

# Leveraging Soil Reflectance Models with Sentinel-2 for Scalable Soil Moisture Mapping to Enhance Agricultural Resilience in South Africa.

YONWABA ATYOSI  
CSIR

Supervisors:

Prof Moses Cho

Prof Abel Ramoelo

Dr Nobuhle Majozi



science, technology  
& innovation  
Department:  
Science, Technology and Innovation  
REPUBLIC OF SOUTH AFRICA



sa air force  
Department:  
Defence  
REPUBLIC OF SOUTH AFRICA





## SHORT BIO

- Yonwaba Atyosi is a candidate researcher at the CSIR's Precision Agriculture Research Group.
- Research focus: Remote sensing of soils; developing innovative techniques for precision agriculture.
- Academic background: Geology, Applied Remote Sensing and GIS
- PhD studies: Geoinformatics, University of Pretoria
- Current project: Enhancing hyperspectral & multispectral soil moisture retrieval by accounting for clay content



**Yonwaba Atyosi**  
Candidate Researcher, CSIR



**science, technology  
& innovation**  
Department:  
Science, Technology and Innovation  
REPUBLIC OF SOUTH AFRICA



**sa air force**  
Department:  
Defence  
REPUBLIC OF SOUTH AFRICA



## Brief background

- ❖ Sensitivity analysis of subtle changes in soil moisture
- ❖ Conducted a canonical analysis –A multivariate statistical measure that helps identify the **most influential soil variables** affecting SMC retrieval.

Developed a hyperspectral based soil reflectance model that consider the interaction between soil moisture and soil clay content through Monte Carlo simulation.

$$R_{soil} = aSMC + bSCC + c(SMC * SCC) + d \quad (1)$$

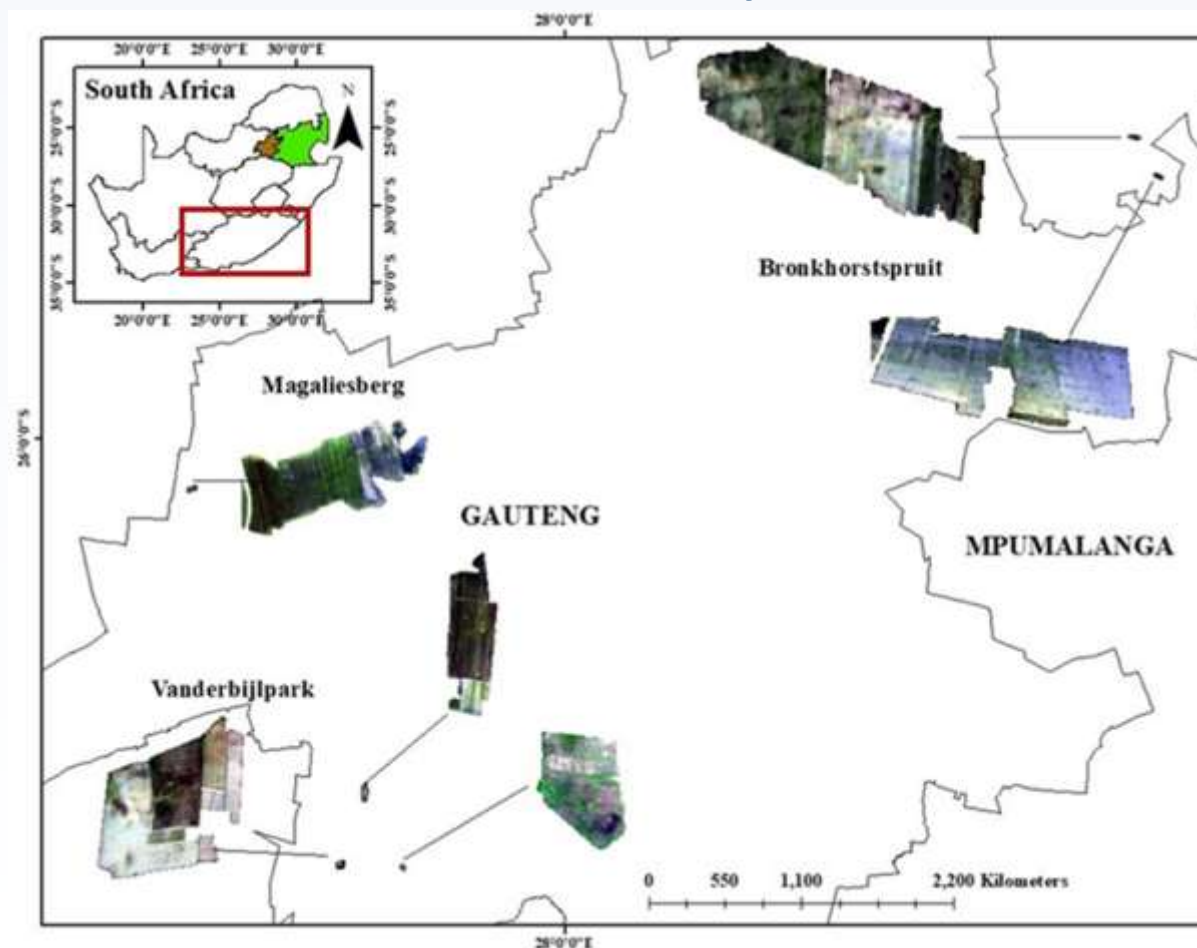
Where coefficients a,b,c were drawn from a Gaussian distribution , and SMC and SCC , uniform distribution.

- ✓ Tested on independent hyperspectral dataset from the EC.
- ✓ We resampled the hyperspectral soil library to Sentinel 2-configuration
- ✓ Tested on resampled hyperspectral independent dataset
- ✓ Tested on actual Sentinel-2 independent datasets from EC and Limpopo (Different growing seasons)
- ❖ Assessed this multivariate soil hyperspectral reflectance model for mapping SMC using Sentinel 2





## Study sites



| Sample ID | Sample picture | Latitude (S)     | Longitude (E)     | Clay content (%)                     |
|-----------|----------------|------------------|-------------------|--------------------------------------|
| 1         |                | 28° 18' 2.64" S  | 27° 40' 7.572" E  | 9.5%<br>(minimum percentile)         |
| 2         |                | 28° 38' 59.04" S | 28° 51' 24.084" E | 15%<br>(20 <sup>th</sup> percentile) |
| 3         |                | 28° 37' 28.76" S | 28° 52' 4.152" E  | 18%<br>(40 <sup>th</sup> percentile) |
| 4         |                | 28° 4' 16.32" S  | 27° 23' 13.272" E | 28%<br>(60 <sup>th</sup> percentile) |
| 5         |                | 28° 4' 31.44" S  | 27° 27' 5.144" E  | 34%<br>(80 <sup>th</sup> percentile) |
| 6         |                | 28° 4' 33.96" S  | 27° 27' 3.744" E  | 43%<br>(maximum percentile)          |



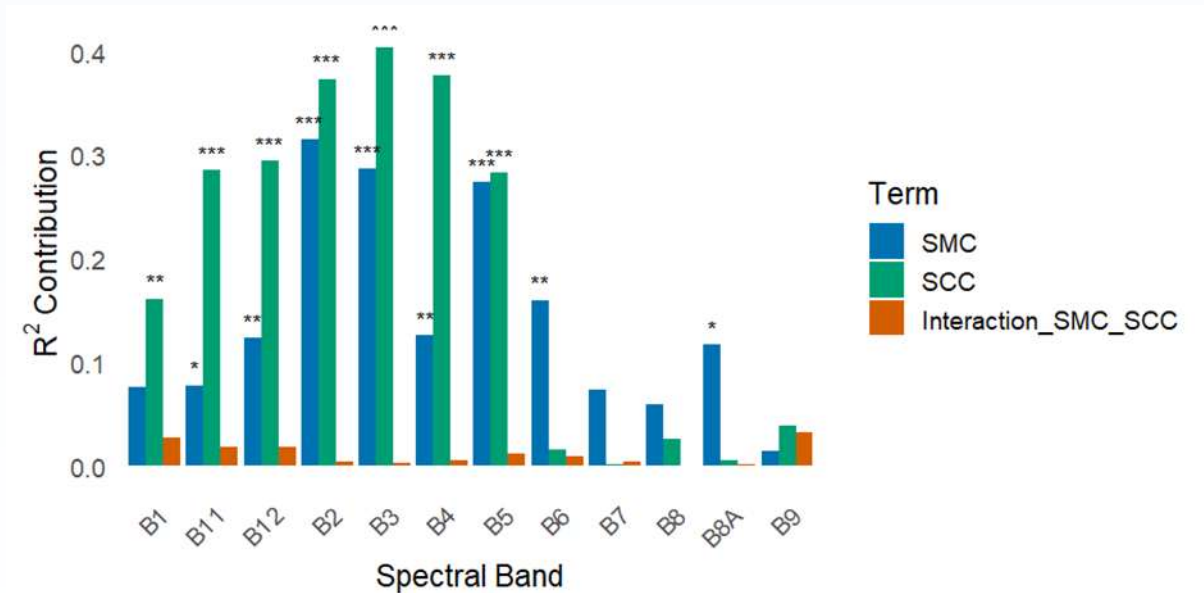
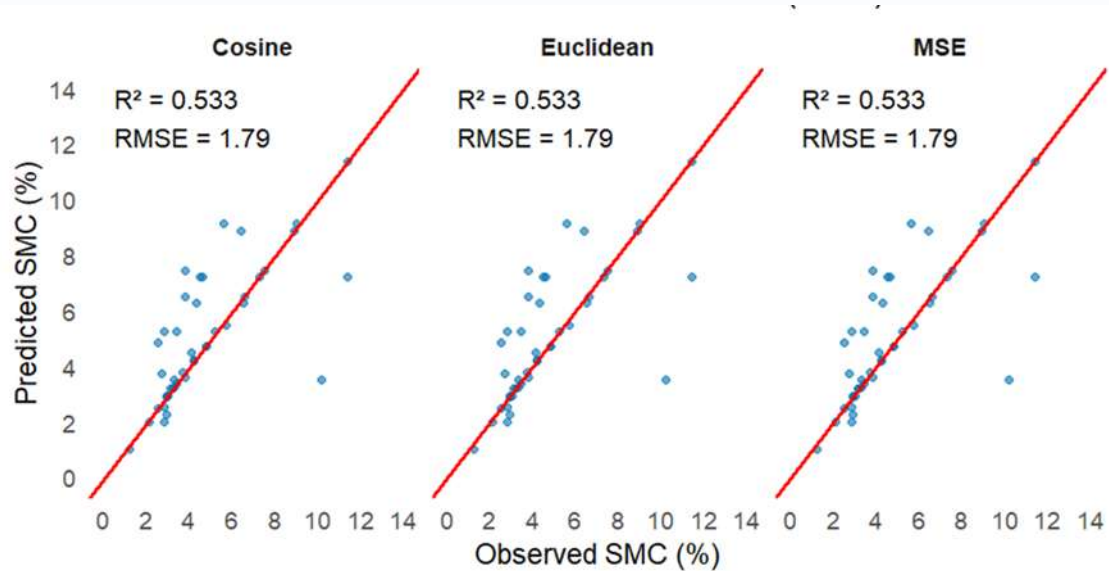
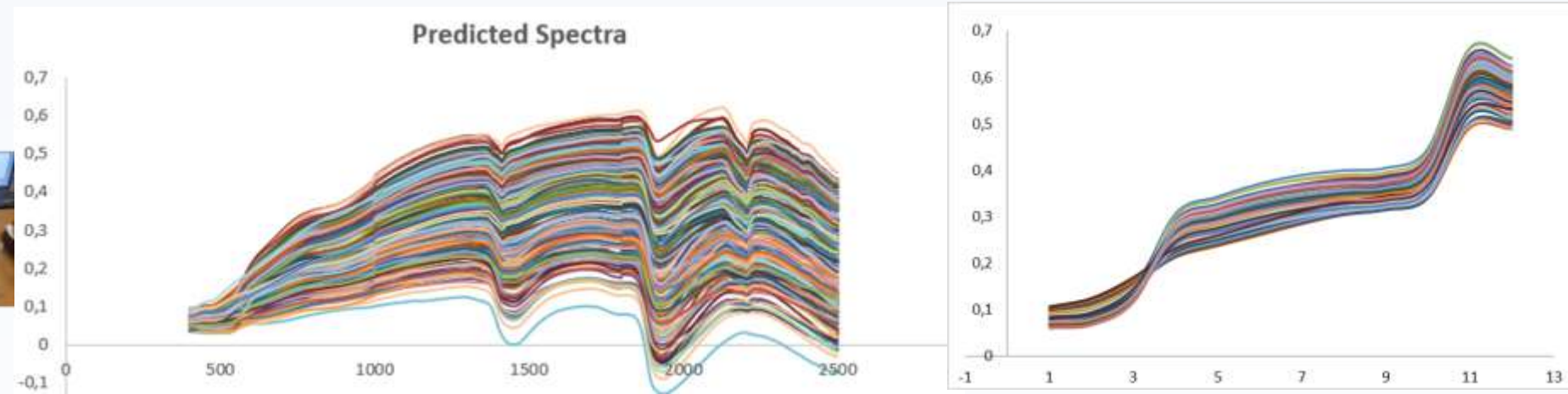
## The problem

- Accurate, timely and transferable SMC retrieval remains a challenge for precision agriculture, drought monitoring, and hydrological modeling.
- Traditional methods (gravimetric sampling, in-situ sensors) provide reliable **point-based soil moisture data**. **However, they are labor intensive, spatial constrained and impractical for large scale monitoring.**
- Remote sensing offers scalable alternatives, yet **retrieving soil moisture from satellite data remains challenging**
  - ❖ **Hyperspectral RS-** capture subtle soil reflectance changes linked to moisture  
**Limited operational use due to scarce, high-resolution, open-access data**
  - ❖ **Multispectral RS-** freely accessible, global coverage, high temporal resolution  
**Broader bands reduce precision in capturing soil moisture dynamics**
- Existing empirical and machine learning approaches often neglect soil property interactions and lack spatiotemporal generalizability.





# Resampling from hyperspectral to multispectral via S2 SRF



# The Solution

| Ratio   | R <sup>2</sup> | RMSE <sub>p</sub> | RMSE <sub>p</sub> % |
|---------|----------------|-------------------|---------------------|
| B3/B1   | 0,85646844     | 1,31309586        | 6,56510241          |
| B12/B1  | 0,76110919     | 1,39941324        | 6,99686626          |
| B12/B4  | 0,71623186     | 1,69385579        | 8,46918447          |
| B12/B5  | 0,70896814     | 1,84626596        | 9,23092127          |
| B12/B6  | 0,70643867     | 1,86979347        | 9,34857727          |
| B12/B7  | 0,69779381     | 1,87776208        | 9,38854749          |
| B12/B9  | 0,69677481     | 1,90507473        | 9,52528708          |
| B12/B8  | 0,68422686     | 1,9084999         | 9,54219589          |
| B12/B2  | 0,6774428      | 1,94740708        | 9,73691695          |
| B12/B8A | 0,67670377     | 1,96829332        | 9,84127391          |
| B8A/B1  | 0,67430007     | 1,97047242        | 9,8522378           |
| B3/B2   | 0,67335584     | 1,97778853        | 9,8888699           |
| B12/B3  | 0,66781178     | 1,98080505        | 9,90379584          |
| B11/B1  | 0,66377546     | 1,99737937        | 9,98675792          |
| B9/B1   | 0,66094162     | 2,00951773        | 10,0473717          |
| B8/B1   | 0,6576905      | 2,01863144        | 10,0921031          |
| B2/B1   | 0,65069057     | 2,02757385        | 10,1376975          |
| B12/B11 | 0,64609479     | 2,04823691        | 10,2410074          |
| B7/B1   | 0,64515605     | 2,0615914         | 10,3079304          |
| B11/B4  | 0,601554       | 2,06446983        | 10,3220059          |

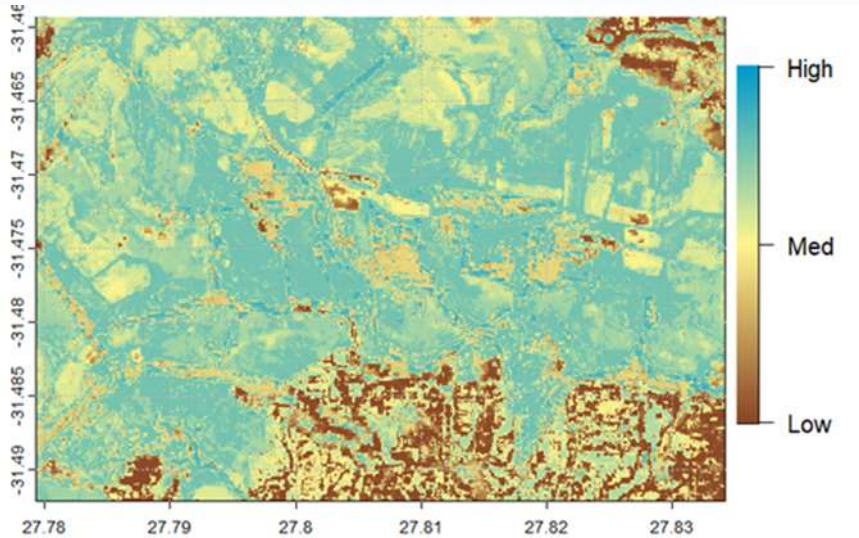
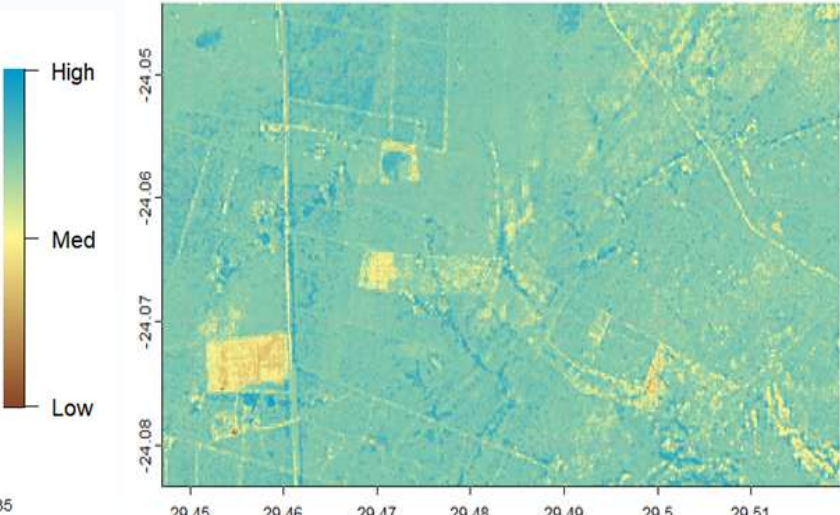
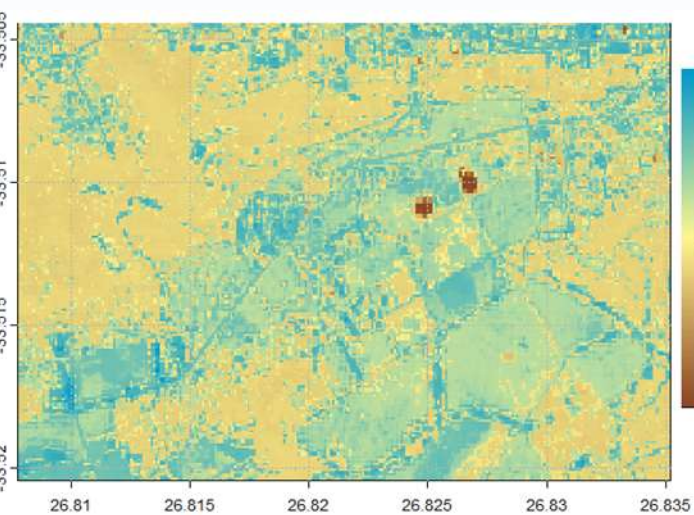
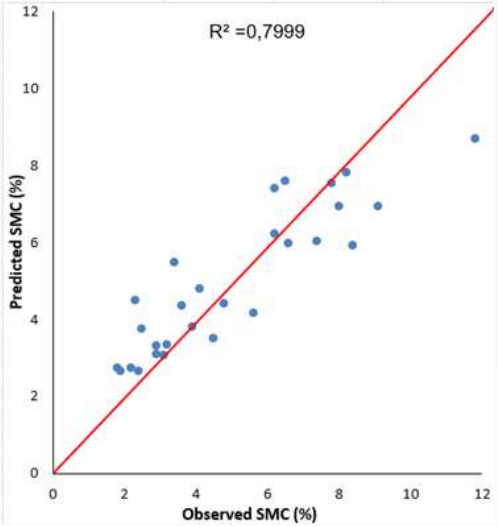
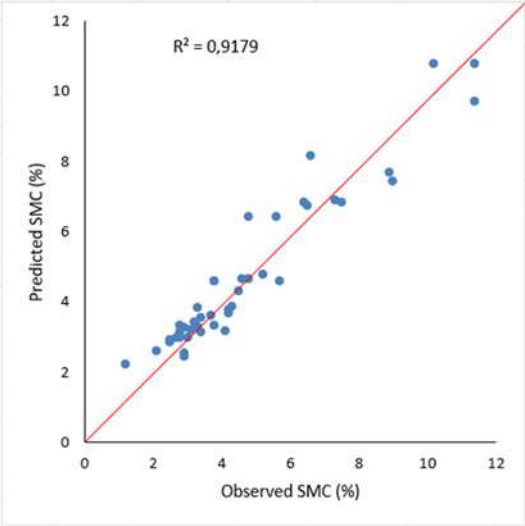
$$R_{soil} = aSMC + bSCC + c(SMC * SCC) + d$$

We developed a hyperspectral synthetic soil database

- Contains hundreds of thousands of spectra with corresponding SMC and SCC
- Representative of a wide range of different possible soil types in South Africa
- Explicitly incorporates SMC–SCC interactions into a multivariate modeling framework
- Enables a transferable, interpretable, and operational pathway for soil moisture retrieval
- Sentinel 2 spectral transformation and feature engineering (band ratios, developed 5 new SMIs)
- **Trained five machine learning models (SLR,MLR, RF, SVM, PLSR) with optimal band ratios and SMIs derived from our robust synthetic hyperspectral soil database to predict SMC**
- Applied the best model to Sentinel-2 imagery to spatio-temporally map SMC , ensuring scalability **across space and time.**

# ENHANCED SMC PREDICTION ACCURACY

| Model  | RMSE       | MAE        | RMSE <sub>p</sub> % | MAPE       | R <sup>2</sup> |
|--------|------------|------------|---------------------|------------|----------------|
| Linear | 1,48256224 | 1,19268966 | 29,0991078          | 33,7612852 | 0,67448176     |
| SVM    | 1,46523953 | 1,13530982 | 28,7591051          | 34,819678  | 0,68204422     |
| PLSR   | 1,48256224 | 1,19268966 | 29,0991078          | 33,7612852 | 0,67448176     |
| RF     | 0,8761717  | 0,67424038 | 17,1971294          | 19,8572694 | 0,88630852     |
| Linear | 1,69341492 | 1,36518149 | 33,2376357          | 37,7790576 | 0,57530589     |
| SVM    | 1,67229466 | 1,31038639 | 32,8230961          | 38,4542239 | 0,5858334      |
| PLSR   | 1,69341492 | 1,36518149 | 33,2376357          | 37,7790576 | 0,57530589     |
| RF     | 1,04154791 | 0,81629979 | 20,4430641          | 23,3875595 | 0,83933992     |
| Linear | 1,04995598 | 0,75829887 | 20,6080943          | 24,4697224 | 0,83673554     |
| SVM    | 1,03098337 | 0,70140427 | 20,2357078          | 26,1918797 | 0,84258258     |
| PLSR   | 1,04995598 | 0,75829887 | 20,6080943          | 24,4697224 | 0,83673554     |
| RF     | 0,58048438 | 0,43356709 | 11,3935033          | 13,7086997 | 0,95009655     |
| Linear | 1,45187678 | 1,15372972 | 28,4968266          | 33,7971762 | 0,68781719     |
| SVM    | 1,4674773  | 1,14348617 | 28,803027           | 35,218368  | 0,68107229     |
| PLSR   | 1,45187678 | 1,15372972 | 28,4968266          | 33,7971762 | 0,68781719     |
| RF     | 0,86071564 | 0,66777073 | 16,8937645          | 20,1808763 | 0,89028428     |
| Linear | 1,59995769 | 1,27733152 | 31,4032963          | 40,9362876 | 0,62088893     |
| SVM    | 1,15526559 | 0,87747005 | 22,675067           | 29,0701767 | 0,80234257     |
| PLSR   | 1,59995769 | 1,27733152 | 31,4032963          | 40,9362876 | 0,62088893     |
| RF     | 0,70856367 | 0,53121581 | 13,9073896          | 15,3228448 | 0,9256455      |





## CONCLUSSIONS

- Accounting for clay content and its interaction with soil moisture content in soil reflectance modelling, improves SMC prediction accuracy.
- Our multivariate hyperspectral model can accurately predict and map SMC with Sentinel 2 across space and time.

The framework bridges hyperspectral richness and multispectral practicality, offering scalable solutions for precision agriculture, water resource management, and environmental monitoring, with

- ❖ Current work is focussed on incorporating vegetation parameters for a fully integrated soil–canopy moisture mapping.



Thank You



science, technology  
& innovation

Department:  
Science, Technology and Innovation  
REPUBLIC OF SOUTH AFRICA



sa air force

Department:  
Defence  
REPUBLIC OF SOUTH AFRICA



CSIR  
Touching lives through innovation