

# PIPEJACKING CLOGGING DETECTION IN SOFT ALLUVIAL DEPOSITS USING MACHINE LEARNING ALGORITHMS

Xue-Dong Bai<sup>1</sup>, Wen-Chieh Cheng<sup>2,3</sup>, Brian B. Sheil<sup>4</sup> and Ge Li<sup>1</sup>

<sup>1</sup>PhD student, School of Civil Engineering, Xi'an University of Architecture and Technology,  
Xi'an 710055, China. Email: baixuedong@xauat.edu.cn (X.-D. Bai), lige@xauat.edu.cn (G.  
Li)

<sup>2</sup>Professor, School of Civil Engineering, Xi'an University of Architecture and Technology,  
Xi'an 710055, China. Email: w-c.cheng@xauat.edu.cn (W.-C. Cheng)

<sup>3</sup>Shaanxi Key Laboratory of Geotechnical and Underground Space Engineering (XAUAT),  
Xi'an 710055, China

<sup>4</sup>RAEng Research Fellow, Department of Engineering Science, University of Oxford, U.K.  
Email: brian.sheil@eng.ox.ac.uk (B.-B. Sheil)

Revision submission,

Main text word count: 7413

Tables: 3

Figures: 19

## ABSTRACT

'Clogging' is a common issue encountered during tunnelling in clayey soils which can impede tunnel excavation, cause unplanned downtimes and lead to significant additional project costs. Clogging can result in a drastic reduction in performance due to reduced jacking speeds and the time needed for cleaning if it cannot be fully mitigated. The data acquired by modern tunnel boring machines (TBMs) have grown significantly in recent years presenting a substantial opportunity for the application of data-driven artificial intelligence (AI) techniques. In this study, a baseline assessment of clogging in slurry-supported pipejacking is performed using a combination of TBM parameters and the semi-empirical diagram proposed in the literature. The potential for one-class support vector machines (OCSVM), isolation forest (IForest) and robust covariance (Robcov) to assess the tendency to clayey clogging is then explored in this work. The proposed approach is applied to a pipejacking case history in Taipei, Taiwan, involving tunnelling in soft alluvial deposit. The results highlight an exciting potential for the use of OCSVM, IForest and Robcov to detect clogging during slurry-supported pipejacking.

**KEYWORDS:** support slurry; pipejacking; clayey clogging; jacking speed; cutterwheel torque

## INTRODUCTION

Fine-grained soils have a strong influence on pipejacking performance due to their tendency to trigger different issues (Tan and Wei 2012; Ong and Choo 2016; Soomro et al. 2020; Zhang et al. 2020a), one of which is 'clogging'. Clogging denotes the adherence of fine-grained soils to cutters at the cutterhead, openings on the cutting wheel, screw conveyor and/or conveyor belt. Clogging can therefore cause unplanned downtimes and, consequently, a significant increase in operation costs (Thewes and Burger 2005; Spagnoli et al. 2011a; Heuser et al. 2012; Hollmann and Thewes 2013; Thewes and Hollmann 2014; Zumsteg et al. 2016). For pipejacking, clogging can be described as the attraction between soil particles and cutters and the adhesion between water in the soil and the cutters (Fountaine 1954; Sass and Burbaum 2008; Kang et al. 2018, 2019). There are four potential mechanisms governing adhesion of clay to a cutter, namely adhesion of clay particles on a cutter surface, inherent cohesion, bridging of clay particles over a cutting wheel opening, and an inability for the clay to dissolve in water (Thewes 1999; Jia 2004; Kang et al. 2019). Previous research has shown that clogging caused by adhesion of clay to a cutter can significantly reduce shield tunnelling performance (van Baalen 1999, 2001; Spagnoli et al. 2011b, 2012a,b, 2014; Zhang et al. 2018).

Various approaches have been proposed to evaluate the potential for clogging to occur such as the use of plasticity index measurements, semi-empirical diagrams, and laboratory-based drilling tests. Hu and Rostami (2020) described the importance of soil conditioning during tunnelling and the role of soil rheology in tuning the desired characteristics of the conditioned soil. Using a novel device, those authors established a relationship between soil rheological parameters, soil type and conditioning parameters for soft ground tunnelling. Using a new framework and new devices, Peila et al. (2015) noted that the effectiveness of a polymer in clay conditioning is strongly dependent on the plasticity index of the clay. For low plasticity clay, the use of polymers can cause an increase in the volume of foam needed because of the water absorption effect of the polymer itself. However, this can also lead to a more homogeneous conditioned soil with long-lasting mechanical properties. Alberto-Hernandez et al. (2017) used the relationship between cohesion (soil-soil strength) and adhesion (soil-structure strength) to evaluate clogging potential, though this method is limited to situations where soil cohesion is greater than adhesion. Hollmann and Thewes (2013) reported relevant factors for the development of clogging and presented a new classification diagram which allows for the quantification of changes in the water content towards estimating changes in the consistency of fine-grained soils under varying availabilities of water. Thewes and Hollmann (2016) explored the risk of clogging in various ground conditions and for different shield types and presented a summary of methods to characterise soil 'stickiness' and

laboratory experiments to assess clogging potential. A newly developed diagram for assessing clogging risks for all types of shields and a new testing scheme for evaluating sedimentary rocks in terms of clogging were also introduced. Feinendegen et al. (2010) recommended a cone pull-out test to detect the adhesion/clogging propensity of a rock or soil, combined with a newly developed scheme for classifying the clogging potential. Following an extended test campaign using soils with different clay contents and minerals, de Oliveira (2018) developed a new device which adds to the first method a kinetic energy impulse via dropping of the beater from a certain height. This combination could give a more reliable evaluation of the potential for clogging to occur along earth pressure balance (EPB) machine tunnel drives. Further, a laboratory routine to characterise the clogging and fluidity of soils, including mixed soils by considering different fractions of clay was proposed by de Oliveira (2019a,b,c). But this routine can still be improved by doing this exercise of preliminary assessment and later backanalysis.

Kang et al. (2019) evaluated the clogging potential of mixed bentonite-kaolin specimens using a combination of the semi-empirical diagram and the drilling test. Kang et al. (2019) also investigated the dependence of the clogging potential on the plasticity indices of the mixtures. The results revealed that mixtures with bentonite had a higher clogging potential than pure kaolin and the drilling tests proved an effective means of quantifying clogging potential and evaluating the performance of additives. Zumsteg et al. (2016) investigated the effect of clay mineralogy and the composition of the supporting slurry on clogging potential using novel 'stickiness' tests including a 'mixing test' and a model tunnel boring machine (TBM) cutterhead test. It was reported that increased slurry strength is likely to increase the possibility of clogging for mixed face conditions. In contrast, polymer additives to slurry can achieve both high slurry resistance with low clogging potential by protecting clay aggregate surface from penetration of water (Zumsteg et al. 2016).

Although previous studies performed over the past decade have greatly enhanced our understanding of the clogging process and the role of key influencing factors (Zumsteg et al. 2013; Ryu et al. 2019), the influence of clogging on the response of tunnelling parameters (e.g. cutterwheel torque and jacking speed) remains unclear. Further, the prevention and mitigation of clogging during shield tunnelling is still highly dependent on the operator's accumulated site experience. The objectives of this paper are: (1) to characterise the response of tunnelling parameters to the development of clogging during slurry-supported pipejacking, (2) to evaluate the clogging potential using a combination of the existing pipejacking parameters and the semi-empirical approach, and (3) to explore the feasibility for artificial intelligence techniques to evaluate the clogging potential.

## CRITICAL FACTORS FOR THE DEVELOPMENT OF CLOGGING

### *Particle-size distribution*

The shear strength of coarse-grained soils (with less than 5% fines) is derived from interparticle friction and geometrical interference (e.g. dilation, particle crushing, and rearrangement) rather than cohesion and they, therefore, minimises clogging issues during tunnel excavation. In contrast, coarse-grained soils (e.g. sands and gravel) with fines content  $> 5\%$  have been shown to increase the potential for clogging (Ni and Cheng 2010). In this context, soil particles  $> 0.075$  mm become part of the clogging material (i.e. causing the blockage), whereas particles  $< 0.075$  mm become 'abrasive' material (i.e. leading the wear of cutting tools). On the other hand, fine-grained soils always promote clogging and a higher concentration of fines content in the slurry during pipejacking (Tombacz and Szekeres 2006). Thus, a comprehensive assessment of the particle size distribution is a prerequisite for a rigorous evaluation of clogging potential.

### *Soil plasticity and consistency*

Three water contents typically used to define the state of fine-grained clayey soils are the natural water content  $\omega_n$ , liquid limit  $\omega_L$  and plastic limit  $\omega_P$ . An increasingly popular means of evaluating clogging potential is through a combination of the plasticity index  $I_p (= \omega_L - \omega_P)$  and the consistency index  $I_c = (\omega_L - \omega_n)/(\omega_L - \omega_P)$ . According to Thewes (1999), clayey soils with  $I_p > 20\%$  and  $I_c = 0.75-1.25$  have the highest potential to cause clogging. However, more recent investigations have shown that extensive clogging can even occur in clayey soils with  $I_c = 1.25-1.50$  (Thewes 1999). Hollmann and Thewes (2013) examined 150 samples of "sticky" material obtained from open and slurry-supported shield tunnelling projects. The samples varied between very soft (23% of samples;  $I_c = 0.4-0.5$ ), soft-medium (58%;  $I_c = 0.5-0.75$ ) and stiff (19%;  $I_c = 0.75-1.0$ ) consistency. Laboratory test results reported by Feinendegen et al. (2011) show good agreement with those documented by Hollmann and Thewes (2013). Encountering clayey soils with "soft" consistency during tunnelling is, therefore considered high-risk for the development of clogging. In contrast, clayey soils with stiff consistency is likely to show relatively low clogging potential because of their lower  $\omega_n$ , whereas clayey soils with very soft consistency also have relatively low clogging potential since their reduced shear strength is not high enough to resist the shear stresses exerted in the excavation chamber.

### *Free water*

Free water, including groundwater inflow and water contained in bentonite slurry and the soil conditioning agent, also plays an important role in clogging potential during tunnel excavation. It is possible for cohesive soil with clay content  $> 10\%$  to be transformed into clogging material. However, the time taken to achieve this transformation is closely related to the natural soil

consistency and the availability of free water which, in turn, is dependent on the type of tunnel excavation method. In slurry-supported shield tunnelling, the large amount of free water (due to the use of bentonite slurry) leads to a higher clogging potential in clayey soils. The authors' experiences are that even very stiff consistency clays can eventually be transformed to a soft consistency. In contrast, in the case of open shield tunnelling the amount of groundwater depends on both the inflow rate and construction downtime. Previously sticky clays in the excavation chamber could transform into non-sticky material when their adhesion reduces over time. Thus, it can be realised that the availability of free water must also be considered when assessing the potential for clogging to occur during shield tunnelling.

## **ASSESSMENT OF CLOGGING POTENTIAL**

### *Methodology*

In this work, three slurry-supported pipejacking drives undertaken in the soft alluvial deposits at Taipei, Taiwan were analysed. Following acquisition of the soil particle size distribution, plasticity  $I_p$  and consistency index  $I_c$ , a baseline assessment of the clogging potential was completed using existing pipejacking records, aided by the semi-empirical approach (Hollmann and Thewes 2013). The time history of pipejacking performance parameters are typically non-stationary which can lead to difficulties in the identification of patterns in the data. Decomposition procedures can be used to alleviate this problem by disaggregating time series data into feature-based sub-series where a weighted moving average dominates data features retained in analysis and eliminates noises.

The large volume of data extracted by modern TBMs presents a substantial opportunity for the application of data-driven anomaly detection (AD) techniques to identify patterns in the data. Instead of simulating the measured system response, AD approaches accentuate characteristics of the system by utilising information extracted from the measured data. Evolutionary Polynomial Regression (EPR) has been drawn more attention because of its more powerful ability in searching the target expression than artificial neural networks (ANNs) and genetic programming (GP) (Gurocak et al. 2012; Alemdag et al. 2016; Yin et al. 2016). Usually the global search for the best expression of the EPR equation is conducted using a genetic algorithm (GA) over the values contained in the user defined vector of components. For high dimensional problems, GAs, however, require a long period of computational time and large memory (Deep and Thakur 2007a,b). Multivariate Adaptive Regression Splines (MARS) primarily aims to organise relationships between a set of input variables and the target dependent (Zhang et al. 2020c). MARS is a nonparametric statistical method based upon 'divide and conquer' strategy where the training datasets are partitioned into separate

piecewise linear splines of different gradients towards representing the integration of additive regression, recursive regression, spline regression and recursive partitioning regression. Compared with the regression AD algorithms, ‘unsupervised’ clustering AD algorithms are more appropriate to process unlabelled data since the measured data (i.e. the outputs) are not labelled in most of practical problems. Unsupervised clustering AD algorithms, applied to infer characteristics in data without reference to known labels, are popular for this reason and employed in this work. A myriad of approaches have been proposed in the literature (Zhang et al. 2020b,c); three of the more popular AD approaches are adopted for this study: (1) one-class support vector machines (OCSVM), (2) isolation forest (IForest) and (3) robust covariance (Robcov). The three AD methods are applied to all case histories where clogging was encountered to assess their ability to identify clogging behaviour during pipejacking. Baseline assessments are also undertaken to benchmark predictions determined using the AD techniques.

#### *One-class support vector machines*

Traditionally many classification problems attempt to solve the two or multi-class situation. The goal of the machine learning task is to discriminate between subclasses of a dataset using a ‘training’ portion of the data. One-class support vector machine (OCSVM) denotes the case where the data comprises only one class and the task is to identify whether new measurements belong to that class. Schölkopf et al. (2001) framed the OCSVM approach by considering the origin as the only member of the second class (Sheil et al. 2020). A hyperplane is constructed in the feature space to separate the dataset from the origin, using a maximal margin (Fig. 1). The hyperplane constructs the following classification rule:

$$f(x) = (\mathbf{w} \cdot x) + b \quad (1)$$

where  $\mathbf{w}$  = adjustable weight vector and  $b$  = bias. The objective is to obtain a function that gives the maximum margin between the data and the origin. To prevent the OCSVM classifier from overfitting, slack variables  $\xi_i$  are used to create a ‘soft margin’ which allows some datapoints to lie within the margin. The optimal separating hyperplane can be obtained by solving the following convex quadratic optimisation problem (Vapnik 1995):

$$\begin{aligned} \text{Minimise} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{\nu n} \sum_{i=1}^n \xi_i - \rho, \quad i = 1, 2, \dots, n \\ \text{Subject to} \quad & (\mathbf{w} \cdot \Phi(x_i)) \geq \rho - \xi_i, \quad i = 1, 2, \dots, n \quad \xi_i \geq 0 \end{aligned} \quad (2)$$

where  $n$  = number of observations,  $\rho$  = margin, and  $\xi_i$  = slack variable, which is penalised in the objective function for nonzero values. The  $\|\mathbf{w}\|^2$  term in Eq. (2) is an L2 regularisation term

to minimise overfitting, and its relative importance is controlled by the parameter  $\nu$ . As the outcome of the decision function relies only on the dot-product of the vectors in the feature space, an explicit mapping to the feature space is not necessary. The ‘kernel trick’ is often adopted which allows the dot-product to be substituted by kernel functions. The decision function (i.e. classification rule) for a datapoint  $x$  can therefore be written as:

$$f(x) = \text{sign} \sum_{i=1}^{N_{sv}} \alpha_i K(x_i, x) - \rho \quad (3)$$

where  $\alpha_i$  = Lagrange multiplier and  $N_{sv}$  = number of support vectors. Every  $\alpha_i$  which is  $> 0$  is weighted in the decision function to ‘support’ the vector machine. Popular selections for the kernel include linear, polynomial, radial basis function (RBF) and sigmoid. RBF is the kernel that was selected here:

$$K(x_i, x) = \exp \left( -\gamma \|x_i - x_j\|^2 \right) \quad (4)$$

where  $\gamma$  = kernel parameter, which is the width of the RBF and typically varies between 0 and 1, and  $\|x_i - x_j\|$  = dissimilarity measure.

#### *Isolation forest*

The IForest is a non-parametric method that constructs classification models in the form of a ‘tree’ structure. An isolation forest, as the name suggests, is an ensemble learning method that operates by constructing  $n_{tree}$  isolation trees and aggregating the results, as illustrated in Fig. 2. The IForest approach detects anomalies by randomly partitioning the data. Partitions are created by randomly selecting a feature and then randomly producing a split value between the maximum and the minimum value of the feature. Partition creation continues until all datapoints are isolated; in most cases a limit is placed on the maximum number of partitions. Multiple training datasets are produced by sampling with replacement randomly from the original dataset, and anomalies are ultimately identified by sorting datapoints according to their corresponding path lengths (Sheil et al. 2020). For a dataset of size  $n$ , the average,  $c(n)$ , of each path length,  $h(x)$ , is calculated as:

$$c(n) = 2H(n-1) - \left( \frac{2(n-1)}{n} \right) \quad (5)$$

where  $H(i)$  = harmonic number (i.e.  $H(i) = \ln(i) + a$ , where  $a$  is Euler-Mascheroni constant). The IForest anomaly score (IF) is then defined as:

$$IF(x, n) = 2^{-\frac{E(h(x))}{c(n)}} \quad (6)$$



where  $E(h(x))$  = average of  $h(x)$  from a collection of isolation trees. A value of  $IF > 1.0$  is classified as anomalous in this study (Liu et al. 2012).

#### *Robust covariance*

The robust covariance ('Robcov') approach was first proposed by Rousseeuw (1984). The concept is to find a given proportion of 'good' observations which are not outliers and compute their empirical covariance matrix. Then this empirical covariance matrix is rescaled towards compensating the performed selection of observations (namely, consistency step). Having computed the Minimum Covariance Determinant (MCD) estimator, one can give weights to observations according to their Mahalanobis distance  $md$ , leading to a reweighed estimate of the covariance matrix of the dataset (namely, reweighting step). For data following a Gaussian distribution, the distance of an observation  $x_i$  to the mode of the distribution can be computed using its  $md$  as follows:

$$md(\mu, \Sigma)(x_i)^2 = (x_i - \mu)' \Sigma^{-1} (x_i - \mu) \quad (7)$$

where  $\mu$  = location of the underlying Gaussian distribution and  $\Sigma$  = covariance of the underlying Gaussian distribution. As the usual covariance maximum likelihood estimate is very sensitive to the presence of outliers in the data, a more robust estimator of covariance is required to minimise the influence of 'erroneous' observations such that  $md$  accurately reflect the true organisation of the observations (Fig. 3). Larger values of  $md$  denote anomalous observations.

#### *Feature selection*

Decomposition techniques, proposed by Persons (1919), isolate salient features of a dataset (e.g. trend (seasonal) component and periodic component). Seasonal-trend decomposition using Loess smoothing (STL; Cleveland et al. 1990) is one of the most popular decomposition techniques and is also used here to partition the global series into three additive components as follows:

$$y_t = P_t + T_t + R_t \quad (8)$$

where  $P_t$  = periodic component,  $T_t$  = trend component and  $R_t$  = residual component. In this work, two feature variables are considered for the application of OCSVM, IForest and Robcov, namely, the residual and trend components of the density of support slurry  $\rho$ , the cutterwheel torque  $T_c$ , and the jacking speed  $V$ . The periodic component  $P_t$  was derived based upon a minimum length of 3 m (i.e. three datapoints). Since the current supplied to the cutterwheel of shield machine is used as a proxy for cutterwheel torque, the residual and trend torque components are, therefore, plotted in 'Amps'.

## IMPLEMENTATION

### *Data screening*

As all case histories involved periods of tunnelling in gravel, the corresponding datapoints were first filtered out of the datasets. The authors' previous research suggests  $T_c > 15$  Amp and  $V \geq 100$  r/min correspond to tunnelling in gravel and they may vary between projects with different geology (Cheng et al. 2017). These data were therefore filtered out from the dataset such that all datapoints pertain to tunnelling in clayey soil. The data was logged at 2 m intervals of jacked distance to capture the clogging-induced decline in  $V$ .

### *Data scaling*

The OCSVM, IForest and Robcov algorithms were implemented using the Python module *Scikit-learn* (Pedregosa et al. 2011). All data were scaled using a 'min-max scaler' such that each feature varies between 0 and 1 (Masters 1993). Thus, given a set of input data  $x_1, x_2, \dots, x_n$ , the scaled dataset  $z_1, z_2, \dots, z_n$  will be:

$$z_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (9)$$

where *min* and *max* are our specified minimum and maximum values of the range to scale. This pre-processing maximises the efficiency and performance of the learning process and equalises the importance of the input dataset.

### *Confusion matrix*

There are four possible outcomes from the analysis, commonly summarised in a 'confusion matrix': (a) true positive (TP – system correctly identified clogging behaviour), (b) false negatives (FN – system has failed to detect clogging behaviour), (c) false positives (FP – system has erroneously detected clogging behaviour) and (d) true negatives (TN). The four outcomes were considered here to evaluate the performance of the AD approaches using the discovery rate (DR), or true positive rate, and the false alarm rate (FAR), or false negative rate.

$$DR = \frac{TP}{TP + FN} \quad (10)$$

$$FAR = \frac{FP}{TN + FP} \quad (11)$$

## PIPEJACKING DRIVES

## *Project overview*

Fig. 4a shows the location of four drives in the soft alluvial deposits located in the Shulin district in Taipei County, Taiwan. Due to the data completeness, three (drives B, C and D) out of the four drives were selected for the present analyses and measured 126 m, 75 m and 102 m in length respectively and were located approximately 10.5 m below ground level. The tunnelling was undertaken using a slurry-supported shield machine with a 1.5 m diameter cutterhead. A 30 mm overcut in the annulus area was created using the 1 m long, 1.44 m wide trailing concrete pipe. Each pipe weighed 12.6 kN. To minimise the frictional resistance between the pipe string and surrounding soil, a highly viscous lubricant with Marsh cone viscosity of 38 mins was injected into the overcut annulus.

## *Engineering Geology*

The geological profile as determined from four geologic boreholes BH1~4 adjacent to the four drives is shown in Fig. 4b. While the soil properties profile as evaluated from both the field and laboratory tests is shown in Fig. 5. It can be observed from Fig. 6 that drives B, C and D were located within gravel and sand in the main. The soils are also classified as clayey gravel and clayey sand for certain sections of drives B, C and D. The minor clay fraction plays a leading role in the macroscale mechanical properties of the clayey gravel or clayey sand. In light of this, the consistency of the clay particles constituted to the clayey gravel would be assessed using a combination of the semi-empirical approach and the machine learning technologies. The phreatic surface is located at a depth of approximately 4.5 m below ordnance datum (BOD). Additional details on the project are available in Cheng et al. (2017, 2018, 2019a,b).

## **RESULTS AND DISCUSSION**

### *Baseline assessment results*

To benchmark predictions, a baseline assessment of clogging potential during pipejacking in the clayey soils was completed using measurements of the support slurry density  $\rho$ , jacking speed  $V$ , and cutterwheel torque  $T_C$ , in conjunction with the semi-empirical approach. During pipejacking in clayey soils, the cutterwheel is pressed into the cutting face and the plastically deformable soil is pushed to both sides in the form of 'lumps'. Water in support slurry can transform the consistency of the cut lumps into a sticky consistency. In this context, uncritical soils turn into sticky material, and parts of the fines content contained in the cut lumps or the soil at the cutting face may disintegrate and accumulate in the support slurry. The negative impacts, induced by the disintegrated fines content, only occur at the final stage of the support slurry flushing process. However, the clogging, induced by the sticky material, happens

throughout process from the excavation at the cutting face (i.e. primary clogging) up to separation and transport for disposal (i.e. secondary clogging). When clogging occurs, there is a substantial reduction in  $V$  and there are also negative impacts on the cutterwheel torque  $T_C$  due to the mechanical forces arising from the excavation-transport-disposal process. Clogging also has a negative impact on the density of support slurry  $\rho$ ; very high  $\rho$  after a flushing process may correspond to the presence of clayey soils in the support slurry (in suspension or cut lumps). In contrast, there is a substantial reduction in  $\rho$  after the flushing process for silty soils. Increasing the jacking force  $F_T$  is a common technique adopted by shield operators to overcome clogging issues but this is not always proportional relation to the severity of clogging. It is therefore likely that clogging is accompanied by a substantial reduction in the jacking speed  $V$ , an increase in the cutterwheel torque  $T_C$  and high  $\rho$  after the flushing process.

At drive B, the clayey gravel at jacking distances of 20-26 m, 45-65 m, 89-111 m, and 122-126 m may pose a risk of clogging (see Fig. 7a). Although the value of  $V$  corresponding to jacking distances of 45-65 m and 122-126 m show significant fluctuations, the variations in  $T_C$  were not significant. These results indicate that the consistency of the excavated material was not critical for clogging and its shear strength was not capable of withstanding the mechanical forces exerted during the transport-disposal process. Further,  $\rho$  after slurry flushing was measured to be significantly below the threshold of 12.6 kN/m<sup>3</sup>, which also corresponds to higher content of silt and low clogging potential, as shown in Fig. 7b. For the section corresponding to a jacking distance of 89-111 m, the variations in  $T_C$  also showed only slight changes despite  $V$  reducing to below 30 r/min on three separate occasions. Corresponding values of  $\rho$  were again significantly below the 12.6 kN/m<sup>3</sup> threshold (see Fig. 7b) indicating that the soil consistency was not critical for clogging. In contrast,  $V$ , corresponding to a jacking distance of 20-26 m, substantially reduced from 112.5 r/min to below 30 r/min despite the discontinuous variations in  $T_C$ . It is noteworthy that  $\rho$  measured at 12.2 kN/m<sup>3</sup> (the highest of  $\rho$  observed in this study), which is derived right after slurry flushing and close to the threshold of 12.6 kN/m<sup>3</sup>. This phenomenon indicates that the excavation was undertaken in the clayey soil and the spoil with high  $\rho$ , induced by the inclusion of clayey soil, caused some difficulty in transporting the spoil towards leading to the inability to secure the jacking speed. Therefore, tunnelling through this section was recognised as a relatively high risk of clogging.

For drive C, clayey gravels were encountered at jacking distances of 35-37 m, 41-53 m, 57-64 m, 67 m, 70 m, and 72-73 m, as shown in Fig. 8a. At jacking distances of 67 m, 70 m, and 72-73 m,  $V$  was measured to be 50 r/min while  $T_C$  measurements reached a maximum of 20 Amp. Further,  $\rho$  dropped to as low as 9.9 kN/m<sup>3</sup> following slurry flushing, which also indicates a higher content of silt, as shown in Fig. 8b. These results suggest low clogging risk in these

sections. In contrast, at jacking distances of 35-37 m, 41-53 m, and 57-64 m the measured value of  $V$  increased to 137.5 r/min despite the similarly high values of  $T_C$  (25 Amp). In addition,  $\rho$  measured significantly below 12.6 kN/m<sup>3</sup> except at a jacking distance of 50 m where  $\rho$  = 11.6 kN/m<sup>3</sup>, as shown in Fig. 8b. It is worth noting that a decrease in  $V$ , accompanied by an increase in  $T_C$ , occurred at jacking distance of 59-63 m, which was the result of a buried wooden log strike, rather than clogging.

At drive D, six sections of the drive involved tunnelling in clayey gravel and were therefore considered for analysis (jacked distances of 18-22 m, 26-47 m, 50 m, 55-62 m, 66-76 m, 81-99 m in Fig. 9a). At a jacking distance of 18-22 m, while  $V$  and  $T_C$  were relatively constant,  $\rho$  was measured to be 12.3 kN/m<sup>3</sup> (Fig. 9b), indicating a higher content of clay soil and therefore a higher tendency to clog. The section of clayey gravel at 50 m jacking distance was too short to cause clogging issues. Although no data were available for  $\rho$  for a jacking distance of 55-62 m, changes in  $V$  and  $T_C$  were not commensurate with clogging suggesting a low clogging risk. For the section 83-99 m,  $V$  reduced to 35 r/min,  $T_C$  increased to 2.5 Amp and  $\rho$  was measured to be 11.3 kN/m<sup>3</sup> (see Figs. 9a and 9b) indicating low clogging risk.  $V$  and  $T_C$  both showed decreasing trends while tunnelling through a jacking distance of 66-76 m, implying that the material consistency could not withstand the mechanical forces exerted during the excavation-transportation-disposal process. Compared to the jacking distance of 81-99 m, there was comparatively less tendency for clogging at a jacking distance of 66-76 m even though  $\rho$  was measured as 12.1 kN/m<sup>3</sup>. However, a high clogging risk was identified for a jacking distance of 26-47 m due to the reduction in  $V$  (from 62 r/min to 25 r/min) and the increase in  $\rho$  (from 9.9 kN/m<sup>3</sup> to 12 kN/m<sup>3</sup>).

The use of the semi-empirical approach is also considered part of the baseline assessment of clogging potential. Avunduk and Copur (2018) reported that specimens extracted from the excavated material of a earth pressure balance (EPB) TBM, used in an Istanbul utility tunnel project, were classified as low-plasticity clay (CL) or high-plasticity clay (CH) and characterised by  $I_p$  = 19-31% and  $I_c$  = 0.23-0.91. Tokgöz (2016) analysed a total of 264 EPB-TBM excavation performance data, when subjected to fine-grained sedimentary materials with  $I_p$  = 40-68% and  $I_c$  = 0.64-0.98. The results revealed that  $F_T$  and  $T_C$  both increase while  $V$  decreases with increasing  $I_p$  and  $I_c$ , and that the increase in  $T_C$  can be explained by the higher shear strength of the clogging material. Further, one can deduce that the effect of fine-grained materials on  $V$  mainly depends upon the clay morphology which affects specific surface area (SSA) and then free water adsorption capability, and an appropriate designed soil conditioning chemical is crucial to prevent clogging and adhesion-related problems. In this study, the fine-grained soil specimens were extracted from nearby geological boreholes (Cheng et al. 2017) and from those in Tamshui T1 area (Woo and Moh 1990) and their clogging potentials, as defined by

Hollmann and Thewes (2013), and those reported in the literature (Tokgöz 2016; Avunduk and Copur 2018) were plotted in the universal classification diagram, as shown in Fig. 10. It can be seen that  $I_p$  and  $I_c$  for the clayey soils spread over the worksite and Tamshui T1 area are in the ranges of 8.3-12.9 and 0.27-0.60 respectively, hence substantiating all the specimens of clogging material with predominantly high and/or little clogging potentials. The results using the semi-empirical approach are in line with the authors' observations corresponding to jacking distances of 20-26 m at drive B and 26-47 m at drive D.

#### *Machine learning results*

In this and subsequent sections, only drives B and D with a risk of clogging were analysed and discussed. The performance results of AD approaches using the transformed slurry density space are shown in Figs. 11 and 12. At drive B, all three AD approaches provided one TP at 21 m jacking distance, with  $R_t = 1.0$  and  $T_t = 1.0$  (Fig. 11). There were two FPs corresponding to jacking distances of 31 m and 41 m. At drive D, Robcov and IForest gave one TP at 38 m jacking distance, whereas OCSVM provided one FN at 38 m jacking distance (Fig. 12). There were two FPs (i.e. 65 m and 79 m) distributed on the right hand side of the feature space (large  $R_t$ ) and the other FP (i.e. 75 m) distributed on the left (small  $R_t$ ). The performance results of AD approaches using the transformed maximum torque space are shown in Figs. 13 and 14. At drive B, IForest and OCSVM produced one TP at 24 m jacking distance (Fig. 13). While Robcov provided one FN at 24 m jacking distance. There were four FPs (i.e. 110-119 m) positioned on the bottom (small  $T_t$ ) in the feature space and another one FP (i.e. 10 m) distributed on the top right corner (large  $R_t$  and  $T_t$ ). The other two (i.e. 54 m and 90 m) sat on the left (small  $R_t$ ). At drive D, OCSVM and Robcov produced one TP at 38 m jacking distance. IForest provided one FN at 38 m jacking distance (Fig. 14). There was one FP (i.e. 89 m) distributed on the right (large  $R_t$ ). The other three FPs (i.e. 19 m, 72 m and 73 m) sat on the top left corner (small  $R_t$  and large  $T_t$ ) and the bottom left corner (small  $R_t$  and  $T_t$ ) as well as the bottom right corner (large  $R_t$  and small  $T_t$ ) respectively. The performance results using the transformed jacking speed space are shown in Figs. 15 and 16. At drive B, IForest produced three datapoints at 20-26 m jacking distance. Robcov and OCSVM gave two datapoints at 20-26 jacking distance. There were two FPs (i.e. 35 m and 49 m) distributed on the top right corner (large  $R_t$  and  $T_t$ ) in the feature space and another one FP (i.e. 36 m) positioned on the top left corner (small  $R_t$  and large  $T_t$ ) (Fig. 15). The other two FPs (i.e. 6 m and 11 m) sat on the right (large  $R_t$ ) and the left (small  $R_t$ ) respectively. At drive D, OCSVM and Robcov produced one datapoint (i.e. 39 m) at 26-47 m. IForest provided one FN (i.e. 39 m) at 26-47 m. There were three FPs (i.e. 62 m, 81 m and 82 m) distributed on the top right corner (large  $R_t$  and  $T_t$ ) in the feature space and another three (i.e. 57 m, 67 m and 68 m) sat

on the bottom left corner (small  $R_t$  and  $T_t$ ) (Fig. 16). The other two FPs (i.e. 75 m and 83 m) distributed on the top left corner (small  $R_t$  and large  $T_t$ ).

#### *Discussion*

Measurements of  $\rho$ ,  $T_C$  and  $V$  and their difference to 'nearby' datapoints can significantly affect the assessment of clogging potential. There are two behaviours that address the formation of clogging in this work: (1) absolute values of  $\rho$ ,  $T_C$  and  $V$ , influence  $T_t$  and (2) the difference between 'current' measurements of  $\rho$ ,  $T_C$  and  $V$  from nearby datapoints influence  $R_t$ . For brevity, only the typical assessment results using the density of support slurry  $\rho$ , the cutterwheel torque  $T_C$  and the jacking speed  $V$  are discussed in sequence as follows. The lower and upper limits of  $\rho$  corresponded to 10.8 kN/m<sup>3</sup> and 12.2 kN/m<sup>3</sup> respectively at drive B and to 9.8 kN/m<sup>3</sup> and 12.8 kN/m<sup>3</sup> respectively at drive D. At drive B, the formation of FP at 41 m jacking distance was attributed to  $T_t = 0$ , induced by  $\rho = 10.9$  kN/m<sup>3</sup> dropping significantly below the threshold of 12.6 kN/m<sup>3</sup>. Similarly, the FP at 65 m distance of drive D was formed due to  $T_t = 1.0$  and  $R_t = 0.87$ , induced by  $\rho = 12.1$  kN/m<sup>3</sup>. These results can be categorised into the 'type-1 behaviour'. In contrast, FP at 31 m distance of drive B was formed by  $R_t = 0$ , induced by a decline in  $\rho$  to 10.8 kN/m<sup>3</sup> from 12.2 kN/m<sup>3</sup> at 21 m jacking distance. The higher  $\rho$  from the nearby datapoint (i.e. 21 m distance) increased  $T_t$  and this, in turn, lowered  $R_t$  to 0. At drive D, FP at 79 m jacking distance was produced as a result of  $R_t = 1.0$ , induced by an increase in  $\rho$  to 11.3 kN/m<sup>3</sup> from 10.1 kN/m<sup>3</sup> at 75 m distance. The lower  $\rho$  from the adjoining datapoint (i.e. 75 m distance) lowered  $T_t$  to approximately 0 and this, in turn, increased  $R_t$  to 1.0. These results can, therefore, be categorised as 'type-2 behaviour'.

The maximum of  $T_C$  (termed  $T_{C,max}$  hereafter) varied from 65 Amp to 80 Amp at drive B, whereas it varied between 45 Amp to 55 Amp at drive D. The main cause to lead to FP at 110 m jacking distance of drive B was due to  $T_t = 0$ , induced by  $T_{C,max} = 65$  Amp (lower limit).  $T_{C,max}$  from surrounding datapoints was similar and the impact on  $R_t$  was minimal. The FP at 26 m jacking distance of drive D was caused by  $T_t = 1.0$  from  $T_{C,max}$  hitting the upper limit of 55 Amp. Surrounding datapoints of similar  $T_{C,max}$  produced a 'platform' leading to a negligible impact on  $R_t$ . These results can therefore be classed as type-1 behaviour. In contrast, the leading cause to form a FP at 10 m jacking distance of drive B was because of  $R_t = 1.0$ , induced by  $T_{C,max} = 80$  Amp (upper limit).  $T_{C,max} = 70$  Amp from the nearby datapoints, i.e. 9 m and 11 m jacking distance, lowered  $T_t$  to 0.75. This, in turn, increased  $R_t$  to 1.0. A FP at 89 m distance of drive D was formed by  $R_t = 1.0$ , induced by  $T_{C,max} = 55$  Amp (upper limit). Nearby datapoints that are featured with  $T_{C,max} = 45$  Amp (lower limit) increased  $R_t$  to 1.0 by lowering  $T_t$  to 0.4. The above two instances can therefore be classed as the type-2 behaviour.

The difference in  $V$  measurements (termed  $\Delta V$  hereafter) varied between -37.5-47.5 r/min and -55-65 r/min for drives B and D respectively. A negative value of  $\Delta V$  represents a descending tendency in  $V$ , while a positive value indicates an ascending tendency in  $V$ . A FP at 35 m jacking distance of drive B ( $R_t = 0.98$  and  $T_t = 1.00$ ) was triggered by  $\Delta V = 47.5$  r/min (upper limit). In addition, the FP at 57 m distance of drive D ( $R_t = 0.02$  and  $T_t = 0$ ) was caused by  $\Delta V$  reaching the lower limit of -55 r/min, indicating type-1 behaviour. The FP at 49 m jacking distance of drive B ( $R_t = 0.97$  and  $T_t = 0.84$ ) was due to  $\Delta V = 37.5$  r/min.  $\Delta V = 0$  r/min from 48 m distance lowered  $T_t$  to 0.84 and this, in turn, increased  $R_t$  to 0.97. The FP at 83 m distance of drive D was formed due to  $R_t = 0$  and  $T_t = 0.66$ , induced by  $\Delta V = -10$  r/min.  $\Delta V = 55$  r/min from 82 m distance increased  $T_t$  to 0.66, lowering  $R_t$  to 0 indicating type-2 behaviour. Table 2 summarises the  $R_t$  and  $T_t$  of datapoints and their behaviour in relation to clogging identification. Datapoints distributed in quadrant II (high  $R_t$  and  $T_t$ ) could produce outlier formed by a clogging-induced increase in the density of support slurry or the cutterwheel torque, while datapoints distributed in quadrant III (low  $R_t$  and  $T_t$ ) could lead to outlier produced by a clogging-induced decline in the jacking speed (see Fig. 17). Although the decision boundary established by the IForest approach presented in an irregular shape which has not been seen from the other two AD approaches and was proven rather sensitive to clogging behaviour, it performed the worst amongst the AD approaches (Table 3). TBM or shield machines, equipped with such an integrated AD system, can be useful to introduce countermeasures in advance of clogging during drives in clayey soils.

For clay soils with low to medium clogging potential, the risk of adhesion is evaluated to be low, according to the practical experience of the authors. However, a risk of clogged openings remains throughout the spoil transport and disposal processes, depending upon the consistency of the material. A clay with medium to high clogging potential can adhere to all parts of the machine. Fig. 18 shows the clay soil adhered to the cutterwheel of machine to produce clayey clogging at 20-26 m distance of drive B. The field observation locates where the clayey clogging occurs and is in line with the clogging baseline assessment using the semi-empirical approach. The Seasonal-Trend decomposition procedure using Loess successfully accentuates the characteristics of data in response to the development of clayey clogging towards allowing the 'slurry density' and 'maximum torque' as well as 'jacking speed' anomaly detectors to detect the clayey clogging. The results obtained from this work provide evidence in proving the reliability of the proposed approach and providing some guideposts as follows for reducing the potential of clayey clogging while tunnelling in soft ground. To use a slurry-supported shield machine in such a soil, small soil chips, narrow passages for the transport of a clay chip from the cutting face to the support slurry line, sharp angles applied to the excavation chamber and clay agglomerations in areas prone to material settlement should be



avoided. Therefore, generation of large soil chips to reduce the adhesion-prone surface of the cut lump is essential to prevent the development of clogging. Further, optimised passages, rounded angles and turbulence (namely, flushing nozzles and agitator) in areas which are prone to material settlement should also be considered. Moreover, manipulating the ratio of the suspension flow rate to the volume of cut lump gives a benefit in preventing clay from accumulation.

Human factor is deemed as a crucial factor in anomaly detection for tunnel construction. The authors have also noted that the torque variation, while pipejacking in 51-58 m distance at drive B, is more distinct than other distances of same soil layer, implying that the human factor may intervene the operation of tunnel boring machine and try to lift the machine out of the fine grained soils by imposing 'breakout' cutterwheel torque. The other distinct variation in torque occurs when jacking in 33-35 m distance at drive D. This is not due to the presence of the fine grained soils but to the presence of the gravel. Despite the breakout torque present in 51-58 m distance at drive B, one of the advantages for the use of artificial intelligence technologies is to accentuate patterns in the data and subsequently disaggregate time series data into (stationary) feature-based sub-series towards preventing the accuracy of anomaly detection from disturbing by unusual data induced by, for example, human factor.

A real-time anomaly detection (AD) scheme for tunnel construction has been proposed following this work, as shown in Fig. 19. Three parts comprising acquisition of real-time data, data pre-processing and clayey clogging assessment are integrated in the real-time AD scheme. The real-time tunnelling data are collected by modern tunnel boring machine, followed by the data pre-processing including data filtering, decomposition, feature selection and scaling. The semi-empirical approach requires a measurement of the natural water content, liquid limit and plastic limit to assess the development of clayey clogging and it cannot be executed in a real time manner. Alternatively, we treat the modern tunnel boring machine as a 'sensor' and the measured cutterwheel torque ( $T$ ) can be converted to the tangential adhesion stress ( $\tau_t$ ) towards establishing its relation with the thrust force  $\sigma_n$  (normal pressure). This relation can then be used to determine the tangential strength  $\tau$  and extended by taking the water content of clayey soil into account. To this end, preliminary laboratory tests should be conducted to investigate the change in  $\tau$  with the water content. The real-time AD scheme, assisted by the proposed anomaly detectors, can thus be realised, alarming the operator of machine and reducing the potential of clogging while tunnelling in soft ground.

## CONCLUSIONS

This paper has established a baseline assessment of clogging potential during slurry-supported pipejacking using a combination of existing tunnelling parameters and the semi-empirical approach and examined the potential for the use of anomaly detection approaches to assess clogging potential. Based on the results and discussion, some main conclusions can be drawn as follows:

- (a) Pre-process procedures including data screening and a seasonal-trend decomposition using Loess smoothing were utilised to transform the density of support slurry-distance relationship, the maximum cutterwheel torque-distance relationship and the jacking speed-distance relationship into stationary features for the application of anomaly detection approaches. For the drives considered in this paper, two feature variables, namely, the residual and trend components were found to be most effective in accentuating the presence of anomalous (clogging) behaviour.
- (b) The baseline assessment of clogging potential suggested there was a comparatively high tendency for clogging to occur during pipejacking at jacking distances of 20-26 m for drive B and 26-47 m for drive D. The assessment of anomalous behaviour using the AD approaches indicated that the density of support slurry and the maximum cutterwheel torque as well as the jacking speed provide an effective means to assessing the risk of clogging. Even though the IForest approach established a decision boundary that was most sensitive to anomalous behaviour, it performed the worst amongst the AD approaches.
- (c) Two behaviours that address the formation of clayey clogging were identified: (1) absolute values of  $\rho$ ,  $T_C$  and  $V$ , influence  $T_t$  and (2) the difference between 'current' measurements of  $\rho$ ,  $T_C$  and  $V$  from nearby datapoints influence  $R_t$ . The proposed real-time anomaly detection scheme can be of great benefit to reduce the potential of clogging during drives in clayey soils.
- (d) Although the AD approaches have shown promising performance for assessing the clogging potential during slurry-supported pipejacking, further validation regarding the applicability of density of support slurry, maximum cutterwheel torque and jacking speed to assess the tendency to clayey clogging is deemed necessary. In addition, performance of the AD approaches requires to be examined when subjected to data from drives with different geological conditions, construction techniques and shield machine geometries.

## ACKNOWLEDGEMENTS

576 The authors convey their thanks and sincerely acknowledge financial supports from the  
577 Special fund for Basic Scientific Research of Central Colleges, Chang'an University, under  
578 Grant No. 300102269502. The third author is funded by the Royal Academy of Engineering  
579 under the Research Fellowship Scheme.

580

#### 581 **DECLARATION OF COMPETING INTEREST**

582 The authors declare that they have no known competing financial interests or personal  
583 relationships that could have appeared to influence the work reported in this paper.

584

#### 585 **AUTHORSHIP CONTRIBUTION STATEMENT**

586 **Xue-Dong Bai**: Data curation, Formal analysis, Validation, Software, Writing - original draft.

587 **Wen-Chieh Cheng**: Conceptualization, Methodology, Writing - review & editing, Supervision,  
588 Funding acquisition. **Brian B. Sheil**: Writing - review & editing. **Ge Li**: Data curation, Formal  
589 analysis, Validation, Writing - original draft.

## REFERENCES

- Alberto-Hernandez, Y., Kang, C., Yi, Y., Bayat, A., 2017. Mechanical properties of clayey soil relevant for clogging potential. *International Journal of Geotechnical Engineering*, 1-8. doi: 10.1080/19386362.2017.1311086
- Alemdag, S., Gurocak, Z., Cevik, A., Cabalar, A.F., Gokceoglu, C., 2016. Modeling deformation modulus of a stratified sedimentary rock mass using neural network, fuzzy inference and genetic programming. *Engineering Geology* 203, 70-82.
- Breunig, M.M., Kriegel, H.P., Ng, R.T., Sander, J., 2000. LOF: Identifying density-based local outliers. In Vol. 29 of ACM SIGMOD record, 93-104. New York: Association for Computing Machinery.
- Cheng, W.C., Ni, J.C., Shen, J.S., 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., 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., 2019b. 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: 2933-2942.
- Cheng, W.C., Wang, L., Xue, Z.F., Ni, J.C., Rahman, M., Arulrajah, A., 2019a. Lubrication performance of pipejacking in soft alluvial deposits. *Tunnelling and Underground Space Technology*, 91: 102991.
- Cleveland, R., Cleveland, W., McRae, J., Terpenning, I., 1990. STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics* 6(1): 3-73.
- Deep, K., Thakur, M., 2007a. A new crossover operator for real coded genetic algorithms. *Applied Mathematics and Computation* 188(1), 895-911.
- Deep, K., Thakur, M., 2007b. A new mutation operator for real coded genetic algorithms. *Applied Mathematics and Computation* 193(1), 211-230.
- de Oliveira, D.G.G., Diederichs, M., Thewes, M., 2019b. EPB excavation of cohesive mixed soils: combined methodology for clogging and flow assessment. In: ITA-WTC 2019, Tunnels and Underground Cities: Engineering and Innovation meet Archaeology, Architecture and Art, Naples, Italy, pp. 2008-2017.
- de Oliveira, D.G.G., Diederichs, M., Thewes, M., 2019c. EPB machine excavation of mixed soils – laboratory characterisation. *Geomech. Tunnel. J.* 12 (4): 373-385.

- de Oliveira, D.G.G., Thewes, M., Diederichs, M., 2019a. Clogging and flow assessment of cohesive soils for EPB tunnelling: Proposed laboratory tests for soil characterisation. *Tunnelling and Underground Space Technology* 94: 103110.
- de Oliveira, D.G.G., Thewes, M., Diederichs, M.S., & Langmaack, L. (2018). EPB tunnelling through clay-sand mixed soils: Proposed methodology for clogging evaluation. *Geomechanics and Tunnelling* 11(4): 375–387.
- Feinendegen, M., Ziegler, M., Spagnoli, G., Fernández-Steege, T., Stanjek, H., 2010. A new laboratory test to evaluate the problem of clogging in mechanical tunnel driving with EPB-shields. In: ISRM International Symposium-eurock, Lausanne, Switzerland, pp. 429-432.
- Feinendegen, M., Ziegler, M., Weh, M., Spagnoli, G., 2011. Clogging during EPB-tunnelling: Occurrence, classification and new manipulation methods. In: Proceedings ITA-AITES World Tunnel Congress, Helsinki, pp. 767-776.
- Fontaine, E.R., 1954. Investigations into the mechanism of soil adhesion. *European Journal of Soil Science* 5(2): 251-263.
- Gurocak, Z., Solanki, P., Alemdag, S., Zaman, M., 2012. New considerations for empirical estimation of tensile strength of rocks. *Engineering Geology* 145-146, 1-8.
- Heuser, M., Spagnoli, G., Leroy, P., Klitzsch, N., Stanjek, H., 2012. Electro-osmotic flow in clays and its potential for reducing clogging in mechanical tunnel driving. *Bulletin of Engineering Geology and the Environment* 71: 721-733.
- Hollmann, F.S., Thewes, M., 2013. Assessment method for clay clogging and disintegration of fines in mechanised tunnelling. *Tunnelling and Underground Space Technology* 37: 96-105.
- Hu, W., Rostami, J., 2020. A new method to quantify rheology of conditioned soil for application in EPB TBM tunneling. *Tunnelling and Underground Space Technology* 96: 103192.
- Jia, X., 2004. Theoretical analysis of the adhesion force of soil to solid materials. *Biosystems Engineering* 87(4): 489-493.
- Kang, C., Wu, Y., Yi, Y., Bayat, A., 2019. Assessment of the clogging potential of two clays. *Applied Clay Science* 178: 105134.
- Kang, C., Yi, Y., Bayat, A., 2018. Performance evaluation of TBM clogging potential for plain and conditioning soil using a newly developed laboratory apparatus. *International Journal of Geotechnical Engineering*: 1-10. doi: 10.1080/19386362.2018.1439671
- Liu, F.T., Ting, K.M., Zhou, Z.H., 2012. Isolation-based anomaly detection. *ACM Transactions on Knowledge Discovery from Data* 6(1): 1-39.
- Masters, T., 1993. Practical neural network recipes in C++. Academic Press, San Diego, CA.

- Ni, J.C., Cheng, W. (2010). Using fracture grouting to lift structures in clayey sand. *Journal of Zhejiang University-SCIENCE A* 11(11): 879-886.
- 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: 2825-2830.
- Pelia, D., Picchio, A., Martinelli, D., Dal Negro, E., 2015. Laboratory tests on soil conditioning of clayey soil. *Acta Geotechnica* 11(5): 1061-1074.
- Persons, W.M., 1919. Indices of business conditions: An index of general business conditions. Cambridge, MA: Harvard University Press.
- Rousseeuw, P.J., 1984. Least median of squares regression. *Journal of the American Statistical Association* 79(388): 871-880.
- Ryu, Y.M., Kwon, Y.S., Kim, T.H., Lee, I.M., 2019. Slurry clogging criteria for slurry shield tunnelling in highly permeable ground. *KSCE Journal of Civil Engineering*: 1-10. doi: 10.1007/s12205-019-2156-x
- Sass, I., Burbaum, U., 2008. A method for assessing adhesion of clays to tunneling machines. *Bulletin of Engineering Geology and the Environment* 68(1): 27-34.
- Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., Williamson, R. C., 2001. Estimating the support of a high-dimensional distribution. *Neural Computation* 13(7): 1443-1471.
- Sheil, B., Suryasentana, S.K., Cheng, W.C., 2020. Assessment of anomaly detection methods applied to microtunneling. *Journal of Geotechnical and Geoenvironmental Engineering*, 146(9): 04020094.
- Soomro, M.A., Kumar, M., Xiong, H., Mangnejo, D.A., Mangi, N., 2020. Investigation of effects of different construction sequences on settlement and load transfer mechanism of single pile due to twin stacked tunnelling. *Tunnelling and Underground Space Technology* 96: 103171.
- Spagnoli, G., Feinendegen, M., Seidl, W., 2019b. The impact of mineralogy and chemical conditioning on the mechanical and adhesive properties of clays. In: XVII European Conference on Soil Mechanics and Geotechnical Engineering 2019, Geotechnical Engineering Foundation of the Future, Reykjavik, Iceland.
- Spagnoli, G., Feinendegen, M., Stanjek, H., Azzam, R., 2011a. Soil conditioning for clays in EPBMs. *Tunnels & Tunnelling International* 43(10): 56-61.
- Spagnoli, G., Feinendegen, M., Ziegler, M., 2019a. Assessing the clogging potential of clay/additive mixtures by cone pull-out tests. *Geomechanics and Tunnelling* 12(4), 362-371.
- Spagnoli, G., Klitzsch, N., Fernandez-Steege, T., Feinendegen, M., Rey, A.R., Stanjek, H., Azzam, R., 2011b. Application of electro-osmosis to reduce the adhesion of clay during mechanical tunnel driving. *Environmental & Engineering Geoscience* 17(4): 417-426.

- Spagnoli, G., Rubinos, D., Stanjek, H., Fernandez-Steege, T., Feinendegen, M., Azzam, R., 2012a. Undrained shear strength of clays as modified by pH variations. *Bulletin of Engineering Geology and the Environment* 71(1): 135-148.
- Spagnoli, G., Stanjek, H., Feinendegen, M., Azzam, R., 2014. Mineralogical, mechanical and electrical properties of the clays and their relation with regard to the clogging during mechanical tunnel driving. In: Geo-Congress 2014, Geo-Characterization and Modeling for Sustainability, Atlanta, USA, pp. 495-504.
- Spagnoli, G., Stanjek, H., Sridharan, A., 2012b. Influence of ethanol/water mixture on the undrained shear strength of pure clays. *Bulletin of Engineering Geology and the Environment* 71(2): 389-398.
- Ong, D.E.L. and Choo, C.S., 2016. Back-analysis and finite element modeling of jacking forces in weathered rocks. *Tunnelling and Underground Space Technology*, 51: 1-10.
- Tan, Y., Wei, B., 2012. Performance of an overexcavated metro station and facilities nearby. *Journal of Performance of Constructed Facilities*, 26(3): 241-254.
- Thewes, M., 1999. Adhesion of clay soil in tunnel drives with slurry shields (In German: Adhäsion von Tonböden beim Tunnelvortrieb mit Flüssigkeitsschilden). Berichte aus Bodenmechanik und Grundbau der Bergischen Universität Wuppertal, Fachbereich Bauingenieurwesen, Bd. 21. Shaker Verlag, Aachen.
- Thewes, M., Burger, W., 2005. Clogging of TBM drives in clay – identification and mitigation of risks. In: Proceedings ITA-AITES World Tunnel Congress, Istanbul, Turkey, pp. 737–742.
- Thewes, M., Hollmann, F.S., 2014. TBM-specific testing scheme to assess the clogging tendency of rock. *Geomechanics and Tunnelling* 7(5): 520-527.
- Thewes, M., Hollmann, F.S., 2016. Assessment of clay soils and clay-rich rock for clogging of TBMs. *Tunnelling and Underground Space Technology* 57: 122-128.
- Tombacz, E., Szekeres, M., 2006. Surface charge heterogeneity of kaolinite in aqueous suspension in comparison with montmorillonite. *Applied Clay Science* 34: 105-124.
- van Baalen, L.R., 1999. Reduction of clay adherence by electroosmosis. Master thesis, Faculty of Civil Engineering and Geosciences, TU Delft.
- van Baalen, L.R., Zimnik, A., Verhoef, P., Ngan-Tillard, D., 2001. Applicability of electroosmosis to reduce clay adhesion in a TBM. In Proceedings of the International Conference on Geotechnical Geological Engineering. Melbourne.
- Vapnik, V., 1995. The nature of statistical learning theory. Springer, New York.
- Yin, Z.Y., Jin, Y.F., Huang, H.W., Shen, S.L., 2016. Evolutionary polynomial regression based modelling of clay compressibility using an enhanced hybrid real-coded genetic algorithm. *Engineering Geology* 210, 158-167.

- Zhang, F., Kong, R., Peng, J., 2018. Effects of heating on compositional, structural, and physicochemical properties of loess under laboratory conditions. *Applied Clay Science* 152: 259-266.
- Zhang, W.G., Li, H.R., Wu, C.Z., Li, Y.Q., Liu, Z.Q., Liu, H.L., 2020b. Soft computing approach for prediction of surface settlement induced by earth pressure balance shield tunneling. *Underground Space*, in Press. doi: 10.1016/j.undsp.2019.12.003
- Zhang, C., Yang, J., Fu, J., Wang, S., Yin, J., & Xie, Y., 2020a. Recycling of discharged soil from EPB shield tunnels as a sustainable raw material for synchronous grouting. *Journal of Cleaner Production* 268: 121947.
- Zhang, W.G., Zhang, R., Wu, C., Goh, A.T.C., Lacasse, S., Liu, Z., Liu, H., 2020c. State-of-the-art review of soft computing applications in underground excavations. *Geoscience Frontiers* 11(4): 1095-1106. doi: 10.1016/j.gsf.2019.12.003
- Zumsteg, R., Plötze, M., Puzrin, A., 2013. Reduction of the clogging potential of clays: new chemical applications and novel quantification approaches. *Geotechnique* 63(4): 276-286.
- Zumsteg, R., Puzrin, A.M., Anagnostou, G., 2016. Effects of slurry on stickiness of excavated clays and clogging of equipment in fluid supported excavations. *Tunnelling and Underground Space Technology* 58: 197-208.

#### List of figures

- Fig. 1 Illustration of one-class support vector machine approach and the construction of the hyperplane by transforming the original input space into a high-dimensional feature space
- Fig. 2 Illustration of isolation forest approach and the construction of separate isolation trees by randomly sampling from the training dataset. Red circles represent outliers, blue circles indicate uncommon inliers and light blue circles represent common inliers
- Fig. 3 Illustration of robust covariance approach showing the use of a robust estimator of covariance to reflect the true organisation of observations using the Mahalanobis distance



- Fig. 4 Pipejacking project description: (a) location of the four drives in Taipei, Taiwan and (b) geological profile derived from four nearby boreholes
- Fig. 5 Soil properties profile
- Fig. 6 Grain-size distribution curves for drives B, C and D
- Fig. 7 Monitored data for drive B: (a) development of jacking force, jacking speed and cutterwheel torque with jacking distance and (b) development of density of support slurry with jacking distance
- Fig. 8 Monitored data for drive C: (a) development of jacking force, jacking speed and cutterwheel torque with jacking distance and (b) development of density of support slurry with jacking distance
- Fig. 9 Monitored data for drive D: (a) development of jacking force, jacking speed and cutterwheel torque with jacking distance and (b) development of density of support slurry with jacking distance
- Fig. 10 Universal classification diagram for the assessment of anomalous clogging behaviour
- Fig. 11 Performance of AD approaches applied to drive B using transformed slurry density space: (a) OCSVM, (b) Robcov and (c) IForest
- Fig. 12 Performance of AD approaches applied to drive D using transformed slurry density space: (a) OCSVM, (b) Robcov and (c) IForest

Fig. 13 Performance of AD approaches applied to drive B using transformed maximum torque space: (a) OCSVM, (b) Robcov and (c) IForest

Fig. 14 Performance of AD approaches applied to drive D using transformed maximum torque space: (a) OCSVM, (b) Robcov and (c) IForest

Fig. 15 Performance of AD approaches applied to drive B using transformed jacking speed space: (a) OCSVM, (b) Robcov and (c) IForest

Fig. 16 Performance of AD approaches applied to drive D using transformed jacking speed space: (a) OCSVM, (b) Robcov and (c) IForest

Fig. 17 Simple diagram for TBM operator to assess tendency to clayey clogging during slurry-supported pipejacking

Fig. 18 Clayey clogging with embedded gravel occurred while spanning through a ground composed of the clayey gravel and the gravel at drive B: (a) before and (b) after the clay soil adhered to cutterwheel

Fig. 19 Real-time anomaly detection scheme for tunnel construction in clayey soil

756 List of tables

Table 1 Summary of geological data and shield machine performance data

Table 2 Summary of residual and trend components of datapoints and their behaviour relevant to anomalous behaviour formation

Table 3 Summary of discovery rate and false alarm rate

757

758

Table 1 Summary of geological data and shield machine performance data

Geological data											Performance data of TBM or slurry-supported shield		
Data source	Particle distribution			Unit weight (kN/m <sup>3</sup> )	Natural water content (%)	Atterberg limits			Consistency $I_c$	USCS symbol	Cutterwheel torque $T_c$ (Amp or kN·m)	Thrust force $F_T$ (kN)	Jacking speed $V$ (r/min)
	Sand	FC	Clay			LL (%)	PL (%)	PI (%)					
Tokgöz (2016)	-	84.0	3.9	17.4	36.0	70	28	42	0.81	CH	70	350	2.8
	-	88.5	4.6	16.8	39.5	65	25	40	0.64	CH			
	-	89.3	6.5	16.8	31.0	88	30	58	0.98	CH	90	255	2.8
	-	92.6	6.5	17.9	33.8	90	30	60	0.94	CH			
	-	99.8	7.7	17.6	39.0	91	36	55	0.95	CH	85	270	2.8
	-	92.7	3.4	18.3	35.1	72	27	45	0.82	CH			
	-	97.5	6.1	17.9	30.0	80	22	58	0.86	CH	87.5	260	2.8
	-	94.5	3.6	17.3	30.2	86	24	62	0.90	CH			
	-	91.0	7.2	17.8	29.4	92	24	68	0.92	CH	87.5	225	2.8
	-	96.6	4.9	17.9	32.3	94	28	56	1.10	CH			
	-	89.8	5.0	17.5	35.0	82	30	52	0.90	CH	75	155	2.8
	-	96.5	6.0	18.0	44.1	81	33	48	0.77	CH			
	-	89.8	7.1	17.7	28.5	86	23	63	0.91	CH	80	250	2.8
Woo and Moh (1990)	67.0	33.0	-	19.7	27.7	34	22	12	0.53	CL	-	-	-
	25.0	75.0	-	18.8	29.6	36	23	13	0.49	CL	-	-	-
	34.0	31.0	26.0	16.5	26.6	33.3	22.1	11.2	0.60	CL	-	-	-
	22.0	43.0	26.0	16.7	30.6	26.9	18.6	8.3	-	CL	-	-	-
	-	-	-	14.8	28.5	31.2	22.8	8.4	0.32	CL	-	-	-

Table 1 Summary of geological data and shield machine performance data (cont'd)

Geological data											Performance data of TBM or slurry-supported shield		
Data source	Particle distribution			Unit weight (kN/m <sup>3</sup> )	Natural water content (%)	Atterberg limits			Consistency $I_c$	USCS symbol	Cutterwheel torque $T_c$ (Amp or kN·m)	Thrust force $F_T$ (kN)	Jacking speed $V$ (r/min)
	Sand	FC	Clay			LL (%)	PL (%)	PI (%)					
This study	85.0	14.0	8.0	20.1	22.4	25	15	10	0.26	SC	50	2254	25
	20.5	5.5	1.5	19.1	22.4	28	17	11	0.51	GP-GC	45	2793	80
Avunduk and Copur (2018)	31	69	-	-	30.0	41	19	22	0.49	CL	797 ± 122	4539 ± 562	-
	35	65	-	-	30.0	46	18	28	0.58	CL			
	38	62	-	-	30.0	48	24	24	0.75	CL			
	39	61	-	-	30.0	56	26	30	0.86	CH			
	42	58	-	-	30.0	57	27	30	0.88	CH			
	39	61	-	-	30.0	58	27	31	0.91	CH			
	21	79	-	-	25.0	37	11	26	0.46	CL	517 ± 153	5705 ± 702	-
	68	32	-	-	25.0	-	-	-	-	SM			
	63	37	-	-	25.0	-	-	-	-	SM			
	57	53	-	-	28.0	33	15	19	0.28	CL	475 ± 189	5750 ± 785	-
	49	51	-	-	28.0	33	12	21	0.23	CL			
	30	70	-	-	28.0	35	13	21	0.31	CL			
	27	73	-	-	28.0	40	15	25	0.47	CL			
	24	76	-	-	28.0	41	14	26	0.48	CL			
Avunduk and Copur (2019)	67	33	-	-	12.5	49	35.6	13.4	2.72	SC	233	1900	3.6

Table 2 Summary of residual and trend components of datapoints and their behaviour relevant to anomalous behaviour formation

Pipejacking drive	Parameter considered	Datapoint	Residual component $R_t$	Trend component $T_t$	Behaviour	Quadrant distribution
B	Density	21	1.00	1.00	1	II
		31	0	0.66	2	I
		41	0.46	0	1	III
	Torque	10	1.00	0.71	1	II
		24	0.77	1.00	1	II
		54	0	0.86	2	I
		90	0.01	0.29	2	III
		110	0.44	0	1	III
		111	0.50	0	1	III
		118	0.34	0	1	III
		119	0.44	0	1	III
	Speed	6	1.00	0.59	2	II
		11	0	0.61	2	I
		23	0.39	0.04	1	III
		24	0.37	0	1	III
		25	0.61	0.16	1	IV
		35	0.98	1.00	1	II
		36	0.20	0.87	2	I
		49	0.97	0.84	2	II

Table 2 Summary of residual and trend components of datapoints and their behaviour relevant to anomalous behaviour formation (cont'd)

Pipejacking drive	Parameter considered	Datapoint	Residual component $R_t$	Trend component $T_t$	Behaviour	Quadrant distribution
D	Density	38	0.94	0.42	2	IV
		65	0.87	1.00	1	II
		75	0	0.12	2	III
		79	1.00	0.17	2	IV
	Torque	19	0.62	0.80	2	I
		26	0.38	1.00	1	I
		37	0.43	1.00	1	I
		38	0.65	1.00	1	II
		72	0.19	0	1	III
		73	0.81	0	2	IV
		89	1.00	0.40	2	IV
	Speed	39	0.13	0.16	2	III
		57	0.02	0	1	III
		62	0.88	0.87	1	II
		67	0.21	0.12	1	III
		68	0.22	0.07	1	III
		75	0.06	0.76	2	I
		81	1.00	1.00	1	II
		82	0.88	0.96	1	II
		83	0	0.66	2	I

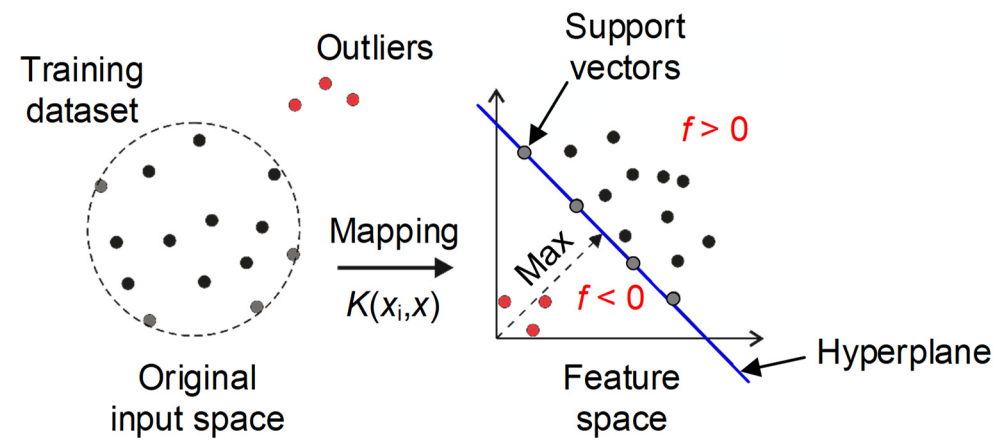
Note: density = density of support slurry; torque = cutterwheel torque; speed = jacking speed; I = distribution quadrant of datapoint as per their values of  $R_t$  and  $T_t$ ; details relevant to behaviours '1' and '2' refer to the beginning of the 'Discussion' section in the main text.

Table 3 Summary of discovery rate and false alarm rate

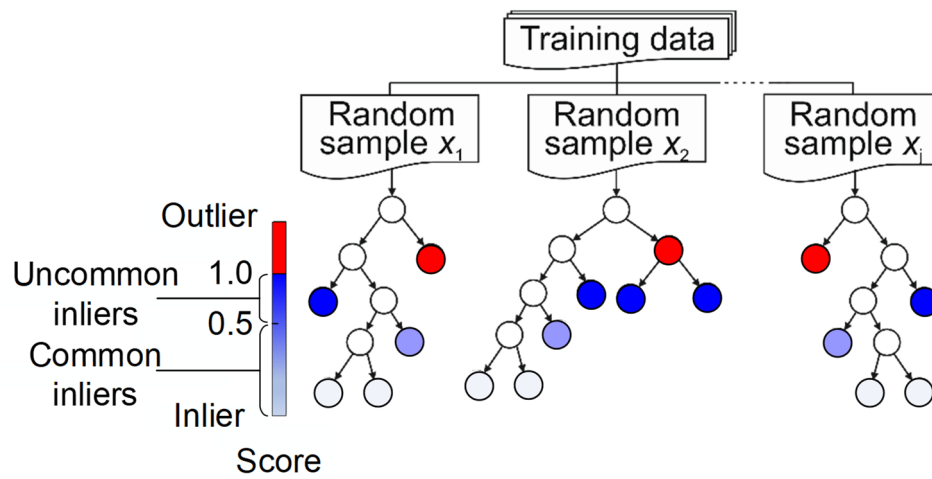
AD approach	Pipejacking drive	Parameter considered	Number of datapoint	Number of TP	Number of TN	Number of FP	Number of FN	Discovery rate, DR (%)	False alarm rate, FAR (%)
OCSVM	B	Density	29	1	26	2	0	100	7.1
		Torque	58	1	52	5	0	100	8.8
		Speed	60	1	55	4	0	100	6.8
	D	Density	25	0	21	3	1	0	12.5
		Torque	39	1	35	3	0	100	7.9
		Speed	69	1	62	6	0	100	8.8
Robcov	B	Density	29	1	28	0	0	100	0
		Torque	58	0	51	6	1	0	10.5
		Speed	60	1	55	4	0	100	6.8
	D	Density	25	1	22	2	0	100	8.3
		Torque	39	1	35	3	0	100	7.9
		Speed	69	1	62	6	0	100	8.8
IForest	B	Density	29	1	28	0	0	100	0
		Torque	58	1	53	4	0	100	7.0
		Speed	60	1	56	3	0	100	5.1
	D	Density	25	1	23	1	0	100	4.2
		Torque	39	0	34	4	1	0	10.5
		Speed	69	0	62	6	1	0	8.8

Note: density = density of support slurry; torque = cutterwheel torque; speed = jacking speed; TP = true positive; TN = true negative; FP = false positive; FN = false negative.

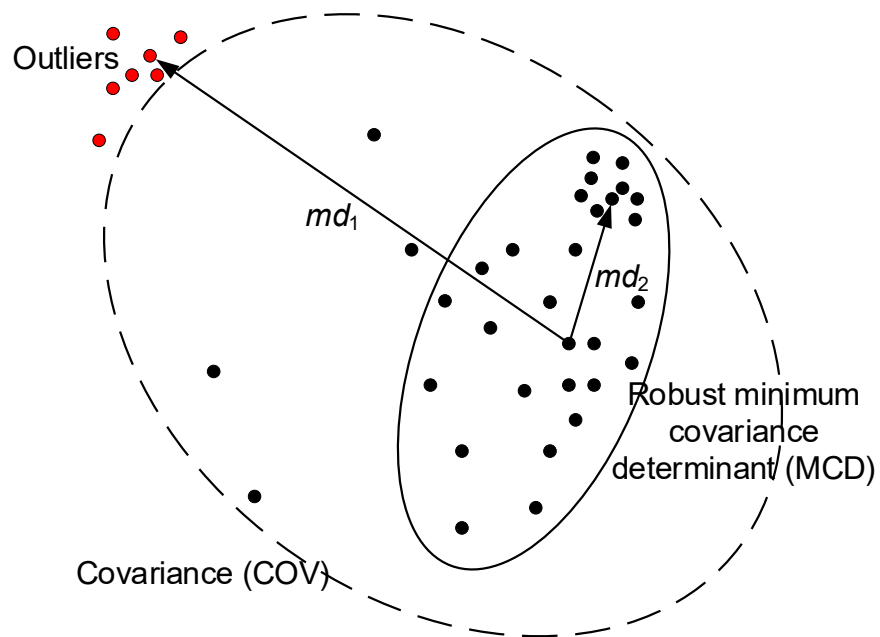




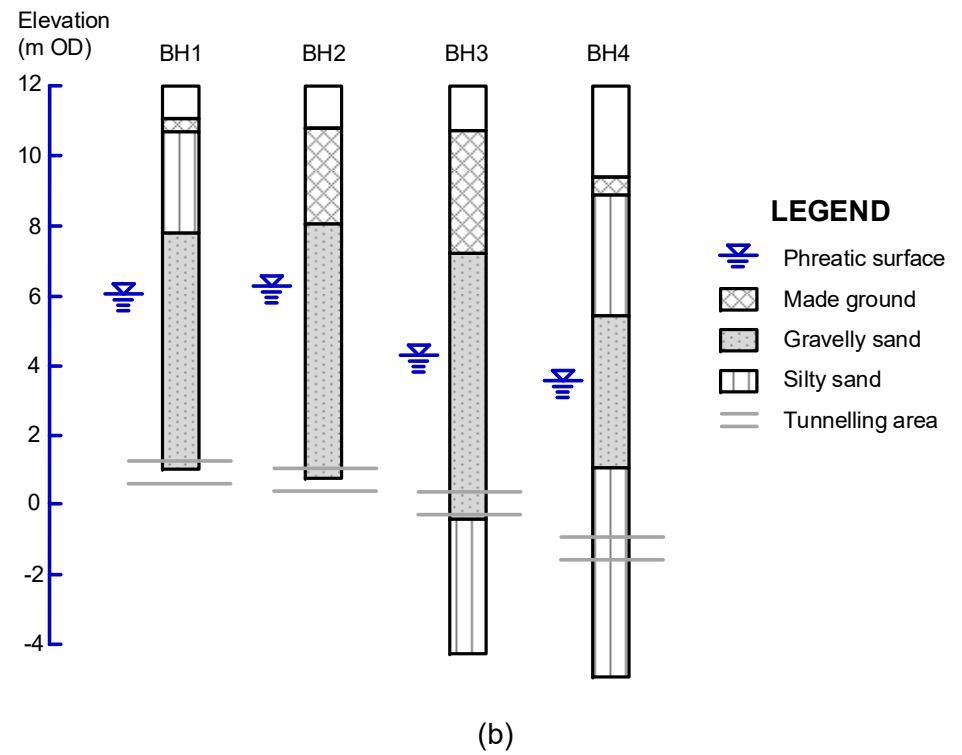
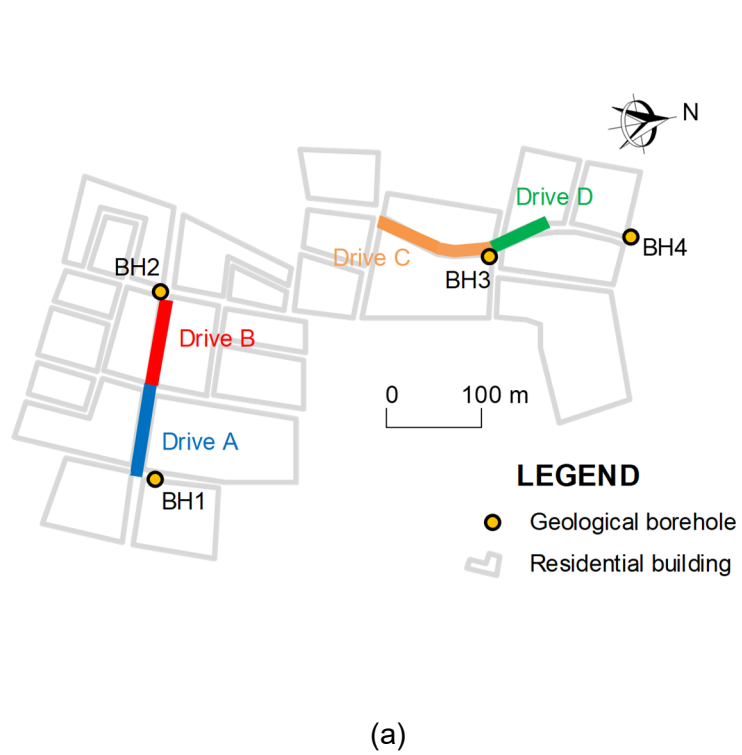
**Fig. 1** Illustration of one-class support vector machine approach and the construction of the hyperplane by transforming the original input space into a high-dimensional feature space



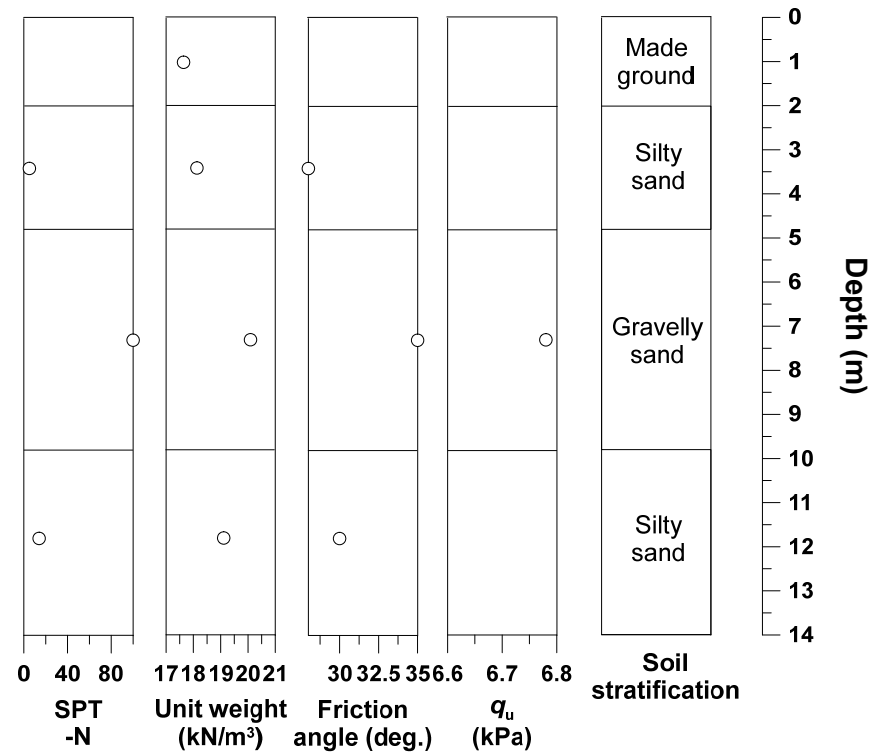
**Fig. 2** Illustration of isolation forest approach and the construction of separate isolation trees by randomly sampling from the training dataset. Red circles represent outliers, blue circles indicate uncommon inliers and light blue circles represent common inliers.



**Fig. 3** Illustration of robust covariance approach showing the use of a robust estimator of covariance to reflect the true organisation of observations using the Mahalanobis distance

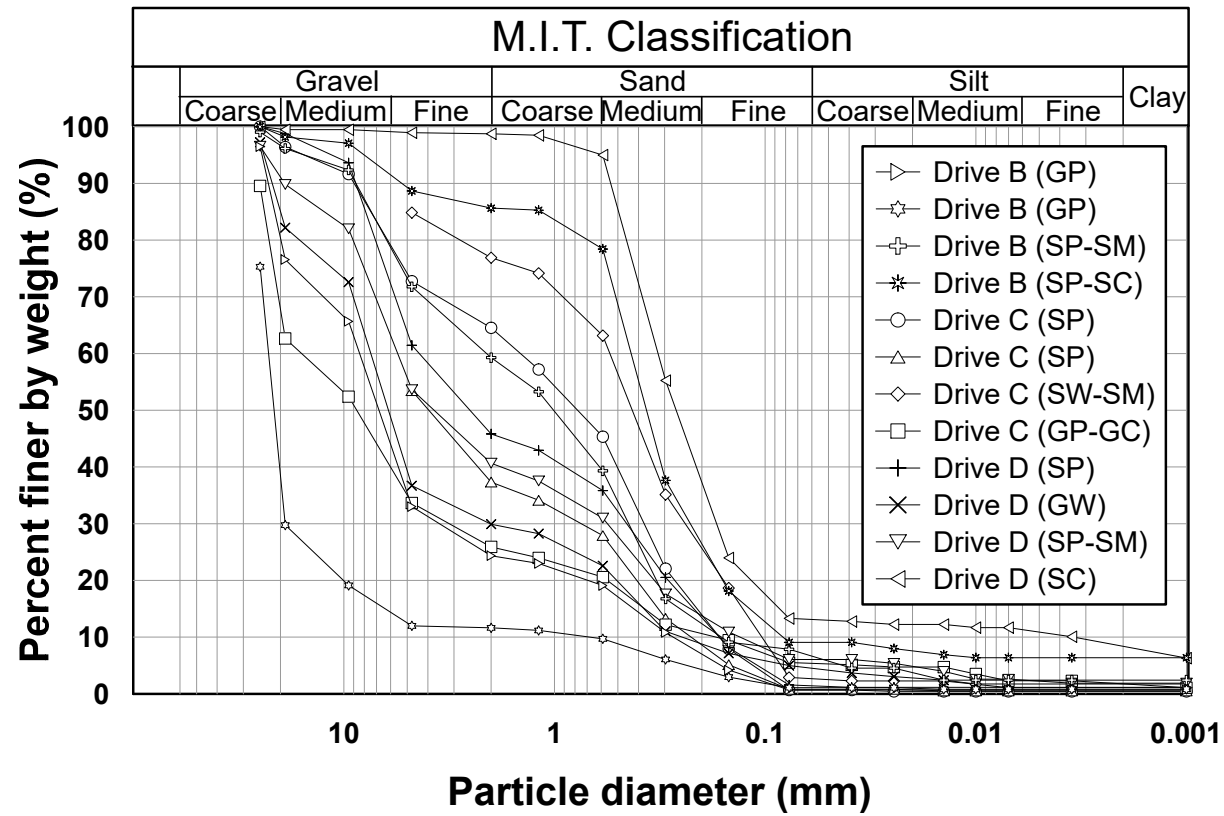


**Fig. 4** Pipejacking project description: (a) location of four drives in Taipei, Taiwan and (b) geological profile derived from four nearby boreholes

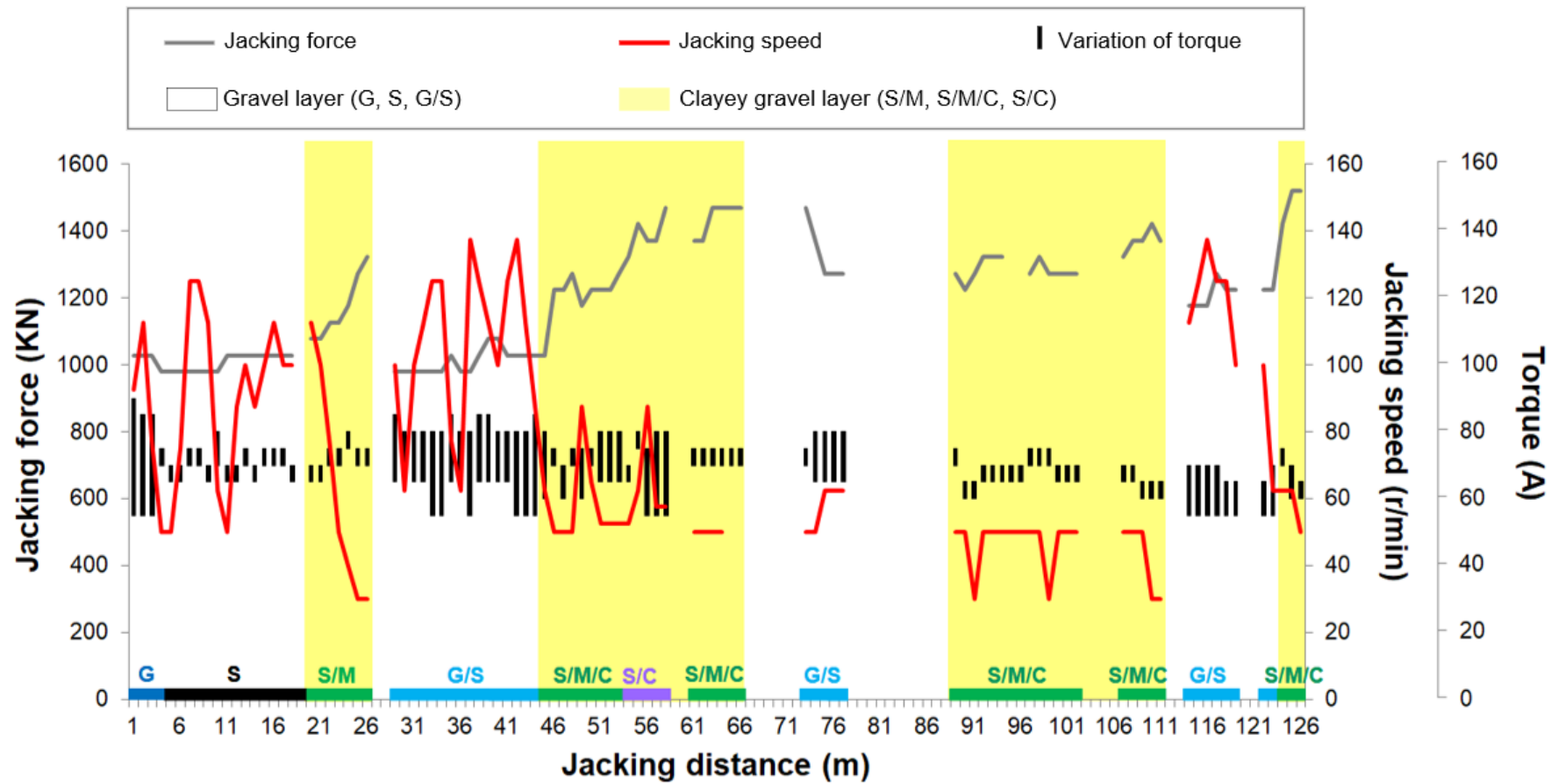


Note:  $q_u$ =Unconfined compressive strength; SPT-N=Blow count N value

**Fig. 5** Soil properties profile

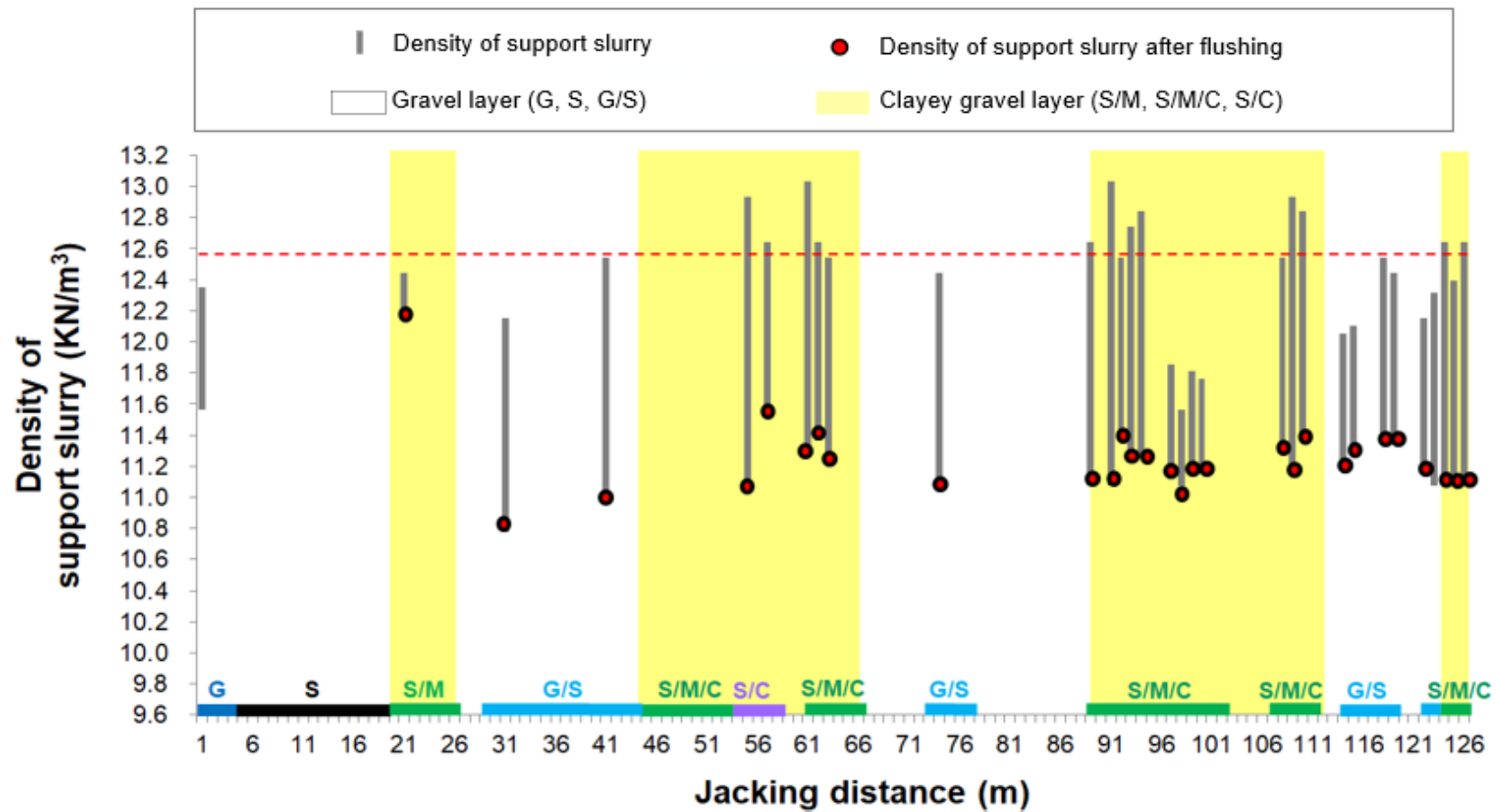


**Fig. 6** Grain-size distribution curves for drives B, C and D



(a)

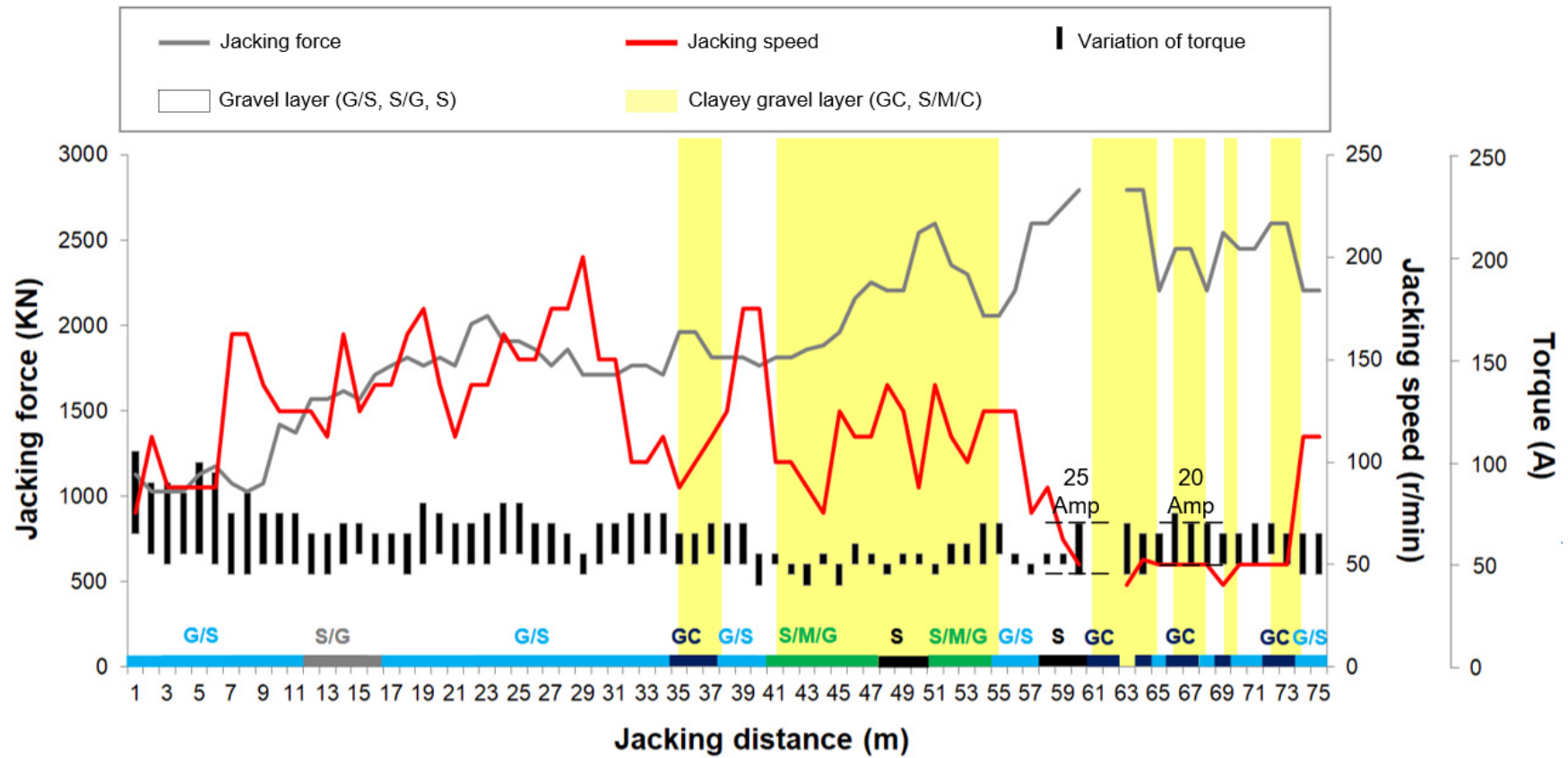
**Fig. 7** Monitored data for drive B: (a) development of jacking force, jacking speed and cutterwheel torque with jacking distance and (b) development of density of support slurry with jacking distance



(b)

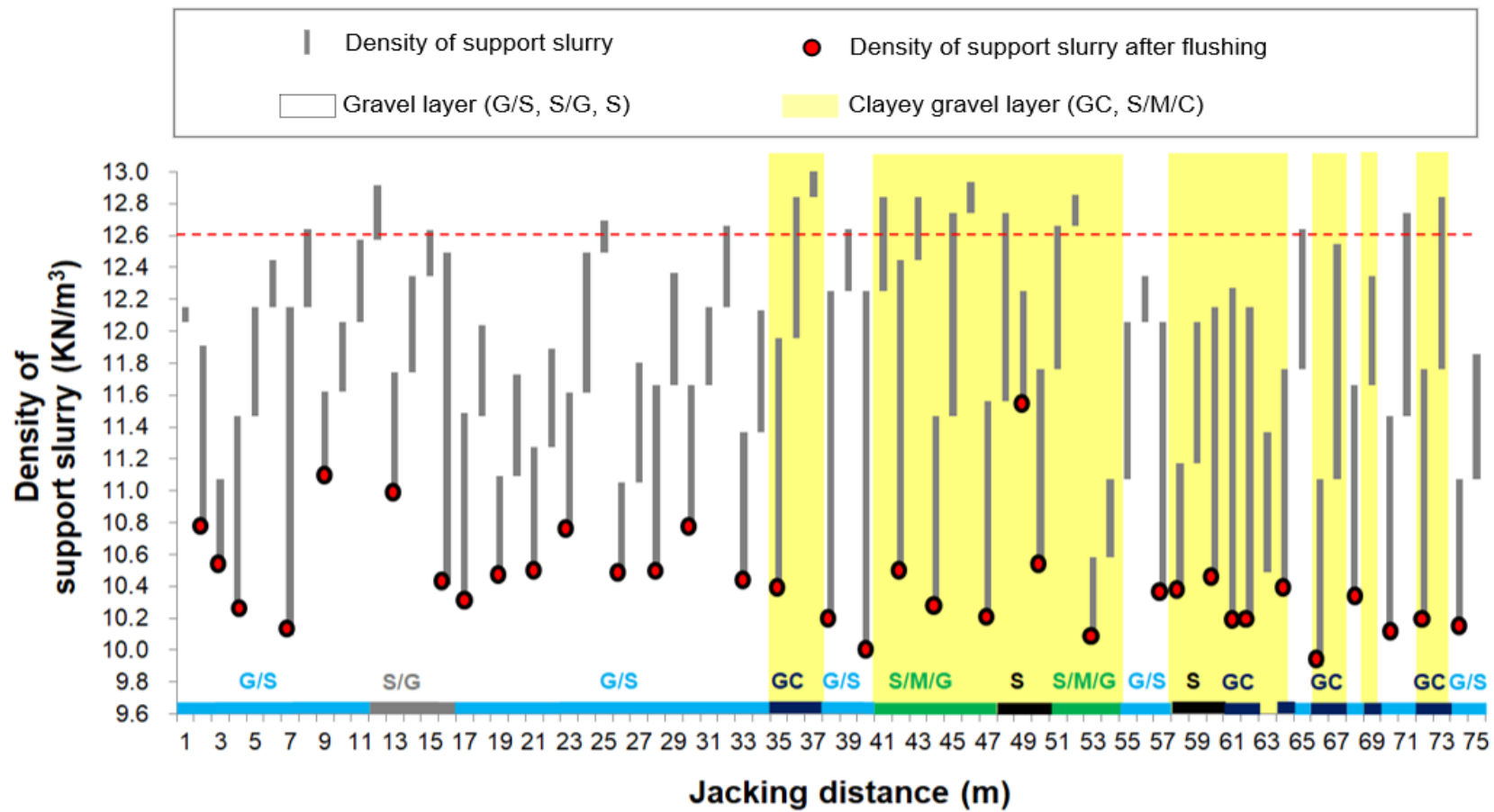
**Fig. 7** Monitored data for drive B: (a) development of jacking force, jacking speed and cutterwheel torque with jacking distance and (b) development of density of support slurry with jacking distance (cont'd)





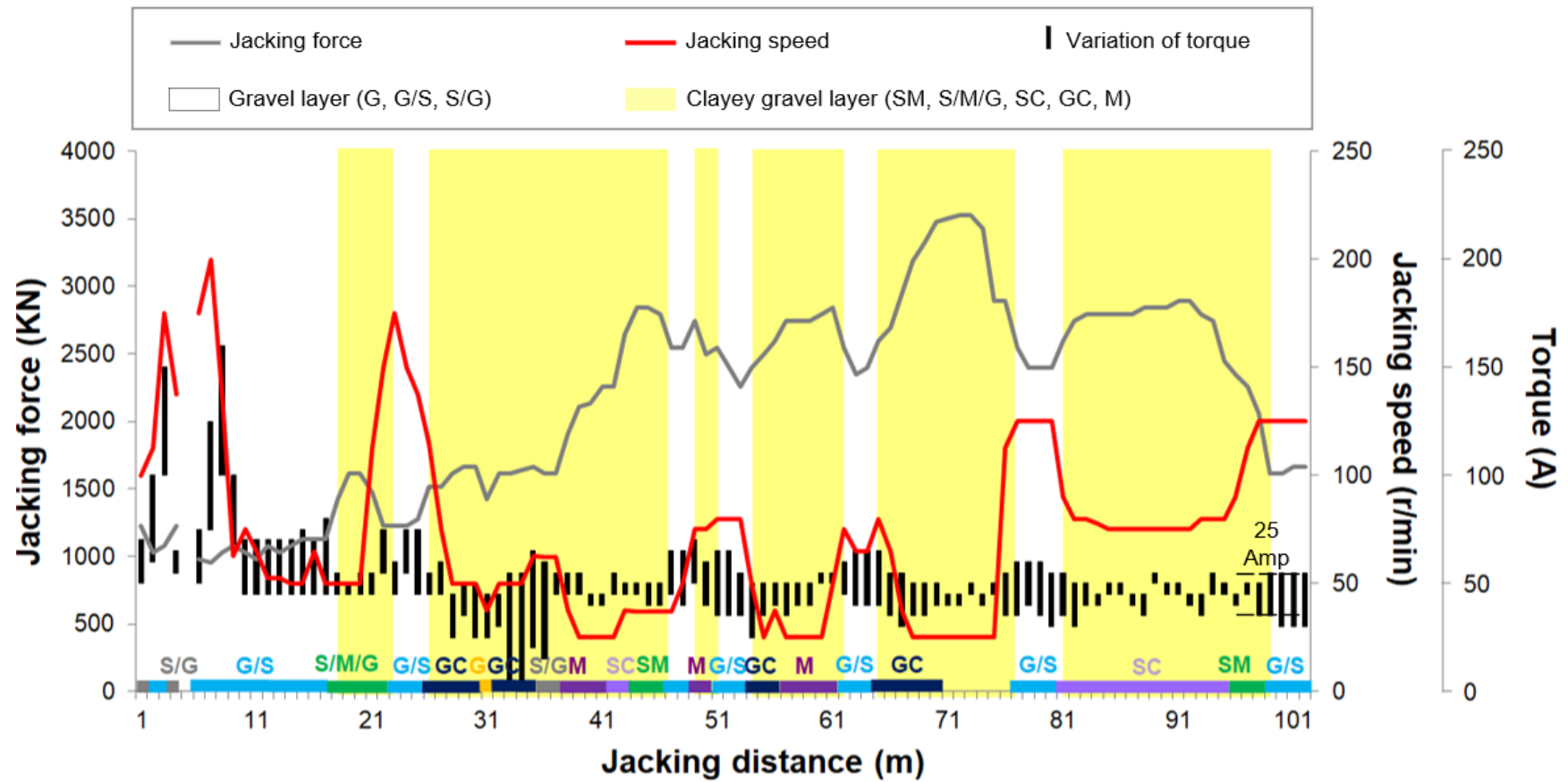
(a)

**Fig. 8** Monitored data for drive C: (a) development of jacking force, jacking speed and cutterwheel torque with jacking distance and (b) development of density of support slurry with jacking distance



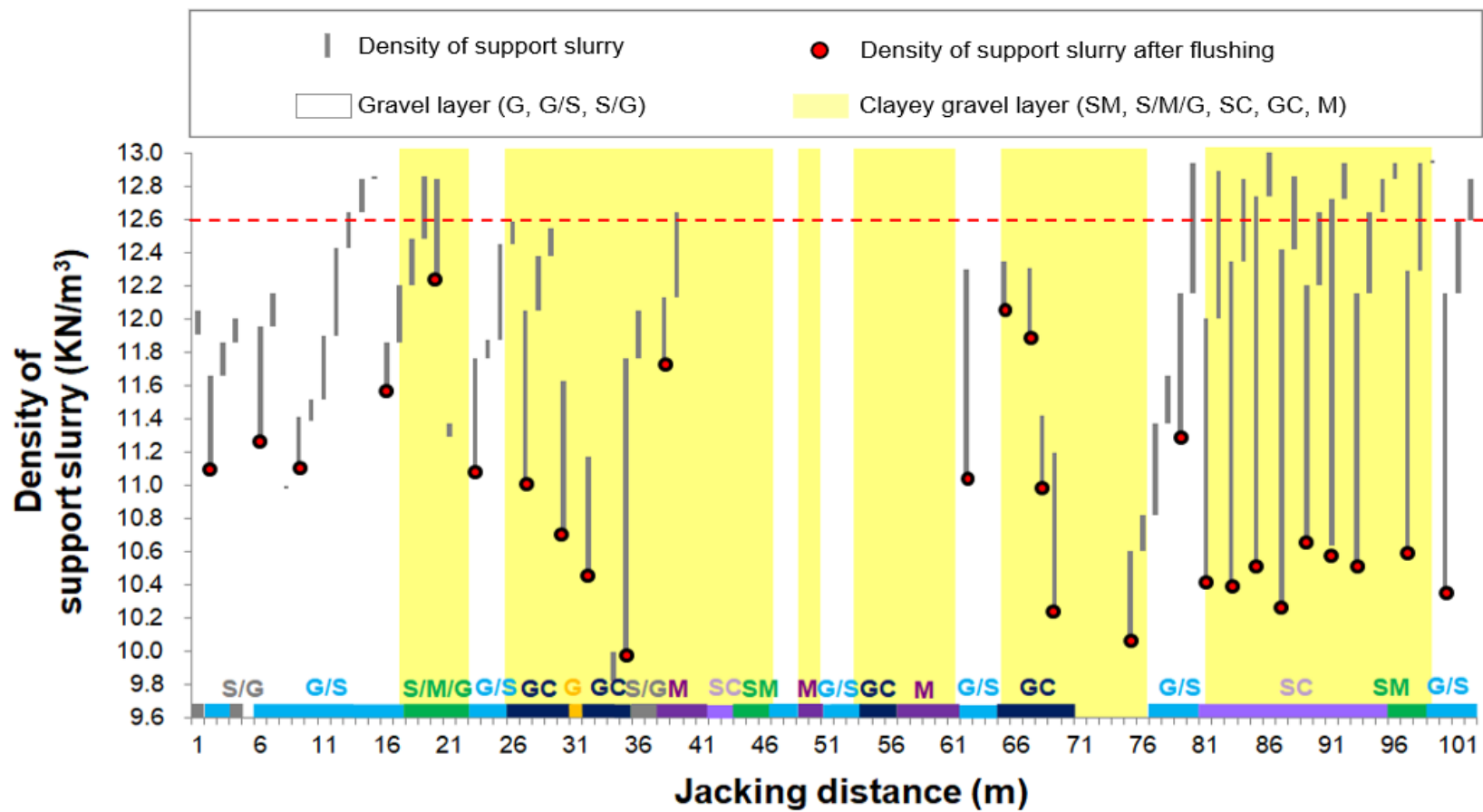
(b)

**Fig. 8** Monitored data for drive C: (a) development of jacking force, jacking speed and cutterwheel torque with jacking distance and (b) development of density of support slurry with jacking distance (cont'd)



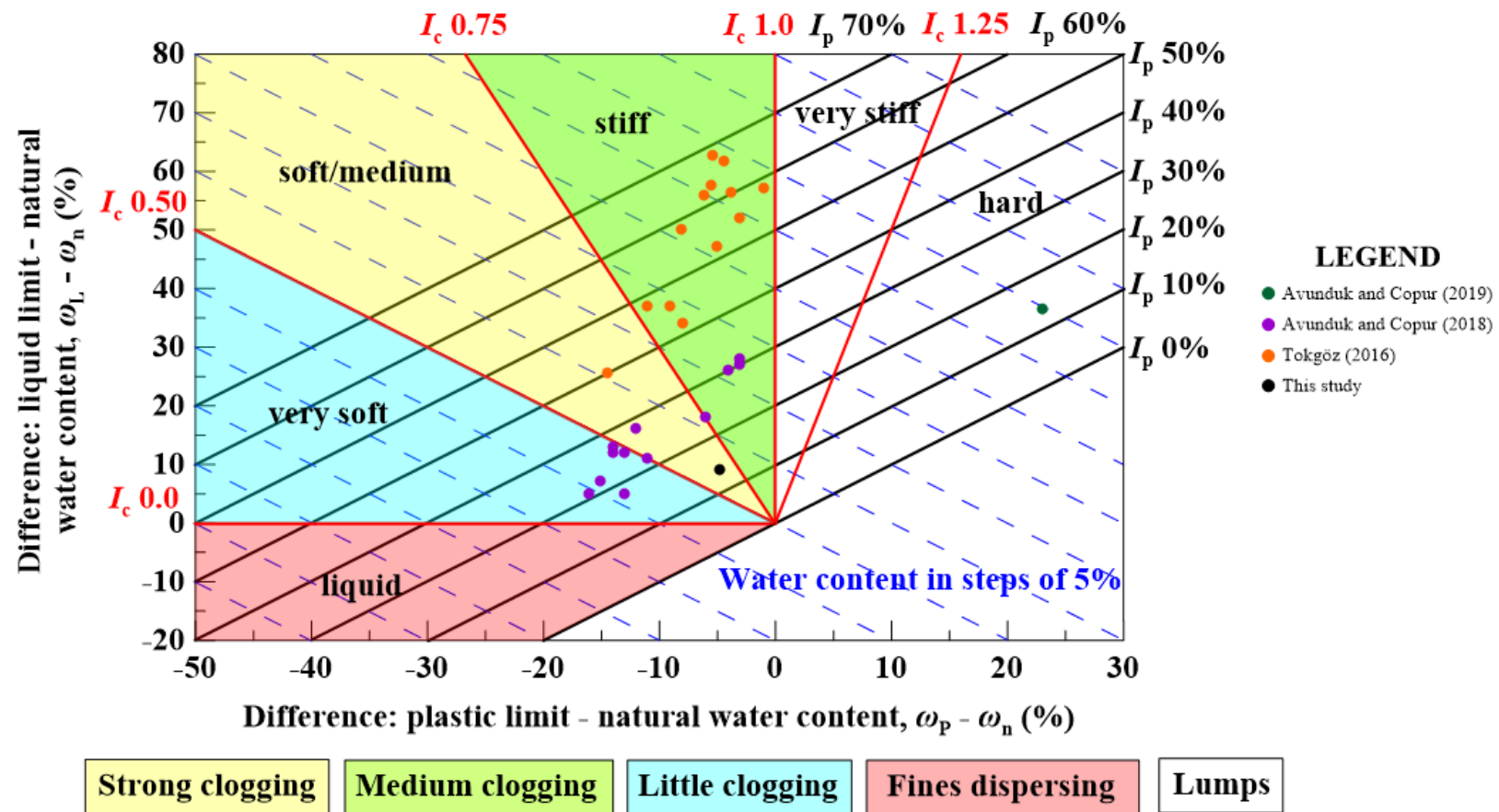
(a)

**Fig. 9** Monitored data for drive D: (a) development of jacking force, jacking speed and cutterwheel torque with jacking distance and (b) development of density of support slurry with jacking distance

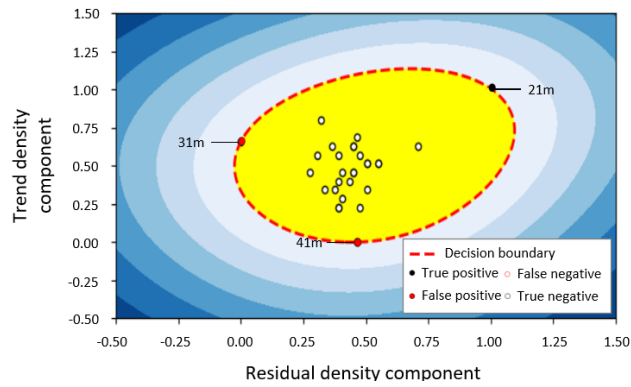


(b)

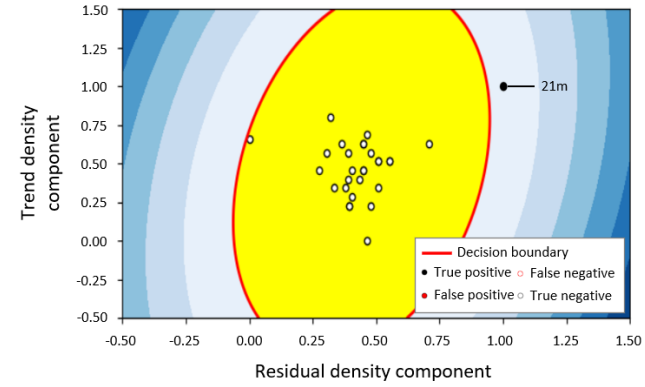
**Fig. 9** Monitored data for drive D: (a) development of jacking force, jacking speed and cutterwheel torque with jacking distance and (b) development of density of support slurry with jacking distance (cont'd)



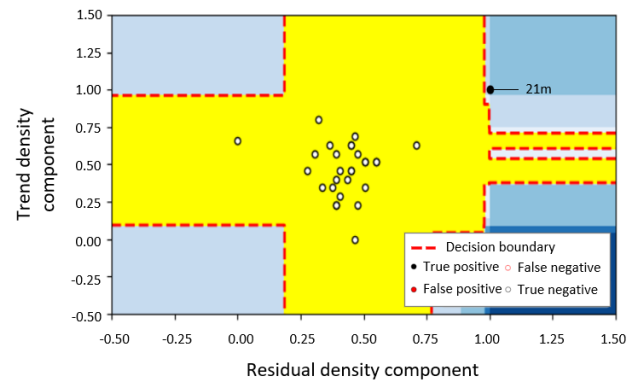
**Fig. 10** Universal classification diagram for the assessment of anomalous clogging behaviour



(a)

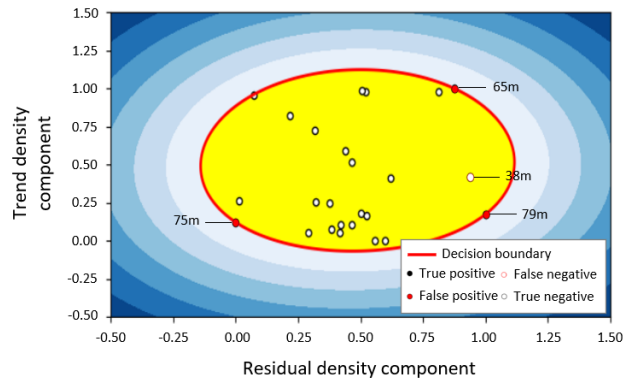


(b)

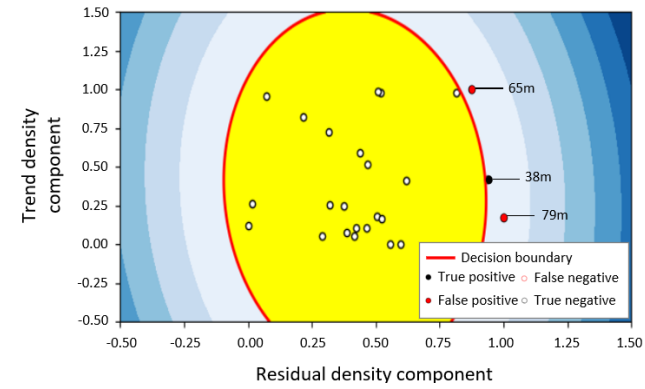


(c)

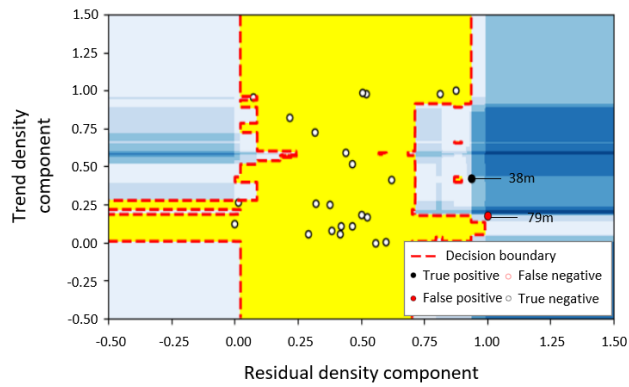
**Fig. 11** Performance of AD approaches applied to drive B using transformed slurry density space: (a) OCSVM, (b) Robcov and (c) IForest



(a)

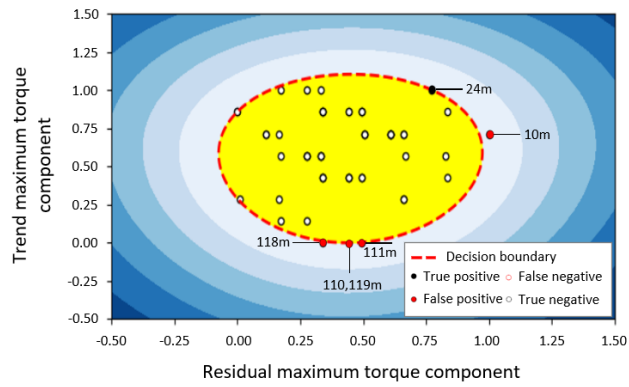


(b)

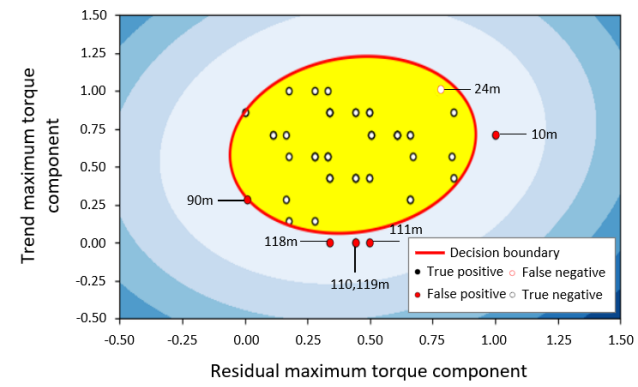


(c)

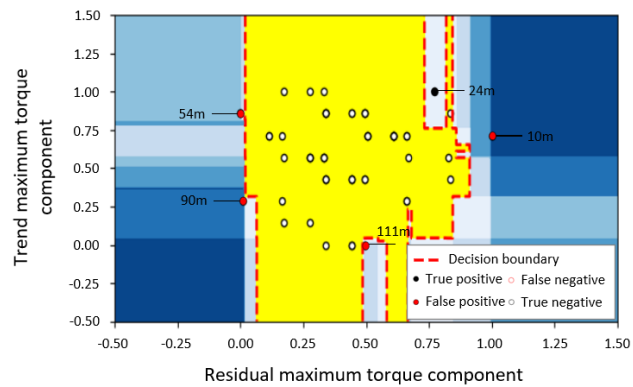
**Fig. 12** Performance of AD approaches applied to drive D using transformed slurry density space: (a) OCSVM, (b) Robcov and (c) IForest



(a)



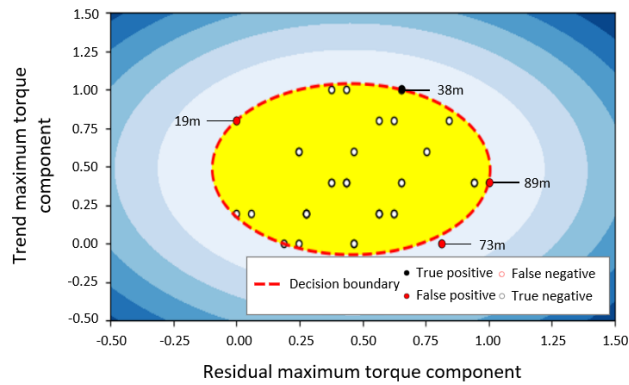
(b)



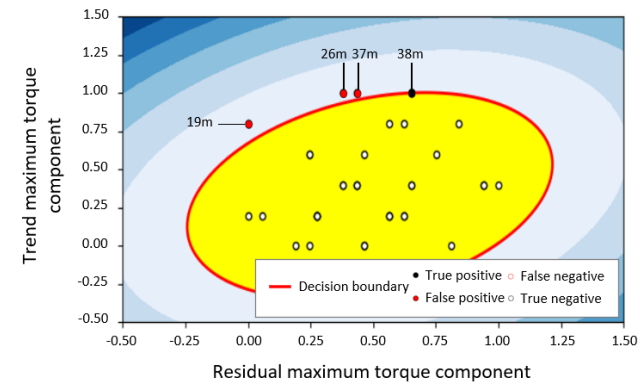
(c)

**Fig. 13** Performance of AD approaches applied to drive B using transformed maximum torque space: (a) OCSVM, (b) Robcov and (c) IForest

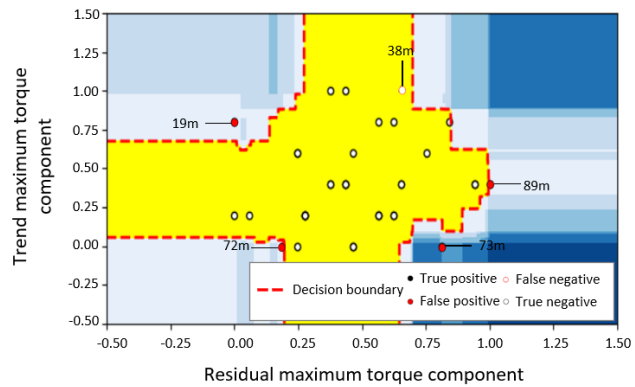




(a)

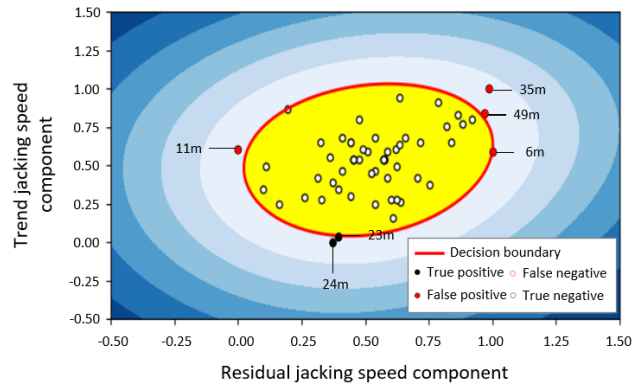


(b)

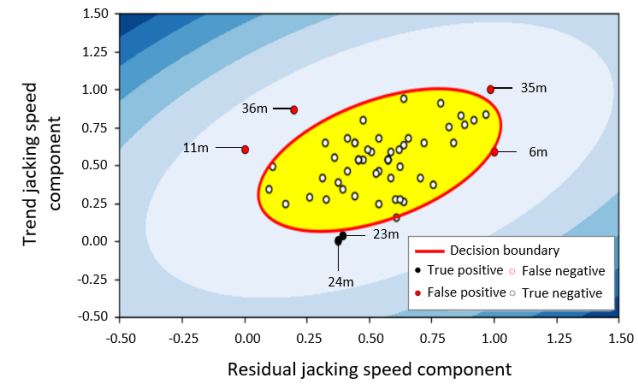


(c)

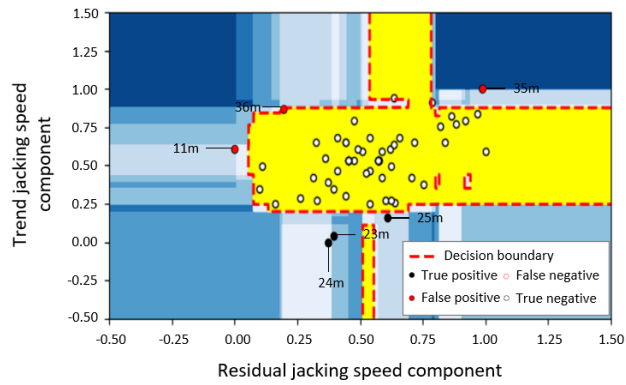
**Fig. 14** Performance of AD approaches applied to drive D using transformed maximum torque space: (a) OCSVM, (b) Robcov and (c) IForest



(a)

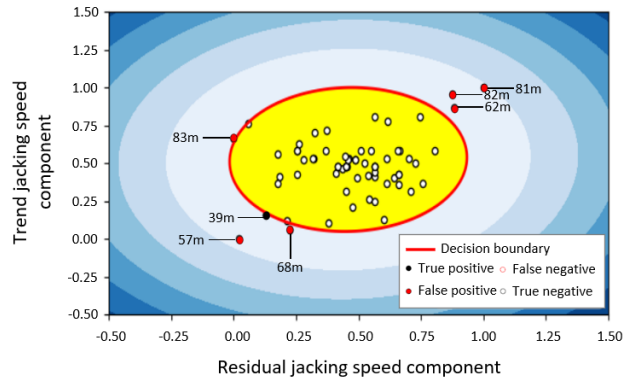


(b)

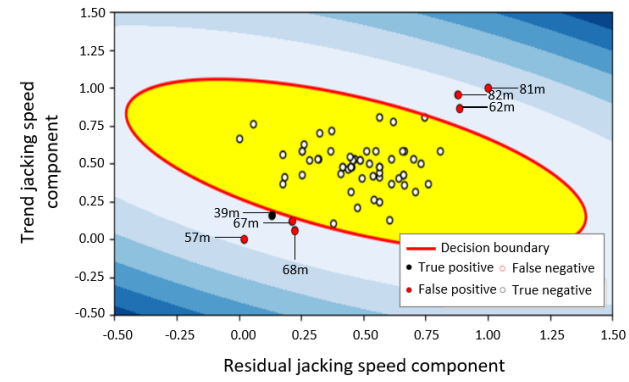


(c)

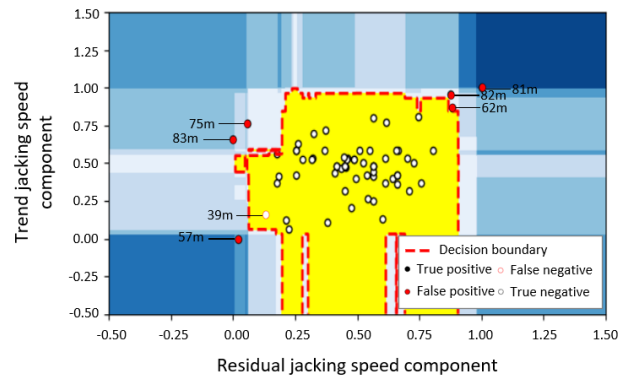
**Fig. 15** Performance of AD approaches applied to drive B using transformed jacking speed space: (a) OCSVM, (b) Robcov and (c) IForest



(a)

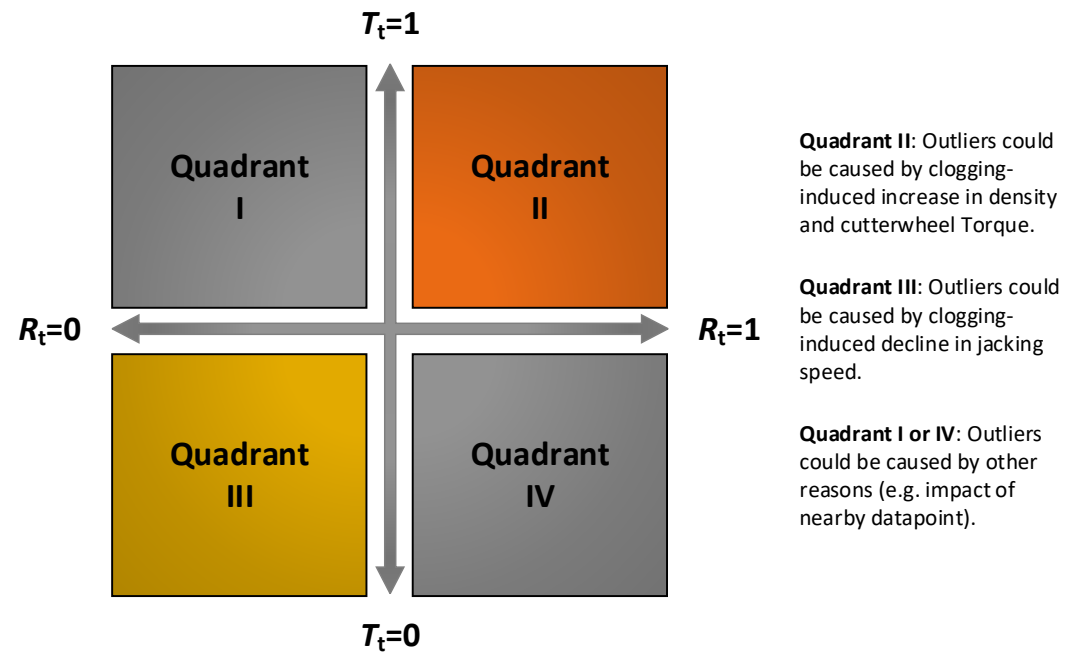


(b)



(c)

**Fig. 16** Performance of AD approaches applied to drive D using transformed jacking speed space: (a) OCSVM, (b) Robcov and (c) IForest



**Fig. 17** Simple diagram for TBM operator to assess tendency to clayey clogging during slurry-supported pipejacking

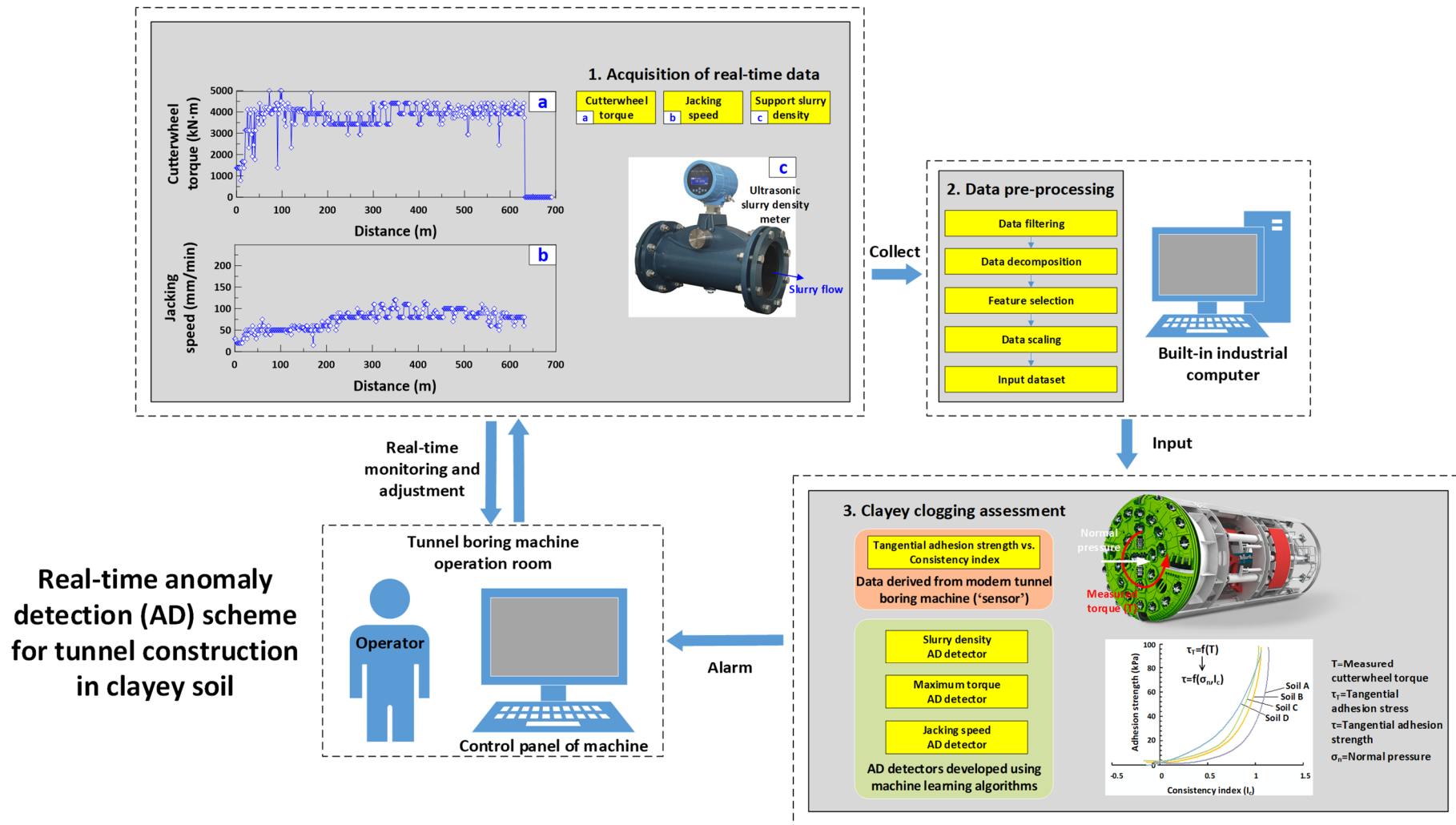


(a)



(b)

**Fig. 18** Clayey clogging with embedded gravel occurred while spanning through a ground composed of the clayey gravel and the gravel at drive  
B: (a) before and (b) after the clay soil adhered to cutterwheel



**Fig. 19** Real-time anomaly detection scheme for tunnel construction in clayey soil