This research has received funding from the Horizon Europe program under Grant agreement No 101086355

Aristotle University of Thessaloniki

Enhanced Soil Property Estimations from Earth Observation Data with Differential Evolution-based Multi-Objective TSK Model

N. L. Tsakiridis, N. Samarinas, E. Kalopesa, J. B. Theocharis, and G. C. Zalidis

Special Session: Evolutionary Multi-Objective Machine Learning

July 2024

Yokohama, Japan

Outline

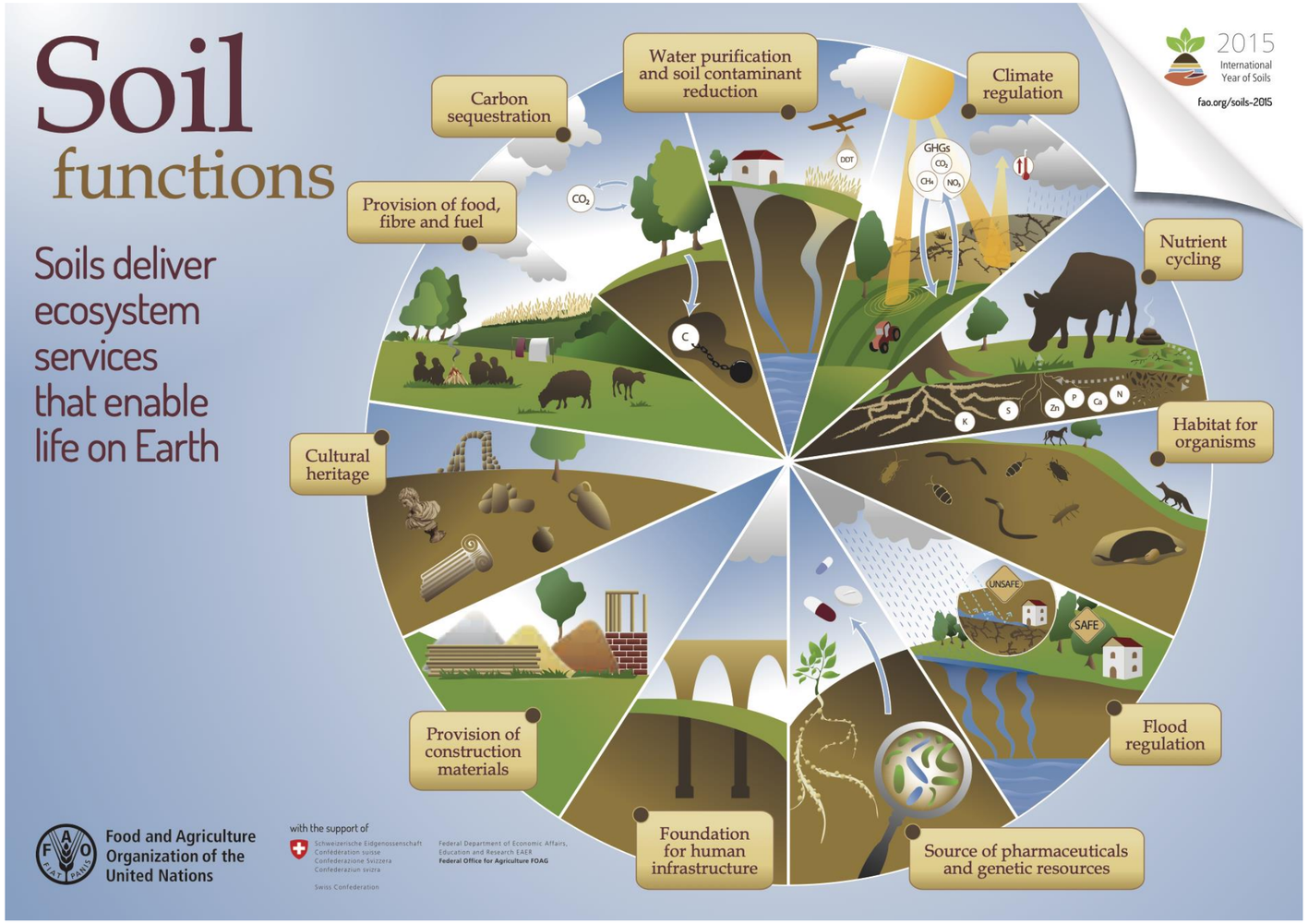
JULY 2024
YOKOHAMA, JAPAN

1. Introduction and problem statement
2. Proposed DE-MO-TSK model
3. Experimental framework
4. Results and discussion
5. Conclusions

SCALEAGDATA
MAKING DATA COUNT

SPECTRA
LAB

Introduction (I)



Soil = foundation of life

- 95% of our food comes from soil
- The most species-rich habitat on earth
- Helps combat and adapt to climate change

Introduction (II)

Soils are degraded

- Erosion, acidification, compaction, ...
- > 60% of EU soils are unhealthy
- Croplands losing C at 0.5% / year
- Cost: 50 B € / year

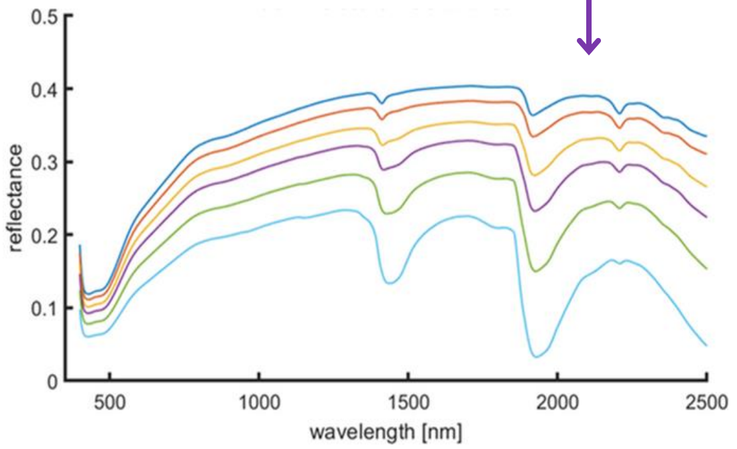
Introduction (III)

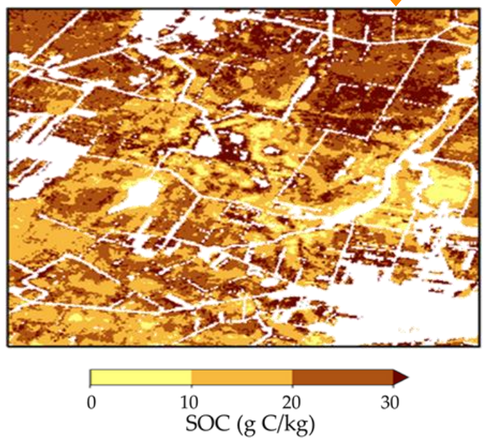
AI modelling

Spaceborne hyperspectral data

Soil spectral signatures

Output





SPECTRA LAB

SCALEAGDATA
MAKING DATA COUNT

The DE-MO-TSK model: FRBS

DE-MO-TSK: A multi-objective TSK type-1 fuzzy system using Differential Evolution

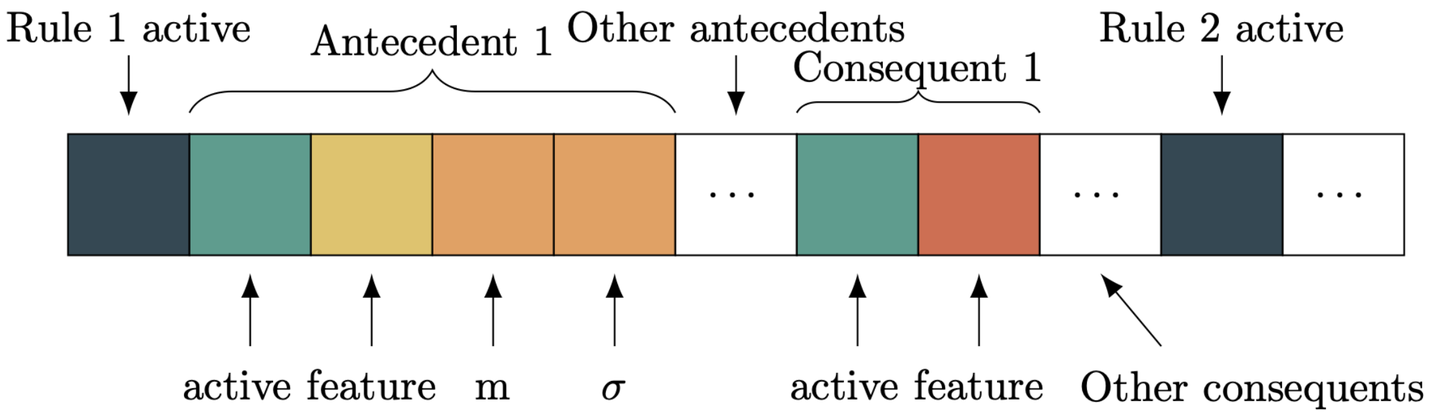
Rules of the form

$$\text{IF } x_{f_1} \text{ is } A_{f_1} \text{ AND } \ldots \text{ AND } x_{f_N} \text{ is } A_{f_N} \text{ THEN } y \text{ is } b_0 + b_{c_1} \cdot x_{c_1} + \ldots + b_{c_M} \cdot x_{c_M}$$

- Different features could be used in input and output
- Gaussian membership functions
- Weighted average for final consequent (weight is the matching degree)

DE-MO-TSK optimizes the whole Knowledge Base (RB + DB) simultaneously

The DE-MO-TSK model: Chromosome



One chromosome = One KB

- Activation gene for a maximum of Nrules
- Feature selection (up to Nantecedents and Nconsequents)
- Gaussian membership functions are also learned
- Length = Nrules * (1 + 4 * Nantecdents + 2 * Nconsequents)

The DE-MO-TSK model: Evolution

⚙️ Initialization

- Random NP chromosomes; consequent weights determined via least squares

🧬 Evolution for a given generation

- Current-to-pbest evolution strategy and binomial crossover
- Memory cells to generate F and Cr
- External archive to promote diversity
- Non-Linear Population Size Reduction
- NSGA-II as selection operator

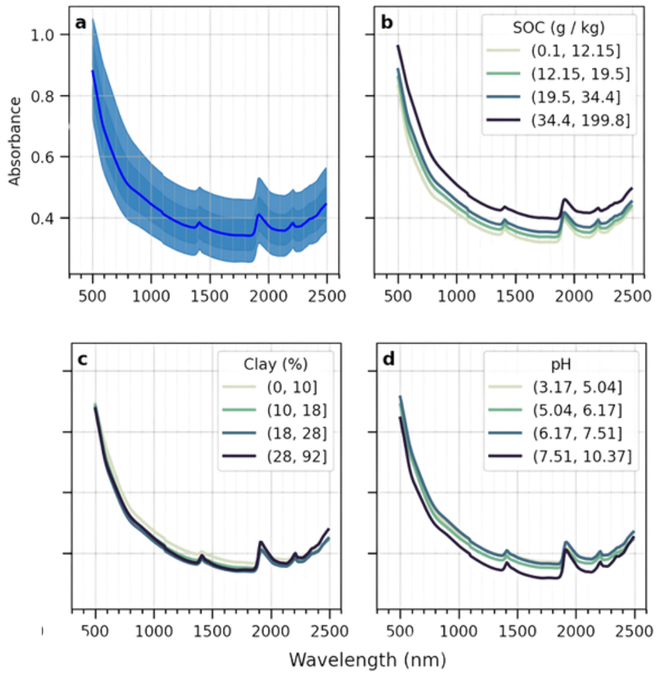
🎯 Multiple objectives

- ⬇️ MSE
- ⬇️ Number of rules
- ⬇️ Number of features

Datasets considered (I)

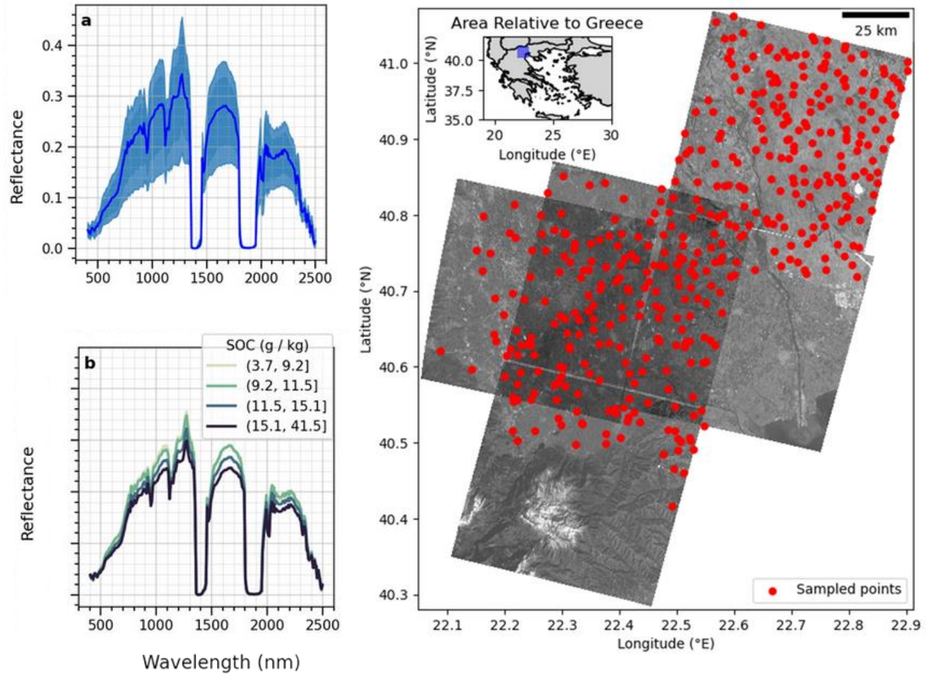
LUCAS 2015
Laboratory data

- Largest open lab hyperspectral database
- EU-28 countries (N = 20,687)
- 3 topsoil properties: SOC, Clay and pH
- 2 spectral sources (original + SG1 pre-treatment)

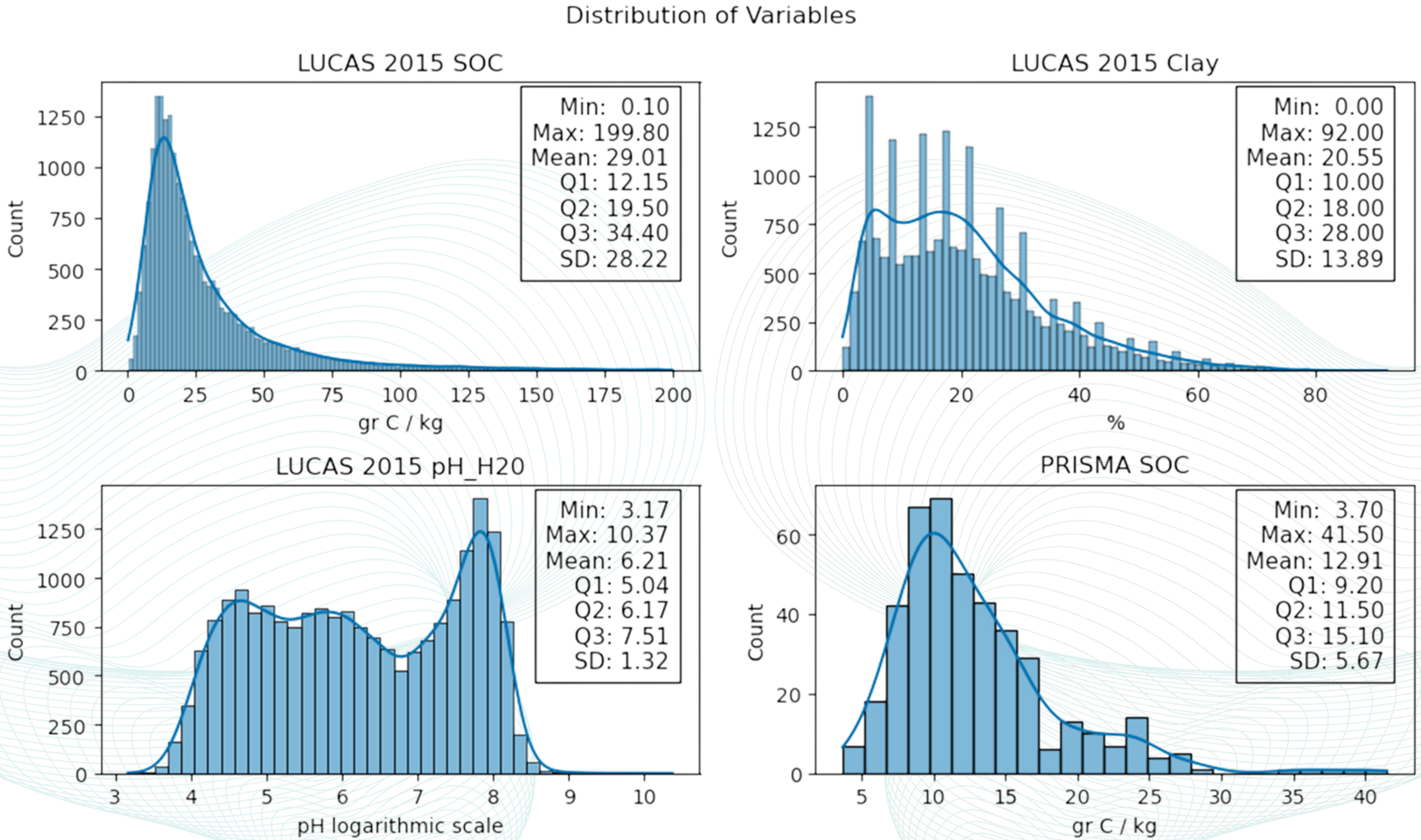


PRISMA
Spaceborne data

- PRISMA (by ASI) hyperspectral images
- Focus on Region of Central Macedonia, GR
- Topsoil SOC, N = 426
- 2 spectral sources (original + SNV pre-treatment)



Datasets considered (II)



Experimental framework

Compare DE-MO-TSK with Random Forest

01

DE-MO-TSK parameters

- **Nrules = 15**
- Nantecedents = 5
- Nconsequents = 5
- 100 generations
- NLPSR: [500, 4000]
- 3 random seeds

02

Random Forest

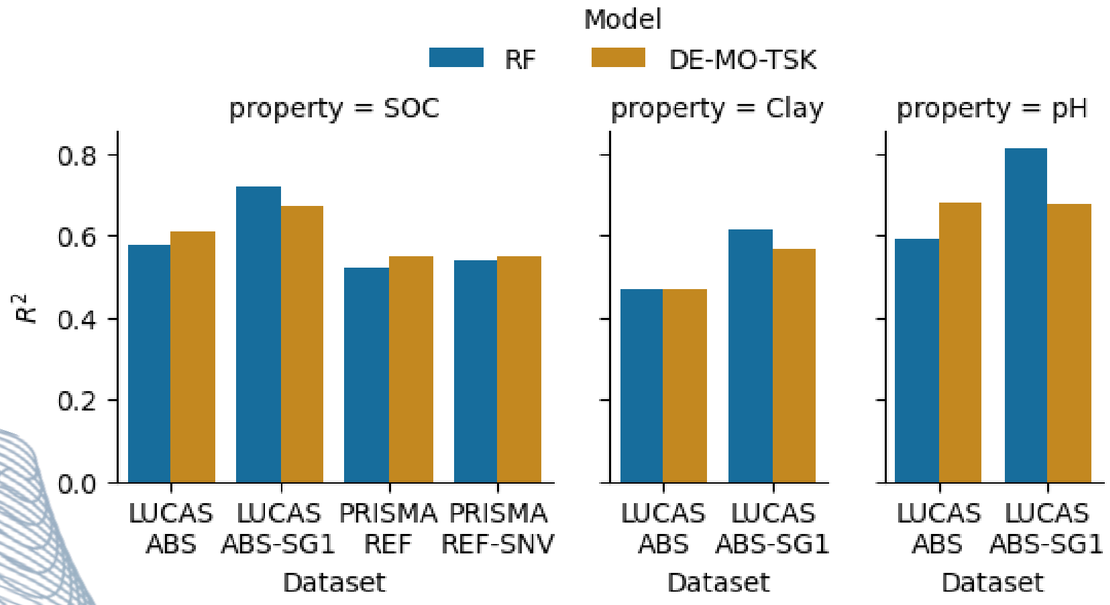
- Hyperparameters optimized via 5-fold CV
- Ntrees, maximum depth of trees, and maximum features considered for splitting

03

Evaluation metrics

- ⬇ RMSE
- ⬆ R-squared
- ⬆ RPIQ = IQR / RMSE
- ⬇ RB complexity

Results – H2H Model comparison



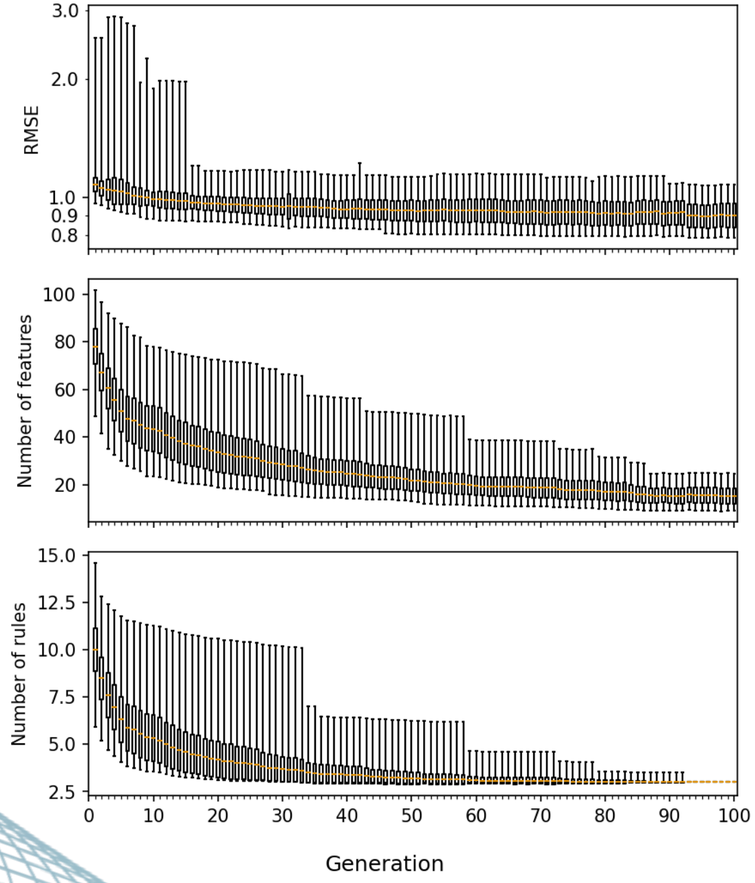
1. RF and DE-MO-TSK are statistically equivalent in terms of model accuracy (Wilcoxon signed rank test)

2. RF uses 200 to 400 nTrees

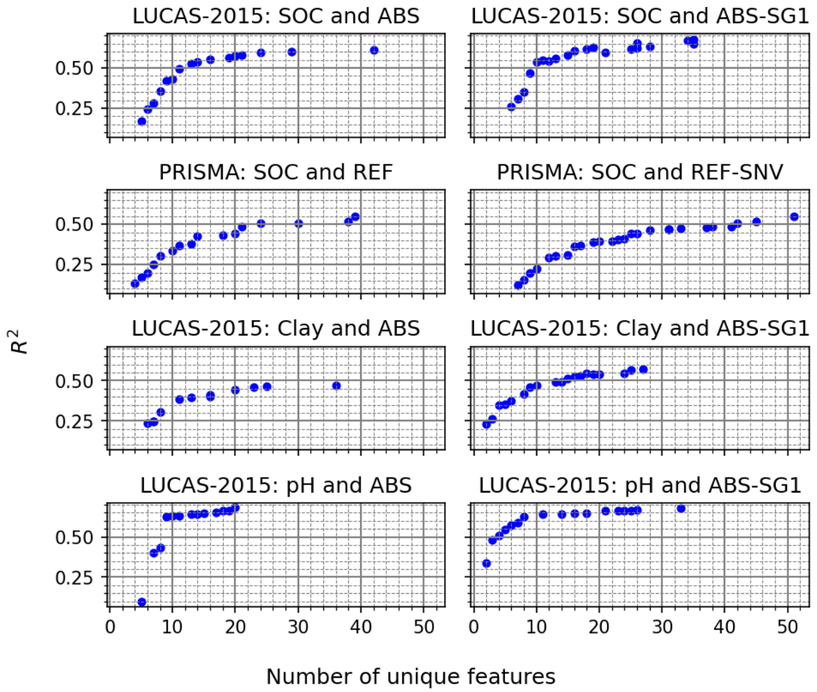
3. DE-MO-TSK far simpler models with
   nRules ~= 4
   nFeatures ~= 35

Results – DE-MO-TSK inspection (I)

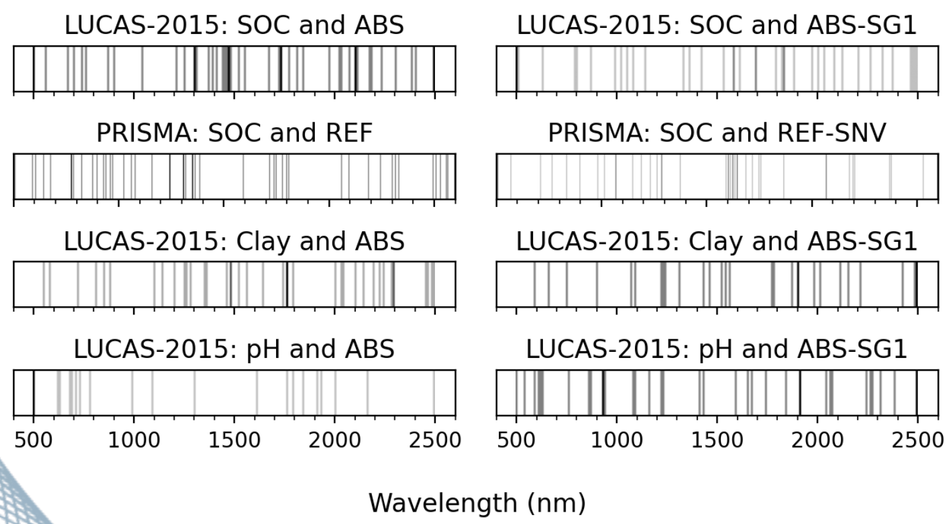
pH prediction from the LUCAS 2015 Absorption spectra



Pareto Fronts



Results – DE-MO-TSK inspection (II)



- Inspection of selected features (premise and consequent) revealed a close correlation with the known absorption bands from the literature

- DE-MO-TSK performs sparse feature selection and sheds light into the underlying associations

Summary and Conclusions

DE-MO-TSK

- Multi-objective Type-1 TSK models
- Differential Evolution
- NSGA-II selection operator
- NLPSR
- Embedded feature selection

Open questions

- Sensitivity analysis
- Feature alignment mechanism
- Use of other EA
- Comparison with other multi-objective FRBSs

Hyperspectral soil monitoring

- Tested two datasets: laboratory LUCAS-2015 and our own space-borne PRISMA dataset in Northern Greece
- Compared to RF: accuracy 〓 complexity ⬇
- Sparse feature selection with physical meaning

Future considerations

- Comparison with other datasets
- … and other models

Get a glimpse into our world:

https://spectralab.gr/

Spectra Lab, AUTH

@spectralabgr

SPECTRA LAB

Thank you for your attention!

Nikolaos L. Tsakiridis
tsakirin@auth.gr
n.tsakiridis@spectralab.gr

Aristotle University
of Thessaloniki