

University of Mississippi

eGrove

---

Electronic Theses and Dissertations

Graduate School

---

1-1-2023

## Evaluating the Relationship Between Vegetation Types and Downscaled Surface Soil Moisture Data

Anupiya Vidarshana Ellepola

Follow this and additional works at: <https://egrove.olemiss.edu/etd>

---

### Recommended Citation

Ellepola, Anupiya Vidarshana, "Evaluating the Relationship Between Vegetation Types and Downscaled Surface Soil Moisture Data" (2023). *Electronic Theses and Dissertations*. 2497.  
<https://egrove.olemiss.edu/etd/2497>

This Thesis is brought to you for free and open access by the Graduate School at eGrove. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of eGrove. For more information, please contact [egrove@olemiss.edu](mailto:egrove@olemiss.edu).

EVALUATING THE RELATIONSHIP BETWEEN VEGETATION TYPES AND  
DOWNSCALED SURFACE SOIL MOISTURE DATA

A Thesis  
presented in partial fulfillment of requirements  
for the degree of Master of Engineering Science  
in the Department of Geology & Geological Engineering  
The University of Mississippi

by

ANUPIYA VIDARSHANA ELLEPOLA

May 2023

Copyright Anupiya Vidarshana Ellepola 2023  
ALL RIGHTS RESERVED

## **ABSTRACT**

Remote sensing systems such as multispectral and radar imaging can provide detailed information about soil moisture levels, vegetation cover, and topography. These data can be used to identify areas of high soil moisture, monitor changes in soil moisture over time, and assess the impact of human activities on watersheds. A commonly used system in orbit for monitoring soil moisture is the Soil Moisture Active Passive Mission (SMAP). For SMAP and all spaceborne systems, one of the major limitations for users to implement satellite-based data is the coarse resolution of the pixels (~9 km). Downscaling approaches are introduced by many researchers to overcome the low resolution of the surface soil moisture data. In this project, the random forest approach is used to downscale surface soil moisture derived from SMAP level 4 root zone soil moisture geophysical (SPL4SMGP) data product to a 1-km spatial resolution for a region in northeastern Mississippi. Normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), and diurnal temperature are used as independent variables for the random forest model. Due to the lack of fine-resolution surface soil moisture data in the study area, 9 km SPL4SMGP product is disaggregated to a 1km spatial resolution without altering original pixel values and used as the dependent variable for the model training. Field data were collected from 25 locations within the study region for six days throughout the year to validate the downscaling output. While results of the downscaling showed poor correlation to the field-collected surface soil moisture data, land cover types and surface geology showed a good match to the downscaled data. According to the results, Croplands and Cropland/Natural Vegetation Mosaics show the

highest surface soil moisture values for any day considered.

## **DEDICATION**

This thesis is dedicated to my parents and my sister, who always believed in me, and to those who believe in science.

## LIST OF ABBREVIATIONS AND SYMBOLS

|               |   |
|---------------|---|
| AMSR2         | Advance Microwave Scanning Radiometer 2                 |
| ASTM          | American Society for Testing and Materials              |
| ASCAT         | Advanced Scatterometer                                  |
| BIOME-BGC     | BIOME-Biogeochemical Cycles                             |
| CART          | Classification and Regression Trees                     |
| EF            | Evapotranspiration fraction                             |
| EOS           | Earth Observing Systems                                 |
| EASE-Grid     | A global cylindrical 9km Equal-area Scalable Earth Grid |
| EVI           | Enhanced Vegetation Index                               |
| European SMOS | European Soil Moisture Ocean Salinity mission           |
| $f$           | Waveform frequency (cycles per second [hertz])          |
| Fr            | Fractional vegetation                                   |
| GEOS-5        | Goddard Earth Observing System Model-version 5          |
| GHz           | Gigahertz (10 <sup>9</sup> hertz)                       |
| IGBP          | International Geosphere-Biosphere Programme             |
| KNN           | K-nearest neighbors                                     |
| LAI           | Leaf Area Index   |
| LST           | Land surface temperature                                |
| MIRAS         | Microwave Imaging Radiometer using Aperture Synthesis   |

|            |   |
|------------|---|
| MOD13A2    | "MODIS/Terra Vegetation Indices 16-Day L3 Global 1km SIN Grid V061" data product                  |
| MODIS      | Moderate Resolution Imaging Spectrometer  |
| MSTM       | Mississippi Transverse Mercator   |
| MYD11A1    | "MODIS/ Aqua Land Surface Temperature/ Emissivity Daily L3 Global 1km SIN Grid V061" data product |
| NDVI       | Normalized Difference Vegetation Index  |
| NSIDC DAAC | NASA National Snow and Ice Data Center Distributed Active Archive Center                          |
| PFT        | Plant Functional Types  |
| SMAP       | Soil Moisture Active Passive Mission  |
| SMOS       | Soil Moisture Ocean Salinity  |
| SPL4SMGP   | SMAP Level 4 Root Zone Soil Moisture Geophysical  |

## **ACKNOWLEDGEMENTS**

First and foremost, I would like to thank Dr. Lance Yarbrough for extending me the opportunity to join the Department of Geology and Geological Engineering at the University of Mississippi, serving as my academic supervisor and being a mentor throughout my Master's degree. I would also like to thank Dr. Greg Easson for giving me the privilege of joining the Mississippi Mineral Resources Institute (MMRI) as a research associate and being a part of my thesis committee. A special thanks also go to Dr. Zahra Ghaffari for joining my thesis committee and helping with my research whenever needed. Dr. Hakan Yasarer, Abdel Rahman, and the rest of the integrated groundwater management (IGM) research team from the University of Mississippi are also deeply appreciated for sharing their insights with me. It was the guidance and positive feedback I got from them that helped me shape this project and navigate myself throughout this project. I would also like to thank all my professors from the Department of Geology and Geological Engineering for educating me and encouraging me to go an extra mile.

A special thanks go to Kristian Macias, Mahesh Wickramasinghe, Madushanka Abeykoon, and Sapumal Witharana for accompanying me during my fieldwork by taking time from their very busy schedules. I could not have done this project without your help. I would also like to thank Mrs. Jaclyn Montoria for having all financial matters covered. I would also like to thank the National Science Foundation for the funds awarded to make this project a possibility. It is thanks to their financial support I was able to focus on my studies and bring this project to fruition.

It is the time I spent with my classmates and friends that really helped me get comfortable with graduate school. Hashindra Herath, Mahesh Wickramasinghe, Jodi Messick, Brightin Blanton, Kristian Macias, Savannah Rathbun, Alex Ward, Luke Campbell, Manuel Eli, Pam Akakpo, Said Al Kindi, Sara Alsarhan, Ilias Mahmud, Robert Andreoli, Pratap Bohara, Santosh Adhikari, and Sapumal Witharana, you guys rock. The list of people I want to thank is longer than I can write down here, I apologize to anyone that I may have left out. You are all deeply appreciated.

Last but not least, I want to thank the free education of Sri Lanka for opening the world beyond the oceans for me and generations of students.

Thanks for the great time, Ole Miss. Hotty Toddy!

## TABLE OF CONTENTS

|  |     |
|--|-----|
| ABSTRACT .....   | ii  |
| DEDICATION .....   | iv  |
| LIST OF ABBREVIATIONS AND SYMBOLS .....                        | v   |
| ACKNOWLEDGEMENTS .....   | vii |
| LIST OF FIGURES .....  | xi  |
| LIST OF TABLES .....   | xiv |
| I. INTRODUCTION .....  | 1   |
| 1.1 Soil moisture .....  | 1   |
| 1.2 Soil moisture measuring methods .....                      | 1   |
| 1.3 Remote sensing methods .....                               | 3   |
| 1.4 Background to the research problem and project goals ..... | 4   |
| 1.5 Introduction to research approach .....                    | 4   |
| II. LITERATURE REVIEW .....                                    | 6   |
| 2.1 Study area .....   | 6   |
| 2.2 Physiography .....   | 8   |
| 2.3 Soil moisture remote sensing methods .....                 | 10  |
| 2.4 Soil Moisture Active Passive Mission (SMAP) .....          | 12  |
| 2.5 Moderate Resolution Imaging Spectrometer (MODIS) .....     | 13  |
| 2.6 Downscaling .....  | 14  |
| III. METHODS AND DATA .....                                    | 19  |
| 3.1 Field data collection .....                                | 19  |
| 3.2 Selecting input data .....                                 | 23  |
| 3.3 Satellite derived surface soil moisture .....              | 25  |
| 3.4 Land surface temperature (LST) .....                       | 25  |
| 3.5 NDVI and EVI .....   | 26  |
| 3.6 Land cover types .....                                     | 26  |
| 3.7 Data preparation .....                                     | 29  |

|  |    |
|--|----|
| 3.8 Random Forest Model.....   | 30 |
| 3.9 Validating results.....  | 32 |
| IV. RESULTS .....  | 34 |
| 4.1 Soil sample testing and field collected data .....   | 34 |
| 4.2 Results from the downscaling .....   | 35 |
| V. DISCUSSION .....  | 46 |
| 5.1 Soil sample testing .....  | 46 |
| 5.2 Visual comparison of pixel patterns in coarse resolution and downscaled images .....                                       | 47 |
| 5.3 Quantitative analysis of the downscaling model.....  | 49 |
| 5.4 Relationship between the land cover types and surface soil moisture .....  | 52 |
| VI. CONCLUSIONS .....  | 54 |
| LIST OF REFERENCES .....   | 56 |
| APPENDICES .....   | 69 |
| APPENDIX A: STATISTICS OF SURFACE SOIL MOISTURE VALUES FOR<br>DIFFERENT LAND COVER TYPES IN THE STUDY REGION ACCORDING TO IGBP | 70 |
| APPENDIX B: STATISTICS OF SURFACE SOIL MOISTURE VALUES FOR<br>DIFFERENT SURFACE GEOLOGIC UNITS IN THE STUDY REGION.....        | 76 |
| VITA .....   | 82 |

## LIST OF FIGURES

|  |    |
|--|----|
| <b>Figure 1.</b> Sensor types and spatial resolution of different soil moisture measuring techniques (Vereecken et al., 2008) .....  | 2  |
| <b>Figure 2.</b> Principal aquifer outcrop map of Mississippi (from Wasson, 1986).....   | 7  |
| <b>Figure 3.</b> The map showing the ecoregions of Mississippi with tributary networks superimposed (from Chapman et al., 2004). .....   | 9  |
| <b>Figure 4.</b> Field locations used for surface soil moisture data collection Field locations used for surface soil moisture data collection. ....   | 21 |
| <b>Figure 5.</b> Surface soil moisture collection locations within the 1 km <sup>2</sup> study site.....   | 22 |
| <b>Figure 6.</b> Soil moisture data collecting using the field soil moisture sensor. (LEFT) in closeup of the instrument interface along with a handheld GPS unit, and (RIGHT) showing position of probe in prepared test excavation. .... | 23 |
| <b>Figure 7.</b> Parameters affecting the spatial variability of surface soil moisture as a function of the spatial extent of the study region considered (from (Crow <i>et al.</i> , 2012) .....  | 24 |
| <b>Figure 8.</b> Figure representation of disaggregating a 9 km cell to 1 km cells using the nearest neighbor method .....   | 30 |
| <b>Figure 9.</b> Flow chart summarizing workflow followed for the downscaling of the coarse resolution surface soil moisture. ....   | 33 |
| <b>Figure 10.</b> Scatter plot showing the comparison of soil moisture data from the field and laboratory tested samples. ....   | 34 |

|   |    |
|---|----|
| <b>Figure 11.</b> Field surface soil moisture measurements in an area of approximately 1 km <sup>2</sup> and statistics for the measurements. Soil moisture values are reported as volume/volume. ....  | 35 |
| <b>Figure 12.</b> 9 km spatial resolution surface soil moisture data from SPL4SMGP product, clipped to the study area for the months of (A) February, (B) March, (C) May, and (D) September. Soil moisture values shown are volume/volume. .... | 37 |
| <b>Figure 13.</b> 9 km spatial resolution surface soil moisture data from SPL4SMGP product, clipped to the study area for the month of October. The soil moisture values shown are volume/volume. ....  | 38 |
| <b>Figure 14.</b> Surface soil moisture downscaling results for the months of (A) February, (B) March, (C) May, and (D) September. Spatial resolution is 1km, and soil moisture values shown are volume/volume. ....                            | 39 |
| <b>Figure 15.</b> Surface soil moisture downscaling results for the month of October. Spatial resolution is 1 km, and soil moisture values shown are volume/volume. ....  | 40 |
| <b>Figure 16.</b> Remotely sensed soil moisture vs. field soil moisture readings for the month of February. ....  | 41 |
| <b>Figure 17.</b> Remotely sensed soil moisture vs. field soil moisture readings for the month of March. ....   | 41 |
| <b>Figure 18.</b> Remotely sensed soil moisture vs. field soil moisture readings for the month of May. ....   | 42 |
| <b>Figure 19.</b> Remotely sensed soil moisture vs. field soil moisture readings for the month of September. ....   | 42 |
| <b>Figure 20.</b> Remotely sensed soil moisture vs. field soil moisture readings for the images from the month of October. ....   | 43 |

|   |    |
|---|----|
| <b>Figure 21.</b> Land cover classification data downloaded from MODIS for the period 2021–2022 visualized over the study area as per IGBP classification. .... | 45 |
| <b>Figure 22.</b> Surface geology of the study region. ....   | 48 |
| <b>Figure B 1.</b> Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in February.....        | 77 |
| <b>Figure B 2.</b> Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in February.....        | 78 |
| <b>Figure B 3.</b> Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in May.....             | 79 |
| <b>Figure B 4.</b> Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in September. ....      | 80 |
| <b>Figure B 5.</b> Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in October. ....        | 81 |

## LIST OF TABLES

|   |    |
|---|----|
| <b>Table 1.</b> Satellites instruments observing surface soil moisture observations .....   | 11 |
| <b>Table 2.</b> Study dates considered for downloading satellite data. ....   | 24 |
| <b>Table 3.</b> International Geosphere-Biosphere Programme (IGBP) classification (Sulla-Menashe & Friedl, 2018).....   | 28 |
| <b>Table 4.</b> Summary of correlation coefficient values for the comparison of field collected surface soil moisture data and remotely sensed surface soil moisture data (values from <b>Figure 16-Figure 19</b> ). .... | 43 |
| <b>Table 5.</b> Parameter importance for the downscaling models constructed for each month.....   | 43 |
| <b>Table 6.</b> Statistics of the downscaled and coarse resolution surface soil moisture data. ....   | 44 |
| <b>Table A 1.</b> Statistics for different land cover types in the study region according to IGBP for the study date in February. Soil moisture values are given as volume/volume.....                                    | 71 |
| <b>Table A 2.</b> Statistics for different land cover types in the study region according to IGBP for the study date in March. Soil moisture values are given as volume/volume.....                                       | 72 |
| <b>Table A 3.</b> Statistics for different land cover types in the study region according to IGBP for the study date in May. Soil moisture values are given as volume/volume. ....  | 73 |
| <b>Table A 4.</b> Statistics for different land cover types in the study region according to IGBP for the study date in September. Soil moisture values are given as volume/volume. ....                                  | 74 |
| <b>Table A 5.</b> Statistics for different land cover types in the study region according to IGBP for the study date in October. Soil moisture values are given as volume/volume. ....                                    | 75 |

**Table B 1.** Statistics for different surface geologic units in the study region for the study date in February. Soil moisture values are given as volume/volume. .... 77

**Table B 2.** Statistics for different surface geologic units in the study region for the study date in March. Soil moisture values are given as volume/volume. .... 78

**Table B 3.** Statistics for different surface geologic units in the study region for the study date in May. Soil moisture values are given as volume/volume. .... 79

**Table B 4.** Statistics for different surface geologic units in the study region for the study date in September. Soil moisture values are given as volume/volume..... 80

**Table B 5.** Statistics for different surface geologic units in the study region for the study date in October. Soil moisture values are given as volume/volume. .... 81

## **I. INTRODUCTION**

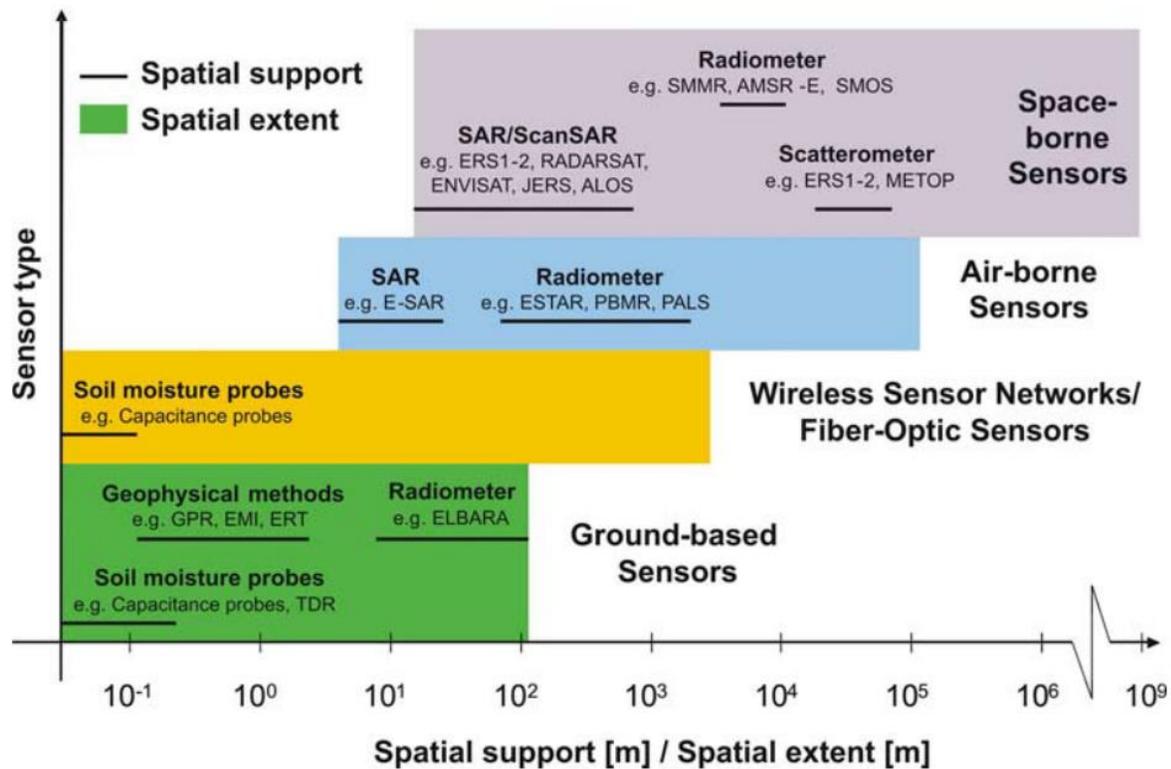
### 1.1 Soil moisture

Soil moisture controls many soil properties and sustains life within soil and is a main source of land evapotranspiration (Seneviratne et al., 2010), a very important portion of the water cycle. Soil moisture is also a key variable controlling the productivity of plants and partitioning rainfall into runoff and infiltration (Daly & Porporato, 2005). According to (Kerr, 2007), soil moisture is the total amount of water in the unsaturated zone. Soil moisture is divided into surface soil moisture (first 5 cm in general) and root zone soil moisture.

### 1.2 Soil moisture measuring methods

There are many examples in the literature introducing soil moisture measuring methods. These soil moisture measuring methods are broadly divided into two main categories; contact-based and contact-free methods (Robinson et al., 2008; Vereecken et al., 2008). Contact-based methods include capacitance sensors, electrical resistivity methods, heat pulse sensors, fiber optic sensors, time domain reflectometry, and destructive sampling methods such as gravimetric methods. Contact-free methods include microwave remote sensing, synthetic aperture radars, scatterometers, and thermal methods (Dorigo et al., 2011; Vereecken et al., 2008). The spatial and temporal resolution of the data obtained by these methods vary based on the technique used (Figure 1). With contact-based methods, it is possible to obtain spatially finer resolution soil moisture data and control the temporal resolution of the data sets based on the need. However, with contact-based methods, it is difficult to obtain a continuous stream of data with a high

spatial resolution for a long period due to the amount of time and human hours needed. Another obstacle in obtaining spatially distributed soil moisture data is understanding the heterogeneity of the soil moisture.



**Figure 1.** Sensor types and spatial resolution of different soil moisture measuring techniques (Vereecken et al., 2008)

Soil moisture is influenced by many soil properties and climate interactions leading to heterogeneity in vertical and horizontal soil profiles (Seneviratne et al., 2010). With the heterogeneity of soil moisture, it becomes more difficult to prepare soil moisture maps using contact-based methods that provide point data. These difficulties associated with contact-based measuring methods bring forward the importance of soil moisture measuring using contact-free methods.

### 1.3 Remote sensing methods

As shown in **Figure 1**, air-borne sensors and space-borne sensors have a wider spatial distribution of the soil moisture measurements. Added to that, it is possible to use these instruments for a longer period giving us the ability to continuously observe the soil moisture. These remote sensing methods are based on the backscatter of electromagnetic waves from the soil. According to Schmugge, (1978) microwave brightness temperature and radar backscatter of the soil have a correlation of approximately 0.9 with the surface soil moisture. The emissivity for soil varies in a range of approximately 0.6 to greater than 0.9. This variation in emissivity corresponds to a soil brightness temperature variation, which covers a range of soil wetness from 40% to 5% moisture by volume. Furthermore, atmospheric effects are small and may be neglected in most cases for lower microwave frequencies (1–3 GHz) (Njoku & Entekhabi, 1996). This makes microwave remote sensing a favorable tool for soil moisture observation.

There are several satellites orbiting Earth that are capable of providing geophysical data for measuring soil moisture. Some of these satellite missions provide soil moisture data as direct products. Soil Moisture Active Passive Mission (SMAP) and Soil Moisture Ocean Salinity (SMOS) are two such satellite missions. There are also open loop models (i.e., without data assimilation) and models with satellite data assimilation that provide soil moisture data (Beck et al., 2021).

## 1.4 Background to the research problem and project goals

One of the major limiting factors for satellite soil moisture data is their coarse spatial resolution. The currently available satellite products have a spatial resolution ranging between 10 km to 50 km (Beck et al., 2021). To overcome this issue, different methods have been utilized to spatially downscale satellite soil moisture data. Spatial downscaling is the procedure followed to bring a spatially coarse resolution image to a finer resolution. According to Peng et al., (2017), downscaling methods are divided into (1) satellite-based methods, (2) methods using geoinformation data, and (3) model-based methods.

The goal of this project was to create a surface soil moisture data set with a spatial resolution of 1km using downscaling and evaluate the relationship of the downscaled data with the prevailing land cover types and surface geology of the terrain. We use a machine learning approach to downscale surface soil moisture data by fusing different satellite products for a study region in northeastern Mississippi, where there are no high-resolution surface soil moisture data for the study period selected. A secondary goal was to evaluate the applicability of a simple machine learning technique to downscale coarse resolution surface soil moisture data despite not having fine-scale target training data and explore how to evaluate the output of the same.

## 1.5 Introduction to research approach

In this current project, random forest model is used to achieve spatial downscaling of SMAP level 4 root zone soil moisture geophysical (SPL4SMGP) data. A study area in northeast Mississippi was selected for testing the models and collecting field data. The accuracy of downscaled results is measured using field soil moisture data collected using a field soil moisture sensor equipment. Additionally, statistics of the pixel values are analyzed and compared for both

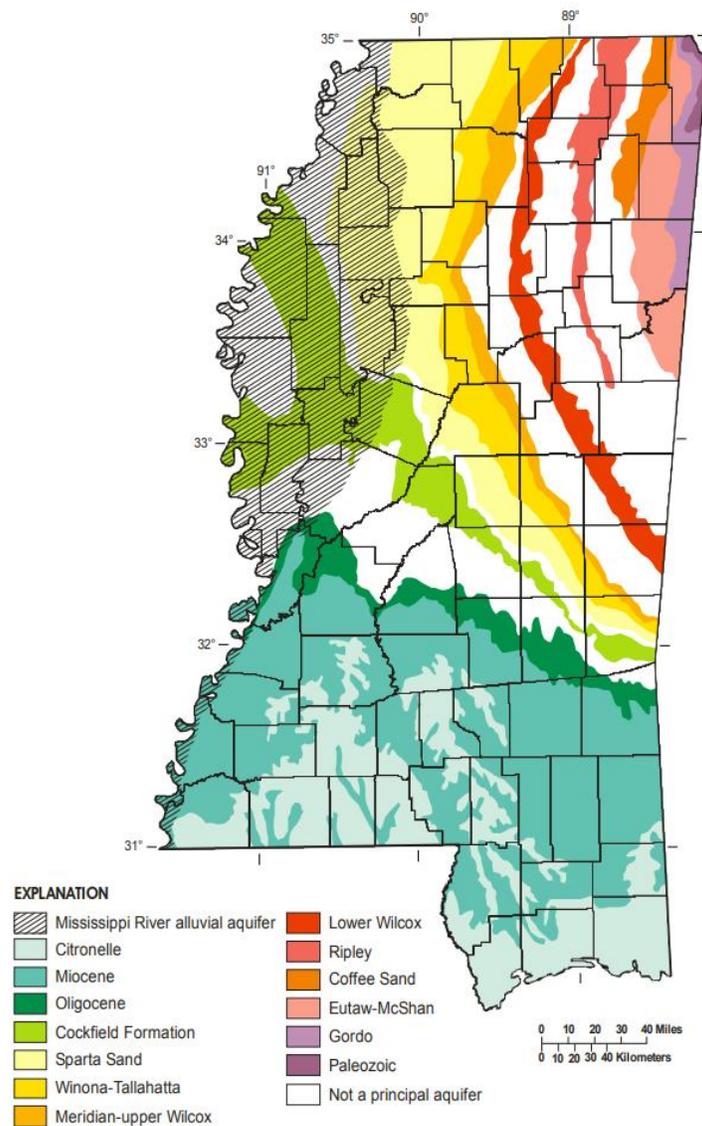
coarse-resolution and fine-resolution images to give a quantitative assessment of the downscaling results. Land cover types and surface geology of the terrain are also used to assess the patterns in surface soil moisture across coarse-resolution and downscaled results.

## **II. LITERATURE REVIEW**

### 2.1 Study area

The study area for this research is a portion of the northeastern Mississippi. The Gulf of Mexico and the relatively flat plains of central North America cover the Southern and Northern boundaries of the state of Mississippi. Thus, Mississippi experiences warm moist air from the Gulf of Mexico and drier continental air masses flowing from the north, which are cold in the winter and warm during the summer. The climate in Mississippi is a combination of relatively mild winters, hot summers, and year round precipitation. Mississippi is also recognized as one of the few areas globally to experience little net warming. Since the beginning of the 20<sup>th</sup> century, Mississippi has experienced a temperature increase by a miniscule 0.05 °C (0.1 °F). The warmest consecutive years recorded so far for Mississippi are between 2016 and 2020. The statewide annual average (1991–2020) precipitation for Mississippi is 1,420 mm (55.9 inches), with an average of 1270 mm (50 inches) for northern Mississippi and 1651 mm (65 inches) for southern Mississippi. According to the records, the wettest period experienced by Mississippi is 2016-2020, with an annual precipitation of 63.1 inches (Runkle et al., 2022). Rainfall is mainly received from January to April and November to December. During the summer, agricultural droughts (inadequate soil moisture levels to meet crop water demands) can be observed in Mississippi (Tang et al., 2018).

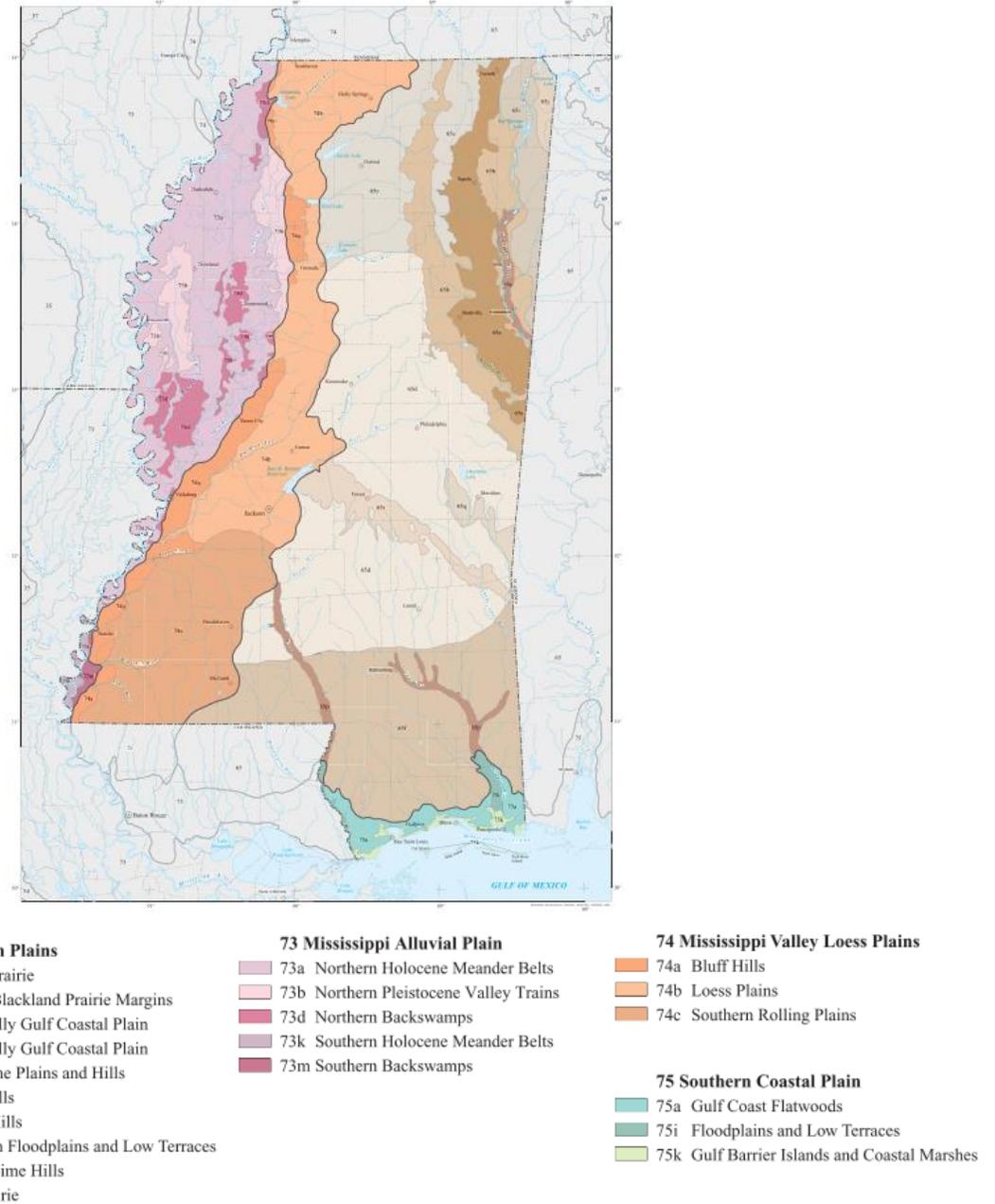
The Mississippi River creates the western boundary for the state of Mississippi, around which the Mississippi River Valley alluvial aquifer is formed. The Mississippi River Valley aquifer is the highest yielding aquifer in Mississippi and provides water for agricultural purposes via wells (Dalsin, 1978). Paleozoic, Gordo, Eutaw-McShan, Coffee Sand, Ripley, and Wilcox are the principal aquifers that are observed to outcrop in Northeastern Mississippi (Wasson, 1986).



**Figure 2.** Principal aquifer outcrop map of Mississippi (from Wasson, 1986).

## 2.2 Physiography

Mississippi is within the Gulf Coastal Plain physiographic region and is characterized by very gradual elevation changes. The summit of Woodall Mountain in Tishomingo County is the highest elevation point in Mississippi, at 245.67 m (806 ft) above mean sea level. The unconsolidated sediments that underlie the surface in Mississippi range in age from recent to Late Cretaceous. Out of the eleven physiographic provinces in the state of Mississippi, six are observed within the northeastern Mississippi: (1) Paleozoic Bottoms; (2) Tombigbee and Tennessee River Hills; (3) Black Prairie; (4) Pontotoc Ridge; (5) Flatwoods; (6) North Central Hills. The six provinces have characteristic vegetation patterns and underlying soil types. Unconsolidated Cretaceous sands, dark soils with high organic matter and smectite, sands of the McNairy Sand, Porters Creek Clay, and sandy units of the Eocene Claiborne Formation are some of the geologic units found in the six physiographic provinces of the northeastern Mississippi. At the level III resolution for the ecosystems, four ecoregions are recognized in Mississippi, with the “Southeastern Plains” ecoregion of the four covering most of the northeastern Mississippi. Southeastern Plains is further subdivided into ten different ecoregions according to level IV resolution for the ecosystems, where ecoregions represent the underlying physiographic units too. Overall the ecoregions in the northeastern Mississippi consist of oak, hickory, pine, pasture lands, mixed forests, prairies, and croplands (Chapman et al., 2004; Dockery & Thompson, 2016).



**Figure 3.** The map showing the ecoregions of Mississippi with tributary networks superimposed (from Chapman et al., 2004).

### 2.3 Soil moisture remote sensing methods

Remote sensing of soil moisture is a widely researched subject. Literature shows that the unique thermal and dielectric properties of soil moisture can be used to remotely sense soil moisture in the surface layer of the soil (Cihlar & Ulaby, n.d.; Curtis et al., 1995; Idso et al., 1975; Patel et al., 2018; Quan et al., 2014; Reginato et al., 1976; Schmugge, 1982; J. R. Wang, 1980). Thermal conductivity and the large heat capacity of the soil moisture allow soils to have a large thermal inertia. Thermal inertia is the square root of the product of thermal conductivity and heat capacity. The diurnal range of soil surface temperature is considered a function of both thermal inertia and external factors such as solar radiation, air temperature, relative humidity, wind, etc. An increase in soil moisture leads to an increase in thermal inertia and ultimately results in a decrease in the diurnal range of surface temperature. This relationship between the soil moisture and the diurnal range of surface temperature creates a proxy to measure surface soil moisture (Idso et al., 1975; Reginato et al., 1976). The experiments that have been conducted show that the relationship between soil moisture content and temperature changes are dependent on the soil type. However, the relationship between the pressure potential and the temperature changes is independent of the soil type (Idso et al., 1975). Thus, moisture values are expressed as a percentage of field capacity. While the method shows good results for the first few centimeters (0-4 cm) of the soil, this method is not applicable where there is a vegetation canopy (Idso et al., 1975; Reginato et al., 1976; Schmugge, 1982).

Microwave remote sensing depends on the large contrasting dielectric properties of the soil moisture for measuring soil moisture content. According to Rayleigh-Jeans approximation, thermal emission measured by microwave radiometers is proportional to the product of the temperature and emissivity of the surface. This product is referred to as the brightness

temperature (Schmugge, 1982; Schmugge et al., 1986). The relationship between the brightness temperature and soil moisture is then used to measure the soil moisture based on the brightness temperature readings from microwave sensors (Njoku & Kong, 1977). Most microwave soil moisture sensors operate in low-frequency bands ( $f < 15$  GHz) due to their lower sensitivity to solar effects and weak sensitivity to atmospheric effects. Advanced Microwave Scanning Radiometer-E (AMSR-E), Advanced Microwave Scanning Radiometer 2 (AMSR2), Advance Scatterometer (ASCAT), SMAP, and Microwave Imaging Radiometer using Aperture Synthesis (MIRAS) are several satellite instruments making surface soil moisture observations.

**Table 1.** Satellites instruments observing surface soil moisture observations

| Acronym | Operating band     | Spatial resolution | Temporal resolution | Temporal coverage | Reference(s)             |
|---------|--------------------|--------------------|---------------------|-------------------|--------------------------|
| AMSR2   | X-band (10.7 GHz)  | ~ 47 km            | 1–3 d               | 2012–present      | (Parinussa et al., 2015) |
| AMSR-E  | C-band (6.9 GHz)   | ~ 25 km            | 1–3 d               | 2002–2011         | (Minnett, 2019)          |
| ASCAT   | C-band (5.255 GHz) | ~ 30 km            | 1–2 d               | 2007–present      | (Wagner et al., 2013)    |
| SMAP    | L–band (1.4GHz)    | ~ 35 km            | 1–3 d               | 2015–present      | (Entekhabi et al., 2010) |
| MIRAS   | L–band (1.4GHz)    | ~ 40 km            | 1–3 d               | 2010–present      | (Kerr et al., 2012)      |

It is natural to come across strong vertical gradients of soil moisture and temperature in the soil. Because of this gradient, a mean value of the soil moisture is considered over a fixed soil thickness known as the sampling depth. Sampling depth is dependent on the frequency of the microwaves used (3 cm at 1.4 GHz, 1 cm at 5 GHz, 0.5 cm if  $f > 5$  GHz). Literature also shows that passive microwave sensors are limited to coarse spatial resolutions of greater than 10 km

and more sensitive to soil moisture while being less affected by surface geometry, whereas the active sensors are capable of reaching meter scale resolution while being very sensitive to surface geometry (Wigneron et al., 1998).

#### 2.4 Soil Moisture Active Passive Mission (SMAP)

Soil Moisture Active Passive Mission (SMAP) was launched in 2015 with the scientific goal of making direct observations of soil moisture and the freeze/thaw state of land surfaces. Weather and climate forecasting, drought monitoring, flood forecasting, identifying agricultural productivity, and national security are several key areas where SMAP data is applicable. SMAP was designed with a 6m diameter, conically scanning, deployable mesh reflector antenna. The antenna is shared by an L-band (1.41GHz) radiometer and an L-band (1.26GHz) radar. The design of SMAP microwave instruments, measurements, and algorithms was heavily influenced by previous L-band missions and experiments. NASA SkyLab missions from the 1970s, European Soil Moisture Ocean Salinity (SMOS) mission, and Aquarius/SAC-D mission, are several such works that contributed to the development of the SMAP design (Entekhabi et al., 2010).

The benefit of an L-band radar along with an L-band radiometer is to integrate the two instruments as a single observation system to produce enhanced soil moisture products. However, due to a technical failure that occurred on July 7, 2015, only the radiometric data is available (Das et al., 2018). The native L-band radiometer (1.41 GHz) produces data at a 40 km resolution (Entekhabi et al., 2010). By using a lower frequency (L-band), the instruments are capable of penetrating up to 5 cm of topsoil, and the measurements are sensitive to soil moisture through the vegetation of up to  $5 \text{ kg}\cdot\text{m}^{-2}$  water content. Whereas the instruments with higher

frequencies only correspond to a soil depth of about 1 cm and are sensitive to soil moisture through the vegetation of up to  $3 \text{ kg} \cdot \text{m}^{-2}$  water content (e.g. C-band; AMSR-E, and WindSat) (Entekhabi et al., 2010).

SMAP baseline data products are produced at four different levels. At each increasing level, data from previous levels are used to process and produce a different product. In this study the Level 4 root zone soil moisture (L4\_SM) data product is used. SMAP L4\_SM is produced by assimilating SMAP radiometer readings with surface meteorological and precipitation observations (Reichle et al., 2019). The Goddard Earth Observing System Model-version 5 (GEOS-5) land data assimilation system is the algorithm used in producing a 9 km gridded SMAP L4\_SM data product (De Lannoy & Reichle, 2016). SMAP science data is available to download from NASA National Snow and Ice Data Center Distributed Active Archive Center (NSIDC DAAC) (O'Neill et al., 2017).

## 2.5 Moderate Resolution Imaging Spectrometer (MODIS)

Moderate Resolution Imaging Spectrometer (MODIS) is a multi-band earth observing sensor designed as an Earth observing systems (EOS) satellite. The first satellite to carry a MODIS instrument into space was Terra (previously named EOS AM-1). Terra was launched in December 1999 along with five instruments onboard, including MODIS. Terra moves from north to south on the daylight side of Earth in a circular sun-synchronous polar orbit. The second MODIS instrument is carried aboard Aqua (previously named EOS PM) which was launched in May 2002. Aqua moves from south to north on the daylight side of the earth in a circular sun-synchronous orbit. Out of the six instruments originally launched aboard Aqua, only MODIS and three other instruments are in operational status as of 2023. Terra and Aqua, both having far

exceeded their original life span, have drifted from their initial orbits. However, the quality of the data from MODIS instruments remains intact and is good for science purposes. Native data from MODIS instruments are available in 250, 500, and 1000 m spatial resolution. MODIS land products have temporal resolutions of daily, 4-day, 8-day, 16-day, monthly, quarterly, and yearly. In this current project, we use MODIS land surface and daily reflectance version 6.1 products by Aqua satellite.

## 2.6 Downscaling

Spatial downscaling is the approach by which a coarse spatial resolution data set is converted to a finer resolution. There is a large body of literature involving spatial downscaling of soil moisture data. The methods employed to downscale soil moisture data are broadly divided into three major categories: (1) methods using geo-information data, (2) model-based methods, and (3) satellite based methods (Peng et al., 2017). Methods using geo-information data rely on establishing a relationship between coarse-scale soil moisture data, geological attributes, and fine-scale soil moisture values. However, geo-information based methods require an extensive amount of in-situ data to build a downscaling relationship and lead to catchment specific results. Therefore, the applicability of geo-information based methods is limited to the spatial scales similar to catchments, while having the potential to further improve (Busch et al., 2012; Peng et al., 2017; Perry & Niemann, 2007).

There are several model-based methods that were developed to downscale soil moisture data. Based on the complexity and approach used, these methods could be further divided into statistical models and land surface models. Statistical models make use of the insights brought from research done on spatial statistics of soil moisture and how statistics vary across scales

(Peng et al., 2017). Kaheil et al., (2008), describe such a statistical method of downscaling soil moisture, where the geostatistics from the spatial distribution of a coarse scale image is used to model soil moisture at a finer resolution. In situ soil moisture observations are used to improve the output of this model further. In land surface models, coarse-scale observations are used to obtain fine-scale soil moisture from a hydrological or land surface model.

Satellite based methods are a means of achieving finer scale soil moisture data by fusing coarse spatial resolution passive microwave (radiometer) soil moisture data with fine resolution satellite data. Njoku et al., (2002), introduced the ability to combine radiometer soil moisture readings with radar soil moisture data through a change detection method to take advantage of the higher spatial resolution of radar observations. It is shown that while the absolute soil moisture observations from the radar are heavily affected by vegetation and surface geometry, the relative changes in soil moisture show similar patterns in both radar and radiometer observations. In addition, the authors show that the effects of vegetation and surface geometry are time-invariant. The relationship formed between radiometer and radar data based on this similarity in soil moisture change detection is introduced as a means of combining radar and radiometer data for improved spatial resolution. The change detection method was further tested and improved by (Narayan et al., 2006) to produce radar/radiometer soil moisture relative change data. Das et al., (2011), have further refined the previous work on change detection method to produce the baseline algorithm for SMAP to combine radar and radiometer data. According to this algorithm, uniform vegetation and surface characters are assumed for each coarse-resolution pixel. This assumption adds to the error in the final result. Authors of this algorithm also show that it is possible to use high-resolution vegetation index data to further improve this algorithm and account for the heterogeneity of the vegetation.

Montzka et al., (2016), presents results of three different approaches to combining radar and radiometer soil moisture data: (1) Using radiometer data to estimate soil moisture and subsequent disaggregation using radar backscatter data; (2) disaggregation of the radiometer brightness temperature data using radar backscatter and subsequent soil moisture estimation; (3) disaggregation of radiometer soil moisture using radar soil moisture residuals. A comparison between the three methods shows that the second method yields the best results in combining radar and radiometer data for higher spatial resolution. It was also shown that auxiliary data such as vegetation and soil characteristics could be used in the second method to improve accuracy. The first and the third method also show reasonable results while being less accurate compared to the second method.

While much research has been carried out in support of radar-radiometer fusion techniques, the biggest challenge is the inconsistent observation times between these sensors. SMAP was designed to overcome this challenge by having both an active radar and a passive radiometer. As the SMAP active sensor is now not operational, the challenge of overcoming inconsistent observation times for radar-radiometer fusion remains. This brings out the importance of downscaling methods that do not involve direct radar-radiometer data fusion.

Many methods introduced for downscaling satellite derived soil moisture data involve the use of land surface parameters that have a correlation and causation to soil moisture. The use of such parameters derived by optical and thermal remote sensing is very popular due to their high spatial resolution and wide availability. Although, there still remains the disadvantage of interference from cloud cover when it comes to optical and thermal remote sensing (Peng et al., 2017). The triangle method for estimating soil moisture is one such method that uses land surface parameters. This method is based on the relationship between the normalized difference

vegetation index (NDVI) and surface radiant temperature. The triangular shape made by the scatter plot of NDVI and surface radiant temperature of an area can be used to calculate fractional vegetation ( $Fr$ ) and the scaled surface radiant temperature ( $T^*$ ) (Chauhan et al., 2003). A polynomial relationship can be then written to find the surface soil moisture ( $Mo$ ) and evapotranspiration fraction ( $EF$ ), as shown in Equation 1 (T. Carlson, 2007; T. N. Carlson et al., 1994).

$$(Mo, EF) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} T^{*i} Fr^j \quad (1)$$

Where  $a_{ij}$  represents the regression coefficient, which can be determined using coarse resolution of the original data. Once the regression coefficient is determined, fine-scale soil moisture data can be generated using the polynomial function and fine-scale ancillary data. This approach has been adapted in other research and has modified the polynomial relationship in some instances (Piles et al., 2011).

Merlin et al., (2008), introduced a deterministic approach to downscale soil moisture data. This approach has three general steps: (1) estimating soil evaporative efficiency (2) link soil evaporative efficiency to near surface soil moisture; (3) building a downscaling relationship to link high resolution soil moisture with coarse-resolution soil moisture data and soil evaporative efficiency. Here the soil evaporative efficiency is calculated using vegetation fraction and feature space for land surface temperature (LST).

While traditional methods have shown promising results in many studies, some authors suggest that the use of machine learning techniques to capture the non-linear nature of predictors such as LST, soil moisture, and NDVI is beneficial (Bartkowiak et al., 2019; Hutengs & Vohland, 2016). Random Forests, boosted regression trees, deep learning, and support vector

regression are a few such methods. Im et al., (2016), present a comparison between the results from downscaling AMSR-E soil moisture ( $25 \times 25$  km) data using random forest, boosted regression trees, and cubist methods. Random forest is a method of creating an ensemble of decision trees and deriving the most voted class by the decision trees (Breiman, 2001). Boosted regression trees is also a method very similar to random forest, where both rely on an ensemble of decision trees. However, boosted regression trees tend to overfit due to the re-weighted scheme that it uses to select training samples. Cubist is also a regression tree-based model that produces rule-based regression models. Vegetation indices, LST, albedo, and evapotranspiration were used to downscale AMSR-E soil moisture data in all three methods. The comparison of the results shows that the random forest method outperforms the other two methods (Im et al., 2016). Similar results were observed in (Liu et al., 2017), where the random forest method shows superior performance over Classification and Regression Trees (CART), K-nearest neighbors (KNN), and the Bayesian method for soil moisture downscaling. Based on the study region, scale, and period, the importance of the independent variables used in random forest models has shown varying results. In most cases, LST, NDVI, topography, albedo, and enhanced vegetation index (EVI) are seen as the most important independent variables (not in order) (Chen et al., 2020a; Im et al., 2016; Liu et al., 2017; Zhang et al., 2022; Zhao et al., 2018).

### III. METHODS AND DATA

#### 3.1 Field data collection

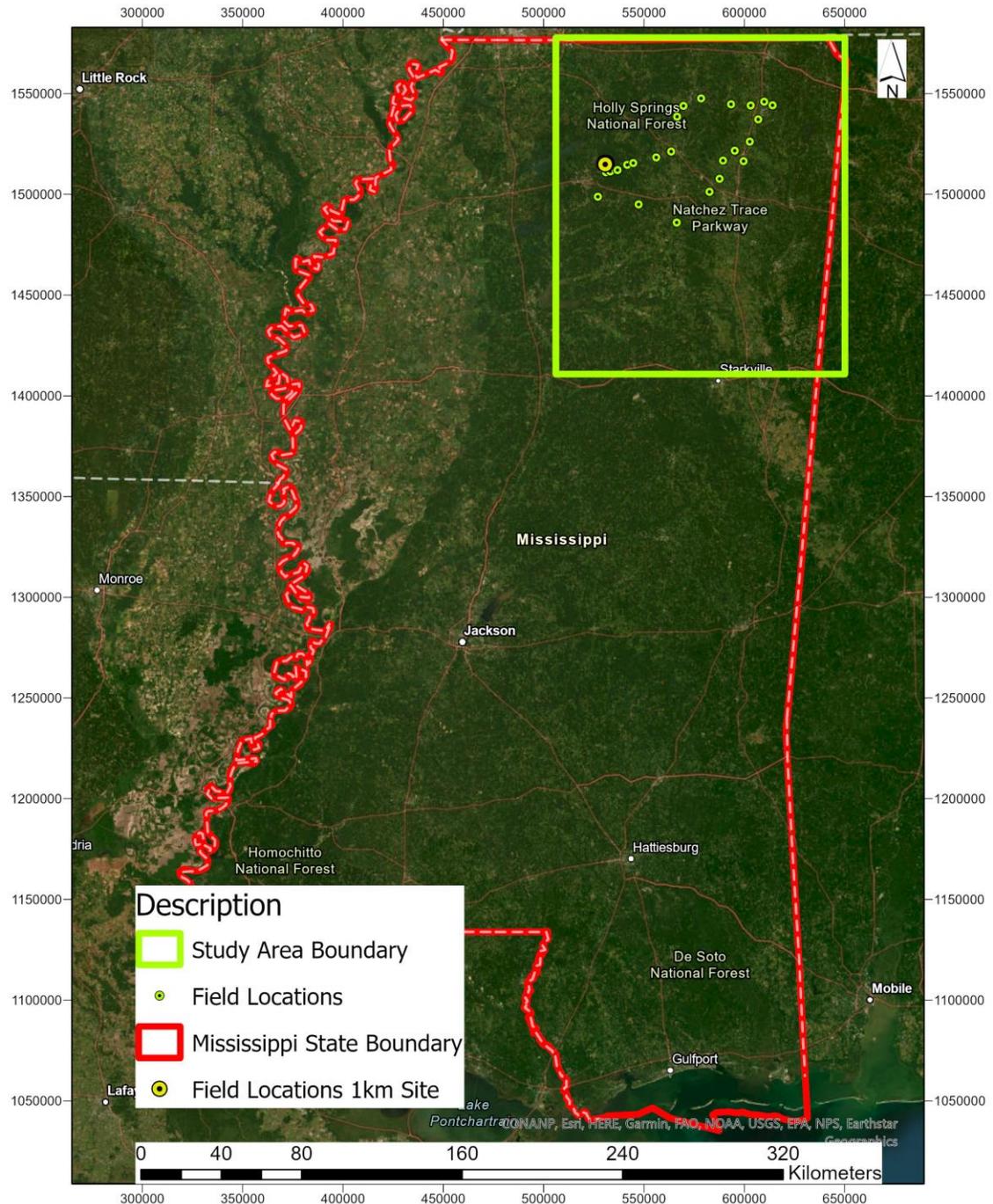
Field surface soil moisture data were collected for the purpose of validating and estimating errors in downscaled soil moisture products. Twenty-five 25 field locations were picked in northeastern Mississippi to collect soil moisture data (**Figure 4**). Field locations were picked to have a wide spatial distribution and include different land cover types. When selecting the locations for field work, it was also taken into consideration the ability to cover all 25 locations within a single day.

On April 2023, one day was spent in a field site (approximately 1 km<sup>2</sup> area) within the study area to collect surface soil moisture with a closer spacing between the sample points. Twenty-three surface soil moisture readings were taken while maintaining a gap of approximately 250 m between the sample locations. Due to the rough terrain, the distance between the points was changed in some locations for easy access. The site itself was mainly covered by non-vegetated cropland with a forest patch and a nearby creek. **Figure 5** shows the locations marked within this study area. This specific field work was carried out to identify the heterogeneity of surface soil moisture in a smaller spatial extent.

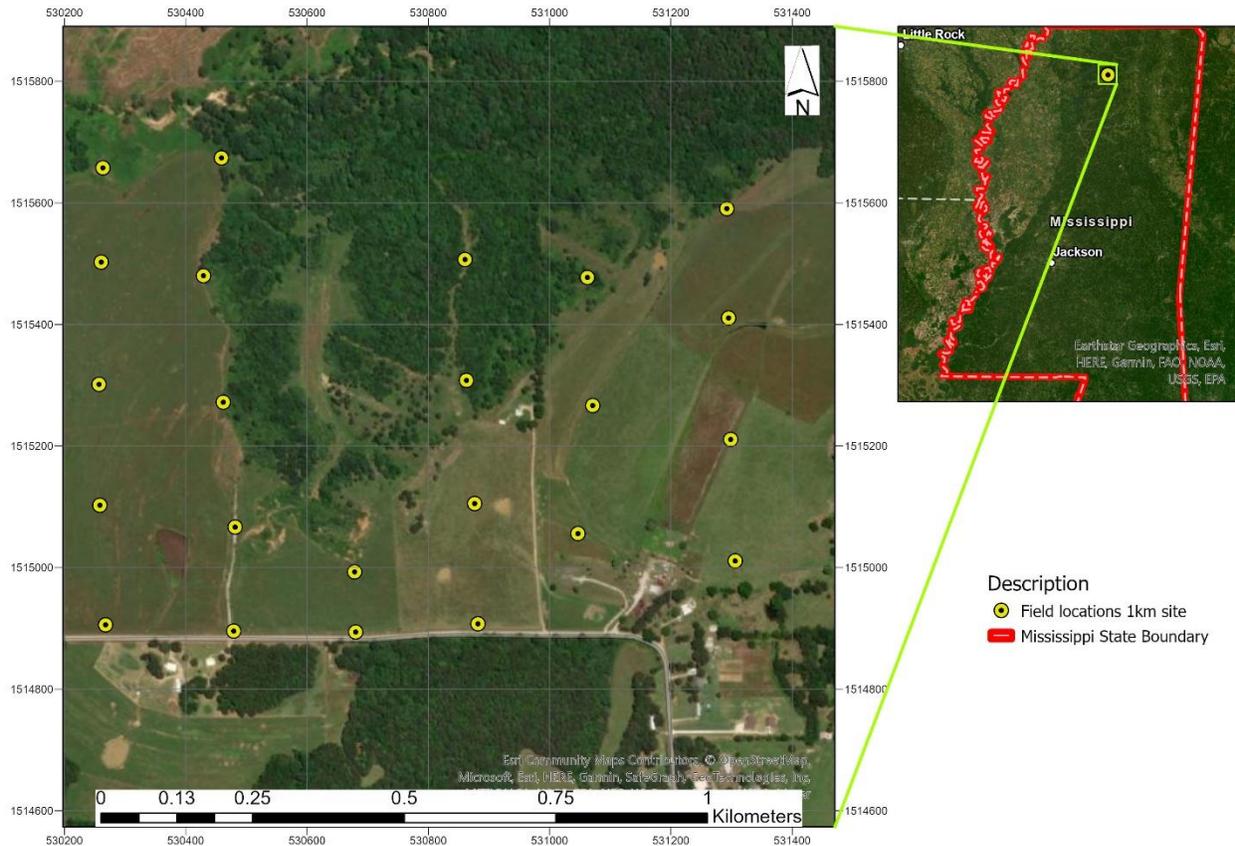
The Fieldscout® soil sensor reader and a portable Waterscout™ SM100 sensor were used to obtain surface soil moisture data. Waterscout™ SM100 sensor is a capacitive sensor that gauges the soil moisture based on the amount of electric charge that can be stored within an

electric potential. The L-band microwave sensing is sensitive to the soil moisture in top the 4 to 5 centimeters of the soil (Kerr, 2007). Therefore, the upper 2–3 cm of soil is removed to expose the soil surface, and the sensor is pressed into the side wall of the dug hole at a 2–3 cm depth to represent the average. In soil where the electrical conductivity is smaller than 8 millisiemens·cm<sup>-1</sup>, the instrument has a nominal accuracy of 0.03 m<sup>3</sup>·m<sup>-3</sup> (Spectrum Technologies, 2022).

Considering the temporal overlap of the satellite images and the least cloud cover, surface soil moisture was measured at each location in northeastern Mississippi on six different days. This surface soil moisture field data was collected to create validation layers for downscaled surface soil moisture data. In addition to collecting surface soil moisture data using the soil moisture probe, soil samples were collected on the first two field visits to ensure the accuracy of the soil moisture probe data. Soil samples were collected into tin containers and sealed at the field to avoid loss of soil moisture.



**Figure 4.** Field locations used for surface soil moisture data collection. Field locations used for surface soil moisture data collection.



**Figure 5.** Surface soil moisture collection locations within the 1 km<sup>2</sup> study site.

Samples brought to the laboratory were measured for the soil moisture content following ASTM D2216-19 standard procedure. As per the standard, soil samples were dried at a temperature of 90 °C for a time period of 12 hours. The samples were weighed before and after drying to calculate the gravimetric soil moisture. The formula used for calculating gravimetric soil moisture is given in Equation 2.

$$w = \frac{\text{mass of water in soil}}{\text{mass of solids in soil}} \times 100 \quad (2)$$



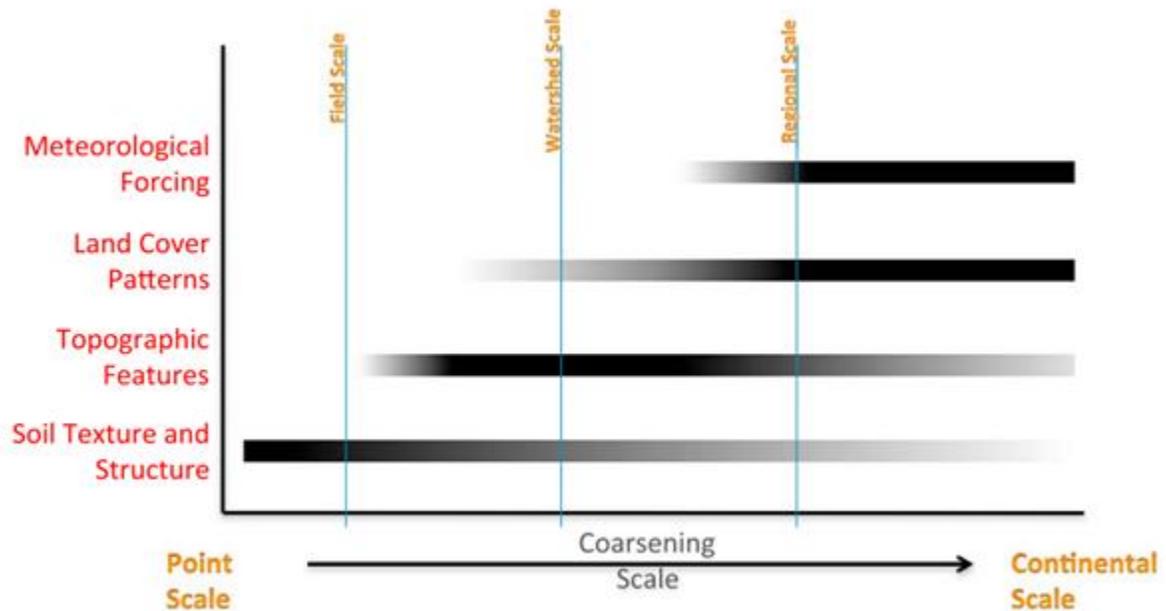
**Figure 6.** Soil moisture data collecting using the field soil moisture sensor. (LEFT) in closeup of the instrument interface along with a handheld GPS unit, and (RIGHT) showing position of probe in prepared test excavation.

### 3.2 Selecting input data

It is important to select parameters that have a correlation with surface soil moisture when using machine learning to downscale satellite derived surface soil moisture data. Based on previous research work and considering the spatial extent of the study region, NDVI, EVI, and diurnal land surface temperature was selected to be used as ancillary data for the downscaling approach (Bai et al., 2019; Chen et al., 2020a; Mao et al., 2022; Peng et al., 2017). As we considered days where there was no precipitation within a window of about 5 to seven days before the field visits, precipitation data were not considered for the downscaling model. **Table 2** shows the dates considered for downloading the data products.

**Table 2.** Study dates considered for downloading satellite data.

| Day Number | Field visit date | MODIS LST  | MODIS NDVI/EVI       | SMAP       |
|------------|------------------|------------|----------------------|------------|
| 1          | 2/7/2022         | 2/6/2022   | 2/2/2022- 2/20/2022  | 2/7/2022   |
| 2          | 3/4/2022         | 3/4/2022   | 2/18/2022-3/7/2022   | 3/4/2022   |
| 3          | 5/11 2022        | 5/10/2022  | 5/9/2022-5/26/2022   | 5/11/2022  |
| 4          | 7/28/2022        | -          | -                    | 7/28/2022  |
| 5          | 9/28/2022        | 9/28/2022  | 9/14/2022-10/1/2022  | 9/28/2022  |
| 6          | 10/21/2022       | 10/21/2022 | 10/16/2022-11/3/2022 | 10/21/2022 |



**Figure 7.** Parameters affecting the spatial variability of surface soil moisture as a function of the spatial extent of the study region considered (from (Crow *et al.*, 2012))

### 3.3 Satellite derived surface soil moisture

Surface soil moisture data used for downscaling were downloaded from the National Snow and Ice Data Center SMAP data repository. Version 7 of the SMAP level four root zone soil moisture geophysical data product (SPL4SMGP) was used in this study. The level four product is produced by assimilating L-band radiometer brightness temperature data into a land surface model (Reichle et al., 2022). Data in this product is gridded into a global cylindrical 9 km equal-area scalable Earth grid (EASE-Grid 2.0). Data granules covering the time period starting at noon and ending at the 15th hour of the day were downloaded for each day considered in the downscaling. Surface soil moisture (0–5 cm vertical average) data from this product were used in this study.

### 3.4 Land surface temperature (LST)

“MODIS/ Aqua Land Surface Temperature/ Emissivity Daily L3 Global 1km SIN Grid V061 ” data product (MYD11A1) downloaded from NASA Earthdata search web portal was used to derive land surface temperature values for the study area. The product contains both night time and daytime land surface temperature values at a spatial resolution of 1km (Wan, 2006). Depending on the cloud cover of the data acquisition time, the pixels may or may not contain LST values for the pixels. As the project makes use of diurnal temperature for the downscaling, it was important to select pixels with both night and daytime LST values for the processing. When data was obscured due to cloud cover, the day before or after the collection of field soil moisture data was considered for downloading LST data products. When there is no precipitation, the diurnal temperature is assumed to be similar for consecutive days.

### 3.5 NDVI and EVI

“MODIS/Terra Vegetation Indices 16-Day L3 Global 1km SIN Grid V061 ” data product (MOD13A2) downloaded from NASA Earthdata Search web portal was utilized in deriving NDVI and EVI data. This product provides the average NDVI and EVI values for a period of 16 days. NDVI and EVI are two standard vegetation indices that are calculated as shown in equation 3 and 4 respectively. The EVI equation given here is a modified version of the usual EVI equation, where the blue band is not used for the calculation. This was done in order to avoid erratic behavior of the conventional EVI equation over bright targets (heavy clouds and snow/ice). This erratic behavior in the original EVI equation is a result of blue band saturation over bright targets. Thus the blue band is removed to maintain the advantages of the EVI while avoiding insensitivity to bright targets (Didan et al., 2015).

$$NDVI = \frac{NIR-Red}{NIR+Red} \quad (3)$$

$$EVI = 2.5 \frac{NIR-Red}{NIR+2.4Red+1} \quad (4)$$

Where NIR is the near infrared band value.

### 3.6 Land cover types

One of the goals of this project is to compare surface soil moisture data with the land cover types of the study area. For this purpose, the MODIS land cover types product made by combining Terra and Aqua observations is used. This product has a 500m spatial resolution and provides data at a yearly intervals. This product contains 5 different land cover type classifications; (1) International Geosphere-Biosphere Programme (IGBP), (2) University of

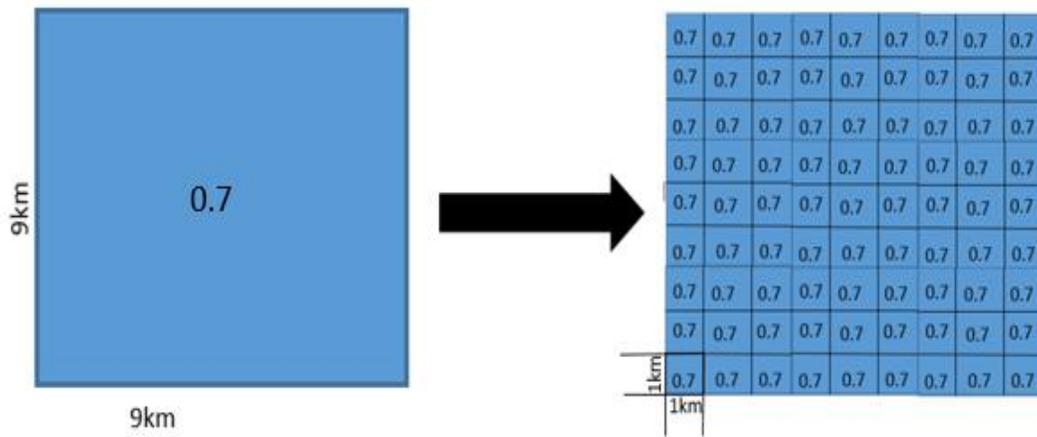
Maryland classification, (3) Leaf Area Index (LAI), (4) BIOME-Biogeochemical cycles (BGC), and (5) Plant Functional Types (PFT). Based on the thick vegetation and its variations in the study region, IGBP classification is used. IGBP classification consists of 17 different land cover types (**Table 3**).

**Table 3.** International Geosphere-Biosphere Programme (IGBP) classification (Sulla-Menashe & Friedl, 2018).

| <b>Name</b>                         | <b>Description</b>   |
|-------------------------------------|--|
| Water Bodies                        | At least 60% of area is covered by permanent water bodies.   |
| Evergreen Needleleaf Forests        | Dominated by evergreen conifer trees (canopy >2m). Tree cover >60%.                                    |
| Evergreen Broadleaf Forests         | Dominated by evergreen broadleaf and palmate trees (canopy >2m). Tree cover >60%.                      |
| Deciduous Needleleaf Forests        | Dominated by deciduous needleleaf (larch) trees (canopy >2m). Tree cover >60%.                         |
| Deciduous Broadleaf Forests         | Dominated by deciduous broadleaf trees (canopy >2m). Tree cover >60%.                                  |
| Mixed Forests                       | Dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy >2m). Tree cover >60%. |
| Closed Shrublands                   | Dominated by woody perennials (1-2m height) >60% cover.  |
| Open Shrublands                     | Dominated by woody perennials (1-2m height) 10-60% cover.  |
| Woody Savannas                      | Tree cover 30-60% (canopy >2m).  |
| Savannas                            | Tree cover 10-30% (canopy >2m).  |
| Grasslands                          | Dominated by herbaceous annuals (<2m).   |
| Permanent Wetlands                  | Permanently inundated lands with 30-60% water cover and >10% vegetated cover.                          |
| Croplands                           | At least 60% of area is cultivated cropland.   |
| Urban and Built-up Lands            | At least 30% impervious surface area including building materials, asphalt, and vehicles.              |
| Cropland/Natural Vegetation Mosaics | Mosaics of small-scale cultivation 40-60% with natural tree, shrub, or herbaceous vegetation.          |
| Permanent Snow and Ice              | At least 60% of area is covered by snow and ice for at least 10 months of the year.                    |
| Barren                              | At least 60% of area is non-vegetated barren (sand, rock, soil) areas with less than 10% vegetation.   |
| Water Bodies                        | At least 60% of area is covered by permanent water bodies.   |
| Unclassified                        | Has not received a map label because of missing inputs.  |

### 3.7 Data preparation

Data preparation for the machine learning models was done in the ArcGIS Pro software. In order to bring all the data into a comparable coordinate system, data products used in the study were clipped to the study area and projected to the Mississippi Transverse Mercator (MSTM) reference coordinate system. For land surface temperature (LST), pixel values were multiplied by a factor of 0.0001 to get the LST value into degrees Kelvin. The diurnal temperature was calculated by subtracting night-time LST values in the data set from daytime LST values. For NDVI and EVI, a scaling factor of 0.0001 was used to obtain proper values. This is because the NASA science team has the original vegetation indices data product scaled to save space and precision during file compression. SMAP soil moisture data were disaggregated into a 1km spatial resolution to match the finest resolution used for the independent variables of the machine learning model. The nearest neighbor method was used in this disaggregation to avoid altering the pixel values (**Figure 8**). In the next step, all raster datasets were converted to point data to prepare them for spatial joining. A fishnet having polygons of 1 km<sup>2</sup> was made covering the study area so that point data for NDVI, diurnal temperature, and surface soil moisture could be spatially joined to the polygons of the fishnet. As a final step, pixels with null values for the variables considered for downscaling were removed from the final data set before running the data in the machine learning model.



**Figure 8.** Figure representation of disaggregating a 9 km cell to 1 km cells using the nearest neighbor method

### 3.8 Random Forest Model

Random forest is a method of creating an ensemble of decision trees and deriving the most voted class by the decision trees (Breiman, 2001). The method is inspired by the previous work done on randomized trees and is introduced as a competitor to boosting method (Cutler et al., 2012). Cutler et al. (2012) also describe random forest as an extension of Breiman’s bagging idea.

Data required for random forests have predictors and dependent variables. A data set consisting of vectors for both of these variables is randomly sampled when developing the forest of regression trees (ensemble of decision trees). This is what gives this method the name “Random Forest”. The randomly selected data set for the regression tree is known as the bootstrap sample. As the data for growing trees are sampled with replacement, each tree gets to use the complete range of the original training data set for random sampling for training. Repeated bootstrap sampling with replacement in the random forest is done through a procedure

named “bootstrap aggregation”, which is given the acronym bagging (Breiman, 1996). Thus the training sets used for tree growth are known as the in-bag data.

For an  $M$  number of input variables, the model decides a number  $m \ll M$ , which is the number of variables picked for splitting at each node of the decision tree. Value for the  $m$  number is kept constant throughout the decision tree growing process. Ultimately predictions are calculated for each decision tree, and an arithmetic average of the predictions is given as a final forecast by the random forest (Bartkowiak et al., 2019). Equation 5 represents how random forest predicts final results based on regression trees.

$$F(x) = \frac{\sum_{j=1}^N T_j(x)}{N} \quad (5)$$

where  $N$  indicates the number of regression trees,  $T_j$  represents each tree, and  $F$  is a prediction for a point  $x$  (Bartkowiak et al., 2019).

According to (Breiman, 2001), there are two things that control the forest error rate. The first is introduced as the correlation between any two trees of the random forest. Higher the correlation between the regression trees, the higher the error rate. The second contributor to the error rate is the individual tree error rate. Both of these errors are influenced by the  $m$  number of variables selected for splitting at each node. While reducing  $m$  reduces the correlation between the trees, it also increases the individual tree error. Therefore, the model should maintain an optimum number to balance the error contributors. The model does this by analyzing the error for the out-of-bag data set. Out-of-bag data is the data left out when in-bag data is selected for model training, meaning the out-of-bag data can give an unbiased estimate of the classification error by the trees.

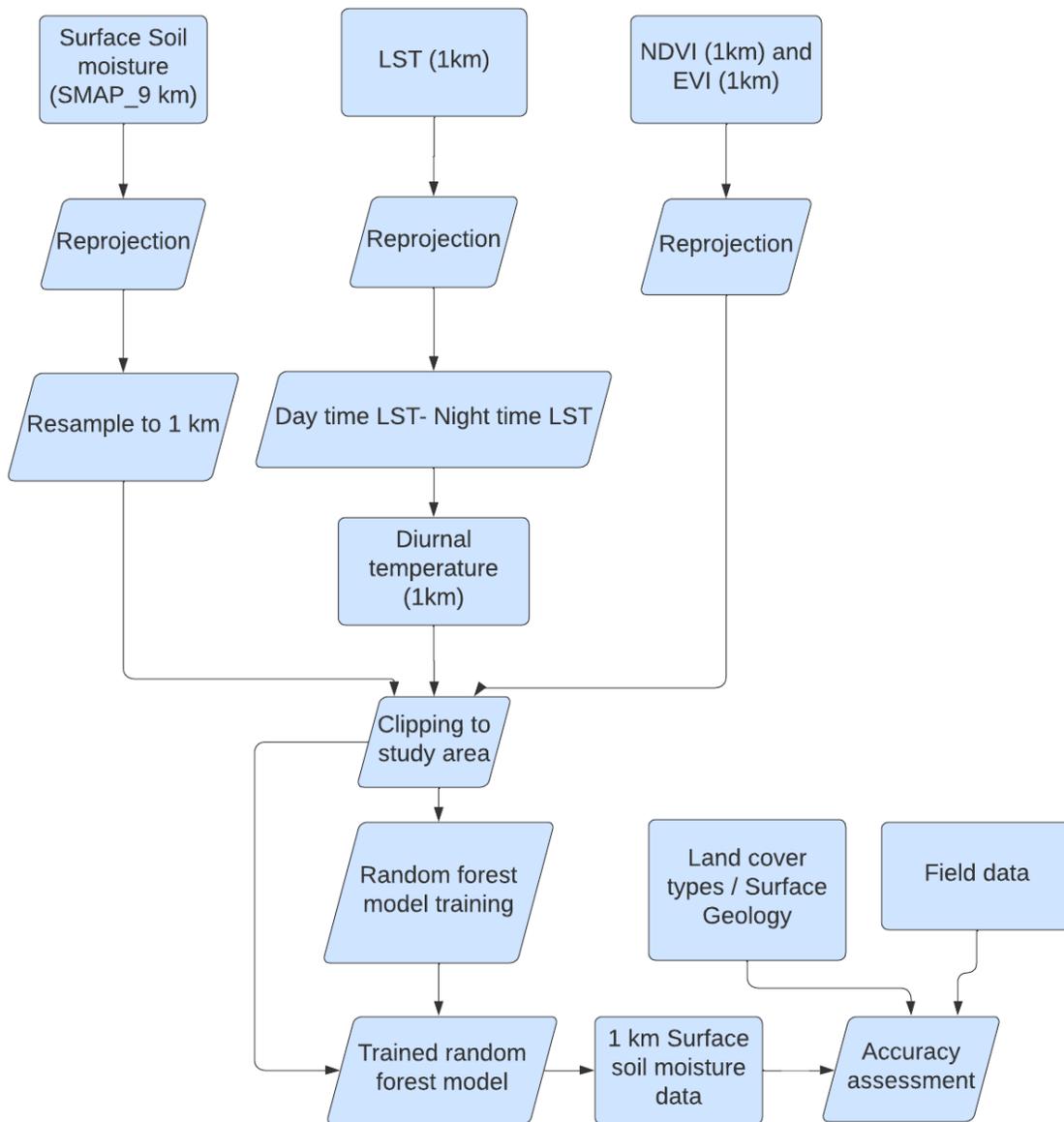
The random forest also calculates variable importance using the out-of-bag data set. To give variable importance, the model uses a method of comparing the number of votes cast by the trees for the correct class when out-of-bag data is used, against the number of votes cast by the trees for the correct class when randomly permuted out-of-bag data is used.

The random forest classification tool from ArcGIS Pro was used to downscale surface soil moisture data. This tool uses the same basic concepts of the random forest technique put forward by (Breiman, 2001). Disaggregated surface soil moisture values from SMAP level four products were used as the variable to predict, while NDVI, EVI, and diurnal temperature values were used as predictors. Due to the cloud cover, a good LST image covering the study area was not found within a window of seven days before and after July 28, 2022. For this reason, the month of July was not considered during the downscaling process.

### 3.9 Validating results

The accuracy of the model output was measured using several approaches. The first approach to validate the model accuracy was to compare the output with the data collected in the field. Correlation coefficient values were obtained for the relationship between field data and coarse resolution/downscaled data for the study dates considered. Furthermore, mean, median, mode, and several other statistics were calculated for the pixels of coarse resolution and downscaled surface soil moisture datasets. These general statistics were used to analyze the key changes between the downscaled and coarse resolution data sets. A visual comparison of the coarse-resolution surface soil moisture, fine-resolution surface soil moisture, surface geology, and land cover types were made by overlaying maps to see if any matching patterns exist between the different data types. As a final step, both the coarse resolution and fine resolution

data sets were clipped based on the land cover types from IGBP, and statistics were calculated for each land cover type. Statistics over different vegetation types were used to show the relationship of the surface soil moisture to different land cover types. **Figure 9** is a summary of the entire workflow followed for the downscaling of the coarse resolution surface soil moisture data.

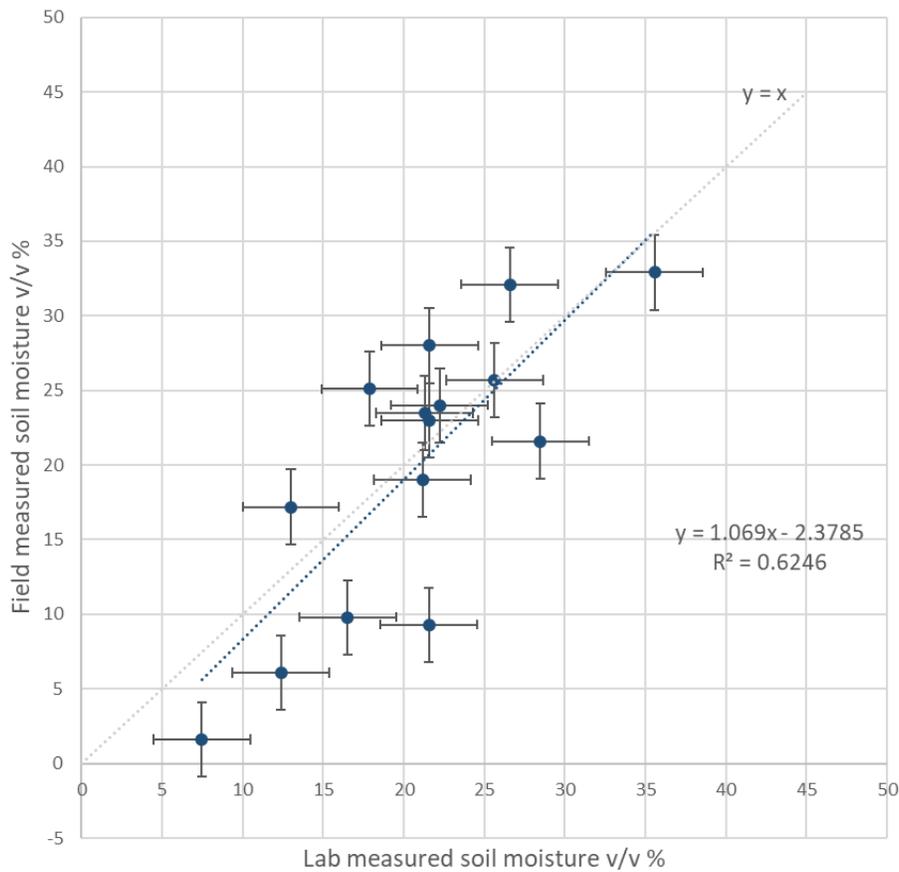


**Figure 9.** Flow chart summarizing workflow followed for the downscaling of the coarse resolution surface soil moisture.

## IV. RESULTS

### 4.1 Soil sample testing and field collected data

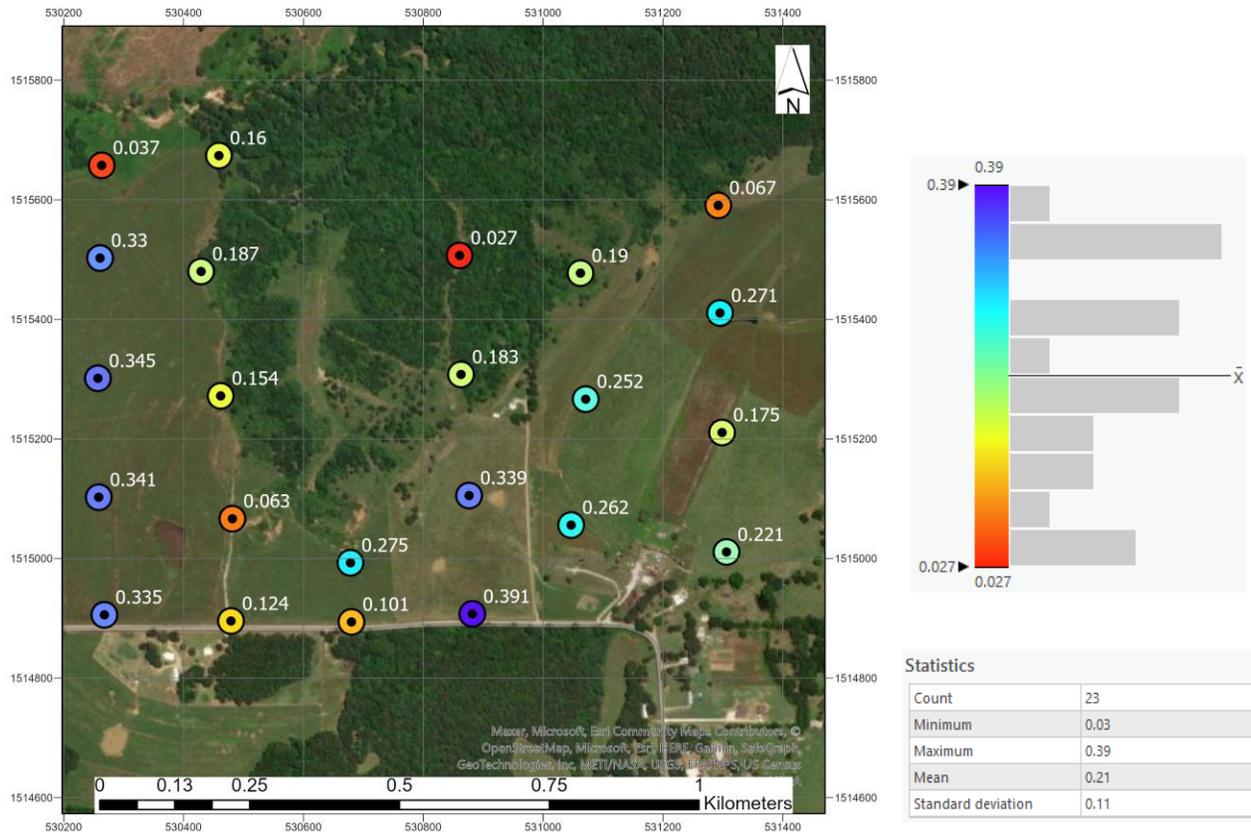
The values of soil moisture measured in the laboratory were compared with the soil moisture probe readings of the same locations from where the soil samples were collected (Figure 10).



**Figure 10.** Scatter plot showing the comparison of soil moisture data from the field and laboratory tested samples.

Field surface soil moisture data collected on April 18, 2023, are plotted on the map to demonstrate the variations in the surface soil moisture values. Minimum, maximum, mean, and

standard deviation were calculated for this data and are shown in

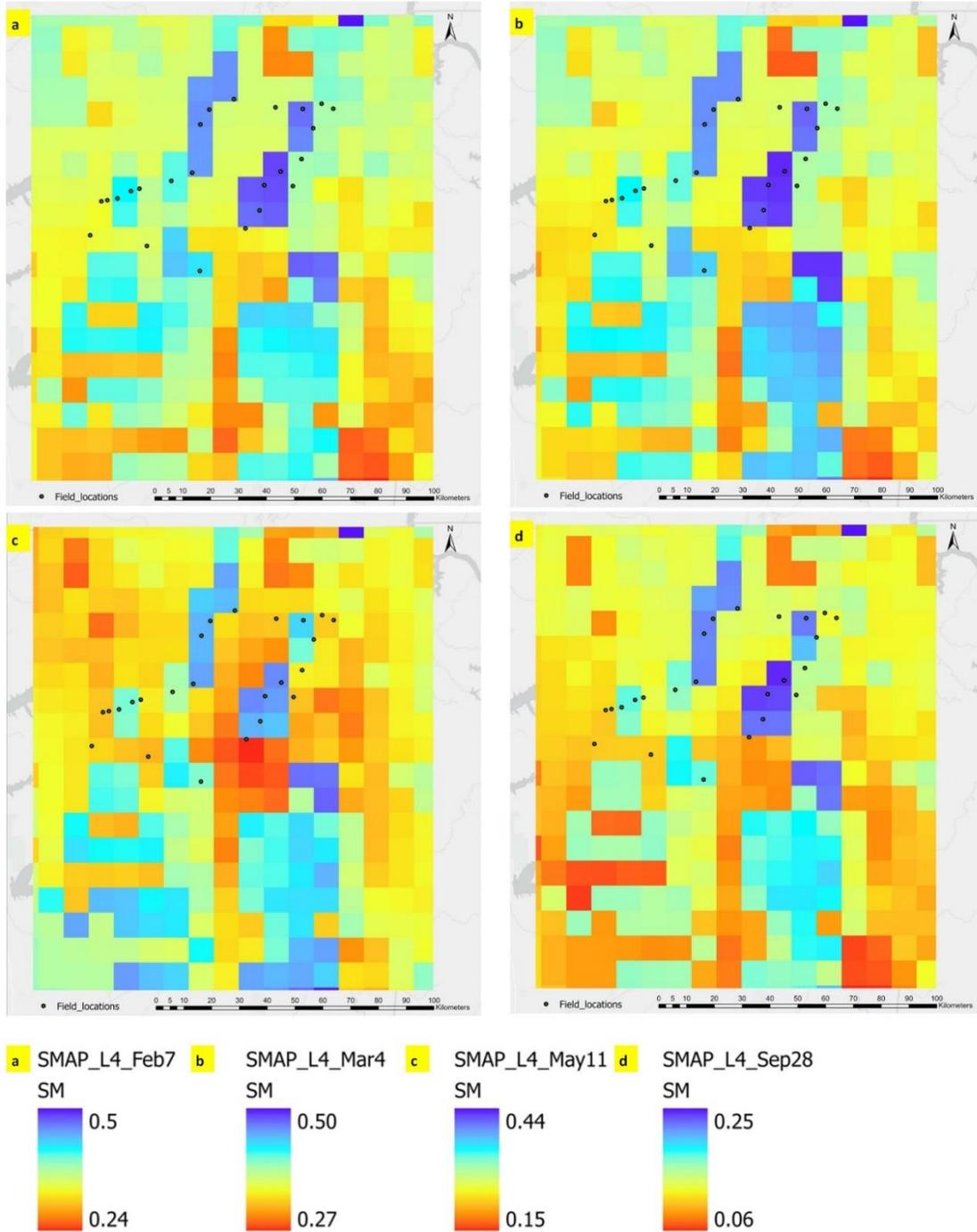


**Figure 11.** Field surface soil moisture measurements in an area of approximately 1 km<sup>2</sup> and statistics for the measurements. Soil moisture values are reported as volume/volume.

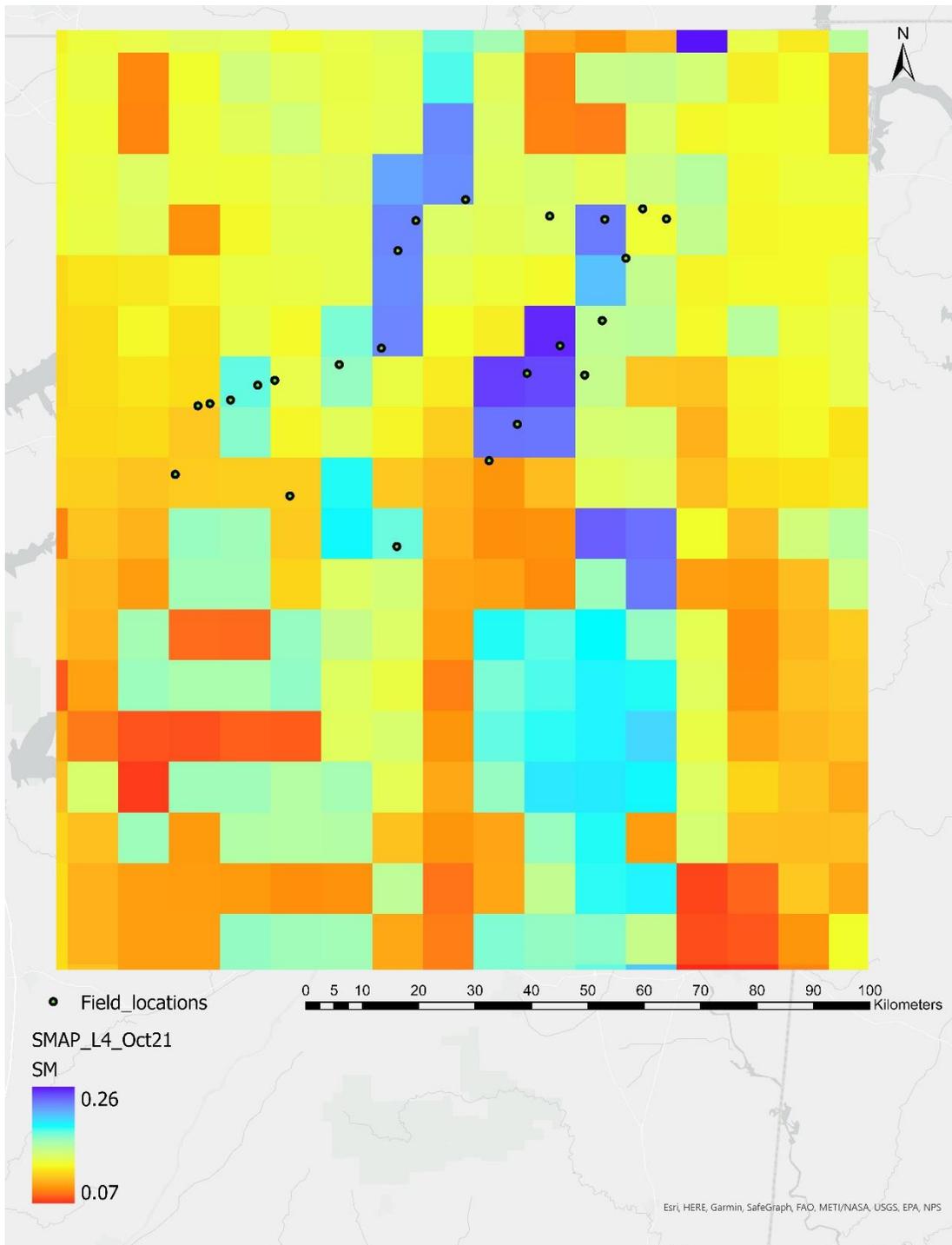
#### 4.2 Results from the downscaling

Both coarse resolution and downscaled data are created over the study region and presented in **Figure 12** through **Figure 15**. Images for October are chosen to present as single large images due to contrasting differences in the surface soil moisture values across the mapped area. All the images are mapped using the same color ramp normalized between the minimum and maximum for the respective image. Scatter plots showing the correlation of field-collected

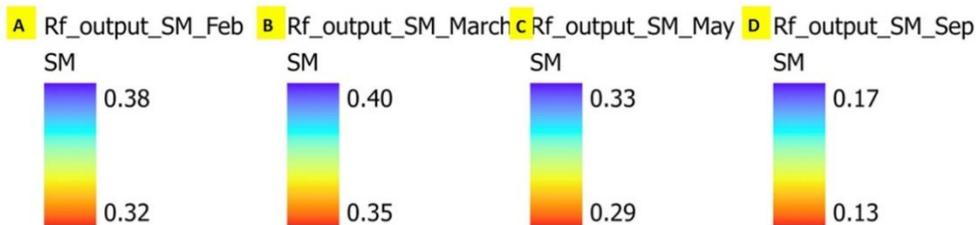
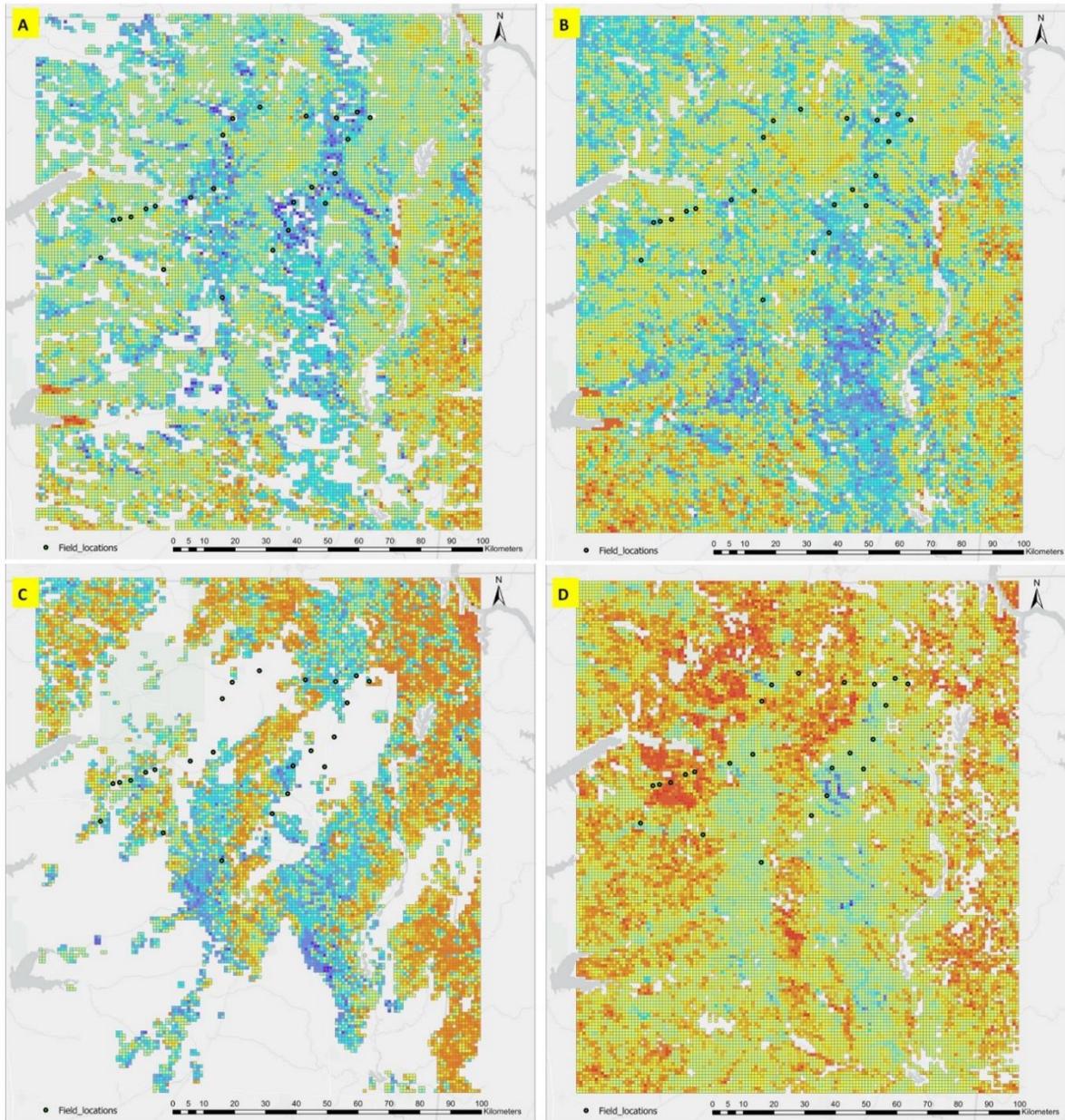
surface soil moisture data to coarse resolution and downscale model predictions are presented in **Figure 16** through **Figure 19**. **Table 4** provides a summary of the correlation values obtained in graphs from **Figure 16-Figure 19**. **Table 5** provides an overview of the importance of the independent variables in the downscaling model. General statistics for the pixel values of field collected, coarse resolution, and downscaled surface soil moisture data are presented in **Table 6**.



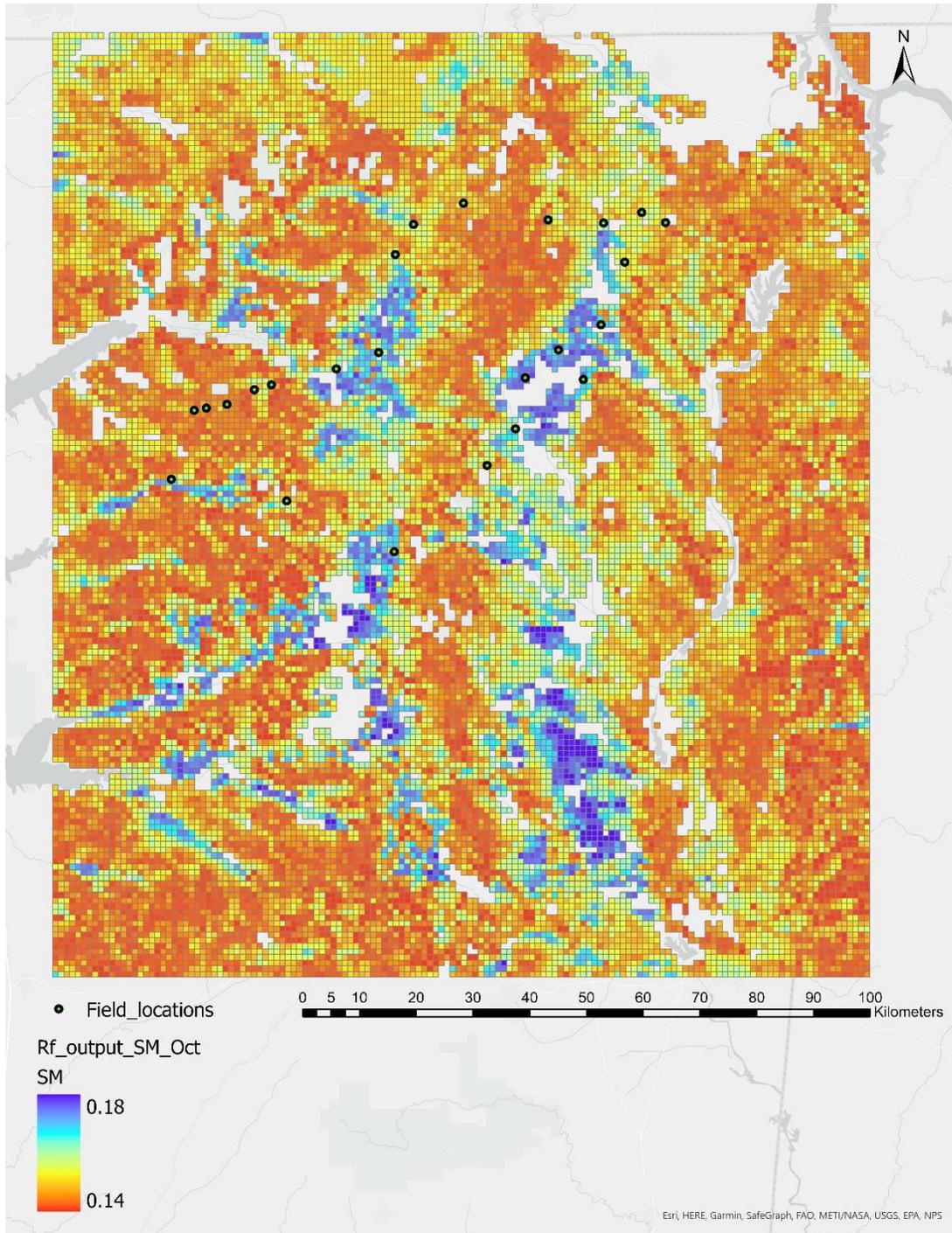
**Figure 12.** 9 km spatial resolution surface soil moisture data from SPL4SMGP product, clipped to the study area for the months of (A) February, (B) March, (C) May, and (D) September. Soil moisture values shown are volume/volume.



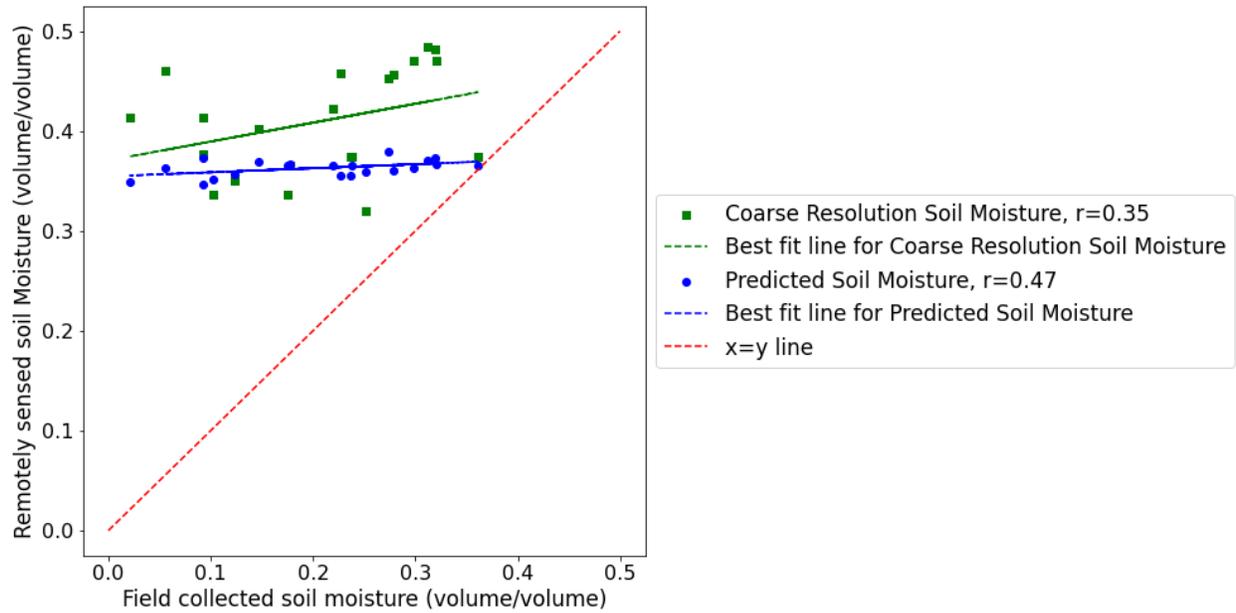
**Figure 13.** 9 km spatial resolution surface soil moisture data from SPL4SMGP product, clipped to the study area for the month of October. The soil moisture values shown are volume/volume.



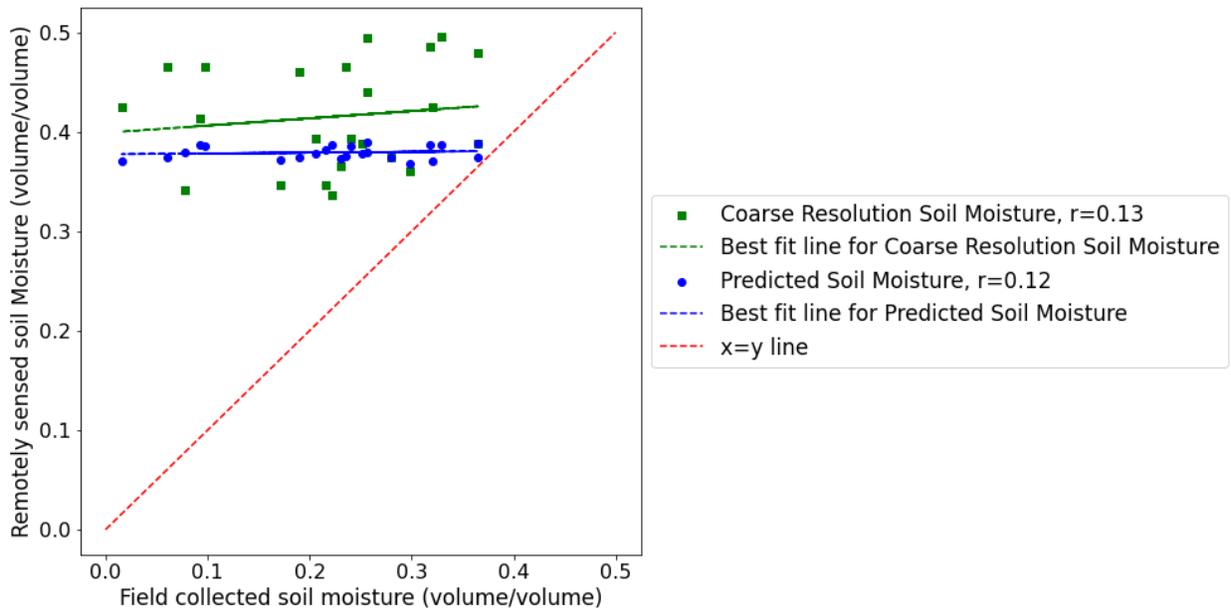
**Figure 14.** Surface soil moisture downscaling results for the months of (A) February, (B) March, (C) May, and (D) September. Spatial resolution is 1km, and soil moisture values shown are volume/volume.



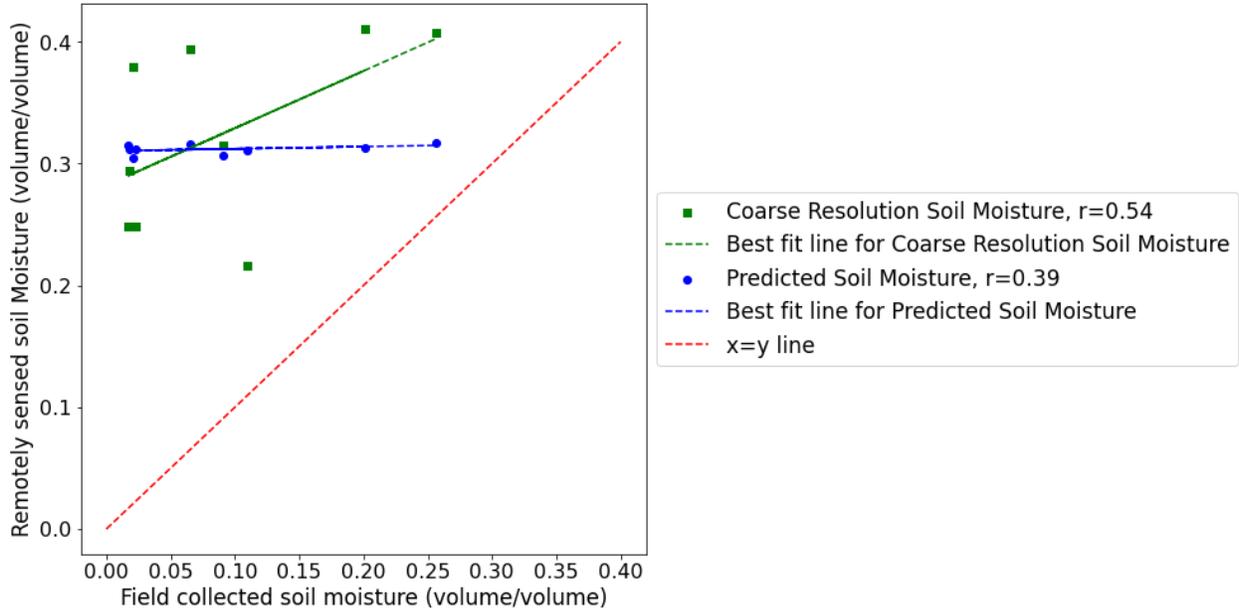
**Figure 15.** Surface soil moisture downscaling results for the month of October. Spatial resolution is 1 km, and soil moisture values shown are volume/volume.



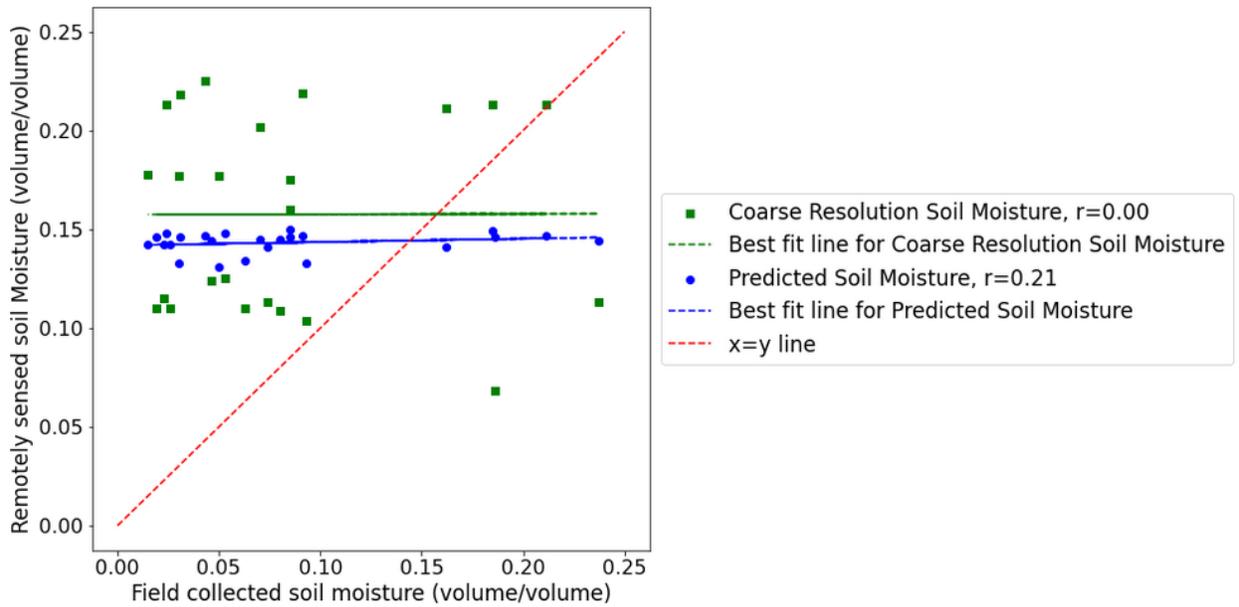
**Figure 16.** Remotely sensed soil moisture vs. field soil moisture readings for the month of February.



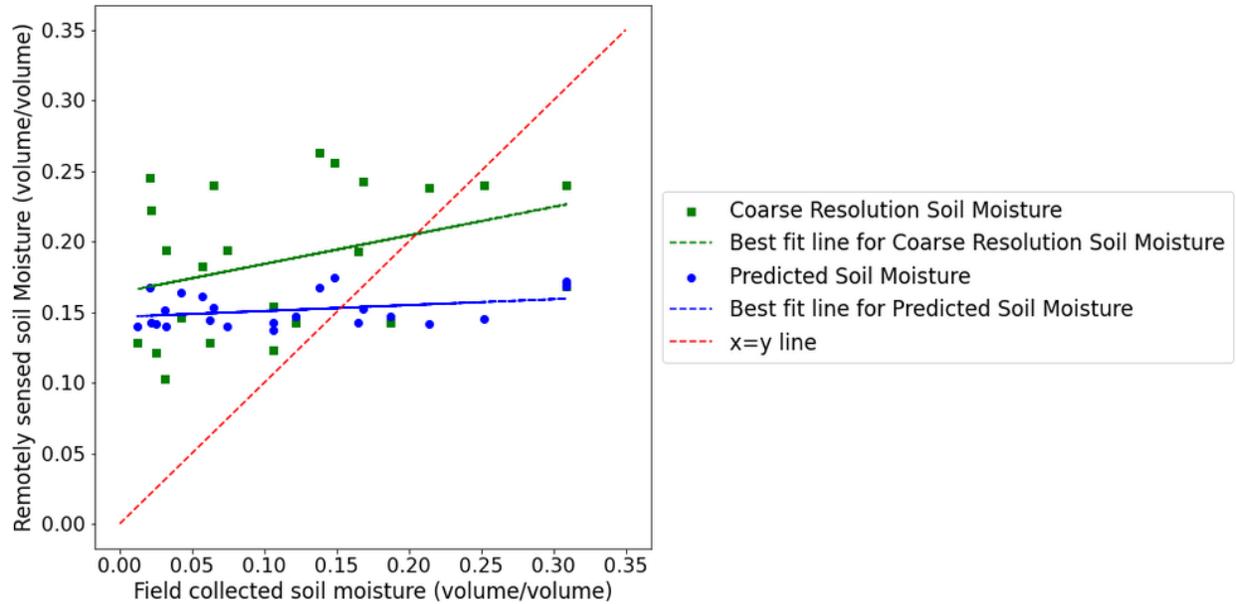
**Figure 17.** Remotely sensed soil moisture vs. field soil moisture readings for the month of March.



**Figure 18.** Remotely sensed soil moisture vs. field soil moisture readings for the month of May.



**Figure 19.** Remotely sensed soil moisture vs. field soil moisture readings for the month of September.



**Figure 20.** Remotely sensed soil moisture vs. field soil moisture readings for the images from the month of October.

**Table 4.** Summary of correlation coefficient values for the comparison of field collected surface soil moisture data and remotely sensed surface soil moisture data (values from **Figure 16-Figure 19**).

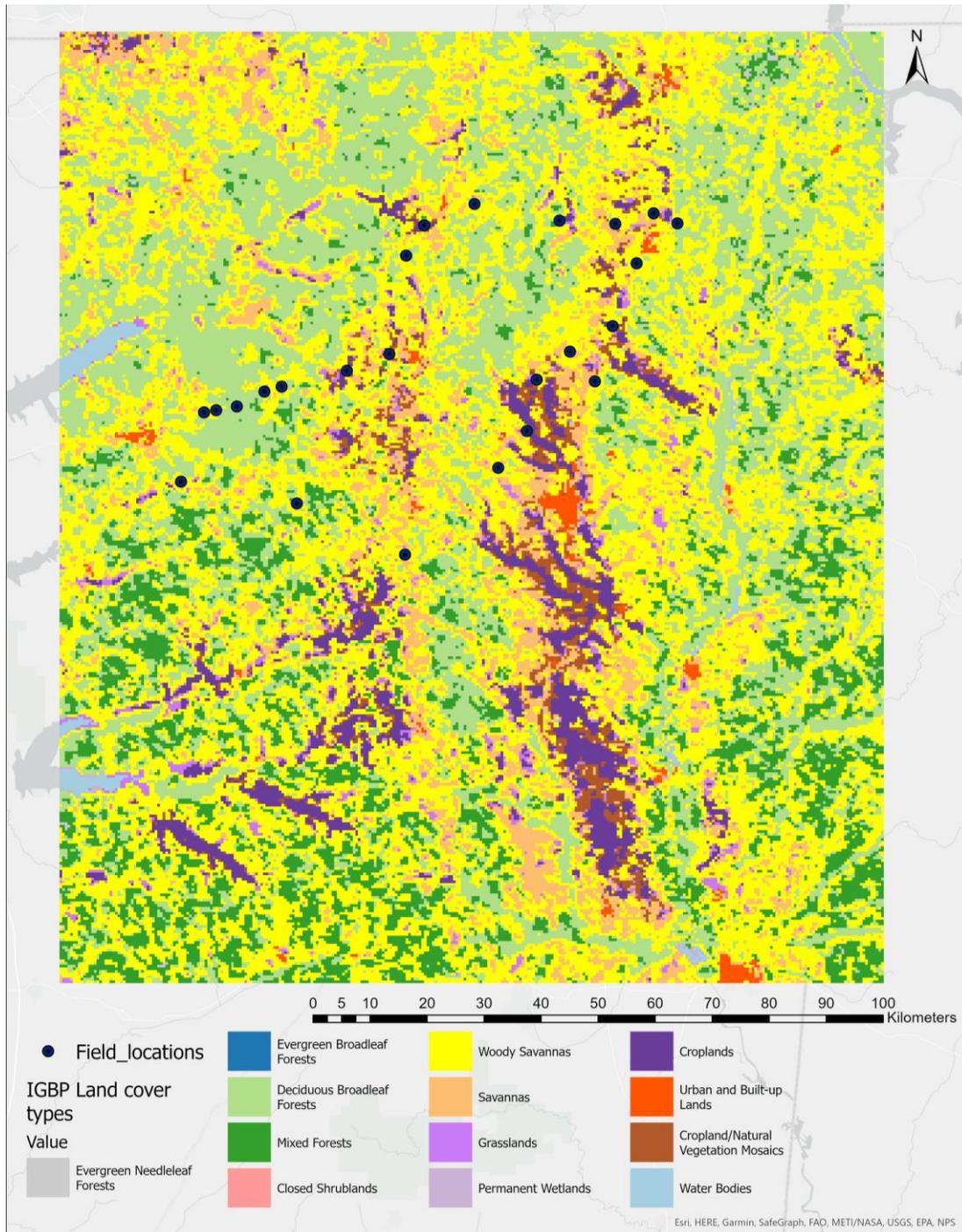
|                          | Correlation of Coefficient |       |      |           |         |
|--------------------------|----------------------------|-------|------|-----------|---------|
|                          | February                   | March | May  | September | October |
| <b>Coarse resolution</b> | 0.35                       | 0.13  | 0.54 | 0         | 0.36    |
| <b>Fine resolution</b>   | 0.47                       | 0.12  | 0.39 | 0.21      | 0.32    |

**Table 5.** Parameter importance for the downscaling models constructed for each month.

| Parameter   | Importance Percentage |       |      |           |         |
|-------------|-----------------------|-------|------|-----------|---------|
|             | February              | March | May  | September | October |
| <b>EVI</b>  | 0.26                  | 0.26  | 0.28 | 0.16      | 0.17    |
| <b>LST</b>  | 0.32                  | 0.26  | 0.35 | 0.25      | 0.25    |
| <b>NDVI</b> | 0.42                  | 0.48  | 0.36 | 0.59      | 0.57    |

**Table 6.** Statistics of the downscaled and coarse resolution surface soil moisture data.

|                  | <b>Field Name</b> | <b>Minimum</b> | <b>Maximum</b> | <b>Mean</b> | <b>Standard Deviation</b> | <b>Median</b> | <b>Count</b> | <b>Number of Unique Values</b> | <b>Mode</b> | <b>Sum</b> |
|------------------|-------------------|----------------|----------------|-------------|---------------------------|---------------|--------------|--------------------------------|-------------|------------|
| <b>February</b>  | SM 9km            | 0.241          | 0.500          | 0.353       | 0.048                     | 0.348         | 18438        | 292                            | 0.346       | 6504.381   |
|                  | Field Data        | 0.021          | 0.361          | 0.206       | 0.097                     | 0.227         | 21           | 21                             |             | 432.300    |
|                  | SM 1km            | 0.000          | 0.381          | 0.353       | 0.010                     | 0.352         | 18439        | 14436                          | 0.340       | 6503.997   |
| <b>March</b>     | SM 9km            | 0.272          | 0.502          | 0.375       | 0.047                     | 0.363         | 22592        | 295                            | 0.333       | 8477.617   |
|                  | Field Data        | 0.016          | 0.365          | 0.222       | 0.098                     | 0.235         | 23           | 21                             |             | 509.780    |
|                  | SM 1km            | 0.352          | 0.400          | 0.375       | 0.008                     | 0.373         | 22592        | 18253                          | 0.355       | 8476.983   |
| <b>May</b>       | SM 9km            | 0.151          | 0.446          | 0.303       | 0.054                     | 0.294         | 11602        | 254                            | 0.299       | 3514.312   |
|                  | Field Data        | 0.017          | 0.256          | 0.089       | 0.087                     | 0.065         | 9            | 9                              |             | 80.200     |
|                  | SM 1km            | 0.286          | 0.327          | 0.303       | 0.010                     | 0.302         | 11602        | 9552                           | 0.320       | 3514.371   |
| <b>September</b> | SM 9km            | 0.068          | 0.254          | 0.141       | 0.039                     | 0.128         | 21437        | 303                            | 0.124       | 3033.142   |
|                  | Field Data        | 0.015          | 0.237          | 0.083       | 0.065                     | 0.070         | 24           | 23                             |             | 198.200    |
|                  | SM 1km            | 0.127          | 0.172          | 0.141       | 0.006                     | 0.142         | 21437        | 17127                          | 0.149       | 3032.162   |
| <b>October</b>   | SM 9km            | 0.076          | 0.263          | 0.146       | 0.038                     | 0.144         | 21212        | 311                            | 0.112       | 3104.134   |
|                  | Field Data        | 0.012          | 0.309          | 0.116       | 0.091                     | 0.106         | 23           | 21                             |             | 266.700    |
|                  | SM 1km            | 0.136          | 0.179          | 0.147       | 0.009                     | 0.144         | 21212        | 14567                          | 0.179       | 3108.614   |



**Figure 21.** Land cover classification data downloaded from MODIS for the period 2021–2022 visualized over the study area as per IGBP classification.

## V. DISCUSSION

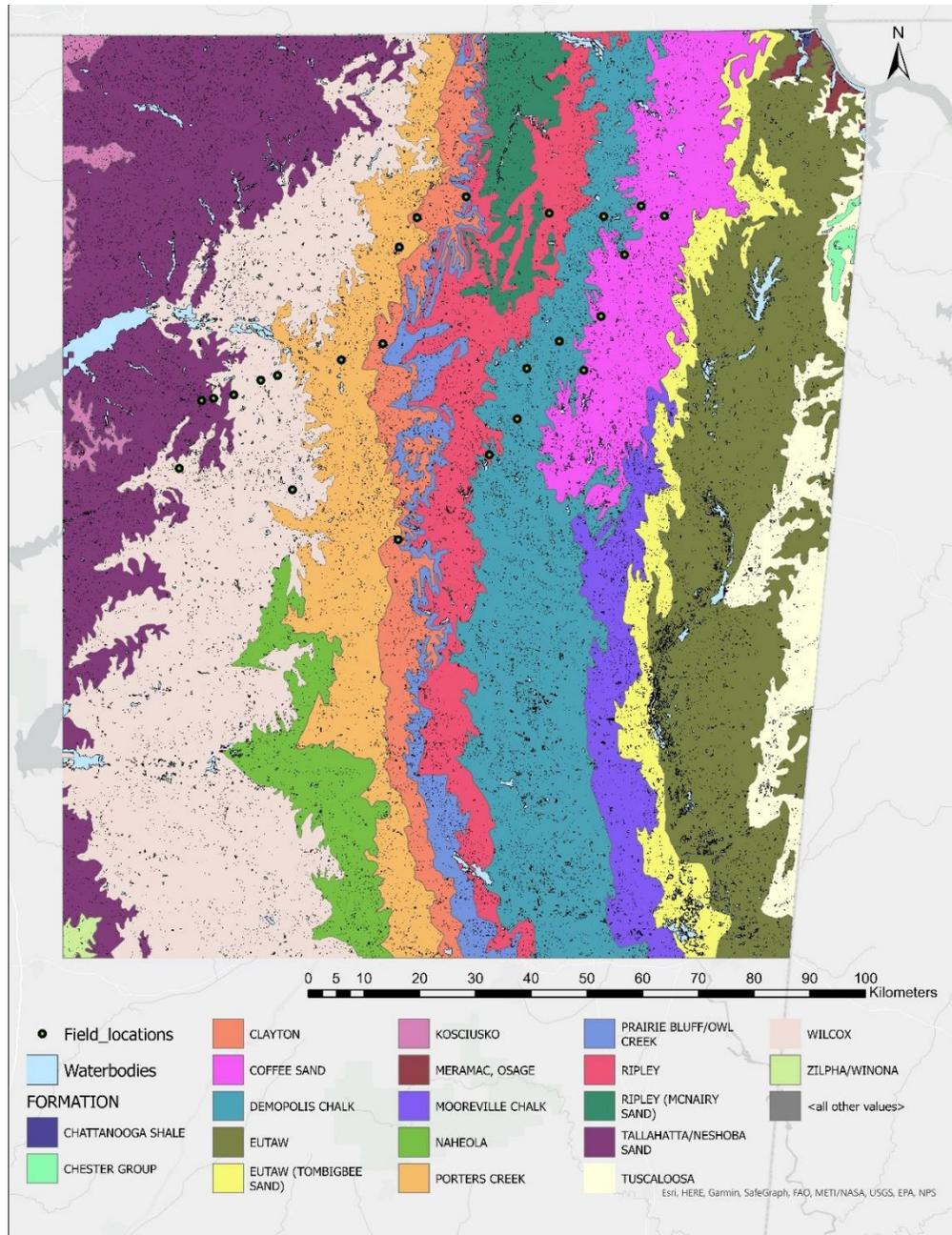
### 5.1 Soil sample testing

As shown in the graph from **Figure 10**, the laboratory measured soil moisture and field measured soil moisture has a coefficient of determination of 0.62 with the best fit line closely aligning to the  $y = x$  line. Ideally, a correlation of determination of 1 along with a best fit line aligning with  $y = x$  would indicate the best performance of the soil moisture probe. However, the readings made in the field using the soil moisture probe could contain an error. One of the most common errors collecting soil moisture with an electronic sensor is poor contact between the sensor and the soil. This is a common error when the soil is loose or very low density. To avoid this, several measurements were taken during the fieldwork, and multiple readings were taken at each site. Added to this, the instrument could contribute a possible error of  $\pm 0.03 \text{ m}^3 \cdot \text{m}^{-3}$  when the soil electrical conductivity is smaller than  $8 \text{ millisiemens} \cdot \text{cm}^{-1}$  (Spectrum Technologies, 2022; Susha Lekshmi et al., 2014). Similarly, an error margin can be expected from laboratory measured soil moisture values, especially when comparing gravimetric and volumetric moisture content. In the field, the volumetric moisture content was estimated using conductivity while in the laboratory the gravimetric was measured. To compare the density of soil must be assumed. Every sample is likely to be different and was assumed to be the same. Hence a source of error.

## 5.2 Visual comparison of pixel patterns in coarse resolution and downscaled images

The coarse-resolution SPL4SMGP images visualized over the study region provide a very interesting observation (**Figure 12, Figure 13**). Despite the dates considered in the five different months, the images have an extremely similar pattern that is only differentiated by the surface soil moisture value. Which means the controls on the distribution of surface soil moisture are spatial. (T. Wang et al., 2020) explains the soil texture is one such spatial control on soil moisture. Since we avoided precipitation events when downloading soil moisture data, we cannot comment on the influence of precipitation on the spatial patterns of surface soil moisture.

As observed in **Figure 22**, surface geology of the terrain presents several geologic units that range from north to south as irregular bands. These irregular bands fit with the pattern observed in the coarse-resolution surface soil moisture images. Through a visual comparison, it is evident that a major portion of the coarse resolution pixels with the highest surface soil moisture coincide with the Demopolis Chalk, Mooreville Chalk, Porters Creek, Ripley, and Coffee Sand. Data presented in **Figure B 1** through **Figure B 5**, and **Table B 1** through **Table B 5** in Appendix B support this observation. The highest values for the maximum and mean surface soil moisture are observed within the same surface geologic units. A similar match between land cover types and surface soil moisture was identified when the images are compared. The highest soil moisture values are observed within the cropland vegetation area as classified by the IGBP. It could be observed that the original pixel pattern observed in the coarse resolution surface soil moisture data has been more refined to match the surface geology and land cover pattern after the downscaling.



**Figure 22.** Surface geology of the study region.

Within the downscaled image, most of the high surface soil moisture pixels coincide with the Demopolis Chalk, Porters Creek Formations, and Mooreville. This could also be observed in **Figure B 1** through **Figure B 5**, and **Table B 1** through **Table B 5** presented in Appendix B. Added to this, the highest surface soil moisture values fit well with the areas covered by

croplands according to the IGBP classification. These similar pixel patterns observed in both coarse-resolution and fine-resolution images verify the accuracy of the results.

### 5.3 Quantitative analysis of the downscaling model

Downscaling was achieved using the random forest technique in this project. In a usual machine learning approach, a data set representing the dependent variable is used as the target data for the model training (Chen et al., 2020b; Im et al., 2016; Liu et al., 2017; Zhao et al., 2018).. In this project, high resolution surface soil moisture data becomes the dependent variable. However, the only available fine scale (1 km) surface soil moisture data for this region for the time period considered were the field collected surface soil moisture data during this project's fieldwork. Since that data set was not big enough to train a machine learning model, coarse-resolution data (SPL4SMGP) was disaggregated and used in training the machine learning model. As a result, the model is led to believe the disaggregated coarse-resolution data to be fine-scale data (target data set) and ends up comparing the predicted values with the coarse-resolution data itself (target data set) to produce statistics for the model accuracy. The purpose of downscaling the coarse-resolution data is to alter the coarse-resolution values to match a finer spatial resolution, this approach leads to a lower correlation of coefficient between the predicted values and the target data set. Therefore, we cannot use the model correlation coefficient value output to quantify the prediction capability of the model. However, having similar correlation coefficient values for both validation and training confirms that the model is stable and overfitting has not occurred. During the training, we were able to observe a maximum correlation of determination of 20 percent with lowest values at 9 percent. Similar values were observed for the validation data set with differences of approximately 1–2 percent from training. The random forest model is also capable of providing the data to show the importance of each independent

variable used for downscaling. In all cases considered, NDVI was observed to be the most important parameter for the downscaling, with diurnal temperature the second and EVI the third most important parameter (**Table 5**).

Graphs shown in **Figure 16** through **Figure 20** represent the relationship observed between the surface soil moisture data collected in the field and the surface soil moisture data derived from coarse resolution (SPL4SMGP) and downscaled data. A very low correlation coefficient was observed for all the months considered, for both coarse resolution and downscaled surface soil moisture data. Also, satellite-based data was observed to be overestimating the surface soil moisture data compared to the field-collected surface soil moisture data. Thus, we are unable to validate the model performance or the accuracy of the coarse-resolution surface soil moisture data based on the field collected surface soil moisture data. It should also be noted despite having field data collected from 25 different locations, for certain months, only a portion of this data could be used due to missing pixels as a result of cloud cover.

There are several reasons that may have led to the poor relationship between surface soil moisture data collected in the field and coarse resolution/downscaled surface soil moisture data. The very fact that we are comparing point data to pixel values that are supposed to represent the average surface soil moisture for a larger spatial extent could be one of them. The heterogeneous nature of the surface soil moisture in the study region may have caused the point data to be poorly correlated to the satellite data. This assumption is supported by the field data collected in a very short time frame (less than 3 hours) for a smaller study area (approximately 1 km<sup>2</sup>). Values observed in the smaller field site have a range of 3% to 39% volumetric surface soil moisture (**Figure 11**). This is evidence that field-collected surface soil moisture data can have a

large variation even within a smaller area closer to 1 km<sup>2</sup>. Impure pixels could be another reason for the poor correlation (having spectrally different materials within the area of a pixel). The land cover and land use of the study area is highly heterogeneous. Being a highly vegetated terrain with many streams and water bodies, the pixels covering the study region may come across different surface soil moisture values for each pixel. Ultimately, this could add to the poor correlation between field data and satellite data. Ideally, collecting surface soil moisture data from several locations within an area similar to that of a pixel (1 km<sup>2</sup>) to come up with an average value for each location would be preferable. However, that was not practical within the context of this project, considering the resources available.

**Table 6** provides an overview of the statistics obtained for the downscaled surface soil moisture data and coarse-resolution surface soil moisture data. It is noticed that despite having different values for minimum, maximum, median, and mode, the Sum and Mean values are the same for both the coarse resolution and downscaled images. This indicates that the random forest model has redistributed and normalized the surface soil moisture data from the coarse-resolution image between a new maximum and minimum for the downscaled image. Furthermore, seeing that the downscaled data matches the patterns of land cover types and surface geology, the model has normalized the data to match the patterns of NDVI, EVI, and diurnal temperature used for the modeling. Since the coarse-resolution data also fits the same pixel patterns with similar relative pixel value differences across the image for all the months (based on visual observation), this normalization achieved by the random forest model could be said to be a good fit for surface soil moisture data downscaling.

#### 5.4 Relationship between the land cover types and surface soil moisture

Tables in **Appendix A** present the statistics of the downscaled and coarse resolution surface soil moisture data over different land cover types observed by IGBP classification. As per the data from these tables, land cover classes can be put into three categories based on mean surface soil moisture. Land cover classes of croplands and cropland/natural vegetation Mosaics have the highest mean surface soil moisture values. Similarly highest values for the mode were observed for Croplands and Cropland/Natural Vegetation Mosaics. The second highest mean surface soil moisture values were observed for savannas, grasslands, and urban and built-up lands. Whereas, third highest surface soil moisture values were observed for the woody savannas, deciduous broadleaf forests, and mixed forests. This is true for both coarse resolution and downscaled data except in the case of urban and urban built-up land cover types. Urban and built-up lands have lowest surface soil moisture in coarse resolution image while having high mean surface soil moisture values similar to grass lands and savannas in the downscaled images. Considering the amount of concrete structures and roads covering the soil, this can be considered an over estimation. These statistics match the visual observations explained in previous sections.

Wang et al., (2013) present very similar results where they put the mean soil moisture for five different land cover types in the order of crop> grass> subshrub> tree> shrub. This relationship of relative soil moisture difference shown in Wang et al., (2013) is described to be stable over time. Yuan et al., (2022) also confirms this relationship where the land cover types based on soil moisture is put in the order of arable land (e.g. crop lands)> grasslands> shrubland> forestland.

Considering that the order of relative differences of mean and mode surface soil moisture

across different land cover types are the same for both coarse-resolution and downscaled data, and the observed relationship matches with literature, we can assume that the random forest model has worked well in predicting fine-resolution surface soil moisture data.

## VI. CONCLUSIONS

From the study results, it is evident that the point surface soil moisture data collected through fieldwork was not sufficient for evaluating the accuracy of downscaled data that has a comparatively much larger spatial resolution of 1 km. The poor correlation seen between the point surface soil moisture data and coarse resolution/fine resolution maps could be attributed to the heterogeneous nature of the surface soil moisture.

The surface geology and the land cover types follow a very similar spatial distribution pattern, which could be observed in both the coarse resolution SPL4SMGP data and downscaled surface soil moisture data set. While using land cover types and surface geology could also improve the downscaling, we decided not to use them for the downscaling to avoid output being biased towards them. Thus we were able to use land cover types and surface geology in evaluating the downscaling results. In the region considered for the study, the highest surface soil moisture values were observed in Demopolis Chalk and Porters Creek geologic formations. When compared to the vegetation types, Croplands, and Cropland/Natural Vegetation Mosaics showed the highest surface soil moisture values. With these results, we can come to the conclusion that it is possible to use the patterns of surface geology and land cover types to evaluate downscaling results when a secondary fine-resolution dataset is not available.

Statistics computed for the pixels of coarse resolution and downscaled surface soil moisture data reveal that during the downscaling process, the random forest model has normalized the surface soil moisture data extracted from coarse-resolution images over the study

region. The random forest model has used NDVI, EVI, and diurnal temperature to do this normalization. Ultimately the results demonstrate the ability to use random forest machine learning technique to downscale surface soil moisture data for a smaller localized area.

## **LIST OF REFERENCES**

Bai, J., Cui, Q., Zhang, W., & Meng, L. (2019). An Approach for Downscaling SMAP Soil Moisture by Combining Sentinel-1 SAR and MODIS Data. *Remote Sensing*, *11*(23), 2736. <https://doi.org/10.3390/rs11232736>

Bartkowiak, P., Castelli, M., & Notarnicola, C. (2019). Downscaling Land Surface Temperature from MODIS Dataset with Random Forest Approach over Alpine Vegetated Areas. *Remote Sensing*, *11*(11), 1319. <https://doi.org/10.3390/rs11111319>

Beck, H. E., Pan, M., Miralles, D. G., Reichle, R. H., Dorigo, W. A., Hahn, S., Sheffield, J., Karthikeyan, L., Balsamo, G., Parinussa, R. M., van Dijk, A. I. J. M., Du, J., Kimball, J. S., Vergopolan, N., & Wood, E. F. (2021). Evaluation of 18 satellite- And model-based soil moisture products using in situ measurements from 826 sensors. *Hydrology and Earth System Sciences*, *25*(1), 17–40. <https://doi.org/10.5194/hess-25-17-2021>

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. <https://doi.org/10.1007/BF00058655>

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(5), 32. <https://doi.org/1010933404324>

Busch, F. A., Niemann, J. D., & Coleman, M. (2012). Evaluation of an empirical orthogonal function-based method to downscale soil moisture patterns based on topographical attributes: DOWNSCALING SOIL MOISTURE PATTERNS BASED ON TOPOGRAPHICAL ATTRIBUTES. *Hydrological Processes*, *26*(18), 2696–2709. <https://doi.org/10.1002/hyp.8363>

Carlson, T. (2007). An Overview of the “Triangle Method” for Estimating Surface Evapotranspiration and Soil Moisture from Satellite Imagery. *Sensors*, *7*(8), 1612–1629.

<https://doi.org/10.3390/s7081612>

Carlson, T. N., Gillies, R. R., & Perry, E. M. (1994). A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. *Remote Sensing Reviews*, 9(1–2), 161–173.

<https://doi.org/10.1080/02757259409532220>

Chapman, S. S., Griffith, G. E., Omernik, J. M., Comstock, J. A., Beiser, M. C., & Johnson, D. (2004). *Ecoregions of Mississippi*. Reston, Virginia, U.S. Geological Survey.

Chauhan, N. S., Miller, S., & Ardanuy, P. (2003). Spaceborne soil moisture estimation at high resolution: A microwave-optical/IR synergistic approach. *International Journal of Remote Sensing*, 24(22), 4599–4622. <https://doi.org/10.1080/0143116031000156837>

Chen, Q., Miao, F., Wang, H., Xu, Z., Tang, Z., Yang, L., & Qi, S. (2020a). Downscaling of Satellite Remote Sensing Soil Moisture Products Over the Tibetan Plateau Based on the Random Forest Algorithm: Preliminary Results. *Earth and Space Science*, 7(6).

<https://doi.org/10.1029/2020EA001265>

Chen, Q., Miao, F., Wang, H., Xu, Z., Tang, Z., Yang, L., & Qi, S. (2020b). Downscaling of Satellite Remote Sensing Soil Moisture Products Over the Tibetan Plateau Based on the Random Forest Algorithm: Preliminary Results. *Earth and Space Science*, 7(6).

<https://doi.org/10.1029/2020EA001265>

Cihlar, J., & Ulaby, F. T. (n.d.). *DIELECTRIC PROPERTIES OF SOILS AS A FUNCTION OF MOISTURE CONTENT*.

Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., De Rosnay, P., Ryu, D., & Walker, J. P. (2012). Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, 50(2), 1–20. <https://doi.org/10.1029/2011RG000372>

Curtis, J. O., Charles A. Weiss, Jr., & Everett, J. B. (1995). *Effect of Soil Composition on Complex Dielectric Properties* (Technical Report December). Waterways Experiment Station, US Army Corps of Engineers. <https://apps.dtic.mil/sti/pdfs/ADA304706.pdf>

Cutler, adele, cutler, D. R., & Stevens, J. R. (2012). Random Forests. In *Ensemble Machine Learning: Methods and Applications* (pp. 157–175). Springer New York. <https://doi.org/10.1007/978-1-4419-9326-7>

Dalsin, G. J. (1978). The Mississippi River Valley alluvial aquifer in Mississippi. In *Water-Resources Investigations Report*. <https://doi.org/10.3133/wri78106>

Daly, E., & Porporato, A. (2005). A Review of Soil Moisture Dynamics: From Rainfall Infiltration to Ecosystem Response. *Environmental Engineering Science*, 22(1), 9–24. <https://doi.org/10.1089/ees.2005.22.9>

Das, N. N., Entekhabi, D., Dunbar, R. S., Colliander, A., Chen, F., Crow, W., Jackson, T. J., Berg, A., Bosch, D. D., Caldwell, T., Cosh, M. H., Collins, C. H., Lopez-Baeza, E., Moghaddam, M., Rowlandson, T., Starks, P. J., Thibeault, M., Walker, J. P., Wu, X., ... Njoku, E. G. (2018). The SMAP mission combined active-passive soil moisture product at 9 km and 3 km spatial resolutions. *Remote Sensing of Environment*, 211(March), 204–217. <https://doi.org/10.1016/j.rse.2018.04.011>

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>

De Lannoy, G. J. M., & Reichle, R. H. (2016). Global assimilation of multiangle and multipolarization SMOS brightness temperature observations into the GEOS-5 catchment land surface model for soil moisture estimation. *Journal of Hydrometeorology*, 17(2), 669–691.

<https://doi.org/10.1175/JHM-D-15-0037.1>

Didan, K., Munoz, A. B., Solano, R., & Huete, A. (2015). *MODIS Vegetation Index User's Guide (MOD13 Series), Version 3.0 (collection 6)*. June, 38.

Dockery, D. T., & Thompson, D. E. (2016). *THE GEOLOGY OF MISSISSIPPI*. University Press of Mississippi, Jackson.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Van Oevelen, P., Robock, A., & Jackson, T. (2011). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, 15(5), 1675–1698.

<https://doi.org/10.5194/hess-15-1675-2011>

Entekhabi, B. D., Njoku, E. G., Neill, P. E. O., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., ... Zyl, J. V. (2010). ( *SMAP* ) *Mission*. 98(5).

Hutengs, C., & Vohland, M. (2016). Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sensing of Environment*, *178*, 127–141.

<https://doi.org/10.1016/j.rse.2016.03.006>

Idso, S., Schmugge, T., Jackson, R., & Reginato, R. (1975). The utility of surface temperature measurements for the remote sensing of surface soil water status. *Journal of Geophysical Research*, *80*(21), 3044–3049.

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), 1120. <https://doi.org/10.1007/s12665-016-5917-6>

Kaheil, Y. H., Gill, M. K., McKee, M., Bastidas, L. A., & Rosero, E. (2008). Downscaling and Assimilation of Surface Soil Moisture Using Ground Truth Measurements. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(5), 1375–1384.

<https://doi.org/10.1109/TGRS.2008.916086>

Kerr, Y. H. (2007). Soil moisture from space: Where are we? *Hydrogeology Journal*, *15*(1), 117–120. <https://doi.org/10.1007/s10040-006-0095-3>

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S. E., Leroux, D., Mialon, A., & Delwart, S. (2012). The SMOS Soil Moisture Retrieval Algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1384–1403. <https://doi.org/10.1109/TGRS.2012.2184548>

Liu, Y., Yang, Y., Jing, W., & Yue, X. (2017). Comparison of Different Machine Learning Approaches for Monthly Satellite-Based Soil Moisture Downscaling over Northeast

China. *Remote Sensing*, 10(2), 31. <https://doi.org/10.3390/rs10010031>

Mao, T., Shangguan, W., Li, Q., Li, L., Zhang, Y., Huang, F., Li, J., Liu, W., & Zhang, R. (2022). A Spatial Downscaling Method for Remote Sensing Soil Moisture Based on Random Forest Considering Soil Moisture Memory and Mass Conservation. *Remote Sensing*, 14(16), 3858. <https://doi.org/10.3390/rs14163858>

Merlin, O., Walker, J., Chehbouni, A., & Kerr, Y. (2008). Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency. *Remote Sensing of Environment*, 112(10), 3935–3946. <https://doi.org/10.1016/j.rse.2008.06.012>

Minnett, P. J. (2019). Satellite Remote Sensing of Sea Surface Temperatures. *Encyclopedia of Ocean Sciences*, 415–428. <https://doi.org/10.1016/B978-0-12-409548-9.04340-2>

Montzka, C., Jagdhuber, T., Horn, R., Bogena, H. R., Hajnsek, I., Reigber, A., & Vereecken, H. (2016). Investigation of SMAP Fusion Algorithms With Airborne Active and Passive L-Band Microwave Remote Sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 54(7), 3878–3889. <https://doi.org/10.1109/TGRS.2016.2529659>

Narayan, U., Lakshmi, V., & Jackson, T. J. (2006). High-resolution change estimation of soil moisture using L-band radiometer and Radar observations made during the SMEX02 experiments. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6), 1545–1554. <https://doi.org/10.1109/TGRS.2006.871199>

Njoku, E. G., & Entekhabi, D. (1996). Passive microwave remote sensing of soil moisture. *Journal of Hydrology*, 184(1–2), 101–129. <https://doi.org/10.1016/0022->

Njoku, E. G., & Kong, J. A. (1977). Theory for Passive Microwave Remote Sensing of Near-Surface Soil Moisture. *J Geophys Res*, 82(20), 3108–3118.

<https://doi.org/10.1029/JB082i020p03108>

Njoku, E. G., Wilson, W. J., Yueh, S. H., Dinardo, S. J., Li, F. K., Jackson, T. J., Lakshmi, V., & Bolten, J. (2002). Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99. *IEEE Transactions on Geoscience and Remote Sensing*, 40(12), 2659–2673. <https://doi.org/10.1109/TGRS.2002.807008>

O'Neill, P., Chan, S., Bindlish, R., Jackson, T., Colliander, A., Dunbar, S., Chen, F., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., ... Kerr, Y. (2017). Assessment of version 4 of the SMAP passive soil moisture standard product. *International Geoscience and Remote Sensing Symposium (IGARSS), 2017-July(8)*, 3941–3944.

<https://doi.org/10.1109/IGARSS.2017.8127862>

Parinussa, R. M., Holmes, T. R. H., Wanders, N., Dorigo, W. A., & de Jeu, R. A. M. (2015). A Preliminary Study toward Consistent Soil Moisture from AMSR2. *Journal of Hydrometeorology*, 16(2), 932–947. <https://doi.org/10.1175/JHM-D-13-0200.1>

Patel, L., Kumar, S. R., Pandey, A., & Shrivastava, A. K. (2018). *Role of Moisture Content and Dielectric Constant in Soil*.

Peng, J., Loew, A., Merlin, O., & Verhoest, N. E. C. (2017). A review of spatial downscaling of satellite remotely sensed soil moisture. *Reviews of Geophysics*, 55(2), 341–366.

<https://doi.org/10.1002/2016RG000543>

Perry, M. A., & Niemann, J. D. (2007). Analysis and estimation of soil moisture at the catchment scale using EOFs. *Journal of Hydrology*, 334(3–4), 388–404.

<https://doi.org/10.1016/j.jhydrol.2006.10.014>

Piles, M., Camps, A., Vall-llossera, M., Corbella, I., Panciera, R., Rudiger, C., Kerr, Y. H., & Walker, J. (2011). Downscaling SMOS-Derived Soil Moisture Using MODIS Visible/Infrared Data. *IEEE Transactions on Geoscience and Remote Sensing*, 49(9), 3156–3166. <https://doi.org/10.1109/TGRS.2011.2120615>

Quan, C., Jiuli, L., Zhihua, T., Jiangyuan, Z., & Yan, L. (2014). Study on the relationship between soil moisture and its dielectric constant obtained by space-borne microwave radiometers and scatterometers. *IOP Conference Series: Earth and Environmental Science*, 17, 012143. <https://doi.org/10.1088/1755-1315/17/1/012143>

Reginato, R., Idso, S., Vedder, J., Jackson, R., Blanchard, M., & Goettelman, R. (1976). Soil water content and evaporation determined by thermal parameters obtained from ground-based and remote measurements. *Journal of Geophysical Research*, 81(9), 1617–1620.

Reichle, R. H., Lannoy, G. D., Koster, R. D., Crow, W. T., Kimball, J. S., Liu, Q., & Bechtold, M. (2022). *SMAP L4 Global 9 km EASE-Grid Surface and Root Zone Soil Moisture , Version 7*. NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/EVKPQZ4AFC4D>

Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J. M., Kimball, J. S., Ardizzone, J. V., Bosch, D., Colliander, A., Cosh, M., Kolassa, J., Mahanama, S. P., Prueger, J.,

Starks, P., & Walker, J. P. (2019). Version 4 of the SMAP Level-4 Soil Moisture Algorithm and Data Product. *Journal of Advances in Modeling Earth Systems*, 11(10), 3106–3130.

<https://doi.org/10.1029/2019MS001729>

Robinson, D. A., Campbell, C. S., Hopmans, J. W., Hornbuckle, B. K., Jones, S. B., Knight, R., Ogden, F., Selker, J., & Wendroth, O. (2008). Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review. *Vadose Zone Journal*, 7(1), 358–389. <https://doi.org/10.2136/vzj2007.0143>

Runkle, J., Kunkel, K. E., Champion, S. M., Frankson, R., Stewart, B. C., & Nielsen-Gammon, J. (2022). Mississippi State Climate Summary 2022. In *NOAA Technical Report NESDIS 150-MS*. NOAA/NESDIS, Silver Spring, MD.

<https://statesummaries.ncics.org/chapter/ms/>

Schmugge, T. (1978). Remote Sensing of Surface Soil Moisture. *Journal of Applied Meteorology*, 17(10), 1549–1557. [https://doi.org/10.1175/1520-0450\(1978\)017<1549:RSOSSM>2.0.CO;2](https://doi.org/10.1175/1520-0450(1978)017<1549:RSOSSM>2.0.CO;2)

Schmugge, T. (1982). Remote Sensing of Surface Soil Moisture. *Journal of Applied Meteorology*, 17(10), 1549–1557.

Schmugge, T., O’neill, P. E., & Wanga, J. R. (1986). Passive Microwave Soil Moisture Research. *IEEE Transactions on Geoscience and Remote Sensing*, GE-24(1), 12–22.

<https://doi.org/10.1109/TGRS.1986.289584>

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing

climate: A review. *Earth-Science Reviews*, 99(3), 125–161.

<https://doi.org/10.1016/j.earscirev.2010.02.004>

*Spectrum Technologies*. (2022). SM100 Soil moisture sensor product manual, Spectrum Technologies. Plainfield, Illinois.

[https://www.especmic.co.jp/download/manual/WatchDog/6460\\_SM1001.pdf](https://www.especmic.co.jp/download/manual/WatchDog/6460_SM1001.pdf)

Susha Lekshmi, S. U., Singh, D. N., & Shojaei Baghini, M. (2014). A critical review of soil moisture measurement. *Measurement*, 54, 92–105.

<https://doi.org/10.1016/j.measurement.2014.04.007>

Tang, Q., Feng, G., Fisher, D., Zhang, H., Ouyang, Y., Adeli, A., & Jenkins, J. (2018). Rain water deficit and irrigation demand of major row crops in the Mississippi delta.

*Transactions of the ASABE*, 61(3), 927–935. <https://doi.org/10.13031/trans.12397>

Vereecken, H., Huisman, J. A., Bogaen, H., Vanderborght, J., Vrugt, J. A., & Hopmans, J. W. (2008). On the value of soil moisture measurements in vadose zone hydrology: A review: SOIL MOISTURE AND HYDROLOGY. *Water Resources Research*, 44(4).

<https://doi.org/10.1029/2008WR006829>

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., ... Steinnocher, K. (2013). The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications. *Meteorologische Zeitschrift*, 22(1), 5–33. <https://doi.org/10.1127/0941-2948/2013/0399>

- Wan, Z. (2006). *MODIS Land Surface Temperature Products Users' Guide* (pp. 1–33). ICES, University of California, Santa Barbara. <http://www.oalib.com/references/9184061>
- Wang, J. R. (1980). The dielectric properties of soil-water mixtures at microwave frequencies. *Radio Science*, *15*(5), 977–985. <https://doi.org/10.1029/RS015i005p00977>
- Wang, S., Fu, B., Gao, G., Liu, Y., & Zhou, J. (2013). Responses of soil moisture in different land cover types to rainfall events in a re-vegetation catchment area of the Loess Plateau, China. *CATENA*, *101*, 122–128. <https://doi.org/10.1016/j.catena.2012.10.006>
- Wang, T., Wua, D., & Zhanc, C. (2020). Comparison of environmental controls on soil moisture spatial patterns at mesoscales: Observational evidence from two regions in China. *Geoderma*, *374*. <https://doi.org/10.1016/j.geoderma.2020.114451>
- Wasson, Billie. E. (1986). Sources for water supplies in Mississippi. *Mississippi Research and Development Center*.
- Wigneron, J. P., Schmugge, T., Chanzy, A., Calvet, J. C., & Kerr, Y. (1998). Use of passive microwave remote sensing to monitor soil moisture. *Agronomie*, *18*(1), 27–43. <https://doi.org/10.1051/agro:19980102>
- Yuan, E., Zhou, Q., Yan, W., Peng, D., & Wang, Y. (2022). Response of the winter soil moisture of different vegetation types to rainfall events in karst slope land. *Hydrology Research*, *53*(10), 1271–1285. <https://doi.org/10.2166/nh.2022.033>
- Zhang, H., Wang, S., Liu, K., Li, X., Li, Z., Zhang, X., & Liu, B. (2022). Downscaling of AMSR-E Soil Moisture over North China Using Random Forest Regression. *ISPRS*

*International Journal of Geo-Information*, 11(2), 101. <https://doi.org/10.3390/ijgi11020101>

Zhao, W., Sánchez, N., Lu, H., & Li, A. (2018). A spatial downscaling approach for the SMAP passive surface soil moisture product using random forest regression. *Journal of Hydrology*, 563, 1009–1024. <https://doi.org/10.1016/j.jhydrol.2018.06.081>

## **APPENDICES**

APPENDIX A: STATISTICS OF SURFACE SOIL MOISTURE VALUES FOR DIFFERENT  
LAND COVER TYPES IN THE STUDY REGION ACCORDING TO IGBP

**Table A 1.** Statistics for different land cover types in the study region according to IGBP for the study date in February. Soil moisture values are given as volume/volume.

| February                                   | Field Name | Minimum | Maximum | Mean  | Median | Count | Mode     | Sum       |
|--|------------|---------|---------|-------|--------|-------|----------|-----------|
| <b>Deciduous Broadleaf Forests</b>         | SM 9km     | 0.241   | 0.500   | 0.348 | 0.346  | 25816 | 0.346    | 8990.622  |
|  | SM 1km     | 0.324   | 0.378   | 0.349 | 0.350  | 25816 | 0.340    | 9022.596  |
| <b>Mixed Forests</b>                       | SM 9km     | 0.241   | 0.500   | 0.335 | 0.325  | 10539 | 0.309    | 3527.903  |
|  | SM 1km     | 0.327   | 0.375   | 0.345 | 0.346  | 10539 | 0.340    | 3632.258  |
| <b>Woody Savannas</b>                      | SM 9km     | 0.241   | 0.500   | 0.354 | 0.350  | 41705 | 0.358    | 14772.225 |
|  | SM 1km     | 0.324   | 0.381   | 0.354 | 0.354  | 41705 | Multiple | 14753.413 |
| <b>Savannas</b>                            | SM 9km     | 0.241   | 0.500   | 0.368 | 0.364  | 7655  | 0.367    | 2816.940  |
|  | SM 1km     | 0.328   | 0.381   | 0.361 | 0.363  | 7655  | 0.363    | 2766.477  |
| <b>Grasslands</b>                          | SM 9km     | 0.253   | 0.485   | 0.362 | 0.356  | 1704  | 0.277    | 616.489   |
|  | SM 1km     | 0.324   | 0.381   | 0.361 | 0.363  | 1704  | Multiple | 615.177   |
| <b>Croplands</b>                           | SM 9km     | 0.259   | 0.482   | 0.383 | 0.392  | 3004  | 0.415    | 1151.368  |
|  | SM 1km     | 0.331   | 0.381   | 0.364 | 0.364  | 3004  | 0.360    | 1094.326  |
| <b>Urban and Built-up Lands</b>            | SM 9km     | 0.241   | 0.471   | 0.333 | 0.329  | 589   | 0.329    | 195.928   |
|  | SM 1km     | 0.328   | 0.376   | 0.359 | 0.361  | 589   | 0.352    | 211.438   |
| <b>Cropland/Natural Vegetation Mosaics</b> | SM 9km     | 0.273   | 0.485   | 0.383 | 0.392  | 1941  | 0.401    | 743.479   |
|  | SM 1km     | 0.331   | 0.380   | 0.364 | 0.365  | 1941  | 0.360    | 706.636   |

**Table A 2.** Statistics for different land cover types in the study region according to IGBP for the study date in March. Soil moisture values are given as volume/volume.

| March                                      | Field Name | Minimum | Maximum | Mean  | Median | Count | Mode  | Sum       |
|--|------------|---------|---------|-------|--------|-------|-------|-----------|
| <b>Deciduous Broadleaf Forests</b>         | SM 9km     | 0.272   | 0.502   | 0.367 | 0.360  | 28124 | 0.356 | 10311.782 |
|  | SM 1km     | 0.353   | 0.396   | 0.372 | 0.371  | 28124 | 0.355 | 10473.318 |
| <b>Mixed Forests</b>                       | SM 9km     | 0.272   | 0.502   | 0.363 | 0.352  | 11717 | 0.333 | 4255.474  |
|  | SM 1km     | 0.352   | 0.397   | 0.368 | 0.368  | 11717 | 0.368 | 4316.212  |
| <b>Woody Savannas</b>                      | SM 9km     | 0.272   | 0.502   | 0.375 | 0.363  | 48109 | 0.333 | 18053.470 |
|  | SM 1km     | 0.353   | 0.400   | 0.376 | 0.375  | 48109 | 0.355 | 18082.089 |
| <b>Savannas</b>                            | SM 9km     | 0.272   | 0.502   | 0.389 | 0.389  | 9679  | 0.391 | 3763.641  |
|  | SM 1km     | 0.354   | 0.397   | 0.381 | 0.382  | 9679  | 0.382 | 3689.393  |
| <b>Grasslands</b>                          | SM 9km     | 0.284   | 0.496   | 0.385 | 0.381  | 2102  | 0.403 | 808.827   |
|  | SM 1km     | 0.354   | 0.400   | 0.380 | 0.383  | 2102  | 0.355 | 799.717   |
| <b>Croplands</b>                           | SM 9km     | 0.284   | 0.494   | 0.406 | 0.416  | 4432  | 0.439 | 1801.576  |
|  | SM 1km     | 0.354   | 0.400   | 0.384 | 0.386  | 4432  | 0.382 | 1702.986  |
| <b>Urban and Built-up Lands</b>            | SM 9km     | 0.272   | 0.491   | 0.358 | 0.347  | 627   | 0.347 | 224.602   |
|  | SM 1km     | 0.355   | 0.395   | 0.380 | 0.381  | 627   | 0.365 | 238.421   |
| <b>Cropland/Natural Vegetation Mosaics</b> | SM 9km     | 0.284   | 0.496   | 0.408 | 0.414  | 2692  | 0.442 | 1098.061  |
|  | SM 1km     | 0.357   | 0.397   | 0.384 | 0.385  | 2692  | 0.386 | 1032.462  |

**Table A 3.** Statistics for different land cover types in the study region according to IGBP for the study date in May. Soil moisture values are given as volume/volume.

| May  | Field Name | Minimum | Maximum | Mean  | Median | Count | Mode     | Sum      |
|--|------------|---------|---------|-------|--------|-------|----------|----------|
| <b>Deciduous Broadleaf Forests</b>         | SM 9km     | 0.151   | 0.446   | 0.295 | 0.290  | 16929 | 0.280    | 4999.402 |
|  | SM 1km     | 0.286   | 0.325   | 0.298 | 0.296  | 16929 | 0.292    | 5050.663 |
| <b>Mixed Forests</b>                       | SM 9km     | 0.171   | 0.446   | 0.293 | 0.286  | 5152  | 0.266    | 1511.568 |
|  | SM 1km     | 0.288   | 0.324   | 0.301 | 0.299  | 5152  | 0.297    | 1548.753 |
| <b>Woody Savannas</b>                      | SM 9km     | 0.151   | 0.446   | 0.305 | 0.295  | 26198 | 0.289    | 7993.558 |
|  | SM 1km     | 0.288   | 0.325   | 0.303 | 0.303  | 26198 | 0.291    | 7947.283 |
| <b>Savannas</b>                            | SM 9km     | 0.167   | 0.446   | 0.316 | 0.317  | 5465  | 0.328    | 1729.598 |
|  | SM 1km     | 0.288   | 0.325   | 0.309 | 0.311  | 5465  | Multiple | 1688.731 |
| <b>Grasslands</b>                          | SM 9km     | 0.167   | 0.437   | 0.320 | 0.314  | 936   | 0.431    | 299.841  |
|  | SM 1km     | 0.289   | 0.324   | 0.313 | 0.314  | 936   | Multiple | 292.510  |
| <b>Croplands</b>                           | SM 9km     | 0.167   | 0.437   | 0.324 | 0.341  | 2193  | 0.356    | 710.050  |
|  | SM 1km     | 0.291   | 0.327   | 0.316 | 0.318  | 2193  | 0.320    | 692.702  |
| <b>Urban and Built-up Lands</b>            | SM 9km     | 0.151   | 0.425   | 0.267 | 0.277  | 366   | 0.226    | 97.828   |
|  | SM 1km     | 0.289   | 0.322   | 0.310 | 0.312  | 366   | 0.306    | 113.613  |
| <b>Cropland/Natural Vegetation Mosaics</b> | SM 9km     | 0.167   | 0.437   | 0.317 | 0.326  | 1627  | 0.167    | 515.951  |
|  | SM 1km     | 0.290   | 0.325   | 0.312 | 0.313  | 1627  | Multiple | 508.132  |

**Table A 4.** Statistics for different land cover types in the study region according to IGBP for the study date in September. Soil moisture values are given as volume/volume.

| September                                  | Field Name | Minimum | Maximum | Mean  | Median | Count | Mode  | Sum      |
|--|------------|---------|---------|-------|--------|-------|-------|----------|
| <b>Deciduous Broadleaf Forests</b>         | SM 9km     | 0.068   | 0.254   | 0.131 | 0.034  | 26598 | 0.099 | 3488.435 |
|  | SM 1km     | 0.127   | 0.162   | 0.137 | 0.006  | 26598 | 0.129 | 3635.428 |
| <b>Mixed Forests</b>                       | SM 9km     | 0.076   | 0.254   | 0.147 | 0.035  | 11015 | 0.161 | 1619.966 |
|  | SM 1km     | 0.127   | 0.160   | 0.140 | 0.006  | 11015 | 0.134 | 1538.758 |
| <b>Woody Savannas</b>                      | SM 9km     | 0.068   | 0.254   | 0.142 | 0.039  | 47306 | 0.131 | 6703.173 |
|  | SM 1km     | 0.127   | 0.169   | 0.142 | 0.005  | 47306 | 0.135 | 6724.350 |
| <b>Savannas</b>                            | SM 9km     | 0.068   | 0.254   | 0.145 | 0.045  | 9678  | 0.211 | 1403.074 |
|  | SM 1km     | 0.128   | 0.170   | 0.145 | 0.004  | 9678  | 0.145 | 1408.041 |
| <b>Grasslands</b>                          | SM 9km     | 0.068   | 0.251   | 0.148 | 0.043  | 2073  | 0.182 | 307.391  |
|  | SM 1km     | 0.128   | 0.168   | 0.146 | 0.005  | 2073  | 0.139 | 302.822  |
| <b>Croplands</b>                           | SM 9km     | 0.068   | 0.251   | 0.163 | 0.043  | 4426  | 0.198 | 721.151  |
|  | SM 1km     | 0.130   | 0.172   | 0.149 | 0.006  | 4426  | 0.149 | 658.741  |
| <b>Urban and Built-up Lands</b>            | SM 9km     | 0.077   | 0.220   | 0.125 | 0.043  | 625   | 0.082 | 78.092   |
|  | SM 1km     | 0.135   | 0.150   | 0.146 | 0.002  | 625   | 0.144 | 90.941   |
| <b>Cropland/Natural Vegetation Mosaics</b> | SM 9km     | 0.068   | 0.233   | 0.157 | 0.048  | 2694  | 0.200 | 422.220  |
|  | SM 1km     | 0.130   | 0.170   | 0.149 | 0.005  | 2694  | 0.145 | 400.208  |

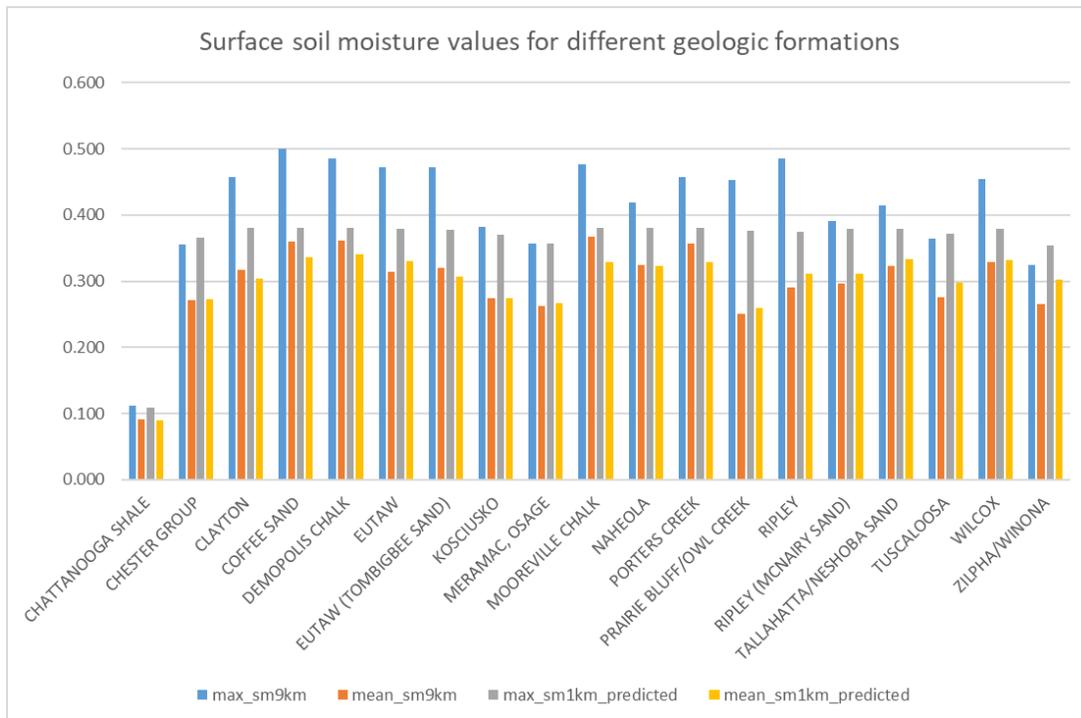
**Table A 5.** Statistics for different land cover types in the study region according to IGBP for the study date in October. Soil moisture values are given as volume/volume.

| October                                    | Field Name | Minimum | Maximum | Mean  | Median | Count | Mode     | Sum      |
|--|------------|---------|---------|-------|--------|-------|----------|----------|
| <b>Deciduous Broadleaf Forests</b>         | SM 9km     | 0.076   | 0.263   | 0.142 | 0.140  | 27238 | 0.147    | 3878.385 |
|  | SM 1km     | 0.136   | 0.176   | 0.142 | 0.141  | 27238 | Multiple | 3863.085 |
| <b>Mixed Forests</b>                       | SM 9km     | 0.078   | 0.258   | 0.135 | 0.121  | 11554 | 0.105    | 1556.046 |
|  | SM 1km     | 0.136   | 0.176   | 0.140 | 0.139  | 11554 | 0.137    | 1620.586 |
| <b>Woody Savannas</b>                      | SM 9km     | 0.076   | 0.263   | 0.147 | 0.144  | 46496 | 0.112    | 6833.569 |
|  | SM 1km     | 0.136   | 0.179   | 0.147 | 0.146  | 46496 | 0.137    | 6838.579 |
| <b>Savannas</b>                            | SM 9km     | 0.076   | 0.263   | 0.157 | 0.150  | 9171  | 0.200    | 1441.330 |
|  | SM 1km     | 0.136   | 0.179   | 0.154 | 0.153  | 9171  | 0.168    | 1411.865 |
| <b>Grasslands</b>                          | SM 9km     | 0.080   | 0.263   | 0.152 | 0.149  | 1941  | 0.194    | 294.575  |
|  | SM 1km     | 0.136   | 0.179   | 0.158 | 0.156  | 1941  | 0.179    | 305.829  |
| <b>Croplands</b>                           | SM 9km     | 0.080   | 0.258   | 0.169 | 0.175  | 3615  | 0.201    | 610.330  |
|  | SM 1km     | 0.136   | 0.179   | 0.165 | 0.166  | 3615  | 0.179    | 595.366  |
| <b>Urban and Built-up Lands</b>            | SM 9km     | 0.076   | 0.251   | 0.131 | 0.116  | 576   | 0.116    | 75.248   |
|  | SM 1km     | 0.138   | 0.172   | 0.153 | 0.154  | 576   | 0.155    | 88.206   |
| <b>Cropland/Natural Vegetation Mosaics</b> | SM 9km     | 0.080   | 0.263   | 0.171 | 0.172  | 2259  | 0.209    | 386.005  |
|  | SM 1km     | 0.137   | 0.179   | 0.161 | 0.158  | 2259  | 0.176    | 363.597  |

APPENDIX B: STATISTICS OF SURFACE SOIL MOISTURE VALUES FOR DIFFERENT  
SURFACE GEOLOGIC UNITS IN THE STUDY REGION

**Table B 1.** Statistics for different surface geologic units in the study region for the study date in February. Soil moisture values are given as volume/volume.

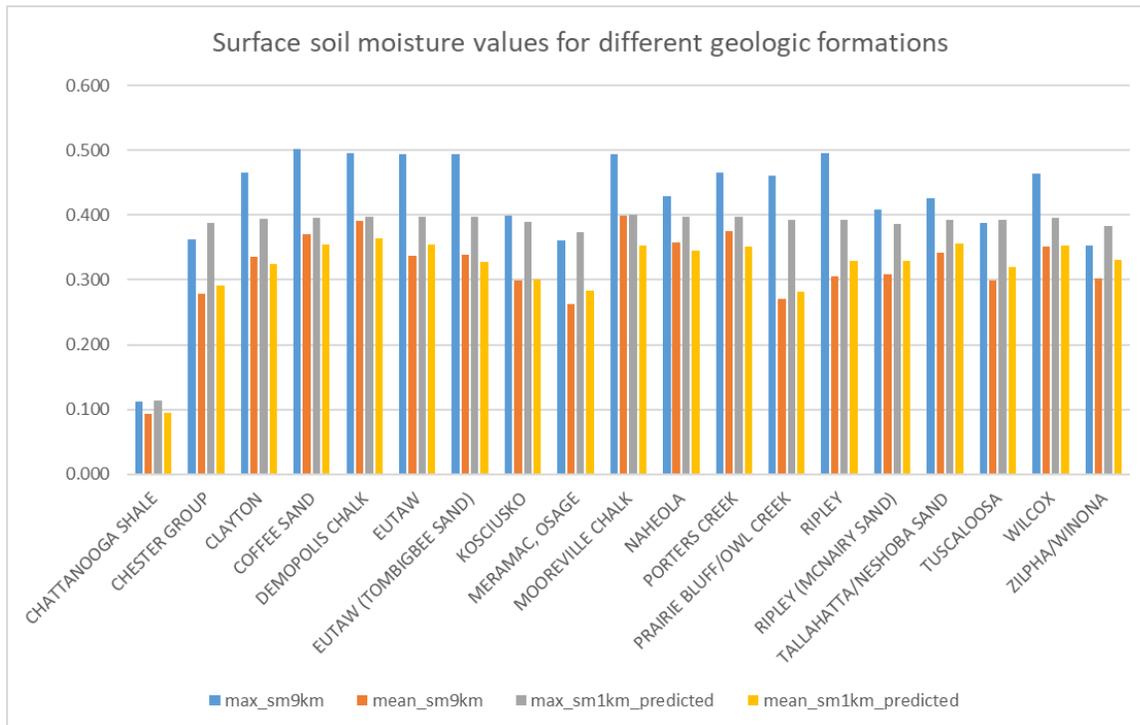
| February | FORMATION               | max_sm9k<br>m | mean_sm9k<br>m | max_sm1km<br>_predicted | mean_sm1km_<br>predicted | sum_Area_SQUA<br>REKILOMETERS |
|----------|-------------------------|---------------|----------------|-------------------------|--------------------------|-------------------------------|
| 1        | CHATTANOOGA SHALE       | 0.112         | 0.092          | 0.109                   | 0.090                    | 0.4                           |
| 2        | CHESTER GROUP           | 0.355         | 0.271          | 0.365                   | 0.272                    | 35.7                          |
| 3        | CLAYTON                 | 0.458         | 0.317          | 0.380                   | 0.304                    | 671.6                         |
| 4        | COFFEE SAND             | 0.500         | 0.360          | 0.380                   | 0.336                    | 1223.2                        |
| 5        | DEMOPOLIS CHALK         | 0.485         | 0.362          | 0.380                   | 0.340                    | 1765.3                        |
| 6        | EUTAW                   | 0.472         | 0.314          | 0.379                   | 0.330                    | 2443.3                        |
| 7        | EUTAW (TOMBIGBEE SAND)  | 0.472         | 0.319          | 0.377                   | 0.306                    | 713.9                         |
| 8        | KOSCIUSKO               | 0.382         | 0.275          | 0.371                   | 0.274                    | 109.8                         |
| 9        | MERAMAC, OSAGE          | 0.357         | 0.262          | 0.357                   | 0.267                    | 27.3                          |
| 10       | MOOREVILLE CHALK        | 0.476         | 0.368          | 0.381                   | 0.328                    | 546.5                         |
| 11       | NAHEOLA                 | 0.418         | 0.324          | 0.380                   | 0.322                    | 504.7                         |
| 12       | PORTERS CREEK           | 0.458         | 0.356          | 0.380                   | 0.329                    | 1217.8                        |
| 13       | PRAIRIE BLUFF/OWL CREEK | 0.453         | 0.251          | 0.376                   | 0.259                    | 468.6                         |
| 14       | RIPLEY                  | 0.485         | 0.291          | 0.375                   | 0.311                    | 1245.7                        |
| 15       | RIPLEY (MCNAIRY SAND)   | 0.391         | 0.297          | 0.379                   | 0.311                    | 406.7                         |
| 16       | TALLAHATTA/NESHOBA SAND | 0.414         | 0.322          | 0.379                   | 0.333                    | 2212.0                        |
| 17       | TUSCALOOSA              | 0.364         | 0.276          | 0.372                   | 0.298                    | 686.0                         |
| 18       | WILCOX                  | 0.454         | 0.329          | 0.380                   | 0.332                    | 3139.8                        |
| 19       | ZILPHA/WINONA           | 0.324         | 0.266          | 0.354                   | 0.302                    | 21.8                          |



**Figure B 1.** Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in February.

**Table B 2.** Statistics for different surface geologic units in the study region for the study date in March. Soil moisture values are given as volume/volume.

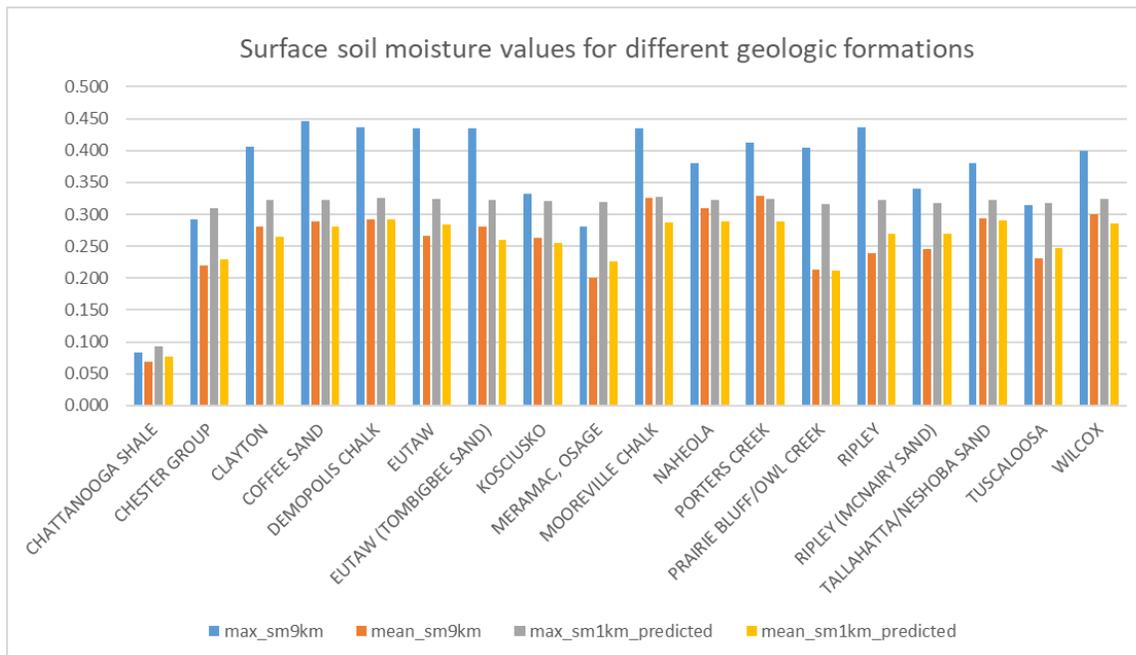
| March | FORMATION               | max_sm9km | mean_sm9km | max_sm1km_predicted | mean_sm1km_predicted | sum_Area_SQUA REKILOMETERS |
|-------|-------------------------|-----------|------------|---------------------|----------------------|----------------------------|
| 1     | CHATTANOOGA SHALE       | 0.113     | 0.093      | 0.114               | 0.094                | 0.4                        |
| 2     | CHESTER GROUP           | 0.362     | 0.278      | 0.387               | 0.291                | 36.6                       |
| 3     | CLAYTON                 | 0.466     | 0.336      | 0.394               | 0.324                | 805.4                      |
| 4     | COFFEE SAND             | 0.502     | 0.370      | 0.396               | 0.355                | 1293.3                     |
| 5     | DEMOPOLIS CHALK         | 0.496     | 0.392      | 0.397               | 0.365                | 2490.1                     |
| 6     | EUTAW                   | 0.494     | 0.337      | 0.397               | 0.354                | 2711.9                     |
| 7     | EUTAW (TOMBIGBEE SAND)  | 0.494     | 0.339      | 0.397               | 0.327                | 779.2                      |
| 8     | KOSCIUSKO               | 0.399     | 0.300      | 0.390               | 0.301                | 149.1                      |
| 9     | MERAMAC, OSAGE          | 0.362     | 0.263      | 0.374               | 0.283                | 27.3                       |
| 10    | MOOREVILLE CHALK        | 0.494     | 0.400      | 0.400               | 0.353                | 709.0                      |
| 11    | NAHEOLA                 | 0.429     | 0.358      | 0.397               | 0.345                | 714.3                      |
| 12    | PORTERS CREEK           | 0.466     | 0.376      | 0.397               | 0.352                | 1640.1                     |
| 13    | PRAIRIE BLUFF/OWL CREEK | 0.461     | 0.271      | 0.392               | 0.282                | 546.9                      |
| 14    | RIPLEY                  | 0.496     | 0.306      | 0.393               | 0.329                | 1321.2                     |
| 15    | RIPLEY (MCNAIRY SAND)   | 0.409     | 0.309      | 0.387               | 0.330                | 414.0                      |
| 16    | TALLAHATTA/NESHOBA SAND | 0.425     | 0.341      | 0.393               | 0.357                | 2973.6                     |
| 17    | TUSCALOOSA              | 0.388     | 0.300      | 0.393               | 0.320                | 770.0                      |
| 18    | WILCOX                  | 0.465     | 0.351      | 0.396               | 0.354                | 4013.9                     |
| 19    | ZILPHA/WINONA           | 0.352     | 0.302      | 0.383               | 0.331                | 38.2                       |



**Figure B 2.** Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in February.

**Table B 3.** Statistics for different surface geologic units in the study region for the study date in May. Soil moisture values are given as volume/volume.

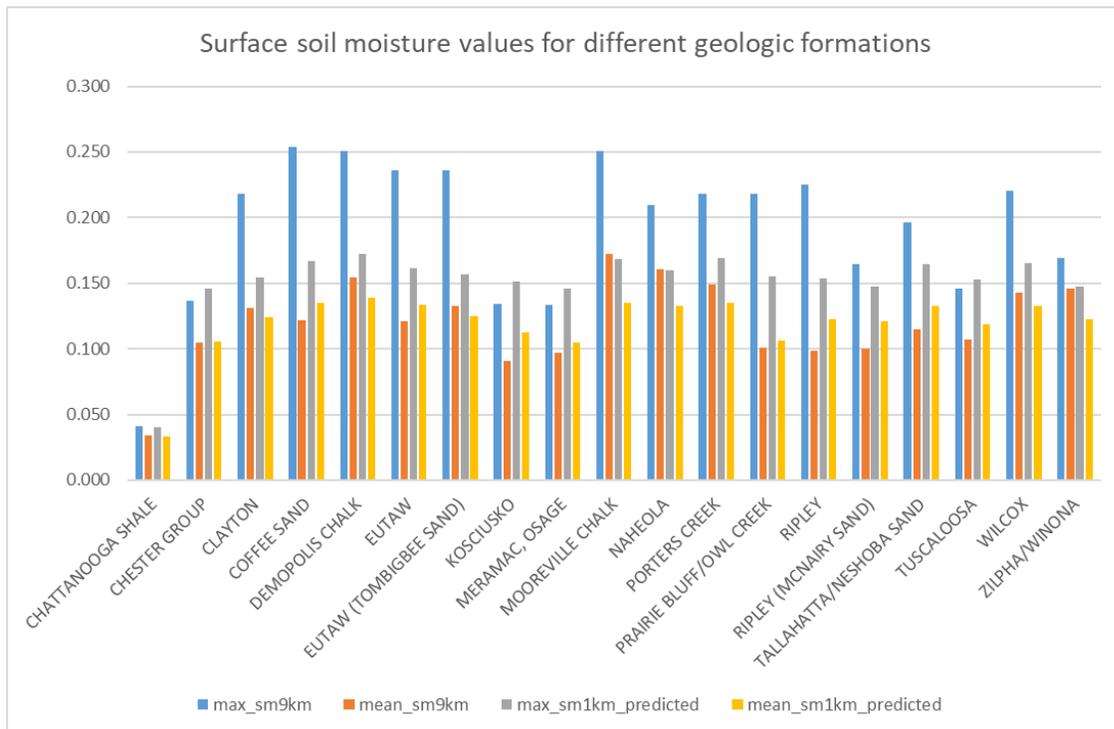
| May | FORMATION               | max_sm9km | mean_sm9km | max_sm1km_predicted | mean_sm1km_predicted | sum_Area_SQUA REKILOMETERS |
|-----|-------------------------|-----------|------------|---------------------|----------------------|----------------------------|
| 1   | CHATTANOOGA SHALE       | 0.083     | 0.068      | 0.092               | 0.076                | 0.4                        |
| 2   | CHESTER GROUP           | 0.292     | 0.220      | 0.310               | 0.229                | 36.6                       |
| 3   | CLAYTON                 | 0.405     | 0.281      | 0.323               | 0.264                | 448.2                      |
| 4   | COFFEE SAND             | 0.446     | 0.288      | 0.323               | 0.280                | 665.7                      |
| 5   | DEMOPOLIS CHALK         | 0.437     | 0.292      | 0.325               | 0.292                | 1192.3                     |
| 6   | EUTAW                   | 0.435     | 0.266      | 0.325               | 0.283                | 2053.6                     |
| 7   | EUTAW (TOMBIGBEE SAND)  | 0.435     | 0.280      | 0.323               | 0.260                | 564.3                      |
| 8   | KOSCIUSKO               | 0.331     | 0.263      | 0.321               | 0.255                | 74.0                       |
| 9   | MERAMAC, OSAGE          | 0.282     | 0.201      | 0.319               | 0.227                | 27.3                       |
| 10  | MOOREVILLE CHALK        | 0.435     | 0.326      | 0.327               | 0.287                | 566.0                      |
| 11  | NAHEOLA                 | 0.381     | 0.309      | 0.323               | 0.288                | 130.8                      |
| 12  | PORTERS CREEK           | 0.413     | 0.328      | 0.324               | 0.288                | 721.5                      |
| 13  | PRAIRIE BLUFF/OWL CREEK | 0.404     | 0.213      | 0.316               | 0.212                | 258.5                      |
| 14  | RIPLEY                  | 0.437     | 0.238      | 0.323               | 0.269                | 852.9                      |
| 15  | RIPLEY (MCNAIRY SAND)   | 0.340     | 0.245      | 0.318               | 0.269                | 304.6                      |
| 16  | TALLAHATTA/NESHOBA SAND | 0.380     | 0.294      | 0.323               | 0.291                | 1308.3                     |
| 17  | TUSCALOOSA              | 0.314     | 0.231      | 0.318               | 0.247                | 470.1                      |
| 18  | WILCOX                  | 0.400     | 0.300      | 0.325               | 0.285                | 987.3                      |
| 19  | ZILPHA/WINONA           |           |            |                     |                      | 0.0                        |



**Figure B 3.** Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in May.

**Table B 4.** Statistics for different surface geologic units in the study region for the study date in September. Soil moisture values are given as volume/volume.

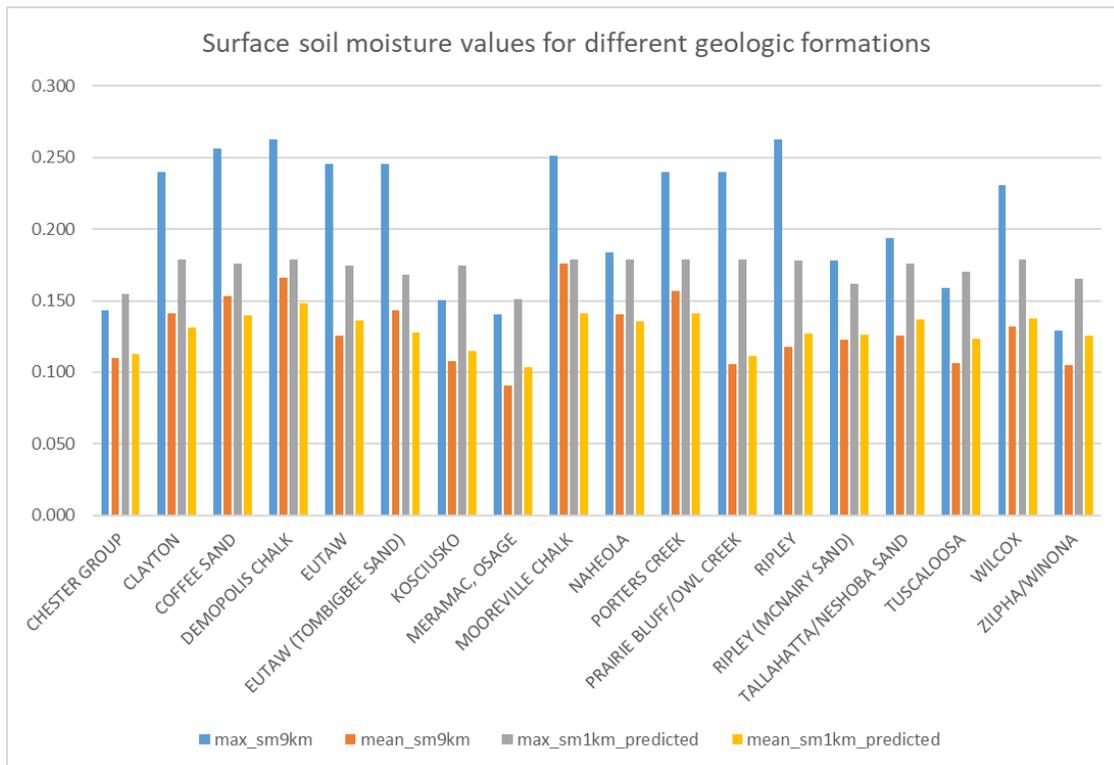
| September | FORMATION               | max_sm9km | mean_sm9km | max_sm1km_predicted | mean_sm1km_predicted | sum_Area_SQUA REKILOMETERS |
|-----------|-------------------------|-----------|------------|---------------------|----------------------|----------------------------|
| 1         | CHATTANOOGA SHALE       | 0.041     | 0.034      | 0.041               | 0.033                | 0.4                        |
| 2         | CHESTER GROUP           | 0.137     | 0.105      | 0.146               | 0.106                | 33.5                       |
| 3         | CLAYTON                 | 0.218     | 0.131      | 0.154               | 0.125                | 802.9                      |
| 4         | COFFEE SAND             | 0.254     | 0.122      | 0.167               | 0.135                | 1191.5                     |
| 5         | DEMOPOLIS CHALK         | 0.251     | 0.154      | 0.172               | 0.139                | 2467.3                     |
| 6         | EUTAW                   | 0.236     | 0.121      | 0.162               | 0.134                | 2351.1                     |
| 7         | EUTAW (TOMBIGBEE SAND)  | 0.236     | 0.133      | 0.157               | 0.125                | 698.9                      |
| 8         | KOSCIUSKO               | 0.134     | 0.091      | 0.151               | 0.113                | 149.8                      |
| 9         | MERAMAC, OSAGE          | 0.133     | 0.097      | 0.146               | 0.105                | 27.3                       |
| 10        | MOOREVILLE CHALK        | 0.251     | 0.173      | 0.168               | 0.135                | 704.0                      |
| 11        | NAHEOLA                 | 0.209     | 0.161      | 0.160               | 0.132                | 674.6                      |
| 12        | PORTERS CREEK           | 0.218     | 0.149      | 0.169               | 0.135                | 1616.9                     |
| 13        | PRAIRIE BLUFF/OWL CREEK | 0.218     | 0.101      | 0.156               | 0.107                | 541.0                      |
| 14        | RIPLEY                  | 0.225     | 0.098      | 0.154               | 0.123                | 1307.9                     |
| 15        | RIPLEY (MCNAIRY SAND)   | 0.164     | 0.100      | 0.147               | 0.121                | 384.2                      |
| 16        | TALLAHATTA/NESHOBA SAND | 0.196     | 0.115      | 0.165               | 0.132                | 2953.7                     |
| 17        | TUSCALOOSA              | 0.146     | 0.107      | 0.153               | 0.119                | 631.7                      |
| 18        | WILCOX                  | 0.220     | 0.143      | 0.165               | 0.133                | 3921.4                     |
| 19        | ZILPHA/WINONA           | 0.169     | 0.146      | 0.147               | 0.123                | 38.2                       |



**Figure B 4.** Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in September.

**Table B 5.** Statistics for different surface geologic units in the study region for the study date in October. Soil moisture values are given as volume/volume.

| October | FORMATION               | max_sm9km | mean_sm9km | max_sm1km_predicted | mean_sm1km_predicted | sum_Area_SQUA REKILOMETERS |
|---------|-------------------------|-----------|------------|---------------------|----------------------|----------------------------|
| 1       | CHATTANOOGA SHALE       |           |            |                     |                      | 0.0                        |
| 2       | CHESTER GROUP           | 0.144     | 0.110      | 0.154               | 0.113                | 36.6                       |
| 3       | CLAYTON                 | 0.240     | 0.141      | 0.179               | 0.131                | 769.9                      |
| 4       | COFFEE SAND             | 0.256     | 0.153      | 0.176               | 0.140                | 949.3                      |
| 5       | DEMOPOLIS CHALK         | 0.263     | 0.166      | 0.179               | 0.149                | 2181.6                     |
| 6       | EUTAW                   | 0.246     | 0.125      | 0.174               | 0.136                | 2581.0                     |
| 7       | EUTAW (TOMBIGBEE SAND)  | 0.246     | 0.143      | 0.168               | 0.128                | 693.9                      |
| 8       | KOSCIUSKO               | 0.150     | 0.108      | 0.174               | 0.115                | 140.9                      |
| 9       | MERAMAC, OSAGE          | 0.140     | 0.090      | 0.151               | 0.104                | 14.8                       |
| 10      | MOOREVILLE CHALK        | 0.251     | 0.176      | 0.179               | 0.141                | 626.7                      |
| 11      | NAHEOLA                 | 0.184     | 0.141      | 0.179               | 0.135                | 686.0                      |
| 12      | PORTERS CREEK           | 0.240     | 0.157      | 0.179               | 0.141                | 1445.4                     |
| 13      | PRAIRIE BLUFF/OWL CREEK | 0.240     | 0.106      | 0.179               | 0.111                | 527.9                      |
| 14      | RIPLEY                  | 0.263     | 0.118      | 0.178               | 0.127                | 1299.7                     |
| 15      | RIPLEY (MCNAIRY SAND)   | 0.178     | 0.123      | 0.161               | 0.126                | 414.0                      |
| 16      | TALLAHATTA/NESHOBA SAND | 0.194     | 0.126      | 0.176               | 0.137                | 2956.8                     |
| 17      | TUSCALOOSA              | 0.159     | 0.106      | 0.171               | 0.123                | 761.2                      |
| 18      | WILCOX                  | 0.231     | 0.132      | 0.179               | 0.137                | 3952.3                     |
| 19      | ZILPHA/WINONA           | 0.129     | 0.105      | 0.165               | 0.125                | 38.2                       |



**Figure B 5.** Bar chart showing the variations of surface soil moisture across different surface geologic units for the study date in October.

## VITA

Anupiya Vidarshana Ellepola was born and raised in Sri Lanka. Anupiya completed his Bachelor of Science degree with a specialization in Geology from the University of Peradeniya, Sri Lanka, in 2018. The title of his honors thesis is “A comparison between resistivity sounding interpretations and borehole logs in evaluating hydrological favorability of water supply boreholes.” Upon graduation, he spent one and a half years as a teaching assistant in the geology and civil engineering departments of the University of Peradeniya. Anupiya’s employment experience also includes working as a field geologist for several government organizations in Sri Lanka. Anupiya moved to Mississippi, USA, in the spring of 2021 to pursue a Master of Engineering Science with an emphasis on Geology. During his studies at the University of Mississippi, Anupiya was employed by the Mississippi Mineral Resources Institute as a research associate. During the program, he was also awarded the Outstanding M.S. Student in Geology and Geological Engineering Award for the year 2022.