Geosyntec consultants

FROM DATA TO INSIGHTS: INTELLIGENT BIG DATA APPROACHES IN GEOTECHNICAL PROJECTS

Presented by

Xin Peng, Ph.D., P.E.

STGEC 2023

November 1st, 2023

OBJECTIVES

- What can you do with standardized data formats in geotechnical engineering?
- Regardless of your current stage, how can you seamlessly leverage the power of big data in your workflows?

BIG DATA IN GEOTECHNICAL ENGINEERING

Big Data

Volume
- The magnitude of data collected/available at a site, e.g., field and lab testing data, instrumentation and surveying data, remote sensing data, etc.

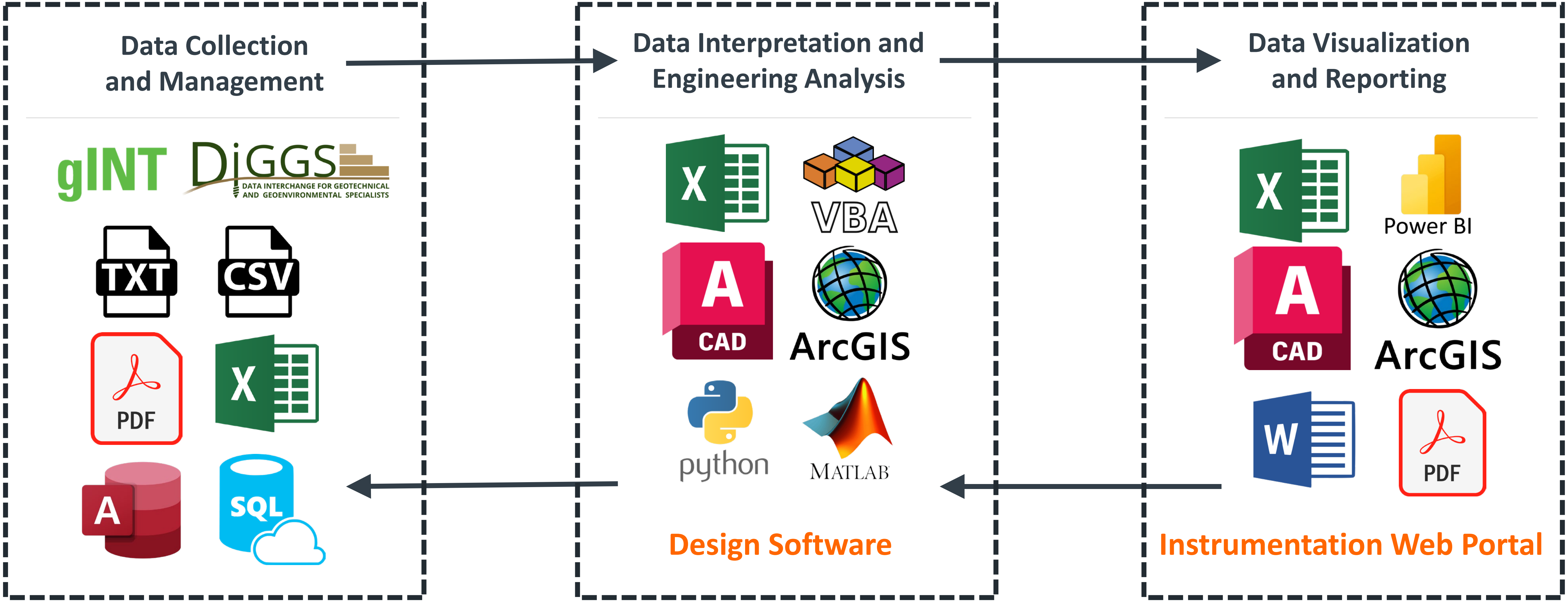
Velocity
- How fast the data can be generated, collected and processed, e.g., Cone Penetration Tests (CPTs), Measurement While Drilling (MWD), real-time instrumentation monitoring, etc.

Value
- How the inherent value of the data can be properly identified and effectively transmuted into actionable insights

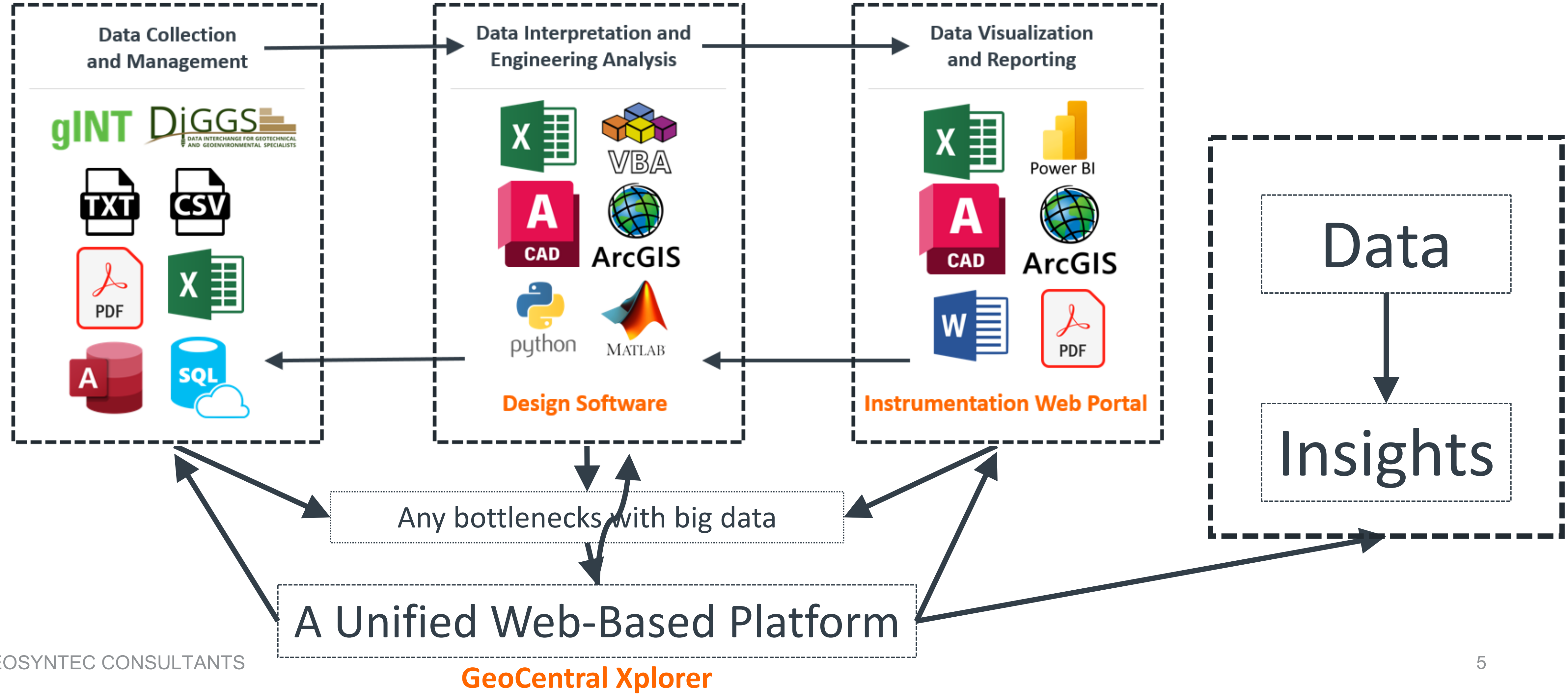
Veracity
- Data validation and quality checking

Variety
- From raw and interpreted data in excel, csv, and traditional databases to unstructured PDF reports and other documents, emails, audios, videos, etc.

TYPICAL DATA WORKFLOW



ENHANCED DATA WORKFLOW



Geotechnical Data

- Field Investigation Planning and Tracking
- Soil Borings
- Geotechnical Laboratory Testing
- Cone Penetration Tests (CPTs)
- Instrumentation Monitoring
- Remote Sensing
- Environmental Sampling Testing
- Survey Data
- CQA Data

QUESTIONS

Have you ever worked on a project that involved more than

20,000 CPT soundings?

Data size for same drilling depth:

20,000 CPTs ≈ **1.2 Million or more** soil borings

PROJECT BACKGROUND

3
Site Areas

A Land Reclamation Project

20,000+
CPTs

In different areas of the project site at multiple phases

400+
BHs

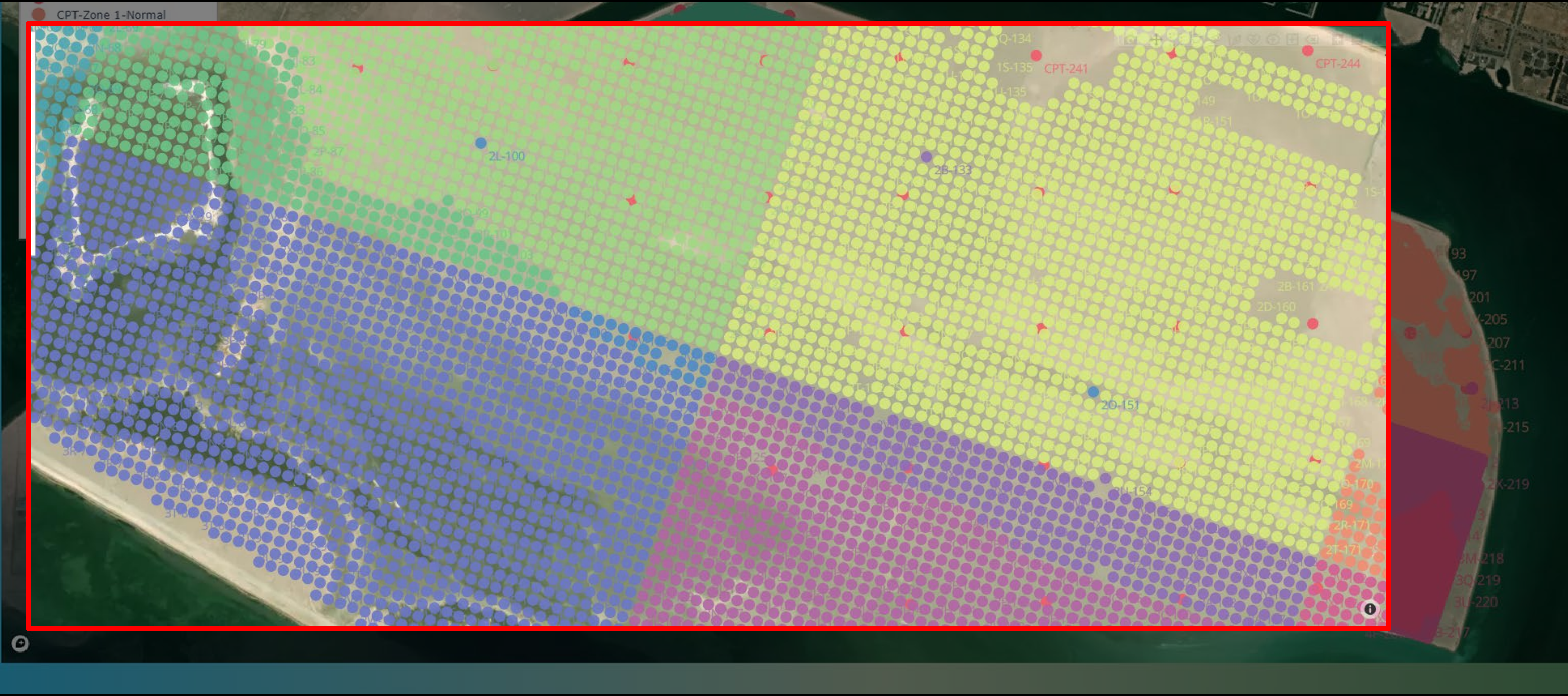
Boring logs and lab testing data for trial pits, marine and land borings.

Tight
Schedule

Multiple teams need to use data to perform engineering analysis

Goal in Phase I of one project site: Use the web-based platform to analyze geotechnical data from **1,800+ CPTs**, **100+ boreholes** with laboratory testing results in **1 week**. Have the selected geotechnical material properties ready for engineering evaluations at the end of **2 weeks**.

Project Site with 11,000+ CPTs and 300+ Soil Borings



Dynamically cut any CPT cross-section in seconds

Efficiently evaluate different CPT interpretation methods

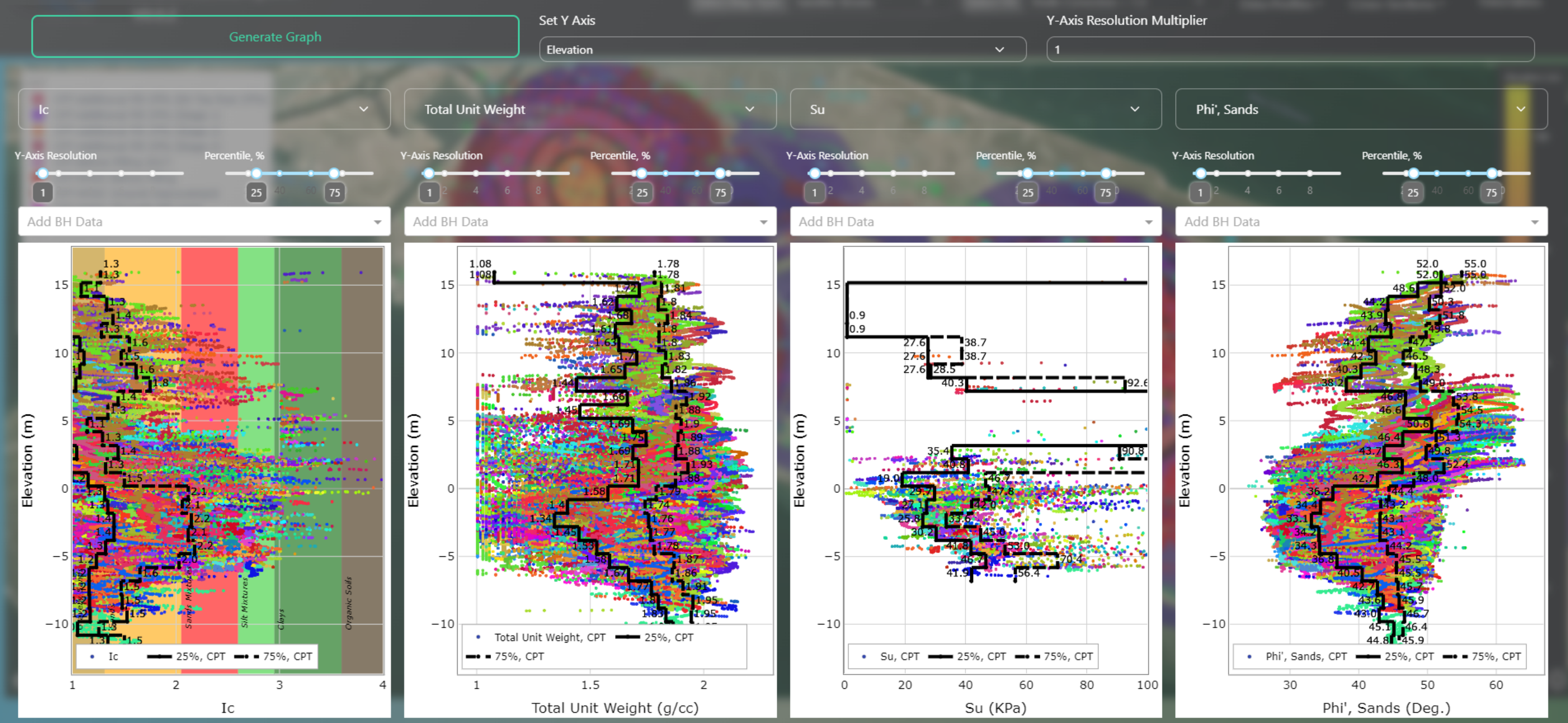
Quickly identify critical soil layers with data filtering on the fly (e.g., layers with phi’ less than 35 deg.)

Dynamically cut cross-sections with both soil borings and CPTs in seconds

Efficiently and interactively compare CPT and soil boring data in a cross-section view

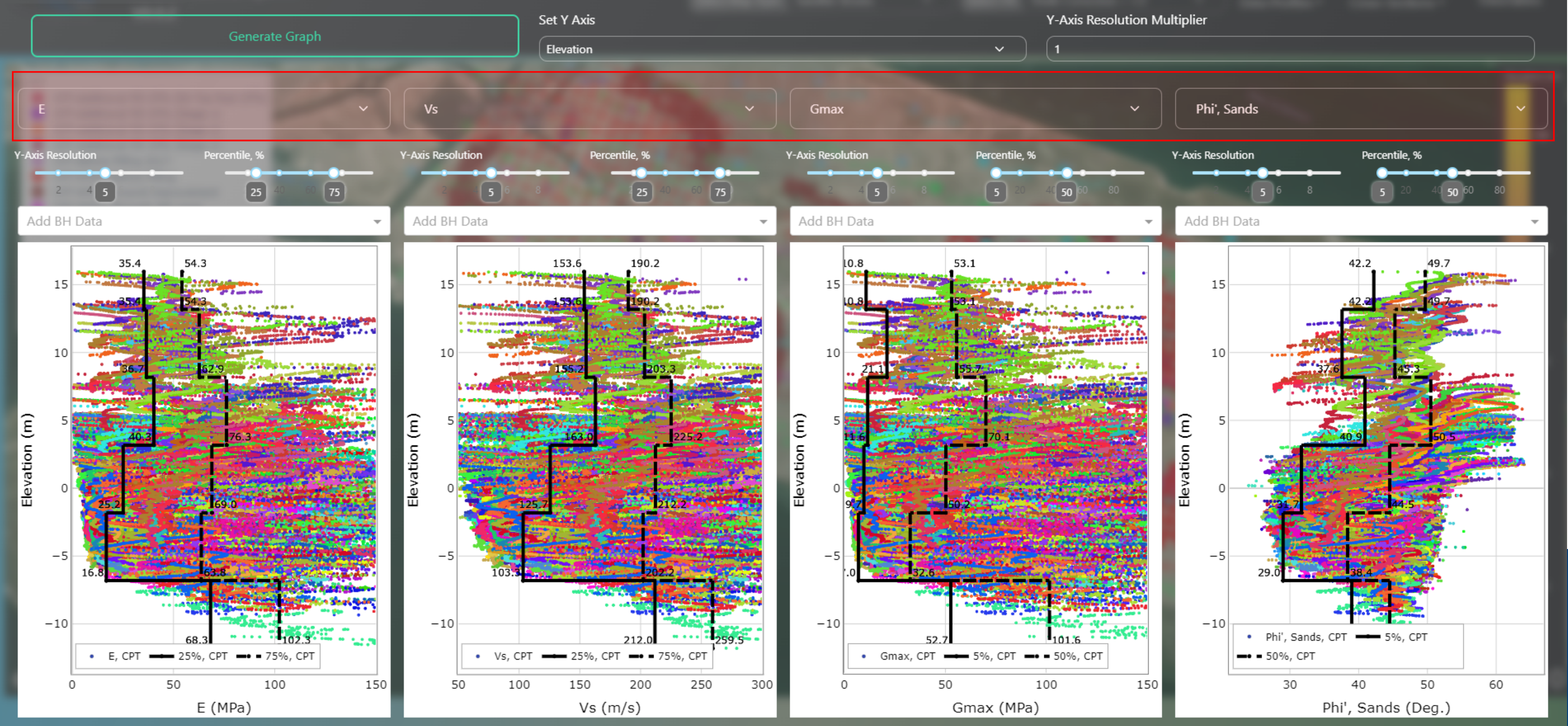
Geospatially identify critical soil layers in plan view at any project area with dynamic data filtering. Select locations with your cursor on the browser

Where are the CPTs with sand layers having Phi’ less than 26 deg.?



Efficiently develop soil design parameter profiles at any area of the site in seconds

900+ CPTs selected from the map in this example



Select and interpret different soil design parameters

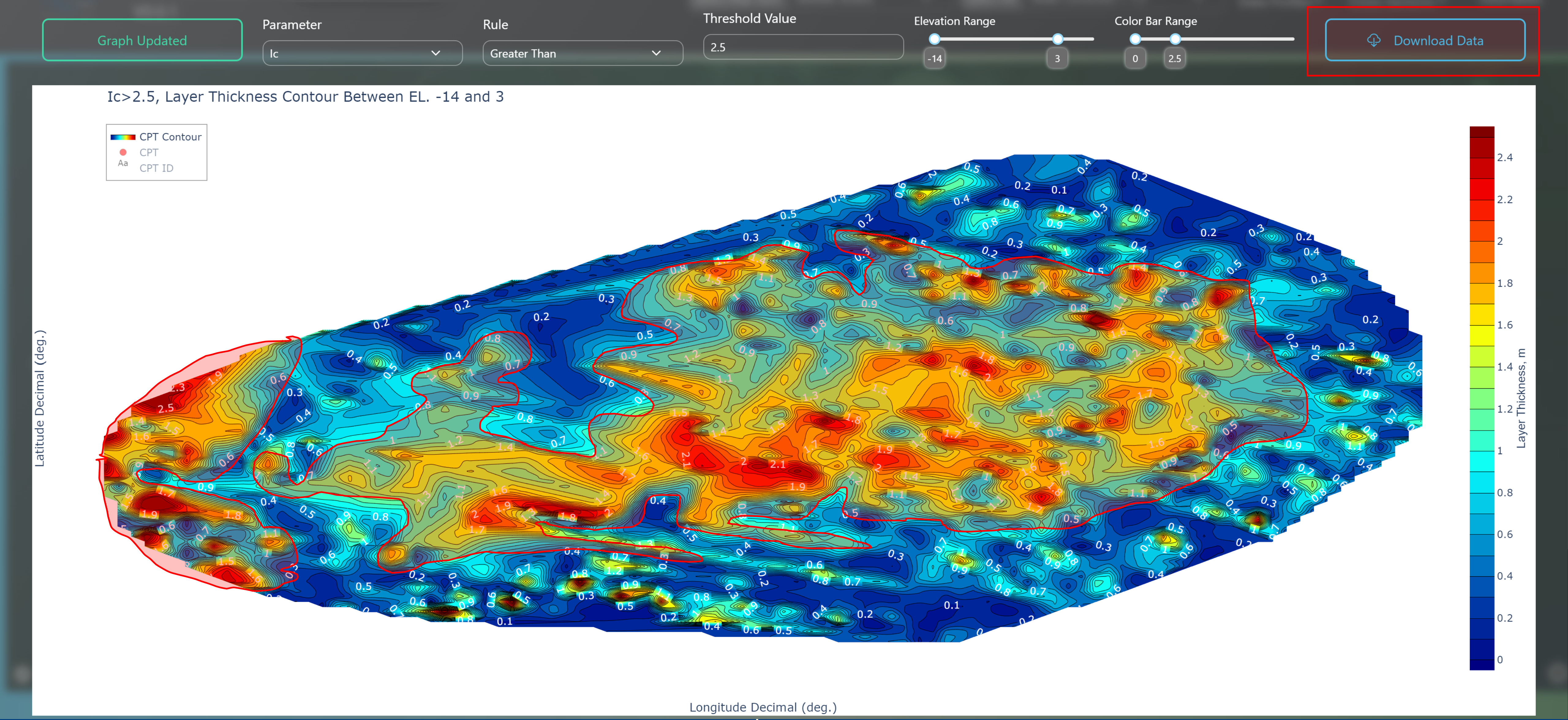
More than 20 soil parameters have been added in this project for slope stability, settlement, seismic, and numerical analyses

Dynamically <u>add lab testing data</u> from nearby soil borings to compare with CPT correlations

Data percentile lines for lab testing data are calculated on the fly for evaluation

INTERACTIVE 2D AND 3D VIEWS

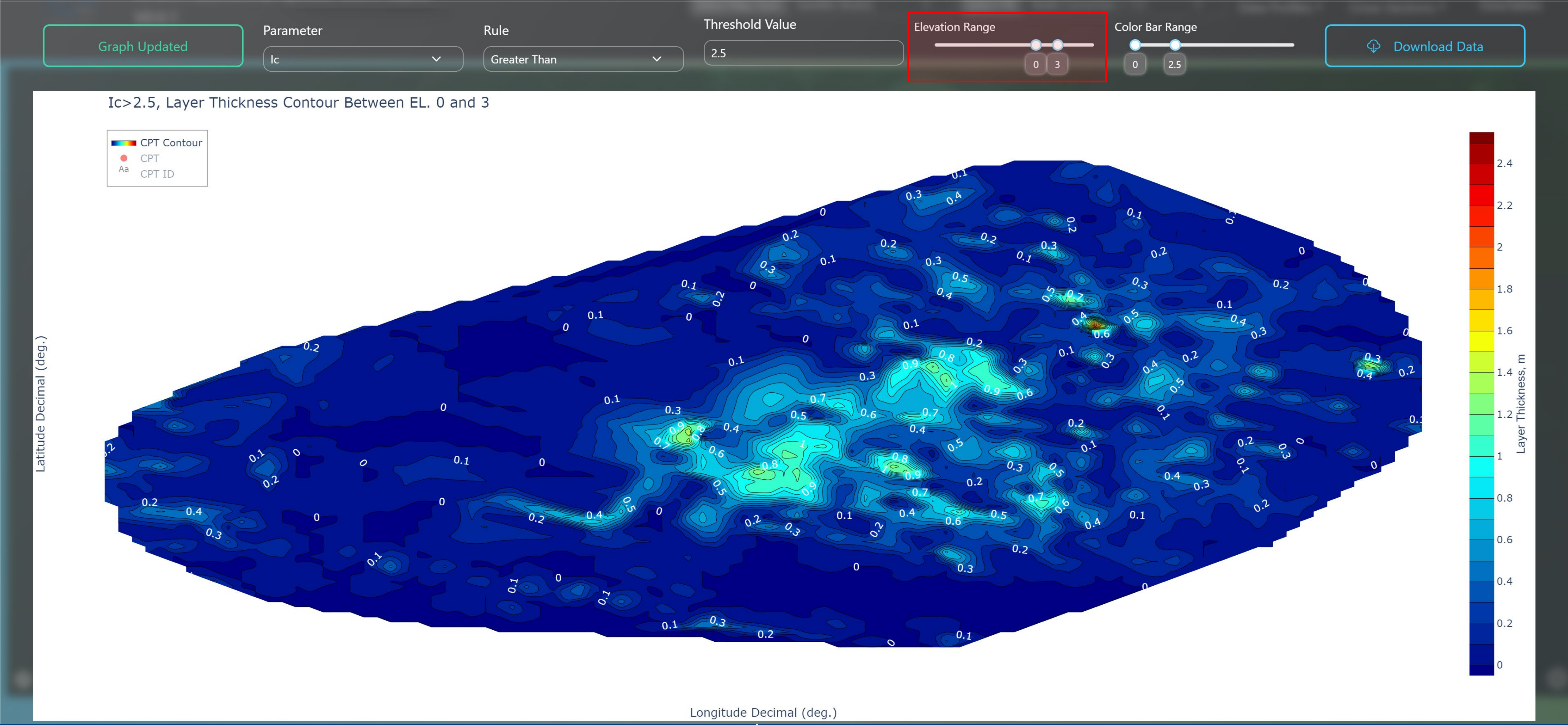
Project Site with
1,300 + CPTs



Dynamically generate interactive 2D contour view to identify the thickness of clay layers throughout the project site

Rule to define clay layers: Ic > 2.5

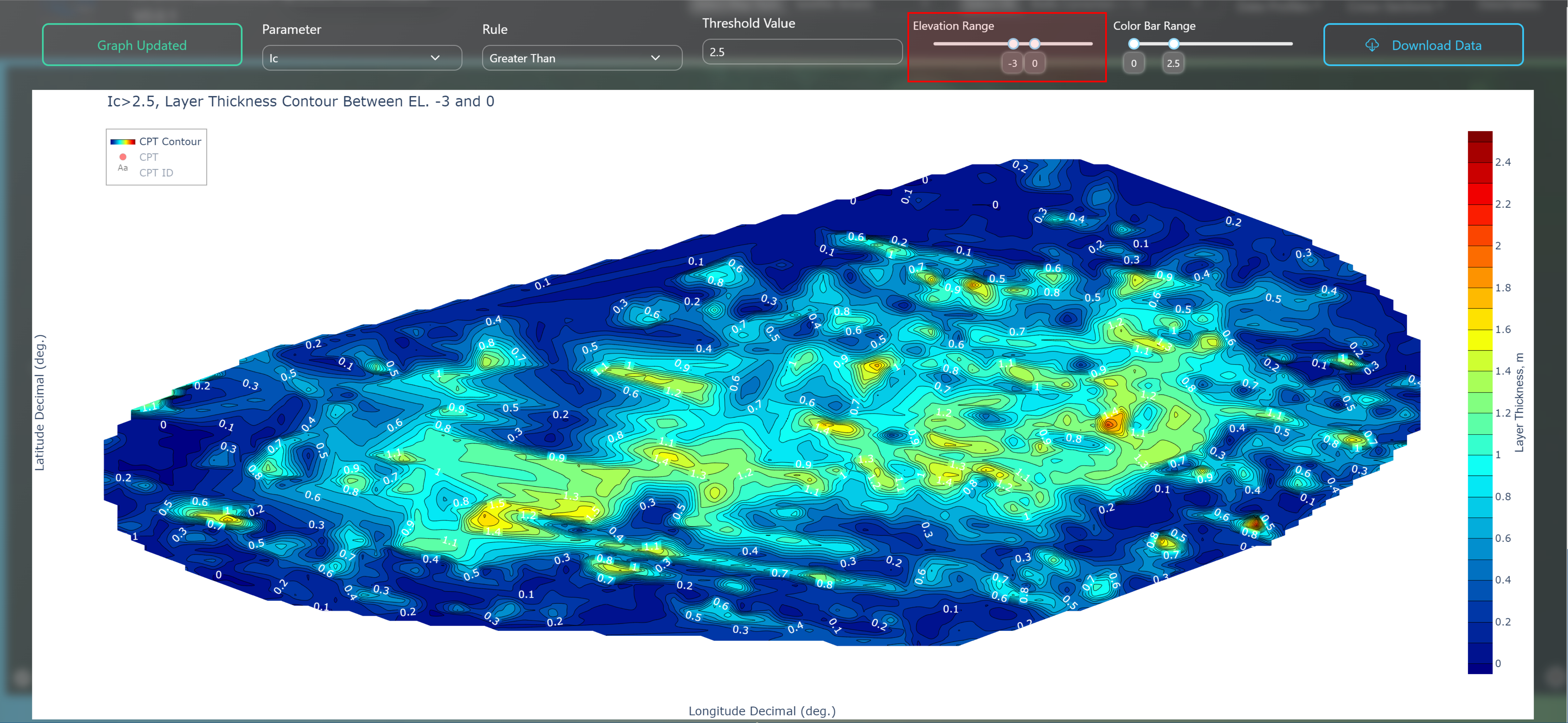
Observation: Two major areas with up to 2.5-m thickness of clay layers



Dynamically cut contour slices to further identify the clay layers in different elevation ranges

Elevation Range: +3 to 0 m

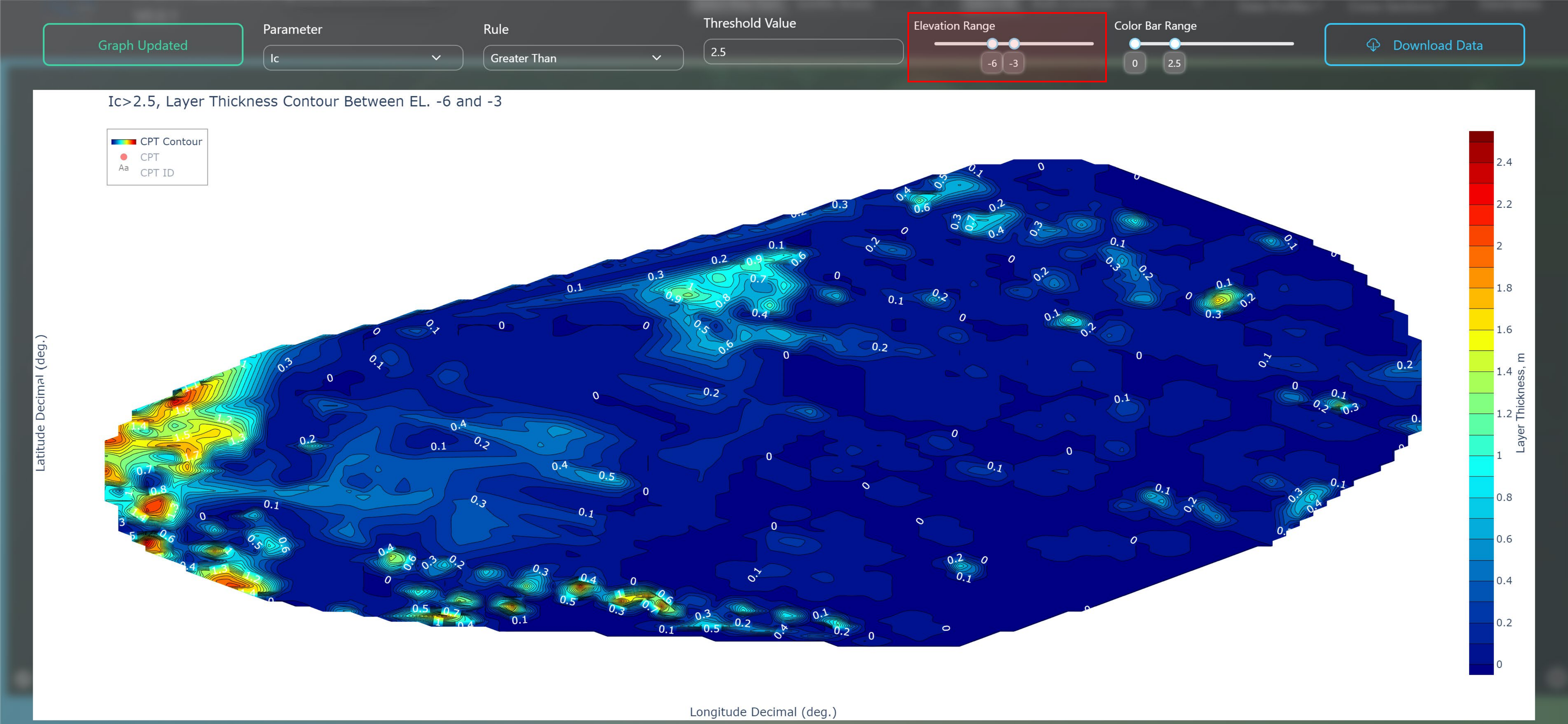
Observation: Approximately up to 1-m thickness clay layers in the **central** area



Dynamically cut contour slices to further identify the clay layers in different elevation ranges

Elevation Range: 0 to -3 m

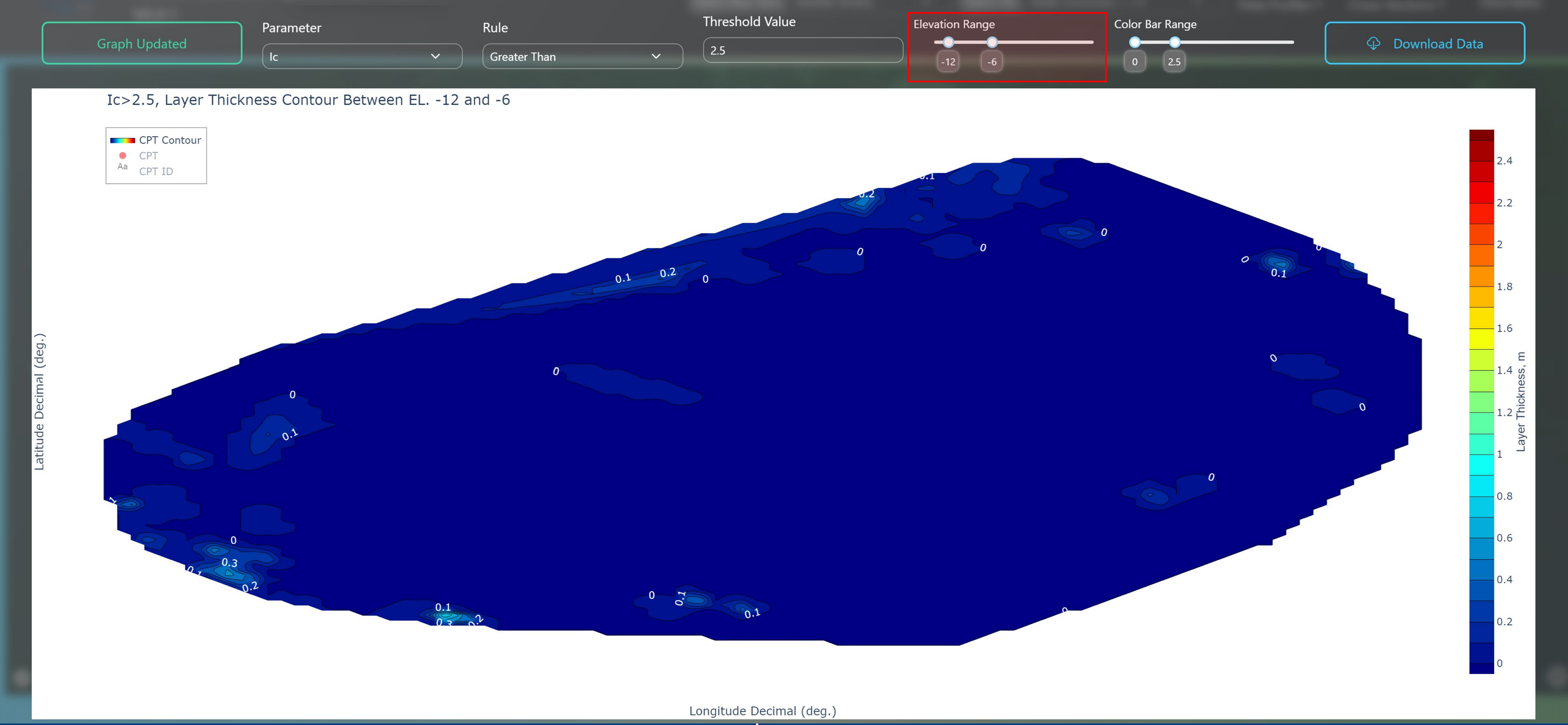
Observation: Approximately 1.5-m thickness clay layers in the **central** area



Dynamically cut contour slices to further identify the clay layers in different elevation ranges

Elevation Range: -3 to -6 m

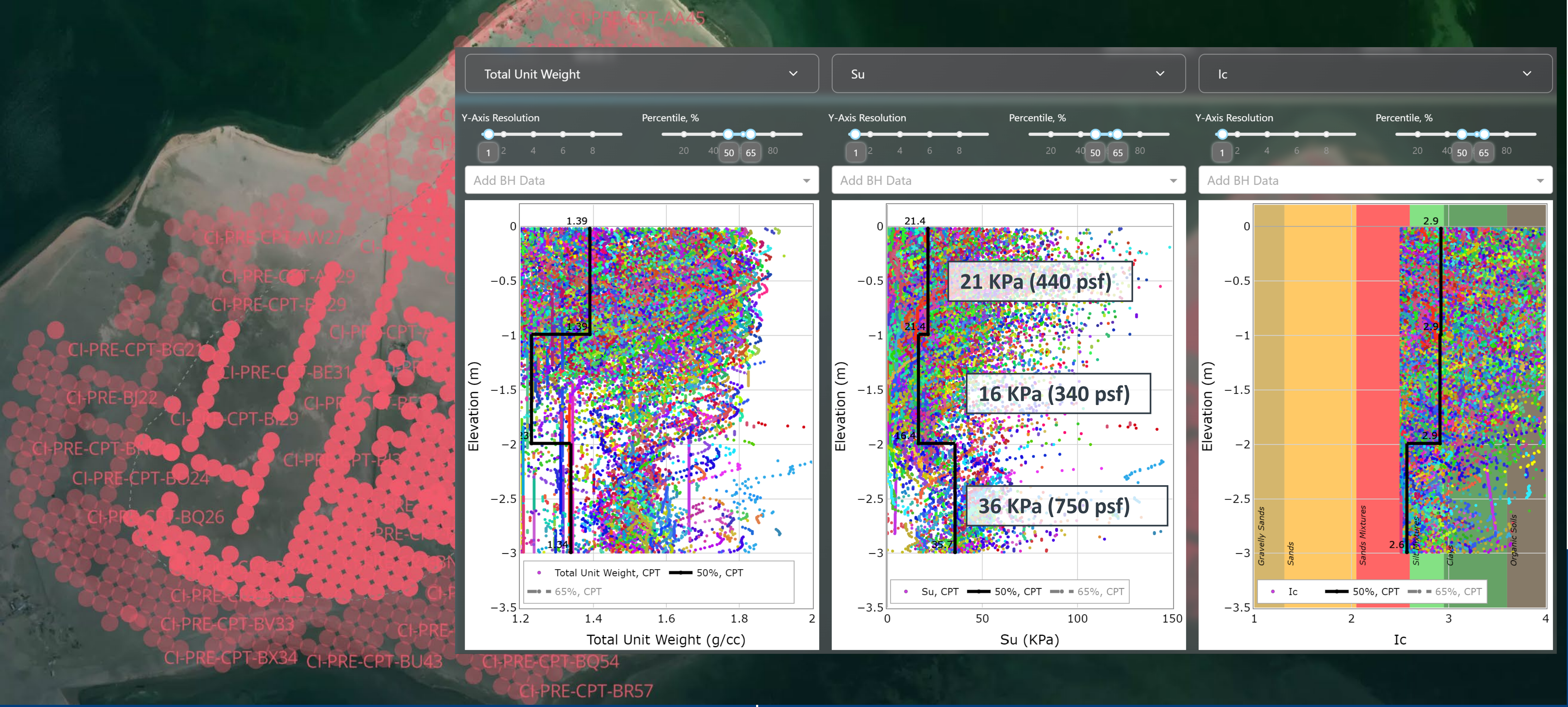
Observation: Approximately up to 2-m thickness clay layers in the **western** area



Dynamically cut contour slices to further identify the clay layers in different elevation ranges

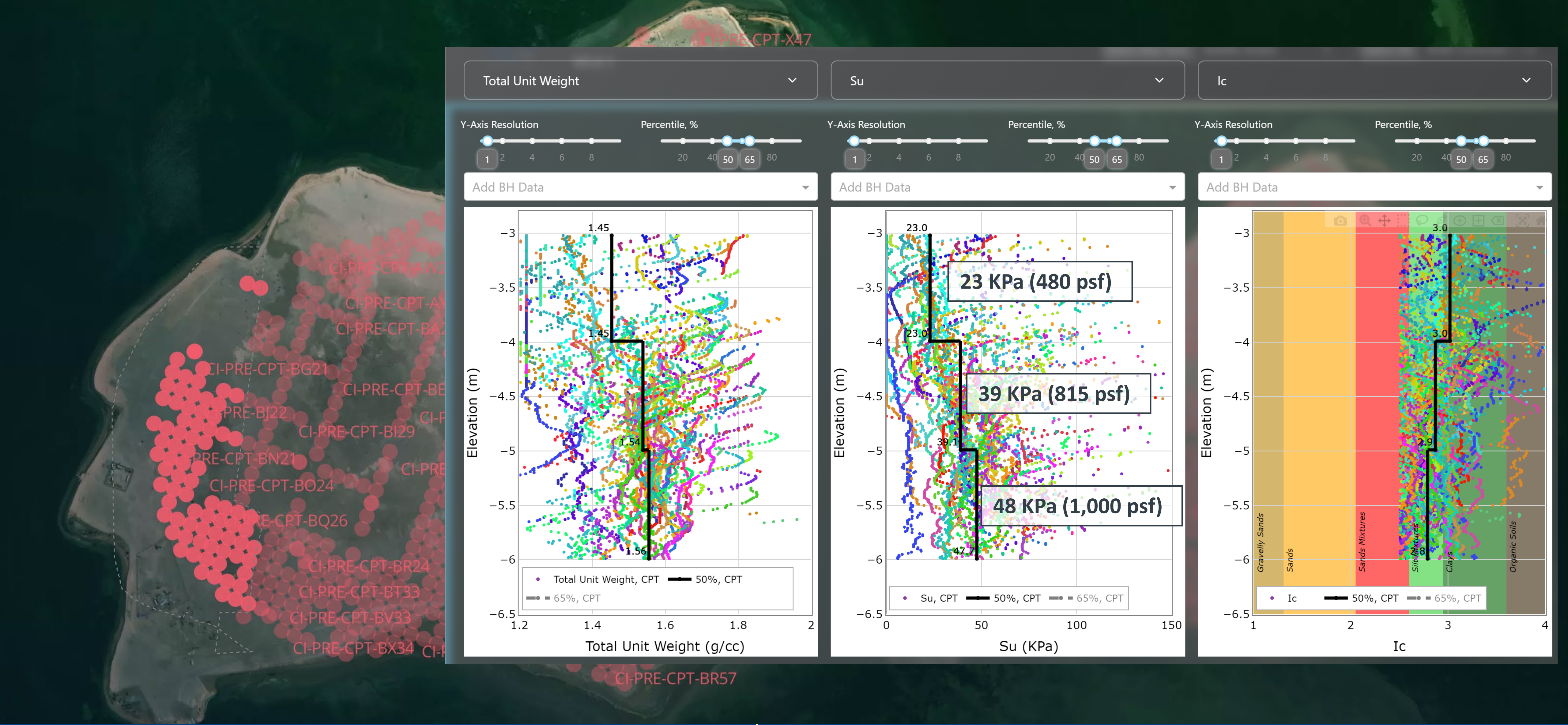
Elevation Range: -6 to -12 m

Observation: No major clay layers identified in the site area



Select **600+** CPTs in the **central** area to develop strength lines for clay layers between **El. 0 and -3 m**

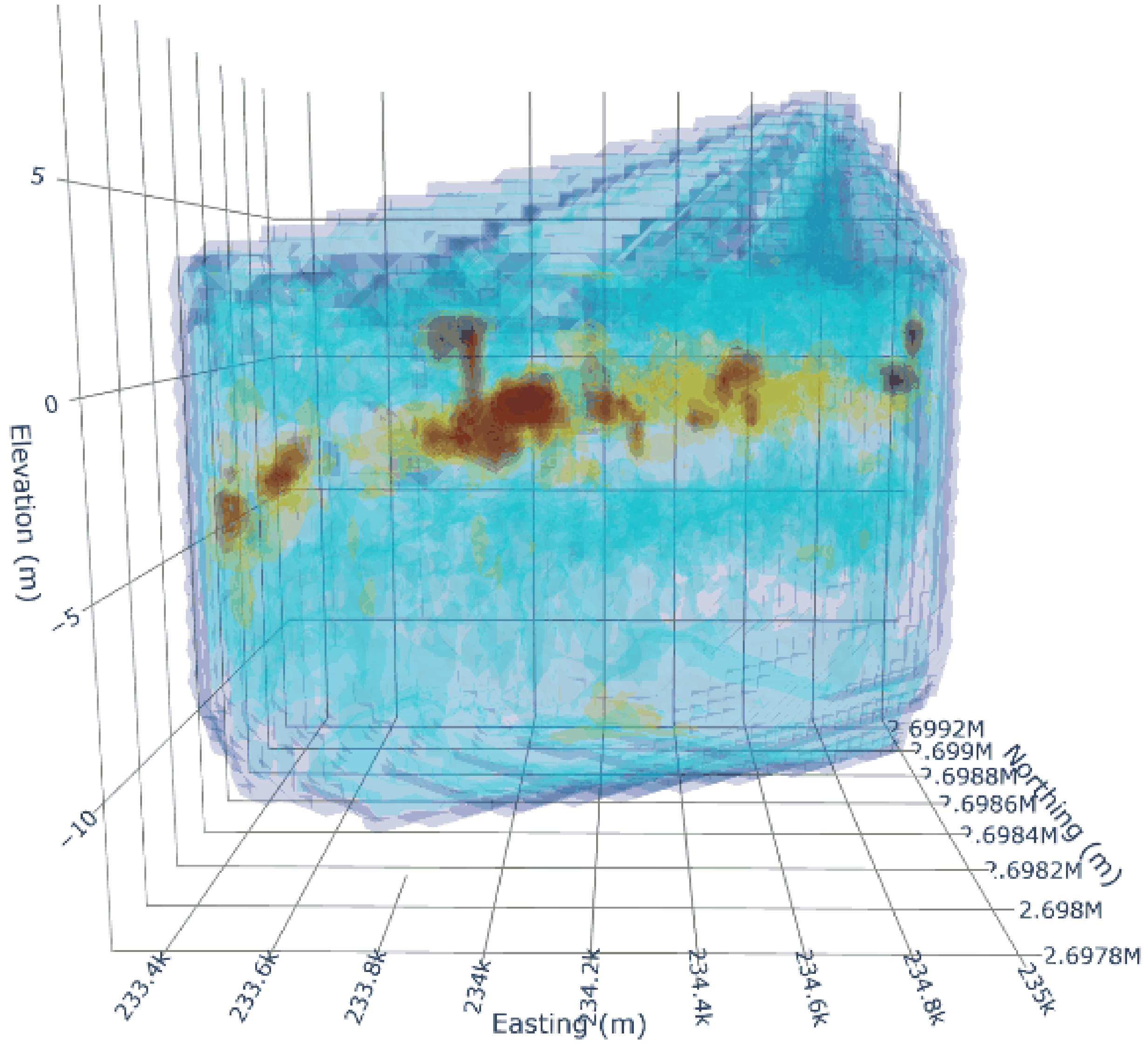
Median Lines of Su: From 16 KPa (340 psf) to 36 KPa (750 psf)



Select **80+** CPTs in the **western** area to develop strength lines for clay layers between El. -3 and -6 m

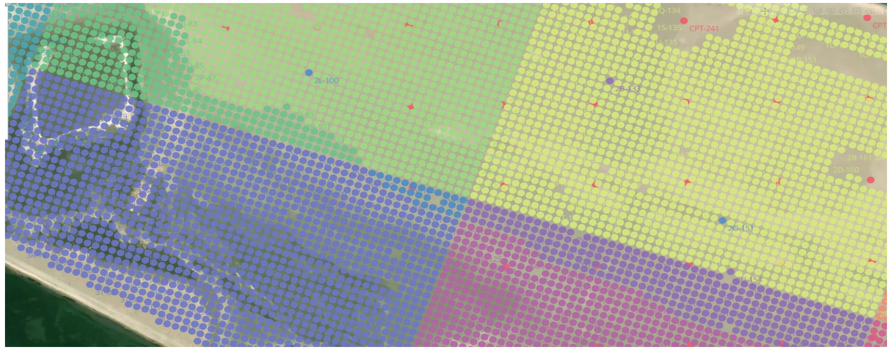
Median Values of Su: From 23 KPa (480 psf) to 48 KPa (1000 psf)

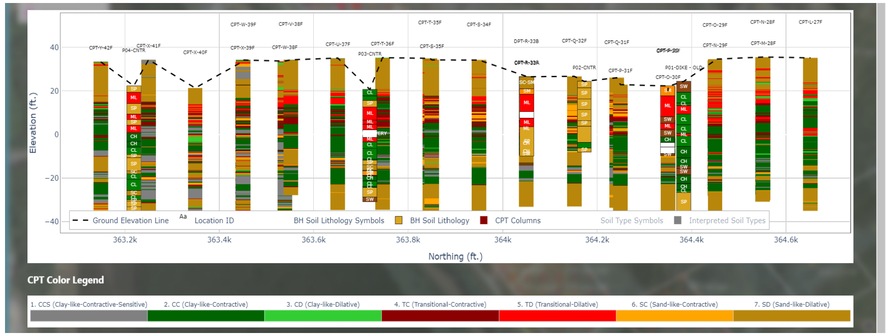
INTERACTIVE 3D VIEWS



Interactive 3D views can be generated on-the-fly: Warm colors (**red** and **yellow**) indicate clay layers

UNDER THE HOOD



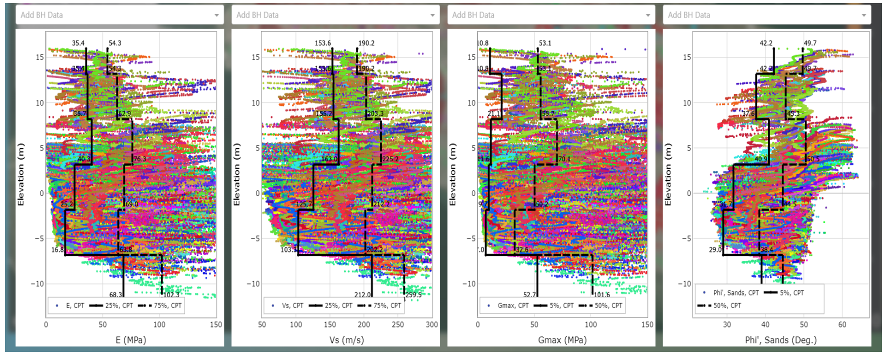


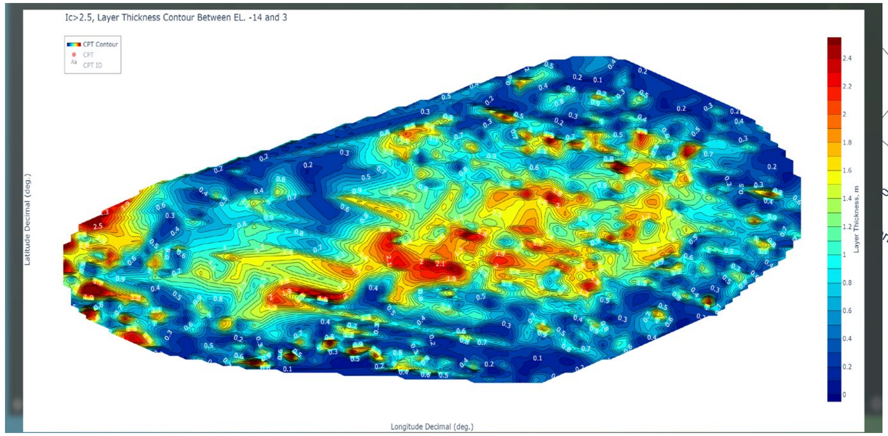
Standardized Digital Data

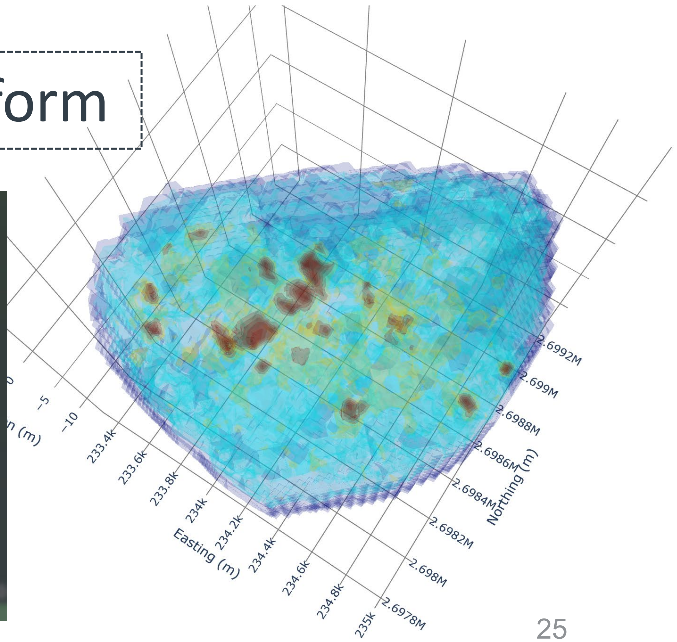
Data Ingestion

GeoCentral Xplorer

A Unified Web-Based Platform





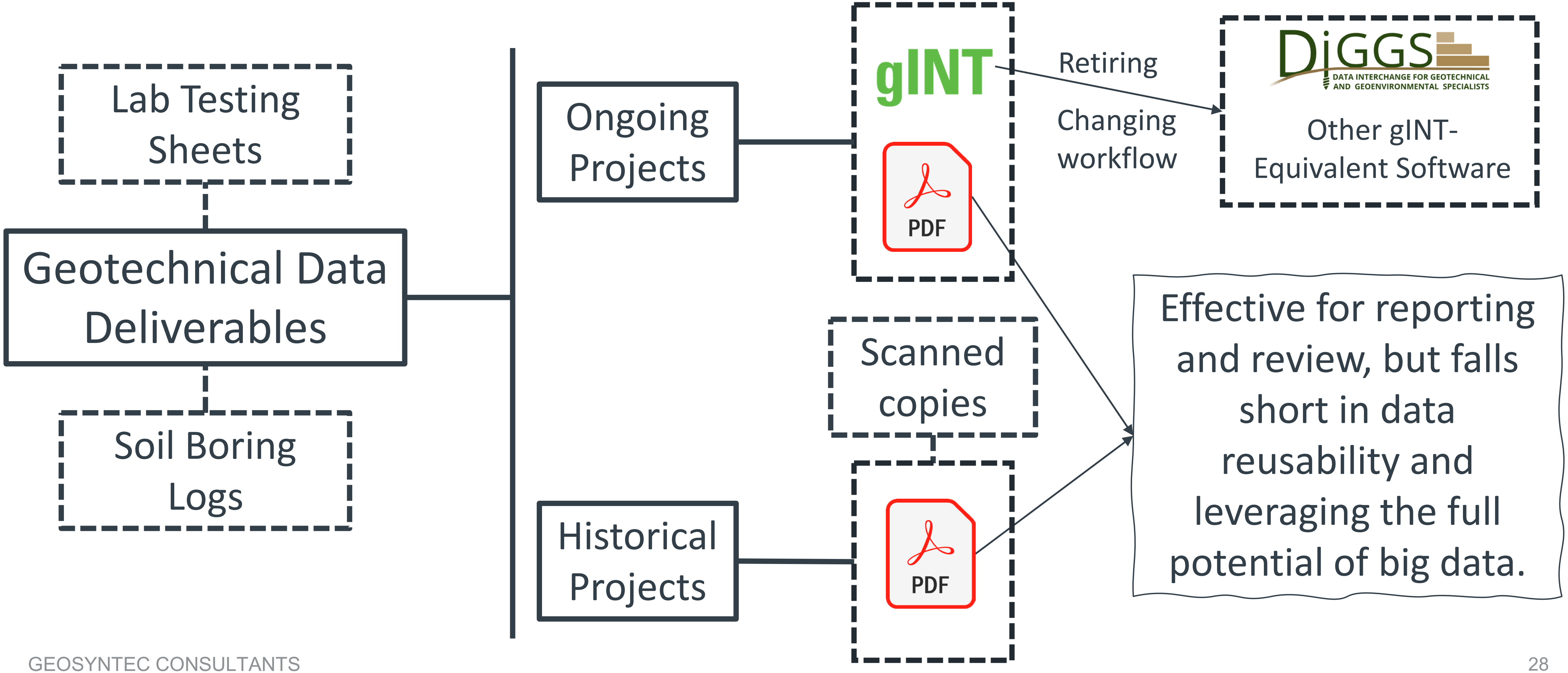


KEY TAKEAWAYS – PART I

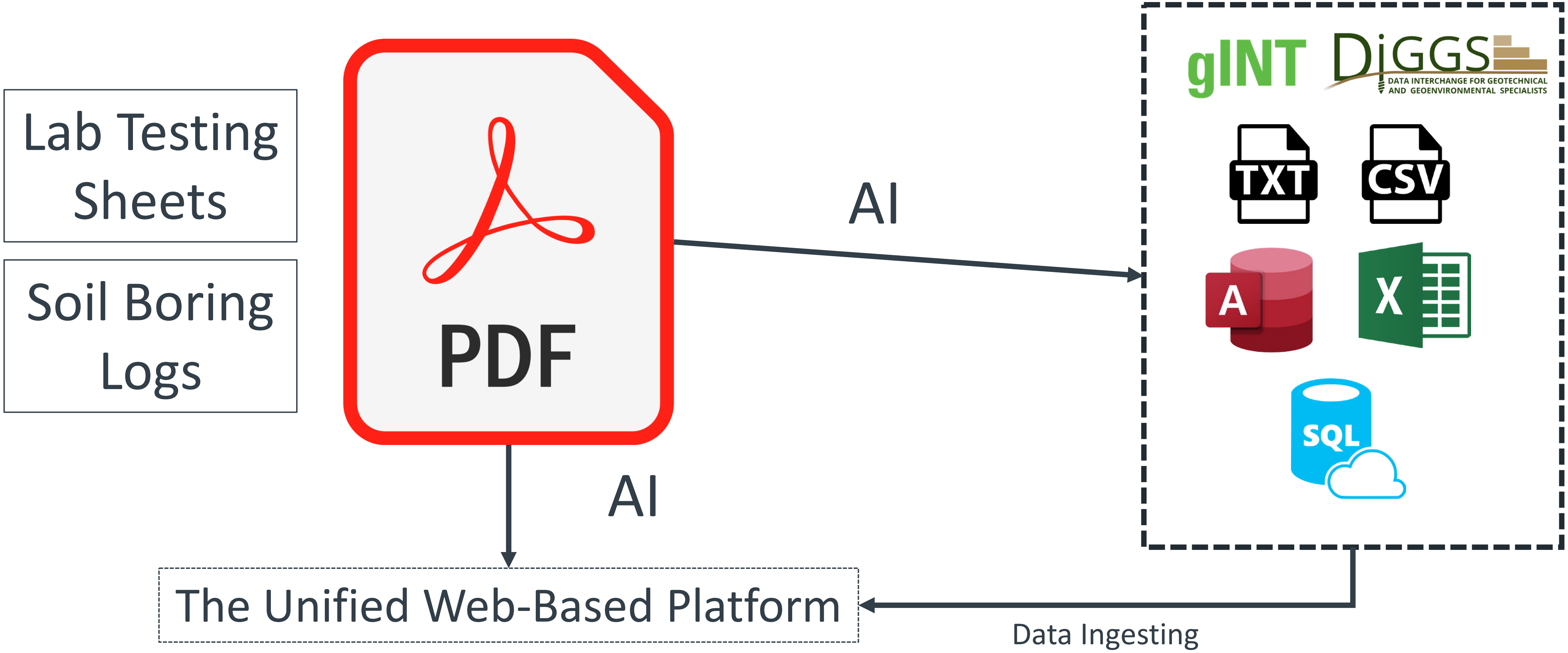
- ✓ The purpose of **data management** extends **beyond merely organizing data**.
- ✓ It's about leveraging the power of big data to **enhance decision-making**, **streamline design and construction processes**, and **achieve cost savings**.
- ✓ We've broadened the use of this **unified web-based platform, GeoCentral Xplorer**, in our current geotechnical projects and plan to integrate it into many upcoming projects at Geosyntec. Its current capabilities extend beyond handling, visualizing, and interpreting data from over **20,000 CPTs**.

BIG DATA STRATEGIES FOR DOT PROJECTS

HOW CAN DOTS INCORPORATE STANDARDIZED DATA FORMATS INTO CURRENT WORKFLOWS?

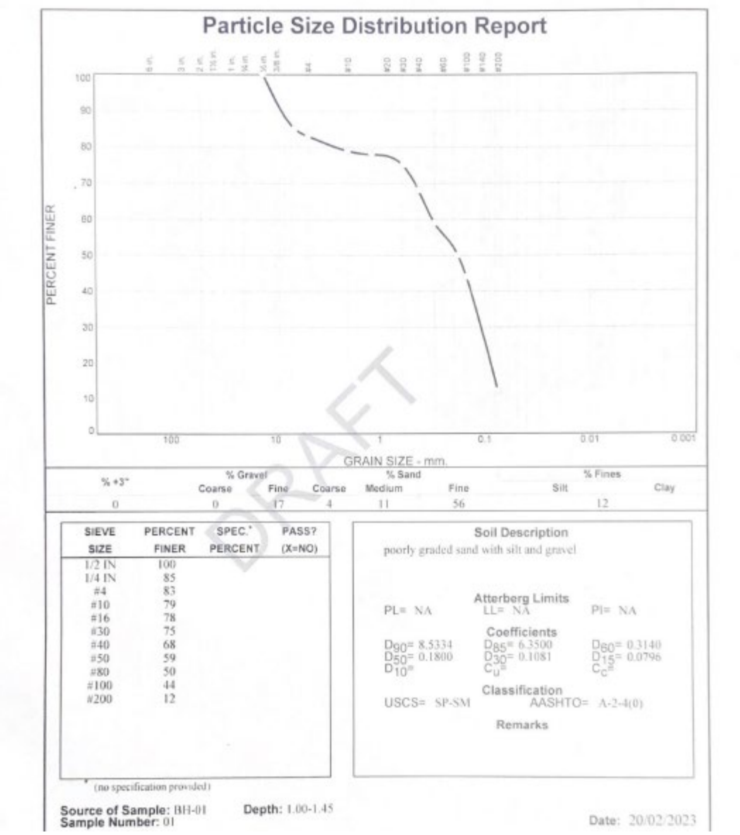


IS IT POSSIBLE TO NOT CHANGE CURRENT WORKFLOWS WITH PDFS IN THE ERA OF BIG DATA?



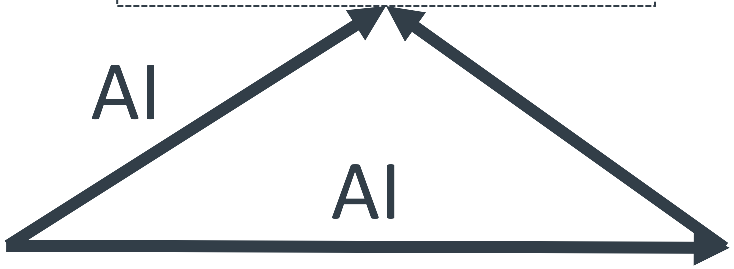
AI IMPLEMENTATION EXAMPLE:
LAB SHEET DIGITIZATION IN A GEOSYNTEC PROJECT

400+ pages from a scanned pdf report



GeoCentral Xplorer

The Unified Web-Based Platform



90+% time savings compared to traditional manual data entry.

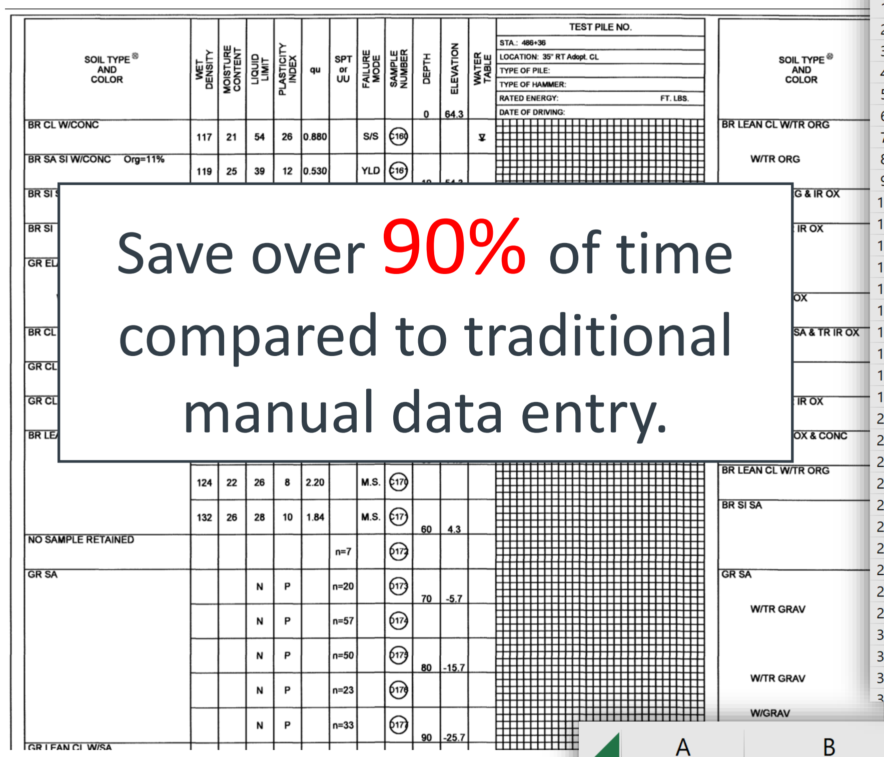
4000+ rows of digital data for engineering interpretation

	A	B	C	D	E	F	G
1	Location ID	Sample ID	Depth Top	Depth Base	USCS	SIEVE SIZE	PERCENT FINER
2	BH-01	1	1	1.45	SP-SM	1/2 IN	100
3	BH-01	1	1	1.45	SP-SM	1/4 IN	85
4	BH-01	1	1	1.45	SP-SM	#4	83
5	BH-01	1	1	1.45	SP-SM	#10	79
6	BH-01	1	1	1.45	SP-SM	#16	78
7	BH-01	1	1	1.45	SP-SM	#30	75
8	BH-01	1	1	1.45	SP-SM	#40	68
9	BH-01	1	1	1.45	SP-SM	#50	59
10	BH-01	1	1	1.45	SP-SM	#80	50
11	BH-01	1	1	1.45	SP-SM	#100	44
12	BH-01	1	1	1.45	SP-SM	#200	12
13	BH-01	2	3	3.45	SM	#10	100
14	BH-01	2	3	3.45	SM	#16	99
15	BH-01	2	3	3.45	SM	#30	99
16	BH-01	2	3	3.45	SM	#40	97
17	BH-01	2	3	3.45	SM	#50	94
18	BH-01	2	3	3.45	SM	#80	90
19	BH-01	2	3	3.45	SM	#100	82
20	BH-01	2	3	3.45	SM	#200	17
21	BH-01	3	8	8.45	SM	#4	100
22	BH-01	3	8	8.45	SM	#10	99
23	BH-01	3	8	8.45	SM	#16	99
24	BH-01	3	8	8.45	SM	#30	95
25	BH-01	3	8	8.45	SM	#40	89
26	BH-01	3	8	8.45	SM	#50	75
27	BH-01	3	8	8.45	SM	#80	60
28	BH-01	3	8	8.45	SM	#100	43
29	BH-01	3	8	8.45	SM	#200	18
30	BH-02	1	1	1.45	SP-SM	1/4 IN	100
31	BH-02	1	1	1.45	SP-SM	#4	99
32	BH-02	1	1	1.45	SP-SM	#10	97
33	BH-02	1	1	1.45	SP-SM	#16	96
34	BH-02	1	1	1.45	SP-SM	#30	91
35	BH-02	1	1	1.45	SP-SM	#40	86
36	BH-02	1	1	1.45	SP-SM	#50	79
37	BH-02	1	1	1.45	SP-SM	#80	70

Sieve Analysis

AI IMPLEMENTATION EXAMPLE: LADOTD BORING LOG DIGITIZATION

Each page contains 3 boring logs, scanned copies, including typed and handwritten with varied scan quality:



Save over **90%** of time compared to traditional manual data entry.

	A	B	C	D	E	F
1	description	depthTop	depthBase	elevationTop	elevationBase	locationID
2	BR CL W/CONC	0	5	64.3	59.3	BORING NO. 4
3	BR SA SI W/CONC Org=11%	5	10	59.3	54.3	BORING NO. 4
4	BR SI SA	10	15	54.3	49.3	BORING NO. 4
5	BR SI	15	20	49.3	44.3	BORING NO. 4
6	GR ELASTIC SI Org=12%	20	30	44.3	34.3	BORING NO. 4
7	W/IR OX	25		39.3		BORING NO. 4
8	BR CL	30	35	34.3	29.3	BORING NO. 4
9	GR CL W/ORG	35	40	29.3	24.3	BORING NO. 4
10	GR CL SA	40	45	24.3	19.3	BORING NO. 4
11	BR LEAN CL W/IR OX	45	60	19.3	4.3	BORING NO. 4
12	NO SAMPLE RETAINED	60	65	4.3	-0.7	BORING NO. 4
13	GR SA	65	90	-0.7	-25.7	BORING NO. 4
14	GR LEAN CL W/SA	90	95	-25.7	-30.7	BORING NO. 4
15	GR SA SI CL	95	100	-30.7	-35.7	BORING NO. 4
16	GR SA LEAN CL	100	105	-35.7	-40.7	BORING NO. 4
17	GR SI SA W/STR CL	105	110	-40.7	-45.7	BORING NO. 4
18	GR SA SI	110	115	-45.7	-50.7	BORING NO. 4
19	GR SA LEAN CL	115	120	-50.7	-55.7	BORING NO. 4
20	GR SI CL SA	120	125	-55.7	-60.7	BORING NO. 4
21	GR SA LEAN CL W/TR ORG	125	130	-60.7	-65.7	BORING NO. 4
22	GR LEAN CL W/SA	130	139	-65.7	-74.7	BORING NO. 4
23	BR LEAN CL WITR ORG	-0.5	9.5	67	57	BORING NO. 5
24	W/TR ORG	5		61.5		BORING NO. 5
25	BR CL WITR ORG & IR OX	9.5	14.5	57	52	BORING NO. 5
26	BR GR CL W/TR IR OX	14.5	24.5	52	42	BORING NO. 5
27	GR CL W/TR IR OX	24.5	30	42	36.5	BORING NO. 5
28	GR LEAN CL W SA & TR IR OX	30	35	36.5	31.5	BORING NO. 5
29	GR ELASTIC SI	35	39.5	31.5	27	BORING NO. 5
30	BR GR CL WITR IR OX	39.5	44.5	27	22	BORING NO. 5
31	GR CL W/TR IR OX & CONC	44.5	49.5	22	17	BORING NO. 5
32	BR LEAN CL W/TR ORG	49.5	55	17	11.5	BORING NO. 5

	A	B	C	D	E	F	G	H	I	J
1	wetDensity	moistureContent	LL	PI	qu	failureMode	SPT	depth	elevation	locationID
2	117	21	54	26	0.88	S/S		2.5	61.8	BORING NO. 4
3	119	25	39	12	0.53	YLD		7.5	56.8	BORING NO. 4
4	130	24			0.63	MS		12.5	51.8	BORING NO. 4
5	129	32			0.35	MS		17.5	46.8	BORING NO. 4
6	113	35	67	34	0.74	S/S		22.5	41.8	BORING NO. 4
7	112	32	75	37	1.09	S/S		27.5	36.8	BORING NO. 4
8	113	52	61	36	0.42	S/S		32.5	31.8	BORING NO. 4
9	109	37	76	44	1.18	S/S		37.5	26.8	BORING NO. 4
10	136	20	41	20	1.12	MS		42.5	21.8	BORING NO. 4
11	126	20	41	20	5.56	MS		47.5	16.8	BORING NO. 4
12	124	22	26	8	2.2	MS		52.5	11.8	BORING NO. 4
13	132	26	28	10	1.84	MS		57.5	6.8	BORING NO. 4
14							7	62.5	1.8	BORING NO. 4
15							20	67.5	-3.2	BORING NO. 4
16							57	72.5	-8.2	BORING NO. 4
17							50	77.5	-13.2	BORING NO. 4
18							23	82.5	-18.2	BORING NO. 4
19							33	87.5	-23.2	BORING NO. 4
20		16	34	19			75	92.5	-28.2	BORING NO. 4
21			28	7			100	97.5	-33.2	BORING NO. 4
22		25	33	15			89	102.5	-38.2	BORING NO. 4
23							100	107.5	-43.2	BORING NO. 4
24		24	29				100	112.5	-48.2	BORING NO. 4
25		24	38	19			80	117.5	-53.2	BORING NO. 4
26		22	26				100	122.5	-58.2	BORING NO. 4
27		22	42	23			88	127.5	-63.2	BORING NO. 4

	A	B	C	D	E	F	G	H	I	J
1	date	locationID	longitude	stationingNo	location	latitude	projectName	projectID_A	projectID_B	elevation
2	7/14/2004	BORING NO. 4	92° 06' 35.02"W	486+36	35" RT Adopt. CL	32° 11' 49.85"N	[redacted]			64.3
3	9/28/2004	BORING NO. 5	92° 06' 36.72"W	488+36	65' LT Adopt. CL	32° 11' 50.00"N	[redacted]			66.5
4	7/21/2004	BORING NO. 6	92° 06' 37.54"W	490+36	42' RT Adopt. CL	32° 11' 52.16"N	[redacted]			64.4

KEY TAKEAWAYS – PART II

- Using standardized digital data formats are important to leverage the power of big data.

- If you are **<u>ready to change</u>** your current data workflows in the era of big data, **DIGGS** is a great tool that allows geotechnical engineers to leverage the power of big data in Geotechnical Engineering:

    ✓ Learn more about **DIGGS**: https://github.com/DIGGSml

    ✓ Learn more about **DIGGS Implementation** using **pyDIGGS**: https://pydiggs.readthedocs.io

- If you are **<u>NOT ready to change</u>** your current data workflows with PDF boring logs and lab testing sheets in the big data age, **AI** can efficiently extract data from both typed and hand-written PDFs. This unlocks immense value from historical archives for any ongoing and future projects.

Geosyntec consultants

SPECIAL THANKS TO

DOTD

LOUISIANA DEPARTMENT OF TRANSPORTATION & DEVELOPMENT

Geosyntec consultants

THANK YOU!

xin.peng@geosyntec.com