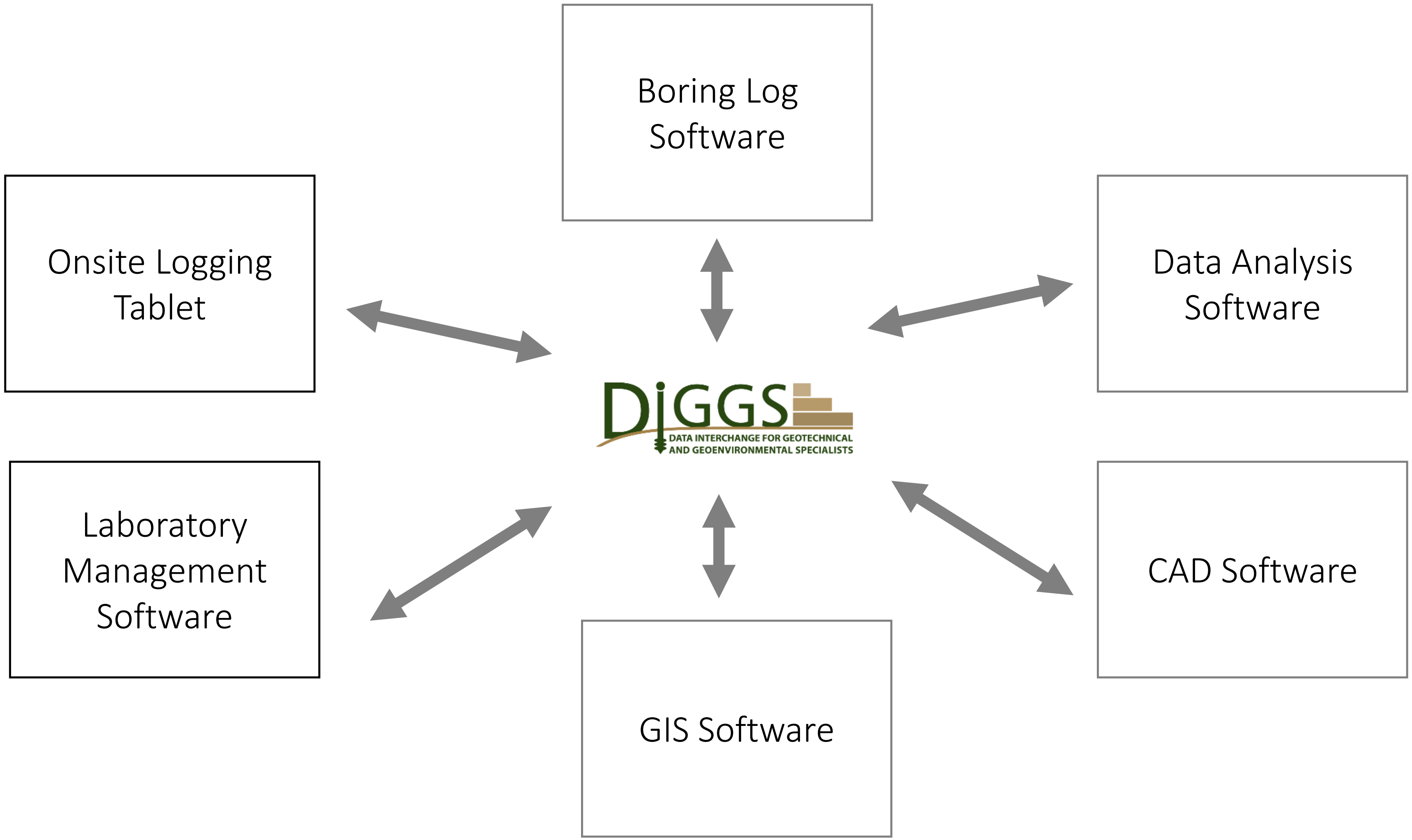
AI Based Methods for Characterization of Geotechnical Site Investigation Data

Leverage Geotechnical Data Asset

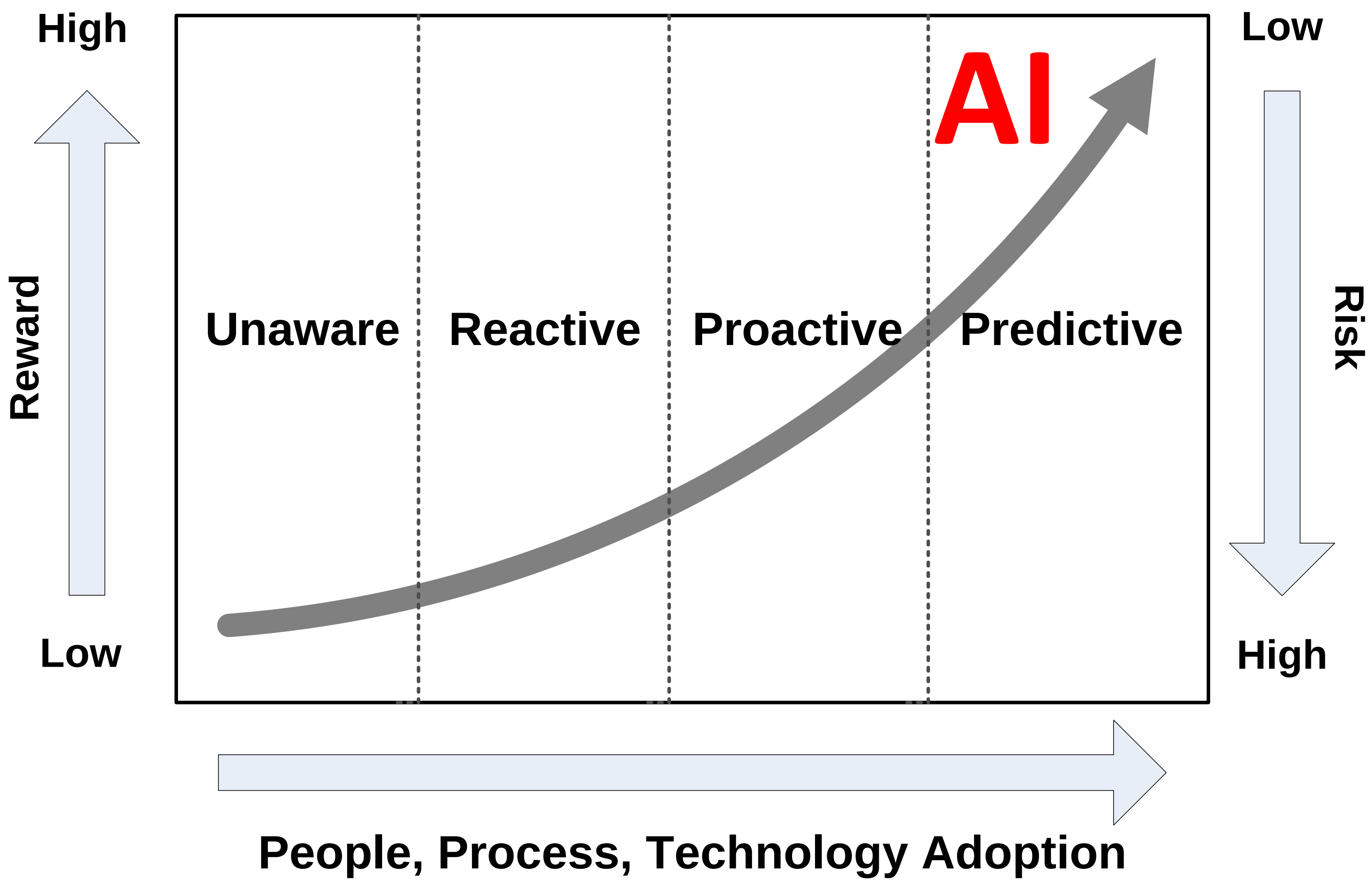
Robert Liang, Ph.D. P.E., Jack (Hui) Wang, Ph.D., and Xiangrong Wang, Ph.D.
The University of Dayton

Streamline Internal Processes



Data Management Maturity Model

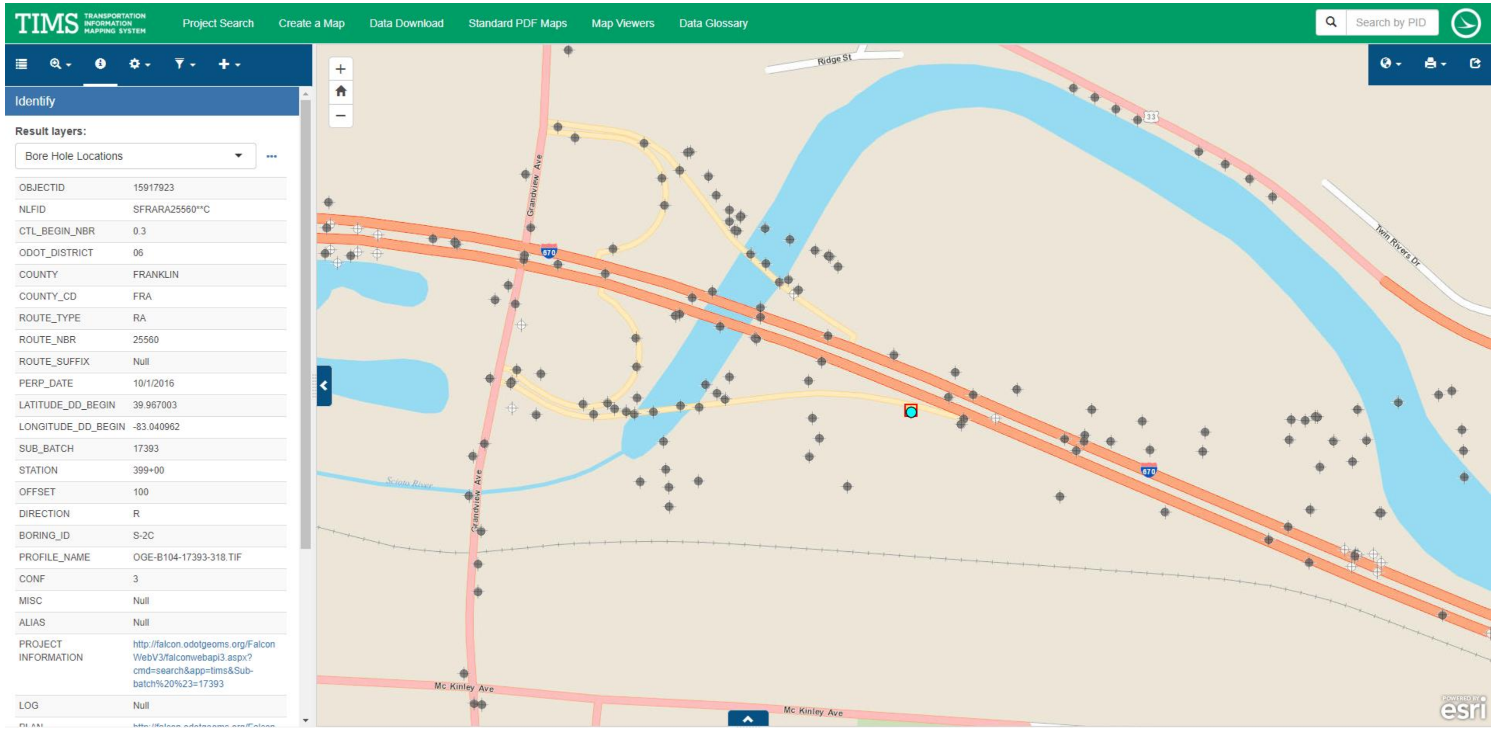
Pareek, D. (2007) “Business Intelligence for Telecommunications”



- Historic geotechnical information/data is an asset
- DIGGS compatible data now required in Ohio

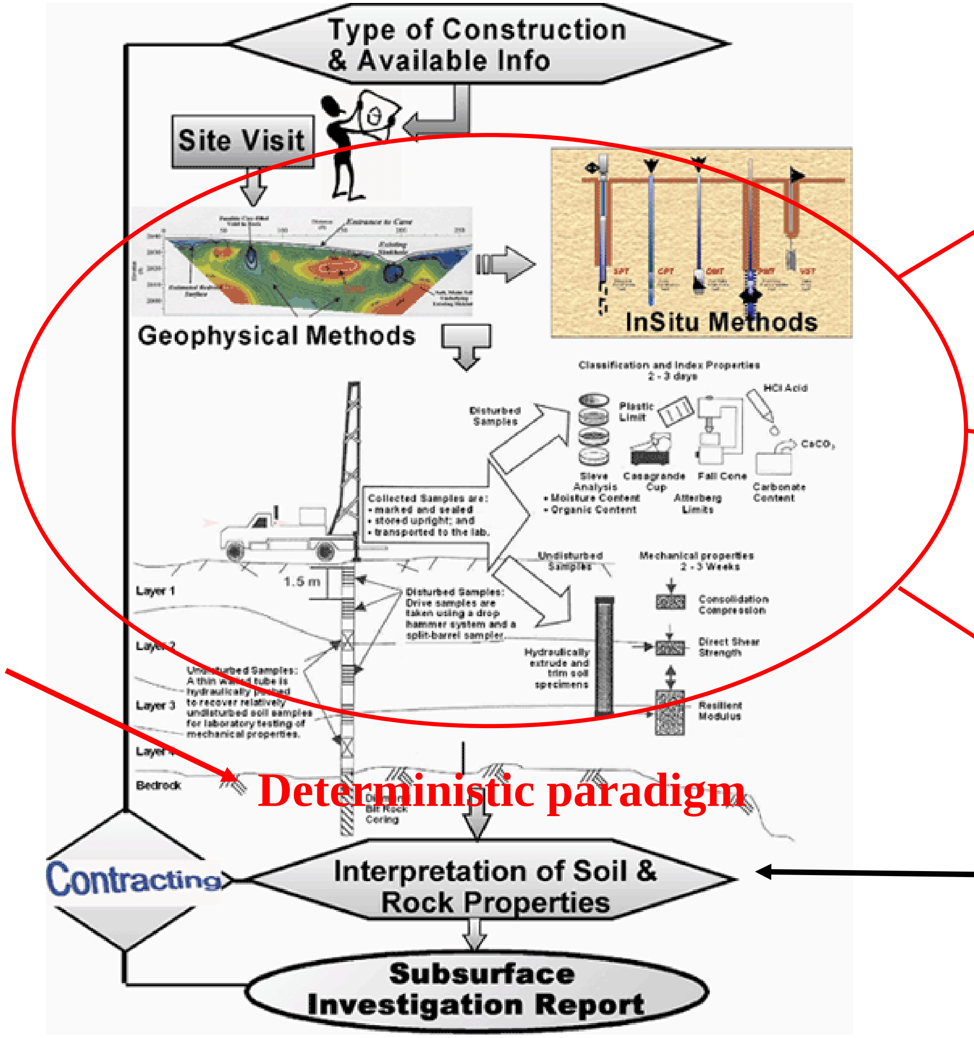
DIGGS XML data file + TIMS = Statewide Digitized Sparse (but rich) Geotechnical Database


```xml
<?xml version="1.0"?>
- <Diggs gml:id="NTM1YTEwMzUtYmRmZS00NTIxLTk1OWItZjk4NGEwOWRjZDcx"
  xmlns:glrov="http://www.opengis.net/gml/3.3/lrov" xmlns:diggs="http://diggsml.org/schemas/2.5.a"
  xmlns:gml="http://www.opengis.net/gml/3.2" xmlns:g3.3="http://www.opengis.net/gml/3.3/ce"
  xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:glr="http://www.opengis.net/gml/3.3/lr"
  xmlns:witsml="http://www.witsml.org/schemas/131" xmlns:diggs_geo="http://diggsml.org/schemas/2.5.a/geotechnical"
  xsi:schemaLocation="http://diggsml.org/schemas/2.5.a https://raw.githubusercontent.com/DIGGSml/diggs-
  schema/2.5.a-dev/Complete.xsd" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="http://diggsml.org/schemas/2.5.a">
  - <documentInformation>
    - <DocumentInformation gml:id="DGS765F-DA6-2E5C-4987-1FA95">
        <creationDate>2018-12-11</creationDate>
      - <author>
        - <BusinessAssociate gml:id="BA_">
            <gml:name/>
          </BusinessAssociate>
        </author>
          <!--<author></author>-->
        <disclaimer>Not specified in gINT file</disclaimer>
      - <auditTrail>
        - <Remark>
            <content>Generated by Keynetix Diggs Feedback Tool</content>
            <remarkDateTime>2018-12-11</remarkDateTime>
          </Remark>
        </auditTrail>
      </DocumentInformation>
    </documentInformation>
  - <project>
    - <Project gml:id="Project_600162">
        <gml:identifier codeSpace="ODOT_CS">100056</gml:identifier>
        <gml:name>600162</gml:name>
          <!--NO roles defined for this file conversion<role> <Role> <rolePerformed>Client</rolePerformed>
          <businessAssociate></businessAssociate> </Role> </role>-->
      - <locality>
        - <Locality gml:id="pl_600162">
          - <address>
            - <Address gml:id="pla_600162">
                <number/>
                <streetAddress/>
                <city/>
                <state/>
                <province/>
                <county>PER</county>
                <country/>
                <postalCode/>
              </Address>
            </address>
          - <plss>
            - <PLSS gml:id="DGS81FA-5C-3CDB-CED2-4A590">
                <township/>
                <range/>
              </PLSS>
            </plss>
            <route>345</route>
            <station/>
            <offsetDirection/>
          </Locality>
        </locality>
        <purpose>Geohazard Exploration</purpose>
      - <otherProjectProperty>
```

Geotechnical site investigation data is a gold mine!

Two Geotechnical Exploration and Testing paradigms

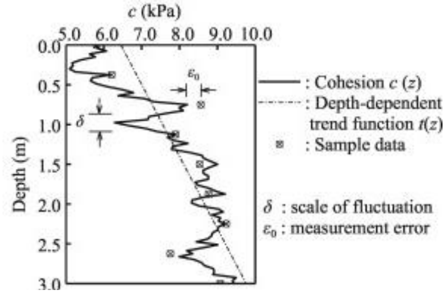


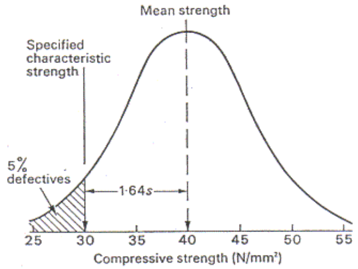
Experience driven

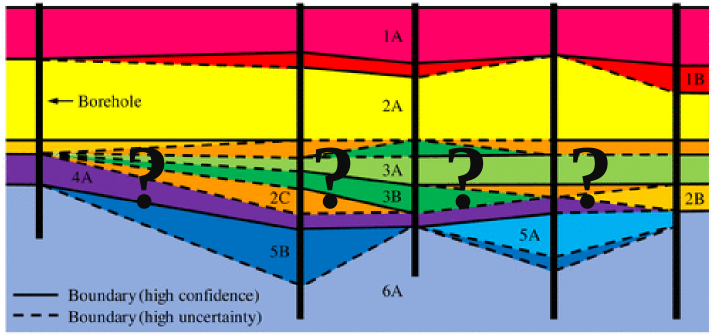
Deterministic paradigm

Courtesy: https://www.fhwa.dot.gov/engineering/geotech/

Modeling & Uncertainty quantification



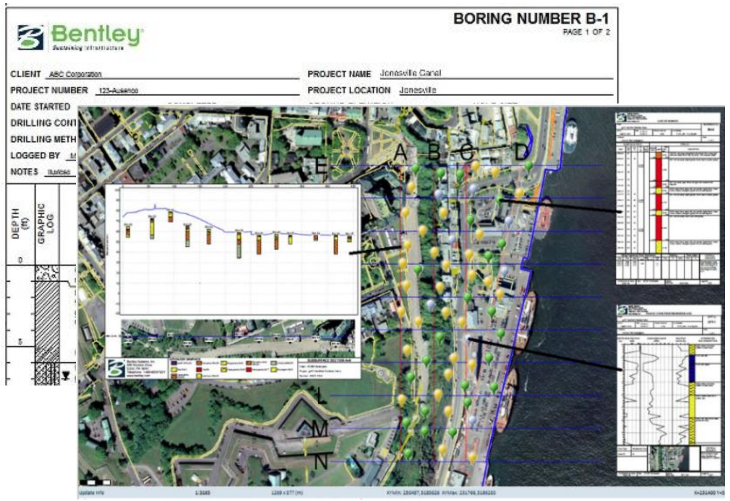




Stochastic paradigm

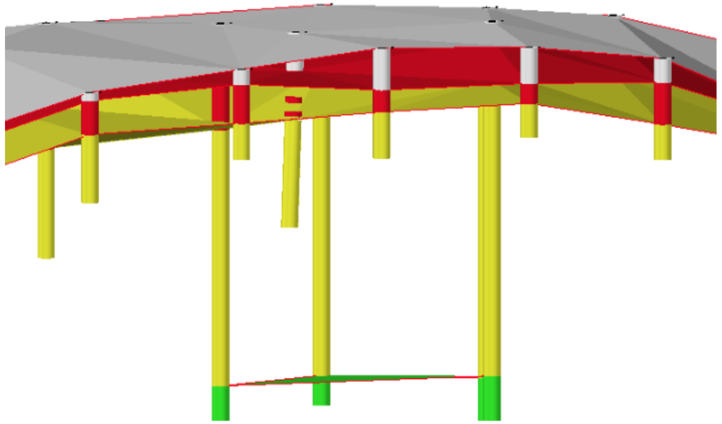
Data driven

DATA RICH ≠ KNOWLEDGE RICH



Digitized site investigation database

No quantitative confidence level



Subsurface model and visualization

Experience-based decision

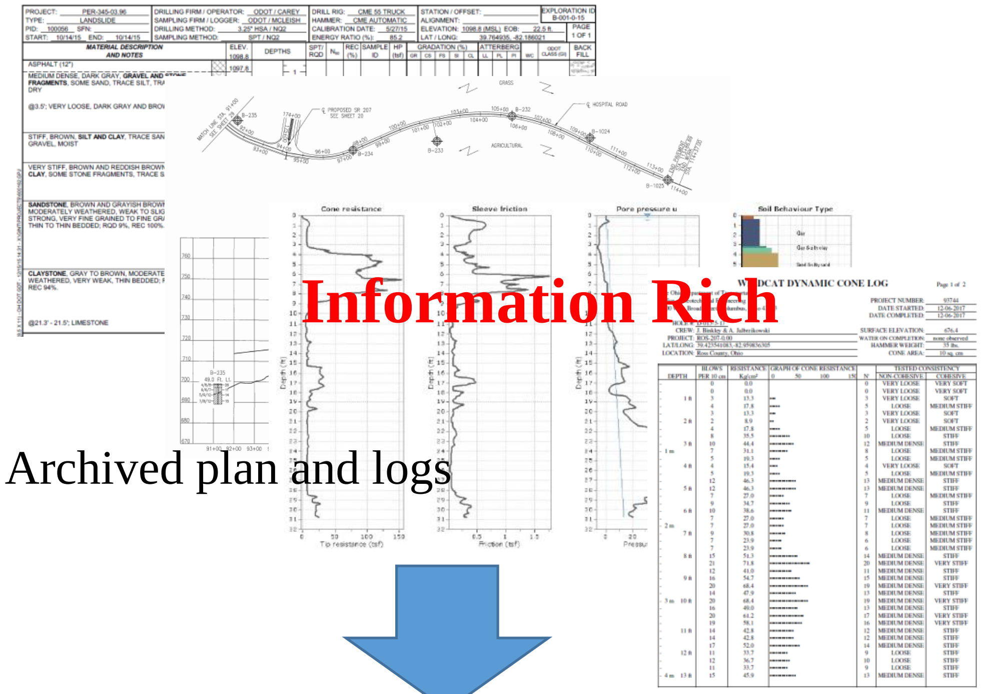
No quantitative cost-benefit evaluation

Additional site investigation layout

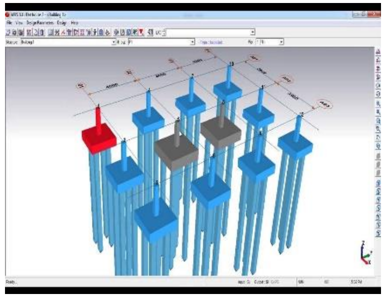
Deficiencies of current practice

- Current practices lack adequate and advanced methodologies for harvesting and harmonizing vast, diverse, and possibly contradicting geotechnical and geophysical investigation information/data (in-situ, laboratory, and derived)
- Engineering experiences-based interpretation involving bias as well as unknown uncertainties (both objective and subjective)
- Low efficiency and less robustness of current geotechnical data transferring and processing
- Deterministic “guess” on the subsurface condition at the unexplored locations (may be with subjective confidence level)

Current practices



Archived plan and logs

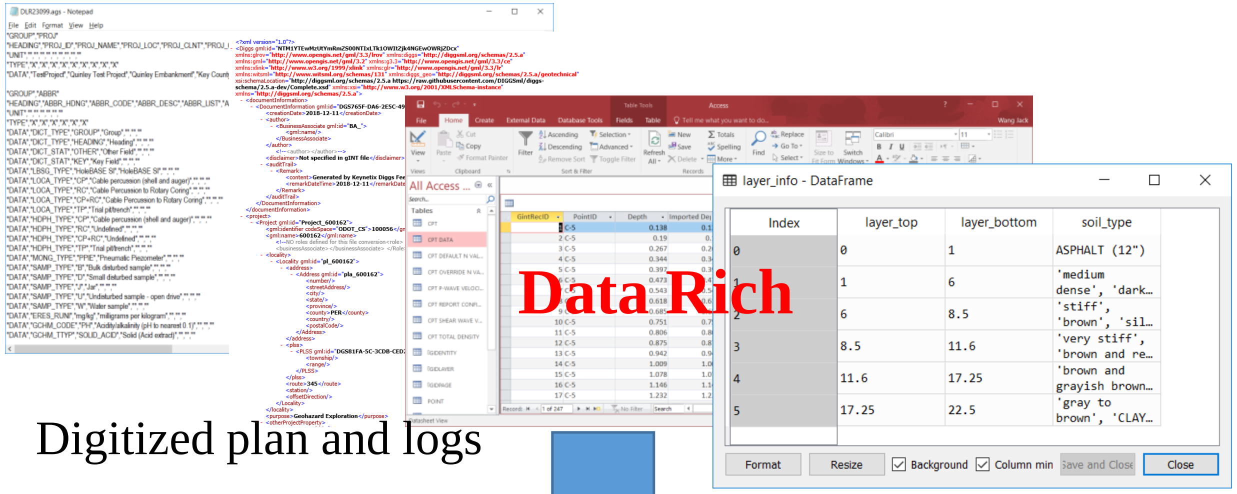


Design software

Engineering judgement

Human Design

Future trend



Digitized plan and logs

Learning Algorithm

Coding implementation

Engineering judgement

Knowledge Rich and Enabling AI Design

University of Dayton (UD) Team is working on data analytics

- Seamlessly integrated into the DIGGS ecosystem (read DIGGS compliant data as part of the data analysis software)

- Automatic geotechnical data interpretation and associated uncertainty quantification (AI: unsupervised learning – machine learning and pattern recognition)

- Stochastic simulation for unexplored locations based on extracted statistical characteristics and spatial correlation from the sampling locations (random field and stochastic simulation)

- Quantify "confidence level" of inferred geotechnical model for informed decision making (e.g., preliminary design and detailed reliability based design)

- Export enhanced DIGGS XML file with above derived data/information for better visualization (using GIS software, AR/VR mixed reality) and/or down stream design (CAD software)

Current UD developed AI based site investigation data interpretation and modeling platform

Module 1: DIGGS compliant data transfer interface

Output: Customized data structure that can be processed by the program

Module 2: Subsurface data fusion and interpretation

Output: Stratification and soil properties interpretations at sampling locations

Module 3: Stochastic subsurface geologic model simulation and uncertainty quantification

Output: Complete 2D/3D subsurface model and uncertainty quantification

Module 4: Downstream design and analysis applications

Output: Enhanced XML file with reliability/risk analysis and design recommendations

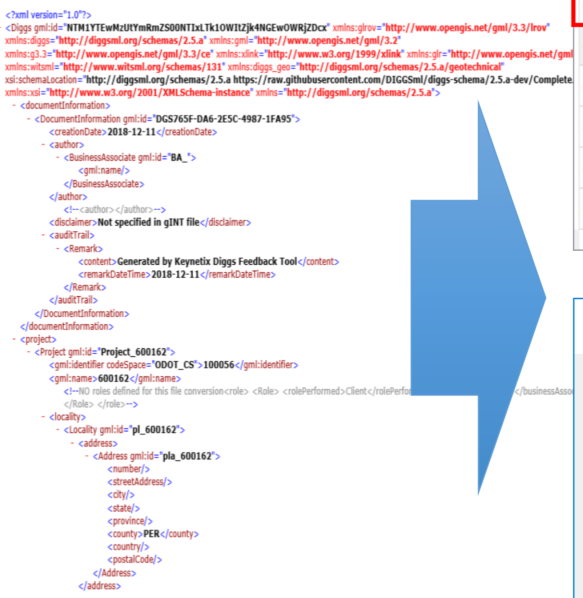
Current UD developed AI based site investigation data interpretation and modeling platform

Module 1: DIGGS compliant data transfer interface

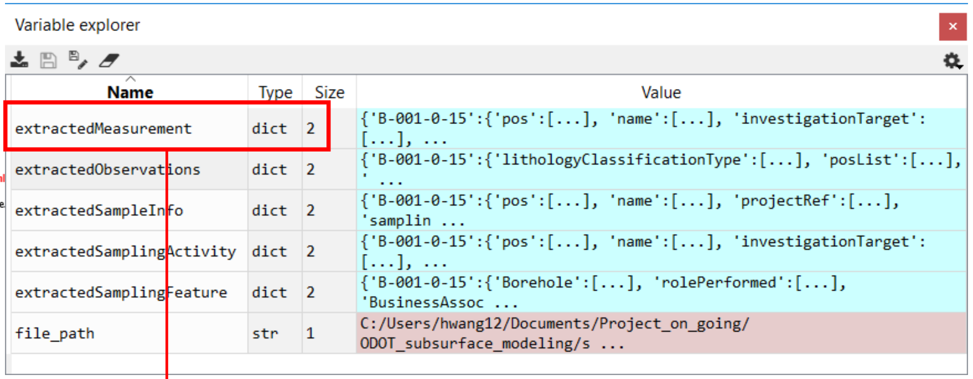
Output: Customized data structure that can be processed by the program

Reading data from DIGGS compliant xml file

XML text

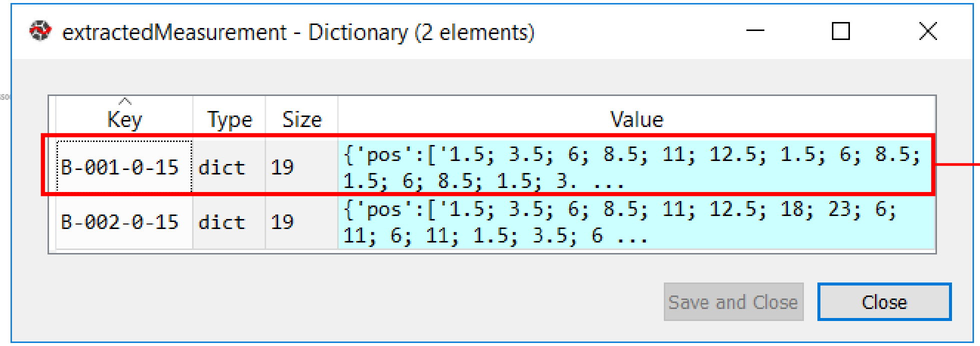


Python variables



Variable explorer

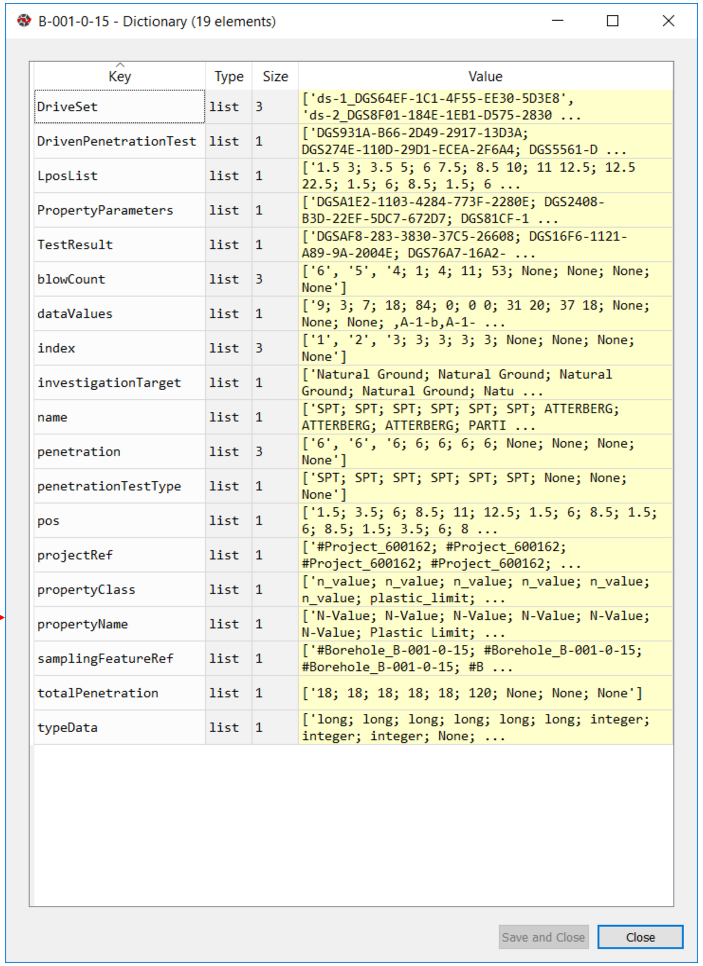
Name	Type	Size	Value
extractedMeasurement	dict	2	{'B-001-0-15':{'pos':[...], 'name':[...], 'investigationTarget':[...], ...
extractedObservations	dict	2	{'B-001-0-15':{'lithologyClassificationType':[...], 'posList':[...], ...
extractedSampleInfo	dict	2	{'B-001-0-15':{'pos':[...], 'name':[...], 'projectRef':[...], 'samplin ...
extractedSamplingActivity	dict	2	{'B-001-0-15':{'pos':[...], 'name':[...], 'investigationTarget':[...], ...
extractedSamplingFeature	dict	2	{'B-001-0-15':{'Borehole':[...], 'rolePerformed':[...], 'BusinessAssoc ...
file_path	str	1	C:/Users/hwang12/Documents/Project_on_going/ODOT_subsurface_modeling/s ...



extractedMeasurement - Dictionary (2 elements)

Key	Type	Size	Value
B-001-0-15	dict	19	{'pos':['1.5; 3.5; 6; 8.5; 11; 12.5; 1.5; 6; 8.5; 1.5; 6; 8.5; 1.5; 3. ...
B-002-0-15	dict	19	{'pos':['1.5; 3.5; 6; 8.5; 11; 12.5; 18; 23; 6; 11; 6; 11; 1.5; 3.5; 6 ...

Save and Close | Close



B-001-0-15 - Dictionary (19 elements)

Key	Type	Size	Value
DriveSet	list	3	['ds-1_DGS64EF-1C1-4F55-EE30-5D3E8', 'ds-2_DGS8F01-184E-1EB1-D575-2830 ...
DrivenPenetrationTest	list	1	['DGS931A-B66-2D49-2917-13D3A; DGS274E-110D-29D1-ECEA-2F6A4; DGS5561-D ...
LposList	list	1	['1.5 3; 3.5 5; 6 7.5; 8.5 10; 11 12.5; 12.5 22.5; 1.5; 6; 8.5; 1.5; 6 ...
PropertyParameters	list	1	['DGSA1E2-1103-4284-773F-2280E; DGS2408-B3D-22EF-5DC7-672D7; DGS81CF-1 ...
TestResult	list	1	['DGSAF8-283-3830-37C5-26608; DGS16F6-1121-A89-9A-2004E; DGS76A7-16A2- ...
blowCount	list	3	['6', '5', '4; 1; 4; 11; 53; None; None; None; None']
dataValues	list	1	['9; 3; 7; 18; 84; 0; 0 0; 31 20; 37 18; None; None; None; ,A-1-b,A-1- ...
index	list	3	['1', '2', '3; 3; 3; 3; 3; None; None; None; None']
investigationTarget	list	1	['Natural Ground; Natural Ground; Natural Ground; Natural Ground; Natu ...
name	list	1	['SPT; SPT; SPT; SPT; SPT; SPT; ATTERBERG; ATTERBERG; ATTERBERG; PARTI ...
penetration	list	3	['6', '6', '6; 6; 6; 6; 6; None; None; None; None']
penetrationTestType	list	1	['SPT; SPT; SPT; SPT; SPT; SPT; None; None; None']
pos	list	1	['1.5; 3.5; 6; 8.5; 11; 12.5; 1.5; 6; 8.5; 1.5; 6; 8.5; 1.5; 3.5; 6; 8 ...
projectRef	list	1	['#Project_600162; #Project_600162; #Project_600162; #Project_600162; ...
propertyClass	list	1	['n_value; n_value; n_value; n_value; n_value; n_value; plastic_limit; ...
propertyName	list	1	['N-Value; N-Value; N-Value; N-Value; N-Value; N-Value; Plastic Limit; ...
samplingFeatureRef	list	1	['#Borehole_B-001-0-15; #Borehole_B-001-0-15; #Borehole_B-001-0-15; #B ...
totalPenetration	list	1	['18; 18; 18; 18; 18; 120; None; None; None']
typeData	list	1	['long; long; long; long; long; long; integer; integer; integer; None; ...

Save and Close | Close

Running time: 0.044s for this example file

DIGGS xml file can be parsed and transferred into our customized data structures efficiently!

Current UD developed AI based site investigation data interpretation and modeling platform

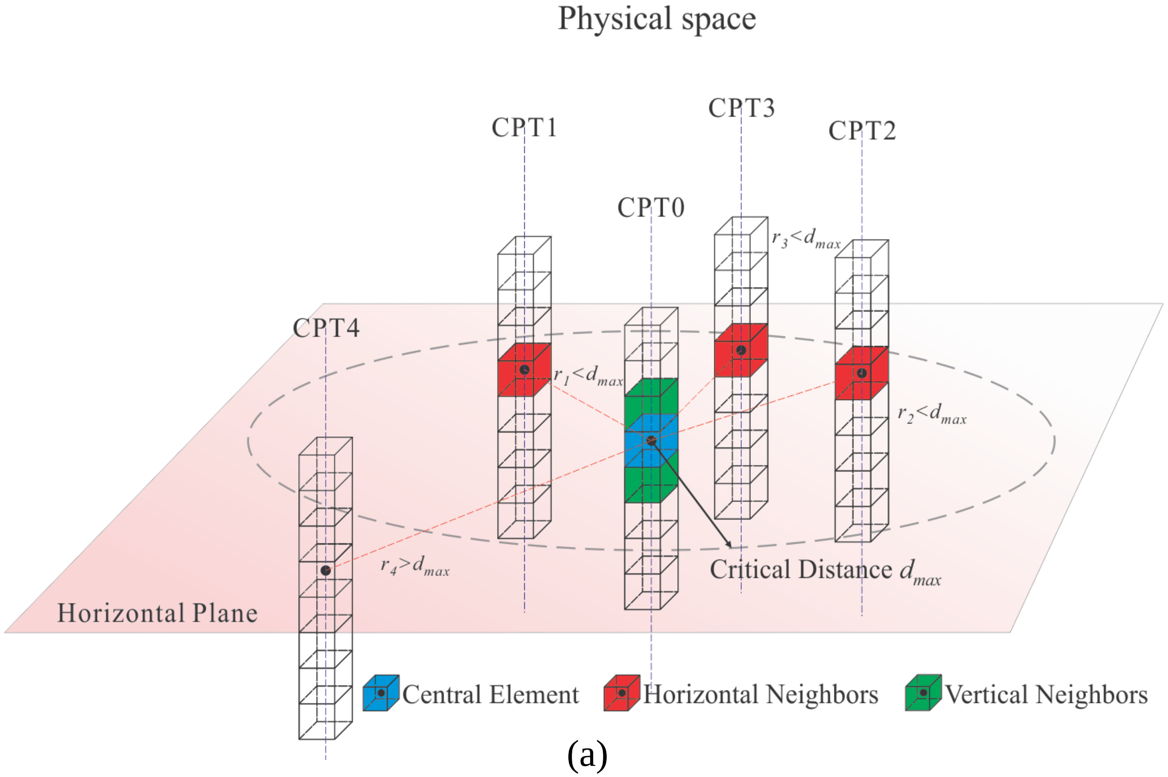
Module 1: DIGGS compliant data transfer interface

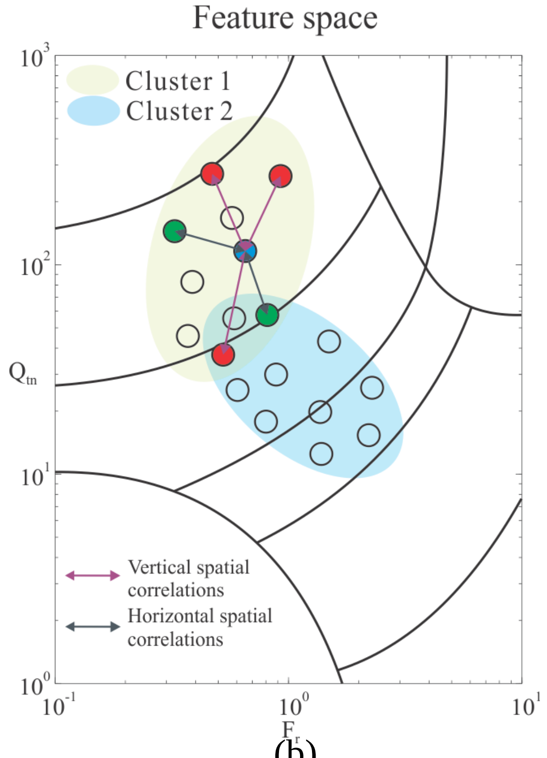
Output: Customized data structure that can be processed by the program

Module 2: Subsurface data fusion and interpretation

Output: Stratification and soil properties interpretations at sampling locations

Joint interpreting multiple CPT sounding data





(a) Neighborhood system of soil elements in the physical space;

(b) associated CPT sounding points subjected to spatial constraints in the feature space.

Real-world example

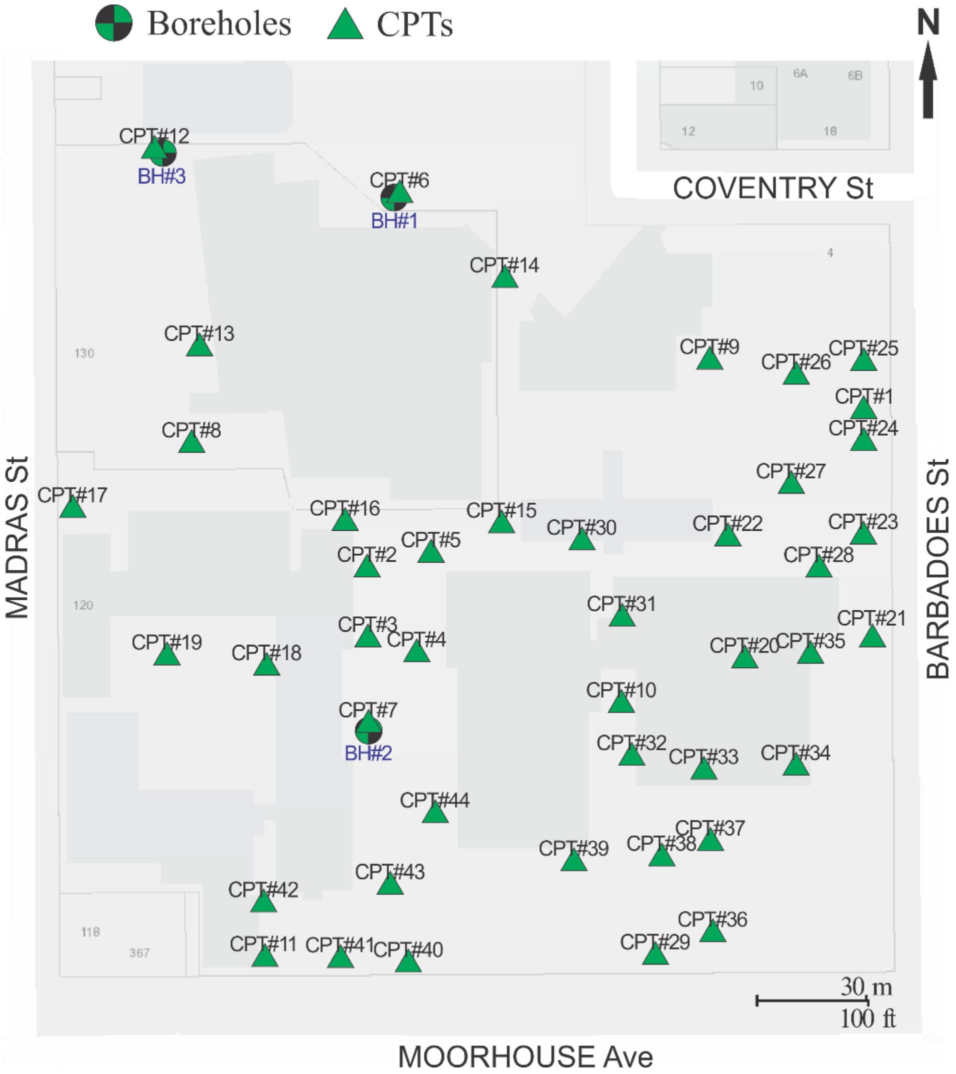
A real-world CPT dataset collected at a project site located within the central business district of the city of Christchurch, New Zealand.

44 CPT soundings and 3 boreholes are sparsely located in a 240 m × 240 m square region.

All 44 CPT soundings are interpreted simultaneously

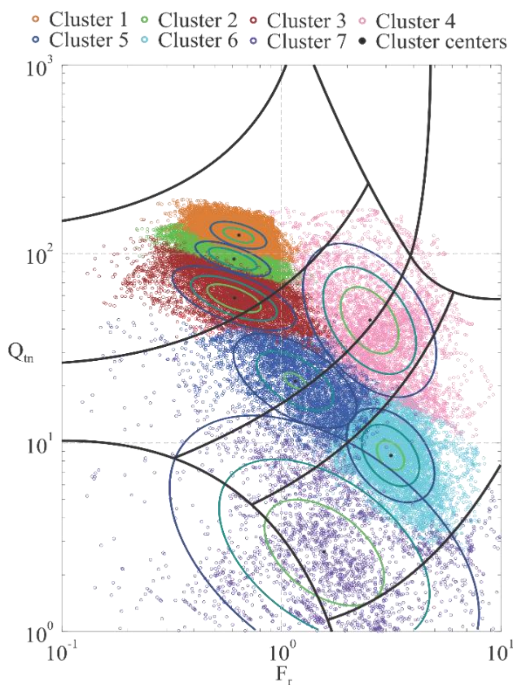
Two validation cases:

1) CPT #6, #7, #12 (Borehole logs validation)
2) CPT #1, #24 (Shortest horizontal distance)

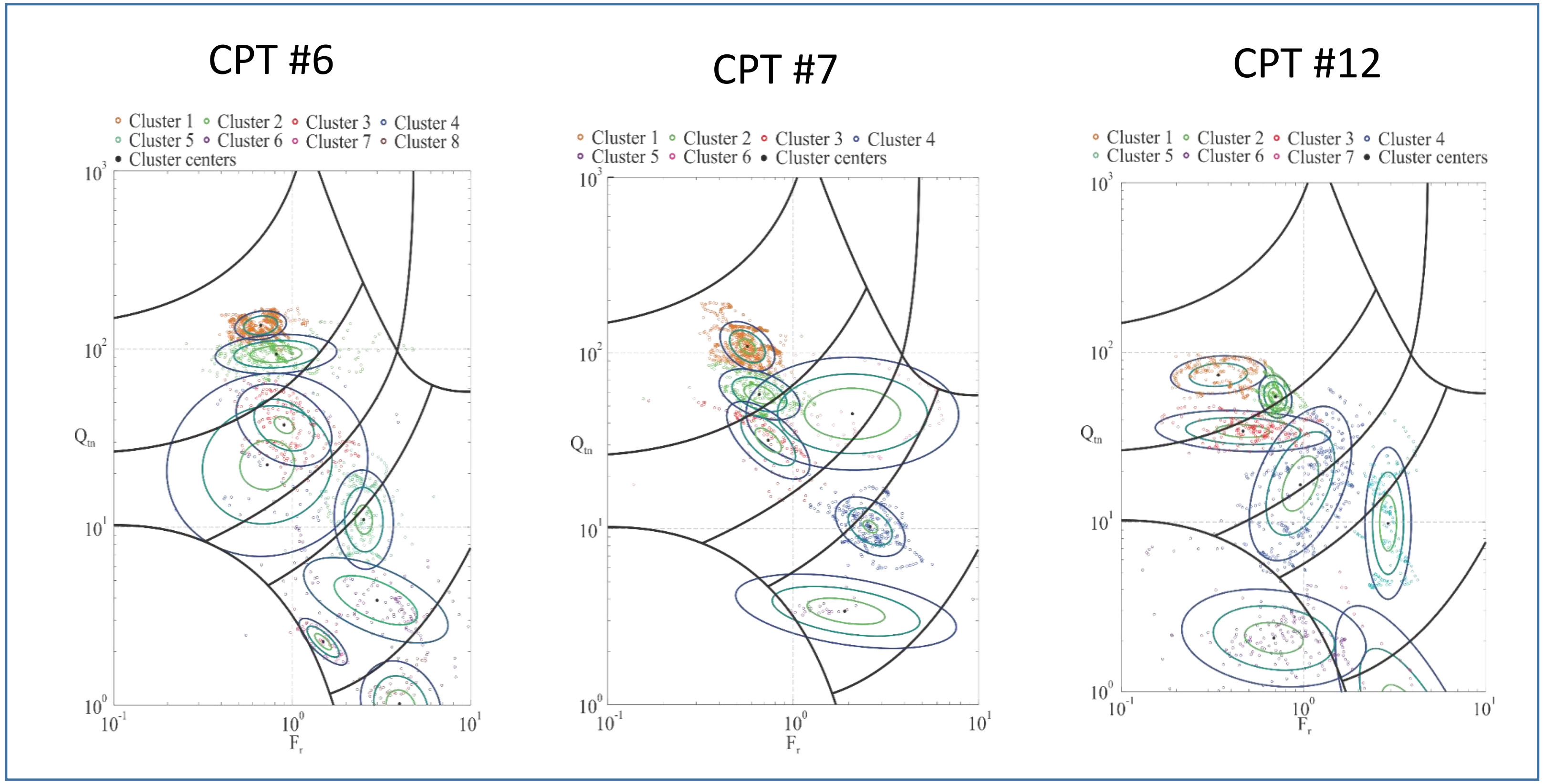


CPT and borehole locations

More data points -> Enhanced clustered pattern



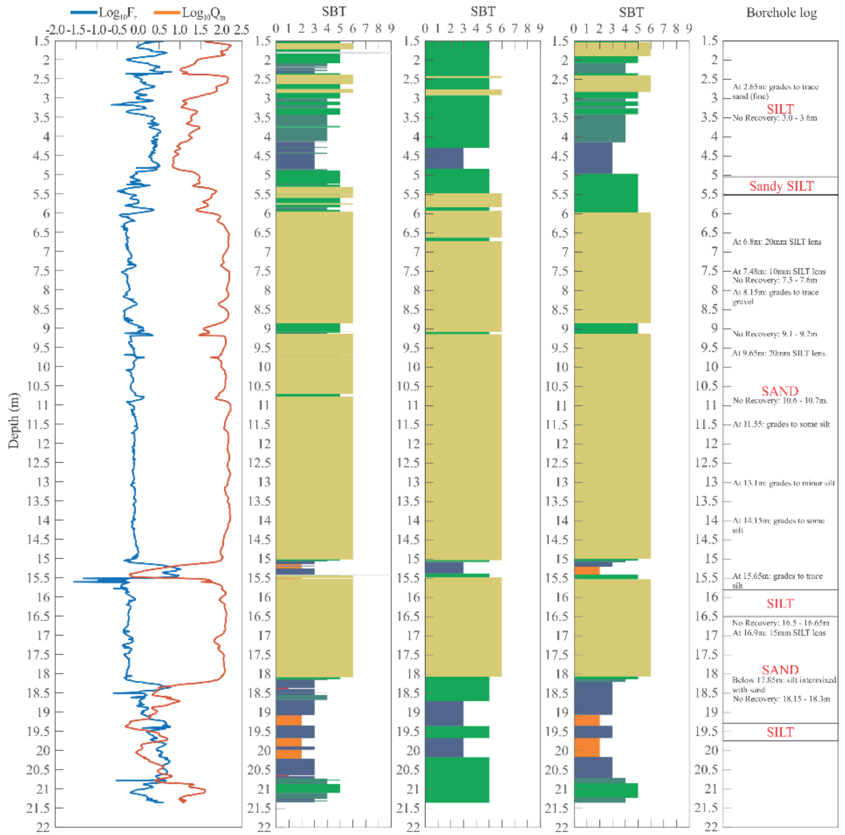
Statistical pattern in Robertson chart from joint interpretation of 44 CPT soundings



Statistical pattern from separate interpretation of three CPT soundings

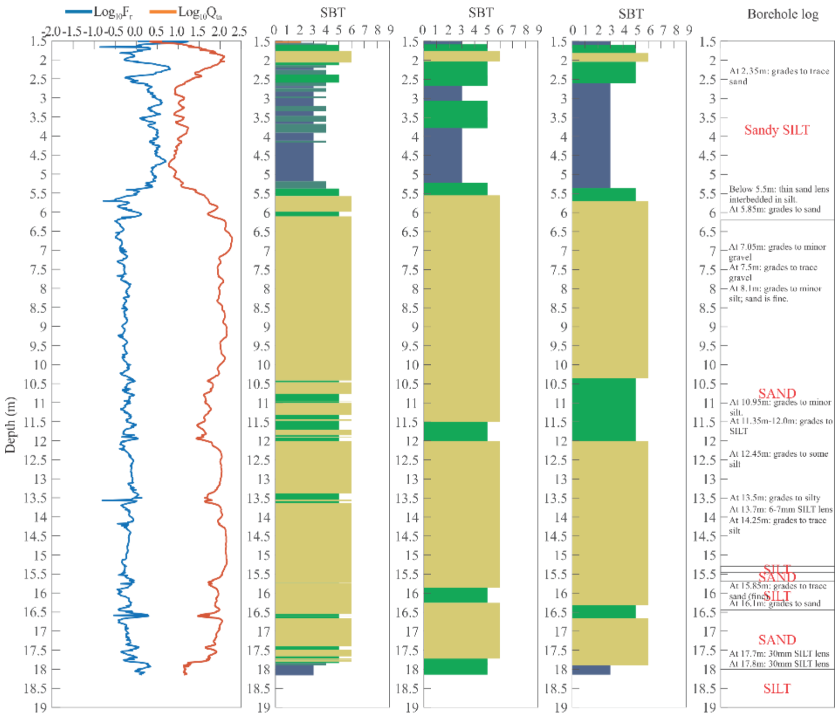
Enhanced clustered pattern -> Enhanced vertical consistency

CPT #6



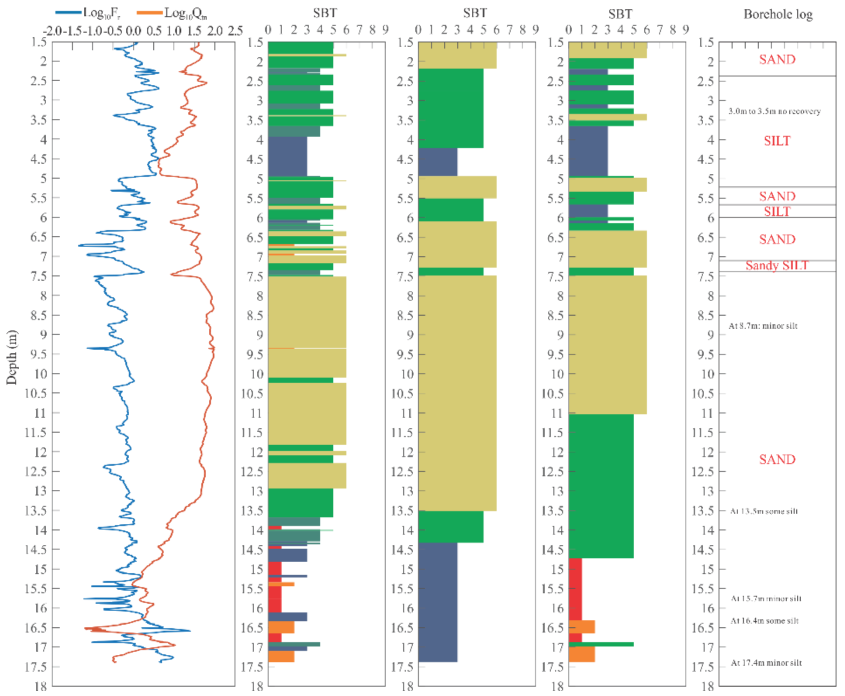
SBT Chart | ML-based jointly | ML-based Separately

CPT #7



SBT Chart | ML-based jointly | ML-based Separately

CPT #12

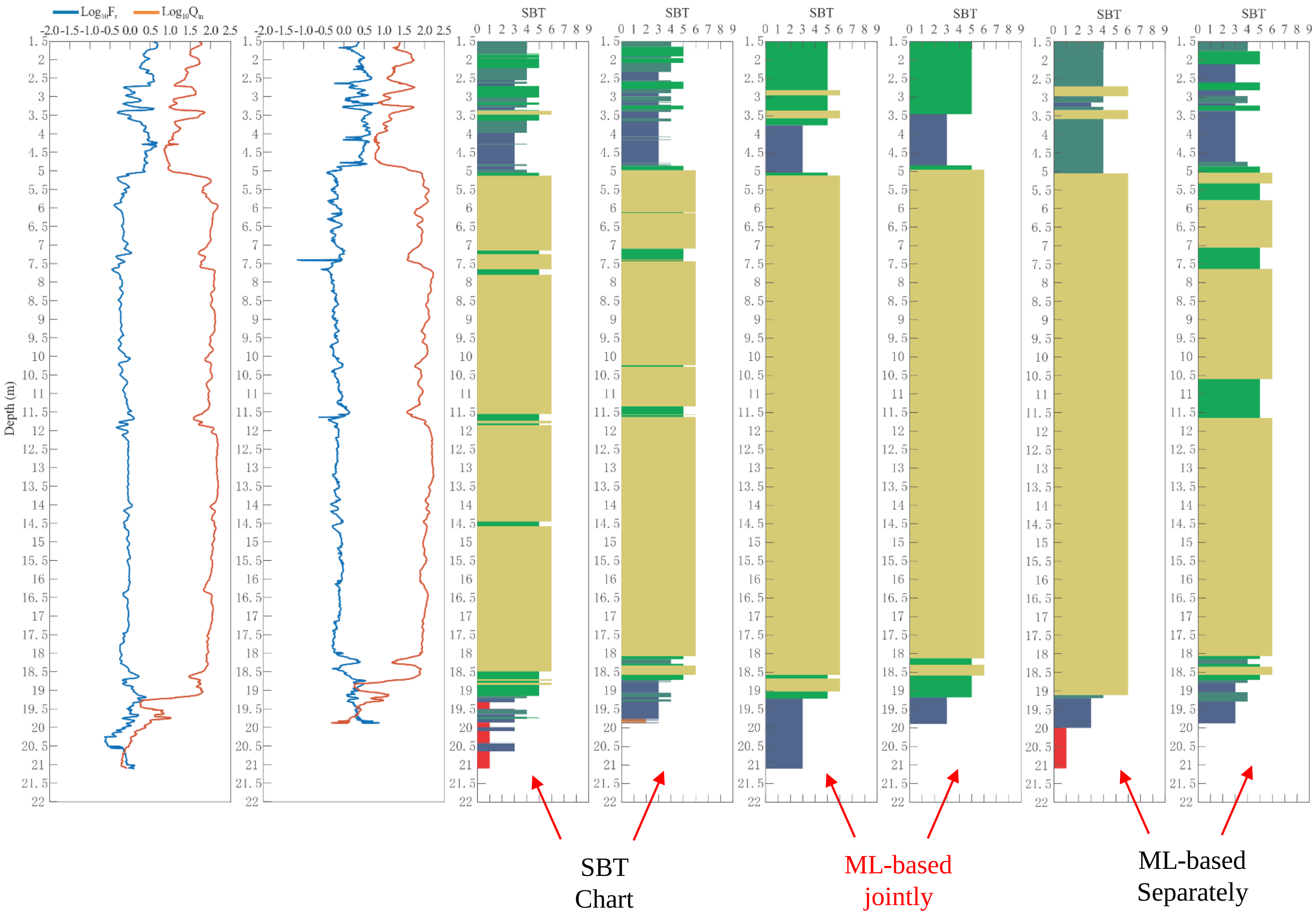


SBT Chart | ML-based jointly | ML-based Separately

Machine learning-based interpretation can take the vertical correlation of the soil physical properties into consideration, and thereby improving the vertical consistency of the interpretation results.

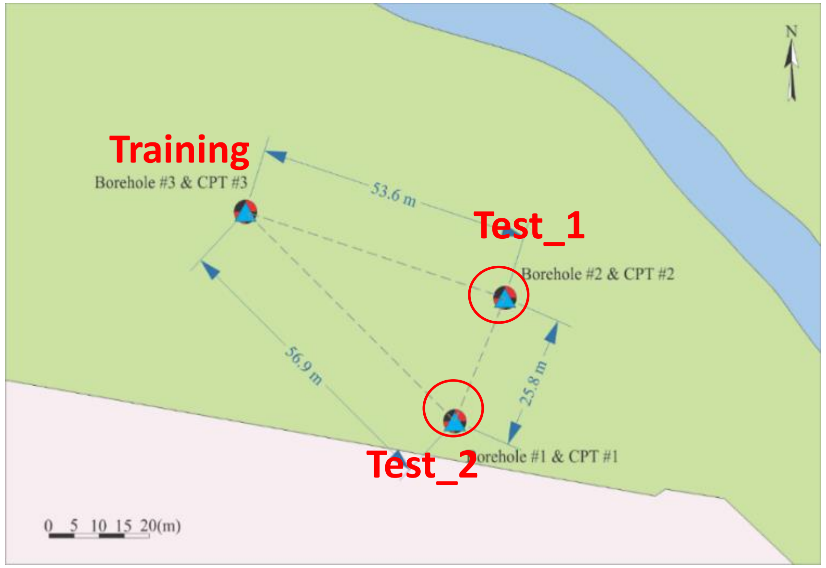
Enhanced clustered pattern -> Enhanced horizontal consistency

CPT #1, #24 (Shortest horizontal distance)

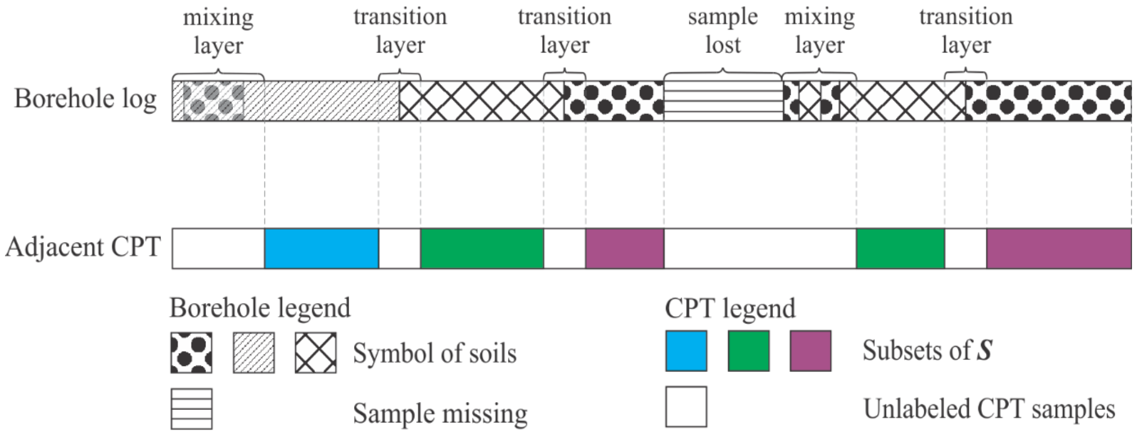


Joint interpretation can eliminate the undesired conflicts among stratification results of nearby CPT records and significantly improve the horizontal interpretation consistency

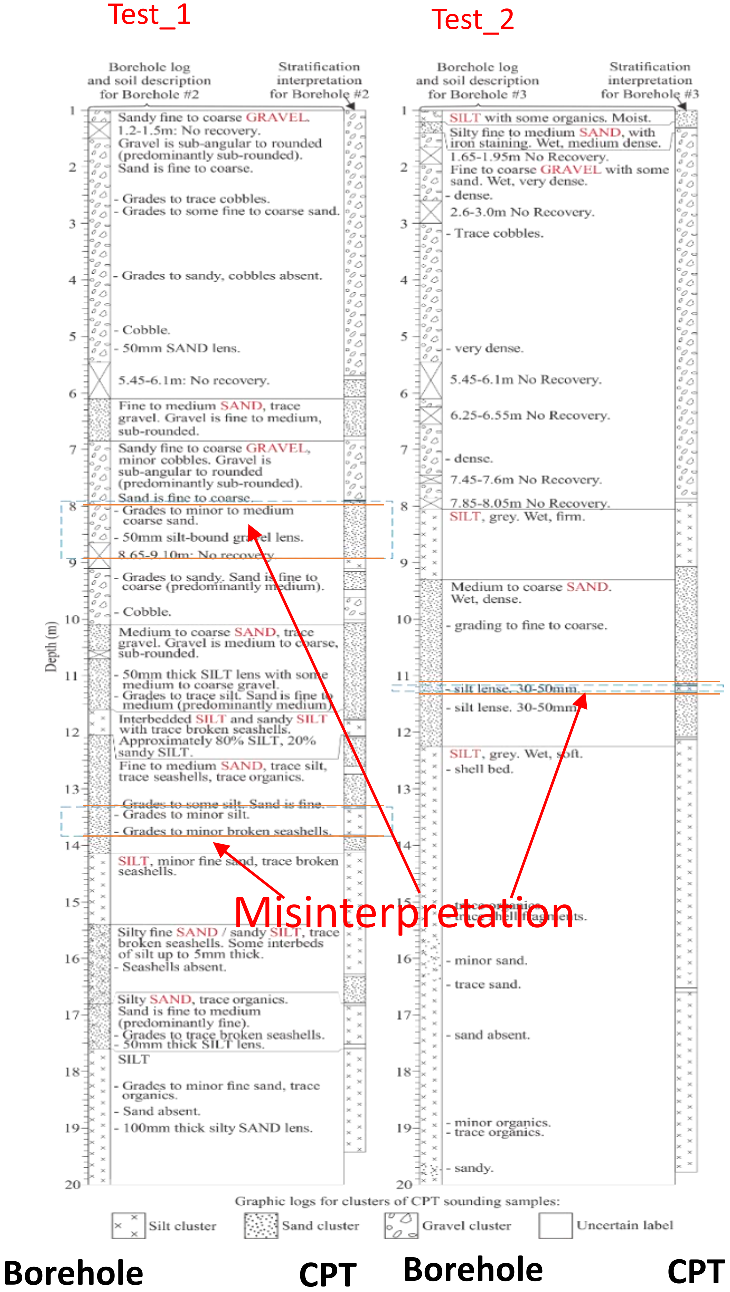
Joint interpretation of borehole logs and CPTs



Boreholes and CPT locations



Schematic diagram for extracting labeled samples from borehole logs



Current UD developed AI based site investigation data interpretation and modeling platform

Module 1: DIGGS compliant data transfer interface

Output: Customized data structure that can be processed by the program

Module 2: Subsurface data fusion and interpretation

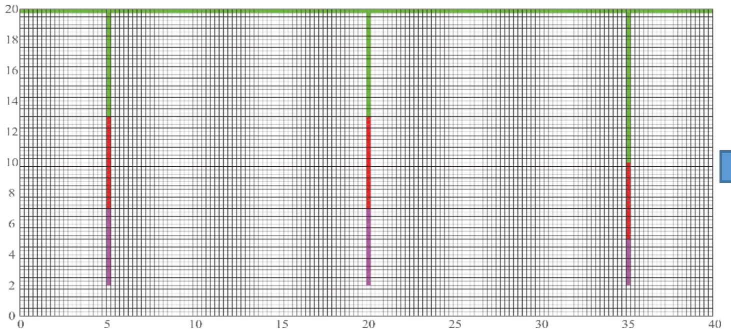
Output: Stratification and soil properties interpretations at sampling locations

Module 3: Stochastic subsurface geologic model simulation and uncertainty quantification

Output: Complete 2D/3D subsurface model and uncertainty quantification

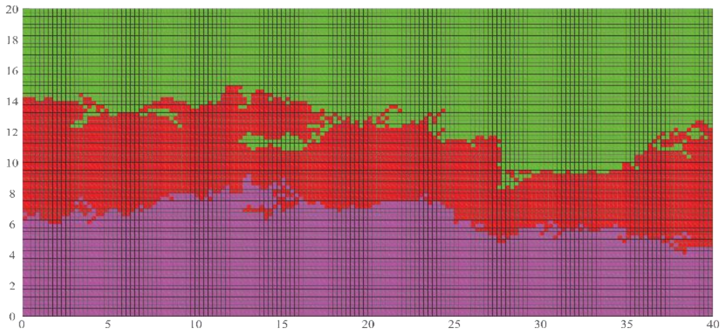
Subsurface geologic model simulation process

1) Random field discretization and sparse stratification data integration



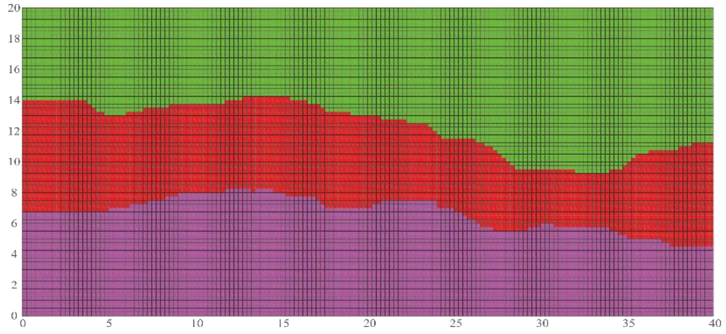
2) Generating stratigraphic realizations

a) Initial configuration



Local optimization

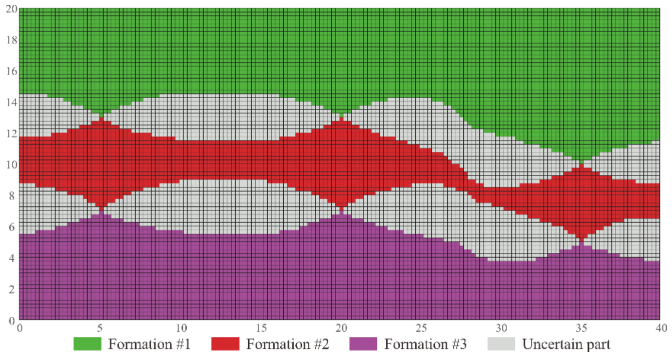
b) Converged MRF– one simulation outcome



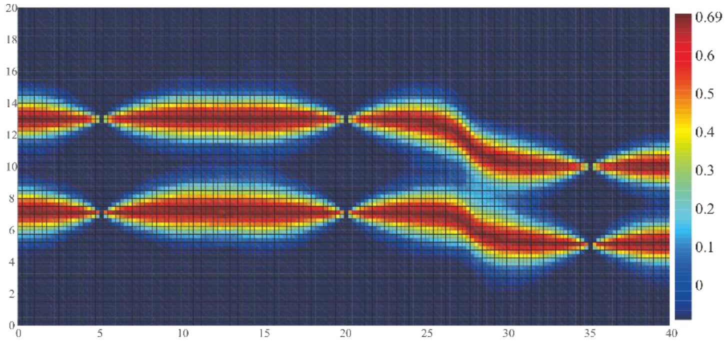
Sufficient number of stratigraphic realizations

3) Uncertainty quantification and visualization

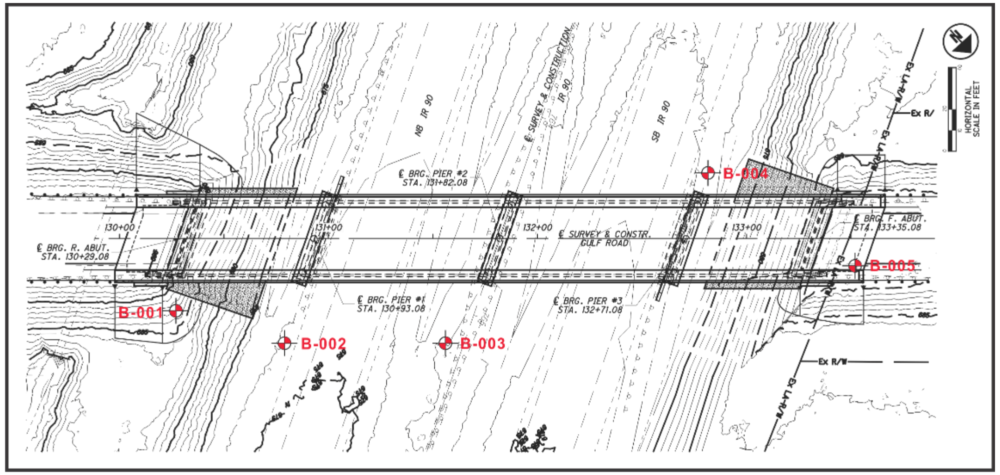
Confidence assignments with 95% confidence level



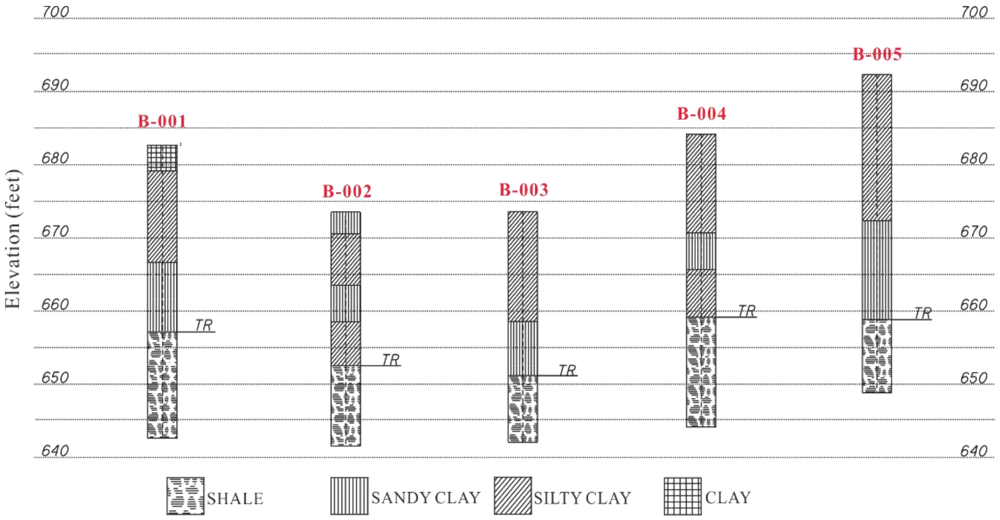
Map of information entropy



2D stratification profile modeling



Available boreholes in the project site



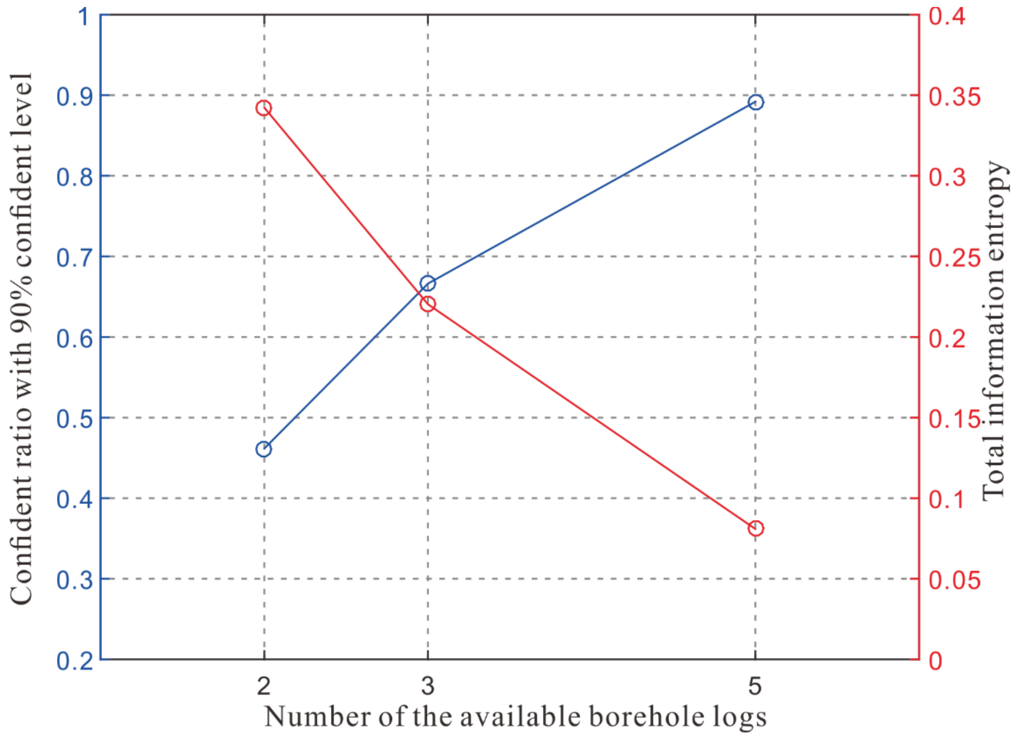
Borehole logs for the two-dimensional project

Three scenarios with different combinations of borehole log data as inputs:

Case 1: Borehole #1 + Borehole #5

Case 2: Borehole #1 + Borehole #3 + Borehole #5

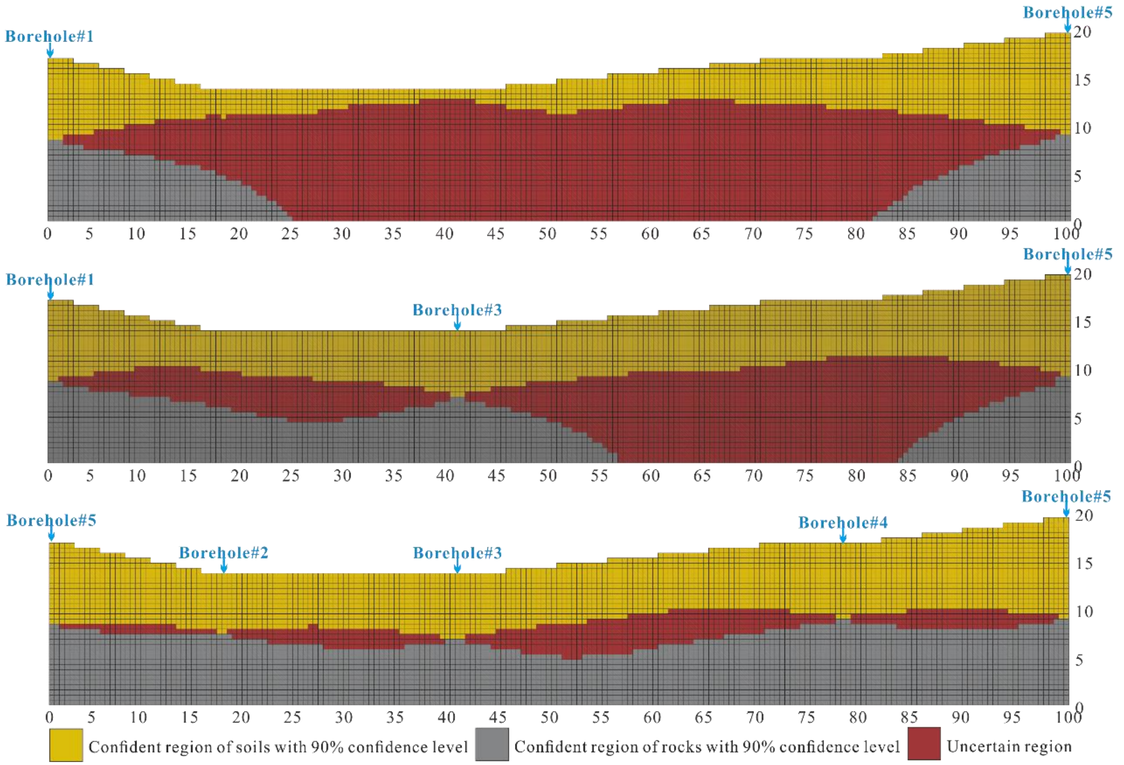
Case 3: Borehole #1 through Borehole #5.



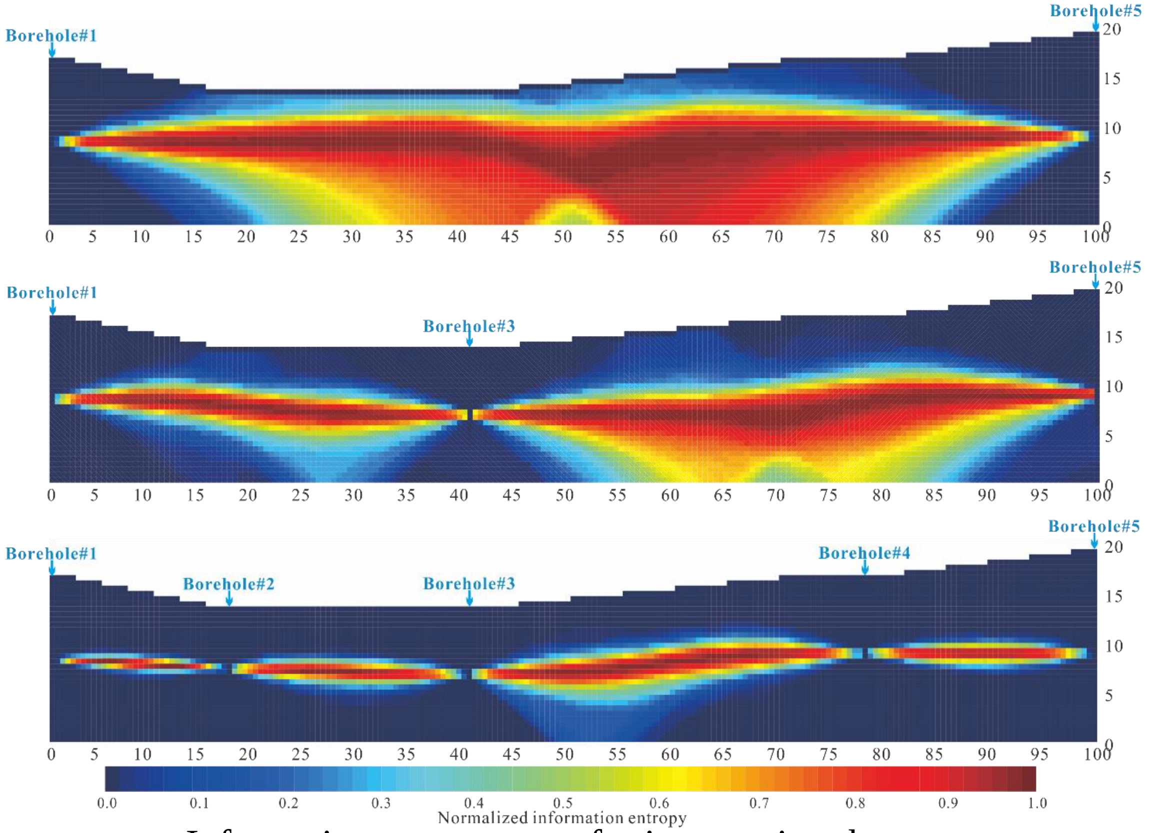
Confidence ratios and average information entropy values of different modeling cases for the two-dimensional project

2D stratification profile modeling

Simulation of the soil-rock interface that is critical for the design of foundation system

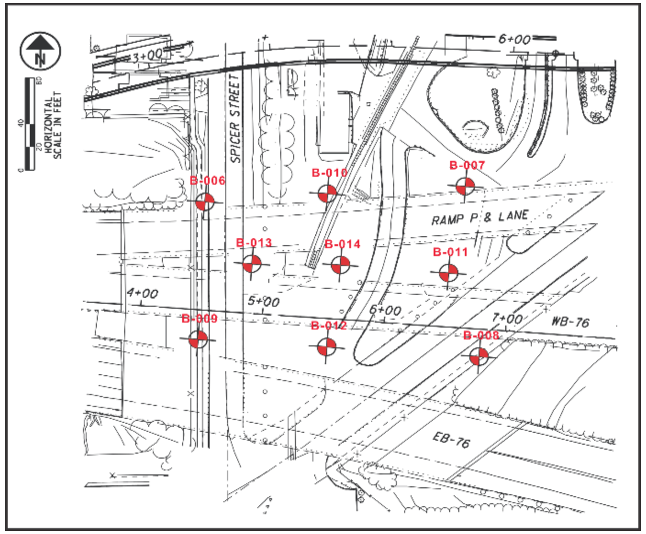


Map of confidence assignments for interpreting the two-dimensional project.

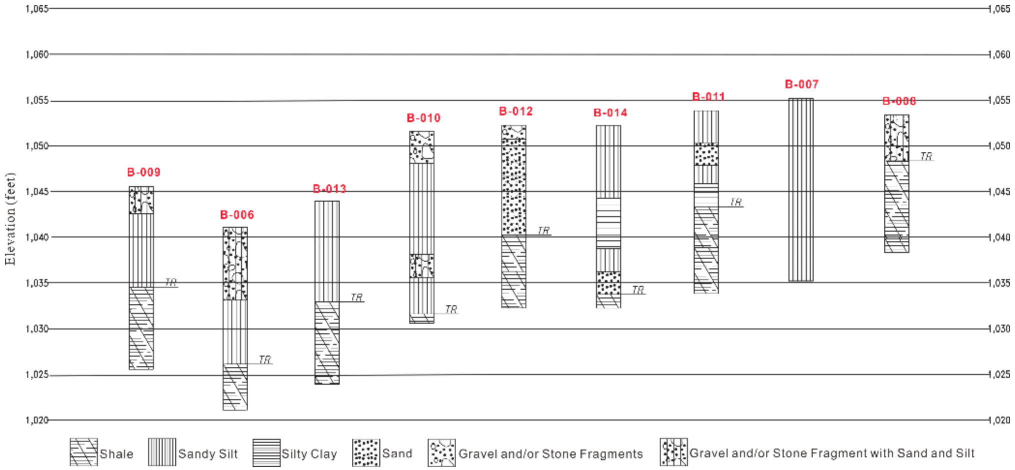


Information entropy map for interpreting the two-dimensional project

3D stratification profile modeling



Available boreholes in the project site



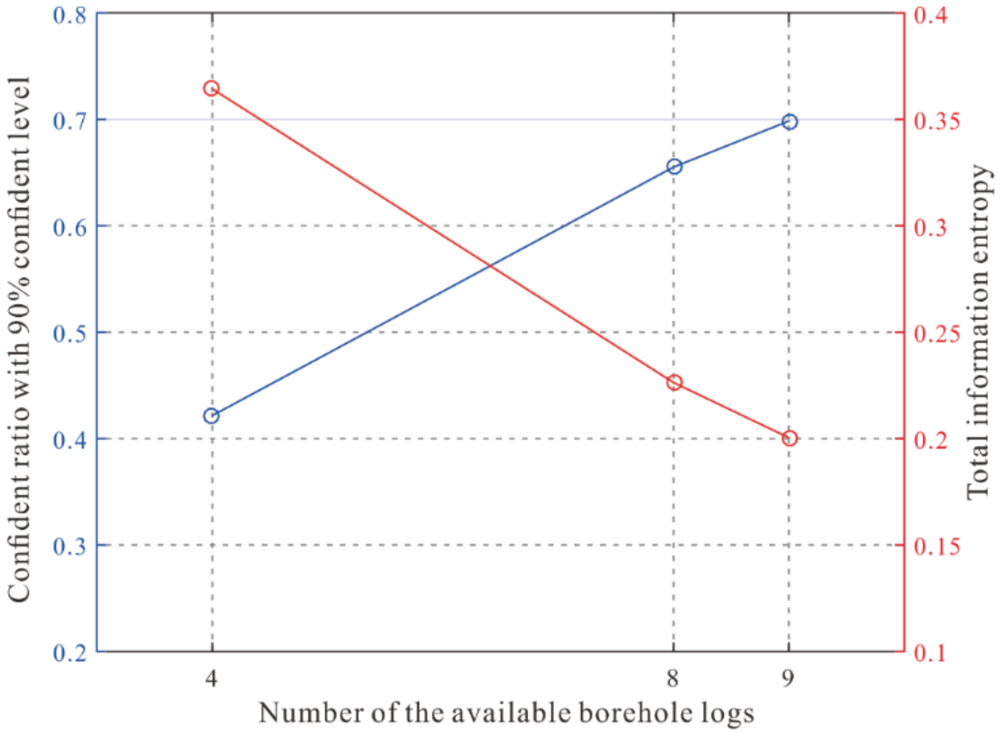
Borehole logs for the three-dimensional project

Three scenarios with different combinations of borehole log data as inputs:

Case 1: Borehole #6 through Borehole #9

Case 2: Borehole #6 through Borehole #13

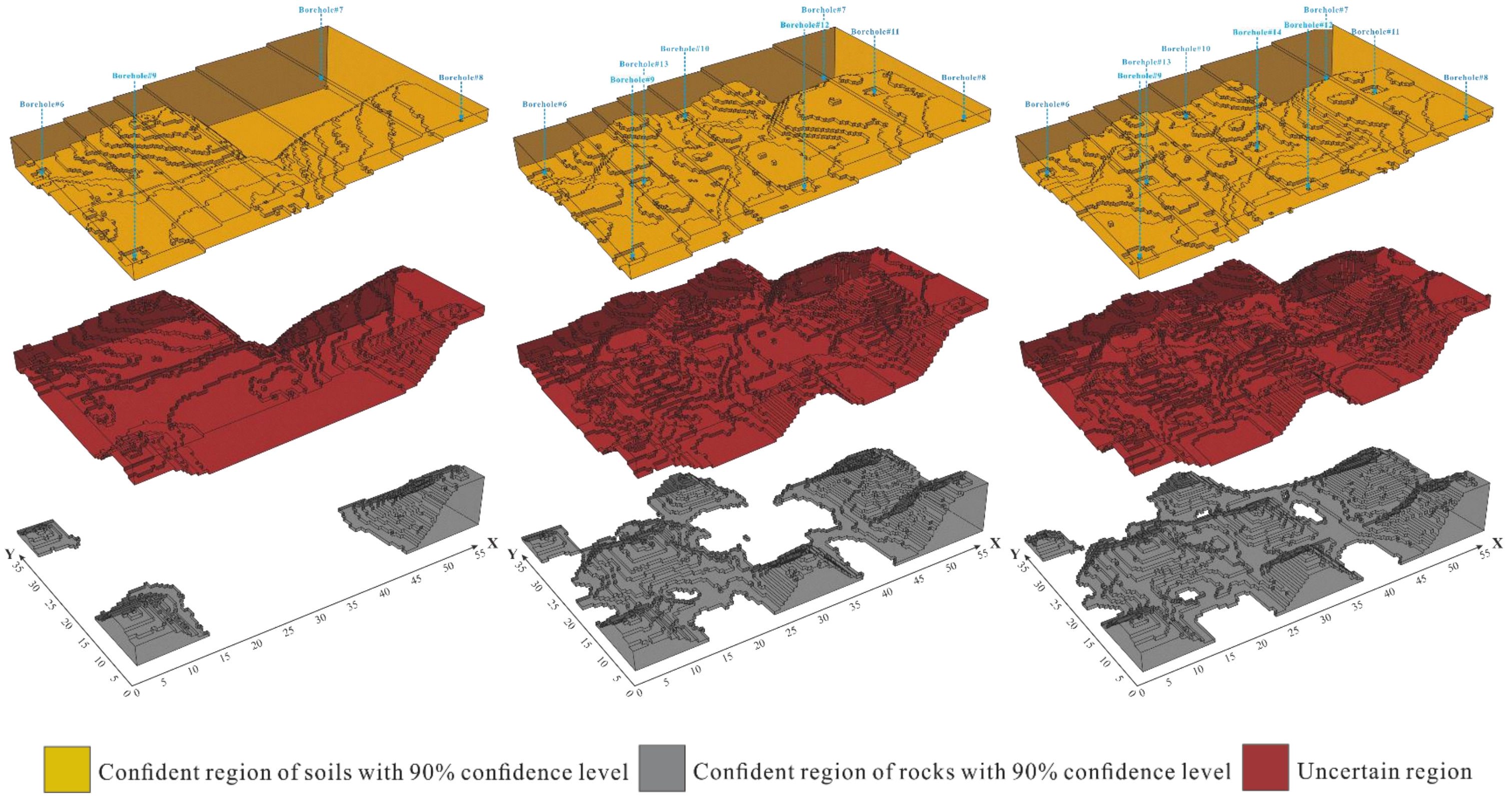
Case 3: Borehole #1 through Borehole #14



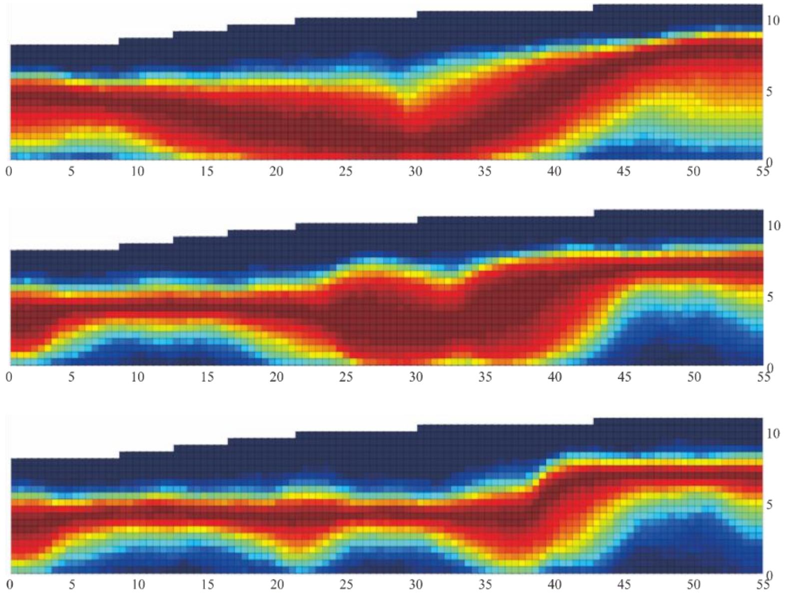
Confident ratios and average information entropy values of different modeling cases for interpreting the three-dimensional project

3D stratification profile modeling

Simulation of the soil-rock interface



Map of confidence assignments for interpreting the three-dimensional project.



Subsurface uncertainty reduction in a vertical section

Current UD developed AI based site investigation data interpretation and modeling platform

Module 1: DIGGS compliant data transfer interface

Output: Customized data structure that can be processed by the program

↓

Module 2: Subsurface data fusion and interpretation

Output: Stratification and soil properties interpretations at sampling locations

↓

Module 3: Stochastic subsurface geologic model simulation and uncertainty quantification

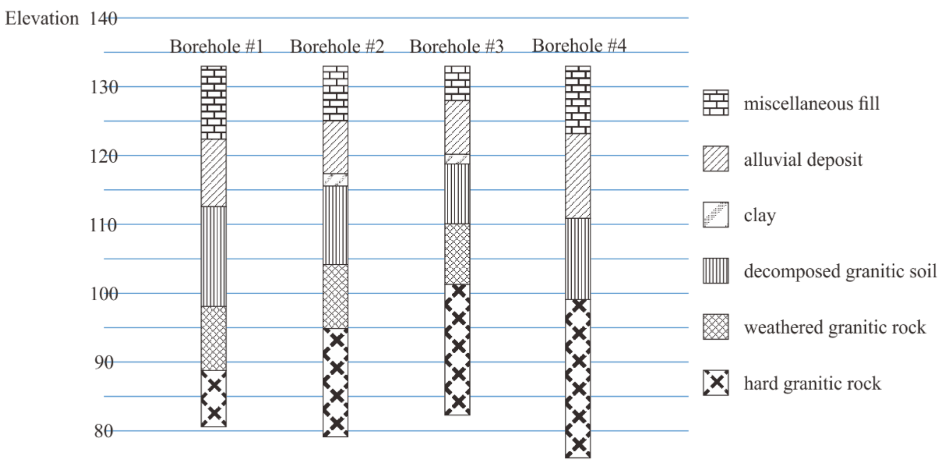
Output: Complete 2D/3D subsurface model and uncertainty quantification

↓

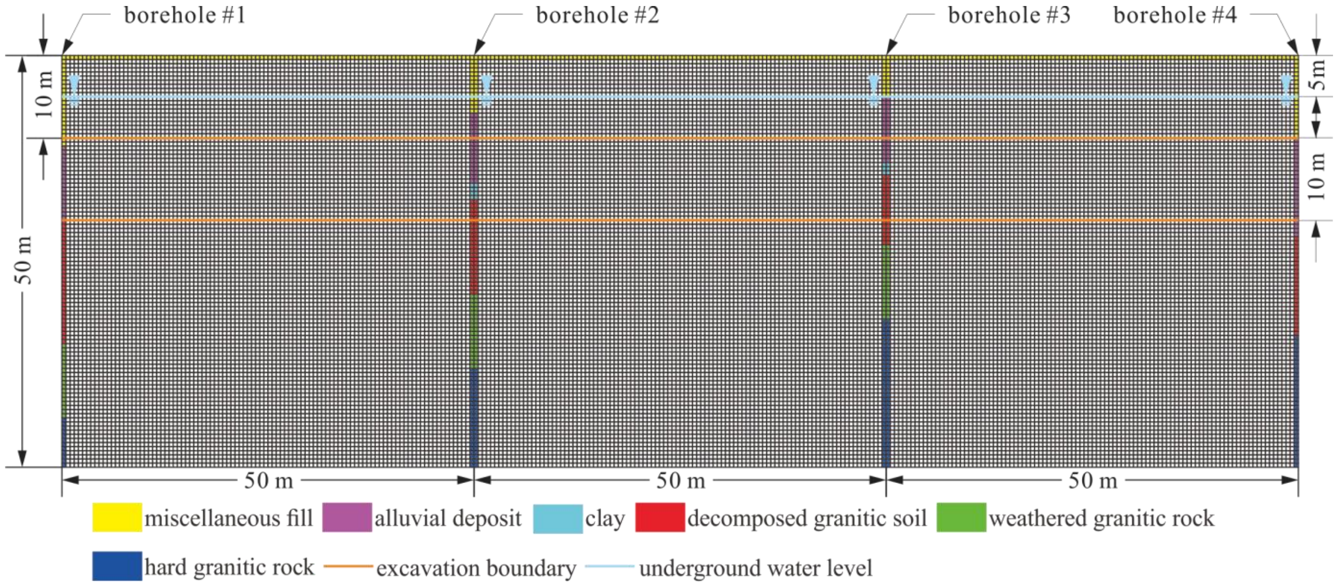
Module 4: Downstream design and analysis applications

Output: Enhanced XML file with reliability/risk analysis and design recommendations

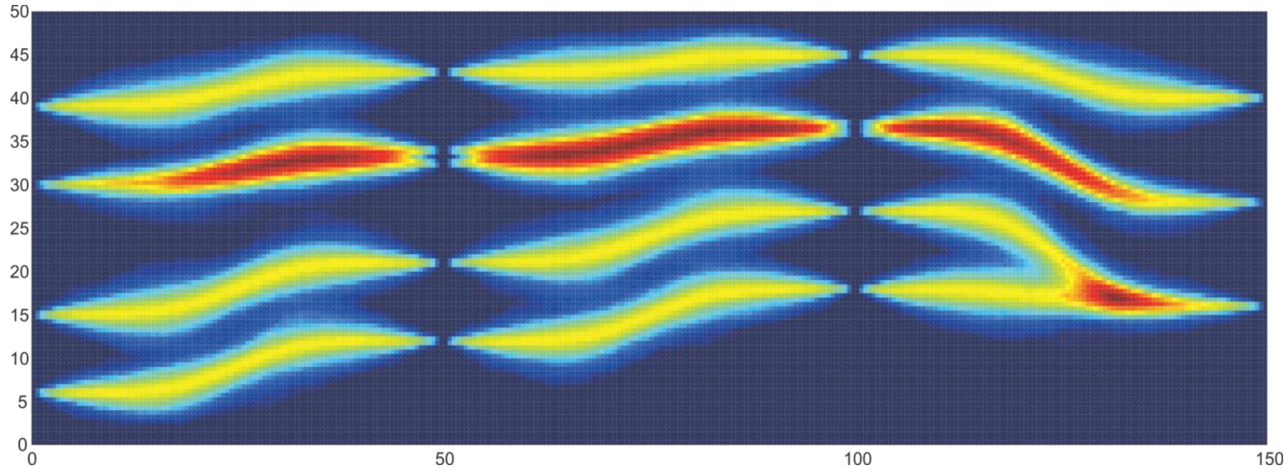
Example 1: Probabilistic analysis of Shield Tunnel in Multiple Strata



Available borehole logs

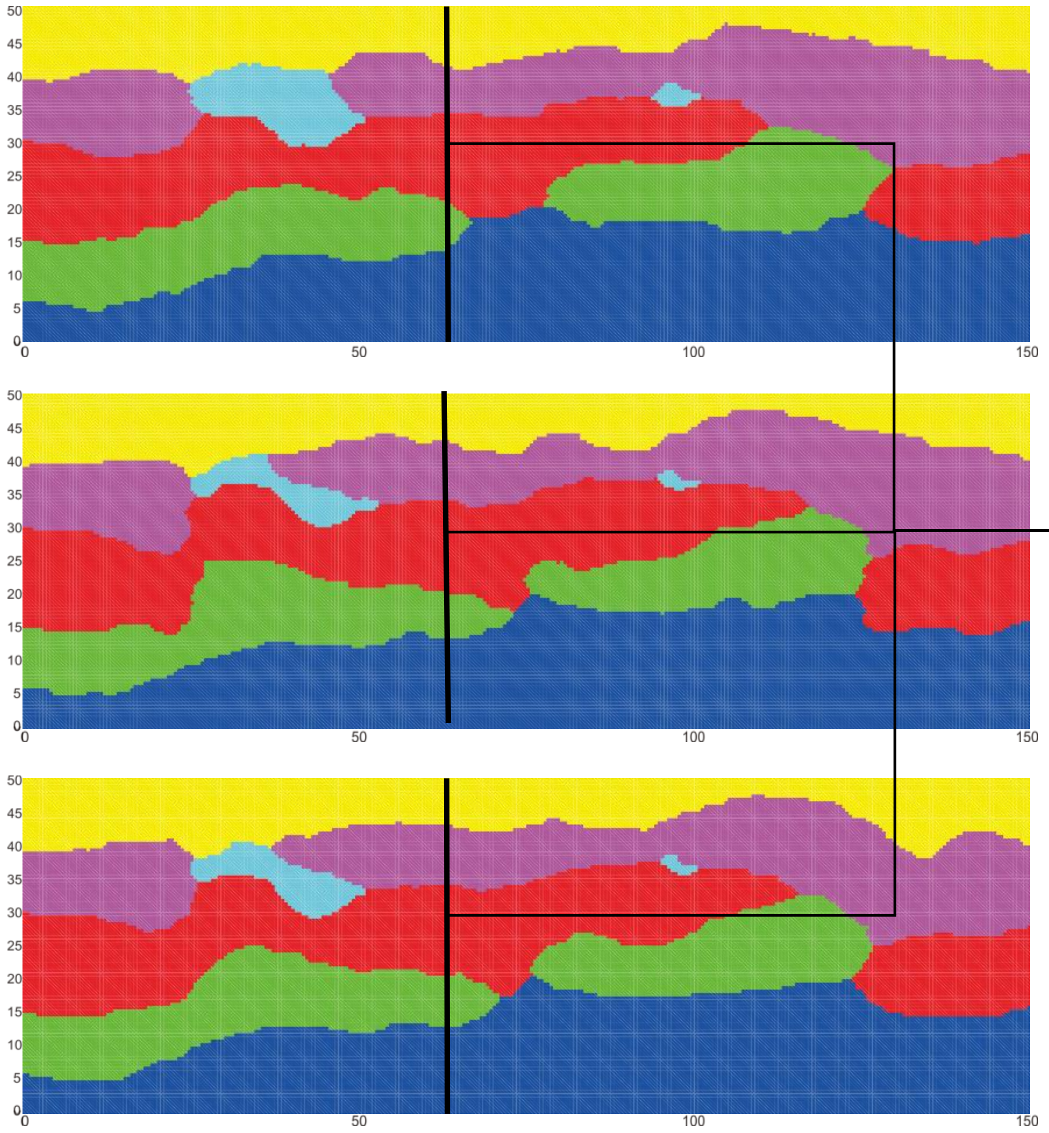


Profile of tunnel alignment and locations of borehole logs



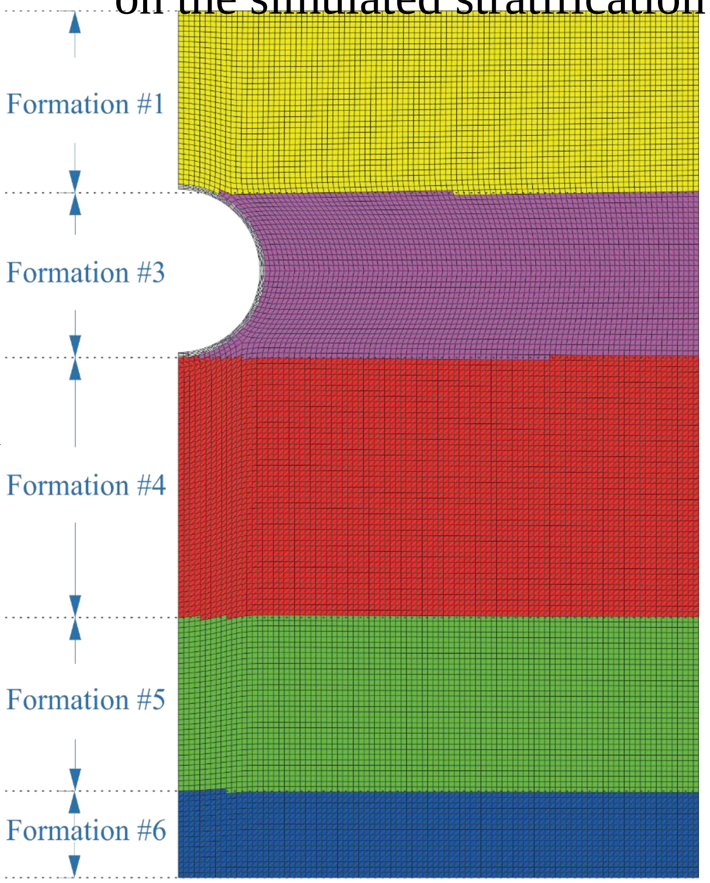
Subsurface uncertainty quantified using information entropy for the tunnel alignment

Probabilistic Analysis of Shield Tunnel in Multiple Strata



Generated stratification realizations along the tunnel alignment

Stochastic Finite Element model can be built based on the simulated stratification profile.

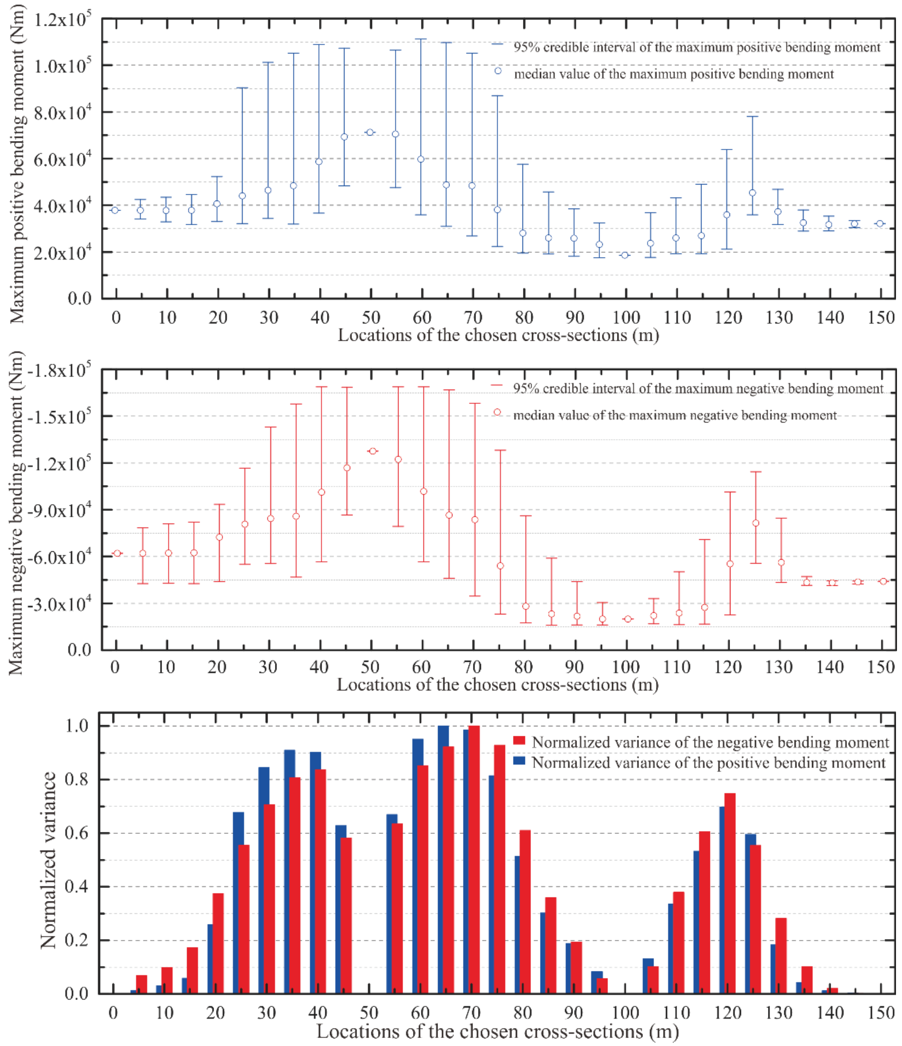


FEA model for tunnel cross-sections built based on the generated realizations

Probabilistic Analysis of Shield Tunnel in Multiple Strata

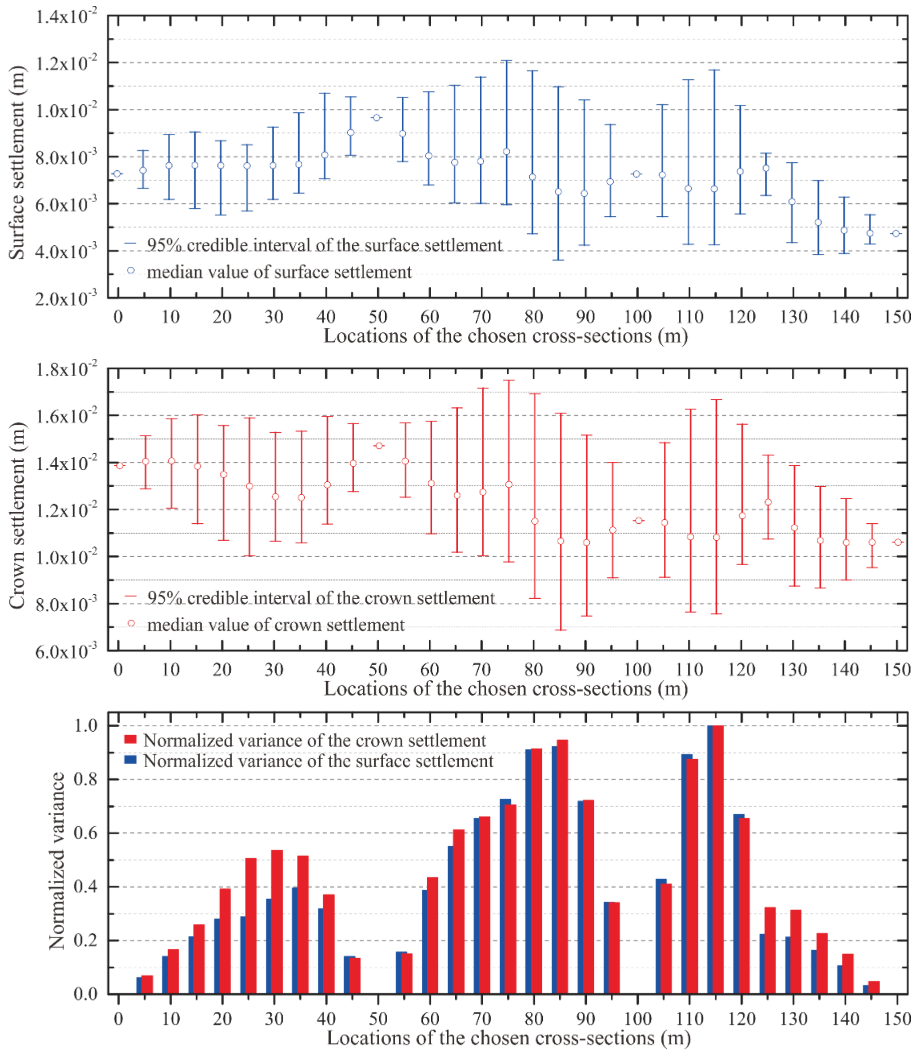
Stochastic Finite Element simulation provides probability/reliability based analysis results

(a)
(b)
(c)



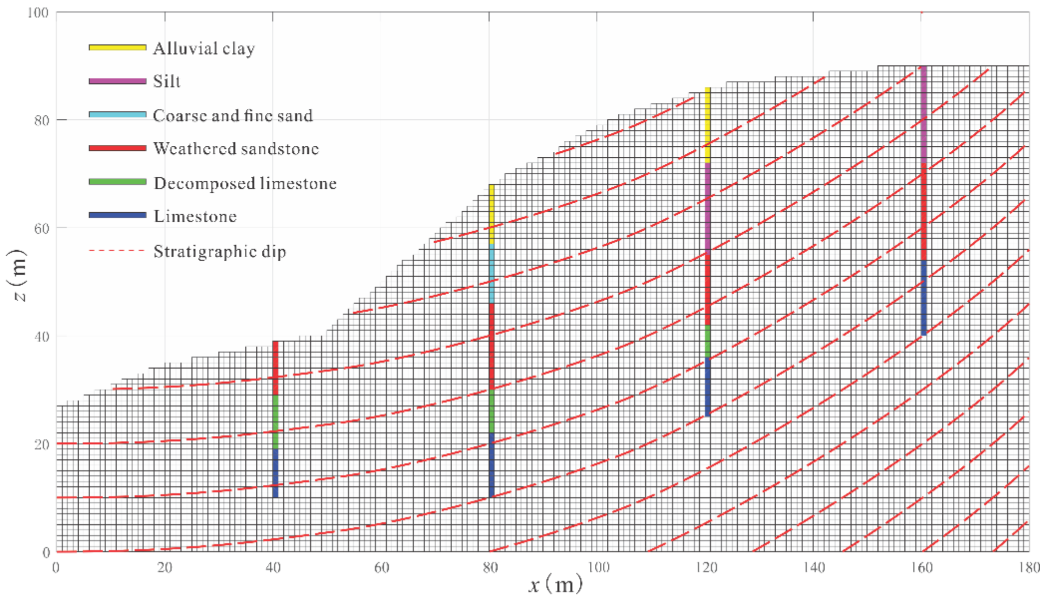
(a) 95% credible intervals of maximum positive bending moment;
(b) 95% credible intervals of maximum negative bending moment;
(c) normalized variance of bending moments.

(a)
(b)
(c)

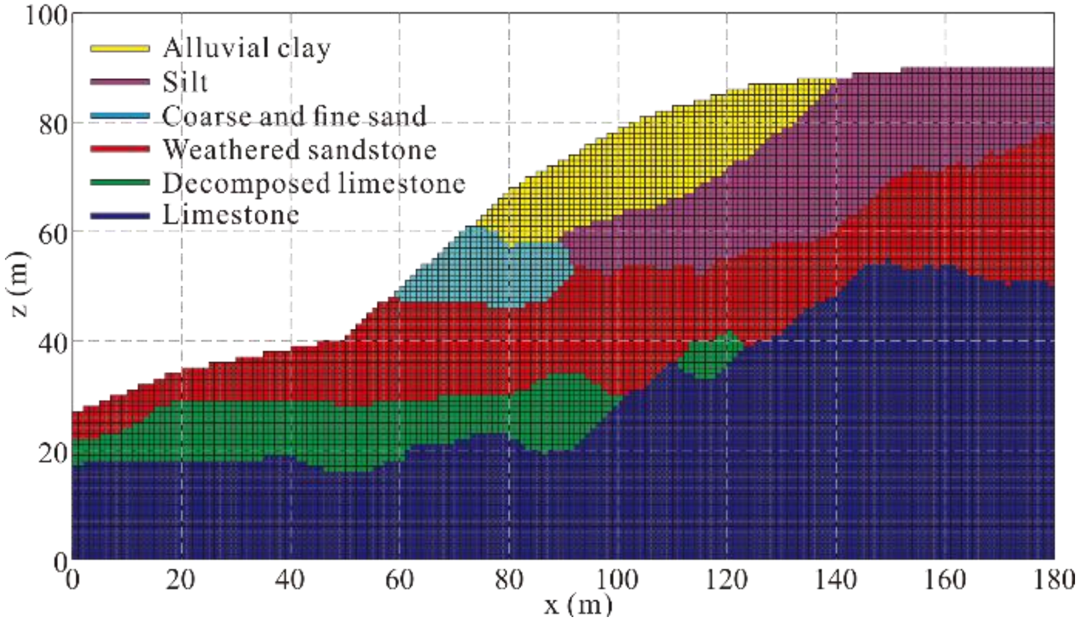


(a) 95% credible intervals of surface settlement;
(b) 95% credible intervals of crown settlement;
(c) normalized variance of settlements.

Example 2: Slope Stability Analysis Considering Subsurface Stratigraphic Uncertainty



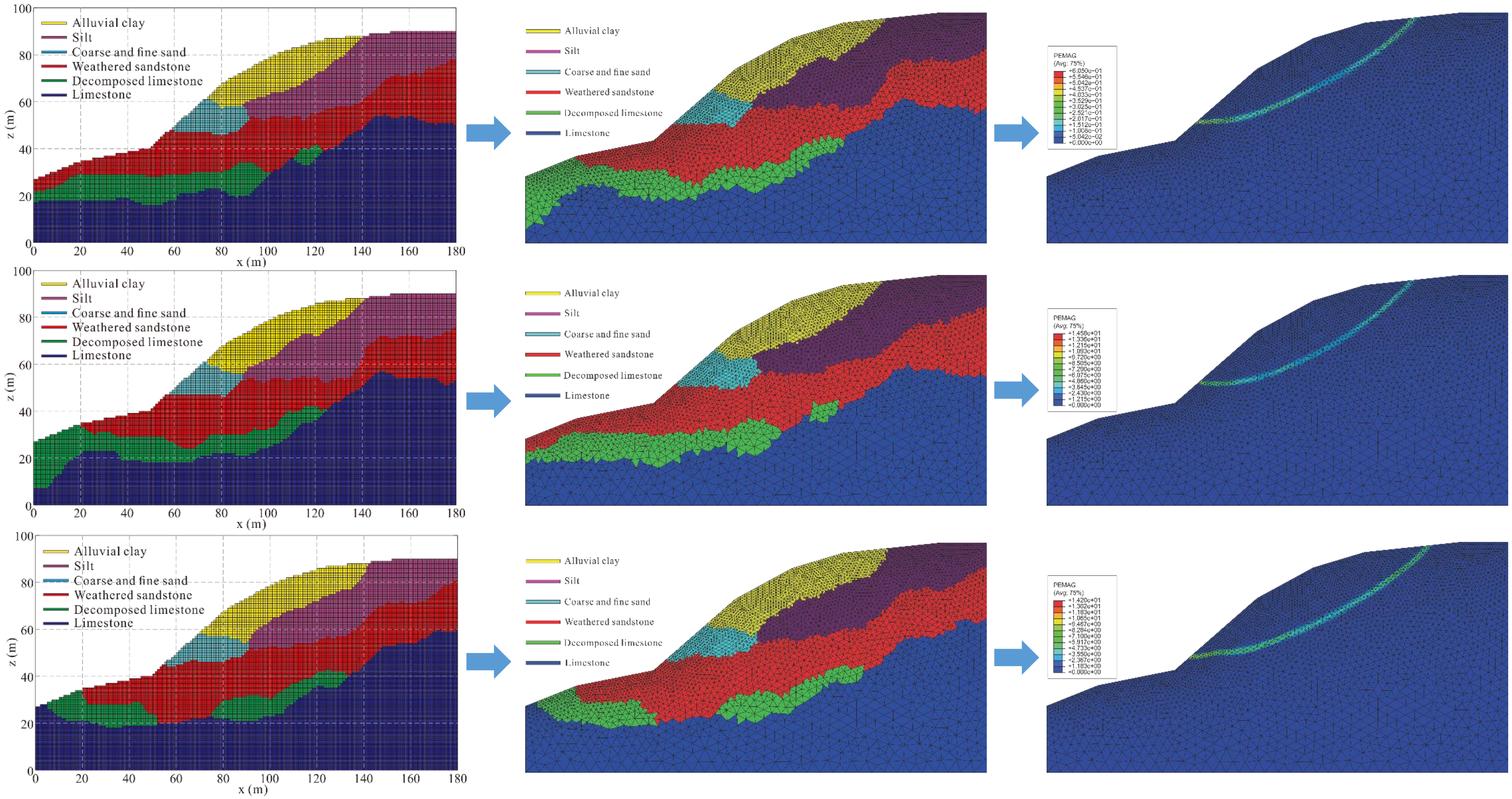
Slope profile and available borehole logs and geologic information



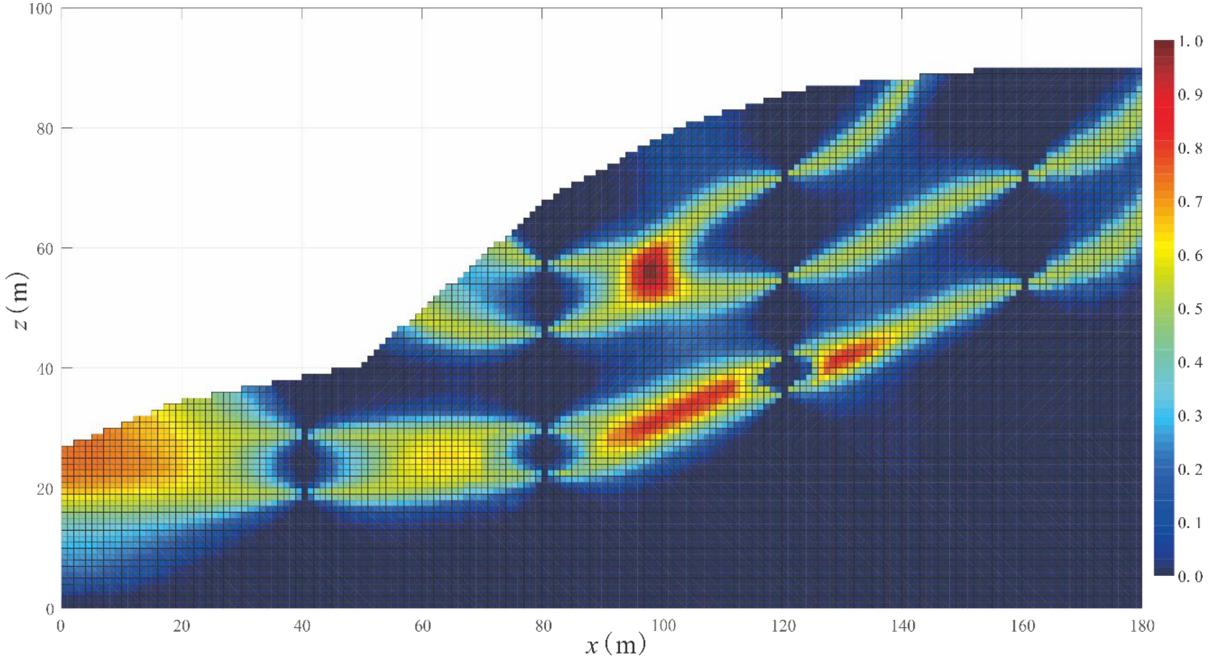
Simulated stratification profile

Example 2: Slope Stability Analysis Considering Subsurface Stratigraphic Uncertainty

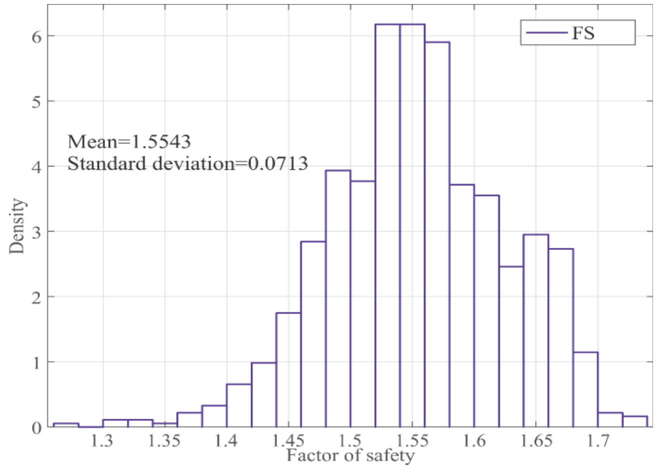
Simulated stratification profiles | Corresponding FEA models | Calculated slip surface



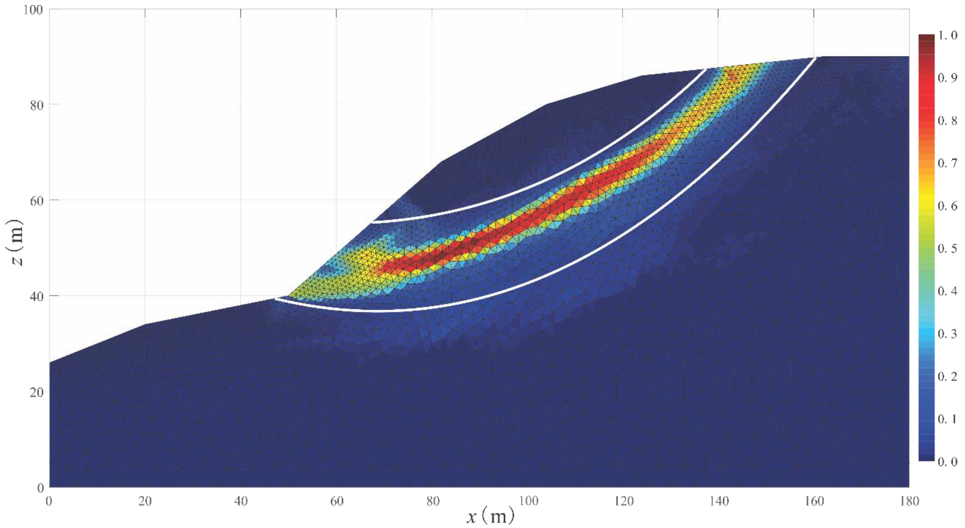
Slope Stability Analysis Considering Subsurface Stratigraphic Uncertainty



Subsurface uncertainty (represented by information entropy) estimated using existing borehole logs



FoS variation caused by the estimated subsurface uncertainty

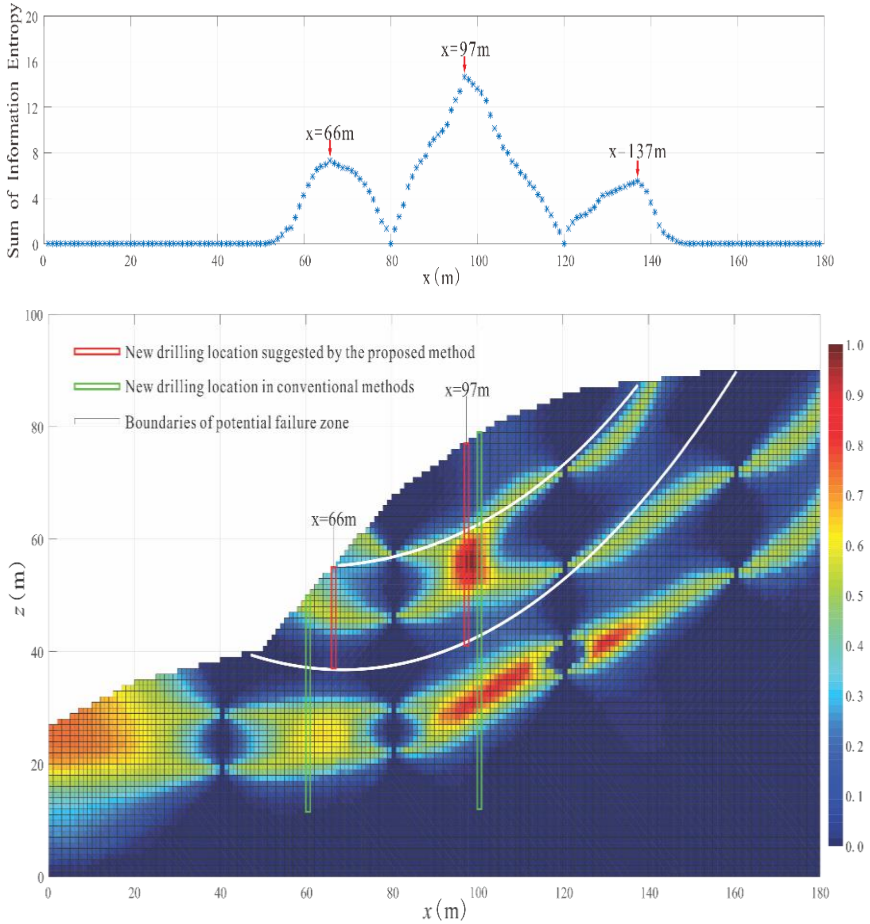


Potential locations of the slip surface calculated based on the estimated subsurface uncertainty

A further question: How to reduce the uncertainty of the FoS and the location of slip surface?

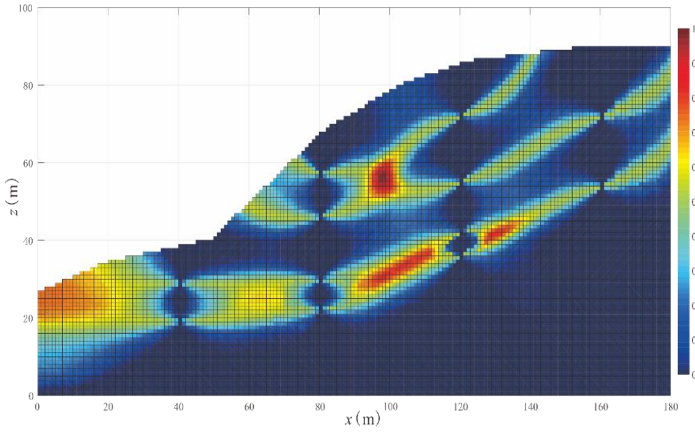
Example 2: Slope Stability Analysis Considering Subsurface Stratigraphic Uncertainty

Analysis flowchart

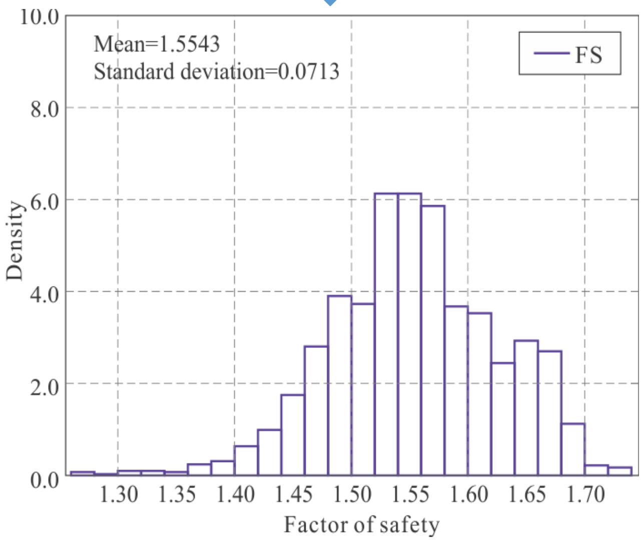


Propose new drilling locations based on the quantification of the subsurface uncertainty

Initial site investigation data

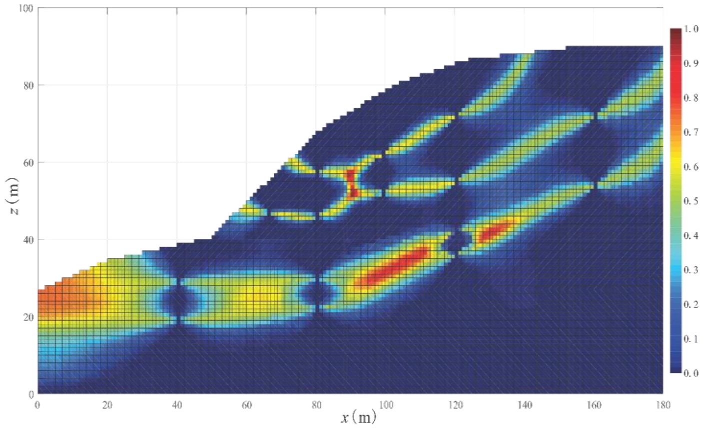


Subsurface uncertainty obtained using initial borehole logs

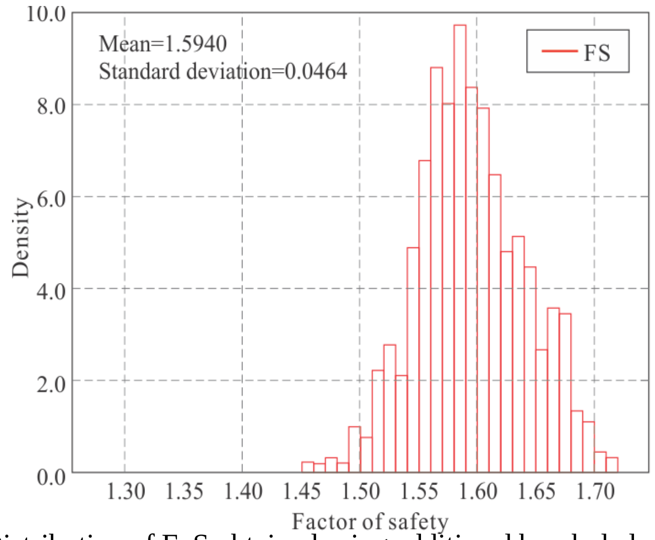


Distribution of FoS obtained using initial borehole logs

Additional site investigation data

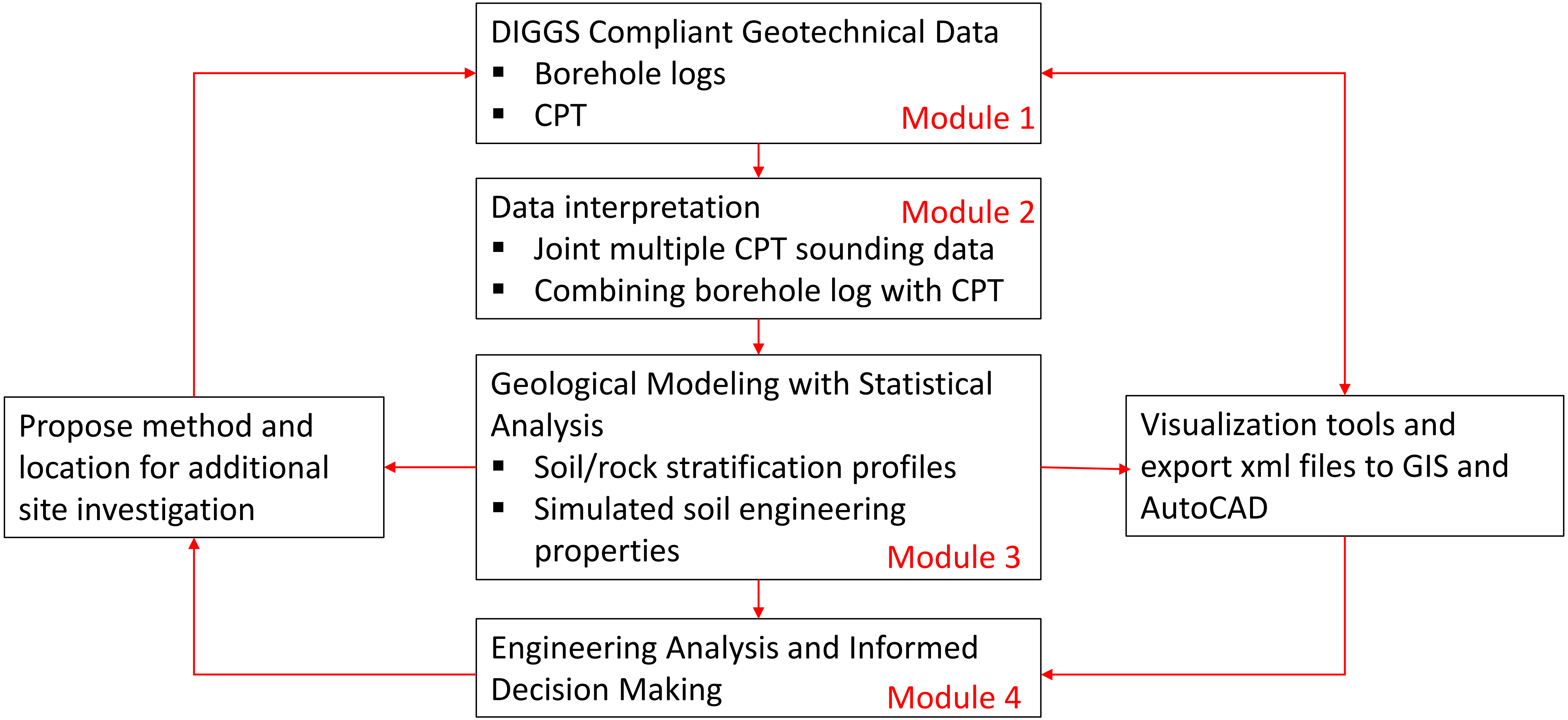


Subsurface uncertainty obtained using additional borehole logs



Distribution of FoS obtained using additional borehole logs

Recap



DIGGS Compliant Geotechnical Data
▪ Borehole logs
▪ CPT
Module 1
Data interpretation
Module 2
▪ Joint multiple CPT sounding data
▪ Combining borehole log with CPT
Geological Modeling with Statistical Analysis
▪ Soil/rock stratification profiles
▪ Simulated soil engineering properties
Module 3
Engineering Analysis and Informed Decision Making
Module 4
Propose method and location for additional site investigation
Visualization tools and export xml files to GIS and AutoCAD

Thank you to all the assistance provided by all Ohio DOT OGE personnel involved during the development of the above methods