Prediction of Box-Jacking Force Using a Probabilistic Observational Approach

YU, Bosong
Department of Earth Resources Engineering, Kyushu University

SHIMADA, Hideki
Department of Earth Resources Engineering, Faculty of Engineering, Kyushu University

SASAOKA, Takashi
Department of Earth Resources Engineering, Faculty of Engineering, Kyushu University

HAMANAKA, Akihiro
Department of Earth Resources Engineering, Faculty of Engineering, Kyushu University

他

出版情報：International Journal of the JSRM. 20 (2), pp.1-12, 2024-01-05. Japanese Society for Rock Mechanics
バージョン：

PROCEEDINGS PAPER
Proceedings of Japan-Korea Joint Symposium on Rock Engineering 2023
International Journal of the JSRM, Vol.20(2024), 240212
International Journal of the JSRM
Japanese Society for Rock Mechanics

Prediction of Box-Jacking Force Using a Probabilistic Observational Approach

Bosong YU[1*], Hideki SHIMADA[1], Takashi SASAOKA[1], Akihiro HAMANAKA[1], Fumihiko MATSUMOTO[2] and Tomo MORITA[2]

[1] Department of Earth Resource Engineering, Kyushu University (Motooka 744, Nishi-ku, Fukuoka, 819-0395, Japan)
[2] Alpha Civil Engineering (1-18, Sanno 1, Hakata-ku, Fukuoka 812-0015, Japan)
*E-mail:yu.bosong.572@s.kyushu-u.ac.jp

Box-jacking is an increasingly popular means for installing underground utilities and infrastructure. Accurately estimating the expected jacking forces in box-jacking is a key design concern, which can ensure the available thrust is not exceeded, to prevent damage to the box-culverts and/or launch shaft, and the construction efficacy of the jacking project. However, prediction of the total jacking force is complicated due to a multitude of influencing factors. The development of jacking force can be influenced by the site geology, the lubricant performance, work stoppages, shape of box culvert, and tunnel boring machine driving style. In this paper, a probabilistic observational approach is introduced aimed at prediction of jacking forces during the box-jacking process. Markov Chain Monte Carlo (MCMC) was adopted for this purpose which allows forecasts to be performed within a probabilistic framework. The proposed framework was applied to a box-jacking case histories completed in Kanagawa: a 150-m drive in fine and medium sands. The forecasts were appraised through comparisons to predictions determined using a classical optimization technique, namely genetic algorithms. The results show that the proposed framework yields highly accurate predictions for the monitored field data, and the prediction accuracy improves obviously as more data are acquired from the drive.

Key Words: *box jacking; Bayesian updating; observational method; jacking force; sampling; prediction*

1. INTRODUCTION

Rectangular box-jacking stands as a mature trenchless technology and an increasingly preferred alternative to the traditional open-cut method, typically utilized in the development of underground pedestrian passageways, subways, and utility tunnels that accommodate various services, including sewer, stormwater, fiber optics, and electricity pipelines. Box-jacking method offers multiple advantages, such as optimum space utilization, cost efficiency, shortened construction duration, and significant adaptability for tunnels with shallow overburden[1)]. In addition, the box-jacking approach is less intrusive to urban locales when compared to traditional methods like cut-and-cover[2)]. Accurately determining the jacking forces during the box-jacking process is a key design concern that dictates the configuration of pipe sections and significantly influences the efficacy of the engineering project[3)].

Presently, site personnel typically combine field experience with ad hoc theoretical relationships to develop prediction models for jacking forces, achieving some progress [4, 5)]. Much of this research has significantly informed the development of guidelines in pipe-jacking standards and handbooks, including those from the Pipe Jacking Association[6)], American Society of Civil Engineers[7)], French Society for Trenchless Technology[8)] and Japan Microtunnelling Association[9)]. However, the uncertainty involved in assessing

geotechnical condition parameters and underground conditions still to be a substantial obstacle to accurately forecasting construction behavior.

Updating uncertain geotechnical parameters during pipe jacking is instrumental in devising rational strategies to optimize construction operations during tunneling. For instance, the prevalent method to address uncertainty in jacking-force predictions includes the deployment of expensive interjacks at specified intervals along the pipe string, which are typically unused during the drive. The observation method is recognized as an effective limit states verification approach, leveraging the vast amounts of data available in the real world to train models and make predictions or insights[10]. It yields substantial savings in terms of materials, time, and costs[11]. This paper introduces a Bayesian updating method that utilizes the monitoring data acquired during the drive to update the uncertain model parameters for jacking force prediction. The proposed framework was employed in a box-jacking case history in Kanagawa, Japan: a 150-m drive in sand layers. In this case history, the key (uncertain) parameters impacting the forecast of jacking forces was determined. Contrasting with existing computation methodologies, the Bayesian updating approach can updates these parameters dynamically during the drive. As more data becomes available, the precision of jacking force predictions significantly improves. The jacking force forecasts using the Bayesian update approach were contrasted with those derived from classical optimization techniques, highlighting the respective advantages of both methodologies. The optimized variables, updated through Bayesian inference for predicting jacking force, can be employed as prior hypothesis for other future projects in similar ground conditions.

2. DETERMINISTIC JACKING FORCE PREDICTION MODEL

During box-jacking construction, the total jacking force can be divided into two distinct elements: the force applied to the face of the TBM and the frictional resistance between the structure, encompassing both the TBM and box string, and the surrounding soil. Hence, the total jacking force encompasses both a face resistance and a frictional resistance, can be represented as follows:

$$F_t = F_{face} + F_{fric} \tag{2a}$$

Prediction models can be categorized into either deterministic or probabilistic types. In pipe-jacked tunnel designs, jacking force estimations employ both deterministic empirical and theoretical models. These applications have yielded satisfactory predictive outcomes.

For the face resistance, JMTA proposed an empirical model for face resistance prediction during microtunnelling process[9]. To provide a better representation of field measurements, Shou et al. later adjusted this model based on empirical data[12], presenting it as:

$$F_{face} = 13.2 \cdot C_s \cdot N' \tag{2b}$$

where C_s is the outer perimeter of the pipe-jacking machine and N' is an empirical coefficient with values of 1 for clayey soils and increasing to 2.5 for sandy soils and 3.5 for gravelly soils.

In calculating the skin friction component, the current literature lacks empirical prediction approaches specifically for box-jacking. To bridge this gap, this paper adapts the established theoretical formula as follows:

$$F_{fric} = B \cdot L \cdot \tan\delta_v \cdot \sigma_v + 2 \cdot H \cdot L \cdot \tan\delta_h \cdot \sigma_h \tag{2c}$$

where $tan\delta_v$ and $tan\delta_h$ denote the interface frictional coefficients; the former on the box crown and the latter on the side surface of the box string; B designates the width, and H the height, of the box culvert; L is the length of the box string; and the vertical and horizontal earth pressures represented by σ_v and σ_h, respectively, can be calculated by modified Terzaghi's classical soil mechanics with considering the lubricant injection as shown in Eqs. (2d), (2e)[13]:

$$\sigma_v = \frac{(B_0\gamma - 2c)}{2K\tan\varphi}\left(e^{2K\tan\varphi\frac{h}{B_0}} - 1\right) + q_0 e^{2K\tan\varphi\frac{h}{B_0}} \tag{2d}$$

$$\sigma_h = K_a \cdot \sigma_v \tag{2e}$$

where c designates soil cohesion, φ the soil friction angle, and γ refers to the soil unit weight. q_0 is the surcharge load, while h represents the burial depth of the box crown. The silo width, B_0, is determined by $B + 2Htan(45° - \varphi/2)$. K denotes the soil pressure ratio, and $K_a = tan^2(45° - \varphi/2)$ is the active earth

pressure coefficient. Note that in the prediction model, the frictional resistance on the box string's bottom is not considered. Box culverts are typically designed to remain buoyant within lubricated overcuts. Continuous lubricant injection during jacking assures this buoyancy, even in unstable overcut scenarios[14]. This results in the bottom of the box string interacting primarily with the lubricant, making the friction between them negligible[15].

In Eqs. (2a)-(2e), the parameters of most interest are N', $tan\delta_v$, $tan\delta_h$, φ and γ. These are viewed as uncertain variables θ and undergo sequential updates during the drive, as shown in **Fig.1**. In practical applications, the update process is sequentially executed with a time delay, waiting for the availability of a new monitored datapoint.

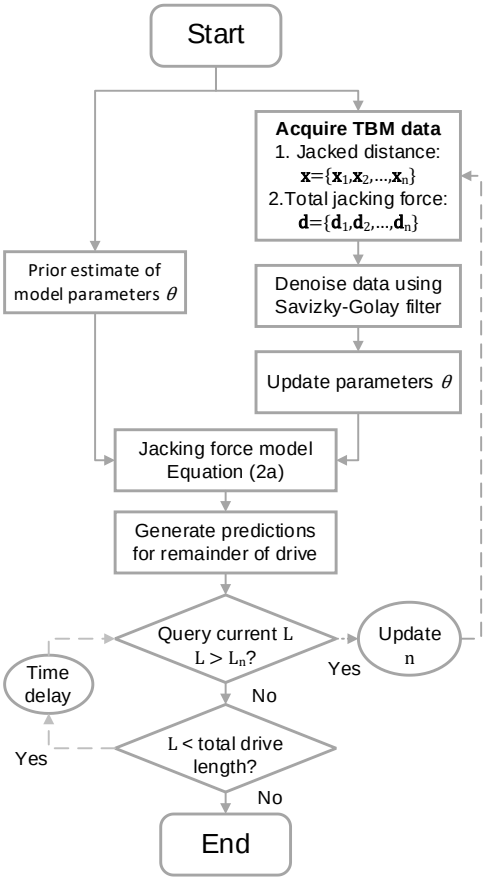


Fig.1 Parameter updating process for jacking force model.

3. METHOD FOR UPDATING THE PARAMETERS OF JACKING FORCE PREDICTION MODEL

(1) Markov Chain Monte Carlo

a) Framework for Bayesian inference

The framework initiates by considering the uncertainty inherent in jacking force prediction models. Deviation of model predictions from reality is typical in a wide range of geotechnical applications due to oversimplification of the real conditions[16]. Due to the model error, the jacking force estimated from any prediction models generally misaligns with actual observations. Therefore, calibrating the model using observed field performance or physical model tests to assess the model error is crucial. Existing literature has extensively documented the updating of geotechnical model parameters within a Bayesian framework to diminish discrepancies between model forecasts and observations from laboratory tests[17] and field monitoring[18]. In the Bayesian updating framework, a common implementation is to generate a deterministic model. In theory, if calculated values align with measurements, the points will lie along the line of equality. Nonetheless, uncertainties due to (1) soil properties, (2) geometry, (3) measurement error, and (4) imperfections and approximations in the calculation model should be incorporated into the updating framework to capture the scatter induced by these uncertainties[19]. These uncertainties are lumped together in a model bias factor, $\mathbf{e}$, applied to the deterministic prediction:

$$\mathbf{d} = f(\theta) + \mathbf{e} \tag{3a}$$

where $\mathbf{d} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_n]$ is the vector of monitored jacking force quantites at jacked distances $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$, with n being the total count of monitored data points utilized for Bayesian updating; $f(\theta)$ signifies the calculated jacking force determined by Eqs. (2a)-(2e), with $\theta = [N', tan\delta_v, tan\delta_h, \varphi, \gamma]$ representing the vector of uncertain model parameters; and the measurement error $\mathbf{e}$ is assumed to be statistically independent of the observation, and is believed to conform a normal distribution, characterized by a mean $\mu_{\mathbf{e}} = 0$ and standard deviation $\sigma_{\mathbf{e}} = 0$, as adopted in many practical fields[20].

The information update centers on the joint probability density function encompassing various parameters. When the likelihood and prior functions exhibit conjugacy, it becomes feasible to derive the posterior distribution analytically[21]. Nevertheless, due to the high dimension of the posterior distribution and the nonlinearity of the numerical model, the analytical

solution for conjugate pairs is usually nonexistent. In cases where probability distributions lack conjugate pairs for posterior distributions, techniques like Markov chain Monte Carlo (MCMC) become indispensable. MCMC adeptly balances computational demands with the versatility to use nonconjugate prior distributions and likelihood functions, ensuring accurate modeling in practical applications. Therefore, a multi-chain MCMC algorithm is applied to estimate and update the posterior joint probability density function (PDF) of the model parameters θ in this study.

Most geotechnical parameters can be regarded as random variables following to the normal, lognormal, or other theoretical distributions. Lognormal distributions are widely adopted in geotechnical engineering because they span the range $(0, +\infty]$, preventing unrealistic negative outputs of the nonnegative geotechnical variables. Thus, we employ lognormal distributions to establish the prior probability distribution of θ. In this paper, uncertainties related to the model parameters θ are represented by a prior distribution. This is presumed to be a multivariate normal distribution with μ'_θ signifying the prior mean and σ'_θ representing the uncertainty of the natural logarithm of θ. In formal terms, $p(\theta') = \text{MVN}(\mu'_\theta,\ \sigma'_\theta)$ with θ' being the natural logarithm of uncertain model parameters θ and MVN denoting a multivariate normal distribution. The prior probability distribution of the random variable θ' is

$$p(\theta' \mid \mu_\theta', \sigma_\theta'^2) = \frac{1}{\sqrt{2\pi\sigma_\theta'^2}} e^{\frac{1}{2}\left[\frac{(\theta'-\mu_\theta')}{\sigma_\theta'}\right]^2} \tag{3b}$$

b) Markov Chain Monte Carlo simulation using No-U-Turn sampling procedure

Using evidence from the monitored data, the prior lognormal distribution is adjusted to obtain the posterior distributions of θ. The data $\mathbf{d}$ affects the posterior distribution via the likelihood function $p(\mathbf{d}|\theta, \mathbf{x})$. This function represents the likelihood of the data observed when applied to the existing model with specific values of the parameter vector θ

$$p(\mathbf{d} \mid \theta, \mathbf{x}) = \prod_i p(\mathbf{d}_i \mid \theta, \mathbf{x}) \tag{3c}$$

The posterior distribution, represented by $p(\theta|\mathbf{d}, \mathbf{x})$, for the model parameters is derived from Bayes' theorem as illustrated below

$$p(\theta \mid \mathbf{x}, \mathbf{d}) = \frac{p(\mathbf{d} \mid \theta, \mathbf{x})\, p(\theta \mid \mathbf{x})}{p(\mathbf{d} \mid \mathbf{x})} \tag{3d}$$

The function $p(\mathbf{d}|\mathbf{x})$ normalizes the joint posterior distribution so that its integral amounts to 1. This normalization is accomplished by marginalizing out θ, detailed as below:

$$p(\mathbf{d} \mid \mathbf{x}) = \int p(\mathbf{d} \mid \theta, \mathbf{x})\, p(\theta \mid \mathbf{x}) d\theta \tag{3e}$$

Analytically resolving Eq.(3d) is often impossible. To address this, the robust sampling method, Markov Chain Monte Carlo (MCMC), is adopted. While the Metropolis-Hastings algorithm is a favored MCMC sampling approach in geotechnical engineering, as evidenced by works like Qi and Zhou[22], the Hamiltonian Monte Carlo (HMC) offers a more efficacious alternative, especially suitable for high-dimensional spaces, as noted by Betancourt[23].

HMC sampling involves of two stages: the proposal and the correction. In the proposal step of HMC, Hamiltonian dynamics are employed to navigate through the parameter space. This involves transforming the posterior density function to a Hamiltonian potential energy function and introducing a momentum variable h_j for each uncertain parameter θ_j. This energy function, termed the canonical distribution, defines a joint distribution of θ and h, represented as $p(\theta|h, \mathbf{d})$. Owing to the independence of these parameters, the joint distribution can be expressed as the product of the posterior density, $p(\theta|\mathbf{d})$ and the momentum density, $p(h)$; that is, $p(\theta|h, \mathbf{d}) = p(\theta|\mathbf{d})p(h)$. Starting from the current parameter state θ and introducing this random momentum, the dynamics, based on the Hamiltonian equations, dictate a trajectory in the parameter space, leveraging the structure and gradients of the potential energy function. This trajectory results in a new sample proposal. However, it's not the dynamics themselves that directly sample from the posterior, but they aid in generating these proposals. After the proposal step, the correction phase comes into play, which rejects any proposal that strays too far from the typical set of the target distribution, ensuring the samples closely approximate the desired posterior.

Due to the absence of the analytical solution for Hamilton's equations in most practical situations, numerical methods become indispensable for their

resolution. The leapfrog method stands out as a preferred choice for this purpose, being a time-reversible and volume-preserving numerical integrator. The leapfrog approach is particularly suited for HMC since it proposes moves to new points in the state space while maintaining certain desirable properties. One of its key advantages is its ability to propose moves to distant states in the parameter space. This characteristic helps reduce the autocorrelation between successive sampled states, resulting in more efficient exploration and faster convergence to the target distribution. Furthermore, by preserving the system's energy through its symplectic nature, leapfrog integration ensures a high acceptance probability for these distant proposals, making the sampling process more efficient. The procedure for MCMC sampling using the HMC algorithm is summarized as follows. Additional details can be found in work of Brian B. Sheil[24].

Step 1: Initialization

A candidate momentum θ is drawn from its posterior distribution, $h \sim N(0, \Sigma)$, where Σ is the covariance matrix of $p(h)$.

Step 2: Leapfrog integration

a) The leapfrog method begins with a half-step to update the momentum:

$$h \leftarrow h + 0.5m \frac{d \log p(\theta \mid \mathbf{d})}{d\theta} \tag{3f}$$

The Eq.(3f) represents the half-step momentum update using the gradient of log-posterior density $p(\theta|\mathbf{d})$.

b) A full step is then taken to update the position (parameter vector θ):

$$\theta \leftarrow \theta + mh \tag{3g}$$

This is the full-step position update which uses the current momentum to propose a new position.

c) Another half-step for momentum (same as the first) completes the leapfrog process. This second half-step uses Eq.(3f) again. This iterative process is executed for L steps, which each step having a size m.

Step 3: Correction phase

After the leapfrog steps, a Metropolis-Hastings acceptance criterion is applied to decide whether to accept the proposed new state:

$$\alpha = \min\left(\frac{p\left(\theta^{(k-1)} \mid \mathbf{d}\right) p\left(h^{(k-1)}\right)}{p\left(\theta^{(k)} \mid \mathbf{d}\right) p\left(h^{(k)}\right)}, 1 \right) \tag{3h}$$

If the proposal is not accepted, then the next position is set to the previous position $\theta^{(k)} = \theta^{(k-1)}$.

Step 4: Iterations

Steps 1-3 undergo repetition for n_{iter} iterations within a chain until convergence achieves. Multiple chains can be utilized for better results; hence, three chains were adopted with $n_{iter} = 5000$. Notably, the initial samples do not rigorously conform to the posterior PDF, as the Markov Chain has not yet reached steady state. Thus, discarding certain initial samples as "burn-in" samples is a rational approach. In this study, an initial “burn-in” of 3,000 samples was used to ensure the Markov chain enters a high probability region, the remaining samples are used to approximate the target distribution. The selected parameters, determined via trial and error, strike an optimal equilibrium between accuracy and computational demands.

Traditional Hamiltonian Monte Carlo (HMC) showcases efficiency in sampling intricate posterior distributions but demands manual tuning of the leapfrog parameters m and L. To address this, this study employs the No-U-Turn Sampler (NUTS), an advanced variant of HMC that automatically adjusts these parameters[25]. Operating through an intricate tree-building algorithm, NUTS dynamically decides the number of leapfrog steps, eliminating the need for a predetermined trajectory length. It embarks on expansive jumps from its initiation, venturing into uncharted regions of the posterior distribution, ensuring a rapid and efficient convergence. In practice, NUTS adjusts the integration step size based on its association with the acceptance ratio depicted in Eq. (3h). This adaptability ensures the trajectory unfolds in a direction opposite from its starting point. The NUTS trajectory begins with a single leapfrog step from the current state, subsequently doubling the number of steps until a U-turn is detected, as illustrated in **Fig.2**. Each doubling phase involves selecting a direction—forward or backward in time—randomly, followed by simulating the Hamiltonian dynamics for 2^j leapfrogs, with j representing the current iteration. This systematic expansion signifies the depth of the tree structure. The framework used in this work was developed using the PyMC (version 5.6.1) programming library in Python (version 3.8)[26].

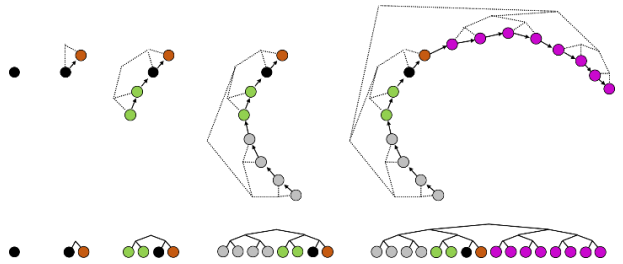


Fig.2 Binary tree construction through repeated doubling with randomized hamiltonian dynamics simulation

(2) Classical Optimization: Genetic Algorithms

To benchmarking, the same uncertain parameters θ are updated employing classical optimization methods. Herein, Genetic Algorithms (GAs) are selected as they are among the most widely used optimization techniques in geotechnical engineering[27]. Holland (1992) first introduced the concept of GA approach[28], which was initially designed to simulate the mechanisms of population genetics and the natural principles of survival, pursuing the ideas of adaptation. The main steps to optimize the uncertain parameters θ in this study are described as follow

Step 1: Initialization

A typical GA begins with an initial set of random solutions, each representing distinct nodes within the search space. This collection is referred to as a population, and in this work, it has a size of $n_{ps} = 20$. To ensure a broad search space for optimization, the initial population was generated by randomly sampling from a 99% confidence interval (CI) of θ's prior distribution.

Step 2: Fitness evaluation

In each iteration, the fitness of every individual was assessed within the current population using a specific fitness function. In this work, the coefficient of determination (R^2) is used as the fitness function, which measures the relationship between the predicted and measured jacking forces up to the current jacked distance. To enhance convergence properties, linear scaling of the fitness values was adopted as recommended by Goldberg[29].

Step 3: Genetic operation

Various genetic operations were used to evolve and optimize solutions over multiple generations. Selection is the process where individuals from the current population are chosen to form the next generation's population. The likelihood of a candidate solution being chosen is proportionate to its scaled fitness value within the population. Once selected, crossover is performed on these individuals. In this operation, portions of individual solutions are randomly swapped based on a crossover probability $p_c = 0.6$. The goal of crossover is to generate a potentially optimal variant of the individual by combining attributes of two parent solutions. To ensure that GA retain some freedom to explore the wider search space, mutation was implemented. During mutation, specific values within a solution are occasionally and randomly changed, governed by a mutation probability $p_m = 0.05$. The common genetic operators used in this study are illustrated in **Fig.3**.

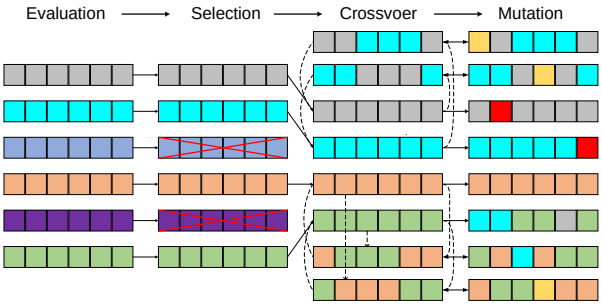


Fig.3 Overview of the genetic operators employed in the GA opptimazation process.

Step 4: Iterations

The above procedure repeats, creating new populations based on the evaluations and genetic operations until reaching a set maximum number of generations m_{iter}.

The GA presented in this study was constructed using Python, and its optimal performance parameters with draw upon recommendations from Goldberg[29] as detailed in Table 1.

Table 1 Performance parameters of GA Based on Goldberg[29].

Parameter	Value
Number of individuals per population, n_{ps}	20
Number of generations, m_i	50
Probability of crossover, p_c	0.6
Probability of mutation, p_m	0.05

4. CASE HISTORY AND TUNNELLING DETAILS

The case history used to assess the proposed

parameter updating framework was a 150-m-long drive undertaken in Tsujido coast of Kanagawa, JP (**Fig.4**). **Fig.5** shows the ground conditions determined from three boreholes (BH5, BH6 and BH7) located on the drive route; profiles of standard penetration test (N) values are superimposed on the figure. The ground conditions comprised a 0.4-0.8 m layer of made ground at the ground surface, followed by fine sand (1.4-3.0 m), medium sand (2.0–3.3 m) and gravel (> 2.0 m). The groundwater table was found to be 2.0 m above mean sea level (MSL). The launch invert for the drive was 4.1 m below ground level (GL), and the overburden depth from the tunnel crown varied between 1.0 and 1.95 m. Tunnelling therefore occurred predominantly through the alternating layers of fine and medium sands. A Earth Pressure Balance (EPB) shield machine with a lining of 4750 * 2230 mm was utilized for the tunnelling, while the outer dimensions of the trailing box culvert string measured 4700 * 2200 mm. This resulted in a 25 mm overcut annulus at both the top and bottom of the box, and a 15 mm overcut on its sides. The overcut gap was filled with a two-component consolidated lubricant named Cantor-SS, widely employed in Japan, during tunnelling to maintain tunnel bore stability and minimize friction between soil and the box sting. Each box culvert had a weight of 150 kN and a length of 1.5 m. The output data from the TBM was recorded at 2-minuts intervals. The development of total jacking force with jacked distance monitored for the drive is presented in **Fig.6**.

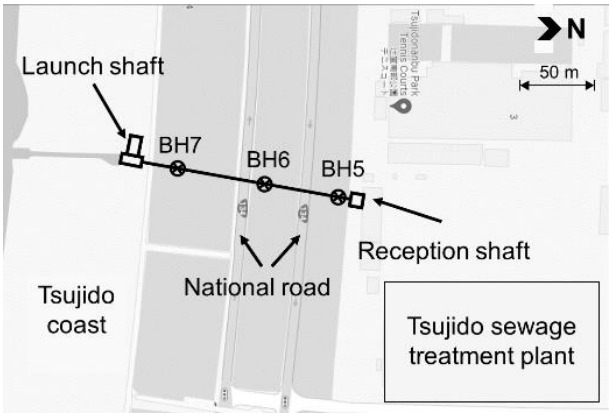


Fig.4 Location of tunnel and shafts at the Tsujido site.

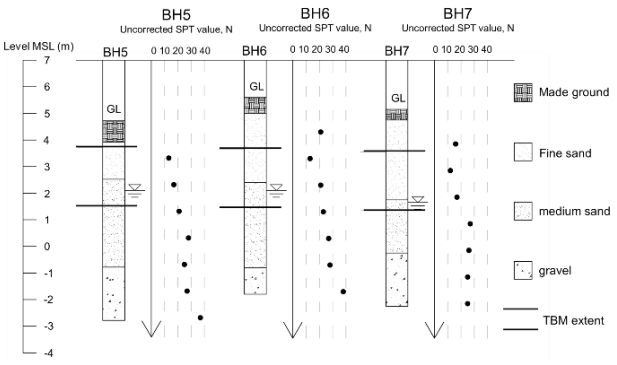


Fig.5 Stratigraphy at the boreholes along the tunnel drive.

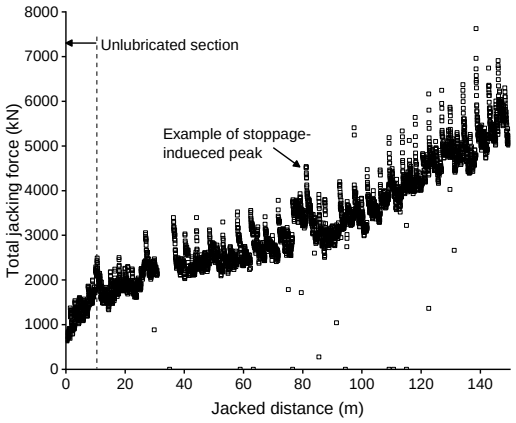


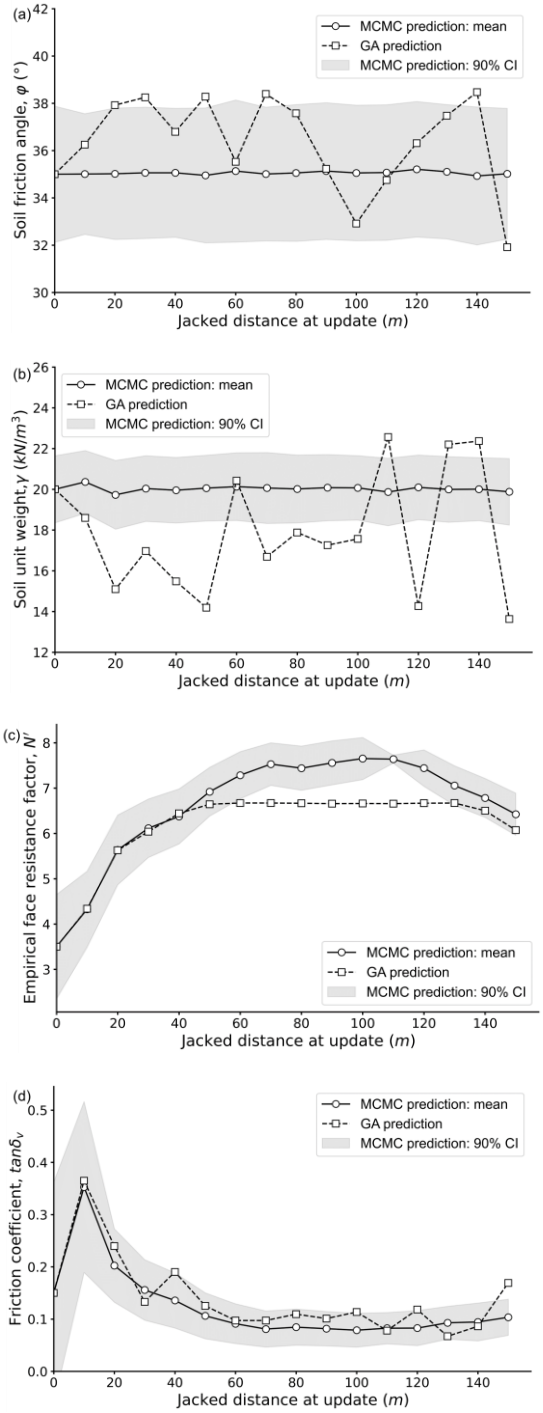
Fig.6 Recorded Jacking force data of Tsujido project.

5. RESULTS AND DISCUSSION

The prior best estimates for the mean and standard deviation of the uncertain model parameters θ [Eq.(3b)] and model error **e** are shown in Table 2. The basis for the prior hypothesis determination is as follows: The mean values for the piror empirical face resistance factor N' are based on Shou et al. with an assumed weak coefficient of variation (CV)[12]. Both the mean and CV for the prior hypothesis of interface friction coefficient $tan\delta$ are derived from Staheli, Shou et al., and Namli and Guler[12,30,31]. The prior soil friction angle φ and soil unit weight γ have their mean values sourced from on-site measurements[12], with their CV is based on Harr and Kulhawy[32,33]. Lastly, the model and measurement error **e** adopts its mean from monitored data by O'Dwyer et al. and presumes a weak CV[14].

Table 2 Prior distribution od model parameters.

Parameter	Mean, μ	CV	Distribution
Soil friction agnle φ (°)	35	0.075	Lognormal
Soil unit weight γ (kN/m^3)	20	0.05	Lognormal
Empirical face resistance factor N'	3.5	0.2	Lognormal
Interface friction coefficient $tan\delta$	0.15	0.86	Lognormal
Model and measurement error $\mathbf{e}$	100	0.2	Normal



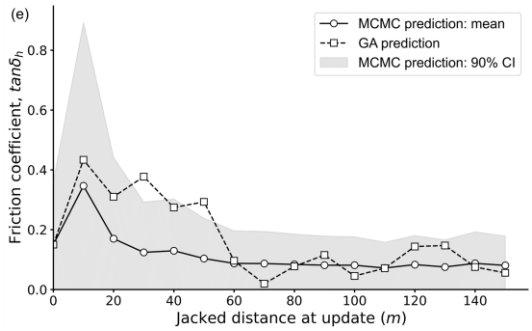
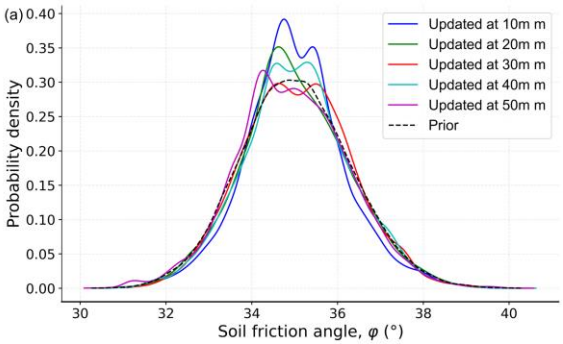


Fig.7 Comparing the variation of updated model parameters using MCMC and GA approaches: (a) friction angle φ; (b) soil unit weight γ; (c) empirical face resistance factor N'; (d & e) interface frictional coefficients $tan\delta_v$ & $tan\delta_h$.

In **Fig.7**, the outcomes of the parameter updating process employing the MCMC method are compared with those determined through the GA optimization method, with the shaded regions delineating the 90% CI for the parameters updated via MCMC. **Fig.7**(a & b) show that the mean of soil friction angle and soil unit weight, estimated using the MCMC method, do not experience significant fluctuations due to the relatively strong prior. Conversely, estimations of φ and γ derived from the GA displaced fluctuation across the 99% CI search space. This occurrence was attributed to the optimization problem being underdetermined, thereby there were multitude of θ combinations that could achieve an optimal solution. In estimating N', the MCMC method showed that the mean of N' gradually increased from a prior of 3.5 to approximately 8 as the drive progressed, while the GA approach tended to underpredict N' value compared to the MCMC approach. This divergence between the two prediction sets became evident after a jacked distance of 40 m, as illustrated in **Fig.7**(c). This was because the limitation of the GA search space boundaries, defined by the 99% CI in this work. The GA lacks the capability to escape the predetermined search space for each parameter, given that the new chromosomes, created for either the initial population or mutation processes, were drawn from the original search space. However, opting to heuristically modify the boundaries of θ to include all feasible values led to significant fluctuations and ultimately, unrealistic parameter estimations. In contrast, a key advantage of the MCMC approach is its ability to ignore the prior distribution of θ when strong evidence is present.

GA and MCMC predictions of $tan\delta$ were in very good agreement and indicated that the prior distribution for $tan\delta$ provides a significant overprediction reflecting the effectiveness of modern pipe-jacking lubrication systems. However, during the initial stage (unlubricated) of the drive, both the GA and MCMC predictions significant increases in $tan\delta$. This phenomenon can be attributed to the combination of the presence of heavier elements, such as the TBM, power pack, and cans, which tend to drag along the bottom of the overcut, and the still relatively small tunnel length. As the drive length increases, coupled with lubricant injection and accumulation in the overcut annulus, these effects get less important and the high levels of $tan\delta_v$ and $tan\delta_h$ decrease significantly [**Fig.7**(d & e)]. Furthermore, the acquisition of additional data led to a decrease in the 90% CIs of both $tan\delta_v$ and $tan\delta_h$. Importantly, the 90% CI of $tan\delta_h$ is markedly larger than that of $tan\delta_v$. This occurs because lateral friction contributes much smaller to the overall friction compared to the vertical friction acting box crown. As such, deviations within a certain range of the $tan\delta_h$ mean are unlikely to significantly affect the prediction of the total jacking force.

Beyond a jacked length of 40 m, the $tan\delta_v$ predicted by the GA was marginally higher than that predicted by the MCMC approach. This difference was because GA was shown to underpredict N' relative to the MCMC approach in **Fig.7**(c) due to the boundary on the search space. Moreover, this observation also indicates that $tan\delta_v$ contributes more significantly to the total jacking force than $tan\delta_h$. The underprediction of N' leads to a divergence in the predictions of $tan\delta_v$, but not $tan\delta_h$, between the GA and the MCMC approaches.



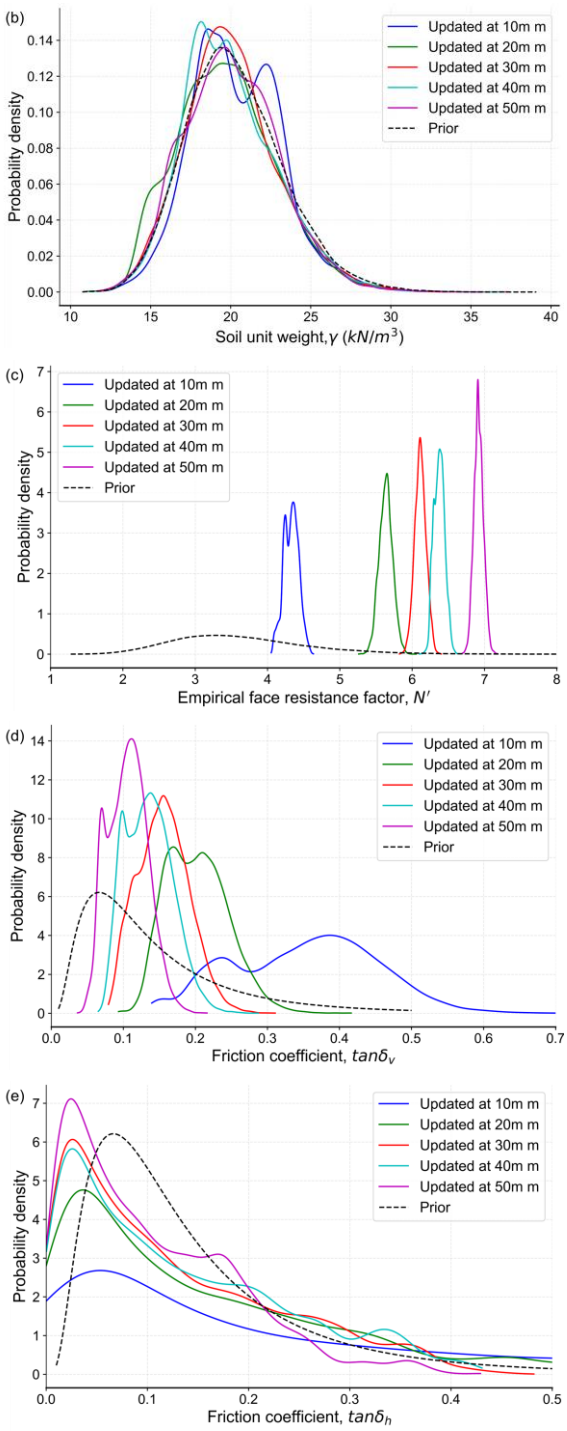
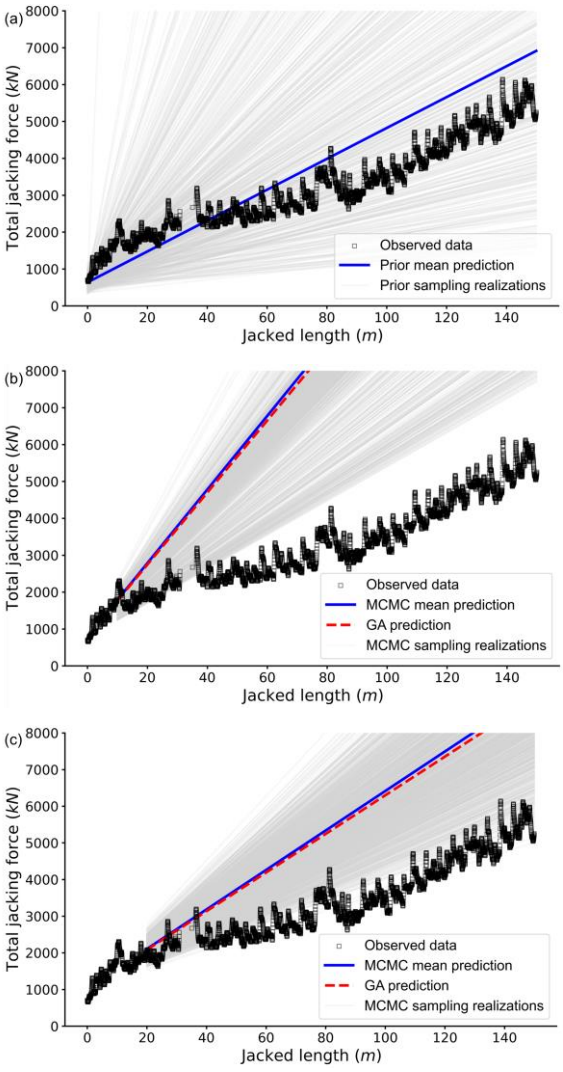


Fig.8 Outcomes of the Bayesian updating process - visualizing the updated uncertain parameter distributions: (a) friction angle φ; (b) soil unit weight γ; (c) empirical face resistance factor N'; (d & e) interface frictional coefficients $tan\delta_v$ & $tan\delta_h$.

The shapes of the updated (posterior) probability distributions of θ are compared with the prior probability distribution in **Fig.8** for selected points during the drive. For N' and $tan\delta_h$, the updating process (using larger and larger data sets) resulted in a stronger posterior in which the distributions appeared to

achieve convergence. **Fig.8**(d & e) highlights the importance of the lognormal distribution to prevent unrealistic negative realizations of $tan\delta$. Given the implementation of robust priors in updating the parameters of φ and γ, the resulting posterior distribution exhibits minimal deviation from its prior distribution. Conversely, the posterior distribution of $tan\delta_v$ and N' undergoes substantial alterations during the updating process, with its standard deviation markedly decreasing as more data was obtained. Reductions of the standard deviation were due to strong consistency of the jacking forces during the initial (unlubricated) stages of the drive (**Fig.8**). Despite a reduction in the posterior mean of $tan\delta_h$, its substantial standard deviation indicates a considerable degree of uncertainty. This implies that $tan\delta_h$ may be less influenced by the dataset and plays a lesser role in predicting jacking force.



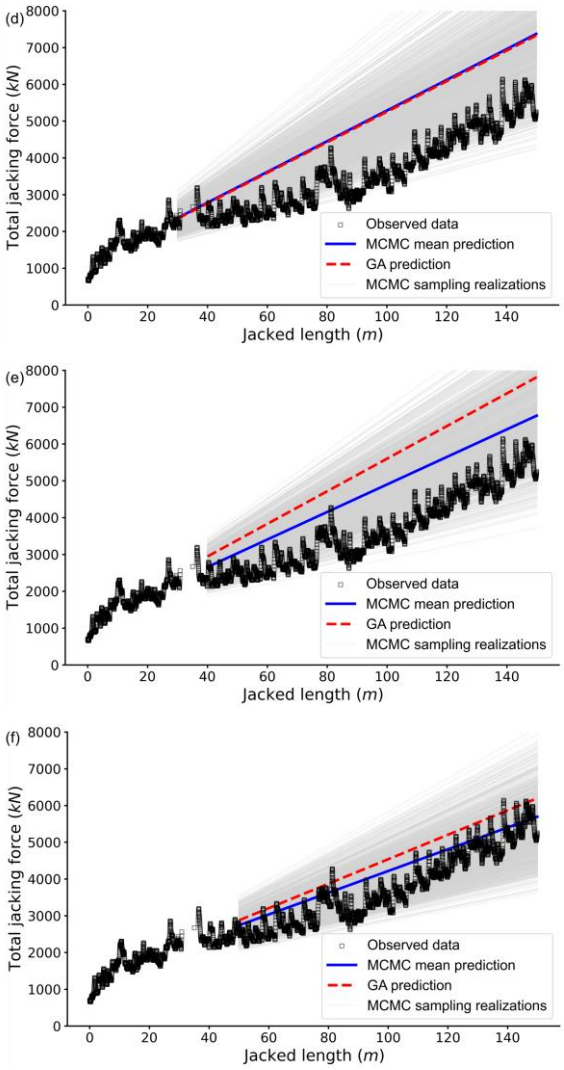


Fig.9 Model predictions of total jacking forces at distances: (a) 0m prior, (b-f) 10m-50m posterior, with 1000 draws from θ's posterior distribution

The estimated total jacking forces for the remainder of the drive are illustrated in **Fig.9**, utilizing parameters back-calculated through both MCMC and GA. For MCMC predictions, besides the mean prediction, 1000 realizations of the jacking force model were generated from sampling values of the posterior distribution of θ, which are depicted as light grey lines. The prior mean prediction derived from the prior distribution of θ substantially overpredicted the observed jacking force, with the scatter of the predictions was also rather significant as shown in **Fig.9**(a). After obtaining the initial 10m of jacking force data, a significant reduction in the scatter of the MCMC sampling predictions was observed. However, the predicted jacking force remained considerably overpredicted. This is because the initial jacking force measurements were taken

from a drive section that lacked sufficient lubrication, leading to an overproduction of the forces required for the subsequent lubricated sections. In this scenario, there is a strong alignment between the predictions from GA optimization and the posterior mean predictions derived from MCMC [**Fig.9**(b)]. With the progression of box-jacking, more data were obtained, improving the alignment between the MCMC posterior predictions and the monitoring data, as depicted in **Fig.9**(c-f). The expansion of datasets enabled the MCMC sampling predictions to accurately include all the peaks in jacking force arising from stoppages [**Fig.9**(f)]. After a jacked length exceeded 40m, there was a notable difference between the predictions determined using parameters optimized by GA and those updated using MCMC due to the GA approach being restricted by its original search space. Notably, despite this limitation, GA still demonstrates certain predictive capabilities, affirming the validity of the prior hypothesis used in this study. For long-drive box-jacking, skin friction is often the main contribution to total jacking force. Thus, even though GA was shown to underestimate N', the reasonable setting of the prior N' ensures that the GA's estimation of the jacking force was only about 7-13% higher than that of the MCMC.

6. CONCLUDING REMARKS

This research introduces a Bayesian updating approach for predicting jacking forces in box-jacking tunnel excavation, illustrated through a 150-meter drive in sandy conditions in Kanagawa, Japan. Performance was evaluated against a genetic algorithm optimization method, highlighting the overprediction issue when using model input parameters based on prior best hypotheses. The Bayesian approach, utilizing Markov Chain Monte Carlo sampling, proved more effective than traditional genetic algorithms, with prediction accuracy improving as more data was acquired from the drive. By using various jacking force model realizations derived from posterior distribution sampled values, all peak forces from stoppages were accurately included. The final updated and optimized parameters can potentially serve as prior estimates for similar future drives.

ACKNOWLEDGMENTS:

The authors would like to extend their sincere appreciation to the China Scholarship Council (CSC) for the generous financial support that facilitated this research. Our heartfelt thanks also go to Alpha Civil Engineering Co., Ltd. in Fukuoka, Japan, for not only supplying the essential machinery details but also significantly enriching the TBM data analysis and performance with their insightful observations and constructive dialogues.

REFERENCES

1) Chen, X., Ma, B., Najafi, M., and Zhang, P. : Long Rectangular Box Jacking Project: A Case Study, Underground Space, Vol. 6, No. 2, pp. 101-125, 2021.
2) Kaushal, V., Najafi, M., and Serajiantehrani, R. : Environmental Impacts of Conventional Open-Cut Pipeline Installation and Trenchless Technology Methods: State-of-the-Art Review, J. Pipeline Syst. Eng. Pract., Vol. 11, No. 2, 2020.
3) Shou, K. J., Jiang, J. M. : A Study of Jacking Force for a Curved Pipejacking, J. Rock Mech. Geotech. Eng., Vol. 2, No. 4, pp. 298-304, 2010.
4) Ma, P., Shimada, H., Sasaoka, T., Hamanaka, A., Moses, D. N., Dintwe, T. K. M., Matsumoto, F., Ma, B., and Huang, S. : A New Method for Predicting the Friction Resistance in Rectangular Pipe-Jacking, Tunnelling Undergr. Space Technol., Vol. 123, 104338, 2022.
5) Wen, K., Shimada, H., Zeng, W., Sasaoka, T., and Qian, D. : Frictional Analysis of Pipe-Slurry-Soil Interaction and Jacking Force Prediction of Rectangular Pipe Jacking, Eur. J. Environ. Civ. Eng., Vol. 24, No. 6, pp. 814-832, 2020.
6) PJA (Pipe Jacking Association) : Guide to Best Practice for the Installation of Pipe Jacks and Microtunnels, PJA, London, 1995.
7) ASCE : Standard Practice for Direct Design of Precast Concrete Pipe for Jacking in Trenchless Construction, ASCE, Reston, VA, 2001.
8) FSTT (French Society for Trenchless Technology) : Microtunneling and Horizontal Drilling: Recommendations, FSTT, Tangier, Morocco, 2006.
9) JMTA (Japan Micro Tunneling Association) : Pipe-Jacking Application, JMTA, Tokyo, 2000.
10) Peck, R. B. : Advantages and Limitations of the Observational Method in Applied Soil Mechanics, Géotechnique, Vol. 19, No. 2, pp. 171-187, 1969.
11) Spross, J., and Johansson, F. : When is the Observational Method in Geotechnical Engineering Favourable?, Struct. Saf., Vol. 66, pp. 17-26, 2017.
12) Shou, K., Yen, J., and Liu, M. : On the Frictional Property of Lubricants and its Impact on Jacking Force and Soil–Pipe Interaction of Pipe-Jacking, Tunnelling Undergr. Space

Technol., Vol. 25, No. 4, pp. 469-477, 2010.

13) Yu, B., Shimada, H., Sasaoka, T., Hamanaka, A., Matsumoto, F., and Morita, T. : A Jacking Force Study Based on Interpretation of Box Jacking Records: A Case Study of Curved Rectangular Box Jacking in Soft Soil in Saitama, Japan, Tunnelling Undergr. Space Technol., Vol. 139, 105228, 2023.

14) O'Dwyer, K. G., McCabe, B. A., and Sheil, B. B. : Interpretation of Pipe-Jacking and Lubrication Records for Drives in Silty Soil, Underground Space, Vol. 5, No. 3, pp. 199-209, 2020.

15) Zhang, Y., Feng, X., Zhou, H., Zhang, P., Ma, B., Tan, L., and Wang, J. : Pressure Characteristics of Rectangular Box Jacking Considering Box-Soil-Lubricant Interaction, Tunnelling Undergr. Space Technol., Vol. 126, 104569, 2022.

16) Zhang, J., Zhang, L. M., and Tang, W. H. : Bayesian Framework for Characterizing Geotechnical Model Uncertainty, J. Geotech. Geoenviron. Eng., Vol. 135, No. 7, 2009.

17) Kelly, R., and Huang, J. : Bayesian Updating for One-Dimensional Consolidation Measurements, Can. Geotech. J., Vol. 52, No. 9, 2015.

18) Zheng, D., Huang, J., Li, D-Q., Kelly, R., and Sloan, S. W. : Embankment Prediction Using Testing Data and Monitored Behaviour: A Bayesian Updating Approach, Comput. Geotech., Vol. 93, pp. 150-162, 2018.

19) Li, J., Hu, P., Uzielli, M., and Cassidy, M. J. : Bayesian Prediction of Peak Resistance of a Spudcan Penetrating Sand-Over-Clay, Géotechnique, Vol. 68, No. 10, pp. 905-917, 2018.

20) Jaynes, E. T., and Bretthorst, G. L. : Probability Theory: The Logic of Science, Cambridge University Press, Cambridge, UK, 2003.

21) Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. : Bayesian Data Analysis, CRC Press, London, 2013.

22) Qi, X-H., and Zhou, W-H. : An Efficient Probabilistic Back-Analysis Method for Braced Excavations Using Wall Deflection Data at Multiple Points, Comput. Geotech., Vol. 85, pp. 186-198, 2017.

23) Betancourt, M. : A Conceptual Introduction to Hamiltonian Monte Carlo, arXiv:1701.02434 [stat.ME], 2017.Sheil, B. B. : Hybrid Framework for Forecasting Circular Excavation Collapse: Combining Physics-Based and Data-Driven Modeling, J. Geotech. Geoenviron. Eng., Vol. 147, No. 12, 2021.

24) Sheil, B. B. : Hybrid Framework for Forecasting Circular Excavation Collapse: Combining Physics-Based and Data-Driven Modeling, J. Geotech. Geoenviron. Eng., Vol. 147, No. 12, 2021.

25) Hoffman, M. D., and Gelman, A. : The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, arXiv:1111.4246 [stat.CO], 2011.

26) Abril-Pla, O., Andreani, V., Carroll, C., Dong, L., Fonnesbeck, C. J., Kochurov, M., Kumar, R., Lao, J., Luhmann, C. C., Martin, O. A., Osthege, M., Vieira, R., Wiecki, T., and Zinkov, R. : PyMC: A Modern, and Comprehensive Probabilistic Programming Framework in Python, PeerJ Comput. Sci., Vol. 9, e1516, https://doi.org/10.7717/peerj-cs.1516, 2023.

27) Juang, C. H., Wang, L. : Reliability-based robust geotechnical design of spread foundations using multi-objective genetic algorithm, Comput. Geotech., Vol. 48, pp. 96-106, 2013.

28) Holland, J. H. : Genetic Algorithms, Sci. Am., Vol. 267, No. 1, pp. 66-72, 1992.

29) Goldberg, D. E. : Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley Longman Publishing, Boston, 1989.

30) Staheli, K. : Jacking force prediction: An interface friction approach based on pipe surface roughness, Georgia Institute of Technology, Atlanta, 2006.

31) Namli, M., Guler, E. : Effect of bentonite slurry pressure on interface friction of pipe jacking, J. Pipeline Syst. Eng. Pract., Vol. 8, No. 2, 04016016, 2017.

32) Harr, M. E. : Reliability-based design in civil engineering, North Carolina State Univ., Raleigh, NC, 1984.

33) Kulhawy, F. H. : On the evaluation of static soil properties, ASCE Geotech., Vol. 31, pp. 95-115, 1992.