CHEMICAL ENGINEERING TRANSACTIONS

VOL. 107, 2023

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: Petar S. Varbanov, Bohong Wang, Petro Kapustenko

ISBN 979-12-81206-07-6; **ISSN** 2283-9216

DOI: 10.3303/CET23107091

The Significance of Data Integration in Geotechnical Engineering: Mitigating Risks and Enhancing Damage Assessment of Expansive Soils

Richard P. Ray, Ammar N. Alnmr*

Széchenyi István University, Department of Structural and Geotechnical Engineering, Egyetem tér 1, Győr, Hungary
alnmr.ammar@hallgato.sze.hu

Geotechnical Engineering is a data-driven specialty that models the behavior of soil and rock as engineering materials. Due to their nature, however, much effort is directed toward defining the material properties and extent of various soil and rock on a site (buildings, bridges, dams) or region (transportation infrastructure). The great extent and variability of soil as an engineering material have led to the development of material, field exploration, and laboratory testing databases. Some have become national in scope, while others have expanded along specialist lines (e.g., earthquakes, landslides). This paper presents a typical application that investigates the impact of swelling clays and discusses its integration into the BENIP framework, with an emphasis on its role in advancing sustainability in geotechnical engineering. The study employs artificial neural networks as a modeling tool to predict crucial parameters such as dry density and in-situ confining stress, which directly influence volumetric changes in the soil. By minimizing these changes, potential damage associated with swelling clays, including ground movement, foundation deterioration, and infrastructure instability, can be mitigated. The results of the study exhibit promising outcomes, signifying the potential effectiveness of the proposed neural network models in promoting sustainability in geotechnical engineering.

1. Introduction

Geotechnical engineers continuously enhance field and lab methods to comprehend the interactions between the built environment and the soils that support it. These improvements include better equipment and sensors, more sophisticated analysis methods, and a greater appreciation for the holistic approach to building. A logical next step would be integrating those improvements into a more comprehensive platform such as BENIP. The Built Environment Information Platform (BENIP) creates a highly contextual information platform that connects city planning and development to traditional domains of architecture, civil, and transportation engineering via Digital Realities (virtual reality, augmented reality, simulations, and real/virtual twins) (Horváth et al., 2023). This paper examines geotechnical engineering's role in an integrated information platform shared among stakeholders, enabled by artificial intelligence and specialized modeling. It presents a geotechnical background, explores the significance of geotechnical databases, and discusses a geotechnical application utilizing artificial neural networks (ANNs) to address essential parameters such as dry density and applied stress. Several studies have demonstrated the effectiveness of ANNs in predicting soil-related phenomena. For instance, (Ashayeri et al., 2009)) estimated the amplitude and swelling pressure of unsaturated clay using ANNs. (Merouane, 2018) used ANNs to estimate the swelling pressure of expansive soils. Additionally, (Dutta et al., 2019) utilized ANNs to predict the free swell index of expansive soil. Notably, previous studies have overlooked the importance of addressing dry density and applied stress, underscoring the novelty and value of this research.
The scientific hypothesis of this research is to investigate the potential of artificial neural networks (ANNs) in predicting crucial geotechnical parameters within the BENIP framework, addressing swelling clays, and promoting sustainability. ANNs were chosen based on prior successful studies. The procedure involves analyzing geotechnical data using ANNs to develop correlations for effective decision-making, mitigating potential damage caused by swelling clays.

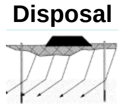
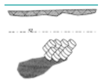
Paper Received: 03 April 2023; Revised: 06 May 2023; Accepted: 13 June 2023
Please cite this article as: Ray R.P., Alnmr A.N., 2023, The Significance of Data Integration in Geotechnical Engineering: Mitigating Risks and Enhancing Damage Assessment of Expansive Soils, Chemical Engineering Transactions, 107, 541-546 DOI:10.3303/CET23107091

2. Geotechnical background

The geotechnical engineering discipline has been data-driven since its early years (1920's) and continues in that direction. The reason for such an approach lies in the nature of soil and rock: their properties and distribution may be highly variable, their behavior changes due to the presence or absence of water, and their sheer volume within a site. Soil hydraulic conductivity may vary over 12 orders of magnitude; particle sizes may be m or nm; strength and stiffness (modulus) may differ by a factor of 10^6. Most of these materials were deposited or arranged through natural processes that take place over a period of millions of years (geologic) or seconds (earthquakes).

Defining material behavior is a crucial task for geotechnical engineers, and a large amount of effort is devoted to that task in every project. Existing data is collected through GIS and historical records, and then field investigations are planned based on the objectives of the project. Such investigations may require only a modest amount of exploration and require less than a week, while others are very extensive and require several years to complete. There are a wide variety of techniques and equipment that may be applied, but they are all focused on answering questions concerning the nature and extent of the soil and rock present on-site. The evolution of field investigations is summarized in Table 1 (Sara, 1994).

Table 1: Evolution of exploration effort and project risk

	High-rise; Dams 	**Offshore; NP Plants** 	**SW Land Disposal** 	**HazWaste Sites RI/PS** 	**RCRA SubD CERCLA Superfund**	**Integrated Engineered Systems**
Borings Monitor Wells	5-10 20-50	1-10 50-100	10-150 40-300	400-1600 200-3400	CPT Geophysics Scanning	Boring/CPT Satellite Geophysics
Catastrophe Risk (Deaths)	Moderate-High (10-5,000)	Moderate Ext. High (10-50,000)	No Immediate	No Immediate	No Immediate	Varies
Environmental Risk	Med-High	High/Ext High	Low Low	High High	High High	CO_2 Climate

The evolution in the table moves from left to right. It shows the increased dependence on data gathered from the field and laboratory to cope with the increased complexity and risk of projects. Projects have become more expensive, and the risks related to errors during construction and operation have grown to almost incalculable levels. Large environmental projects were especially difficult because they were numerous, and their extent and impact were often undefined at the start of exploration. In the 1990s, geotechnical projects were integrated with environmental regulation and a broader base of expertise from other disciplines. This forced geotechnical professionals to formalize procedures that had been more intuitive in the past. Conceptual models were defined, investigations were separated into phases, and the goals for each phase were more clearly described. While such formality was sometimes frustrating and time-consuming, it developed into a clear framework for planning, executing, and managing an investigation. It also defined more clearly the process of data gathering and processing and its use in analysis. This process is illustrated in Table 2, where data evolves from raw measurements into more and more refined stages from left to right.

Table 2: Generations of data from initial raw lab and field (left) to model predictions (right)

First Generation Data	**Second Generation Data**	**Third Generation Data**	**Fourth Generation Data**
Raw field Data Borings Test Pits Geophysical	Edited Data Sets Indicator Parameters	Point in Space (x, y, z) Time Trend Values	Geographical Engineering Data
Groundwater Depths	Cross-Sections	Statical Data Analysis	Prediction of Future Events
Location and Topographic Info	Structure Contour Maps	GIS Layering Constructs	Sensitivity Analysis
Lab Index and Classification Tests	Correlation Index to Performance	Evaluation of Boundary and Initial Conditions	Evaluation of Exposure/Risk

With each generation, the data becomes more focused on answering specific design questions. This has led to a more direct relationship between exploration, design, and analysis. Simply put (Sara, 1994): The most important data is that which leads to making decisions. Therefore, only collect data that is part of the decision-making process.

The data collection/analysis process will often cycle back on itself; more data gives rise to more questions that require additional investigation. The circular nature of investigate-design-construct was first practiced by Peck for tunnel design and construction and is known as the Observational Method (Self et al., 2012). It became popular due to the highly variable nature of soil and rock; the designer/builder simply could not know enough prior to the start of tunneling. The approach had to be holistic since design, construction, management, and finance had to sign off on all changes to the project scope as it progressed. If one of the parties had not agreed that modifications were the norm rather than the exception, any changes in conditions would be met with a serious slowdown or complete shutdown of the project.

3. Geotechnical databases

Database development in geotechnics began as early as the 1960s. It existed on paper (tabular or maps) and was narrowly focused on specific classes of problems such as earthquakes, landslides, construction materials, and laboratory testing. Some were shared between similar agencies (e.g., highway departments) within a region (cold regions research laboratories) or between professional societies. More recently, attempts to broaden these databases have focused on ways to better exchange often disparate data. The US Federal Highway Administration developed middleware to encourage better data exchange (Lefchik et al., 2007). One such database is the National Geotechnical Properties Database maintained by the British Geological Survey (Lefchik et al., 2007). The database contains records for 7,370 projects (site investigation reports). Linked to the reports are 180,000 holes. There is a total of 3,600,000 in-situ field records, with 880,000 samples linked to the holes and a total of 5,200,000 laboratory test data records. Some of the database structure is shown in Figure 1. Records are grouped by exploration or test method and by material property. Figure 2 shows various examples of data presentation and results of analysis at the third-generation level (see Table 2). The data analysis is performed within the database software.

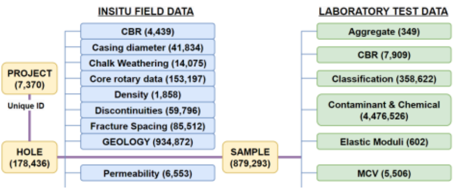


Figure 1: Structure of BGS National Geotechnical Database

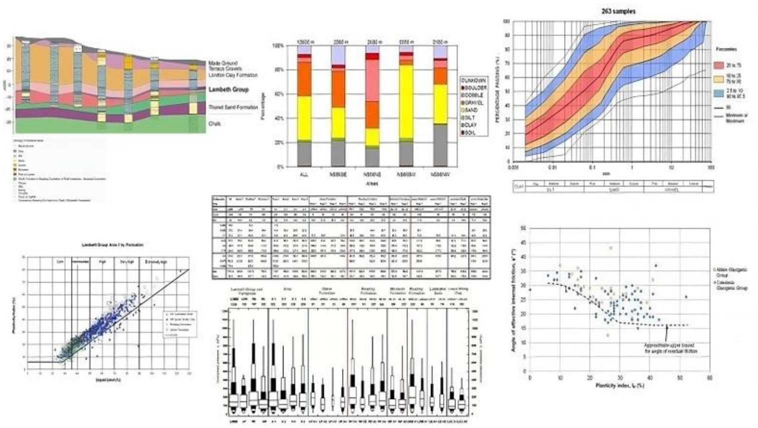


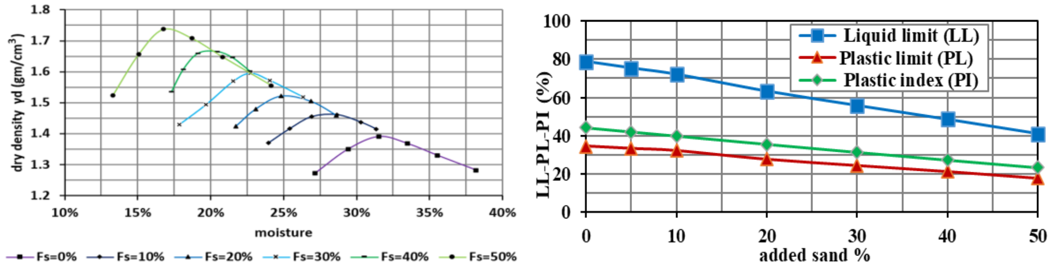
Figure 2: Extracting processed data from NGD

Being a public agency, BGS encourages the submission of additional data by providing an online portal. The database resides within a larger network of databases, which are maintained by BGS. A common issue faced by the developers has been reconciling overlapping data. Since several of the databases are designed to run as stand-alone, they may contain the same data. When combined, the overlapping data must be properly accounted for.

4. Example Geotechnical Application

The example discussed here is a geotechnical design problem that is experienced throughout the world. Certain types of soils containing specific clay minerals exhibit very large expansion and contraction when subjected to wetting and drying. These swelling clays are present near the ground surface and may be treated in a variety of ways. However, the greatest challenge is to develop an economical solution that could be applied to residential homes and other light buildings. To address the problem, the engineer must perform a variety of tests that require time and expense. A better solution would be to correlate results from simpler, quicker tests to determine the risk of swelling in a particular soil found on-site. One approach that is gaining popularity is to use artificial neural networks (ANN) to develop these correlations and determine the level of risk the candidate soil presents. Simple remedies, such as mixing a percentage of sand with the swelling soil, may be evaluated as well (Najjar et al., 2019). The utilization of artificial neural network (ANN) models enables informed decision-making to address and mitigate problems effectively (Anibal et al., 2023).

The candidate swelling clay soil comes from Damsarkho city in the Latakia region of Syria. The Unified Soil Classification System (ASTM, 2017) classifies it as CH. The sand used to modify its swelling behavior is a fine marine sand common to the coastal area. Various soil mixtures were prepared by incorporating different proportions of sand (10 %, 20 %, 30 %, 40 %, and 50 %) into the expansive soils based on dry weight. Proctor experiments were conducted (Figure 3(a)) to determine the maximum dry density and optimum moisture for each mixing percentage. The results demonstrate a continuous increase in the maximum dry density, which is consistent with Gupta and Sharma's findings (Gupta and Sharma, 2014).



(a) (b)

Figure 3: (a) Proctor experiments for clay samples with different percentages of added sand (b) Beneficial effect of adding sand to swelling clay

The benefits of adding sand are shown in Figure 3(b), where the liquid limit is reduced from an original value of 80 % to 40 %. Liquid limit is an index indicator of swelling potential where values above 50 % are considered to have high swell potential. The clay mineralogy of soil can be associated with the position of the Atterberg limits plot on a plasticity chart (Casagrande, 1948). To this end, the liquid limit and plasticity index were plotted on a plasticity chart for each percentage of added sand, which included ranges for montmorillonite, illite, kaolinite, and chlorite (Holtz and Kovacs, 1981). The present study found that the samples with varying percentages of added sand consisted predominantly of illite based on the plasticity chart. The free swell was determined for clear clayey soil (i.e., without any sand content) using Prakash and Sridharan's (Prakash and Sridharan, 2004) method and was found to be 127 %.

The other two parameters (PL, PI) are indexes that reflect improved behaviour with lower values. Additional parameters that influence the degree of swelling are the compacted dry density of the clay/sand mixture and the amount of compaction energy. Optimizing these two main parameters, in addition to accounting for many other conditions, became a very complex operation where traditional approaches such as regression analysis were not adequate to predict behavior.

Artificial neural networks were used to help predict the dry density that minimized volumetric changes (swell/shrink), required compaction effort, and in-situ confining stress. Confining stress represents the load due to buildings constructed over the swelling soil. Additional conditions and constraints were also included, but they are beyond the scope of this paper. MATLAB was utilized as the primary tool for implementing the ANN model, with performance metrics including coefficient of correlation (R) and mean square error (MSE) used to evaluate the model's accuracy. We developed prediction models for dry density ANN(γ_{max}) and stress state for minimum volume change ANN(σ_{min}). Inputs came from the results of simple experiments and combinations of sand/clay mixtures. In this work, 648 sets of data were used for the dry density and for applied stress that achieves minimum volumetric change. For training, 70 % of the data was used, while the remaining 30 % was used for validation and testing. A schematic of the network is shown in Figure 4. Constraints of the models were set to minimums and maximums determined from prior laboratory measurements, as shown in Table 3.

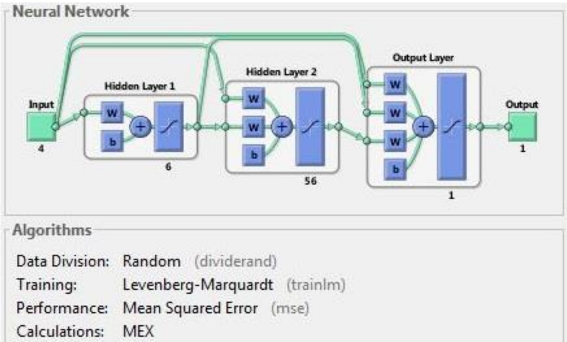


Figure 4: Schematic of neural network models

Table 3: Constraints used in ANN process

Input parameters	ANN(γ_{max})		ANN(σ_{min})	
	Min	**Max**	**Min**	**Max**
LL (%)	41	79	41	79
SR (%)	45	100	45	100
σ (kN/m²)	25	300		
γ_d **(kN/m³)**			12	17.8
γ_{bro}	13.95	17.4	13.95	17.4
Output parameters	**Min**	**Max**	**Min**	**Max**
γ_{min} **(kN/m³)**	12	17.8		
σ_{min} **(kN/m²)**			25	300

Training, validation, and testing phases showed excellent agreement between target values and outputs. The results are encouraging, although since the initial evaluations were somewhat contrived, a very high level of agreement would be expected. One would expect less agreement for data sets that are dispersed over a region or include a wider variety of clays. Additional analyses were performed on datasets from other authors, and similar agreement was found.

Integrating expansive soil modeling with detailed data on the built environment is critical for increasing sustainability in geotechnical engineering. The adoption of this integrated methodology enables better planning and risk assessment in relation to damage caused by swelling clays. It is crucial to note that the annual damage to structures caused by swelling clays in the United States is roughly \$ 10^{10}. Because of their propensity to undergo significant volume changes, expansive soils provide severe sustainability challenges, resulting in ground movement, foundation damage, and infrastructure instability. By implementing the methodology proposed in this paper, which focuses on predicting applied stress or dry density at specific site locations, the volumetric changes of expansive soils can be minimized. This, in turn, facilitates the design of appropriate foundations that apply the predicted stress, mitigating issues such as cracks, settlements, and structural instability in buildings constructed on expansive soil. By addressing these critical concerns, the integration of such modeling techniques and data enhances sustainability in geotechnical engineering.

5. BENIP and Geotechnical Engineering

Software platforms that combine databases, rigorous analysis, statistical evaluations, and sophisticated visualization are already common in the geotechnical community. Connecting those software platforms has been a much greater challenge. The de facto standard finite element software for geotechnical engineers is Plaxis which incorporates a material database, inverse laboratory simulations to extract model parameters, and methods to conveniently import and export data and results. Other software, such as Midas GTS, performs similar analyses and is part of a larger suite of structural engineering and construction software that is integrated (to say seamlessly integrated would be too generous to the software). Integrated site design software from Autodesk and Bentley covers a wide range of design methods and specialized solutions. However, a truly integrated platform has remained elusive.

Part of the challenge is hardware, part of it is software, and a third component is the very distinct and specialized nature of engineering. The first two are left for another discussion; the third is especially important to BENIP. Specialized language and tools have evolved in engineering because they increase the working efficiency of the engineer. This has been evolving in Geotechnical Engineering since its inception in the 1920s, and it has allowed for a tremendous increase in understanding and clearly communicating problems, ideas, and solutions. Integrating the specialty with other specialties will require compromise and new ideas. This situation is very similar to the growth of Geo-environmental Engineering in the 1990s as mentioned earlier. Disciplines that were outside the realm of geotechnics, such as aquatic chemistry, biology, and contaminant transport, became integral parts of site investigations and remediation design. Today, the newly developed methods and instruments have become fully integrated into an entire spectrum of field and laboratory investigations that go well beyond environmental work. The integration of disciplines is critical for promoting sustainability in the field of geotechnical engineering.

6. Conclusions

This paper has presented a Geotechnical Engineering perspective of BENIP ideas by presenting the ideas and methods we have been using to find solutions to difficult problems in civil engineering projects. The specialty has been data-driven since its beginning due to the highly variable nature of the materials used in the design. The increasing effort in the field and the laboratory to better define material behavior has led to projects with increasing complexity and reduced tolerance for error. The profession has already made significant progress in data collection, processing, and organization. More and more standards for data storage and retrieval, as well as testing and design methods, have allowed for a wider integration of data and workflow. However, challenges remain due to computational demand, specialized problem-solving, and impediments in communication. The application of artificial neural networks (ANN) in this study has yielded promising results. By utilizing ANN models, we have successfully predicted critical parameters such as dry density and applied stress, contributing to more accurate and informed decision-making processes. The utilization of ANN demonstrates its potential to enhance the efficiency and effectiveness of geotechnical engineering practices. The ongoing efforts to overcome computational challenges, enhance problem-solving capabilities, and improve communication will further propel the integration of data and foster interdisciplinary collaboration. With continued progress, geotechnical engineering will play a pivotal role in driving innovation, optimizing project outcomes, and ensuring the long-term sustainability of civil engineering projects.

References

Anibal A.L., Edwin T.S., Silvia G.M., Halyn A.V., Juan O.A., Dario P.B., 2023, Machine Learning Models for Natural Gas Consumer Decision Making: Case Study of a Colombian Company. Chemical Engineering Transactions, 100, 169–174, DOI: 10.3303/CET23100029.

Ashayeri I., Yasrebi S., 2009, Free-Swell and Swelling Pressure of Unsaturated Compacted Clays; Experiments and Neural Networks Modeling. Geotechnical and Geological Engineering, 27, 137–153.

ASTM International, ASTM D2487-17e1, 2017, Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System), DOI: 10.1520/D2487-17E01.

Casagrande A., 1948, Plasticity Chart for the Classification of Cohesive Soils. Transactions of the American Society of Civil Engineering, 113(1), 901–930

Dutta R.K., Singh A., Gnananandarao T., 2019, Prediction of Free Swell Index for the Expansive Soil Using Artificial Neural Networks. Journal of Soft Computing in Civil Engineering, 3, 47–62.

Gupta C., Ravi K.S., 2014, Influence of Marble Dust, Fly Ash and Beas Sand on Sub Grade Characteristics of Expansive Soil. Journal of Mechanical and Civil Engineering, 13, 13–18.

Holtz R D., Kovacs W D., 1981, An introduction to geotechnical engineering. Prentice-Hall, Englewood Cliffs, NJ., USA, ISBN: 0-13-484394-0.

Horváth B., Szép J., Borsos A., 2023, The Concept of BENIP-Built Environment Information Platform. Infocommunications Journal, A Publication of the Scientific Association for Infocommunications (HTE), no. SP: 29–34, DOI: 10.36244/ICJ.2023.SI-IODCR.5.

Lefchik T.E., Bobba C., Beach K., 2007, DIGGS: Setting the Standard for Geotechnical and Geoenvironmental Data Management, In 58th Highway Geology Symposium. Pocono Manor, Pennsylvania, United States.

Merouane F.Z., Mamoune S.M.A., 2018, Prediction of Swelling Parameters of Two Clayey Soils from Algeria Using Artificial Neural Networks. Mathematical Modelling in Civil Engineering, 14, 11–26.

Najjar Y.M., Imad A.B., Richard M., 2019, Neural Modeling of Kansas Soil Swelling. Journal of the Transportation Research Board, 1526(1), 14–19, DOI: 10.1177/0361198196152600103.

Prakash K., Sridharan A., 2004, Free Swell Ratio and Clay Mineralogy of Fine-Grained Soils. Geotechnical Testing Journal, 27(2), 220–25, DOI: 10.1520/GTJ10860.

Sara M.N., 1994. Standard handbook for solid and hazardous waste facility assessments. Lewis Publishers, Boca Raton, FL, United States, ISBN: 0873713184.

Self S., Entwisle D., Northmore K., 2012, The Structure and Operation of The BGS National Geotechnical Properties DATABASE. Version 2, British Geological Survey, Nottingham, United Kingdom.