applied sciences

MDPI

Article

Application of an Artificial Neural Network (ANN) Model to Determine the Value of the Damping Ratio (D) of Clay Soils

Marzena Lendo-Siwicka, Karina Zabłocka, Emil Soból, Anna Markiewicz and Grzegorz Wrzesiński *

Institute of Civil Engineering, Warsaw University of Life Sciences, Nowoursynowska 159 St., 02-776 Warsaw, Poland; marzena_lendo_siwicka@sggw.edu.pl (M.L.-S.)
* Correspondence: grzegorz_wrzesinski@sggw.edu.pl; Tel.: +48-22-59-35-210

Abstract: The properties and behavior of soils depend on many factors. The interaction of individual factors is difficult to determine by traditional statistical methods due to their interdependence. The paper presents a procedure of creating an artificial neural network (ANN) model to determine the value of the damping ratio (D) of clay soils. The main purpose of this paper is to compare the appropriateness of ANN model application with empirical formulas described in the literature. The ANN model was developed using a series of laboratory tests of the damping ratio performed in the Resonance Column. Predicted values of the damping ratio of clay soils obtained from the ANN model are characterized by high convergence (coefficient of determination R^2 = 0.976). In comparison with other published empirical formulas, the ANN model showed an improvement in the prediction accuracy. What is more, ANN models proved to be more flexible compared to formulas and relationships with a predetermined structure, and they were well suited to modeling the complex behavior of most geotechnical engineering materials, which, by their very nature, exhibit extreme variability. In conclusion, ANNs have the potential to predict the damping ratio (D) of clay soils and can do much better than traditional statistical techniques.

Keywords: artificial neural networks (ANNs); damping ratio; cohesive soil; resonant column

check for updates

Citation: Lendo-Siwicka, M.; Zabłocka, K.; Soból, E.; Markiewicz, A.; Wrzesiński, G. Application of an Artificial Neural Network (ANN) Model to Determine the Value of the Damping Ratio (D) of Clay Soils. *Appl. Sci.* **2023**, *13*, 6224. https://doi.org/10.3390/app13106224

Academic Editors: Daniel Dias and Wei Gao

Received: 30 March 2023
Revised: 12 May 2023
Accepted: 18 May 2023
Published: 19 May 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Properties and behavior of soils are determined by the following factors: fabric, mineralogy, and pore water. Due to the interdependence of individual factors, it is difficult to determine their interaction using only traditional statistical methods [1]. Classical constitutive modeling based on the elasticity and plasticity theories is unable to properly simulate the behavior of geomaterials. This is due to the formulation complexity, the idealization of the behavior of the material, and excessive empirical parameters [2]. Therefore, in many papers, artificial neural networks (ANNs) have been proposed as a reliable and practical alternative to modeling the constitutive monotonic and hysteretic behavior of geomaterials [3–28].

This article concerns the use of artificial neural networks (ANNs) to determine the value of the damping ratio (D) of clay soils. Parameter D is one of the basic soil parameters. It characterizes the behavior of soil subjected to dynamic loads, such as the foundations of machines, the traffic of city trams and subways, and the driving of precast pile or sheet pile. These geotechnical problems are associated with significant damping, so its impact has to be taken into account in the dynamic analysis [29]. Unfortunately, the damping ratio, although important, is not easy to determine either on the basis of empirical formulas or in laboratory tests. The authors present a model of an artificial neural network (ANN) and compare the appropriateness of its application with empirical formulas described in the literature. The ANN model was developed on the basis of a series of laboratory tests of the damping ratio performed in the Resonance Column manufactured by GDS Instruments Ltd. Neural networks were to increase the accuracy of prediction. They provide an attractive

solution for determining the behavior of a complex system and they are widely used for nonlinear pattern recognition and regression.

The paper also presents the procedure of creating an artificial neural network (ANN) model to determine the value of the damping ratio (D) of clay soils. The main purpose of this paper is to compare the appropriateness of ANN model application with empirical formulas described in the literature, and it also presents a method of selecting the structure of the ANN with the best predictive quality. To determine the architecture of the neural network, the following were determined: the number of input variables, the number of output variables, the number of hidden layers, the number of neurons in hidden layers, the type of activation function in the neurons of hidden layers and the output layer, and the number of learning epochs [30]. The Python programming language and the following libraries: TensorFlow, Keras, NumPy, Pandas, and Scikit-Learn, were used to create the ANN model. This article can be used as a guideline for creating other ANN models in the field of geotechnics.

2. Materials and Methods

Soils for the tests were collected from 3 test sites in Warsaw (Poland) from different depths (from 1.5 to 9.5 m). The soil cores were extracted in their undisturbed state using Shelby tubes.

Before performing the damping tests in the Resonance Column, the particle size distribution of the soil and physical properties were studied. The obtained results are presented in Table 1.

Table 1. Soil characteristics and mean effective stress range for tested soils.

Test Site	Test No.	Depth (m)	Soil Type	w_c (%)	LL (%)	PL (%)	IP (%)	p′ (kPa)	e_o (-)	OCR (-)
Bartycka	S1	2.5	sandy clay	11.75	22.1	11.5	10.6	50–200	0.41	2.80–1.00
	S2	1.5	sandy silty clay	18.85	32.8	16.0	16.8	30–390	0.56	1.00
	S3	1.7	clayey sand	14.43	21.2	12.2	8.9	35–210	0.44	1.00
	S4	6.0	clayey sand	11.06	17.7	11.1	6.6	60–240	0.32	3.33–1.00
Pelczynskiego	S5	4.5	sandy silty clay	17.41	36.5	14.1	22.4	90	0.48	1.00
	S6	7.5	sandy silty clay	10.76	24.5	12.5	12.0	75–415	0.30	2.13–1.00
	S7	6.0	silty clay	17.53	37.3	17.1	20.1	120-410-120	0.47	1.00–3.42
Pory	K8	8.5	siCl	19.84	44.60	19.	25.11	85–310	0.5	2.00–1.00
	S9	2.2	sandy clay	12.23	37.0	11.5	25.6	45–315	0.37	1.00
Pelczynskiego	S10	2.2	clayey sand	15.57	41.7	14.3	27.4	90–315	0.40	1.00
	S11	2.2	clayey sand	10.50	18.2	9.1	9.1	45–315	0.43	1.00
	S12	7.2	silty clay	21.98	51.3	23.7	27.6	145–290	0.60	1.38–1.00
Pory	S13	8.0	silty clay	22.95	63.5	26.8	36.7	160–320	0.63	1.25–1.00
	S14	9.5	clay	26.04	71.0	33.1	37.8	95–285	0.75	6.32–2.11
Pelczynskiego	S15	2.7	sandy silty clay	12.68	27.1	12.3	14.8	55-165-55	0.39	1.00–3.00

Notes: w_c—initial water content, LL—liquid limit, PL—plasticity limit, IP—plasticity index, p′—mean effective stress, e_o—initial void ratio, OCR—overconsolidation ratio.

After the examination of the properties of the soil, the core sample was pushed out from the Shelby tube by a hydraulic press. Subsequently, from the undisturbed core, a cylindrical specimen at 140 mm high and 70 mm diameter was cut out. The prepared sample was placed on the resonant column pedestal. After that, the proper test was started. To prepare the specimen for testing, the resonant column chamber was closed, and the soil sample was saturated to a Skempton parameter B of at least 0.95. The consolidation process was then initiated, with dynamic tests conducted at every stage of the consolidation. First, the resonant frequency at torsional mode excitation was determined, followed by the application of a sinusoidal wave at the resonant frequency on top of the soil sample using the drive system. After seconds of excitation, the coils were switched off, and the specimen was left to vibrate freely. The damping of the soil sample movement was recorded by an accelerometer and displayed on the screen as a free vibration decay curve. The scientist then carefully selected the appropriate number of damping cycles at which specimen motions

were completely damped based on the free vibration decay curve. A detailed description of the performed research is presented in [31].

To generate a damping ratio curve, the amplitude of the applied voltage on the drive system was increased, and the same processes as described above were repeated. The damping tests were performed following the procedures outlined in [32]. Furthermore, every damping test was repeated 10 times at each shear strain. Later, from ten results of the parameter D, the average value was evaluated and used in the performed analysis. More details about materials and methods are presented in [29].

3. Artificial Neural Network

Artificial neural networks (ANNs) are computational methods based on the principle of operation of human neurons or nerve cells. When certain information reaches them, they process it and remember it for future use. Artificial systems work in this way as well. Data are inserted into the system, processed, stored, and corrected.

In geotechnics, the use of neural networks allows to achieve progress in the interpretation of research results, especially in relation to such specific materials as soil. When analyzing problems related to geotechnics, we most often deal with regression issues, which means a problem in which we want to predict a numerical value. Regression ANNs predict an output variable as a function of the inputs. It is required that numeric variables be dependent in regression ANNs. The use of ANNs often allows to develop better predictive models than statistical ones and to improve the models as new data become available.

According to Geron [33], in order to create any machine learning project, the following series of steps should be performed:

- Analyze the whole task;
- Get the data;
- Discover and visualize the data to recognize patterns and additional information—Exploratory data analysis;
- Prepare data in terms of machine learning algorithms—Data preparation;
- Select and teach a model;
- Tune the model—Model adjustment;
- Present the solution—Results.

In this article, we will go through all these steps to create the ANN model which is used to predict the value of the damping ratio determined based on the basic geotechnical parameters of cohesive soils.

3.1. Analyze the Whole Task

The first step is to define the purpose for which we create the project. In our case, the goal is to determine the value of the damping ratio based on the available soil parameters. This will allow us to know the value of the coefficient of a given soil without the need to perform tedious, expensive, and not always available laboratory tests. The next step is to define the problem we are dealing with. This is essential to start designing the system. We faced a classic problem of supervised learning. The task was a multiple regression task because our system would use many features to forecast the result (shear strain, effective stress, soil type, etc.). We also used a batch learning mechanism (offline learning).

The root mean square error (RMSE) was chosen as the performance metric, which, in regression problems, is a classic measure of performance. The mean absolute error (MAE) and coefficient of determination (R^2) were also determined, but the RMSE metric was more sensitive to outliers than the MAE. Equation (1) shows the mathematical formula used to compute the RMSE error:

$$RMSE(\boldsymbol{X}, h) = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(h(\boldsymbol{x}^{(i)}) - y^{(i)}\right)^2} \quad (1)$$

where:

- m—the number of elements of the training set;
- $\boldsymbol{x}^{(i)}$—the value vector of all the features of the i-th example;
- $y^{(i)}$—the label of i-th example;
- $\boldsymbol{X}$—the matrix containing the value of all features;
- h—the predictive function.

3.2. Get the Data

The data that we used to create the ANN model comes from the laboratory tests described in Section 2 herein; "Materials and Methods". The test results were saved in the .xlsx format in Excel.

The data were loaded using the Jupyter notebook and the Python Pandas module. The first five lines of "raw" data are presented below (Table 2):

Table 2. First 5 lines of "raw" data from the damping ratio laboratory test.

Test No.	Cl (%)	Si (%)	Sa + Gr (%)	LL (%)	PL (%)	IP (%)	IL (%)	p′ (%)	eo (-)	OCR (-)	γ (%)	G (MPa)	G/Gmax (-)	D (%)
S1	14	26	60	22.1	11.5	10.6	0.024	50	0.40	2.8	0.0006	53.35	1.00	4.99
S1	14	26	60	22.1	11.5	10.6	0.024	50	0.40	2.8	0.0009	54.63	1.00	4.99
S1	14	26	60	22.1	11.5	10.6	0.024	50	0.40	2.8	0.0012	54.20	1.00	5.01
S1	14	26	60	22.1	11.5	10.6	0.024	50	0.40	2.8	0.0018	53.63	0.98	5.06
S1	14	26	60	22.1	11.5	10.6	0.024	50	0.40	2.8	0.0024	53.35	0.98	5.11

Notes: Cl—clay content, Si—silt content, Sa + Gr—sand + gravel content, LL—liquid limit, PL—plasticity limit, IP—plasticity index, IL—liquidity index, p′—mean effective stress, e_o—initial void ratio, OCR—overconsolidation ratio, γ—shear strain, G—shear modulus, G/G_{max}—normalized shear modulus.

Our dataset consists of 1227 examples obtained from testing 15 soil samples. All attributes (except "Test No.") have numeric values. A summary of all numerical attributes is provided in Table 3.

Table 3. Summary of all numeric attributes.

	Cl (%)	Si (%)	Sa + Gr (%)	LL (%)	PL (%)	IP (%)	IL (%)	p′ (%)	e_o (-)	OCR (-)	γ (%)	G (MPa)	G/Gmax (-)	D (%)
count	1227	1227	1227	1227	1227	1227	1227	1227	1227	1227	1227	1227	1227	1227
mean	17.83	35.38	46.79	33.77	15.28	18.49	0.04	193.96	0.440	1.37	0.0117	120.75	0.87	4.57
std	7.26	17.01	21.83	11.74	4.86	7.58	0.11	104.61	0.104	0.85	0.0220	76.70	0.20	2.78
min	10.00	20.00	6.00	17.70	9.10	6.60	−0.19	30.00	0.280	1.00	0.0001	6.45	0.16	1.57
25%	14.00	23.00	17.00	24.50	12.24	12.03	0.01	110.00	0.375	1.00	0.0008	69.82	0.81	2.57
50%	14.00	26.00	60.00	32.75	14.26	16.79	0.02	180.00	0.433	1.00	0.0031	108.61	0.97	3.73
75%	21.00	60.00	62.00	37.25	17.14	25.11	0.07	270.00	0.510	1.37	0.0111	152.26	1.00	5.45
max	50.00	69.00	70.00	70.95	33.11	37.84	0.24	415.00	0.747	6.32	0.2219	431.76	1.02	19.04

Notes: std—standard deviation, 25%—25th percentile, 50%—median, 75%—75th percentile.

After the first look at the data, the next particularly important, but often underestimated, step is to create a test set. The separation of the test set at such an early stage is to prevent the occurrence of a phenomenon called the "data snooping bias". When we look at the test set, we can see some interesting pattern (not always right) that will make us choose a specific machine learning model based on it. When dealing with exceptionally large datasets, it is enough to randomly select some examples and put them back to create a test dataset. This is a satisfactory solution, but it does not work well with small datasets. Next time the program runs, a completely different test set is generated. After some time, the person creating the model or the machine learning algorithm itself will see the entire dataset, and this is what we want to avoid. This is why the random seed was used, which allowed to obtain the same pseudorandom test set every time. The test set accounts for 20% of the total data.

3.3. Exploratory Data Analysis

The objectives of exploratory data analysis (EDA) and exploratory data analysis (EDA) methods are classified according to [34].

3.3.1. Searching for Correlation

At this stage of the project, we only deal with training data. The first step in regression problems should be looking for the correlation between the features and the label (damping ratio, D). To do this, we calculate a linear correlation coefficient (Pearson correlation coefficient) between each pair of values (Table 4).

Table 4. Correlation matrix.

	Cl	**Si**	**Sa + Gr**	**LL**	**PL**	**IP**	**IL**	**p′**	**e_o**	**OCR**	**γ**	**G**	**G/G_{max}**	**D**
Cl	1.00	0.52	−0.75	0.85	0.84	0.78	−0.48	0.12	0.65	0.41	0.02	−0.02	0.01	−0.13
Si	0.52	1.00	−0.96	0.57	0.61	0.48	−0.25	0.29	0.48	0.11	−0.01	0.09	0.11	−0.31
Sa + Gr	−0.75	−0.96	1.00	−0.73	−0.76	−0.64	0.36	−0.26	−0.60	−0.22	0.00	−0.06	−0.09	0.28
LL	0.85	0.57	−0.73	1.00	0.91	0.96	−0.33	0.14	0.74	0.23	0.05	0.10	−0.02	−0.21
PL	0.84	0.61	−0.76	0.91	1.00	0.77	−0.33	0.15	0.86	0.34	−0.05	0.06	0.09	−0.25
IP	0.78	0.48	−0.64	0.96	0.77	1.00	−0.30	0.12	0.59	0.13	0.11	0.11	−0.10	−0.16
IL	−0.48	−0.25	0.36	−0.33	−0.33	−0.30	1.00	−0.23	0.08	−0.42	−0.04	−0.18	0.03	−0.05
p′	0.12	0.29	−0.26	0.14	0.15	0.12	−0.23	1.00	−0.08	−0.29	−0.19	0.63	0.21	−0.35
e_o	0.65	0.48	−0.60	0.74	0.86	0.59	0.08	−0.08	1.00	0.16	−0.02	−0.19	0.06	−0.18
OCR	0.41	0.11	−0.22	0.23	0.34	0.13	−0.42	−0.29	0.16	1.00	−0.03	−0.11	0.00	0.07
γ	0.02	−0.01	0.00	0.05	−0.05	0.11	−0.04	−0.19	−0.02	−0.03	1.00	−0.42	−0.85	0.83
G	−0.02	0.09	−0.06	0.10	0.06	0.11	−0.18	0.63	−0.19	−0.11	−0.42	1.00	0.48	−0.53
G/G_{max}	0.01	0.11	−0.09	−0.02	0.09	−0.10	0.03	0.21	0.06	0.00	−0.85	0.48	1.00	−0.88
D	−0.13	−0.31	0.28	−0.21	−0.25	−0.16	−0.05	−0.35	−0.18	0.07	0.83	−0.53	−0.88	1.00

Thanks to the correlation matrix, we can determine the degree of correlation of individual features with the damping ratio (Table 5).

Table 5. Degree of correlation (Pearson's correlation coefficient) of individual features with the damping ratio.

D	**γ**	**Sa + Gr**	**OCR**	**IL**	**Cl**	**IP**	**e0**	**LL**	**PL**	**Si**	**p′**	**G**	**G/Gmax**
1.00	0.83	0.28	0.07	−0.06	−0.13	−0.16	−0.18	−0.21	−0.25	−0.31	−0.35	−0.53	−0.88

The values of the correlation coefficient range from −1 to 1. Values close to 1 indicate a strong positive correlation, values close to −1 say that there is a strong negative correlation, while values close to 0 indicate no linear correlation. However, it should be remembered that the Pearson correlation coefficient measures only a linear relationship. It may completely ignore the nonlinear relationship. The highest positive correlation with the damping factor is shown by the γ, while the highest negative correlation with the G/G_{max} is shown by the G and p′.

Two variables may also be related by a nonlinear relationship or may have a non-Gaussian distribution. In these cases, Spearman's correlation coefficient can be useful to determine the strength and direction between the two data samples. Spearman's rank correlation can also be used if there is a linear relationship between the variables, but will have slightly less power [35]. Table 6 presents the result of the Spearman correlation test.

Table 6. Degree of correlation (Spearman's correlation coefficient) of individual features with the damping ratio.

D	**γ**	**Sa + Gr**	**OCR**	**IL**	**e_o**	**IP**	**Cl**	**LL**	**p′**	**Si**	**PL**	**G**	**G/G_{max}**
1.00	0.62	0.39	0.11	−0.04	−0.29	−0.30	−0.32	−0.40	−0.40	−0.48	−0.48	−0.65	−0.71

From Spearman's correlation, we see a moderately strong negative correlation of the damping ratio (D) with PL, Si, p′, LL, Cl, and IP. We can observe that the other features are also characterized by a higher degree of correlation than in the Pearson's correlation, which may indicate that the dependence of these features with D is nonlinear. On the other hand, a decrease of degree of correlation of values, such as the G/G_{max} and γ, can be noticed.

Another way to check the correlation between the attributes is a scatter matrix. The scatter matrix of all numerical attributes would be unreadable (and unjustified). Figure 1 shows a scatter matrix for some of the most promising attributes. Both the color and size of the markers relate to the value of the damping ratio (D). The plot shows a matrix of scatter plots where each variable is plotted against all other variables, resulting in a grid of scatter plots. The diagonal of the plot shows the kernel density estimation (KDE) plot for each variable, which provides a visual representation of the distribution of each variable.

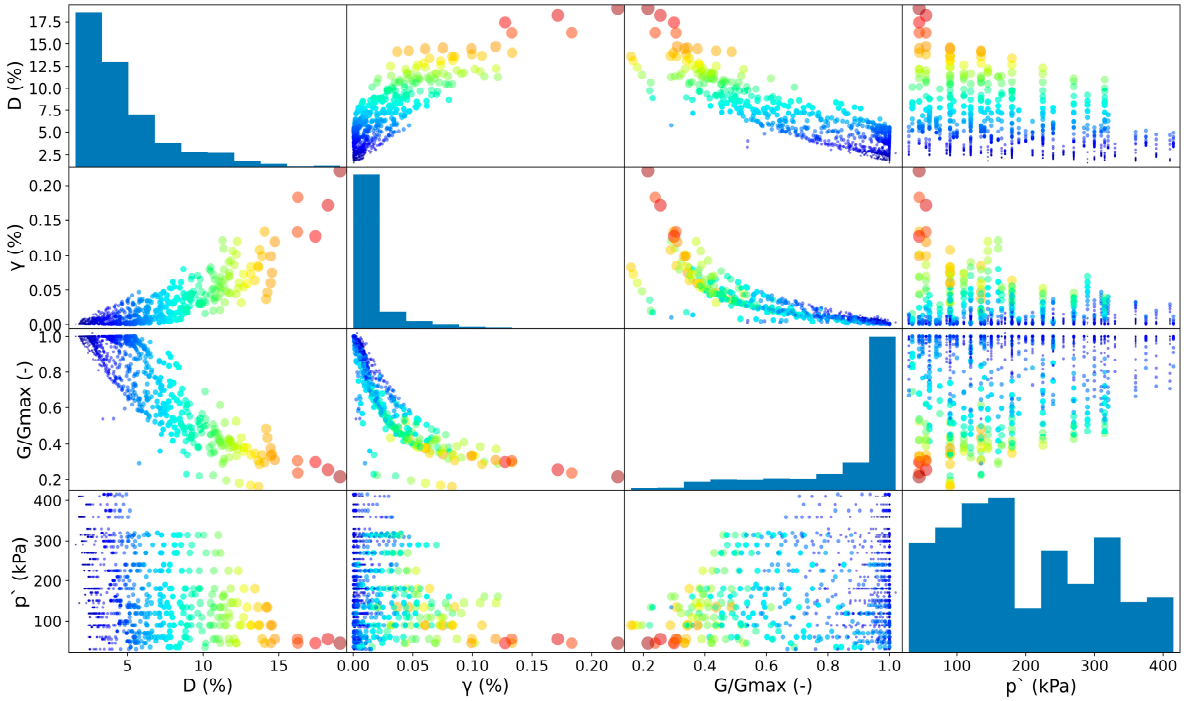


Figure 1. Scatter matrix of the damping ratio (D), shear strain (γ), normalized shear modulus (G/G_{max}), and mean effective stress (p′).

The above charts show that the correlation between the shear strain (γ), normalized shear modulus (G/G_{max}), mean effective stress (p′), and the damping ratio (D) is strong and the individual points are quite close to each other.

3.3.2. Detect Outliers and Anomalies

Some data batches include outliers. An outlier is a data point that differs significantly from other observations. Some outliers may be caused by measuring, recording, or copying errors, or by errors in entering the data into the computer. When such errors occur, it is important to detect and correct them, if possible. If not, excluding the erroneous values from further analysis is another possibility. Not all outliers are erroneous. Some may merely reflect unusual circumstances or outcomes; so, having these outliers called to our attention can help to uncover valuable information [36].

Two methods of the graphical determination of outliers are presented below. The first option is to create histograms of individual features. Histograms, in addition to visualizing outliers, allow researchers to gain insight into data, including the distribution, central tendency, spread, and modality. Figure 2 shows the histograms of the G distribution with outliers (on the left) and without outliers (on the right).

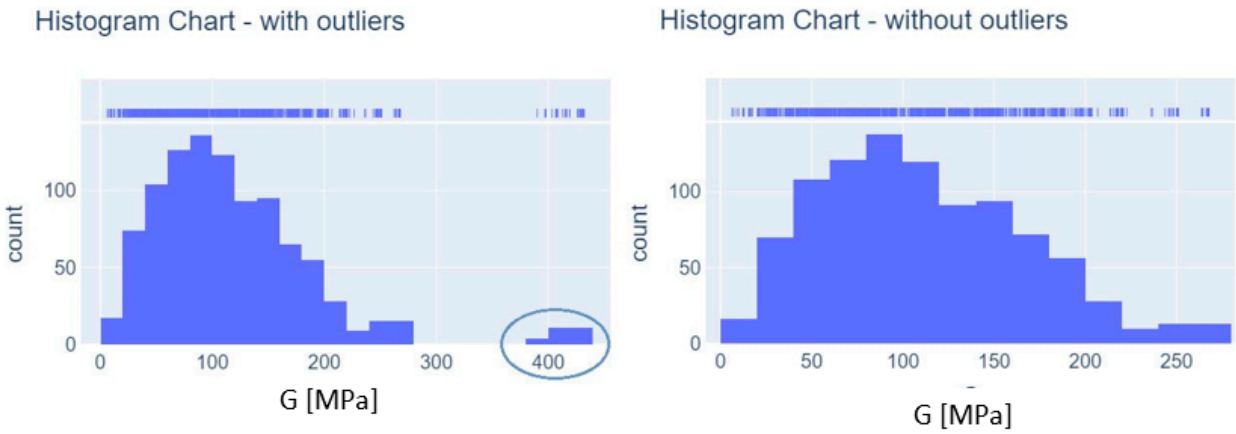


Figure 2. Histogram charts of G with and without outliers.

Another way to determine the outliers is to create boxplots. Boxplots are interesting for representing information about the symmetry, central tendency, skew, and outliers. It is a good EDA technique since it relies on robust statistics such as the median and IQR (interquartile range). Figure 3 presents an annotated boxplot which explains how it is constructed. The central rectangle is limited by Q1 and Q3. The middle line represents the median of the data. The whiskers are drawn in each direction, to the most extreme point that is less than 1.5 IQR beyond the corresponding hinge. Values beyond 1.5 IQR are considered outliers [34]. Figure 4 shows the boxplots of the G feature with outliers (on the left) and without outliers (on the right).

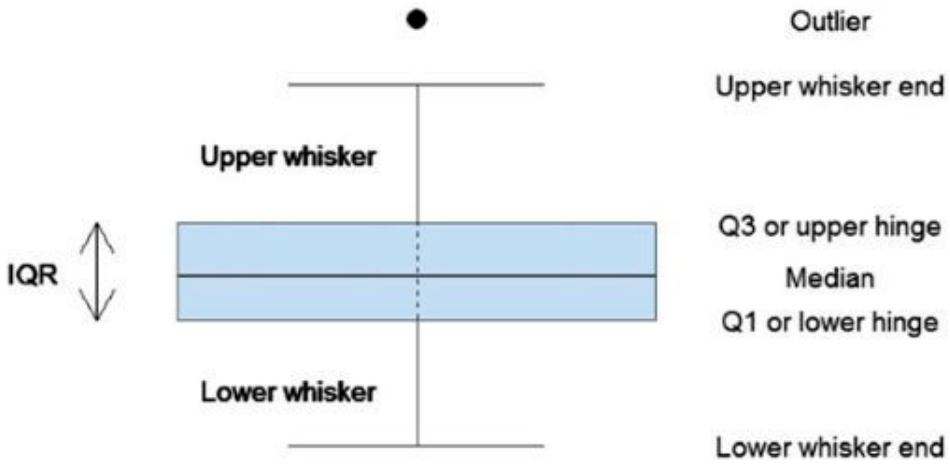


Figure 3. Example of boxplot with annotations.

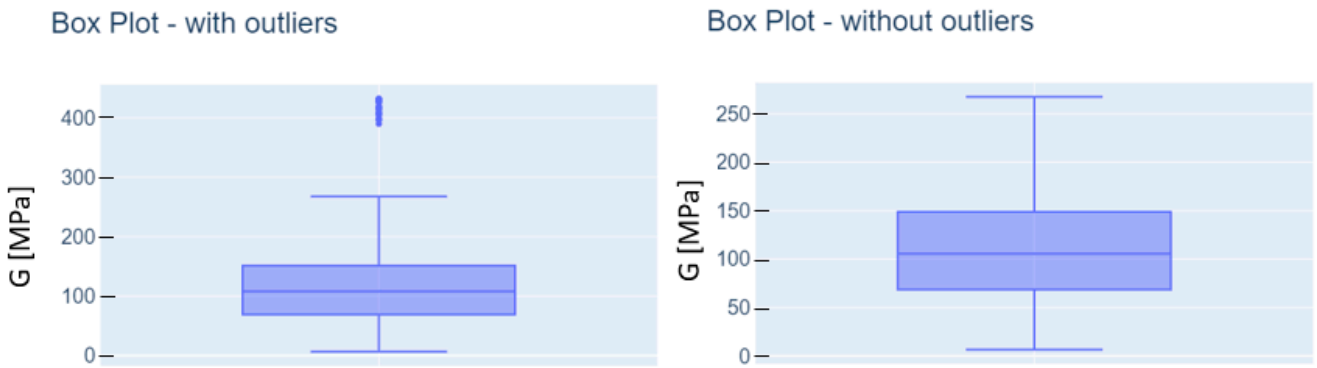


Figure 4. Boxplots of G with and without outliers.

Boxplots are an exploratory data analysis technique, and a researcher should consider designating a boxplot as an outlier as a suggestion that the points might be mistakes or otherwise unusual. Outliers should not be eliminated in the first instance, since depending on their context may provide a lot of information.

3.4. Data Preparation

Based on basic information about the dataset and exploratory data analysis (EDA), we can draw some conclusions which can determine the basic steps for data preparation:

- All data are numeric;
- There are no empty values—no data cleansing is needed;
- Data are of different orders of magnitude—feature scaling may be required;
- γ and G/G_{max} show high linear correlation with the damping ratio—features may be useful in the ANN model;
- LL, p′, Si, PL, Cl, and IP show moderately strong nonlinear correlation with the damping ratio—features may be useful in the ANN model;
- Features have outliers—removing outliers should be considered.

Each dataset is different; therefore, there is no one pattern of proceeding in exploratory data analysis or data preparation. When looking at the data, the researcher has to decide what steps should be taken to obtain the best possible machine learning (ML) model. Both exploratory data analysis and data preparation will probably have to be repeated several times during ANN model improvement.

3.4.1. Feature Scaling

Scaling features is one of the most important data transformations. Since the range of values of data may vary widely, it becomes a necessary step in data preprocessing while using machine learning algorithms. Most machine learning algorithms are poor at dealing with numerical attributes that fall within different ranges of the scale. This also applies to the analyzed dataset, where, e.g., the p′ feature has values in the range from 30 to 415 and where the values of the G/G_{max} feature range from 0.16 to 1.02.

One of the two most common types of scaling are min–max scaling and standardization. Min–max scaling is commonly known as normalization. Normalization transforms the data in such way that the features are within a specific range, e.g., (0, 1). To do this, it subtracts the minimum value from the given value and divides the result by the difference between the maximum and minimum values. Standardization (also called z-score normalization) transforms data in such way that the resulting distribution has a mean of 0 and a standard deviation of 1. The standardization mechanism is as follows: we subtract the mean value from a given value and then divide it by the standard deviation, thanks to which the resulting distribution has a unit variance. Standardization does not limit the scaled values to a certain range, as is the case, for example, during normalization. It is also much less sensitive to outliers.

The type of value scaling which will be the most appropriate in the analyzed task can be checked by substituting the normalized values, and then the standardized values, into the model. Target values usually do not need scaling. An important issue, that scientists often do incorrectly, is what values are scaled. As with all transformations, the scaling functions should be adjusted only to the training data, not to the entire dataset. Only after fitting against the training data can these functions be used to transform test data (and new cases).

Figures 5 and 6 show the results already obtained from the ANN model. The figures show the MAE and RMSE values of the training and validation set (constituting 20% of the training set) for the data subjected to normalization and standardization.

For the normalized training data: MAE = 0.2809, RMSE = 0.5069, and R^2 = 0.961; for the standardized data: MAE = 0.2032, RMSE = 0.4352, and R^2 = 0.979. For the ANN model aimed at determining the damping ratio based on other soil parameters, the standardization works better, and this scaling method will be applied to data in the final model version.

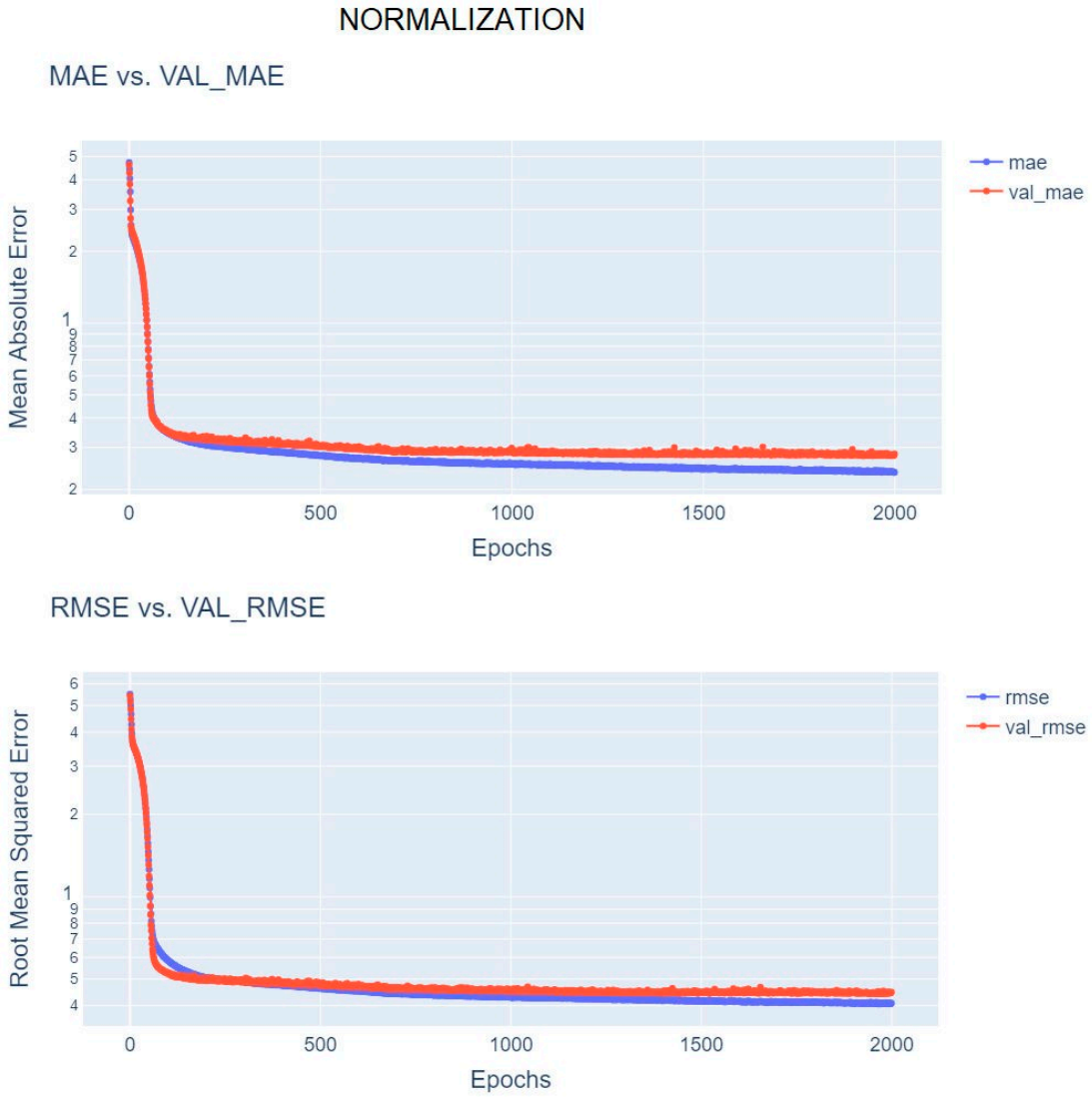


Figure 5. MAE and RMSE from the ANN model for normalized data.

3.4.2. Architecture of Artificial Neural Networks (ANNs)—Model Adjustment

The structure of the designed neural network depends on the complexity of the problem to be solved and the type of independent and dependent variables. In order to define the ANN architecture, the following should be determined: the number of inputs, the number of output neurons, the number of hidden layers, the number of neurons in each hidden layer, the type of activation function in the hidden and output layers, and the weight values of individual neurons. There are also parameters related to training the neural network, such as the batch size, training epochs, learning rate, and momentum. We can operate with all these parameters to improve the accuracy of our model.

This can be done in several ways. One of them is to manually select the hyperparameter values until obtaining the best combination of them. However, this is a very tedious task. Instead, by taking advantage of all the benefits of Python programming, we can outsource the search for the best parameters to the GridSearchCV class (scikit-learn). Grid search is a model hyperparameter optimization technique. For this purpose, it is enough to provide the desired parameters and their proposed values, and all combinations will be assessed using a cross-validation test.

Table 7 below shows the optimized parameters and their values that resulted in the lowest MAE and MSE errors.

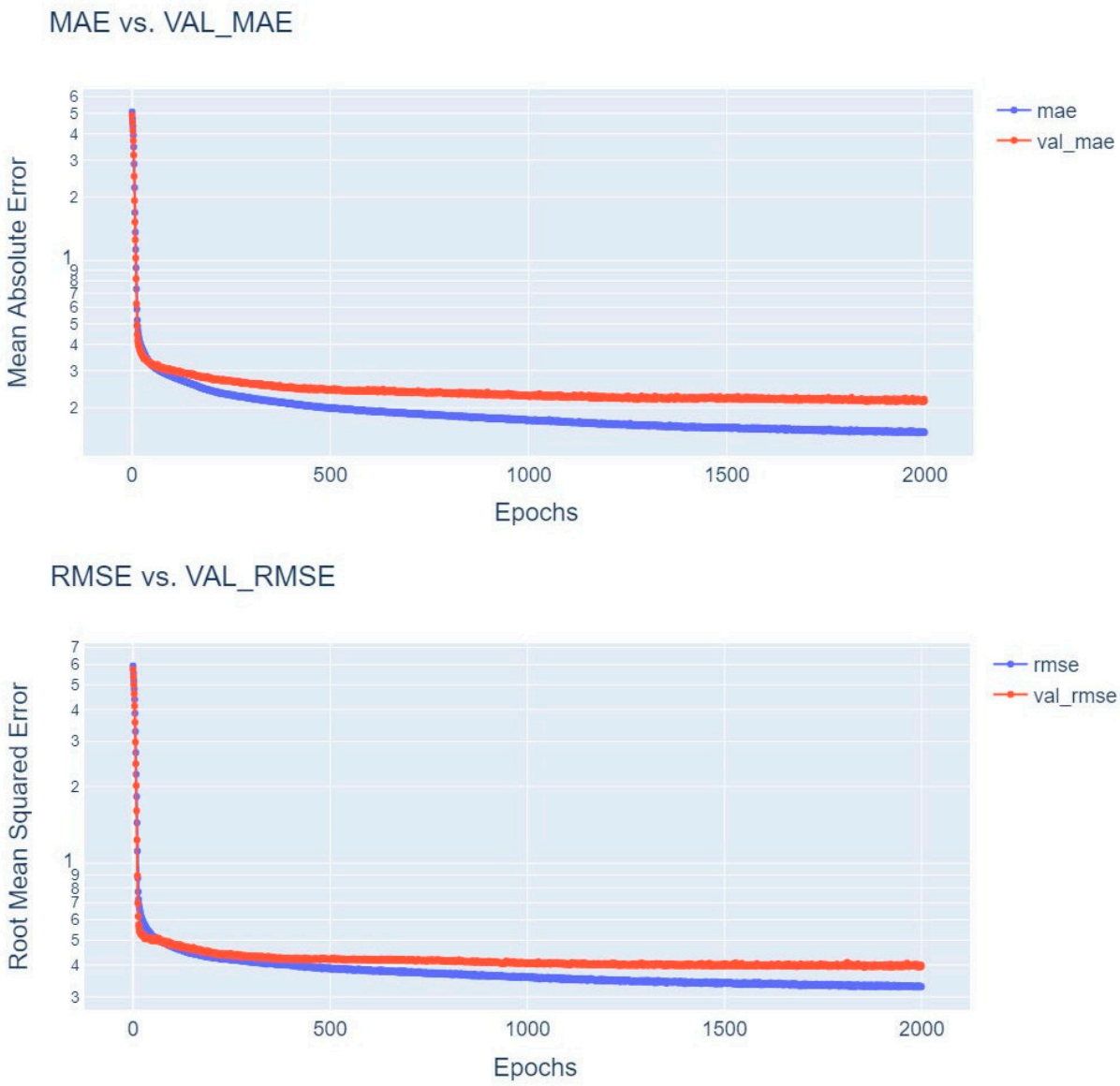


Figure 6. MAE and RMSE from the ANN model for standardized data.

Table 7. Optimized parameter values.

Parameter	Value	Range
Batch Size	10	(1:100)
Number of Epochs	1000	No significant difference when increasing the value
Training Optimization Algorithm	Adam	('SGD', 'RMSprop', 'Adagrad', 'Adadelta', 'Adam', 'Adamax', 'Nadam')
Network Weight Initialization	he_uniform	('uniform', 'lecun_uniform', 'normal', 'zero', 'glorot_normal', 'glorot_uniform', 'he_normal', 'he_uniform')
Neuron Activation Function	relu	('softmax', 'softplus', 'softsign', 'relu', 'tanh', 'sigmoid', 'hard_sigmoid', 'linear')
Dropout Regularization	0	(0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)
Number of Hidden Layers	2	(0, 1, 2, 3)
Number of Neurons in the 1 Hidden Layer	95	(0:200)
Number of Neurons in the 2 Hidden Layers	80	(0:200)
Learning Rate	0.0001	(0.00001, 0.0001, 0.001, 0.01, 0.1)

4. Results

4.1. ANN Model Results

Using the data from laboratory tests, an artificial neural network was created to predict the damping ratio.

The input parameters included:

- γ—the shear strain;
- G/G_{max}—the normalized shear modulus;
- LL—the liquid limit;
- p′—the mean effective stress;
- Si—the silt content;
- PL—the plasticity limit;
- Cl—the clay content.

The output data of the neural network were the damping ratio determined based on laboratory tests. To determine the value of the damping ratio (D) of clay soils in this study, deep feed-forward (DFF) network was used, which means that we applied more than one hidden layer in the model. The main problem with using only one hidden layer is overfitting. Therefore, by adding more hidden layers, we may achieve (not in all cases) reduced overfitting and improved generalization. The network used has the following architecture: 7-95-80-1, which means an artificial neural network with seven inputs and two hidden layers; the first with 95 neurons, the second with 80 neurons, and an output layer with 1 neuron. The relu function was used as the activation function in the first and second hidden layers. The Adam optimization algorithm was chosen to train the neural network. The input data were divided accordingly for the data used to train the network—80% and 20%—for network testing. From the training set, 20% was used to create a validation set which was used to tune the parameters of a classifier.

The correlation between the damping ratio predicted by the ANN and their values measured in the laboratory is shown in Figure 7. The predicted values of the damping ratio of clay soils are characterized by high convergence (R^2 = 0.976) with their values measured in the laboratory, which justifies the use of the ANN to predict their damping ratios.

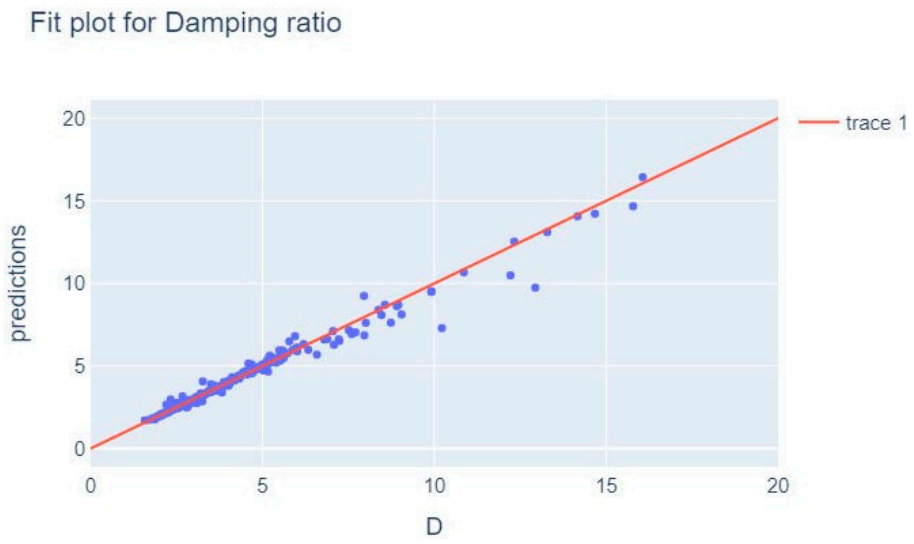


Figure 7. Correlation of the damping ratios of the clay soils predicted by the artificial neural network (test cases) with the values measured in the laboratory.

Figure 8 illustrates the histogram of errors made by a neural network when predicting the damping ratio. The error histogram is the histogram of errors between the target values and predicted values after training a feedforward neural network. These error values indicate how the predicted values differ from the target ones. The mean absolute error (MAE) for the test set was 0.1854, while the mean square error (MSE) was 0.1704.

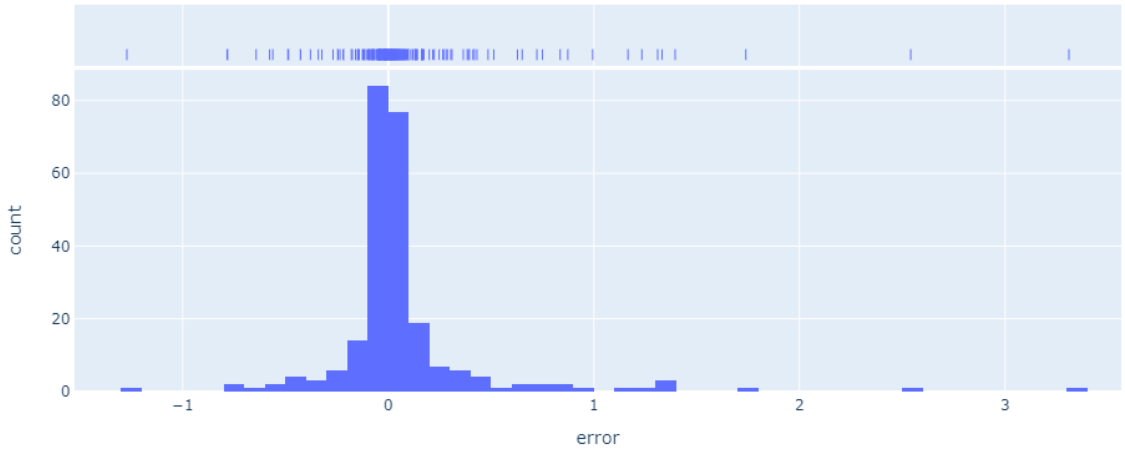


Figure 8. Error histogram.

4.2. Comparison of ANN Model with Other Empirical Formulas Available in the Literature

The obtained results were compared with the results obtained from earlier tests for cohesive soils. The authors chose five formulae for the damping ratio (D) in the wide shear strain range, summarized as follows:

Ishibashi and Zhang [36]:

$$D = \left(\frac{0.333\left(1 + e^{-0.0145I_P^{1.3}}\right)}{2}\right) \cdot \left[0.586\left(\frac{G}{G_{max}}\right)^2 - 1.547\left(\frac{G}{G_{max}}\right) + 1\right]; \quad (2)$$

Park and Stewart [37]:

$$D = 17.83\left[0.56\left(\frac{G}{G_{max}}\right)^2 - 1.39\left(\frac{G}{G_{max}}\right) + 1\right]; \quad (3)$$

Michaelides et al. [38]:

$$D = 2 + [18 - 0.08(I_P - 15)] \cdot \left[1 - \left(\frac{G}{G_{max}}\right)\right]; \quad (4)$$

Zhang et al. [39]:

$$D = 10.6\left(\frac{G}{G_{max}}\right)^2 - 31.6\left(\frac{G}{G_{max}}\right) + 21 + \left[(0.008I_P + 0.82)\left(\frac{\sigma'_0}{P_a}\right)^{-0.5k}\right]; \quad (5)$$

Soból et al. [29]:

$$D = \left[a\left(\frac{G}{G_{max}}\right)^2 - b\left(\frac{G}{G_{max}}\right) + c\right] + dI_P + e\left(\frac{p'}{P_a}\right)^f; \quad (6)$$

where a, b, c, d, e, and f are constants, different for both analyzed soil groups, and Pa is the atmospheric pressure equal to 100 kPa. For low- and medium-cohesive soils with an IP < 20%: a = 14.8, b = 34.3, c = 26, d = −0.31, e = 1.36, and f = −0.32; for very cohesive soil with an IP > 20%: a = 6.32, b = 20.36, c = 14.43, d = 0.062, e = 0.75, and f = −1.49.

Each considered model in the literature was suitable only for the tested soils. The equation presented by Ishibashi and Zhang [36] was calibrated for cohesive soil with an IP below 50% and an effective stress below 1000 kPa. Park and Stewart [37] presented an equation for the average damping ratio curve based on nine publications and considered different types of clay, e.g., offshore clayey silt, Edger plastic kaolin, San Old Bay clay, Mexico City clay, etc. Michaelides et al. [38] created an empirical model for cohesive soil with an IP from 0% to 100%. The formula shown by Zhang et al. [39] was appropriated for Tertiary, Quaternary, and older soils with an IP from 0% to 132%, collected from a depth

from 0 to 326 m. The equation proposed by Soból et al. [29] was calibrated on the same soil samples as the ANN model. It should be noted, however, that the formula proposed there was different for soils with an IP < 20% and different for those with an IP > 20%. The use of the ANN model allowed to avoid this division, further increasing the accuracy of the damping ratio predictions. Table 8 presents the comparison between the mean absolute error and determination coefficient.

Table 8. The comparison of mean absolute error (MAE) and determination coefficient (R^2) for analyzed models.

	Ishibashi and Zhang [36], Equation (2)	Park and Stewart [37], Equation (3)	Michaelides et al. [38], Equation (4)	Zhang et al. [39], Equation (5)	Soból et al. [29] Equation (6)	Authors ANN Model
MAE	3.63	1.02	1.06	1.77	0.45	0.19
R^2	0.21	0.79	0.78	0.77	0.96	0.98

The conducted comparisons indicate that the best model for cohesive soils taken from the analyzed site is the ANN model, which was proposed by the authors of this article.

The ANN model tested in the presented research can be useful to other researchers in several ways compared to traditional equations, such as Equation (6). Firstly, ANN models have the ability to learn from data and improve their predictive performance over time, which can result in more accurate and reliable predictions. Additionally, ANN models are more flexible than traditional equations and can capture more complex relationships between input and output variables. This can be especially useful in cases where the underlying relationships between variables are not well understood or are nonlinear. Finally, ANN models can be trained on large datasets, allowing for the inclusion of a wide range of input variables, which can lead to more robust predictions. Therefore, ANN models can provide researchers with a powerful tool for predicting the behavior of cohesive soils and can potentially lead to new insights and discoveries in this field.

5. Conclusions

The article presents the procedure for developing an artificial neural network (ANN) model. The potential of using an ANN to increase the confidence level of the prediction of the damping ratio (D) of normally and lightly overconsolidated Quaternary cohesive soil was described. Careful selection of the input variables describing the problem at hand is necessary in order to balance the complexity of the issue at hand with the amount of information available. Minor variables were omitted. The mean absolute error (MAE) for the performed tests was 0.1854, while the mean square error (MSE) was 0.1704. The predicted values of the damping ratio of clay soils are characterized by high convergence (coefficient of determination, $R^2 = 0.976$).

Comparing the ANN model with other published empirical formulas shows an improvement in the prediction accuracy. It is not possible to reliably evaluate the damping ratio for the tested cohesive soil with the use of commonly known empirical equations. A comparison of the measured and calculated value with equations from the literature gave a very large spread of the damping ratio and large error values. The conducted analysis shows that it can be concluded that an empirical model should be applied to the soil, with a similar genesis and mineral composition as the soil for which they were created. Considering the problem, it was shown that artificial neural networks are more flexible compared to formulas and relationships with a predetermined structure. Additionally, another advantage of artificial neural networks is the possibility to further model generalization by inserting newer records and the automatic adaptation of the architecture.

Artificial neural networks are well suited to modeling the complex behavior of most geotechnical engineering materials, which, by their very nature, exhibit extreme variability. The use of artificial neural networks to predict the damping ratio of clay soils is an effective, highly efficient, and easy-to-use method.

Author Contributions: Conceptualization, M.L.-S. and K.Z.; methodology, K.Z. and E.S.; laboratory tests, K.Z.; data analysis, K.Z. and E.S.; writing—original draft preparation, K.Z., A.M. and G.W.; supervision, A.M. and G.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that were used to support the findings of this study are included within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, Y.; Rosenbaum, M.S. The artificial neural network as a tool for assessing geotechnical properties. *Geotech. Geol. Eng.* **2002**, *20*, 149–168. [CrossRef]
2. Adeli, H. Neural networks in civil engineering: 1989–2000. *Comput.-Aided Civ. Infrastruct. Eng.* **2001**, *16*, 126–142. [CrossRef]
3. Basheer, I.A. Selection of methodology for neural network modeling of constitutive hysteresis behavior of soils. *Comput.-Aided Civ. Infrastruct. Eng.* **2000**, *15*, 445–463. [CrossRef]
4. Basheer, I.A. Stress-strain behavior of geomaterials in loading reversal simulated by time-delay neural networks. *J. Mater. Civ. Eng.* **2002**, *14*, 270–273. [CrossRef]
5. Basheer, I.A.; Najjar, Y.M. Modeling cyclic constitutive behavior by neural networks: Theoretical and real data. In Proceedings of the 12th Engineering Mechanics Conference, La Jolla, CA, USA, 17–20 May 1998; pp. 952–955.
6. Ellis, G.W.; Yao, C.; Zhao, R. Neural network modeling of the mechanical behavior of sand. In *Engineering Mechanics*; ASCE: Reston, VA, USA, 1992; pp. 421–424.
7. Ellis, G.W.; Yao, C.; Zhao, R.; Penumadu, D. Stress-strain modeling of sands using artificial neural networks. *J. Geotech. Eng.* **1995**, *121*, 429–435. [CrossRef]
8. Ghaboussi, J.; Sidarta, D.E. New nested adaptive neural networks (NANN) for constitutive modeling. *Comput. Geotech.* **1998**, *22*, 29–52. [CrossRef]
9. Gorai, A.K.; Raval, S.; Patel, A.K.; Chatterjee, S.; Gautam, T. Design and development of a machine vision system using artificial neural network-based algorithm for automated coal characterization. *Int. J. Coal Sci. Technol.* **2021**, *8*, 737–755. [CrossRef]
10. Habibagahi, G.; Bamdad, A. A neural network framework for mechanical behavior of unsaturated soils. *Can. Geotech. J.* **2003**, *40*, 684–693. [CrossRef]
11. Haj-Ali, R.; Pecknold, D.A.G.J.; Voyiadjis, G.Z. Simulated micromechanical models using artificial neural networks. *J. Eng. Mech.* **2001**, *127*, 730–738. [CrossRef]
12. Hashash, Y.M.A.; Jung, S.; Ghaboussi, J. Numerical implementation of a neural network based material model in finite element analysis. *Int. J. Numer. Methods Eng.* **2004**, *59*, 989–1005. [CrossRef]
13. Huang, F.; Xiong, H.; Chen, S.; Lv, Z.; Huang, J.; Chang, Z.; Catani, F. Slope stability prediction based on a long short-term memory neural network: Comparisons with convolutional neural networks, support vector machines and random forest models. *Int. J. Coal Sci. Technol.* **2023**, *10*, 18. [CrossRef]
14. Lawal, A.I.; Aladejare, A.E.; Onifade, M.; Bada, S.; Idris, M.A. Predictions of elemental composition of coal and biomass from their proximate analyses using ANFIS, ANN and MLR. *Int. J. Coal Sci. Technol.* **2021**, *8*, 124–140. [CrossRef]
15. Najjar, Y.M.; Ali, H.E. Simulating the stress-strain behavior of Nevada sand by ANN. In Proceedings of the 5th U.S. National Congress on Computational Mechanics (USACM), Colorado, CA, USA, 4–6 August 1999.
16. Najjar, Y.M.; Huang, C. Simulating the stress-strain behavior of Georgia kaolin via recurrent neuronet approach. *Comput. Geotech.* **2007**, *34*, 346–362. [CrossRef]
17. Penumadu, D.; Chameau, J.L. Geo-material modeling using artificial neural networks. In *Artificial Neural Networks for Civil Engineers: Fundamentals and Applications*; ASCE: New York, NY, USA, 1997; pp. 160–184.
18. Penumadu, D.; Zhao, R. Triaxial compression behavior of sand and gravel using artificial neural networks (ANN). *Comput. Geotech.* **1999**, *24*, 207–230. [CrossRef]
19. Romo, M.; Garcia, S.R.; Mendoza, M.J.; Taboada-Urtuzuastegui, V. Recurrent and Constructive-Algorithm Networks for Sand Behavior Modelling. *Int. J. Geomech.* **2001**, *1*, 371–387. [CrossRef]
20. Shahin, M.A.; Indraratna, B. Modeling the mechanical behavior of railway ballast using artificial neural networks. *Can. Geotech. J.* **2006**, *43*, 1144–1152. [CrossRef]
21. Shahin, M.; Jaksa, M.; Maier, H. State of the Art of Artificial Neural Networks in Geotechnical Engineering. *Electron. J. Geotech. Eng.* **2008**, *13*, 1–26.
22. Tutumluer, E.; Seyhan, U. Neural network modeling of anisotropic aggregate behavior from repeated load triaxial tests. *Transp. Res. Rec. Natl. Res. Counc.* **1998**, *1615*, 86–93. [CrossRef]

23. Wrzesiński, G.; Sulewska, M.J.; Lechowicz, Z. Evaluation of the Change in Undrained Shear Strength in Cohesive Soils due to Principal Stress Rotation Using an Artificial Neural Network. *Appl. Sci.* **2018**, *8*, 781. [CrossRef]
24. Wrzesiński, G.; Markiewicz, A. Prediction of Permeability Coefficient k in Sandy Soils Using ANN. *Sustainability* **2022**, *14*, 6736. [CrossRef]
25. Zhou, J.; Lin, H.; Jin, H.; Li, S.; Yan, Z.; Huang, S. Cooperative prediction method of gas emission from mining face based on feature selection and machine learning. *Int. J. Coal Sci. Technol.* **2022**, *9*, 51. [CrossRef]
26. Zhu, J.H.; Zaman, M.M.; Anderson, S.A. Modeling of soil behavior with a recurrent neural network. *Can. Geotech. J.* **1998**, *35*, 858–872. [CrossRef]
27. Zhu, J.H.; Zaman, M.M.; Anderson, S.A. Modeling of shearing behaviour of a residual soil with recurrent neural network. *Int. J. Numer. Anal. Methods Geomech.* **1998**, *22*, 671–687. [CrossRef]
28. Zhu, J.H.; Zaman, M.M.; Trafalis, T.B. Prediction of shear stress-strain behavior of soil with recurrent neural network. *Intell. Eng. Syst. Artif. Neural Netw.* **1996**, *6*, 809–814.
29. Soból, E.; Głuchowski, A.; Szymański, A.; Sas, W. The New Empirical Equation Describing Damping Phenomenon in Dynamically Loaded Subgrade Cohesive Soils. *Appl. Sci.* **2019**, *9*, 4518. [CrossRef]
30. Drozd, W.; Kozioł, P.; Zima, K. *Cyfryzacja W Budownictwie I Architekturze*; PWN: Warsaw, Poland, 2019.
31. Geron, A. *Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow*, 2nd ed.; Helion: Everett, WA, USA, 2019.
32. *D4015-07*; Standard Test Methods for Modulus and Damping of Soils by Resonant-Column Method. ASTM International: West Conshohocken, PA, USA, 2007.
33. Komorowski, M.; Marshall, D.C.; Crutain, Y. Exploratory Data Analysis. In *Secondary Analysis of Electronic Health Records*; Springer: Cham, Switzerland, 2016.
34. Brownlee, J. Machine Learning Mastery. Available online: https://machinelearningmastery.com/how-to-use-correlation-to-understand-the-relationship-between-variables (accessed on 8 September 2022).
35. Velleman, P.F.; Hoaglin, D.C. *Applications, Basics, and Computing of Exploratory Data Analysis*; Duxbury Press: London, UK, 2004.
36. Ishibashi, I.; Zhang, X. Unified dynamic shear moduli and damping ratios of sand and clay. *Soils Found* **1993**, *33*, 182–191. [CrossRef]
37. Park, D.; Stewart, H. Suggestion of Empirical Equations for Damping Ratio of Plastic and Non-Plastic Soils Based on the Previous Studies. In Proceedings of the International Conferences on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, San Diego, CA, USA, 29 March 2001.
38. Michaelides, O.; Gazetas, G.; Bouckovalas, G.; Chrysikou, E. Approximate non-linear dynamic axial response of piles. *Geotechnique* **1998**, *48*, 33–53. [CrossRef]
39. Zhang, J.; Andrus, R.; Juang, C. Normalized shear modulus and material damping ratio relationships. *J. Geotech. Geoenviron. Eng.* **2005**, *131*, 453–464. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.