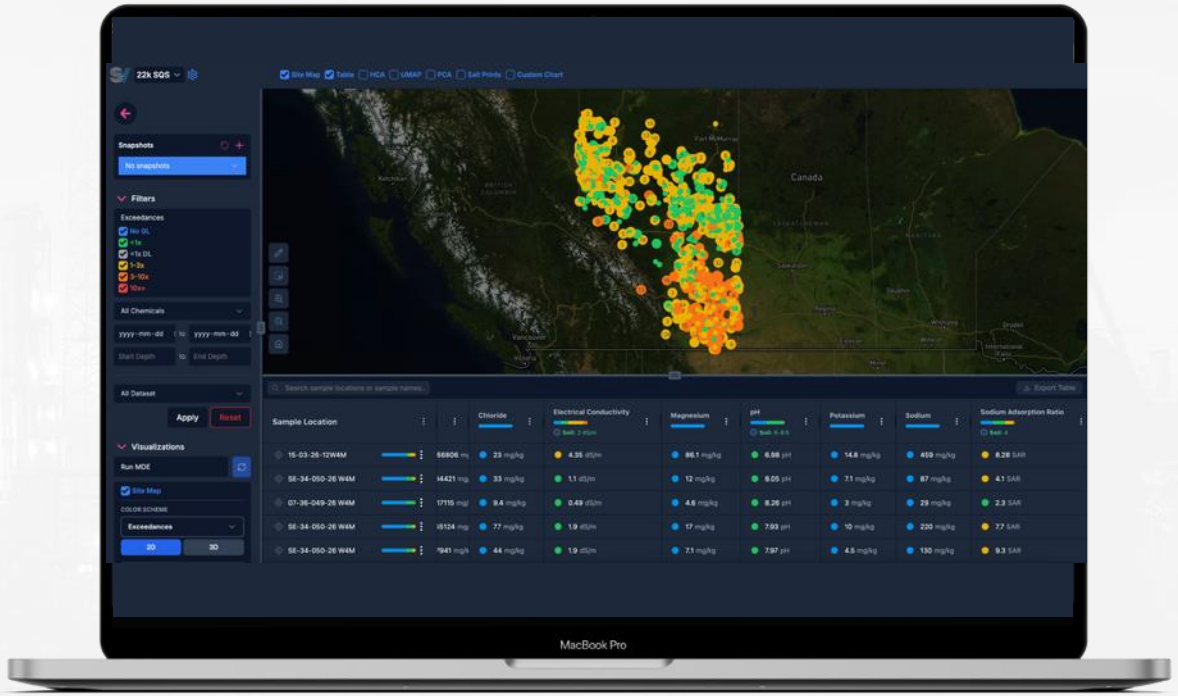
statvis Trust Through Data

Machine Learning: The Secret Weapon in Unraveling Salinity Source Mysteries at Contaminated Sites to Support Regulatory Closure

Court Sandau, PhD, PChem, FRSC
Co-Founder and Principal Scientist



From raw data to actionable insight.
In a fraction of the time.

Remediation Technologies Symposium 2025

REMTECH 2025, Banff, Alberta

statVIS Trust Through Data

Chemistry Matters Inc.

- Niche chemistry consulting company specializing in environmental forensics, geoforensics, and biomonitoring.
- Big environmental datasets have become our specialty to merge chemistry understanding with statistical interpretation
- Virtual office based, located in Canada, US, and UK
- Established in 2011

Statvis Analytics Inc.

- Software company specializing in data and visualization software
- Turning all data and statistics into easy-to-understand graphics and pictures
- *Environmental data intelligence* and our EDI pipeline
- Knowledgebase vs. database
- Helping people communicate science
- Established in 2018

Environmental Data Intelligence Pipeline

Digitize

Create a comprehensive knowledgebase

Extract

Hard data and contextual information

Query

Ask questions to fill or identify data gaps

Analyze

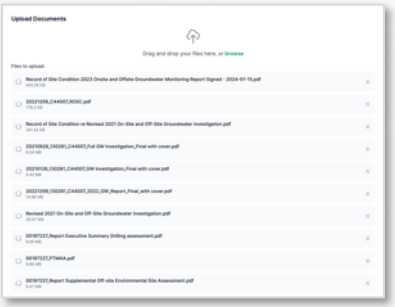
Determine responsibility and risk

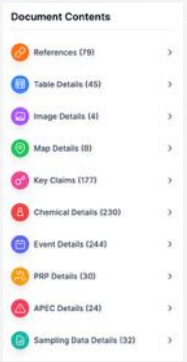
Visualize

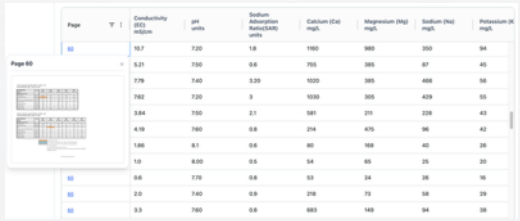
Clear messages for stakeholder buy-in

Report

Automatically generate built-for-purpose deliverables







statVIS Trust Through Data


DATA JAIL
PDF

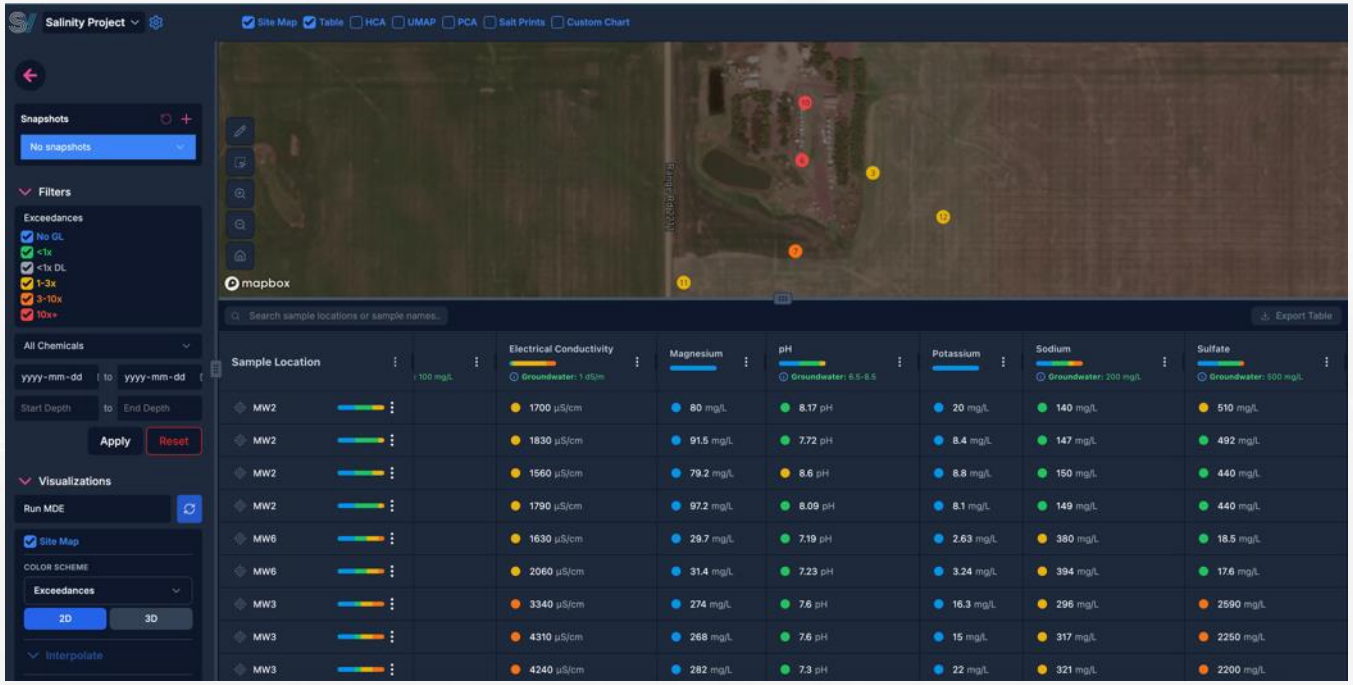
How are you collecting your data?

- All information from your site could be important/useful
- All file types
- Human memory is inefficient and can be lost when people leave the company
- Leaving it in your inbox means it could be lost and not part of company knowledge
- Comprehensive as possible
- You have paid for chemical analysis and consulting work and opinions – data should always be available
- **Do you have it all?**

statVIS Trust Through Data

The Site

- Demonstrative dataset only
- Sources and most of the data from a real-world site
- Location of monitoring wells changed to protect the privacy of the site

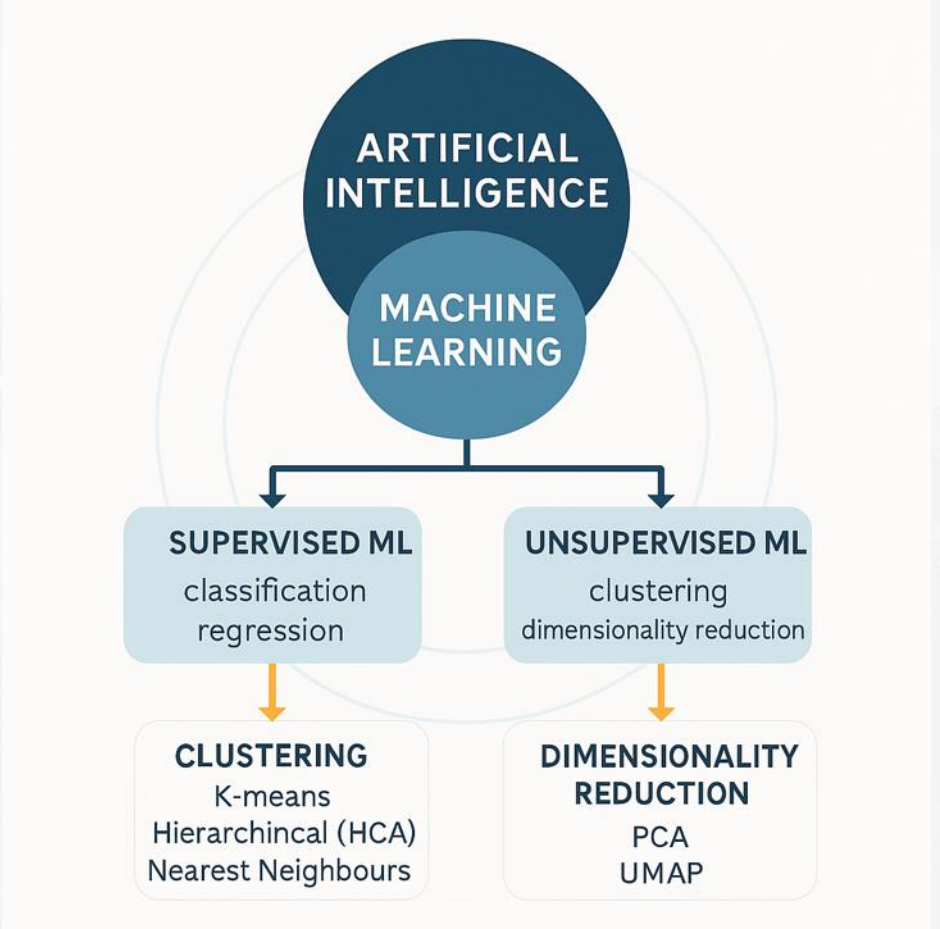


First off – What is Machine Learning?

Artificial Intelligence – computer systems capable of performing tasks that normally require human intelligence – reasoning, learning, decision making

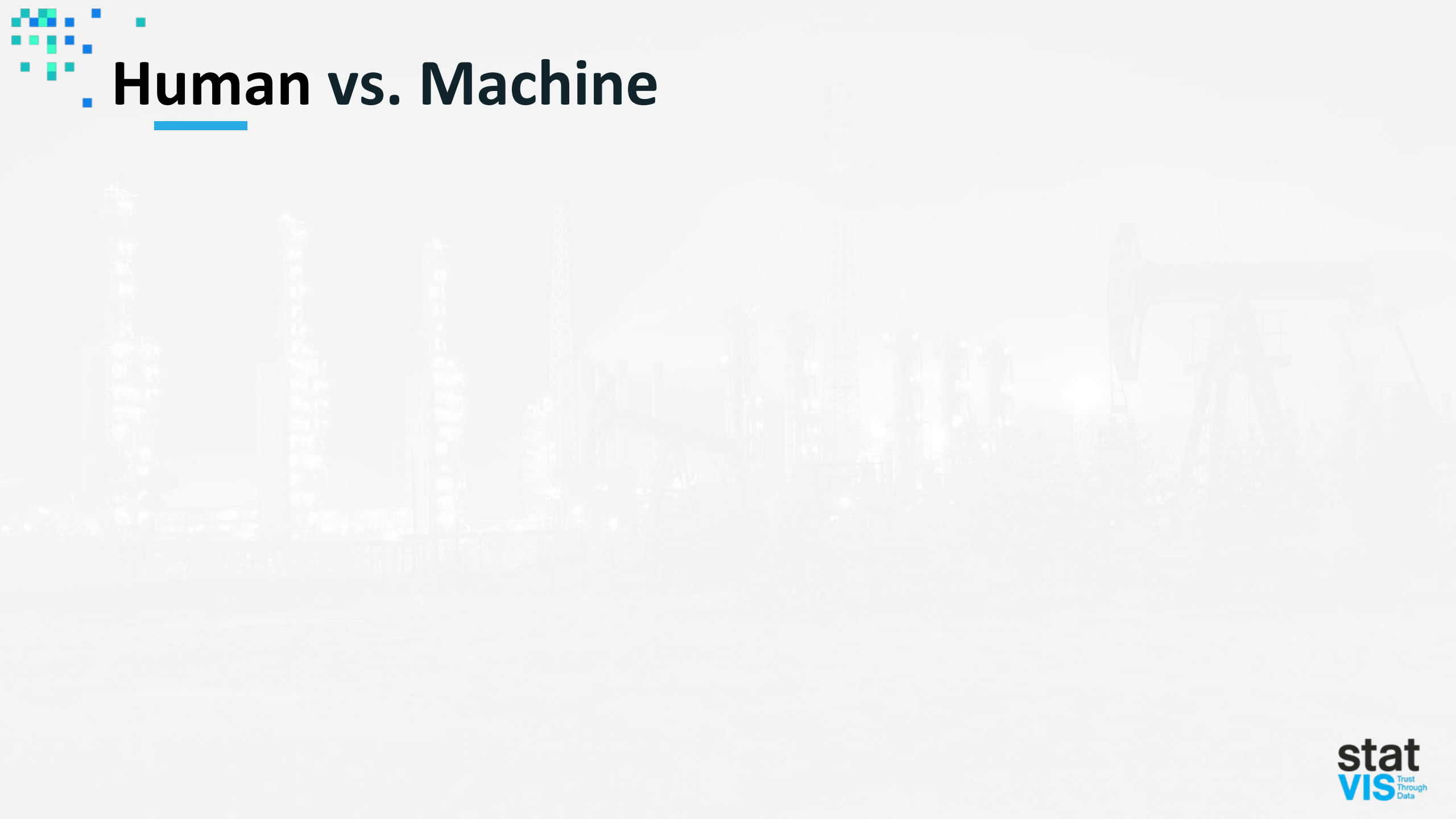
Machine Learning – subset of AI focused on algorithms that learn from data and can make predictions without explicit programming. There can be supervised and unsupervised learning.

Unsupervised does not mean unattended and on its own!



statVIS Trust Through Data

Human vs. Machine



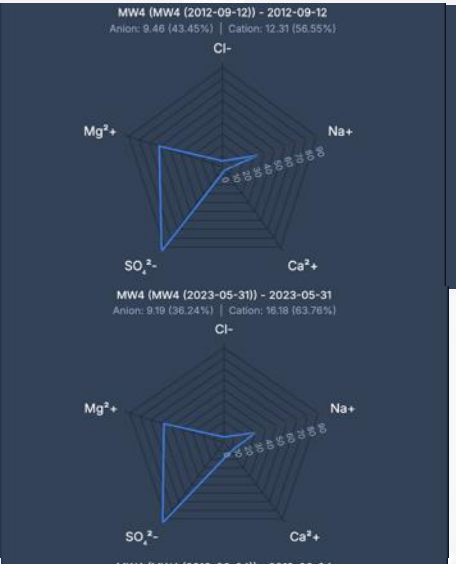
Human vs. Machine

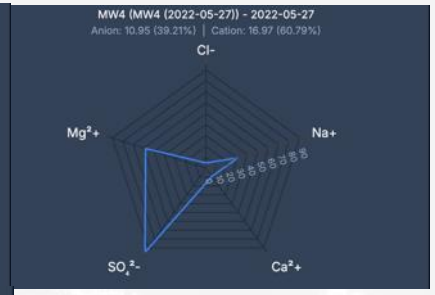
MW1	14.36 mg/L	1813 mg/L	2600 µS/cm	2.5 mg/L	8 pH	99.4 mg/L	1048 mg/L	15 mg/L
MW1	14 mg/L	1800 mg/L	1119 µS/cm	1.5 mg/L	8.31 pH	99 mg/L	980 mg/L	0.24 mg/L
MW1	15 mg/L	1800 mg/L	1119 µS/cm	1.4 mg/L	7.83 pH	84 mg/L	1000 mg/L	30 mg/L
MW1	13 mg/L	1900 mg/L	1119 µS/cm	2.9 mg/L	8.37 pH	82 mg/L	910 mg/L	32 mg/L

statVIS Trust Through Data

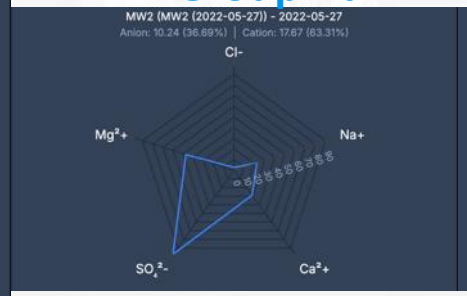
Human vs. Machine

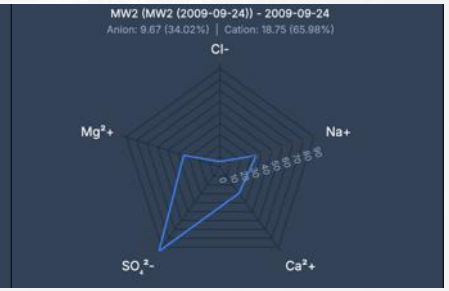
Group 1



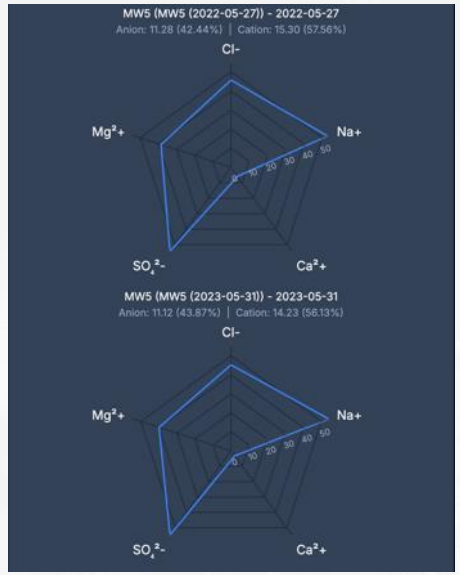


Group 1b

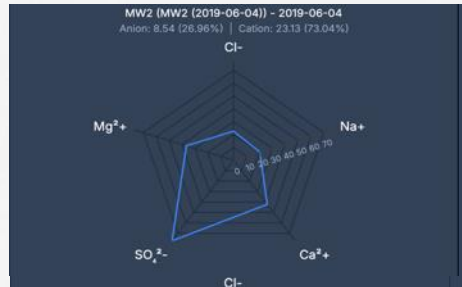


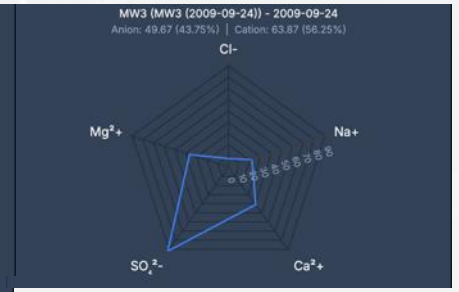


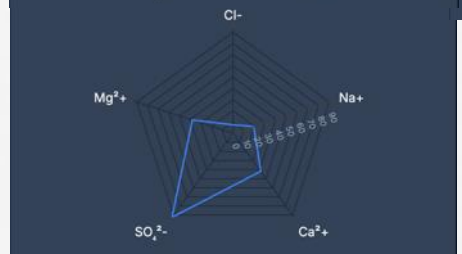
Group 2



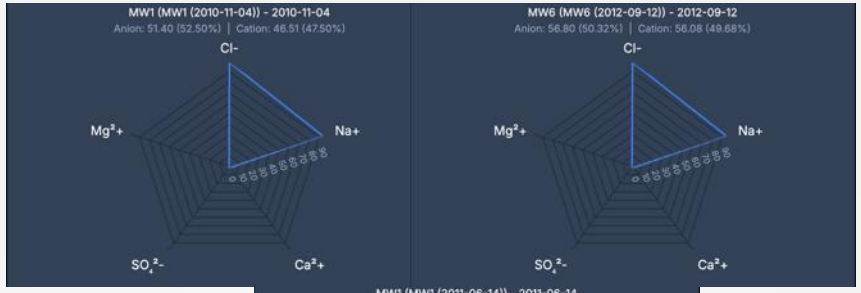
Group 3

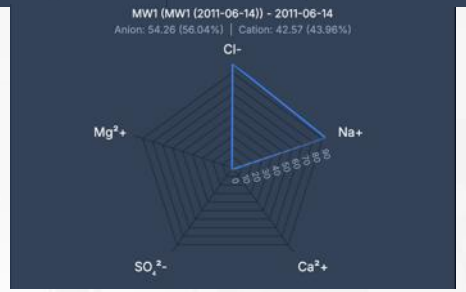






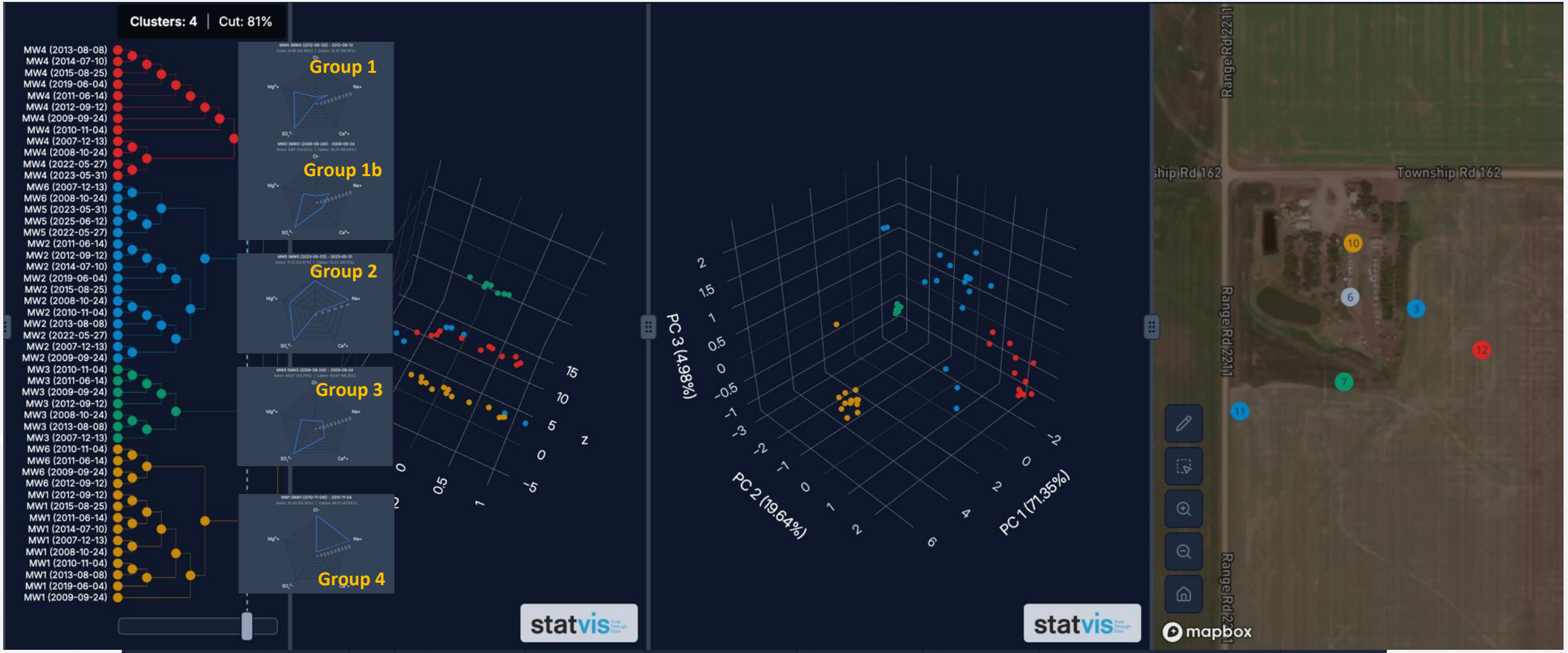
Group 4





statVIS Trust Through Data

Human vs. Machine



MW1	14 mg/L	1800 mg/L	1119 µS/cm	1.5 mg/L	8.31 pH	99 mg/L	980 mg/L	0.24 mg/L
MW1	15 mg/L	1800 mg/L	1119 µS/cm	1.4 mg/L	7.83 pH	84 mg/L	1000 mg/L	30 mg/L
MW1	13 mg/L	1900 mg/L	1119 µS/cm	2.9 mg/L	8.37 pH	82 mg/L	910 mg/L	32 mg/L

statVIS Trust Through Data

Environmental Data Intelligence Pipeline

Digitize
Create a comprehensive knowledgebase

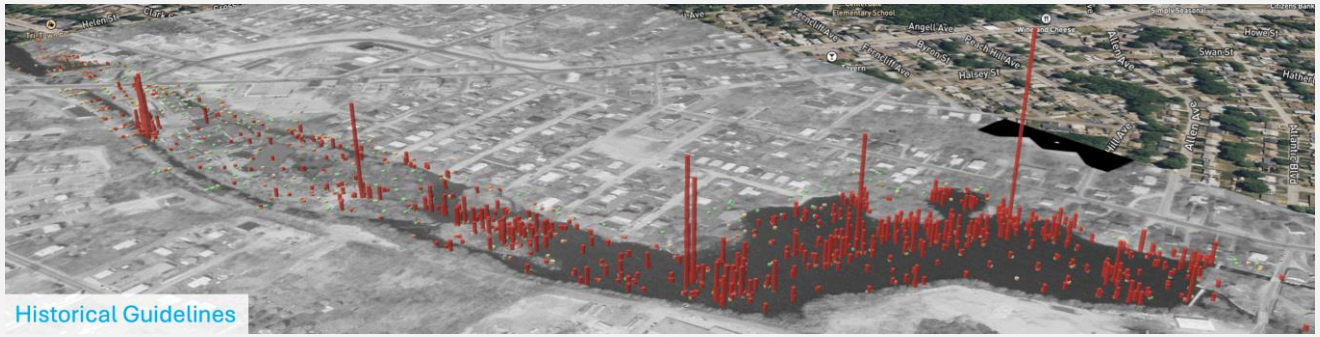
Extract
Hard data and contextual information

Query
Ask questions to fill or identify data gaps

Analyze
Determine responsibility and risk

Visualize
Clear messages for stakeholder buy-in

Report
Automatically generate built-for-purpose deliverables



statVIS Trust Through Data

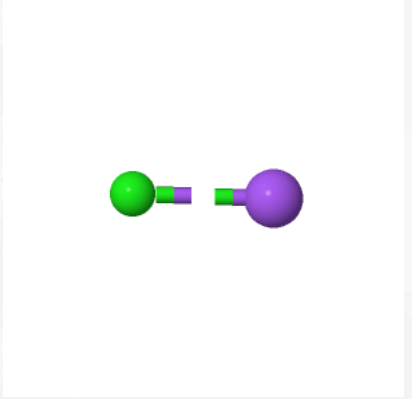
Anthropogenic and Natural Salinity Sources

- **Oil and Gas** – produced water, drilling fluids, leaks – mostly NaCl
- **Road Maintenance and Deicing** – NaCl and $CaCl_2$ applied for ice control and end up near roads and waterways due to runoff
- **Agricultural Fertilizers** – application of K, Ca, NH_4 chloride fertilizers
- **Natural Saline Soils** – dissolved salts from glacial till and bedrock deposits carried to surface by water and water evaporates

statVIS Trust Through Data

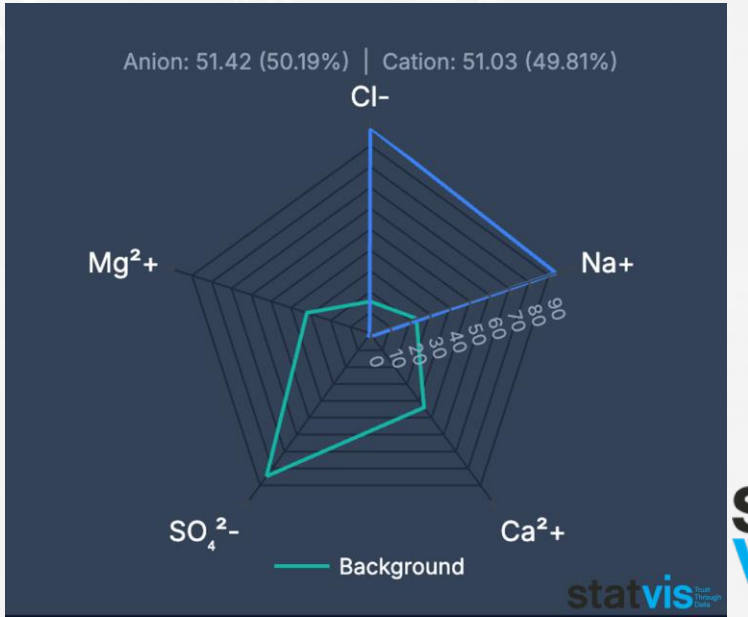
What is there to Fingerprint?

- Using normalized (major) cations and anions and simple radar plot



Radar plots can be used to:

- Can overlay most likely background samples for comparison
- Visualize fingerprints of background and impacted samples (various sources)
- Human Validation of UMAP and HCA branches based on radar plot similarities
- Identify major salinity cations/anions in various sources



statVIS Trust Through Data

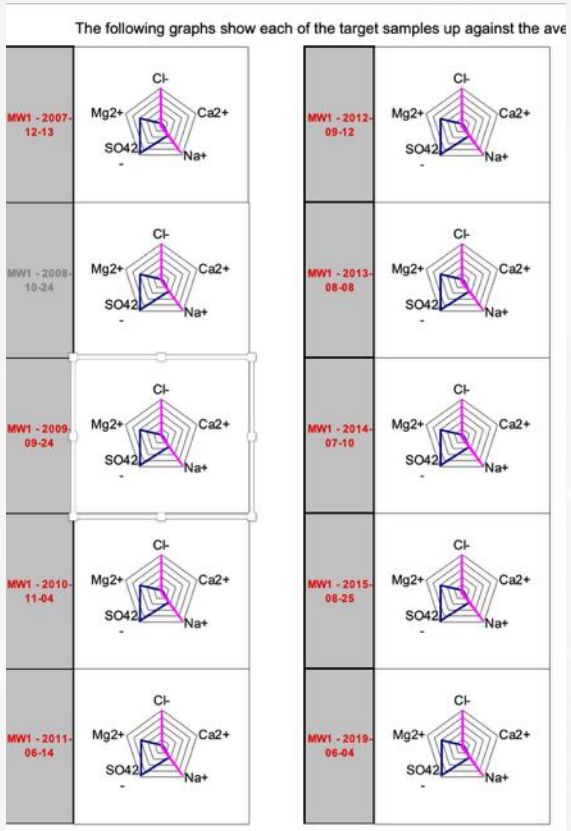
Began Using Salinity Fingerprints in 2006

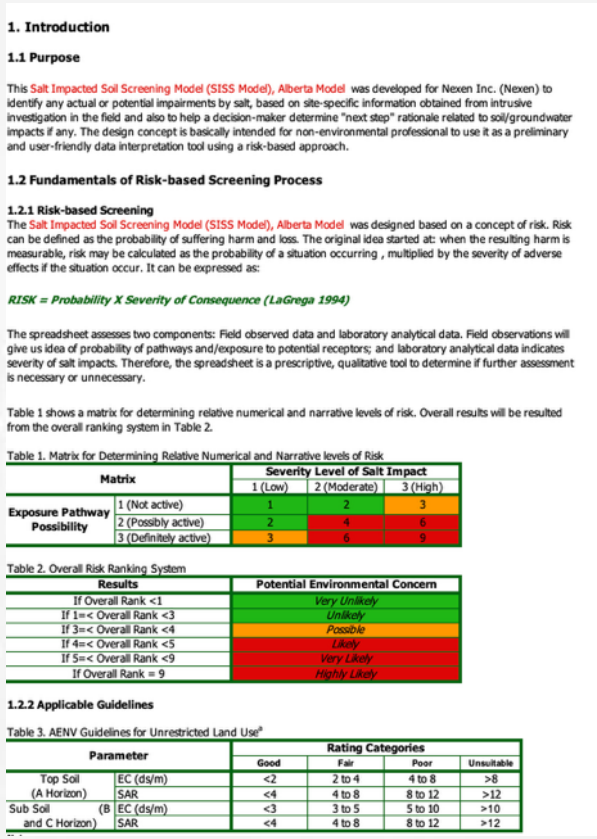
SISS - MODEL

SALT-IMPACTED SOIL SCREENING MODEL
ALBERTA
(AB-SISS MODEL - B)

SITE LOCATION	
LEGAL DESCRIPTION	
DATE OF DATA ENTRY	

Date of Version: November 13, 2006





1. Introduction

1.1 Purpose

This Salt Impacted Soil Screening Model (SISS Model), Alberta Model was developed for Nexen Inc. (Nexen) to identify any actual or potential impairments by salt, based on site-specific information obtained from intrusive investigation in the field and also to help a decision-maker determine "next step" rationale related to soil/groundwater impacts if any. The design concept is basically intended for non-environmental professional to use it as a preliminary and user-friendly data interpretation tool using a risk-based approach.

1.2 Fundamentals of Risk-based Screening Process

1.2.1 Risk-based Screening

The Salt Impacted Soil Screening Model (SISS Model), Alberta Model was designed based on a concept of risk. Risk can be defined as the probability of suffering harm and loss. The original idea started at: when the resulting harm is measurable, risk may be calculated as the probability of a situation occurring , multiplied by the severity of adverse effects if the situation occur. It can be expressed as:

RISK = Probability X Severity of Consequence (LaGrega 1994)

The spreadsheet assesses two components: Field observed data and laboratory analytical data. Field observations will give us idea of probability of pathways and/exposure to potential receptors; and laboratory analytical data indicates severity of salt impacts. Therefore, the spreadsheet is a prescriptive, qualitative tool to determine if further assessment is necessary or unnecessary.

Table 1 shows a matrix for determining relative numerical and narrative levels of risk. Overall results will be resulted from the overall ranking system in Table 2.

Table 1. Matrix for Determining Relative Numerical and Narrative levels of Risk

Matrix		Severity Level of Salt Impact		
		1 (Low)	2 (Moderate)	3 (High)
Exposure Pathway Possibility	1 (Not active)	1	2	3
	2 (Possibly active)	2	4	6
	3 (Definitely active)	3	6	9

Table 2. Overall Risk Ranking System

Results	Potential Environmental Concern
If Overall Rank <1	*Very Unlikely*
If 1=< Overall Rank <3	*Unlikely*
If 3=< Overall Rank <4	*Possible*
If 4=< Overall Rank <5	*Likely*
If 5=< Overall Rank <9	*Very Likely*
If Overall Rank = 9	*Highly Likely*

1.2.2 Applicable Guidelines

Table 3. AENV Guidelines for Unrestricted Land Use[6]

Parameter		Rating Categories			
		Good	Fair	Poor	Unsuitable
Top Soil (A Horizon)	EC (ds/m)	<2	2 to 4	4 to 8	>8
	SAR	<4	4 to 8	8 to 12	>12
Sub Soil (B and C Horizon)	EC (ds/m)	<3	3 to 5	5 to 10	>10
	SAR	<4	4 to 8	8 to 12	>12

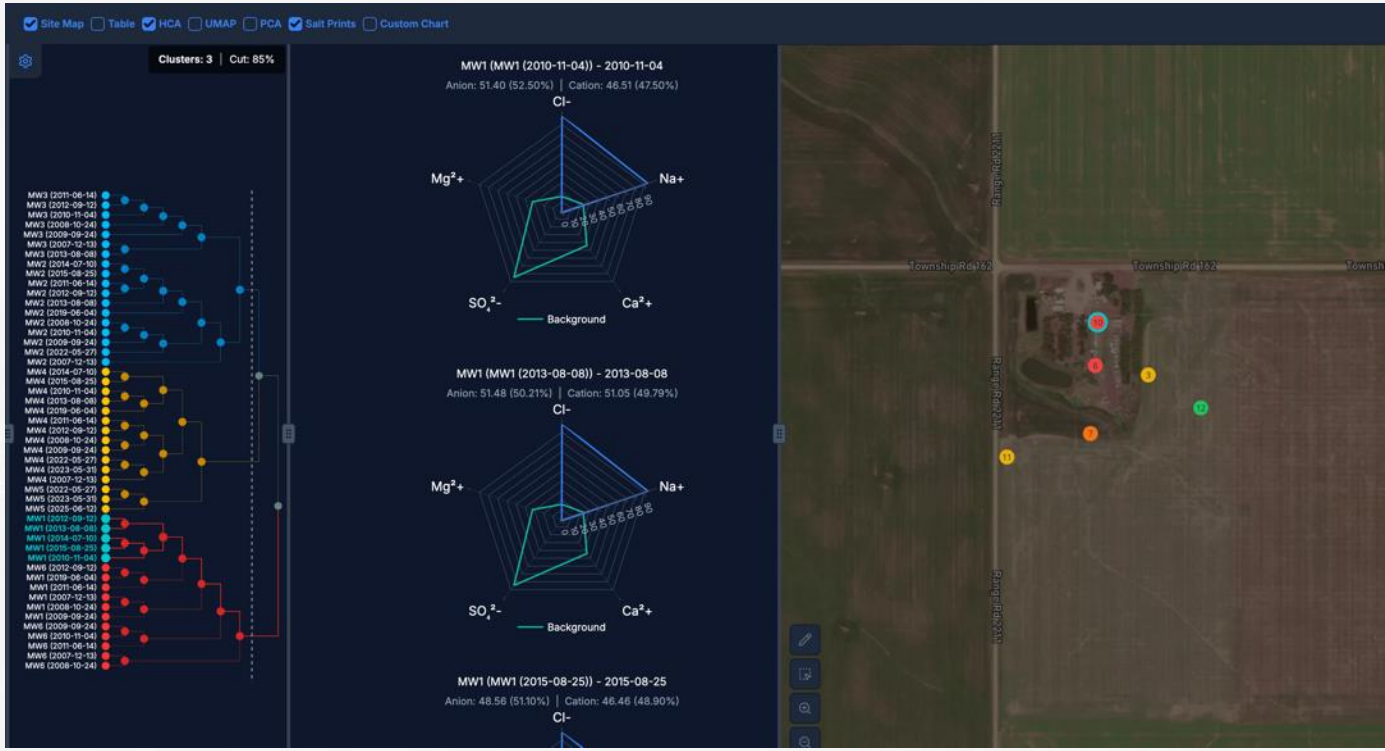
- Developed salinity fingerprinting to evaluate risk at salt impacted sites
- First time using radar plots as a tool to identify background impacts (above guidelines) from produced water impacts

statVIS Trust Through Data

Fingerprint Salinity Sources

Know your sources for your site:

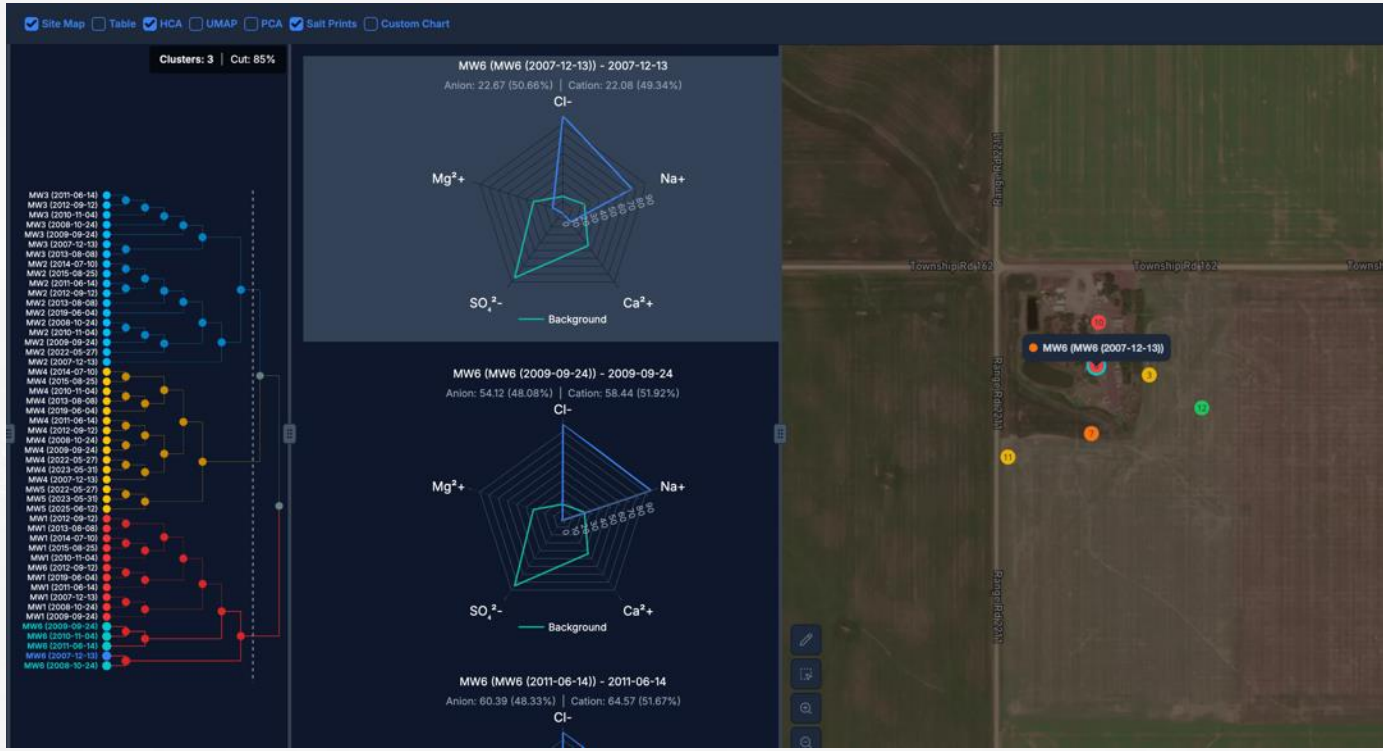
- Impacted NaCl spill / release with a dominant Na/Cl presence



Fingerprint Salinity Sources

Know your sources for your site:

- Impacted NaCl spill / release with a dominant Na/Cl presence
- Less Na, More Ca/Mg



Fingerprint Salinity Sources

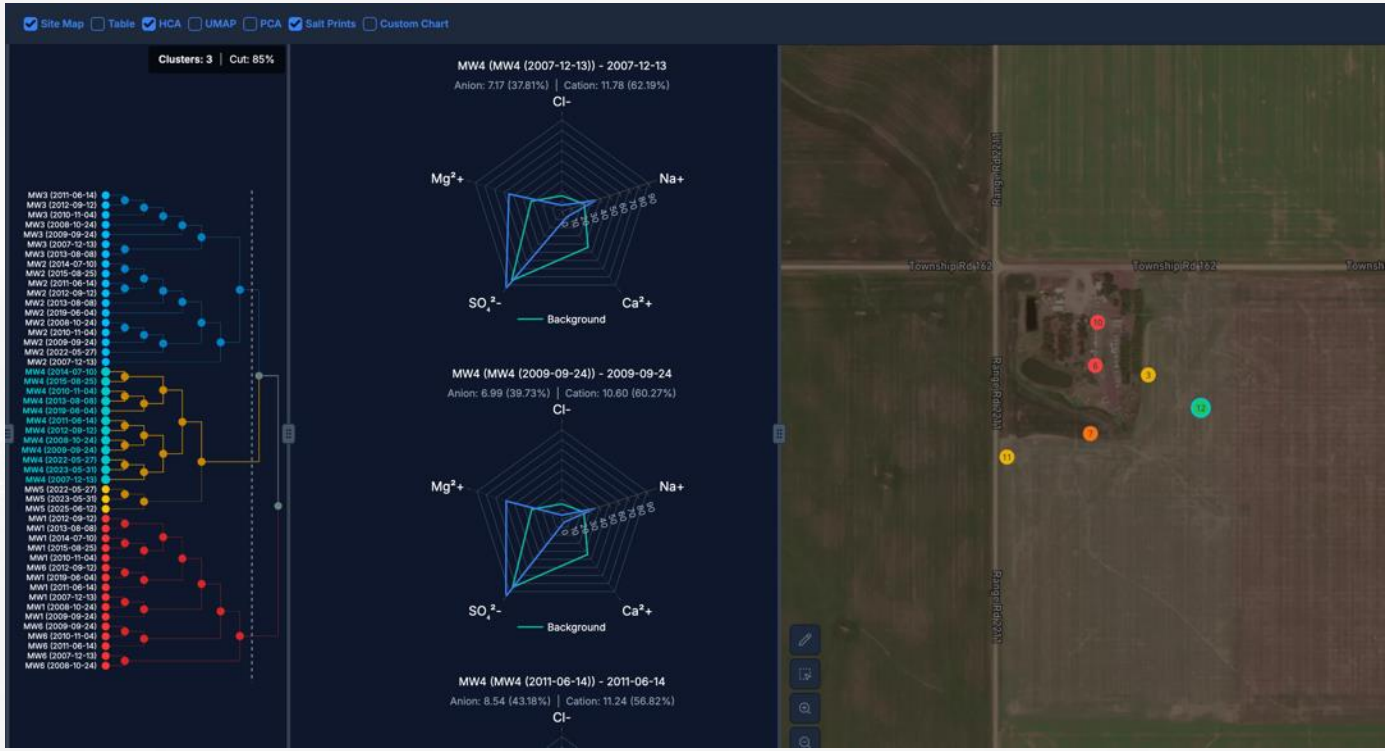
Know your sources for your site:

- Background salinity pattern with a Ca/Mg/Na Sulfate dominant pattern

Fingerprint Salinity Sources

Know your sources for your site:

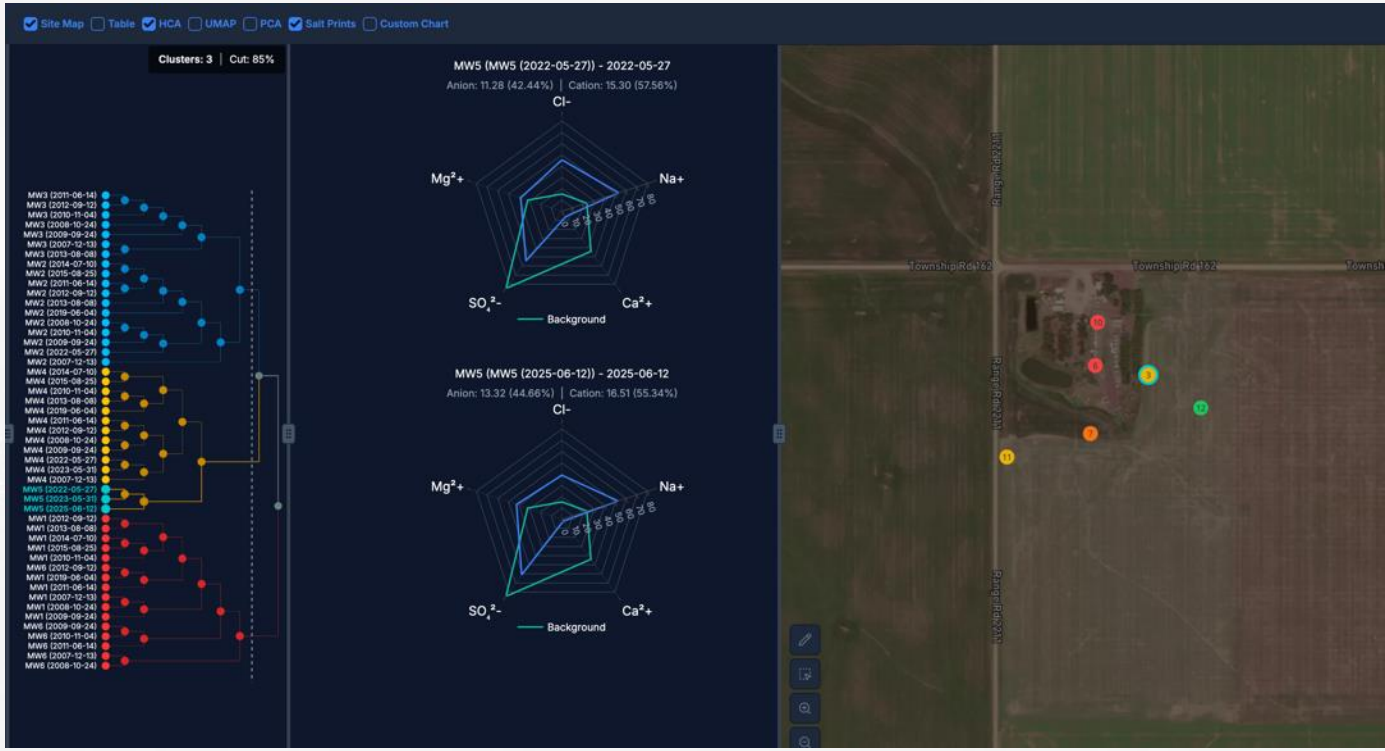
- Mg/Na sulfate dominant pattern



Fingerprint Salinity Sources

Know your sources for your site:

- Mixed pattern having NaCl on top of background

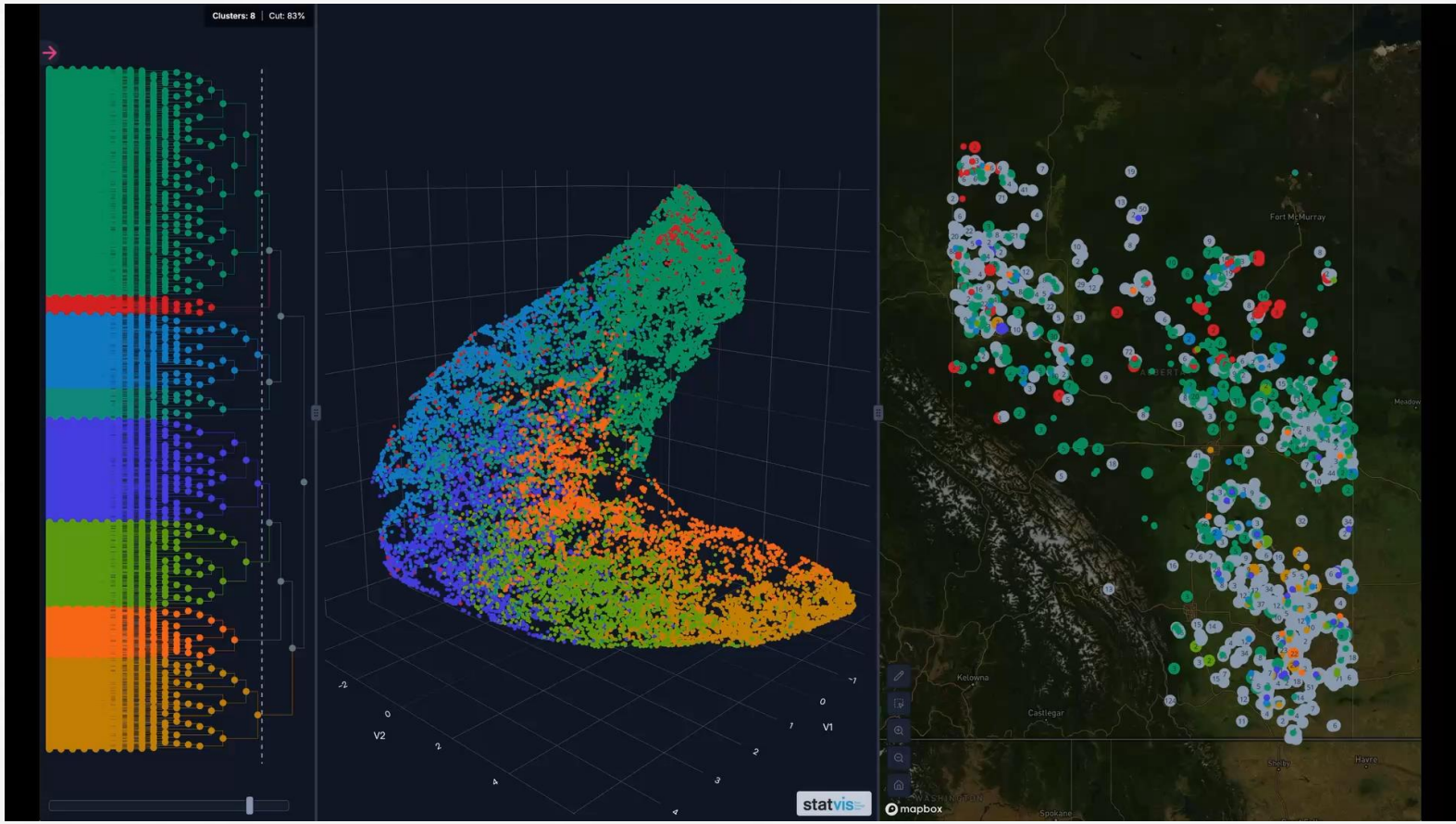


Interactivity of HCA, UMAP/PCA, fingerprint graphs (radar plots/histograms) and geospatial awareness

statVIS Trust Through Data

Alberta Background Database Project

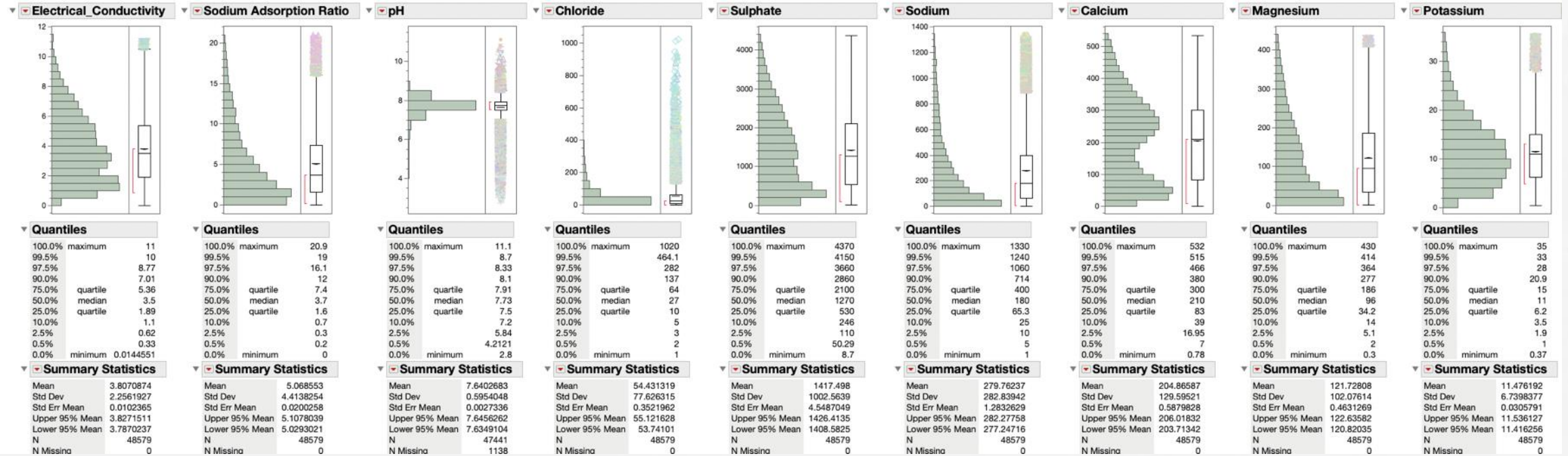
- Master dataset for the project contains 224,902 records with EC, SAR, Cl, SO4, Mg, Ca, K, Na.
- 74,943 samples remained after data exclusions
- Applied ML salinity forensic workflow
  - Identified chloride data node, which was removed to be conservative



statVIS Trust Through Data

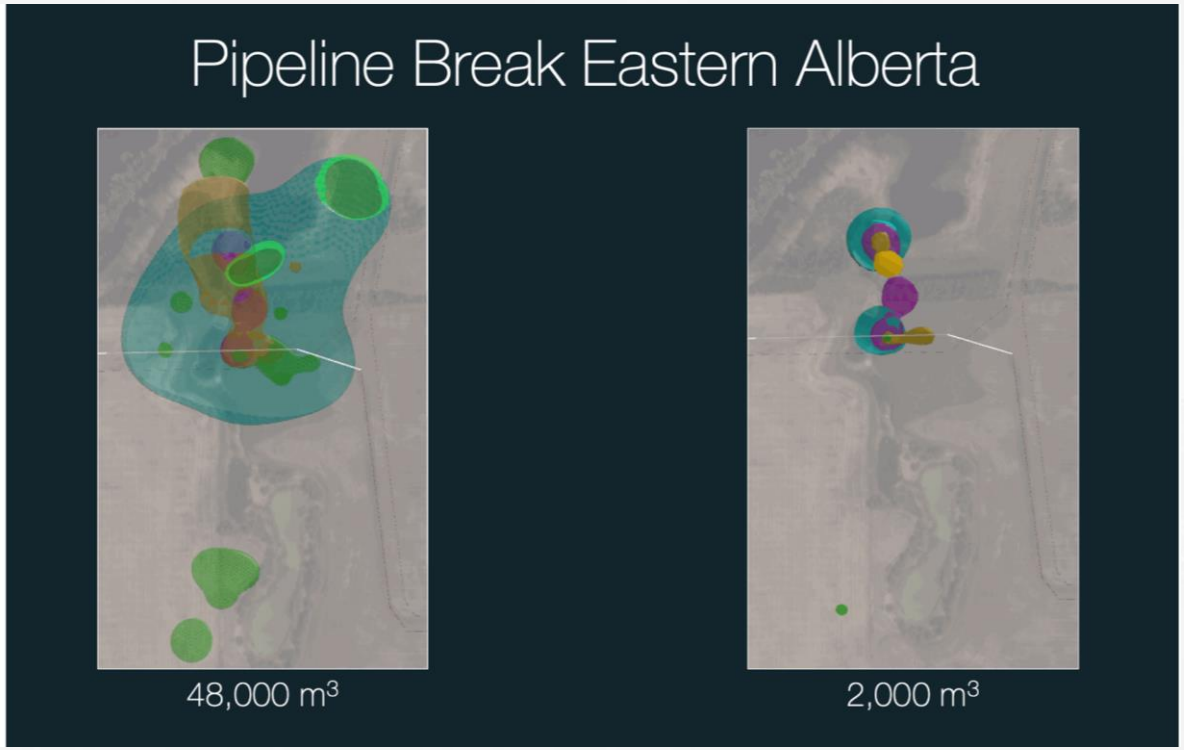
Alberta Background Database

- 48,579 data records remained to characterize background salinity patterns and levels in Alberta
- 5 primary patterns of natural salinity in Alberta



Keys to Regulatory Closure

- Transparency
  - Is anything missing (*hiding anything*)?
  - Statistics, summaries, graphs all validated
- Communication
  - What is written matches the data supplied
  - Any data gaps or anything left to interpret?

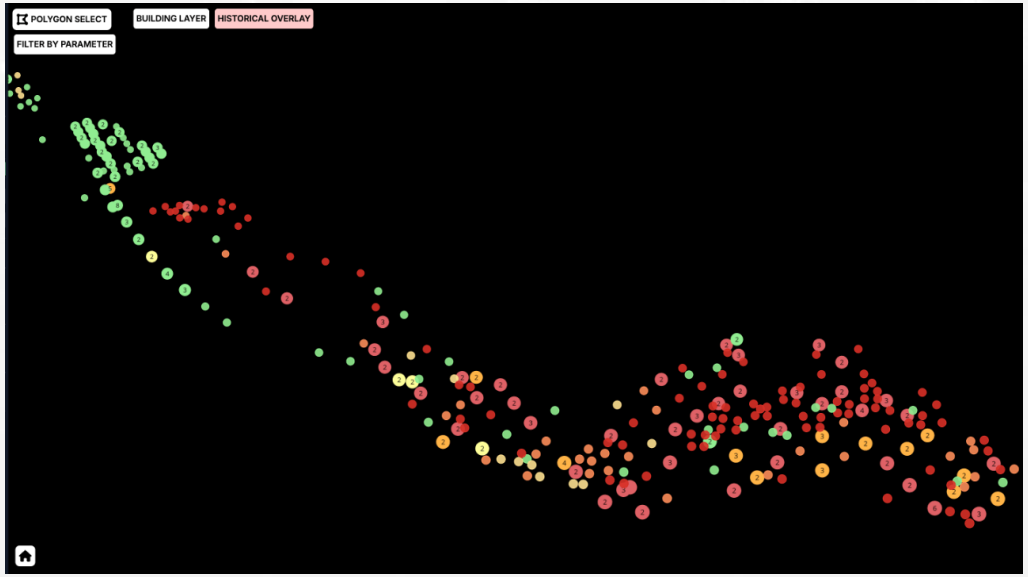


Statvis can provide portal (website) to regulator/client/opposing counsel and experts for maximum transparency

statVIS Trust Through Data

Conclusions

- All contaminants have chemical fingerprints that can be used to identify sources, even salinity
- ML tools can process large (and small) datasets through workflows that can confidently identify sources
- Users of ML need to understand (and potentially defend) those identifications



- It needs to be communicated effectively to client, regulators or opposing parties

statVIS Trust Through Data

Thank you for your attention

Contact Info:

Court Sandau

phone: +1.403.669.8566

email: csandau@chemistry-matters.com

or: court@statvis.com

www.chemistry-matters.com

www.statvis.com

Follow Statvis Analytics on LinkedIN for the latest.