

# **Estimating Daily Soil Moisture at High Spatial Resolution for Drought Monitoring by Fusing Multi-Source Data Based on Random Forest**

XINRONG LI

July, 2023

SUPERVISORS:

Dr. Y. Zeng

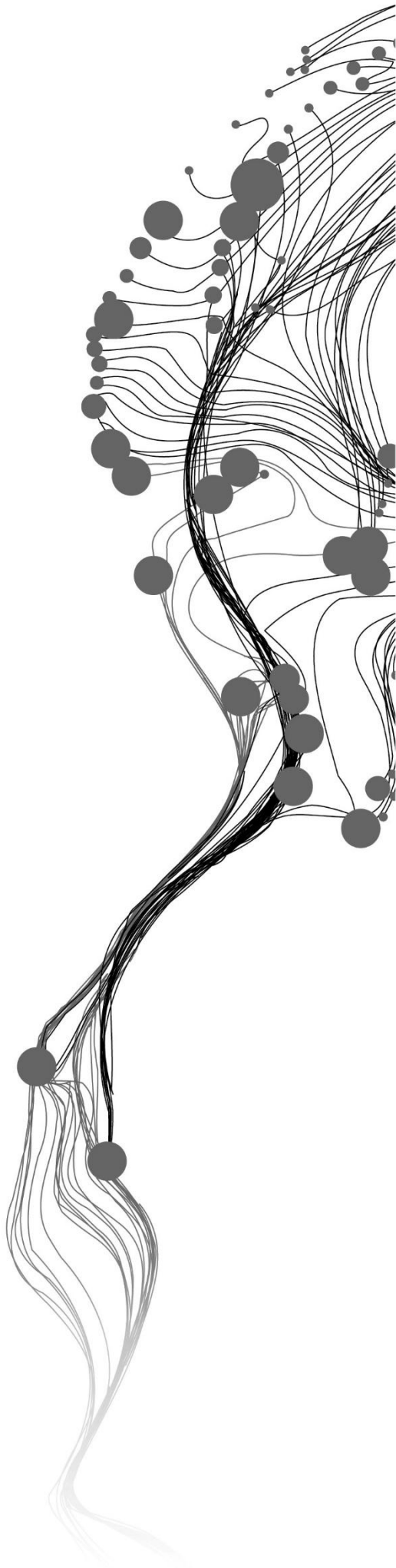
Prof.dr. Z. Su

ADVISORS:

MSc. Q Han

Ir. A.M. van Lieshout





# **ESTIMATING DAILY SOIL MOISTURE AT HIGH SPATIAL RESOLUTION FOR DROUGHT MONITORING BY FUSING MULTI-SOURCE DATA BASED ON RRANDOM FOREST**

XINRONG LI

Enschede, The Netherlands, [July, 2023]

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Water Resource and Environmental Management

## **SUPERVISORS:**

Dr. Y. Zeng  
Prof.dr. Z. Su

## **ADVISORS:**

MSc. Q Han  
Ir. A.M. van Lieshout

## **THESIS ASSESSMENT BOARD:**

Prof.dr. D. van der Wal (Chair)  
Dr. Lijie Zhang (Forschungszentrum Juelich, Germany)

#### DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

## ABSTRACT

The Netherlands is grappling with increasingly severe drought conditions as a result of climate change. Soil moisture (SM) is a significant indicator for drought monitoring, but available local/regional SM products lack high spatiotemporal resolution. In this study, we generate a high-resolution (10m) SM product for the Twente region in the Netherlands by downscaling the original 1km SM product with auxiliary variables (day time land surface temperature, night land surface temperature, Ground Range Detected dataset, Enhanced Vegetation Index, Leaf Area Index, precipitation, soil texture, groundwater level) using a random forest model. The RF model showed an acceptable result with correlation coefficient of 0.8 and RMSE of  $0.0418(\text{cm}^3 \text{ cm}^{-3})$ . Feature importance of the RF model indicates that day time land surface temperature and groundwater level are the most significant features, followed by clay and precipitation. The spatial and temporal distribution of downscaled 10m SM product is consistent with the original 1km SM product. Applying the 10m SM product for drought monitoring indicates the downscaled soil moisture product can effectively capture drought conditions.

Keywords: soil moisture, high resolution, random forest, sentinel 1/2, drought monitoring, Twente, the Netherlands

## ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who have supported and contributed to the completion of my thesis.

I am immensely grateful to my supervisor, Dr. Y. Zeng and Prof.dr. Z. Su, for their guidance and support during the research process. Their expertise and insightful feedback have been instrumental in shaping the thesis.

I would also like to extend my appreciation to my advisors, MSc. Q. Han and Ir. A.M. van Lieshout, for their valuable input. And also many thanks to the external examiner Dr. Lijie Zhang and the chair Prof. Dr. D. van der Wal. Their suggestions in the proposal defense and midterm have contributed to improving the quality of my thesis.

Furthermore, I would like to thank my friends and family. Their constant support and belief in my abilities have been a source of motivation.

Last but not the least, I would like to thank myself for the efforts and perseverance. Especially during the COVID-19 pandemic, it was not easy to get through these two years.



# TABLE OF CONTENTS

---

List of figures.....	v
List of tables.....	v
1. Introduction.....	1
2. Objectives and Research Questions.....	5
2.1. Objectives.....	5
2.2. Research Questions.....	5
3. Study Area and dataset.....	6
3.1. Study Area.....	6
3.2. Dataset.....	7
4. Methodology.....	9
4.1. Data Preprocessing and Harmonization.....	9
4.2. Machine Learning Model.....	12
4.3. SM Product Generation and Evaluation.....	12
5. Result AND Discussion.....	13
5.1. Random Forest Model.....	13
5.2. Product Validation.....	14
5.3. Comparison with original SM datasets.....	18
5.4. Application.....	19
6. Conclusion and recommendation.....	27
List of references.....	28



## LIST OF FIGURES

---

Fig 3.1 Study Area.....	6
Fig 4.1 Flowchart.....	9
Fig 4.2 Distribution of In Situ SM Stations.....	12
Fig 5.1 Model Performance and Feature Importance.....	16
Fig 5.2 Comparison of SM from in-situ, 10m product and 1km product in Twente stations.....	17
Fig 5.3 Mean SM map from 2018 to 2019.....	18
Fig 5.4 Time-latitude diagram over the Twente.....	18
Fig 5.5 Time-longitude diagram over Twente.....	20
Fig 5.6 Drought distribution evaluated by SSMI in Twente, 2018.....	21
Fig.5.7 Comparison of precipitation and LST in the Twente region in 2018 and 2019 with the long-term average.....	23
Figure 5.8 GWD distribution in Twente, 2018.....	24

# LIST OF TABLES

---

Table 3.1 Required Dataset..... 7

Table 5.2 Hyperparameter tuning..... 13

Table 5.3 Evaluation metrics of 1km SM product(original) and 10m SM product(downscaled) in Twente stations..... 15

Table 5.4 Drought classification of SSMI..... 20

# 1. INTRODUCTION

In the context of global warming, the acceleration of the water cycle increases the odds of extreme weather events. The Netherlands experienced one of the most severe droughts on record in 2018, which posed threats to agriculture, ecology, etc., and caused huge economic losses to the entire country. Furthermore, many studies indicate there is a trend of worsening drought in the Netherlands under climate change (Cook et al., 2018a; Philip et al., 2020a; Zhao et al., 2020).

Soil moisture (SM) is an indicator for drought monitoring because of its essential role in the water and energy cycle (André et al., 1986; Idso et al., 1975). Drought monitoring aims to track the severity and location of drought. Detailed drought descriptions help decision-makers develop measures to reduce drought-related losses. However, local/regional drought monitoring is facing difficulties with the lack of high spatiotemporal resolution SM data. This is mainly due to the fact that in situ measurements provide reliable point-scale high temporal resolution SM observation but lack the spatial coverage. While remotely sensed SM data provide local/regional spatial monitoring they are at very coarse spatial resolution (e.g., 25km, 9km, 1km). Therefore, we need a high-resolution SM product for better field water condition monitoring in the Netherlands to provide more accurate drought monitoring across the country.

In situ measurement is regarded as the most reliable and robust method to obtain SM (Yang et al., 2021). The in situ measurement networks in Raam region and Twente in the Netherlands provide references for remote sensing SM product validation (Benninga et al., 2018; Dente et al., 2012). Because of the high investment and maintenance cost of sensors, in situ measurements can only detect SM in a limited area. As an alternative, remote sensing technique has been introduced to SM monitoring for years. Passive and active remotely sensed SM is obtained by inverting the measured emitted or reflected microwave radiation from the land surface. Passive microwave radiometers provide frequent but coarse resolution SM observations while active microwave sensors retrieve SM at higher spatial resolution with low frequency. Although microwave technology has been greatly improved to better retrieve SM in recent years, it is still facing the difficulty to remove vegetation effects on the microwave emission of soils. Optical remotely sensed SM is derived from the relationships between land surface parameters, and thermal approaches rely on variations in soil surface temperature caused by different SM. Compared with passive and active remotely sensed SM, optical and thermal methods have better spatiotemporal resolutions, but they are sensitive to both vegetation and the weather conditions (Zhuang et al., 2023).

Coarse remotely sensed SM limits the application of SM at finer spatial resolution. Fortunately, data fusion approach can help address partially this issue. Data fusion is the process of assembling multiple data to represent a more useful real-world object (Zeng et al., 2016; Zhuang et al., 2020). Zhan et al. (2006) merged 36 km radiometer brightness temperature and 3 km radar backscatter with ancillary data including surface roughness, vegetation water content and surface temperature to generate the SM data at 9 km. Das et al. (2011) generated the gridded 9 km SM by integrating 36 km near-surface soil moisture L-band radiometer retrievals and 3km L-band radar observations. Same as Zhan et al.'s study, Das et al. took advantage of the passive and active SM observations. Estimating SM by fusing optical image with microwave data was proposed in SM retrieval studies which combined the strengths of microwaves and

optical sensing techniques (He et al., 2014; Mattar et al., 2012; Notarnicola et al., 2006). The arrival of Sentinel 1 and Sentinel 2 give us the chance to map SM with higher details. The twin Sentinel 1 satellites (Sentinel 1a and Sentinel 1b) provide high spatial resolution SAR images every 6 days while Sentinel 2a and Sentinel 2b provide high spatial resolution Visible (VNIR) and Near Infra-Red (NIR) to the Short Wave Infra-Red (SWIR) images with a 5-day revisit frequency (Drusch et al., 2012; Potin, 2019). The studies synthesizing sentinel 1 SAR images with sentinel 2 can generate SM with better resolution (Hajj et al., 2017; Ma et al., 2020), so do other SM fusion studies using multiple sensor. However, the high computational cost of the fusion algorithms restricts the generation of near real-time SM.

Lately, the machine learning method was applied to downscale SM (Su et al., 2020). The computationally-efficient machine learning (ML) provides a possibility to better establish nonlinear relationships between environmental factors and SM. Abbaszadeh et al. (2019) proposed a random forest (RF) based model with predictors including normalized difference vegetation index (NDVI), land surface temperature (LST), precipitation, elevation and soil texture to generate 1 km resolution SM. Zhang et al. (2021) applied in-situ observation-constrained RF for the SM estimates at a global scale with the same predictors, and the spatiotemporal patterns of the SM were compared against the in-situ observations, showing reasonable agreements. Lv et al. (2021) used environmental variables such as LST, the NDVI, the normalized shortwave-infrared difference bare soil moisture indices (NSDSI), the digital elevation model (DEM), and calculated slope data (SLOPE) to estimate SM at 1 km resolution with artificial neural networks (ANN). In the study of estimating SM with gradient boosting decision tree regression (GBDT) over the Tibetan plateau, 26 indices frequently used in SM downscaling were analyzed by filter method for their relative importance and high relative importance indices Enhanced Vegetation Index (EVI), Distance Drought Index (DDI), Modified Perpendicular Drought Index (MPDI), Modified Shortwave Infrared Perpendicular Water Stress (MSPSI), NDVI, Modified Perpendicular Drought Index (PDI), Shortwave Single Slope Index (SASI) (Wei et al., 2019). Although the above studies lead to acceptable predictions, feature selection is still a challenge in ML-based SM downscaling (Sachindra & Kanae, 2019). Appropriate variables not only reduce the complexity of the model but also result in a better estimation. Apart from predictors, the choice of algorithm also affects the SM estimate. There are many studies that compared the performance of different algorithms in generating SM products. Im et al. (2016b) compared the performance of random forest (RF), boosted regression trees, and cubist in generating 1 km SM in Korea and Australia. The result indicates the superior performance of RF with high accuracy and robustness. In the comparison study of six different algorithms in estimating SM, the effectiveness of RF was confirmed again (Liu et al., 2020). The same result was seen in the Yan and Bai (2020)'s comparison study. In summary, RF is the best-performing ML algorithm in the existing SM studies.

In this study, we use sentinel 1 Ground Range Detected (GRD) dataset and vegetation indices (EVI and LAI) from sentinel 2 with environmental variables (precipitation, soil texture, Surface Reflectance and ground water level (GWL)) and 1km SM to generate a high-resolution (10m) SM product for the Netherland by RF. The 1km dataset consist of aggerated 1 km predictors and 1km SM will be used to train, validate and test the model. Then we use the 10m predictors to generate the 10m SM product for the Netherlands. The generated high resolution SM product is evaluated by the in situ measurement and the importance of the predictors are analyzed. We expect the generated 10 m SM product can better reveal the field water conditions to achieve more accurate drought monitoring in the Netherlands under global warming.

Due to the computational limitations, the 10m soil moisture data we generated is only for the Twente region in this study, the workflow of which can be easily scaled up to the whole Netherlands.



## 2. OBJECTIVES AND RESEARCH QUESTIONS

### 2.1. Objectives

Main objective: To generate a high-resolution (10m) SM product for the Netherlands

Sub-objective 1: Selecting the appropriate predictors for the model

Sub-objective 2: Completion of preprocessed and harmonized multi-source data at 1km daily and 10m daily resolution

Sub-objective 3: Build the RF based SM downscaling model.

Sub-objective 4: Generate and evaluate the gridded daily SM at 10m in the Netherlands

### 2.2. Research Questions

RQ1: What variables can be considered as predictors?

RQ2: How to downscale the auxiliary data into 10m resolution with acceptable accuracy?

RQ3: How to split the dataset?

RQ4: How is the performance of the ML based model?

RQ5: What is the spatial and temporal variation of the 10m SM product?

### 3. STUDY AREA AND DATASET

#### 3.1. Study Area

The Netherlands is located between latitude 51°-54°N, and longitude 3°-7°E. Influenced by the North Sea and the Atlantic Ocean, the country enjoys the moderate maritime climate, which is cloudy and humid for most of the year. The Netherlands' average annual temperature and precipitation is around 10 °C and 790mm and the general features of the climate are mild winters and cool summers.

The Netherlands experienced extreme drought in 2018 which caused huge economic losses and many studies indicate there is a more significant trend of worsening drought in the Netherlands under climate change (Cook et al., 2018b; Philip et al., 2020b; Zhao et al., 2020).

Twente is the region located in the eastern Netherlands. There are fourteen municipalities that make up the area, including Almelo, Borne, Dinkelland, Enschede, Haaksbergen, Hellendoorn, Hengelo, Hof van Twente, Losser, Oldenzaal, Rijssen-Holten, Tubbergen, Twenterand, and Wierden.

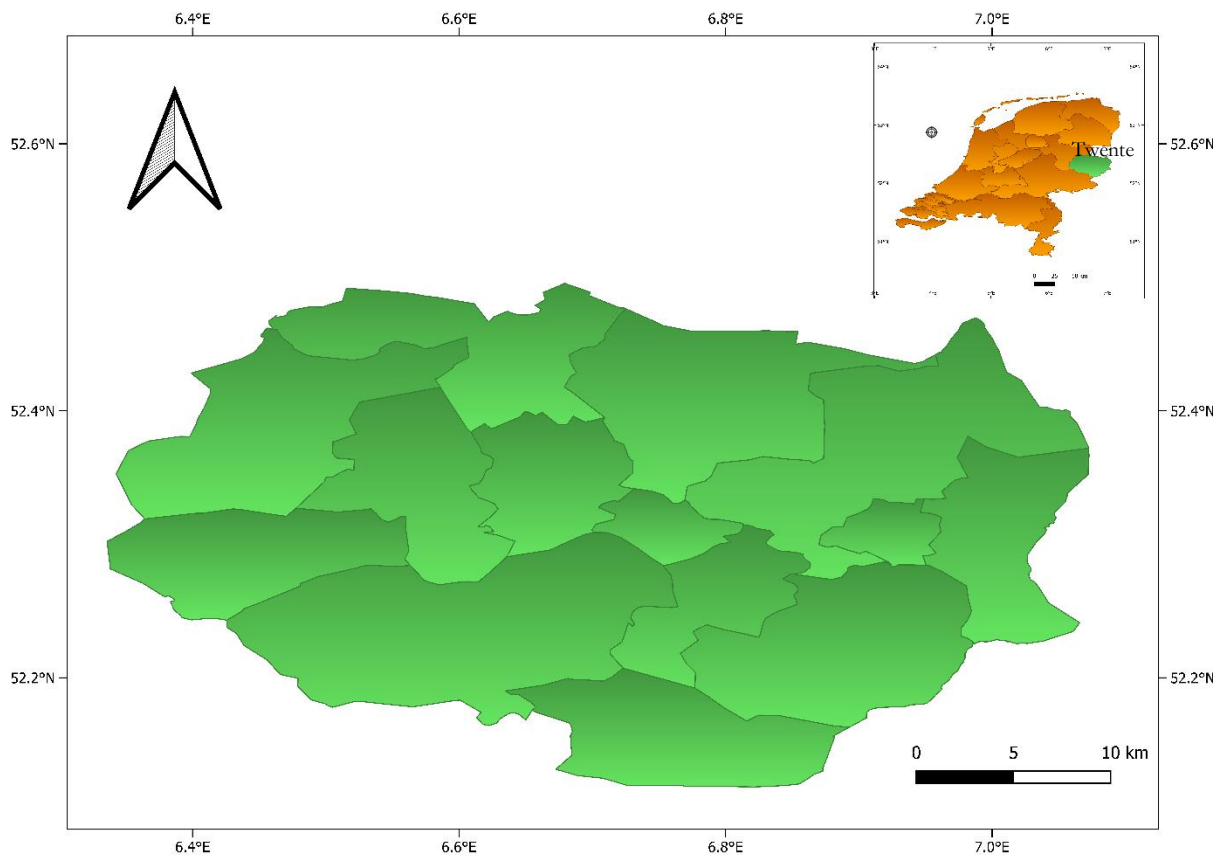


Fig 3.1 Study Area



### 3.2. Dataset

In this research, auxiliary data including precipitation, LST, soil texture, surface reflectance, GWL will be selected as predictors with sentinel 1 level 1 GRD and vegetation indices (EVI and LAI) from sentinel 2 and 1 km SM simulated from in situ measurement to generate daily SM at 10 m. Other details of the dataset are shown in the following table.

Table 3.1 Required Dataset

Variable	Description	Source	Data Duration	Resolution	Unit
Precipitation	The KNMI precipitation dataset provides gridded 24 hour precipitation accumulations from climatological gauge-adjusted radar dataset for The Netherlands	<a href="#">KNMI Data Platform</a>	01/01/2008-01/01/2023	1 km/Hourly	mm
LST	LST_day and LST_night are obtained through interpolation from MODIS 1km LST_day and LST_night data.	(Han et al., 2023)	01/01/2000-01/01/2020	1 km/ Daily	°C
Soil Texture	Soil texture is the proportion of sand, silt and clay content. sized particles that make up the mineral fraction of the soil.	<a href="#">OpenGeoHub</a>	/	30m	clay % sand % bulk 10kg/m <sup>3</sup>
LAI	The Leaf Area Index is an important indicator reflecting the growth status of the plant population. It is the one-sided green leaf area over a unit of land.	<a href="#">Sentinel-2</a>	28/03/2017-present	10m/ 5d	/
		<a href="#">Modis</a>	18/02/2000-present	500m/8d	/
EVI	Enhanced vegetation index is an 'optimized' vegetation index. It is calculated by NIR, red, C1, C2, blue and L bands observed by Sentinel-2.	<a href="#">Sentinel-2</a>	28/03/2017-present	10m/ 5d	/
		<a href="#">Modis</a>	18/02/2000-present	250m/16d	/ K

ESTIMATING DAILY SOIL MOISTURE AT HIGH SPATIAL RESOLUTION FOR DROUGHT MONITORING BY FUSING MULTI-SOURCE DATA BASED ON RANDOM FOREST

GWL	The 1km GWL is generated from The DINO groundwater database.	<a href="#">DINO</a>	Depend on station	pointed	m
GRD	The sentinel-1 L1 Ground Range Detected dataset provided by GEE is a calibrated, ortho-corrected product.	<a href="#">Sentinel-1</a>	01/10/2014-present	10m/6d	dB
SM	The in situ measurement datasets are from the Raam region and Twente.	In situ measurements( Dente et al., 2012)	01/01/2016-31/12-2019	15 minutes	cm <sup>3</sup> /cm <sup>3</sup>
SM	The 1km SM was generated by global in situ dataset.	(Han et al., 2023)	01/01/2000-01/01/2020	1km/daily	cm <sup>3</sup> /cm <sup>3</sup>

## 4. METHODOLOGY

The methodology in this study consists of three steps: data processing and harmonization, build the SM model and generate and evaluate the 10m SM product. Each step corresponds to one or two sub-objectives while answering several research questions. The flowchart (see Fig 4.1) shows the process of the study.

### 4.1. Data Preprocessing and Harmonization

There are many variables that affect SM. The commonly used predictors in the existing SM simulation studies include precipitation, Surface Reflectance (SR), LST and soil texture (see details in the introduction). Except for these variables, our study includes groundwater level (GWL) as the predictor. Groundwater effects on soil moisture is not explicitly treated in most of the existing SM simulation studies. However, GWL has a strong effect on SM especially in areas with shallow groundwater (Chen & Hu, 2004). Inspired by the precipitation simulation studies which used auxiliary information as independent covariables (Baez-Villanueva et al., 2020; Wu et al., 2020; Ling Zhang et al., 2021), we select GRD from sentinel 1 and EVI and LAI as predictor variables.

In the following steps, we train the RF model at 1km and then generate the 10m SM by the trained model with the predictors at 10m. Therefore, we need to resample predictors at 1km daily and 10m daily resolution respectively. Before the aggregation, we preprocessed the Sentinel 2 data and GWL data. For the data from optical satellite Sentinel 2, we remove the clouds by the characteristics of the QA band and then calculate EVI and LAI follow the equations below (R. Zhang et al., 2022) (Boegh et al., 2002). GWL from DINO is point observation and we interpolate it into raster by kriging. GRD data from Sentinel-1 is calibrated and ortho-corrected. We have selected the VV polarization, which is more sensitive to soil moisture (Alexakis et al., 2017).

$$EVI = G \frac{N-R}{N+C_1R-C_2B+L} \quad \text{eq.1}$$

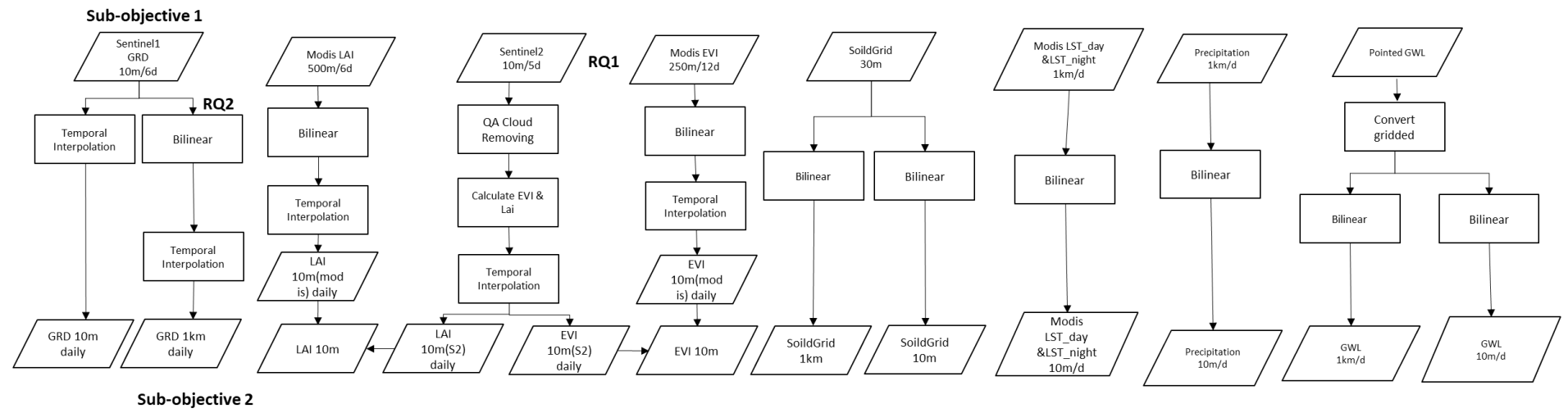
$$LAI = 3.618 * EVI - 0.118 \quad \text{eq.2}$$

Where N, R and B are the atmospherically corrected surface reflectance of near-infrared, red and blue bands; G is the gain factor; C<sub>1</sub>, C<sub>2</sub> are the coefficients of the aerosol attenuation term; L is Canopy background brightness correction factor (Jiang et al., 2008).

We use bilinear method to resample predictors into 1km and 10m daily resolution. Bilinear interpolation is the image scaling algorithm that utilizes the four real pixel values surrounding a virtual point in the source image to determine a pixel value in the target image. Its scaling effect is much better than nearest neighbor interpolation, and the complexity is lower than that of bicubic interpolation (Kian et al., 2008). [Temporal linear interpolation](#) is used to harmonize all the predictors into daily resolution.

There are a large number of non-value areas after cloud-removing in sentinel 2 LAI and EVI, we use 10m EVI and LAI resample from 250m resolution to fill the gaps. Additionally, 1km LAI and EVI are resampled from Modis EVI and LAI at 250m respectively.

**Step1:Data preprocessing and harmonization**



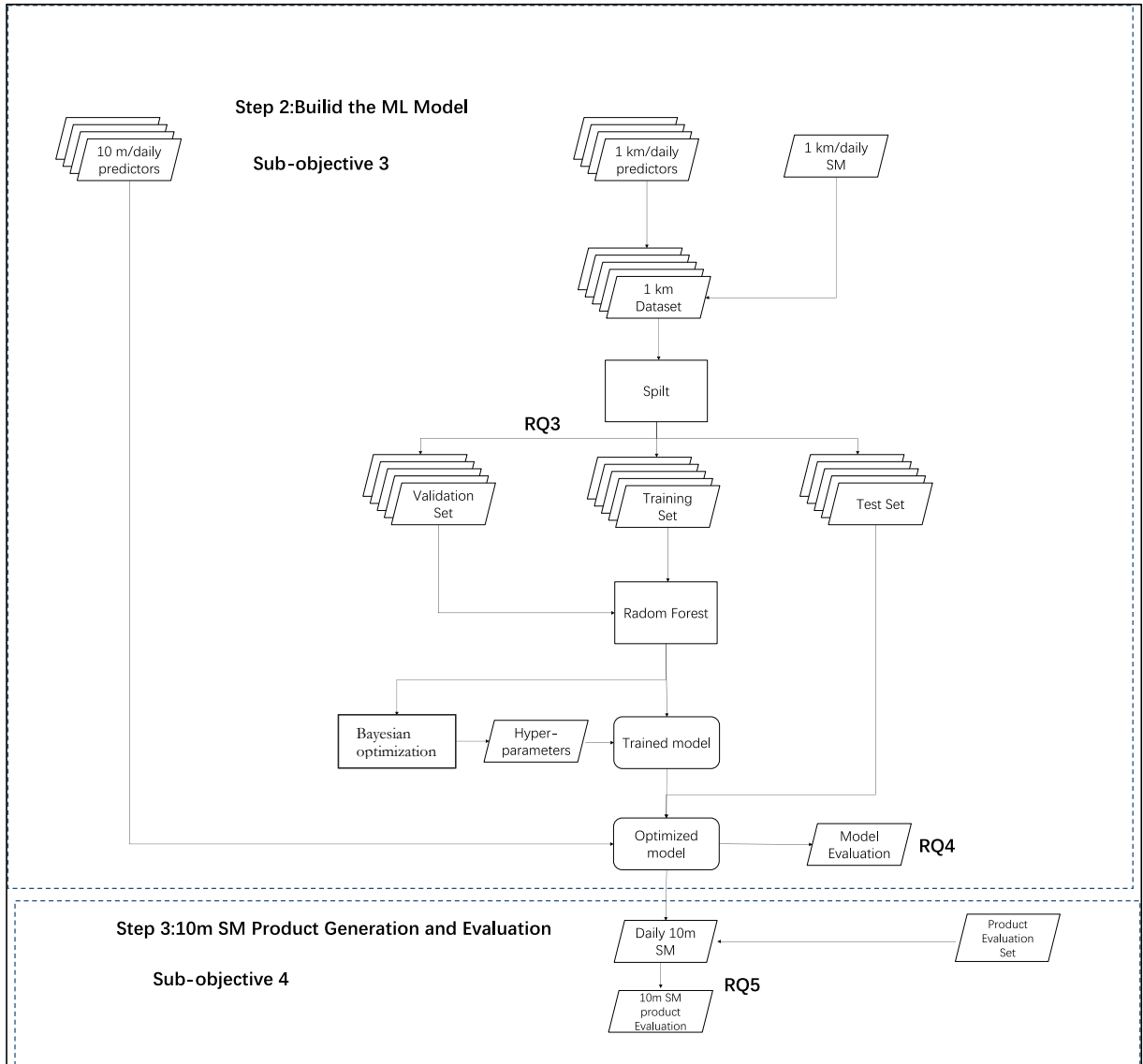


Fig 4.1 Flowchart

## 4.2. Machine Learning Model

### 4.2.1. Spilt the dataset

Our 1km dataset covers the period from 2018 to 2020. Based on the seasonal variation of SM, we spilt the dataset by time that validation set consists of data from January, April, July, and October of 2018 (first month of each season), while the test set includes data from the same months in 2019. The remaining data are assigned to the training set. Training set is used to train the model and validation set and test set is used to tune the hyperparameters and test the model performance respectively.

### 4.2.2. Build the RF model

Random Forest is an ensemble learning algorithm consisting of multiple decision trees (Liaw & Matthew, 2004). In regression tasks, it makes the best prediction by averaging the results of each individual tree (Ho, 1995, 1998). Based on the Bootstrap Aggregation (Bagging) strategy, RF samples with random replacements make the algorithm less likely to overfit and less sensitive to noise (Altman & Krzywinski, 2017). Such advantages lead to the wide application of RF.

We train the RF model with the python library scikit-learn. The training set is used to train the model. Validation set is used to tune the hyperparameters: number of trees, min sample leaf, min sample spilt and max depth, which influence the accuracy and complexity of the RF model. We apply the [bayesian optimization](#) to tune the model. Bayesian optimization is an efficient method for hyperparameter tuning which works by building a posterior distribution of functions (Gaussian process) that accurately models the function to be optimized (Snoek et al., 2012). The tuned hyperparameters form the optimal model with the correspondingly trained model.

After getting the optimal model, we will test the model performance by comparing the predicted result with the in situ SM. Several statistical metrics for testing include Root Mean Square Error (RMSE), unbiased Root Mean Square Error (ubRMSE), Pearson Correlation Coefficient ( $r$ ) and Mean Difference (MD) (Zhang et al., 2021). The equations for the metrics can be seen below. Before generating the 10m SM, we analyze the importance of the predictors.

$$RMSE = \sqrt{\frac{\sum_i^N (y_{pred,i} - y_{ref,i})^2}{N}} \quad \text{eq.3}$$

$$r = \frac{\sum (y_{pred} - \bar{y}_{pred})(y_{ref} - \bar{y}_{ref})}{\sqrt{\sum (y_{pred} - \bar{y}_{pred})^2} \sqrt{\sum (y_{ref} - \bar{y}_{ref})^2}} \quad \text{eq.4}$$

$$MD = \frac{\sum_i^N (y_{pred,i} - y_{ref,i})}{N} \quad \text{eq.5}$$

$$ubRMSE = \sqrt{RMSE^2 - MD^2} \quad \text{eq.6}$$

## 4.3. SM Product Generation and Evaluation

After building the RF-based model, we apply the 10m resolution predictor variable to predict the SM for the Twente region in the Netherlands. And eight SM in situ measurement stations were selected to evaluate the generated SM product with the above metrics. The temporal and spatial distribution of the product is analysed during the evaluation process.

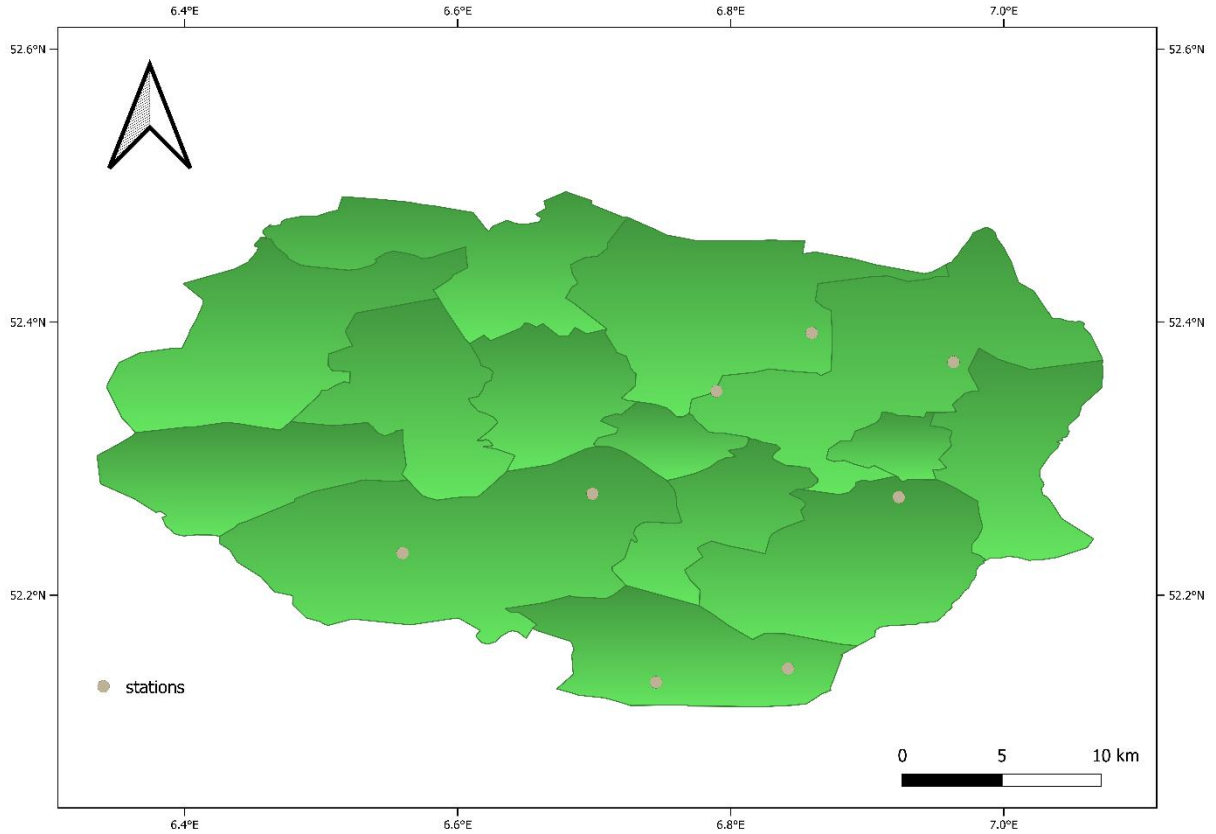


Fig 4.2 Distribution of In Situ SM Stations

## 5. RESULT AND DISCUSSION

### 5.1. Random Forest Model

#### 5.1.1. Model Tuning

We tuned the hyperparameters `n_estimators`, `min_samples_leaf` in the range of 10 to 200, and 10 to 50 respectively by [bayesian optimization](#). Bayesian optimization is the method used in machine learning for hyperparameter tuning and model optimization which combines Bayesian inference and optimization algorithms (Snoek et al., 2012). The target for an optimization problem depends on the specific case. Our study, prediction of 10m SM, is a regression problem, therefore we select RMSE as the target. Furthermore, we add a minus to RMSE for application since bayesian optimization aims to solve maximization problems (in the context of model performance, a lower RMSE indicates better performance). The ten iterations result were shown in the following table.

Table 5.1 Hyperparameter tuning

iter	target(-RMSE)	min_samples_leaf	n_estimators
1	-0.02211	11	180
2	-0.02868	41	154

3	-0.0221	12	179
4	-0.02776	33	200
5	-0.0233	13	10
6	-0.02905	42	13
7	-0.02871	42	83
8	-0.02169	10	40
9	-0.02158	10	98
10	-0.02157	10	126

In the 10th iteration the target reached the optimized result with the hyperparameters `min_samples_leaf = 10`, and `n_estimators=126`.

### 5.1.2. Model testing

There are 4076385 records in the test set and the set is used to test the model performance. The RF model shows an acceptable result with  $r$  of 0.8, RMSE of  $0.0418(\text{cm}^3 \text{cm}^{-3})$ , MD of  $-0.0019(\text{cm}^3 \text{cm}^{-3})$  and ubRMSE of  $0.0418(\text{cm}^3 \text{cm}^{-3})$ . Fig 5.1(b) indicates the importance of 12 features. LST\_day and GWL are the most significant features followed by clay and precipitation while EVI and LAI show the least importance. The possible reason that EVI and LAI are least important is that they are resampled from 16 days/ 8 days resolution.

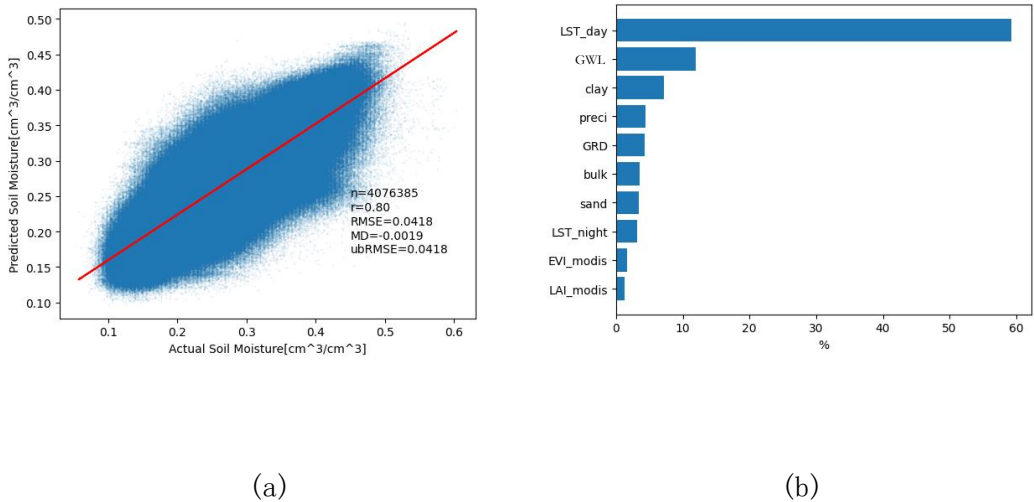


Fig 5.1 Model Performance and Feature Importance

## 5.2 Product Validation

We calculate the evaluation metrics among the original SM product(1km) and downscaled SM product(10m). The results demonstrate that the original SM product has a better relationship with the in situ measurements than the downscaled SM product because measured SM at Twente stations (Han et al., 2023). The 10m SM time series have been analyzed with the in-situ measurement, 1km SM and the most significant predictor variable LST\_day at Twente stations. The completely opposite trend between the 10m SM time series and LST\_day across all stations demonstrates that the model is consistent with the underlying physical mechanisms (e.g., wetter soil is cooler). The diagrams illustrate that the 10m SM product is able to capture the dynamic of 1km SM product and suggest that the 10-meter product performs acceptably, although not capturing the peaks and valleys. It may be related to the representativeness of Modis LST\_day. Because of lack of in situ measurement of daily LST, we compare

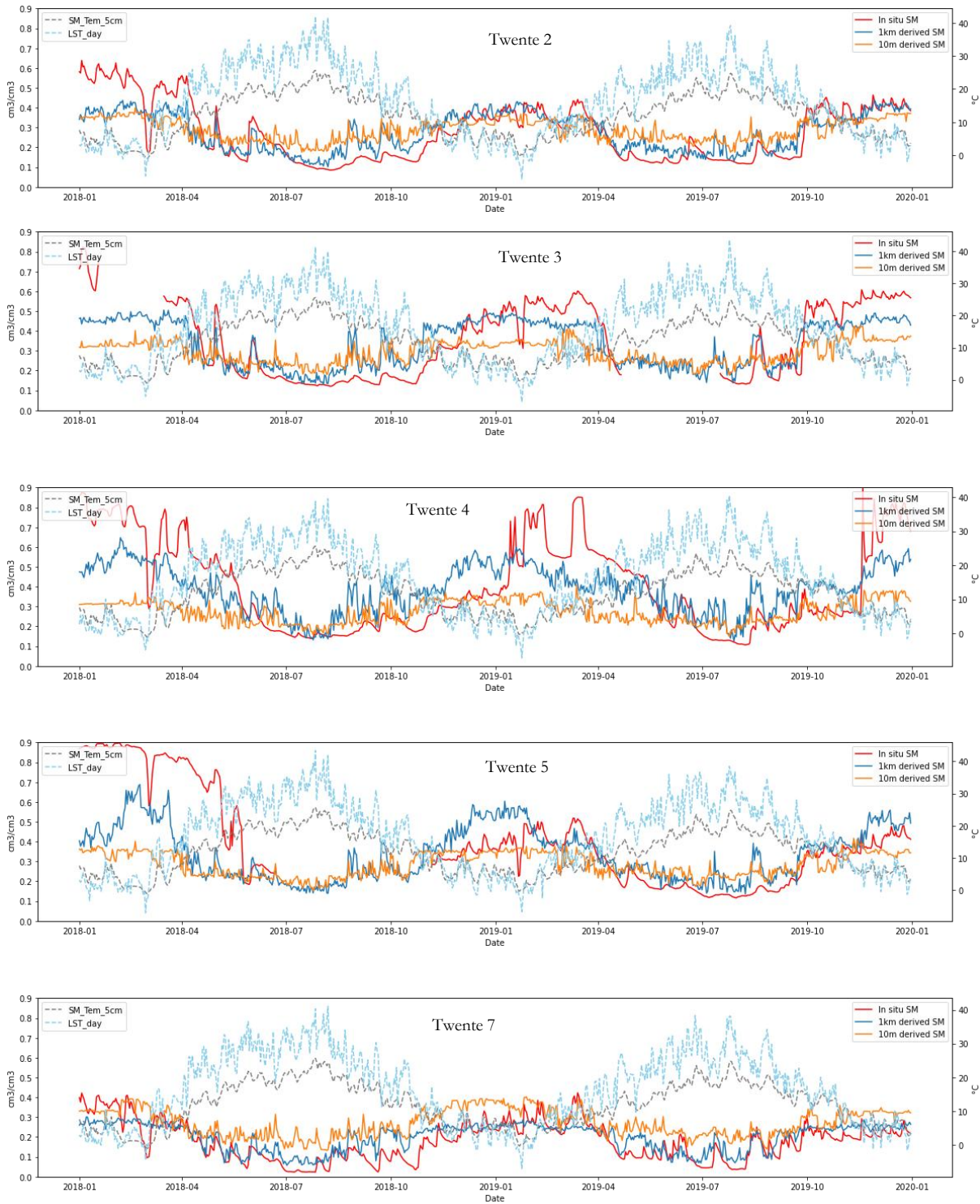


modis lst\_day with measured soil temperature at 5cm(SM\_Tem\_5cm). To have the same temporal resolution as MODIS LST\_day, soil temperature at 5 cm is averaged of daytime and night-time. During winter period, soil temperature at 5cm is higher than LST\_day at Twente stations. While at summer period, the soil temperature at 5cm is lower than that of MODIS LST\_day. This is plausible. Nevertheless, MODIS LST\_day is at 1km resolution, which represents a much smoother LST at regional scale than at point scale. Therefore, we infer that the MODIS LST\_day is smoother than the actual measured land surface temperature at 10 m resolution or at Twente stations, which leads to a smoother predicted soil moisture at 10m resolution. Duan et al. (2019) also noted MODIS LST\_day lower than the in situ LST\_day in the Modis LST validation study in the United State. LST\_day as the most significant predictor variable and showed a negative correlation with soil moisture. Resampling LST\_day from 1km to 10m resulted in information loss and the loss of information caused the predicted 10m SM in a smaller range, thereby failing to capture the peaks and valleys.

Table 5.3 Evaluation metrics of 1km SM product(original) and 10m SM product(downscaled) in Twente stations

station	metrics	1km SM/in situ SM	10m SM/in situ SM
T2	RMSE	0.08	0.11
	R	0.84	0.74
	MD	0.00	0.01
	ubRMSE	0.08	0.11
T3	RMSE	0.18	0.19
	R	0.85	0.68
	MD	0.02	-0.06
	ubRMSE	0.18	0.18
T4	RMSE	0.16	0.23
	R	0.73	0.60
	MD	0.01	0.13
	ubRMSE	0.16	0.20
T5	RMSE	0.21	0.25
	R	0.57	0.47
	MD	0.07	-0.13
	ubRMSE	0.20	0.21
T7	RMSE	0.06	0.11
	R	0.80	0.74
	MD	-0.02	0.09
	ubRMSE	0.06	0.07
T8	RMSE	0.07	0.12
	R	0.57	0.49
	MD	-0.02	0.10
	ubRMSE	0.07	0.07
T9	RMSE	0.07	0.10
	R	0.79	0.67
	MD	-0.02	0.06
	ubRMSE	0.06	0.08
T11	RMSE	0.16	0.29

R	0.54	0.44
MD	0.04	-0.24
ubRMSE	0.15	0.16



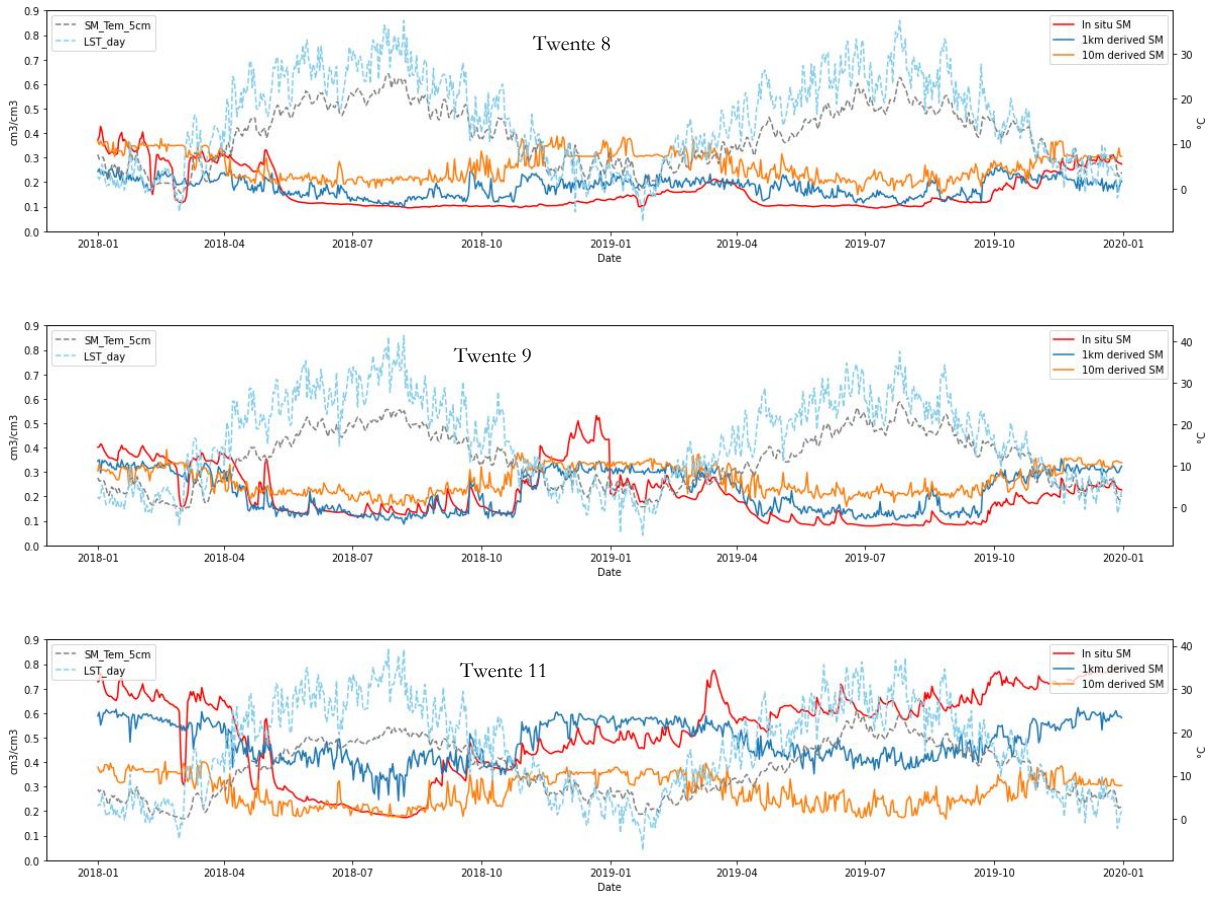


Fig 5.2 Comparison of SM from in-situ, 10m product and 1km product in Twente stations

### 5.3 Comparison with original SM datasets

#### 5.3.1 Spatial Patterns

Figure 5.3 shows the average 1km SM(original) and 10m SM (downscaled) from 2018 to 2019 over the Twente region. 1km SM is in the range of 0.15 to 0.5 while 10m SM in the range of 0.2 to 0.35. Compared to the original SM product, the downscaled product has less variation. As we mentioned in 5.2, it is related to the loss during resampling MODIS\_LST. Nevertheless, a similar spatial pattern, west Twente is wetter than east Twente, can be observed in the two maps. Compared with the 1km SM map, the 10m SM map shows much more details of relatively wet areas in the western parts of Twente.

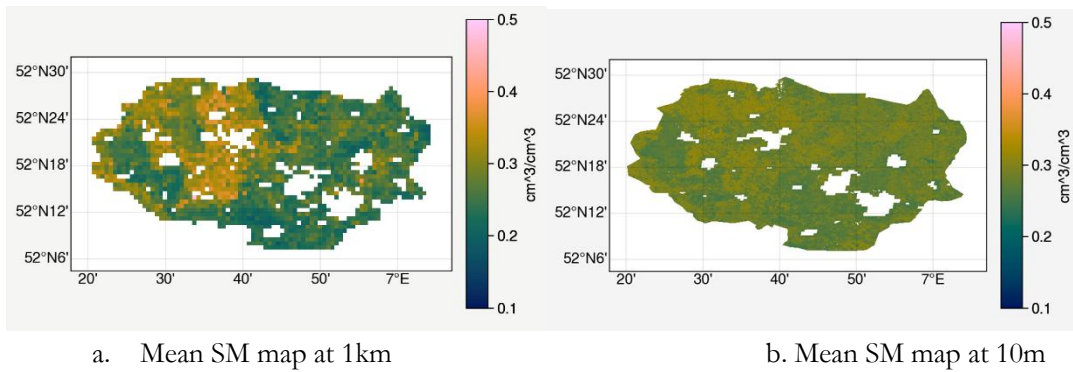


Fig 5.3 Mean SM map from 2018 to 2019 (The white areas indicate urban areas)

#### 5.3.2 Spatio-Temporal Patterns

The time-latitude diagram represents the average pixel value along the latitude for the years 2018 and 2019 in Twente region. The diagram shows an obvious seasonal variability: the soil is dry during the summer and autumn months, while it is moist during the spring and winter seasons. Compared with the year 2019, SM has a longer duration with lower values in dry season (summer and autumn) in year 2018. The similar seasonal variability shows in 1km time-latitude diagram. In contrast to the diagram at 10m resolution, relatively low SM values can be observed at 52.49N in 1km time-latitude, and this is because 1km SM maps have less information at 52.49N while 10m maps show the variability at this latitude.

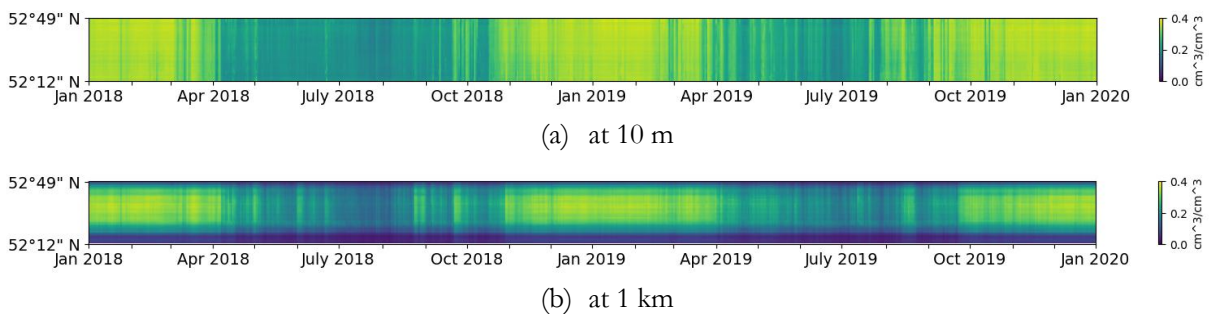


Figure 5.4 Time-latitude diagram over the Twente

The time-longitude diagram represents the average pixel value along the longitude for the years 2018 and 2019. Seasonal variability, similar to that observed in the time-latitude diagram, can be observed in the

time-longitude diagram. Furthermore, the spatial pattern of western Twente being wetter than eastern Twente can be observed in Fig.5.5. Compared to the diagram at 10m resolution, relatively low SM values can be observed at 6.33E and 7.10E in 1km diagram and the limited amount of information attributed to this pattern.

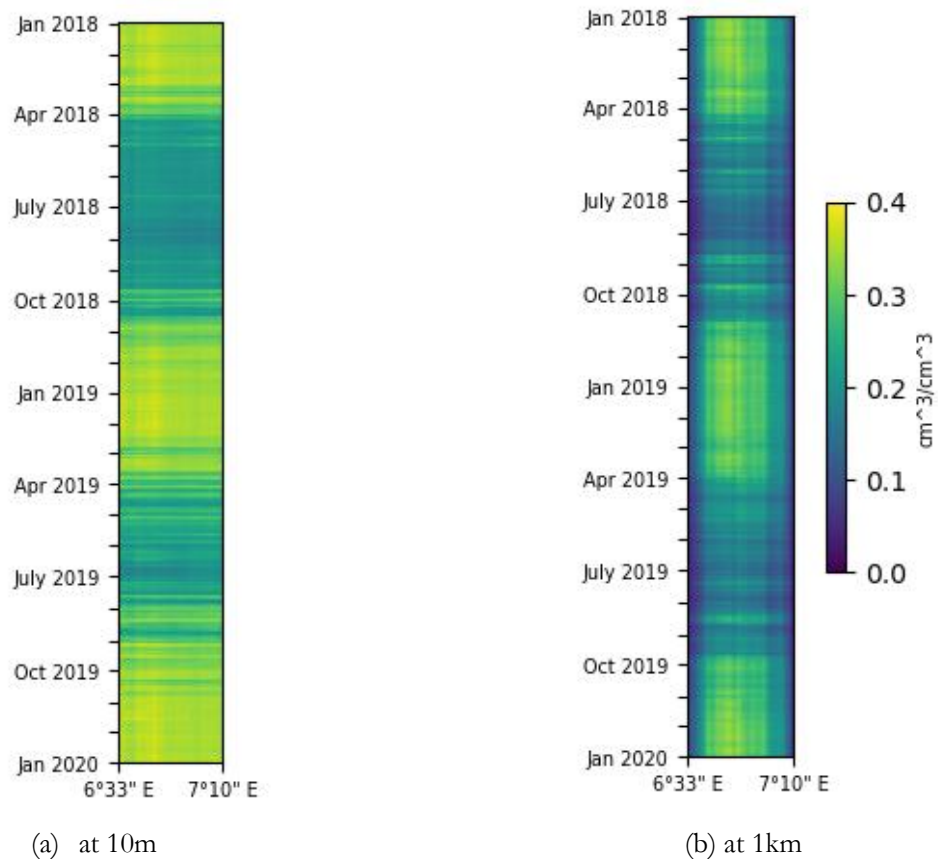


Figure 5.5 Time-longitude diagram over Twente

## 5.4 Application

In this section, we analyse the drought distribution at 10m by Standardized Soil Moisture Index (SSMI) with the generated 10m SM product. It would be ideal to analyse the drought events (and its severity) with 30 years of high resolution soil moisture data. However, Sentinel 1/2 data are only available from 2014/2017, which limit the approach in this study for generating long-term 10m soil moisture data. Additionally, it is computationally expensive to generate SM even just for 2016-2020. Therefore, we only presented the 10m SM for 2018 and 2019, and analyse the drought distribution and spread at Twente region only.

### 5.4.1 Drought index:SSMI

The Standardized Soil Moisture Index (SSMI) is a widely used monthly-scale drought index that monitors drought conditions by comparing current soil moisture to historical records for a specific location (Xu et al., 2018). In order to better monitor drought conditions in the year 2018 and 2019, we calculated the single-month-scale SSMI at 10m resolution in Twente region. The equation of SSMI and the Drought classification of SSMI are presented below (Su et al., 2017).

$$SSMI = \frac{SM_i - \mu}{\sigma} \quad \text{eq.7}$$

where  $SM_i$  is the average SM of month  $i$ ,  $\mu$  is the mean value of soil moisture content for multiple years, and  $\sigma$  is the standard deviation.

Table 5.4 Drought classification of SSMI

Value range	Categories	Return period (1 in ...)
$F \leq \mu - 3\sigma$	Extremely dry	15787
$\mu - 3\sigma < F \leq \mu - 2\sigma$	Severely dry	370
$\mu - 2\sigma < F \leq \mu - 1\sigma$	Moderate dry	22
$\mu - 1\sigma < F \leq \mu + 1\sigma$	Near normal	3
$\mu + 1\sigma < F \leq \mu + 2\sigma$	Moderate wet	22
$\mu + 2\sigma < F \leq \mu + 3\sigma$	Severely wet	370
$\mu + 3\sigma < F$	Extremely wet	15787

Due to the imitation of 10m SM dataset, we calculate  $\mu$  (mean) and  $\sigma$  (standard deviation) using 1km SM data from 2004 to 2019, and then resample them to a 10m resolution.

#### 5.4.2 Drought Distribution

Fig.5.6 (a) demonstrates the monthly scale SSMI in 2018. In the beginning two months of 2018, the SSMI distribution of Twente region indicated a mainly non-drought with sparkling mild drought in the western part. From March, drought is prevalent throughout the region that extreme and severe drought started spreading from the western parts and covered almost the entire region until July, with 80% of the region experiencing extreme drought, while the rest was primarily affected by severe drought. From August, the drought situation began to have a turn. In the western parts of Twente region, it transitioned from extreme drought to moderate and mild drought in September and October, while it shifted from severe drought to mild drought in the eastern areas. The SSMI spatial distribution in November returned to the distribution observed in the Jan and Feb in 2018, and the distribution remained until to Feb, 2019. From March, drought was prevalent throughout the region again and the extreme and severe drought started spreading from the west Twente as 2018. However, the most severe drought occurred in June in 2019 rather than July in 2018, and the areas experiencing extreme drought are much smaller than the peak of the drought in 2018. The areas experiencing extreme drought decrease from June and returned to the spatial pattern observed in the first three months of 2018 in October.

From time perspective, the entire area entered a state of drought from March-April in 2018 and 2019 and experienced the most extreme drought during late spring and early summer. In 2019, most of the areas experienced alleviation of drought conditions as the fall season progressed. However, in 2018, the drought persisted until late fall. From a spatial perspective, western Twente suffered more severe drought than eastern Twente and eastern Twente recovered more quickly than Western Twente when facing extreme drought conditions.

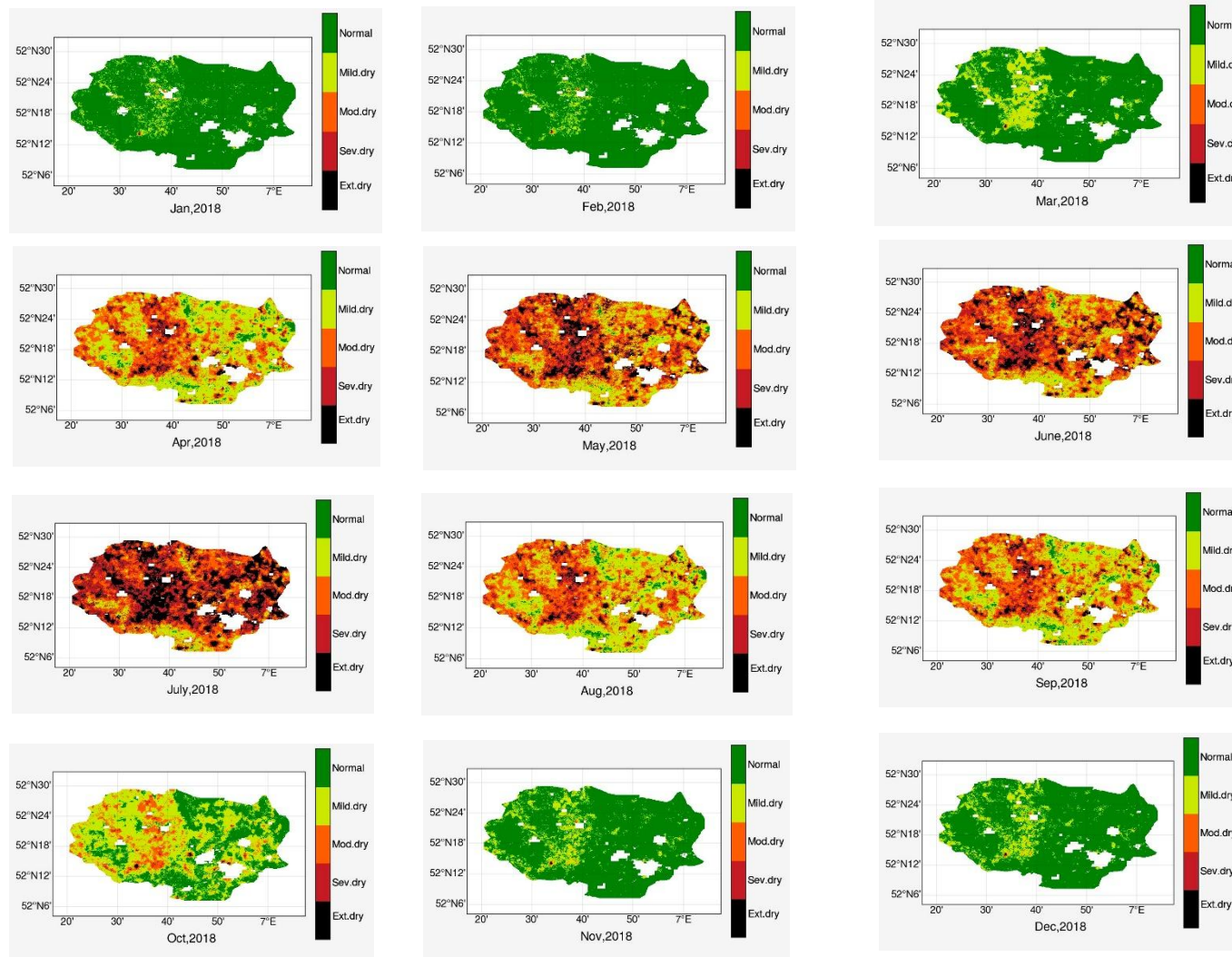


Figure 5.6 (a) Drought distribution evaluated by SSMI in Twente, 2018

ESTIMATING DAILY SOIL MOISTURE AT HIGH SPATIAL RESOLUTION FOR DROUGHT MONITORING BY FUSING MULTI-SOURCE DATA BASED ON RANDOM FOREST

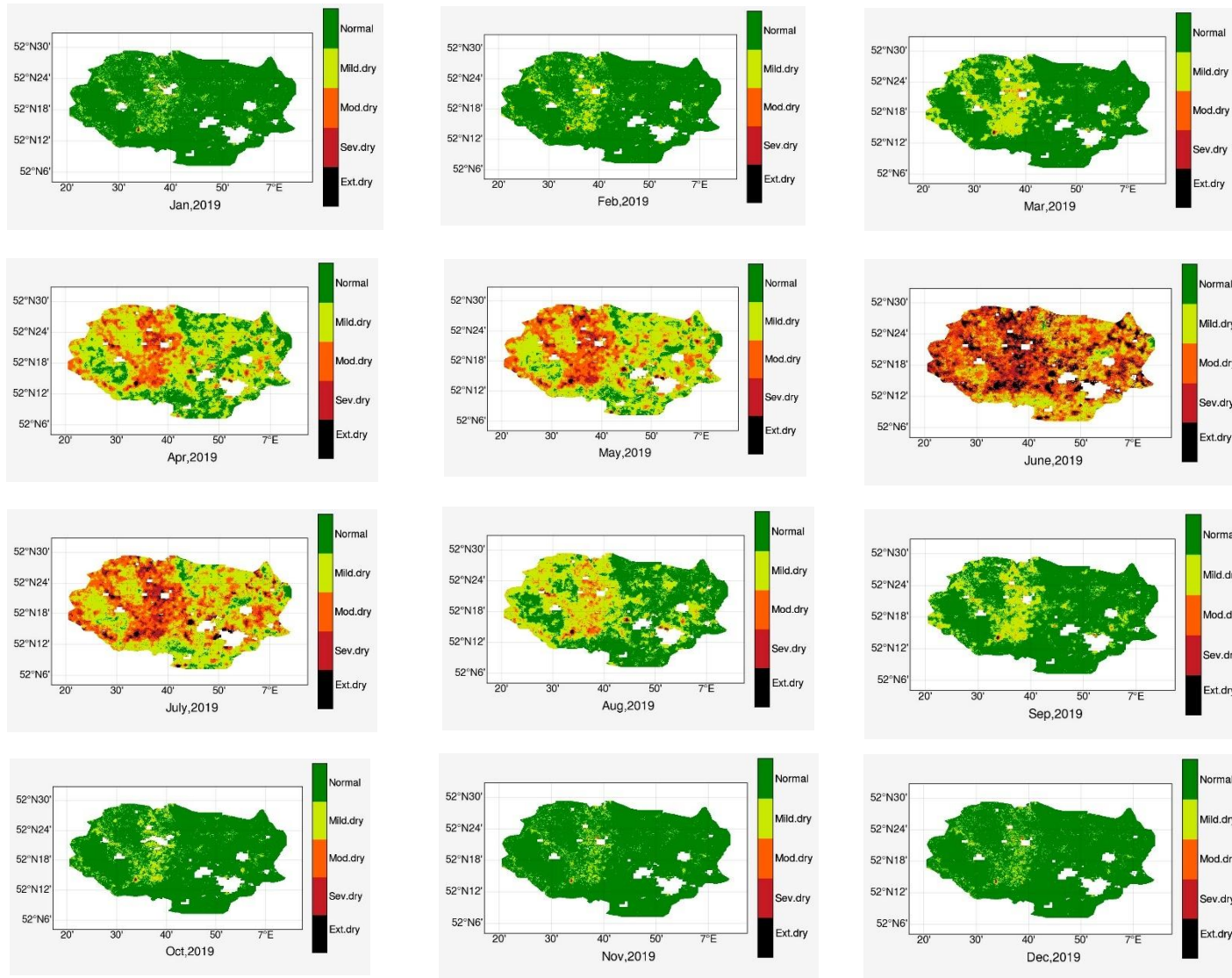


Figure 5.6 (b) Drought distribution evaluated by SSMI in Twente, 2019



### 5.4.3 Driving factors of drought distribution

SSMI quantifies the SM deficit and the SM deficit results from insufficient precipitation or/and excessive evaporation demand (Luo et al., 2017). We analyze the monthly precipitation and LST (i.e., reflecting the evaporation demand) from 2018 to 2019 in Twente with the multiple years average (2004-2017). The results indicate that precipitation below the multi-year average in Feb, May to Nov and LST above the multi-year average from Apr to Nov in 2018, while precipitation below the multi-year average in Jan, May to Aug and LST above the multi-year average from Jan to Sep(except May) in 2019. The occurrence of drought is highly consistent with insufficient precipitation and excessive evapotranspiration that drought begins when rainfall is below the average and LST exceed the average. During the most severe drought in July 2018 and June 2019, there is a significant deficiency in rainfall and excessive LST. Therefore, meteorological factors are the causes of the temporal distribution of drought.

Groundwater can be a soil water source, especially in shallow groundwater areas like the Netherlands. Shallow groundwater replenishes soil moisture in the root zone through capillary rise. Due to the fact that groundwater stations vary in observation periods and most of the stations do not have continuous observation data, it is hard to conclude how much extent groundwater contributes to the temporal distribution of drought. In order to understand the influence of groundwater on the distribution of droughts, we analyse the groundwater depth(GWD) distribution in Twente from 2018 to 2019. GWD distribution exhibits temporal variations throughout the year. Starting in February 2018, GWD declines from the central area of Twente and spread to the entire region by October 2018. From November 2018 to March 2019, GWD gradually recovers. In April 2019, GWD experiences a decline from the central area again, and it recovers by September, 2019. The decline of GWD is similar to the occurrence of drought in the Twente region, as they both spread from the central area to the entire region. It can be inferred that groundwater contributes to the distribution of drought. Additionally, GWD declines during the drought period and recovers after the end of drought. It indicates that drought affects GWD while GWD influences the drought distribution.

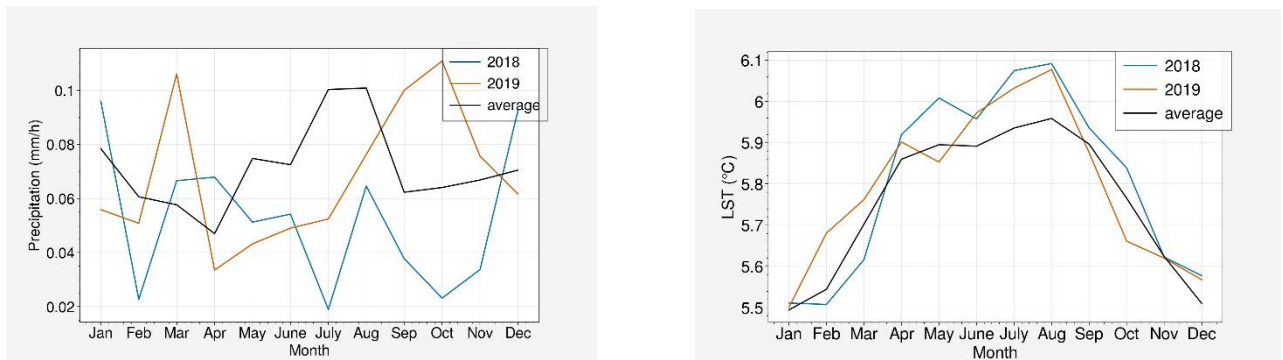


Fig.5.7 Comparison of precipitation and LST in the Twente region in 2018 and 2019 with the long-term average

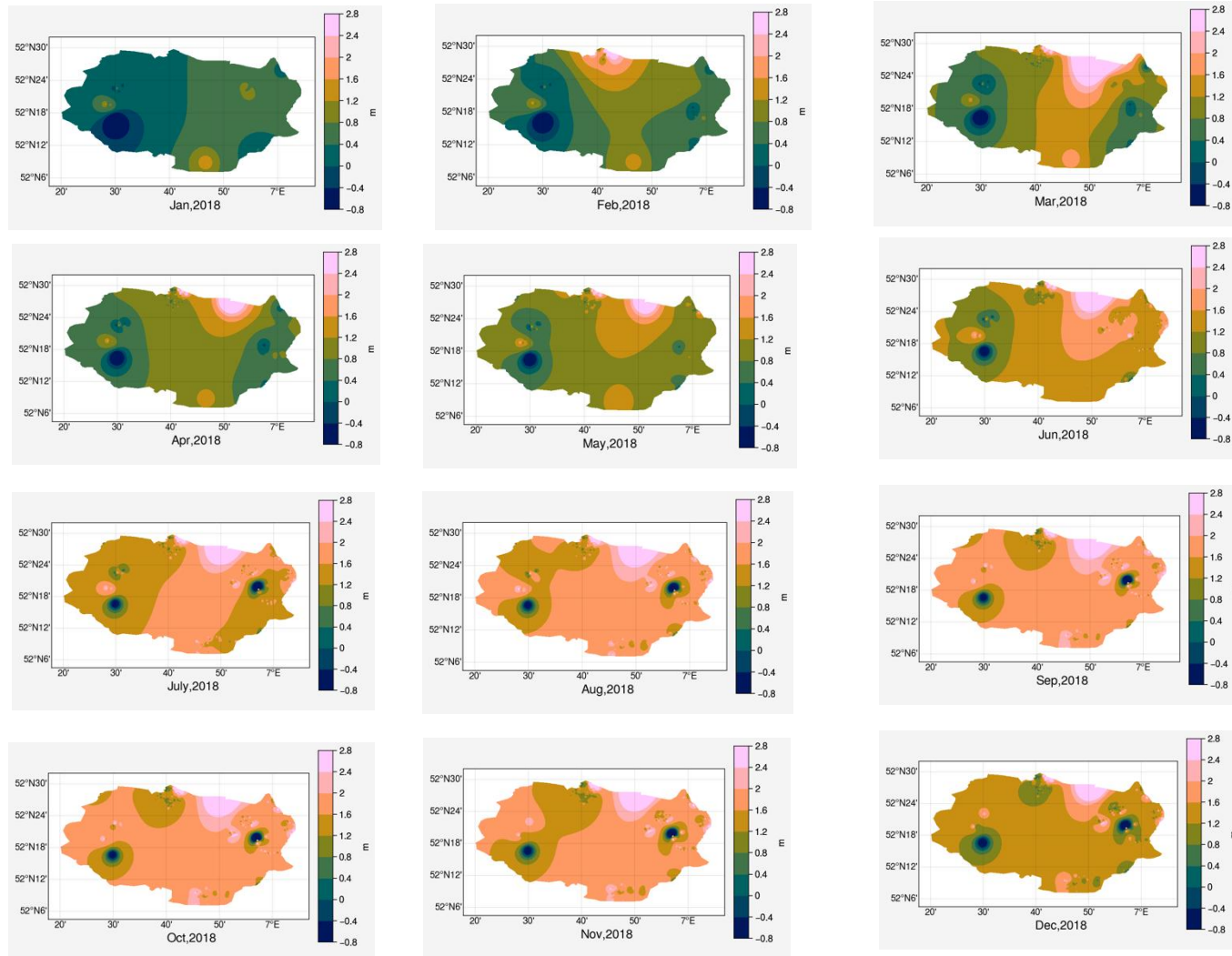


Figure 5.8 (a) GWD distribution in Twente, 2018

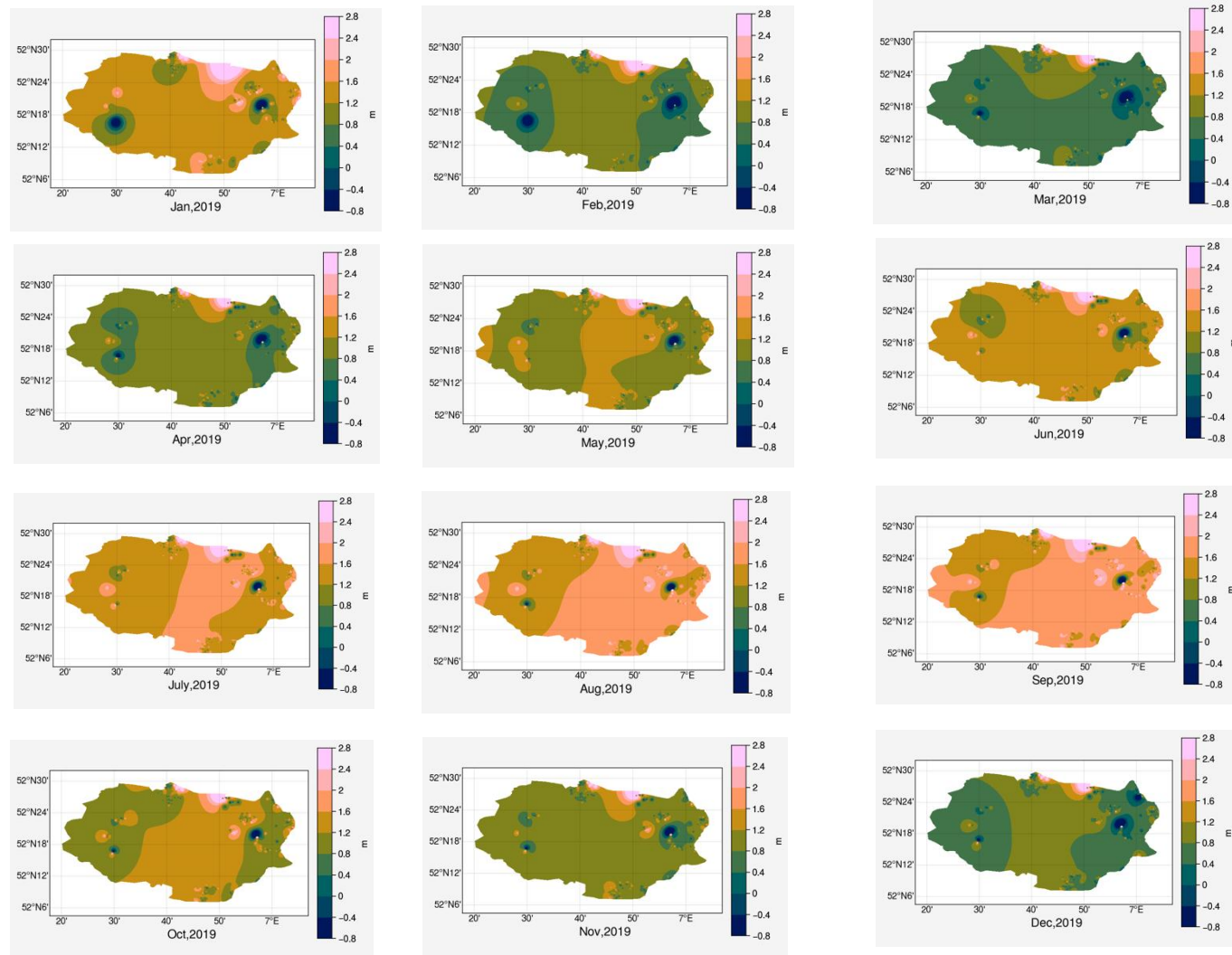


Figure 5.8 (b) GWD distribution in Twente, 2019



## 6. CONCLUSION AND RECOMMENDATION

In this study, we present a scalable workflow that could be applied to generate 10m SM not only at regional scale, but also at the Netherlands, Europe, and even at Global scale if the data and computational resources are available. Due to the limitation of data availability and computation resources, we generate the high-resolution (10m) SM product from 2018 to 2019 for Twente region only. The original 1km SM product is downscaled by RF with auxiliary variables (LST<sub>day</sub>, LST<sub>night</sub>, GRD, EVI, LAI, precipitation, soil texture, GWL). The RF model showed an acceptable result with  $r$  of 0.8, RMSE of  $0.0418 \text{ (cm}^3 \text{ cm}^{-3}\text{)}$ , MD of  $-0.0019 \text{ (cm}^3 \text{ cm}^{-3}\text{)}$  and ubRMSE of  $0.0418 \text{ (cm}^3 \text{ cm}^{-3}\text{)}$ . LST<sub>day</sub> and GWL are the most significant features followed by clay and precipitation while EVI and LAI show the least importance. The 10m SM product is able to capture the dynamic of 1km SM product and suggest that the 10 meter product performs acceptably.

The spatial and temporal patterns of the 10m SM product was compared with the original 1km SM product. The similar spatial pattern, western Twente is wetter than eastern Twente, can be observed in 1km and 10m SM product maps. Furthermore, 10m SM map with more abundant details that relatively wet areas can be recognized in the western parts of Twente at 10m resolution while 1km SM map illustrate west Twente is dry for these areas. The obviously seasonal variability can be observed in time-longitude diagram and time-latitude at 1km and 10m resolution that the soil is dry during the summer and autumn months, while it is moist during the spring and winter seasons.

We monitor the drought in Twente region from 2018 to 2019 by calculating SSMI with generated 10m SM product. From the temporal perspective, the entire area entered a state of drought from March-April in 2018 and 2019 and experienced the most extreme drought during late spring and early summer. Compared to 2019, drought in 2018 lasted longer and was more severe. From a spatial perspective, western Twente suffers a more severe drought than eastern Twente. Additionally, we analyze the driving factors behind the distribution of drought in Twente. Meteorological factors are the causes of the temporal distribution of drought and the spatial distribution of GWD contributes to the drought distribution.

Our 10m SM product demonstrates acceptable performance. However, future studies have opportunities for further improvement: 1) Retrieval SM from 10m GRD before resampling it into 10m daily resolution. GRD is sensitive to vegetation and underlying soils' dielectric and geometric characteristics. As such, the direct pre-processing of GRD data into 10m may introduce some errors; 2) The model's accuracy can be improved.  $R$  of 0.8 is relatively good but can still be improved by including more predictor variables related to SM.; 3) Higher spatiotemporal resolution of land surface temperature data should be used as predictors; 4) Validated the product with the independent in situ SM stations. Twente stations have been used as the training data for 1km SM product generation. Independent validation stations can better illustrate the 10m product performance.

## LIST OF REFERENCES

---

- Abbaszadeh, P., Moradkhani, H., & Zhan, X. (2019). Downscaling SMAP Radiometer Soil Moisture Over the CONUS Using an Ensemble Learning Method. *Water Resources Research*, 55(1), 324–344. <https://doi.org/10.1029/2018WR023354>
- Alexakis, D. D., Mexis, F. D. K., Vozinaki, A. E. K., Daliakopoulos, I. N., & Tsanis, I. K. (2017). Soil Moisture Content Estimation Based on Sentinel-1 and Auxiliary Earth Observation Products. A Hydrological Approach. *Sensors* 2017, Vol. 17, Page 1455, 17(6), 1455. <https://doi.org/10.3390/S17061455>
- Altman, N., & Krzywinski, M. (2017). Points of Significance: Ensemble methods: Bagging and random forests. *Nature Methods*, 14(10), 933–934. <https://doi.org/10.1038/nmeth.4438>
- André, J.-C., Goutorbe, J.-P., & Perrier, A. (1986). HAPEX—MOBILHY: A Hydrologic Atmospheric Experiment the Study of Water Budget and Evaporation Flux at the Climatic Scale. *Bulletin of the American Meteorological Society*, 67(2), 138–144. <https://doi.org/10.1175/1520-0477-67.2.138>
- Baez-Villanueva, O. M., Zambrano-Bigiarini, M., Beck, H. E., McNamara, I., Ribbe, L., Nauditt, A., Birkel, C., Verbist, K., Giraldo-Osorio, J. D., & Xuan Thinh, N. (2020). RF-MEP: A novel Random Forest method for merging gridded precipitation products and ground-based measurements. *Remote Sensing of Environment*, 239, 111606. <https://doi.org/10.1016/J.RSE.2019.111606>
- Benninga, H. J. F., Carranza, C. D. U., Pezij, M., Van Santen, P., Van Der Ploeg, M. J., Augustijn, D. C. M., & Van Der Velde, R. (2018). The Raam regional soil moisture monitoring network in the Netherlands. *Earth System Science Data*, 10(1), 61–79. <https://doi.org/10.5194/ESSD-10-61-2018>
- Boegh, E., Soegaard, H., Broge, N., Schelde, K., Thomsen, A., Hasager, C. B., & Jensen, N. O. (2002). Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. *Remote Sensing of Environment*, 81(2–3), 179–193. [https://doi.org/10.1016/S0034-4257\(01\)00342-X](https://doi.org/10.1016/S0034-4257(01)00342-X)
- Chen, X., & Hu, Q. (2004). Groundwater influences on soil moisture and surface evaporation. *Journal of Hydrology*, 297(1–4), 285–300. <https://doi.org/10.1016/j.jhydrol.2004.04.019>
- Cook, B. I., Mankin, J. S., & Anchukaitis, K. J. (2018a). Climate Change and Drought: From Past to Future. *Current Climate Change Reports*, 4(2), 164–179. <https://doi.org/10.1007/s40641-018-0093-2>
- Cook, B. I., Mankin, J. S., & Anchukaitis, K. J. (2018b). Climate Change and Drought: From Past to Future. *Current Climate Change Reports* 2018 4:2, 4(2), 164–179. <https://doi.org/10.1007/S40641-018-0093-2>
- Das, N. N., Entekhabi, D., & Njoku, E. G. (2011). An algorithm for merging SMAP radiometer and radar data for high-resolution soil-moisture retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 49(5), 1504–1512. <https://doi.org/10.1109/TGRS.2010.2089526>
- Dente, L., Su, Z., & Wen, J. (2012). Validation of SMOS Soil Moisture Products over the Maqu and Twente Regions. *Sensors* 2012, Vol. 12, Pages 9965-9986, 12(8), 9965–9986. <https://doi.org/10.3390/S120809965>
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA’s Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25–36. <https://doi.org/10.1016/J.RSE.2011.11.026>
- Hajj, M. El, Baghdadi, N., Zribi, M., & Bazzi, H. (2017). Synergic use of Sentinel-1 and Sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas. *Remote Sensing*, 9(12), 1–28. <https://doi.org/10.3390/rs9121292>
- Han, Q., Zeng, Y., Zhang, L., Wang, C., Prikaziuk, E., Niu, Z., & Su, B. (2023). Global long term daily 1 km surface soil moisture dataset with physics informed machine learning. *Scientific Data* 2023 10:1, 10(1), 1–12. <https://doi.org/10.1038/s41597-023-02011-7>
- He, B., Xing, M., & Bai, X. (2014). A synergistic methodology for soil moisture estimation in an alpine prairie using radar and optical satellite data. *Remote Sensing*, 6(11), 10966–10985. <https://doi.org/10.3390/rs61110966>
- Ho, T. K. (1995). Random decision forests. *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, 1*, 278–282. <https://doi.org/10.1109/ICDAR.1995.598994>
- Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), 832–844. <https://doi.org/10.1109/34.709601>

- Idso, S., Jackson, R., Reginato, R., Kimball, B., & Nakayama, F. (1975). The Dependence of Bare Soil Albedo on Soil Water Content. *J Appl Meteorol*, *14*, 109–113.
- Im, J., Park, S., Rhee, J., Baik, J., & Choi, M. (2016). Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. *Environmental Earth Sciences*, *75*(15). <https://doi.org/10.1007/S12665-016-5917-6>
- Jiang, Z., Huete, A. R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, *112*(10), 3833–3845. <https://doi.org/10.1016/j.rse.2008.06.006>
- Kian, K. T., Ibrahim, H., & Bejo, S. K. (2008). Investigation on several basic interpolation methods for the use in remote sensing application. *Proceedings of the 2008 IEEE Conference on Innovative Technologies in Intelligent Systems and Industrial Applications, CITISIA*, 60–65. <https://doi.org/10.1109/CITISIA.2008.4607336>
- Liaw, A., & Matthew, W. (2004). Classification and Regression by random Forest. *Journal of Dental Research*, *83*(5), 434–438. <https://doi.org/10.1177/154405910408300516>
- Liu, Y., Jing, W., Wang, Q., & Xia, X. (2020). Generating high-resolution daily soil moisture by using spatial downscaling techniques: a comparison of six machine learning algorithms. *Advances in Water Resources*, *141*, 103601. <https://doi.org/10.1016/J.ADVWATRES.2020.103601>
- Luo, L., Apps, D., Arcand, S., Xu, H., Pan, M., & Hoerling, M. (2017). Contribution of temperature and precipitation anomalies to the California drought during 2012–2015. *Geophysical Research Letters*, *44*(7), 3184–3192. <https://doi.org/10.1002/2016GL072027>
- Lv, A., Zhang, Z., & Zhu, H. (2021). A neural-network based spatial resolution downscaling method for soil moisture: Case study of qinghai province. *Remote Sensing*, *13*(8). <https://doi.org/10.3390/RS13081583>
- Ma, C., Li, X., & McCabe, M. F. (2020). Retrieval of high-resolution soil moisture through combination of Sentinel-1 and Sentinel-2 data. *Remote Sensing*, *12*(14), 1–28. <https://doi.org/10.3390/rs12142303>
- Mattar, C., Wigneron, J. P., Sobrino, J. A., Novello, N., Calvet, J. C., Albergel, C., Richaume, P., Mialon, A., Guyon, D., Jimenez-Munoz, J. C., & Kerr, Y. (2012). A combined optical-microwave method to retrieve soil moisture over vegetated areas. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5 PART 1), 1404–1413. <https://doi.org/10.1109/TGRS.2011.2179051>
- Notarnicola, C., Angiulli, M., & Posa, F. (2006). Use of Radar and Optical Remotely Sensed Data for Soil Moisture Retrieval Over Vegetated Areas. *Ieee Transactions on Geoscience and Remote Sensing*, *44*(4), 925–935.
- Philip, S. Y., Kew, S. F., Van Der Wiel, K., Wanders, N., Jan Van Oldenborgh, G., & Philip, S. Y. (2020a). Regional differentiation in climate change induced drought trends in the Netherlands. *Environmental Research Letters*, *15*(9). <https://doi.org/10.1088/1748-9326/ab97ca>
- Philip, S. Y., Kew, S. F., Van Der Wiel, K., Wanders, N., Jan Van Oldenborgh, G., & Philip, S. Y. (2020b). Regional differentiation in climate change induced drought trends in the Netherlands. *Environmental Research Letters*, *15*(9), 094081. <https://doi.org/10.1088/1748-9326/AB97CA>
- Potin, P. (2019). *Copernicus Sentinel-1 Constellation Mission Operations Status; Copernicus Sentinel-1 Constellation Mission Operations Status. 1*. <https://sentinels.copernicus.eu/web/sentinel/missions/sentin>
- Sachindra, D. A., & Kanae, S. (2019). Machine learning for downscaling: the use of parallel multiple populations in genetic programming. *Stochastic Environmental Research and Risk Assessment*, *33*(8–9), 1497–1533. <https://doi.org/10.1007/S00477-019-01721-Y/FIGURES/13>
- Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in Neural Information Processing Systems*, *4*, 2951–2959. <https://arxiv.org/abs/1206.2944v2>
- Su, Z., He, Y., Dong, X., & Wang, L. (2017). *Drought Monitoring and Assessment Using Remote Sensing*. 151–172. [https://doi.org/10.1007/978-3-319-43744-6\\_8](https://doi.org/10.1007/978-3-319-43744-6_8)
- Su, Zhongbo, Zeng, Y., Romano, N., Manfreda, S., Francés, F., Ben Dor, E., Szabó, B., Vico, G., Nasta, P., Zhuang, R., Francos, N., Mészáros, J., Dal Sasso, S. F., Bassiouni, M., Zhang, L., Rwasoka, D. T., Retsios, B., Yu, L., Blatchford, M. L., & Mannaerts, C. (2020). An Integrative Information Aqueduct to Close the Gaps between Satellite Observation of Water Cycle and Local Sustainable Management of Water Resources. *Water* *2020, Vol. 12, Page 1495, 12*(5), 1495. <https://doi.org/10.3390/W12051495>

- Wei, Z., Meng, Y., Zhang, W., Peng, J., & Meng, L. (2019). Downscaling SMAP soil moisture estimation with gradient boosting decision tree regression over the Tibetan Plateau. *Remote Sensing of Environment*, 225, 30–44. <https://doi.org/10.1016/J.RSE.2019.02.022>
- Wu, H., Yang, Q., Liu, J., & Wang, G. (2020). A spatiotemporal deep fusion model for merging satellite and gauge precipitation in China. *Journal of Hydrology*, 584, 124664. <https://doi.org/10.1016/J.JHYDROL.2020.124664>
- Xu, Y., Wang, L., Ross, K. W., Liu, C., & Berry, K. (2018). Standardized soil moisture index for drought monitoring based on soil moisture active passive observations and 36 years of North American Land Data Assimilation System data: A case study in the Southeast United States. *Remote Sensing*, 10(2). <https://doi.org/10.3390/rs10020301>
- Yan, R., & Bai, J. (2020). A New Approach for Soil Moisture Downscaling in the Presence of Seasonal Difference. *Remote Sensing 2020*, Vol. 12, Page 2818, 12(17), 2818. <https://doi.org/10.3390/RS12172818>
- Yang, Y., Zhang, J., Bao, Z., Ao, T., Wang, G., Wu, H., & Wang, J. (2021). Evaluation of multi-source soil moisture datasets over central and eastern agricultural area of China using in situ monitoring network. *Remote Sensing*, 13(6), 1–26. <https://doi.org/10.3390/rs13061175>
- Zeng, Y., Su, Z., Van Der Velde, R., Wang, L., Xu, K., Wang, X., & Wen, J. (2016). Blending satellite observed, model simulated, and in situ measured soil moisture over Tibetan Plateau. *Remote Sensing*, 8(3), 1–22. <https://doi.org/10.3390/rs8030268>
- Zhan, X., Houser, P. R., Walker, J. P., & Crow, W. T. (2006). A method for retrieving high-resolution surface soil moisture from hydros L-band radiometer and radar observations. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6), 1534–1544. <https://doi.org/10.1109/TGRS.2005.863319>
- Zhang, Lijie, Zeng, Y., Zhuang, R., Szabó, B., Manfreda, S., Han, Q., & Su, Z. (2021). In situ observation-constrained global surface soil moisture using random forest model. *Remote Sensing*, 13(23). <https://doi.org/10.3390/rs13234893>
- Zhang, Ling, Li, X., Zheng, D., Zhang, K., Ma, Q., Zhao, Y., & Ge, Y. (2021). Merging multiple satellite-based precipitation products and gauge observations using a novel double machine learning approach. *Journal of Hydrology*, 594, 125969. <https://doi.org/10.1016/J.JHYDROL.2021.125969>
- Zhang, R., Jia, M., Wang, Z., Zhou, Y., Mao, D., Ren, C., Zhao, C., & Liu, X. (2022). Tracking annual dynamics of mangrove forests in mangrove National Nature Reserves of China based on time series Sentinel-2 imagery during 2016–2020. *International Journal of Applied Earth Observation and Geoinformation*, 112(July), 102918. <https://doi.org/10.1016/j.jag.2022.102918>
- Zhao, Y., Weng, Z., Chen, H., & Yang, J. (2020). Analysis of the evolution of drought, flood, and drought-flood abrupt alternation events under climate change using the daily SWAP index. *Water (Switzerland)*, 12(7). <https://doi.org/10.3390/w12071969>
- Zhuang, R., Manfreda, S., Zeng, Y., Su, Z., Ben Dor, E., & Petropoulos, G. P. (2023). Soil moisture monitoring using unmanned aerial system. *Unmanned Aerial Systems for Monitoring Soil, Vegetation, and Riverine Environments*, 179–200. <https://doi.org/10.1016/B978-0-323-85283-8.00003-5>
- Zhuang, R., Zeng, Y., Manfreda, S., & Su, Z. (2020). Quantifying long-term land surface and root zone soil moisture over Tibetan plateau. *Remote Sensing*, 12(3), 1–20. <https://doi.org/10.3390/rs12030509>