Data-Driven Uncertainty Reduction in Geotechnical Engineering: Optimal Preloading of a Road Embankment

Dafydd Cotoarbă
PhD Student, TUM Georg Nemetschek Institute Artifical Intelligence for the Built World, Technische Universität München, Germany

Elizabeth Bismut
PhD, Engineering Risk Analysis Group, Technische Universität München, Germany

Johan Spross
PhD, Division of Soil and Rock Mechanics, KTH Royal Institute of Technology, Stockholm, Sweden

Daniel Straub
Professor, Engineering Risk Analysis Group, Technische Universität München, Germany

ABSTRACT: Decision making in geotechnical engineering is characterized by considerable uncertainty. To find optimal solutions in such an environment, an iterative decision-making process, which includes new information as it becomes available, is required. In this contribution we extend the risk-based framework of Bismut et al. (2023) for optimal planning of a geotechnical construction to include predictions, which are based on a linear regression analysis of monitoring data. This approach is illustrated for the design of a surcharge on an embankment in clayey soil, as the optimal preloading sequence is searched. We demonstrate that by increasing the amount of information considered more cost efficient strategies can be identified and outcome uncertainties can be reduced.

1. Introduction

Predicting the consolidation behavior of soils represents a major challenge to geotechnical engineers, as it is subject to significant uncertainty. Reasons are the lack of knowledge of prevailing soil conditions, limitations in understanding the behaviour of soil over time and predicting ground-structure interactions. Considering the design of geotechnical structures, it is critical that such uncertainties are understood and considered. In practice, the observational method offers a pragmatic approach to dealing with uncertainties in sequential manner (Spross and Larsson, 2021). In some areas of engineering, e.g., optimal inspection and maintenance planning, sequential decision problems under uncertainty are approached by formal decision analysis (e.g., Straub and Faber, 2005; Memarzadeh et al., 2014; Wang et al., 2022). An example for the design of a geotechnical structure under uncertainty is the design of preloading with surcharge and prefabricated vertical drains (PVD) for an embankment on soft soil (see Figure 1). The prediction of the long-term settlement of preloaded soils is challenging. Geotechnical investigations are required to increase the knowledge of the prevailing soil conditions. If epistemic knowledge uncertainty is neglected in the design phase, excessive residual settlements might occur after it has been taken into service. In a risk-based approach to this problem, engineers try to find an optimal preloading strategy, which balances the risk of settlements with the expected construction costs.

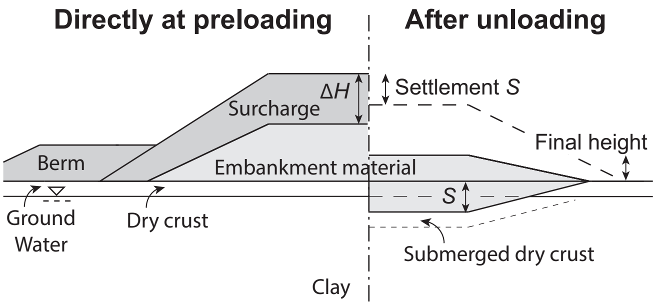

Figure 1: Preloading of an embankment with a surcharge of total height ΔH to accelerate consolidation. Prefabricated vertical drains are omitted for clarity (Bismut et al., 2023, CC-BY-4.0).

To address this challenge, Bismut et al. (2023) proposed a risk-based decision-theoretic framework to optimal sequential planning in geotechnical engineering. It allows to compare different strategies, by quantifying the consequences of actions using a cost model and combining them with a probabilistic geotechnical model. To optimise the sequence of decisions a heuristic approach is used (Bismut and Straub, 2021). This methodology was applied to the embankment preloading problem, which can be modeled using the following decision sequence: 1) initial surcharge height and 2) later additions to the surcharge depending on the observed settlements. The optimisation should ensure that the desired settlement is reached at minimal expected costs. It results in a preloading strategy that defines the initial surcharge height and a possible increase in the surcharge height at a later time dependent on a measurement outcome. The approach uses the probabilistic geotechnical model developed by Spross and Larsson (2021), which describes the long-term primary compression settlement of an embankment with initial pre-loading and PVDs.

While in Bismut et al. (2023) only a single measurement is considered, in practice, settlement measurements can be available at regular time intervals. We refer to this case as continual monitoring.

In this work, the methodology is extended to include predictions, which are based on continual monitoring. We investigate the effect of this modified framework for the same embankment preloading problem that was previously studied in Bismut et al. (2023). The new approach is then compared to the old framework.

2. Elements of the decision analysis

The underlying basic theory for the decision analysis framework follows (Raiffa and Schlaifer, 1961). It formalises decision problems under uncertainty with varying information, which can be used to find the optimal preloading strategy. The optimisation of the surcharge height strategies is framed as a sequential decision problem influenced by measurements of the embankment settlement during the preloading phase.

A decision analysis under uncertainty is based on a probabilistic model of the system, a model of the decision alternatives and a utility or cost function. In the following each model is introduced.

2.1. Probabilistic Model

A complete probabilistic model should not only model the system behaviour, but also account for the effects of actions affecting the system and reflect uncertainty in information collection, through a likelihood function (Bismut and Straub, 2022).

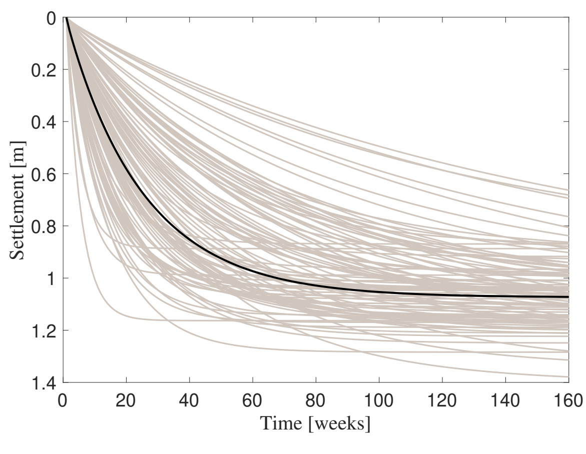

Figure 2: Sample soil settlement trajectories for an initial surcharge $h_0 = 0$[m] (no additional surcharge). One such trajectory is highlighted in black.

2.1.1. Geotechnical Model

In the investigated engineering problem, we use the soil consolidation model defined by Spross and Larsson (2021). This geotechnical model describes for an embankment with an added surcharge on top of prefabricated vertical drains (PVD), 1) how the primary compression settlement develops with time, and 2) the effect of the preloading on the overconsolidation ratio (OCR). Sample trajectories of the settlement obtained using the model are depicted in Figure 2.

The evolution of the consolidation depends on parameters for the soil properties and PVD design. In this soil consolidation model the soil parameters are modelled as random variables with an associated probability distribution, which is obtained by means of constant-rate-of-strain (CRS) oedometer tests or, in cases of lack of data, defined based on engineering judgement. The PVD model is based on the work of Hansbo (1979). The detailed description of the model is provided in Spross and Larsson (2021).

2.1.2. *Settlement and OCR requirements*

Spross and Larsson (2021) also define performance criteria, based on which preloading decisions can be evaluated. To ensure sufficient consolidation such that the embankment can be taken into service, settlement and OCR targets are defined.

The settlement target s_{target} is defined such that the long term settlement S_∞ caused by the load of the completed embankment $\Delta\sigma_{emb}$ does only exceed a residual (post-construction) primary consolidation settlement with an acceptable probability p_{FT}. In the numerical investigations, p_{FT} is set to 5% to reflect a serviceability limit state. Thereby, s_{target} to be reached within the available preloading time, can be obtained as the quantile value corresponding to p_{FT} of samples of the $S_\infty(\Delta\sigma_{emb})$. If s_{target} is not reached for a preloading strategy, penalties due to projects delays or structure damages occur.

The overconsolidation ratio quantifies the effect of secondary consolidation in soils. To avoid any negative impact from this type of consolidation, the OCR_{target} should be reached in the middle of the soft soil stratum at the moment of unloading the surcharge. In this paper, we define $OCR_{target} = 1.10$ in line with the general technical requirements and guidance for geotechnical works issued by the Swedish Transport Administration (2013a,b).

2.1.3. *Observations*

The probabilistic consolidation model described above can be used to obtain information on the state of the system in form of a measurement M_{t_i} of the settlement S_{t_i} at time t_i. The value of M_{t_i} includes an additive measurement error ε:

$$M_{t_i} = S_{t_i} + \varepsilon \tag{1}$$

2.2. *Cost function*

To represent the preferences of the decision maker, the outcome of a decision is evaluated in terms of utility (Fishburn, 1970). In this work, we express utility in terms of costs associated with the actions and the systems performance. The optimal decision can be identified as the one which minimizes the total expected cost.

For the considered engineering problem, we have identified three cost components, the sum of which results in the total cost C_{tot}

$$C_{tot} = \sum_i C_{sur,i} + C_{delay} + C_{OCR}. \tag{2}$$

The first cost component $C_{sur,i}$ quantifies the cost of adding a surcharge of the height ΔH_i (see Figure 1). The term accounts for all the costs surrounding the construction of the surcharge and includes material costs, mobilization costs, material availability at the time of the decision, and the need for berms for slope stability.

The second cost component C_{delay} is used to express the penalty for cases when the settlement target s_{target} is not reached within a required time period. As it is crucial that s_{target} is reached before the embankment can be taken into service, the penalty is proportional to the additional time required for reaching it.

To quantify the consequences of reaching insufficient overconsolidation at time of unloading, we introduce the third cost component C_{OCR}. By not reaching OCR_{target} the probability for residual secondary consolidation settlement (creep) during the lifetime of the embankment is high. This damage is not critical for structural safety, but negatively impacts the serviceability of the superstructure.

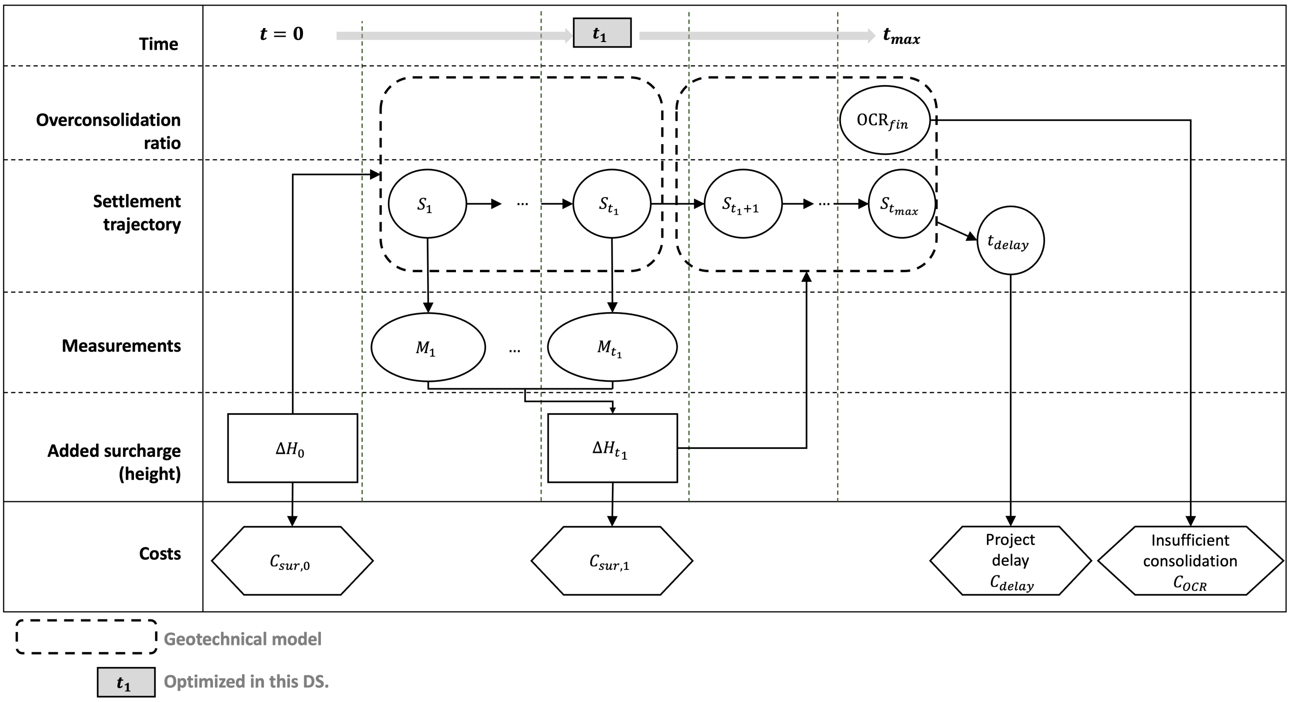

Figure 3: Influence diagram for the DS presented in Section 2.3. We note that interactions between different decisions on surcharge height are represented in a simplified manner.

2.3. Decision setting and decision alternatives

Decision settings (DS) for finding the optimal surcharge strategy can be defined. For this, operational constrains and the complexity of the decision sequence have to be considered.

In this contribution, a decision setting is analysed in which the initial surcharge of the height ΔH_0 is fixed in the beginning. Afterwards, there is one opportunity at time t_1 to increase the surcharge height to ΔH_1. The adjustment is based on a series of measurements taken up to the point t_1. In the numerical example, the time between measurements is set to one week. The influence diagram of the DS can be seen in Figure 3. This DS will be compared to the decision settings from our previous work, where only a single measurement was considered.

2.4. Optimal decision

The goal of the optimisation is to identify the decision sequence with the lowest expected cost. This is a non-trivial task, because of the uncertain nature of the soil which results in an uncertain total cost.

For this particular sequential decision problem, the optimal action can be identified after observing the evolution of the system state over time. A function is required, which extracts information from the observations and transforms them into (optimal) actions. This function, which is a mapping from observations to actions, is called decision rule or policy. Finding optimal policies is computationally expensive and exact solutions become intractable with increasing number of decision steps (Papadimitriou and Tsitsiklis, 1987).

One approach to solving general sequential decision problem, is by defining preloading strategies $\mathcal{S}$, which describe a sequence of decisions for the proposed problem. A strategy contains a set of rules for any time step, which specify by how much the surcharge height should be increased, based on information extracted from observations. For our specific decision setting, it prescribes the initial surcharge height at time $t = 0$ and the time t_1 at which a decision on the adjustment of the surcharge height should be made. It also includes rules based on measurements that specify if the surcharge height should be increased and by what amount.

From Equation (2) one obtains the expected total cost associated to a preloading strategy $\mathbf{E}\left[C_{tot}(\mathcal{S})\right]$

as

$$\mathbf{E}[C_{tot}(\mathcal{S})] = \mathbf{E}[C_{sur}(\mathcal{S})] + \mathbf{E}[C_{delay}(\mathcal{S})] + \mathbf{E}[C_{OCR}(\mathcal{S})] \quad (3)$$

where for the preloading strategy $\mathcal{S}$: $\mathbf{E}[C_{sur}(\mathcal{S})]$ is the total expected surcharge cost, $\mathbf{E}[C_{delay}(\mathcal{S})]$ expected delay cost and $\mathbf{E}[C_{OCR}(\mathcal{S})]$ expected cost of creep.

The optimal preloading strategy $\mathcal{S}^*$ is the preloading strategy which minimizes the expected total cost

$$\mathcal{S}^* = \arg\min_{\mathcal{S}} \mathbf{E}[C_{tot}(\mathcal{S})]. \quad (4)$$

The expected total cost of a strategy can usually not be evaluated analytically. Instead we utilize a Monte Carlo (MC) approximation. Using the geotechnical model of (Spross and Larsson, 2021), we generate n_{MC} random settlement trajectories $S_t{}^{(k)}$, and the overconsolidation ratio at unloading, $OCR_{fin}^{(k)}$, in function of the surcharge heights. Using the cost model defined in Equation (2), the cost of each trajectory can be computed. Therefore, the MC approximation of the expected total cost is

$$\mathbf{E}[C_{tot}(\mathcal{S})] \simeq \frac{1}{n_{MC}} \sum_{k=1}^{n_{MC}} C_{tot}^{(k)}(\mathcal{S}). \quad (5)$$

2.5. *Heuristic Approach*

As the solution space of Equation (4) is infinite-dimensional, an exact solution is generally intractable. Therefore, to find the optimal sequence of decision, we employ a heuristic approach following Bismut and Straub (2022). In the heuristic approach, the problem complexity is reduced by restricting the number of considered strategies. By means of predefined parameters $\mathbf{w} = [w_1; w_2; \ldots; w_n]$ a set of rules can be defined, to which a strategy must adhere. The heuristic parametrisation of a problem enables the encoding of engineering knowledge and operational constraints. The simplicity of the heuristic rules has the added benefit that the resulting strategies can be interpreted from a geotechnical point of view. However, the heuristics should be defined carefully to ensure that the identified solutions are near-optimal strategies.

The optimal heuristic strategy $\mathbf{w}^*$ is given as the set of parameter values which minimise the expected cost

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \mathbf{E}[C_{tot}(\mathcal{S}(\mathbf{w}))]. \quad (6)$$

To solve this problem, we employ an efficient cross-entropy (CE) based optimisation (Rubinstein and Kroese, 2004; Bismut et al., 2022). It enables to find solutions for noisy optimisation problems. Here, the objective function evaluations are subject to noise because of the MC approximation error.

2.6. *Linear Regression*

To extract information from the observation data, linear regression analysis is performed. A multiple linear regression model is defined as follows

$$y = \beta_0 + \sum_i \beta_i x_i + \varepsilon \quad (7)$$

with β_i as the regression coefficients and an error term ε.

For the embankment problem, the regression analysis is used to predict the settlement at time of unloading t_{max} (the response variable y) based on settlement measurements (the predictors x_i). An example is illustrated in Figure 4, where a regression model is used to predict $S_{t_{max}}$ based on 5 measurement points. In this specific case, there is a prediction error of $0.1\,m$. One common way to increase the accuracy of the prediction is to increase the number of used predictors x_i.

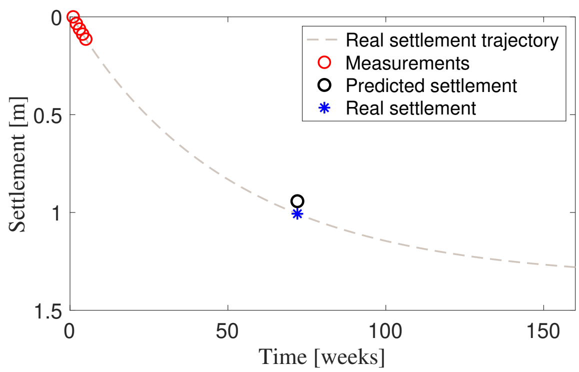

Figure 4: Prediction of the settlement at $t = 72[week]$ based on 5 measurement points using a regression model.

3. Numerical Investigation

3.1. Problem setting

The numerical investigation in this work is performed on the example presented in Spross and Larsson (2021). Using $p_{FT} = 5\%$ we obtain a settlement target $s_{target} = 1.27[m]$.

Table 1: Parameters of the cost model

Cost factor	Value
c_{sur}	$3.45 \cdot 10^6 [SEK/m]$
c_{delay}	$3 \cdot 10^5 [SEK/week]$
c_{OCR}	$2 \cdot 10^7 [SEK]$
$f_{add,0}$	1
$f_{add,1}$	1
t_{max}	72[weeks]

3.2. Cost model

The cost components of the cost function described in Section 2.2 for this specific problem, defined in Bismut et al. (2023), are presented here.

The cost of adding surcharge is quantified in the component $C_{sur,i}$. It includes the construction costs, as well as the cost of ensuring slope stability with berms. If one considers the total surcharge height H_{tot} the cost is defined as

$$C_{sur}(H_{tot}) = \begin{cases} H_{tot} \cdot c_{sur} & \text{if } H_{tot} \leq 1m \\ 1.25 \cdot H_{tot} \cdot c_{sur} & \text{otherwise.} \end{cases} \quad (8)$$

To account for potential added costs in the case of adding surcharge ΔH_i at a later time $t > 0$, we introduce the factor $f_{add,i} \geq 1$ such that

$$C_{sur,i}(\Delta H_i) = (C_{sur}(H_{tot} + \Delta H_i) - C_{sur}(H_{tot})) \cdot f_{add,i}. \quad (9)$$

The second cost component, C_{delay} quantifies the cost of project delay, which occurs when settlement does not reach s_{target} within the project specific preloading time, t_{max}, ($t_{target} > t_{max}$) or is unable to meet s_{target} at all in reasonable time ($t_{target} > t_{lim}$). Thus, the cost is defined as

$$C_{delay}(t_{target}) = \begin{cases} 0 & \text{if } t_{target} \leq t_{max} \\ c_{delay} \cdot (\min(t_{lim}, t_{target}) - t_{max}) & \text{otherwise,} \end{cases} \quad (10)$$

where c_{delay} represents the penalty per week of delay.

The third cost component C_{OCR} quantifies the cost associated to residual creep settlement due to insufficient OCR after the embankment has been taken into service. We use a logistic function to encode the correlation between overconsolidation ratio at time of unloading OCR_{fin} and the amount of residual settlement to occur. Thus, we have

$$C_{OCR}(OCR_{fin}) = \frac{c_{OCR}}{1 + \exp\left(-\frac{1.075 - OCR_{fin}}{4.5 \cdot 10^{-3}}\right)}. \quad (11)$$

The cost factors c_{sur}, c_{delay} and c_{OCR} and the available preloading time t_{max} for the numerical investigation are given in Table 1.

3.3. Heuristic parametrisation

The novelty of this work is in investigating the heuristic for the decision setting where continual monitoring of the settlement in considered. We refer to this heuristic as Heuristic 4 and compare it to the heuristics 1 and 3 defined in (Bismut et al., 2023).

Heuristic 4

Optimised parameters: $h_0 \geq 0$, $s_{th} \geq 0$, $a \geq 0$, $b \geq 0$, $t_1 \in \{1, 2, 3, ..., t_{max}\}$

1. At time $t = 0$, add surcharge of height $\Delta H_0 = h_0$
2. Obtain weekly measurements $m_i + \varepsilon$ up to time t_1.
3. Predict settlement $s^{pred}_{t_{max}}$ at time t_{max} using measurements m_i as an input to the linear regression model
4. Compute $d = s_{target} - s^{pred}_{t_{max}}$
5. Add surcharge
$$\Delta H_1 = \begin{cases} 0, & d \leq 0 \\ \ln(d * a + 1) * b, & d > 0. \end{cases}$$

In this case the preloading strategy starts with specifying an initial surcharge height h_0. Measurements m_i are taken in weekly time steps up until the decision time t_1. The optimisation searches a decision time which balances the collection of information with performing the mitigation action early enough. For example, the later the decision time, the more information is collected. However, the remaining time until unloading at t_{max} is shorter. Based on these measurements, a predictor function, obtained using regression analysis, predicts the settlement at t_{max}. The prediction is compared to s_{target}. Depending on the difference d between the two, the strategy adjusts the height of the added surcharge using a exponential function, characterized by the parameters a and b.

3.4. Linear regression

Regression analysis requires a data-set of response variables and associated predictors to be trained. To create this a data-set we use the probabilistic model presented in Section 2.1.1. To learn the behaviour of the soil we simulate 10 000 sample settlement trajectories for an embankment without preloading. In this specific case, the response variable y is the settlement at time t_{max} and the predictors x_i are settlement measurements. Thus, the linear regression model gives a prediction of the settlement $s^{pred}_{t_{max}}$ at time t_{max} based on measurements M_i up to the point t_1.

4. RESULTS

The optimal parameter values for the heuristic for the presented cost model are obtained with the CE method. In this work we analysed two variations: first, we consider observations without measurement errors (results in Table 2). This can be compared to our previous work, where the same approach was used. Secondly, we consider observations with measurement errors (see Table 3) and analyse their impact on the optimisation problem.

4.1. Comparison to previous work

The results of the optimisation for measurements without measurement errors are presented in Table 2. We consider two heuristics from our previous work. Heuristic 1 is the reference case from our previous work, which optimizes the initial surcharge height without allowing any surcharge adjustment at a later stage. Heuristic 4 is a simplified version of the new heuristic proposed in this study, which uses a single measurement at t_1 to decide whether ti adjust the surcharge height instead of considering all measurements up to that point. The coefficient of variation of the total cost for the optimal strategies varies between $80-90\%$, which results in a standard error of the MC estimated expected cost of $\leq 1\%$. This is a high enough accuracy, to evaluate the heuristics according to the estimated expected costs.

The optimal surcharge strategy for our new approach, Heuristic 4, shows a clear improvement to our previous best heuristic, Heuristic 3. By including more information in the decision process, a strategy with 2% less cost is obtained. However, more interesting is that the coefficient of variation is reduced by about 10%. This indicates that by increasing the amount of considered data in the decision making process, uncertainties of the final expect cost can be significantly reduced.

4.2. Measurement error

In the second analysis the influence of measurement errors is analysed. The results of this optimisation are presented in Table 3. To simulate imperfection during measurements, we define a Gaussian distribution $\varepsilon \sim \mathcal{N}(\mu_\varepsilon, \sigma_\varepsilon^2)$ with $\mu_\varepsilon = 0[cm]$ and a standard deviation $\sigma_\varepsilon \in [0.01, 0.02, 0.03]$. The predictor is trained using measurements which contain measurement errors. Table 3 shows that even by introducing a measurement error, preloading strategies that are at least 10% better than the reference case (no adjustment of the surcharge) are obtained. With a larger measurement error the expected cost of the identified optimal strategy also increases. The increasing noisiness of the data, increases the difficulty for the predictor to learn the underlying behaviour of soil. For best results with this method, adequate data quality is required.

5. CONCLUSION

In this paper, we extended the methodology based on heuristics for finding optimal preloading strategies for a geotechnical problem defined by Bismut et al. (2023), such that a continual data stream can be integrated in the decision making process. The analysis showed the potential of data-driven reduction of uncertainties in the decision making process for geotechnical problems.

Table 2: Optimal heuristic parameters and associated expected costs

Parameter	Unit	Heuristic 1	Heuristic 3	Heuristic 4
h_0	[m]	1.05	0.95	0.99
a	[m]	-	-0.28	2.95
b	[m]	-	-	1.98
t_1	[weeks]	-	20	13
Expected cost	$[10^6 SEK]$	8.11	6.06	5.94
Std. dev. cost	$[10^6 SEK]$	7.4	5.6	4.84

Table 3: Optimal heuristic parameters and associated expected costs for different measurement errors.

Parameter	Unit	Heuristic 1	Heuristic 4			
σ_ε	[cm]	-	0	0.01	0.02	0.03
h_0	[m]	1.05	0.99	0.99	1.08	1.00
a	[m]	-	2.95	1.57	3.22	1.34
b	[m]	-	1.98	3.1	1.80	1.17
t_1	[weeks]	-	13	14	18	13
Expected cost	$[10^6 SEK]$	8.11	5.94	7.01	7.11	7.33
Std. dev. cost	$[10^6 SEK]$	7.4	4.84	4.81	4.91	5.56

6. REFERENCES

Bismut, E., Cotoarbă, D., Spross, J., and Straub, D. (2023). "Optimal adaptive decision rules in geotechnical construction considering uncertainty [manuscript submitted to a scientific journal].

Bismut, E. and Straub, D. (2021). "Optimal adaptive inspection and maintenance planning for deteriorating structural systems." *Reliability Engineering & System Safety*, 215, 107891.

Bismut, E. and Straub, D. (2022). "A unifying review of NDE models towards optimal decision support." *Structural Safety*, 97, 102213.

Bismut, E., Straub, D., and Pandey, M. (2022). "Inspection and maintenance planning of a feeder piping system." *Reliability Engineering & System Safety*, 224, 108521.

Fishburn, P. C. (1970). "Utility theory for decision making." *Report no.*, Research analysis corp McLean VA.

Hansbo, S. (1979). "Consolidation of clay by band-shaped prefabricated drains." *Ground Engineering*, 12(5), 16–25.

Memarzadeh, M., Pozzi, M., and Kolter, J. Z. (2014). "Optimal planning and learning in uncertain environments for the management of wind farms." *Journal of Computing in Civil Engineering*, 29(5), 04014076.

Papadimitriou, C. H. and Tsitsiklis, J. N. (1987). "The complexity of Markov decision processes." *Mathematics of operations research*, 12(3), 441–450.

Raiffa, H. and Schlaifer, R. (1961). *Applied Statistical Decision Theory*. Harvard Business Review Press, Boston, MA.

Rubinstein, R. Y. and Kroese, D. P. (2004). *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning*. Springer Science & Business Media.

Spross, J. and Larsson, S. (2021). "Probabilistic observational method for design of surcharges on vertical drains." *Géotechnique*, 71(3), 226–238.

Straub, D. and Faber, M. H. (2005). "Risk based inspection planning for structural systems." *Structural Safety*, 27(4), 335–355.

Swedish Transport Administration (2013a). "TK Geo 13: Trafikverkets tekniska krav för geokonstruktioner." *Report No. TDOK 2013:0667*, Borlänge, Sweden: Trafikverket (Swedish Transport Administration).

Swedish Transport Administration (2013b). "TR Geo 13: Trafikverkets tekniska råd för geokonstruktioner." *Report No. TDOK 2013:0668*, Borlänge, Sweden: Trafikverket (Swedish Transport Administration).

Wang, Y., Zechner, M., Mern, J. M., Kochenderfer, M. J., and Caers, J. K. (2022). "A sequential decision-making framework with uncertainty quantification for groundwater management." *Advances in Water Resources*, 166, 104266.