59% and 77.31% success rate, respectively. Also, k-NN modelling was set on regression mode to predict permeability value and produced similar success rates in regression similarity with the actual value. Although, fecal coliforms and fecal streptococci microorganisms had neutral or negative contribution on analyses. For prediction accuracy and regression analysis, the k-NN method was considered for modeling the data. The results of the k-Nearest Neighbors method proved that it is a promising tool for predicting permeability of compacted clay by using microbial activity.

Key words | compacted clay soil, k-Nearest Neighbors, leachate, microbial activity, permeability

M. Sukru Ozcoban
Faculty of Civil Engineering, Department of Geotechnical,
Yildiz Technical University,
34349 Istanbul, Turkey

M. Erdem Isenkul (corresponding author)
Faculty of Engineering, Department of Computer Engineering,
Istanbul University,
34320 Avcilar-Istanbul, Turkey
E-mail: *eisenkul@istanbul.edu.tr*

S. Güneş-Durak
Faculty of Engineering-Architecture, Department of Environmental Engineering,
Nevsehir Haci Bektas Veli University,
50300 Nevsehir, Turkey

T. Ormanci-Acar
N. Tüfekci
Faculty of Engineering, Department of Environmental Engineering,
Istanbul University,
34320 Avcilar-Istanbul,Turkey

S. Övez
Faculty of Civil Engineering, Environmental Engineering Department,
Istanbul Technical University,
34469 Maslak-Istanbul, Turkey

INTRODUCTION

The amount of wastes in the world have reached dangerous levels due to growing trend in the consumption and increasing production. 1.3 billion tons of solid wastes are generated all over the world per year. Since Istanbul is the largest and most crowded city of Turkey, approximately 16,000 tons of waste are collected daily. There are two different sanitary landfill facilities which are in the European and the Asian sides of city. 10,500 tons of wastes are disposed of in the European side of the city, and the rest are disposed of in the Asian side of the city (Kömürcüoda Solid Waste Landfill Facility).

Deposited wastes have a complex blend and consist of organic/inorganic wastes, plastics, glass, metals or pharmaceutical wastes. It is known that leachate has a high concentration of organic and inorganic contaminants (Raghab *et al.* 2013). Permeation of deposits waste generates landfill leachate. Leachate contains high levels of chemical oxygen demand (COD), 5-day biochemical oxygen demand (BOD$_5$), ammonium nitrogen (NH$_4$N), boron, lead, cadmium, nickel, copper, cobalt, iron, manganese, phenols, hardness, sodium, potassium, etc. (Brennan *et al.* 2017). The landfills that closed for less than 5 years or operational landfills generate highly biodegradable leachate and it is named as 'young landfill leachate' (Renou *et al.* 2008). Young landfill leachate exposes approximately 80,000 mg/L COD, 3,100 mg/L NH$_4$N and BOD$_5$: COD $= 0.7$ (Brennan *et al.* 2017). Also, in a study conducted in Turkey, high COD and ammonia concentrations in Komurcuoda Landfill

leachate ranging from 12,350 to 47,800 mg/L, and from 1,500 to 2,680 mg/L were reported, respectively (Inanc *et al.* 2000).

In the case of leachate leakage, the storage of soil waste must be kept under control to prevent the contamination of groundwater and soil. Compacted clay or composite clay and geomembranes are used as barrier to prevent the leakage of leachate (Carey & Swyka 1991). Compacted clay soils are commonly used in solid waste landfills to protect the soil and groundwater from pollution originating in landfills because of their cost effectiveness and large capacity of attenuation (Mohamedzein *et al.* 2005; Aldaeef & Rayhani 2014). The content of leachate is quite important, because it may affect the properties of compacted clay soils. Permeability of clay soils could be reduced due to the quantity of suspended solid matters, acidic waste solutions and microorganisms in the leachate could fill the spaces between the particles of clay (Griffin *et al.* 1976; Francisca & Glatstein 2010; Hamdi & Srasra 2013).

Compacted clay soil with high plasticity can absorb upward of water that has multiples of its own mass. Permeability of compacted clay soil which is put under water pressure, increases over time and it is possible for the compacted clay soil to become instable because of its expansive capacity. It is stated that the main conditions of compacted clay soil are to secure the reduction of pollutant migration for long periods of time, shrinkage and low swelling, and shearing resistance (Brandl 1992; Kayabali 1997; Cazaux & Didier 2000; Hamdi & Srasra 2013).

Statistical machine learning methods (support vector machines (SVM), k-Nearest Neighbor (k-NN) algorithms, decision trees, artificial neural networks (ANN), etc.), although primarily used in the field of computer science, are successfully implemented in many different disciplines. Since the use of these methods is so popular in environmental engineering, the collected environmental data are persistent and verifiable, leading to multidisciplinary research (Khandelwal & Singh 2005; Hanbay *et al.* 2008). There are many statistical analysis studies on fouling membrane in literature. Gao *et al.* have compared the fouling membrane model results developed by using the SVM-based network structure to the ANN models. In conclusion, they have shown that SVM model works better than ANN models because of the ability to generalize and improve classification performance (Gao *et al.* 2007). Aya *et al.* have investigated the effects of Fe^{2+} and Mn^{2+} metals on pollution using the SVM regression model (Aya *et al.* 2016). Bouamar and Ladjal have built a model that classifies the water quality by using SVM classifier (Bouamar & Ladjal 2007).

Many studies have been conducted based on ANN and SVM. Although they are very popular methods in regression and classification problems, it is very difficult to apply these methods due to parameter optimization and the complexity (Table 1). Because of that, we have investigated that the k-NN method on permeability analysis, which is one of the most basic statistical classification and regression analysis method and it is non-parametric. According to the known literature on permeability, no research has yet been conducted on the use of the k-NN methods.

k-NN algorithm is a simple learning algorithm based on similarity calculations which is used frequently in the field of machine learning and statistics. k-NN is a nonparametric technique used in statistical determinations and pattern recognition problems (such as classification, clustering and regression) which are quite long (Altman 1992).

Table 1 | Comparison of different statistical methods

	Advantages	Limits
k-Nearest Neighbor (k-NN) Algorithm	- supports non-linear separation - suitable for multimodal classes - non parametric - no cost on learning process	- time complexity of calculate nearest neighbor for large datasets - sensitivity on noisy data - algorithm performance depends on the number of features
Support Vector Machine (SVM) Algorithm	- high accuracy - kernel support - memory efficiency - high dimensional space transformation	- number of parameters - parameter optimization - black-box models - kernel selection
Artificial Neural Network (ANN) Algorithm	- adapt to unsupervised learning - powerful - easy of use, learn by example	- overlearning - large complexity of network - long training time

This algorithm produces output according to the class of the nearest neighbors up to the given k value. When used for regression analysis, the given test case is equal to the average of the nearest k neighbor values. In classifying problems, the given test case is classified according to the sample if it is from the class label as the majority by looking at the class labels of the k closest neighbors (Figure 1).

The distance between classes in our work is calculated using Euclidean distance (Equation (1)).

$$D_{euc} = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2} \qquad (1)$$

where D_{euc} is the distance of between x and y samples, i is the current feature and the n is the total number of features. With 1-nearest neighbor rule, the predicted class of test sample x is set equal to the true class of its nearest neighbor. For k-NN the predicted class of test sample x is set equal to the most frequent true class among k nearest training samples. In practice, k is usually chosen to be odd ($k = 1$, 3, 5, 7 …) so as to avoid ties.

In statistical applications and analyzes, values are often spread at different intervals in the data space. This causes the attributes or values that are compared and participating in the calculation to affect the analysis negatively. In order to increase the classification and regression performance of the k-NN method that was applied, the data must first be subjected to the normalization process. There are many different methods of normalization, which vary according to the complexity of the application and the application areas. Standard score normalization method is generally preferred for regression and classification analysis. Accordingly, all values in the dataset are subtracted from the average value of the data set, then the difference is divided by the standard deviation (Equation (2)).

$$Z = \frac{X - \mu}{\sigma} \qquad (2)$$

where Z is the standard score of raw X, μ is the mean of the population, and σ is the standard deviation of the population. All values normalized so that the distribution can be converted to -1, $+1$ are calculated as shown in Equation (3) so that all microorganism values and permeability values were normalized to the desired range.

$$Z' = (-1) + \frac{(Z - Z_{min})(+1 - (-1))}{Z_{max} - Z_{min}} \qquad (3)$$

In this study, k-NN model was applied to the permeability data of leachate obtained from the filtration tests of different microorganisms in order to estimate the most prevalent microorganism. For this purpose, five different microorganisms were chosen that present in leachate as *Total HBac*, *Total Coli*, *Fecal Coli*, *Fecal Strep* and *Fungi*.
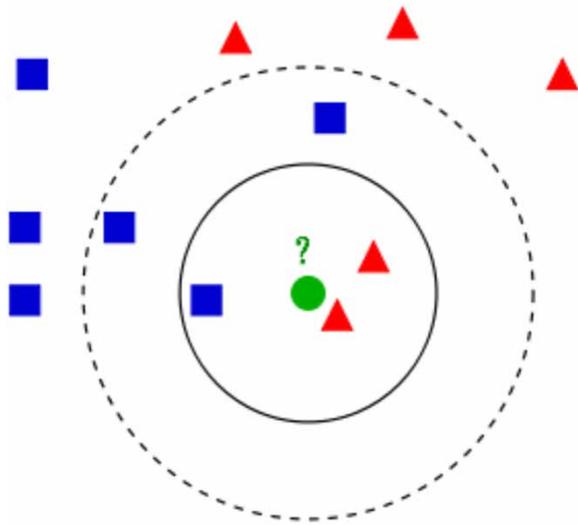
MATERIALS AND METHODS

Materials

Properties of the clay soil and the leachate

The Şile-Kömürcüoda Landfill Site soil samples contains 68 to 71% kaolinite, 6 to 9% free quartz, 15 to 18% illite, and 2 to 5% others. The kaolinite and illite have been considered to be true clay soil minerals. The soil samples had a coefficient of permeability $k = 1 \times 10^{-8}$ m/s a discharge loss of 8.5 to 9%, and a water absorption of 0.2 to 0.4% (Ozcoban 2008a).

Used leachate is a young leachate and has a dark brown color and very small granules and also contains large amounts of organic, inorganic contaminants and a high concentration of metals. Essential parameters of leachate were given in Table 2 (Ozcoban 2008b; Tüfekci *et al.* 2010). All

Figure 1 | Example of k-NN classification. The test sample (circle) should be classified as triangles if k = 3. If k = 5 (dashed line circle), the class will be square (Coomans & Massart 1982).

Table 2 | Properties of the leachate

Parameter/Averages	pH	COD (mg/L)	BOD (mg/L)	TKN (mg/L)	Total P (mg/L)	SS (mg/L)	VSS (mg/L)
Ozcoban (2008b)	6.5	9,500	3,500	3,000	5	1,159	
Tüfekci *et al.* (2010)	6.8	10,000	1,010	1,635	5	1,010	855
This study	6.7	13,526	6,235	2,876	7	1,283	987

analyses were conducted according to Standard Methods (Gilcreas 1966). Microbial properties of leachate was determined by measuring three different samples and the results are presented in Table 3. The concentration of microorganisms were found to be around $2.52 \times 10^6 \pm 0.29$, $7.63 \times 10^5 \pm 3.15$, $3.67 \times 10^5 \pm 0.99$, $2.36 \times 10^5 \pm 1.83$ and $3.10 \times 10^5 \pm 1.79$ for total bacteria, total coliform, fecal coliform, fecal streptococci and fungi, respectively. Microbial activity (total heterotrophic bacteria, total coliform, fecal coliform, fecal streptococcus, and fungi) have been measured according to standard APHA methods both in the influent and effluent of the continuous reactor (APHA 2005). These analyses have been conducted on the samples taken from the influent of the solid waste leachate and effluent of the reactors treating the leachate, using membrane filtration technique. Membrane filtration method has been applied to detect microorganisms under aseptic techniques. Results were reported as a colony forming unit (CFU) per 100 mL of sample.

Experimental methods

Permeability tests. As seen from Figure 2, a continuous filtration system was operated. The reactor tests have been performed by flowing the liquid downwards through 100 mm diameter compacted specimens. The height of the compacted clay soil was 110 mm. The soil was constrained against swelling. The clay soil has been saturated under a 0.3 bar pressure (Zimmie 1981; Daniel *et al.* 1985). The total discharge (Q) of filtration test was continuously

recorded. Constant–head tests have been performed to find the permeability of the clay soil which is calculated using the following equation:

$$k = \frac{QL}{At(h_1 - h_2)} \tag{4}$$

where

k: Coefficient of permeability, cm/s;
A: Surface area of the specimen, cm^2;
L: Distance between the manometers, cm;
$(h_1$-$h_2)$: Differential head across the sample, cm;
Q: Total discharge, cm^3/s; t: elapsed time, s.

   Modified Proctor (ASTM D1557/AASHTO T180) (Standard 2006) method is commonly applied in the laboratory at different water contents in a mound (0.102 m. ID X 0.117 m. H) and vary only in the amount of applied energy to determine the maximum dry density–water content relationship (ASTM 2005).

Modeling

In this study, the effects of microorganisms on the permeability of clay samples taken at different time intervals were estimated using k-NN regression method. Two different problems have been solved in the analysis of the permeability effects and the relations between the microorganisms and the permeability have been tried to be determined according to the performance of the k-NN

Table 3 | Microbial properties of the leachate

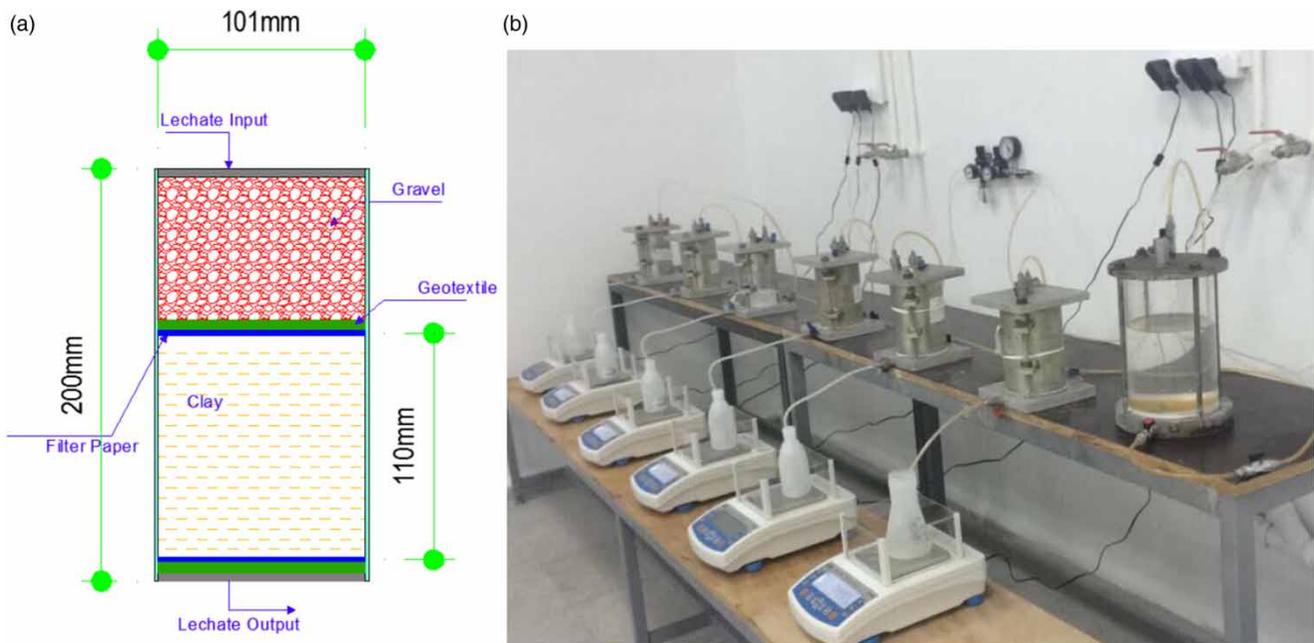| Microorganism types/date	Total bacteria (cfu/100 mL)	Total coliform (cfu/100 mL)	Fecal coliform (cfu/100 mL)	Fecal streptococci (cfu/100 mL)	Fungi (cfu/100 mL)
Sample 1	2.85×10^6	4.75×10^5	3.26×10^5	3.68×10^5	1.45×10^5
Sample 2	2.31×10^6	7.13×10^5	2.96×10^5	3.12×10^5	2.85×10^5
Sample 3	2.4×10^6	1.1×10^6	4.8×10^5	2.7×10^4	5.0×10^5
Average values	$2.52 \times 10^6 \pm 0.29$	$7.63 \times 10^5 \pm 3.15$	$3.67 \times 10^5 \pm 0.99$	$2.36 \times 10^5 \pm 1.83$	$3.10 \times 10^5 \pm 1.79$

Figure 2 | Schematic (a) and real image (b) of experimental set up.

regression method. Developed problems are presented in two different sub-chapters:

- The effects of microorganisms on k-NN regression similarity
- The effects of microorganisms on the classification performance of k-NN

k-NN regression similarity approach

In the first developed problem, regression similarity, which is the most basic approach, has been examined. Randomly selected 246 soil samples which contained microbial activities at different time intervals were divided into train and test set (75% train set, 25% test set) to generate regression permeability values based on microorganisms. Randomization is required to select the different time intervals. Also, the separation process of the data into train and test set is to test the trained regression model without overfitting. What is required here is the similarity between the regression and the actual permeability values. In the regression similarity approach, the root mean square error (RMSE) of the two values is subtracted from 1 and the error rate is converted to the similarity ratio (Equation (5)).

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

where n is the number of test size, y_i is the observed value for the ith observation and $\hat{y}_i$ is the predicted value.

In addition, R^2 similarity analysis, which is frequently used in regression analysis, was also applied in our study. The most general definition formula for the identification coefficient (R^2) is:

$$R^2 \equiv 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \qquad (6)$$

where y_i is the observed value for the ith observation, $\hat{y}_i$ is the predicted value and $\bar{y}$ is the mean value of the response variable.

It has been observed that microorganisms selected individual and together (double, triple, quadruple or quintile groups) as an attribute have different effects on the similarity ratio (in the range of 0 to 1 or non-similar to similar) between the value produced by regression and the actual permeability value. This procedure was repeated 100 times and the average of similarity ratios was taken as the basis for the analysis.

k-NN classification approach

Our second approach to study is to create class labels according to the difference between the selected permeability value at time t and the selected permeability value at time $t + 1$

(0: no change, +1: increase, −1: decrease). Randomly selected 246 soil samples at different time intervals were assigned to the k-NN classifier, separated by training and test set (75% training set, 25% test set). The classification success of the k-NNs of selected microorganisms, individual or together as an attribute (binary, triple, quadruple or quintile groups), demonstrates the effects on the increase or decrease of permeability values. As in the previous problem, this process was repeated 100 times and the classification success average was taken as the basis.

RESULTS

Regression and correlation analysis

The question underlying the regression analysis is how microorganisms or groups of microorganisms affect the permeability value. Thus, the similarity between actual and predicted permeability value should be calculated by using microorganism values. The similarities between measured permeability and predicted permeability would be calculated using only microorganism values would give some information about the effects of microorganisms on permeability. It is known that the presence of some microorganism can occlude the pores and cause the decrease of leachate's permeability (Glatstein & Francisca 2014). For this reason, each microorganism was selected as an attribute as individual, double, triple, quadruple and quintile groups and used for the k-NN regression. Then, regression similarities were examined between the actual and predicted permeability values. This process was repeated 100 times with randomly selected samples every time. Similarity ($1 - RMSE$), R^2 regression similarity and correlation (ρ) analysis results are shown in the tables given below (Tables 4–8).

When the regression values of selected microorganisms as an individual feature are examined, it is seen that '*Total HBac*' and '*Fungi*' microorganisms produce highly permeability regression similarity rates (>75%). Also, correlation analysis shows that the two microorganisms are highly correlated with the permeability (>78%). However, although the '*Fecal Coli*' and '*Fecal Strep*' and '*Total Coli*' microorganisms are correlated with permeability at the average of ~60%, k-NN regression produces permeability value that are inconsistent with the actual permeability value according to the R^2 regression analysis.

In addition, the same process was examined among the other groups. The similarity ratios of the regression values calculated by using double attributes in Table 5, triple in

Table 4 │ Regression similarity and correlation with permeability (ρ) measurements of individually selected features

	'Total HBac'	'Total Coli'	'Fecal Coli'	'Fecal Strep'	'Fungi'
$1 - RMSE$	0.7731	0.6808	0.6402	0.6561	0.7616
R^2	0.5933	0.1938	0.0148	0.0143	0.5506
ρ	0.8091	0.6313	0.5853	0.5971	0.7827

Table 5 │ Regression similarity and correlation with permeability (ρ) measurements of two selected features

	'Total HBac' 'Total Coli'	'Total HBac' 'Fecal Coli'	'Total HBac' 'Fecal Strep'	'Total HBac' 'Fungi'
$1 - RMSE$	0.7823	0.7757	0.7796	0.7804
R^2	0.6359	0.6153	0.6263	0.6443
ρ	0.8262	0.8167	0.8241	0.8293
	'Total Coli' 'Fecal Coli'	**'Total Coli' 'Fecal Strep'**	**'Total Coli' 'Fungi'**	
$1 - RMSE$	0.7441	0.7220	0.7739	
R^2	0.4953	0.3916	0.6021	
ρ	0.7690	0.7303	0.8083	
	'Fecal Coli' 'Fecal Strep'	**'Fecal Coli' 'Fungi'**		
$1 - RMSE$	0.6561	0.7726		
R^2	0.0858	0.5941		
ρ	0.5996	0.8052		
	'Fecal Strep' 'Fungi'			
$1 - RMSE$	0.7778			
R^2	0.6047			
ρ	0.8129			

Table 6, quartiles in Table 7, and all features used together in Table 8, given below.

If the attributes were considered as a group instead of being singly selected, the groups of microorganisms producing high affinity in Tables 2 and 3 are performed well, both in correlation values and in regression similarity. Maximum correlation and similarity performances were produced in the triple group with '*Total HBac*', '*Fecal Strep*' and '*Fungi*' (in Table 5: Similarity: 0.7863, $R^2 = 0.6485$, $\rho = 0.8304$). The permeability regression values produced by this group and the actual permeability values graph is shown in Figure 3 below for 20 randomly selected samples (Figure 3).

Table 6 | Regression similarity and correlation with permeability (ρ) measurements of three selected features

	'Total HBac' 'Total Coli' 'Fecal Coli'	'Total HBac' 'Total Coli' 'Fecal Strep'	'Total HBac' 'Total Coli' 'Fungi'	'Total HBac' 'Fecal Coli' 'Fecal Strep'	'Total HBac' 'Fecal Coli' 'Fungi'	'Total HBac' 'Fecal Strep' 'Fungi'
$1 - RMSE$	0.7839	0.7795	0.7805	0.7786	0.7764	0.7863
R^2	0.6448	0.6147	0.6255	0.6198	0.6222	0.6485
ρ	0.8270	0.8185	0.8212	0.8187	0.8161	0.8304
	'Total Coli' 'Fecal Coli' 'Fecal Strep'	**'Total Coli' 'Fecal Coli' 'Fungi'**	**'Total Coli' 'Fecal Strep' 'Fungi'**			
$1 - RMSE$	0.7486	0.7850	0.7801			
R^2	0.5036	0.6369	0.6308			
ρ	0.7706	0.8213	0.8265			
	'Fecal Coli' 'Fecal Strep' 'Fungi'					
$1 - RMSE$	0.7749					
R^2	0.6096					
ρ	0.8119					

Table 7 | Regression similarity and correlation with permeability (ρ) measurements of four selected features

	'Total HBac' 'Total Coli' 'Fecal Coli' 'Fecal Strep'	'Total HBac' 'Total Coli' 'Fecal Coli' 'Fungi'	'Total HBac' 'Total Coli' 'Fecal Strep' 'Fungi'	'Total HBac' 'Fecal Coli' 'Fecal Strep' 'Fungi'
$1 - RMSE$	0.7855	0.7764	0.7745	0.7783
R^2	0.6346	0.6113	0.6065	0.6119
ρ	0.8301	0.8178	0.8128	0.8232
	'Total Coli' 'Fecal Coli' 'Fecal Strep' 'Fungi'			
$1 - RMSE$	0.7740			
R^2	0.6046			
ρ	0.8124			

Table 8 | Regression similarity and correlation with permeability (ρ) measurements of all features together

	'Total HBac' 'Total Coli' 'Fecal Coli' 'Fecal Strep' 'Fungi'
$1 - RMSE$	0.7823
R^2	0.6284
ρ	0.8235

Classification analysis

Similar operations, similar to regression analysis, have also been used in k-NN classification. The parameter k value for k-NN is chosen according to best performance, respectively. The highest average performance of k-NN with $k = 1$ is 76.69%, $k = 3$ is 78.33%, $k = 5$ is 76.44% and $k = 7$ is 73.61% after 100 iterations. Thus, for the $k = 3$ neighborhood, the attributes were selected in groups of single, double, triple, quadruple, and quintile to analyze which attributes or attribute groups gave the highest classification rates. Figure 4 shows that the classifier performance of the k-NN of individually selected features.

In the problem of classification of increase–decrease variation of permeability, 'Total Hbac' and 'Fungi' microorganism attributes correctly classify this problem as 78.59% and 77.31% success rate, respectively. However, it is likely that the effect of permeability change will be different, as microorganisms are selected together. For this reason, classification performance of microorganisms in double, triple, quadruple and quintile groups is also important. The following table shows that the average classification performances of microorganisms included groups (Table 9).

When all groups of microorganisms are examined, 'Total HBac' and 'Fungi' microorganisms, the general group gives positive contribution for triple, quadruple and quintile groups. However, 'Fecal Coli' and 'Fecal Strep'
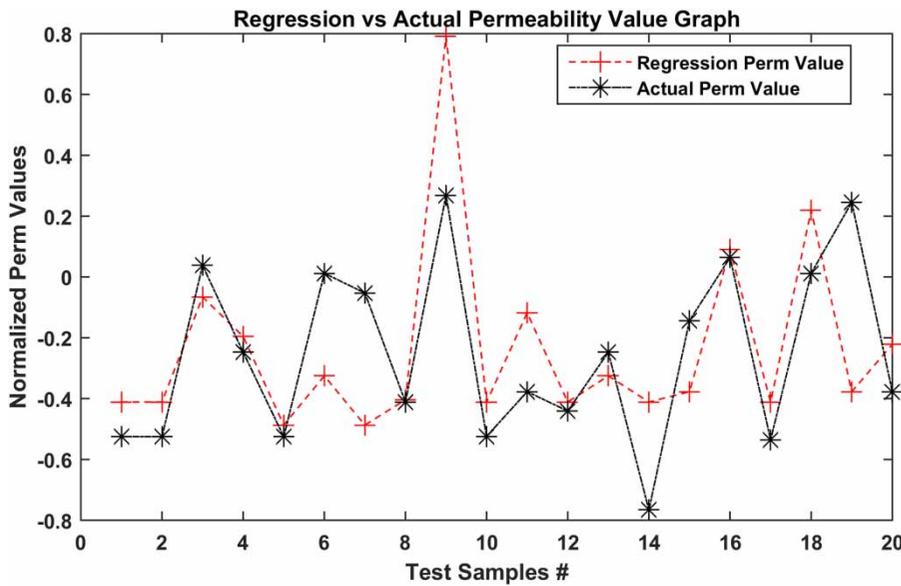
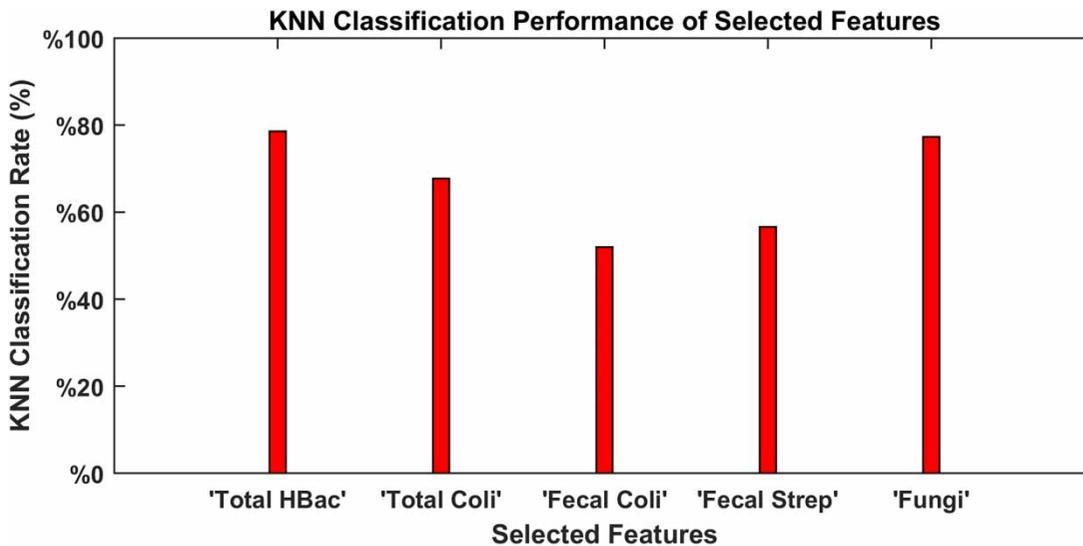**Figure 3** │ Graph of actual and regression permeability values.

Figure 4 │ k-NN classifier performance of ($k = 3$) of individually selected attributes.

Table 9 │ k-NN classifier average performances of microorganisms included feature groups ($k = 3$)

	Individual groups	Double groups	Triple groups	Quadruple groups	Quintile group
'Total HBac'	**0.7859**	**0.7928**	**0.7962**	**0.7958**	0.80
'Total Coli'	0.6774	0.7691	0.7868	0.7947	
'Fecal Coli'	0.5198	0.7260	0.7840	0.7943	
'Fecal Strep'	0.5659	0.7255	0.7846	0.7945	
'Fungi'	**0.7731**	**0.7873**	**0.7934**	**0.7952**	

microorganisms are neutral or no effect on the permeability activities. This results for 'Total HBac' and 'Fungi' microorganisms can be attributed to the fact that they can multiply under favorable conditions easily. Especially, 'Total HBac' can easily multiply and clog pores in the system as it contains many bacterial species. Likewise, 'Fungi' can easily enter and fill pores with proliferation spores (ascospore, basidiospore, etc.) owing to their wide variety species. Also, 'Fungi' respond to slightly acidic pH values (between 5.5–6.5 that were used in this study) and can rapidly multiply

in this conditionals. The spores that have 1–2 μm pores can fill and clog pores, developing more hyphae to the bush to form nets that cause clogging (Thullner *et al.* 2002).

CONCLUSIONS

In the case of leachate leakage, the storage of soil waste must be kept under control to prevent the contamination of groundwater and soil. Compacted clay soils are commonly used in solid waste landfills to protect the soil and groundwater from pollution originating in landfills because of their cost effectiveness and large capacity of attenuation. As the leachate enhances the permeability of compacted clay soils, it is desired to decrease the permeability to protect the groundwater resources.

In this study, the effects of microbial activity in leachate, on permeability of clay soils that were compacted with Modified Proctor compaction tests method, are investigated. The results taken out from this experimental process are monitored and tested using (k-NN) method. In the short term, permeability of compacted clay soil decreased because of the suspended solids and clogging caused by microorganisms in the leachate fills the spaces in the clay soil. On the other hand, in the long term, the permeability increase due to the distortions of structure of the clay soil. After investigating all groups of microorganisms, permeability changes are most accurately predictable in the groups 'Total HBac' and 'Fungi', but 'Fecal Coli' and 'Fecal Strep' are not contributed to the classification of the same parameter. Obtained model results validates the microorganism roles in the leachate permeability. Especially, it can be said that 'Total HBac' and 'Fungi' are connected to the permeability changes. For future studies, this finding can be a verified reference to understand the correlation between the microorganism in the leachate and their effect of permeability.

REFERENCES

Aldaeef, A. & Rayhani, M. T. 2014 Hydraulic performance of Compacted Clay Liners (CCLs) under combined temperature and leachate exposures. *Waste Management* **34** (12), 2548–2560.

Altman, N. S. 1992 An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician* **46** (3), 175–185.

APHA 2005 *Standard Methods for the Examination of Water and Wastewater*. American Public Health Association (APHA), Washington, DC, USA.

ASTM 2005 *Standard Test Methods for Laboratory Determination of Water (Moisture) Content of Soil and Rock by Mass*.

American Society for Testing and Materials, West Conshohocken, PA, USA, 2216.

Aya, S. A., Ormanci Acar, T. & Tufekci, N. 2016 Modeling of membrane fouling in a submerged membrane reactor using support vector regression. *Desalination and Water Treatment* **57** (51), 24132–24145.

Bouamar, M. & Ladjal, M. 2007 Multisensor system using support vector machines for water quality classification. In: *Signal Processing and Its Applications, 2007. ISSPA 2007. 9th International Symposium*, IEEE, pp. 1–4.

Brandl, H. 1992 Mineral liners for hazardous waste containment. *Geotechnique* **42** (1), 57–65.

Brennan, R., Clifford, E., Devroedt, C., Morrison, L. & Healy, M. 2017 Treatment of landfill leachate in municipal wastewater treatment plants and impacts on effluent ammonium concentrations. *Journal of Environmental Management* **188**, 64–72.

Carey, P. J. & Swyka, M. A. 1991 Design and placement considerations for clay and composite clay/geomembrane landfill final covers. In: *Landfill Closures* (R. M. Koerner, ed.), Elsevier, London, UK, pp. 133–140.

Cazaux, D. & Didier, G. 2000 Field evaluation of hydraulic performances of geosynthetic clay liners by small and large-scale tests. *Geotextiles and Geomembranes* **18** (2), 163–178.

Coomans, D. & Massart, D. L. 1982 Alternative k-nearest neighbour rules in supervised pattern recognition: part 1. k-Nearest neighbour classification by using alternative voting rules. *Analytica Chimica Acta* **136**, 15–27.

Daniel, D. E., Anderson, D. C. & Boynton, S. S. 1985 Fixed-wall versus flexible-wall permeameters. *Hydraulic Barriers in Soil and Rock* **874**, 107–126.

Francisca, F. M. & Glatstein, D. A. 2010 Long term hydraulic conductivity of compacted soils permeated with landfill leachate. *Applied Clay Science* **49** (3), 187–193.

Gao, M., Tian, J. & Li, J. 2007 The study of membrane fouling modeling method based on support vector machine for sewage treatment membrane bioreactor. In: *Industrial Electronics and Applications, 2007. ICIEA 2007. 2nd IEEE Conference*, IEEE, pp. 1393–1398.

Gilcreas, F. 1966 Standard methods for the examination of water and wastewater. *American Journal of Public Health and the Nations Health* **56** (3), 387–388.

Glatstein, D. A. & Francisca, F. M. 2014 Hydraulic conductivity of compacted soils controlled by microbial activity. *Environmental Technology* **35** (15), 1886–1892.

Griffin, R. A., Shimp, N. F., Steele, J. D., Ruch, R. R., White, W. A. & Hughes, G. M. 1976 Attenuation of pollutants in municipal landfill leachate by passage through clay. *Environmental Science & Technology* **10** (13), 1262–1268.

Hamdi, N. & Srasra, E. 2013 Hydraulic conductivity study of compacted clay soils used as landfill liners for an acidic waste. *Waste Management* **33** (1), 60–66.

Hanbay, D., Turkoglu, I. & Demir, Y. 2008 Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks. *Expert Systems with Applications* **34** (2), 1038–1043.

Inanc, B., Calli, B. & Saatci, A. 2000 Characterization and anaerobic treatment of the sanitary landfill leachate in Istanbul. *Water Science and Technology* **41** (3), 223–230.

Kayabali, K. 1997 Engineering aspects of a novel landfill liner material: bentonite-amended natural zeolite. *Engineering Geology* **46** (2), 105–114.

Khandelwal, M. & Singh, T. 2005 Prediction of mine water quality by physical parameters. *Journal of Scientific & Industrial Research* **64**, 564–570.

Mohamedzein, Y. E.-A., Al-Rawas, A. A., Al-Aghbari, M. Y., Qatan, A. & Al-Rawas, A.-H. 2005 Assessment of crushed shales for use as compacted landfill liners. *Engineering Geology* **80** (3), 271–281.

Ozcoban, M. 2008a Permeability and removal rate of compacted clays to organic and inorganic matters. *Journal of Scientific & Industrial Research* **67** (2), 136.

Ozcoban, M. 2008b Permeability and removal rate of compacted clays to organic and inorganic matters. *Journal of Scientific & Industrial Research* **67**, 136–140.

Raghab, S. M., El Meguid, A. M. A. & Hegazi, H. A. 2013 Treatment of leachate from municipal solid waste landfill. *HBRC Journal* **9** (2), 187–192.

Renou, S., Givaudan, J. G., Poulain, S., Dirassouyan, F. & Moulin, P. 2008 Landfill leachate treatment: review and opportunity. *Journal of Hazardous Materials* **150** (3), 468–493.

Standard A. D2434-68 2006 *Standard Test Method for Permeability of Granular Soils (Constant Head)*. ASTM International, West Conshohocken, PA, USA.

Thullner, M., Zeyer, J. & Kinzelbach, W. 2002 Influence of microbial growth on hydraulic properties of pore networks. *Transport in Porous Media* **49** (1), 99–122.

Tüfekci, N., Özçoban, M., Yalçİn, S., Aşcİ, Y. & Akgüner, C. 2010 Adsorption and permeability of clays permeated with ferrous iron and manganese. *Fresenius Environmental Bulletin* **19** (8b), 1703–1714.

Zimmie, T. 1981 Geotechnical testing considerations in the determination of laboratory permeability for hazardous waste disposal siting. *ASTM Special Technical Publication No.* **760**, 293–304.