arag✳rn

Integration of soil health indicators into ARAGORN's holistic and tiered approach to assessment of management options for contaminated sites.

Hans Peter H. Arp (NGI) and Ali Kanso (Microhumus)

Contact: hans.peter.arp@ngi.no and a.kanso@microhumus.fr

https://aragorn-horizon.eu/

SOILveR Webinar, March 20, 2026

aragorn

Achieving Remediation And GOverning Restoration of contaminated soils Now

Goal

Guidance for land managers in Europe dealing with urban, peri-urban and rural contaminated soils with the focus on **hotspot** pollution:

- PFAS
- Organochlorine/bromides [OCBs]
- Petroleum and coal compounds (PETCO)
- Metals

PC

WPLs

Partners

Project Outcomes:
- Mapping and monitoring tools for source identification and effective remediation
- From regrettable to restorative remediation of land (holistic socio-economic and resilience assessment)
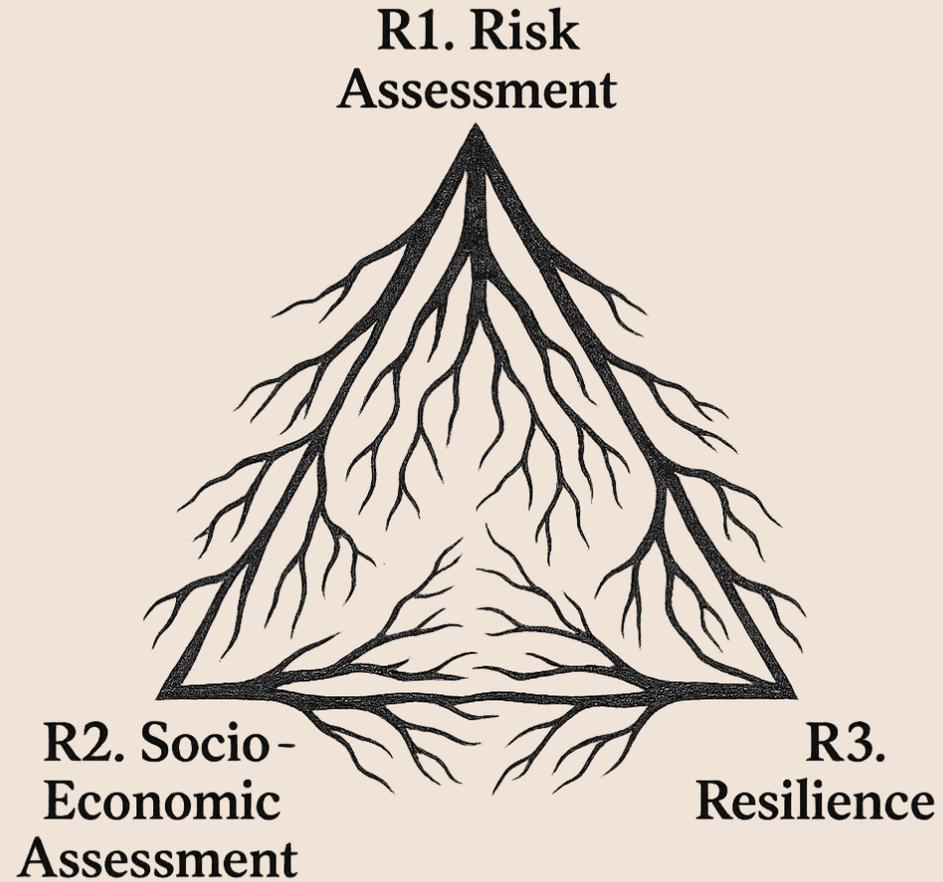- Planning and decision-making tools

Field trials:

9x test sites for remediation: Including field-testing, human and environmental Risk assessments (AFFF site and Fluoropolymer production)

- **Budget:** € 7.1 M

- **Duration:** October 1st 2023 – September 30th 2027

- **18 partners & 12 countries** Research Institutes, SMEs, Industry, Public Authorities

Three Frameworks

Three Tiers

| Tier level of analytic complexity	Definition	R1. Risk Assessment	R2. SEA	R3. Resilience
T1	Qualitative analysis.	Just concentration and assumptions on soil properties	No parameters in monetary terms, only some parameters quantified.	Identification of key resilience indicators for qualitative decision making
T2	Semi-quantitative analysis.	Bioavailability, more detailed site assessment	Many parameters quantified, but only some in monetary terms	Some resilience indicators quantified and modelled
T3	Fully quantitative analysis.	Detailed data on risk	All parameters quantified and can be used in monetary terms.	Multi-dimensional modelling on resilience indicators

Co-Creation towards stakeholder empowerment as the goal

➢ *Empower*: stakeholders are the final decision-makers. The goal is to support **decision-making and self-governance** among the community of stakeholders.

➢ *Collaborate*: stakeholders are a key partner in the design process; activities are focuses on co-generating ideas, co-designing, co-creating, co-implementing solutions. The goal is to **integrate ideas** from stakeholders into the decision-making as much as possible.

➢ *Involve*: stakeholders are directly involved in the project; project decisions reflect their ideas and concerns. The goal is to **gather requirements, build consensus, co-design** of elements of a broader solution.

➢ *Consult*: stakeholders provide feedback on alternatives **or solutions. The goal is receiving actionable feedback** on available options by stakeholders.

➢ *Inform*: stakeholders are informed of the project and solutions. The goal is to **generate interest** and increase awareness among stakeholders.
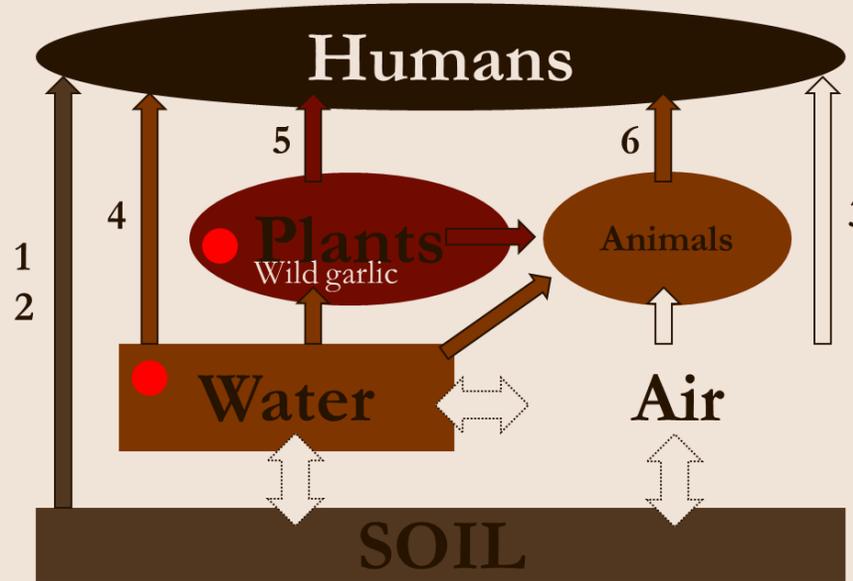
R1. Risk Assessment tools

| Spreading model | Human Health Risk assessment Modell | Terrestrial Toxicity |
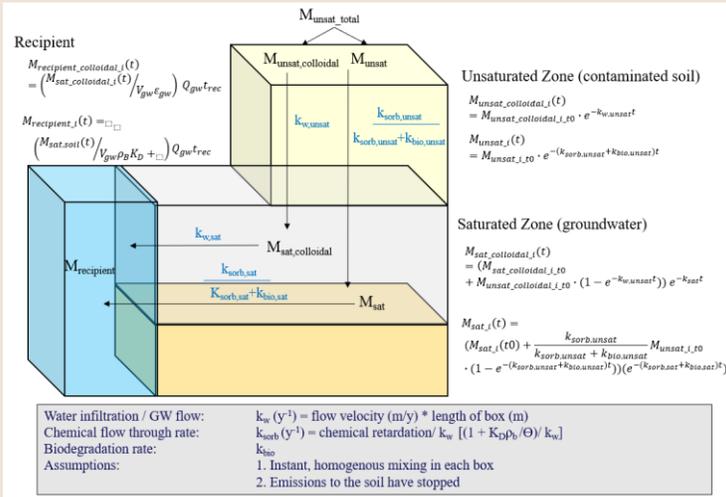
- C_{soil}, $C_{porewater}$ in the unsaturated and saturated zone,
- Concentration in the recipient
- As a function of time (t)

Exceedance of MTDI for different groups based on different exposure scenarios.

Determination of ecotoxicological thresholds of metals for different protection objectives

R2. Socioeconomic Assessment

Impact components (examples)	T1 Qualitative assessment		T2 Partially quantified assessment		T3 Fully quantified assessment, including monetisation of impacts	
	Baseline	Intervention	Baseline	Intervention	Baseline	Intervention
Risk (environment)	Environmental score	Change of the score	Env. concentration or stock estimate; any other observed effect	Change of concentration, stock, or observed effect	Non-monetary description of the env. risk	WTP to reduce the env. risk as described under the baseline
Risk (health)	Health score	Change of the score	Number of people with certain health effects	Change of the number of people with certain health effects	Non-monetary description of the health risk	WTP to reduce the health risk as described under the baseline
Resilience (environment)	Functionality score	Change of the score	Functionality score	Change of the score	Non-monetary description of env. resilience	WTP for improving env. resilience
Resilience (socio-ecological)	Functionality score	Change of the score	Functionality score	Change of the score	Non-monetary description of socio-ecol. resilience	WTP for improving socio-ecol. resilience
Employment	Employment score	Change of the score	Number of people (un-)employed	Change of the number of (un-)employed people	Number of people (un-)employed	WTP to increase employment
Housing	Housing score	Change of the score	Number of houses/apartments	Change of the number of houses	Housing values	Change of housing values
Infrastructure	Infrastructure score	Change of the score	Number of highways/train stations	Change of the number of highways/train stations	Non-monetary description of the benefits of infrastructure	WTP to improve infrastructure
Transport and logistics effort	Transport score	Change of the score	Number of working hours/amount of time/fuel needed	Change of the number of working hours/time/fuel needed	Costs for transport and logistics	Change of costs for transport and logistics
Monitoring effort	Monitoring score	Change of the score	Number of working hours/fte needed for monitoring	Change of the number of working hours/fte needed for monitoring	Costs for monitoring a certain area	Change of costs for monitoring a certain area

aragⰟrn

Funded by
The European Union

R3. Resilience
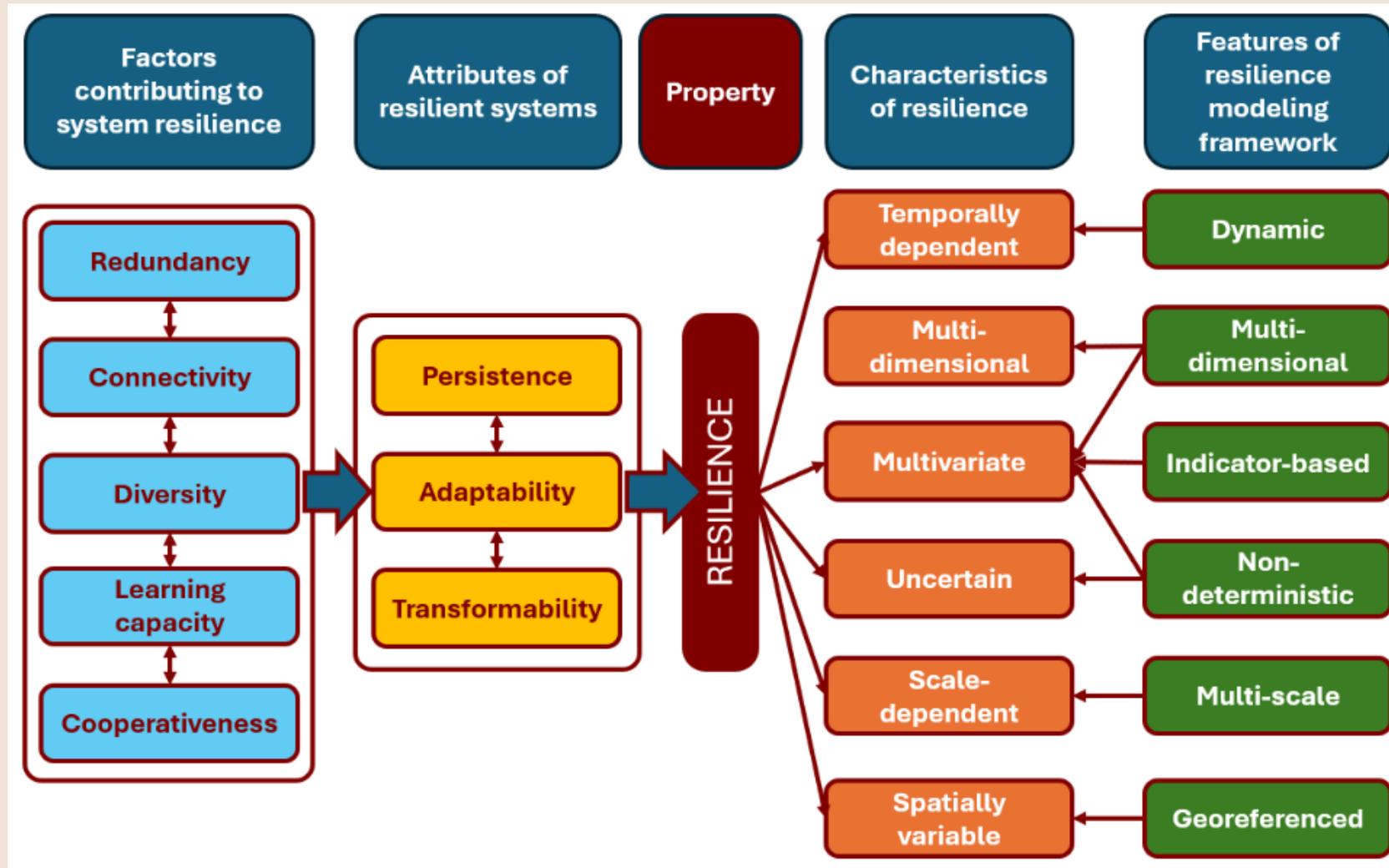
Holistic and Tiered Approacj
The ARAGORN Decision Tree

Step 0: When to investigate if soil is contaminated?
National regulation/practices
Step 1: Initial analysis of boundary conditions. Intentional
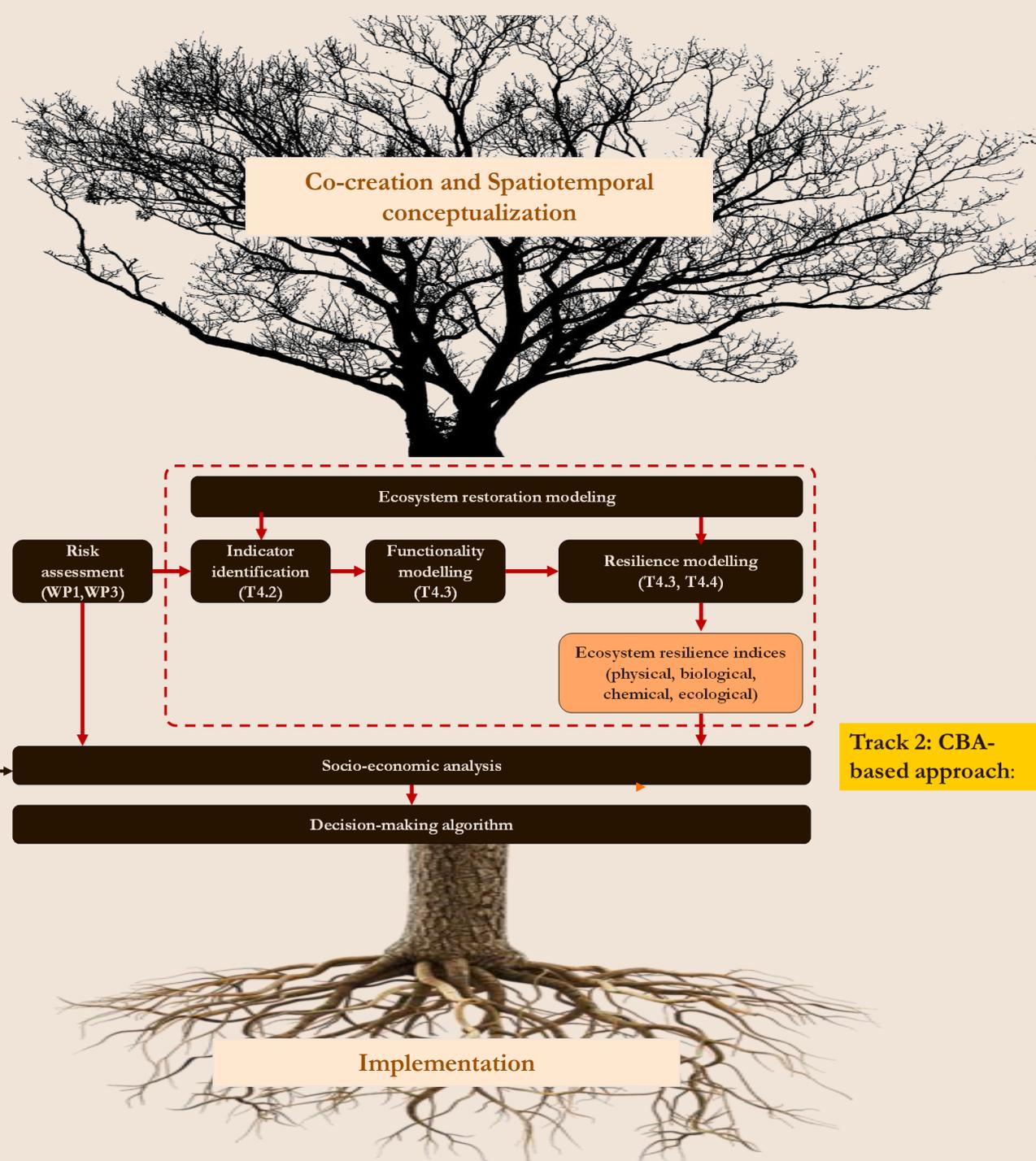future use of area, time frame, economy, stakeholder analysis etc.
Step 2: How to tell if risk is worth remediating

Decision for the need for remediation and restoration (the needs, benefits and goals)

Step 3: The engine of the R1 Risk- R2 Resilience- R3
SEA framework towards an R2 SEA based decision.

Following through on the SEA decision on best action

Step 4: Selection of indicators for evaluation of
remediation/restoration with stakeholders

Co-creation and Spatiotemporal conceptualization

Ecosystem restoration modeling

| Risk assessment (WP1,WP3) | Indicator identification (T4.2) | Functionality modelling (T4.3) | Resilience modelling (T4.3, T4.4) |

Ecosystem resilience indices (physical, biological, chemical, ecological)

Track 1: CEA-based approach:

Track 2: CBA-based approach:

Socio-economic analysis

Decision-making algorithm

Implementation

Example R3 Methodology : A 2-Step Indicator Selection

Tier1   Expert Based Approach (qualitative indicator selection)

- Key indicators will be aligned with international and European initiatives.
- Key indicator sets developed by the European Environment Agency (EEA) and by major EU research programmes such as the LANDMARK project.
- These initiatives define soil health through functional metrics that capture the soil's capacity to sustain ecosystem, like nutrient cycling, contaminant reduction, water retention, soil health, biodiversity index.

Tier 2/3   A Data-Driven Approach (quantitative)

- A fully data-driven pipeline integrates 34 heterogeneous studies into a harmonized, multi-parameter dataset, enabling cross-site comparisons.
- This unified, evidence-based framework turns scattered case studies into a coherent foundation for evaluating remediation outcomes with statistical rigor.

The dataset: A "Meta-Analysis" challenge

Objective

- To statistically model the impact of soil remediation on these soil functions over time.

Source

A compilation of 34 diverse case studies distributed as
follows:
- 25 experiments from the literature
- 4 sites from ARAGORN
- 5 sites from Microhumus
- This resulted in a dataset of 350 total samples.

Challenge

- Heterogeneity data : Different studies measured different parameters.
- Missing parameters of some samples. For example, few studies conducted microbial analysis.

Methodology : The analysis pipeline

Phase 1: Data Harmonization

Cleaning and standardizing the multi-study dataset. Consolidation into a unified 18-parameter.

Phase 2: Indicator Selection

Identification of underlying structures and relationships among the 18 parameters

Phase 3: Metric Construction

Development of composite soil function metrics of selected indicators.

Indicators were clustered together based on the MFA analysis.

Phase 4: Modeling

4A — Trajectory Analysis: OLS regression to evaluate whether soil remediation improves soil functions over time

4B — Predictive Modeling: Random Forest models to identify which indicators most strongly drive metric scores

Do Remediation Trajectories Differ Over Time?

We used Ordinary Least Squares (OLS) regression to see if the 'remediated' group behaved differently than the reference and control groups over time.

Trajectory (OLS Cross-Sectional): Metric_Contaminant_Reduction
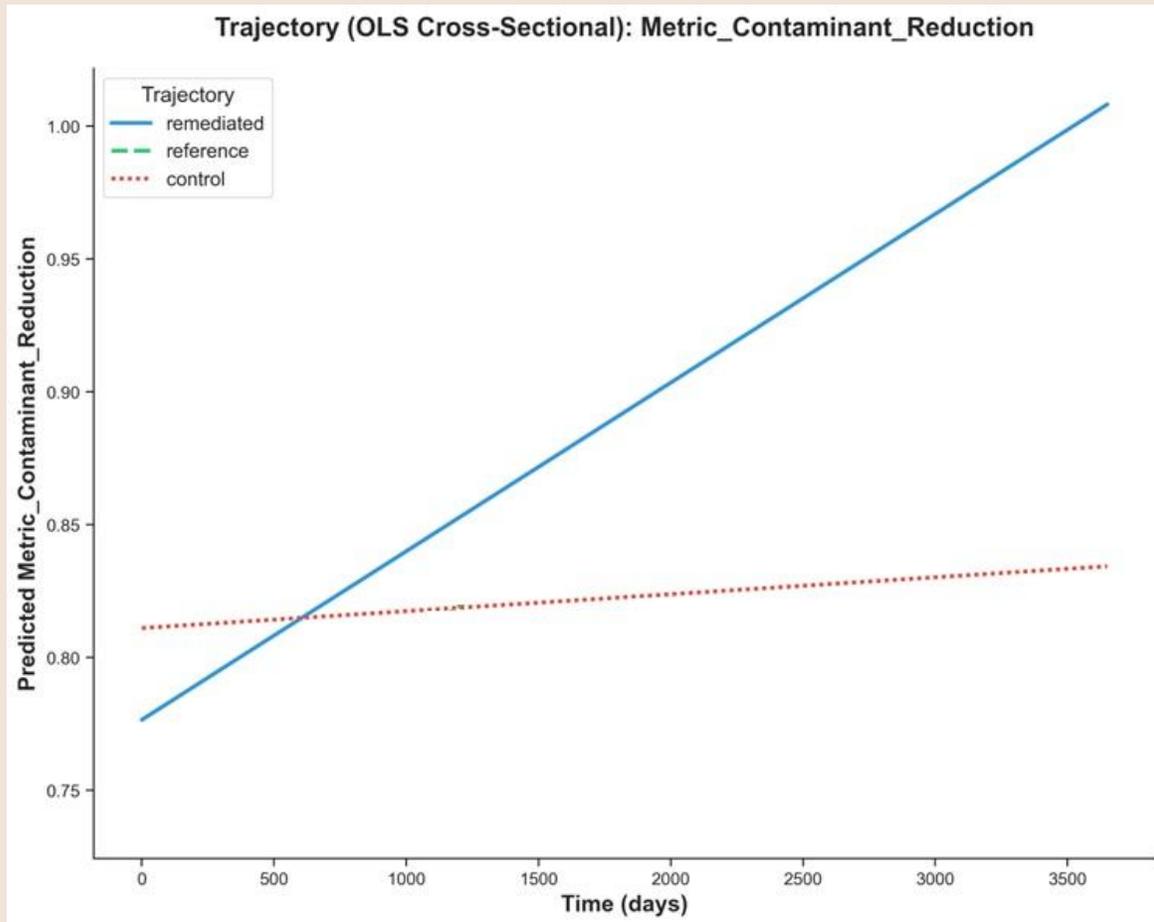
- The model was significant ($p = 0.0235$).

- However, the main driver was just Time_days ($p = 0.061$).

- Interpretation: Contamination decreases over time across *all* groups (likely due to natural attenuation), but the remediated trajectory wasn't statistically different from the others.

- The remediated (blue) and control (green) lines starts lower than the reference (orange) line.

- The OLS model confirmed that the blue line's upward slope is statistically significant, showing a trend towards recovery.

What Indicators *Predict* Soil Function? (Random Forest)

- We built Random Forest models to predict the *Metrics score* using the *Core Indicators* as features.
- This tells us which indicators are the most important.

Random Forest Feature Importances for Metric_Contaminant_Reduction
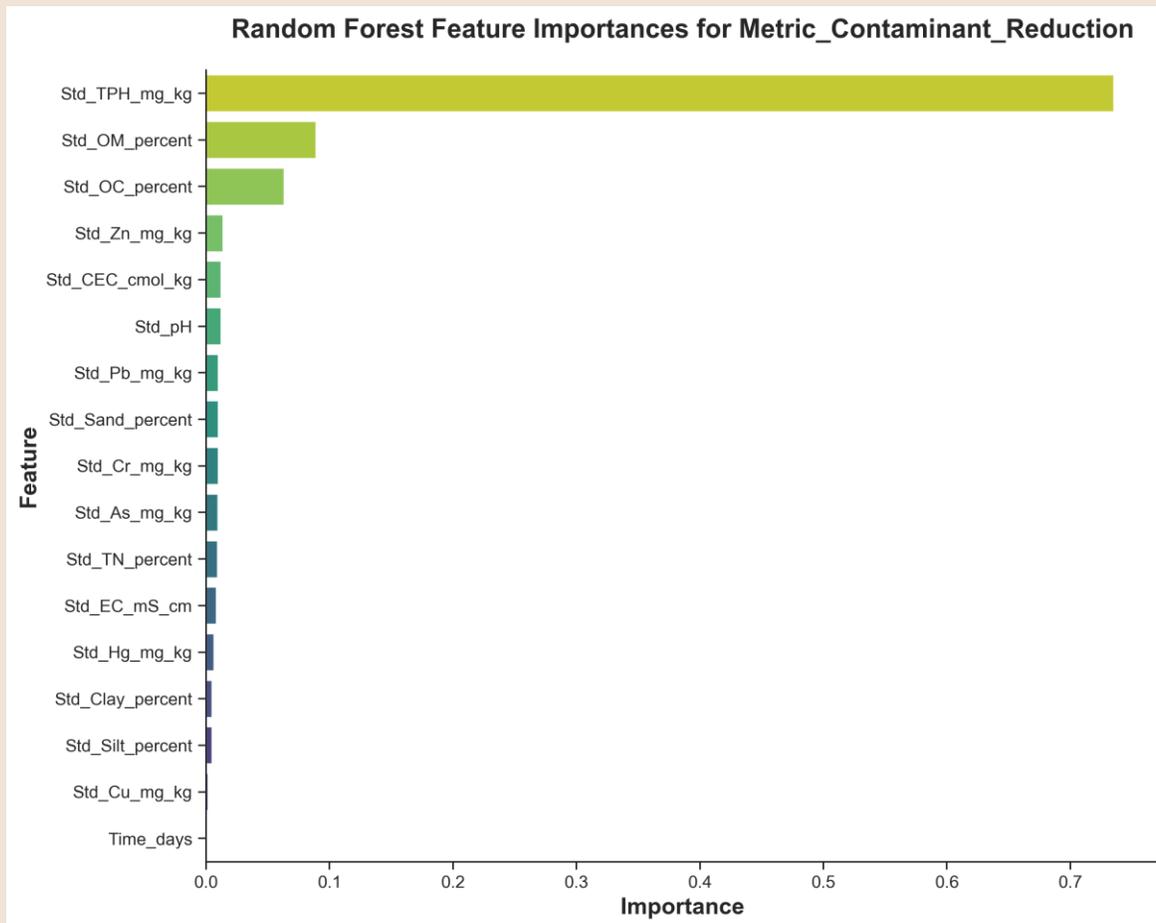
Metric_Contaminant_Reduction: Good Model

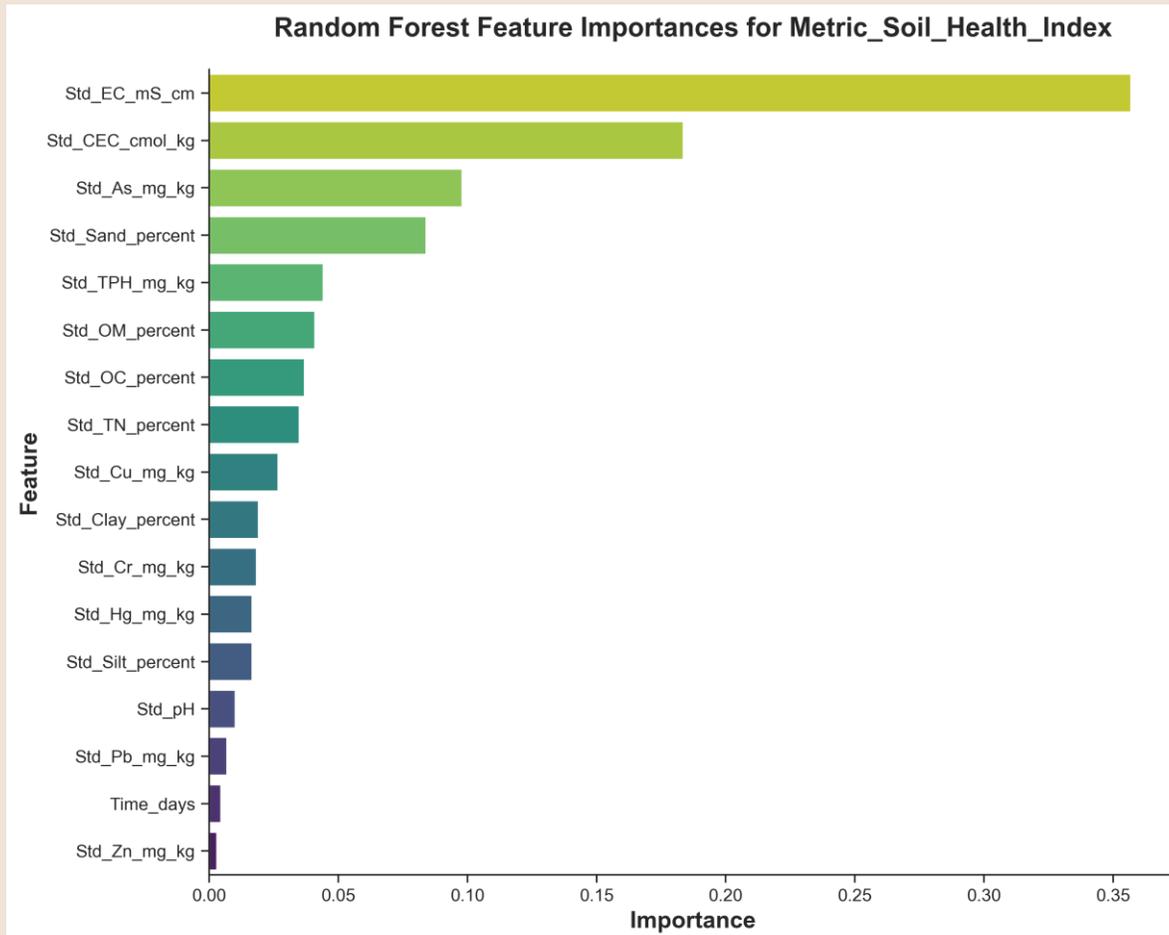Test R^2: *0.96* | Cross-Validated R^2: *0.81*

This model is accurate and reliable.

Interpretation:

- **Validation:** The model's accuracy is confirmed by identifying **TPH** (the metric's primary component) as the #1 predictor.

- **Discovery:** The true finding is that **Organic Matter (OM)** and **Organic Carbon (OC)** are the 2nd and 3rd most important predictors. This is a critical insight, showing that soils high in organic matter are associated with better contaminant reduction.

What Indicators *Predict* Soil Function? (Random Forest)

- We built Random Forest models to predict the *Metrics score* using the *Core Indicators* as features.
- This tells us which indicators are the most important.

Random Forest Feature Importances for Metric_Soil_Health_Index

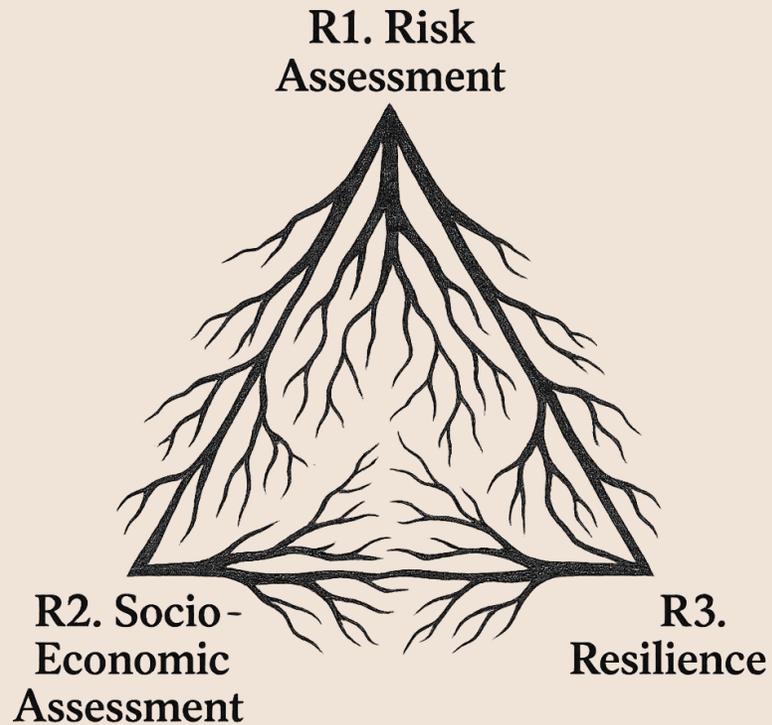**Metric_Soil_Health_Index:** Fair MODEL

Test *R^2: 0.85* | Cross-Validated *R^2: 0.5*

Interpretation: The model is predictive, but its generalizations are fair.

Interpretation:

- **Validation:** The model correctly identifies **Electrical Conductivity (EC)**, one of the metric's components, as the most important predictor.

- **Discovery:** The key discovery is that **CEC (Cation Exchange Capacity)**, which was *not* in the metric's formula, is the **#2 most important predictor**. This demonstrates that CEC is a powerful, independent indicator of overall soil health as we've defined it.

Thank-you! Any questions?

R1. Risk Assessment

R2. Socio-Economic Assessment

R3. Resilience

arag☀rn

NGI