deposits of the Shulin district in Taipei County, Taiwan. One out of the four drives namely drives C was considered here to assess the selected AI techniques, . The length for Drives C was 75 m. Overburden depth relative to the tunnel crown for Drives C was measured at 10.8 m. The pipejacking was undertaken using a slurry shield machine with a 1.5 m diameter cutterhead. The trailing concrete pipe of 1.44 m in diameter and 1 m in length ensured a 30 mm overcut was created in the annulus area. The self-weight of each pipe was 12.6 kN. A highly viscous lubricant with Marsh cone viscosity of 38 mins was injected into the overcut annulus, thereby reducing friction resistance to viscous resistance.

4.2 Engineering geology

Fig. 1 presents the geological profile as determined from four geological boreholes (BH1 – BH4) installed close to drives C. The phreatic surface was located at approximately 4.5 m below ordnance datum (BOD). Fig. 2 shows the soil properties profile as determined from both field and laboratory tests. In addition, Drives C was rammed in a ground predominantly composed of gravel and clayey gravel (Fig. 3). Additional information on the project is available in Cheng et al. (2017, 2018, 2019a,b).

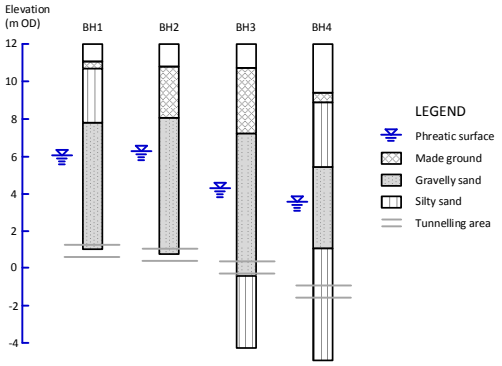

Fig. 1 Geological profile along the tunnel alignment

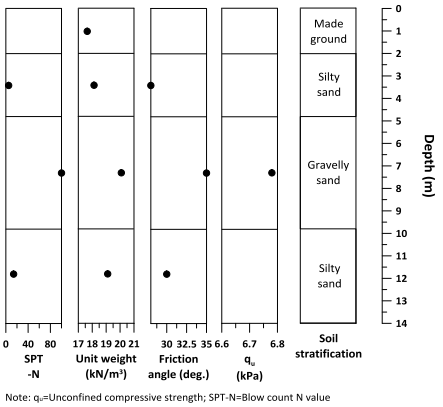

Fig. 2 Soil properties profile

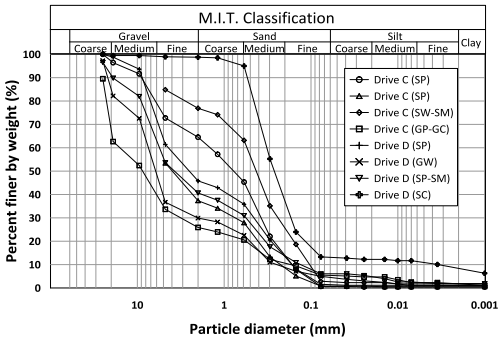

Fig. 3 Grain-size distribution curves for drive C

5. RESULTS AND DISCUSSION

5.1 Classification results

The three hyperparameter optimisation algorithms were first considered to evaluate their performance for this problem. To assess the feasibility of the optimisation algorithms, the fitness at each generation was traced, as shown in Fig. 4.

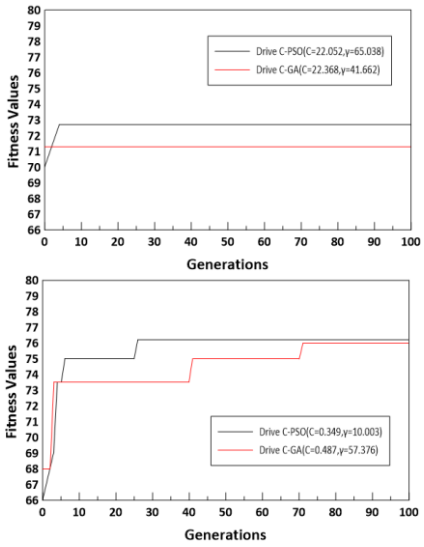

Fig. 4 Optimised results of the PSO and GA algorithms: (a) gravel identification and (b) clayey gravel identification

It can be observed that for gravel identification at drive C (Fig. 4a), the GA achieved an optimal solution immediately whereas for PSO it was achieved within four generations. For the identification of clayey gravel at drive C (Fig. 4b), GA required 71 generations to converge to a final solution whereas PSO required 26. In general, the PSO approach achieved faster convergence than GA. The wider the parameter range is, the more possibilities GS has of finding the best combination parameter. Notwithstanding that, GS is extremely time consuming especially when the number of possible different combinations of variables is rather high. Therefore, a grid search space ranging from 0.001 to 100 which is divided by a multiple of 10 was adopted here

to tackle the indicated issue, hence sacrificing the accuracy of prediction. The results of the parameter optimisation are tabulated in Table 1.

Table 1 Results of the parameter optimisation: (a) gravel identification and (b) clayey gravel identification

Hyperparameters used in identifying gravel at drive C	GA	PSO	GS
C	22.368	22.052	100
γ	41.662	65.038	100

Hyperparameters used in identifying clayey gravel at drive C	GA	PSO	GS
C	0.487	0.349	1
γ	57.376	10.003	1

The classification results present henceforth correspond to the optimised hyperparameters. We therefore use the more general terminology, 'SVM classifier', in the following sections. Further, we only discuss the cases that failed to correctly class the soil i.e. 'FN' in the residual-trend torque plot and 'FP' in the residual-trend jacking force plot. The results of the SVM classifier applied to the transformed data for drive C are shown in Fig. 5.

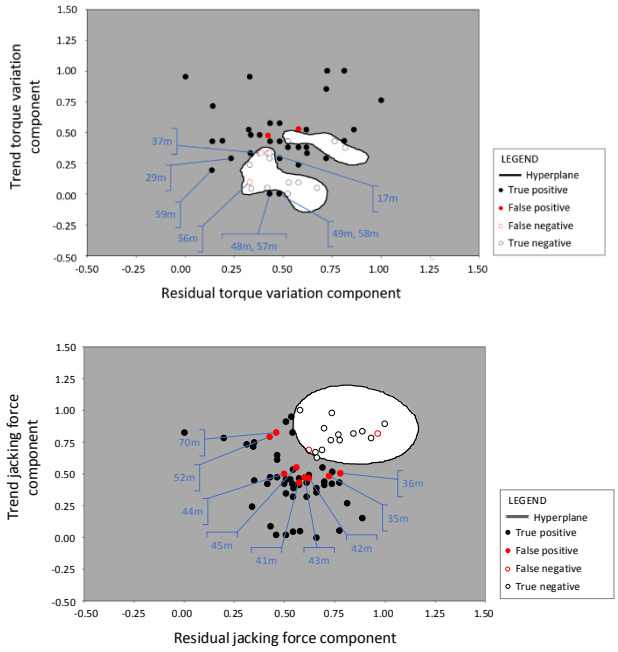

Fig. 5 Performance of the PSO-SVM model applied to drive C mapped to transformed parameter space: (a) identification of gravel (*C*=22.052, *γ*=65.038) and (b) identification of clayey gravel (*C*=0.349, *γ*=10.003)

Fig. 6 provides the reader with useful context in relation to the mapping of the TP, TN, FP and FN results from the transformed feature space back to the original (raw) parameter space. The SVM classifier provided excellent predictions with three FNs (see Fig. 5a). In this case, the FNs indicate that the gravels were erroneously classed as clayey gravels. Their locations in the original space are at jacked distances of 17 m, 37 m, and 56 m respectively. In contrast, several FPs appear in the predictions presented in Fig. 5b. FPs denote that the clayey gravels were erroneously identified as gravels. They appear at jacked distances of 35-36 m, 41-45 m, 52 m, and 70 m respectively.

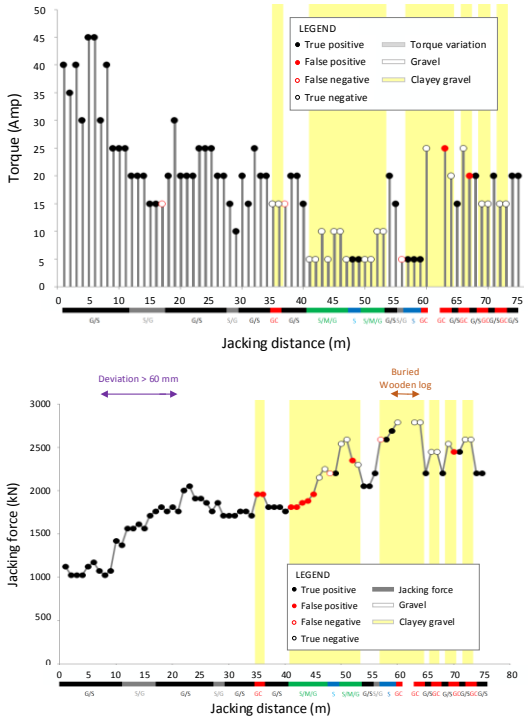

Fig. 6 Performance of the PSO-SVM model applied to drive C remapped back to initial parameter space: (a) identification of gravel (C=22.052, γ=65.038) and (b) identification of clayey gravel (C=0.349, γ=10.003)

5.2 Discussion

Most of the gravels present at drive C were successfully classed (Figs. 5a and 6a). The main cause to lead to the formation of the three FNs could be due to the fact that the occasional presence of sands, reported by Cheng et al. (2017), reduced their trend components Tt to nearly identical to or below 0.33 (16 Amp) (T_t > 0.33 can be classed as gravel) towards causing the misclassifications. Further, the clayey gravel is not frequently encountered along the tunnel alignment, which caused difficulties for the SVM classifier to define the hyperplane. However, the gravels at 29 m, 48-49 m, and 57-59 m jacking distances were correctly classed. The datapoints do not exactly correspond to gravels but sands with T_t below 0.33. The reason to explain why the

gravels can be correctly classed is that the lack of clayey gravel datapoints caused some difficulty in defining the hyperplane and the SVM classifier defined the noncontinuous boundaries by bypassing the sand datapoints in order to pursue improved predictive performance.

In contrast, nine FPs appear in the predictions in relation to identification of clayey gravel (Figs. 5b and 6b). A clayey gravel can be identified by satisfying two conditions; that are, the trend component $T_t > 0.58$ (2052 kN) and the residue component $R_t > 0.58$ (-5 kN). In spite that the FPs at 35-36 m and 41-45 m jacking distances were featured with $R_t > 0.58$, their T_t were below 0.58, induced by the gravels surrounding. Although the FPs at 52 m and 70 m jacking distances were featured with $T_t > 0.58$, the misclassifications occurred because of their $R_t < 0.58$, induced by traversing into the gravel from the clayey gravel. It is noteworthy that the predictions were surprisingly not affected by the effect of pipe deviation being greater than the threshold of 60 mm between 8 m and 21 m jacking distance.

It can be concluded that the discovery rate (DR) regarding identification of gravel being 94.1% at drives C, hence verifying the applicability of the trend component T_t, decomposed from the variation in cutterwheel torque, to class the gravel. The sands at drive C and striking buried wooden log led to misleading interferences about classification of gravel (inconsistent two regions) , as shown in Fig. 5a. Some FNs (i.e. 29 m, 48-49 m, and 57-59 m jacking distances at drive C) would have appeared if the SVM classifier did not pursue improved predictive performance by bypassing the datapoints. DR in relation to identification of clayey gravel was 59% at drives C, which also indicated a fairly good ability for the SVM classifier to class the clayey gravel. The clayey gravel can be successfully identified by higher T_t and R_t, in accordance with the established hyperplanes. In addition to the FPs induced by the reduced T_t, FP can also be induced by the reduced R_t resulting from jacking into the gravel from the clayey gravel.

The relative merits of the three optimisation algorithms were evaluated. Here the DRs and FARs determined using the three hyperparameter optimisation algorithms are summarised in Table 2.

Table 2 Performance of the optimisation algorithms: (a) gravel identification and (b) clayey gravel identification

Drive	Gravel identification					
	Discovery rate, DR (%)			False alarm rate, FAR (%)		
	GS	GA	PSO	GS	GA	PSO
C	94.1	92.2	94.1	9.1	18.2	9.1

Drive	Clayey gravel identification					
	Discovery rate, DR (%)			False alarm rate, FAR (%)		
	GS	GA	PSO	GS	GA	PSO
C	50.0	50.0	59.0	7.8	3.9	3.9

The PSO algorithm outperformed the other two algorithms. Although the PSO algorithm is more prone to becoming ‘trapped’ in local optima, it provided the best

balance between exploration and exploitation tendencies. In contrast, the GS algorithm provided the worst performance. The present analyses revealed that while the GS optimization was robust for the identification of optimal hyperparameter combinations, the required computational time was excessive. It is therefore highly reliable but suitable for only low dimensional datasets.

6. CONCLUSIONS

This paper has examined the potential for the use of AI techniques to identify geological conditions encountered during pipejacking. A selection of the most popular parameter optimisation algorithms was considered to improve the accuracy and efficiency of the AI predictions, namely genetic algorithms, particle swarm optimisation and grid search. Based on the results and discussion, some main conclusions can be drawn as follows:

(1) Decomposition of the data was implemented to transform the cutterwheel torque-jacking distance relationship and the jacking force-jacking distance relationship into feature-based sub-series for direct application of the SVM classifier. The optimal features were found to be the trend and residual components of both the cutterwheel torque and the total jacking force; the trend component dominated the identification of gravel, whereas the trend component T_t and the residual component R_t controlled the identification of clayey gravel. Further, the particle swarm optimisation algorithm provided the best performance due to its best balance between exploration and exploitation tendencies.

(2) Clayey gravels were not encountered frequently during pipejacking at drive C which caused some difficulty in defining the hyperplane. In addition, the surrounding geology could have implications on Tt, leading to false negatives. Some datapoints with low T_t did not cause false negatives because the SVM classifier bypassed them in order to pursue improved predictive performance. Similarly, surrounding geology could also have implications on T_t and R_t producing false positives. Further selection of highly-discriminant features, while classing the silt and/or sand, is deemed necessary to improve the accuracy of prediction.

(3) The performance of the optimisation algorithms was assessed using four performance pressures, namely, TP, FN, FP and TN, using the DR and the FAR indices. The PSO algorithm performed the best amongst the three optimisation algorithms because it provided the best balance between exploration and exploitation tendencies. While the GS sacrified the accuracy of predictions with a grid of low fineness.

REFERENCES

Breiman, L. (2001), “Random forests,” *Machine Learning*, **45**(1), 5-32.

Persons, W.M. (1919), *Indices of Business Conditions: An Index of General Business Conditions*, Harvard University Press.

Cleveland, R., Cleveland, W., McRae, J., Terpenning, I. (1990), “STL: a seasonal-trend decomposition procedure based on loess,” *Journal of Official Statistics*, **6**, 3-73.

Pellet-Beaucour, A.L. and Kastner, R. (2002), "Experimental and analytical study of friction forces during microtunneling operations," *Tunnelling and Underground Space Technology*, **17**(1), 83-97.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J. (2011), "Scikit-learn: Machine learning in Python," *Journal of machine learning research*, **12**(Oct), 2825-2830.

Masters, T. (1993), *Practical neural network recipes in C++*, Academic Press, San Diego, CA.

Kennedy, J. and Eberhart, R. (1995), "Particle Swarm Optimization," *Proceedings of IEEE International Conference on Neural Networks*, **4**, 1942-1948.

Cheng, W.C., Ni, J.C., Shen, J.S.L. and Huang, H.W. (2017), "Investigation into factors affecting jacking force: a case study," *Proceedings of the Institution of Civil Engineers - Geotechnical Engineering*, **170**(4), 322-334.

Cheng, W.C., Ni, J.C., Arulrajah, A. and Huang, H.W. (2018), "A simple approach for characterising tunnel bore conditions based upon pipe-jacking data," *Tunnelling and Underground Space Technology*, **71**, 494-504.

Cheng, W.C., Ni, J.C., Huang, H.W., Shen, J.S. (2019a), "The use of tunnelling parameters and spoil characteristics to assess soil types: a case study from alluvial deposits at a pipejacking project site," *Bulletin of Engineering Geology and the Environment*, **78**(4), 2933-2942.

Cheng, W.C., Wang, L., Xue, Z.F., Ni, J.C., Rahman, M., Arulrajah, A. (2019b), "Lubrication performance of pipejacking in soft alluvial deposits,' *Tunnelling and Underground Space Technology*, **91**, 102991.