Spatial and temporal analysis of road deformation based on remote sensing and subsurface exploration

Melika Sajadian

Mentor #1: Dr.ir. Mathias Lemmens
Mentor #2: Dr. Faraz Sadeghi Tehrani

Supervisor (Deltares): Dr. Ana Teixeira

Co-reader: Amin Askarinejad

TU Delft

Deltares
Enabling Delta Life

Outline

- Introduction and related work
- Methodology and implementation
- Results and discussions
- Conclusion and future works

TUDelft Deltares Enabling Delta Life

Outline

- Introduction and related work
- Methodology and implementation
- Results and discussions
- Conclusion and future works

Problem and motivation


Image courtesy: www.thefiscaltimes.com

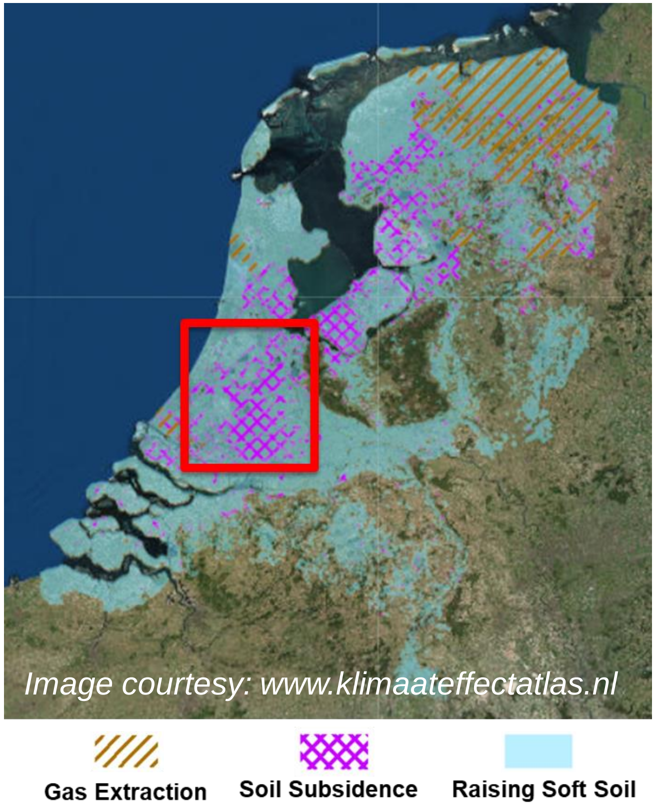


Image courtesy: www.klimaateffectatlas.nl

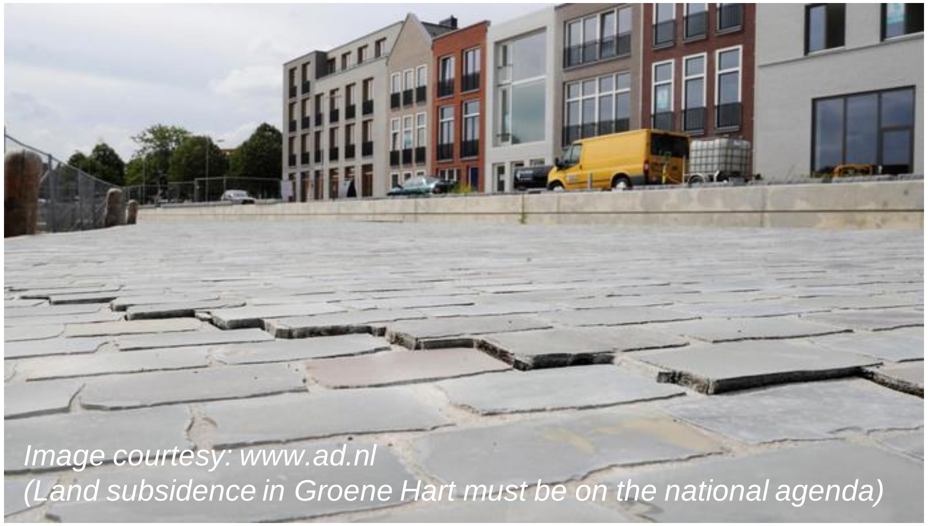


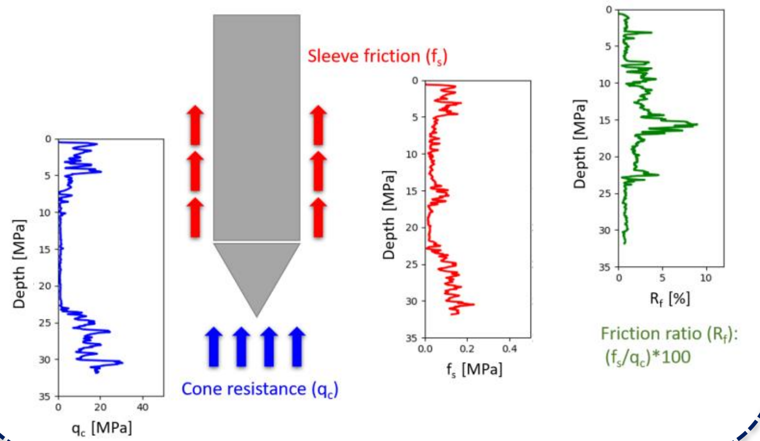
Image courtesy: www.ad.nl
(Land subsidence in Groene Hart must be on the national agenda)


Image courtesy: www.tno.nl

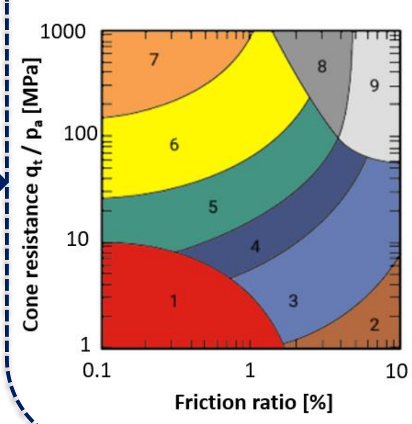
TU Delft
Deltares
Enabling Delta Life

Solutions: deformation prediction

In-situ measurements

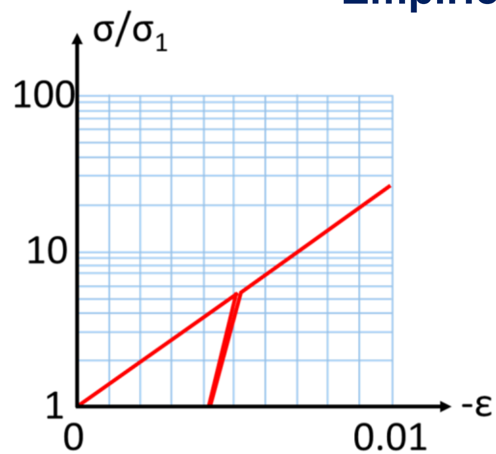


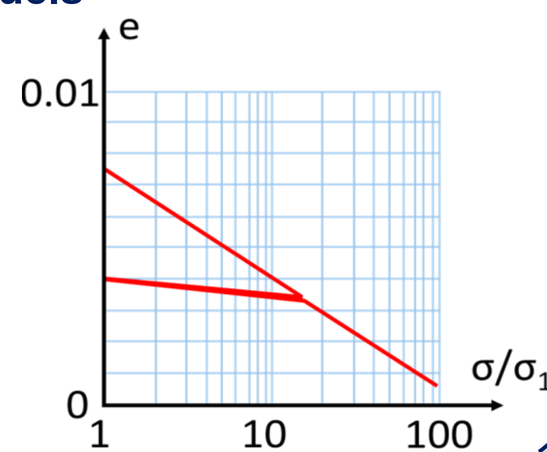
Empirical charts and tables



Zone	Soil type
1	Sensitive fine - grained
2	Clay – organic soil
3	Clays: clay to silty clay
4	Silt mixtures: clayey silt and silty clay
5	Sand mixtures: Silty sand to sandy silt
6	Sands: clean sands to silty sands
7	Dense sand to gravelly sand
8	Stiff sand to clayey sand
9	Stiff fine -grained

Empirical models



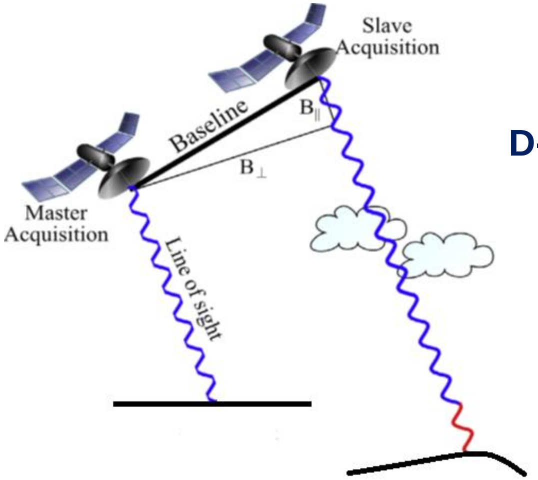


TU Delft
Deltares
Enabling Delta Life

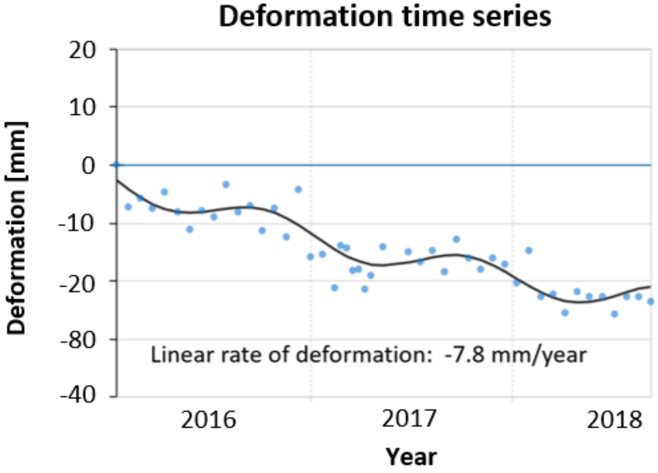
Solutions: deformation monitoring

Leveling

GNSS measurements



D-InSAR techniques



Challenges

- Deformation prediction (geotechnical)
  - Empiricism in defining soil classes, models, parameters
  - Diversity of models
  - Subjective results

- Deformation monitoring (D-InSAR)
  - Accurate measurements but cannot be used for prediction

Related works

- Soil types, transport infrastructure data and climate classes to monitor the response of roads and railways to ground deformation *(North et al. [2017])*
- D-InSAR techniques for monitoring road deformation *(North et al. [2017])*
- Borehole data and loading history for studying subsidence *(Asselen et al. [2018])*
- Mapping subsidence potential using CPT *([Koster et al., 2018])*
- Different machine learning algorithms used to model the relationship between the influential parameters and ground deformation *(Tien Bui et al. [2018], Ilia and Loupasakis [2018], Zhou et al. [2019])*
- Relationship between seasonal deformation and temperature and precipitation *(Ozer et al. [2019])*

Research questions

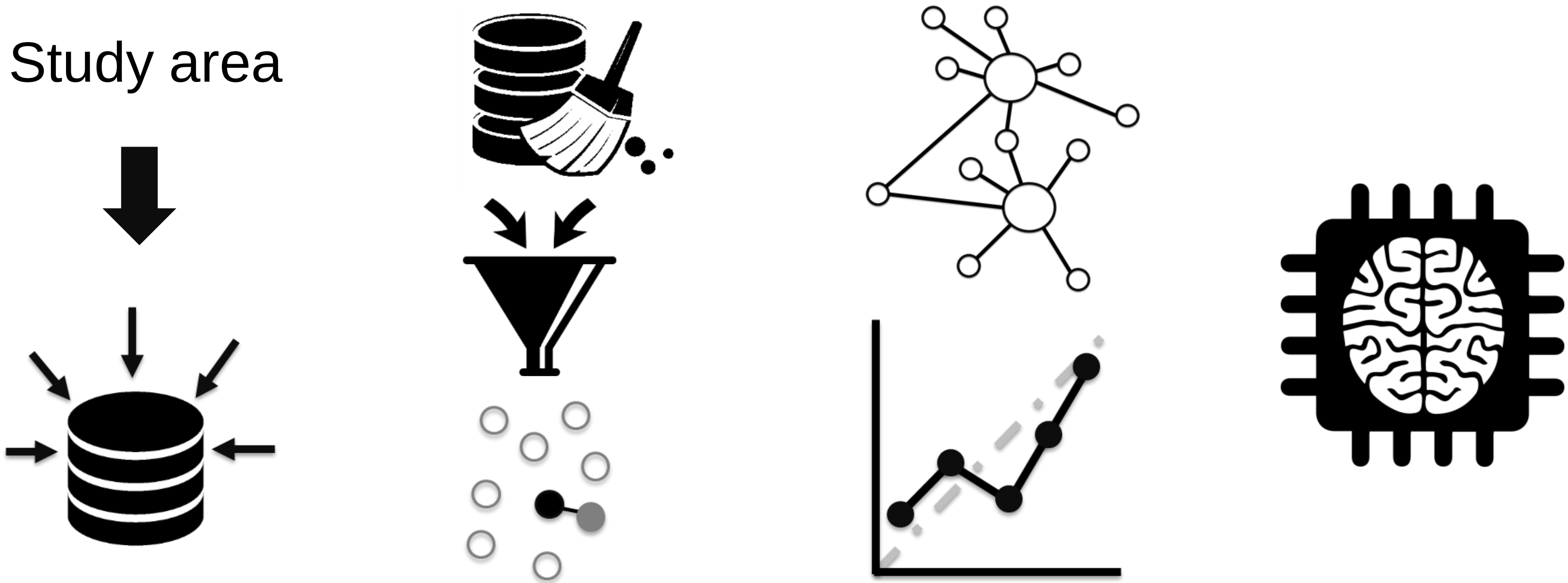
*Using machine learning techniques, is it possible to model spatial-temporal relationship between the **deformation measurements,** the **soil properties**, and **loading/unloading** on roads?*

- What are the data sources needed for studying soil properties, loading/unloading and deformation measurements?
- Is there a correlation between soil properties, loading/unloading and deformation measurements?
- What parameters/features should be included from the available data sets?
- What machine learning algorithm(s) are more suitable in establishing the relationship?
- What is the accuracy of the chosen machine learning technique and is it satisfactory?

Outline

- Introduction and related work
- Methodology and implementation
- Results and discussions
- Conclusion and future works

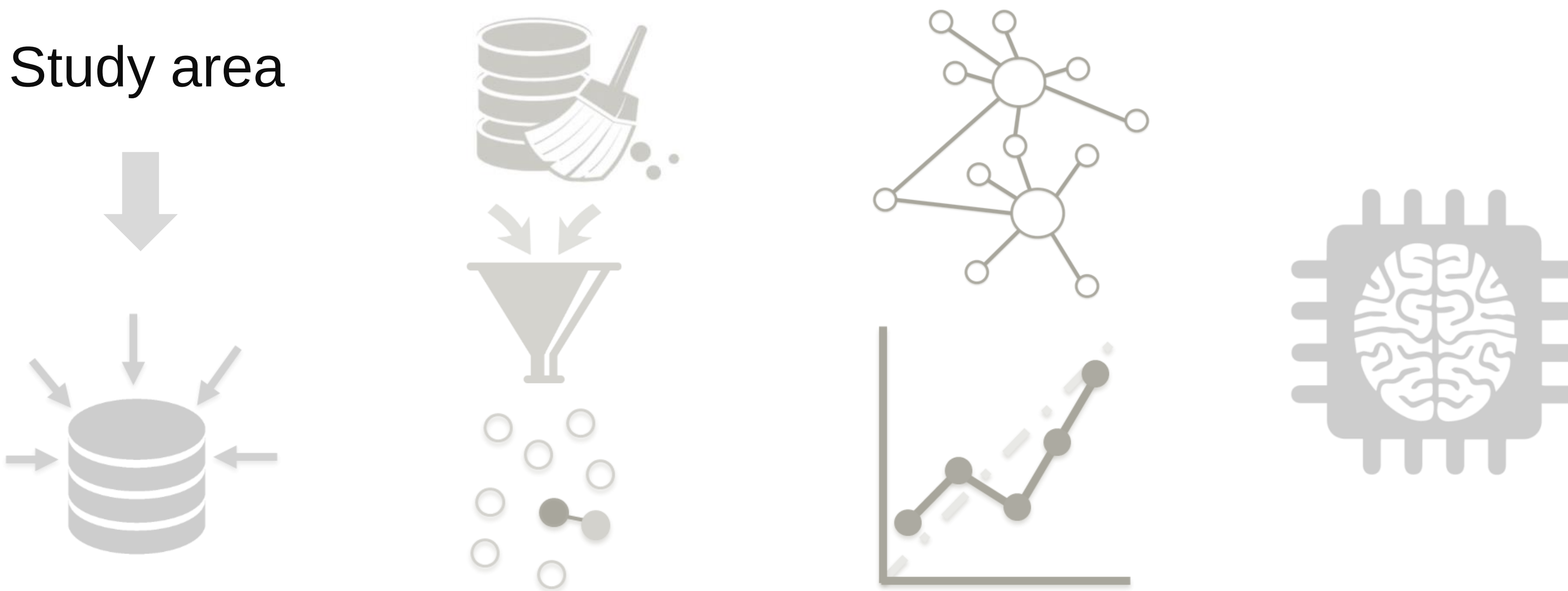
Overview of the methodology



Data collection → Pre-processing → Correlations and similarities → Modeling through machine learning

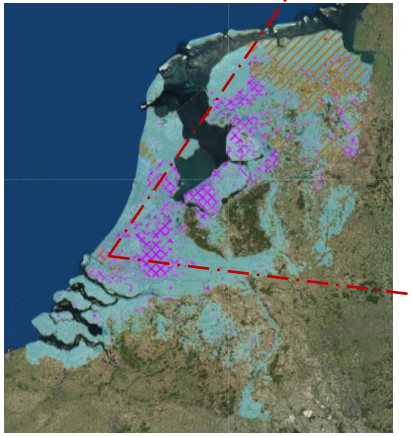
Overview of the methodology

Study area

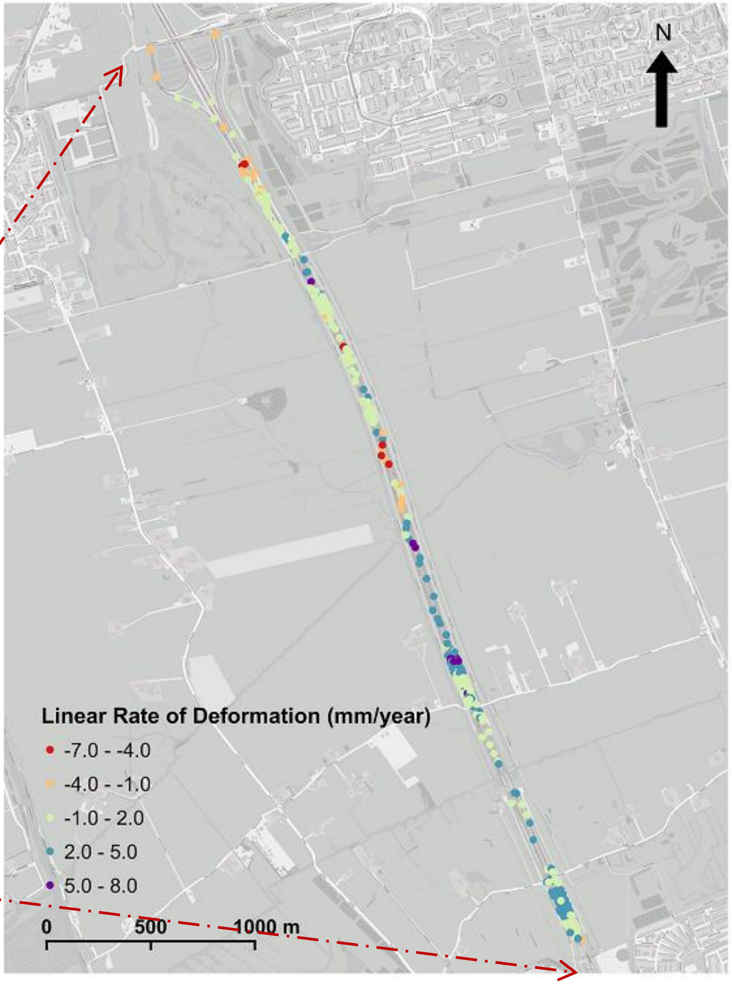


Data collection → Pre-processing → Correlations and similarities → Modeling through machine learning

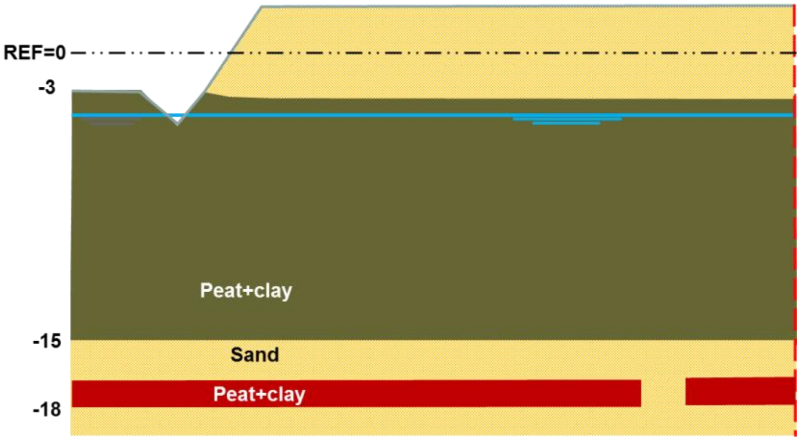
Study area: A4 highway (Schiedam-Delft)



Linear Rates of Deformation on the A4 Highway

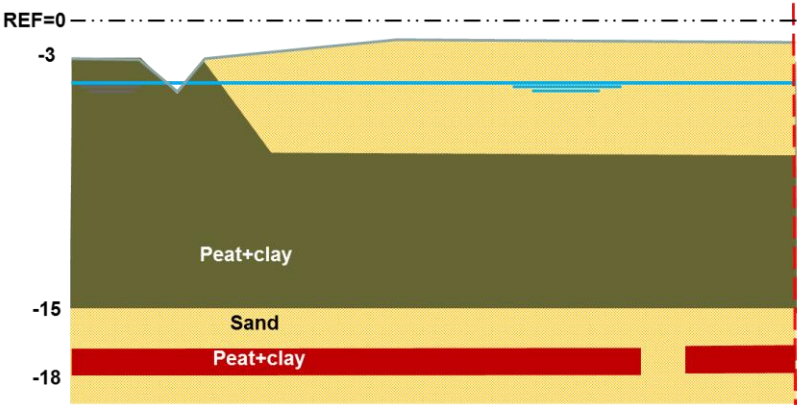


Study area: A4 highway (Schiedam-Delft)



Design in 1970s

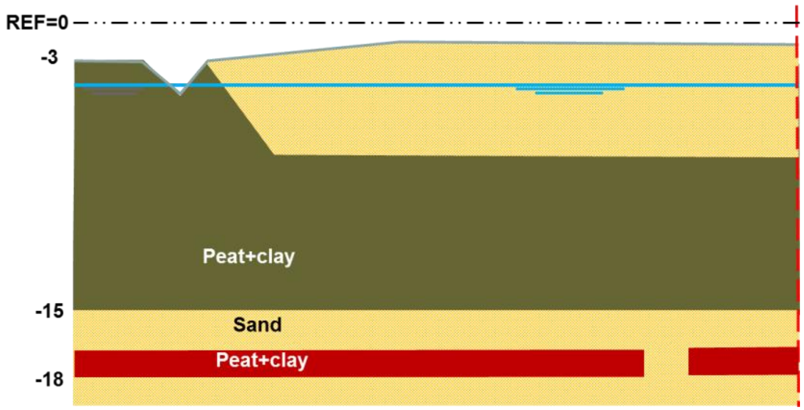
Study area: A4 highway (Schiedam-Delft)

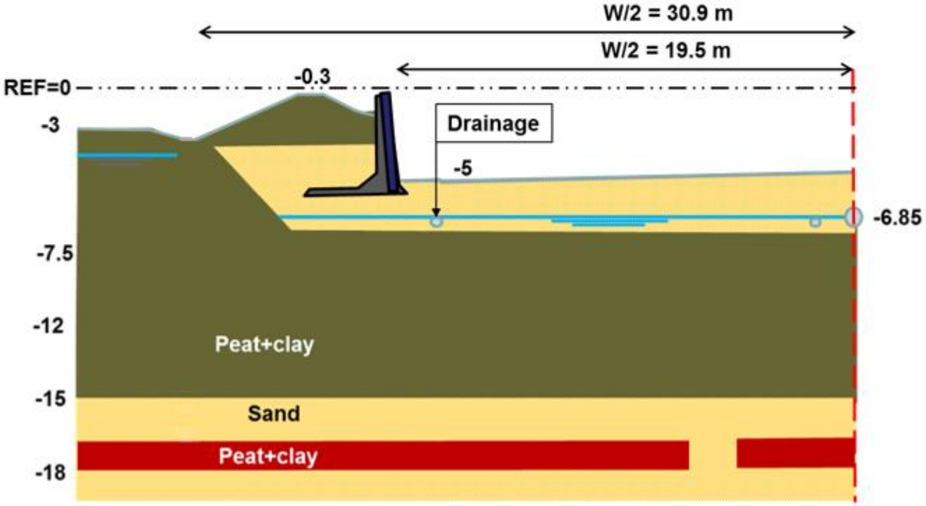


Design in 1970s

Consolidation of soil

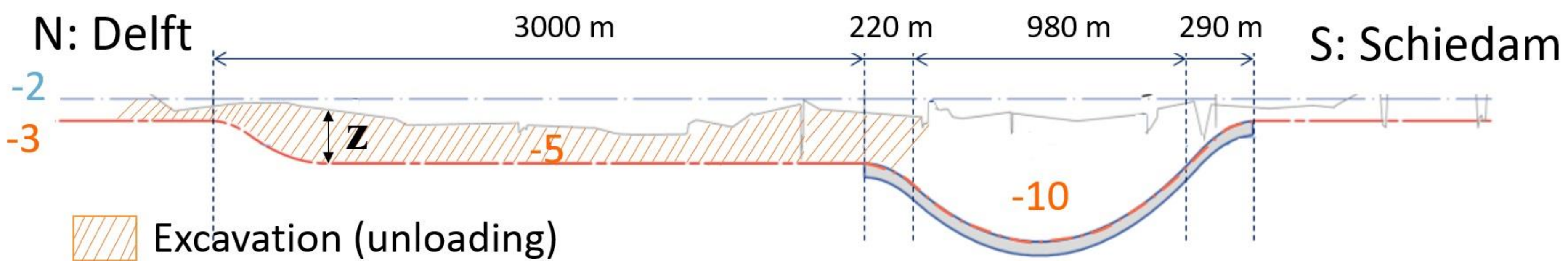
Study area: A4 highway (Schiedam-Delft)





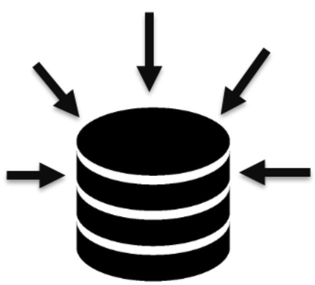
Design in 1970s
Consolidation of soil

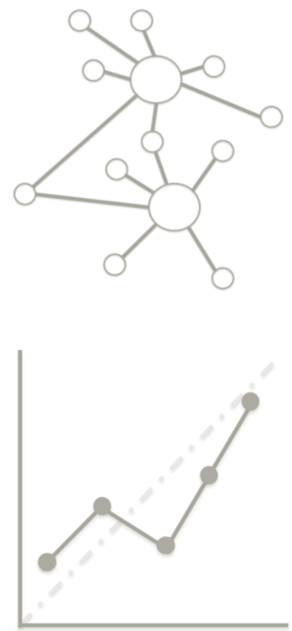
Design A4 2011
Shallow cutting L = 3 km



Overview of the methodology

Study area





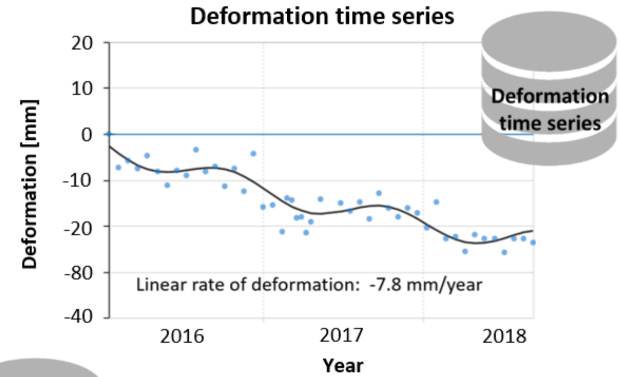
Data collection

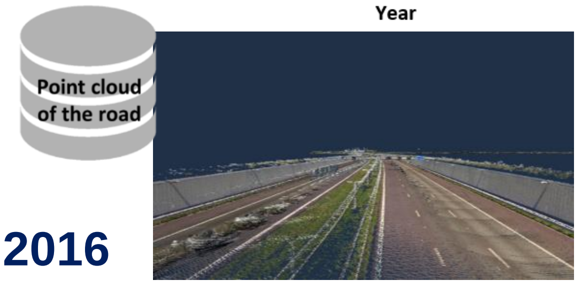
Pre-processing

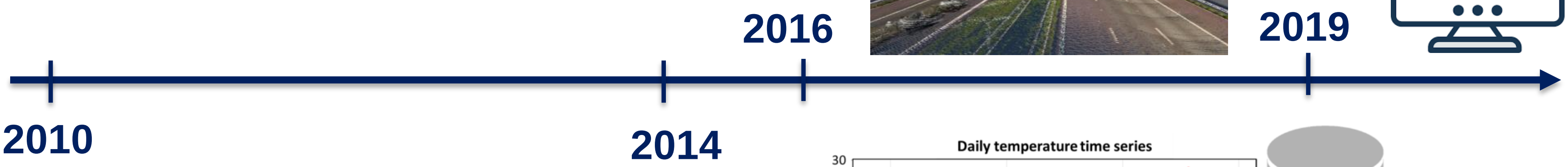
Correlations and similarities

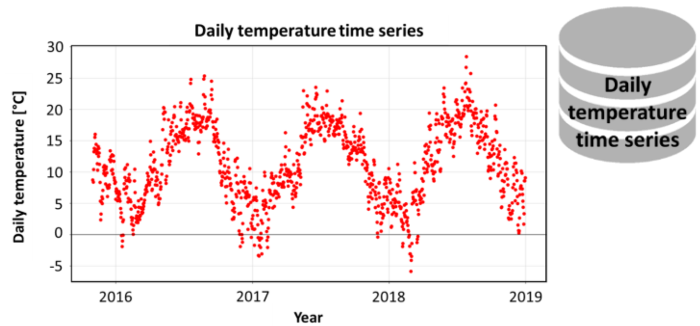
Modeling through machine learning

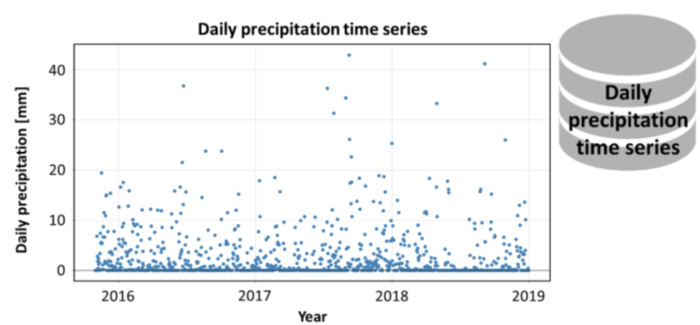
Data set

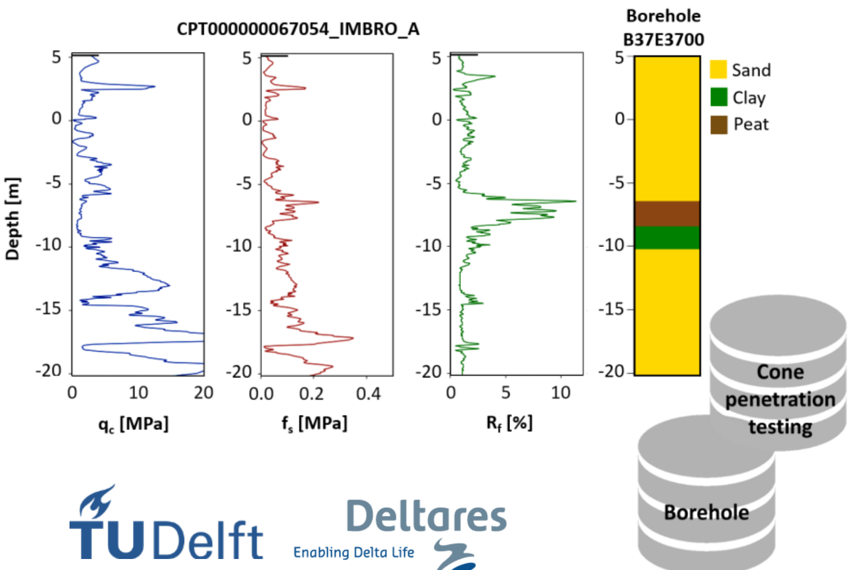

Deformation time series
Deformation time series
20
10
0
-10
-20
-80
-40
Deformation [mm]
Linear rate of deformation: -7.8 mm/year
2016
2017
2018
Year


Point cloud of the road


2010
2014
2016
2019


Daily temperature time series
Daily temperature time series
Daily temperature [°C]
30
25
20
15
10
5
0
-5
2016
2017
2018
2019
Year

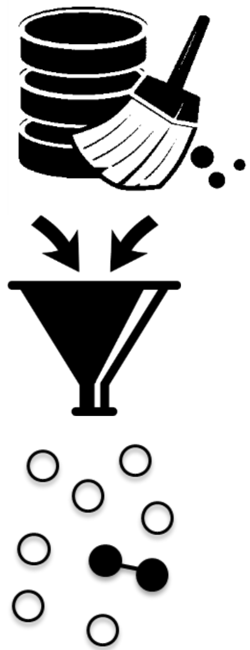

Daily precipitation time series
Daily precipitation time series
Daily precipitation [mm]
40
30
20
10
0
2016
2017
2018
2019
Year


CPT000000067054_IMBRO_A
Depth [m]
5
0
-5
-10
-15
-20
0
10
20
q_c [MPa]
0.0
0.2
0.4
f_s [MPa]
0
5
10
R_f [%]
Borehole
B37E3700
Sand
Clay
Peat
Cone penetration testing
Borehole

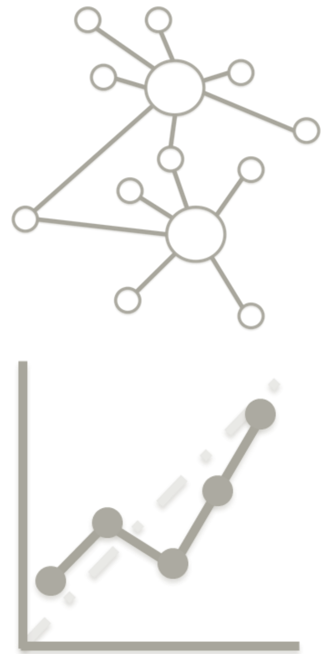
Overview of the methodology

Study area

Data collection



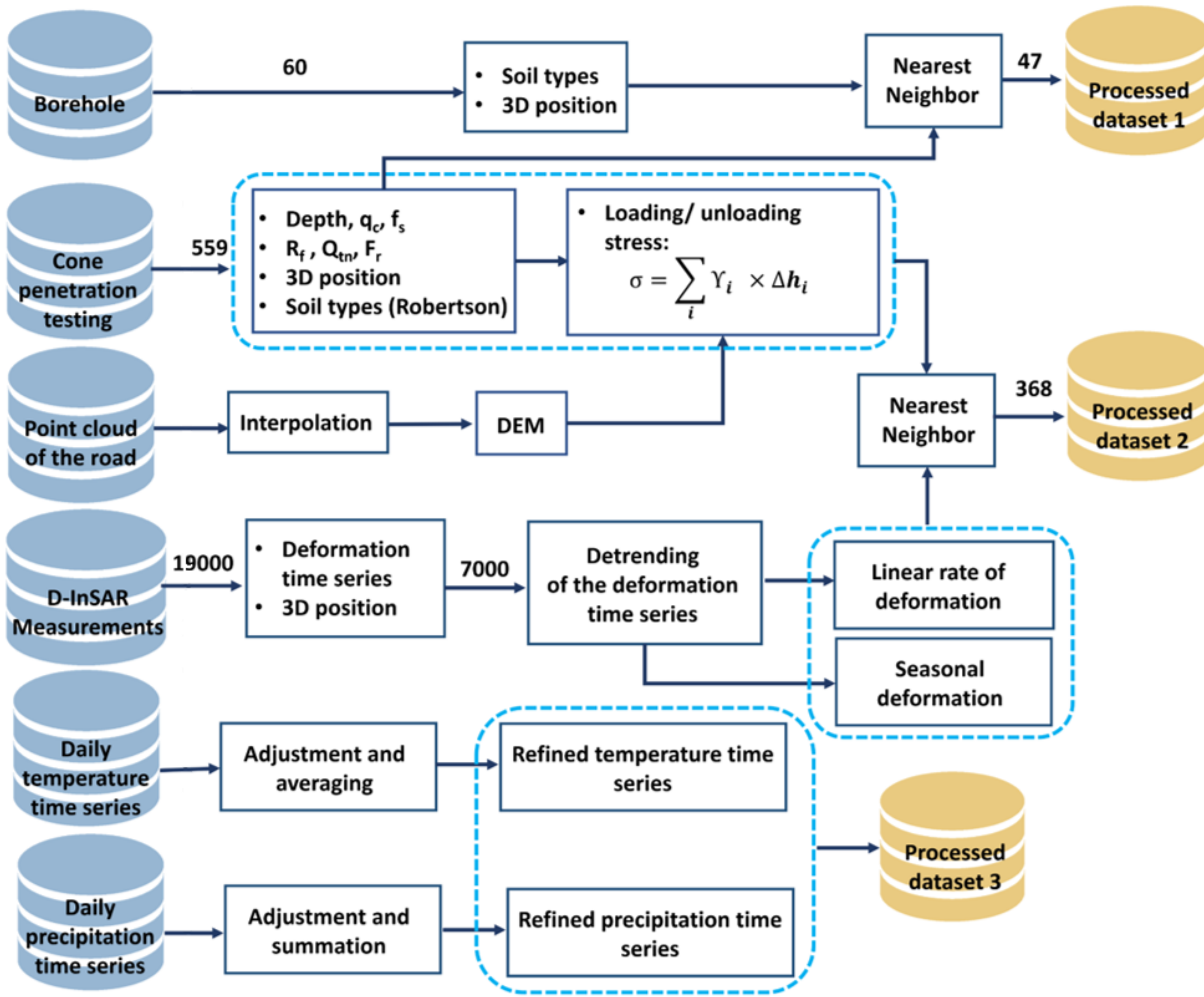
Pre-processing



Correlations and similarities

Modeling through machine learning

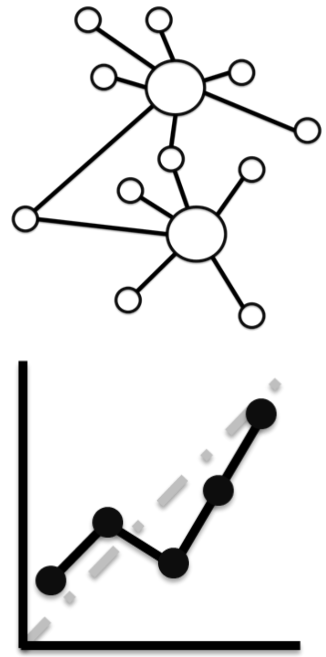
Step 1: Pre-processing



TU Delft Deltares Enabling Delta Life

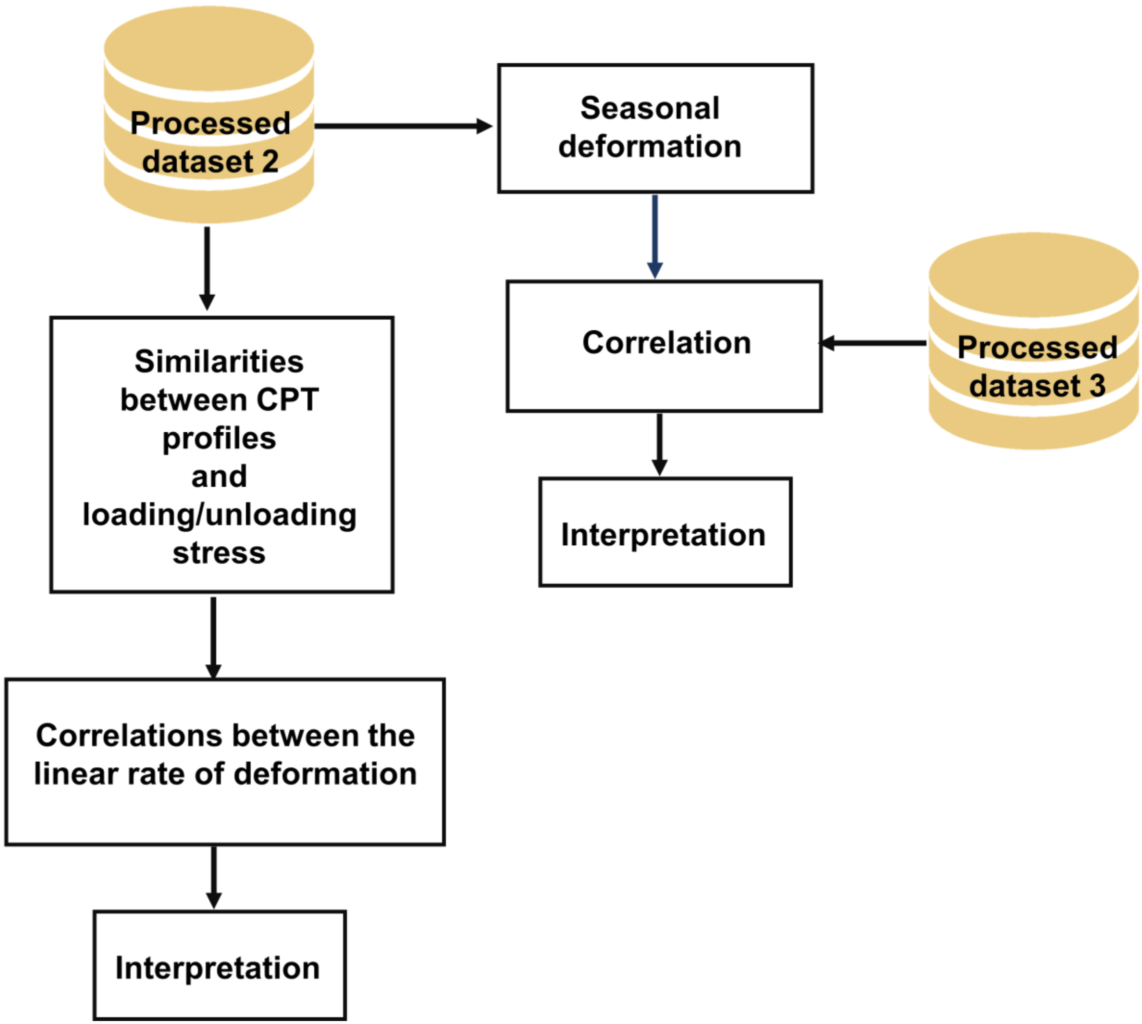
Overview

Study area



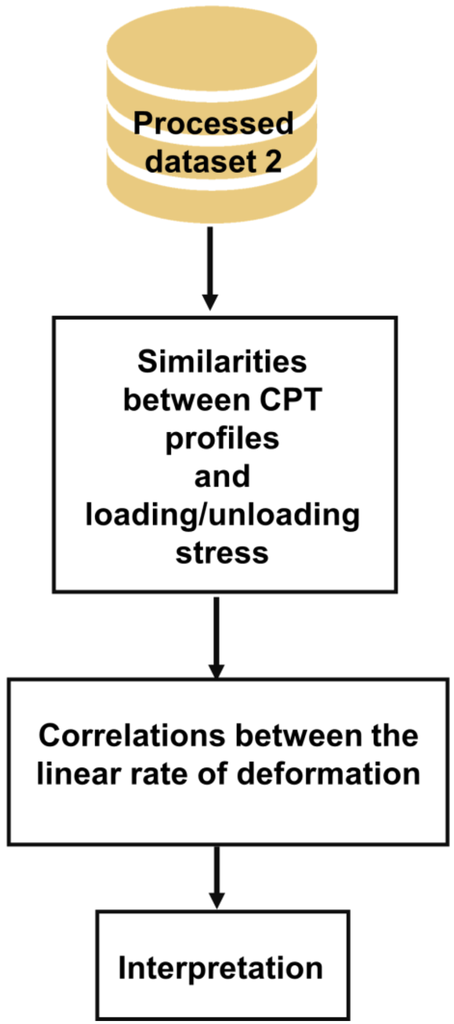
Data collection → Pre-processing → **Correlations and similarities** → Modeling through machine learning

Step 2: Correlations and similarities



Step 2: Correlations and similarities

$$d_{L_2}(x, y) = \left(\sum_{i=1}^{M} (x_i - y_i)^2\right)^{\frac{1}{2}}$$



Step 2: Correlations and similarities

$$d_{L_2}(x,y) = \Big(\sum_{i=1}^{M} (x_i - y_i)^2\Big)^{\frac{1}{2}}$$

Data point *i* and *j* are similar if:

$$\begin{cases} N(N(d_{L_2}(q_{c_i}, q_{c_j})) + N(d_{L_2}(f_{s_i}, f_{s_j}))) < 0.02 \\ \Delta\sigma_{ij} < 10 \text{ [kPa]} \end{cases}$$

N stands for normalization

Processed dataset 2

Similarities between CPT profiles and loading/unloading stress

Correlations between the linear rate of deformation

Interpretation

TU Delft Deltares Enabling Delta Life

Step 2: Correlations and similarities

Let

x_i be the linear rate of deformation of point *i*

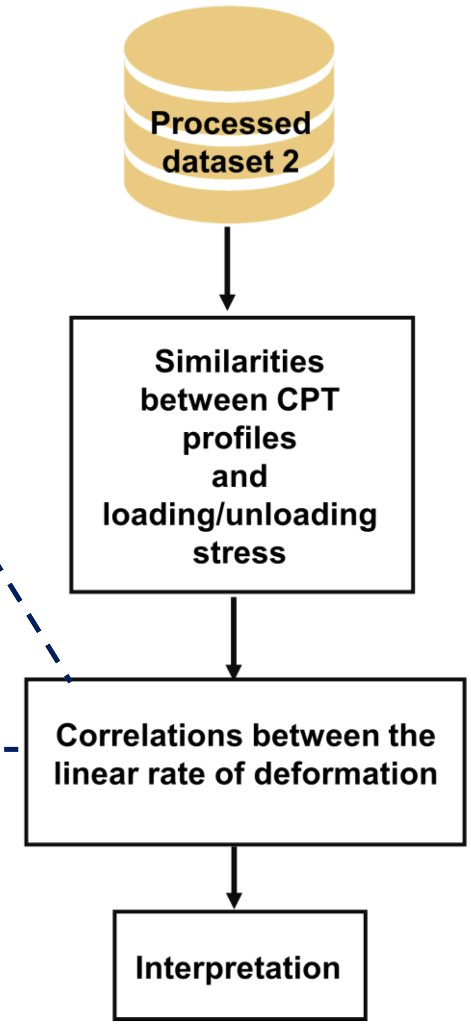
and

y_i be the mean of linear rate of deformation of its similar points

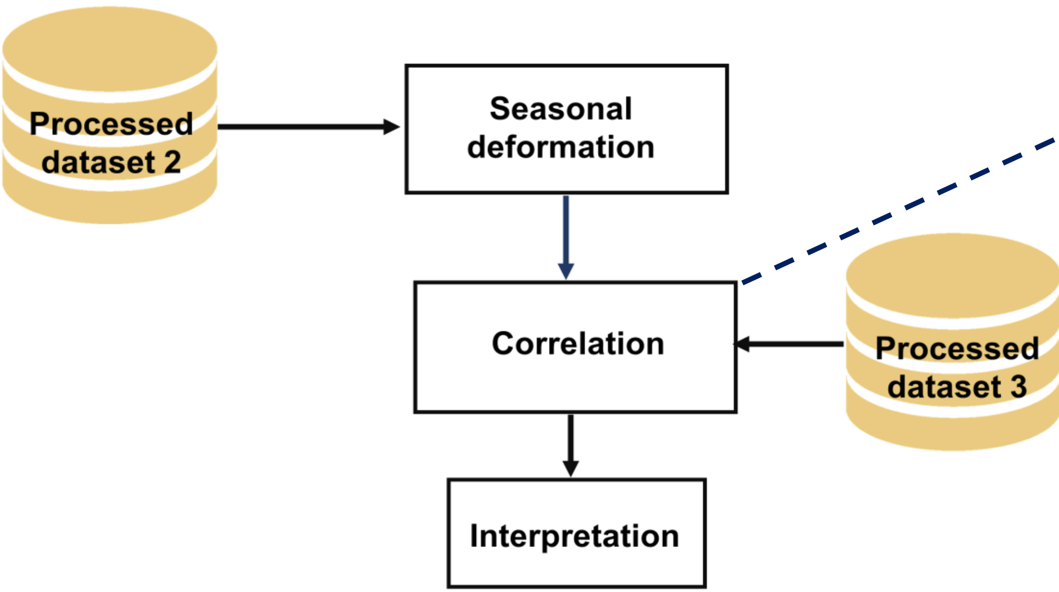
Correlation measures can be:

$$\boldsymbol{\rho_{xy}(k) = \frac{E[(x - \mu_x)(y - \mu_y)]}{\sigma_x \sigma_y}}$$

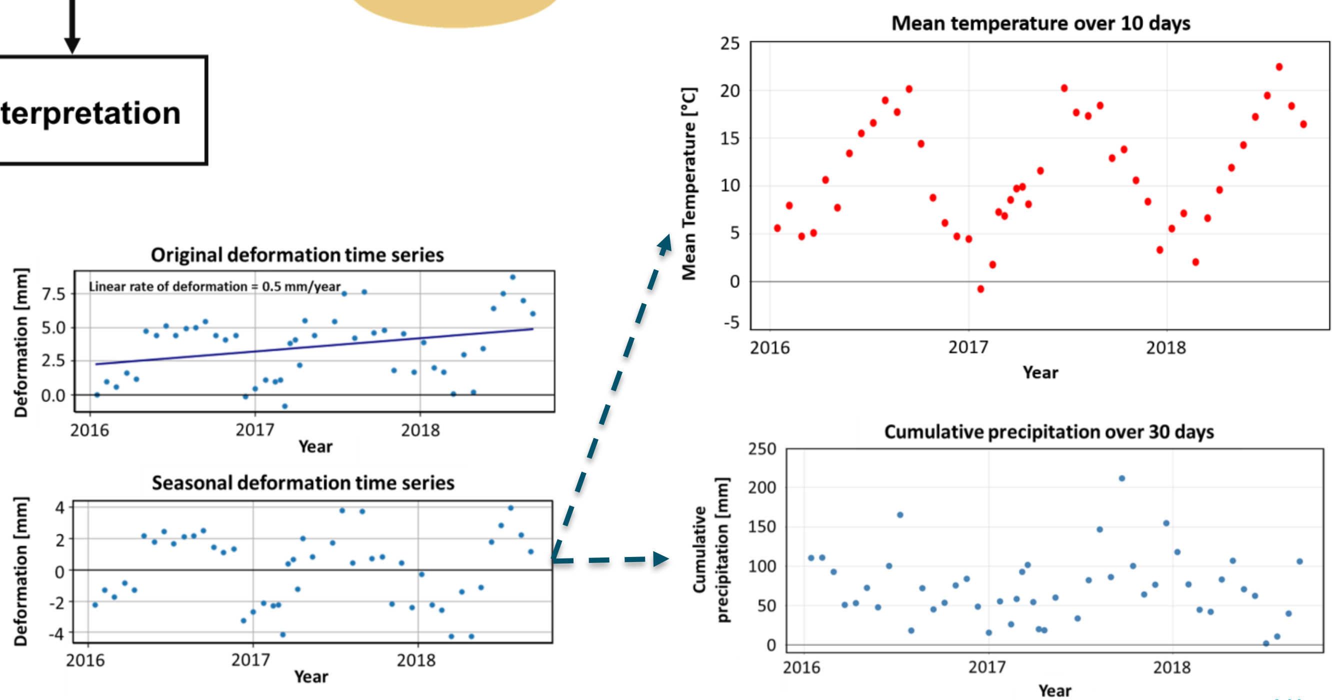
$$\boldsymbol{R_{xy}(k) = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \mu_x)^2}}$$



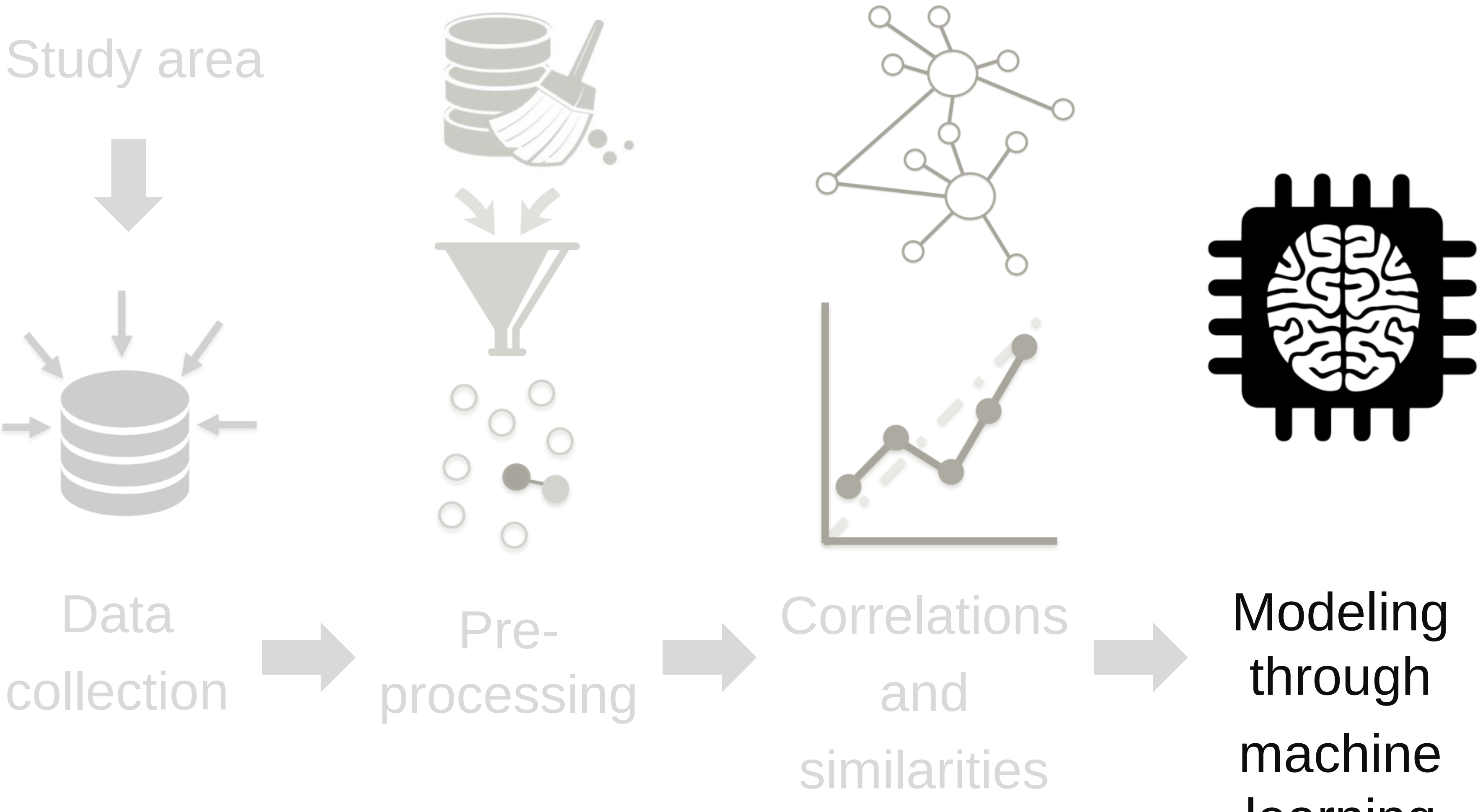
Step 2: Correlations and similarities



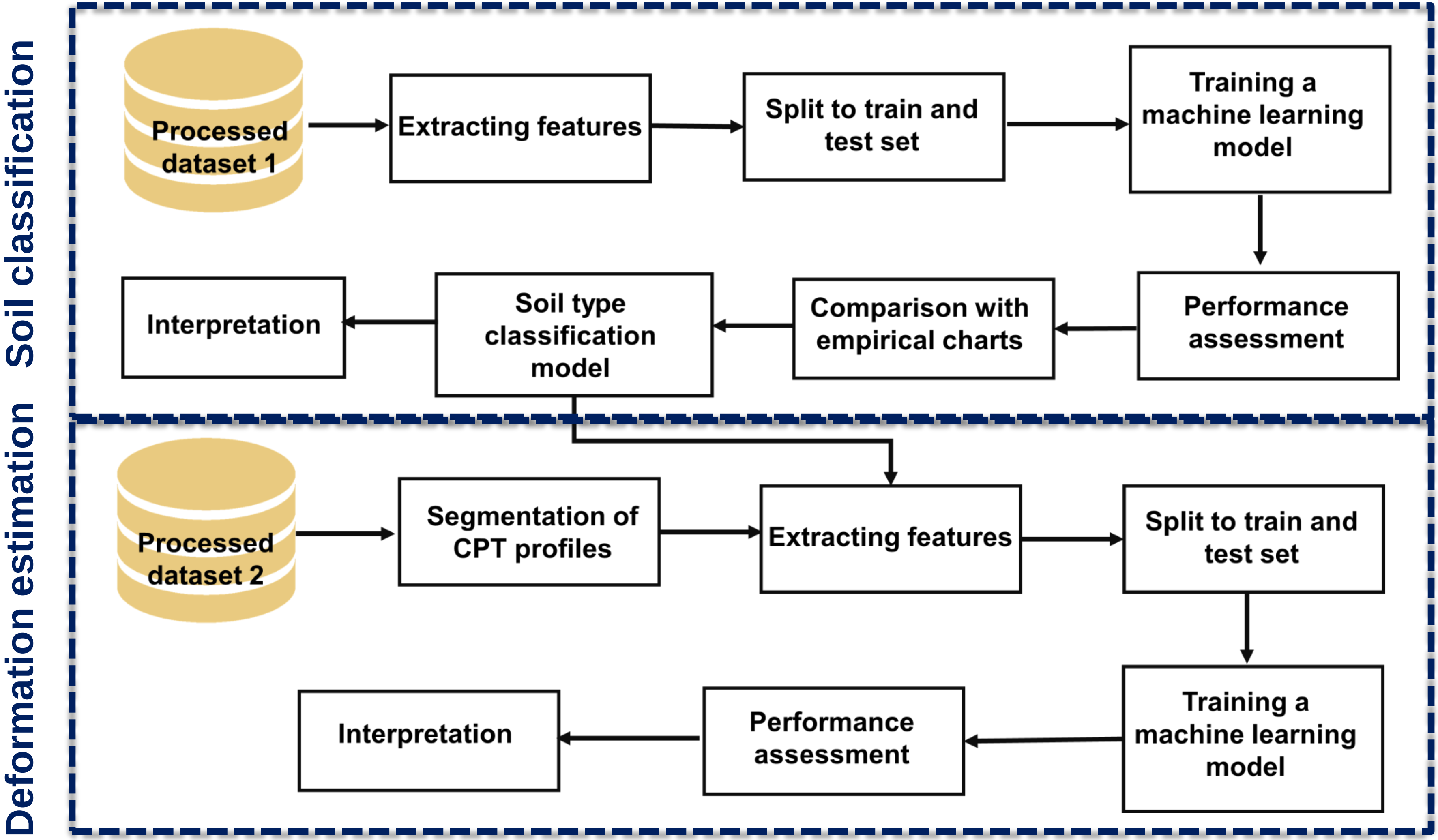
$$\rho_{xy}(k) = \frac{E[(x_t - \mu_x)(y_{t-k} - \mu_y)]}{\sigma_x \sigma_y}$$



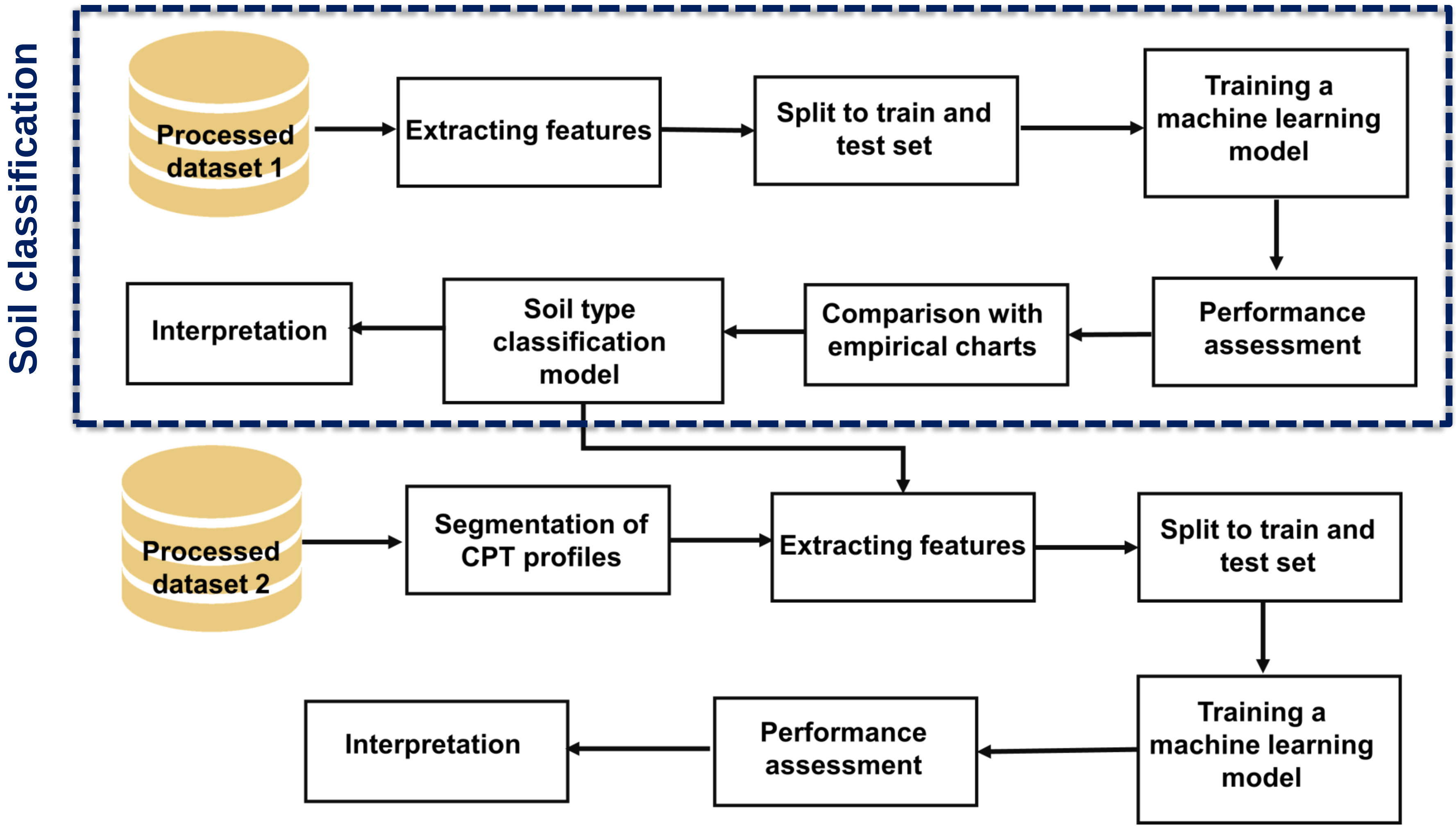
Overview of the methodology



Step 3: Modeling through machine learning

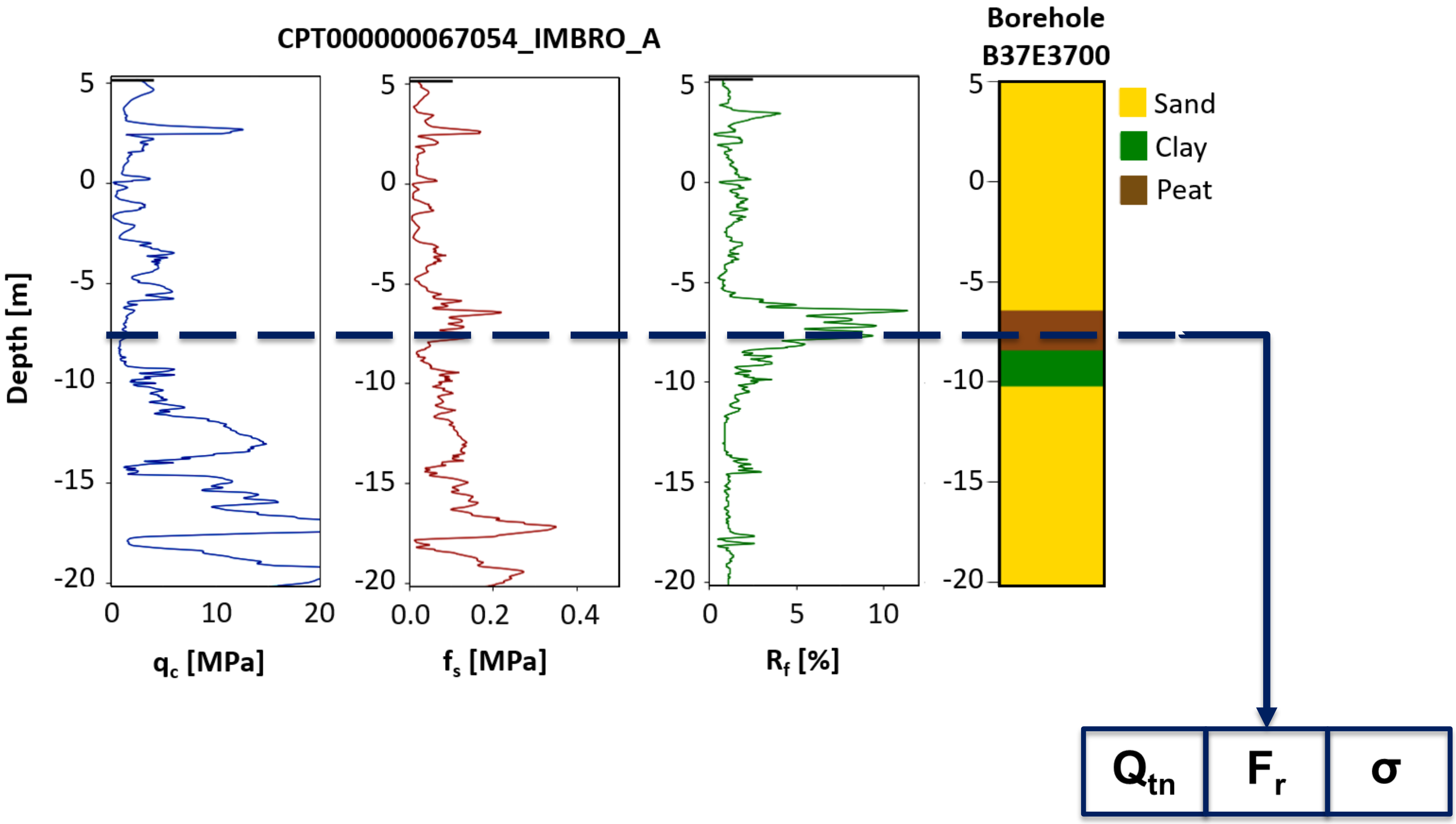


Step 3: Modeling through machine learning

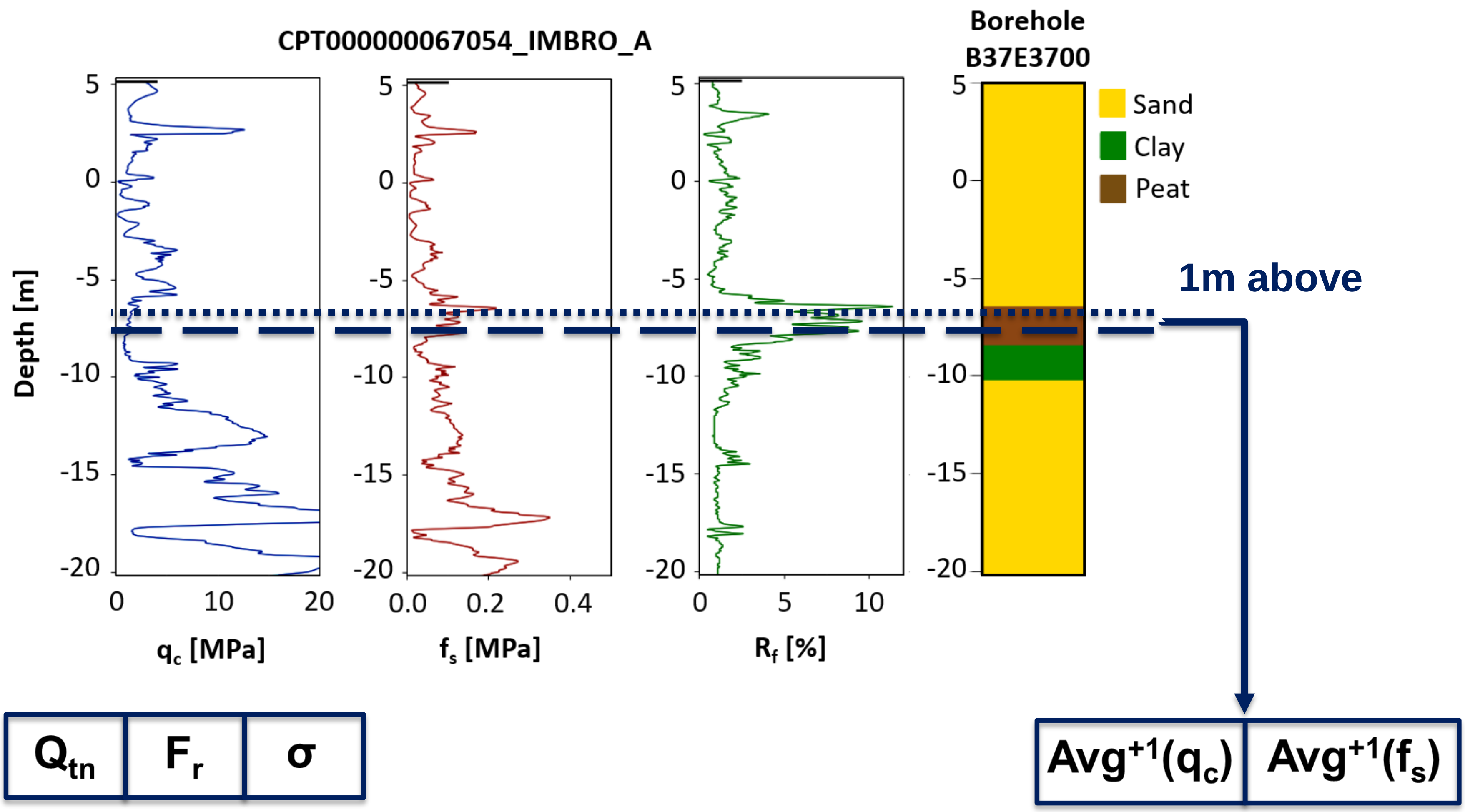


TU Delft Deltares Enabling Delta Life

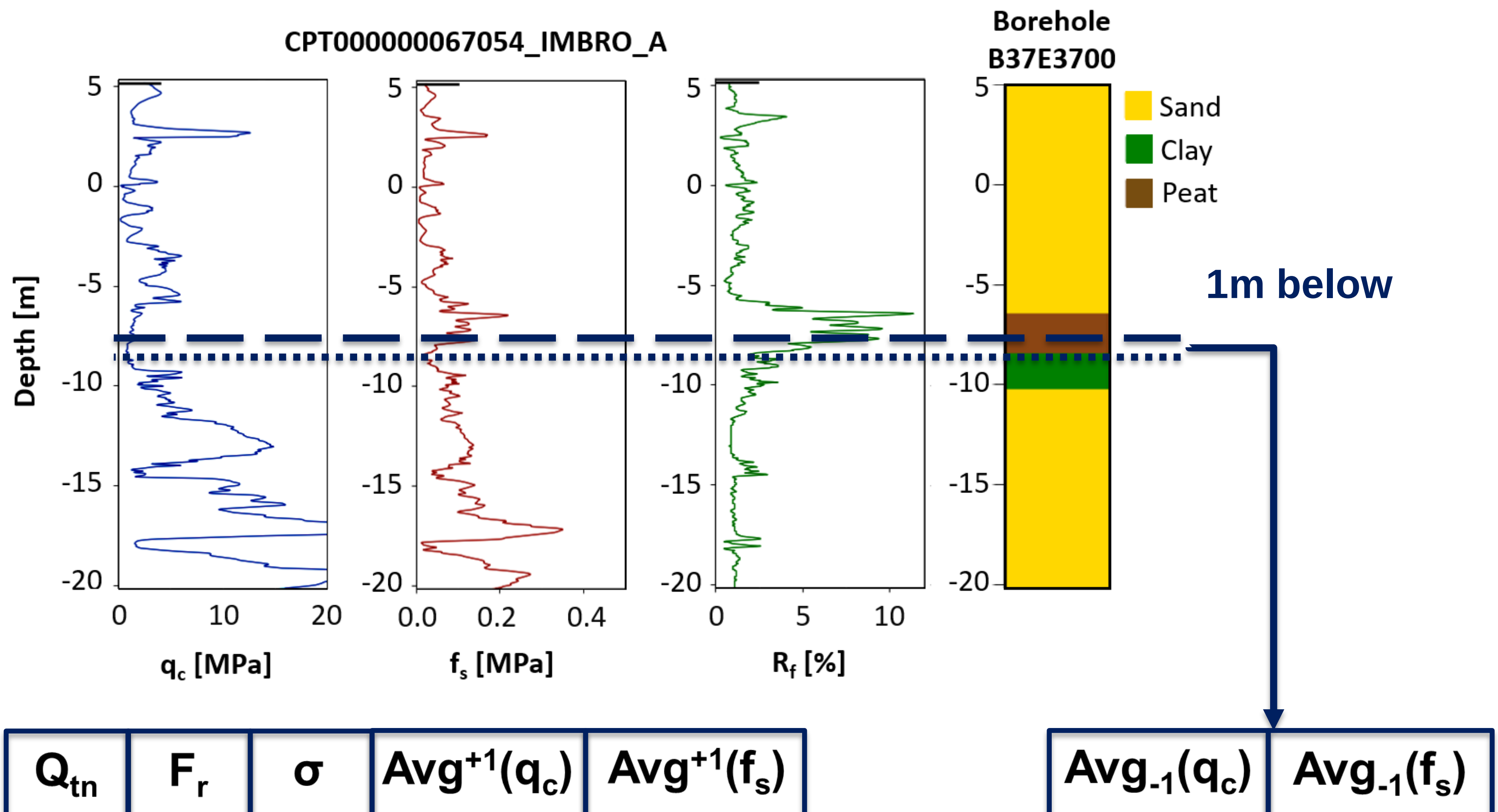
Feature extraction



Feature extraction

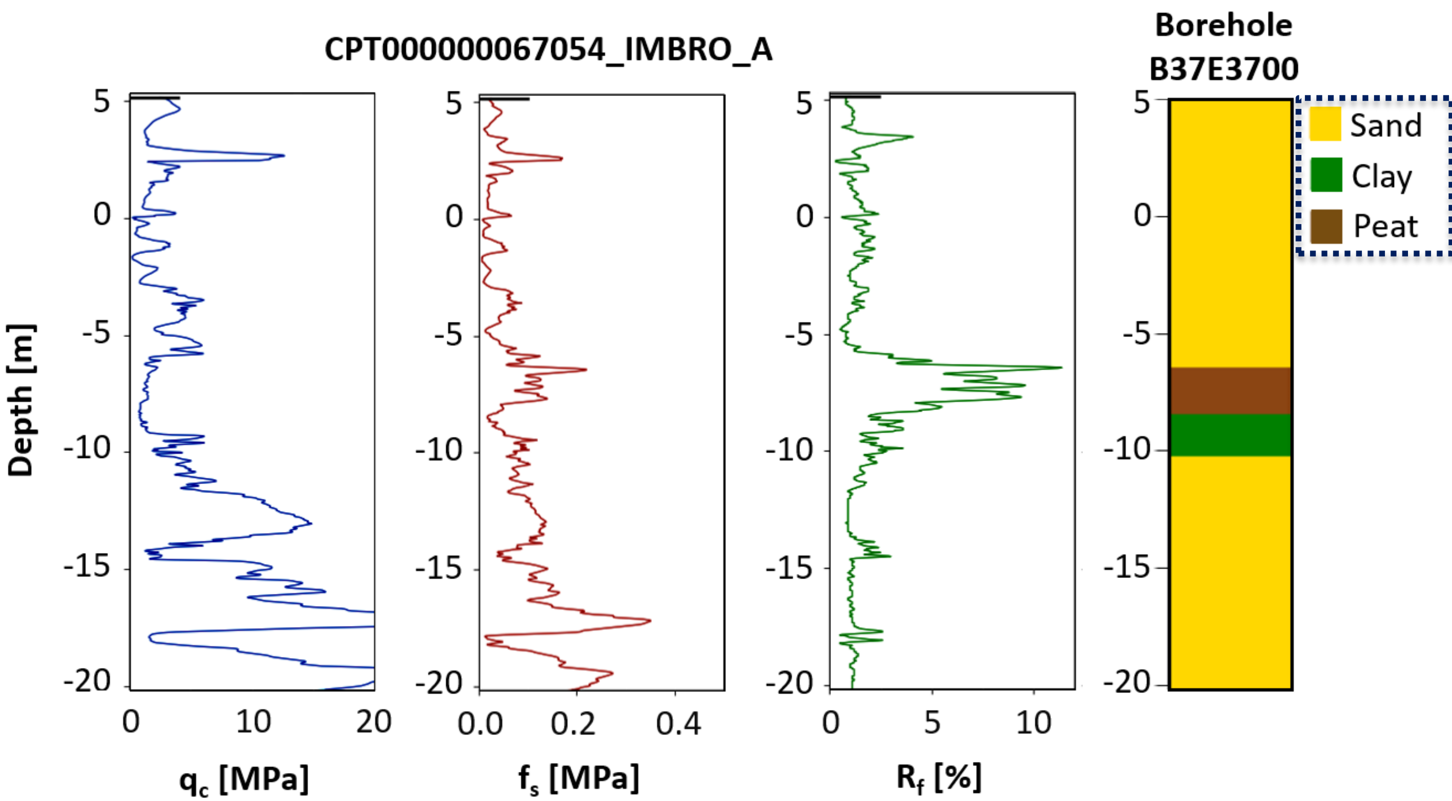


Feature extraction



TU Delft
Deltares
Enabling Delta Life

Feature extraction

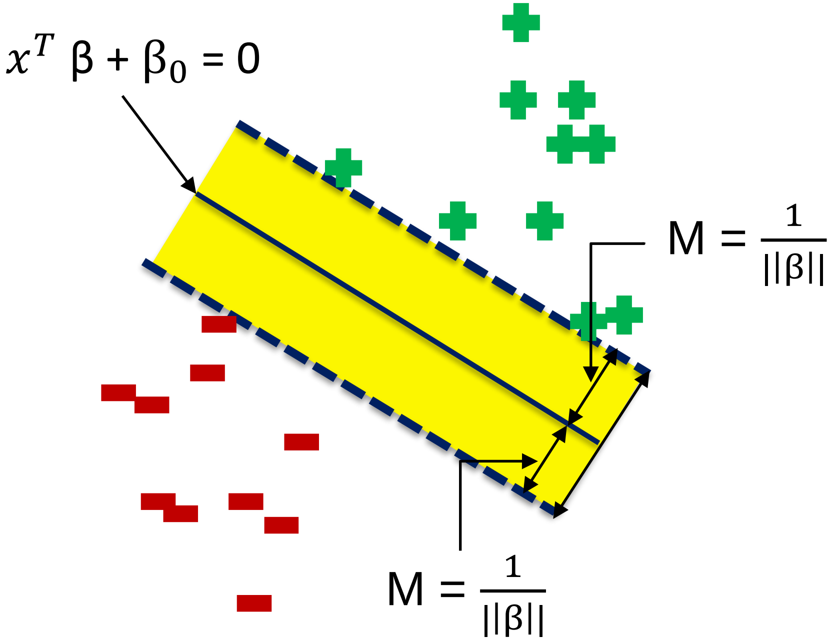


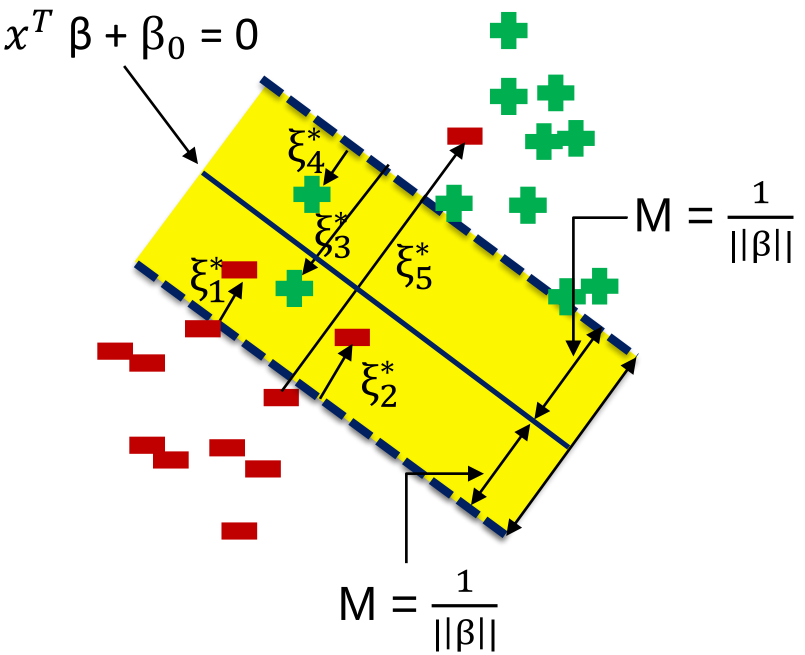
Q_{tn}	F_r	σ	$Avg^{+1}(q_c)$	$Avg^{+1}(f_s)$	$Avg_{-1}(q_c)$	$Avg_{-1}(f_s)$	Soil type

Feature vector (7 features)

Label

Supervised Learning: SVM





$Where\ \ \xi_1^* = M\xi_1$

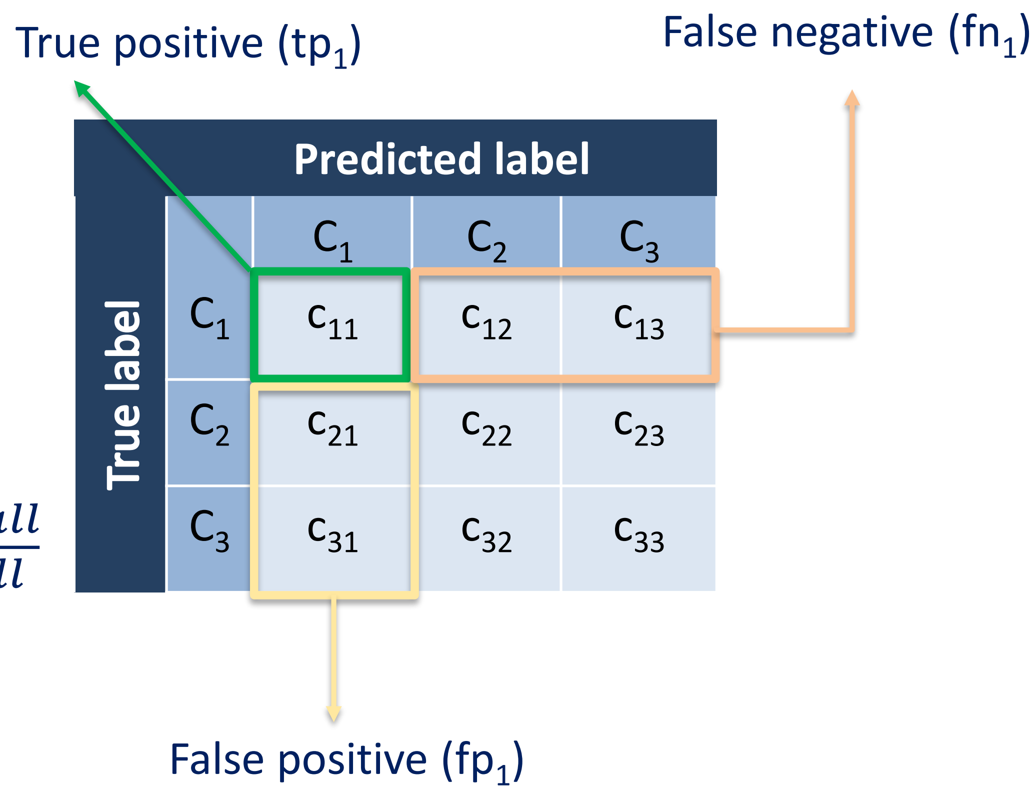
Performance metrics

$$\text{Precision} = \frac{tp_i}{tp_i + fp_i}$$

$$\text{Recall} = \frac{tp_i}{tp_i + fn_i}$$

$$\text{F1-score} = \frac{2\, Percision \times Recall}{Percision + Recall}$$

$$\text{Accuracy} = \frac{\sum_{i=1}^{n} tp_i}{n}$$

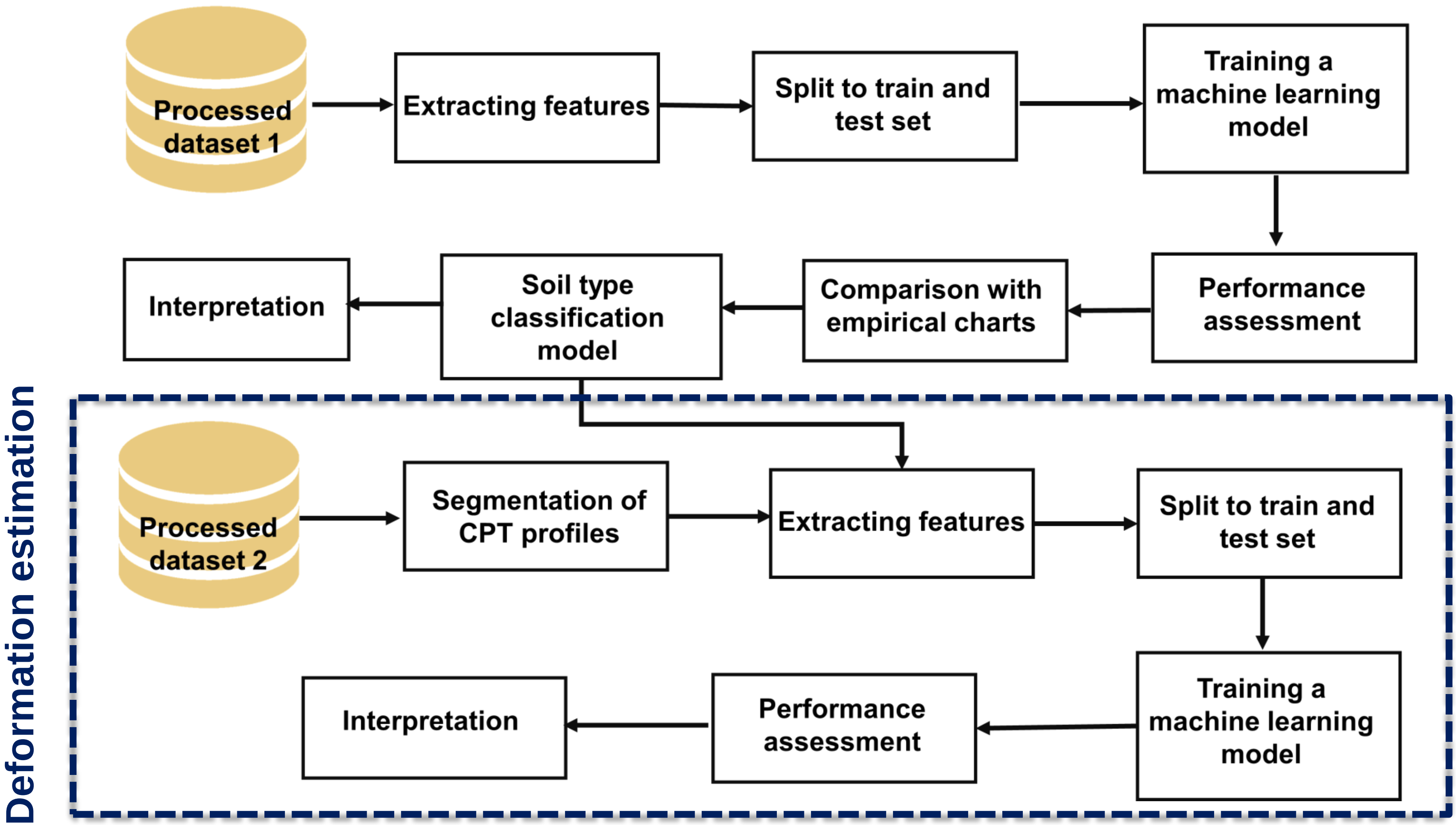


Performance metrics

		Predicted label			
True label		C_1	C_2	C_3	Σ
	C_1	c_{11}	c_{12}	c_{13}	c_{1+}
	C_2	c_{21}	c_{22}	c_{23}	c_{2+}
	C_3	c_{31}	c_{32}	c_{33}	c_{3+}
	Σ	c_{+1}	c_{+2}	c_{+3}	

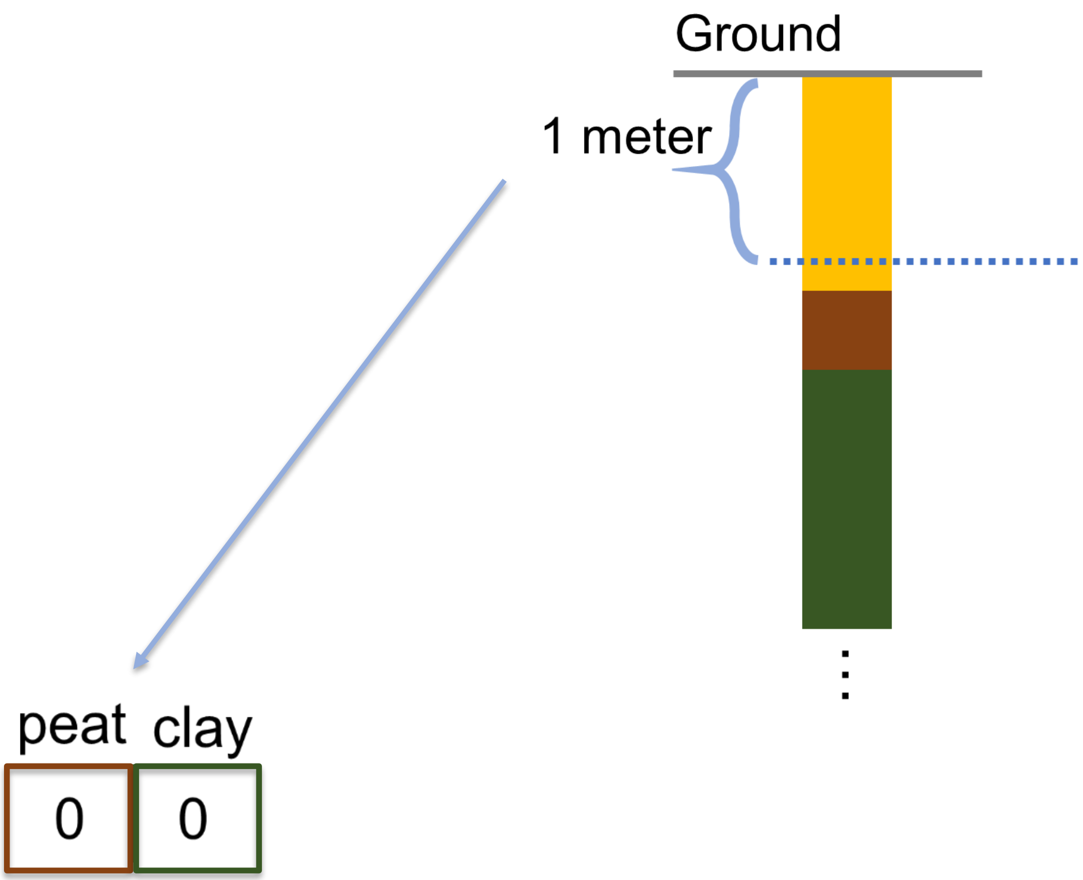
$$\text{Cohen Kappa} = \frac{N \sum_{i=1}^{n} c_{ii} - \sum_{i=1}^{n} c_{i+} c_{+i}}{N^2 - \sum_{i=1}^{n} c_{i+} c_{+i}}$$

Step 3: Modeling through machine learning

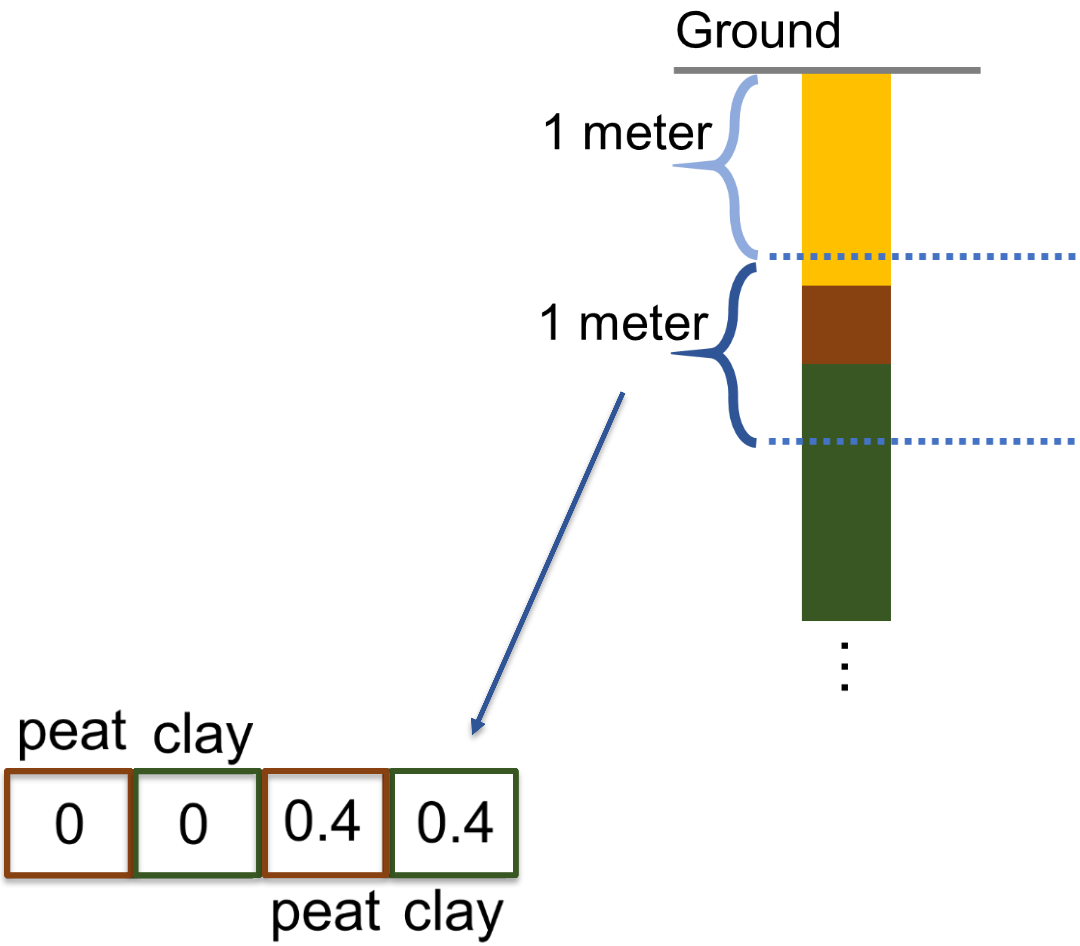


TU Delft Deltares Enabling Delta Life

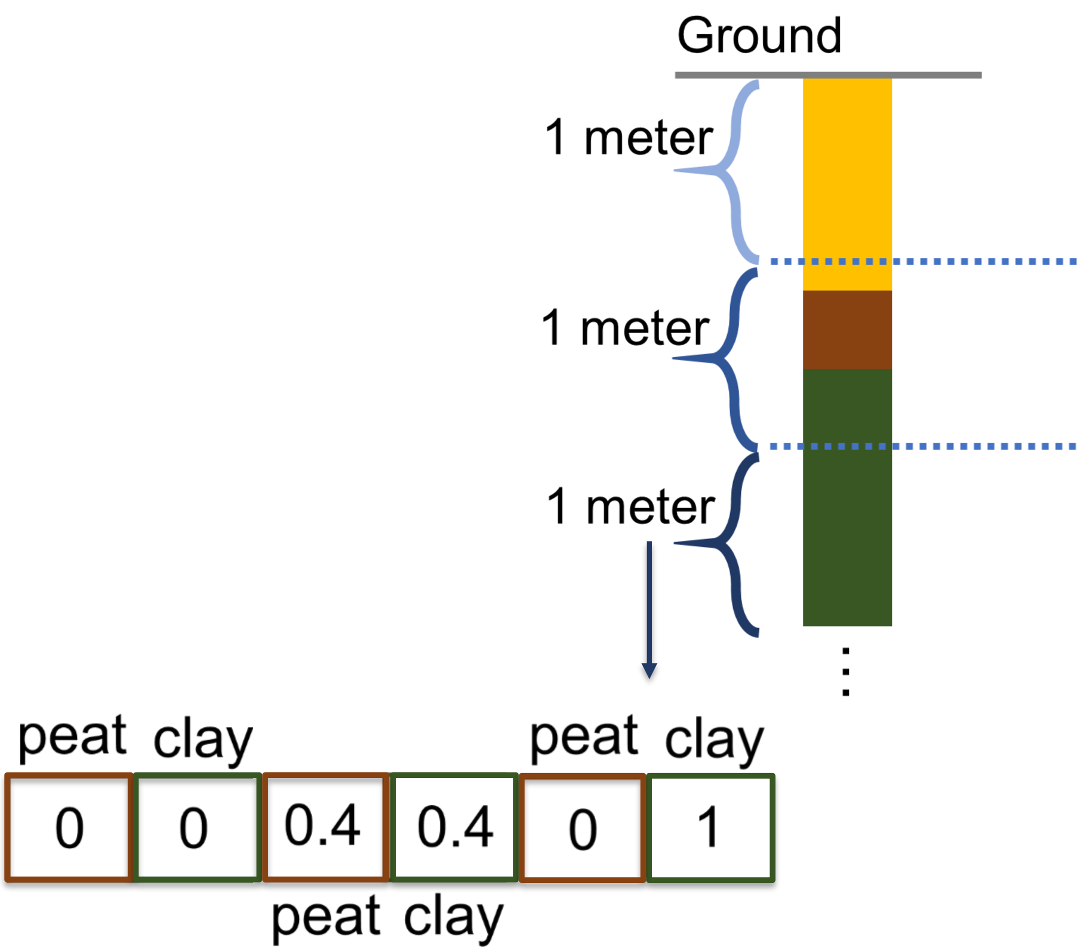
Feature extraction: qualitative descriptors



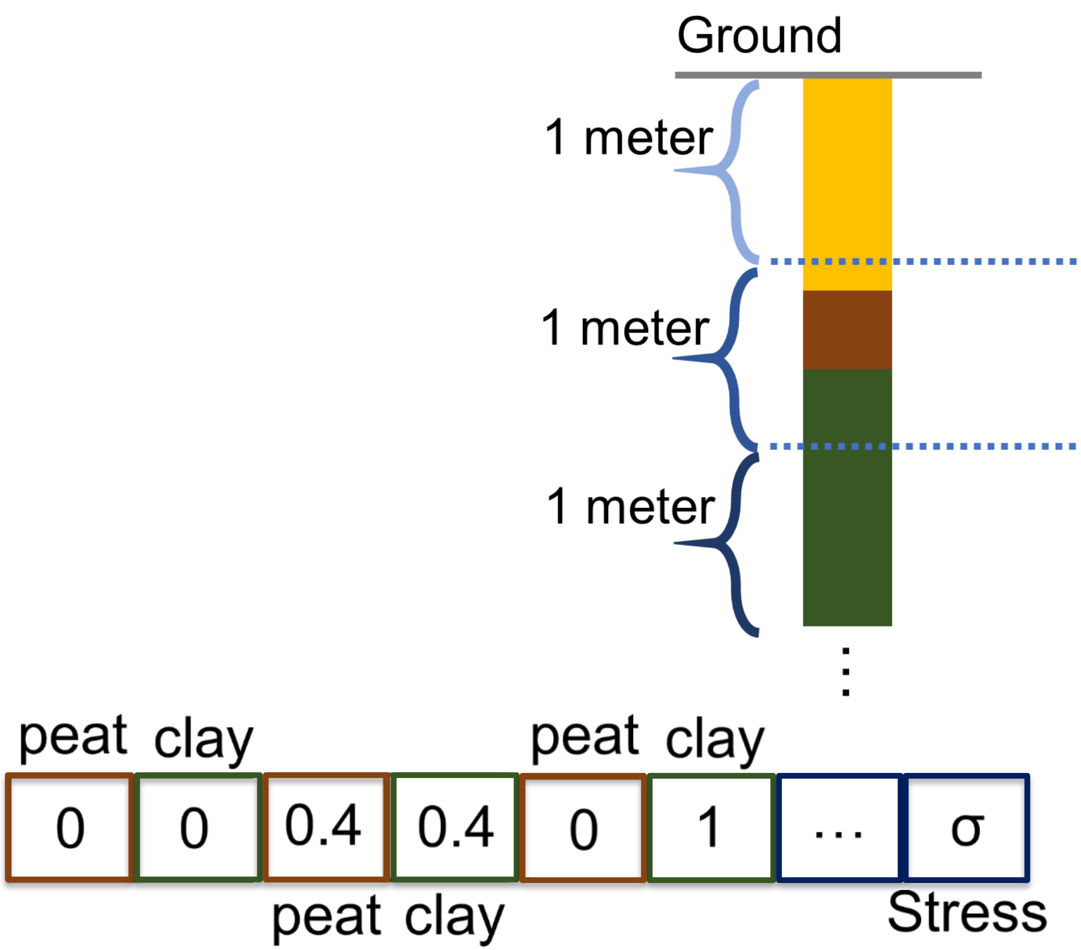
Feature extraction : qualitative descriptors



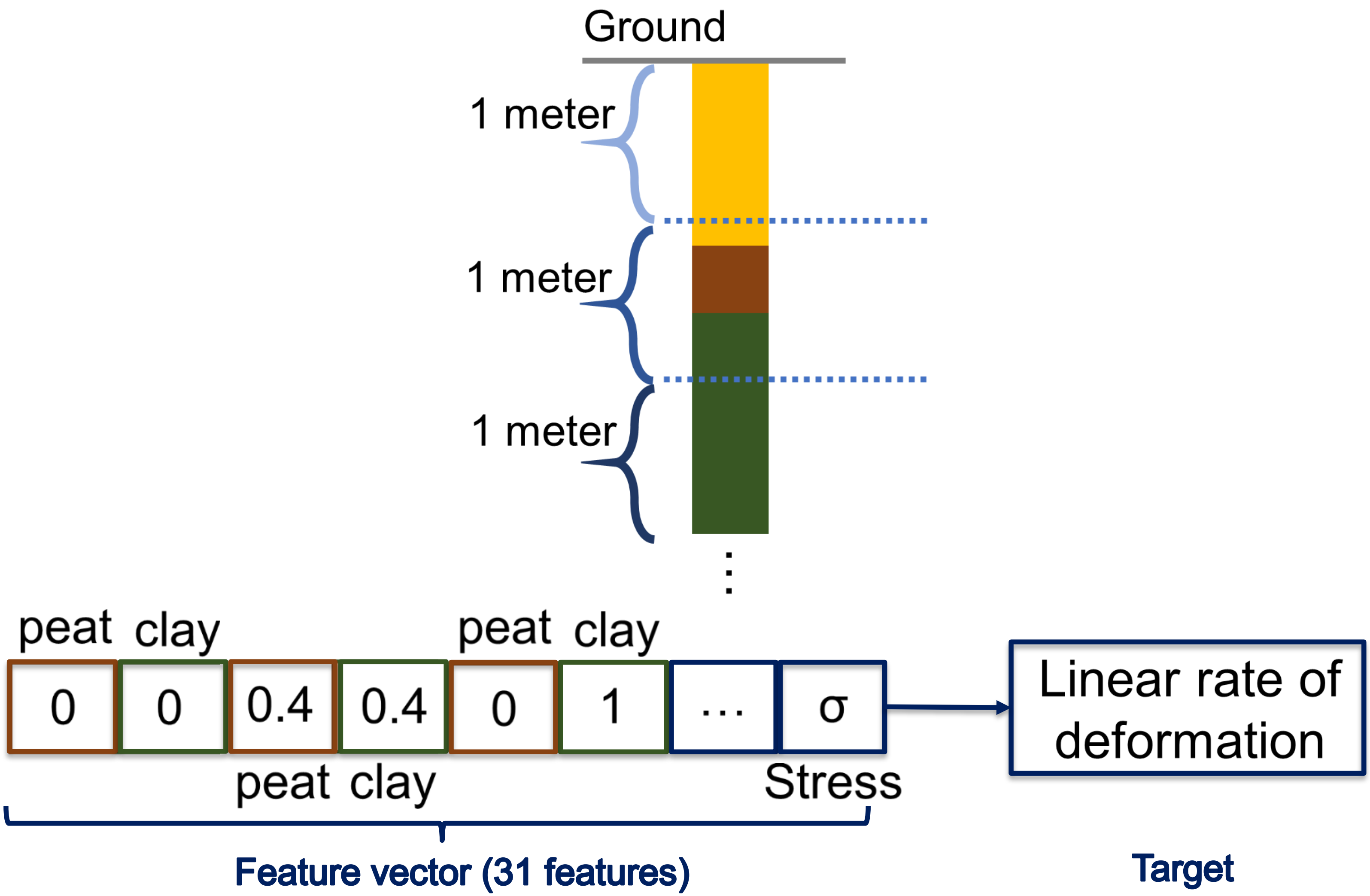
Feature extraction : qualitative descriptors



Feature extraction : qualitative descriptors



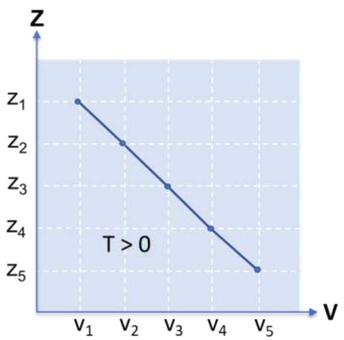
Feature extraction : qualitative descriptors

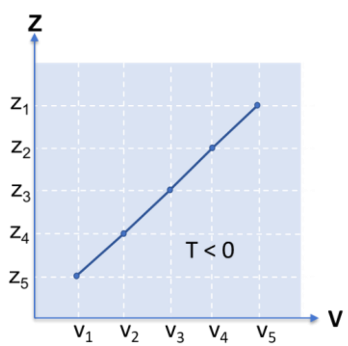


Feature extraction : quantitative descriptors

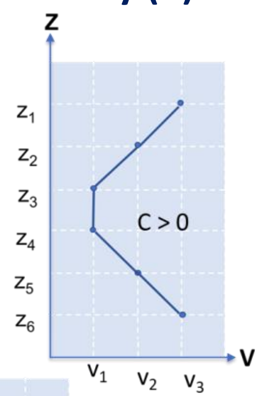
Median	Standard deviation (STD)	Skewness	Minimum	Maximum	Interquartile range (IQR)

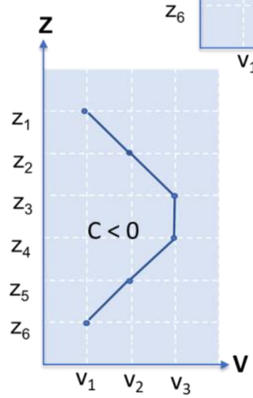
Indicator of simple trend (T)





Indicator of convexity or concavity (C)





Normalized number of fluctuations about the median (R)

$$\frac{\textbf{Number of runs}}{\textbf{Number of data points}}$$

Run = an uninterrupted series of values lower than median or an uninterrupted series of values higher than median

Sharpness of upper boundary of the segment (B)

$$\frac{\mathbf{ub - lb}}{\mathbf{ub + lb}}$$

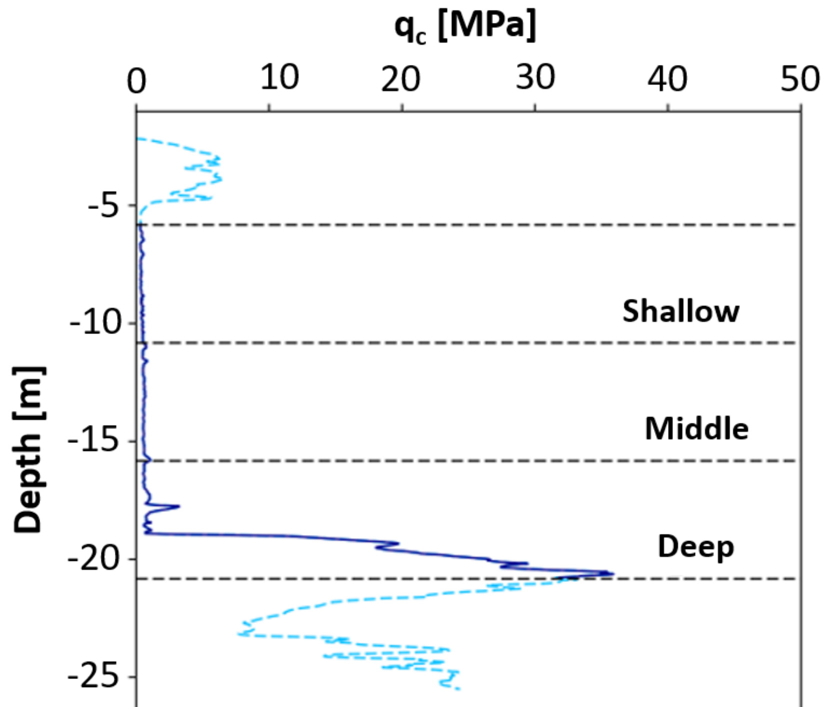
ub , lb are sum of two records, respectively, above and below the segment boundary

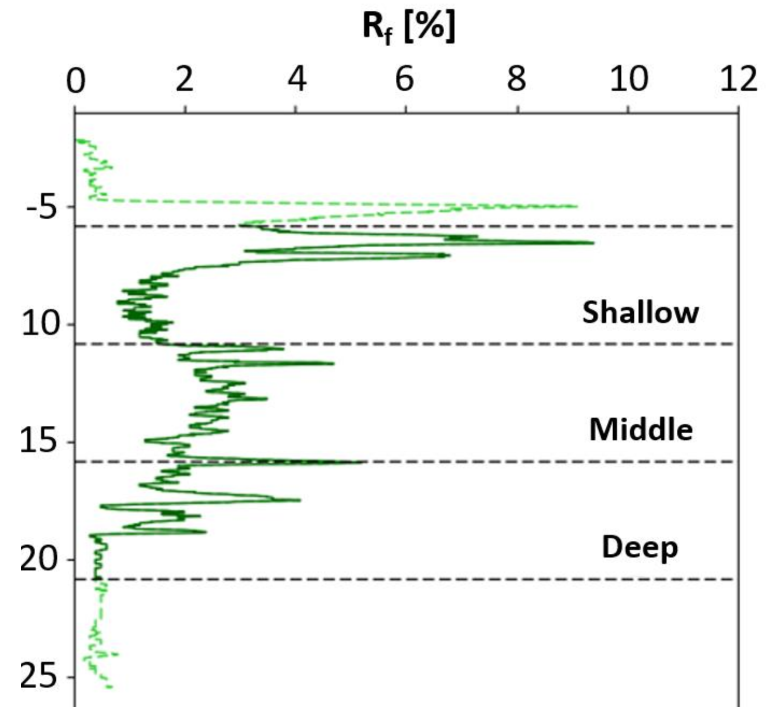
-1 < b < 0 : upward decrease
0 < b < 1 : upward increase

TUDelft Deltares Enabling Delta Life

Coerts [1996]

Feature extraction : quantitative descriptors



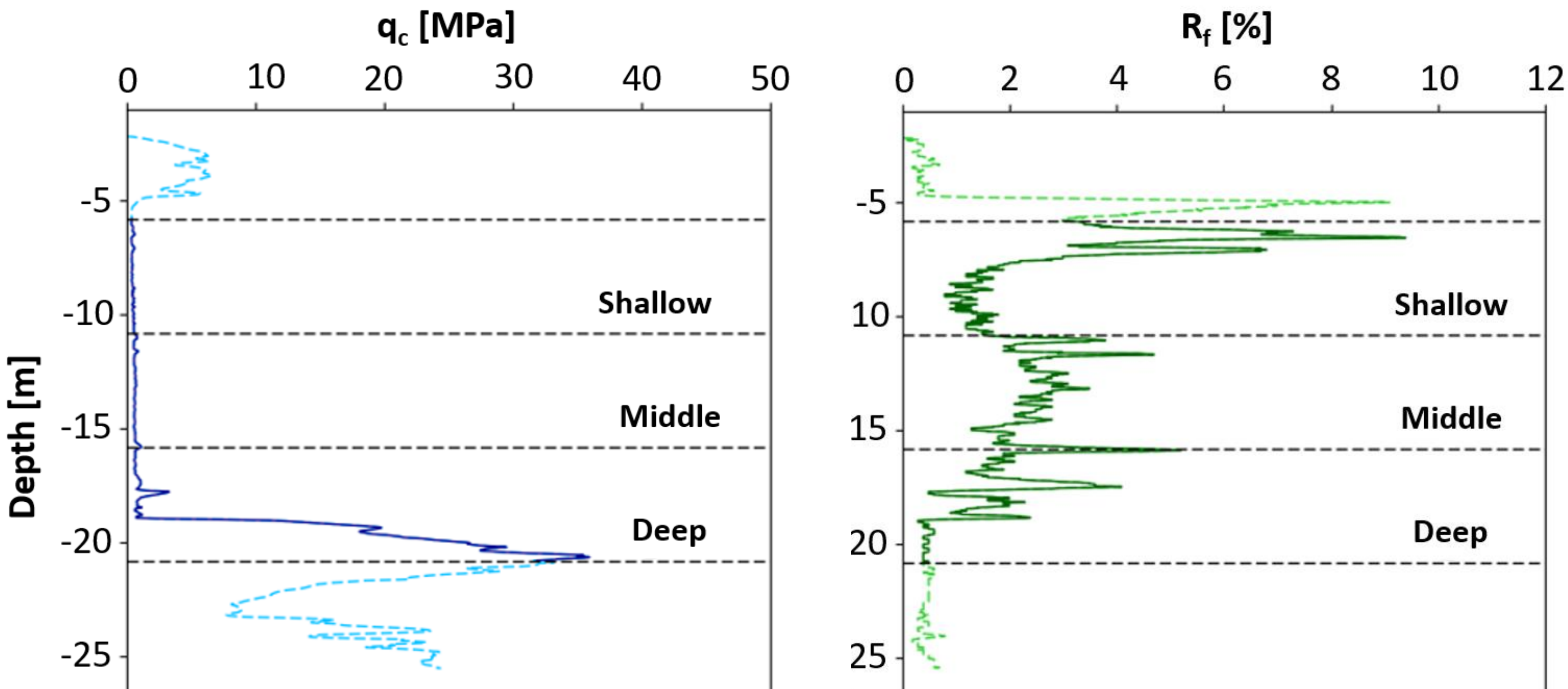


	Median	STD	Skew	Min	Max	IQR	T	C	R	B
Shallow	0.45	0.07	0.39	0.36	0.64	0.13	0.39	0.22	0.08	-

Median	STD	Skew	Min	Max	IQR	T	C	R	B
1.60	2.07	1.45	0.80	9.40	2.17	-0.56	0.20	0.07	-

q_c Shallow	R_f Shallow

Feature extraction : quantitative descriptors



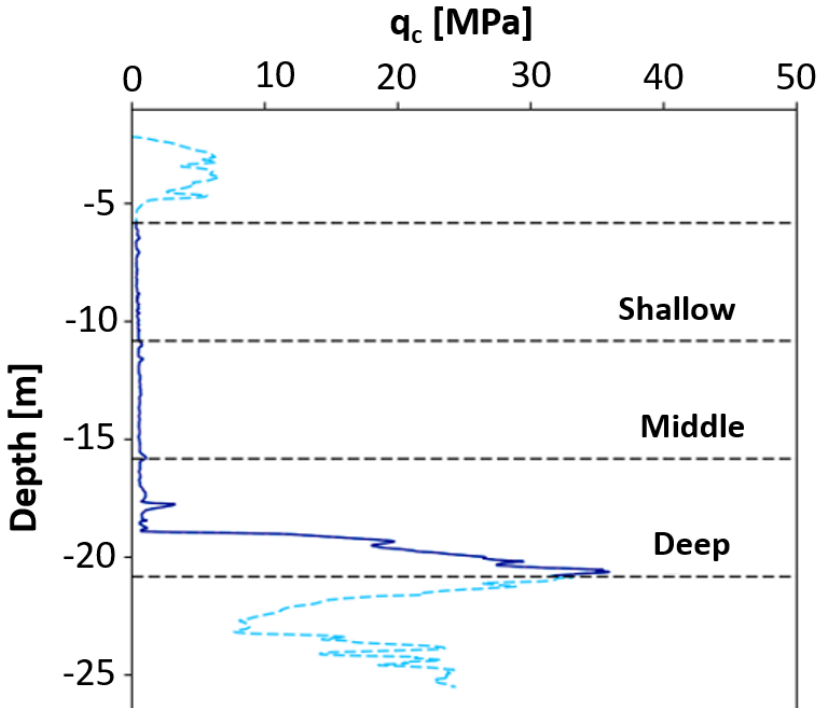
	Median	STD	Skew	Min	Max	IQR	T	C	R	B
Shallow	0.45	0.07	0.39	0.36	0.64	0.13	0.39	0.22	0.08	-
Middle	0.63	0.09	2.37	0.53	1.09	0.07	0.10	0.01	0.08	-0.01

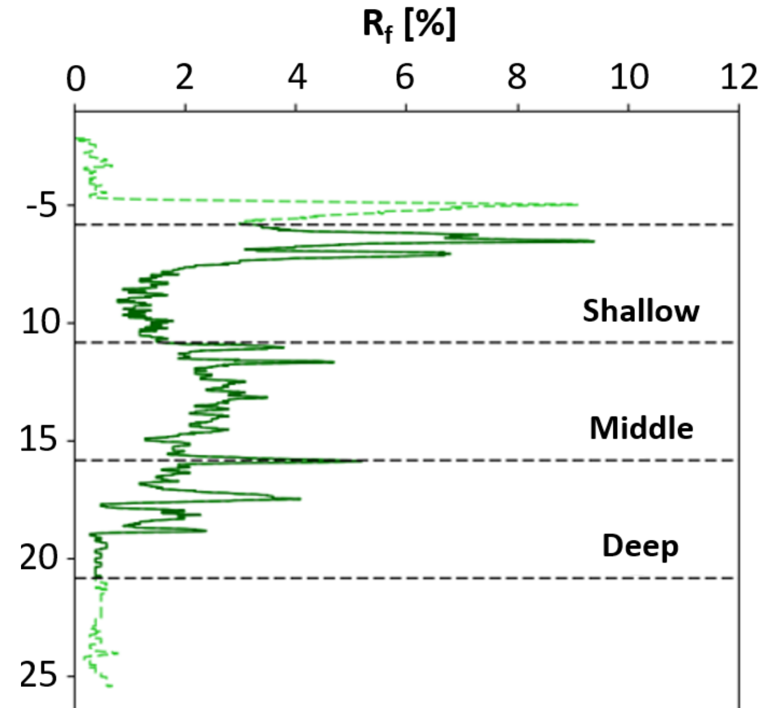
	Median	STD	Skew	Min	Max	IQR	T	C	R	B
Shallow	1.60	2.07	1.45	0.80	9.40	2.17	-0.56	0.20	0.07	-
Middle	2.40	0.58	0.96	1.30	4.70	0.70	-0.24	-0.27	0.08	-0.02

q_c Shallow | R_f Shallow

q_c Middle | R_f Middle

Feature extraction : quantitative descriptors





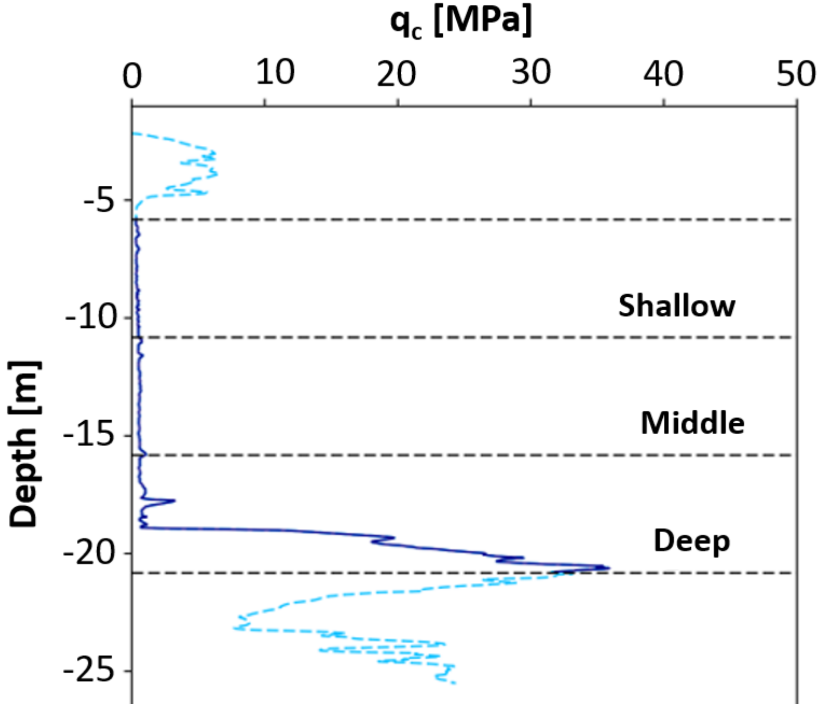
	Median	STD	Skew	Min	Max	IQR	T	C	R	B
Shallow	0.45	0.07	0.39	0.36	0.64	0.13	0.39	0.22	0.08	-
Middle	0.63	0.09	2.37	0.53	1.09	0.07	0.10	0.01	0.08	-0.01
Deep	1.07	11.84	0.93	0.57	35.93	18.57	0.71	0.11	0.05	-0.01

	Median	STD	Skew	Min	Max	IQR	T	C	R	B
Shallow	1.60	2.07	1.45	0.80	9.40	2.17	-0.56	0.20	0.07	-
Middle	2.40	0.58	0.96	1.30	4.70	0.70	-0.24	-0.27	0.08	-0.02
Deep	1.30	1.06	1.27	0.30	5.20	1.40	-0.56	-0.12	0.03	-0.08

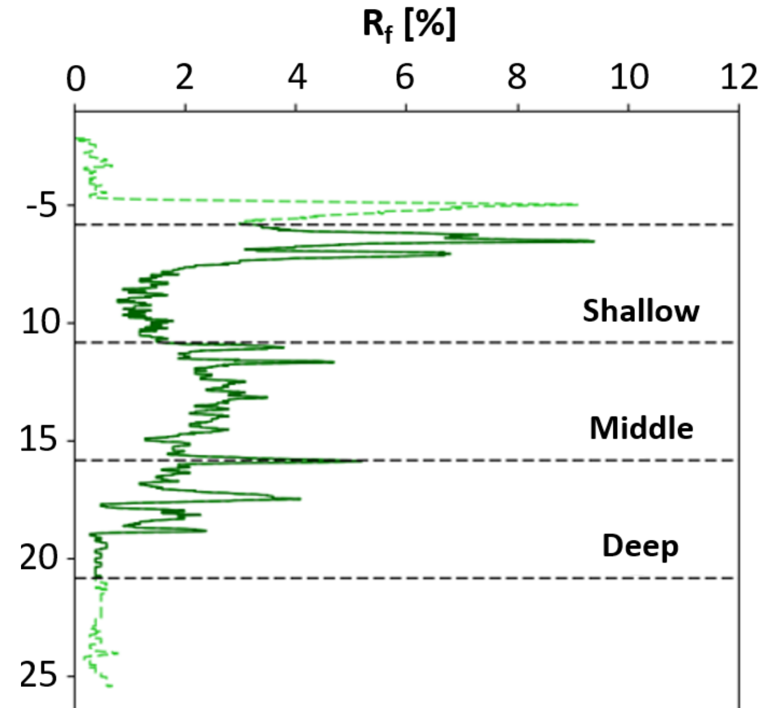
q_c Shallow	R_f Shallow	q_c Middle	R_f Middle	q_c Deep	R_f Deep

TU Delft Deltares Enabling Delta Life

Feature extraction : quantitative descriptors



	Median	STD	Skew	Min	Max	IQR	T	C	R	B
Shallow	0.45	0.07	0.39	0.36	0.64	0.13	0.39	0.22	0.08	-
Middle	0.63	0.09	2.37	0.53	1.09	0.07	0.10	0.01	0.08	-0.01
Deep	1.07	11.84	0.93	0.57	35.93	18.57	0.71	0.11	0.05	-0.01



Median	STD	Skew	Min	Max	IQR	T	C	R	B
1.60	2.07	1.45	0.80	9.40	2.17	-0.56	0.20	0.07	-
2.40	0.58	0.96	1.30	4.70	0.70	-0.24	-0.27	0.08	-0.02
1.30	1.06	1.27	0.30	5.20	1.40	-0.56	-0.12	0.03	-0.08

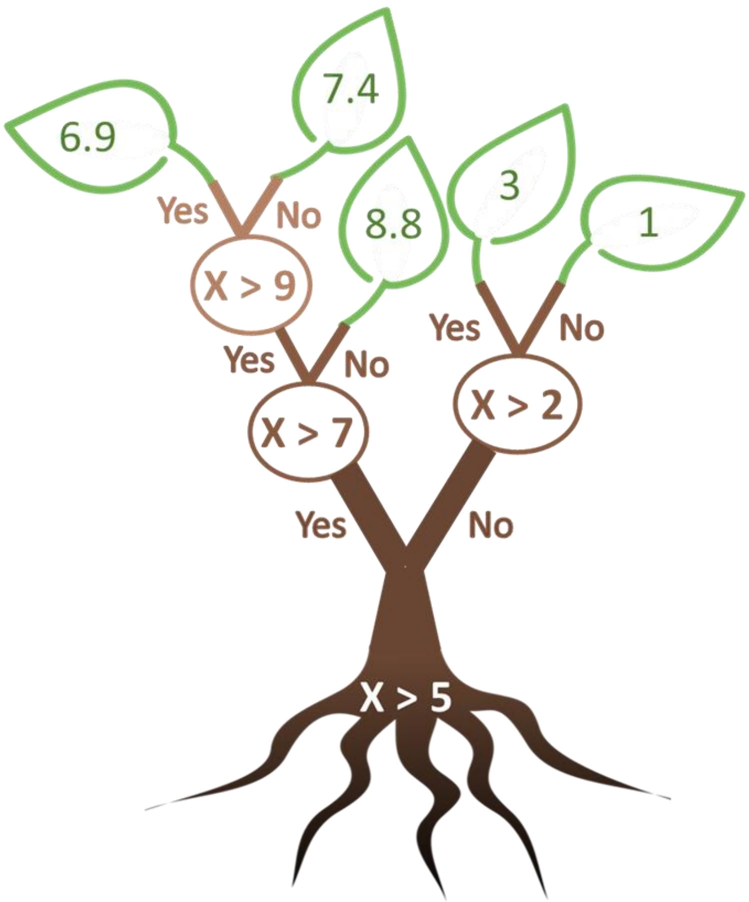
q_c Shallow	R_f Shallow	q_c Middle	R_f Middle	q_c Deep	R_f Deep	σ

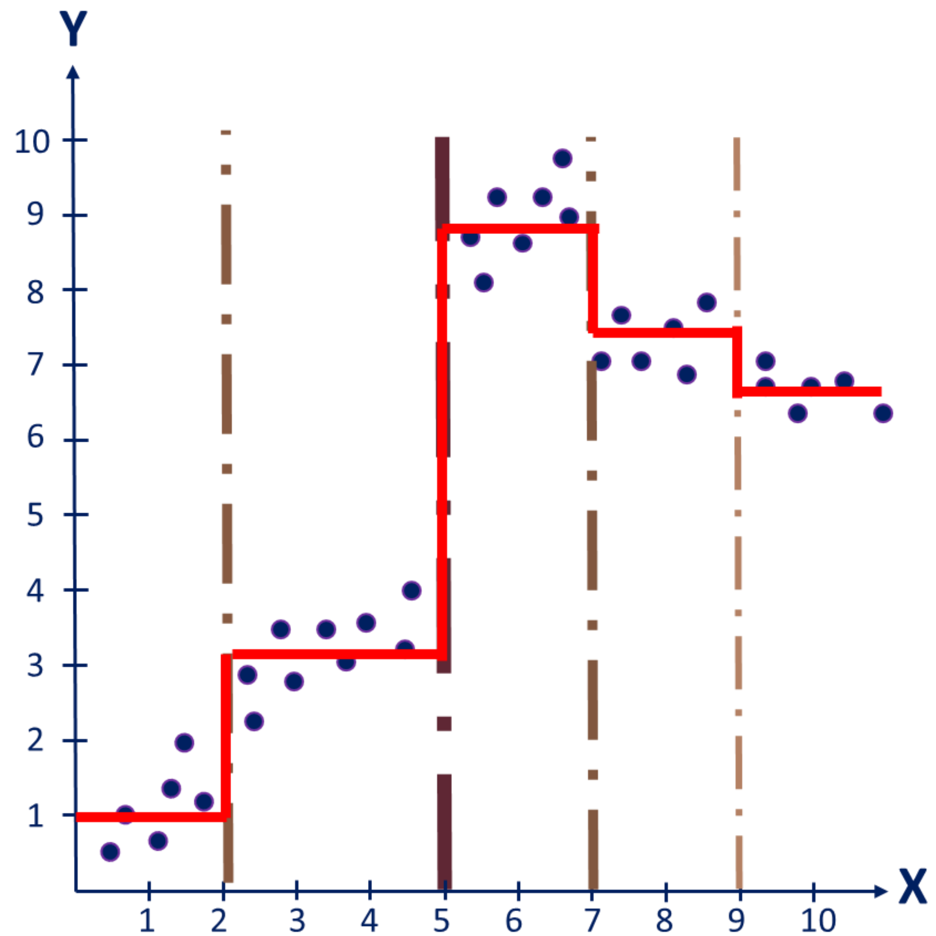
$\underbrace{\text{Feature vector}}_{\text{Feature vector (59 features)}} \rightarrow$ Linear rate of deformation

Feature vector (59 features)

Target

Supervised Learning: Decision tree

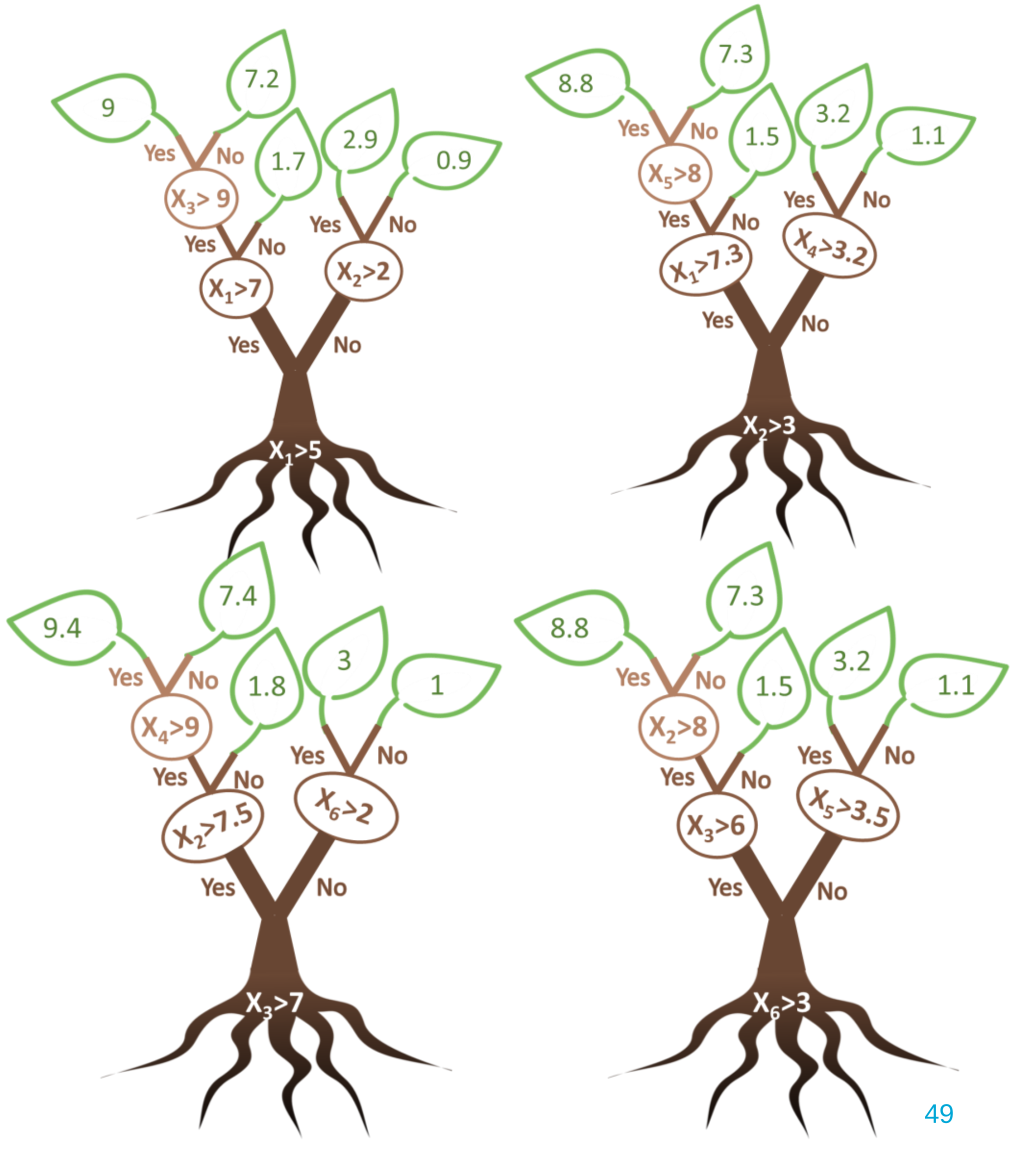




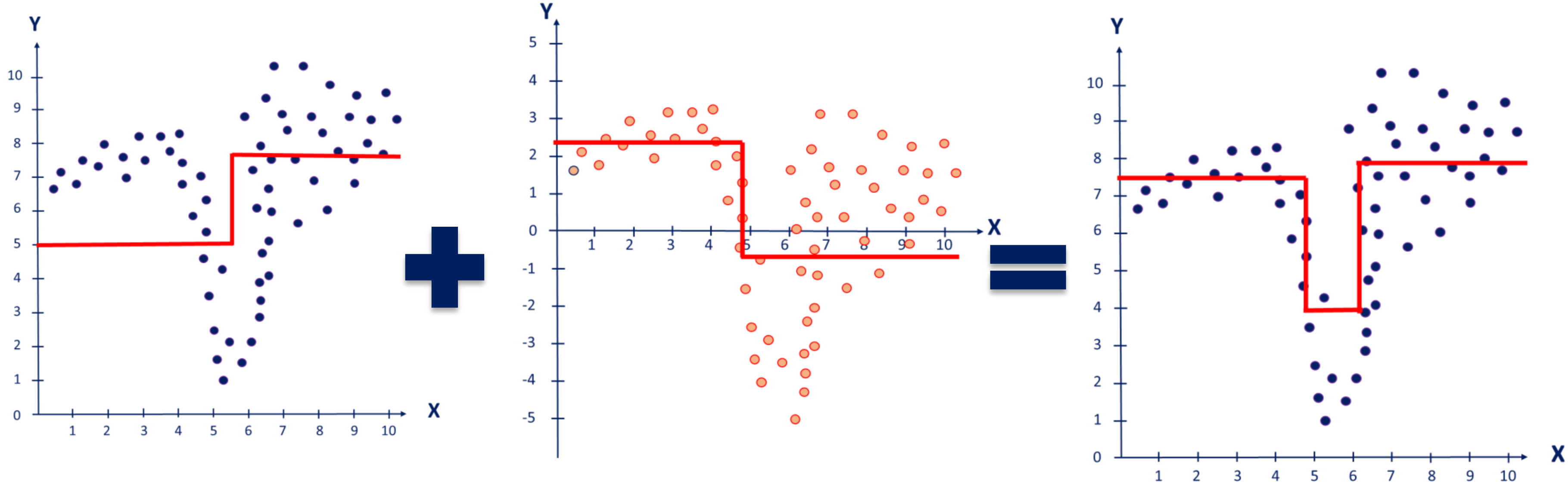
Supervised Learning: Random forest

- Constructing multiple decision tress from subsamples during training phase

- Final decision is the unweighted average of the decision by each tree



Supervised Learning: Gradient boosting



Fitting a simple model **data points**

Fitting a model to the **error residuals**

Combining the two models for creating a more complicated model

Performance metrics

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \mu_y)^2}$$

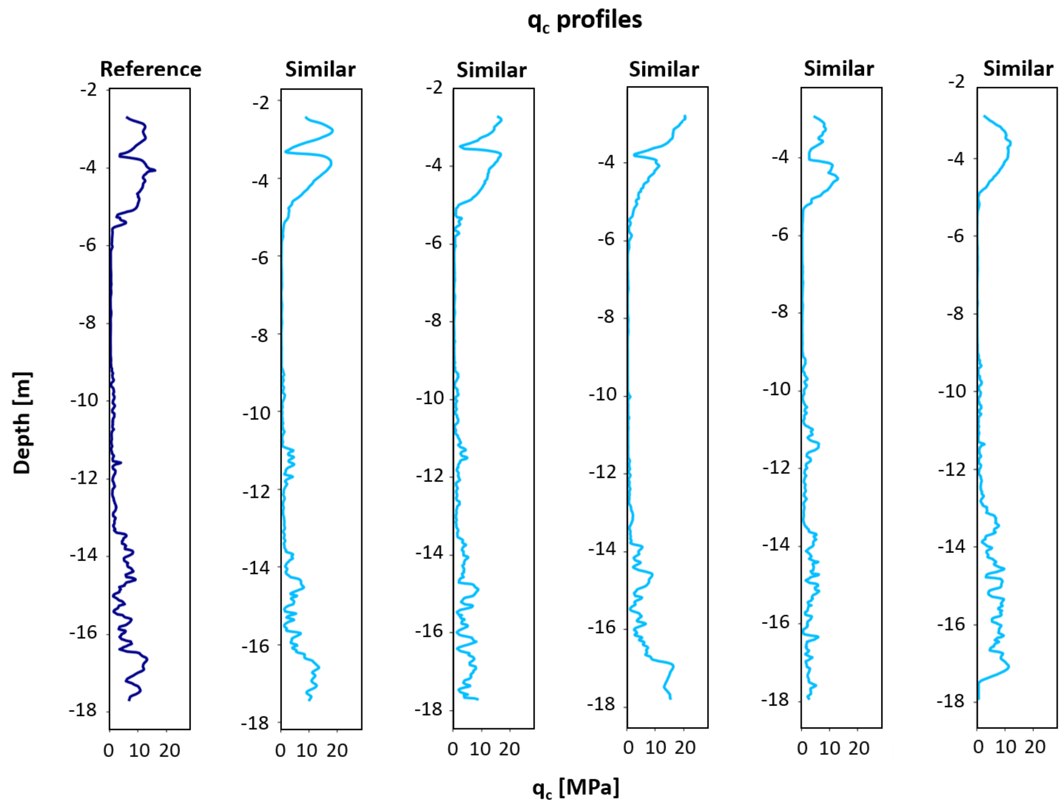
TU Delft Deltares Enabling Delta Life

Outline

- Introduction and related work
- Methodology and implementation
- Results and discussions
- Conclusion and future works

Results: correlations

q_c profiles (MPa)



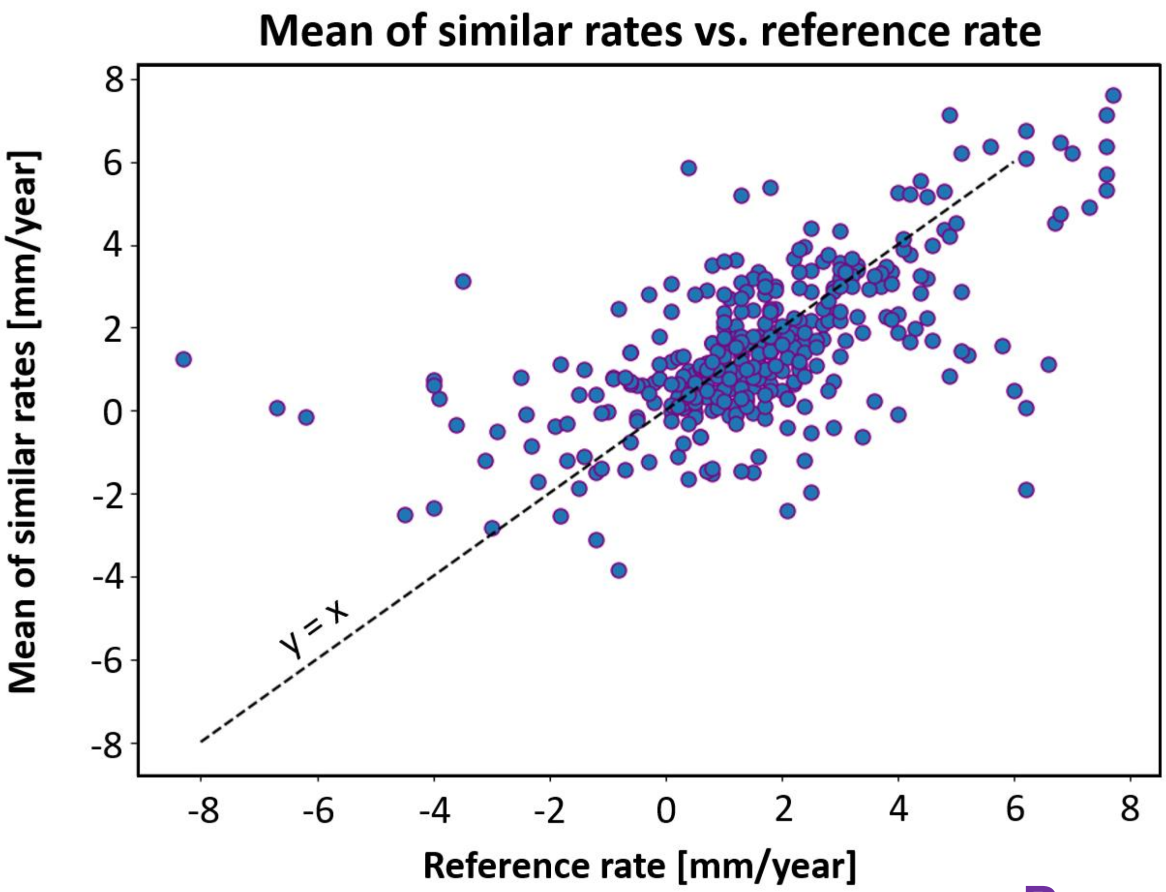
Loading/unloading stress (kPa)	- 8.7	- 8.8	- 9.4	- 16.5	- 8.2	- 8.1
Linear rate of deformation (mm/year)	- 4	- 4.5	- 3	- 0.8	- 2.2	- 1.2

Mean = - 2.3 mm/year

TU Delft

Deltares
Enabling Delta Life

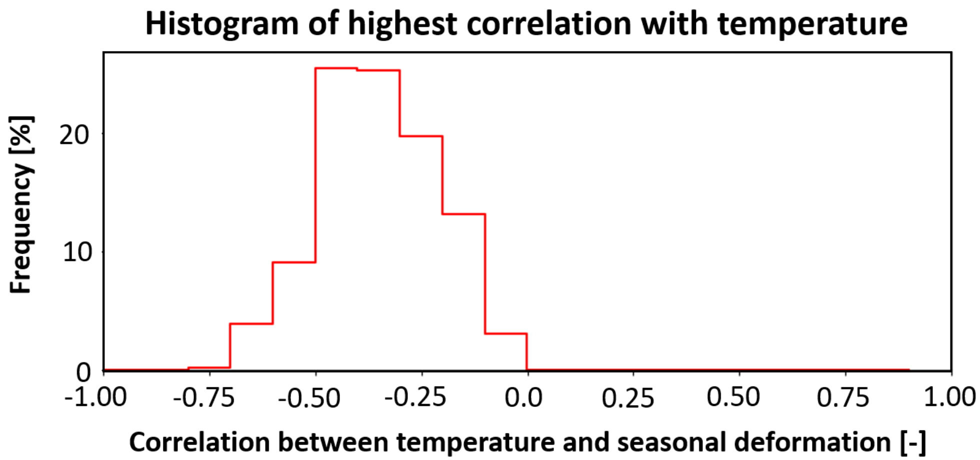
Results: correlations

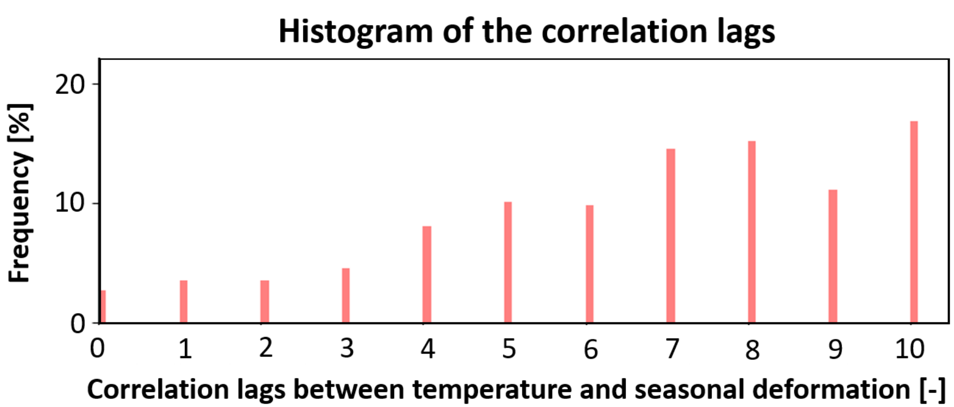


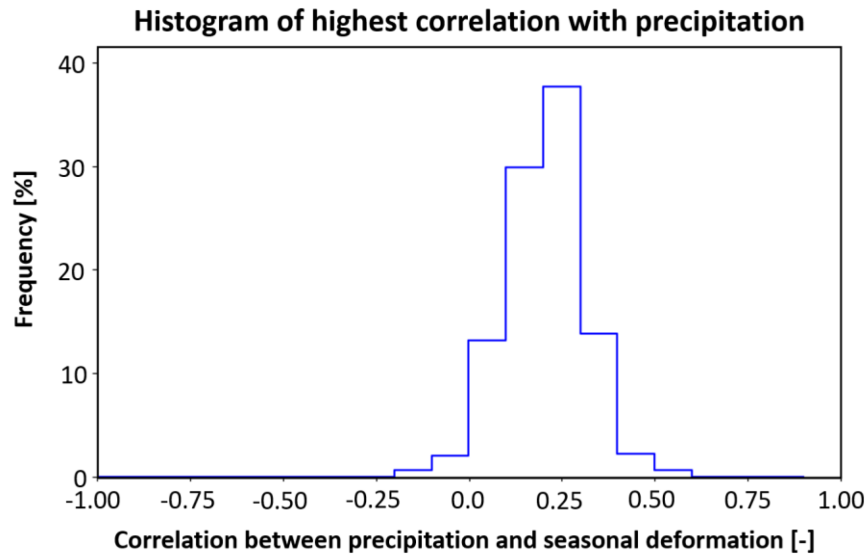
Pearson Correlation = 0.6

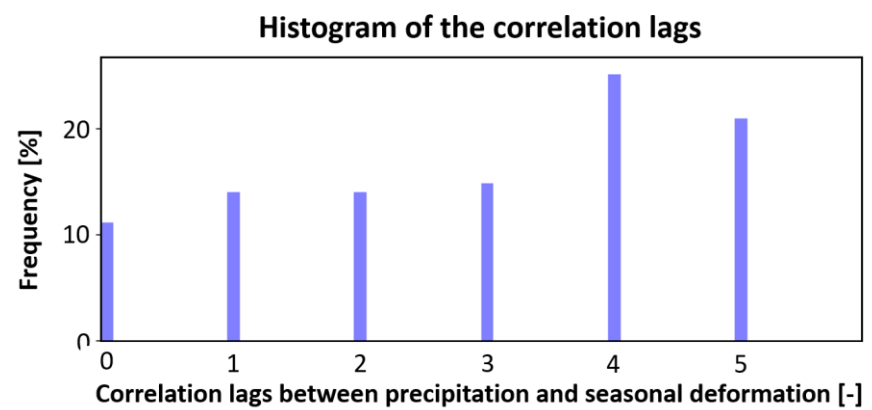
Coefficient of Determination = 0.4

Results: correlations

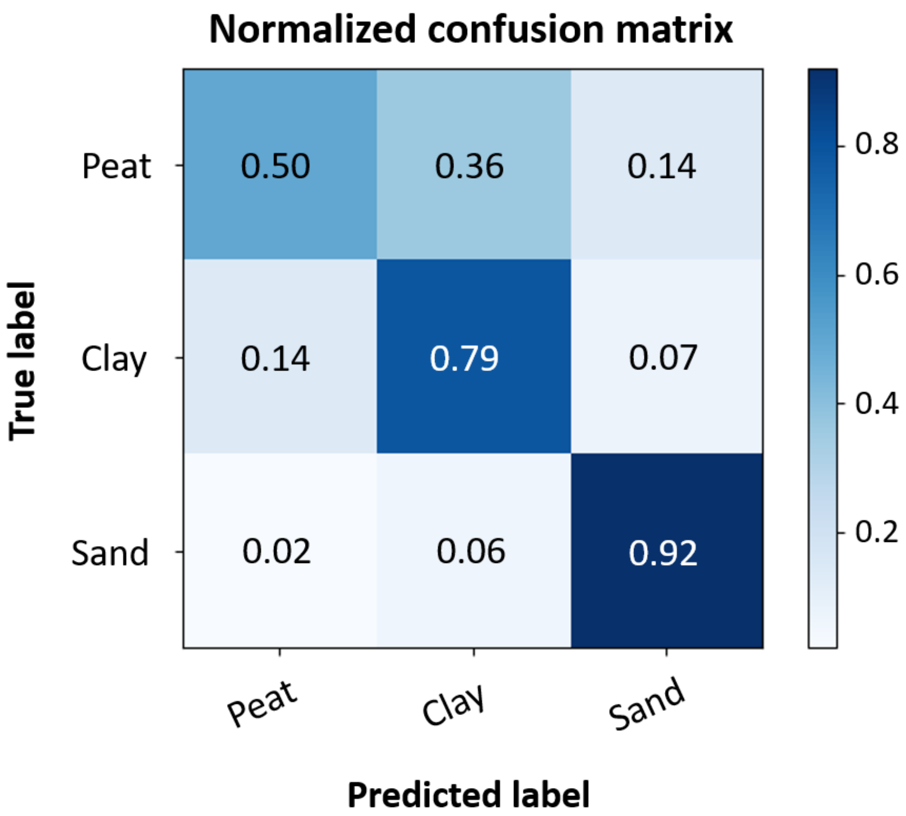




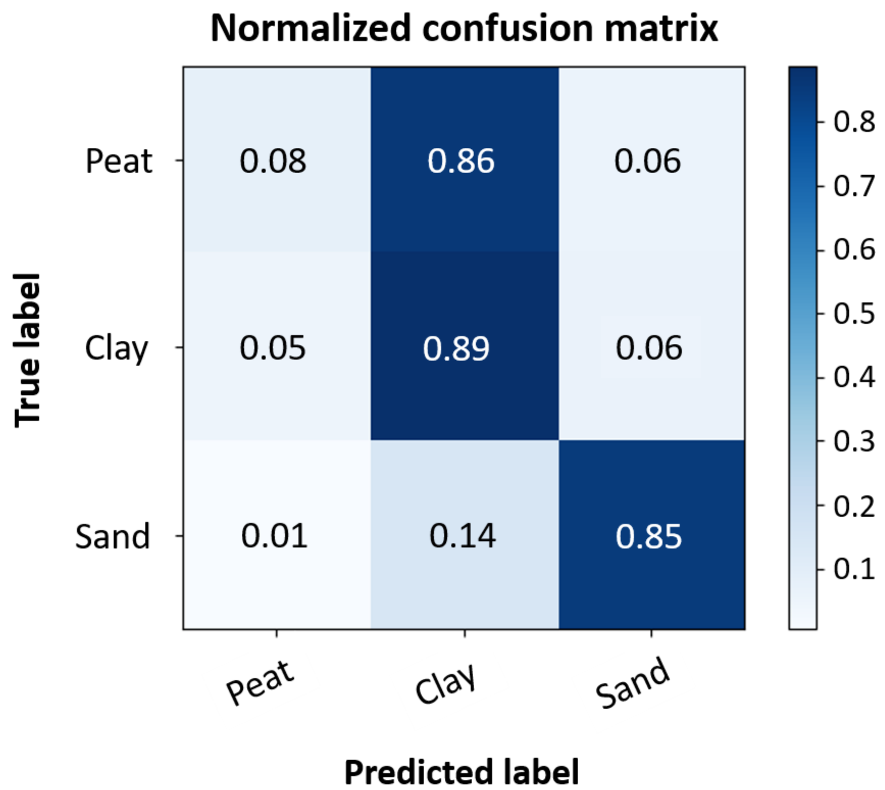




Results: soil classification



SVM



Robertson

Results: soil classification

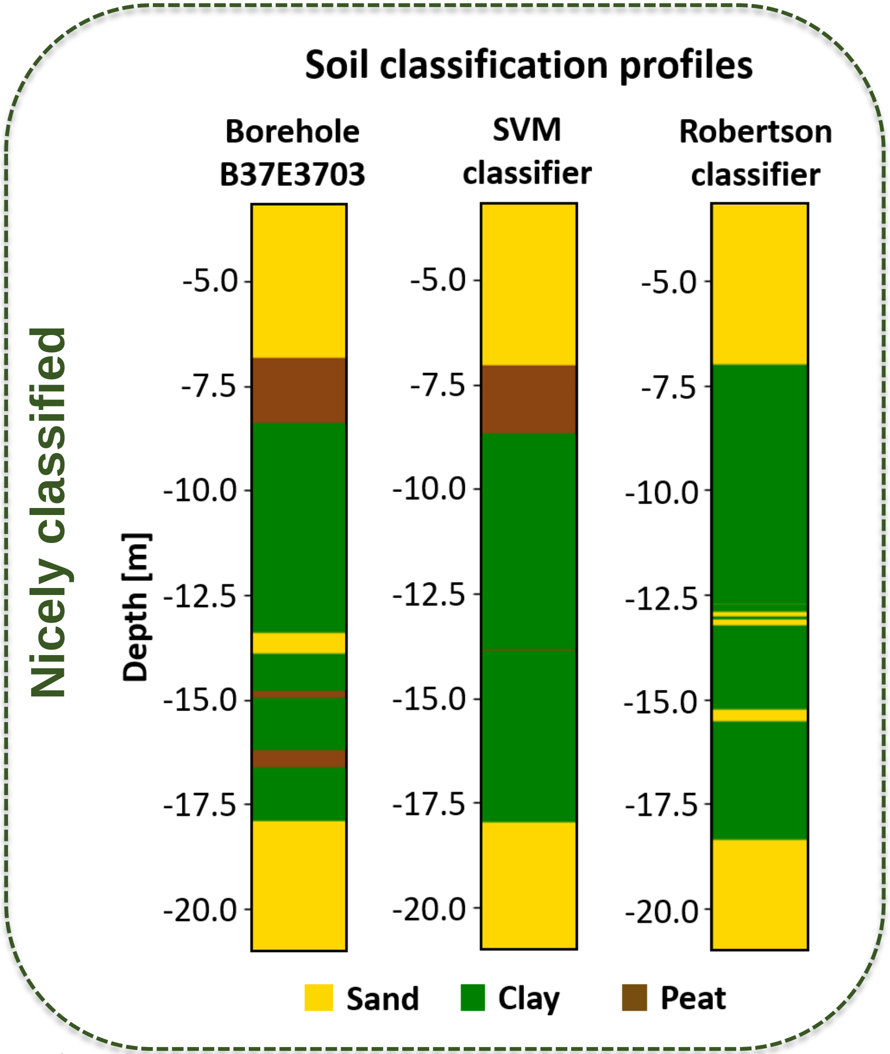
SVM

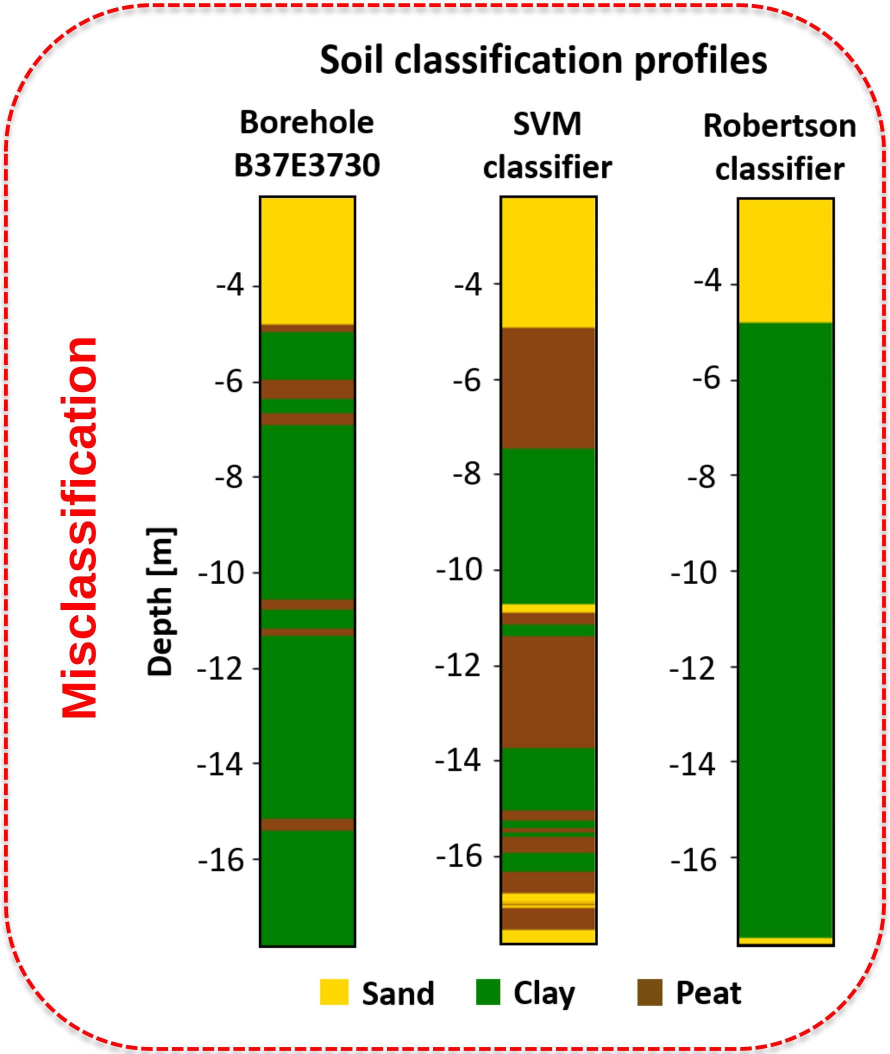
	Precision	Recall	F1-Score	
Peat	0.35	0.5	0.41	
Clay	0.87	0.79	0.83	
Sand	0.91	0.92	0.91	
Micro Average	0.83	0.83	0.83	**Overall accuracy: 0.83**
Macro average	0.85	0.74	0.72	**Kappa : 0.71**

Robertson

	Precision	Recall	F1-Score	
Peat	0.2	0.08	0.11	
Clay	0.75	0.89	0.81	
Sand	0.92	0.85	0.88	
Micro Average	0.81	0.81	0.81	**Overall accuracy: 0.81**
Macro average	0.62	0.61	0.60	**Kappa : 0.65**

Results: soil classification

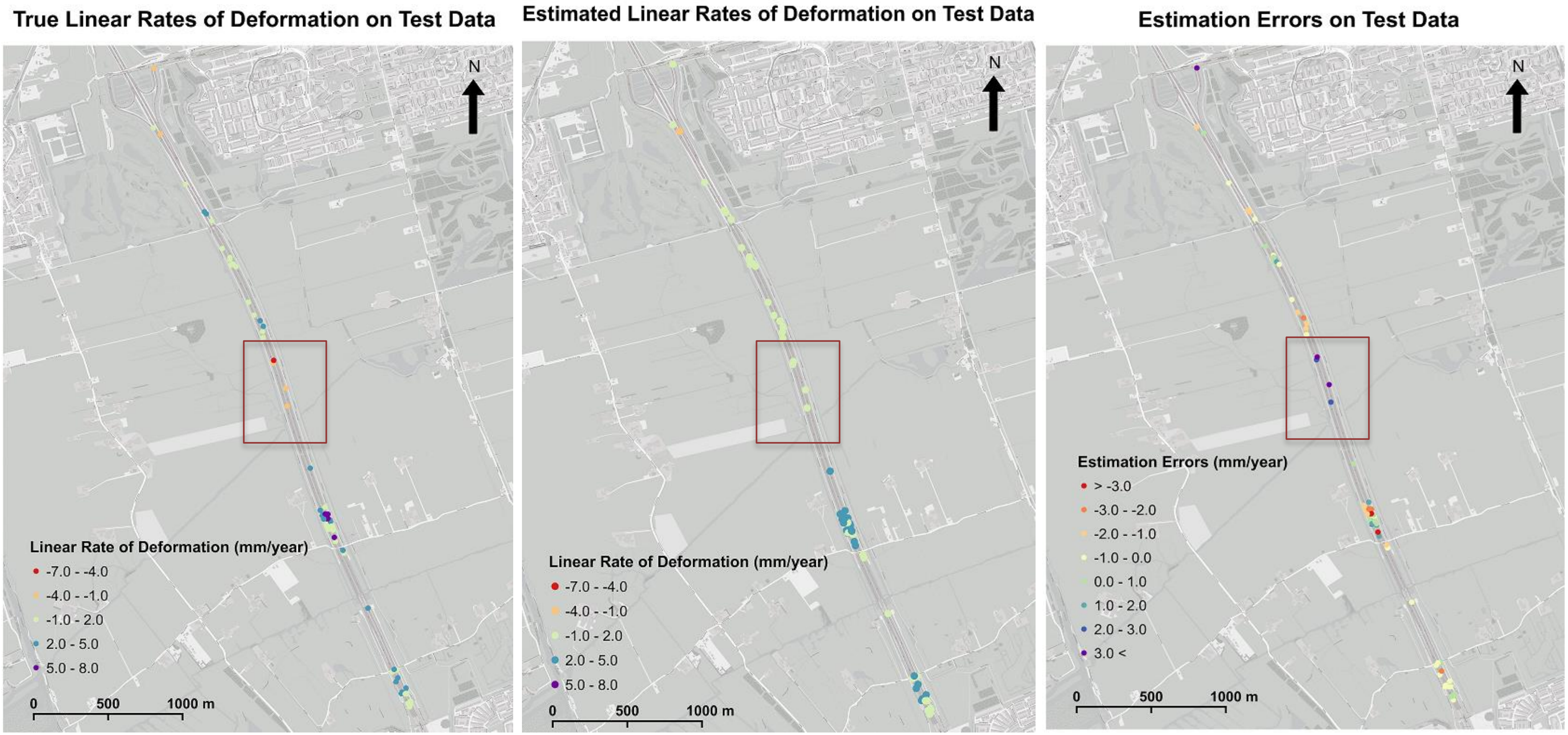




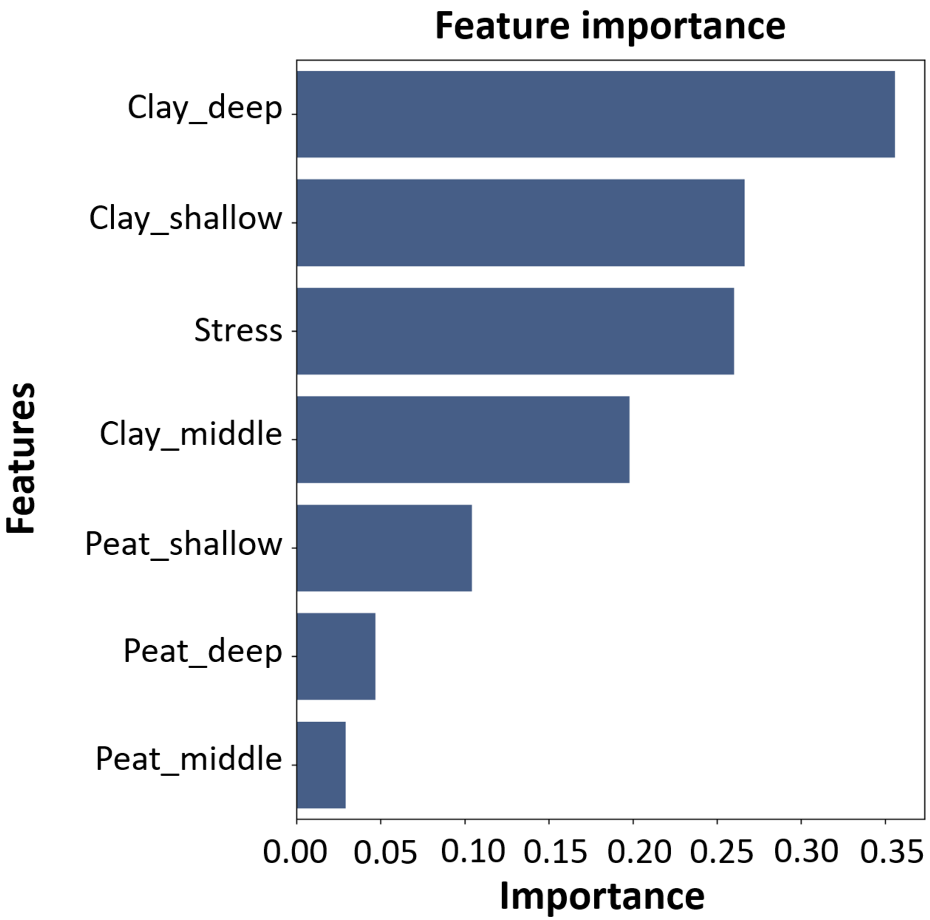
Results: deformation estimation

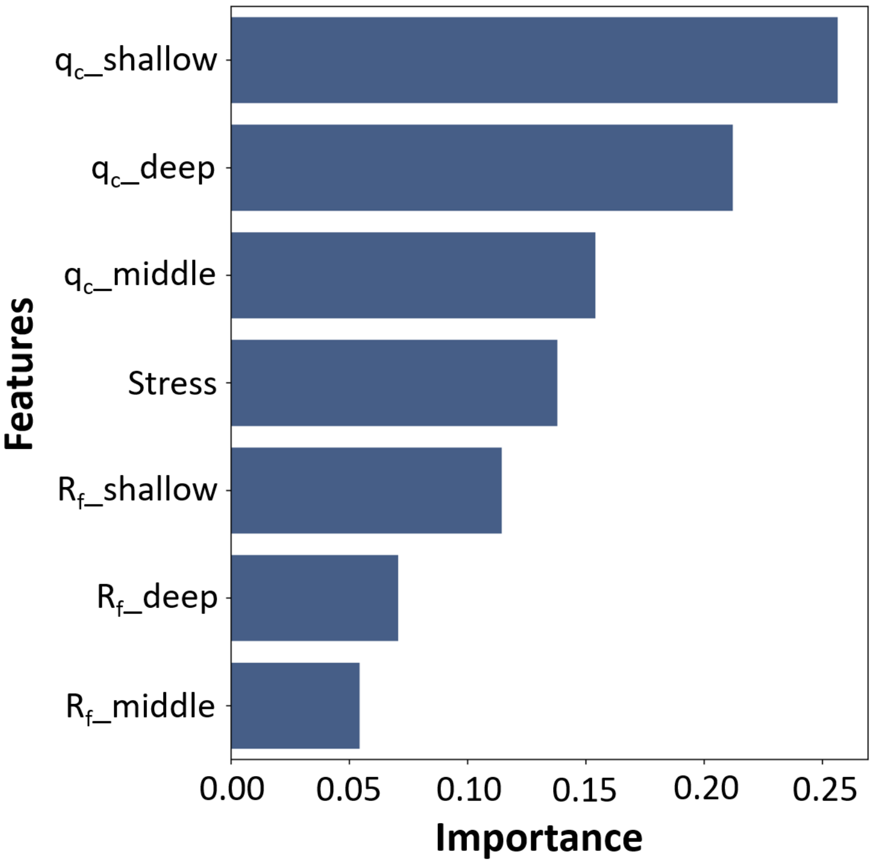
Features	Qualitative descriptors		Quantitative descriptors	
Algorithm	Gradient boosting	Random forest	Gradient boosting	Random forest
MAE (mm/year)	1.3	1.3	1.1	1.2
RMSE (mm/year)	1.8	1.8	1.6	1.6
R^2 (-)	0.3	0.3	0.5	0.4

Results: deformation estimation



Results: deformation estimation





TU Delft
Deltares
Enabling Delta Life

Outline

- Introduction and related work
- Methodology and implementation
- Results and discussions
- Conclusion and future works

Conclusions

- What are the data sources needed for studying soil properties, loading and deformation measurements?
  - Soil properties: CPT, borehole, temperature and precipitation as indicators of soil moisture
  - Loading/unloading: Elevation before and after construction
  - Deformation: D-InSAR time series

Conclusions

- Is there a correlation between soil properties, loading/unloading and deformation measurements?
  - Moderate correlation between soil properties and loading/loading and deformation
  - No meaningful correlation between temperature and precipitation and the seasonal deformation

Conclusions

- What parameters/features should be included?
  - Features for soil classification include Q_{tn}, F_r, total stress, the average q_c and f_s of 1 meter above and below of the measurement point
  - Features for deformation estimation
    - Qualitative descriptors (thickness of clay and peat) : more interpretable and less accurate
    - Quantitative descriptors (median, STD, skewness, Min, Max, IQR, T, C, R, B): more accurate and less interpretable

Conclusions

- What machine learning algorithm(s) are more suitable in establishing the relationship?
  - SVM for soil classification: low dimensionality of feature space
  - Tree-based algorithms for deformation estimation: high dimensional feature space and more interpretable

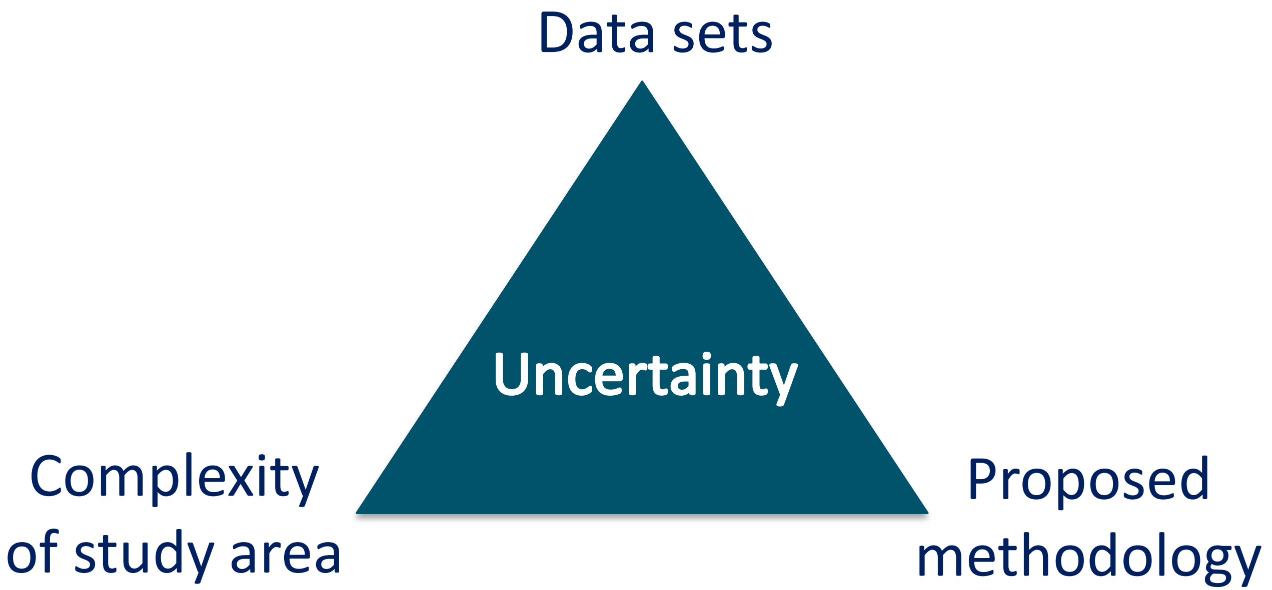
Conclusions

- What is the accuracy of the chosen machine learning technique and is it satisfactory?
  - For soil classification:
    - The performance is better compared to empirical charts
  - For deformation estimation:
    - The accuracy is not high.
    - It can be seen that the uncertainty of the model may not be desirable.
    - Improvements are needed for applications in which high accuracy is required.

Conclusions

*Using machine learning techniques, is it possible to model spatial-temporal relationship between the **deformation measurements,** the **soil properties,** and **loading/unloading** on roads?*

- It is possible to develop a fully data-driven model
- However, the accuracy is moderate due to sources of uncertainty:



Future work

- Applying the methodology to other study areas and land uses (with less complexity)
- Soil classification based on data driven approaches can be explored on the country scale
- Including more features and/or exploring other feature extraction methods

Thank You!