Application of Data Mining Techniques to Soil Stabilization

Markus Nikolai Berner

Agenda

- Målsetninger for masteroppgaven
- Hva er Data Mining?
- Presentasjon av prosjektet E6 Kvithammar-Åsen
- CPTU korrelasjonsanalyse
- Lab korrelasjonsanalyse
- Norsk database
- Svensk database
- Sammenligning
- Konklusjon

Målsetninger for masteroppgaven

1. Study data mining techniques applied to geotechnical engineering.
2. Acquire and systemize real-life DDM data to obtain a DDM database.
3. Apply relevant data mining techniques to the derived DDM database.
4. Test if reliable correlations in the DDM database can be detected.
5. Translate the obtained findings into practical guidance informing future DDM projects.

Hva er Data Mining?

- Data Mining er en prosess der man sorterer store datasett for å identifisere mønstre og sammenhenger som kan bidra til å finne ny kunnskap.

E6 Kvithammar-Åser

Informasjon om prosjektet:

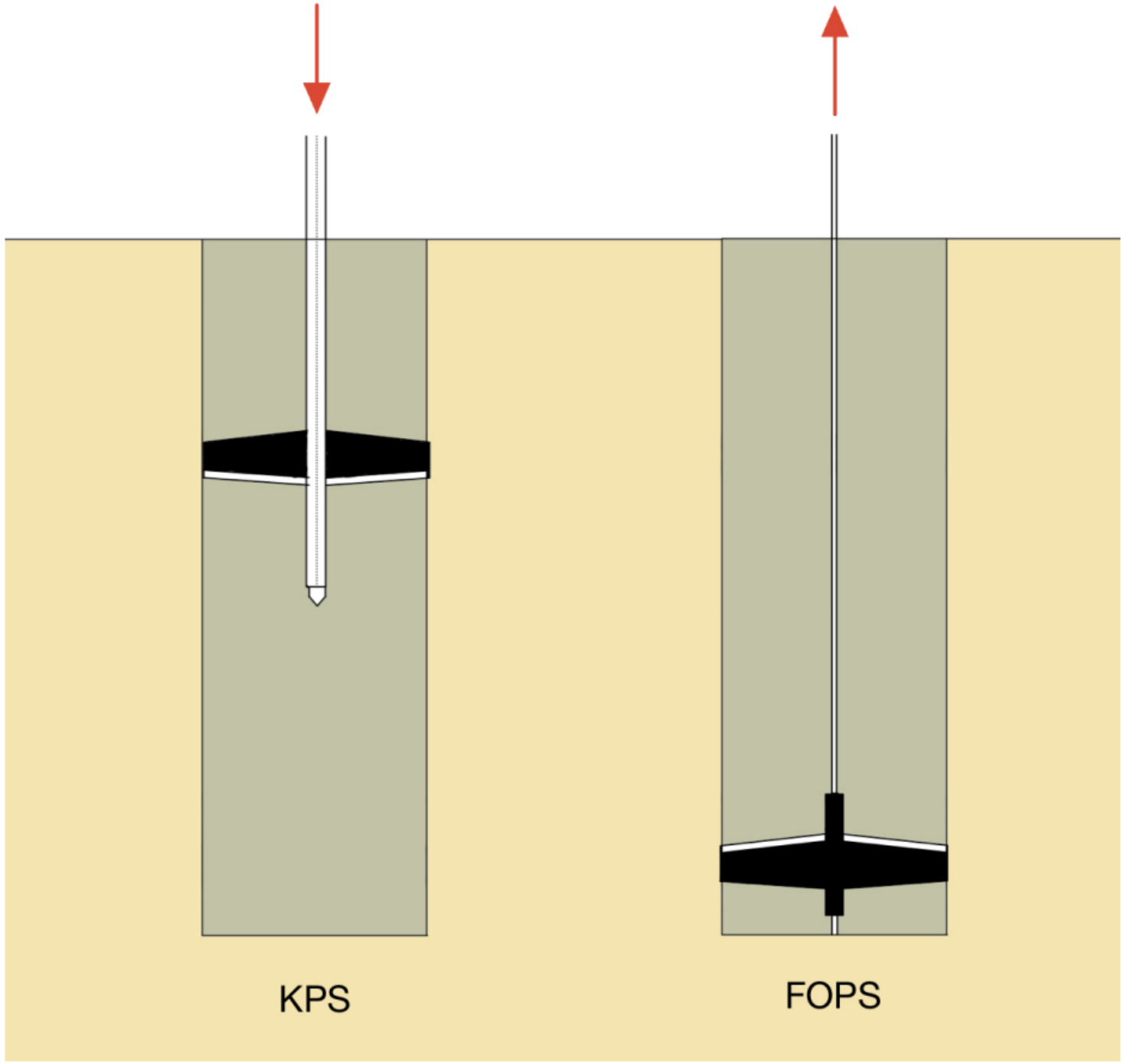
- Ny 19 km firefelts motorvei
- 108 000 KS-peler installert

Installasjonsparametere:

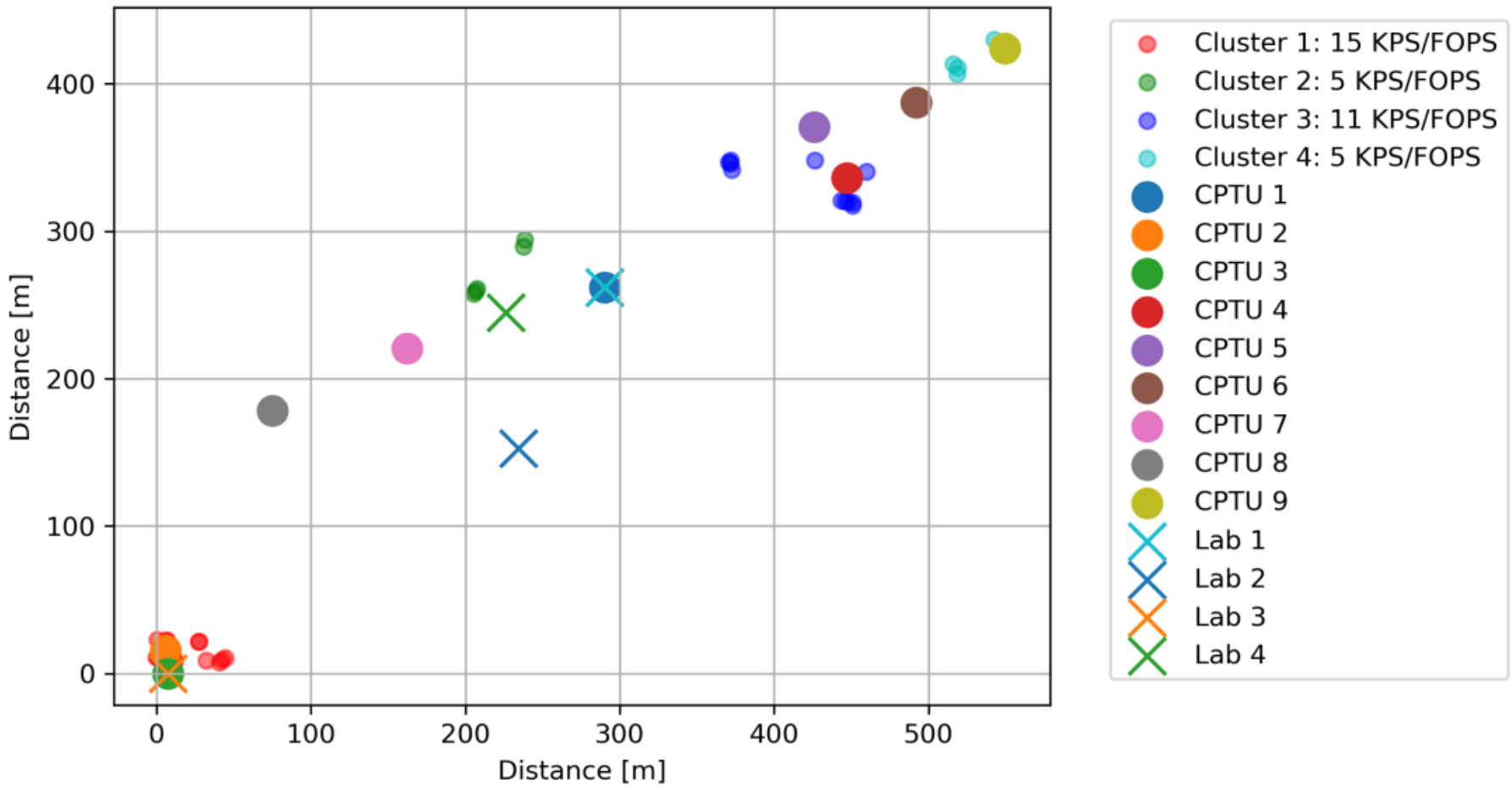
- 50 kg/m^2 bindemiddel
- BRN > 250
- Diameter = 0,6-0,8m

Tilgjengelig data fra 5 parseller:

- 167 KPS/FOPS
- 33 CPTU
- 14 Prøveserier

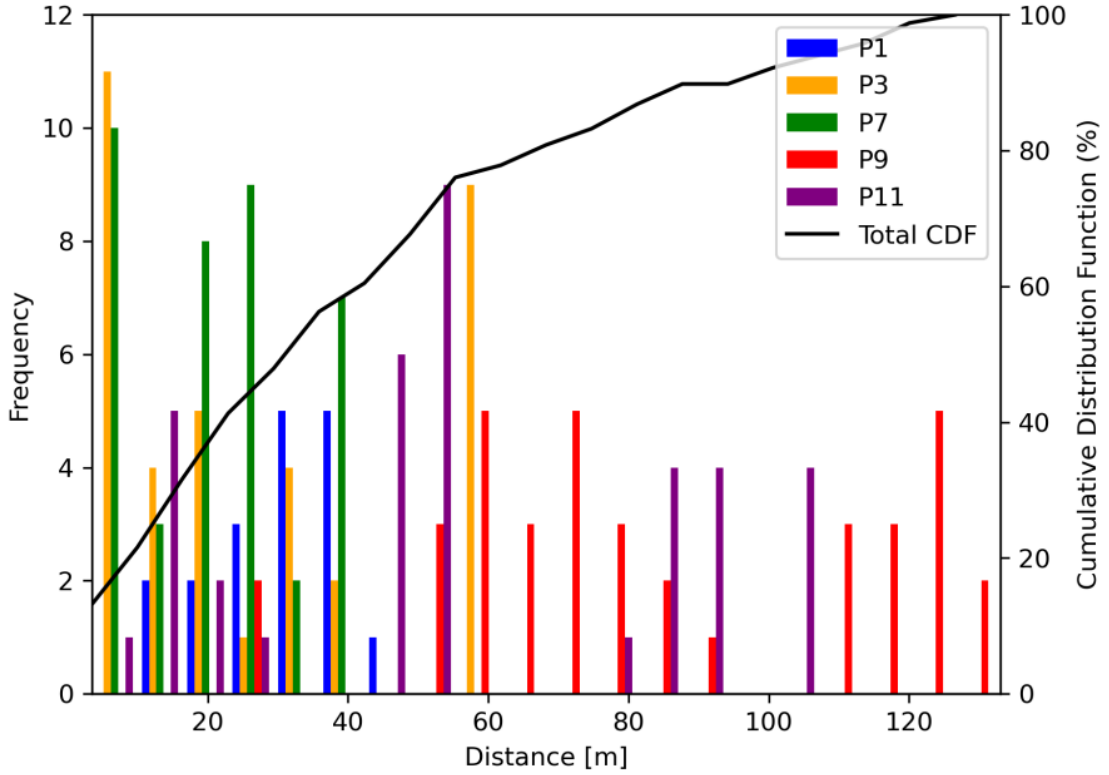


Oversikt P3

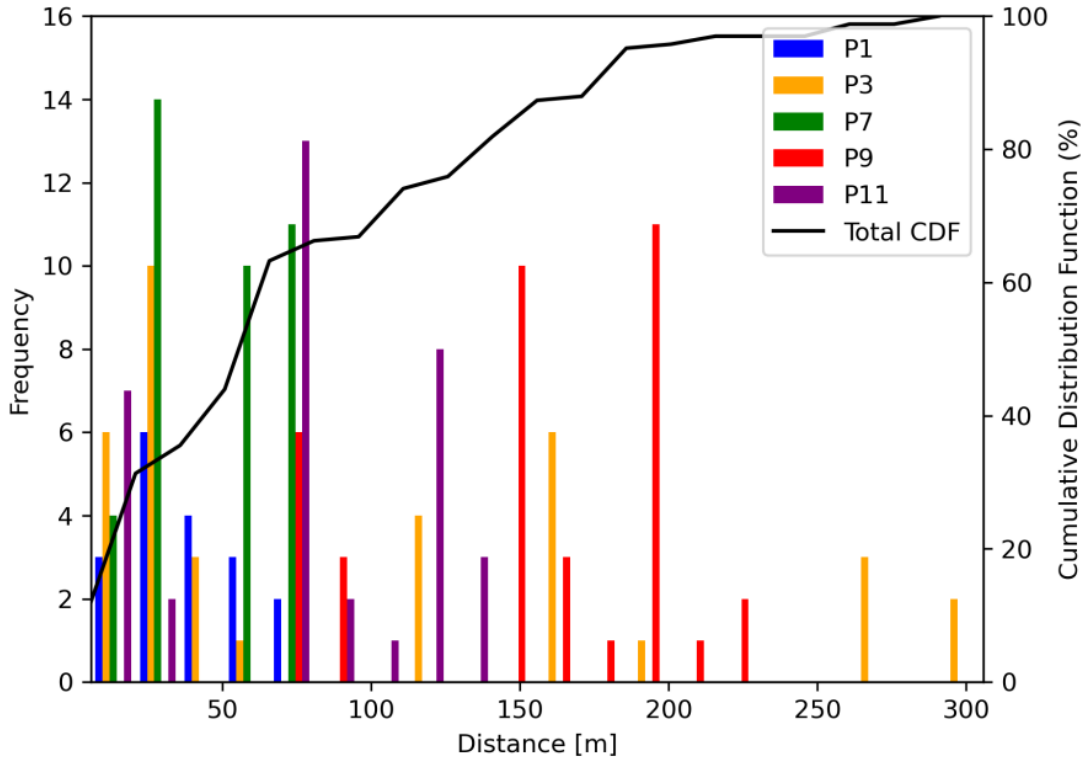


SWECO

Avstand mellom KPS/FOPS og CPTU



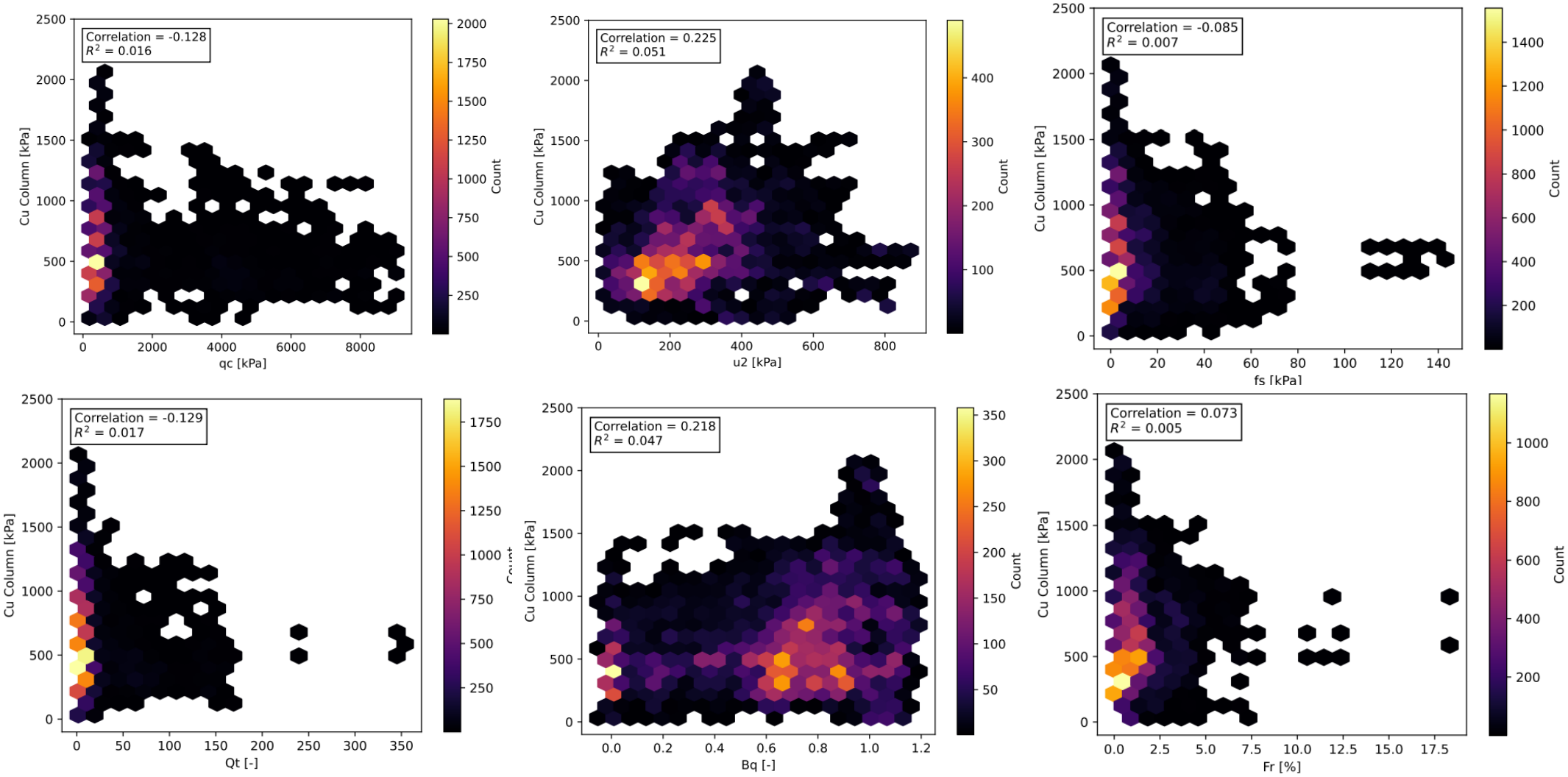
Avstand mellom KPS/FOPS og Lab



Korrelasjonsanalyse CPTU Data

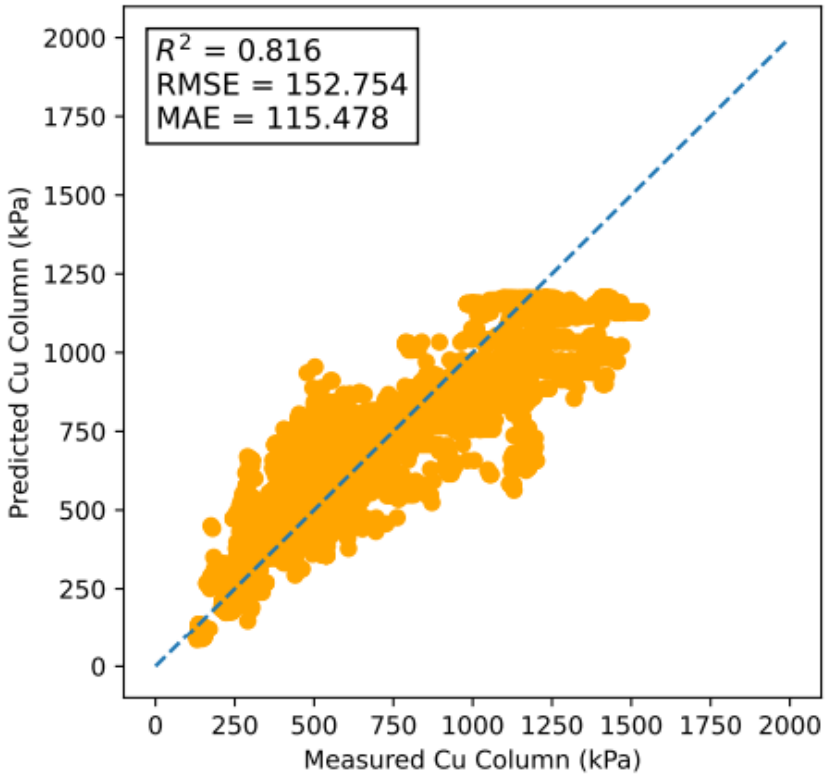
Variable	Symbol	Count	R	R^2
Cone resistance	q_c	17071	-0.13	0.02
Pore pressure	u_2	17071	0.23	0.05
Sleeve friction	f_s	17071	-0.09	0.01
Cone resistance number	Q_t	17071	-0.13	0.02
Friction ratio	F_r	17071	0.07	0.01
Pore pressure ratio	B_q	17071	0.22	0.05

Korrelasjonsanalyse CPTU Data

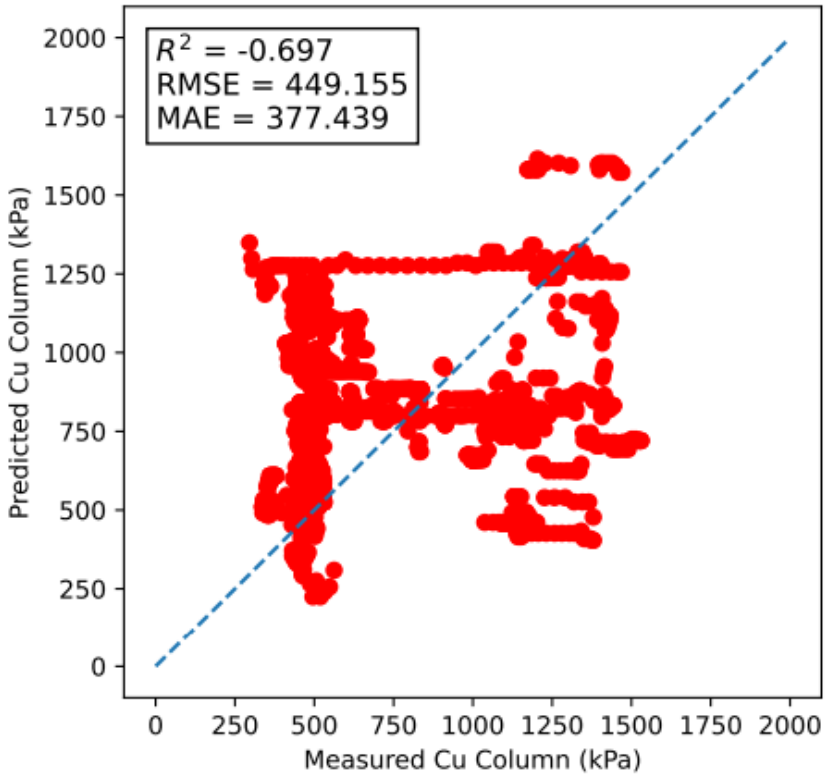


SWECO

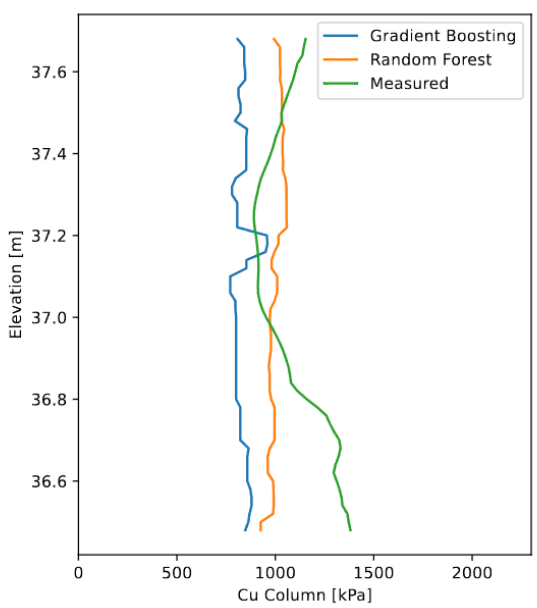
Prediksjon av Skjærstyrke med Maskinlæring

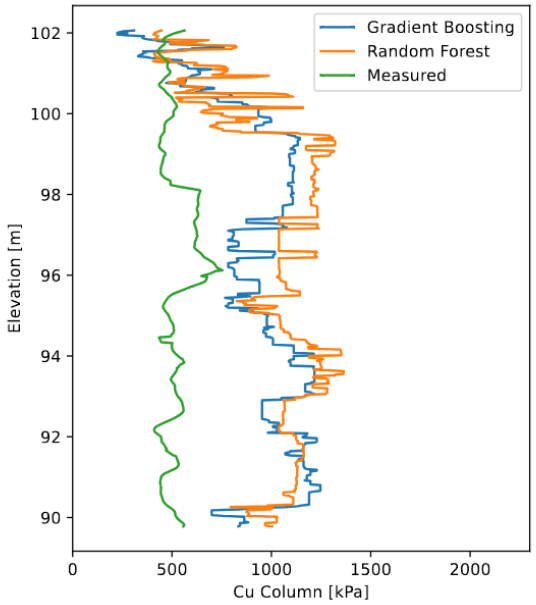


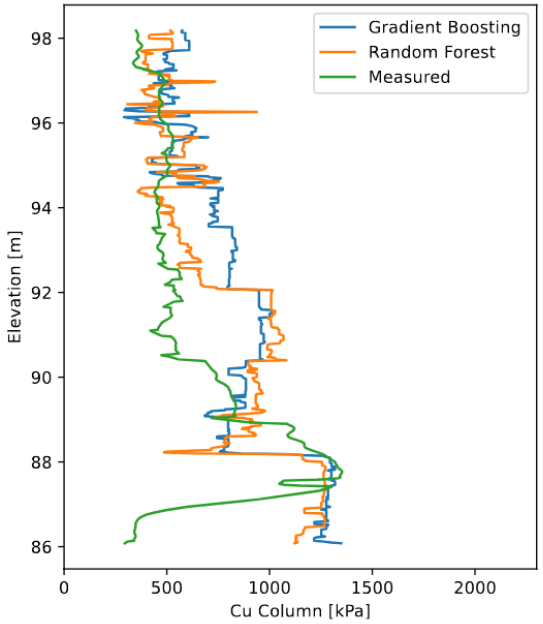
(a) Training data CPTU E6.

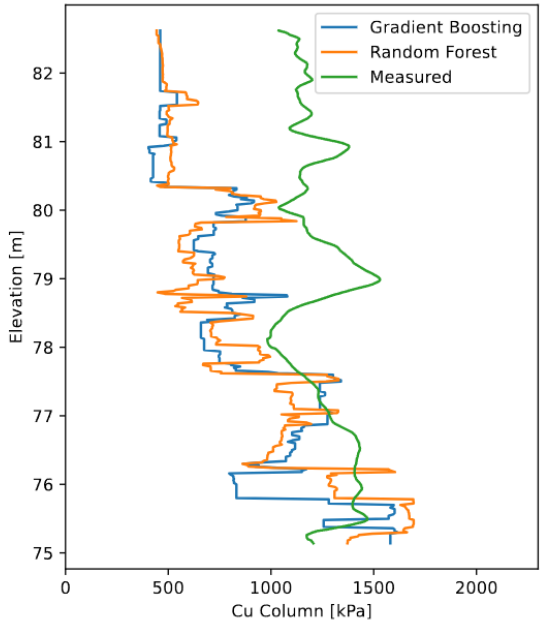


(b) Testing data CPTU E6.


Gradient Boosting
Random Forest
Measured
Elevation [m]
37.6
37.4
37.2
37.0
36.8
36.6
0
500
1000
1500
2000
Cu Column [kPa]


Gradient Boosting
Random Forest
Measured
Elevation [m]
102
100
98
96
94
92
90
0
500
1000
1500
2000
Cu Column [kPa]


Gradient Boosting
Random Forest
Measured
Elevation [m]
98
96
94
92
90
88
86
0
500
1000
1500
2000
Cu Column [kPa]

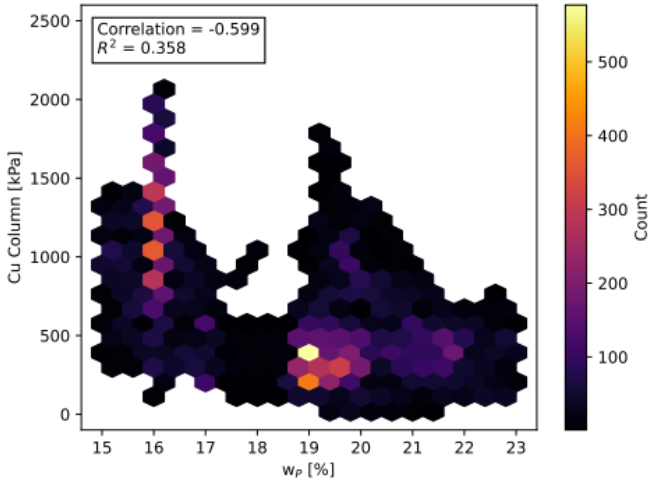

Gradient Boosting
Random Forest
Measured
Elevation [m]
82
81
80
79
78
77
76
75
0
500
1000
1500
2000
Cu Column [kPa]

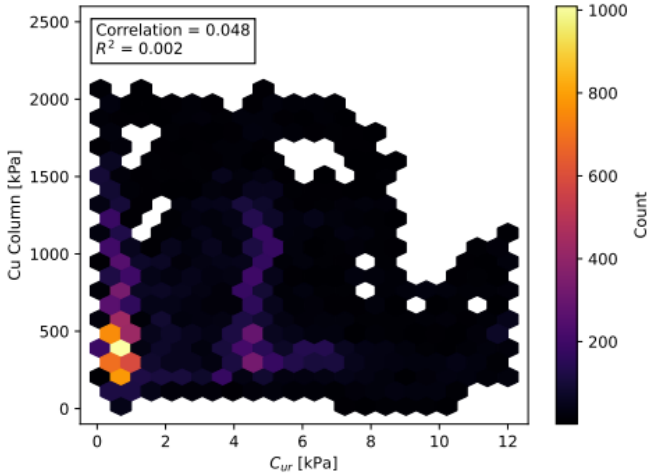
SWECO

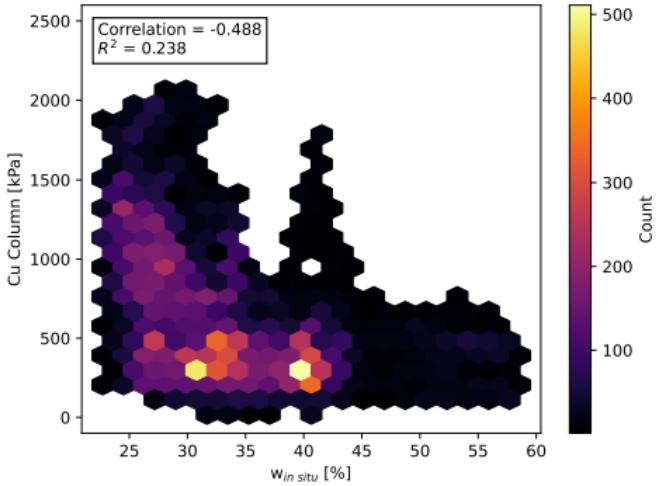
Korrelasjonsanalyse Laboratoriedata

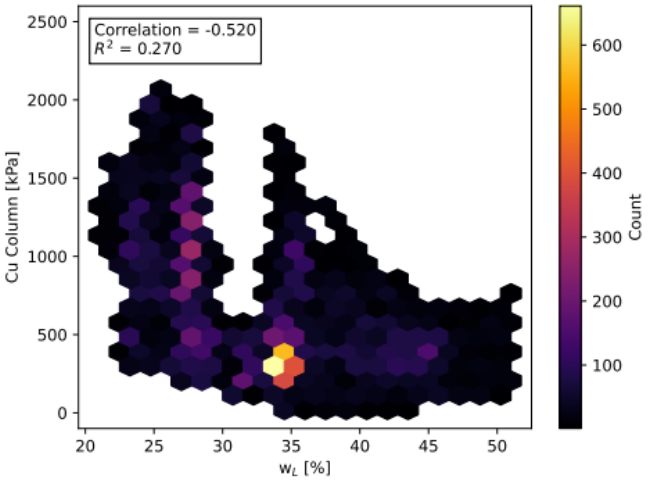
Variable	Symbol	Count	R	R^2
Water content	$w_{\text{in situ}}$	18137	-0.49	0.24
Shear strength	C_u	18042	-0.04	0.00
Remolded shear strength	C_{ur}	18042	0.05	0.00
Sensitivity	S_t	18042	-0.03	0.00
Plastic limit	w_P	12706	-0.60	0.36
Liquid limit	w_L	12706	-0.52	0.27

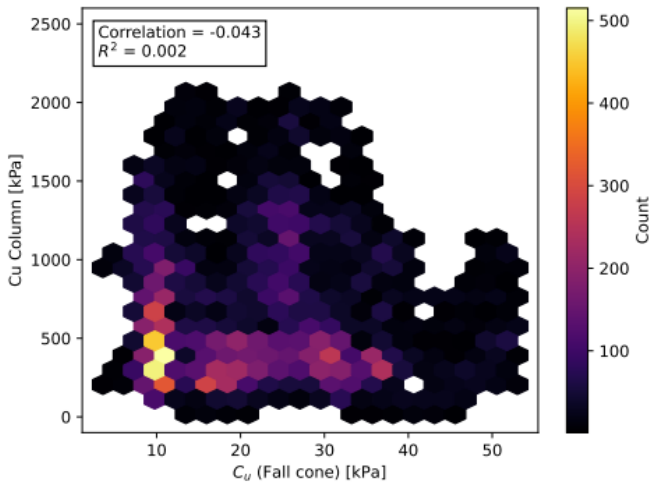
Korrelasjonsanalyse Laboratoriedata

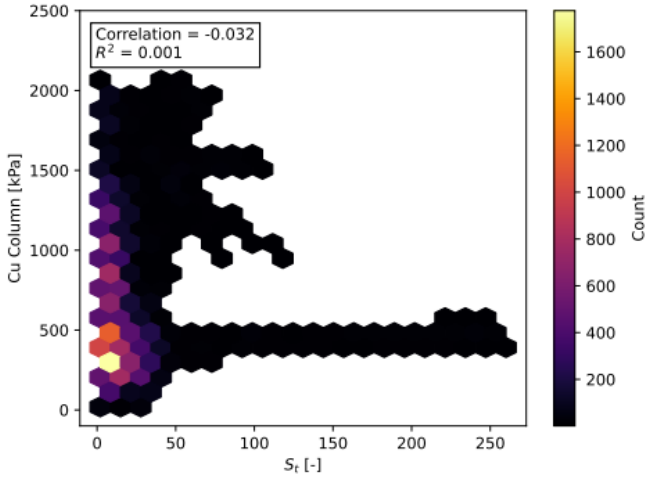






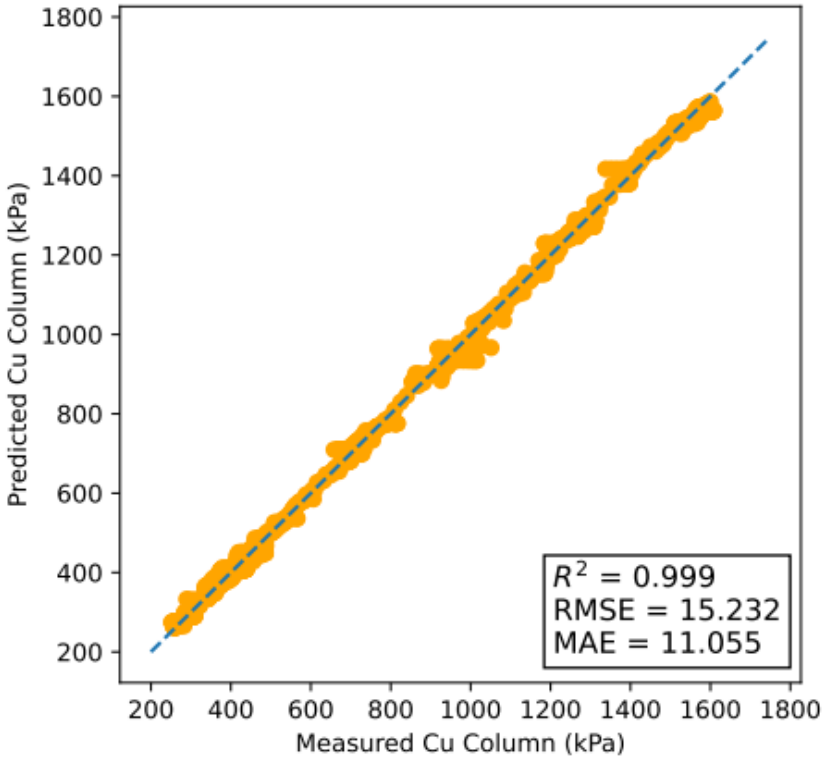




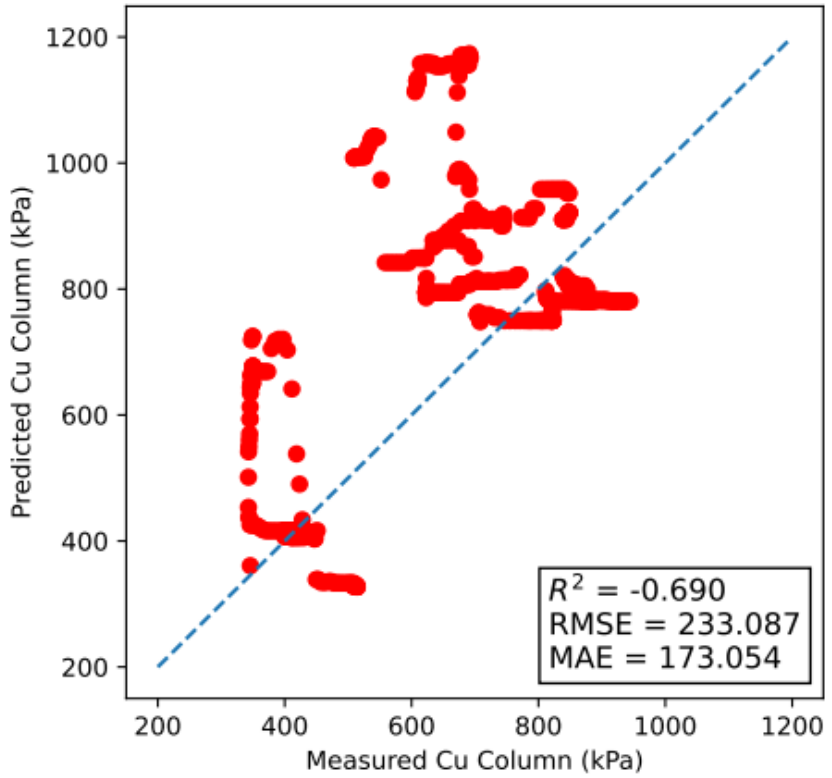


SWECO

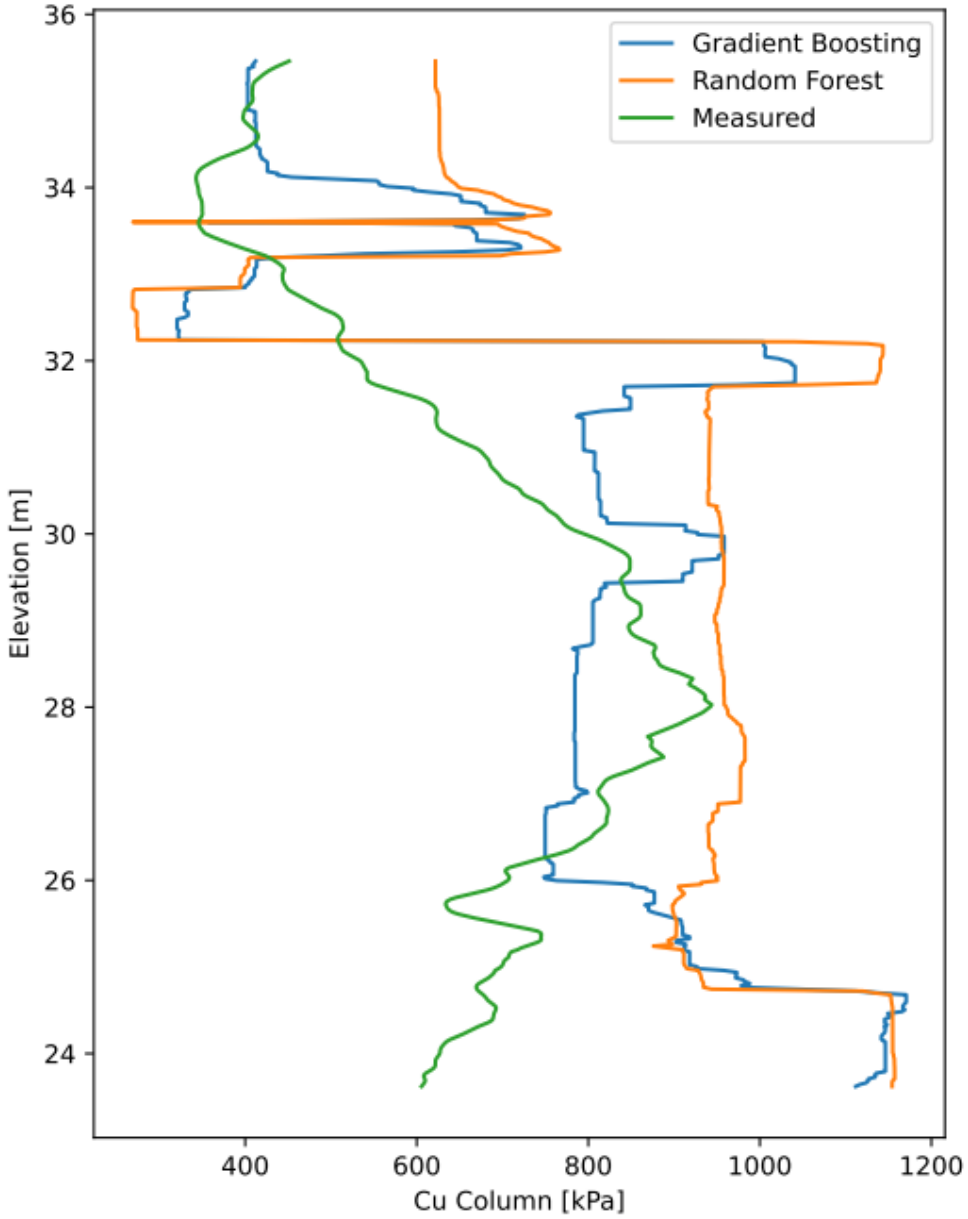
Prediksjon av Skjærstyrke med Maskinlæring



(a) Training data Lab E6.



(b) Testing data Lab E6.

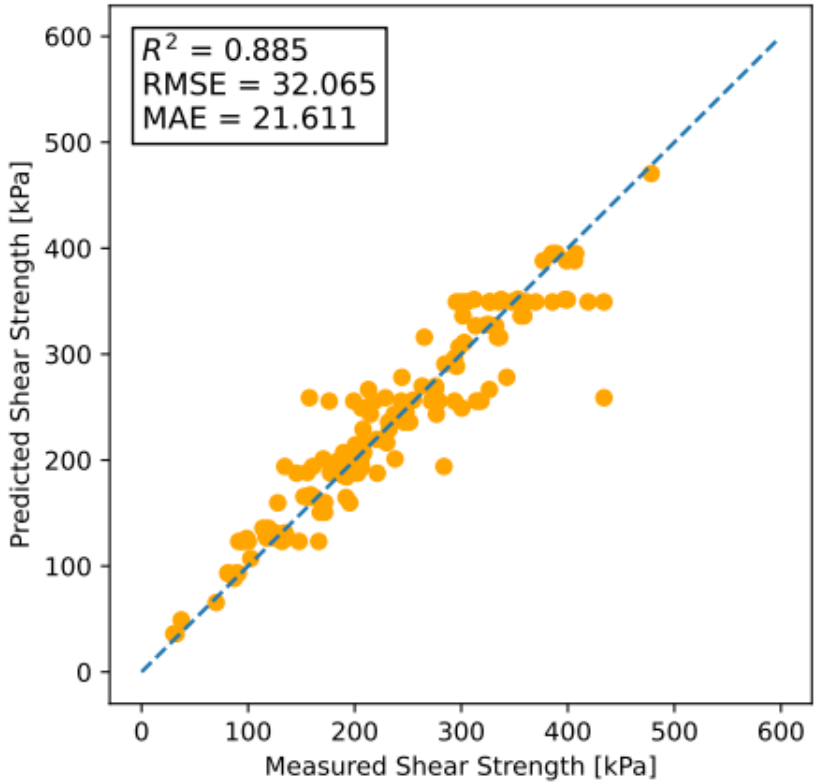

Gradient Boosting
Random Forest
Measured
Elevation [m]
36
34
32
30
28
26
24
400
600
800
1000
1200
Cu Column [kPa]

SWECO

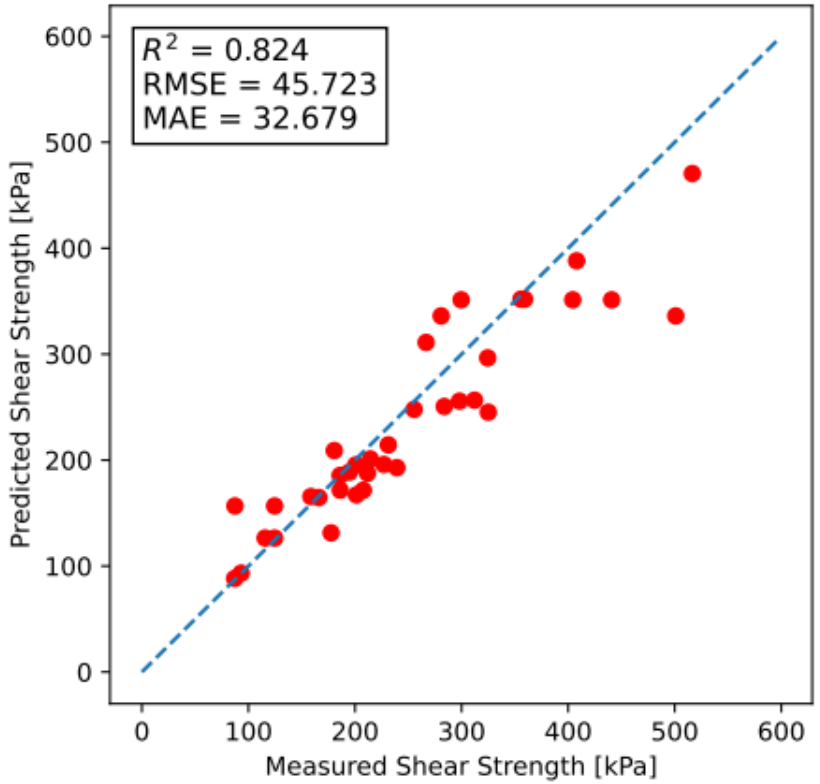
Norsk Database

Variable	Symbol	Count	R	R^2
Water content	$w_{in\,situ}$	256	-0.67	0.45
Sensitivity	S_t	86	0.19	0.04
Unit weight	γ	246	0.72	0.52
Plastic limit	w_P	195	-0.58	0.34
Liquid limit	w_L	198	-0.66	0.44
Plasticity index	I_P	195	-0.60	0.36
Liquidity index	I_L	195	0.25	0.06
Lime content	-	192	0.30	0.09
Cement content	-	292	0.19	0.04
CKD content	-	108	-0.07	0.01
Porosity unstabilized	-	246	-0.71	0.50
Water binder ratio	wbr	246	-0.59	0.35

Prediksjon av Skjærstyrke med Maskinlæring



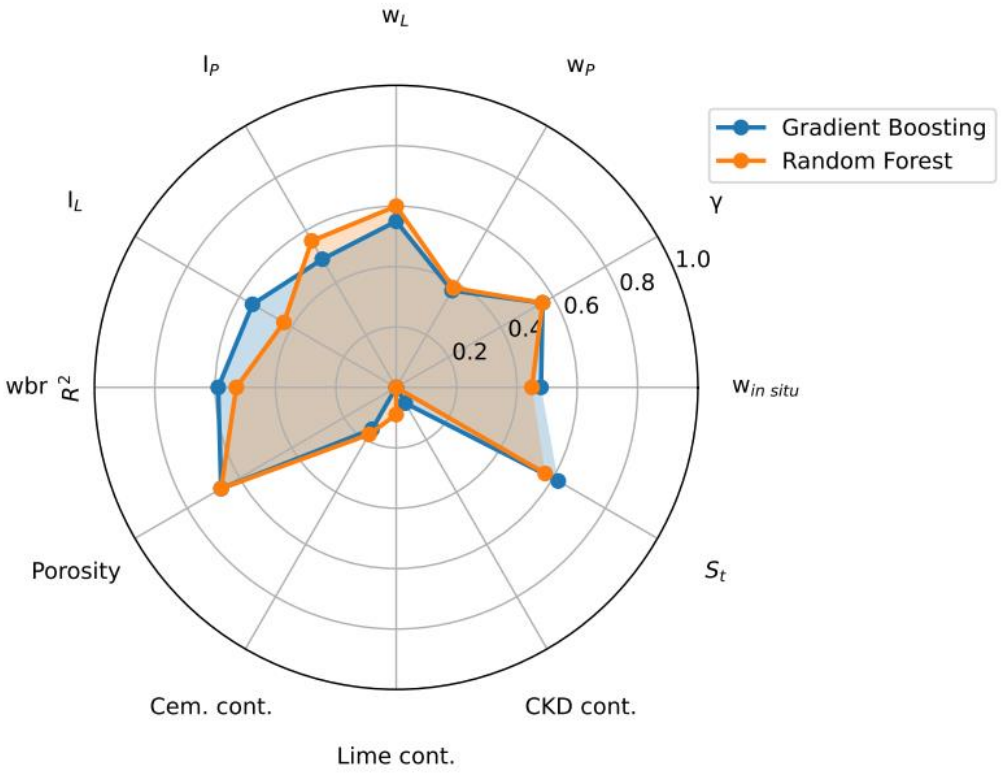
(a) Training data Norwegian database.

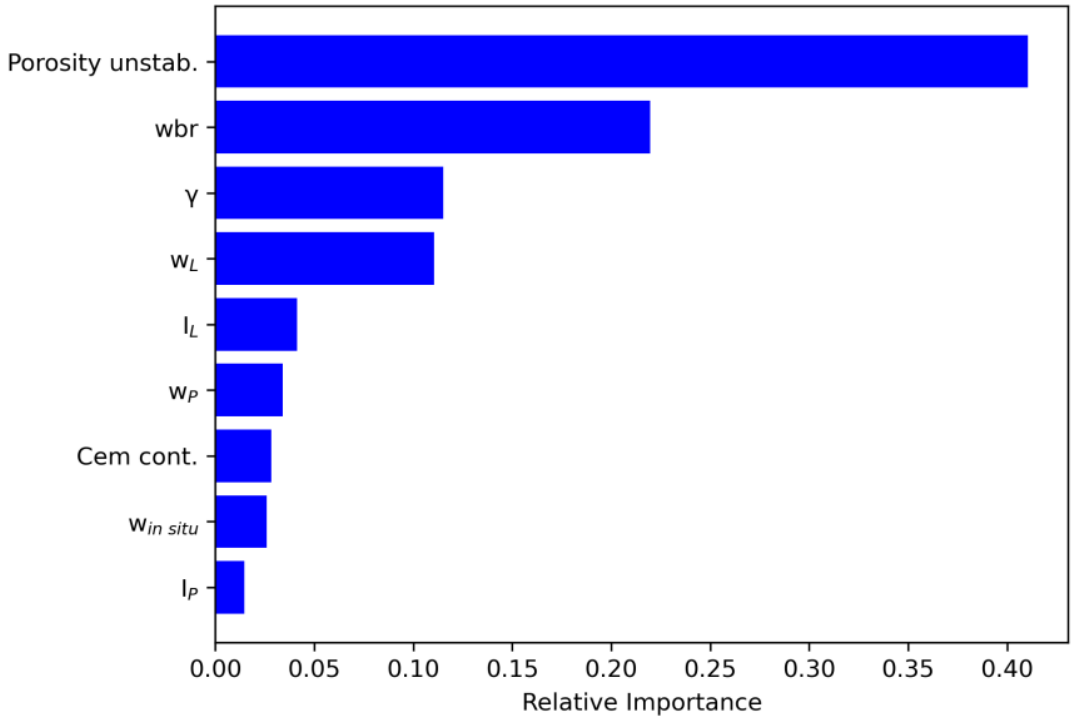


(b) Testing data Norwegian database.

SWECO

Viktighet av hver parameter



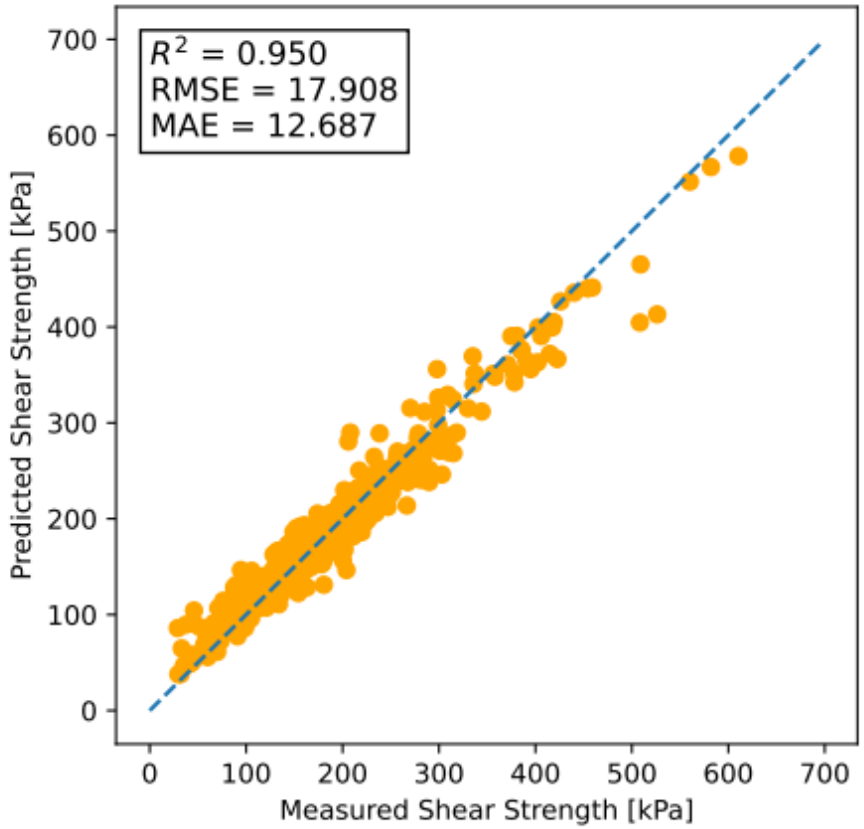


SWECO

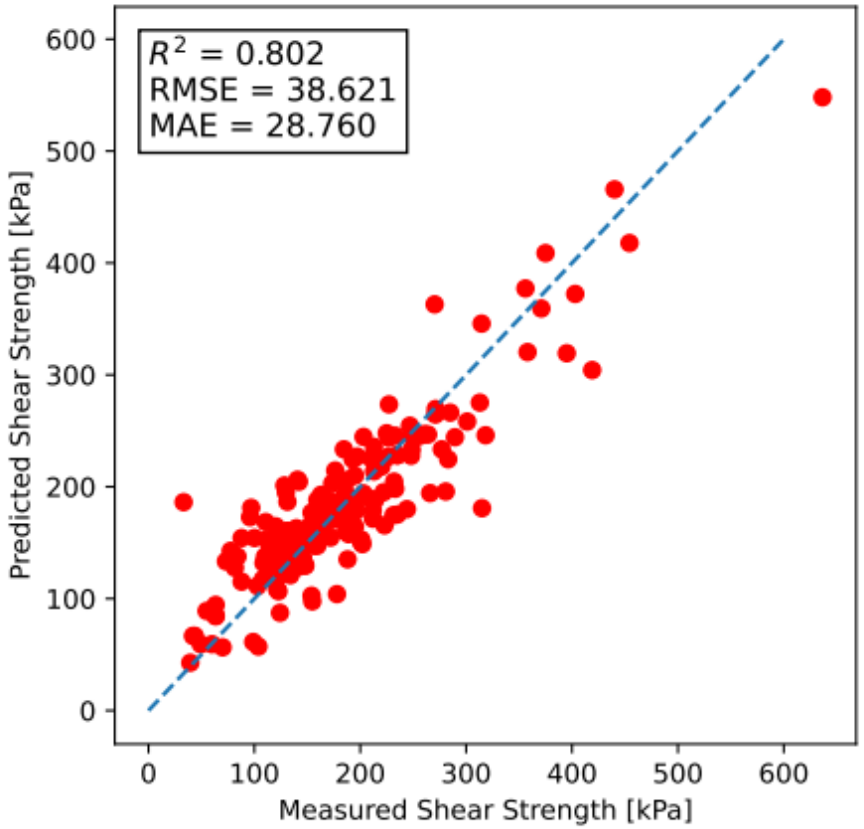
Svensk Database

Variable	Symbol	Count	R	R^2
Bulk density	ρ	874	0.09	0.01
Water content	$w_{\text{in situ}}$	874	0.08	0.01
Lime content	-	473	0.07	0.01
Cement content	-	874	0.41	0.17
CKD content	-	383	0.10	0.01
Porosity unstabilized	-	874	-0.07	0.01
Dry weight clay	-	874	0.12	0.01
Water binder ratio	wbr	874	-0.48	0.23

Prediksjon av Skjærstyrke med Maskinlæring



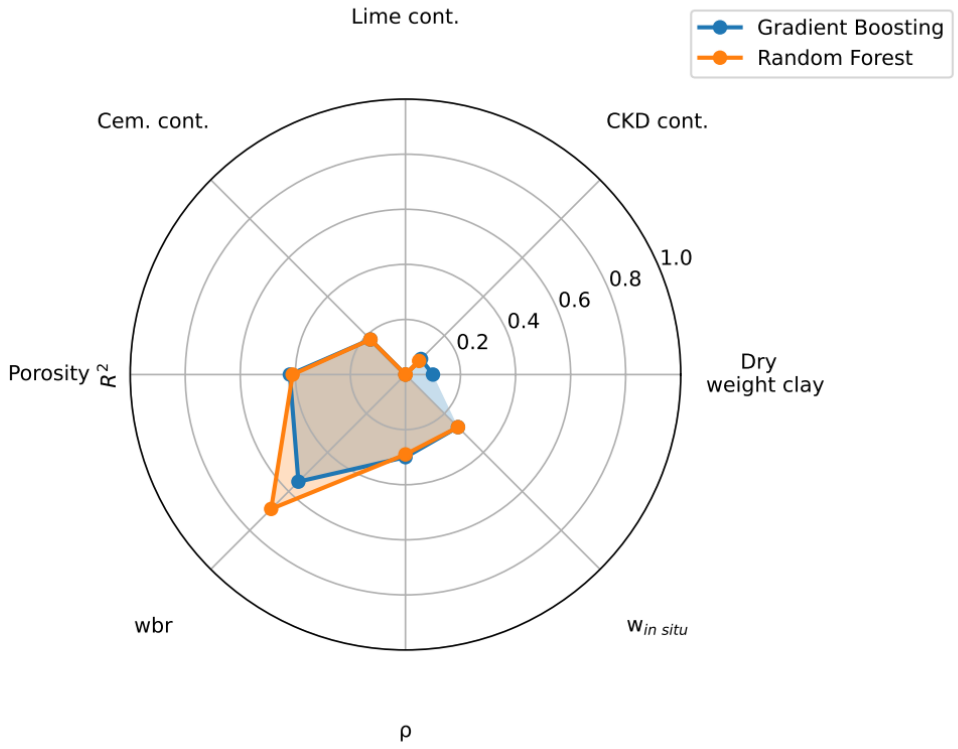
(a) Training data Swedish database.

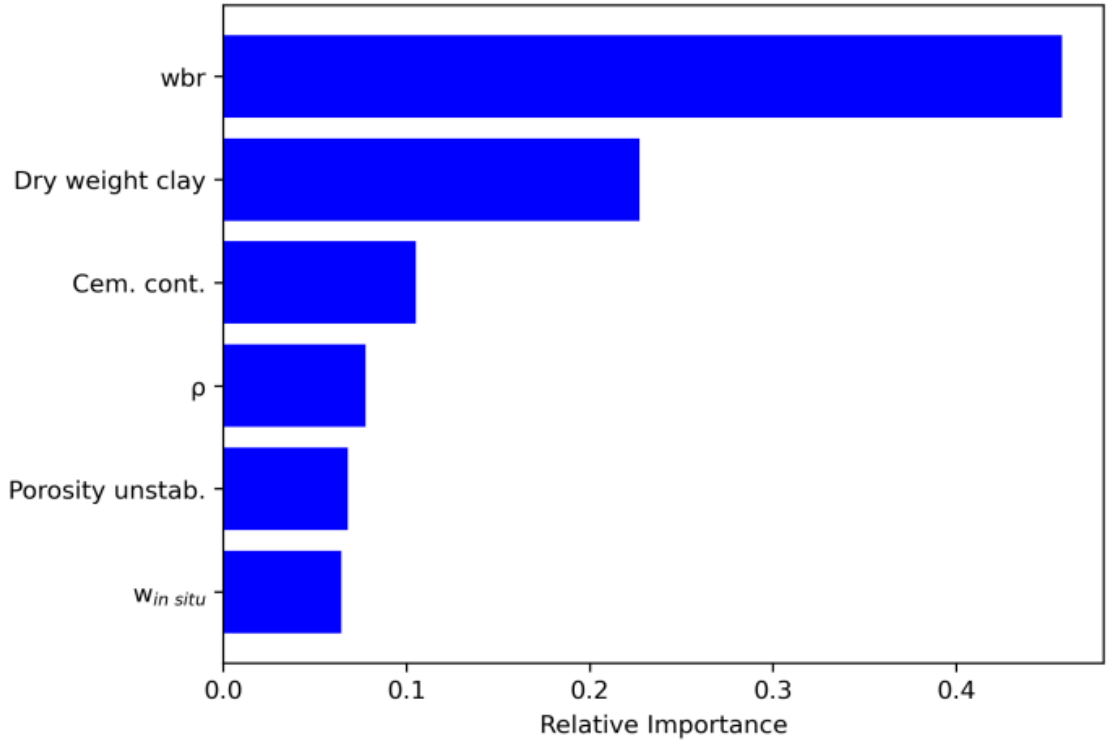


(b) Testing data Swedish database.

SWECO

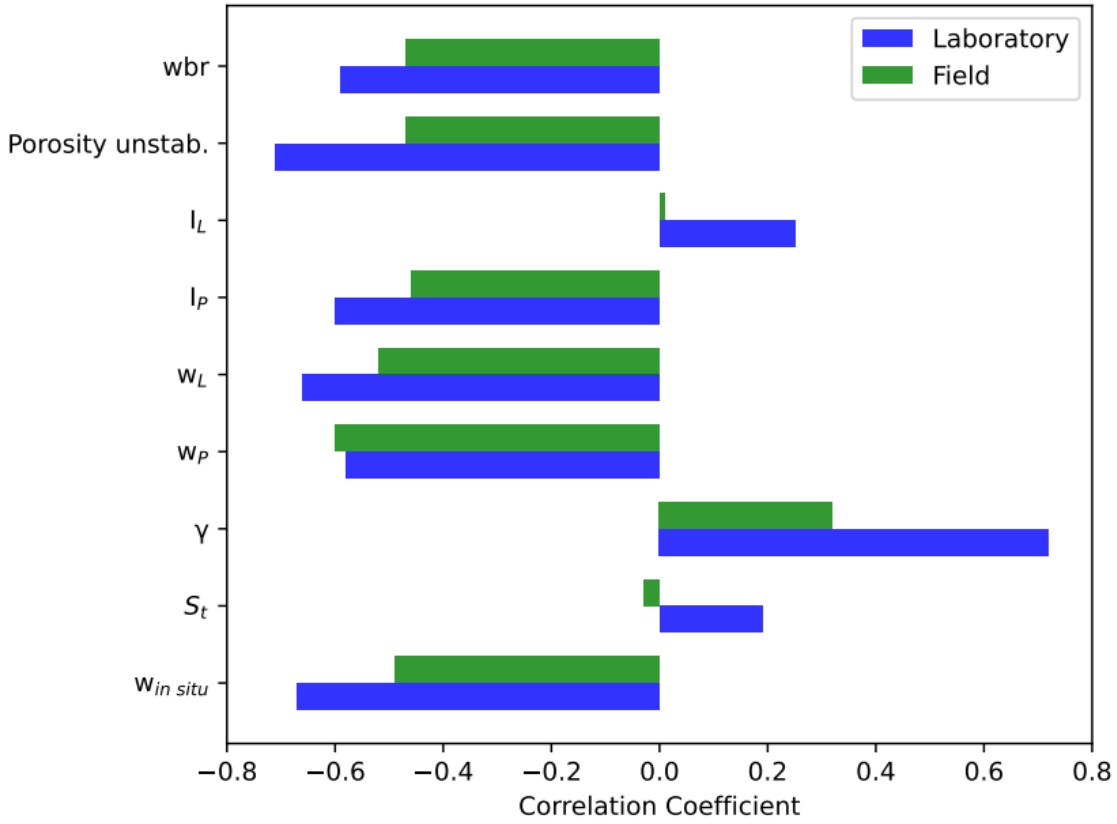
Viktighet av hver parameter

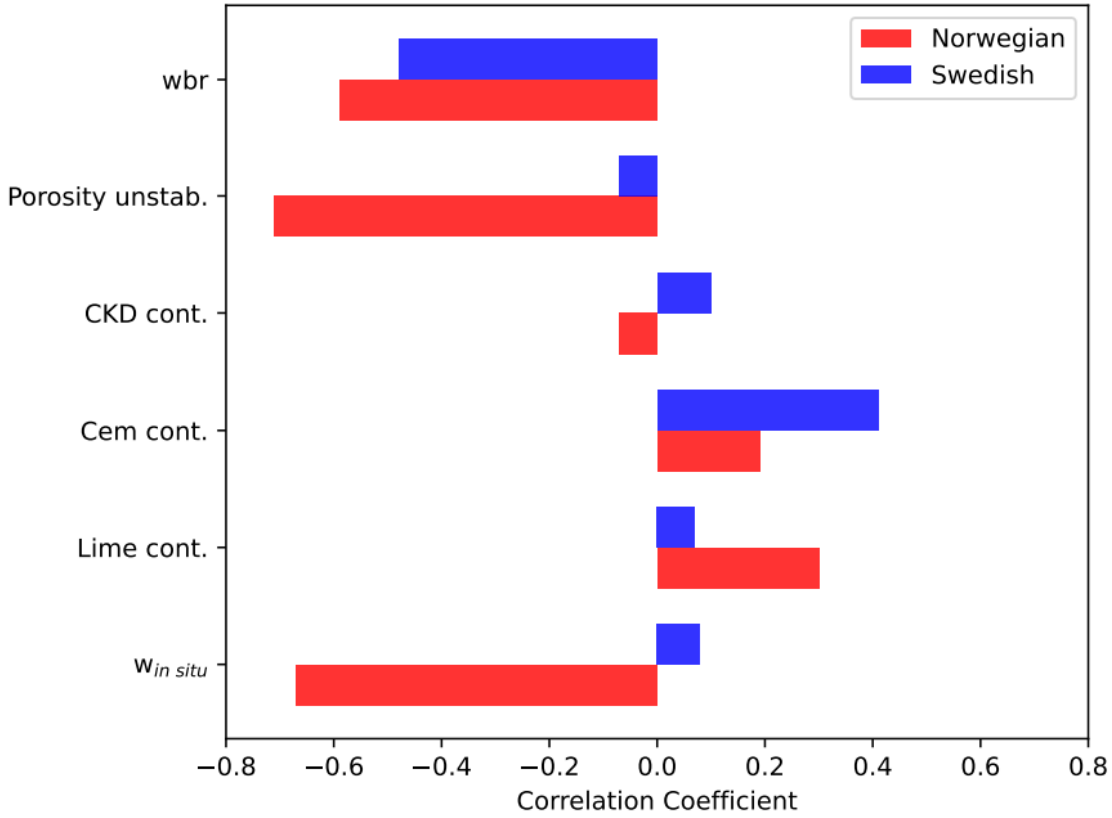




SWECO

Sammenligning av Korrelasjoner





SWECO

Konklusjon

E6 Kvithammar-Åsen:

- Ingen korrelasjon mellom styrken til KS pelene og CPTU data.
- Derimot en svak korrelasjon mellom styrken til KS pelene og laboratoriedata.
- Det har vist seg vanskelig å benytte maskinlæring til å predikere styrken til KS peler basert på feltdata.

Norsk og Svensk database:

- Den norske databasen oppnådde den høyeste ytelsen med en R^2 = 0.82
- Vann-bindemiddel tallet klarte å predikere styrken til den stabiliserte jorden relativt bra på egenhånd gitt ved en R^2 = 0.69.