CHAPTER 7

QUALITY CONTROL AND QUALITY ASSURANCE OF SOIL MOISTURE DATA

Mike Cosh, Ed Ayres, Todd Caldwell, Vinit Sehgal, Zamir Libohova, Nandita Gaur

Learning Outcomes

Properly identifying and flagging data errors helps to ensure the quality of a dataset and support viability of data use.

Quality assurance (QA) (Campbell et al., 2013) of soil moisture data refers to the set of processes that are employed to improve confidence and decrease errors in the production of soil moisture data. Quality control (QC) processes are measures taken after data are collected to improve or remove data points so that the final product is of high quality.

QUALITY ASSURANCE

Pre-requisites of a well-functioning soil moisture sensor include proper installation of the sensor at the correct depth and use under conditions that the sensor is designed for. Proper installation methods can be accessed in Caldwell et al. (2022), and recommendations for siting methods, sensor calibration, and other important installation processes have been discussed in Chapters 2-6 of this document. This chapter describes methods to identify data errors, trace their origins, and report or fix them.

CATEGORIES OF DATA ERRORS

There are two primary types of errors (described below) that can be flagged. Type I errors can be easily flagged using automated tests during QC, but Type II errors require manual verification and are often detected during QA activities. Type II errors should not be considered errors until proper manual inspection of each erroneous data point is conducted.

The International Vocabulary of Basic and General Terms in Meteorology (VIM) also provides a classification system for identifying Uncertainty in Measurement (GUM): Type A (systematic) and Type B (random) data errors (VIM International Vocabulary of Metrology – Basic and General Concepts and Associated Terms, 2006). In this Soil Moisture Data Quality Guidance document, a different approach than the GUM approach is used to categorize data errors. In this document, error type is based on the skill set necessary to identify and address potential data errors: Type I (errors that can be identified via a simple automated algorithm) and Type II (errors that require additional skillsets to identify and address).

Visually Observable, Easy-to-Identify Errors (Type I)

Visually observable Type I errors are relatively straightforward to detect. QA tests that detect sensor failure, recorder failure, or disruptive environmental events can be discovered and corrected by automated methods since they leave distinctly identifiable signatures in the data (Campbell et al., 2013; Dorigo et al., 2013; Dorigo et al., 2021). These tests can be easily automated and should be used for flagging. Examples of Type I errors in data are shown in Figure 7, taken from Caldwell et al. (2022).

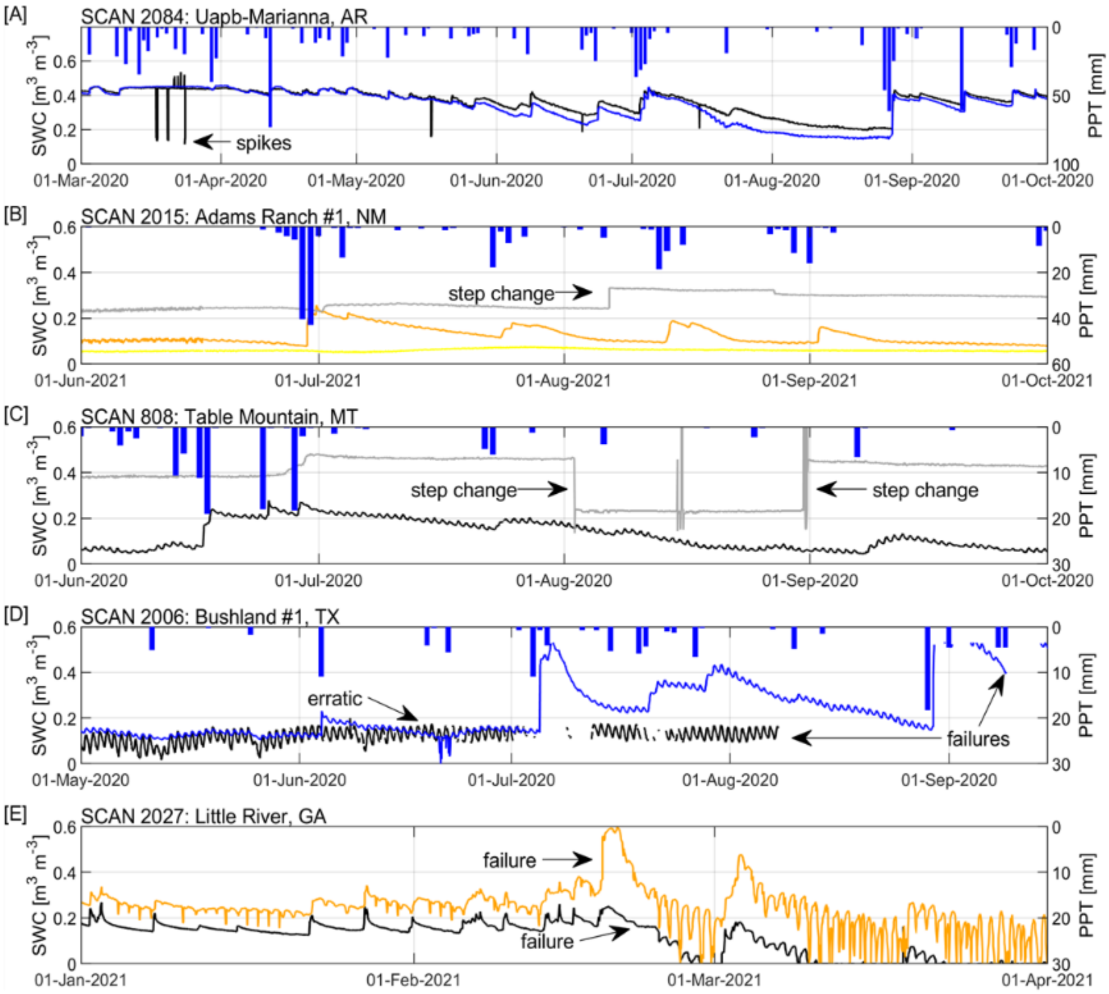

Figure 7. Examples of Type I errors in data. Data from (A) SCAN 2084, Uapb-Marianna, Arkansas, that shows periodic dips at 5 cm. (B) SCAN 2015, Adams Ranch #1, New Mexico, shows a positive step change at 50 cm depth without changes in the upper depths. (C) SCAN 808, Table Mountain, Montana, with a downward step change, spikes, and even recovery at 50 cm depth that does not correspond to rainfall events. (D) SCAN 2006, Bushland #1, Texas, showing no response to precipitation events at the 5 or 10 cm sensor, with some recovery of the 10 cm sensor followed by the eminent failure of both. (E) SCAN 2027, Little River, Georgia, with a glitching sensor at 20 cm and failure at both the 5 and 20 cm depths. Sensor depths are denoted as 5 cm (black), 10 cm (blue), 20 cm (orange), 50 cm (dark gray), and 100 cm (yellow). Abbreviations: SWC = soil water content; PPT = precipitation. Figure Credit: Caldwell et al., 2022.

Ideally, automated QA tests should be performed at the native sensor sampling frequency. A list of automated tests (Hubbard et al., 2005) that are used by existing networks and projects to flag soil water content data include:

1. **Range test or high/low range limit test:** The range test (sometimes called the high/low range limit or "upper and lower" threshold test) checks whether each soil moisture value is less than the porosity of the soil and more than 0 (or the lower measurement limit of the sensor). One occurrence of an off value would need to be flagged as an individual point in the dataset, but repeated or periodic observations of such values will require flagging of the sensor itself. Under such conditions, the sensor(s) will need manual inspection and may require replacement or additional tests before data can be reported.

2. **Constant value test:** A failed sensor may often report a constant value. While in some cases (for example, saturated conditions or very dry conditions) constant values may be legitimate, an unexplained constant value of soil moisture over an extended period must be flagged and manually examined.
3. **Spike test:** A spike test pertains to observing spikes (Figure 7c) in the data. Spikes in soil moisture must only correspond with rainfall or snowmelt events, and typically the spikes will be observed first in the near surface soil moisture and subdued replicas of the spikes may be observed after a lag in the subsequent depths. It is possible to observe a spike in the deeper soil before the sensors above in cases of shrink-swell clay soils. A spike in data that is not occurring concurrently with rainfall must be automatically flagged and subject to manual testing of the sensor since it may imply a failed or temporarily failing sensor.
4. **Break test:** Breaks in soil moisture measurements must be flagged as missing data.
5. **Temperature test:** Most soil moisture sensors are built to function within specific temperature ranges and cannot detect frozen soil moisture. If the soil temperature at any depth falls outside of that range or below 0 °C, the data must be flagged and not reported until a manual verification is performed.

Complex and Hard-to-Identify Errors (Type II)

Type II errors are more challenging to detect as the data might seem plausible at first glance, but the underlying dynamics do not align with the expected patterns of moisture redistribution or unsaturated flow processes. Addressing these errors requires a higher level of investigation, and they are better identified using QA procedures described later in this chapter. Any flags assigned to Type II errors should be verified by an expert. Some of the hydrologic principles that can be employed to assess inconsistencies include:

1. **Correlations in temporal relationship of soil moisture with ancillary variables**: It is expected that changes in soil moisture values will correlate to changes in other ancillary variables, such as rain events or some changes in air and soil temperature. Data errors wherein these relationships are not present can be identified by utilizing known rainfall-runoff-storage-soil moisture redistribution relationships that correspond to the specific soil, geomorphic, and cover conditions. Soil moisture values that do not conform to long-established relationships can be temporarily flagged and investigated during QA activities to ensure that the variability in the relationships is a true data error and not driven by changes in the landscape that justify the deviation in correlative relationships. Advanced multiscale signal-processing techniques, like Wavelet Transform and Empirical Mode Decomposition, can be used to assess the temporal variability in the soil moisture observations across multiple time scales and to detect potential anomalies and outliers (Mallat and Hwang, 1992; Geng et al., 2011; Thill et al., 2017).
2. **Moisture redistribution in unsaturated flow process:** This data quality test relies on assessing the consistency of time series data to determine if moisture redistribution patterns align with expected unsaturated flow processes. These include verifying that profile soil moisture sensors respond in order by depth to rainfall, with those nearest to the soil surface responding first. These checks would need to be investigated after immediate flagging, since in cases of preferential flow, some sensors below may respond

earlier than those above them. Such types of evaluation may require developing site-specific tools.

3. **Calibration shift or sensor fouling:** Sensors may change in the accuracy of their readings over time. A slow drift in a sensor is often also very hard to detect. Identifying drift of this type would require careful evaluation of the data and recalibration of the sensor (Wagner et al., 2006). One way of observing this may be to identify the ultimate minimum soil moisture value after a significantly long drydown (> 60 days). It is reasonable to assume that two periods of 60 days of drying would end at the same number if there were a similarity in the temperature or season. Consistency of readings is not as readily observed at maximum soil moistures because the salinity of the soil can change more easily with rainfall/overland flow, and salinity may be a dominant portion of the dielectric constant, if a dielectric based probe is being used.

Correcting Source of Data Errors

Upon error detection, the cause of errors in the data should be identified and corrected. Random Type I errors in data can occur as a result of voltage fluctuations, inappropriate temperature conditions, or other random or natural occurrences beyond the operator's control, but systematically occurring errors must be corrected. Errors of either type can occur as a result of four causes.

1. **Instrument malfunction:** Instrument malfunction will typically lead to Type I errors and can be corrected by repair or replacement of the sensor.
2. **Personnel errors**: Personnel errors may include incorrect metadata for error detection (for example, incorrect upper and lower limits for the range test). Such errors will also be a systematic and can be corrected by training personnel or establishing protocols that do not leave scope for errors.
3. **Transmission errors**: Transmission errors will typically lead to increased latency in data and so are easily detected. Correcting the source of these types of errors would require expert inspection.
4. **Data processing errors**: Data processing errors can be Type I or II. Once identified, these must be corrected by updating algorithms or incorporating additional manual verification.

Considerations for Automating Soil Moisture Data Quality Checks

Automated QA tests are often the most effective way to identify suspect data for large networks and/or long-term projects where the quantity of data is typically too large relative to staffing resources to allow effective manual data reviewing and flagging[10]. One advantage of automated tests is that they can usually be run shortly after data collection, facilitating near-real-time provisionally QA data availability. Other advantages of automated tests include their consistency and reliability. QA checks performed by humans are more prone to biases, person-to-person variability, competing time commitments, variable focus, and other disruptions. However, there are also disadvantages to automated QA tests, including the large upfront development effort and

[10] Resources for automating soil moisture data flags, include discussion of this topic in a paper from the International Soil Moisture Network (ISMN) (Dorigo et al., 2021, Section 3.1) and a GitHub repository managed by the Group on Earth Observations (GEO): https://github.com/TUW-GEO/flagit.

ongoing code and database maintenance, as well as the effort involved in optimizing test thresholds for a range of sites and measurement depths. Moreover, automated tests are typically designed within the context of recent historical data under “normal” operating conditions, which may result in inappropriate flagging when real but unusual or unprecedented events occur (e.g., extreme rain events, burrowing animals), likely requiring human intervention in data flagging for these edge cases.

Previous work has developed broadly applicable automated QA tests that can be applied to a wide range of sensors, including soil water content sensors (e.g., Hubbard et al., 2005), but soil water content sensors also require some sensor-specific tests due to their measurement principle. In particular, most, if not all, moisture sensors are unable to detect frozen water, and sensors typically output very low water contents under frozen conditions, regardless of the actual soil water content. **As a result, flagging data as unreliable when the soil is frozen or close to freezing is an important QA procedure for soil moisture sensors in most temperate, polar, and high elevation regions.**

Since any given data point may pass, fail, or be unassessed (e.g., due to missing thresholds and/or missing data) by any combination of the automated QA tests, it can be useful to synthesize the results into an all-encompassing final quality flag at the published data’s resolution (i.e., averaging interval) so that users have the option to quickly filter for valid data without having to inspect the results of each test. For example, some networks, such as NEON, assign a final quality flag of 0 (i.e., good). For NEON, this flag is applied when less than 10% of data points fail any QA tests and the tests are performed on more than 80% of data points are within the averaging interval (Smith et al., 2014).

Types of Flags and Flagging Frequency

Type I errors should be flagged in near real time, while data must be assessed for Type II errors at least once a year and flagged accordingly. A note should be made on the public-interfacing website describing the type of flags and flagging frequency for each flag.

QUALITY CONTROL

QC may include comparisons or correlations with existing data sources or water balance studies including: (1) correlations or relationships with ancillary data (like temperature, humidity, etc.) and soil properties (like porosity), (2) comparisons with other measured or modelled soil moisture data, (3) checking for expected trends based on long-term temporal analysis of data, and (4) checking for expected relationships between neighboring soil moisture stations (time stability analysis). Each of the four methods could be used in isolation but when used in combination will create the best quality soil moisture data.

QA is especially necessary to identify Type II errors. These can be identified in several ways described below.

Triple Collocation

Triple collocation is a method to understand the ability of a location to provide a representative soil moisture estimate. Three independent methods of estimating soil moisture (usually remote sensing, modelling, and in situ monitoring) are needed to determine the random error of estimation (Chen et al., 2016). Random error estimation is a means of comparing multiple estimates of the same metric to determine if the values match or differ from one another[11]. With a well-functioning sensor, random errors should be low. Random error estimation from in situ monitoring has been estimated as low as 0.01 m^3/m^3, although it is more often in the range of 0.02-0.05 m^3/m^3 for sensors with a soil-specific calibration (Table 5, Chapter 6). There are challenges related to the differences in observation scale because there is a significant difference between the multiple data series; however, this is still a useful tool for estimating error budgets.

Correlations with Ancillary Data and Soil Properties

Various ancillary data can be measured along with soil moisture to assess its quality. These datasets include soil temperature, soil permittivity, potential or actual evapotranspiration, and precipitation. Expected correlations between soil moisture and these variables will often vary by pedon and be hydro-climate specific. Hence, they cannot be borrowed from neighboring stations. Correlations between soil moisture and ancillary data are best developed by long-term data from the same site or by recommendations from an expert. Additionally, all possible response variables measured by a soil moisture sensor like bulk electrical conductivity (BEC), temperature, voltage ratios etc. must be recorded. This ancillary sensor-based information can be used during manual verification of errors in times of uncertainty. Examples of comparisons of soil moisture data with ancillary data (e.g., normalized difference vegetation index [NDVI], evapotranspiration, and temperature) can be found in Zhang et al. (2018), Engstrom et al. (2008), Wang et al. (2007), Dong et al. (2022), and Ford and Quiring (2014).

Comparison with Satellite/Remote Sensing

Remote sensing calibration and validation rely upon ground truth data from in situ stations that monitor the near surface, as this depth is the limit of current soil moisture sensing from L and C band radiometers/radars. Therefore, very often the critical installation depth for a point sensor is at 5 cm, with a sensing volume that can be calibrated to the top 5 cm, the approximate monitoring depth of L-band radiometry. Examples of remote sensing data that can be used for comparisons include Zhang et al. (2017), Colliander et al. (2017), Li et al. (2022), and Wang et al. (2021).

Examples of Datasets for Performing QC Activities

Open-source hydrological, meteorological or vegetation datasets can be accessed through cloud and web-based platforms that provide reference datasets for QA of soil moisture observations. One such application, Application for Extracting and Exploring Analysis Ready Samples (AppEEARS), developed by NASA, provides a convenient on-demand extraction tool for point and area samples. A survey of some relevant datasets available through AppEEARS is given in Table 7.

[11] In some cases, a triple collocation approach can yield a lumped estimate of sensor measurement and representativeness uncertainty. This challenge is described more explicitly in Gruber et al. (2013) and Miralles et al. (2010).

Table 7. A brief sampling of hydrometeorological and vegetation indices products accessible through the AppEEARS platform for point-based comparison with network sensors.

Suite	Full Name	Identifier	Variable	Resolution	Source
MODIS	MODIS/Aqua Vegetation Indices Monthly L3 Global 1 km SIN Grid	MYD13A3	NDVI	1 km, monthly	Satellite
	MODIS/Aqua Vegetation Indices 16-Day L3 Global 250 m SIN Grid	MYD13Q1	NDVI	250 m, 16-day	Satellite
	MODIS/Aqua Net Evapotranspiration Gap-Filled 8-Day L4 Global 500 m SIN Grid	MYD16A2GF	Total ET Total PET	500 m, 8-day	Satellite
DAYMET	Daily Surface Weather Data on a 1-km Grid for North America, Version 4 R1	Daymet_v4	Precipitation Air temp Shortwave radiation	1 km, daily	Ground-based observations and statistical interpolating/extrapolating
SMAP	SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture, Version 5	SPL3SMP_E v005	Soil moisture descending (6 AM) and ascending (6 PM)	9 km, daily	Satellite
	SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 8	SPL3SMP v008	Soil moisture descending (6 AM) and ascending (6 PM)	36 km, daily	Satellite
	SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, Version 6	SPL4SMGP v006	Rootzone (0-100) soil moisture Top layer (0-5cm) soil moisture	9 km, 3-hour	Land surface model with satellite data assimilation

Comparison with AI/Machine Learning (AI/ML) Driven Tools

Spatial estimation of soil moisture can be accomplished with artificial intelligence (AI) combined with hydrologic modeling. AI is ideal for this task because of the complex nature of the relationships between different processes and variables. However, this methodology is largely driven by training data. While it is possible to forecast outside of the observed training domain, the artificial intelligence will rely upon the mechanisms that are observed within the training domain. This can be a particular challenge for extreme or rare events. Drought and flood events are often the most important features that a soil moisture network needs to detect/forecast, but many statistics are not optimized for performance at these extremes.

Gap Filling for Missing Values of Soil Moisture Continuous Data

Gap filling of missing soil moisture data involves following scientific procedures to estimate soil moisture values during times of periodic sensor failure. It is often an important exercise for several stakeholders but, if attempted, must be done with caution. This is especially true if gap filling is conducted over long time periods. Care must be taken to closely follow literature and adhere to all mentioned conditions before incorporating gap filling into a network's protocol. **Gap filled data must also be flagged**.

Many hydrological models require high temporal resolution of inputs such as weather data (precipitation, temperature, wind, solar radiation, etc.). Because high temporal resolution data are rare, many methods have been developed to fill in the gaps with values from other sources, when available, or by different interpolation techniques of the actual incomplete data (Waichler and Wigmosta, 2022; Libohova et al., 2024; Owens et al., 2024). Missing data (gaps) from soil moisture sensors are not uncommon and can happen for many reasons (instrument failure, low battery, accidental damage, funding, etc.). This may in turn result in time series gaps spanning from few hours to days depending on the sensors setting.

Simple techniques, such as linear interpolation, or more complex techniques, such as random forest and other machine learning techniques, can be used successfully to fill in gaps or intervals with missing data. The selection of the techniques depends on the temporal resolution of the sensor and width of the gaps. For example, gaps of a few hours can be filled in through linear interpolation or by calculating the rolling average from the five-hour period centered on the missing time point, across all years. Random forest can be used to fill in wider gaps consisting of multiple days or weeks. Similar techniques can be used to increase or decrease the temporal resolution (timestep) of the moisture data. The coarsening, or decrease, in the temporal resolution is usually more accurate than the opposite, although often finer temporal resolutions are preferred. Data from different sensors can be combined to create a complete dataset, and factors such as sensors type, depth, location, soils, or landscape position need to be considered for pairing the correct appropriate sensors.

In Figure 8, sensors within the watershed boundaries have gaps in data that can be filled out with data from sensors outside of the watershed boundaries. However, sensors need to be grouped based on slope positions (Summit; sideslope (SS); toeslope (TS)); by the stream). For example, the sensor with missing data located on a summit within the watershed (26) needs to be paired with sensors located in the same or similar slope position outside the watershed boundaries (12, 18, 21, and 29). Plotting the moisture data grouped by slope position (Figure 9) shows the gaps and provides the first visual assessment of the potential to fill in the gaps for the sensors within the watershed using sensors outside of the watershed boundaries. However, not all the gaps can be filled: some gaps might be too large, and any approach would not yield accurate results (Figure 10).

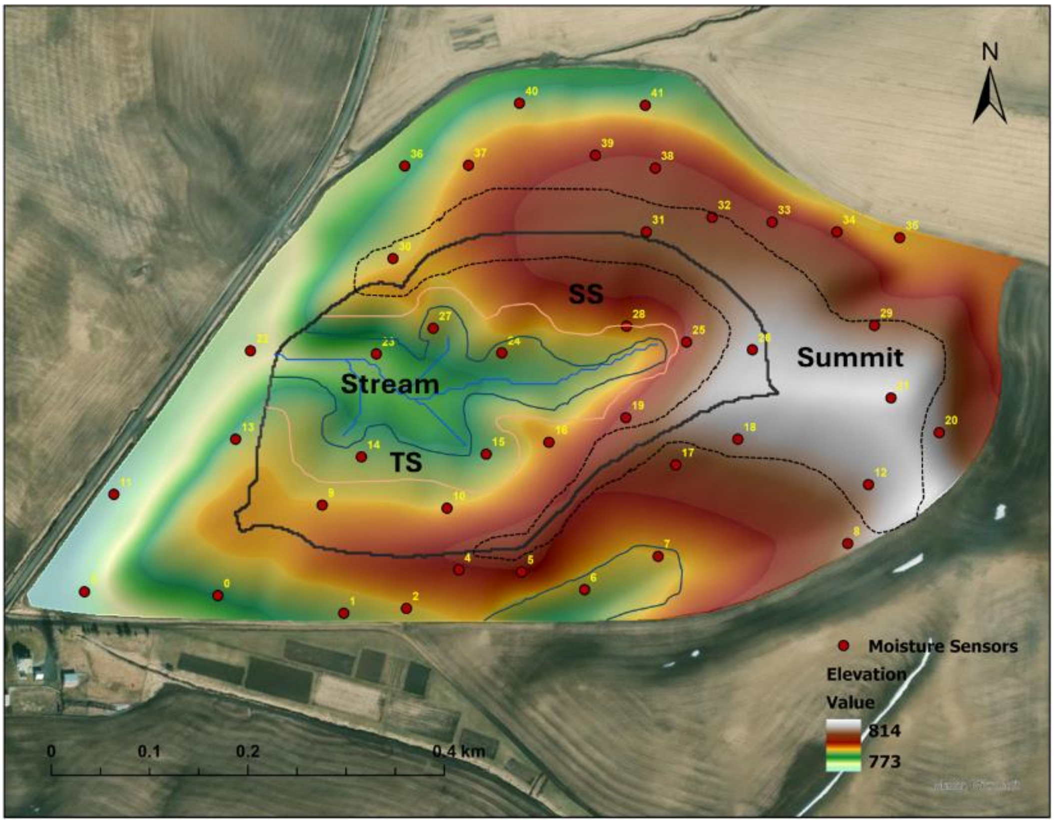

Figure 8. Layout of soil moisture sensors within and outside watershed boundaries and grouped by slope position (Summit; SS – sideslope; TS – toeslope; by the stream). Figure Credit: The Long-Term Agroecosystem Research (LTER) Network site of USDA-ARS Northwest Sustainable Agroecosystems Research, at Cook Farm, Washington State University, Department of Crop and Soil Sciences, Pullman, Washington.

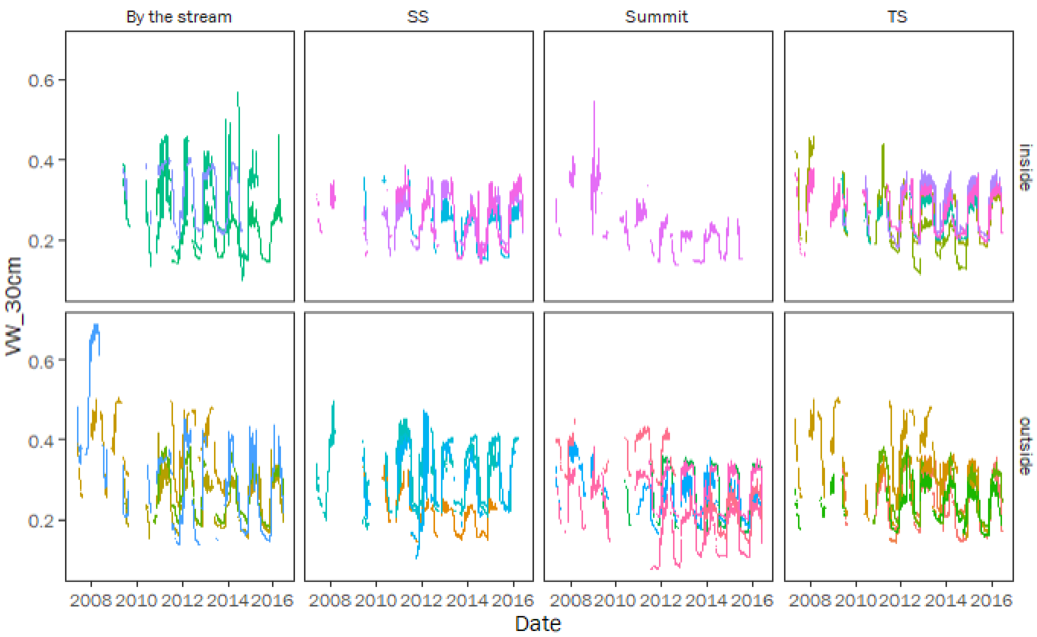

Figure 9. Soil moisture data plotted over time for sensors inside and outside of the watershed grouped by slope position (Summit; SS – sideslope; TS – toeslope: and by the stream). Colors indicate different sensors. Figure Credit: USDA-ARS Northwest Sustainable Agroecosystems Research, at Cook Farm, Washington State University, Department of Crop and Soil Sciences, Pullman, Washington Cook Farm. Data compiled by Caley Gasch, under supervision of David Brown, Department of Crop and Soil Sciences, Washington State University, Pullman, Washington.

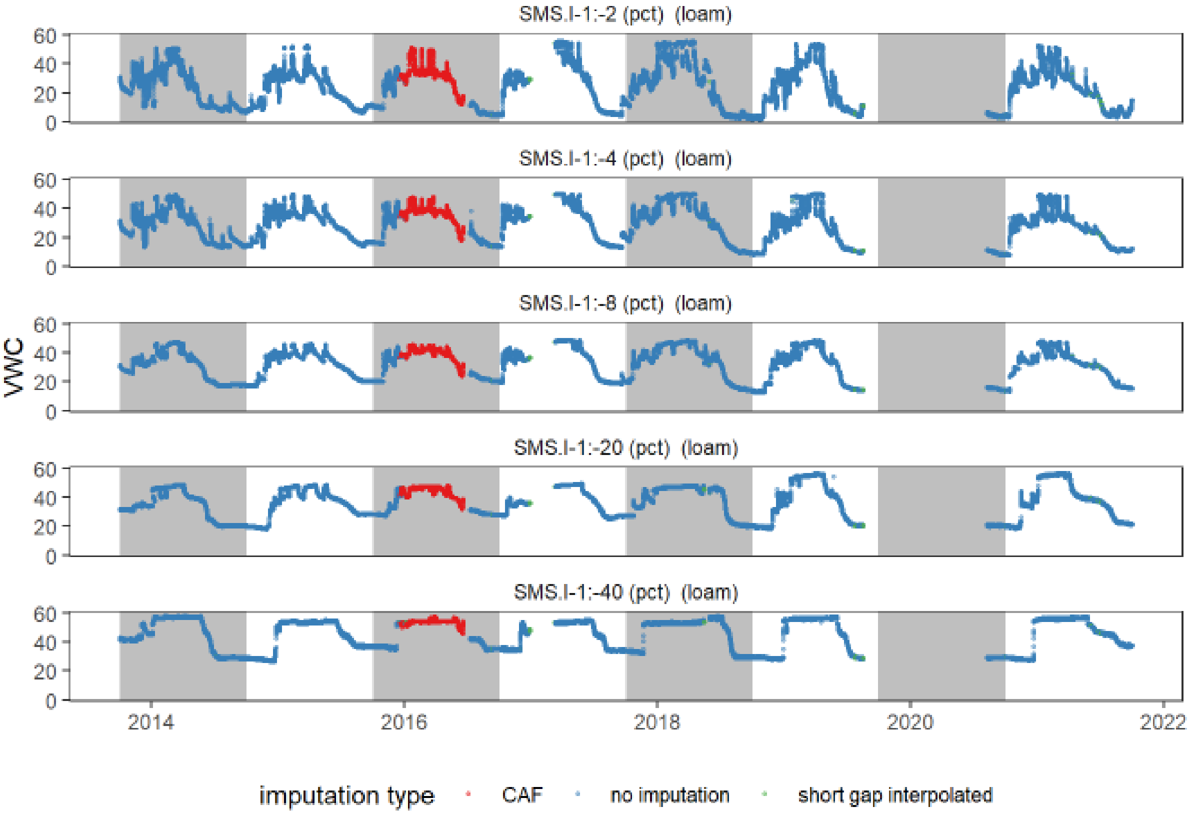

Figure 10. A soil moisture sensor filled in with data from different techniques as described earlier. The red line represents filled in gaps. Figure Credit: USDA-ARS Northwest Sustainable Agroecosystems Research, at Cook Farm, Washington State University, Department of Crop and Soil Sciences, Pullman, Washington Cook Farm. Data compiled by Caley Gasch, under supervision of David Brown, Department of Crop and Soil Sciences, Washington State University, Pullman, Washington.