KSCE Journal of Civil Engineering 30 (2026) 100438

Contents lists available at ScienceDirect

KSCE Journal of Civil Engineering

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

Development of an ANN-based model for predicting unfrozen water content in sandy soils: Insights from experimental data and GUI implementation

Hee Won Kim, Dinh-Viet Le, Gyu-Hyun Go *

School of Architecture, Civil and Environmental Engineering, Kumoh National Institute of Technology, Gumi, Gyeongbuk 39177, South Korea

ARTICLE INFO

Keywords:
Frozen soil
Freezing chamber system
Unfrozen water content
Artificial neural network
GUI software

ABSTRACT

Predicting the unfrozen water content curve is essential to understanding the thermal–hydraulic–mechanical behavior of frozen soil. However, existing methods often require extensive experiments. When experimental data is lacking, these methods rely on other empirical models, resulting in overly complex workflows. To overcome these issues, we developed two artificial neural network (ANN) models to predict the unfrozen water content of saline sandy soils for two distinct soil types. The trained model files were then implemented in a Graphic User Interface (GUI) program for end-users. The training datasets were collected through a series of laboratory experiments that varied void ratio, pore water salinity, and soil type using a newly designed freezing chamber system. A total of 186 datasets were collected and divided into training and testing sets. The training results have a coefficient of determination (R^2) greater than 99 % for both ANN models. Meanwhile, the R^2 of test results were over 98 % demonstrating the accuracy of predicting the unfrozen water content, closely matching the measured values. In addition, the trained ANN models are embedded in a GUI program for wide application in coastal and permafrost environments.

1. Introduction

When saturated soil is frozen, a three-phase structure of ice, water, and soil particles is formed. Although the soil reaches a sub-zero temperature, there remains water that does not freeze, which is called unfrozen water. In numerical analyses simulating the freezing process of soil, the unfrozen water content is evaluated as one of the main input factors, and it affects the strength, deformation, and thermal–hydraulic behavior of frozen soil (Burt and Williams, 1976; Ohrai, 1986; Kweon, 2003; Shin and Park, 2003; Zhou and Meschke, 2013; Liu, 2014; Kim et al., 2016; Chen et al., 2021b; Woo and Go, 2024). Understanding the unfrozen water content under various conditions is essential for the reliable prediction and assessment of frozen soil behavior.

Standard methods for measuring unfrozen water in frozen soil include Nuclear Magnetic Resonance (NMR), Differential Scanning Calorimetry (DSC), frequency-domain reflectometry (FDR), and time-domain reflectometry (TDR). NMR measures unfrozen water content by detecting atomic nuclei absorbing and re-emitting electromagnetic waves in a magnetic field. It is considered the most accurate non-destructive method for measuring unfrozen water content (Tice et al., 1976; Watanabe and Mizoguchi, 2002; Chen et al., 2021a; Chen et al., 2021b; Kruse et al., 2018, Woo and Go, 2024) used NMR data to propose a prediction model based on the experimental data. The proposed model considered adsorption and capillary phenomena during the freezing process and showed relatively good agreement with independent experimental data. Kruse et al. (2018) developed a normalization method to improve the P-NMR approach for quantifying unfrozen water content. The traditional method has the potential to overestimate the unfrozen water content because it incorporates the reactions of both ice and water. The experimental results demonstrated that the conventional method yielded an unfrozen water content estimate 3.3 times higher than that obtained using the normalization method. NMR methods have been widely used to study unfrozen water content (Tice et al., 1976; Watanabe and Mizoguchi, 2002; Chen et al., 2021a; Chen et al., 2021b; Kruse et al., 2018, Woo and Go, 2024), but significant financial investment is required to build the necessary equipment.Kozlowski and Nartowska (2013) investigated the effect of freeze-thaw cycles on unfrozen water content profiles using differential scanning calorimetry (DSC). They found that the profiles changed significantly depending on the number of cycles. FDR analyzes the frequency shift of the reflected wave by exploiting the difference in the material's relative permittivity. This method allows for continuous measurements and is relatively

* Corresponding author.
E-mail address: gyuhyungo@kumoh.ac.kr (G.-H. Go).

https://doi.org/10.1016/j.kscej.2025.100438

Available online 5 November 2025

Table 1
Summary of previous prediction models for unfrozen water content analysis.

Model	Prediction formula	Note
Chai et al. (2018)	$\theta_u(T_i) = \theta_{cw}(T_i) + \theta_{bw}(T_i)$	Max: θ_u Min: θ_r
Li et al. (2020)	$\theta_u = \theta_{uu} + \theta_{uf}$	
Kurylyk and Watanabe (2013)	$\theta_u = \theta_r + (\theta_0 - \theta_r)\exp\left[-(T/\gamma)^2\right]$	
Ren et al. (2017)	$\theta_u = \theta_r + (\theta_s - \theta_r)\left[1 + \left(-a_v L_f \rho_w \ln\frac{T+273.15}{T_0+273.15}\right)^{n_v}\right]^{-m_v}$ where, a_v, n_v, m_v are curve fitting parameters.	
Anderson and Tice (1972)	$\omega_u = \alpha(-T)^{\beta}$ where, α, β are curve fitting parameters.	Max: ω_u Min: ω_r
Michalowski and Zhu (2006)	$\omega_u = \begin{cases} \omega_0 & T \geq T_f \\ \omega_r + (\omega_0 - \omega_r)\exp[\mu(T - T_f)] & T < T_f \end{cases}$ where, $\mu$is curve fitting parameters.	
Kozlowski (2007)	$\omega_u = \begin{cases} \omega_0 & T > T_f \\ \omega_r + (\omega_0 - \omega_r)\exp\left[-3.35\left(\frac{T_f - T}{T - T_m}\right)^{0.37}\right] & T_m < T \leq T_f \\ \omega_r & T \leq T_m \end{cases}$ where, $T_f = -0.0729 w_p^{2.462} w^{-2}, \omega_r = 0.042SSA + 3, T_m = -12^{\circ}C$	
Kong et al. (2020)	$\omega_u = \begin{cases} \omega_0 & T > T_f \\ \omega_0 (T/T_f)^g & T \leq T_f \end{cases}$ where, $T_f = -0.015 I_p - 0.063$, $g = 0.300\ln I_p - 2.232$	
McKenzie et al. (2007)	$S_w = S_r + (1 - S_r)\exp\left[-((T - T_f)/\gamma)^2\right]$	Max: 1 Min: S_u
Tice et al. (1976)	$S_w = [1 - (T - T_f)]^{a}$	
Zhou and Meschke (2013)	$S_w = \left(1 + ((T_f - T)/\Delta T_{ch})^{1/(1-m)}\right)^{-m}$	
Teng et al. (2021)	$S_w = D_r \cdot S_{w,TH} + (1 - D_r) \cdot S_{w,SC}$	

*$\theta$is volumetric water content, **$\omega$is gravimetric water content, ***S is degree of saturation

inexpensive. However, it has been shown to have limitations in terms of accuracy because it cannot consistently estimate values depending on pore water salinity and soil type (Yoshikawa et al., 2004). Similarly, TDR estimates the content of unfrozen water using the concept of relative permittivity and the time-series analysis of the reflected wave. The reported measurement error is within 0.5 % compared to the nuclear magnetic resonance (NMR) method (Smith and Tice, 1988), meaning that the NMR method has excellent measurement performance relative to its cost. Shin and Park (2003) measured the changes in freezing pressure and unfrozen water content under various time and temperature conditions using TDR equipment. They established a relationship between temperature and water content and derived the characteristics of unfrozen water content that dominate the expression of freezing pressure. Xin et al. (2023) measured the soil freezing characteristic curve (SFCC) of silty clay using TDR. Then, based on the experimental

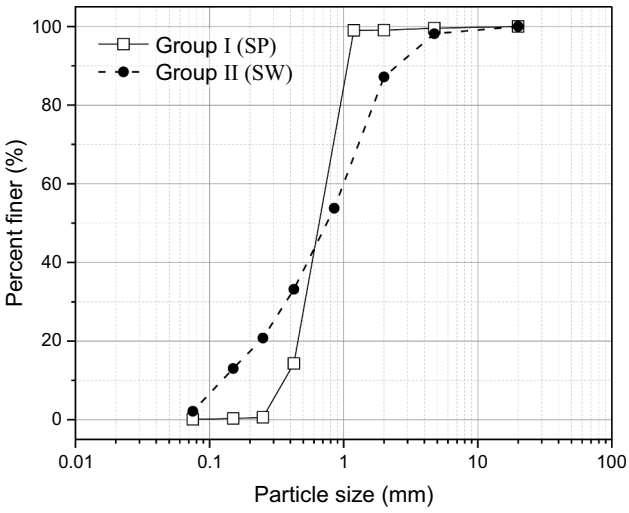

Fig. 1. Particle size distribution for SP and SW sand.

results, they constructed a freezing point model that considers the weight and dry density of the water content.

Numerous empirical models have been proposed to estimate the unfrozen water content properties of frozen soils based on the findings of previous studies (Anderson and Tice, 1972; Tice et al., 1976; Michalowski and Zhu, 2006; Kozlowski, 2007; Zhou and Meschke, 2013; Ren et al., 2017; Chai et al., 2018; Li et al., 2020; Teng et al., 2021; Kong et al., 2020). The most widely used models are summarized in Table 1. These models were developed by analyzing the relationship between unfrozen water content and subzero temperatures observed in laboratory experiments. Most empirical equations include fitting curve coefficients derived from the experimental data. However, using empirical formulas without experimental data is challenging, prompting some studies to propose formulas based on easily obtainable physical properties such as the specific surface area and plasticity index of the soil as alternatives to fitting curve coefficients (Anderson and Tice, 1972; Kozlowski, 2007; Kong et al., 2020). However, these models are limited to silt and clay soil. For sandy soil, challenges remain in determining the fitting curve coefficients through experimental data or arbitrary assumptions of these coefficients.

In light of these considerations, this study proposes a novel artificial neural network (ANN) model capable of predicting the unfrozen water content of sandy soils during freezing based solely on their fundamental physical properties. Some data mining models, such as ANN models, are widely used in civil engineering as a prediction tool (Omar et al., 2018; Chen et al., 2019; Yoon et al., 2021; Samadi et al., 2023; Yoon et al., 2023; Al-Swaidani et al., 2024; Tariq et al., 2024; Yoon et al., 2024). In this study to achieve this, the physical properties of soil that affect the unfrozen water content of sandy soil were evaluated, and numerous experiments were conducted to measure the unfrozen water content at subzero temperatures under various experimental conditions. Subsequently, an ANN model was developed from the experimental results, and an independent evaluation was conducted to verify the reliability of the model using data excluded from model development.

2. Measurement

Various factors affect the unfrozen water content. It is closely related to the pore size and shape, type and concentration of electrolyte, surface characteristics of solid particle (Wang et al., 2022), which in turn significantly influence moisture transport and thermal properties during the freezing process. Based on the results of these previous studies, this study selected void ratio (e), salinity concentration (ζ) of the pore water, and particle size distribution as the main factors affecting the unfrozen

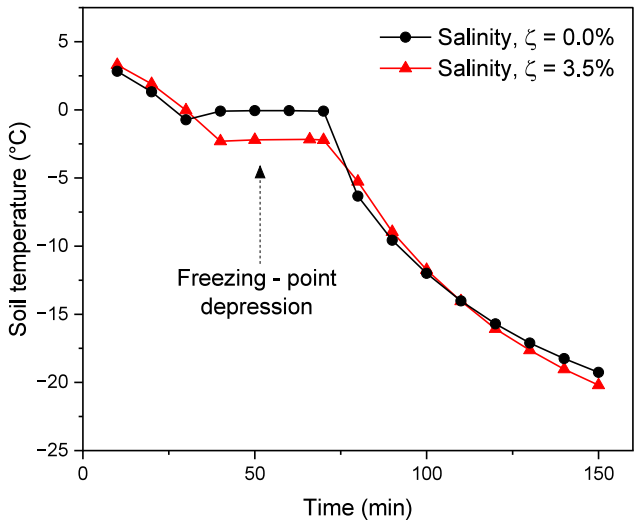

Fig. 2. Freezing point depression occurred in saline soil.

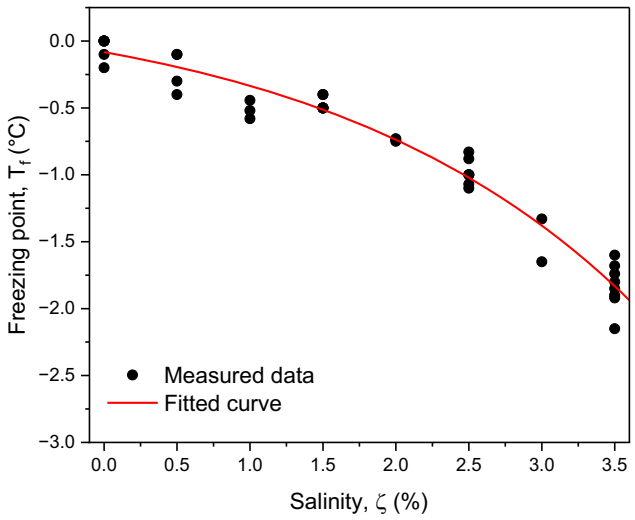

Fig. 3. Curve fitting results of freezing point based on salinity changes.

water content of soil. Subsequently, the indoor experiments were conducted under varied conditions. The target samples were standard sand from Jumoonjin and soil samples from Pohang. The fundamental characteristics of the target samples were determined by sieve analysis, and the results of the particle size distribution analysis are illustrated in Fig. 1. Following the Unified Soil Classification System (USCS) classification (ASTM, 1985), the two samples were identified as sand. Among them, Jumoonjin sand was determined to be poorly graded sand (SP) with particles sized concentrated between 1.19 mm and 0.425 mm. The Pohang sample was found to be well-graded sand (SW) with a suitable particle size distribution. Artificial salt was used to prepare the brine, simulating natural seawater. The preparation method followed ASTM D1141–98 (2013), an international standard for reproducing seawater composition in the laboratory. The artificial brine was prepared using Instant Ocean artificial sea salt. Distilled water and artificial salt were mixed to achieve the desired salinity and then heated using a hotplate with a magnetic stirrer.

2.1. *Freezing point measurement experiment*

Calculating the unfrozen water content requires determining the freezing point of the soil. However, in soils where salt is present in the pore water, it is often difficult to reliably estimate the freezing point. When pure water freezes, the water molecules are bonded in a specific arrangement; however, the presence of impurities, such as salt, hinders this arrangement and lowers the freezing point. The higher the salinity, the greater the decrease in the freezing point and the more delayed freezing occurs, resulting in more unfrozen water content at the same temperature, which has a significant impact on the calculation of unfrozen water content. Fig. 2 shows the temperature changes of some samples during a freezing point measurement experiment. The freezing point of sandy soil saturated with fresh water (salinity, $\zeta = 0$ %) is similar to the freezing point of water (0°C), but if the pore water is salt water (salinity, $\zeta = 3.5$ %), the freezing point of the sandy soil is lower. In this study, the freezing point was defined as the temperature at which phase change begins, i.e., the initial freezing point (Joudieh et al., 2024).

Therefore, an indoor freezing experiment was conducted to estimate the freezing point of saturated sandy soils where salt was present in the pore water. The sand was saturated at different salinity levels (salinity, $\zeta = 3.5$ %, 3.0 %, 2.5 %, 2.0 %, 1.5 %, 1.0 %, 0.5 %, 0 %) in a mold measuring 50 mm in diameter (D) × 119 mm in height (H), and the freezing point was measured during the freezing process using TEMPOS and TR-3 sensors from the METER Group. Since TEMPOS can simultaneously measure temperature, it was used as a temperature sensor in this study. The data were used to estimate the freezing point according to the salinity of the pore water; the experimental results are presented in Fig. 3. It was shown that the freezing point depression increased in strength in proportion to the salinity of the pore water. Furthermore, a formula for estimating the freezing point of sand samples based on the salinity concentration of pore water was proposed based on the results of the laboratory experiments. The rationale behind choosing the exponential function over its linear and polynomial counterparts pertains to its superior ability to explain the variation in the data, as evidenced by its higher coefficient of determination (R^2).

$$T_f = -0.4425 \cdot \exp(0.46798 \cdot \varsigma) + 0.3395 \quad (1)$$

2.2. *Unfrozen water content measurement experiment*

An indoor experiment was conducted to evaluate the effects of void ratio, pore water salinity, and particle size distribution on the unfrozen water content during freezing (Fig. 4). The target samples, comprising Jumoonjin standard sand and specimens procured from Pohang, were saturated with various salinity concentrations within a mold measuring 140 mm in diameter (D) and 130 mm in height (H). These samples placed in a freezing chamber measuring 500 mm in wide (W), 500 mm in deep (D) and 600 mm in height (H), capable of cooling the internal air temperature to maximum of -90°C. Experiments were conducted until the internal temperature of the sample reached -15°C or lower, and the amount of unfrozen water content within the sample was measured using a TEROS 11. The TEROS 11 sensor measures the sample's dielectric permittivity in real time, calculates the volumetric water content based on the known relationship between dielectric permittivity and water content (i.e., unfrozen water content in this study). The sensor's reader was connected to the outside of the chamber, allowing real-time monitoring of the temperature and unfrozen water content within the sample even during freezing.

The experimental conditions are listed in Table 2. As illustrated in Fig. 5 (a), the effect of the void ratio on the unfrozen water content was more pronounced in the SP sample. As the void ratio increased, the unfrozen water content decreased at a more rapid rate during the initial phase of freezing (0 to 5°C). The effect of the salinity concentration of the pore water on the unfrozen water content in the SW group (Fig. 5 (b)); the higher the salinity concentration of the pore water, the slower the decrease in the unfrozen water content, and the residual water content decreased. This is thought to be because the salinity concentration of pore water continues to increase as freezing progresses, which delays freezing, and the higher the salinity concentration, the smaller

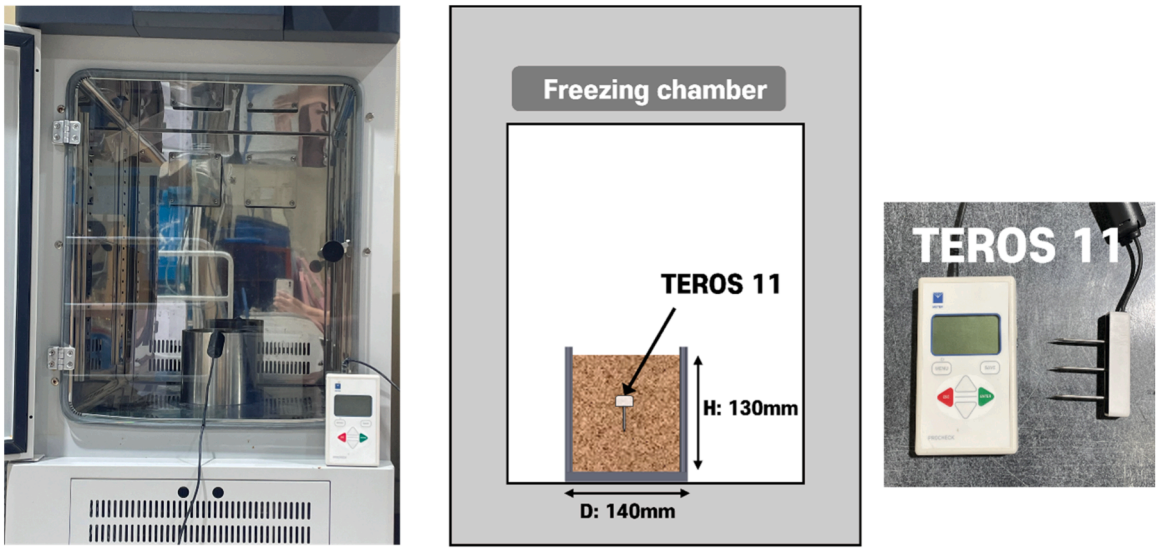

Fig. 4. Experimental setup for measurement of the unfrozen water content during freezing.

Table 2
Specimen types and properties for unfrozen water content measurement.

Soil type	Void ratio	Salinity
SP	0.55 ~ 0.67	0.0 % ~ 3.5 %
SW	0.55 ~ 0.64	0.0 % ~ 3.5 %

the proportion of pure water.

The discrepancy in the effects of the void ratio and salinity concentration on the unfrozen water content, which depends on the particle size distribution of the sandy soil, is believed to be attributable to the divergence in the fine particle content and soil particle size. Previous research confirmed that as the fine particle content increases, the specific surface area of the particles increases because of the diffusion double layer and the freezing process of pore water is hindered due to its electrical properties (Mitchell, 2005) . Soil particles typically carry a negative charge on their surface, and to neutralize this charge, cations are absorbed onto the particle surface to form a diffused double layer. The cation concentration increases with increasing salinity, as indicated by the growth of Na+ and K+ cations. The cation concentration inversely correlates with the thickness of the diffuse double layer. The negative charge on the particle surface cannot be neutralized sufficiently if there are insufficient cations present. Thus, the cations spread widely around the soil particles, forming a thick, diffused double layer. Conversely, when there is an excess of cations, they are concentrated on the surfaces of the soil particles, resulting in a thinner diffuse double layer. The SP samples were less sensitive to changes in salinity and showed a heightened sensitivity to void ratios. This is due to the lower content of fine particles. This occurs because the diffuse double layer was not significantly compressed even when the salinity increased. Instead, they were sensitive to changes in the void ratio. Because the pore size was relatively large compared with that of the SW samples, the effect was thought to be more direct and pronounced when the void ratio changed. In contrast, the SW samples were susceptible to salinity and were less affected by the void ratio. The SW samples exhibited a higher fine particle content than the SP samples; therefore, the salinity increased, and more pronounced compression of the diffuse double layer occurred as the salinity increased. This, in turn, led to a discernible decline in the residual water content. Furthermore, the presence of salt resulted in a slower rate of decrease in unfrozen water than in the absence of salt. Additionally, the rate of decline became increasingly slow as the salinity of the pore water increased. This is attributed to the continuous increase in salinity within the pore water as the freezing process progressed.

3. ANN model

In general, the ANN model consists of three layers: input, hidden, and output. The input layer is responsible for obtaining data from the database sets and transmitting the information to the neurons in the hidden layer. The hidden layer contains neurons that connect the input and output layers. These neurons receive data as a weighted sum from the previous layer and are then transformed through a transfer function to transmit information to the neurons in the next layer. Finally, the returned value in the output layer represents the value predicted using the ANN model. Generally, the output value of each neuron is expressed as follows:

$$y = f(Wx + b) \tag{2}$$

where W is a weight matrix that stores the connection weights between the current and previous layers. Vectors (x) and (b) represent the input signal and bias value of the neuron in the current layer, respectively. The transfer function f is a nonlinear function that enhances ANN performance and ensures the continuity between the input and output layers. Various transfer functions are used in ANN models; however, the tangent sigmoid function $f(x) = 2/(1+e\text{-}2x)\text{-}1$ is a commonly used transfer function in the ANN model. Therefore, the sigmoid function was applied to the neurons in the ANN, except for the output layer, where the linear function $f(x) = x$ was used.

ANN training generally involves two steps: forward and backward propagation. Before training, the weight matrix and bias values were randomly initialized. However, to enhance ANN performance, the Nguyen–Widrow method (Nguyen and Widrow, 1990) was applied to determine the initial weights and biases of the ANN model. The data transmitted from the input layer through each neuron in the hidden layer were processed using a transfer function. In the output layer, the loss function measures the difference between the two predicted values from the ANN model and the expected value from the dataset and is expressed by the following formula:

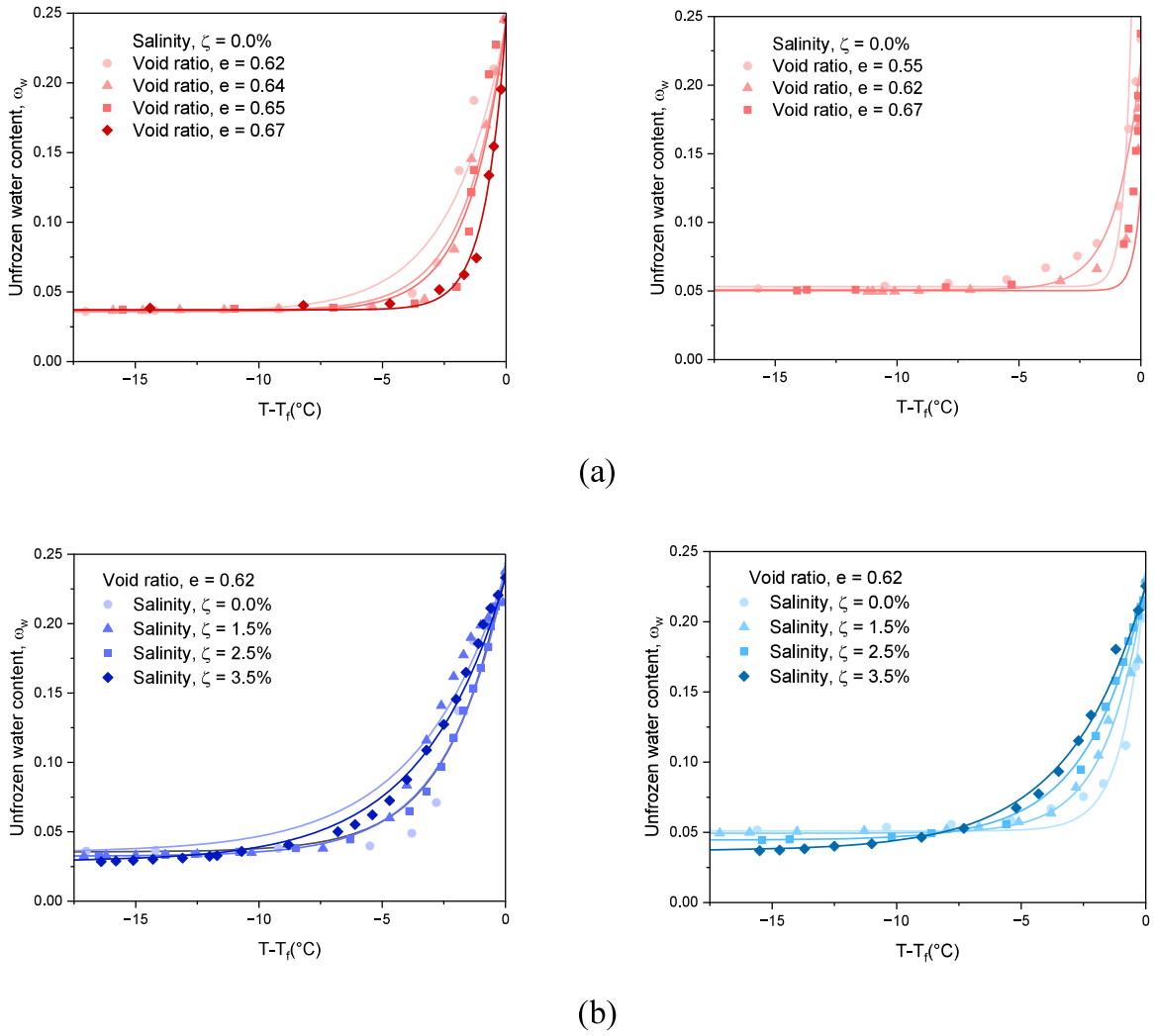

Fig. 5. Unfrozen water content during freezing based on (a) void ratio, (b) salinity.

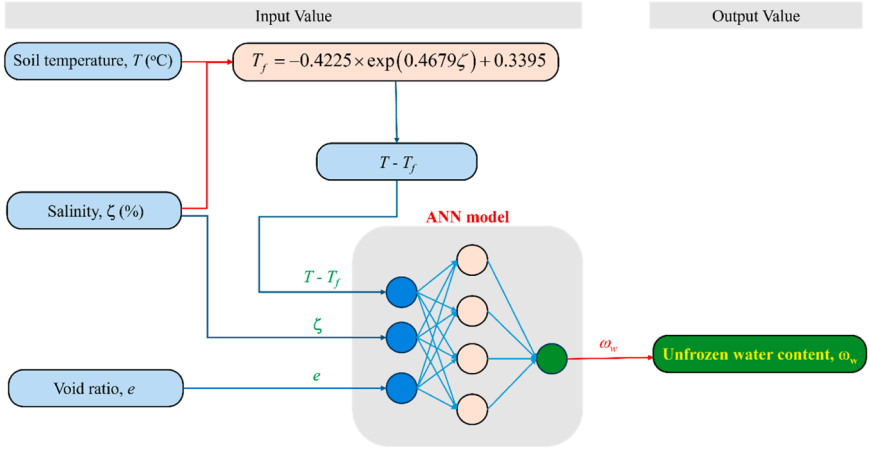

Fig. 6. Schematic view of ANN architecture to estimate unfrozen water content.

$$C = \sum (y_i - \exp_i)^2 \tag{3}$$

where y_i and exp_i are the values predicted by the ANN and the expected values from the dataset, respectively.

During the backpropagation phase, the weight matrix W^+ is updated through each training step based on the current weight matrix W and linear gradient function of the loss function ∇C. The matrix is expressed as follows:

$$W^+ = W - \eta \times \nabla C = W - \eta \times [\partial C/\partial \omega_1 \cdots \partial C/\partial \omega_n]^T \tag{4}$$

In which η is learning rate.

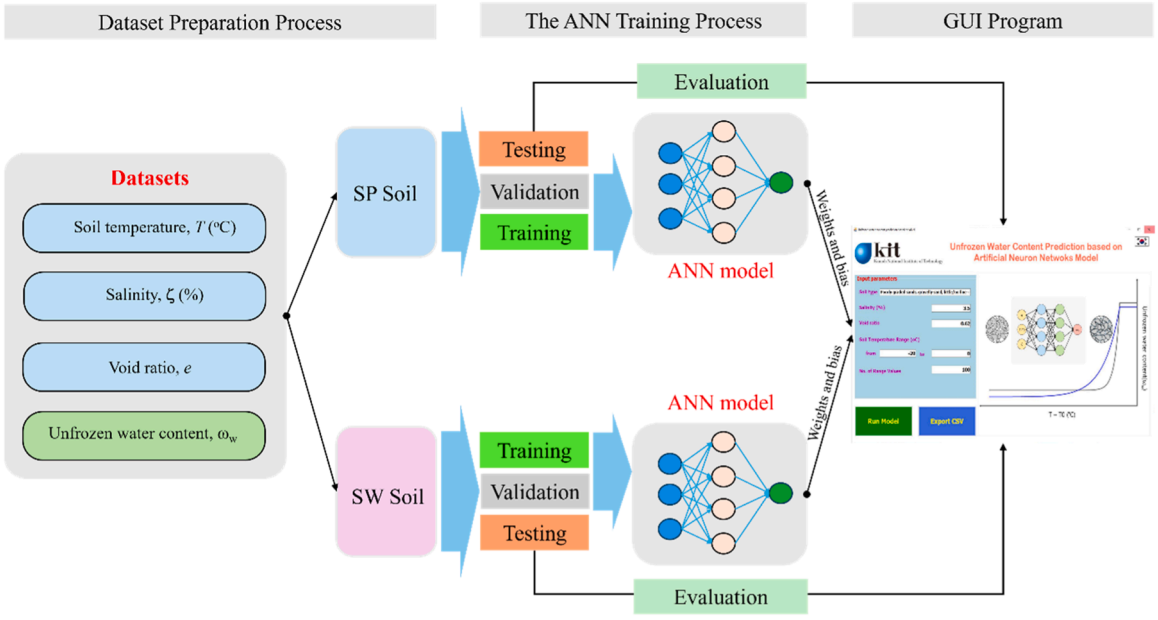

Fig. 7. A combination of the ANN models and end user GUI program.

Although the weight matrix of an ANN model can be easily updated through the training steps, there are two problems: overfitting and underfitting. These two essential problems must be addressed during training of the ANN model. Underfitting occurs when the ANN model cannot learn essential features from the training data, leading to poor predictions for both training and testing data. The leading cause of underfitting was that the model was too simple and insufficiently complex to fully represent the relationships between the input parameters and target values. This can be overcome by increasing the number of neurons in the hidden layer or the number of hidden layers. Conversely, overfitting occurs when the model becomes too complex, learning specific details of the training data, including bias, which leads to poor performance of the testing data. Regularization methods, such as Bayesian regularization, dropout, and dataset augmentation, can be applied to minimize overfitting. In this study, Bayesian regularization was adopted because it is simple and easy to use during the training process. According to the Bayesian method, the Eq. (3-2) for calculating the loss function can be rewritten as follows:

$$C = \chi E_D + \delta E_W \tag{5}$$

where E_D and E_w are the sums of the squared errors and squared weights of the ANN model, respectively. Additionally, χ and δ are parameters that can be calculated using the Gauss–Newton approximation method.

4. Result and discussion

This study aimed to predict the unfrozen water content in sandy soil samples using only basic physical properties. To achieve this, a new ANN model that predicts the unfrozen water content based on an experimental database is proposed. The ANN models were divided into two sample groups (SW and SP) according to the particle size distribution of the sand samples. Each ANN model had three input parameters: T- T_f, salinity, and void ratio. The database included 26 datasets that collectively encompass 273 experimental data pairs. Of these, 17 sets were classified into the SP group, while the remaining nine were assigned to SW(see Fig. 6).

Each data group was split into training and testing datasets. Notably, the testing dataset was independent and did not participate in the ANN training. Therefore, three datasets were selected randomly to test each

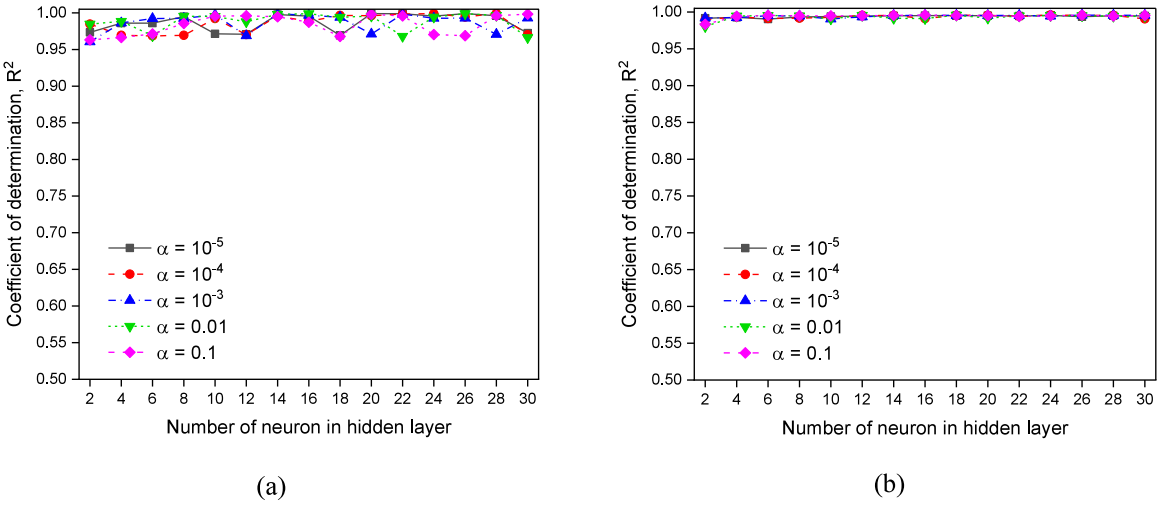

Fig. 8. Relationship between R2 value and number of neurons in the hidden layer for training process: (a) SP soil; (b) SW soil.

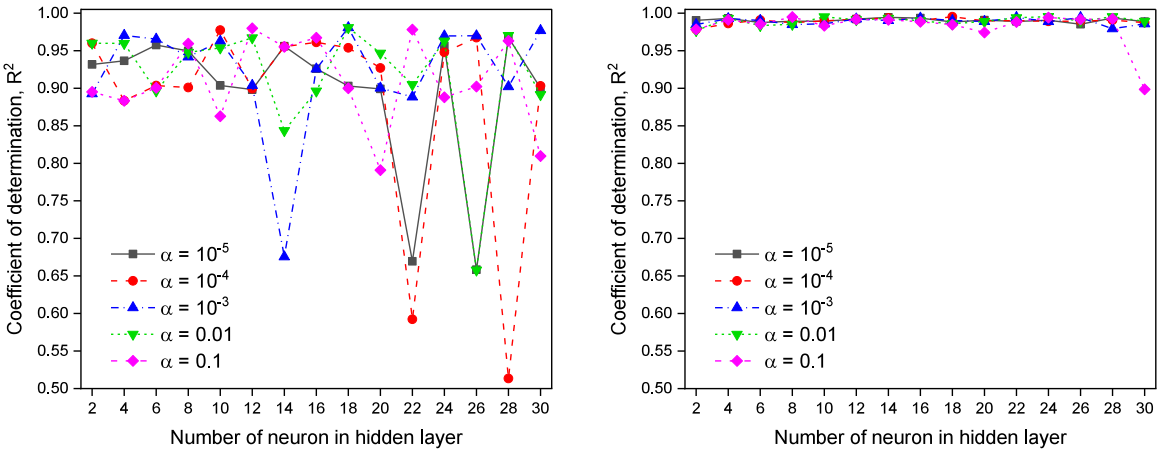

Fig. 9. Relationship between R2 value and number of neurons in the hidden layer for validation process: (a) SP soil; (b) SW soil.

soil type. In contrast, the number of datasets used in the ANN model training process for the SP and SW soil types were 14 and 6, respectively. Additionally, the training dataset was randomly arranged and divided into two sub-datasets: training and validation data at ratios of 75 % and 25 %, respectively. These datasets had different numbers of input–target data pairs and were randomly shuffled before being divided into two sub-datasets for training purposes. As a result, 141 training and 47 validation data pairs were used for the SP ANN model. There were 64 training and 21 validation data pairs for the ANN model for the SW soil type. After training, the weight and bias matrices of each ANN model for each soil type were implemented in a GUI object-oriented software programmed in C# to enable users to quickly calculate the unfrozen water content from the desired input parameters. The combination of the ANN model and the GUI software is shown in Fig. 7.

The selected ANN architecture comprises three layers: input, hidden, and output layers. ANN models can include multiple hidden layers. However, in the context of the ANN architecture, a single hidden layer is typically employed to reduce the computational burden and accommodate an increased number of neurons. It is worth noting that the number of neurons in the hidden layer and the learning rate were investigated to evaluate their influence on the performance of the ANN network model for each soil type. Five learning rates were investigated, specifically 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, and 0.1. The number of neurons in the hidden layer varied from 2 to 30, with 15 values. Following the previous study (Yoon et al., 2021), the tangent sigmoid function *f(x) = 2/(1+e-2x)-1* was adapted to the neurons in the ANN, except for the output layer, where the linear function f(x) = x was used. The tangent sigmoid function maps input values to a bounded output range of (-1, 1), which is particularly effective for normalizing neuron activations and facilitating stable gradient-based learning during backpropagation. In contrast, the linear transfer function was applied to the output layer to produce unbounded output values, which is appropriate for regression tasks in this study. Additionally, early stopping and Bayesian Regularization techniques were applied to the ANN network model. The maximum number of training steps required is 10,000. The ANN network model for each soil type was developed based on statistics and the machine-learning toolbox library in MATLAB. The results of the determination coefficient (R^2) calculations in the training and validation processes for each soil type are shown in Figs. 8 and 9, respectively.

Fig.8 shows the change in the R^2 coefficient during the training process of the ANN model for the SP and SW soil types. Overall, the R^2 coefficient was relatively high (above 0.95) for both ANN models, proving that the training process was excellent. In particular, the influence of the learning rate on the ANN model for the SW soil type did not differ, even with an increase in the number of neurons. Meanwhile, the R^2 coefficient during the training process of the ANN model for the SP soil type fluctuated when the number of neurons in the hidden layer and the learning rate were changed. Therefore, it is challenging to select

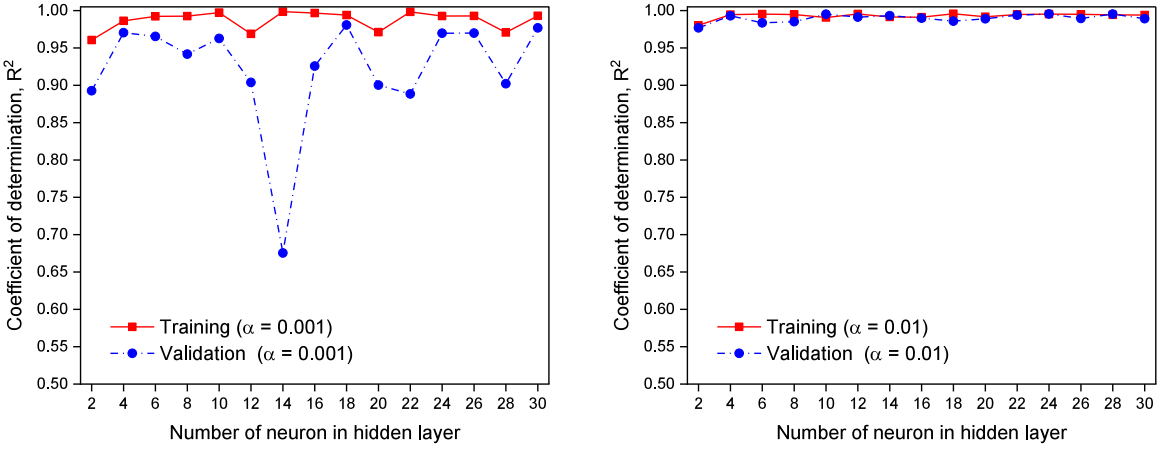

Fig. 10. Relationship between R2 value and number of neurons in the hidden layer for training and validation process: (a) SP soil; (b) SW soil.

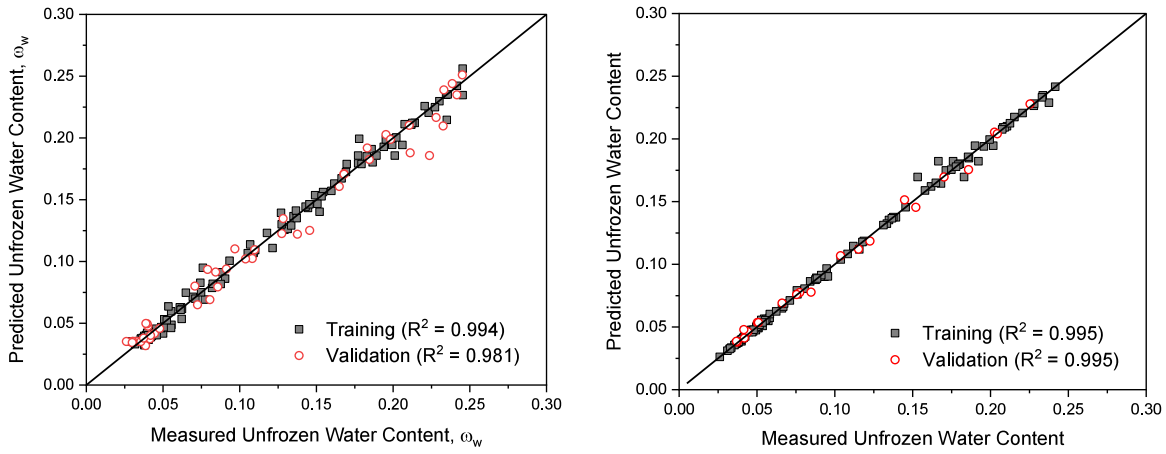

Fig. 11. Comparison of measured and predicted unfrozen water content: (a) SP soil; (b) SW soil.

the best model based only on the R^2 coefficient during the training process of the ANN model for each soil type. Consequently, the R^2 coefficient during the validation process was examined, considering the influence of the number of neurons and the change in the learning rate.

Fig. 9 shows the change in the R^2 coefficient with the validation data of the two ANN models for the two soil types, SP and SW, considering the changing number of neurons in the hidden layer and the learning rate. It can be observed that the learning rate exerts a considerable influence on the performance of the ANN model for soil type SP, whereas no discernible impact is evident for soil type SW. Based on the result of the R^2 coefficient, the selected learning rates of the ANN model for the SP and SW soil types were 0.001 and 0.01, respectively.

Fig. 10 shows the change in the R^2 coefficient with the training and validation data of the ANN model for SP and SW soil types with the selected learning rate. Based on the results mentioned above, 18 and 10 neurons were chosen in the hidden layer for the ANN model of the SP and SW soil types, respectively. Thus, the optimal ANN model for the SP soil type had 18 neurons in the hidden layer with a learning rate of 0.001, whereas the optimal ANN model for the SW soil type had ten neurons in the hidden layer with a learning rate of 0.01.

Fig. 11 shows a comparison of the predicted and measured values of the unfrozen water content of the two soil types, SP and SW. The R^2 coefficients of both models with the training and testing files were very high (above 0.98), indicating that the predicted values of the unfrozen water content of the soil were close to the experimentally measured values. It is worth noting that the R^2 coefficients of the training data files and the testing data files of the ANN model for the SW soil type are equal (R^2=0.995), indicating that the prediction performance of the ANN model works well for this soil type.

To evaluate the sensitivity and importance of the input parameters used in building the ANN model, Garson's analysis (Garson, 1991) was adapted to analyze the weight and bias data from the pre-trained ANN model for both the SP and SW soil types. The results from the Garson analysis technique would represent the distribution of connection weights between the layers of the ANN model, providing valuable information about the essential features corresponding to the input variables of the ANN model. According to this method, the distribution of the connection weights from the input to the output of an ANN model can be determined as follows:

$$\frac{\sum_j^{N_H}\left(\frac{I_{V_j}}{\sum_k^{N_V} I_{V_j}} O_j\right)}{\sum_i^{N_V}\sum_j^{N_H}\left(\frac{I_{V_j}}{\sum_k^{N_V} I_{V_j}} O_j\right)} \tag{6}$$

where N_H and N_V are the numbers of neurons in the hidden layers and variables (input parameters), respectively. I_V represents the sum of the products of the input connection weights in the hidden layer, and O represents the connection weight of the output node.

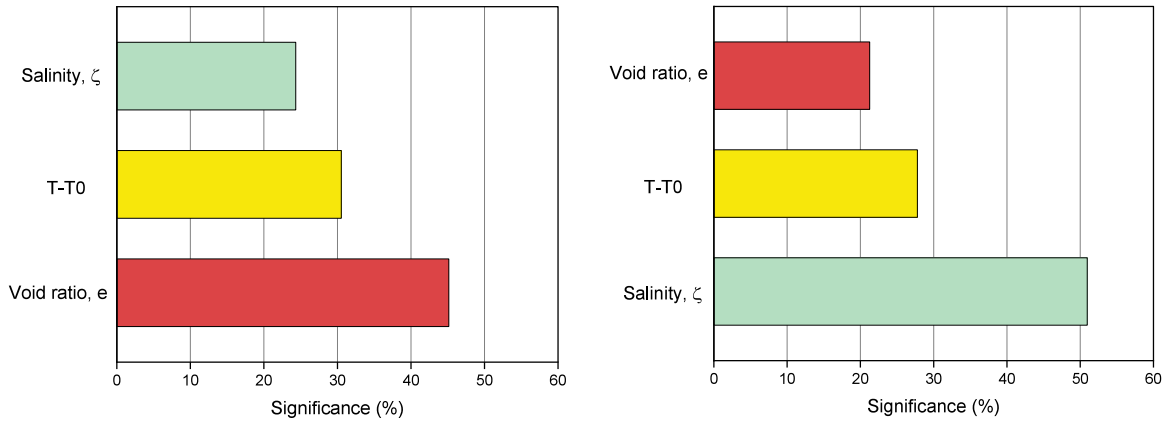

Fig. 12. Importance Plot based Garson Analysis: (a) SP soil; (b) SW soil.

Table 3
Garson Analysis for SP soil model.

		Connection weights (Input to hidden)				Connection shares × hidden node input		
		Para 01	Para 02	Para 03	OUT	Para 01	Para 02	Para 03
Hidden node	1	-2.3923	6.8311	-2.3151	3.5505	0.7361	2.1020	0.7124
	2	-7.1559	-1.7993	6.1832	3.6191	1.7107	0.4302	1.4782
	3	3.4549	4.6626	-4.8915	5.1534	1.3686	1.8470	1.9377
	4	4.9120	7.9103	-1.6638	3.0307	1.0277	1.6549	0.3481
	5	-0.4763	0.2960	-0.1794	0.6281	0.3144	0.1954	0.1184
	6	-0.4778	0.2981	-0.1801	0.6309	0.3153	0.1967	0.1188
	7	-0.4782	0.2986	-0.1802	0.6316	0.3156	0.1971	0.1189
	8	2.9351	-2.5147	-3.5049	-5.5215	1.8098	1.5506	2.1611
	9	6.6996	-4.5106	-3.7648	3.2759	1.4656	0.9867	0.8236
	10	1.4310	5.7097	-0.7648	-7.5556	1.3677	5.4570	0.7309
	11	-0.4759	0.2954	-0.1792	0.6273	0.3141	0.1950	0.1183
	12	0.6705	-6.8135	3.9232	-3.0780	0.1809	1.8385	1.0586
	13	-1.5248	-3.4163	-0.7439	-3.5769	0.9594	2.1495	0.4681
	14	-0.4758	0.2953	-0.1792	0.6272	0.3140	0.1949	0.1183
	15	-0.4397	0.2528	-0.1651	0.5657	0.2900	0.1667	0.1089
	16	-0.4758	0.2953	-0.1792	0.6272	0.3140	0.1949	0.1183
	17	-3.9655	5.2822	-2.3693	-2.9847	1.0189	1.3571	0.6087
	18	-0.4765	0.2962	-0.1795	0.6284	0.3145	0.1955	0.1185
Sum:						14.1373	20.9097	11.2657
	Input node share of output layer connections, excluding bias ones							
						Para 01	Para 02	Para 03
						30.53 %	**45.15 %**	**24.33 %**

Table 4
Garson Analysis for SW soil model.

		Connection weights (Input to hidden)				Connection shares × hidden node input		
		V1	V2	V3	OUT	V1	V2	V3
Hidden node	1	-0.0135	0.0118	-0.0116	-0.0212	-0.0135	0.0118	-0.0116
	2	-0.0584	0.0512	-0.0504	-0.0925	-0.0584	0.0512	-0.0504
	3	-0.0242	0.0212	-0.0208	-0.0382	-0.0242	0.0212	-0.0208
	4	-5.6246	-0.8116	2.9081	-5.6021	-5.6246	-0.8116	2.9081
	5	0.0139	-0.0122	0.0120	0.0222	0.0139	-0.0122	0.0120
	6	0.4466	-0.3807	0.4088	0.7192	0.4466	-0.3807	0.4088
	7	0.0066	-0.0058	0.0057	0.0106	0.0066	-0.0058	0.0057
	8	0.0987	-0.0865	0.0854	0.1568	0.0987	-0.0865	0.0854
	9	0.0877	-0.0769	0.0758	0.1392	0.0877	-0.0769	0.0758
	10	-8.9425	0.3056	7.6823	4.8287	-8.9425	0.3056	7.6823
	11	-0.0568	0.0498	-0.0490	-0.0899	-0.0568	0.0498	-0.0490
	12	-0.0581	0.0510	-0.0502	-0.0921	-0.0581	0.0510	-0.0502
	13	-6.4104	-7.6575	7.6079	5.1711	-6.4104	-7.6575	7.6079
	14	-0.0810	3.2557	3.7867	3.8526	-0.0810	3.2557	3.7867
	15	0.0906	-0.0794	0.0783	0.1439	0.0906	-0.0794	0.0783
	16	-0.0545	0.0479	-0.0471	-0.0864	-0.0545	0.0479	-0.0471
	17	0.0996	-0.0873	0.0862	0.1582	0.0996	-0.0873	0.0862
	18	0.0861	-0.0755	0.0745	0.1368	0.0861	-0.0755	0.0745
	19	0.0876	-0.0768	0.0757	0.1391	0.0876	-0.0768	0.0757
	20	-0.7837	-2.4879	10.6356	-8.5023	-0.7837	-2.4879	10.6356
	21	-0.0540	0.0474	-0.0466	-0.0855	-0.0540	0.0474	-0.0466
	22	2.5460	2.7849	-7.4843	5.1820	2.5460	2.7849	-7.4843
	23	-0.0568	0.0499	-0.0490	-0.0900	-0.0568	0.0499	-0.0490
	24	-0.0556	0.0488	-0.0480	-0.0881	-0.0556	0.0488	-0.0480
Sum:						9.8444	7.5391	18.0653
	Input node share of output layer connections, excluding bias ones							
						V1	V2	V3
						27.77 %	**21.27 %**	**50.96 %**

Fig. 12 shows the importance chart based on Garson's analysis of the input parameters of the ANN model for SP and SW soil types. Consequently, void ratio played the most crucial role in the SP soil type, whereas salinity played the most vital role in the SW soil type. In both ANN models, the T- T_f difference plays a neutral role, with an importance level ranging from 28 to 30 % compared to the other two parameters. Thus, salinity plays an essential role in the development of unfrozen water content for fine-grained soils (SW soil type). In contrast, the void ratio plays a decisive role in the development of unfrozen water content for large-grained soils, such as SP soil type.

Tables 3 and 4 present the connection weights and bias values for each layer of the ANN model for the SP and SW soil types. According to these tables, there are two subsections: Connection Weights and Sharing Connection Weights with Hidden Node Inputs. The Connection Weights section describes the weight matrices of the ANN model for the soil types. In contrast, the remaining section shows the calculation results according to Garson using Eq. (6). The columns "Para 1" to "Para 3" correspond to the connections derived from the input parameters of the ANN network. These parameters correspond to three input parameters: T- T_f difference, salinity, and void ratio e. Because the two ANN models for the two soil types, SP and SW, had the same ANN network architecture, both ANN models had two weight matrices, although they had

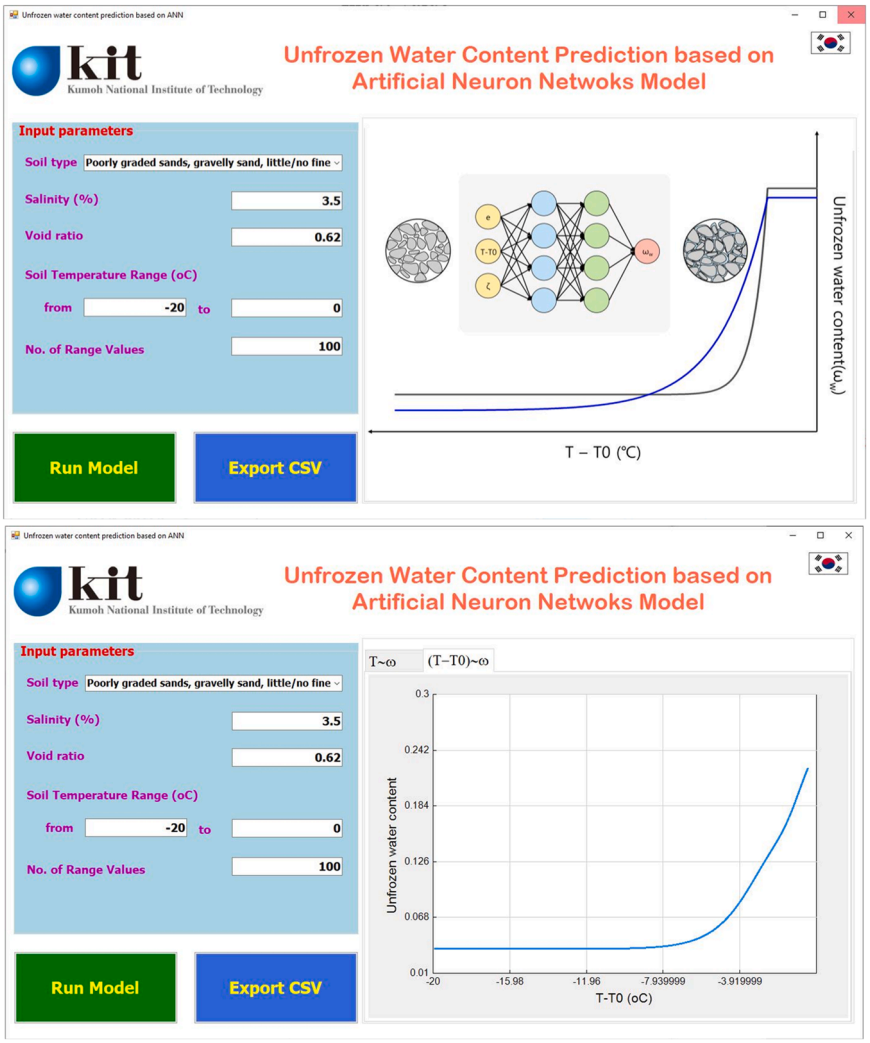

Fig. 13. Graphic user interface program for predicting the unfrozen water content.

different rows due to the other numbers of neurons in the hidden layer. The first weight matrix, connecting the input layer and the hidden layer with the number of rows equal to the number of neurons in the hidden layer, is shown from columns "Para 1" to "Para 3" in the Connection Weights section. The second weight matrix connecting the hidden and output layers appears in the "OUT" column, with the number of rows equal to the number of hidden-layer neurons. The difference in the input weight distributions of the two artificial neural network (ANN) models demonstrates the influence of the mechanical properties and particle size distribution of the two soils. SP soil, with its uniform particle size distribution and narrow particle size distribution, is significantly affected by void ratio, as evidenced by the highest weight (45.15 %). Void ratio determines the interparticle space, which directly affects the unfrozen water content through capillary action and particle surface area. Conversely, SW soil, with its wide and diverse particle size distribution, is significantly affected by salinity, with the highest weight (50.96 %). The fine-grained particles in SW soil increase the specific surface area, enhancing the interaction between salinity and soil water. This, in turn, affects the unfrozen water content through physical and chemical processes such as ion adsorption and capillary pressure changes. The temperature gradient (T- T_f) shows similar weights in both soils, suggesting that temperature plays an important role in controlling the frozen state of water. Therefore, the differences in soil particle composition and structure between SP and SW reasonably explain the different weight distributions in the ANN models, reflecting the soil mechanical mechanisms governing the unfrozen water content.

After the training process was completed, the weight and bias matrices of the two ANN models for the SP and SW soil types were implemented using GUI software developed in the C# language. The program interface was designed as shown in Fig. 13, and it includes input parameters such as soil type, salinity, void ratio, and soil temperature value domain, along with the desired number of calculation points. The software integrated an empirical formula to calculate T- T_f from two input parameters: salinity and soil sample temperature. When the user selects the soil type, the weight and bias matrices corresponding to the pretrained ANN model for that soil type are used and calculated to provide the predicted value of the unfrozen water content throughout the soil temperature value domain. Two graphs depicting the relationship between unfrozen water content and soil specimen temperature, as well as the relationship with T- T_f, are shown in the graphic area of the program. Users can export the corresponding prediction data as CSV files to aid the detailed analysis of the calculation results and use them as inputs for numerical simulation models in commercial software.

With the many outstanding features mentioned above, the GUI software can be used to predict the unfrozen water content with any input parameter that was not used during the training process. Therefore, this study used GUI software to evaluate the two ANN models for SP and SW soil types with testing datasets that were not used during the training process. The results of the predicted unfrozen water content and its relationship with T- T_f corresponding to the testing datasets are

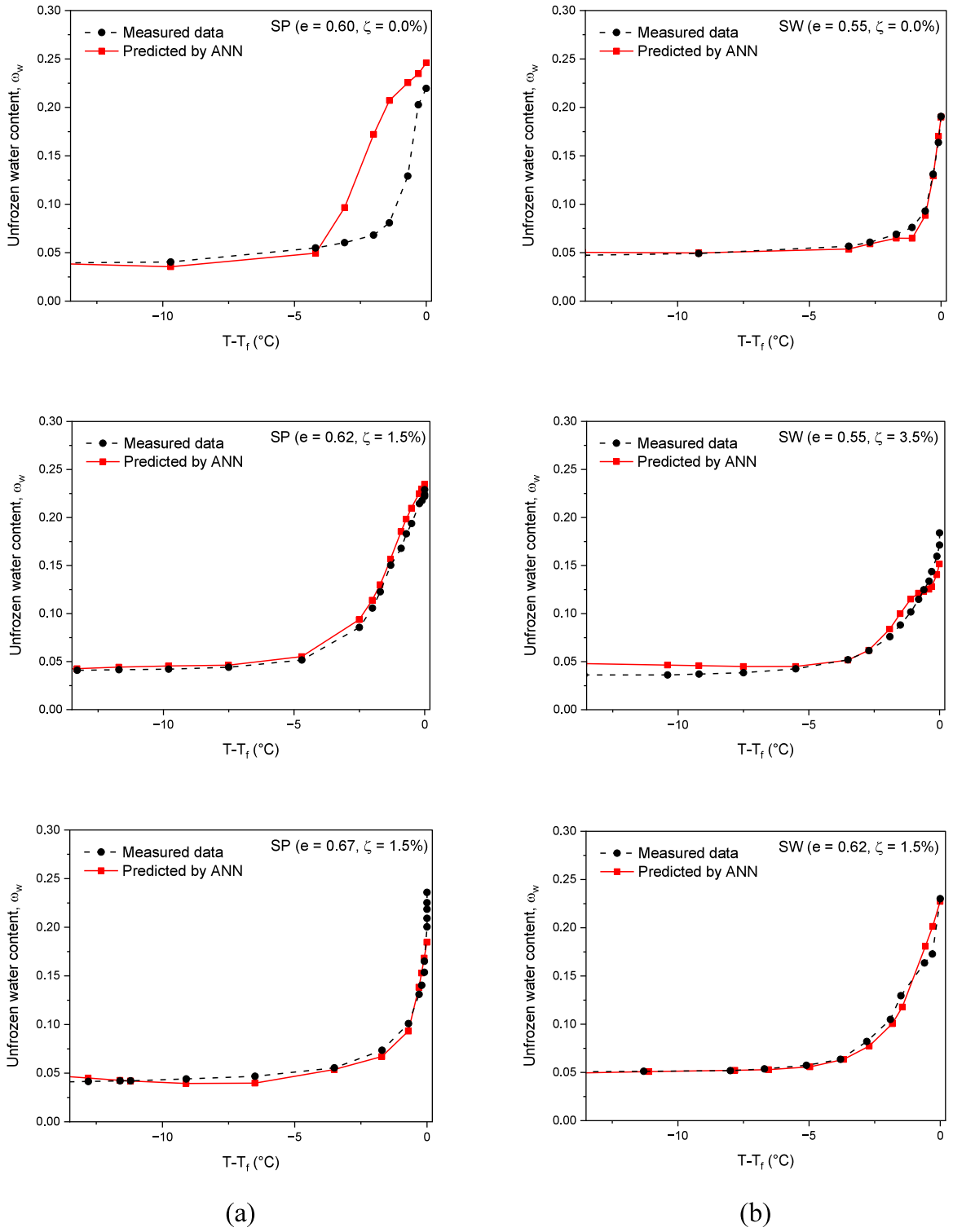

Fig. 14. Comparison of predicted and measured unfrozen water content for testing datasets: (a) SP soil; (b) SW soil.

shown in Fig. 14 (a) for the SP soil type and Fig. 14 (b) for the SW soil type. In general, a nonlinear relationship exists between the predicted unfrozen-water content and T- T_f. The results show that the ANN model for the SP soil type predicts accurately for two cases with a void ratio of 0.62 and a void ratio of 0.67. In contrast, the ANN model predicts a significant error for the datasets with a void ratio of 0.6 and a salinity of 0 %. However, the model still gives accurate prediction results when the T- T_f is less than -4. Meanwhile, the ANN model for the SW soil type provided accurate predictions of unfrozen water content for all datasets with different void ratios and salinities.

5. Conclusions

To make reliable predictions and evaluate frozen sandy soils, it is essential to understand the unfrozen water content under various conditions. In this study, an ANN model was developed to predict the unfrozen water content in frozen sandy soils using only their physical characteristics. Laboratory experiments were conducted under various

conditions to generate data to train the ANN model. To evaluate the reliability of the model, an independent dataset not used during training was employed for validation. The main conclusions of this study are summarized as follows:

(1) The void ratio and salinity, both of which affect the amount of unfrozen water content, varied depending on the quantity and size of the fine particles in the soil. In SP samples with low fine particle content and large pore sizes, void ratio had a greater impact on the amount of unfrozen water content, indicating that changes in void ratio led to more significant changes in the amount of unfrozen water content.
(2) In contrast, the effect of salinity concentration was more pronounced in the SW sample. It is believed that the high finer grain content of the sample led to greater compression of the diffuse double layer due to salinity concentration, which also contributed to the decrease in residual unfrozen water content. Furthermore, salinity delayed freezing at the initial stage. This delay is believed to be due to the continued increase in salinity concentration in the pore water as freezing progresses.
(3) The ANN model developed based on the experimental data showed high performance in predicting the unfrozen water content of the SP and SW samples. Additionally, a user-friendly GUI software was developed, and the results confirmed that the ANN model could effectively predict the unfrozen water content.
(4) The GUI software allows easy estimation of unfrozen water content. This software provides an intuitive user interface, and users can quickly and accurately calculate the unfrozen water content by simply inputting the basic physical properties of the sample. It is expected to be practically used in various fields, including geotechnical engineering, as it can be easily used by researchers and non-experts.

Data availability

The authors do not have permission to share data.

CRediT authorship contribution statement

Hee Won Kim: Writing – original draft, Methodology. **Dinh-Viet Le:** Software, Methodology, Investigation. **Gyu-Hyun Go:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) (No. 2022R1C1C1006507).

References

Al-Swaidani, A. M., Meziab, A., Khwies, W. T., Al-Bali, M., & Lala, T. (2024). Building MLR, ANN and FL models to predict the strength of problematic clayey soil stabilized with a combination of nano lime and nano pozzolan of natural sources for pavement construction. *International Journal of Geo-Engineering, 15*(1), 2.

Anderson, D. M., & Tice, A. R. (1972). Predicting unfrozen water contents in frozen soils from surface area measurements. *Highway Research Record, 393*(2), 12–18.

Burt, T. P., & Williams, P. J. (1976). Hydraulic conductivity in frozen soils. *Earth Surface Processes, 1*(4), 349–360.

Chai, M., Zhang, J., Zhang, H., Mu, Y., Sun, G., & Yin, Z. (2018). A method for calculating unfrozen water content of silty clay with consideration of freezing point. *Applied Clay Science, 161*, 474–481.

Chen, R. P., Zhang, P., Kang, X., Zhong, Z. Q., Liu, Y., & Wu, H. N. (2019). Prediction of maximum surface settlement caused by earth pressure balance (EPB) shield tunneling with ANN methods. *Soils Found, 59*(2), 284–295.

Chen, Y., Zhou, Z., Wang, J., Zhao, Y., & Dou, Z. (2021a). Quantification and division of unfrozen water content during the freezing process and the influence of soil properties by low-field nuclear magnetic resonance. *Journal of Hydrology, 602*, Article 126719.

Chen, Z., Guo, X., Shao, L., Li, S., & Gao, L. (2021b). Sensitivity analysis of thermal factors affecting the nonlinear freezing process of soil. *Soils Found, 61*(3), 886–900.

Garson, G. D. (1991). A comparison of neural network and expert systems algorithms with common multivariate procedures for analysis of social science data. *Social Science Computer Review, 9*(3), 399–434.

Joudieh, Z., Cuisinier, O., Abdallah, A., & Masrouri, F. (2024). Artificial ground freezing—On the soil deformations during freeze–thaw cycles. *Geotechnics, 4*(3), 718–741.

Kim, S. Y., Hong, W. T., Hong, S. S., Baek, Y., & Lee, J. S. (2016). Unfrozen water content and unconfined compressive strength of frozen soils according to degree of saturations and silt fractions. *Journal of the Korean Geotechnical Society, 32*(12), 59–67.

Kong, L., Wang, Y., Sun, W., & Qi, J. (2020). Influence of plasticity on unfrozen water content of frozen soils as determined by nuclear magnetic resonance. *Cold Regions Science and Technology, 172*, Article 102993.

Kozlowski, T. (2007). A semi-empirical model for phase composition of water in clay–water systems. *Cold Regions Science and Technology, 49*(3), 226–236.

Kozlowski, T., & Nartowska, E. (2013). Unfrozen water content in representative bentonites of different origin subjected to cyclic freezing and thawing. *Vadose zone Journal : VZJ, 12*(1). vzj2012-0057.

Kruse, A. M., Darrow, M. M., & Akagawa, S. (2018). Improvements in measuring unfrozen water in frozen soils using the pulsed nuclear magnetic resonance method. *Journal of Cold Regions Engineering, 32*(1), Article 04017016.

Kurylyk, B. L., & Watanabe, K. (2013). The mathematical representation of freezing and thawing processes in variably-saturated, non-deformable soils. *Advances in Water Resources, 60*, 160–177.

Kweon, G. C. (2003). Design of anti-frost layer considering frost heaving characteristics of subgrade soils. *Symposium Korean Society of Civil Engineers*, 1786–1791.

Li, Z., Chen, J., & Sugimoto, M. (2020). Pulsed NMR measurements of unfrozen water content in partially frozen soil. *Journal of Cold Regions Engineering, 34*(3), Article 04020013.

Liu. (2014). Influence of freeze-thaw cycles on resilient modulus of different plasticity index subgrade soil. *Chinese Journal of Geotechnical Engineering, 36*(4), 633–639.

McKenzie, J. M., Voss, C. I., & Siegel, D. I. (2007). Groundwater flow with energy transport and water–ice phase change: numerical simulations, benchmarks, and application to freezing in peat bogs. *Advances in Water Resources, 30*(4), 966–983.

Michalowski, R. L., & Zhu, M. (2006). Frost heave modelling using porosity rate function. *International Journal for Numerical and Analytical Methods in Geomechanics, 30*(8), 703–722.

Mitchell JK (2005) Fundamentals of soil behavior.

Nguyen, D. H., & Widrow, B. (1990). Neural networks for self-learning control systems. *IEEE Control Systems Magazine, 10*(3), 18–23.

Ohrai, T. (1986). *Experimental studies on the effects of ice and unfrozen water on the compressive strength of frozen soil*. Sapporo, Japan: PhD dissertation, Hokkaido Univ.

Omar, M., Shanableh, A., Mughieda, O., Arab, M., Zeiada, W., & Al-Ruzouq, R. (2018). Advanced mathematical models and their comparison to predict compaction properties of fine-grained soils from various physical properties. *Soils Found, 58*(6), 1383–1399.

Ren, J., Vanapalli, S. K., & Han, Z. (2017). Soil freezing process and different expressions for the soil-freezing characteristic curve. *Sciences in Cold and Arid Regions, 9*(3), 221–228.

Samadi, H., Hassanpour, J., & Rostami, J. (2023). Prediction of earth pressure balance for EPB-TBM using machine learning algorithms. *International Journal of Geo-Engineering, 14*(1), 21.

Shin, E. C., & Park, J. J. (2003). An experimental study on frost heaving pressure characteristics of frozen soils. *Journal of the Korean Geotechnical Society, 19*(2), 65–74.

Smith, M. W., & Tice, A. R. (1988). Measurement of the unfrozen water content of soils: comparison of NMR and TDR methods. *CRREL-88-18*. Hanover, NH: Cold Regions Res Eng Lab.

Tariq, A., Abualshar, B., Deliktas, B., Song, C. R., Al-Nimri, B., Barret, B., & Glennie, N. (2024). ANN-based evaluation system for erosion resistant highway shoulder rocks. *International Journal of Geo-Engineering, 15*(1), 17.

Teng, J., Zhong, Y., Zhang, S., & Sheng, D. (2021). A mathematic model for the soil freezing characteristic curve: the roles of adsorption and capillarity. *Cold Regions Science and Technology, 181*, Article 103178.

Tice, A. R., Anderson, D. M., & Banin, A. (1976). The prediction of unfrozen water contents in frozen soils from liquid limit determinations. *Cold Regions Research and Engineering Laboratory, US Army Corps of Engineers*.

Wang, C., Li, S., Lai, Y., Chen, Q., He, X., Zhang, H., & Liu, X. (2022). Predicting the soil freezing characteristic from the particle size distribution based on micro-pore space geometry. *Water Resources Research, 58*(1), Article e2021WR030782.

Watanabe, K., & Mizoguchi, M. (2002). Amount of unfrozen water in frozen porous media saturated with solution. *Cold Regions Science and Technology, 34*(2), 103–110.

Woo, H. J., & Go, G. H. (2024). Mechanical behavior assessment of retaining wall structure due to frost heave of frozen ground. *International Journal of Geo-Engineering, 15*(1), 7.

Xin, Q., Su, Y., Cao, Y., She, X., Wang, Z., Ling, X., & Yang, T. (2023). Experimental and modeling investigation of freezing characteristic curve of silty clay using TDR. *Cold Regions Science and Technology, 205*, Article 103715.

Yoon, S., Le, D. V., & Go, G. H. (2021). Artificial neural network-based model for prediction of frost heave behavior of silty soil specimen. *Applies Science, 11*(22), Article 10834.

Yoon, S., Jeong, H., Lee, H. L., Kim, T., Hong, C. H., & Kim, J. S. (2023). Evaluation of uniaxial compression and point load tests for compacted bentonites. *Acta Geotechnica, 18*(9), 4633–4644.

Yoon, S., Lee, G. J., Lee, D. H., Kim, M. S., Kim, J. T., & Kim, J. S. (2024). Evaluation of thermal properties for the bentonil-WRK bentonite. *Journal of Nuclear Fuel Cycle and Waste Technology, 22*(1), 9–16.

Yoshikawa, K., Overduin, P. P., & Harden, J. W. (2004). Moisture content measurements of moss (Sphagnum spp.) using commercial sensors. *Permafrost and Periglacial Processes, 15*(4), 309–318.

Zhou, M. M., & Meschke, G. (2013). A three-phase thermo-hydro-mechanical finite element model for freezing soils. *International Journal for Numerical and Analytical Methods, 37*(18), 3173–3193.