UNIVERSITY OF TWENTE.

GLOBAL LONG-TERM DAILY 1KM SURFACE SOIL MOISTURE DATASET WITH PHYSICS-INFORMED MACHINE LEARNING (GSSM1KM)

REPORTER: **QIANQIAN HAN**, YIJIAN ZENG, LIJIE ZHANG, BOB SU

18 JAN 2023

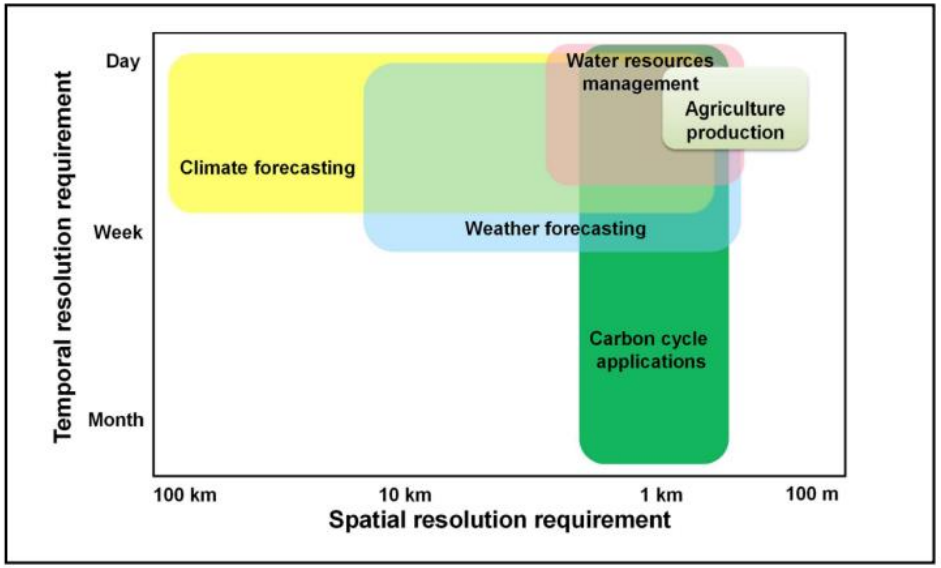
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

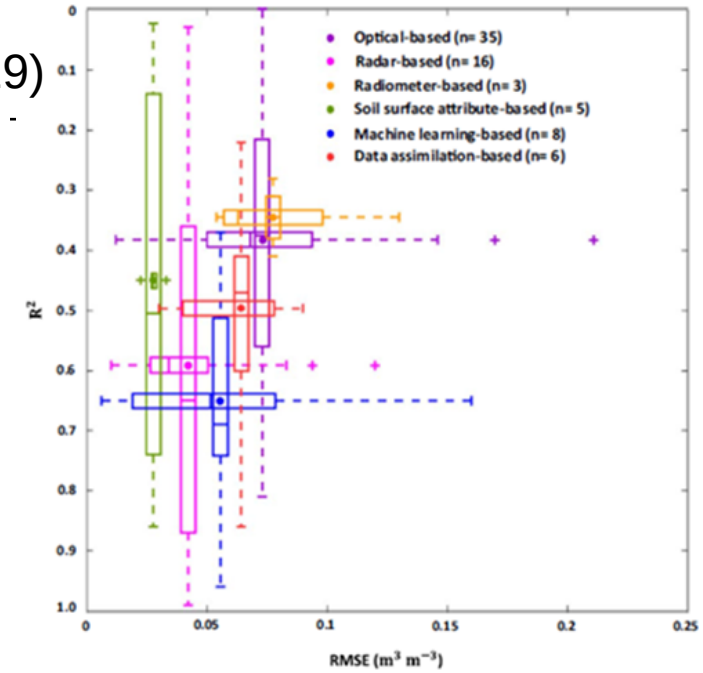
OUTLINE

1. Introduction
2. Materials and Methodology
3. Results

1. INTRODUCTION

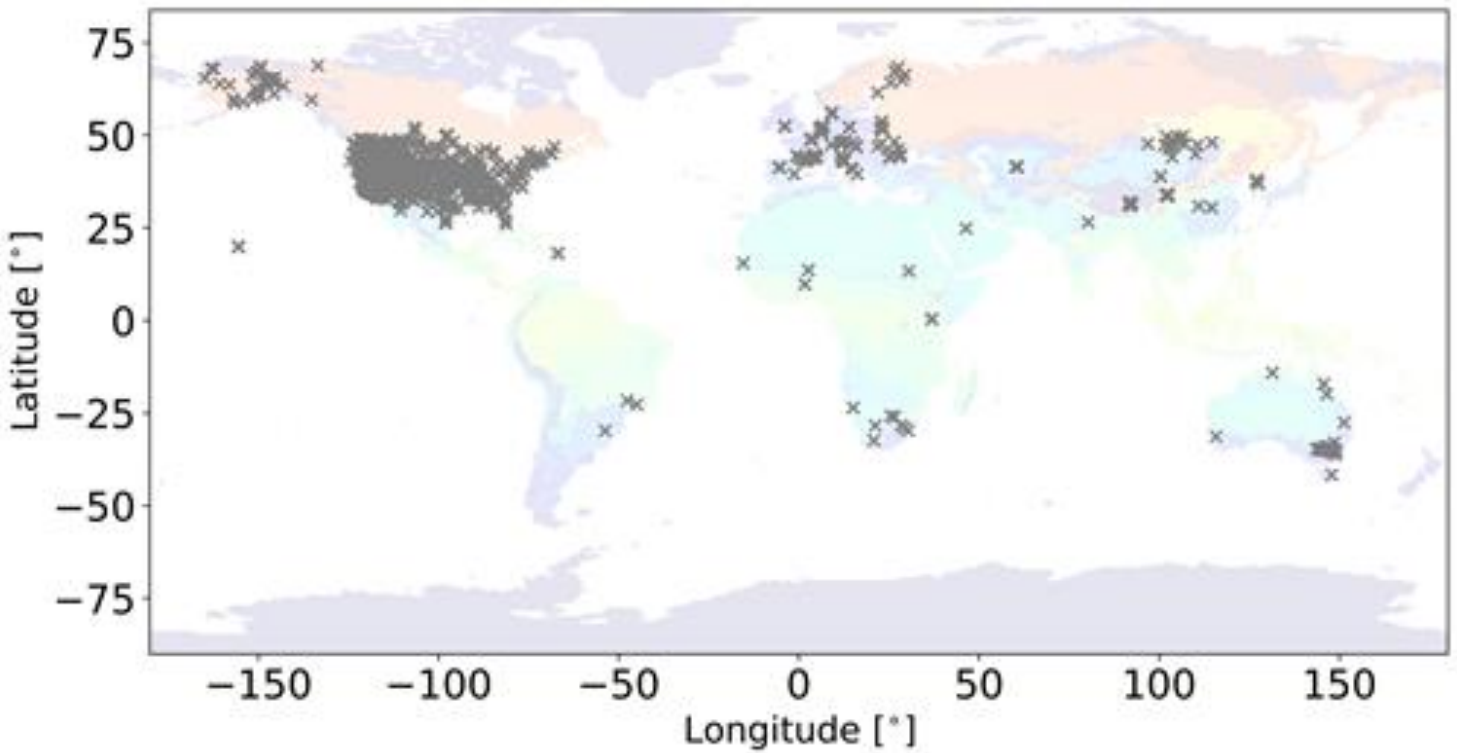
(Sabaghy, 2019)





- Surface soil moisture (SSM) is one of the essential climate variables that play a fundamental role in the water and heat exchanges between the land and atmosphere.
- A long-term and high resolution SSM is missing.
- Random forest (RF) provide a possibility to facilitate the understanding of the relationship between the available in-situ SSM and land surface (atmospheric) features at the global scale.

2. MATERIALS AND METHODOLOGY



International soil moisture network (ISMN) delivers quality controlled long-term in-situ soil moisture observations. By the end of 2020, the database consisted of 2678 stations from 65 networks around the world, and ISMN is still growing.

2. MATERIALS AND METHODOLOGY

- In-situ Soil Moisture: International Soil Moisture Network (ISMN)
- Land surface features from Satellites (**available on GEE**):

① Antecedent Precipitation Evaporation Index (APEI), Evaporation, Air Temperature (Tair) – ERA5

② Daily Land Surface Temperature (LST), Daily LST Difference – MOD11A1

③ Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) – MOD13A1

④ Soil Texture, porosity, organic matter content - Soilgrids

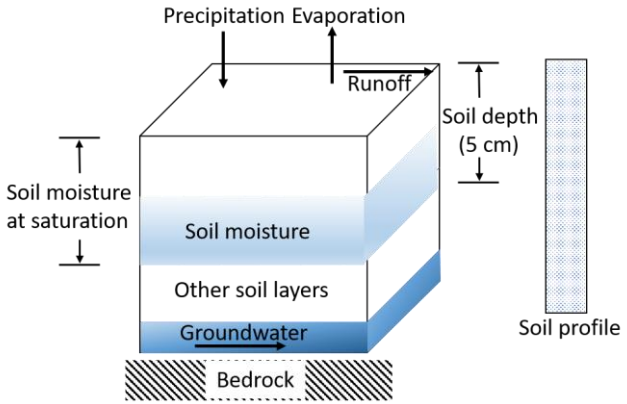
⑤ Topographic Index

⑥ Digital Terrain Model (DEM)

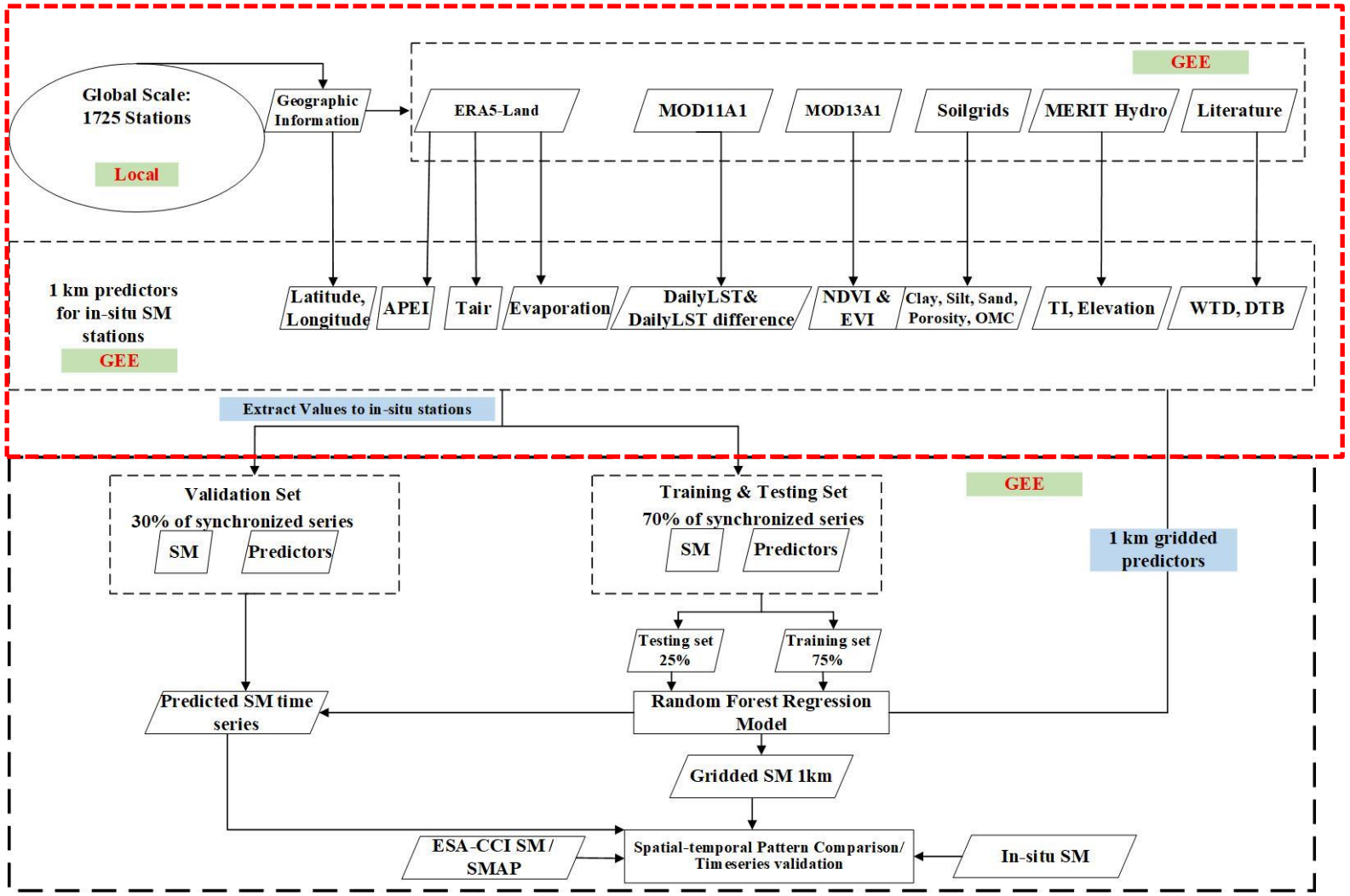
⑦ Water Table Depth

⑧ Depth to Bedrock

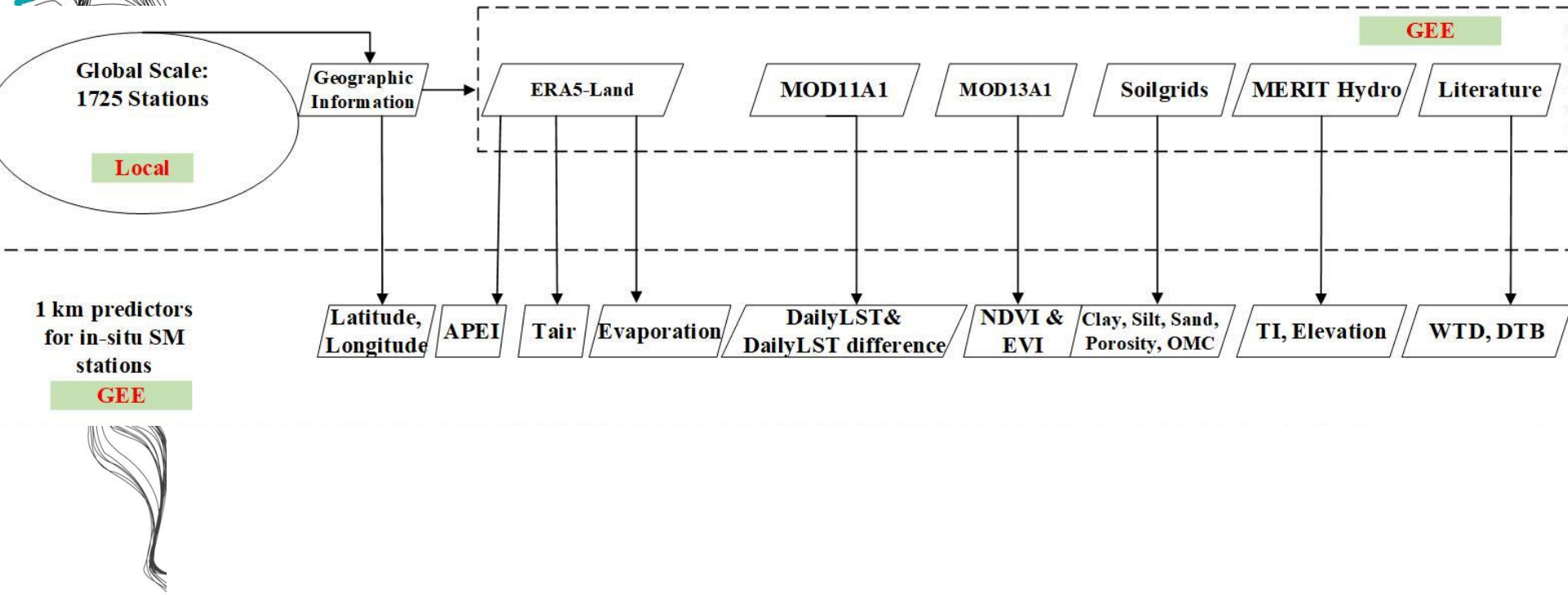
⑨ Geographical location



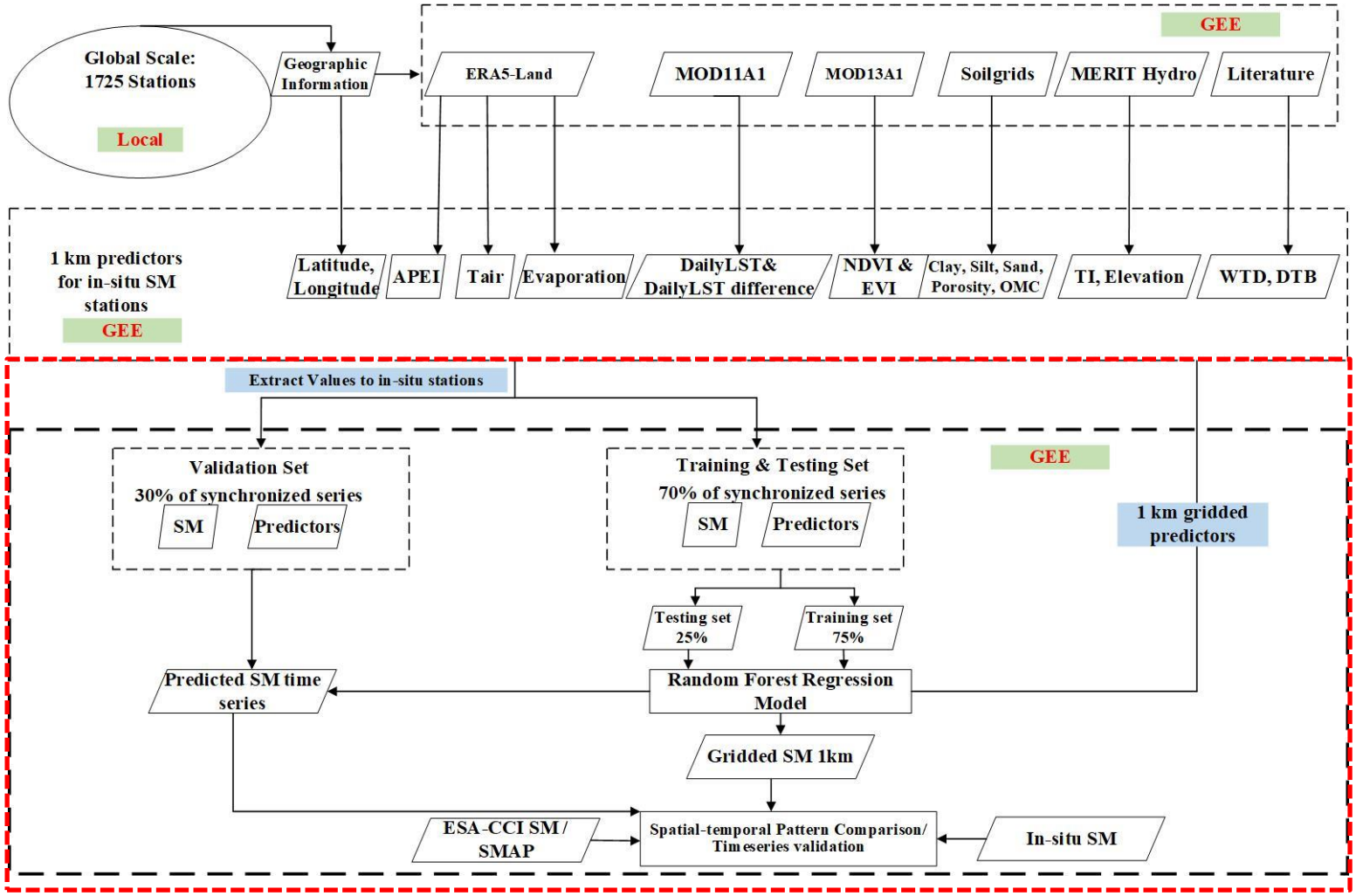
3. RESULTS


GEE
Global Scale:
1725 Stations
Local
Geographic Information
ERA5-Land
MOD11A1
MOD13A1
Soilgrids
MERIT Hydro
Literature
1 km predictors for in-situ SM stations
GEE
Latitude, Longitude
APEI
Tair
Evaporation
DailyLST& DailyLST difference
NDVI & EVI
Clay, Silt, Sand, Porosity, OMC
TI, Elevation
WTD, DTB
Extract Values to in-situ stations
Validation Set
30% of synchronized series
SM
Predictors
Training & Testing Set
70% of synchronized series
SM
Predictors
GEE
1 km gridded predictors
Testing set 25%
Training set 75%
Predicted SM time series
Random Forest Regression Model
Gridded SM 1km
ESA-CCI SM / SMAP
Spatial-temporal Pattern Comparison/ Timeseries validation
In-situ SM

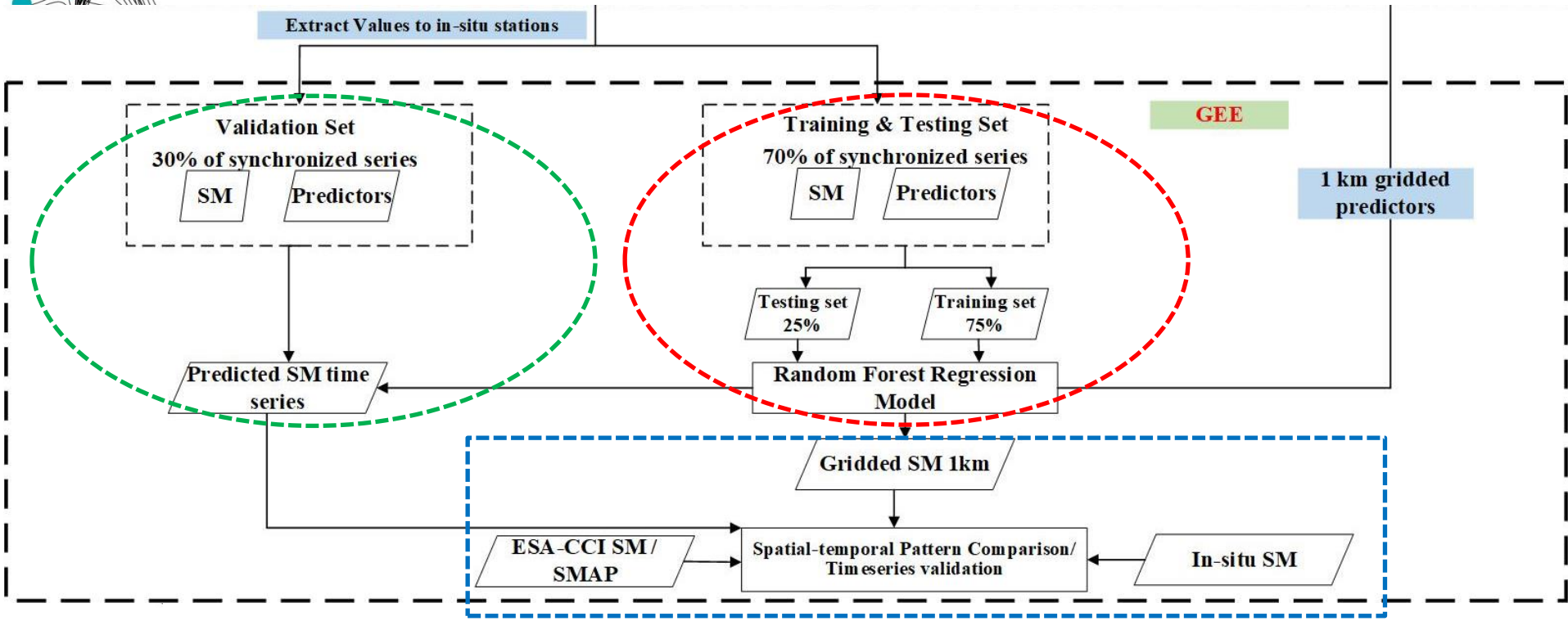
3. RESULTS


GEE
Global Scale:
1725 Stations
Local
Geographic Information
ERA5-Land
MOD11A1
MOD13A1
Soilgrids
MERIT Hydro
Literature
1 km predictors for in-situ SM stations
GEE
Latitude, Longitude
APEI
Tair
Evaporation
DailyLST& DailyLST difference
NDVI & EVI
Clay, Silt, Sand, Porosity, OMC
TI, Elevation
WTD, DTB

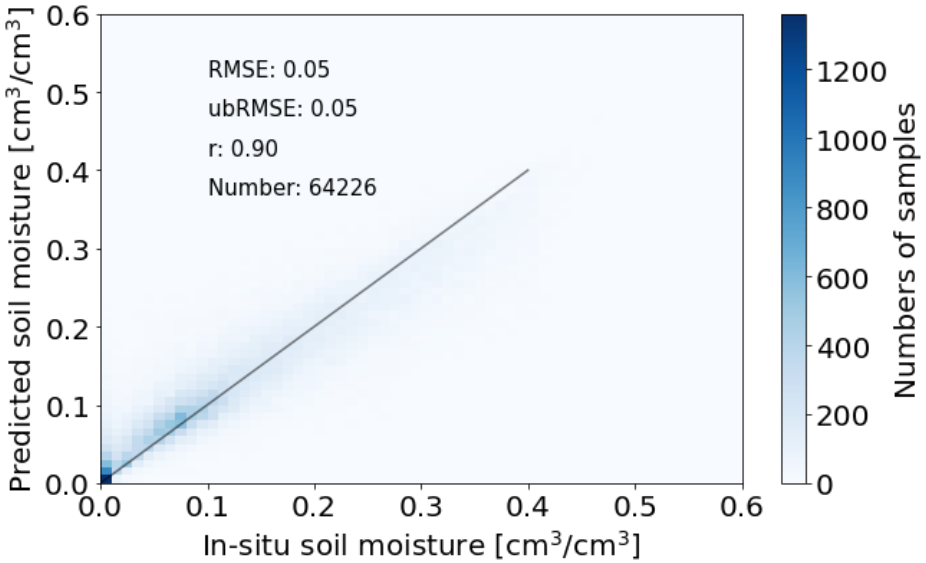
3. RESULTS


Global Scale:
1725 Stations
Local
Geographic Information
GEE
ERA5-Land
MOD11A1
MOD13A1
Soilgrids
MERIT Hydro
Literature
1 km predictors for in-situ SM stations
GEE
Latitude, Longitude
APEI
Tair
Evaporation
DailyLST& DailyLST difference
NDVI & EVI
Clay, Silt, Sand, Porosity, OMC
TI, Elevation
WTD, DTB
Extract Values to in-situ stations
Validation Set
30% of synchronized series
SM
Predictors
Training & Testing Set
70% of synchronized series
SM
Predictors
GEE
1 km gridded predictors
Testing set 25%
Training set 75%
Predicted SM time series
Random Forest Regression Model
Gridded SM 1km
ESA-CCI SM / SMAP
Spatial-temporal Pattern Comparison/ Timeseries validation
In-situ SM

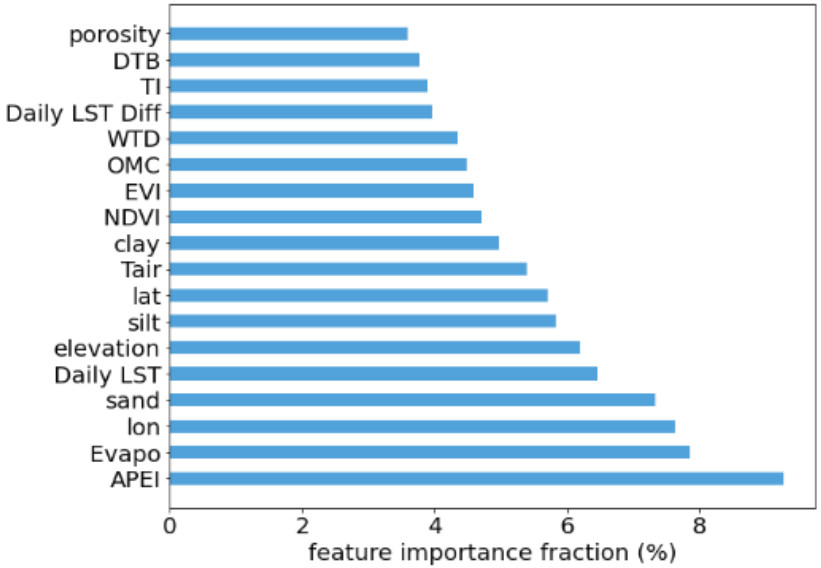
3. RESULTS


Extract Values to in-situ stations
Validation Set
30% of synchronized series
SM
Predictors
Training & Testing Set
70% of synchronized series
SM
Predictors
GEE
1 km gridded predictors
Testing set 25%
Training set 75%
Predicted SM time series
Random Forest Regression Model
Gridded SM 1km
ESA-CCI SM / SMAP
Spatial-temporal Pattern Comparison/ Timeseries validation
In-situ SM

3. RESULTS

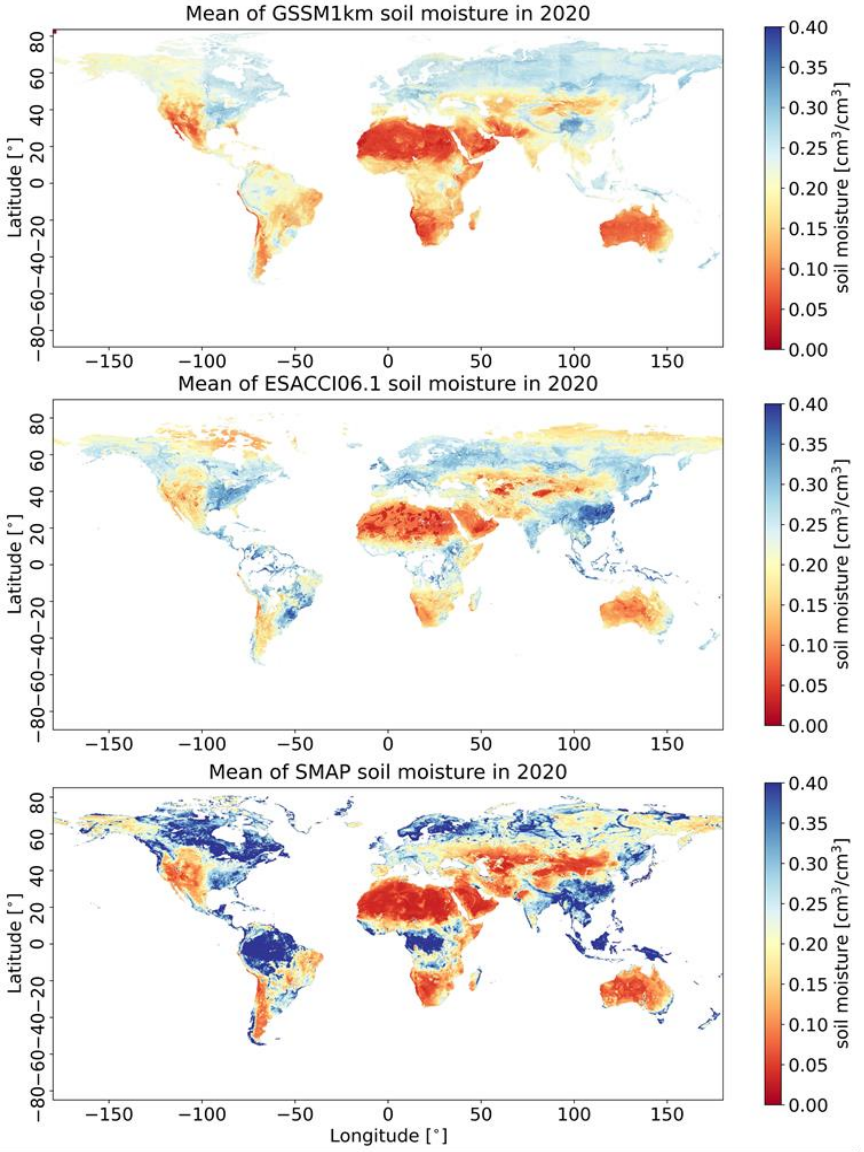


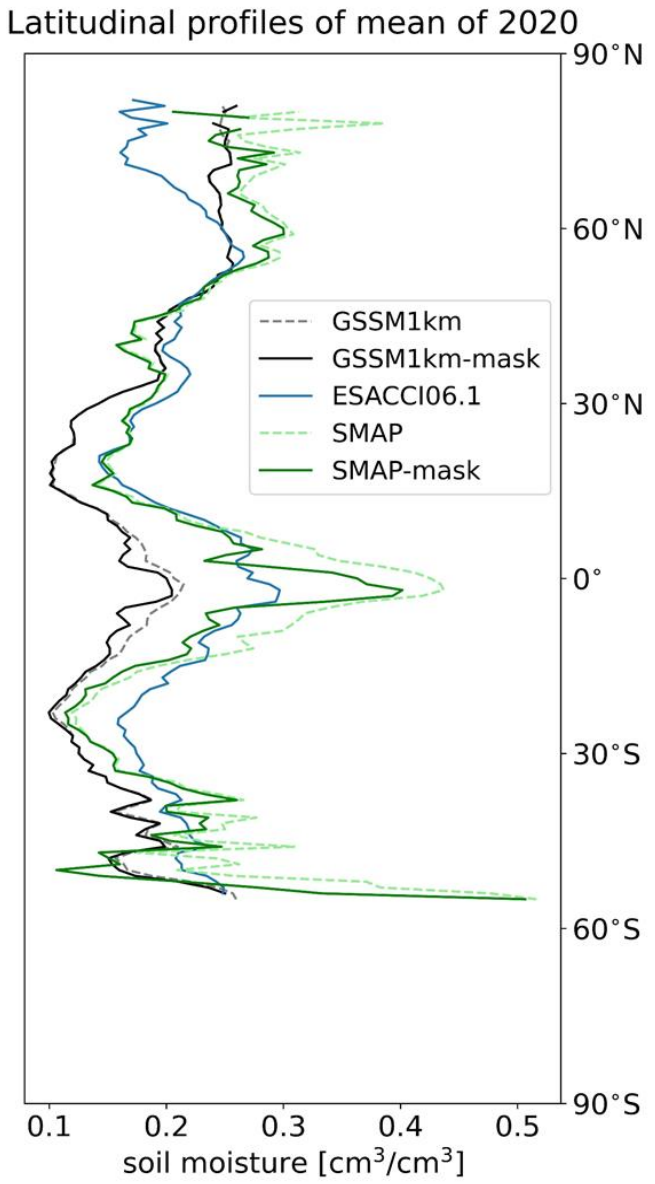
(a) Model testing



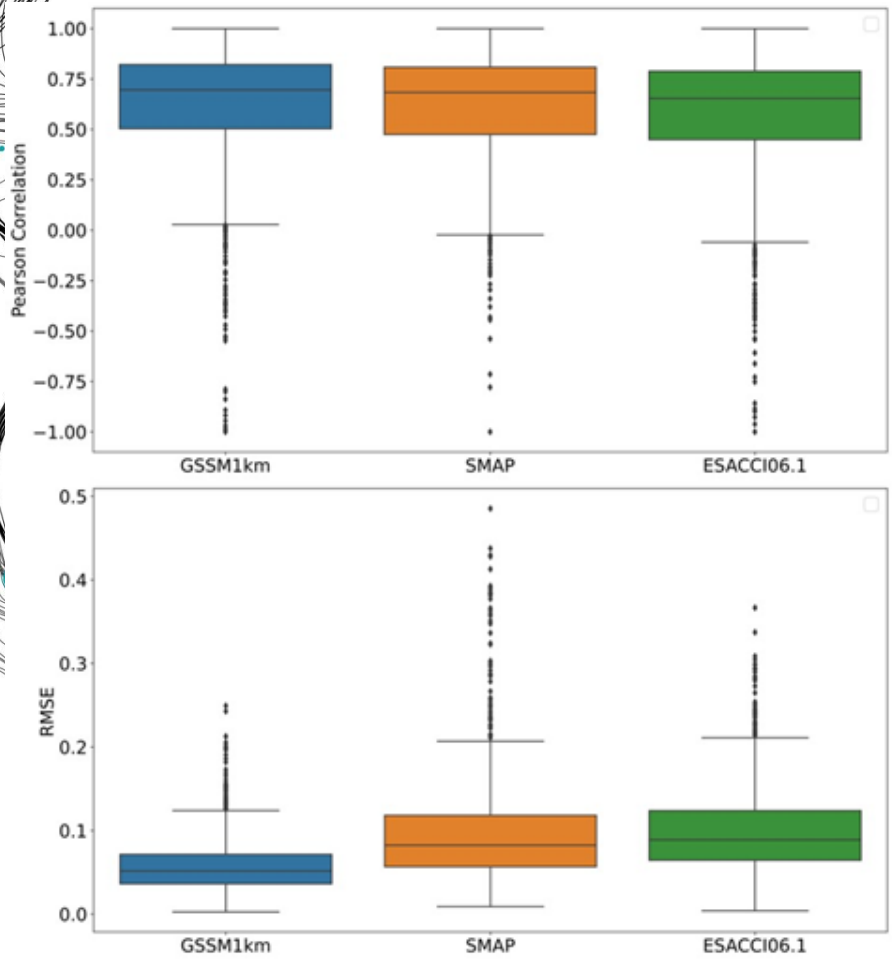
(b) Feature Importance

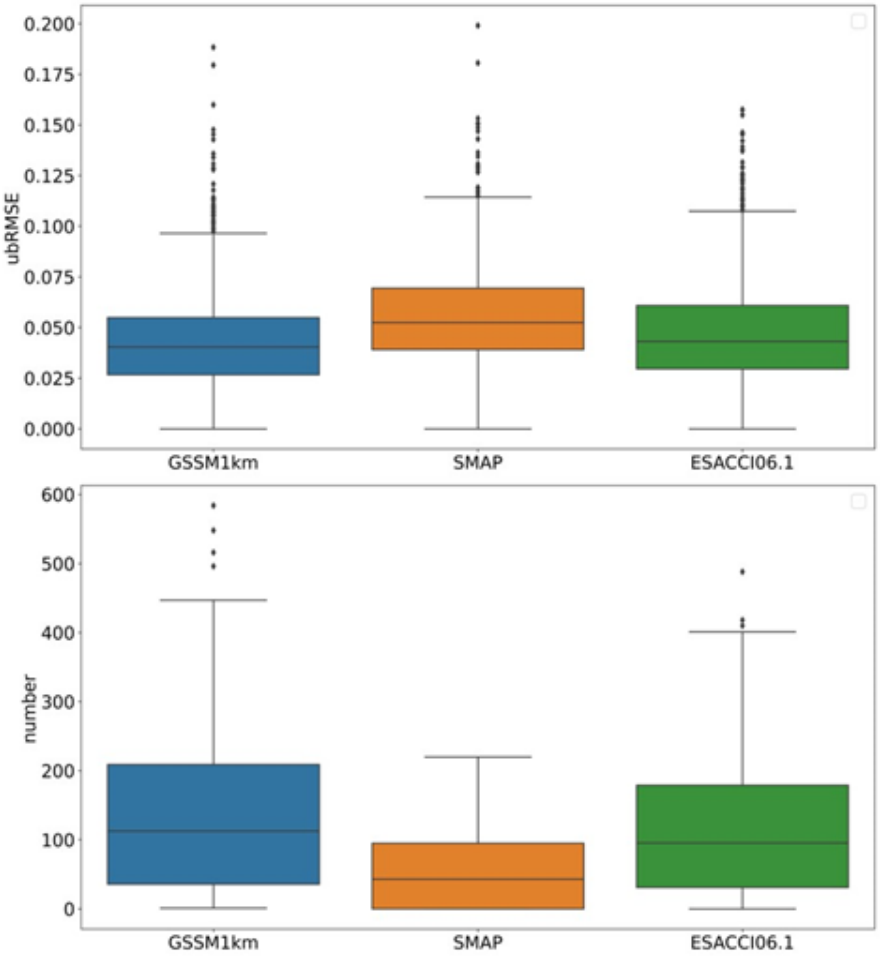
3. RESULTS


Mean of GSSM1km soil moisture in 2020
Mean of ESACCI06.1 soil moisture in 2020
Mean of SMAP soil moisture in 2020
Latitude [°]
Longitude [°]
soil moisture [cm³/cm³]


Latitudinal profiles of mean of 2020
GSSM1km
GSSM1km-mask
ESACCI06.1
SMAP
SMAP-mask
soil moisture [cm³/cm³]

3. RESULTS

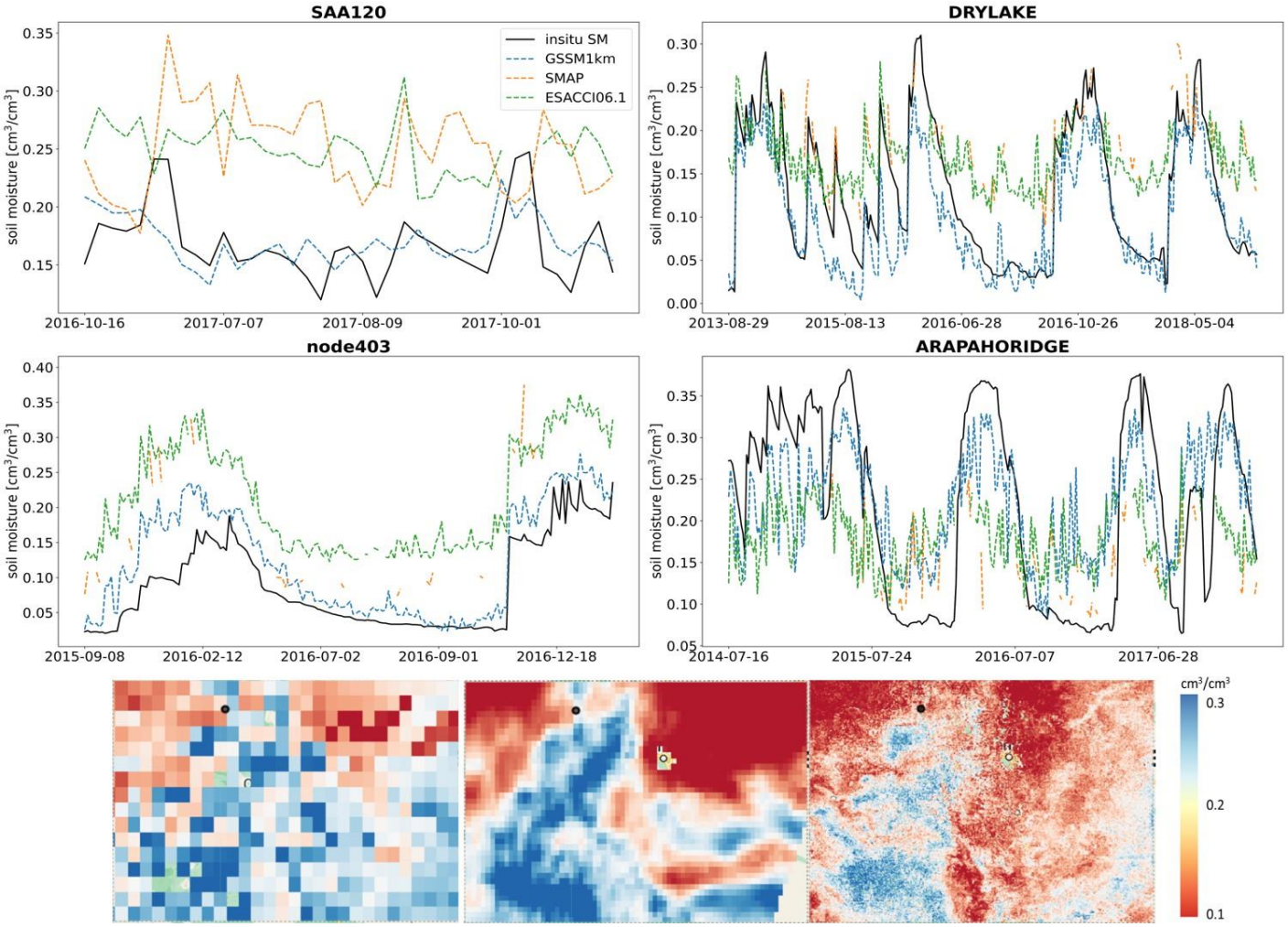

Pearson Correlation
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
GSSM1km
SMAP
ESACCI06.1


ubRMSE
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
GSSM1km
SMAP
ESACCI06.1

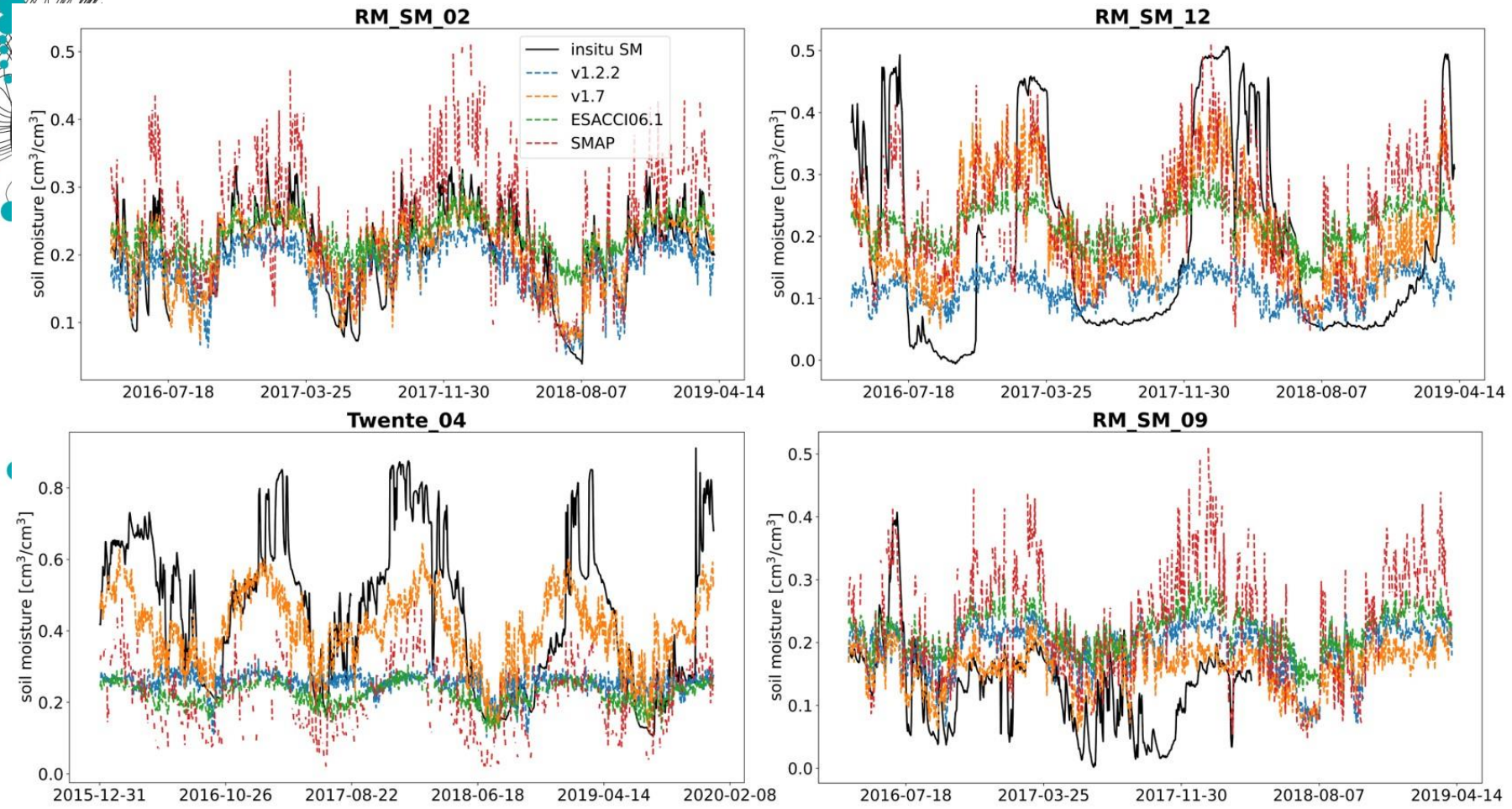
RMSE
0.5
0.4
0.3
0.2
0.1
0.0
GSSM1km
SMAP
ESACCI06.1

number
600
500
400
300
200
100
0
GSSM1km
SMAP
ESACCI06.1

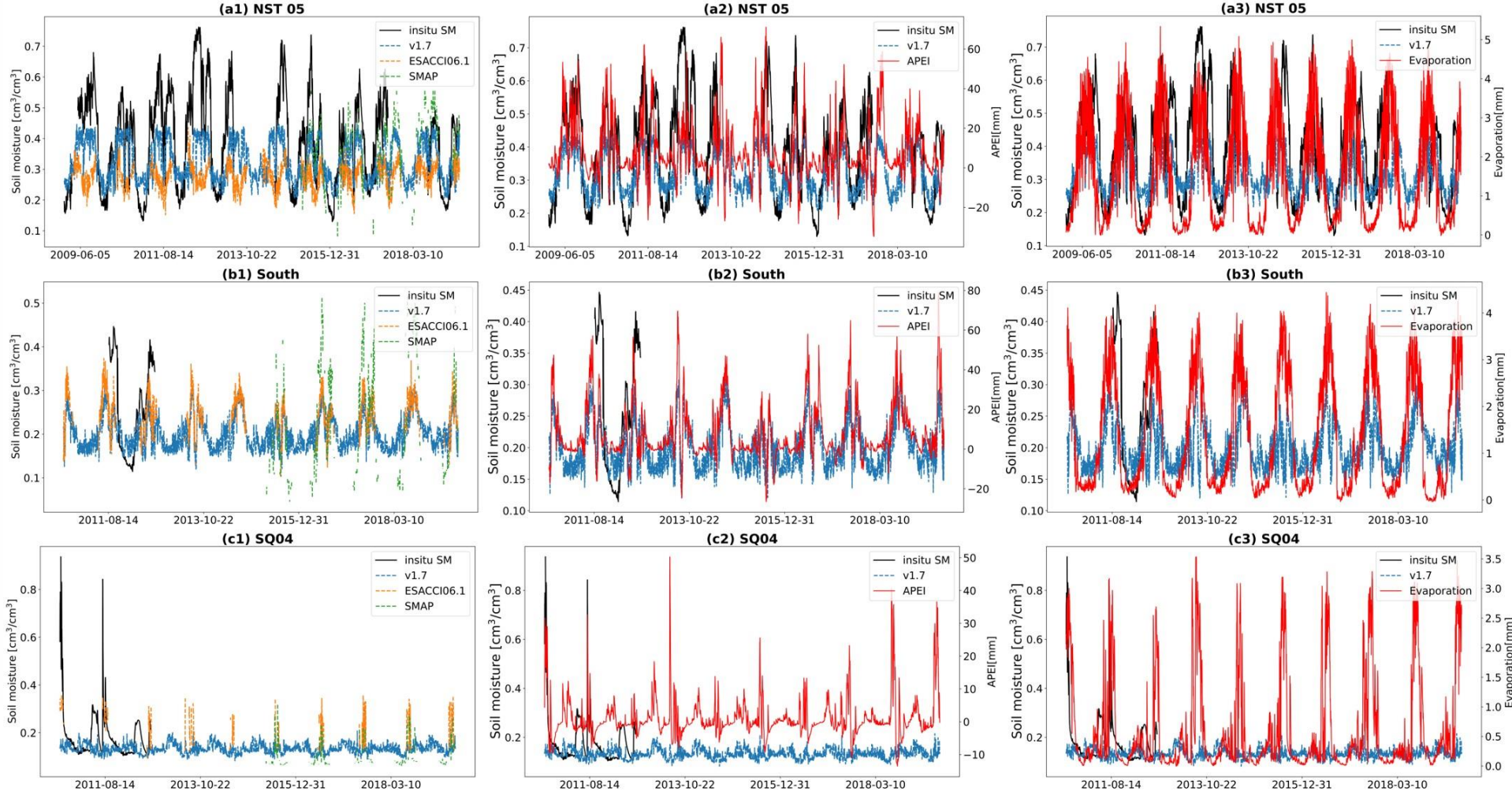
3. RESULTS


SAA120
insitu SM
GSSM1km
SMAP
ESACCI06.1
soil moisture [cm³/cm³]
2016-10-16
2017-07-07
2017-08-09
2017-10-01
DRYLAKE
2013-08-29
2015-08-13
2016-06-28
2016-10-26
2018-05-04
node403
2015-09-08
2016-02-12
2016-07-02
2016-09-01
2016-12-18
ARAPAHORIDGE
2014-07-16
2015-07-24
2016-07-07
2017-06-28
cm³/cm³

3. RESULTS


RM_SM_02
RM_SM_12
Twente_04
RM_SM_09
insitu SM
v1.2.2
v1.7
ESACCI06.1
SMAP
soil moisture [cm³/cm³]
2016-07-18
2017-03-25
2017-11-30
2018-08-07
2019-04-14
2015-12-31
2016-10-26
2017-08-22
2018-06-18
2020-02-08

3. RESULTS



UNIVERSITY OF TWENTE.

5. CONCLUSION AND FUTURE WORK

- The RF regression was trained based on the synchroised in-situ SM and other land surface features (e.g. API, LST, NDVI/EVI, TI…)
- The testing results show that the RMSE is 0.05 cm3/cm3 and Pearson Correlation Coefficient is 0.88 at the global scale.
- The evaluation results of the RF regression model at in-situ stations also show satisfactory performance.

THANK YOU FOR YOUR LISTENING!

q.han@utwente.nl