CHAPTER- VI

PREDICTION OF DYNAMIC SHEAR MODULUS OF UNREINFORCED AND REINFORCED MSW FINES: MACHINE LEARNING APPLICATION

6.1 INTRODUCTION

The mechanical or geotechnical properties of the MSW fines are very unpredictable due to the heterogeneity of the material and these properties are site-specific. The physical characteristics of the waste dumped in the landfills very much depend on the climate, culture, composition, consumption patterns of the community, etc. This physical composition governs the mechanical characteristics of the waste. The determination of these properties in the laboratory itself is a very challenging task and requires an ample amount of time, which makes this data very expensive (concerning time and cost of instrumentation) to generate. Other than this, excavation, segregation, handling, and placing of waste make it more challenging and costly. Machine Learning (ML) based mathematical or numerical models can be used to save cost and time to estimate the static and dynamic properties of the MSW and can cover a broader range of problems. As technology advances, predictive models are gravitating toward artificial intelligence (AI) approaches. AIs are computationally designed methods that are expected to mimic human thinking or cognitive skills in the solution of engineering problems. This technique is appropriate for solving engineering problems with many inputs or random variables where the correlation between input(s) and output(s) is unknown. Machine learning (ML) (a branch of AI) is used to create a mathematical model that makes predictions without being explicitly programmed to do so (Kandiri et al., 2020). There are limited ML-based models used for MSW.

Table 6.1 Machine learning application for MSW model prediction.

ML model used for MSW	Objective	Reference
Artificial neural network (ANN), k-nearest neighbors (kNN), adaptive neuro-fuzzy inference system (ANFIS), and support vector machine (SVM)	Prediction of monthly waste generation (Australia)	Abbasi and El Hanandeh (2016)
Decision trees and neural networks	MSW generation and diversion in terms of socioeconomic and demographic variables	Kannangara et al. (2018)
Least-squares support vector machine (LS-SVM)	Computing higher heating value concerning elemental compositions.	Rostami and Baghban (2018)
Feedforward neural network (FFNN) and an SVM.	Estimating high heating value for various MSW	Bagheri et al. (2019)
ML-based fuzzy probabilistic model	Compost usability index and its quality	Mohurle and Devare (2019)
ANN	Applications for sustainable development, such as waste management	Gue et al. (2020)
Artificial Neural Networks (ANN), Multivariate Adaptive Regression Splines (MARS), Multi-Gene Genetic Programming (MGGP), and M5 model Tree (M5Tree)	Prediction of the shear modulus of MSW	Alidoust et al. 2021

There have been some past studies on predicted models for the MSW (shown in Table 6.1). Although there are studies, where ML is used for predicting some physical characteristics of the waste or waste generation, only a few studies that are recently been introduced to evaluate the static or dynamic characteristics of the MSW. Even in the case of soils, ML applications are very limited and have been introduced in the last two decades.

6.1.1 ML Applications in Constitutive Modeling of Soils

For the last many decades, researchers are proposing consecutive models to predict the complex behaviour of soils. These models can be categorized as linear elastic perfectly plastic models, nonlinear models, critical state-based models, micromechanical models, etc. There are limitations attached to this model, certain assumptions underpin all constitutive models proposed, and each model is only appropriate for a subset of soil types, with numerous parameters, mathematical formulas become increasingly complicated. Although the mathematical equation in a constitutive model is derived from the experimental observations, and the formula's form provides excellent accuracy for only selected tests, the model's predictive ability for other tests with different stress paths is limited. The ML models can overcome these drawbacks with the following advantages (Zhang et al., 2021b):

- Without making any assumptions, the ML model can learn the stress-strain relationship directly from the raw data.
- As long as the experiments of such soils are included in a database, the ML can develop a uniform model for simulating the behaviours of various soils.
- With an increasing number of datasets, the predictive accuracy and application scopes of ML-based models can be improved.
- ML-based model is data-driven, no parameter calibration is required once the ML configurations are determined.

Many ML algorithms have been used by researchers, including genetic programming (GP) (Cabalar and Cevik, 2011), evolutionary polynomial regression (EPR) (Javadi and

Rezania, 2009; Cuisinier et al., 2013; Asr et al., 2013), support vector machine (SVM) (Zhao et al. 2014; Kohestani and Hassanlourad, 2016), backpropagation neural network (BPNN) (Ghaboussi and Sidarta, 1998; Basheer, 2002; Banimahd et al., 2005; Hashash and Song, 2008; Hi and Li, 2009; Sezer, 2011; Araei, 2014; Lin et al., 2019), radial basis function (RBF) (Peng et al., 2005), recurrent neural network (RNN) (Romo et al., 2001), long short-term memory (LSTM) (Wang et al., 2018; Zhang et al., 2020), and gate recurrent unit (GRU) (Wang and Sun, 2019) for clay, sand, gravel, ballast, rockfill, frozen soil, reinforced soil, and soils with various mixtures to simulate the stress-strain responses.

6.2 PROBLEM SETTING

Machine learning modeling is a proven effective measure to efficiently study the behavior of various systems. These systems otherwise can be resources intensive in terms of time required to perform these experiments. The machine learning models can be learned from the data obtained by computer simulations or physical experiments. These simulations or experiments are obtained in the form of system inputs and corresponding system outputs and approximate the behavior of any underlying real-life system. The input-output relationship of these underlying processes can be derived from the data, which makes it computationally efficient to run these models. A machine learning model takes input and output from a computationally expensive model or process.

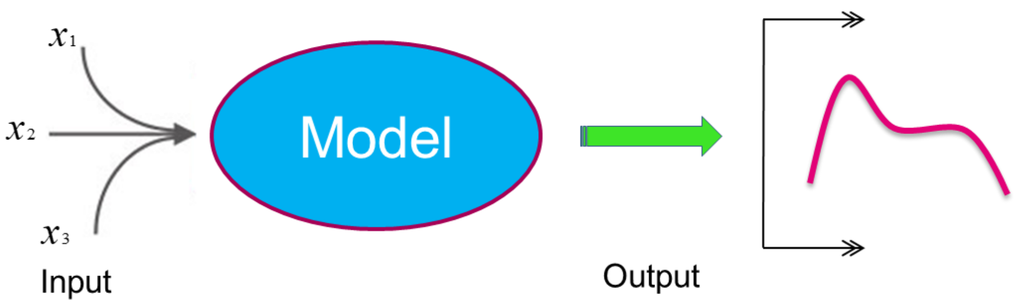


Figure 6.1 Machine learning model

A typical machine-learning model learning requires the following steps:

a) **Model formulation:** it involves proper parameterization of the study problem and defining process variables.
b) **Sample Design:** the choice of sampling strategy influences the performance of the machine learning model. This is also called data collection (sampling).
c) **Model fitting/Training:** a model can be used that best fits the data. In the model training, the modeling parameters are optimized to best represent the input-output relationship.
d) **Model validation:** a trained model is validated for accuracy and robustness with the help of performance metrics (error measures).
e) **Prediction:** the system response can be predicted at a new point using a trained machine learning model.

6.2.1 Prediction Models

To study the properties of the MSW fines, physical experiments were performed in the laboratory, and data is recorded. The detail of the input variables are Input variables: Relative compaction, confining pressure, frequency, and shear strain. The output quantity of interest is the dynamic shear modulus. Two cases are considered in the analysis: a) fiber-reinforced MSW fines and, b) MSW fines (unreinforced).

6.2.1.1 Artificial Neural Network (ANN)

The neural networks are inspired by biological neural networks and derived their ability by processing data utilizing parallelism. These processing elements are called neurons, which function together in a group as their biological counterparts to solve a problem utilizing organizational principles believed to be used in humans. A biological neuron is a special cell that processes information (in the form of impulses, and signals). Similarly, a computational or artificial neuron is a building block of the neural network and processes information. A neural network is composed of a group of neurons (single unit perceptron) and is created by assembling these neurons in a suitable architecture. The most common form of these networks is a multi-layer feed-forward structure. However, specifying the architecture and then training the network are still the main issues in creating such a neural network.

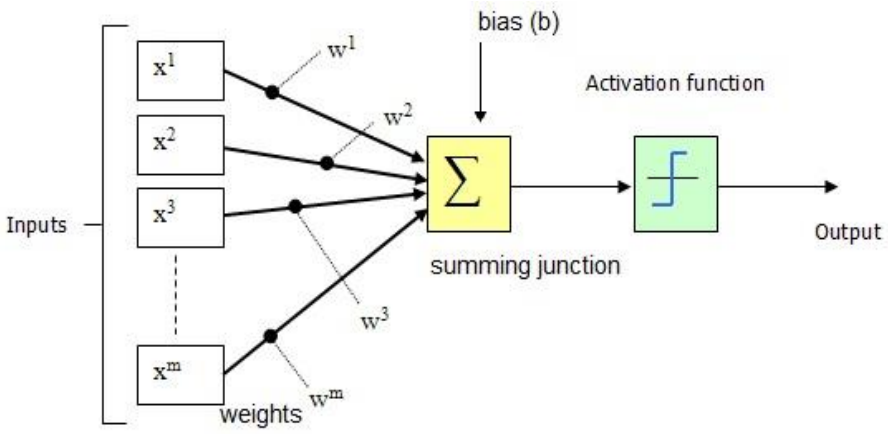


Figure 6.2 Constitution of an artificial neurons

Moreover, a binary threshold unit as a computational model is shown in Figure 6.2, (McCulloch and Pitts, 1943). Where, [x^1, x^2, …, x^m] are the inputs that are parameterized by suitable weights [w^1, w^2, …, w^m] to generate an output. In a computational neuron, an activation function is used to control the amplitude of the output of a neuron. Also, known as a mathematical neuron, it can be described as (Equation 6.1)

$$y = F\left(\sum_{j=1}^{n} wj\, xj + b\right) \quad (6.1)$$

where, x_j is the input space, w_j is the weight value, b is the bias, F is a transfer function and y is the output value.

A neural network architecture is formed when two or more artificial neurons are combined to form an artificial neural network. The way the network is connected is called as topology/architecture or graph of a network. The optimum topology of the network defines the network's ability to solve a problem efficiently and accurately. However, as the parent biological neural networks learn responses to given inputs, in a similar way these networks need to learn accurately. This is achieved by learning proper responses through neural network training. The mathematical objective of learning is to optimize the values of weights and biases on the given inputs to minimize the error or maximize the learning. Levenberg-Marquardt training algorithm has proven successful in terms of realization accuracy and improved convergence in a budget time. It forms a good transition between the Gauss-Newton algorithm and the steepest descent method. Its basic idea is to combine training around the area with complex curvature. Once the network is trained accurately, it can be used to predict the output at any new input in the problem domain. An accurate neural network not only depends on the above factors but also on the statistical normalization of the data, training time, and activation function type.

6.2.1.2 Gaussian Process Regression

In regression analysis, a model can be fitted to define a function on a given input (training data) and corresponding outputs. In a nonlinear problem, a model outcome will be a single function that represents the data and its relationship. There may be the cases where a series of functions (at least more than one) defines the input-output relationship precisely. Extending it, a Gaussian process (Regression) model can be described by a probability distribution over possible functions that describes these input-output relationships. This is a prior (assumption) over infinite possible functions. The prior distribution of these functions is the multivariate Gaussian distribution. The prior distribution consists of an infinite number of functions which reduces as the number of data points increases (keeping only the best that fit the present data well).

A Gaussian distribution is represented by its mean and variance. Similarly, a Gaussian process which is a distribution over the functions is defined by a mean function and a covariance function (Equation 6.2).

$$f(x) \sim GP(m(x), k(x, x')) \tag{6.2}$$

where, f(x) is the set of functions, m(x) is a mean function, and k(x,x′) is a covariance function that represents the covariance between points (x,x′) in the input domain. “x” are the observed data points. The shape of the Gaussian processes is defined by the covariance function. The covariance function relates one observation to another observation and is the heart of the Gaussian process regression (GPR) model. The covariance is described by a kernel that is parameterized (called kernel parameters or hyperparameters). For example, a radial basis kernel function (RBF) can be defined as (Equation 6.3):

$$k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{1}{2l}(x_i - x_j)^T (x_i - x_j)\right) \tag{6.3}$$

The RBF function is parameterized by σ_f (vertical scale) and l (horizontal scale) which are the hyperparameters. These hyperparameters need to be optimized to obtain the best representation of the data by a GPR model. Once the hyperparameters are tuned, the predictions at a new point can be performed.

6.2.1.3 Sensitivity Analysis

In any functional model, the output quantity of interest can be influenced differently by the input variables of the system. The output can be more sensitive to some input variables or may not be affected at all by variables. The sensitivity of each input variable can be calculated to study its impact on the output. This value is often represented by a numerical value. These interactions can come in several forms such as:

- Sensitivity contribution of a single input variable on the output variance.
- The sensitivity contribution is because of the interaction of the two input variables (higher-order sensitivity indices).
- The total sensitivity contribution to the output by the individual and interaction of the two input variables.

Several methods are available to perform the sensitivity analysis (SA), here SOBOL (SOBOL, 2001) indices have been considered for the sensitivity analysis (SA), which is a widely used method to determine the sensitivity of the input variables. SOBOL indices can be represented as (Equations 6.3 and 6.4)

$$S_i = \frac{Var(M_i(X_i))}{Var(Y)}, \quad i = 1, \dots, n \tag{6.3}$$

$$S_{ij} = \frac{Var(M_{ij}(X_{ij}))}{Var(Y)}, \quad 1 \leq i < j \leq n \tag{6.4}$$

where, S_i is the first-order SOBOL indices of an individual input variable and S_{ij} is the second-order indices representing interactions between the variables. M is a random variable that depends on the input observations of X, and Y is the output.

6.3 RESULTS AND DISCUSSION

6.3.1 Test Setup

In this section, the performance of the prediction of the machine learning models and sensitivity analysis is demonstrated. A total number of 100 samples are used for the unreinforced MSW fines and 21 for the fiber-reinforced MSW fines respectively. The data is divided into training (70%) and testing sets (30%). Two prediction models are considered to predict the dynamic shear modulus, i.e., artificial neural network and Gaussian process regression.

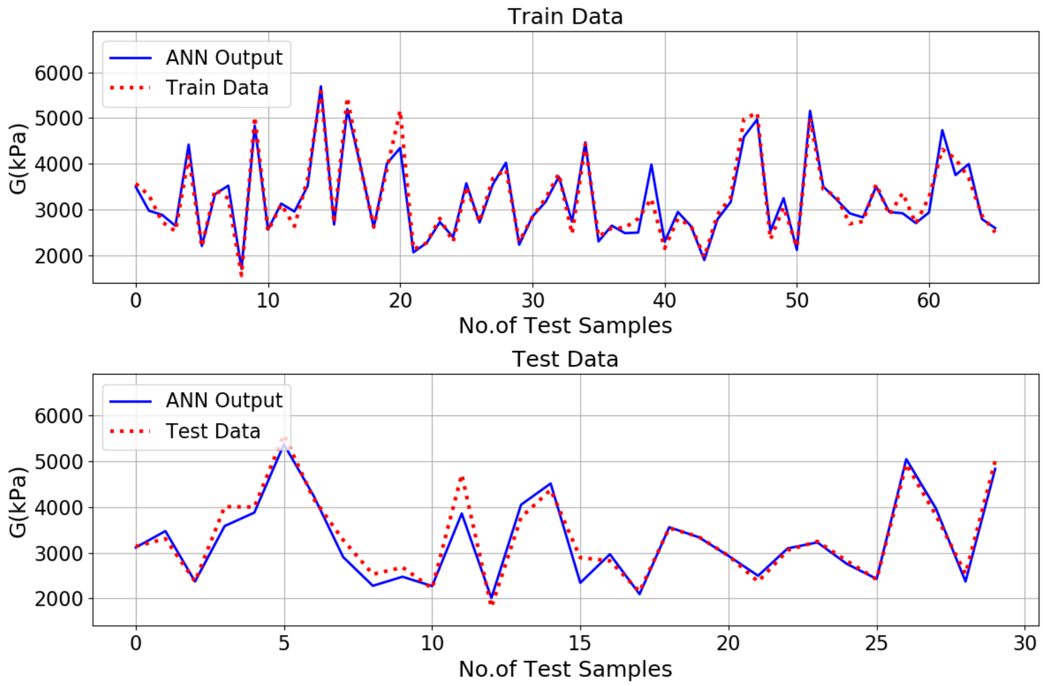
A network architecture of [4,5,5,1], i.e., 4 inputs, 2 hidden layers with 5 neurons each, and one output layer are used for the unreinforced MSW fines case. In the case of reinforced MSW fines, a network architecture [2,2,2,1], i.e., 2 inputs, 2 hidden layers with 2 neurons each, and one output layer is used. The training of the neural network is performed with the Levenberg-Marquardt algorithm for both cases. To learn a Gaussian process regression model on the input data, the Matern52 kernel function is used. The modeling ability and fit of the predicting models are assessed using the root mean squared error. The ANN model is written in python and MATLAB is used for the GPR models (Equation 6.5).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \|y(i) - \hat{y}(i)\|^2}{N}} \quad (6.5)$$

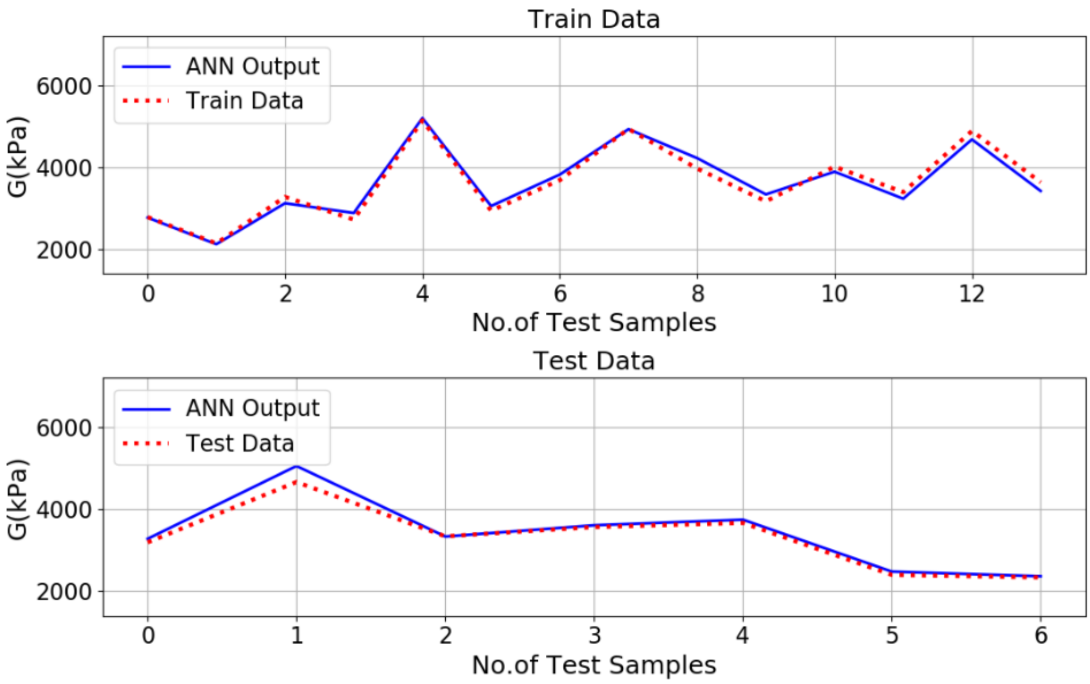
where, N is the number of samples, y(i) is the observed output and $\hat{y}$(i) is the corresponding predicted output value.

6.3.2 Prediction Using Artificial Neural Network

In Figure 6.3, the prediction results of dynamic shear modulus are presented for both the cases of unreinforced and reinforced MSW fines using the Artificial Neural Network (ANN) model. An ANN model results in good approximation for both test and trained data. It captures the dynamic behaviour, i.e., the dynamic shear modulus satisfactorily. At a few locations, the model seems to be overtrained and results in a satisfactory outcome. The RMSE values of 226 and 244 are observed for the training and testing data which is significant considering the dynamic shear modulus value of the MSW fines(Figure 6.3(a)). In the case of reinforced MSW (Figure 6.3(b)), the approximation is better, and all variations are captured by the model. The RMSE values come out to be 106 and 159 for the training and testing data.



(a)

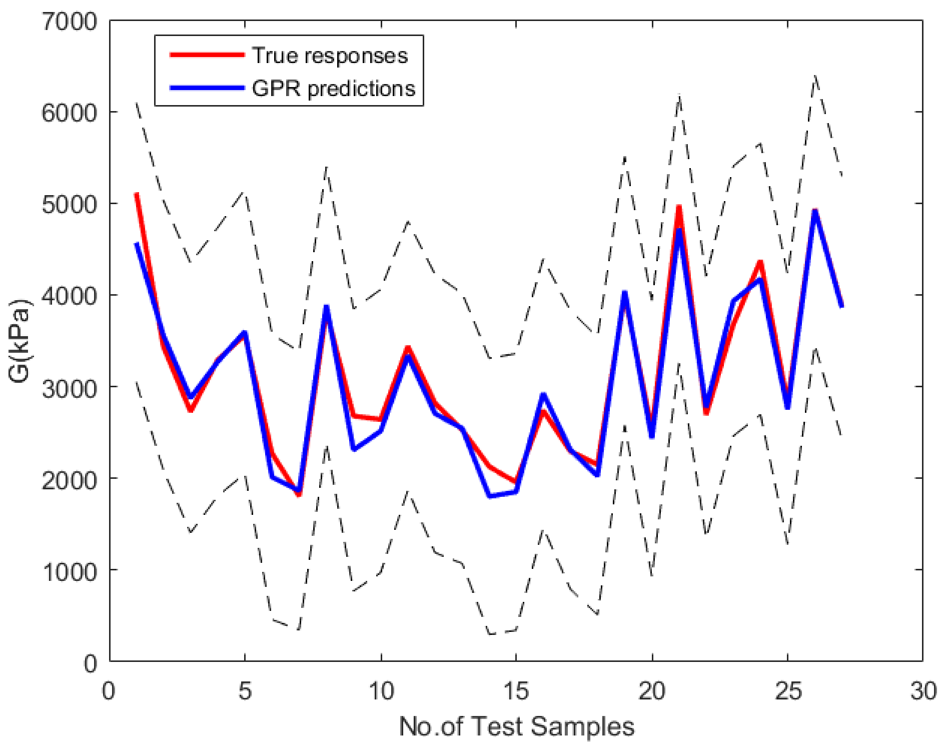


(b)

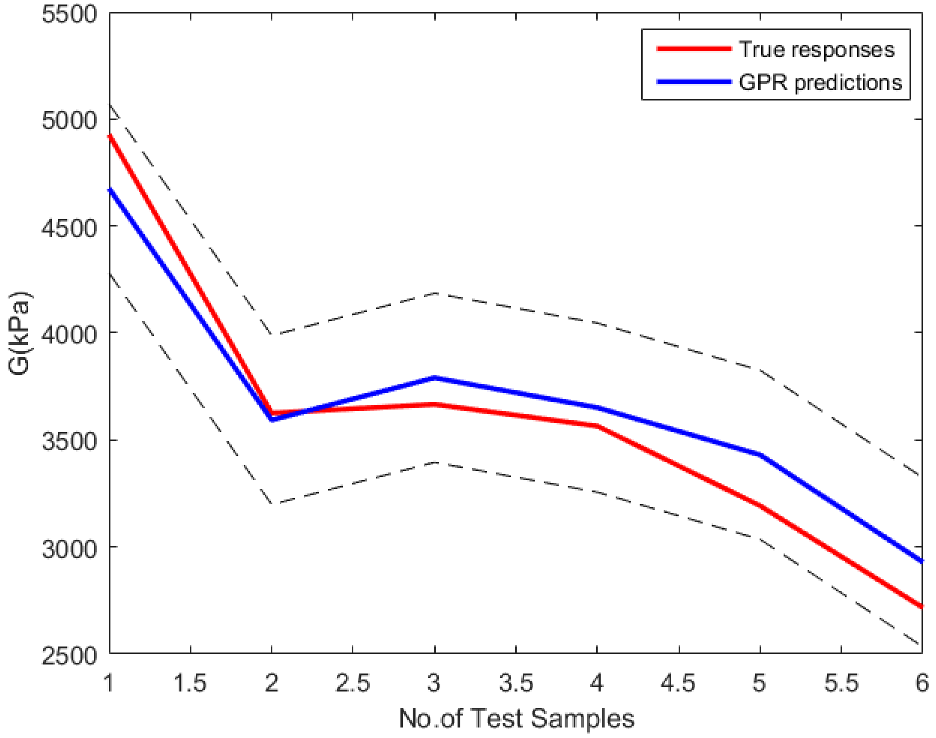
Figure 6.3 Prediction of dynamic shear modulus for **(a)** Unreinforced MSW fines, and (**b)** Fiber reinforced MSW fines using ANN model

6.3.3 Prediction Using Gaussian Process Regression

In Figure 6.4 (a), the results of dynamic shear modulus from the Gaussian process regression (GPR) model are demonstrated for the unreinforced and reinforced MSW fines. In both cases, the GPR model captures all significant trends in the dynamic shear modulus (G) behaviour. The RMSE value of 27 and 187 were observed in the case of unreinforced MSW fines for the training and testing data set (Figure 6.4 (a)). Similarly, the RMSE value of 0.42 and 177 were observed in the case of fiber-reinforced MSW fines for the training and testing data set (Figure 6.4 (b)).



(a)



(b)

Figure 6.4 Prediction of dynamic shear modulus for **(a)** Unreinforced MSW fines, and (**b)** Fiber reinforced MSW fines using GPR model

Table 6.2 Root Mean Squared Error (RMSE) values obtained from prediction model analysis.

Material	RMSE value			
	ANN Model		GPR Model	
	Train	Test	Train	Test
Unreinforced MSW fines	221	246	27	187
Reinforced MSW fines	147	158	0.42	177

Moreover, in Table 6.1, the results are compared for all the cases. In both cases, the GPR model shows improvement over the ANN model. On the test data set, an improvement of over 24% and 10% is observed by the GPR model over the ANN model for the unreinforced and reinforced MSW fines respectively. The choice of kernel function and its initial parameters highly affects the GPR model.

6.3.4 Sensitivity Analysis

The results of the sensitivity analysis of the considered parameters in the case of unreinforced MSW fines on dynamic shear modulus are presented in Table 6.2. The first-order interactions highlight that the input variable has a higher impact on the dynamic shear modulus. While the highest individual influence is from shear strain (77.5%) and the lowest is from the frequency (0.2%). The relative compaction and effective confining pressure have limited influence in comparison to shear strain.

Table 6.3 Sensitivity indices for the first-level interactions.

Variable	Sensitivity Index (First level Interactions)	Contribution %
Relative Compaction	0.045	4.5
Effective Confining Pressure	0.112	11.2
Frequency	0.002	0.2
Shear Strain	0.775	77.5

6.4 SUMMARY

The dynamic shear modulus (G) computed from the cyclic triaxial test for the unreinforced and reinforced MSW fines was predicted through two machine learning techniques, i.e., the ANN model and the GPR model. It was observed that the GPR method produced better results for the unreinforced case, whereas both ANN and GPR model works well for the reinforced MSW fines. The sensitivity analysis shows the highest influence of shear strain on the dynamic parameter (G) with a 77.5% contribution, with the negligible effect of loading frequency, which has also been concluded in chapter IV through experimental analysis.