Use of machine learning in determining Gmax from bender element tests

Wenzhang Xu[1], *Truong* Le[2]

Imperial College London, Civil and Environmental Engineering, Imperial College Rd, South Kensington, London SW7 2BB, UK

[#]*wenzhang.xu21@imperial.ac.uk, truong.le@imperial.ac.uk*

ABSTRACT

The use of bender element is one of the most popular methods of determining shear wave velocity, and hence elastic shear modulus due to its relatively straightforward experimental set-up. While several analysis methods have been proposed, manual interpretation using the first arrival continues to be favoured owing to its simplicity. This paper presents a novel automated program for determining the shear wave velocity and associated maximum shear modulus. The proposed new method involves the use of Convolutional Neural Networks (CNNs) to predict the most probable shear wave velocity using a range of input frequencies as the inputs. Estimates made by the trained CNN are compared to values determined using more traditional interpretation methods (first-arrival, cross-correlation and frequency domain). The program is able to autonomously determining the shear modulus in the three principal orientations (G_{vh}, G_{hv}, and G_{hh}) at a range of stress levels. The shear modulus determined using the range of techniques showed great agreement. Statistical analysis of the determined shear modulus regression of over 0.99 between interpretations made using first arrival and that estimated using the new CNN approach.

Keywords: Bender element, small-strain stiffness, machine learning, convolutional neural networks

1. Introduction

It is now well known that the stress-strain behaviour of soil is highly non-linear and the shear modulus, G, may decay with strain by orders of magnitude. This has led to the development of methods of analysis for geotechnical structures that routinely takes into stiffness nonlinearity into account. The shear modulus at small strains near the start of loading is G_0. While it is possible to measure soil stiffness in the small strain range using local instrumentation, it is much easier to measure soil stiffness using dynamic methods. One of the most widely used methods for estimating the small strain shear modulus is to make use of bender elements (BE), small piezoceramic devices that can generate and detect waves in the soil. Using BE, the maximum shear modulus can be indirectly determined by estimating the travel time, t_{tra}, of a shear wave between two piezo-ceramic plates embedded in the sample. It is well known however that the accuracy of determined G_0 is largely dependent on the interpretation method of the travel time.

Since the development of BE test in the late 1970s (Shirley and Hampton, 1978), many researchers have been dedicated to seeking a scientific interpretation method of the shear wave velocity. The most currently used methods can be divided into two main parts: time domain (TD) and frequency domain (FD). TD method directly measures the time lag among input and output signals, whereas the FD method determines the arrival time from the inclination of the unwrapped phase spectrum by assuming a non-dispersion Soil-BE system. An alternatively TD method, the cross-correlation (CC) method has been developed to measure the correlation between the input excitation signal and the received signal. These approaches do however give a range of travel times due to a number of issues related to signal interpretation, such as NFE, wave reflection, wave dispersion in the specimen, and the discrepancies in frequency component and waveform between two signals (Greening and Nash, 2004; Gu, X. et al., 2015; Viana da Fonseca et al., 2009; Yamashita et al., 2009; Yang and Gu, 2013). More recently, new approaches have been developed to overcome some of these deficiencies, such as applying the Akaike information criterion to automatically detect the shear wave arrival time (Finas et al., 2016) or using wavelets to transform signals into an easy-handle state (Fernández Lavín and Ovando Shelley, 2020). While these newer approaches are recognised as having the potential to improve the accuracy of the determined shear wave travel time, the interpretation process becomes increasingly complex.

Machine learning (ML) algorithms are seen as an alternative, more robust method for determining the travel time of a shear wave in routine tests using BEs. As the use of neural networks for signal data processing have been shown to be highly effective in other fields, such as speech recognition, real-time electrocardiogram monitoring, and damage detection based on vibration signals (Kiranyaz et al., 2019), this paper will present an implementation of a Convolutional Neural Network (CNN) algorithm in the determination of shear wave velocity using BEs in a triaxial apparatus. It will be shown that the CNN algorithm provides a reliable and consistent estimate of the shear wave velocity using

independent sets of training, validation and testing input/output BE signals.

2. Description of laboratory test

In BE test, the small-strain shear modulus of the soil sample is simply related to the velocity of the shear wave, V_s, and the bulk density of the soil sample, ρ_t:

$$G_0 = \rho_t \cdot V_s^2 \quad (1)$$

The shear wave velocity, V_s, is given by:

$$V_s = L/t_{tra} \quad (2)$$

where L is shear wave travel distance (tip-to-tip distance) and t_{tra} is the interpreted shear wave travel time. The shear wave travel time can be determined using either a time domain or frequency domain approach. For more details on these methods, refer to Viggiani and Atkinson (1995) or Yamashita et al., (2009).

Bender element tests were performed in a 100mm Bishop and Wesley triaxial apparatus equipped with pairs of BEs mounted in the base pedestals and across the mid-height of the specimen to generate and detect shear waves in three orientations, namely, waves propagating vertically with horizontal vibration (V_{vh}), waves propagating horizontally with vertical vibration (V_{hv}) and waves propagating horizontally with horizontal vibration (V_{hh}). The bender elements are measured to have a length of 5mm beyond the face of the housing. A TiePie Handyscope HS5 High Resolution Oscilloscope with arbitrary function generator provided excitation of the bender elements. Input signals, with an amplitude of 20V, were used for to generate the shear wave. The TiePie Handscope was controlled using a bespoke Python script which automatically varied the input waveform and stacked a user defined number of outputs received signal. In this study, 10 stacked received signals was found to result in a sufficiently clear waveform for interpretation. A variety of continuous and pulsed input waveforms were used in this study. Four different types of input signals at a range of frequencies, as shown in Fig. 1, are used: a single sine wave, square wave, triangle wave and a sweep of sine waves, commonly referred to as a chirp waveform.

A 100mm diameter by 200mm high intact sample of London Clay was fitted with the three pairs of bender elements and isotopically consolidated from an initial isotropic effective stress of 92 kPa to a maximum of 400 kPa by slowly ramping the cell pressure at a rate of 1kPa/hr. Hold stages at 100, 150, 200, 250, 300 and 400 kPa isotropic effective stress were defined in order for the bender element data to be collected.

3. Convolutional neural networks (CNNs)

3.1. Data samples preparation

Bender element test data was obtained by recording and pairing the input and received signals from opposing elements in the three orientations. At the end of each consolidation stage, 15 different input frequencies were tested to form a single set of data for each signal waveform with 4 sets of data being collected for each type of input signal. The signal data were subsequently divided into a training, validation and test data set in a ratio of 2:1:1 in order to calibrate and test the ML algorithm.

Besides the paired signals, a labelled value was required for each data sample to enable the CNN to learn and modify its weights and biases during the training. To develop a CNN that can make predictions on the G_0 of the soil sample, the attached label value should reflect the *best-estimated* G_0 of the soil sample.

As G_0 is an intrinsic property of the soil specimen and is independent of the type of input signals and interpretation technique employed, a constant *best-estimated* G_0 using the first arrival method at an optimal frequency was assigned to all paired signals for a given consolidation stage and measured orientation.

An appropriate input sine frequency for the first arrival technique in order to minimise signal distortion effects, near field effects (NFE) and wave attenuation (Dyvik and Madshus, 1985; Blewett et al., 2000, Kawaguchi et al., 2016; Ogino et al., 2015; Sanchez-Salinero et al., 1986; Styler and Howie, 2013; Viggiani and Atkinson, 1995; Yamashita et al., 2009). Fig. 2 presents three sets of bender element signals at increasing frequencies in the horizontal orientation with vertical propagated at 300kPa. The clarity of the received signal is shown to vary with input frequency, as expected. The most stable received signal is shown in Fig. 2 (b). In contrast, Fig. 2 (a) shows an intense vibration around the shear wave arrival point caused by NFE, while the output signal in Fig. 2 (c) fails to maintain a reasonably sinusoidal waveform due to wave attenuation. The clarity of the received signal is noted to remain approximately constant with additional stacking. In this paper, 4 kHz, 7 kHz, and 7 kHz are determined as the optimal frequency for *vh, hh, and hv* orientations respectively. Signals measured at these frequencies resulted in the clearest signals and therefore provided the greatest confidence in estimated shear wave velocity. The travel distance to wavelength ratio was noted to be approximately 4 for all orientations. The V_s, and hence G_0, values determined at these frequencies are assumed to be the *best-estimated* V_s and should be taken as label values assigned to data samples.

3.2. Architecture of CNNs

Traditional CNNs are generally 2D and are used for image recognition, with 2D image data serving as the input layer. However, as data samples collected in the BE test are composed of 1D amplitude data based on time series 1D CNNs are introduced to receive the signal data measured from the BE test.

3.2.1. Input and output layer

The time range for both the input and received signal was set to 2ms with a total 10,000 data points being recorded for each signal. In this study, input and output signals were included as 1D datasets containing 10,000 data points. Accordingly, the input layer is supplied with signal information of size 10,000 × 2. The “10,000”

refers to the length of signal data while "2" stands for the two channels involved, the input signal channel (Channel-I) and output signal channel (Channel-O), as shown in Fig. 3.

The output layer of CNNs adopted in this paper is designed to estimate the travel time of shear wave for each given input-output signal pair so that it is constituted by one single output unit.

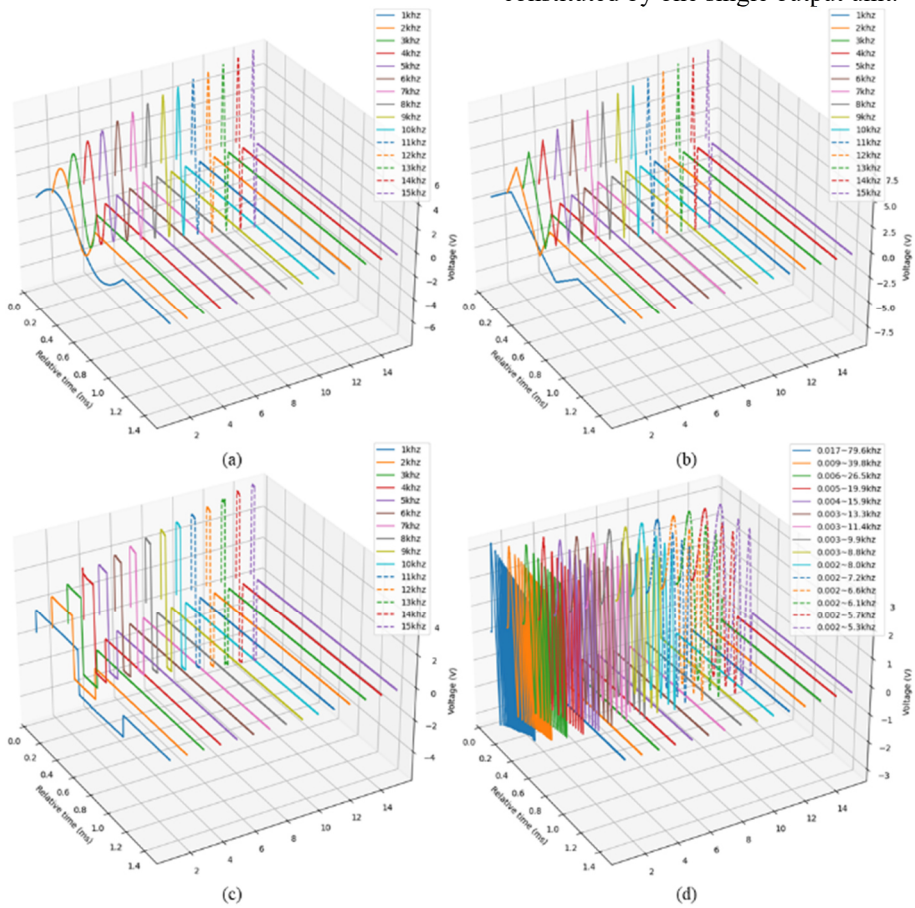


Figure 1. Schematic diagram of the input signals. (a): sin-wave signal, (b): triangle-wave signal, (c): square-wave signal, and (d): chirp-wave signal

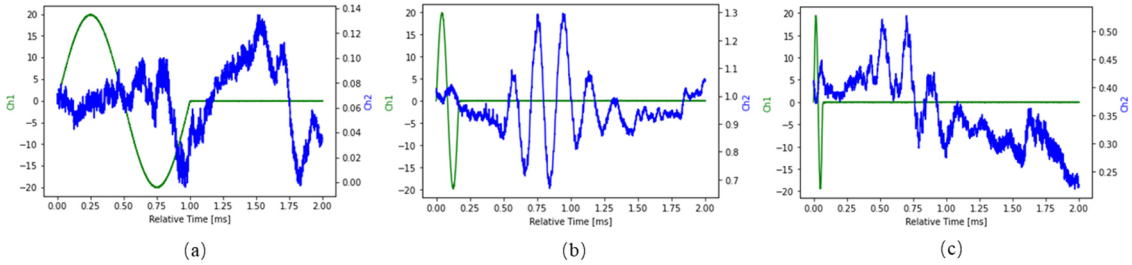


Figure 2. Comparison on G_{hv} signal data with different input frequencies. (a): 1 kHz, (b): 6 kHz, and (c): 15 kHz

3.2.2. Other layers

The main advantage of choosing an CNN algorithm has been the availability of the hidden layers to aid in the interpretation complex data sets over traditional ML methods (Tang et al. 2020). In general, the components of a CNN are the input layer, convolutional layer, pooling layer, activation layer, fully-connected layer and output layer, as shown in Fig. 4. Table 1 describes the hidden layers employed in the CNN algorithm for feature extraction and imputation. The exact implementation of the CNN can be found in Wenzhang (2022).

To evaluate the extent to which the predicted V_s of the CNNs differs from the *best-estimated* V_s, Mean Squared Errors (MSE) is adopted as the loss function for CNN training in this study.

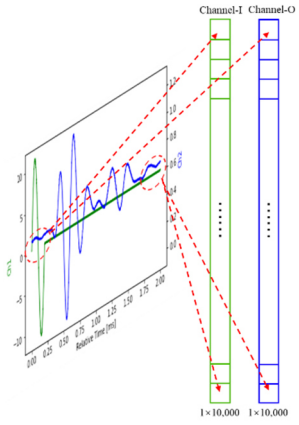


Figure 3. The composition of the input layer for CNN

3.3. Training of CNNs

Due to simple computation and high sparsity of generated output, rectified linear unit (ReLU) was one of the most notable activation functions in the last decade. This study employs ReLU as the activation function for CNN training since it works more efficiently on training than others, such as sigmoid and tanh (Maas, Hannun & Ng, 2013; Russakovsky et al., 2015; Zeiler et al., 2013). The mathematic expression of ReLU is defined as:

$$a_{i,k} = max(z_{i,k}, 0) \quad (3)$$

where $z_{i,k}$ is the input of the activation at location i on the k channel.

Overfitting is recognized as a key factor in the training and subsequent performance of the neural network. Dropout is an established method to overcome overfitting of data (Hinton et al., 2012). For the CNNs in this paper, dropout layers are employed behind each fully-connected layer to prevent overfitting. The output of dropout layer is given by

$$\boldsymbol{y} = \boldsymbol{r} * a(\boldsymbol{W}^T \boldsymbol{x}) \quad (4)$$

where $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$ is the input to the fully-connected layer, $\boldsymbol{W} \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{d}}$ is the weight matrix and $\boldsymbol{r} = [r_1, r_2, \ldots, r_d]$ is dropout vector whose elements satisfy Bernoulli distribution with parameter p ($r_i \sim Bernoulli(p)$) (Gu, J. et al., 2018).

The CNN algorithm is additionally optimised using an Adaptive Moment Estimation (Adam) as it has proven to effectively converge the less loss function as well as its relative insensitivity to initialized hyperparameters through the introduction of the 1st moment (mean) and the 2nd raw moment (uncentered variance) of the gradient. The Adam algorithm is given by:

$$m_t = \beta_1 \cdot m_{t-1} + (1-\beta_1) \cdot g_t \quad (5)$$

$$v_t = \beta_2 \cdot v_{t-1} + (1-\beta_2) \cdot g_t^2 \quad (6)$$

$$\widehat{m}_t = m_t/(1-\beta_1^t) \quad (7)$$

$$\hat{v}_t = v_t/(1-\beta_2^t) \quad (8)$$

$$\theta_t = \theta_{t-1} - lr \cdot \widehat{m}_t/(\sqrt{\hat{v}_t} + \epsilon) \quad (9)$$

where g_t represents the gradients with reference to the stochastic objective at timestep t (Kingma & Ba, 2014). m_t and v_t are biased first moment estimate and second raw moment estimate respectively, while $\widehat{m}_t$ and $\hat{v}_t$ are corrected estimate for m_t and v_t. To initialize moment vector, m_0=0 and v_0= 0. θ_t is the updated parameters at timestep t. β_1, $\beta_2 \in [0,1)$, are exponential decay rates for the moment estimates. In Eq. 9, lr is the learning rate and ϵ>0 is a constant and introduced to avoid zero denominator. A value of β_1=0.9, β_2=0.999 and ϵ=10^{-8} is typically recommended in practice, and was used in this study.

Batch normalisation (BN) layers are applied and attached behind the convolutional layers and are fully connected layers so that the weights and biases can be updated as expected, improving the overall performance of CNNs. Suppose the input layer has n dimension, *i.e.,* $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$, the normalized kth dimension shows as follows:

$$\hat{x}_k = (x_k - \mu_{\mathcal{B}})/\sqrt{\delta_{\mathcal{B}}^2 + \epsilon} \quad (10)$$

where the $\mu_{\mathcal{B}}$ and $\delta_{\mathcal{B}}^2$ are the mean and variance of the mini-batch, and ϵ is a constant. The normalized input $\hat{x}_k$ is further transformed into:

$$y_k = BN_{\gamma,\beta}(x_k) = \gamma \hat{x}_k + \beta \quad (11)$$

where γ and β are learned parameters. Compared with global data normalization, BN reduces the internal covariant shift as well as the dependence of gradients on the scale of the parameters and of their initial values which is beneficial to the gradient flow through the neural network (Gu et al., 2018).

As detailed previously, 4 sets of data are collected at the end of each consolidation stage. During the training of CNNs, 50% of the data samples are used for model training, 25% of data samples are for validation, allowing the real-time evaluation of the prediction accuracy and overfitting and simplifying the optimization of CNN. The remaining 25% is left for final testing to evaluate the performance of the entire CNN architecture.

By referring to the architecture of the VGG net (Simonyan & Zisserman, 2014), a trained CNN that meets the accuracy requirements has been developed.

As shown in Fig. 4, A pair of two-channel data formed by an input and output signal with a size 2 × 10,000 is fed into CNN algorithm. The data are processed by a series of convolution blocks, each comprising of two convolution layers and one attached pooling layer. The setting on filter size, stride, and padding of these layers, the data length is compressed down to 1 before being further processed by a fully-connected layer. In order to prevent feature loss on the signals during compression, the output channels were increased up to 1,024 to obtain more feature maps.

Table 1. Types and functions of employed hidden layers

Categories	Functions
Convolutional layer	Extracting and mapping features
Pooling layer	Sparsifying the feature map Reducing the amount of data computation Reducing overfitting and improving the fault tolerance of CNNs
Activation layer	Introducing non-linear mapping
Fully-connected layer	Refitting and reducing the loss of feature information
Dropout layer	Applying to fully-connected layer to reduce overfitting.
Batch normalisation layer	Preventing gradient dispersion and accelerating network training

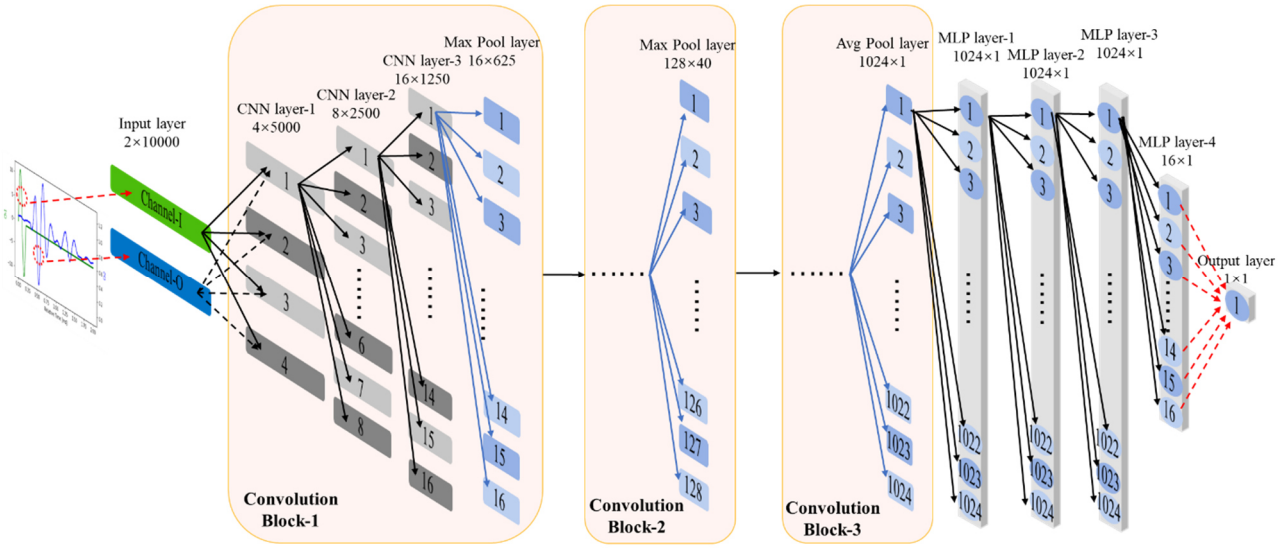


Figure 4. The architecture of the CNN

Finally, the fully-connected layer integrates the feature information and regresses to determine the output with respect to t_{tra}, and hence the G_0.

Table 2 lists the required hyperparameters for training CNNs, which are determined by trial and error until a steadily decreasing loss could be observed on the validation dataset during training. To monitor the training, the MSE is recorded at each five epochs, and the deviation between training and validation datasets is plotted. The value of the MSE indicates the degree of overfitting of the CNN. Fig. 5 plots the variance of MSE as the training epoch progresses. The training was terminated when the MSE on validation datasets was below 10^{-4}, as this is seen to be sufficient to satisfy the experimental requirements for accurate predictions.

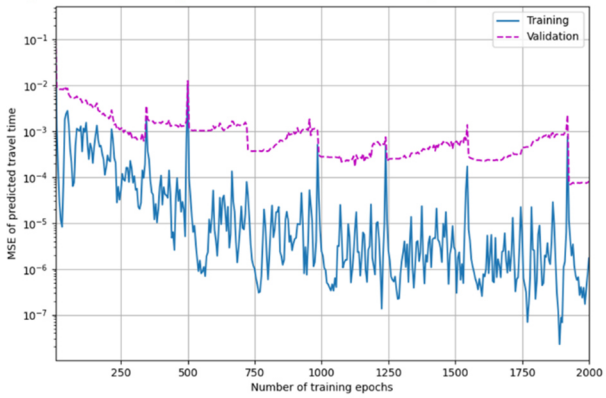


Figure 5. Variance in MSE with epoch progressing

Table 2. Hyperparameters employed for training CNNs

Batch size	Number of epochs	Learning rate
32	2,000	0.0003

4. Result analysis and discussion

4.1. Performance of trained CNN

As discussed in Section 3.1, a single sine wave of optimal frequency was used to determine first arrival, and hence V_s. The relative effectiveness of the signal type was assessed by comparing the predictions made by the CNN for the range of signals with that determined using the first arrival approach. An example of relative performance of the different input signals is given in Fig. 6 where the *best-estimated* G_0 and the G_0 calculated by the CNN-predicted V_s using the four different input signal waveforms. Results suggest a general agreement between the CNN-predicted and *best-estimated* G_0, despite small discrepancies observed at both lower (<3kHz) and higher (>12kHz) input frequencies independent of the input waveform. The square and chirp waves were however found to perform slightly better in these problematic ranges. A single CNN-predicted G_0 value was obtained as an average of determined values across a range of frequencies.

To further assess the reliability and consistency of CNN-predicted G_0, the coefficient of determination (r^2) and root mean squared error (RMSE) are used to evaluate the regression. A comparison of the CNN-predicted and *best-estimated* G_0 values is shown in Fig. 7 for the four different input waveforms. A total of 270 data points were collected for each of the four waveforms. All inferred r^2 values of the four signal types are greater than 0.99 with the RMSEs smaller than 3, indicating a strong correlation between CNN-predicted G_0 and *best-estimated* G_0.

4.2. Continuous prediction of G_0

As with other methods for travel time interpretation, CNNs can monitor the G_0 of a soil sample in real time during the consolidation process by analysing the generated and received shear wave data. A comparison of the CNN-predicted G_0 and *best-estimated* G_0 values in three measured orientations is plotted in Fig. 8 at the end of each consolidation stage. Each point represents an average value of predictions on all signal data collected. In all cases the CNN-predicted and *best-estimated* G_0 values are in good agreement. The variation of very small strain-stiffness with stress, as determined using the CNN algorithm, is shown to agree with the power function, as suggested by previous researchers (e.g. Hardin, 1978). Differences in between the G_{vh} and G_{hv} are likely to be a result of experimental variability.

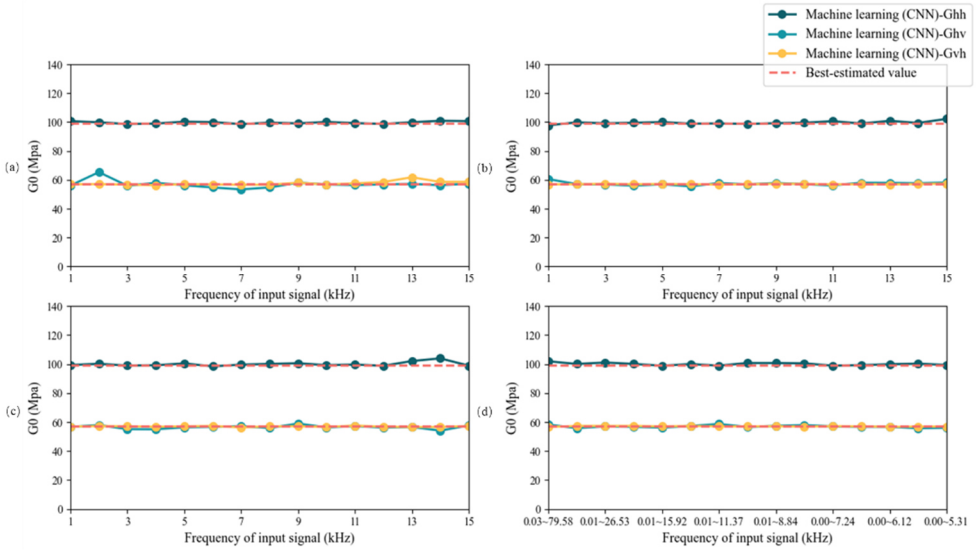


Figure 6. Comparison of CNN-predicted and *best-estimated* G_0 for four signal types at 150 kPa. (a): sin-wave signal, (b): square-wave signal, (c): triangle-wave signal, and (d): chirp-wave signal

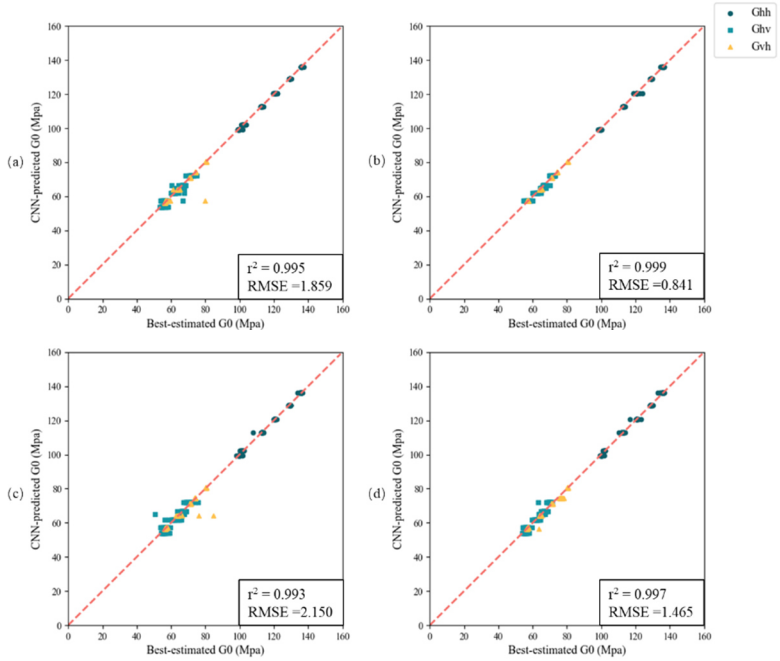
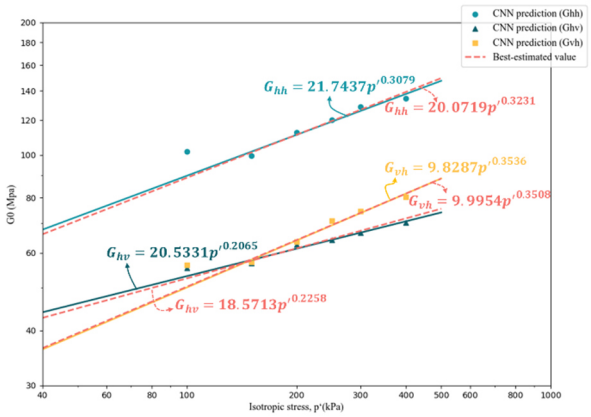


Figure 7. Reliability analysis on CNN-predicted G_0 for four signal types. (a): sin-wave signal, (b): square-wave signal, (c): triangle-wave signal, and (d): chirp-wave sign

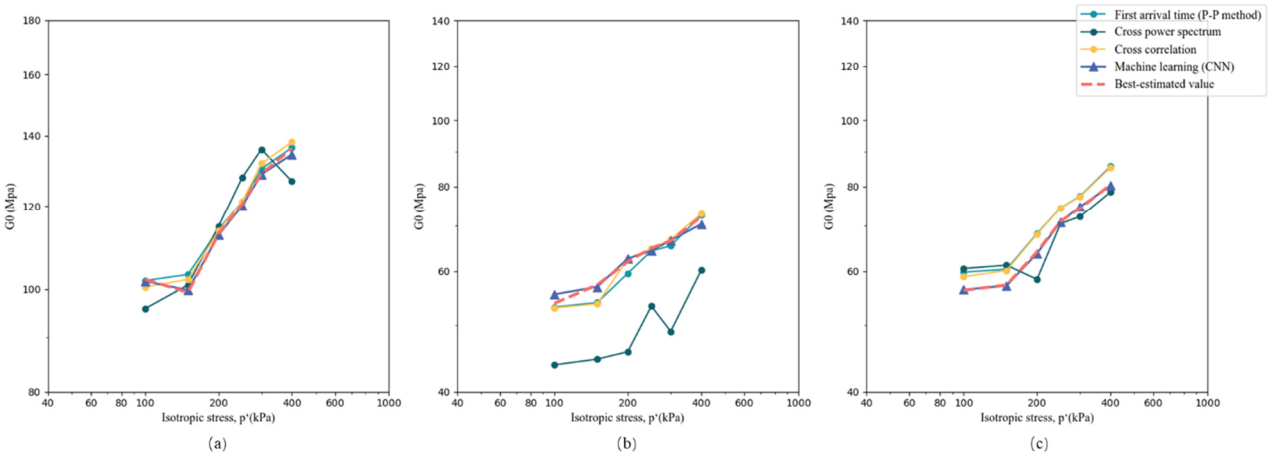


*Only signal data collected in 3 ~ 10 kHz are included

Figure 8. Variation of elastic stiffnesses G_{hh}, G_{vh}, and G_{hv} with stress

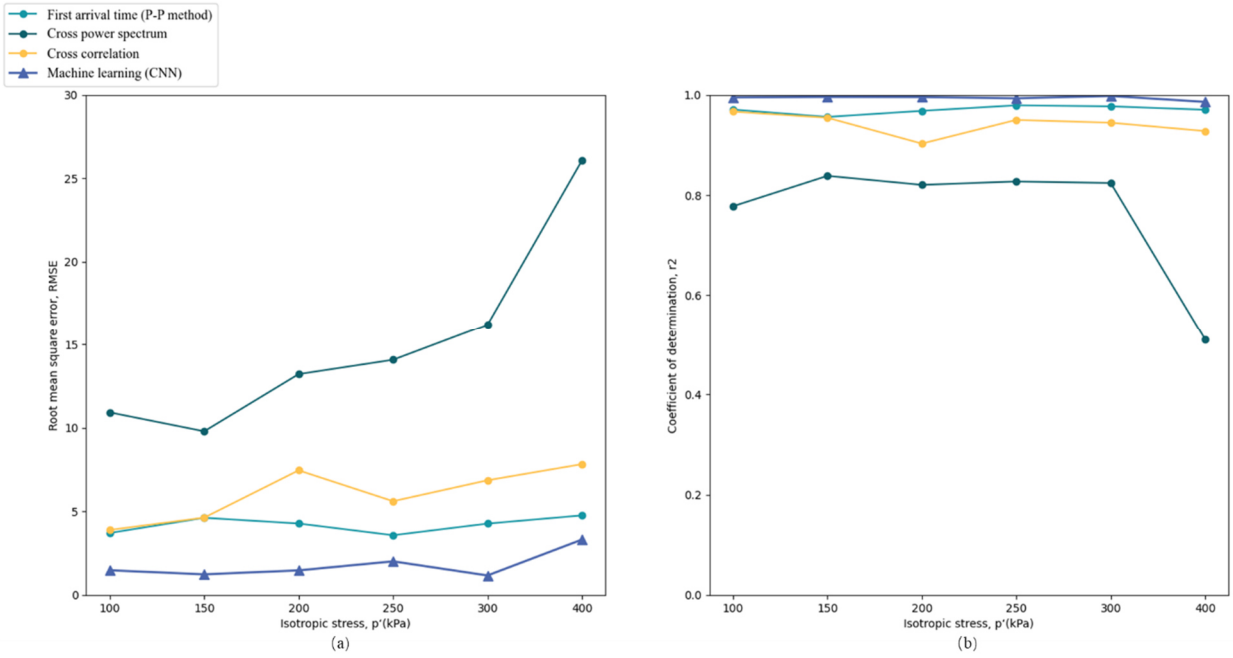
A comparison of the predicted G_0 values at the end of the various consolidation stages using the time domain, frequency domain, cross-correlation, *best-estimated* (manual first arrival) and CNN is made to assess the reliability of the various approaches in capturing the change of stiffness with stress. Fig. 9 (a), (b), and (c) show the variation in predicted shear modulus with stress for three measured orientations using the 5 methods. Similar to previous studies, good agreement can be found for the most approaches except for frequency domain method. As discussed by Viana da Fonseca et al. (2009), the frequency domain method is generally unreliable or require considerable user intervention to provide reasonable results and are therefore generally less successful when fully automated, such as in this study.

Fig. 10 (a) and (b) measure the performance of different methods at each consolidation stage precisely using the RMSE and r^2, respectively. The values of the RMSE and r^2 are the average values in three measured orientations. Assuming the *best-estimated* G_0 value an accurate representation of the soil stiffness, it can be said that the CNN-predicted G_0 value shows the greatest agreement of the available techniques as seen in the high r^2 and lower RMSE values. Further research on repeat samples and new soils is needed to accurately assess the method's potential to replace traditional interpretation methods.



*Only signal data collected in 3 ~ 10 kHz are included

Figure 9. Comparison between predictions of different interpretation methods, (a): G_{hh}, (b): G_{hv}, and (c): G_{vh}



*Only signal data collected in 3 ~ 10 kHz are included

Figure 10. Comparison on RMSE and r^2 of different interpretation methods, (a): RMSE, (b): r^2

5. Conclusion

Traditional methods for determining the shear wave velocity using bender elements are usually either time-consuming, unreliable or require considerable user intervention. While notable attempts have been made to overcome these limitations, considerable difficulty is still encountered in automating a reliable approach for determining the shear wave velocity from bender element tests. This paper presents a new automated method for interpreting input/output signals from bender elements using a convolutional neural network algorithm. With the proper training and validation, CNNs are able to detect the amplitude features behind the signal data and provide predictions for the shear wave velocity for given input-output signal pairs. The trained CNNs are validated to meet the accuracy requirements for practical application and show better consistency in its prediction than any other travel time interpretation methods. Some conclusions are summarised as follows:

a) CNNs can receive the input and output signals simultaneously for a given BE test and learn the patterns behind them to estimate travel time.
b) The trained CNNs can efficiently predict the travel time for a BE test, and its predictions reasonably agree with the results of other methods.

c) A higher consistency in travel time interpretation was obtained using CNNs when compared to other interpretation methods.
d) The accuracy of CNN-predicted travel times is sensitive to the frequency of input signals, and the level of such effect varies in different measured orientations.
e) More research is necessary to validate the performance of CNNs for repeat tests and other soil types.

References

Blewett, J., Blewett, I. J. and Woodward, P. K. 2000. "Phase and amplitude responses associated with the measurement of shear-wave velocity in sand by bender elements." In *Canadian Geotechnical Journal; Revue Canadienne De Géotechnique.* 37 (6), 1348-1357. https://doi.org/10.1139/t00-047

Dyvik, R. and Madshus, C. 1985. "Lab measurements of G_{max} using bender elements." In *Proc., ASCE Convention on Advances in the Art of Testing Soils under Cyclic Conditions,* 186-196. Available at: [https://cedb.asce.org/CEDBsearch/record.jsp?dockey=0046357]

Fernández Lavín, A. and Ovando Shelley, E. 2020 "Haar wavelet transform for arrival time identification in bender element tests." In *Geotechnical Testing Journal.* 43 (4), 937-949. http://doi.org/10.1520/GTJ20180400

Finas, M., Ali, H., Cascante, G. and Vanheeghe, P. 2016. "Automatic shear wave velocity estimation in bender element testing." In *Geotechnical Testing Journal.* 39 (4), 557-567. http://doi.org/10.1520/GTJ20140197

Greening, P. D. and Nash, D. F. T. 2004. "Frequency domain determination of G_0 using bender elements." In *Geotechnical Testing Journal.* 27 (3), 288-294. http://doi.org/10.1520/GTJ11192

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. and Chen, T. 2018. "Recent Advances in Convolutional Neural Networks." In *Pattern Recognition.* 77 354-377. http://doi.org/10.1016/j.patcog.2017.10.013

Gu, X., Yang, J., Huang, M. and Gao, G. 2015 "Bender element tests in dry and saturated sand: signal interpretation and result comparison." In *Soils and Foundations.*55 (5), 951-962. https://doi.org/10.1016/j.sandf.2015.09.002

Hardin, B. O. 1978. "The nature of stress-strain behaviour for soils." Proc. Geot. Div. Specialty Conf. on Earthquake Engineering and Soil Dynamics, Pasadena 1, 3-39. Available at: [https://trid.trb.org/view/74543]

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. R. 2012. "Improving neural networks by preventing co-adaptation of feature detectors". https://doi.org/10.48550/arXiv.1207.0580

Kawaguchi, T., Ogino, T., Yamashita, S. and Kawajiri, S. 2016. "Identification method for travel time based on the time domain technique in bender element tests on sandy and clayey soils." In *Soils and Foundations.* 56 (5), 937-946. http://doi.org/10.1016/j.sandf.2016.08.017

Kingma, D. P. & Ba, J. 2014. "Adam: A method for stochastic optimization." http://doi.org/arXiv:1412.6980

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M. and Inman, D. J. 2021. "1D Convolutional neural networks and applications: a survey." In *Mechanical Systems and Signal Processing.* 151 107398. http://doi.org/10.1016/j.ymssp.2020.107398.

Kiranyaz, S., Ince, T., Abdeljaber, O., Avci, O. and Gabbouj, M. 2019. "1-D Convolutional neural networks for signal processing applications." ICASSP, IEEE. pp.8360-8364. http://doi.org/10.1109/ICASSP.2019.8682194

Maas, A. L., Hannun, A. Y. and Ng, A. Y. 2013. "Rectifier nonlinearities improve neural network acoustic models." Proc. Icml, Volume 30. Available at: [https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.693.1422&rank=1&q=Rectifier%20nonlinearities%20improve%20neural%20network%20acoustic%20models&osm=&ossid=]

Ogino, T., Kawaguchi, T., Yamashita, S. and Kawajiri, S. 2015. "Measurement deviations for shear wave velocity of bender element test using time domain, cross-correlation, and frequency domain approaches." In *Soils and Foundations.* 55 (2), 329-342. https://doi.org/10.1016/j.sandf.2015.02.009

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C. and Fei-Fei, L. 2015. "ImageNet large scale visual recognition challenge." In *International Journal of Computer Vision.* 115 (3), 211-252. https://doi.org/10.1007/s11263-015-0816-y

Sanchez-Salinero, I., Roesset, J. M., Stokoe, I. I. and Kenneth, H. 1986. "Analytical studies of body wave propagation and attenuation", report GR 86-15. Civil Engineering Department, University of Texas at Austin. Available at: [https://apps.dtic.mil/sti/citations/ADA179487]

Shirley, D. J. and Hampton, L. D. 1978. "Shear-Wave Measurements in Laboratory Sediments." In *The Journal of the Acoustical Society of America.*63 (2), 607-613. https://doi.org/10.1121/1.381760

Simonyan, K. and Zisserman, A. 2014. "Very deep convolutional networks for large-scale image recognition." https://doi.org/10.48550/arXiv.1409.1556

Styler, M. A. and Howie, J. A. 2013. "Combined time and frequency domain approach to the interpretation of bender-element tests on sand." In *Geotechnical Testing Journal.* ASTM, 36 (5), 1-11. http://doi.org/10.1520/GTJ20120081.

Tang, W., Long, G., Liu, L., Zhou, T., Jiang, J. and Blumenstein, M., 2020. "Rethinking 1d-cnn for time series classification: A stronger baseline." https://doi.org/10.48550/arXiv.2002.10061

Viana da Fonseca, A., Ferreira, C. and Fahey, M. 2009. "A framework interpreting bender element tests, combining time-domain and frequency-domain methods." In *Geotechnical Testing Journal.* 32 (2), 1-17. http://doi.org/10.1520/GTJ100974

Viggiani, G. and Atkinson, J. H. 1995. "Interpretation of bender element tests." In *Géotechnique.* 45 (1), 149-154. http://doi.org/10.1680/geot.1995.45.1.149

Wenzhang, X. 2022. "Use of machine learning in determining Gmax from bender element tests." MSc dissertation. Department of Civil and Environmental Engineering, Imperial College London

Yamashita, S., Kawaguchi, T., Nakata, Y., Mikami, T., Fujiwara, T. and Shibuya, S. 2009 "Interpretation of international parallel test on the measurement of Gmax using bender elements". In *Soils and Foundations.* 49 (4), 631-650. http://doi.org/10.3208/sandf.49.631

Yang, J. and Gu, X. Q. 2013. "Shear stiffness of granular material at small strains: Does It Depend on Grain Size?" In *Géotechnique.* 63 (2), 165-179. https://doi.org/10.1680/geot.11.P.083

Zeiler, M. D., Ranzato, M., Monga, R., Mao, M., Yang, K., Le, Q. V., Nguyen, P., Senior, A., Vanhoucke, V. and Dean, J. 2013. "On rectified linear units for speech processing." 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE. pp.3517-3521. http://doi.org/10.1109/ICASSP.2013.6638312