*GeoVadis: The Future of Geotechnical Engineering –
Juneja, Joseph & Dasaka (Eds)*
 ISBN: 978-1-041-08547-8

Use of machine learning for granular geomaterials: A short critical review

H. Hunt, B. Indraratna, R. Malisetty and Y. Qi
Transport Research Centre, School of Civil and Environmental Engineering, University of Technology Sydney, Sydney, New South Wales, Australia

R. Kelly
Priority Research Centre for Geotechnical Science and Engineering, School of Engineering, the University of Newcastle, Australia and Snowy Mountains Engineering Corporation, SMEC, Australia

ABSTRACT: Knowing a soil's strength and deformation characteristics is crucial for load-bearing applications, and they are often determined through costly and onerous laboratory procedures. However, these characteristics can vary based on a multitude of factors, meaning new tests are needed each time the material properties or loading environment changes. Empirical correlations can be used to remedy this, though extrapolation inaccuracies may arise due to their simplistic modelling frameworks. This paper presents a short critical review of selected studies using machine learning (ML) techniques to predict the shear strength characteristics and resilient modulus of granular materials. Given the popularity of ML, these studies are scrutinized from a geotechnical perspective, rather than focusing on specific modelling intricacies. From this review, it is evident that there is a clear divide between studies that use ML as a tool to either fit a given set of data, or to enhance well-established geotechnical relationships and concepts.

1 INTRODUCTION

Artificial intelligence (AI) and machine learning (ML), a subset of AI, is one of, if not the most popular and expanding current research areas. This is attributed primarily to the significant improvement in computing efficiency and the ability ML techniques have in extrapolating complex nonlinear patterns and relationships within a given dataset (Indraratna et al. 2025; Liu et al. 2024). For example, a multitude of parameters (e.g., material type, fabric, particle shape, hardness, density, porosity, water content, breakage susceptibility, loading conditions), and their several interdependencies, can influence the behavior and properties of a granular material (Yang and Rosenbaum 2002; Xiao et al. 2014), with varying degrees of nonlinearity. These factors increase modelling complexities, making it difficult to apply conventional statistical and analytical constitutive methods without introducing numerous assumptions and/or simplifications (Wallace & Ng 2016; Shahin 2024).

In contrast, ML techniques are more adept in this regard as they are data-driven and no prior assumptions or simplifications are required (Carter et al. 2000; Shahin 2024). Examples of these techniques include the artificial neural network (ANN), adaptive neuro-fuzzy inference systems (ANFIS), random forest (RF), as well as the more conventional regression techniques such as linear and nonlinear multiple regression (MR). However, as they are predominately data-driven, this surge in interest warrants caution to be taken by both those developing these ML models, and of course those implementing them for practical application. Ultimately, ML is a practical tool that should be used to enhance or refine, rather than violate, the underlying geotechnical engineering concepts that are well-established. For these ML models to be considered viable from a geotechnical standpoint, it is crucial to assess how well these established concepts are incorporated based on their input variables predictions. The purpose of this review is not to

DOI: 10.1201/9781003645917-15

discuss in-depth the specific modelling techniques and intricacies used by previous ML studies in the literature, but rather scrutinize them from a purely geotechnical perspective (e.g., selection of input parameters). For brevity, two main areas are explored with respect to granular geomaterials: predicting their (i) shear strength characteristics, and (ii) resilient modulus, with comparisons between conventional regression and the more advanced ML techniques.

2 SHEAR STRENGTH CHARACTERISTICS

Reliably knowing the shear strength of a geomaterial is imperative for loading applications as shear failure is often the governing failure mechanism and design criterion (Hunt et al. 2024). For a granular matrix where cohesion is treated as negligible, the peak friction angle (ϕ'_{peak}) is often used as a representative shear strength parameter in place of the failure shear stress.

2.1 *Conventional regression*

Numerous equations have been proposed by researchers in an attempt to efficiently predict the shear strength of granular materials, with selected examples summarized in Table 1. Using a power function of the normal stress (σ_n), with regression coefficients as a function of uniaxial compressive strength (σ_c) and void ratio of rockfill material, Andjelkovic et al. (2018) returned a strong predictive performance on both the calibration and validation datasets. Although it was noted that the coefficient of uniformity (C_u) could not be included to its heterogeneity within the implemented database, sufficient variance is apparent. Leslie (1969) reported an increase in shear strength with a respective increase of C_u and, therefore, the successful inclusion of this parameter would likely enhance the accuracy and viability of the proposed model. Hunt et al. (2023) also modelled the shear strength of coal wash-rubber crumb mixtures as a power function, though the two regression coefficients were expressed empirically as functions of the rubber content. However, due to the small size of calibration data and model simplicity, the authors noted that the regression constants would require recalibration to be more suitable when considering other mixture types. This in essence nullifies the purpose of developing predictive models since determining these empirical constants requires laboratory testing in the first place. Similar findings were reported by Indraratna et al. (1998) for ballast, where the values of each regression constant vary based on the material type and gradation characteristics. Sharma et al. (2019) attempted to predict the friction angle of sand with varying median particle size, C_u, and relative density with moderate success, though considerable prediction errors were present within the validation dataset.

Table 1. Examples of conventional regression shear strength models within the literature.

Model	Material/Mixture	Reference
$\tau_f/\sigma_c = a(\sigma_n/\sigma_c)^b$	Ballast (latite basalt)	Indraratna et al. (1998)
$\tau_f = a(\sigma_n)^b$	Rockfill	Andjelkovic et al. (2018)
$\tau_f = a(\sigma_n)^b$	Coal wash + rubber	Hunt et al. (2023)
$\phi'_{peak} = \phi'_b + a \cdot \exp(b \cdot \sigma'_n)$	Ballast (latite basalt)	Indraratna et al. (1998)
$\phi = a(D_{50})^b (C_u)^c (R_d)^d$	Sand	Sharma et al. (2019)
$\phi'_{peak} = a + \Sigma b_i(x_i) + \Sigma b_{ii}(x_i)^2 + \ldots$ $\ldots + \Sigma\Sigma b_{ij}(x_i x_j) + c \cdot \ln(\sigma'_3)$	Granular waste, sand, ballast (+ rubber)	Hunt et al. (2024)

Where σ_n = normal stress; σ_c = uniaxial compressive strength; *a-d* = regression coefficients; ϕ'_b = true interparticle friction angle from tilt table test; D_{50} = median particle size; C_u = coefficient of uniformity; R_d = relative density; x = (rubber content, void ratio, dry unit weight, D_{50}); σ'_3 = minor principle stress (or confining pressure).

A glaring trend within the literature is that these predictive models are often constrained to a single material type, as evidenced in Table 1. This was investigated by Hunt et al. (2024) where a nonlinear regression model was developed with the aim to predict ϕ'_{peak} for a variety of granular material/mixture types, including conventional (sand, ballast) and waste (CW, SFS) materials with and without rubber inclusion. Several material properties, as well as the effective confining pressure (minor principle stress), were used in order to better describe the material type and its physical state of compaction, which are easily obtainable or measurable without the need for triaxial testing methods. Two-way interaction between interrelated variables were also incorporated by the authors. Although a strong predictive performance was achieved for most materials, this model was less accurate with sand and sand-rubber mixtures, highlighting the limitations in conventional regression modelling techniques.

2.2 *Other machine learning techniques*

A major shortcoming among the models previously discussed is that the nature/degree of nonlinearity needs to be known prior to modelling in order for traditional regression to be effective (Shahin et al. 2024). This is manageable for basic models that relate shear strength to the normal stress since the nonlinear failure envelope is fairly well-established. However, the latter needs to be clearly established in order for these models to be practically viable. Models that are based solely on known, or at least easily obtainable, material properties and stress conditions, such as Hunt et al. (2024), enable the shear strength to be efficiently approximated. However, it is difficult to establish the degree of nonlinearity with respect to each of these parameters, both individually and combined.

Using ANN, Dutta et al. (2015) predicted the peak deviatoric stress (q_{peak}) of sand-plastic mixtures under static triaxial compression. Although several input parameters were used to characterize the soil blend, key geotechnical parameters that represent the material properties and physical state are absent. This is also similar with the axial stress-strain predictive model developed by Edincliler et al. (2012) for sand-rubber mixtures. Particle sizes, gradation, and relative density parameters were included by Abozraig et al. (2022) to predict of sand and gravel using various ML techniques. However, since ϕ'_{peak} is stress-dependent, the choice of ϕ'_{peak} is fundamentally invalid as no stress parameter (e.g., σ'_3) is present.

In contrast, Hunt et al. (2024) developed their ANN model for various granular materials blended with rubber based on key geotechnical parameters, including particle size, dry density, and void ratio, in an attempt to increase its versatility in predicting ϕ'_{peak}. Indraratna et al. (2025) extended this by including the coefficient of curvature (C_c) to improve its predictive performance and generalization. In order to further validate the model, parametric analysis was carried out to verify that fundamental geotechnical relationships and concepts were preserved, similar to Shahin & Indraratna (2006) with the axial stress-strain response of ballast. This further validation step is lacking in the majority of geotechnical ML studies, where validation is often based solely on statistical performance metrics. Although, in the two aforementioned studies, properties such as particle hardness and shape were not considered due to the limited availability of data. Hardness was however included by Zhou et al. (2019) for τ_f of rockfill materials using RF and cubist algorithms, while Penumadu and Zhao (1999) successfully included both hardness and shape to forecast the axial stress-strain curve of sand and gravel using recurrent neural networks (RNN). While numerous studies in the literature are limited due to data scarcity and incompleteness, the implementation of a more extensive database with greater variance enabled the two aforementioned studies to investigate these additional parameters more thoroughly. In the current age of big data and AI/ML, the importance of more extensive recording and publishing of various material properties is hereby reflected.

3 RESILIENT MODULUS

The resilient modulus (M_R), defined as the ratio of the cyclic deviatoric stress (q_{cyc}) and recoverable axial strain ($\varepsilon_{a,rec}$), represents the stiffness of the elastic (recoverable) component of a material when subjected to dynamic loading-unloading cycles. It is a key parameter in dynamic loading problems such as the design and maintenance of unbound pavement and track substructure layers, and is typically determined through cyclic or repeated load triaxial (RLT) tests (Zaman et al. 1994; Indraratna et al. 2023). As these laboratory methods are onerous and expensive, relying on them may jeopardize project budgets and timelines, particularly for smaller projects and/or where relatively large testing suites are required. This problem is exacerbated when considering large granular materials, such as railway ballast, where large-scale test samples and equipment are desired. Therefore, it is favorable for practitioners to be able to accurately predict M_R based on correlations with key material and loading properties that are either already known or easily measurable.

3.1 *Conventional regression*

Numerous models have been proposed by prior researchers using conventional statistical regression, each with their own advantages and disadvantages. Of these, there is generally a clear distinction with respect to their versatility and validity as a result of the parameters they are based on. Arguably the simplest and most popular models are those where M_R is directly correlated to the California Bearing Ratio (CBR) (e.g., Heukelom & Klomp (1962), which is a parameter widely used in road and pavement design. Although simple and efficient to use as the CBR test can be easily performed, these models can grossly underpredict or overpredict M_R (Zaman et al. 1994; Plati & Tsakoumaki 2023). This is due primarily to the model lacking material property parameters since different granular specimens may have the same CBR but differing resilient stiffnesses. In fact, almost all of these correlations are restricted to fine-grained materials (Plati & Tsakoumaki 2023). Other models in the literature correlate M_R with loading parameters such as bulk stress, octahedral shear stress, deviator stress, and confining pressure (e.g., Seed et al. 1967; Shackel 1973; Brown et al. 1975; Uzan 1985). However, as these models contain regression constants, these models again lack transferability for use in cases where material and/or loading values differ from those in each specific study. To alleviate this, these regression constants require recalibration experimentally, thus nullifying their practicality. The resilient modulus is also influenced by the aggregate type, roughness, size/shape, and grading (Lackenby 2006), though, again, it can be difficult to accurately capture these relationships using conventional methods.

3.2 *Other machine learning techniques*

Several researchers (e.g., Kim et al. 2014; He et al. 2024) have utilized ML techniques to incorporate material properties such as moisture content (%), dry density, particle size and gradation, unconfined compressive strength, as well as the loading parameters previously mentioned. Of these, the use of ANN is the most popular (Sadik 2024), though the majority are concerned with soft and fine-grained soils (Indraratna et al. 2023) and thus often include Atterberg limits that are less relevant for coarser particles. ANFIS and ANN have been used to predict M_R for coarser construction and demolition wastes (Ghorbani et al. 2021) and blends with rubber (Ghorbani et al. 2024). Both models used σ_3, q_{cyc}, and $w\%$, with the latter arguably being more robust as it also includes CBR, dry density, and optimum moisture content, though gradation characteristics were not considered. Some gradation characteristics were also excluded by Hao & Pabst (2023) for course-grained waste rock due to their interdependencies with other input parameters. For geotechnical problems, it is uncommon for one material property to be completely independent of another. For example, particle size and gradation influence the density and porosity of a matrix, all of which affect a material's loading behavior. Since ML techniques are more adept at extrapolating complex relationships within data, there is a strong case for geotechnical researchers to diverge from the traditional way of thinking where complete independency between model input parameters is desired. In fact, incorporating more of these interrelated parameters can help distinguish between different material types and thus increase a model's versatility.

Lastly, while it is beneficial for practitioners to forecast the evolution of M_R with number of loading cycles () over the lifetime of a granular layer, most studies are concerned with only the final value upon reaching plastic shakedown. Using ANN and ANFIS, Indraratna et al. (2023) were the first researchers to develop (to the authors' knowledge) predictive ML models to forecast M_R of railway ballast as a function of N. Although, due to limited data variance, this model contains no material-defining input parameters and hence its applicability is limited to the specific ballast material and physical properties used. Loading frequency was included as an input, however, making it more versatile to changes in train speeds compared to the earlier regression model proposed by Lackenby (2006), which is a function of just q_{cyc}, σ_3, and N. At present, there are little to no ML studies to date that effectively incorporate particle shape and roughness parameters to predict M_R for coarse granular materials, which is again a direct result of data scarcity.

4 CONCLUSIONS

This study has provided a short critical review of selected studies using of ML techniques to predict shear strength characteristics and resilient moduli of granular materials. In lieu of analyzing specific modelling intricacies such as learning algorithms and architectures used, this review was focused on the choice of input parameters and their adherence to geotechnical engineering concepts that are well-established.

From this analysis, it is clear that ML can facilitate the inclusion of numerous material properties that are otherwise difficult to model using conventional regression and mathematical methods. However, it is difficult to extrapolate meaningful relationships using datasets with low variance, which is often a result of data scarcity and lack of extensively reported material properties. With this, the authors encourage researchers to increase the recording and publishing of these properties for use in future ML studies.

Furthermore, there is a division between studies that use ML as a tool to either fit a specific dataset, or to develop more generalized models that enhance existing geotechnical relationships and concepts for practical use. Given the increasing popularity of ML, coupled with the hesitancy of practitioners to apply it, it is evident that more emphasis is needed with respect to model validation against these established concepts, rather than using statistical performance indices alone. As and M_R are state parameters, practitioners also need to know the range of inputs (e.g., confining stress) an ML model has been calibrated to as this limits the range of applicability. Moreover, researchers are encouraged to diverge from the traditional way of thinking and further investigate input variables that are interdependent, which may help increase model versatility.

Lastly, this review identified the lack of ML studies that predict M_R of coarse granular materials, with the majority pertaining to soft and fine-grained materials. Of those that do exist, only the final M_R value is usually predicted while its evolution over time is lacking in the current literature.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support provided by the Australian Research Council Linkage Project (ARC-LP200200915), as well as the financial and technical support provided by industry partners ARTC, Sydney Trains, SMEC, Bentley, and Bestech.

REFRENCES

Abozraig, M., Ok, B. & Yildiz, A. 2022. Determination of shear strength of coarse-grained soils based on their index properties: a comparison between different statistical approaches. *Arabian Journal of Geosciences* 15: 593.

Andjelkovic, V., Pavlovic, N., Lazarevic, Z. & Radovanovic, S. 2018. Modelling of shear strength of rockfills used for the construction of rockfill dams. *Soils and Foundations* 58: 881–893.

Back, M. & Szendefy, J. 2019. Comparison of the resilient modulus and the California Bearing Ratio in case of stabilized soils. In *Proceedings of the XVII ECSMGE-2019*

Brown, S.F., Lashine, A.K.F. & Hyde, A.F.L. 1975. Repeated load triaxial testing of a silty clay. *Géotechnique* 25(1): 95–114.

Carter, J.P., Chandra, D., Potts, D., Schweiger, H. & Sloan, S. 2000. Computing and computer modelling in geotechnical engineering. In *Proceedings of the International Conference on Geotechnical & Geological Engineering (GeoEng2000) Melbourne, 19-24 November 2000*: 1157–1252.

Dutta, R.K., Dutta, K. & Jeevanandham, S. 2015. Prediction of deviator stress of sand reinforced with plastic strips using neural network. *International Journal of Geosynthetics and Ground Engineering* 1: 11.

Edincliler, A., Cabalar, A.F., Cagatay, A. & Cevik, A. 2012. Triaxial compression behavior of sand and tire wastes using neural networks. *Neural Computing and Applications* 21: 441–452.

Ghorbani, B., Arulrajah, A., Narsilio, G., Horpibulsuk, S. & Leong, M. 2021. Resilient moduli of demolition wastes in geothermal pavements: Experimental testing and ANFIS modelling. *Transportation Geotechnics* 29: 100592.

Ghorbani, B., Yaghoubi, E., Wasantha, P.L.P, van Staden, R., Guerrieri, M. & Fragomeni, S. 2024. Machine learning-based prediction of resilient modulus for blends of tire-derived aggregates and demolition wastes. *Road Materials and Pavement Design* 25(4): 694–715.

Hao, S. & Pabst, T. 2023. Mechanical characterization of coarse-grained waste rocks using large-scale triaxial tests and neuroevolution of augmenting topologies. *Journal of Geotechnical and Geoenvironmental Engineering* 149(6): 04023039.

He, B., Armaghani, D.J., Tsoukalas, M.Z., Qi, C., Bhatawdekar, R.M. & Asteris, P.G. 2024. A case study of resilient modulus prediction leveraging an explainable metaheuristic-based XGBoost. *Transportation Geotechnics* 45: 101216.

Heukelom, W. & Klomp, A.J.G. 1962. Dynamic testing as a means of controlling pavements during and after construction. In *Proceedings of the International Conference on the Structural Design of Asphalt Pavements Ann Arbor*, *20-24 August 1962*, Volume 203: 495–510.

Hunt, H., Indraratna, B. & Qi, Y. 2023. Ductility and energy absorbing behaviour of coal wash-rubber crumb mixtures. *International Journal of Rail Transportation* 11(4): 508–528.

Hunt, H., Indraratna, B. & Qi, Y. 2024. Predicting the shear strength of granular waste materials using machine learning. In *Proceedings of the 5th International Conference on Transportation Geotechnics (ICTG) Sydney*, *20-22 November 2024*, Volume 1: 155–163.

Indraratna, B., Ionescu, D. & Christie, H.D. 1998. Shear behavior of railway ballast based on large-scale triaxial tests. *Journal of Geotechnical and Geoenvironmental Engineering* 124(5): 439–449.

Indraratna, B., Armaghani, D.J., Gomes Correia, A., Hunt, H. & Ngo, T. 2023. Prediction of resilient modulus of ballast under cyclic loading using machine learning techniques. *Transportation Geotechnics* 38: 100895.

Indraratna, B., Hunt, H., Malisetty, R.S., Alagesan, S., Qi, Y. & Rujikiatkamjorn, C. 2025. Optimization of inputs for the application of ANN to rail track granular materials. *Canadian Geotechnical Journal* 62: 1–22.

Kim, S.H., Yang, J. & Jeong, J.H. 2014. Prediction of subgrade resilient modulus using artificial neural network. *KSCE Journal of Civil Engineering* 18: 1372–1379.

Lackenby, J. 2006. Triaxial behaviour of ballast and the role of confining pressure under cyclic loading. PhD thesis, University of Wollongong, Wollongong.

Leslie, D.D. 1969. Relationship between shear strength, gradation and index properties. In *Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering Mexico City 1969*, Specialty Session No.1: 212–222.

Liu, H., Su, H., Sun, L. & Dias-da-Costa, D. 2024. State-of-the-art review on the use of AI-enhanced computational mechanics in geotechnical engineering. *Artificial Intelligence Review* 57(196).

Penumadu, D. & Zhao, R. 1999. Triaxial compression behavior of sand and gravel using artificial neural networks (ANN). *Computers and Geotechnics* 24: 207–230.

Plati, C. & Tatsoumaki, M.A. 2023. Critical comparison of correlations for rapid estimation of subgrade stiffness in pavement design and construction. *Construction Materials* 3: 127–142.

Sadik, L. 2024. Developing prediction equations for soil resilient modulus using evolutionary machine learning. *Transportation Infrastructure Geotechnology* 11: 1598–1620.

Seed, H.B., Mitry, F.G., Monismith, C.L. & Chan, C.K. 1967. Factors influencing the resilient deformation of untreated aggregate base in two layer pavements subjected to repeated loading. *Highway Research Board, Record No. 190*.

Shackel, B. 1973. The derivation of complex stress-strain relations. In *Proceedings of the 8th International Conference on Soil Mechanics and Foundation Engineering*, Volume 1: 353–359.

Shahin, M. & Indraratna, B. 2006. Modeling the mechanical behavior of railway ballast using artificial neural networks. *Canadian Geotechnical Journal* 43: 1144–1152.

Shahin, M. 2024. Progression of artificial intelligence/machine learning in geotechnical engineering. *Machine Learning and Data Science in Geotechnics* 1(1): 1–5.

Sharma, V., Kumar, A., Priyadarshee, A. & Chhtou, A.K. 2019. Prediction of shear strength parameter from the particle size distribution and relative density of granular soil. In *Environmental Technology. Proceedings of EGRWSE Jalandhar*, *29-31 March 2018*: 185–191.

Uzan, J. 1985. Characterization of granular material. *Transportation Research Record* 1022: 52–59.

Wallace, M. & Ng, K. 2016. Development and application of underground space use in Hong Kong. *Tunnelling and Underground Space Technology* 55: 257–279.

Xiao, Y., Liu, H., Chen, Y. & Jiang, J. 2014. Strength and deformation of rockfill material based on large-scale triaxial compression tests. II: Influence of particle breakage. *Journal of Geotechnical and Geoenvironmental Engineering* 140(12): 04014071.

Yang, Y. & Rosenbaum, M.S. 2002. The artificial neural network as a tool for assessing geotechnical properties. *Geotechnical and Geological Engineering* 20: 149–168.

Zaman, M., Chen, D.H. & Laguros, J. 1994. Resilient moduli of granular materials. *Journal of Transportation Engineering* 120(6): 967–988.

Zhou, J., Li, E., Wei, H., Li, C., Qiao, Q. & Armaghani, D.J. 2019. Random forests and cubist algorithms for predicting shear strengths of rockfill materials. *Applied Sciences* 9: 1621