



## **TASK 8**

### **Final Report**

**PGA # 582-21-11425-009**

**Deliverable:**

8.2: FINAL Report

**Submitted on:**

January 31, 2023

**Submitted by:**

Michael H. Young  
Hassan Dashtian  
Tyson McKinney  
Bissett Young  
Shray Mathur

**Bureau of Economic Geology  
University of Texas at Austin**

---

## Table of Contents

Executive Summary.....	ES-1
1. Introduction, Rationale and Scope of the Study.....	1
1.1 Project Specifics .....	1
1.2 Task Description.....	3
2. Background Information and Literature Review.....	4
2.1 In-Situ Techniques for Measuring Soil Water Content .....	5
2.2 Soil Monitoring Networks .....	5
2.3 Relevance of Soil Water Content to Extreme Weather Conditions.....	6
2.4 Limitations of Satellite Remote Sensing of Soil Water Content .....	7
3. Monitoring Station Deployment and Data.....	7
3.1 Site Selection.....	7
3.1.1 Evaluation Process .....	7
3.1.2 Evaluation Results and Proposed Sites .....	8
3.2 Instrumentation and Data Collection .....	12
3.2.1 Description of Instruments and Logger Programming.....	12
3.2.2 Data Collection and Quality Control .....	16
3.2.3 Data Collection, Cleansing and Assimilation .....	21
3.3 Data Archiving and Distribution to TRACER.....	21
3.4 What Worked and What Didn't Work .....	22
4. Development of Gridded Soil Moisture Fields .....	23
4.1 Background on Approaches.....	23
4.1.1 Comparison of SMAP, CYGNSS and In-Situ Data.....	24
4.1.2 SMAP .....	28
4.1.3 Comparison of SMAP and In-Situ Data over Houston Area .....	30
4.2 Issues related to latency, why it matters, and how it impacts weather modeling.....	34
5. Nowcasting of Soil Water Content.....	35
5.1 Nowcasting SMAP Using Deep Learning Approach .....	35
5.2 Results of Approach .....	35
6. Conclusions and Discussions .....	37
6.1 Value and Need for Ground-based Soil Water Sensors .....	37
6.2 Need for Nowcasting and Forecasting Soil Water Content .....	37
6.3 Future Use of Data and Approaches.....	38
7. References.....	39
8. Appendices.....	43

## Table of Tables

3.1. Metadata of monitoring sites, with sensors measuring soil water content, currently being operated by Harris County Flood Control District (HCFCD). .....	11
3.2. Metadata of soil water content monitoring sites installed by UT and added to TxSON. ....	12
3.3. Instruments and model numbers used in the BEG’s participation in TRACER. ....	14
3.4. Meteorological QA screening†. ....	20
3.5. Completeness data and adjusted flags for meteorological and soil water content measurements, adjusted flags.....	22

## Table of Figures

ES-1. Houston, Texas area of interest for this project (upper left), photograph showing deployed soil water and weather station at the Guy, TX site (upper right), and example showing results of ground-truth and nowcasted SMAP water content data (bottom). ....	ES-2
1.1. Map of study area that includes Harris County with the location of HCFCD monitoring sensors (blue markers) and TxSON monitoring stations installed for this study (orange markers). The gridded layer shows the last available SMAP L4 data. Interactive map is available at: <a href="https://coastal.beg.utexas.edu/soilmoisture2/TRACER_SM_P.html">https://coastal.beg.utexas.edu/soilmoisture2/TRACER_SM_P.html</a> .....	<b>Error! Bookmark not defined.</b>
3.1. Location of proposed soil moisture sites, including those maintained by Harris County Flood Control District sites where soil moisture is currently being monitored. ....	9
3.2. Metadata of soil moisture sites currently active (Harris County Flood Control District) and the proposed sites include [a] soil texture at 0-5 cm from POLARIS (Chaney et al., 2019); [b] landcover from the USDA National Agriculture Statistics Service Cropland Data Layer (USDA National Agricultural Statistics Service Cropland Data Layer, 2019); and [c] elevation from the 3D Elevation Program Digital Elevation Model (DEM) (U.S. Geological Survey, 2017). ....	10
3.3. Soil water content monitoring station with additional meteorological monitoring equipment located in Guy, TX (TRACER station). Dashed arrows indicate orientation of buried sensors relative to the tower. ....	13
3.4. Communication and data transfer between dataloggers and computers/storage in TxSON network. ....	14
4.1. Study area to compare SMAP, CYGNSS and in-situ data. ....	24
4.2. Spatial correlation heatmaps for pair-wise comparisons between CYGNSS, SMAP and TxSON. Grids outlined in black represents grid with highest R value. For comparisons between CYGNSS vs TxSON (grids in second column), the level of significance for each grid is shown (# significant at 0.1 level, * - significant at 0.05 level, ** - significant at 0.001 level). F) shows the grid numbering followed for all heatmaps.....	<b>Error! Bookmark not defined.</b>
4.3. Scatter plots for the grid (outlined in black in Figure 4.2) with the highest R value in each heatmap in Figure 4.2. Data point within box in h) represents an outlier. ....	28
4.4. Comparison of SM retrieved from various methods including in-situ (TxSON) and satellite (CYGNSS, CYGNSS L3 and SMAP).....	28
4.5. Example gridded soil water content images obtained by NASA’s SMAP assimilated product for January 1, 2021. Each image represents water content at 6-hour time increments as shown above each image. ....	29

4.6. Index map showing grid designations for subsetted SMAP data and nowcasted SMAP data. ....	31
4.7. Soil moisture data collected at all 25 HCFC stations at 20 cm depth during 2022. ....	31
4.8. Soil water content at 5, 10, 20 and 50 cm depths for the UH Coastal Center (LAMQ) and UH Sugar Land (SUGL) stations, along with the hourly and daily precipitation. The charts show data from January to December, 2022. ....	32
4.9. Example of a comparison of soil moisture derived from the average of four sensors at the UH Sugar Land location and the SMAP L4 grid data over this location.....	33
4.10. Violin plot of the SMAP derived (red) and in-situ derived (blue) soil moisture time series.....	34
5.1. a) The front end of the web-tool for nowcasting soil water rasters over Houston area; b) example of observed raster from SMAP L4, predicted (nowcasted) and its associated absolute error raster over Houston; c) comparison of machine learning nowcasted soil water content data and SMAP-derived data for a specific 9 km x 9 km grid in the study area.....	36

## Executive Summary

The main purpose of this project was to provide ground-truth and satellite-based soil water content data in the Houston, TX, area to support the Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) user facility, and the 2022 field campaign for the Tracking Aerosol Convection Interactions Experiment (TRACER) (DOE, 2020). This involved installing four soil monitoring stations at different locations in Houston area, and assimilating the data with similar data collected by the Harris County Flood Control District (HCFCD), to improve the spatial coverage of the real-time monitoring network. The ground-truth sensor data were compared to the satellite-based (NASA's Soil Moisture Active Passive [SMAP] mission) data in the TRACER area of interest, and analyzed further to nowcast gridded data over Houston area. The nowcasted, gridded data product was made available to TRACER researchers for use in climate and land-atmosphere interaction modeling that can help understand formation and persistence of convective storm, and predict possible environmental events (e.g., floods).

This project was subdivided into several distinct parts: field collection of soil water content and meteorological data at select locations around Houston, and use of these data for improving the quality and timeliness of gridded soil water content data products. The monitoring stations were installed in the area east, south and west of the City of Houston, with two stations installed each in 2021 and 2022 (the different time lines were in part to project delays stemming from the COVID-19 pandemic). Data were collected every 5 minutes. Averaged data were streamed hourly to UT Austin, quality checked and uploaded in bulk to password-protected folders for Texas Commission on Environmental Quality (TCEQ) and TRACER collaborators. In general, uptime of these stations was excellent, though communications issues were noted at one station.

We tested two satellite products that could provide gridded soil water content data, including the Cyclone Global Navigation Satellite System (CYGNSS) and Soil Moisture Active Passive (SMAP) system, specifically the Level 4 (L4) product, both available through NASA. These systems were tested against one another and against data from a long-term soil monitoring network operated by UT Austin, known as the Texas Soil Observation Network (TxSON). The bulk of the TxSON stations are located in the Texas Hill Country; thus, we tested these systems where the ground-truth data were dense. The results showed that SMAP was more stable and better reflected the variability of ground conditions better than CYGNSS, which tended to mute the peaks and troughs seen in the in-situ sensor data. For this reason, we chose to pursue use of SMAP in the Houston area of interest.

As with nearly all satellite products available to the scientific research community, SMAP data release contains between 2-5 days of latency, which is the time delay between satellite overflight and data availability. This latency limits operationalization of these data for some types of geologic hazards, especially those involving floods and convective cell generation (the subject of the TRACER experiment). In this project, we tried to address some of the challenges of streaming, assimilating and using real-time soil water content data from various sources. The activities to accomplish this project included installing new soil monitoring stations, assimilating and harmonizing soil water content data from two different monitoring networks, quality checking of the data, downloading satellite-derived, gridded soil water data for the study area and using these data for nowcasting soil water content in real-time. The nowcasting tool reduced/eliminated the latency in the SMAP data, with a very small level of error ( $\sim 0.003 \text{ m}^3/\text{m}^3$ ). SMAP gridded products are now available every 6 hours, and are being transmitted to

researchers at the Pacific Northwest National Laboratory, where process-based regional climate modeling is being undertaken.

Significant progress in assimilating ground-based and satellite-based soil monitoring data was shown, and could be used for various environmental hazards such as flood/drought monitoring and risk mitigation. The existence of regional soil monitoring networks in the state of Texas and across the U.S. could provide soil water content in real time, in some cases at depths up to 100 cm. Tied together, these data could be used for various operational (such as calibration and validation of satellite-derived data) and application purposes (such as real-time forecasting of flash floods). Assimilating various soil water content and meteorological data from various regional networks and satellite systems for nowcasting and forecasting regional is making progress; however, significant challenges related to spatiotemporal variability and data accessibility, along with real-time training and scoring (error estimation) of data-driven models and machine learning model structures remain. Little progress has been made in the use of real-time large-scale soil moisture observations (both ground-based and satellite-derived) within the context of land-atmosphere interactions modeling.

Opportunities abound to advance the science and practice of use of large-scale soil moisture monitoring for the sake of improved Earth system monitoring and modeling, land-atmosphere interactions, nowcasting and forecasting. High-resolution nowcasted and forecasted soil water data over urban areas is an important parameter that is currently not available for any regional, national or global scale.

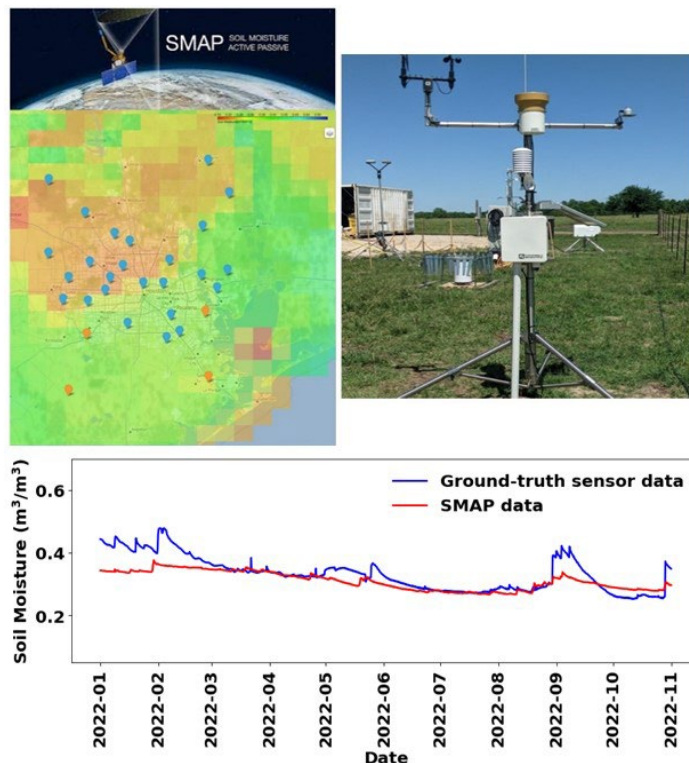


Figure ES-1. Houston, Texas area of interest for this project (upper left), photograph showing deployed soil water and weather station at the Guy, TX site (upper right), and example showing results of ground-truth and nowcasted SMAP water content data (bottom).

## 1. Introduction, Rationale and Scope of the Study

Soil moisture (SM) or volumetric water content (VWC) is an essential climate variable influencing land-atmosphere interactions, an essential hydrologic variable impacting rainfall-runoff processes, an essential ecological variable regulating net ecosystem exchange, and an essential agricultural variable constraining food security. Large-scale soil moisture monitoring has advanced in recent years, creating opportunities to transform scientific understanding of soil moisture and related processes. These advances are being driven by researchers from a broad range of disciplines, but this complicates collaboration and communication; and, for some applications, the science required to utilize large-scale soil moisture data, assimilate various data from numerous sources, validate and calibrate data from various sources, is poorly developed.

The primary purpose of this project was to indicate various soil properties including soil type, moisture content, and temperature at multiple locations in the Houston area in support of the Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) user facility, and the 2022 field campaign for the Tracking Aerosol Convection Interactions Experiment (TRACER) (DOE, 2020). The soil moisture data collected by our experiment were assimilated with the SM collected by the Harris County Flood Control District (HCFCD) to improve and the ground-truth sensor data were compared with the satellite-based data. The assimilated ground-truth sensor data were used to nowcast SM gridded data over Houston area.

### 1.1 Project Specifics

**Study area:** The study area includes Harris County, TX (see Figure 1.1), which contains the city of Houston. Harris County is the most populous county in Texas and the third-most populous county in the United States. Due to its humid subtropical climate and proximity to the coast, the county receives large amounts of rainfall. The annual average precipitation in Harris County is 1200 mm (Awal et al., 2019). Harris County has a robust network of rain gauges observations as part of its flood warning system. According to Sikder et al. (2019), Harris County Flood Warning System (FWS) has one of the densest precipitation gauge networks in the U.S., with a rain gauge density as high as one gauge per  $5 \times 5$ -km grid.

**Data:** We developed data collection, assimilation and validation processes for quality-controlled data validation and reporting. All the processes have been designed to work in real-time from data collection to data streaming. We have used various datasets including soil water content data at various depths (from both UT and HCFCD), meteorological data and satellite data. We have installed four soil moisture and meteorological data collection stations in four different locations in Houston area. Our stations collect soil water content and soil temperature data at four depths (5, 10, 20, 50 cm) along with air temperature, humidity, solar radiation and wind speed and direction. We have used both soil water content data from ground-truth sensor-based data and also satellite-based water content data. Initially six potential monitoring locations were considered to be installed by UT but instead of installing new stations, we decided to use the ready-to use water content and monitoring stations by HCFCD. A specific workflow has been designed and implemented to obtain HCFCD data and assimilate the data with the data from UT stations.

NASA's Soil Moisture Active Passive (SMAP) satellite data are being downloaded and subset over the study area in real-time to enable validation and comparison of the in-situ and satellite derived SM data and also nowcasting of SM.

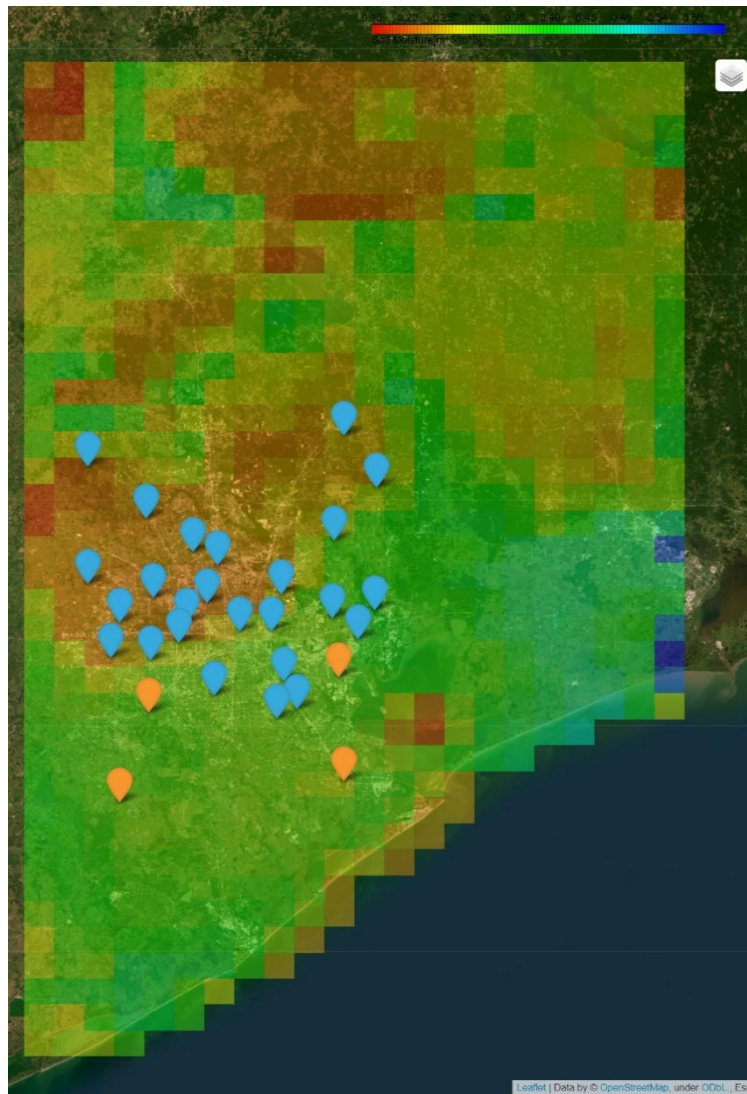


Figure 1.1. Map of study area that includes Harris County with the location of HCPCD monitoring sensors (blue markers) and TxSON monitoring stations installed for this study (orange markers). The gridded layer shows the last available SMAP L4 data. Interactive map is available at: [https://coastal.beg.utexas.edu/soilmoisture2/TRACER\\_SM\\_P.html](https://coastal.beg.utexas.edu/soilmoisture2/TRACER_SM_P.html)



## 1.2 Task Description

The project was divided into eight tasks:

Task 1 – Grant Activities Document: UT will develop the Grant Activities Description and Quality Assurance Project Plan.

Task 2 – Progress Reports: UT will submit monthly progress reports and answer any questions from the TCEQ project manager and contract specialist.

Task 3 – Monitoring Site Evaluation: UT evaluated potential monitoring sites in the Houston area, including those from the TCEQ, University of Houston, Harris County, United States Geological Survey (USGS), etc., for suitable soil parameter sampling locations to support the TRACER field campaign. Potential locations were based on benefits to the TRACER research program, site accessibility, and representative soil type and geomorphic setting, including soil thickness, bedrock geology, and terrain. We identified three priority TRACER field campaign sites (La Porte Airport; Guy, TX; and University of Houston Coastal Center in La Marque). We note that delays in the Intensive Observational Period (IOP) due to COVID restrictions led to an extension of UT's work, and additional funds for a fourth station, which was located at the University of Houston Sugar Land campus, southwest of Houston city.

Task 4 – Site Preparation, Monitor Installation, and Sampling: This task required UT to obtain site access from appropriate organizations (e.g., DOE ARM, University of Houston, etc.) and prepare the sites to accommodate the sampling equipment. Each site was to be equipped with instruments to measure weather parameters (precipitation, wind direction, wind speed, air temperature, relative humidity, incident solar radiation) and basic soil parameters including soil water content and soil temperature at depths of 5-, 10-, 20-, and 50 centimeters (cm). Stations were to be maintained throughout the IOP, and data were to be downloaded, quality checked and uploaded to a storage location available to the TCEQ Project Manager, as well as the TRACER Research team.

Task 5 – TRACER Surface Site Data Validation: UT shall validate the collected soil water content and meteorological data using standard methods and operating procedures, as described by Caldwell et al. (2019) and that conform to the National Aeronautics and Space Administration (NASA) Committee on Earth Observing Satellites Working Group on Calibration and Validation (NASA, 2020), and TCEQ Monitoring Division Standard Operating Procedure (SOP) Data Quality Review Process for meteorology (DQRP-016). All raw data collected from the dataloggers and processed data using the QA/QC protocols (described below) were made available to TCEQ, including accuracy and precision metrics for all collected parameters.

Task 6 – Harris County Flood Control District Soil Moisture Data Evaluation: The Harris County Flood Control District (HCFCD) operates a network of monitoring stations that are intended for identifying flood status, but 25 of these stations are also equipped with sensors to measure soil water content. For those 25 stations, UT obtained access to the data streams through an agreement with HCFCD and included the data in TRACER analyses. Data quality was assessed, reported as a separate deliverable, and used in the Task 7 analyses.

Task 7 – Comparison of In-Situ Soil Moisture Data to Satellite-Based Gridded Fields: In this task, UT compared soil water content estimates from two different satellite systems to ground-based measurements collected under Tasks 4 and 6. The first system was the Soil Moisture Active Passive

(SMAP) system, specifically the Level 4 (L4) product output representing 0-100 cm depths over the Houston area, and the second system was NASA's Cyclone Global Navigation Satellite System (CYGNSS), which was used over an intensively monitored research area in central Texas. Both systems were scaled to 9 km x 9 km resolution. The intention was to obtain and refine maps of soil water content values across the TRACER study area, so that process-based modeling done by the TRACER research team could use more accurate ground conditions. We anticipated the potential for using NASA's Short-term Prediction Research and Transition Center, Land Information Systems (SPoRT LIS). This will be discussed below in more detail.

Task 8 – Draft and Final Reports: In this task, we provide a concise overview of activities undertaken and data collected and analyzed during the contract period. We highlight key findings, provide pertinent analysis, describe encountered problems and associated corrective actions, and detail relevant statistics including data accuracy and precision.

All data have been transmitted to TCEQ, in accordance with Tasks 5 and 7 requirements, which are included below for completeness.

## **2. Background Information and Literature Review**

Monitoring soil water content at large spatiotemporal scales has the potential to transform scientific understanding of the patterns and dynamics of soil moisture and soil-moisture-related processes (Robinson et al., 2008). This can improve our understanding of how soil water correlates with water, energy and carbon fluxes between land and atmosphere, which is essential to meteorological forecasting and assessing environmental hazard and risk (Seneviratne et al., 2010). Soil water content measurements are also key in assessing flooding (Nied et al., 2013) and monitoring drought (Sohrabi et al., 2015). Knowledge gained from large-scale soil water content observations can help mitigate (or at least plan for) these natural hazards, yielding potentially great economic and societal benefits.

Soil water content can be defined and measured at different spatial scales that range upward from point measurements (Miralles et al., 2010) to the global scale using satellite-derived measurements (Naeimi et al., 2009). Much equipment and many methods have been developed to measure soil water content under field conditions. Also, these measurements have various classifications. Here, we divide the measurements into two categories: 1) ground-based soil water content sensors, and 2) satellite-derived soil water content measurements.

Depending on the application and field of study, soil moisture can be a vague term and thus defining it first is important. We also compare it to the more quantitative term soil water content, which is used mostly herein. Soil moisture is the level of wetness in the soil, often expressed qualitatively or gravimetrically. The most common understanding of the term is “the total amount of water in the unsaturated zone.” It is now, today, a commonly used term, trending from colloquial usage to scientific literature. Soil water content, on the other hand, is a quantitative term expressed either gravimetrically (g water / g soil), or volumetrically, using units of  $\text{cm}^3/\text{cm}^3$  or  $\text{m}^3/\text{m}^3$ . The soil bulk density converts the two. The two terms are often used interchangeably, but soil water content is the generally accepted term when expressed quantities of water in soil.

The spatial scales of different types of measurements varies. Western and Blöschl (1999) proposed a scale definition based on spacing, extent, and support to distinguish soil water content measurement

approaches. The term “spacing” refers to the distance between measurement points, the term “extent” indicates the overall coverage, and “support” refers to the measurement or integration volume or area. As an example, for a network of sensors that measure soil water content deployed in a small farm field, the scale may be defined based on 10-cm spacing (distance between sensors), 100-m<sup>2</sup> extent (area of the field), and 1000-m<sup>3</sup> support (volume of the soil sampled by the sensor). Similarly, for satellite-derived measurements over Houston, the scale would be defined based on 9 km spacing (distance between the centers of two neighboring grids in SMAP, for example), 342 km × 198 km extent (the footprint area), and 81 km<sup>2</sup> support (the grid size).

Missions such as the Soil Moisture and Ocean Salinity (SMOS) operated by the European Space Agency (ESA; launched in 2009), and the Soil Moisture Active Passive (SMAP), launched in 2015 by NASA, have parallel goals of providing soil water content estimates within an error limit of  $\pm 0.04 \text{ m}^3/\text{m}^3$ , within the first 5 cm of soil, in regions where the vegetation water content is  $< 5 \text{ kg}/\text{m}^2$  (Entekhabi et al., 2010). The SMOS mission is currently providing soil water content estimates at approximately a 40-km resolution, with a 2–3 day repeat time (Kerr et al., 2010). The SMAP mission has a repeat time similar to SMOS, with radiometer measurements provide soil water content estimates near the same 40-km resolution (Entekhabi et al., 2010), although subsequent analyses and assimilation of other products has improved this resolution to 1 km by 1 km (see Fang et al., 2022).

## **2.1 In-Situ Techniques for Measuring Soil Water Content**

Point observations are commonly directly obtained with in situ techniques such as simple (but destructive) gravimetric sampling, various electromagnetic sensors (e.g., time domain reflectometry, frequency domain reflectometry), or application of neutron moderation techniques with sensing depths ranging from about 2-5 cm to ~60-90 cm. In general, electromagnetic (EM) sensors are considered as the most reliable means for direct and accurate determination of moisture within the soil profile, assuming no anthropogenic interferences. The time domain reflectometry (TDR) method was popularized by the work on Topp et al. (1980), who introduced a widely used calibration curve that converted dielectric constant (or, permittivity) of porous media. Dielectric constant is a measure of the electrical potential energy of a substance under the influence of an EM field. The EM sensor electric field is generally directed into the soil along 2-, 3-, or 4-parallel electrodes or an adjacent pair of rings. The measurement of permittivity provides a highly accurate determination of water content in soil because the permittivity of water is about 80, while permittivities of solids and air are around 4 and 1, respectively. Interestingly, the most accurate water content determination method remains the original TDR method operating at around 1-GHz frequency.

## **2.2 Soil Monitoring Networks**

To calibrate and validate such satellite- and model-based soil water content estimates, in-situ measurements are an indispensable source of information (e.g., Caldwell et al., 2019). As remote sensing platforms become more strategic for global monitoring of Earth resources, various ground-based invasive and noninvasive soil moisture measurement techniques and their monitoring networks are being utilized for their validation around the globe. Perhaps the oldest state-run meteorological network in the US is found in Oklahoma, and represents today the gold standard of soil monitoring. Brock et al. (1995) described the approach and justification of creating the Oklahoma Mesonet, in which at least one station is available in each county for use in predicting extreme weather events, from tornados to wildfires, and for assisting their agricultural community. The Texas Soil Observation Network

(Caldwell et al., 2019) was justified given the success of the OK Mesonet. Since then, Texas has assessed the benefits of creating a Texas Mesonet (Nielsen-Gammon et al., 2017), which has since yielded the funding of the TexMesonet, now operated by the Texas Water Development Board. Data from these networks are now being compiled by international organizations for use in continental and global studies. For example, the International Soil Moisture Network (ISMN) was initiated in 2009, funded by the European Space Agency, to serve as a centralized data hosting facility for globally available in situ soil moisture measurements (Dorigo et al., 2011). The ISMN brings together in situ soil moisture measurements collected and freely shared by a multitude of organizations globally, and applies advanced quality control protocols to ensure that data stored in their database achieves similar levels of data quality. Users can freely retrieve the data from this database through an online web portal (<https://ismn.earth/en/>).

### **2.3 Relevance of Soil Water Content to Extreme Weather Conditions**

Extreme weather events, an important driver for the TRACER experiment, are often interpreted as the result of large-scale atmospheric circulation and sea surface temperature anomalies (Cook et al., 2010). For example, El Niño Southern Oscillation and atmospheric blocking are tropospheric warming that considered key processes for drought and heatwave initiations (Helama et al., 2009). In addition to large-scale atmospheric circulation anomalies forced by the ocean, soil moisture anomalies, as a land-atmosphere feedback parameter, can strongly modulate near-surface heat and aridity (Berg et al., 2016), and promote large-scale extreme conditions. Soil moisture thus is considered a key variable of the climate system (Zhou et al., 2019). It constrains plant transpiration and photosynthesis, and impacts water, energy and biogeochemical cycles. Moreover, it is a storage component for precipitation and radiation anomalies, inducing persistence in the climate system. Soil moisture is involved in a number of feedbacks at the local, regional and global scales, and plays a major role in climate-change projections.

Heat and aridity events have recently received more attention because of their devastating impacts on the environment, economy and society (Saini et al., 2016). Ecosystem productivity during droughts is believed to be impacted by two main factors, including: low soil water content and high atmospheric vapor pressure deficit (Novick et al., 2016). These factors cause stress on ecosystems, which can substantially reduce terrestrial carbon uptake and food production (see Lobell et al., 2013) and can drive widespread tree mortality (Adams et al., 2017). Hong and Kalnay (2000) reported the role of SM in the 1998 Oklahoma–Texas drought. Drought and tree mortality can cause frequent wildfire as a consequence of low soil water content (Stephens et al., 2018).

Conversely, elevated soil water content can also result in extreme weather conditions, including flash and regional-scale flooding. A number of factors are assumed to impact severity of a flooding event including precipitation intensity, percentage of sealed catchment area, soil permeability, water holding capacity, topographic slopes and antecedent soil water content (Grillakis et al., 2016). In contrast to these other factors, changes in soil water content between events are generally much smaller in magnitude, though changes can occur quickly, even on minute to hourly time scales. Soil water content can vary from near to the wilting point (the point of cavitation in plant systems) to effective saturation (nearly all soil pores are filled with water). Water content is considered the most important soil factor influencing rapid runoff and flash flooding. Soil moisture can, in fact, control whether a given precipitation event produces surface runoff or not, due to the non-linear nature of runoff response to rainfall (Wasko & Nathan, 2019). Both satellite-derived and in-situ soil water content data are used to

predict the risk of flooding (Wanders et al., 2014; Kim et al., 2019). Many of the monitoring stations maintained by HCFCD are equipped to measure soil water content for monitoring risk flood in Harris County. However, despite their widespread use and essential roles in extreme weather conditions, data derived from satellite measurements and from in-situ sensors do have some limitations that should be considered in site selection and use of the data.

## **2.4 Limitations of Satellite Remote Sensing of Soil Water Content**

At large-scale, to assess effects of soil water content on magnitudes of extreme weather events (e.g., floods), data are needed at the catchment scale. However, this is difficult to measure using in-situ sensors, often requiring a combination of monitoring station and satellite derived soil water content data. Satellite derived data can provide a partial solution, in the sense that coverage is global and the uniqueness of the instruments favors low bias in the data collection procedure. However, there is a tradeoff on spatial resolution, depth of penetration and revisits. Moreover, these measurements need to be calibrated to *in-situ* ground-based measurements. Finally, specific surface conditions (dense rain forest, snow covered soils, surface water, etc.), land cover and land use, can significantly hamper or bias the measurements; global coverage of high resolution, accurate soil water content data is a difficult goal to attain.

Despite its numerous advantages and accuracy, EM based methods (like the TDR) and, in general, point measurements may have some limitations, including high soil salinity levels, loss of contact between sensors and the surrounding soil from shrinkage or biological activity, or lack of calibration, especially for soils with high specific surface areas. Nonetheless, point-scale measurements are vital components in any monitoring program that relies on near-real time interrogation of soil conditions.

## **3. Monitoring Station Deployment and Data**

In this section, we describe field and data management activities that occurred mostly under Tasks 3, 4 and 5.

### **3.1 Site Selection**

In this task, BEG evaluated potential soil monitoring sites in the Houston area to support the TRACER field campaign. Potential locations were based mostly on benefits to the TRACER research program, site accessibility, and representative soil type and geomorphic setting. Our evaluation began by considering the three TRACER field campaign sites (University of Houston Coastal Center (UHCC) in La Marque, TX; La Porte Texas Airport, and a site near Guy, TX), the latter two sites were already been selected by TRACER for the project. In late 2020, we submitted proposals to the TRACER program to stage our equipment at these sites, and received approval by DOE ARM.

#### **3.1.1 Evaluation Process**

We initialized the evaluation of potential sites by including the three sites listed above, and existing sites instrumented and maintained by the Harris County Flood Control District (HCFCD; Figure 1.1 and 3.1). The latter sites were included because their instrumentation platforms for measuring soil water content are the same as those used in the Texas Soil Observation Network (TxSON), and their data are available to our research program, augmenting data we will be collecting for TRACER.

To evaluate the sites, soil textural data were obtained from the Probabilistic Remapping of SSURGO (POLARIS) data set (Chaney et al., 2019), which derives terrain-based, 30 m probabilistic soil properties,

and from the Soil Survey Geographic Database (SSURGO) and gridded SSURGO database (Soil Survey Staff, 2014). POLARIS consists of six soil layers with boundaries at depths of 5, 15, 30, 60, 100, and 200 cm. Using the USDA texture triangle, soil texture for the 0-5 cm layer was obtained. Second, landcover information was obtained from the U.S. Department of Agriculture (USDA, National Agricultural Statistics Service Cropland Data Layer, 2019), which uses image classification of satellite imagery collected during the growing season, producing a 30 m resolution product. Finally, elevation data were obtained from the U.S. Geological Survey (USGS) 3D Elevation Program (3DEP) Digital Elevation Models (DEMs) at a resolution of 1/3 arc-second (USGS, 2017).

### **3.1.2 Evaluation Results and Proposed Sites**

It is clear (Figure 3.2a) that soil textures across the Houston area and the HCFCD sites cover a wide range, with sandier soils in northern Harris County and clayey soils to the south. The cropland data layer (Figure 3.2b) shows that most of Harris County is classified as developed area with forests to the north, wetlands near the coast and at the bayous, grasslands to the west, and a mixture of grassland and agriculture to the south. Most HCFCD sites are considered to be in developed areas. We note that TRACER sites at the La Porte Airport and the UHCC are also classified as developed, while the site in Guy, TX is classified as a grassland. Elevation (above mean sea level) at the HCFCD sites range from 2 m to 81 m. Elevation is less than 10 m at the La Porte and UHCC TRACER sites near the coast, and 22 m at the site in Guy, Texas (Figure 3.2c). Table 3.1 shows site-specific information for all sites belong to HCFCD and Table 3.2 includes the information related to four stations installed by UT.

In summary, we suggested Site BEG1 to be located at the ARM Mobile Facility (AMF1) near La Porte, Texas at La Porte Municipal Airport, a main site for TRACER and where our monitoring station will include soil water content measurements. Site BEG2 was proposed to be located at the University of Houston Coastal Center in La Marque, TX, located southeast of Houston near the coast. This site was discussed as potentially hosting other TRACER sensors; thus, collocating our soil water content sensors from our monitoring station would be advantageous. The site also filled a gap in coverage, being south of Harris County. Site BEG3 was proposed to be located at a site in Guy, Texas, another major staging area for TRACER, and where soil water content data will be valuable. We note that the COVID pandemic delayed by a year the IOP, thus also delaying the deployment of the BEG3 station. To maximize the amount of data collected, and at the suggestion of TCEQ, we initiated discussions with University of Houston about deploying a station at the UH Sugar Land (UHSL) campus, northeast of BEG3, and west of downtown Houston. The UHSL site was installed on September 10, 2021. By early 2022, as the pandemic was subsiding, TRACER personnel identified the Guy, TX site, and installation of BEG3 occurred on May 26, 2022. This site is located southwest of Houston, on the mainland and fills a gap in soil monitoring locations. Given that other TRACER sensors will be deployed at these sites, they are ideal locations for our monitoring platforms as well.

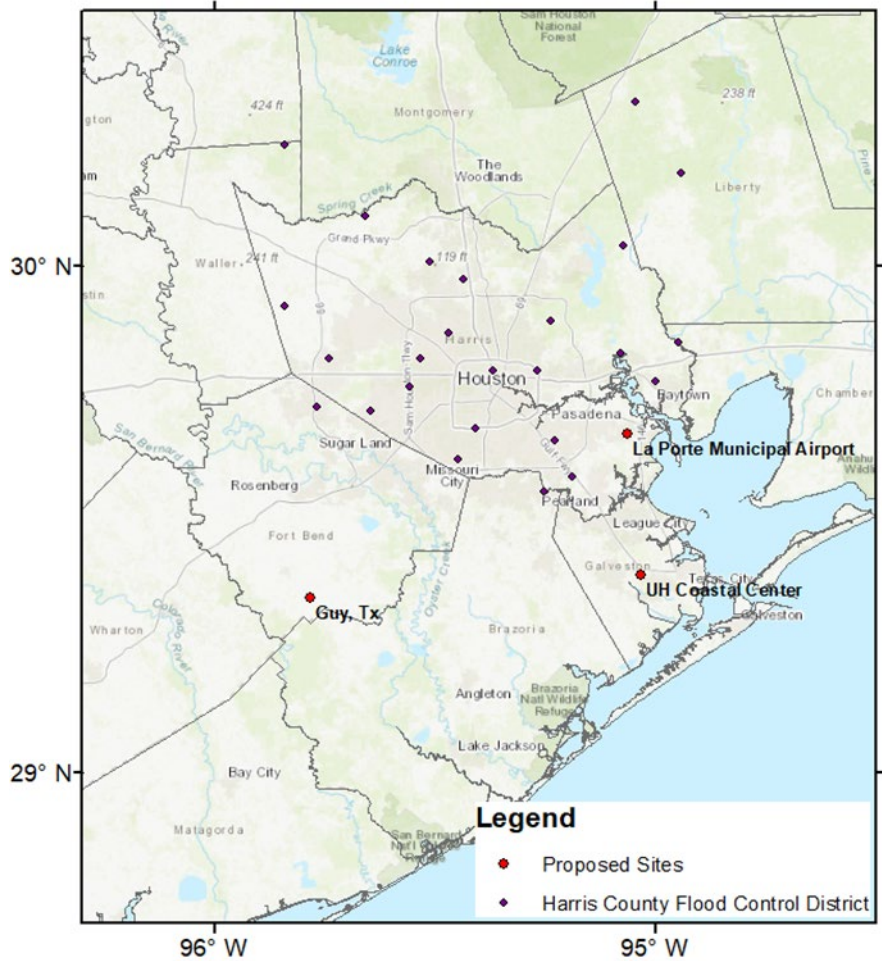


Figure 3.1. Location of proposed soil moisture sites, including those maintained by Harris County Flood Control District sites where soil water content is currently being monitored.



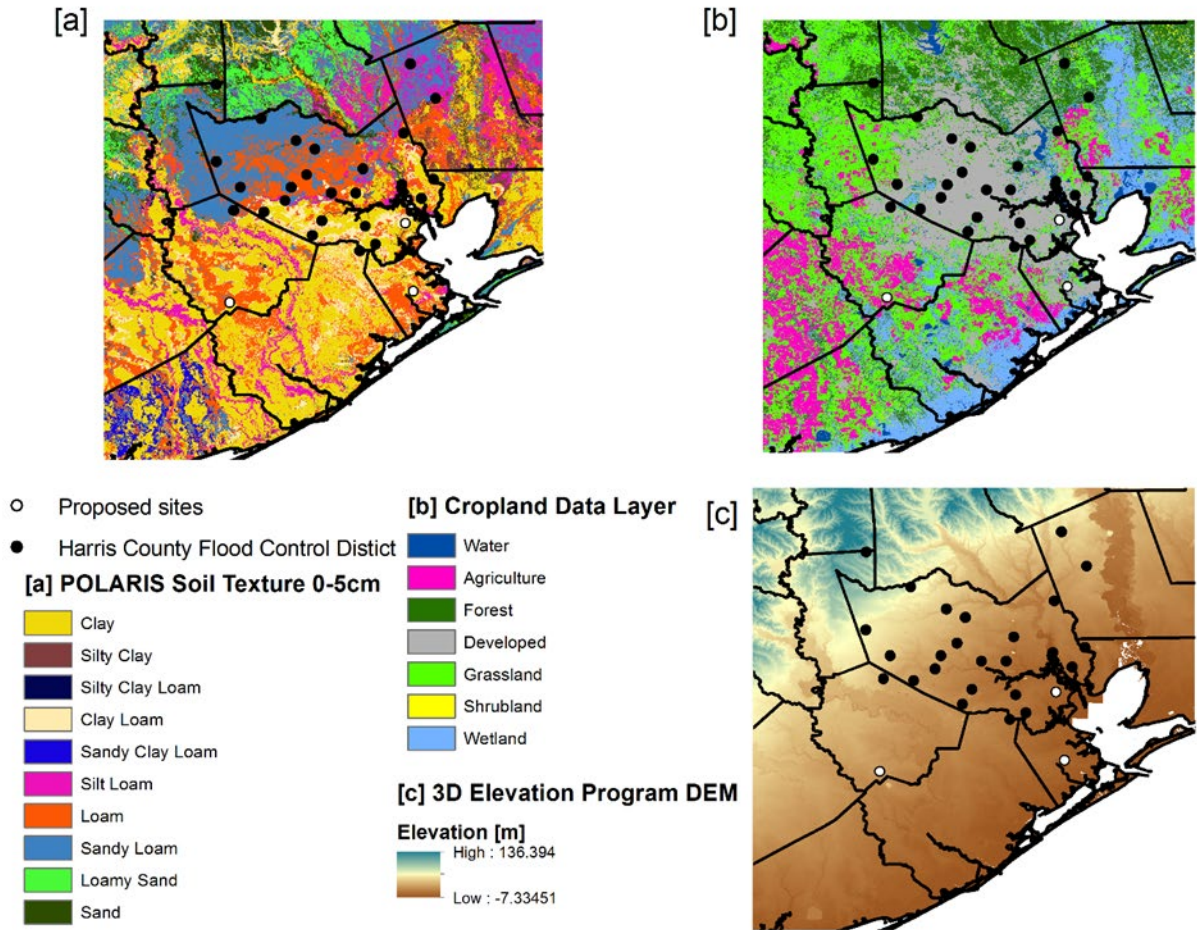


Figure 3.2. Metadata of soil moisture sites currently active (Harris County Flood Control District) and the proposed sites include [a] soil texture at 0-5 cm from POLARIS (Chaney et al., 2019); [b] landcover from the USDA National Agriculture Statistics Service Cropland Data Layer (USDA National Agricultural Statistics Service Cropland Data Layer, 2019); and [c] elevation from the 3D Elevation Program Digital Elevation Model (DEM) (U.S. Geological Survey, 2017).



Table 3.1. Metadata of monitoring sites, with sensors measuring soil water content, currently being operated by Harris County Flood Control District (HCFCD).

Site Name	Station ID	Latitude	Longitude	Start Date	Texture	Landcover	Elevation (m)
Turkey Creek @ FM 1959	140	29.5839	-95.1876	08/21/19	Clay	Developed	7.73
Clear Creek @ Country Club Drive	150	29.5561	-95.2520	08/05/20	Clay	Developed	11.67
Berry Bayou @ Nevada Avenue	310	29.6563	-95.2289	08/12/20	Clay	Developed	8.02
Sims Bayou @ Hiram-Clarke Road	380	29.6193	-95.4459	08/12/20	Clay	Developed	15.79
Brays Bayou @ SH 6	485	29.7151	-95.6439	08/05/20	Sandy loam	Developed	25.25
Little White Oak Bayou @ Trimble Street	560	29.7930	-95.3678	05/14/20	Silt loam	Developed	14.18
Vogel Creek @ Victory Drive	585	29.8681	-95.4692	01/19/20	Loam	Developed	18.42
San Jacinto River @ Rio Villa	710	29.8290	-95.0777	08/19/19	Sand	Wetland	2.00
Hunting Bayou @ Loop 610 East	830	29.7936	-95.2680	05/20/20	Clay loam	Developed	11.05
Tomball Repeater	1075	30.0965	-95.6573	01/01/19	Sand	Developed	71.87
Birch Creek @ Riley Road	1076	30.2379	-95.8395	08/20/19	Sand	Developed	80.74
Cypress Creek @ Stuebner-Airline Road	1140	30.0063	-95.5116	08/12/20	Sand	Developed	35.26
Cypress Creek @ Sharp Road	1185	29.9210	-95.8402	01/01/20	Sandy loam	Developed	50.48
Goose Creek @ Baker Road	1540	29.7709	-94.9996	05/15/20	Silty clay	Developed	5.69
Greens Bayou @ Mount Houston Parkway	1600	29.8919	-95.2376	08/05/20	Loam	Developed	15.43
North Fork Greens Bayou @ Ella	1655	29.9726	-95.4350	08/20/19	Loam	Developed	26.65
Cedar Bayou @ FM1942	1730	29.8494	-94.9475	01/30/20	Clay	Forest	8.47
Huffman Repeater	1930	30.0386	-95.0720	01/01/19	Loam	Developed	21.32
Luce Bayou @ SH 321	1960	30.1805	-94.9399	07/06/20	Sandy loam	Wetland	30.24
Tarkington Bayou @ SH 105	1975	30.3213	-95.0437	02/24/20	Loam	Developed	38.40
Hosepen Creek @ Trailside Drive	2130	29.8839	-95.6347	08/12/20	Sandy loam	Developed	35.10

Site Name	Station ID	Latitude	Longitude	Start Date	Texture	Landcover	Elevation (m)
Buffalo Bayou @ Peek Road	2025	29.7222	-95.7674	08/12/20	Sandy loam	Developed	32.57
South Mayde Creek @ Morton Road	2170	29.8169	-95.7405	05/18/20	Sandy loam	Developed	38.80
Buttermilk Creek @ Moorberry Lane	2253	29.8159	-95.5327	09/03/19	Loam	Developed	26.51
Buffalo Bayou @ West Beltway 8	2270	29.7620	-95.5575	06/13/19	Sandy loam	Developed	21.92

Table 3.2. Metadata of soil water content monitoring sites installed by UT and added to TxSON.

Site Name	Station ID	Latitude	Longitude	Start Date	Texture	Landcover	Elevation (m)
University of Houston Coastal Center	UHCC	29.388	-95.0425	05/05/2021	Clay	Developed	5.00
La Porte Airport	LAPT	29.6697	-95.0584	05/06/2021	Clay	Developed	7.00
University of Houston Sugar Land	SUGL	29.5734	-95.6495	11/10/2021	Clay	Developed	22.00
Guy, TX	GUYT	29.3292	-95.7403	05/26/2021	Clay Loom	Ranchland	22.00

## 3.2 Instrumentation and Data Collection

### 3.2.1 Description of Instruments and Logger Programming

On May 6 and 6, 2021, the first two stations were installed in Houston area (UHCC and La Porte) and added to the TxSON network. Two more stations were installed later in 2021 and 2022 (UHSL on September 10, 2021 and Guy, TX on May 26, 2022). The stations record and monitor soil water content and temperature, precipitation, wind speed and direction, air temperature and relative humidity, and solar radiation. The installation and data collection in these stations follow the same procedure of TxSON (Caldwell et al., 2019). Each station is powered by a mast-mounted 10-watt solar panel along with a back-up 12V, 8 amp-hour battery. The battery, modem, and datalogger are housed inside a mast-mounted enclosure, bolted to a tripod frame, and protected by lightning rod. Each monitoring station is equipped with sensors buried at 5, 10, 20, and 50 cm depths, and measures volumetric soil water content (VWC), temperature, and electrical conductivity every 5 minutes. Figure 3.3 shows an image of the station at Guy, TX, and highlights the specific instruments. Maximum height of instruments at this site is approximately 7-10 feet.

All monitoring sites were secured by fencing at each of the different facilities, so no station-specific fencing was necessary.

Data were collected using the proprietary software of Campbell Scientific, Inc. (Logan, UT), called LoggerNet, which communicates with the Campbell Scientific dataloggers used in this project and pulls data to the main TxSON Hub (“server”) at the Bureau of Economic Geology at UT Austin. Data were stored on the individual loggers continuously, and uploaded hourly. The data retrieval schedule was meant to reduce the power draw on the modem and increase the longevity of the monitoring system. New data were appended to existing data files. Figure 3.4 shows the flowchart of data transfer and storage between dataloggers and computers.

The details of the instruments used in each station are provided in Table 3.3. All sensors are commercially available and used by research and scientific organizations worldwide. Each logger is programmed using Campbell Scientific-based programming language called CRBasic. The editor is part of a PC-compatible platform that provides significant flexibility for using instrument data acquisition and signal conditioning for their loggers. An example of the program is shown in Appendix A.

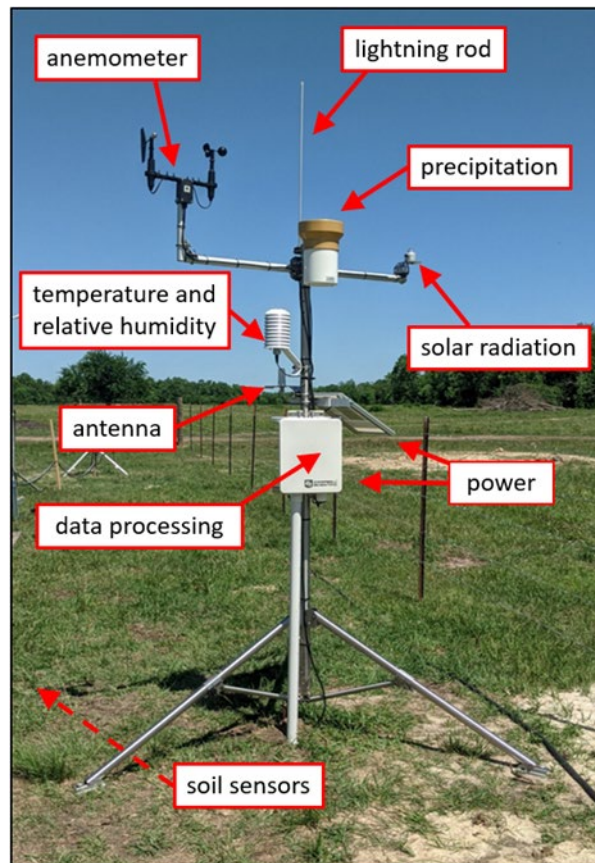


Figure 3.3. Soil water content monitoring station with additional meteorological monitoring equipment located in Guy, TX. Dashed arrow indicates orientation of buried sensors relative to tower.

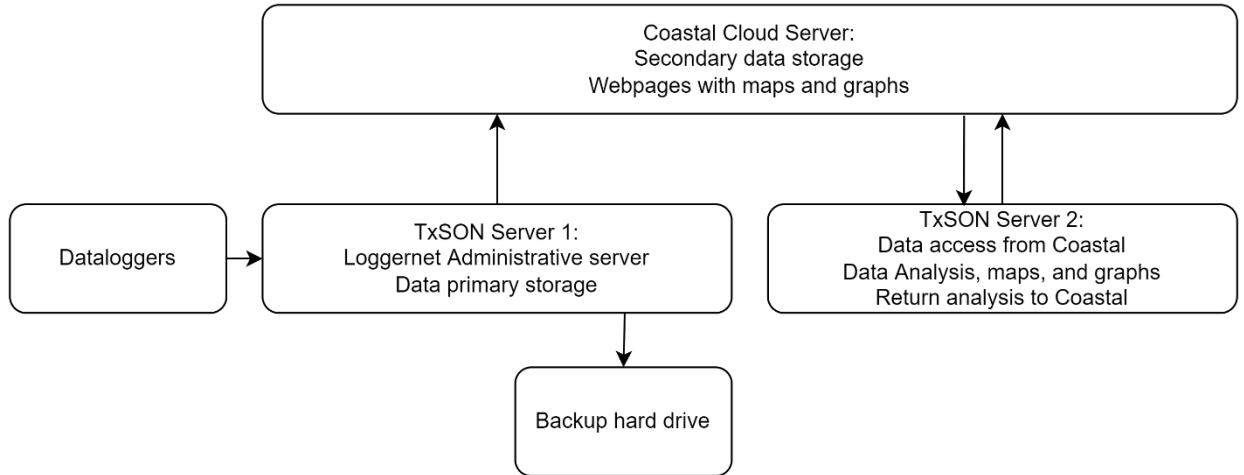
















Figure 3. 4. Communication and data transfer between dataloggers and computers/storage in TxSON network.

Table 3. 3. Instruments and model numbers used in the BEG’s participation in TRACER.

#	Photo	Model	Description	Quantity
1		CS655-33-PT-VS	12cm Water Content Reflectometer Plus	4
2		03002-L15-PT	RM Young Wind Sentry Set	1
3		TE525WS-L15-PT	Texas Electronics Rain Gage 0.01 inch Tip w/8 inch Orifice	1

#	Photo	Model	Description	Quantity
4		HygroVUE10-10PT	CSL Digital Temperature /RH Sensor	1
5		RAD10E	METSPEC 10-Plate Solar Radiation Shield for Larger Sensors	1
6		ENC10/12-SC-MM	Weather Resistant Enclosure, 10 x 12 inches	1
7		7BP7	12V Sealed Rechargeable Battery	1
8		COAXSMA-L8	Antenna Cable LMR195 w/SMA & Type N Male	1
9		CM110	Stainless Steel, 10ft Tripod w/Grounding Kit	1
10		CR300-NA	Datalogger (-40 to +70C) - NA No Additional Coms	1

#	Photo	Model	Description	Quantity
11		CM225	Solar Sensor Mounting Stand	1
12		SR05-L10-PT	Hukseflux ISO 9060 Second Class Pyranometer w/ Analog & RS-485 MODBUS	1
13		CELL210-ND	4G LTE Cat1 Cellular Module for Verizon	1
14		CM220	Right Angle Mounting Kit	1

### 3.2.2 Data Collection and Quality Control

All collected data during the TRACER effort were archived and stored on secure computers within UT Austin. Standard procedures for data quality assurance and control (QA/QC) were applied to each TRACER station installed in Houston. Specifically, standard procedures adopted for soil water content data for this project are described by Caldwell et al. (2019) and conform to the National Aeronautics and Space Administration (NASA) Committee on Earth Observing Satellites Working Group on Calibration and Validation (NASA, 2020) and Montzka et al. (2020). All meteorological data QA/QC protocols conform to the TCEQ Monitoring Division Standard Operating Procedure (SOP) Data Quality Review Process for Meteorology (DQRP-016). Raw and validated data were uploaded regularly to a password-protected folder available to the TCEQ Project Manager. We are also collaborating with TRACER researchers at the Pacific Northwest National Laboratory, who are also accessing data for their process-level meteorological modeling.

All logger-based data are saved as ASCII data files (.dat). A python script then pre-processes and combines the data for streaming, transfer, and archiving. Data are saved into hourly and daily tables that append to separate files for each station. The processing code takes the ASCII file as input, reads the column names, and converts the names to a general name applicable for each sensor, while ascertaining

the number and type of instruments for a given site. Missing or “NaN” (Not a Number) and outlier data are determined and replaced with the (local) median value of the data. The output is a Pandas DataFrame (a popular python data structure for fast and versatile computation) that is used to create near real-time interactive graphs that are published to the public facing website.

Data from SMAP are downloaded using both R Studio and Python scripting, as soon as they are published by NASA. We then subset the data to cover only the Houston area. Another Python script processes and visualizes the data. A separate Matlab script acquires and pre-processes the HCFCD data and sends it to another Python script used for data visualization. All scripts are publicly available on Github ([https://github.com/begtxson/TRACER\\_Soil\\_Moisture](https://github.com/begtxson/TRACER_Soil_Moisture)).

### 3.2.2.1 Quality Assurance/Quality Control (QA/QC) for Soil Moisture Data

Built into the Python processing code are automated quality assurance (QA) flags that are used for soil water content data. Flags being raised do not necessarily indicate that sensors are failing, only that data need to be reviewed more carefully. Four different flags were designed into the QA check for sensors at each depth, some of which are based on whether physical quantities are realistic and others on possible mismatches between environmental conditions and soil responses:

1. Constant soil temperature values using a 10-hr moving time window (Dorigo et al., 2013)
2. Geophysical constraints using sensor-specific, operating ranges, including physically realistic boundaries for volumetric water content (VWC) (also called soil moisture) and temperature (T), which TxSON sets at (a)  $0.04 \leq VWC \leq 0.70$  and (b)  $0 \leq T \leq 60^\circ\text{C}$  (Dorigo et al., 2013)
3. Spike detection outside of a rainfall event using a Hampel filter (Caldwell et al., 2019)
4. Outlying spikes using a Hampel filter (Caldwell et al., 2019)

In reality, sensors measure soil permittivity and electrical conductivity (EC). An empirical equation converts these parameters to soil water content. Ledieu et al. (1986) suggested the following relationship between apparent relative permittivity ( $K_a$ ) and SWC:

$$SWC = C_0 + C_1\sqrt{K_a}$$

where,  $C_0 = -0.081$  and  $C_1 = 0.093$ . Considering that time domain reflectometry (TDR) travel time also increases with the square root of EC, Evett et al. (2005) added another parameter to Ledieu et al. (1986) equation, resulting in:

$$SWC = C_0 + C_1\sqrt{K_a} + C_2\sqrt{EC}$$

where EC is soil bulk electrical conductivity and  $C_2 = 0.031$ . Throughout this project, we have focused on the more comprehensive equation Evett et al. (2005), and have provided data to TCEQ using the Evett curve.

For all soil water content data, we apply the following QA criteria to the resulting time-series.

- 1) Constant water content: Significant rises ( $x_t > x_{t-1}$ ) are flagged if the following conditions at time step, t, apply:

$$x_t - x_{t-24} > 2\sigma_{x[t-24,t]}$$

where  $x_t$  is the soil moisture value at timestep  $t$  in hours and  $\sigma_{x[t-24,t]}$  is the standard deviation of  $x$  over the preceding 24 h.

- 2) Soil temperature below zero.
- 3) Spike Detection: If a substantial change in soil water content is detected between two consecutive time steps; that is, a minimum increase or decrease of 15% compared to the previous value, the reference time step  $t$  is identified as potential spike:

$$\frac{x_t}{x_{t-1}} < 0.85 \text{ or } \frac{x_t}{x_{t-1}} > 1.15$$

The above equation is valid for the time in which no precipitation has been observed. Thus, second derivative of soil water content time series is used to identify the spikes:

$$0.8 < \frac{x''_{t-1}}{x''_{t+1}} < 1.2$$

We added a third criterion based on the coefficient of variation over a 24-hr interval centered at  $t$ :

$$\frac{\sigma^2(x_{t-12}, x_{t+12})}{\mu(x_{t-12}, x_{t+12})} < 1$$

where  $\sigma^2$  is the variance and  $\mu$  the average over the interval  $x_{t-12}, x_{t+12}$ . The threshold results from the properties of the coefficient of variation, where a value above 1 symbolizes very noisy data. An observation is flagged as a spike only if all three above conditions are fulfilled. We note this check cannot be performed in real time, because a 24-hour time window covering observations “from the future” is needed.

- 4) Break detection: To be flagged as a break, an observation needs to fulfill three criteria. The relative change of soil water content must be at least 10%. Moreover, to prevent overall flagging of low absolute moisture values, the absolute change in soil water content needs to be at least 0.01, leading to:

$$\left| \frac{x_t - x_{t-1}}{x_t} \right| > 0.1 \text{ and } |x_t - x_{t-1}| > 0.01$$

Another factor considered was the change in the first derivative of soil water content time series, where a negative (positive) break (i.e., a sudden change in the value of the data) is expressed by a strong negative (positive) change of the first derivative,  $x'_t$ .

$$x'_t > 10 \frac{1}{n} \sum_{k=-12}^{12} x'_{t+k}$$

Second derivative criterion: A negative (positive) break results in a large negative (positive) second derivative at  $t$  followed by a large positive (negative) value at  $t + 1$ : the peaks in the second derivative are approximately of the same size (though opposite in sign) resulting in a ratio around one. At  $t + 2$  the second derivative returns to a value close to zero; hence, the ratio of the absolute second derivative between at  $t + 1$  and  $t + 2$  is very large. This results in the following conditions:



$$\left| \frac{x''_{t-1}}{x''_{t+1}} \right| = 1 \text{ and } \left| \frac{x''_{t-1}}{x''_{t+1}} \right| > 10$$

Thus, for soil water content time series, for a specific sensor at a defined depth, we have defined four flags ([flag1, flag2, flag3, flag4]), each with value either 0 or 1 indicating flag low and flag high, respectively; for example, [0,1,0,0] means only the second flag is high. As each station has four sensors, we therefore have 16 elements for each station, the first 4 elements represent the flags for the sensor at 5 cm depth, the next 4 elements represent the flags for sensor at 10 cm depth, and so on. The following code represents the flagging elements and their location in the binary number that we create.

```
Sflag = char([
    num2str(VWC_E_Flag_hr(:,1,1,2)) num2str(VWC_E_Flag_hr(:,1,1,3)) num2str(VWC_E_Flag_hr(:,1,1,4)) ...
    num2str(VWC_E_Flag_hr(:,1,2,2)) num2str(VWC_E_Flag_hr(:,1,2,3)) num2str(VWC_E_Flag_hr(:,1,2,4)) ...
    num2str(VWC_E_Flag_hr(:,1,3,2)) num2str(VWC_E_Flag_hr(:,1,3,3)) num2str(VWC_E_Flag_hr(:,1,3,4)) ...
    num2str(VWC_E_Flag_hr(:,1,4,2)) num2str(VWC_E_Flag_hr(:,1,4,3)) num2str(VWC_E_Flag_hr(:,1,4,4)) ...
```

Writing a 16-digit number may be confusing, especially when a small portion of the data is typically flagged and the rest are clean data. Therefore, instead of writing a 16-digit binary number, we convert the binary number to a decimal value. Nearly all data collected during the field work were within QA criteria, thus yielding only one element (that is, 0). To convert a decimal flag to a binary flag, we use the dec2bin() function in Matlab; for example dec2bin(0,16) yields '0000000000000000' meaning that no flags are shown as high for any sensors. The data file that we provide after flagging the data has the "flagged" term in its name. The last column in the data file is called "Flag." This column has decimal numbers generated from the binary representation of flags of all sensors and four criteria for each time steps. For example, a value of 65535 in Flag column means that the original binary number was '1111111111111111'; that is, all for flags are high for all sensors. A flag value of 4095 represents binary number of '0000111111111111', meaning data from sensor at 5 cm depth are within threshold, but data from the other three sensors are outside of threshold for all four criteria. A flag value of '3855' in binary format is '0000111100001111', meaning that data from sensors at 5 cm and 20 cm depth are within threshold, but data from sensors at 10 cm and 50 cm depths are outside threshold. More complex cases: for example, a flag value of 143 represents binary number of '0000000010001111', meaning that sensor at 20 cm depth has recorded almost constant soil water content for the past 24 hours, exceeding the first criteria ( $x_t - x_{t-24} < 2\sigma_{x[t-24,t]}$ ), and that all flags for sensor at 50 cm depth were raised high.

### 3.2.2.2 QA/QC for Meteorological Monitoring

For stations installed and operated by BEG, screening criteria (Table 3.4) were used to accept data collected with additional meteorological instruments.

The approach to identify flagged meteorological data assigns an 18-digit code to each data collection interval, very similar to that described above for the soil water content, only simpler. Each digit and value of each digit correspond the flag listed above. For example:

- Decimal number is 0 in Flag column represents binary number of '000000000000000000' – all flags are low
- 139810 represents '100010001000100010', meaning all values (wind speed, wind direction, temperature (air and dew) and solar radiation) are 'NaN'.

- 64 represents '000000000001000000', meaning that air temperature has not changed over last 12 hours.
- 16384 represents '000100000000000000', meaning that wind speed has been constant over 12 hours.

Table 3.4. Meteorological QA screening†.

Parameter	Data Screening
Wind speed (WS)	0 m/s $\geq$ WS $\leq$ 25 m/s, WS varies $\geq$ 0.1 m/s for 3 consecutive hours‡, WS varies $\geq$ 0.5 m/s for 12 consecutive hours, or per site specific climatology criteria 1/week or more frequent
Wind direction (WD)	0° $\geq$ WD $\leq$ 360°, WD varies $\geq$ 1°/3 consecutive hours, or per site specific climatology criteria
Temperature	Local record low $\geq$ Temperature $\leq$ local record high, Temp $\leq$ 5°C from previous hourly record, Temp varies $\geq$ 0.5°C during 12 consecutive hours, or per site specific climatology criteria
Relative Humidity (RH) Dew Point	Dew Point Temperature $\leq$ Ambient temperature for time period, Dew Point Temperature $<$ 5°C change from previous hour, Dew Point Temperature $\geq$ 0.5°C from previous hour, and Dew Point Temperature $<$ Ambient temperature for 12 consecutive hours.
Solar Radiation (SR)	Night time SR = 0, Day time SR $<$ max SR for date and latitude

† - These criteria are taken from the EPA *Quality Assurance Handbook for Air Pollution Measurement Systems, Volume IV, Appendix C* (U.S. EPA, 2008) based upon hourly data.

‡ - Criteria was changed from 3 consecutive hours to 6 consecutive hours, and reported to TCEQ in August 2021 monthly progress report.

The approach to identify flagged meteorological data assigns an 18-digit code to each data collection interval, very similar to that described above for the soil water content, only simpler. Each digit and value of each digit correspond the flag listed above. For example:

- Decimal number is 0 in Flag column represents binary number of '000000000000000000' – all flags are low
- 139810 represents '100010001000100010', meaning all values (wind speed, wind direction, temperature (air and dew) and solar radiation) are 'NaN'.
- 64 represents '000000000001000000', meaning that air temperature has not changed over last 12 hours.
- 16384 represents '000100000000000000', meaning that wind speed has been constant over 12 hours.

This approach is preferred over using alphanumeric or special symbols when flags are high, because numeric data saved in CSV format can be more easily imported into a spreadsheet program or any other post-processing code platform, without requiring special coding. In practice, a multi-digit code imported into a spreadsheet will lose leading zeros. The conversion of decimal flags in the data file can be done using `dec2bin(flag value,18)`.

### **3.2.3 Data Collection, Cleansing and Assimilation**

In addition to the data collected by BEG at the four stations installed around Houston, through a data use agreement with the Harris County Flood Control District (HCFCD), BEG was able to access and use data from 25 HCFCD stations located throughout the Houston region, but generally north of downtown Houston. These data were obtained using Application Programming Interface (API) calls, basically, scripts that contact and pull specific types of data from the HCFCD server. These API calls were conducted every 6 hours, and data were stored on secure computers located at BEG. Task 6 of our contract with TCEQ stated that BEG would assess the quality of soil water content data collected from these 25 stations and conduct revisions of calibration curves that incorporate site-specific soil texture, rather than the assumed sandy soil type that the District has been using throughout their network. The task description further states that the latest published calibration curve from Caldwell et al. (2018) would be applied to the collected data.

Toward the end of September 2020, we were provided with an API protocol needed to access their dataset, which was designed by a data management vendor. The stations accessed (Figure 2.1) were equipped with sensors for measuring water content, with three sensors per station, installed at 5, 20, and 40 cm (2, 8, and 16 inches) depth. To further analyze the soil water content data and calibrate HCFCD stations, all data were reanalyzed using the Ledieu et al. (1986) form of the calibration curve, obtained using electrical permittivity data from five soils of a variety of textures, ranging from loamy sand to loam. As described in Caldwell et al. (2018), soils were packed into cylinders at field bulk density at water contents increasing from air dry to near saturation, in steps of approximately  $0.10 \text{ m}^3/\text{m}^3$ , and assessed for water content with the same sensor used by HCFCD (i.e., model CS655, Campbell Scientific, Inc., Logan, UT). A universal calibration curve (water content as a function of relative permittivity) was then fitted, based on the measured laboratory values. This calibration curve (Ledieu equation) was then used to re-calculate water content for all sensor data from the HCFCD stations.

Primary QA/QC protocols, similar to TXSON stations, were performed on the data and possible outliers were replaced with local averages. During the performance period, a long period of outliers was reported from only one station; data for that period also were removed by HCFCD. In general, very few flags were raised after processing the data through the QA protocols, indicating that the data are within physical ranges of the soil and are reacting to environmental changes (i.e., ambient temperature). What is less obvious or apparent is the connection between the soil environment and the potential for groundwater upwelling into the range of the deeper (40 cm) sensor, or potential anthropogenic noise from the built infrastructure (e.g., runoff from impervious surfaces toward the sensors, shading from buildings, etc.). By all visual inspections, the data look useful and within the ranges expected for this area of the country and the soil texture estimated from POLARIS dataset.

### **3.3 Data Archiving and Distribution to TRACER**

Soil water content and weather data collected and stored on the data logger by LoggerNet (Campbell Scientific, Inc. Logan, UT) were uploaded hourly to BEG and stored locally in appended data files in a

workstation in the BEG offices. Files were then archived on an external hard-drive and on an encrypted cloud-based server operated by UT Austin. The data were also sent to the TRACER team using a password-protected cloud folder.

The subsetted SMAP data, over the Houston area of interest, and the nowcasted SMAP data were stored and streamed through a cloud-based server, as discussed in deliverable 7.4 (for Task 7). A copy of the SMAP data covering the area of interest was made available to the TRACER team, also through a password-protected cloud folder. The data were updated as soon as new data were available from NASA. Each file was named for the date and time that the data represent. The current version of machine learning model that produced the nowcasted SMAP data (v1) was also appended to the name of each file. A copy of the nowcasted SMAP data for each date and time is also available to the TRACER team.

### 3.4 What Worked and What Didn't Work

Data collection using the monitoring systems and technologies chosen for this project are designed for remote and environmentally harsh conditions. In general, these systems are extremely resilient and uptime is often near 100%. For the most part, remote data collection from the four stations maintained by BEG did perform as expected. However, we did experience instability in the communications between the station installed at La Porte Airport and our offices in Austin. During March 2022, data drops were detected and troubleshooting required several trips to La Porte, resulting in a swapping of modems, loggers, etc., in order to isolate the source of the instability. The swapping out of equipment resulted in loss of some data during April 2022. By April 20, the problem with the modem was identified, and communications was consistently reestablished. Fortunately, the data loss did not occur during the Intensive Operational Period. Other field issues included the presence of bird nests in the rain gauge at the La Porte site. The organic matter in the nest slowed the response of the gauge to precipitation events. Table 3.5 shows the completeness data for each station, most parameters in the mid-upper 90% complete.

Table 3.5. Completeness data and adjusted flags for meteorological and soil water content measurements, adjusted flags.

Monitoring site	UHCC	La Porte	UHSL	GUYT
Wind speed	94.93	93.79	98.9	95.43
Wind direction	96.50	93.65	98.72	95.56
Air temperature	98.89	94.27	90.53	98.64
Dew point Temperature	97.82	93.73	92.13	94.55
Solar radiation	99.68	94.38	99.74	94.01
Soil Water Content - 5 cm depth	96.64	93.07	95.76	96.23
Soil Water Content - 10 cm depth	96.45	94.12	96.45	97.01
Soil Water Content - 20 cm depth	97.90	95.81	98.30	97.34

#### 4. Development of Gridded Soil Moisture Fields

Satellites specifically designed to measure soil water content include the NASA Soil Moisture Active and Passive (SMAP) mission and the European Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al., 2001), both carrying an L-band microwave radiometer payload (Entekhabi et al., 2010). In both cases, measurements are available with a spatial resolution on the order of 40 km, with a global coverage achieved every 2–3 days. The level 4 (L4) soil moisture algorithm merges the SMAP observations with the soil water content estimates from NASA Catchment Land Surface Model (Catchment-LSM) to produce a high-resolution (9 km) remotely sensed dataset (Reichle et al., 2012). In addition to satellite-based microwave water content estimate missions, researchers have also tested the viability of surface-reflected Global Navigation Satellite System (GNSS) signals to estimate surface soil moisture (Camps et al., 2016). The first dedicated space-borne Global Positioning System (GPS) reflectometry (GPS-R) receiver on board the UK-Disaster Monitoring Constellation satellite (also known as BNSCSAT-1, launched in September 2003) proved that GNSS signals can reliably describe surface conditions, including ocean, snow, and land surface properties (Gleason et al., 2005). NASA’s new weather prediction project, the Cyclone GNSS (CYGNSS) was started in December 2016 and included eight CYGNSS microsattellites that were launched from a single launch vehicle. CYGNSS was designed to measure ocean surface wind fields using a bistatic scatterometer technique with GPS-R receivers, but it can simultaneously measure changes in soil water content around 5 times per day (Kim & Lakshmi, 2018). CYGNSS allows measurements from 32 channels, spread across 8 small satellites, each with 4 channels. Mean revisit time can be as short as seven hours over the ocean. However, because of the ever-changing geometries of the GNSS and CYGNSS satellites, CYGNSS observations have a quasi-random characteristic, which is different from the repeatable swath-like sampling of most remote sensing satellites, such as SMAP. For a 9 km × 9 km grid (the resolution of SMAP L4), the CYGNSS sample rate can fall to 0.1 per day. Spatial or temporal averaging could provide full temporal or spatial coverage. However, this degrades the high-resolution potential of CYGNSS-based soil water estimates, both in space and time. Thus, the soil water algorithm for CYGNSS uses collocated retrievals from SMAP to calibrate CYGNSS observations from the same day. For a given location, a linear relationship between the SMAP soil moisture and CYGNSS reflectivity is determined and used to transform the CYGNSS observations into soil water content. The differences in SMAP-based and CYGNSS-based data comes down to spatial resolution, temporal resolution and accuracy/reliability of the data. In our study, we tested both SMAP and CYGNSS data against the in-situ water content data and decided to use SMAP data.

##### 4.1 Background on Approaches

The goal of our work was to retrieve accurate estimates of soil water content over a target area that included Houston, Texas, at fine grid resolutions (in this case, 9 km). We compared several sources of water content data including satellite- and in-situ based. Our approach includes: 1) Download and harmonize the CYGNSS, SMAP and in-situ water content data at the required spatial and temporal scales; 2) carry out a set of statistical steps to study the relationships between these data sources; and 3) study the effects of a set of ancillary variables on CYGNSS-derived surface reflectivity. We investigated

whether a linear relationship can sufficiently describe the relationship between CYGNSS reflectivity and water content, and derived insights into the spatiotemporal complexities of these relationships.

#### 4.1.1 Comparison of SMAP, CYGNSS and In-Situ Data

We studied a 36x36 km region, which is home to the Texas Soil Observation Network (TxSON). The site is located in the central Texas Hill Country near Fredericksburg, Texas (Figure 4.1).

This region is representative of the terrain of the semiarid rangelands of Texas Hill Country. Vegetation over this region includes oak trees (red, live, and post), woody plants (honey mesquite), and a mixture of

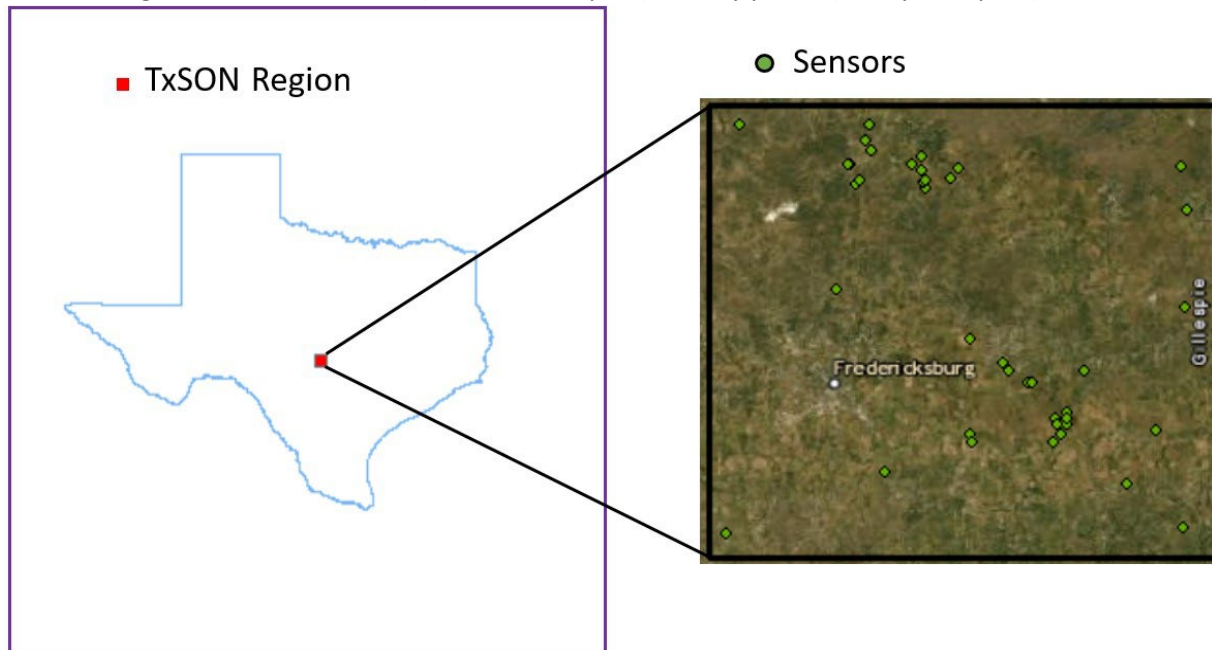


Figure 4.1. Study area to compare SMAP, CYGNSS and in-situ data.

short and mid-height grasses (switchgrass, bluestem, curly mesquite). The soils are generally not appropriate for small grain or row crop production due to high erosion rates, shallow depths, and low water retention capacity, but they are well suited for grazing and viticulture. The 30-yr mean annual precipitation is 807 mm and air temperature is 18.4°C. Because the area has few development/buildings, this location is more suitable for comparing the results of SMAP and CYGNSS, than the more urban areas around Houston. Moreover, comparison of SMAP and CYGNSS is not representative in regions with high surface water fraction, dense vegetation and urban and mountainous areas. O’Neill et al. (2016) and Chew & Small (2018) also argued that comparing water content estimates using various products within urban areas can be difficult. NASA’s Short-term Prediction and Transition Center – Land Information System (SpoRT-LIS) also does not provide the nowcasted soil water content over urban areas because of the uncertainty in measurements over urban areas.

In this study, we used 3 months of CYGNSS acquisitions: January, April and July 2019, to develop and test the retrieval algorithms. SMAP Enhanced, L3 Radiometer, Global, daily 9-km Equal Area Scalable Earth-Grid (EASE Grid, v2.0) data are also used for the same months. Data include water content data, quality control flags, and other auxiliary information, with ascending and descending passes averaged together to form a single daily pass.

Here, hourly in-situ soil water content data are collected and used as ground truth from 40 monitoring sites, which constitute the Texas Soil Observation Network (TxSON) region. The spatial distribution of sensors within the 36 km TxSON region (Figure 4.1) are nested into 3, 9 and 36 km areas and provides mean hourly water content values at 4 separate depths (5, 10, 20 and 50 cm). Precipitation is also measured at all locations with six meteorological stations also providing air temperature and humidity, wind speed and direction, and solar radiation. These data are publicly available through the Texas Data Repository (Dashtian and Young, 2023).

In this study, we only use 5 cm water content readings provided because the CYGNSS-derived reflectivity originates from the topsoil layer (0–5 cm). To conduct this analysis, the three data sources were harmonized at the same spatial and temporal scales (daily 9km grid resolution). While SMAP data were available at the required resolution, the raw CYGNSS and TxSON data were upscaled and aggregated appropriately. In both cases, we employ the Voronoi upscaling method (as done by Caldwell et al. (2019)) to bring both data sources to a 9 km grid resolution. For the CYGNSS data, individual observations at sub-daily scale were ignored and daily observations were directly gridded to the same SMAP grid. Because the TxSON data provides water content estimates every hour, the Voronoi method was first applied to obtain hourly gridded in-situ measurements. Then, within each grid, hourly estimates were aggregated to daily estimates for each 9 km grid over the TxSON region.

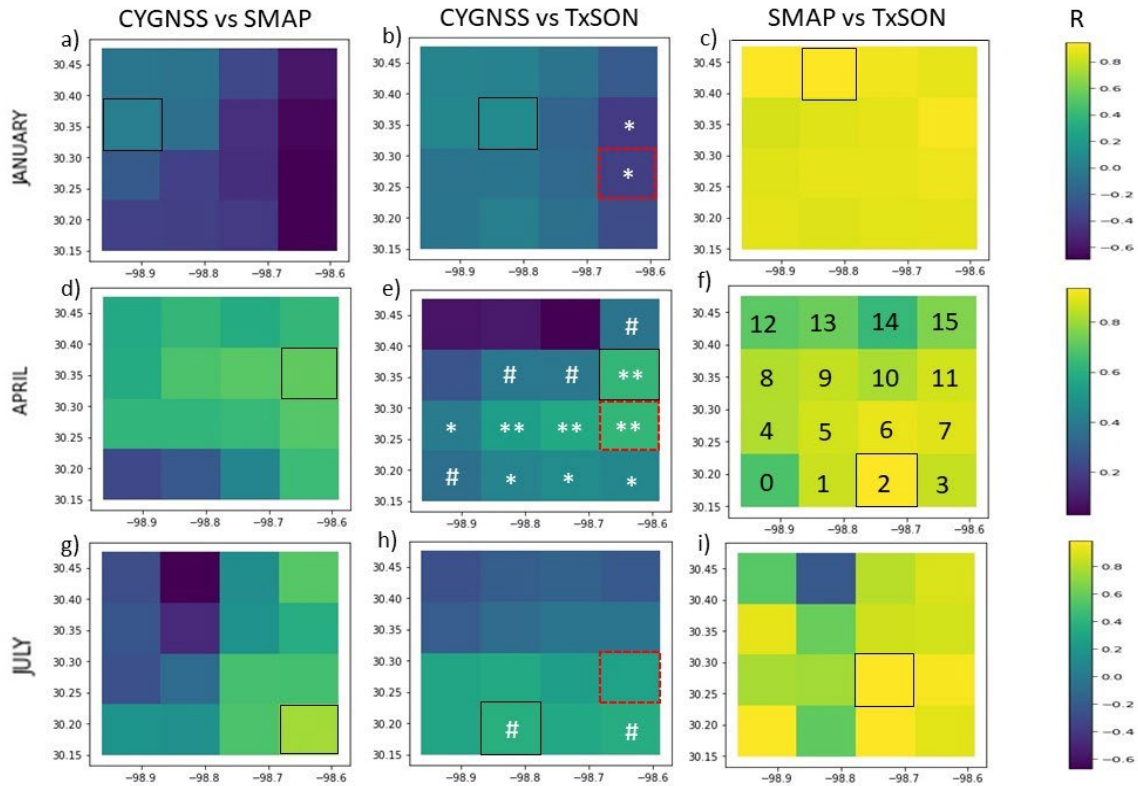


Figure 4.2. Spatial correlation heatmaps for pair-wise comparisons between CYGNSS, SMAP and TxSON. Grids outlined in black represents grid with highest R value. For comparisons between CYGNSS vs TxSON (grids in second column), the level of significance for each grid is shown (# significant at 0.1 level, \* - significant at 0.05 level, \*\* - significant at 0.001 level). F) shows the grid numbering followed for all heatmaps.

The correlation coefficient for each grid (16 grids total, each 9 km x 9 km) (Figure 4.2) shows, as expected, high correlation between TxSON and SMAP soil water content data ( $r \geq 0.5$ ) (Figure 4.2 c, f, i) for nearly all 9km grids across the three months. However, spatial variation and monthly shifts in correlation trends in the CYGNSS-SMAP and CYGNSS-TxSON comparisons seem to exist.

We first analyze the correlations from a spatial perspective. Using the CYGNSS vs TxSON comparisons, for example, at a monthly scale, grid cells in the eastern region in January are strongly negatively correlated (grids 7 and 11 significant at 0.05 level). As the vegetation leafs-out during the spring, we noted positive correlations in the central and eastern regions (grid cells 5, 6, 7, and 11 significant at the 0.001 level) in April, in the south and south east regions (grids 1 and 3 significant at the 0.01 level) during July. At a sub-monthly (approximately daily) scale, correlation values between the 9 km grids are not stable and changed drastically. Differences between the maximum and minimum correlations for the three months is 0.498, 0.600 and 0.638. We also noted substantial changes in absolute values of R for grid cells and changes in significance levels over the three months. For example, the R values for cell 7 (highlighted in red in Figure 4.2h) are -0.39 (significant at the 0.05 level), 0.628 (significant at the 0.01 level) and 0.279 (not significant at the 0.1 level) for January, April, and July, respectively.



In Figure 4.4, the CYGNSS observations are compared to the Soil Moisture Active Passive (SMAP), in-situ Texas Soil Observation Network (TxSON) and NASA's CyGNSS L3 soil moisture (SM) measurements for the entire 2019 year for the area shown in Figure 4.1.

To further analyze the daily variability, we plot scatter plots in Figure 4.3 for grids with the highest correlation value in each heatmap (outlined in black in Figure 4.2). Although there is a good agreement between the data, given the different scales, instruments and methods that data were acquired, the SMAP seems to better represent the changes in soil water content due to precipitation. For example, the increase in soil water content (as shown in TxSON data) in May, was captured by SMAP while the magnitude of change in CYGNSS data is smaller than that shown in ground-truth data. Therefore, a linear correlation between CYGNSS-derived surface reflectivity and soil water content varies significantly both spatially and temporally. This demonstrates that a simple linear model to estimate soil water content from CYGNSS reflectivity in this area may not be sufficient. Either other land physical parameters need to be considered or a more complex non-linear relations needs to be used. As a result, we decided to use SMAP data over the Houston area, given that it is an urban area and the variation of soil water content is considerable. Furthermore, while the use of CYGNSS to derive soil water content is still underdevelopment, SMAP has been providing water content data for various application since 2010.

#### **4.1.2 SMAP**

After comparing various SM products, we selected SMAP L4 data as the main source of satellite data to perform other tasks. We obtained processed SMAP data (Level 4) for surface and root zone soil moisture (100 cm depth) from the National Snow and Ice Data Center (nsidc.org). The data were generated by assimilating the atmospheric forcing from SMAP sensors (e.g., brightness temperature data) into the precipitation data from NASA-IMERGE product and rescaled using the Global Precipitation Climatology Project (GPCP) v2.3 product. R Studio and Python-based scripts were used to download SMAP data from NASA as soon as they are available, though data were often latent by 2-5 days, sometimes much longer. We then subsetted the data to cover only AOI over the Houston area. A separate Python script processed and visualized the data, and prepare them for the machine learning algorithm described briefly below and more fully in the Task 7.3 report submitted to TCEQ. Figure 4.5 shows an example of SMAP-derived water content for January 1, 2022, which was processed using the machine learning approach to address latency. Data formats and file designations are described below.

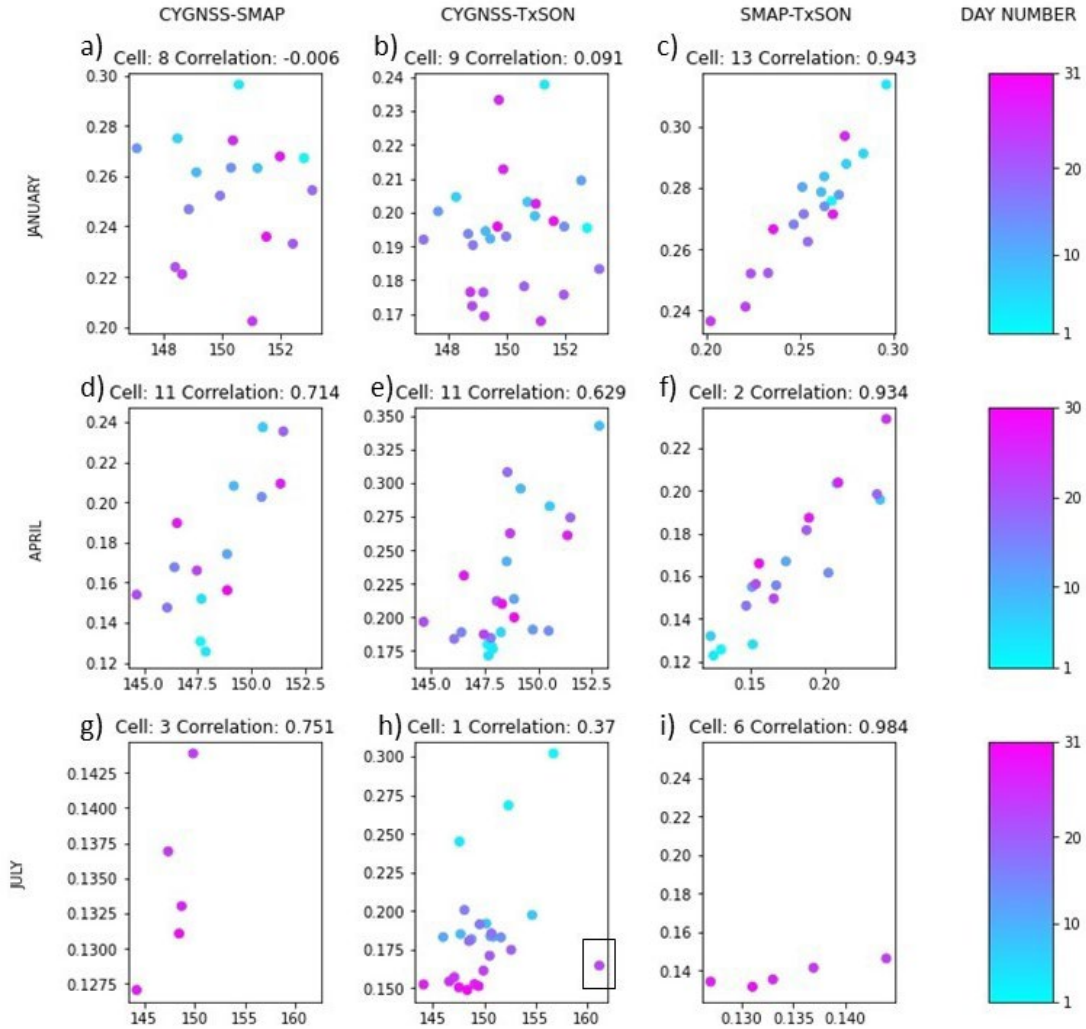


Figure 4.3. Scatter plots for the grid (outlined in black in Figure 4.2) with the highest R value in each heatmap in Figure 4.2. Data point within box in h) represents an outlier.

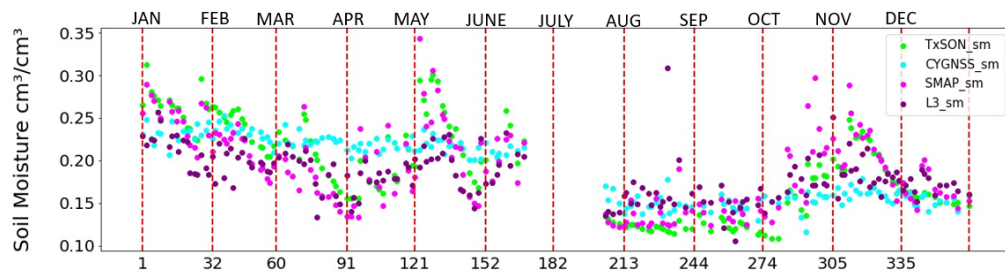


Figure 4.4. Comparison of SM retrieved from various methods including in-situ (TxSON) and satellite (CYGNSS, CYGNSS L3 and SMAP).

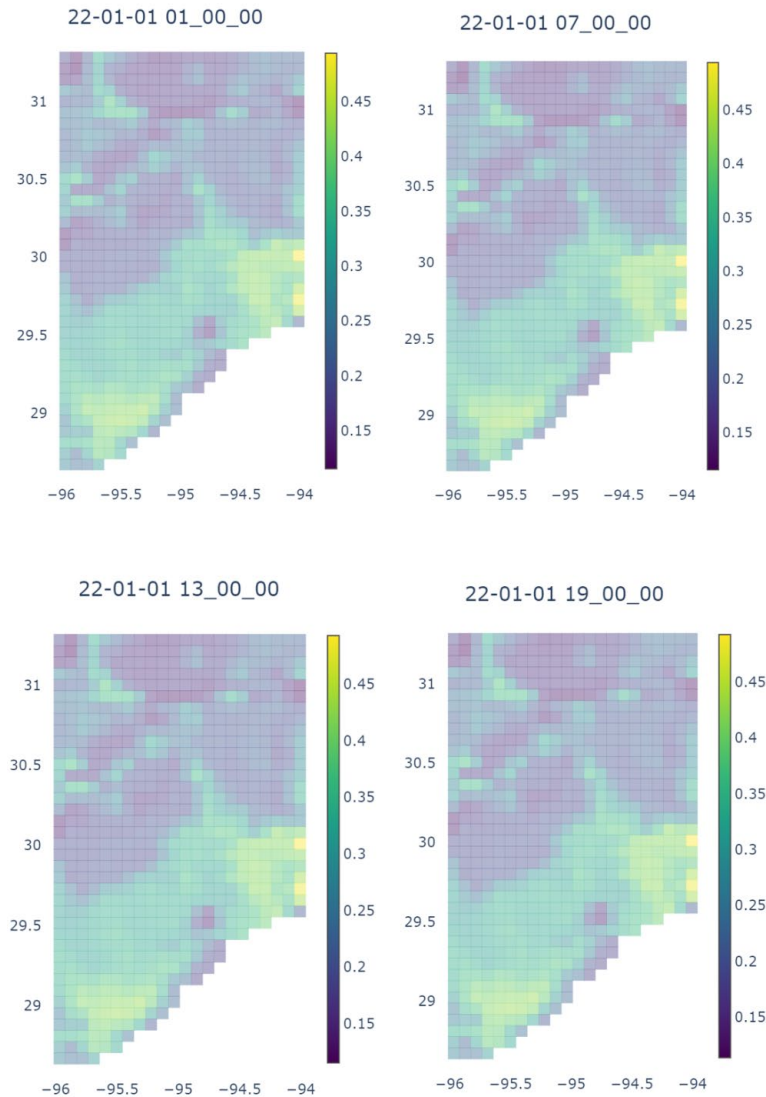


Figure 4.5. Example gridded soil water content images obtained by NASA’s SMAP assimilated product for January 1, 2021. Each image represents water content at 6-hour time increments as shown above each image.

Two versions of SMAP data are available: publicly available data and nowcasted data, both of which are subsetted to the Houston AOI. Regarding the former, subsetted publicly available SMAP data are saved in 6-hr increments, in comma-separated value (CSV) format. CSV is the most flexible format available, allowing users to easily import into mapping or other graphically based applications. File names are in the format of WWWW-XX-YY-ZZZZ\_vi, where WWWW is the year; XX is the month; YY is the day; ZZZZ is the time in 24-hr format and vi indicates the version of our code that is used to produce the results. The current version is v1. Each file contains four columns: (1) SMAP cell number in subsetted region, as shown in Figure 4.6; (2) latitude in WGS-84 reference system; (3) longitude WGS-84 reference system; and, (4) soil moisture value in units of m<sup>3</sup>/m<sup>3</sup>.

Regarding the nowcasted version of SMAP data, information is saved in 6-hr increments, in comma-separated value (CSV) format. File names are in the format of WWWW-XX-YY-ZZZZ\_vi, where WWWW is the year; XX is the month; YY is the day; ZZZZ is the time in 24-hr format; and, vi is the version number of the Python code that created the nowcasted data set for that time. Each file contains four columns: (1) SMAP cell number in subsetted region, as shown in Figure 4.3; (2) latitude in WGS-84 reference system; (3) longitude WGS-84 reference system; and, (4) soil water content value in units of  $m^3/m^3$ .

#### **4.1.3 Comparison of SMAP and In-Situ Data over Houston Area**

Figure 4.7 shows the soil moisture recorded from 25 stations during 2022 (to date) operated by HCFC. The data are highly variable when compared against one another because the stations are installed across a variety of urban, suburban and non-urban areas. That said, the results do show strong correspondence to precipitation events, and similar decreases in value due to prolonged periods without rainfall.

At the conclusion of the TRACER Intensive Observation Period (IOP), two of the four BEG-installed stations were demobilized (e.g., La Porte Airport and Guy, TX sites). The remaining two sites (University of Houston (UH) Coastal Center and UH Sugar Land) are still in operation. Those data are being streamed through the TxSON website (<https://www.beg.utexas.edu/research/programs/txson/map>), and will continue to be operated through BEG-internal funds. Figure 4.8 shows the soil water content and precipitation data collected at those two stations for calendar year 2022. The interactive graphs of these data are also available. Below are webpages with a more interactive web-based visualization of the soil water content and soil temperature data for the following sites:

Soil Water Content:

UH Coastal Center: [https://coastal.beg.utexas.edu/soilmoisture2/data/CR300\\_17/SM\\_P\\_plot.html](https://coastal.beg.utexas.edu/soilmoisture2/data/CR300_17/SM_P_plot.html)

La Porte Airport: [https://coastal.beg.utexas.edu/soilmoisture2/data/CR300\\_18/SM\\_P\\_plot.html](https://coastal.beg.utexas.edu/soilmoisture2/data/CR300_18/SM_P_plot.html)

UH Sugar Land: [https://coastal.beg.utexas.edu/soilmoisture2/data/CR300\\_19/SM\\_P\\_plot.html](https://coastal.beg.utexas.edu/soilmoisture2/data/CR300_19/SM_P_plot.html)

Guy TX: [https://coastal.beg.utexas.edu/soilmoisture2/data/CR300\\_20/SM\\_P\\_plot.html](https://coastal.beg.utexas.edu/soilmoisture2/data/CR300_20/SM_P_plot.html)

Soil Temperature:

UH Coastal Center: [https://coastal.beg.utexas.edu/soilmoisture2/data/CR300\\_17/ST\\_P\\_plot.html](https://coastal.beg.utexas.edu/soilmoisture2/data/CR300_17/ST_P_plot.html)

La Porte Airport: [https://coastal.beg.utexas.edu/soilmoisture2/data/CR300\\_18/ST\\_P\\_plot.html](https://coastal.beg.utexas.edu/soilmoisture2/data/CR300_18/ST_P_plot.html)

UH Sugar Land: [https://coastal.beg.utexas.edu/soilmoisture2/data/CR300\\_19/ST\\_P\\_plot.html](https://coastal.beg.utexas.edu/soilmoisture2/data/CR300_19/ST_P_plot.html)

Guy TX: [https://coastal.beg.utexas.edu/soilmoisture2/data/CR300\\_20/ST\\_P\\_plot.html](https://coastal.beg.utexas.edu/soilmoisture2/data/CR300_20/ST_P_plot.html)

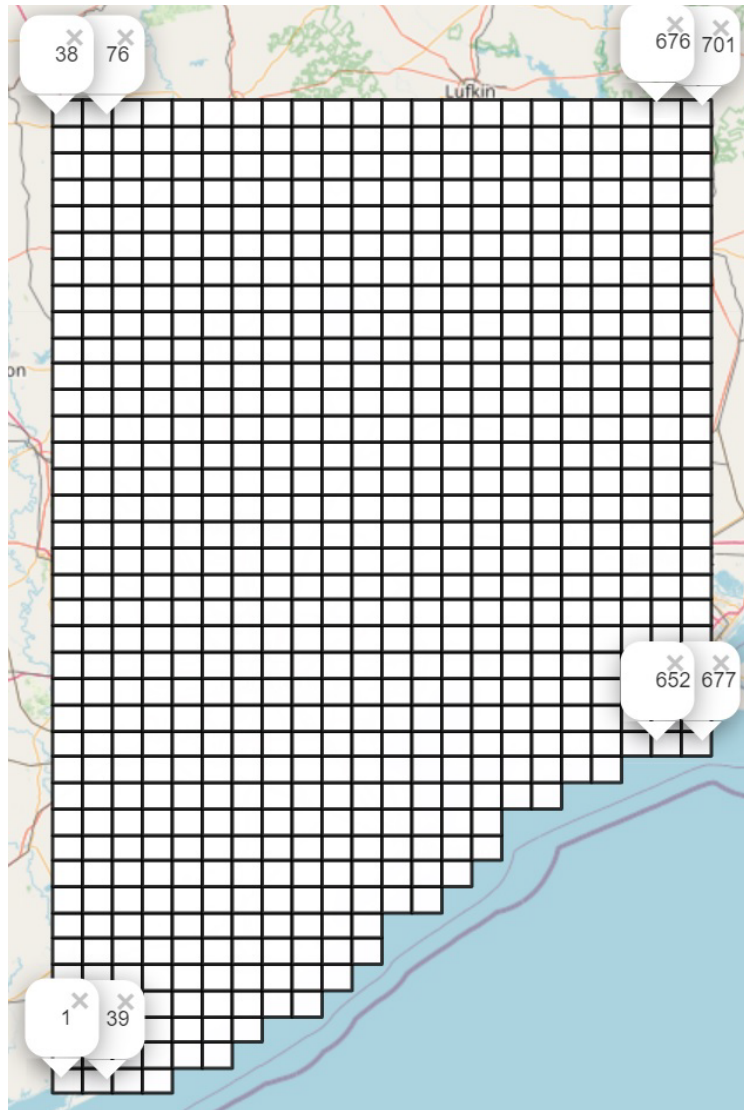


Figure 4.6. Index map showing grid designations for subsetting SMAP data and nowcasted SMAP data.

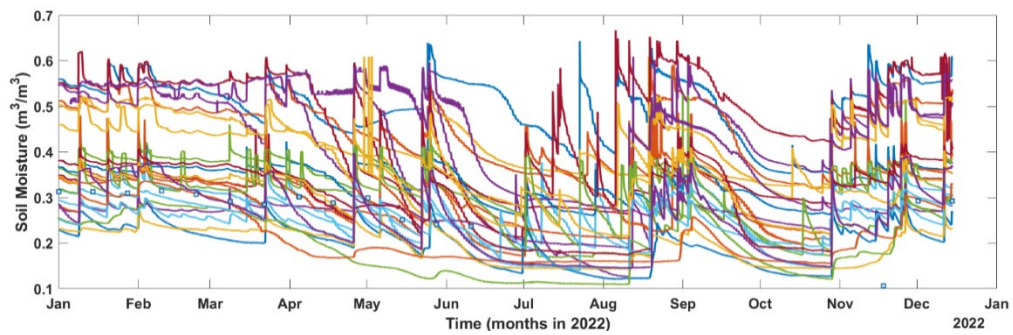


Figure 4.7. Soil moisture data collected at all 25 HCFC stations at 20 cm depth during 2022.



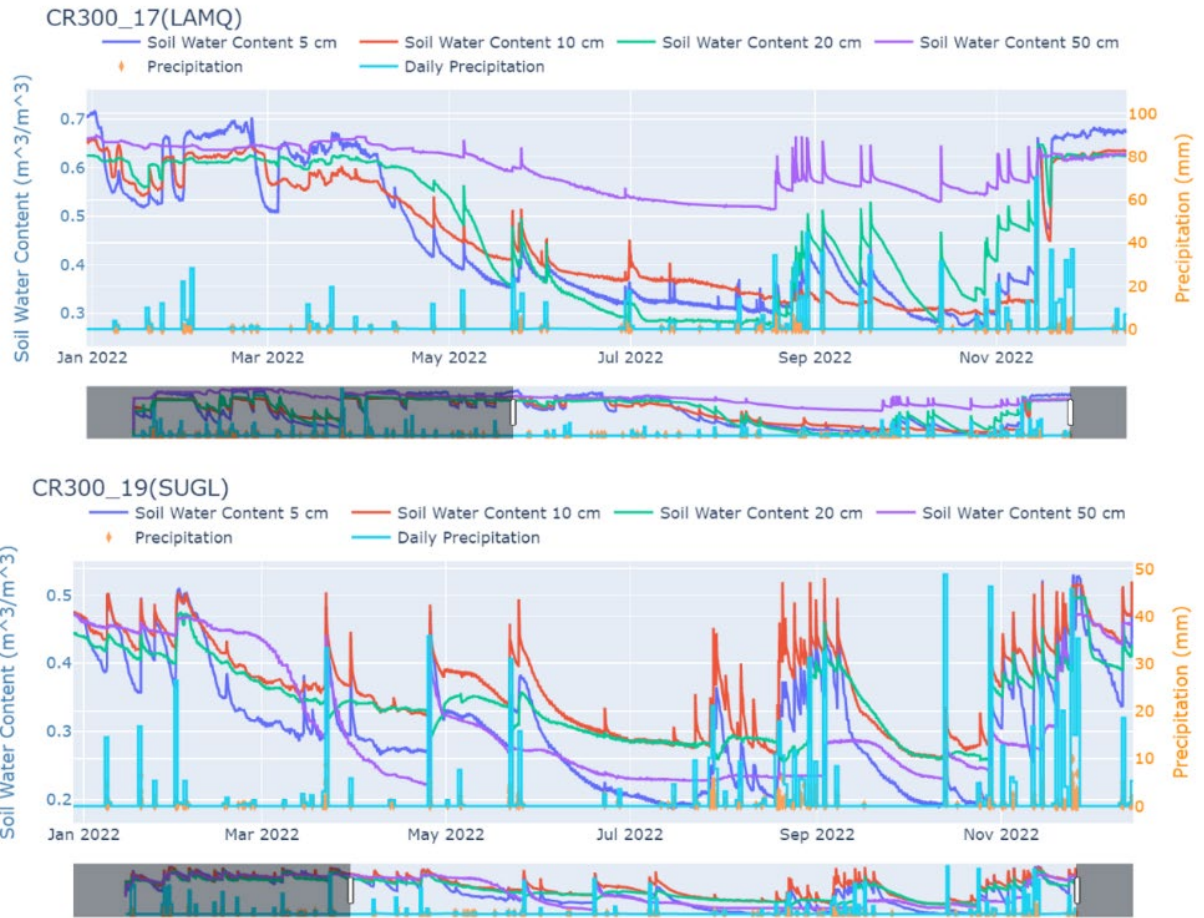


Figure 4.8. Soil water content at 5, 10, 20 and 50 cm depths for the UH Coastal Center (LAMQ) and UH Sugar Land (SUGL) stations, along with the hourly and daily precipitation. The charts show data from January to December, 2022.

We also compared the in-situ SM data with the SMAP L4 data. We developed source code to download and compare SMAP data for each grid to in-situ data from the geographically closest station. Each station has four sensors at 5, 10, 20 and 50 cm depths. For each collection time,  $t$ , we use the following equation to calculate the depth-averaged water content at each station from ground surface to 75 cm depth:

$$SM_m(t) = \frac{1}{\sum_{i=1}^4 h_i} \sum_{i=1}^4 SM_i(t) * h_i, h_1 = 5, h_2 = 10, h_3 = 20, h_4 = 50$$

SMAP L4 water content data expressed as the average over 100 cm depth over a 9\*9 km<sup>2</sup> area. Note that the area, the depth and the method that satellite based and in-situ data represent and use to report the SM is different and the comparison between these two data sets includes some levels of uncertainty due to these factors. The SMAP measurements of land surface microwave emission (or brightness temperature) and radar backscatter at L-band frequencies provide information on surface soil moisture (top 5 cm of the soil column) and on the freeze-thaw state of the land surface. To provide estimates of root zone (top 100 cm soil column) a specific algorithm is used that combines SMAP

observations with soil moisture estimates from the NASA Catchment land surface model. The algorithm is based on the ensemble Kalman filter (EnKF) and is called L4\_SM. Various versions of the algorithm has been developed with the version 7 of the if being the latest one that was released on November 2022. This report includes data from version 6 and 7 of L4\_SM. The NASA Catchment model describes the vertical transfer of soil moisture between the surface and root zone reservoirs. The model is driven with observation-based surface meteorological forcing data, including precipitation, and runs on a global 9 km Earth-fixed grid with a 20 min model time step (Reichle et al., 2022).

The in-situ sensors, however, usually consist of two 12-cm-long stainless-steel rods connected to a printed circuit board. The circuit board is encapsulated in epoxy and a shielded cable is attached to the circuit board for data logger connection. The sensor measures propagation time, signal attenuation, and temperature. Dielectric permittivity, volumetric water content, and bulk electrical conductivity are then derived from these raw values.

Figure 4.9 compares the ground-based sensor derived SM and satellite-based SM in UHSL area. The data shows good correspondence between ground-based and satellite derived SM data. The response to the precipitations is more obvious (with higher fluctuations) in ground-based data. This can be due to the fact that ground-based data represents an average of a few square meters while satellite data represents 9\*9 square kilometers of area. The area is considered an urban area and includes buildings, pavements and concrete and open areas.

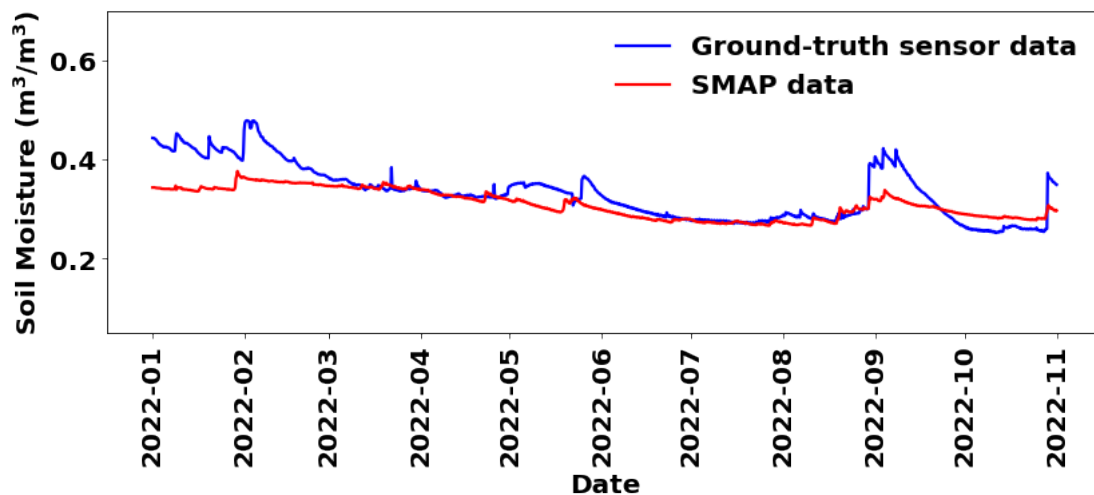


Figure 4.9. Example of a comparison of soil moisture derived from the average of four sensors at the UH Sugar Land location and the SMAP L4 grid data over this location.

Another way to compare the two time-series is through their probability distribution functions. In Figure 4.10, we use violin plots to show the distribution of SMAP and in-situ sensor data. These plots include a marker (small white circle) for the median of the SMAP and ground-truth sensor data and a small box (thick black line) indicating the interquartile ranges of the time series. Both distributions for SMAP and sensor data show bimodal distribution indicating that data represents two different categories associated with wet and dry periods. The spread of in-situ sensor data is much higher than the SMAP

data, showing more variation over time. For this specific SMAP grid, the median soil moisture is 0.312 m<sup>3</sup>/m<sup>3</sup> and the median for the sensor is 0.336 m<sup>3</sup>/m<sup>3</sup>.

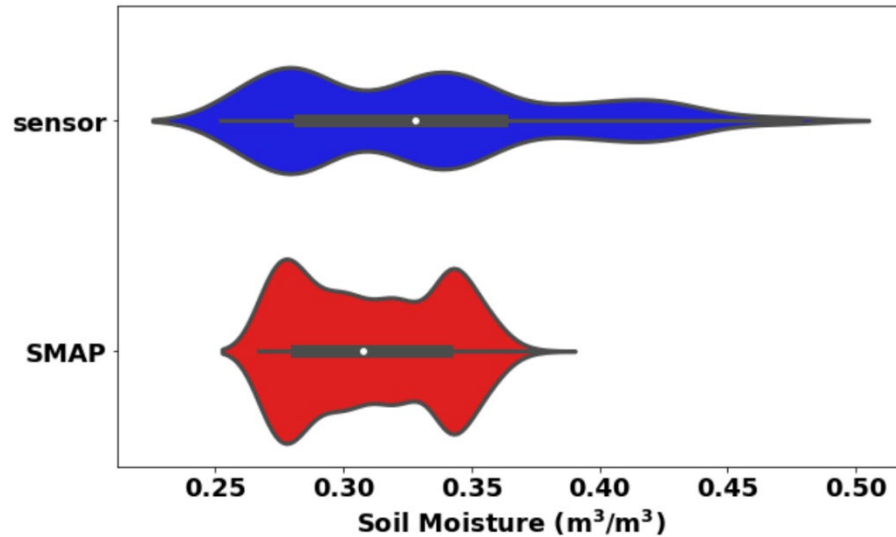


Figure 4.10. Violin plot of the SMAP derived (red) and in-situ derived (blue) soil moisture time series.

#### 4.2 Issues related to latency, why it matters, and how it impacts weather modeling

SM data with high spatiotemporal resolution are valuable for many purposes such as agricultural product irrigation (Eeswaran et al., 2021), drought and flood monitoring (Souza et al., 2021; Wasko & Nathan, 2019). However, current SM products (including ground-based, satellite-based and model-based data) have various limitations. Ground-based SM data are accurate but are limited in spatial coverage (few inches each sensor). Satellite-derived SM data products such as SMAP and SMOS missions have resolutions too coarse to resolve the soil moisture heterogeneity below the 10 km scale. Land surface models can simulate soil moisture at high resolution, but errors in model structure and parameters can limit their predictive accuracy and prevent them from fully exploiting the information contained in the observations (Nearing et al., 2021). The SM data are needed in climate models as water balance considerations at the soil surface lead to an equation that relates the autocorrelation of soil moisture in climate models various land-atmosphere. Assimilated satellite- and ground-based SM data are valuable source of SM that can be used in climate models in regional studies such the area of interest over Houston, TX. Nevertheless, there exist temporal gaps often in SMAP products, which can limit their application. This issue is often more pressing in areas where soil moisture retrieval fails, or is flagged as unreliable, due to radio frequency interference (RFI), dense vegetation, or intense rainfall. To enhance the accuracy and spatiotemporal coverage SMAP L4 has been deployed to produce a time-continuous data assimilation product based on the assimilation of SMAP measurements into a land surface model (LSM). However, these advantages come at the expense of slightly increased data latency (average of about 2.5 days) due to a time lag incurred by the use of gauge-based precipitation as a required input for the SMAP L4 analysis. Thus, the SMAP data temporal gaps limit the use of its values in near-real-time climate model applications such as flood risk assessment. Considering this, we have developed a product that uses available SMAP L4 data and ground-based SM data in Harris county, and



generates near real-time SM gridded data (similar to SMAP L4) that can be used in climate models for real-time forecasting and modeling.

## **5. Nowcasting of Soil Water Content**

### **5.1 Nowcasting SMAP Using Deep Learning Approach**

We have developed a machine-learning algorithm to assimilate in-situ and satellite-derived soil water content data, and to train the model to nowcast soil water maps at the same scale as the satellite. Usually, satellite-derived data are not available for the current time, but rather are 2-5 days delayed between the overflight and when data are published. This is known as latency, which restricts applications of satellite-derived soil moisture data in the forecasting of environmental events, such as flash floods. We developed a specific algorithm that combines Long Short-Term Memory (LSTM) (Fang et al., 2017; Liu et al., 2021; Liu et al., 2022), Principal Component Analysis (PCA) (Amato et al., 2020) and a fully connected neural network (FCNN) (Amato et al., 2020) to nowcast soil moisture in space and time. The LSTM is used mainly to nowcast time series in time, and the PCA and FCNN are used to nowcast the soil moisture in space. The input data includes historical SMAP data over the Houston area (we use 701 SMAP grids in our analyses) and in-situ data from 29 monitoring stations (25 stations from HCFC and 4 stations from BEG). The most recently published SMAP data with latency is then parameterized in  $n$  steps, with each step being 6 hours, but the in-situ soil water content data from each station are up-to-date. We use the ground truth soil water content data as a guiding time series for the SMAP data, up to the most recent availability of SMAP data. For example, if the current time is  $t$ , then the 29 time series datasets, from  $(t-n)$  to current time, are used as auxiliary information to nowcast (estimate) soil moisture for all 701 SMAP grids.

### **5.2 Results of Approach**

The nowcasted data includes  $n$  estimated rasters over the Houston area, corresponding to the missing (not yet available) SMAP data. The resulting processing essentially brings latent SMAP data to near-real time, which allows it to be operationalized for flood risk estimates, or other environmental risks. The numerical approach and the raw data from the in-situ stations were shared with TRACER collaborators from the Pacific Northwest National Laboratories, who were tasked with process-level atmospheric modeling. The nowcasted data over Houston is being published as an online tool ([https://coastal.beg.utexas.edu/soilmoisture2/TRACER\\_SM\\_P.html](https://coastal.beg.utexas.edu/soilmoisture2/TRACER_SM_P.html)).

Figure 5.1 (a) shows the front end of the web tool, which allows the user to toggle on/off specific layers as time approaches current. Figure 5.1 (b) shows an example of a water content raster over Houston area, the predicted (nowcasted) raster using our method and the absolute error of estimation. Errors below  $0.01 \text{ m}^3/\text{m}^3$  highlight the accuracy of the method. Figure 5.1 (c) shows the comparison of actual and nowcasted water content time series, throughout most of 2022, at grid #420 in our study area. We have been nowcasting the data since July 2022 and the error of estimation remains very low. Usually, precipitation affect the error of estimation, and other meteorological factors, land cover and land use, vegetation, and soil type are also important variables. These are being addressed through ongoing research. Such a tool and assimilated data sets will enable us to better understand the interplay of these factors over long time periods and in different land surface conditions. Furthermore, real-time training

on streaming water content data and near real-time scoring of machine learning models (i.e., by minimizing errors in estimates) helped us to highlight the spatiotemporal variability of nowcasting. This means that, most likely, a unique machine learning model will not perform well at larger (subcontinent) scales, though region-specific models are possible for nowcasting and forecasting soil water content.

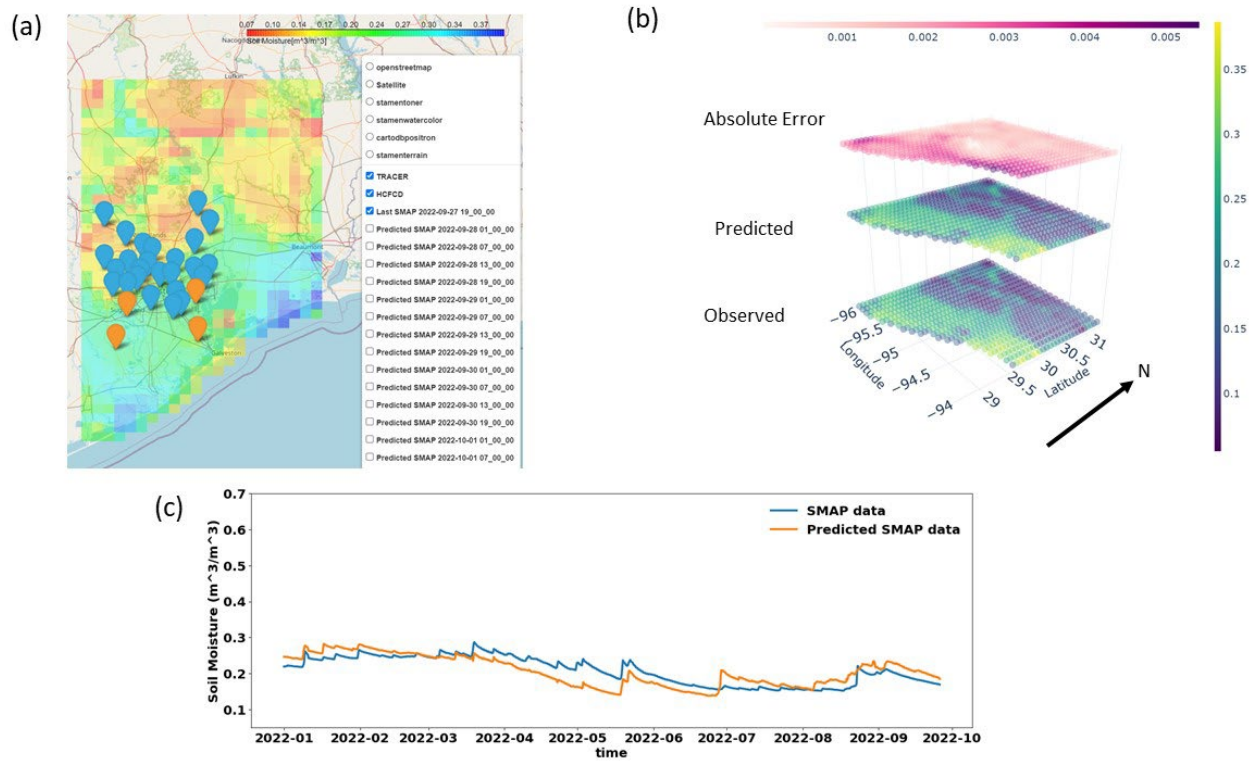


Figure 5.1. a) The front end of the web-tool for nowcasting soil water rasters over Houston area; b) example of observed raster from SMAP L4, predicted (nowcasted) and its associated absolute error raster over Houston; c) comparison of machine learning nowcasted soil water content data and SMAP-derived data for a specific 9 km x 9 km grid in the study area.

## 6. Conclusions and Discussions

### 6.1 Value and Need for Ground-based Soil Water Sensors

Ground-based, soil water content data are usually considered “true” or ground-truth data because, in contrast to satellite-based data, the soil water sensors are in contact with soil bulk, reducing noise and possible interferences from environmental factors. Of course, because ground-truth data are at significantly smaller scale than satellite-based data—often at the 10s cm<sup>3</sup> volume versus 10s km<sup>3</sup>—sensor placement can be chosen to avoid most interferences. Nowcasting soil water content requires in-situ data, not only for ground-truth but also as forcing factors in models. For example, we found that errors in nowcasted soil water content on SMAP grids increases with distance between the center of the grid to the guiding ground-based sensor. Therefore, densifying in-situ sensor networks would help to minimize errors for specific regions. Assimilated and harmonized data from various ground-based monitoring networks is an important step, and a vital source of ground-truth soil water content data for validation and calibration of other data products, especially satellite-derived and model-based. Depending on the area of interest, this could require regional and national collaboration between various agencies.

To maintain a monitoring presence in the areas southeast and southwest of Houston, which is sparsely covered by HCFCD, we opted to maintain two stations after the end of the TRACER project. These two stations provide soil water content data of deeper soils compared to HCFCD stations. We have noted that data collected at the UHSL stations is very similar to that recorded by SMAP. Maintaining these two stations also is in line with the long-term goal of expanding the TxSON network to cover various regions of Texas, thus expanding the spatial calibration/validation range of satellite-derived products. These monitoring stations can serve as calibration/validation data sets for future satellite-derived products, including NASA CYGNSS, or other field campaigns by different research groups.

### 6.2 Need for Nowcasting and Forecasting Soil Water Content

Soil water content is widely recognized as a key parameter in the mass and energy balance between land surface and the atmosphere. The potential societal benefits of accurate nowcasting and forecasting of soil water content are immense. Soil water content forecasts are vital for environmental monitoring, early warning of rapidly occurring geological hazards (e.g., floods, extreme weather events). Over the past 10 years or so, significant improvement in data assimilation has led to regional and global data products that include soil physical and hydraulic properties (Hengl et al., 2017; Chaney et al., 2019), from which soil water forecasting can be improved and operationalized, benefitting hydrology, agriculture, and community well-being. While measurements of soil water content require instruments and infrastructure, the nowcasting and forecasting part needs algorithms and mathematical models. The two approaches provide significant leveraging opportunities (Byun et al., 2011).

At the same time, we note that the benefits of availability of nowcasted soil water content are limited by the ability of the modeling community to operationalize the data. During the TRACER project, for example, real-time soil water and meteorological data were not used directly in predictive models, in part because the significant computational load and data requirements limit real-time, place-based simulations. Furthermore, the focus on process-based climatology by TRACER modelers is not well suited for operationalizing model outcomes. This type of process-based modeling, while vital for

understanding climate systems, benefits from longer time series datasets for specific regions. Our stations near Houston were operating for between 6 months to 1.5 years, likely too short for understanding temporal variability in weather patterns and responses. Improved statistical techniques can provide extrapolated spatio-temporal data for future studies.

Writ large, as the scientific and engineering communities further develop sensor technologies, communication protocols between sensors and models, and the computational architecture to simulate regional atmospheric processes in real-time (all of which are happening quickly), complete connection between data and early warning will be realized. When these technologies are better developed and available, results can be used for rapid decision making.

### **6.3 Future Use of Data and Approaches**

Besides the current use of data from monitoring stations installed in Houston, used in this project for calibrating and nowcasting satellite-derived data, these types of data can be used in any application that involves soil-atmosphere interactions. For example, as Joung & Buie (2015) demonstrated, the soil water content can change the mechanism and amount of aerosols that raindrops may generate on the soil surface. Other variables affecting the rate of aerosol production include degree of wettability, infiltration rate, and precipitation rate, among other things, several of which were monitored in the stations deployed in this study. Long-term records of soil water content data in a specific region can also provide valuable information when used with statistical methods to transfer raw water content data into more useful metrics; for example, by comparing hourly water content observations against longer-term historical conditions in Houston, we can provide the drought status for each site at a specific time and date. Such data and information are useful for U.S. Drought Monitor (USDM), irrigation scheduling, potential for dust emission, etc. Outside of this project, we have attempted to forecast soil water content by blending point-scale meteorological data with forecast weather information from the National Blend of Models (NBM), a NOAA product designed for use by the transportation sector, especially for air travel. The four stations installed in Harris County included instruments that simultaneously recorded meteorological data. Such data with that frequency is rare to find and we anticipate that our data will be used by others to develop machine learning and deep learning models for forecasting purposes. The data provides the opportunity to study correlations and causality between soil moisture and meteorological data.

As we discussed before, the current SMAP data resolution is 9 km x 9 km. Ground-based data can be used to down-scale satellite-derived data to achieve higher resolutions data (e.g., Fang et al., 2021). As the scientific and engineering communities continue downscaling further, products that nowcast soil water content over urban areas will become vital for flood risk mitigation; even NASA's SPORT-LiS product excludes urban and developed areas. Improving nowcasting tools by incorporating land cover land use data in and around urban areas would immediately leverage future downscaled SMAP data and focus these products and predictions on areas currently unavailable.

## 7. References

- Adams, H. D., Zeppel, M. J., Anderegg, W. R., Hartmann, H., Landhäusser, S. M., Tissue, D. T., ... & McDowell, N. G. (2017). A multi-species synthesis of physiological mechanisms in drought-induced tree mortality. *Nature ecology & evolution*, 1(9), 1285-1291.
- Amato, F., Guignard, F., Robert, S., & Kanevski, M. (2020). A novel framework for spatio-temporal prediction of environmental data using deep learning. *Scientific Reports*, 10(1), 1–11.
- Awal, R., Fares, A., & Habibi, H. (2019). Optimum turf grass irrigation requirements and corresponding water-energy-CO<sub>2</sub> Nexus across Harris County, Texas. *Sustainability*, 11(5), 1440.
- Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S. I., Van Den Hurk, B., ... & Milly, P. C. (2016). Land–atmosphere feedbacks amplify aridity increase over land under global warming. *Nature Climate Change*, 6(9), 869-874.
- Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts (1995): The Oklahoma Mesonet: A Technical Overview. *J. Atmos. Oceanic Technol.*, 12, 5-19. (DOI: 10.1175/1520-0426(1995)012<0005:TOMATO>2.0.CO;2).
- Byun, D., H.-C. Kim, and F. Ngan, 2011: Improvement of meteorological modeling by accurate prediction of soil moisture in the Weather Research and Forecasting (WRF) Model. NOAA, 46 pp.
- Caldwell, T.G., M. Cosh, T. Bongiovanni, C. Halley, M.H. Young 2018. Field and Laboratory Evaluation of the CS655 Soil Water Content Sensor. *Vadose Zone J.* doi.org/10.2136/vzj2017.12.0214.
- Caldwell, T. G., Bongiovanni, T., Cosh, M. H., Jackson, T. J., Colliander, A., Abolt, C. J., ... Young, M. H. (2019). The Texas Soil Observation Network: a comprehensive soil moisture dataset for remote sensing and land surface model validation. *Vadose Zone Journal*, 18(1), 1–20.
- Camps, A., Park, H., Pablos, M., Foti, G., Gommenginger, C. P., Liu, P. W., & Judge, J. (2016). Sensitivity of GNSS-R spaceborne observations to soil moisture and vegetation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(10), 4730-4742.
- Chaney, N.W., B. Minasny, J.D. Herman, T.W. Nauman, C. Brungard, C.L.S. Morgan, et al. (2019). POLARIS soil properties: 30-m probabilistic maps of soil properties over the contiguous United States. *Water Resources Research*, 55, 2916–2938. Doi: 10.1029/2018WR022797.
- Chew, C. C., & Small, E. E. (2018). Soil moisture sensing using spaceborne GNSS reflections: Comparison of CYGNSS reflectivity to SMAP soil moisture. *Geophysical Research Letters*, 45, 4049–4057
- Cook, E. R., Anchukaitis, K. J., Buckley, B. M., D’Arrigo, R. D., Jacoby, G. C., & Wright, W. E. (2010). Asian monsoon failure and megadrought during the last millennium. *Science*, 328(5977), 486-489.
- Dashtian, H., Young, M. H. (2023). Soil Water Content and Meteorological Data to Support Tracking Aerosol Convection Interactions Experiment (TRACER) over Harris County, Texas, <https://doi.org/10.18738/T8/FN3RWZ>, Texas Data Repository, V1.
- Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, 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.
- Eeswaran, R., Nejadhashemi, A. P., Alves, F. C., & Saravi, B. (2021). Evaluating the applicability of soil moisture-based metrics for gauging the resiliency of rainfed agricultural systems in the midwestern United States. *Soil and Tillage Research*, 205, 104818.
- Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., ... & Van Zyl, J. (2010). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, 98(5), 704-716.

- [EPA] U.S. environmental Protection Agency. 2008. Quality Assurance Handbook for Air Pollution Measurement Systems, Volume IV: Meteorological Measurements Version 2.0 (Final). EPA-454/B-08-002, March 2008. Research Triangle Park, North Carolina.
- Fang, K., Shen, C., Kifer, D., & Yang, X. (2017). Prolongation of SMAP to spatiotemporally seamless coverage of continental us using a deep learning neural network. *Geophysical Research Letters*, 44(21), 11–030.
- Fang, B., Lakshmi, V., Cosh, M., Liu, P.-W., Bindlish, R., & Jackson, T. J (2022). A global 1-km downscaled SMAP soil moisture product based on thermal inertia theory. *Vadose Zone Journal*, e20182. <https://doi.org/10.1002/vzj2.20182>
- Fang, B., Lakshmi, V., Cosh, M.H., C. Hain. (2021) Very High Spatial Resolution Downscaled SMAP Radiometer Soil Moisture in the CONUS Using VIIRS/MODIS Data. *IEEE J. of Selected Topics in Applied Earth Observations and Remote Sensing*. Vol. 14, 4946-4965.
- Grillakis, M. G., Koutroulis, A. G., Komma, J., Tsanis, I. K., Wagner, W., & Blöschl, G. (2016). Initial soil moisture effects on flash flood generation—A comparison between basins of contrasting hydro-climatic conditions. *Journal of Hydrology*, 541, 206-217.
- Gleason, S., Hodgart, S., Sun, Y., Gommenginger, C., Mackin, S., Adjrad, M., & Unwin, M. (2005). Detection and processing of bistatically reflected GPS signals from low Earth orbit for the purpose of ocean remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 43(6), 1229–1241
- Helama, S., Meriläinen, J., & Tuomenvirta, H. (2009). Multicentennial megadrought in northern Europe coincided with a global El Niño–Southern Oscillation drought pattern during the Medieval Climate Anomaly. *Geology*, 37(2), 175-178.
- Hengl T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić A., et al. 2017) SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* 12(2): e0169748. doi:10.1371/journal.pone.0169748.
- Hong, S. Y., & Kalnay, E. (2000). Role of sea surface temperature and soil-moisture feedback in the 1998 Oklahoma–Texas drought. *Nature*, 408(6814), 842-844.
- Joung, Y. S., & Buie, C. R. (2015). Aerosol generation by raindrop impact on soil. *Nature communications*, 6(1), 1-9.
- Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., & Wood, E. F. (2017). Four decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons. *Advances in Water Resources*, 109, 236–252.
- Kim, S., Zhang, R., Pham, H., & Sharma, A. (2019). A review of satellite-derived soil moisture and its usage for flood estimation. *Remote Sensing in Earth Systems Sciences*, 2(4), 225-246.
- Kim, H., & Lakshmi, V. (2018). Use of Cyclone Global Navigation Satellite System (CYGNSS) observations for estimation of soil moisture. *Geophysical Research Letters*, 45(16), 8272-8282.
- Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Delwart, S., Cabot, F., Boutin, J., ... & Mecklenburg, S. (2010). The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE*, 98(5), 666-687.
- Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J. A. M. J., Font, J., & Berger, M. (2001). Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE transactions on Geoscience and remote sensing*, 39(8), 1729-1735.
- Kerr, Y. H., Wigneron, J. P., Al Bitar, A., Mialon, A., & Srivastava, P. K. (2016). Soil moisture from space: Techniques and limitations. In *Satellite Soil Moisture Retrieval* (pp. 3-27). Elsevier.

- Liu, J., Rahmani, F., Lawson, K., & Shen, C. (2022). A multiscale deep learning model for soil moisture integrating satellite and in situ data. *Geophysical Research Letters*, 49(7), e2021GL096847.
- Liu, L., Morton, Y. J., & Liu, Y. (2021). Machine learning prediction of storm-time high-latitude ionospheric irregularities from GNSS-derived roti maps. *Geophysical Research Letters*, 48(20), e2021GL095561.
- Lobell, D. B., Hammer, G. L., McLean, G., Messina, C., Roberts, M. J., & Schlenker, W. (2013). The critical role of extreme heat for maize production in the United States. *Nature climate change*, 3(5), 497-501.
- Miralles, D. G., Crow, W. T., & Cosh, M. H. (2010). Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations. *Journal of Hydrometeorology*, 11(6), 1423–1429.
- Montzka, C., M. Cosh, B. Bayat, A. Al Bitar, A. Berg, R. Bindlish, H. R. Bogena, J. D. Bolten, F. Cabot, T. Caldwell, S. Chan, A. Colliander, W. Crow, N. Das, G. De Lannoy, W. Dorigo, S. R. Evett, A. Gruber, S. Hahn, T. Jagdhuber, S. Jones, Y. Kerr, S. Kim, C. Koyama, M. Kurum, E. Lopez-Baeza, F. Mattia, K. McColl, S. Mecklenburg, B. Mohanty, P. O'Neill, D. Or, T. Pellarin, G. P. Petropoulos, M. Piles, R. H. Reichle, N. Rodriguez-Fernandez, C. Rüdiger, T. Scanlon, R. C. Schwartz, D. Spengler, P. Srivastava, S. Suman, R. van der Schalie, W. Wagner, U. Wegmüller, J.-P. Wigneron, F. Camacho, and J. Nickeson (2020): Soil Moisture Product Validation Good Practices Protocol Version 1.0. In: C. Montzka, M. Cosh, J. Nickeson, F. Camacho (Eds.): Good Practices for Satellite-Derived Land Product Validation (p. 123), Land Product Validation Subgroup (WGCV/CEOS), doi:10.5067/doc/ceoswgcv/lpv/sm.001
- Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., & Wagner, W. (2009). An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 1999–2013.
- NASA, Goddard Space Flight Center. 2020. CEOS Working Group on Calibration and Validation, Land Product Validation Subgroup. [https://lpvs.gsfc.nasa.gov/SM/SM\\_home.html](https://lpvs.gsfc.nasa.gov/SM/SM_home.html), accessed 15Dec2022.
- Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame, J. M., et al. (2021). What role does hydrological science play in the age of machine learning? *Water Resources Research*, 57(3), e2020WR028091
- Nied, M., Hundecha, Y., & Merz, B. (2013). Flood-initiating catchment conditions: A spatio-temporal analysis of large-scale soil moisture patterns in the Elbe River basin. *Hydrology and Earth System Sciences*, 17, 1401–1414.
- Nielsen-Gammon, J., G. Fipps, T.G. Caldwell, D.B. McRoberts, D. Conlee. 2017 Feasibility Study for Development of Statewide Evapotranspiration Network. Final Report submitted to Final report Texas Water Development Board under contract number 1613581995. 97 pp.
- Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., ... & Phillips, R. P. (2016). The increasing importance of atmospheric demand for ecosystem water and carbon fluxes. *Nature climate change*, 6(11), 1023-1027.
- O'Neill, P. E., Chan, S., Njoku, E. G., Jackson, T., & Bindlish, R. (2016). SMAP enhanced L3 radiometer global daily 9 km EASE-grid soil moisture, version 1. Boulder, CO: NASA National Snow and Ice Data Center Distributed Active Archive Center.

- Reichle, R. H., Ardizzone, J. V., Kim, G. K., Lucchesi, R. A., Smith, E. B., & Weiss, B. H. (2022). Soil Moisture Active Passive (SMAP) mission level 4 surface and root zone soil moisture (L4\_SM) product specification document.
- Robinson, D. A., Campbell, C. S., Hopmans, J. W., Hornbuckle, B. K., Jones, S. B., Knight, R., ... & Wendroth, O. (2008). Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review. *Vadose Zone Journal*, 7(1), 358-389.
- Saini, R., Wang, G., & Pal, J. S. (2016). Role of soil moisture feedback in the development of extreme summer drought and flood in the United States. *Journal of Hydrometeorology*, 17(8), 2191-2207.
- 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(3-4), 125–161.
- Sikder, M. S., Ahmad, S., Hossain, F., Gebregiorgis, A. S., & Lee, H. (2019). Case study: rapid urban inundation forecasting technique based on quantitative precipitation forecast for Houston and Harris county flood control district. *Journal of Hydrologic Engineering*, 24(8), 05019017.
- Sohrabi, M. M., Ryu, J. H., Abatzoglou, J., & Tracy, J. (2015). Development of soil moisture drought index to characterize droughts. *Journal of Hydrologic Engineering*, 20(11).
- Souza, A. G. S. S., Ribeiro Neto, A., & de Souza, L. L. (2021). Soil moisture-based index for agricultural drought assessment: SMADI application in pernambuco state-Brazil. *Remote Sensing of Environment*, 252, 112124.
- Stephens, S. L., Collins, B. M., Fettig, C. J., Finney, M. A., Hoffman, C. M., Knapp, E. E., ... & Wayman, R. B. (2018). Drought, tree mortality, and wildfire in forests adapted to frequent fire. *BioScience*, 68(2), 77-88.
- Wanders, N., Karszenberg, D., De Roo, A., De Jong, S. M., & Bierkens, M. F. P. (2014). The suitability of remotely sensed soil moisture for improving operational flood forecasting. *Hydrology and Earth System Sciences*, 18(6), 2343-2357.
- Wasko, C., & Nathan, R. (2019). Influence of changes in rainfall and soil moisture on trends in flooding. *Journal of Hydrology*, 575, 432-441.
- Western, A. W., & Blöschl, G. (1999). On the spatial scaling of soil moisture. *Journal of Hydrology*, 217(3–4), 203–224.
- Zhou, S., Williams, A. P., Berg, A. M., Cook, B. I., Zhang, Y., Hagemann, S., ... & Gentile, P. (2019). Land–atmosphere feedbacks exacerbate concurrent soil drought and atmospheric aridity. *Proceedings of the National Academy of Sciences*, 116(38), 18848-18853.



## 8. Appendices

### Appendix A: Sample code for QA/AC of raw data

```
clear; close all;

dirname='E:\Soil_Moisture\TRACER_Report\TCEQ_read';
load ([dirname, '\TCEQ_sites_v1_3.mat']) %from read_TCEQ_v1

%%
tic

Ledieu_param=[-0.081 0.093];
Evetts_param=[-0.081 0.085 0.031];
Topp_param=[-1.04E-06 -1.15E-04 1.75E-02 3.83E-02];
%%

%%
t_start = datenum('05/01/21'); % plotting limits
t_end=datenum(datetime('now','TimeZone','-06:00'));

%time range to keep the data on the same time
ts_same=datenum('05/01/21 00:00:00'):datenum('00/00/0000 00:05:00'):datenum(clock);
tS_hr=datenum('05/01/21 00:00:00'):datenum('00/00/0000 01:00:00'):datenum(clock);
ts_hr_dv=datevec(tS_hr);
ts_same=datevec(ts_same); %this fixes rounding problems
ts_same=datenum(ts_same);

numsen=[4 4 4 4]; %number of soil moisture sensors for each station

T=NaN(length(ts_same),max(numsen),length(temp));
VWC=NaN(length(ts_same),max(numsen),length(temp));
Ka=NaN(length(ts_same),max(numsen),length(temp));
EC=NaN(length(ts_same),max(numsen),length(temp));

PPT=NaN(length(ts_same),length(temp));
Batt=NaN(length(ts_same),length(temp));

Tair=NaN(length(ts_same),length(temp));
RH=NaN(length(ts_same),length(temp));
Ws=NaN(length(ts_same),length(temp));
Wd=NaN(length(ts_same),length(temp));
Srad=NaN(length(ts_same),length(temp));
ETo=NaN(length(ts_same),length(temp));
Rso=NaN(length(ts_same),length(temp));
Tdew=NaN(length(ts_same),length(temp));

T_hr=NaN(length(tS_hr),max(numsen),length(temp));
VWC_hr=NaN(length(tS_hr),max(numsen),length(temp));
Ka_hr=NaN(length(tS_hr),max(numsen),length(temp));
EC_hr=NaN(length(tS_hr),max(numsen),length(temp));

PPT_hr=NaN(length(tS_hr),length(temp));

Tair_hr=NaN(length(tS_hr),length(temp));
RH_hr=NaN(length(tS_hr),length(temp));
Ws_hr=NaN(length(tS_hr),length(temp));
Wd_hr=NaN(length(tS_hr),length(temp));
Srad_hr=NaN(length(tS_hr),length(temp));
ETo_hr=NaN(length(tS_hr),length(temp));
```

```

Rso_hr=NaN(length(ts_hr),length(temp));
Tdew_hr=NaN(length(ts_hr),length(temp));

for i =1:4%length(station)
clear ts_temp vwc_temp temp_temp ec_temp ka_temp
ts_temp=ts{i};
ts_temp_dv=datevec(ts_temp);
vwc_temp=vwc{i};
temp_temp=temp{i};
ec_temp=ec{i};
ka_temp=ka{i};
ppt_temp=ppt{i};
batt_temp=batt{i};

ts_met_temp=ts_met{i};
ts_met_temp_dv=datevec(ts_met_temp);
Tair_temp=tair{i};
RH_temp=rh{i};
Ws_temp=ws{i};
Wd_temp=wd{i};
Srad_temp=srad{i};
ETo_temp=eto{i};
Rso_temp=rso{i};
Tdew_temp =
243.12*(log(RH_temp/100)+17.625*Tair_temp./(243.12+Tair_temp))./(17.625-
log(RH_temp/100)-17.625*Tair_temp./(243.12+Tair_temp));

for j=1:length(ts_same)
ind=find(ts_same(j)==ts_temp);
ind2=find(ts_same(j)==ts_met_temp);
if isempty(ind)==0

T(j,1:numsen(i),i)=temp_temp(ind(1),:); %had to use ind(1)... repeated
data
VWC(j,1:numsen(i),i)=vwc_temp(ind(1),:);
Ka(j,1:numsen(i),i)=ka_temp(ind(1),:);
EC(j,1:numsen(i),i)=ec_temp(ind(1),:);
PPT(j,i)=ppt_temp(ind(1),:);
Batt(j,i)=batt_temp(ind(1),:);

Tair(j,i)=Tair_temp(ind2(1),:); %had to use ind(1)... repeated data
RH(j,i)=RH_temp(ind2(1),:);
Ws(j,i)=Ws_temp(ind2(1),:);
Wd(j,i)=Wd_temp(ind2(1),:);
Srad(j,i)=Srad_temp(ind2(1),:);
ETo(j,i)=ETo_temp(ind2(1),:);
Rso(j,i)=Rso_temp(ind2(1),:);
Tdew(j,i)=Tdew_temp(ind2(1),:);
end
end

for j=1:length(ts_hr)
ind=find(ts_hr_dv(j,1)==ts_temp_dv(:,1) & ts_hr_dv(j,2)==ts_temp_dv(:,2) &
ts_hr_dv(j,3)==ts_temp_dv(:,3) & ts_hr_dv(j,4)==ts_temp_dv(:,4));
ind2=find(ts_hr_dv(j,1)==ts_met_temp_dv(:,1) &
ts_hr_dv(j,2)==ts_met_temp_dv(:,2) & ts_hr_dv(j,3)==ts_met_temp_dv(:,3) &
ts_hr_dv(j,4)==ts_met_temp_dv(:,4));
if isempty(ind)==0

T_hr(j,1:numsen(i),i)=mean(temp_temp(ind,:)); %had to use ind(1)...
repeated data
VWC_hr(j,1:numsen(i),i)=mean(vwc_temp(ind,:));

```

```

Ka_hr(j,1:numsen(i),i)=mean(ka_temp(ind,:));
EC_hr(j,1:numsen(i),i)=mean(ec_temp(ind,:));
PPT_hr(j,i)=sum(ppt_temp(ind,:));

v_east = mean(Ws_temp(ind2).*sin(Wd_temp(ind2) * pi/180));
v_north = mean(Ws_temp(ind2).*cos(Wd_temp(ind2) * pi/180));
Tair_hr(j,i)=mean(Tair_temp(ind2,:)); %had to use ind(1)... repeated data
RH_hr(j,i)=mean(RH_temp(ind2));
Ws_hr(j,i)=mean(Ws_temp(ind2));
Wd_hr(j,i)=mod(360+(atan2(v_east,v_north)*180/pi()),360);
Srad_hr(j,i)=mean(Srad_temp(ind2));
ETo_hr(j,i)=sum(ETo_temp(ind2));
Rso_hr(j,i)=mean(Rso_temp(ind2));
Tdew_hr(j,i)=mean(Tdew_temp(ind2));
    end
end
last_collection(i,1)=tS{i}(end);
last_battery(i,1)=batt{i}(end);

end
%%

%%

tS=ts_same;
%%

%%

legendnames={['5 cm'] ['10 cm'] ['20 cm'] ['50 cm']};... %CR300_17_TRACER1
    ['5 cm'] ['10 cm'] ['20 cm'] ['50 cm']};... %CR300_18_TRACER2
    ['5 cm'] ['10 cm'] ['20 cm'] ['50 cm']};... %CR300_19_TRACER3
    ['5 cm'] ['10 cm'] ['20 cm'] ['50 cm']};... %CR300_20_TRACER4
};

%%
ind = VWC > 0.52;
VWC(ind) = -0.053 + 0.0292.*Ka(ind) - 0.00055.*Ka(ind).^2 + 0.0000043*Ka(ind).^3;

%% filters
Flag=repmat(Ka*0,[1,1,1,10]);
Flag(tS<datetime('05-May-2021 14:55:00'),:,1,1)=1;
Flag(tS<datetime('06-May-2021 11:15:00'),:,2,1)=1;
Flag(tS<datetime('10-Sept-2021 13:15:00'),:,3,1)=1;
Flag(tS<datetime('26-May-2022 14:20:00'),:,4,1)=1;

for i=1:4
    for j=1:4% it was 3 for 3 stations
        ind = isnan(T(:,i,j));
        Flag(ind,i,j,1) = 1;

        ind = Ka(:,i,j) <1 | Ka(:,i,j) > 80;
        Flag(ind,i,j,2) = 1;

        ind = EC(:,i,j) <=0;
        Flag(ind,i,j,3) = 1;
    end
end

```

```

ind = VWC(:,i,j) <0 | VWC(:,i,j) > .9;
Flag(ind,i,j,4) = 1;

ind = T(:,i,j) <0 | T(:,i,j) > 60;
Flag(ind,i,j,4) = 1;
% if temperature is the same for 10 hours timesteps then sensor is likely
broken
for k=100:length(tS)
    if all(squeeze(T(k-99:k,i,j))==squeeze(T(k,i,j)))
        Flag(k,i,j,5)=1;
    end
end

end
end
%%
VWC_L= Ledieu_param(1) + (Ledieu_param(2).*sqrt(Ka));
VWC_E= Evett_param(1) + (Evett_param(2).*sqrt(Ka))+(Evett_param(3).*sqrt(EC));

VWC_L_hr=Ledieu_param(1) + (Ledieu_param(2).*sqrt(Ka_hr));
VWC_E_hr=Evett_param(1) + (Evett_param(2).*sqrt(Ka_hr))+(Evett_param(3).*sqrt(EC_hr));

%%
VWC_L_Flag_hr= repmat(VWC_L_hr*0,[1,1,1,10]);
VWC_L_Flag_hr(tS_hr<datetime('05-May-2021 14:55:00'),:,1,1)=1;
VWC_L_Flag_hr(tS_hr<datetime('06-May-2021 11:15:00'),:,2,1)=1;
VWC_L_Flag_hr(tS_hr<datetime('10-Sept-2021 13:15:00'),:,3,1)=1;
VWC_L_Flag_hr(tS_hr<datetime('26-May-2022 14:20:00'),:,4,1)=1;
D=[.05 .10 .20 .50] %measurement depth of the sensor (m)
A=0.05 %Accuracy of the sensor (m3m-3)
p=.5 %soil porosity
for i=25:length(tS_hr)-25
    for k=1:4
        for j=1:4
            x=VWC_L_hr(:,k,j);
            P=PPT_hr(:,j);
            xprime=diff(x);
            xdoubleprime=diff(xprime);
            xT=T_hr(:,k,j);
            %Equation 1 and 2
            if x(i)>x(i-1) %Eq 1
                if x(i)-x(i-24)>2*std(x(i-24:i)) %Eq 2
                    Pmin=D(k)*A*p; %Eq 3
                    if sum(P(i-24:i))<Pmin
                        VWC_L_Flag_hr(i,j,k,2)=1;
                    end
                end
            end
            if xT(i)<0% Soil temperature below 0
                VWC_L_Flag_hr(i,j,k,3)=1;
            end
            %Spectrum-Based Approaches
            %Spike Detection
            % % can't be done in real time
            if x(i)/x(i-1)>1.15 || x(i)/x(i-1)<.85 %Eq 4
                if abs(xdoubleprime(i-1)/xdoubleprime(i+1))>=1.2 ||
abs(xdoubleprime(i-1)/xdoubleprime(i+1))<=.8 %Eq 5
                    if abs(var(x(i-12:i+12))/mean(x(i-12:i+12)))<1 %Eq 6
                        VWC_L_Flag_hr(i,j,k,4)=1;
                    end
                end
            end
        end
    end
end
end

```



```

        if (xprime(i))> 10*(1/n)*sum(x(i-12:i+12)) %Eq8 first derivative and
I think n is 24 but need to check
        eq9term1(i)=abs(xdoubleprime(i)/xdoubleprime(i+1)); % Eq 9 not
sure how to implement???
        if abs(xdoubleprime(i+1)/xdoubleprime(i+2))>10 %Eq 9
            VWC_E_Flag_hr(i,j,k,5)=1;
        end
    end
end
end
end
end
end

%%

%% filters
%%met
%Windspeed and direction variability of 3hrs
for i = 3:length(Ws_hr)
    Ws_hr3(i,:) = abs((max(Ws_hr(i-2:i,:)) - min(Ws_hr(i-2:i,:))));
    Wd_hr3(i,:) = abs((max(Wd_hr(i-2:i,:)) - min(Wd_hr(i-2:i,:))));
end
%Windspeed and direction variability of 6hrs
for i = 6:length(Ws_hr)
    Ws_hr6(i,:) = abs((max(Ws_hr(i-5:i,:)) - min(Ws_hr(i-5:i,:))));
    Wd_hr6(i,:) = abs((max(Wd_hr(i-5:i,:)) - min(Wd_hr(i-5:i,:))));
end
%Windspeed and temperature variability of 12hrs
for i = 12:length(Ws_hr)
    Ws_hr12(i,:) = abs((max(Ws_hr(i-11:i,:)) - min(Ws_hr(i-11:i,:))));
    Tair_hr12(i,:) = abs((max(Tair_hr(i-11:i,:)) - min(Tair_hr(i-11:i,:))));
    Tdew_hr12(i,:) = abs((max(Tdew_hr(i-11:i,:)) - min(Tdew_hr(i-11:i,:))));
    Tdew_Tair_hr12(i,:) = max(abs(Tdew_hr(i-11:i,:) - (Tair_hr(i-11:i,:))));
end

%%

Ws_hr_flag= repmat(zeros(size(Ws_hr)), [1,1,5]);
Wd_hr_flag= repmat(zeros(size(Wd_hr)), [1,1,5]);
Tair_hr_flag= repmat(zeros(size(Tair_hr)), [1,1,5]);
Tdew_hr_flag= repmat(zeros(size(Tdew_hr)), [1,1,5]);
Srad_hr_flag= repmat(zeros(size(Srad_hr)), [1,1,5]);

Ws_hr_flag(tS_hr<datetime('05-May-2021 14:55:00'),1,1)=1;
Wd_hr_flag(tS_hr<datetime('05-May-2021 14:55:00'),1,1)=1;
Tair_hr_flag(tS_hr<datetime('05-May-2021 14:55:00'),1,1)=1;
Tdew_hr_flag(tS_hr<datetime('05-May-2021 14:55:00'),1,1)=1;
Srad_hr_flag(tS_hr<datetime('05-May-2021 14:55:00'),1,1)=1;

Ws_hr_flag(tS_hr<datetime('06-May-2021 11:15:00'),2,1)=1;
Wd_hr_flag(tS_hr<datetime('06-May-2021 11:15:00'),2,1)=1;
Tair_hr_flag(tS_hr<datetime('06-May-2021 11:15:00'),2,1)=1;
Tdew_hr_flag(tS_hr<datetime('06-May-2021 11:15:00'),2,1)=1;
Srad_hr_flag(tS_hr<datetime('06-May-2021 11:15:00'),2,1)=1;

Ws_hr_flag(tS_hr<datetime('10-Sept-2021 13:15:00'),3,1)=1;
Wd_hr_flag(tS_hr<datetime('10-Sept-2021 13:15:00'),3,1)=1;
Tair_hr_flag(tS_hr<datetime('10-Sept-2021 13:15:00'),3,1)=1;
Tdew_hr_flag(tS_hr<datetime('10-Sept-2021 13:15:00'),3,1)=1;

```

```

Srad_hr_flag(tS_hr<datenum('10-Sept-2021 13:15:00'),3,1)=1;

Ws_hr_flag(tS_hr<datenum('26-May-2022 14:20:00'),4,1)=1;
Wd_hr_flag(tS_hr<datenum('26-May-2022 14:20:00'),4,1)=1;
Tair_hr_flag(tS_hr<datenum('26-May-2022 14:20:00'),4,1)=1;
Tdew_hr_flag(tS_hr<datenum('26-May-2022 14:20:00'),4,1)=1;
Srad_hr_flag(tS_hr<datenum('26-May-2022 14:20:00'),4,1)=1;

%
Srad_max_hr=Srad_hr+1; %figure this out
for i=1:4

    Ws_hr_flag(isnan(Srad_hr(:,i)),i,1)=1;
    Wd_hr_flag(isnan(Srad_hr(:,i)),i,1)=1;
    Tair_hr_flag(isnan(Srad_hr(:,i)),i,1)=1;
    Tdew_hr_flag(isnan(Srad_hr(:,i)),i,1)=1;
    Srad_hr_flag(isnan(Srad_hr(:,i)),i,1)=1;

    %0 m/s ? WS ? 25 m/s,
    Ws_hr_flag(Ws_hr(:,i)<0,i,2)=1;
    Ws_hr_flag(Ws_hr(:,i)>25,i,3)=1;
    %Ws_hr_flag(Ws_hr3(:,i)<.1,i,4)=1; %WS varies ? 0.1 m/s for 3 consecutive hours
    Ws_hr_flag(Ws_hr6(:,i)<.1,i,4)=1; %WS varies ? 0.1 m/s for 6 consecutive hours
    Ws_hr_flag(Ws_hr12(:,i)<.5,i,5)=1; %WS varies ? 0.5 m/s for 12 consecutive hours

    %0°? WD ? 360°,
    Wd_hr_flag(Wd_hr(:,i)<0,i,2)=1;
    Wd_hr_flag(Wd_hr(:,i)>360,i,3)=1;
    Wd_hr_flag(Wd_hr3(:,i)<1,i,4)=1; %WD varies ? 1°/3 consecutive hours

    %Local record low? Temp? local record high
    Tair_hr_flag(Tair_hr(:,i)<-14,i,2)=1; %record low
    Tair_hr_flag(Tair_hr(:,i)>43,i,3)=1; %record high
    Tair_hr_flag(abs(diff(Tair_hr(:,i)))>=5,i,4)=1; %Temp ? 5°C from previous hourly
record
    Tair_hr_flag(Tair_hr12(:,i)<0.5,i,5)=1; %Temp varies ? 0.5°C over 12 consecutive
hours

    Tdew_hr_flag(Tdew_hr(:,i)>Tair_hr(:,i),i,2)=1; %Dew Pont Temp ? Ambient temp for
time period
    Tdew_hr_flag(abs(diff(Tdew_hr(:,i)))>=5,i,3)=1; %Temp ? 5°C from previous hourly
record
    Tdew_hr_flag(Tdew_hr12(:,i)<0.5,i,4)=1; %Temp ? 0.5°C from previous hourly record
    Tdew_hr_flag(Tdew_Tair_hr12(:,i)<.01,i,5)=1; %Dew Pont Temp ? Ambient Temp for 12
consecutive hrs. (.01 tolerance)

    %Temp ? 0.5°C from previous hourly record %Dew Pont Temp ? 0.5°C over 12
consecutive hours
    %Dew Pont Temp ? Ambient Temp for 12 consecutive hrs.

    Srad_hr_flag(Srad_hr(:,i)>Srad_max_hr(:,i),i,2)=1; %Temp ? 0.5°C from previous
hourly record
end
%%

%%
cd(dirname);
save TCEQ_final_v1_3 tS T EC VWC PPT Ka filename legendnames station VWC_E VWC_L
last_collection last_battery Flag ...
    Ws Wd Tdew Tair Srad ETo ...

```

```
tS_hr T_hr EC_hr VWC_hr PPT_hr Ka_hr Ws_hr Wd_hr Tdew_hr Tair_hr Srad_hr ETo_hr
...
Ws_hr_flag Wd_hr_flag Tdew_hr_flag Tair_hr_flag Srad_hr_flag ...
VWC_L_hr VWC_E_hr VWC_L_Flag_hr VWC_E_Flag_hr
```



## Appendix B: Write clean data and flags

```

VWC_L_Flag_hr(isnan(VWC_L_Flag_hr))=1;
VWC_E_Flag_hr(isnan(VWC_E_Flag_hr))=1;

sflag = char([
    %num2str(VWC_L_Flag_hr(:,1,1,2)) num2str(VWC_L_Flag_hr(:,1,1,3))
num2str(VWC_L_Flag_hr(:,1,1,4)) num2str(VWC_L_Flag_hr(:,1,1,5)) ...
    %num2str(VWC_L_Flag_hr(:,1,2,2)) num2str(VWC_L_Flag_hr(:,1,2,3))
num2str(VWC_L_Flag_hr(:,1,2,4)) num2str(VWC_L_Flag_hr(:,1,2,5)) ...
    %num2str(VWC_L_Flag_hr(:,1,3,2)) num2str(VWC_L_Flag_hr(:,1,3,3))
num2str(VWC_L_Flag_hr(:,1,3,4)) num2str(VWC_L_Flag_hr(:,1,3,5)) ...
    %num2str(VWC_L_Flag_hr(:,1,4,2)) num2str(VWC_L_Flag_hr(:,1,4,3))
num2str(VWC_L_Flag_hr(:,1,4,4)) num2str(VWC_L_Flag_hr(:,1,4,5)) ...
    num2str(VWC_E_Flag_hr(:,1,1,2)) num2str(VWC_E_Flag_hr(:,1,1,3))
num2str(VWC_E_Flag_hr(:,1,1,4)) num2str(VWC_E_Flag_hr(:,1,1,5)) ...
    num2str(VWC_E_Flag_hr(:,1,2,2)) num2str(VWC_E_Flag_hr(:,1,2,3))
num2str(VWC_E_Flag_hr(:,1,2,4)) num2str(VWC_E_Flag_hr(:,1,2,5)) ...
    num2str(VWC_E_Flag_hr(:,1,3,2)) num2str(VWC_E_Flag_hr(:,1,3,3))
num2str(VWC_E_Flag_hr(:,1,3,4)) num2str(VWC_E_Flag_hr(:,1,3,5)) ...
    num2str(VWC_E_Flag_hr(:,1,4,2)) num2str(VWC_E_Flag_hr(:,1,4,3))
num2str(VWC_E_Flag_hr(:,1,4,4)) num2str(VWC_E_Flag_hr(:,1,4,5)) ...
    %num2str(Wd_hr_flag(:,1,1,1)) num2str(Wd_hr_flag(:,2,1,1))
num2str(Wd_hr_flag(:,3,1,1)) num2str(Wd_hr_flag(:,4,1,1))...
    %num2str(Ws_hr_flag(:,1,1,1)) num2str(Ws_hr_flag(:,2,1,1))
num2str(Ws_hr_flag(:,3,1,1)) num2str(Ws_hr_flag(:,4,1,1))...
    %num2str(Tair_hr_flag(:,1,1,1)) num2str(Tair_hr_flag(:,2,1,1))
num2str(Tair_hr_flag(:,3,1,1)) num2str(Tair_hr_flag(:,4,1,1))
    %num2str(flag_c(:,2)) num2str(flag_GC(:,2)) num2str(flag_out2(:,2))
num2str(flag_out(:,2)) ...
    %num2str(flag_c(:,3)) num2str(flag_GC(:,3)) num2str(flag_out2(:,3))
num2str(flag_out(:,3)) ...
    1]);
%sflag(isnan(sflag))=1;
ssflag = bin2dec(sflag);

out_file = sprintf('%s.dat' , 'CR300_17_Soil_flagged');
fid = fopen(sprintf('%s', out_file), 'wt');
fprintf(fid, '%s.dat, %s factory calibration \n', 'CR300_17_Soil', 'version 1');
fprintf(fid, 'Date, Ppt, VWC_5, VWC_10, VWC_20, VWC_50, T_5,
T_10, T_20, T_50, Flag\n');
for j = 1:length(tS_hr)
    fprintf(fid, '%s, %6.2f, %6.3f, %6.3f, %6.3f, %6.3f, %6.2f, %6.2f, %6.2f,
%6.2f, %8i\n', datestr(tS_hr(j), 'mm-dd-yyyy HH:MM:SS'), PPT_hr(j,1),...
        VWC_hr(j,1,1,1), VWC_hr(j,2,1,1),
VWC_hr(j,3,1,1),VWC_hr(j,4,1,1),T_hr(j,1,1,1), T_hr(j,2,1,1), ...
        T_hr(j,3,1,1),T_hr(j,4,1,1), ssflag(j));
end
fclose(fid);

sflagw=char([
    num2str(Ws_hr_flag(:,1,1)) num2str(Ws_hr_flag(:,1,2))
num2str(Ws_hr_flag(:,1,3)) num2str(Ws_hr_flag(:,1,5))...
    num2str(Wd_hr_flag(:,1,1)) num2str(Wd_hr_flag(:,1,2))
num2str(Wd_hr_flag(:,1,3)) num2str(Wd_hr_flag(:,1,4))...
    num2str(Tair_hr_flag(:,1,1)) num2str(Tair_hr_flag(:,1,2))
num2str(Tair_hr_flag(:,1,3)) num2str(Tair_hr_flag(:,1,5))...
    num2str(Tdew_hr_flag(:,1,1)) num2str(Tdew_hr_flag(:,1,2))
num2str(Tdew_hr_flag(:,1,3)) num2str(Tdew_hr_flag(:,1,5))...
    num2str(Srad_hr_flag(:,1,1)) num2str(Srad_hr_flag(:,1,2))
]);

```

```

ssflagw = bin2dec(sflagw);

out_file = sprintf('%s.dat' , 'CR300_17_Meteoro_flagged');
fid = fopen(sprintf('%s', out_file), 'wt');
fprintf(fid, '%s.dat, %s factory calibration \n', 'CR300_17_Meteorological',
'version 1');
fprintf(fid, 'Date, Ppt, Wind_speed, Wind_direction, T_air, T_dew,
Solar_radiation Flag\n');
for j = 1:length(tS_hr)
    fprintf(fid, '%s, %6.2f, %6.3f, %6.3f, %6.2f, %6.2f, %6.2f, %8i\n',
datestr(tS_hr(j), 'mm-dd-yyyy HH:MM:SS'), PPT_hr(j,1),...
Ws_hr(j,1), Wd_hr(j,1), Tair_hr(j,1),Tdew_hr(j,1),Srad_hr(j,1),
ssflagw(j));
end
fclose(fid);

```

### Appendix C: Example of portion of QA/QC meteorological data

```

CR300_17_Meteorological.dat, version 1 factory calibration
Date, Ppt, Wind_speed, Wind_direction, T_air, T_dew, Solar_radiation Flag
05-05-2021 14:00:00, 0.00, 2.528, 27.356, 25.84, 16.36, 790.70, 157282
05-05-2021 15:00:00, 0.00, 3.306, 40.309, 26.13, 16.73, 689.26, 64
05-05-2021 16:00:00, 0.00, 2.893, 43.451, 26.29, 16.62, 483.69, 64
05-05-2021 17:00:00, 0.00, 2.849, 49.060, 25.86, 16.06, 260.17, 64
05-05-2021 18:00:00, 0.00, 1.787, 42.453, 25.01, 15.47, 56.33, 0
05-05-2021 19:00:00, 0.00, 1.352, 80.171, 22.48, 16.54, 0.39, 0
05-05-2021 20:00:00, 0.00, 0.002, 58.230, 19.97, 17.88, 0.00, 0
05-05-2021 21:00:00, 0.00, 0.331, 49.100, 18.89, 17.58, 0.00, 0
05-05-2021 22:00:00, 0.00, 2.091, 52.481, 20.83, 16.14, 0.00, 0
05-05-2021 23:00:00, 0.00, 2.039, 46.822, 21.00, 16.05, 0.00, 0
05-06-2021 00:00:00, 0.00, 1.678, 42.366, 20.30, 16.54, 0.00, 0
05-06-2021 01:00:00, 0.00, 2.086, 41.289, 20.07, 17.59, 0.00, 0
05-06-2021 02:00:00, 0.00, 1.286, 49.524, 19.18, 17.37, 0.00, 0
05-06-2021 03:00:00, 0.00, 1.447, 49.436, 18.58, 16.96, 0.00, 0
05-06-2021 04:00:00, 0.00, 1.588, 52.071, 18.41, 16.75, 0.00, 0
05-06-2021 05:00:00, 0.00, 2.340, 45.718, 18.55, 16.50, 5.64, 0
05-06-2021 06:00:00, 0.00, 3.277, 45.741, 19.40, 16.27, 126.32, 0
05-06-2021 07:00:00, 0.00, 4.283, 44.377, 20.68, 15.41, 348.67, 0
05-06-2021 08:00:00, 0.00, 4.566, 40.385, 21.68, 14.76, 575.40, 0
05-06-2021 09:00:00, 0.00, 4.215, 39.531, 22.56, 14.86, 773.53, 0
05-06-2021 10:00:00, 0.00, 3.757, 47.899, 23.66, 14.53, 919.92, 0
05-06-2021 11:00:00, 0.00, 3.205, 41.405, 24.78, 13.92, 1006.83, 0
05-06-2021 12:00:00, 0.00, 3.370, 31.363, 26.04, 12.37, 1029.17, 0
05-06-2021 13:00:00, 0.00, 3.439, 33.649, 26.93, 12.31, 978.67, 0
05-06-2021 14:00:00, 0.00, 3.187, 62.547, 27.05, 13.56, 861.92, 0
05-06-2021 15:00:00, 0.00, 2.906, 49.490, 27.00, 14.08, 694.60, 0
05-06-2021 16:00:00, 0.00, 2.374, 49.350, 26.82, 14.55, 482.72, 0
05-06-2021 17:00:00, 0.00, 2.313, 55.320, 26.31, 14.86, 259.68, 0
05-06-2021 18:00:00, 0.00, 1.450, 67.254, 24.43, 15.18, 56.77, 0
05-06-2021 19:00:00, 0.00, 0.300, 160.678, 21.29, 15.62, 0.28, 0
05-06-2021 20:00:00, 0.00, 0.009, 250.200, 18.62, 15.85, 0.00, 0
05-06-2021 21:00:00, 0.00, 0.000, 0.000, 16.90, 15.81, 0.00, 0
05-06-2021 22:00:00, 0.00, 0.000, 0.000, 15.89, 15.52, 0.00, 0
05-06-2021 23:00:00, 0.00, 0.000, 0.000, 15.27, 15.15, 0.00, 1024

```

## Appendix D: Example of portion of raw meteorological data

```
"TOA5","CR300_17_TRACER1","CR300","26683","CR300.Std.10.03","CPU:CR300_17_TRACER1.CR300",  
"30484","SubHourly_met"  
"TIMESTAMP","RECORD","Rain_mm_Tot","AirTC_Avg","RH_Avg","WS_ms_S_WVT","WindDir_D1_WVT",  
"SlrW_Avg","ETos","Rso"  
"TS","RN","mm","Deg C","%", "meters/second", "Deg", "W/m^2", "mm", "MJ/m^2"  
"","Tot","Avg","Avg","WVc","WVc","Avg","ETXs","Rso"  
"2021-05-06 14:00:00",278,0,27.13,37.32,3.493,35.98,927,0.239,2.959  
"2021-05-06 14:05:00",279,0,27.14,44.77,3.03,64.13,913,0.198,2.885  
"2021-05-06 14:10:00",280,0,27.02,44.07,2.54,46.8,901,0.179,2.844  
"2021-05-06 14:15:00",281,0,27.03,44.27,3.653,71.94,891,0.221,2.796  
"2021-05-06 14:20:00",282,0,26.98,45.12,2.792,52.89,883,0.185,2.746  
"2021-05-06 14:25:00",283,0,27.24,39.19,2.563,38.25,874,0.19,2.716  
"2021-05-06 14:30:00",284,0,27.1,44.8,3.105,82.6,859,0.197,2.647  
"2021-05-06 14:35:00",285,0,26.96,45.94,3.078,73.76,844,0.192,2.593  
"2021-05-06 14:40:00",286,0,26.93,44.73,3.688,83.5,830,0.217,2.545  
"2021-05-06 14:45:00",287,0,26.98,43.93,3.313,57.68,819,0.205,2.494  
"2021-05-06 14:50:00",288,0,27.16,44.44,3.548,61.42,809,0.213,2.436  
"2021-05-06 14:55:00",289,0,26.89,43.59,3.435,70.3,793,0.209,2.385  
"2021-05-06 15:00:00",290,0,26.85,44.68,3.365,83.3,780.6,0.202,2.326  
"2021-05-06 15:05:00",291,0,27.02,44.99,2.683,55.88,767.1,0.175,2.266  
"2021-05-06 15:10:00",292,0,26.96,44.77,3.402,51.73,752.4,0.203,2.209  
"2021-05-06 15:15:00",293,0,26.87,44.91,3.103,57.87,738.7,0.19,2.15  
"2021-05-06 15:20:00",294,0,27.08,44.45,3.04,47.57,721.1,0.188,2.089  
"2021-05-06 15:25:00",295,0,27.05,44.67,2.33,32.75,704,0.158,2.028  
"2021-05-06 15:30:00",296,0,27.03,46.62,2.923,52.41,685.9,0.176,1.962  
"2021-05-06 15:35:00",297,0,27.14,46.25,2.808,34.15,671.1,0.172,1.9  
"2021-05-06 15:40:00",298,0,26.92,44.56,3.475,55.65,654.9,0.201,1.843  
"2021-05-06 15:45:00",299,0,27.01,44.31,2.983,35.51,638.6,0.181,1.779  
"2021-05-06 15:50:00",300,0,27.06,44.87,2.615,34.55,619.3,0.164,1.714  
"2021-05-06 15:55:00",301,0,26.96,46.17,2.14,39.61,601.5,0.14,1.647  
"2021-05-06 16:00:00",302,0,27.06,46.34,2.642,48.19,583.7,0.16,1.581  
"2021-05-06 16:05:00",303,0,26.91,45.82,2.735,19.16,568.7,0.164,1.518
```

### Appendix E: Example of portion of QA/QC soil data

CR300\_17\_Soil.dat, version 1 factory calibration

Date,	Ppt,	VWC_5,	VWC_10,	VWC_20,	VWC_50,	T_5,	T_10,	T_20,	T_50,	Flag
05-26-2022 11:00:00,	0.00,	0.416,	0.455,	0.435,	0.601,	25.02,	23.68,	23.81,	23.84,	15
05-26-2022 12:00:00,	0.00,	0.417,	0.455,	0.433,	0.600,	25.94,	24.24,	23.83,	23.84,	15
05-26-2022 13:00:00,	0.00,	0.417,	0.454,	0.430,	0.600,	26.64,	24.84,	23.89,	23.83,	15
05-26-2022 14:00:00,	0.00,	0.416,	0.453,	0.428,	0.599,	27.21,	25.39,	24.00,	23.81,	15
05-26-2022 15:00:00,	0.00,	0.416,	0.453,	0.426,	0.599,	27.52,	25.79,	24.12,	23.79,	0
05-26-2022 16:00:00,	0.00,	0.414,	0.452,	0.425,	0.599,	27.51,	25.99,	24.27,	23.78,	0
05-26-2022 17:00:00,	0.00,	0.412,	0.451,	0.425,	0.598,	27.25,	26.07,	24.41,	23.76,	0
05-26-2022 18:00:00,	0.00,	0.411,	0.451,	0.424,	0.598,	26.84,	26.04,	24.56,	23.76,	0
05-26-2022 19:00:00,	0.00,	0.409,	0.450,	0.424,	0.598,	26.36,	25.92,	24.67,	23.75,	0
05-26-2022 20:00:00,	0.00,	0.408,	0.450,	0.424,	0.598,	25.85,	25.72,	24.76,	23.73,	0
05-26-2022 21:00:00,	0.00,	0.407,	0.449,	0.424,	0.598,	25.34,	25.50,	24.81,	23.73,	0
05-26-2022 22:00:00,	0.00,	0.406,	0.448,	0.424,	0.598,	24.92,	25.23,	24.83,	23.73,	0
05-26-2022 23:00:00,	0.00,	0.405,	0.447,	0.423,	0.599,	24.67,	25.00,	24.82,	23.72,	0
05-27-2022 00:00:00,	0.00,	0.405,	0.446,	0.424,	0.598,	24.47,	24.81,	24.80,	23.74,	0
05-27-2022 01:00:00,	0.00,	0.404,	0.446,	0.423,	0.598,	24.34,	24.66,	24.76,	23.72,	0
05-27-2022 02:00:00,	0.00,	0.403,	0.445,	0.423,	0.597,	24.18,	24.53,	24.71,	23.73,	0
05-27-2022 03:00:00,	0.00,	0.403,	0.445,	0.423,	0.598,	24.05,	24.41,	24.66,	23.74,	0
05-27-2022 04:00:00,	0.00,	0.403,	0.444,	0.423,	0.597,	23.92,	24.30,	24.61,	23.75,	0
05-27-2022 05:00:00,	0.00,	0.403,	0.443,	0.423,	0.597,	23.68,	24.20,	24.56,	23.75,	0
05-27-2022 06:00:00,	0.00,	0.402,	0.443,	0.422,	0.598,	23.62,	24.08,	24.51,	23.75,	0
05-27-2022 07:00:00,	0.00,	0.402,	0.443,	0.422,	0.598,	23.76,	24.01,	24.46,	23.75,	0
05-27-2022 08:00:00,	0.00,	0.402,	0.443,	0.421,	0.598,	24.11,	24.07,	24.41,	23.76,	0
05-27-2022 09:00:00,	0.00,	0.403,	0.442,	0.421,	0.598,	24.70,	24.27,	24.37,	23.75,	0

## Appendix F: Example of portion of raw soil data

```
"TOA5","CR300_17_TRACER1","CR300","26683","CR300.Std.10.03","CPU:CR300_17_TRACER1.CR300","30484","SubHourly_soil"
"TIMESTAMP","RECORD","Rain_mm_Tot","VWC_1_Avg","VWC_2_Avg","VWC_3_Avg","VWC_4_Avg","T_1_Avg","T_2_Avg","T_3_Avg","T_4_Avg","EC_1_Avg",
,"EC_2_Avg","EC_3_Avg","EC_4_Avg","P_1_Avg","P_2_Avg","P_3_Avg","P_4_Avg","BattV_Min"
"TS","RN","mm","m^3/m^3","m^3/m^3","m^3/m^3","m^3/m^3","Deg C","Deg C","Deg C","Deg C","dS/m","dS/m","dS/m","dS/m","unitless","unitless","unitless","unitless","Volts"
","","Tot","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Avg","Min"
"2021-05-26 11:00:00",6002,0,0.69,0.63,0.621,0.7,28.44,25.72,24.27,22.67,0.936,0.837,0.736,1.288,63.98,58.06,57.07,64.82,13.52
"2021-05-26 11:05:00",6003,0,0.691,0.63,0.621,0.7,28.57,25.77,24.27,22.66,0.979,0.838,0.736,1.291,64.04,58.06,57.07,64.81,13.47
"2021-05-26 11:10:00",6004,0,0.691,0.63,0.621,0.699,28.67,25.84,24.26,22.66,0.979,0.839,0.735,1.372,64.04,58.06,57.07,64.71,13.47
"2021-05-26 11:15:00",6005,0,0.691,0.63,0.621,0.7,28.76,25.89,24.27,22.67,0.979,0.839,0.735,1.294,64.04,58.06,57.07,64.81,13.49
"2021-05-26 11:20:00",6006,0,0.69,0.63,0.621,0.699,28.83,25.91,24.26,22.67,0.944,0.84,0.735,1.372,63.99,58.06,57.07,64.71,13.5
"2021-05-26 11:25:00",6007,0,0.691,0.63,0.621,0.7,28.96,26.01,24.29,22.67,0.985,0.84,0.735,1.292,64.05,58.06,57.07,64.81,13.5
"2021-05-26 11:30:00",6008,0,0.691,0.63,0.621,0.7,29.03,26.01,24.28,22.67,0.986,0.843,0.735,1.289,64.05,58.07,57.07,64.81,13.49
"2021-05-26 11:35:00",6009,0,0.691,0.631,0.621,0.7,29.17,26.09,24.27,22.66,0.988,0.842,0.735,1.294,64.05,58.16,57.07,64.81,13.5
"2021-05-26 11:40:00",6010,0,0.691,0.631,0.621,0.699,29.31,26.13,24.3,22.67,0.992,0.843,0.736,1.371,64.06,58.16,57.07,64.71,13.49
"2021-05-26 11:45:00",6011,0,0.692,0.631,0.621,0.7,29.42,26.19,24.28,22.65,0.991,0.845,0.735,1.289,64.15,58.17,57.07,64.82,13.5
"2021-05-26 11:50:00",6012,0,0.691,0.631,0.621,0.7,29.55,26.23,24.28,22.69,0.953,0.845,0.734,1.29,64.11,58.17,57.06,64.81,13.5
"2021-05-26 11:55:00",6013,0,0.692,0.631,0.621,0.7,29.67,26.3,24.27,22.68,0.997,0.845,0.735,1.29,64.16,58.17,57.07,64.81,13.5
"2021-05-26 12:00:00",6014,0,0.692,0.631,0.621,0.7,29.81,26.34,24.28,22.65,0.999,0.848,0.736,1.293,64.16,58.17,57.07,64.81,13.5
"2021-05-26 12:05:00",6015,0,0.692,0.631,0.621,0.699,29.89,26.4,24.28,22.67,0.997,0.848,0.735,1.372,64.16,58.17,57.07,64.71,13.44
"2021-05-26 12:10:00",6016,0,0.692,0.632,0.621,0.7,30.01,26.46,24.27,22.68,0.962,0.837,0.735,1.291,64.12,58.25,57.07,64.81,13.45
"2021-05-26 12:15:00",6017,0,0.692,0.631,0.621,0.7,30.08,26.51,24.29,22.68,1.005,0.85,0.735,1.29,64.17,58.17,57.07,64.81,13.48
"2021-05-26 12:20:00",6018,0,0.693,0.632,0.621,0.7,30.23,26.56,24.3,22.69,1.005,0.849,0.735,1.291,64.27,58.27,57.07,64.81,13.48
```



```

Units WindDir=degrees
Units AirTC=Deg C
Units RH=%
Units VWC_1=m^3/m^3: Units EC_1=dS/m: Units T_1=Deg C: Units P_1=unitless:' Units
PA_1=nSec: Units VR_1=unitless
Units VWC_2=m^3/m^3: Units EC_2=dS/m: Units T_2=Deg C: Units P_2=unitless:' Units
PA_2=nSec: Units VR_2=unitless
Units VWC_3=m^3/m^3: Units EC_3=dS/m: Units T_3=Deg C: Units P_3=unitless:' Units
PA_3=nSec: Units VR_3=unitless
Units VWC_4=m^3/m^3: Units EC_4=dS/m: Units T_4=Deg C: Units P_4=unitless:' Units
PA_4=nSec: Units VR_4=unitless

```

```

Units ETos = mm
Units Rso = MJ/m^2

```

```

'//////////////////////////////////// OUTPUT SECTION //////////////////////////////////////
DataTable(SubHourly_met,True,-1)
  DataInterval(0,5,Min,1)
  Totalize(1,Rain_mm,FP2,False)
  Average(1,AirTC,FP2,False)
  Average(1,RH,FP2,False)
  WindVector (1,WS_ms,WindDir,FP2,False,0,0,1)
  FieldNames("WS_ms_S WVT,WindDir_D1_WVT")
  Average(1,SlrW,FP2,False)
  ETsz(AirTC,RH,WS_ms,SlrMJ,site_lon,site_lat,site_z,2.5,0,FP2,False)
  FieldNames("ETos,Rso")
EndTable

```

```

DataTable(SubHourly_soil,True,-1)
  DataInterval(0,5,Min,1)
  Totalize(1,Rain_mm,FP2,False)
  Average(1,VWC_1,FP2,False)
  Average(1,VWC_2,FP2,False)
  Average(1,VWC_3,FP2,False)
  Average(1,VWC_4,FP2,False)

  Average(1,T_1,FP2,False)
  Average(1,T_2,FP2,False)
  Average(1,T_3,FP2,False)
  Average(1,T_4,FP2,False)

  Average(1,EC_1,FP2,False)
  Average(1,EC_2,FP2,False)
  Average(1,EC_3,FP2,False)
  Average(1,EC_4,FP2,False)

  Average(1,P_1,FP2,False)
  Average(1,P_2,FP2,False)
  Average(1,P_3,FP2,False)
  Average(1,P_4,FP2,False)

  Minimum (1,BattV,FP2,False,False)
EndTable

```

```

DataTable(Daily,True,-1)
  DataInterval(0,1440,Min,10)
  Sample (1,Site_ID,FP2)
  Minimum(1,BattV,FP2,False,False)
  Average(1,PTemp_C,FP2,False)
  Totalize(1,Rain_mm,FP2,False)
  Maximum(1,AirTC,FP2,False,False)
  Minimum(1,AirTC,FP2,False,False)

```



```

Maximum(1,RH,FP2,False,False)
Minimum(1,RH,FP2,False,False)
Totalize(1,SlrMJ,IEEE4,False)
EndTable

'\////////////////////////////////\ PROGRAM //////////////////////////////////
BeginProg
  'Main Scan
  Site_ID = Site
  SW12 (1 )

Scan(10,Sec,1,0)
  'Default Datalogger Battery Voltage measurement 'BattV'
  Battery(BattV)
  RealTime(RTime)

  'Default Wiring Panel Temperature measurement 'PTemp_C'
  PanelTemp(PTemp_C,60)
  'TE525/TE525WS Rain Gauge measurement 'Rain_mm'
  PulseCount(Rain_mm,1,P_SW,2,0,0.254,0)

      'LP02 Pyranometer measurements 'SlrMJ' and 'SlrW'
      VoltDiff(SlrW,1,mV34,2,True,0,60,Slr_S,0)      'Based on shortcut
recommendations ()
      'VoltDiff(SlrW,1,mV250,1,True,0,_60Hz,Slr_S,0) 'From CR1000 program
      If SlrW<0 Then SlrW=0
      SlrMJ=SlrW*10*0.000001 ' converstion = scan_rate*1e-6

  '03002 Wind Speed & Direction Sensor measurements 'WS_ms' and 'WindDir'
  PulseCount(WS_ms,1,P_LL,1,1,0.75,0.2)
  If WS_ms<0.21 Then WS_ms=0
      BrHalf(WindDir,1,mV2500,1,VX1,1,2500,False,20000,60,352,0) 'Based on
shortcut recommendation

  If WindDir>=360 OR WindDir<0 Then WindDir=0

  'HygroVUE10Temperature&RelativeHumiditySensormeasurements'AirTC''and'RH'
  SDI12Recorder(AirTCRH(),C1,"0","M!",1,0,-1)
  ' CS65X Water Content Reflectometer measurements
  '+++++++ UPDATE SDI12 ADDRESSES BELOW FOR EACH STATION
+++++++
  If IfTime(0,5,Min) Then
    SDI12Recorder(CS65X_1(),C2,VWC_1_add,"M4!",1,0,-1)
    SDI12Recorder(CS65X_2(),C2,VWC_2_add,"M4!",1,0,-1)
    SDI12Recorder(CS65X_3(),C2,VWC_3_add,"M4!",1,0,-1)
    SDI12Recorder(CS65X_4(),C2,VWC_4_add,"M4!",1,0,-1)
  EndIf

  'SW12 Timed Control
  'Turn ON SW12 between 0600 hours and 2100 hours CST
  'for 10 minutes every 60 minutes
  If TimeIsBetween(0,1440,1440,Min) AND TimeIsBetween(0,10,60,Min) Then
    SW12State=True
  'Turn OFF SW12 only if time runs out and RS-232 is not active
  ElseIf (ComPortIsActive(ComRS232)=False) Then
    SW12State=False
  EndIf
  'Always turn OFF SW12 if battery drops below 11.5 volts
  If BattV<11.5 Then SW12State=False
  'Set SW12-SW12V to the state of 'SW12State' variable
  SW12(SW12State)

  'Call Data Tables and Store Data

```

```
    CallTable(SubHourly_soil)
    CallTable(Daily)
    CallTable(SubHourly_met)
NextScan
```

```
EndProg
```