# SOIL BEHAVIOUR TYPE CLASSIFICATION FOR TOTAL SOUNDINGS

**Sigurður Már Valsson[1], Samson A. Degago[2], Gudmund Reidar Eiksund[3]**

**KEYWORDS**

Soil investigations, SBT classification, Total sounding, Machine learning

**ABSTRACT**

The total sounding method is the most frequently used site investigation method in Norway; however, only a few classification models exist to interpret its results. This work proposes a new soil behavior type classification model for total soundings. A nearest neighbor classifier with the dynamic time warping distance is proposed, utilizing archived soundings and laboratory results from about 4500 boreholes across Norway. The presented model capability in classifying different soil types is shown to be promising.

## 1. INTRODUCTION

The Norwegian Total Sounding is a soil investigation method, developed in cooperation between the Research council of Norway, the Norwegian Public Roads Administration (NPRA), and the Norwegian Geotechnical Institute. It combines elements from the rotary pressure sounding and rock control drilling and is usually conducted prior to any advanced test to get an idea of the soil stratification and depth to bedrock. This method has been the topic of previous papers in this conference (e.g. [1], [2] and [3]). In Sweden, selected phases from the total sounding have been combined with elements from Jb-2 method to define the Jb-tot method (e.g. [1] and [4]).

The total sounding has 5 phases: a normal-, increased rotation-, flushing-, hammering- and rock drilling phase. The method definition given in [5] gives details on when each phase shall be applied, but the first two are identical to the rotary pressure sounding procedure. Interpreting soil stratification with total soundings is done qualitatively based on visual observations of sounding profiles and examples of interpretation procedures are given in the method

[1] Norwegian Public Roads Administration, Trondheim, Norway. sivals@vegvesen.no
[2] Norwegian Public Roads Administration, Trondheim, Norway
[3] Norwegian University of Science and Technology, Trondheim, Norway

definition. Despite its widespread use in Norway, only a handful of classification models have so far been proposed (e.g. [3] and [6]). The penetration resistance is measured at terrain level and includes both the tip-resistance as well as the accumulated rod friction from the surface to the tip. Resistance curves from all material types are found to span the entire presentation range for the push force 0-30kN, so defining models to identify materials based on resistance values alone is problematic. This paper defines a machine learning approach for Soil Behavior Type (SBT) classification for total soundings, using short intervals of the push force as a basis for the classification, combining the dynamic time warp distance [7] with a k-nearest neighbor classifier [8]. Field and laboratory results from over 4500 boreholes were collected to define the model, but this dataset was later reduced to representative soundings for each class to make the model implementation more practical.

## 2. DATASET

Total sounding registrations were linked by depth to grain size analysis (GSA) results in the same position from numerous NPRA projects across Norway. The depth interval around each soil sample was selected as 1m giving about 6500 short push force curve segments. Figures 1A and 1B show the locations of boreholes and how curves were extracted from depths around samples.

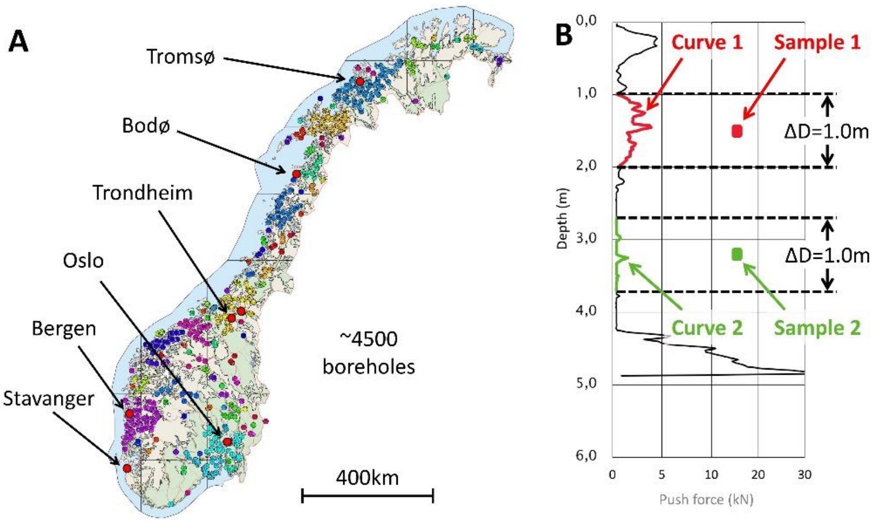


*Figure 1 A) Location of boreholes in the training set. B) Push force curve segments were collected in 1m intervals around sample depths.*

The classes are generated by interpreting GSA curves according to standard Norwegian practice [9]. The number of classes were reduced to a practical limit of 10 by only allowing main fractions (clay, silt, sand and gravel), possibly with one neighboring secondary fraction as an adjective (e.g. clayey silt, gravel, gravelly sand). Curves corresponding to other labels are disregarded (e.g. clayey sand, gravelly silt). About 4100 curves fit these criteria and are illustrated in Figure 2 colored by class.

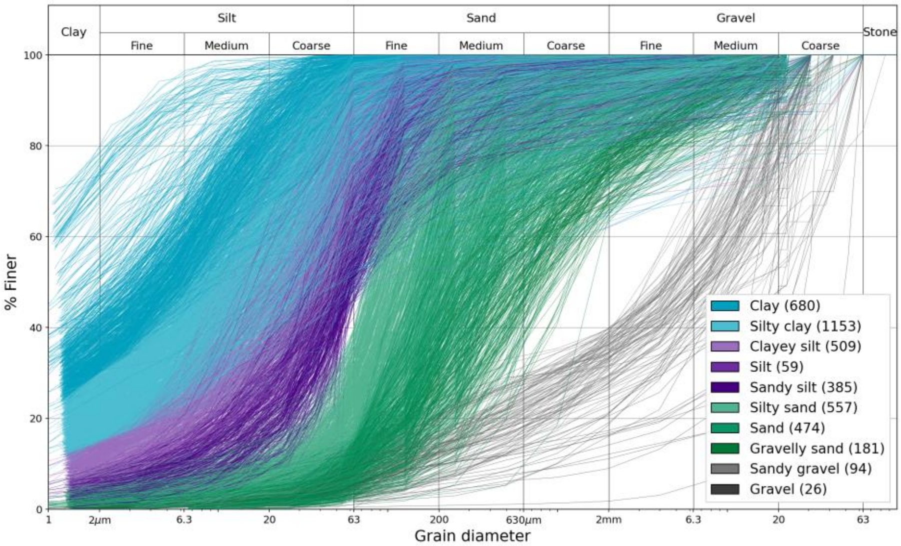


*Figure 2 GSA curves used to generate labels for the total sounding data. Each color is a model class and numbers in parenthesis show the number of samples for the class.*

## Dynamic time warping (DTW)

To compare curve segments from new tests to those in our dataset, the dynamic time warping (DTW) distance is used. This method was proposed by Vintsyuk [7] to identify similarities between audio signals for speech discrimination. A key ingredient of DTW is the alignment process, where optimal alignment of series is identified before their similarity is evaluated. In the original context, the alignment process was useful to account for different dialects or cadence in spoken language, but the algorithm has been found useful across a host of other domains, a few are listed in [10].

Let $x = (x_1, x_2, \cdots, x_m)$ and $y = (y_1, y_2, \cdots, y_n)$ be two sequences that are not in sync, (Figure 3B upper). The warp cost between any element $i$ and $j$ of $x$ and $y$ can be calculated as

$$cost(i,j) = \left(x_i - y_j\right)^2 \qquad (1)$$

for $i = 1, \dots, m$ and $j = 1, \dots, n$. A D-matrix is constructed as the accumulated cost between all elements in both series (Figure 3 A). The value of D for elements $x_i$ and $y_j$ is defined as

$$D(i,j) = min \begin{Bmatrix} D(i, j-1) \\ D(i-1, j-1) \\ D(i-1, j) \end{Bmatrix} + cost(i,j) \qquad (2)$$

Eq. (2) defines a dynamic programming scheme, where the matrix is constructed from the lowest index ($i = j = 1$) and expanded along either rows or columns towards the end. Backtracking indexes from the final step to the origin reveals the optimal alignment path (Figure 3B lower). The DTW distance between the two sequences is then calculated as

$$DTW(x, y) = \sqrt{D(m, n)} \tag{3}$$

Variants of DTW include the endpoint relaxation, where alignment can start at index $p$, in either sequence and can finish $q$ indexes from the end of either sequence with $p, q \in \{0, 1, \ldots, r\}$, (e.g. [7] and [10]). It has been suggested that better classification performance can be obtained by windowing the warping path, disallowing paths that stray far from the diagonal (e.g. [11]). This method also reduces the number of calculations needed, thus speeding up the classification. Figure 3 shows an example of a DTW D-matrix for two sequences and their resulting alignment. The DTW distance in this example is $DTW(x, y) = \sqrt{D(10, 9)} = \sqrt{171} \approx 13.08$.

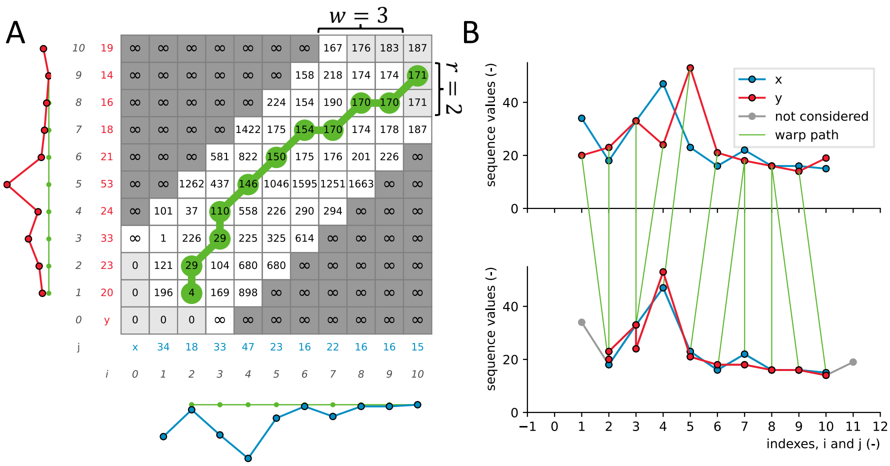


*Figure 3 shows both a A) DTW D-matrix using endpoint relaxation r=2 and window w=3 and the B) original (upper) and aligned (lower) sequences.*

The DTW algorithm is flexible and can produce degenerate alignments. It is therefore important to evaluate appropriate settings for the model constraints for each problem (e.g. [12]).

### k-Nearest neighbors

We have used a k-Nearest neighbor (k-NN) classifier to produce classifications for new soundings using the DTW distance described in the previous section. The k-NN method was first proposed by Fix and Hodges [8], for binary classification problems using the Euclidean distance as a metric. Figure

4A shows how the original definition can be used to decide the class of the blue cross, which is set to belong to the majority class for points within a given radius.

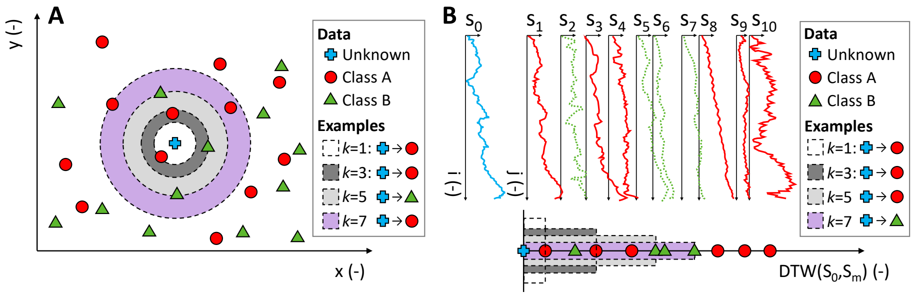


*Figure 4 k-NN classification using A) Euclidean distance in 2D and B) DTW distance in 1D.*

If the radius in Figure 4A is chosen so the circle contains 3 neighbors ($k = 3$), then 2 are red and 1 is green. The blue cross is then assigned the red class. Increasing the radius so that the circle contains 5 neighbors, then 3 are green and 2 are red. The blue cross is in this case assigned the green class.

The same logic can be applied using DTW, as illustrated in Figure 4B. First all curve segments in the training dataset are compared to the blue curve, ordered by ascending DTW distance. The decision threshold can be adjusted as before. If it is set to contain 3 neighbors ($k = 3$), then 2 are red and 1 is green. Here the blue curve gets assigned the red class. Rather than adjusting the value of the decision boundary (radius or distance in the above examples) in k-NN, the neighbors are simply sorted by the chosen measure and the majority class of the k-nearest neighbors is used as the answer.

## 3. DATASET REDUCTION

With the implementation of the DTW algorithm given earlier, the calculations are slow. For practical purposes it is desirable to save calculation time by reducing the dataset to representative segments for each class.

Our initial study agreed with the conclusion in [3], where normalizing the push force was found to give a better basis for classification. It is defined as

$$q_n = \frac{\mathrm{F}_{DT}}{A \cdot \sigma'_{v0}} \approx \frac{\mathrm{F}_{DT}}{A \cdot \gamma' \cdot z} \tag{4}$$

where $\mathrm{F}_{DT}$ is the push force, $A$ is the drill-bit cross section area, $A = 2.55 \cdot 10^{-3} m^2$. $\sigma'_{V0}$ is the effective vertical stress, $z$ is the depth and $\gamma'$ is the effective unit soil weight, which is here set to $\gamma' = 7\ kN/m^3$.

The set reduction is implemented by first defining two sets for each soil class.

One containing all $q_n$ curves, $S_A$, and another empty set, $S_B$. Members are iteratively moved from $S_A$ to $S_B$, in each step selecting the one that minimizes

$$d_{S.i} = \sum_{a \in S_A} \inf\{\, DTW(a,b) \mid b \in S_B \,\} \tag{5}$$

for steps $i = 1, 2, \ldots, n$. Figure 5A shows the first 100 steps of the reduction for each class. A threshold was set at $t = 20$, where all curves still had members and were past their elbow point. It is beyond the scope of this paper to discuss model parameter tuning, but the $AUC_{total}$-, $F_1$- and accuracy measures are used as described in [13]. As the soil classes are interrelated, predicting the nearest neighboring class during validation is counted as the correct answer. Figure 5B is presented to show that all the measures used support selecting the parameters $w = r = 20\%$ of the curve length and $k = 17$.

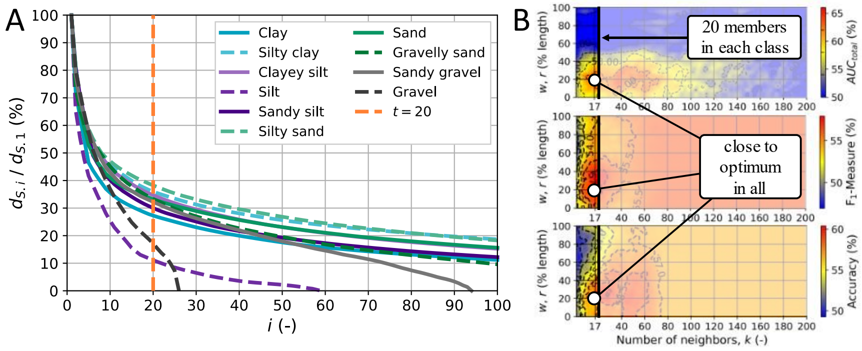


*Figure 5 A) Normalized average distance between all elements in $S_A$ to any element in $S_B$ as a function of the iteration step. B) $AUC_{total}$-, $F_1$ score and classification accuracy as a function model parameters k, w and r.*

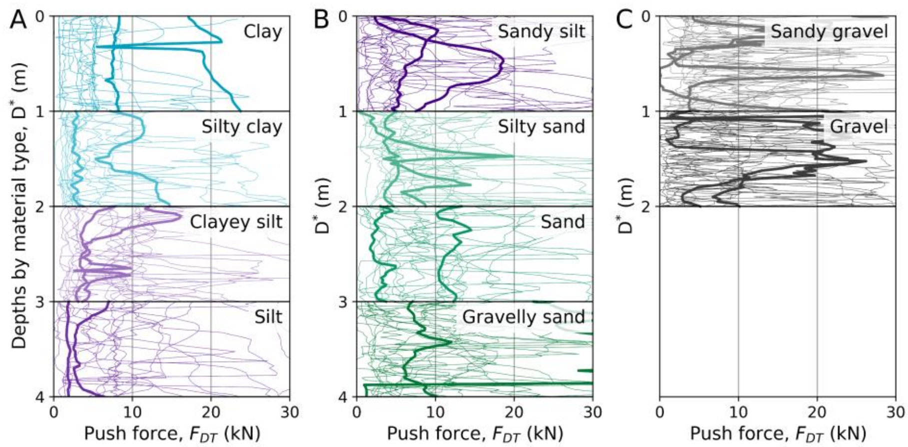


*Figure 6 A-C) Reduced training set with 20 representative push force curves for each material type. Each material plotted in 1m depth window. The first two curves selected for each class in Figure 5A are drawn with thicker lines.*

## 4. CLASSIFICATION EXAMPLES

Two total soundings are selected from the NPRA database to showcase the presented model capability in different soil types. Intervals with flushing or hammering are washed out as the model is not defined for these phases.

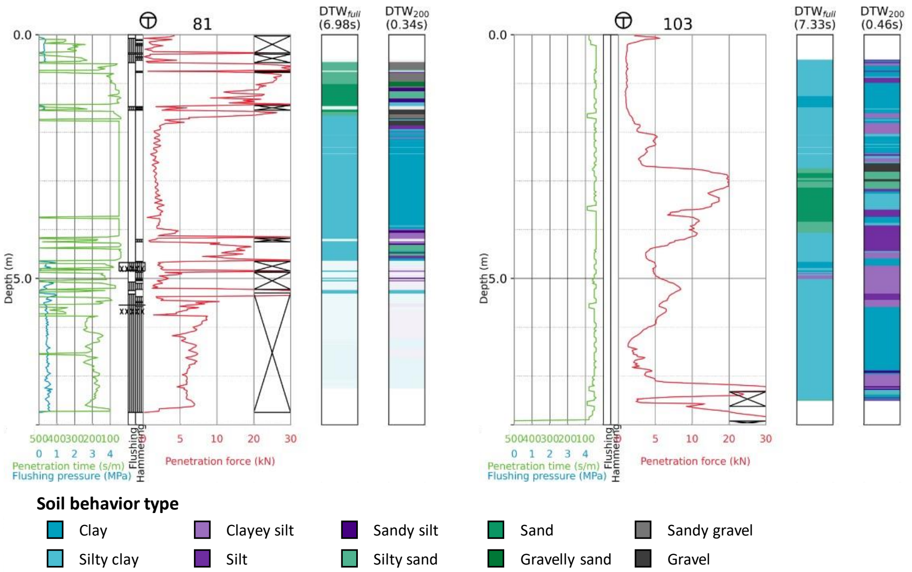


*Figure 7 Classification examples using the proposed model with the full- and reduced version of the dataset. Time needed to calculate each profile is given in parentheses.*

## 5. CONCLUSIONS

This work demonstrated that an SBT classifier for the total sounding can be defined with k-NN and the DTW distance using sequences of $q_n$ as features. This approach shows promise and should be investigated further.

The k-NN DTW approach should be applicable for other methods as well, such as the jb-tot, rotary pressure sounding and cone penetration tests.

Classification can be significantly sped up using representative datasets for each class. Examples are given on classification of soundings for the full- and trimmed dataset along with calculation times needed to generate both.

The quality of machine learning models is fully dependent on the quality of the training data. These could be improved using proper estimates of $\sigma_{v0}$ and $u_0$, rather than a fixed $\gamma'$. Efforts could be put into removing outliers or curves from improperly performed tests from the training set.

There is room for improvement on the total sounding method as well. A fundamental method improvement would be to implement a load cell at- or close to the tip, thus removing interference from friction against the rod system.

## ACKNOWLEDGEMENT

The authors would like to thank the Research Council of Norway (RCN), NPRA, Bane NOR and The Norwegian Water Resources and Energy Directorate (NVE) for supporting this work.

## REFERENCES

[1] F. Fredriksen, F. Oset, N. Rygg: Totalsondering – en rasjonell metode for kartlegging av lösmasser. NGM, 1992

[2] G. Nilsson and G. Forssman. Total sounding – Some experience in Sweden and development potential. NGM, 2004.

[3] E. Haugen, S.A. Degago, O.V. Kirkevollen, D.Nigussie, X. Yu: A Preliminary attempt towards soil classification chart from total sounding. NGM,2016.

[4] SGF: Rapport 4. Metodbeskrivning för jord-bergsondering. Utförande, utrustning och kontroll. Svenska Geotekniska Föreningen, 2009

[5] NGF: Melding Nr. 9. Veiledning for utførelse av totalsondering. Norsk Geoteknisk Forening, 2018.

[6] T. Kydland, H. Firman, K. Aunaas: Geotolk – Tolkning av totalsondering med maskinlæring. Geoteknikkdagen, 2021.

[7]T.K. Vintsyuk: Speech Discrimination by Dynamic Programming. Cybernetics and Systems Analysis, vol. 4, no. 1. 1968.

[8] E. Fix, J.L. Hodges, Jr.: Discriminatory Analysis-Nonparametric Discrimination: Consistency Properties. Report No. 4. Project Number 21-49-004, USAF School of Aviation Medicine, Randolph Field, Texas. 1951.

[9] NGF: Melding Nr. 2. Veiledning for symboler of definisjoner I geoteknikk. Identifisering og klassifisering av jord. Norsk Geoteknisk Forening, 2011.

[10] D.F. Silva, G.E.A.P.A. Batista, E. Keogh: On the Effect of Endpoints on Dynamic Time Warping. SIGKDD Workshop on Mining and Learning from Time Series, II, 2016.

[11] H. Sakoe, S. Chiba: Dynamic Programming Algorithm Optimization for Spoken Word Recognition. IEEE Transactions on speech, and signal processing, vol ASSP-26 no. 1, 1978.

[12] C. A. Ratanamahatana and E. Keogh, "Three Myths about Dynamic Time Warping Data Mining," in Proceedings of the 2005 SIAM International Conference on Data Mining, Newport Beach, 2005.

[13] T. Fawcett: An introduction to ROC analysis. Pattern Recognition Letters, vol. 27, no. 8, 2006.