

A Deep Learning Data Fusion Model using Sentinel-1/2, SoilGrids, SMAP- USDA, and GLDAS for Soil Moisture Retrieval



Vishal Batchu^a, Grey Nearing^a, Varun Gulshan^a.

^a *Google Research*

Corresponding author: Vishal Batchu, vishalbatchu@google.com

Early Online Release: This preliminary version has been accepted for publication in *Journal of Hydrometeorology*, may be fully cited, and has been assigned DOI 10.1175/JHM-D-22-0118.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

ABSTRACT

We develop a deep learning based convolutional-regression model that estimates the volumetric soil moisture content in the top ~5 cm of soil. Input predictors include Sentinel-1 (active radar), and Sentinel-2 (multispectral imagery) as well as geophysical variables from SoilGrids and modelled soil moisture fields from SMAP-USDA and GLDAS. The model was trained and evaluated on data from ~1000 in-situ sensors globally over the period 2015 - 2021 and obtained an average per-sensor correlation of 0.707 and ubRMSE of 0.055 m³/m³, and can be used to produce a soil moisture map at a nominal 320m resolution. These results are benchmarked against 14 other soil moisture evaluation research works at different locations, and an ablation study was used to identify important predictors.

SIGNIFICANCE STATEMENT

Soil moisture is a key variable in various agriculture and water management systems. Accurate and high resolution estimates of soil moisture have multiple downstream benefits such as reduced water wastage by better understanding and managing the consumption of water, utilising smarter irrigation methods and effective canal water management. We develop a deep learning based model that estimates the volumetric soil moisture content in the top ~5 cm of soil at a nominal 320m resolution. Our results demonstrate that machine learning is a useful tool for fusing different modalities with ease, while producing high resolution models that are not location specific. Future work could explore the possibility of using temporal input sources to further improve model performance.

1. Introduction

Soil moisture is one of the primary hydrological state (memory) variables in terrestrial systems (Brocca et al., 2017; Laiolo et al., 2016; Robinson et al., 2008), and is one of the primary controls for agriculture and water management (Dobriyal et al., 2012; Rossato et al., 2017). Soil moisture affects evapotranspiration and vegetation water availability, which are at the core of the climate-carbon cycle (Falloon et al., 2011) and play an important role in hydrological risks such as floods, drought, erosion, and landslides (Kim et al., 2019; Legates et al., 2011; Trambly et al., 2012). Accurate measurement of soil moisture has numerous downstream benefits (Moran et al., 2015) including reduced water wastage by better understanding and

managing the consumption of water (Brocca et al., 2018; Foster et al., 2020), utilising smarter irrigation methods (Kumar et al., 2014) and effective canal water management (Zafar et al., 2021) where soil moisture estimates are used to track and reroute canals to distribute water more effectively.

The most accurate way to measure soil moisture is via ground-based methods such as direct gravimetric measurements (Klute, 1986) or indirect methods such as dielectric reflectometry, capacitance charge, etc. (Bittelli, 2011), which in-situ sensors utilize (Walker et al., 2004). However, in-situ sensors are difficult to scale spatially, and are expensive to install and maintain. Remote sensing-based methods scale globally and provide modestly accurate estimates of top-soil moisture (typically 0-5 cm) (Ahmed et al., 2011), while lowering deployment and maintenance costs relative to ground based methods. While remote sensing soil moisture estimates generally have lower accuracy than in-situ measurements, they scale well spatially.

A large body of remote sensing-based methods use microwave band radiometric reflectance to quantify soil moisture. These sensors can either be located on aerial or satellite platforms, and can be broadly classified into two types - passive and active. Passive remote sensing mainly uses L/C-band brightness temperatures. Early retrieval methods were site specific, semi-empirical models such as Oh (Oh et al., 1992), Dubois/Topp (Dubois et al., 1995) and the IEM models (Baghdadi et al., 2004; K. S. Chen et al., 2003). These methods provide reasonably accurate estimates of soil moisture, however, they are sensitive to site-specific parameters such as soil roughness (Mattia et al., 1997), and are only capable of estimating the soil moisture on bare-ground soils (Verhoest et al., 2008) or soils with low vegetation content. These models have been extensively tested and evaluated (Choker et al., 2017; Ma et al., 2021; MirMazloumi & Sahebi, 2016; Panciera et al., 2014), showing that they are generally not reliable when scaled globally. More generalizable retrieval methods were developed on top of remote sensing systems such as SMAP (Entekhabi et al., 2010) - resolution of 36 km/9 km, SMOS (Y. H. Kerr et al., 2001) - resolution of 30-50 km and AMSR-E (Njoku et al., 2003) - resolution of 25km. Baseline algorithms developed for each of these systems (Chan et al., 2012; Y. Kerr et al., 2012; Njoku et al., 2003) were physics-based models built on top of the microwave brightness temperatures validated with in-situ sensor grids (Al-Yaari et al., 2017; Cai et al., 2017; Q. Chen et al., 2018;

Colliander et al., 2017). In addition to the brightness temperatures, these models also use various globally available static land surface parameters such as soil type, land cover, etc. (Chan et al., 2012). The main limitation of passive microwave products is their coarse spatial resolution, for example the SMAP L2_SM_P_E has a resolution of 9 km x 9 km (Entekhabi et al., 2010) which is too coarse for many agriculture related tasks, e.g., field-scale monitoring.

Active remote sensing methods mainly utilise L/C-band Synthetic Aperture Radar (SAR), and allow for more high-resolution estimates (Haider et al., 2004; Shi et al., 1997). However, the retrieval accuracy of these methods is not as high as passive remote sensing methods (Njoku et al., 2002). Active and passive methods can be combined (Das et al., 2018, 2019; Wu et al., 2017) to improve resolution and maintain accuracy which was an impetus for the SMAP mission. Passive and active methods are complementary – the former is more accurate but has a coarse spatial resolution while the latter is less accurate but has a higher spatial resolution.

Recent work has looked at the combination of passive and active radar with other remote sensing sources, such as optical/thermal imagery (Gao et al., 2017; Ojha et al., 2021). A SMAP + Sentinel-1 product was also developed from the perspective of disaggregating SMAP radiometer brightness temperatures based on Sentinel-1 readings (Das et al., 2019), however, this method has not yet achieved accuracies similar to SMAP at a higher resolution.

In recent years, the use of machine learning (ML) has shown promise in various fields of geoscience (Lary et al., 2016). ML has been used specifically to help disaggregate microwave-based soil moisture estimates into higher resolution products (Abbaszadeh et al., 2019; Kolassa et al., 2018; Liu et al., 2020; Mao et al., 2019). Machine learning provides a unique set of abilities that work well for soil-moisture estimation — enabling the use of large datasets such as the ISMN database (W. Dorigo et al., 2011) without the need for site-specific tuning, modelling non-linear relationships between multiple predictors (remote sensed inputs) and the target (SM), and fusing multiple sources of inputs which could potentially aid in handling vegetation and canopy covers (Greifeneder et al., 2021; Karthikeyan & Mishra, 2021; Lee et al., 2018; Vergopolan et al., 2021). While some of these ML efforts have shown improvements in the estimation of soil moisture compared to previous methods, these techniques mostly depend on a

small number of inputs (Liu et al., 2020) and face issues with speckle in active remote sensing data (Oliver & Quegan, 1998) and are unable to perform scene understanding which is required in soil moisture estimation (Davenport et al., 2008).

To address the aforementioned issues and continue pushing the boundaries of soil moisture models used in the community, we develop, train, and test a deep learning model that provides high resolution (nominal resolution of 320m - discussed in detail in Section 3e) and accurate (average Pearson correlation of 0.727) estimates of soil moisture globally. A few things which combinedly make our approach unique are - the use of high resolution inputs (Sentinel-1/2) capable of scene understanding, the use of deep learning (which doesn't require feature engineering), the development of location-agnostic models that are not location/regional specific, the use of a large scale dataset and the ease of addition of new input sources to the model for future exploration.

2. Data

We generate large datasets for training, evaluation, and testing. Each data point consists of (i) a set of model inputs from various sources (described presently), (ii) soil moisture labels that the model is trained to estimate, and (iii) additional metadata such as timestamp, geographical coordinates etc. These data, their sources, and pre-processing are described in the following subsections.

a. Labels

1) INTERNATIONAL SOIL MOISTURE NETWORK

The largest repository (to our knowledge) of in-situ sensor data for soil moisture is from the ISMN 'network of networks' (W. Dorigo et al., 2011). This dataset has been used extensively for calibrating, training, and evaluating models, and we use it to maintain consistency with previous studies. At each sensor location, ISMN provides measurements in volumetric units at hourly intervals at various depths. The data are quality checked and flagged for anomalies/inconsistencies (W. A. Dorigo et al., 2013) in accordance with NASA's good validation practices (Montzka et al., 2020). We use ISMN soil moisture tagged with a depth of 0-

5 cm and 5-5 cm. Both of these correspond to the top layer of soil moisture, however, the notation differs across different provider networks.

Network	Countries	# Sites/Sensors	# Datapoints	Temporal availability
AMMA-CATCH (Galle et al., 2018)	Benin, Niger, Mali	4	277	2006-2018
BIEBRZA_S-1 (Dąbrowska-Zielińska et al., 2018)	Poland	25	3036	2015-2018
DAHRA*	Senegal	1	4	2002-2016
FR_Aqui (Alyaari et al., 2018)	France	3	1164	2012-Present
MAQU (Su et al., 2011)	China	15	674	2008-2019
NAQU (Su et al., 2011)	China	7	137	2010-2019
REMEDHUS (González-Zamora et al., 2019)	Spain	20	7736	2005-Present
RSMN (Sandric et al., 2016)	Romania	19	9929	2014-Present
SCAN (Schaefer et al., 2007)	USA	179	21210	1996-Present
SMOSMANIA (Calvet et al., 2016)	France	10	3457	2007-2021
SNOTEL (Leavesley et al., 2010)	USA	372	40334	1980-Present
SOILSCAPE	USA	60	394	2011-2017
USCRN (Bell et al., 2013)	USA	114	15551	2000-Present
WSMN*	UK	8	9	2011-2016
Total		828	103912	

Table 1. List of sensor networks from ISMN that we use as a part of our study. The number of data points correspond to the number of samples present in the final data post merge with remote sensed data as described in Section 2c. A * indicate networks where a very small number of data points were available.

The area of study is limited by the presence of in-situ sensors. Although the ISMN repository has a global presence, sensors are primarily located in the US, Europe, Australia and China (W. Dorigo et al., 2021). This limits the applicability of models trained on this data to similar geographies. The sensor networks that we use from ISMN are listed in Table 1.

2) SMAP CORE VALIDATION SITES

In addition to the ISMN network, we also obtain data from the SMAP core validation sites (Colliander et al., 2017) listed in Table 2 and combine it with the ISMN data to train and validate our models.

Network	Countries	# Sites/Sensors	# Datapoints	Temporal availability
FORTCOBB (Starks et al., 2014)	USA	17	4152	2014-Present
KYEAMBA (Smith et al., 2012)	Australia	8	432	2001-Present
LITTLEWASHIT A (Starks et al., 2014)	USA	21	3621	2014-Present
TONZIRANCH*	USA	11	11	2014-2020
TWENTE (Velde & Benninga, 2021)	Netherlands	22	7771	2008-2019
TXSON (Caldwell et al., 2019)	USA	40	3364	2015-2020
YANCO (Smith et al., 2012)	Australia	12	944	2003-Present
Total		120	20295	

Table 2. List of SMAP core validation sites that we use as a part of our study. We only use a subset of all the core validation sites (*SMAP/In Situ Core Validation Site Land Surface Parameters Match-Up Data*, n.d.) due to data access limitations. The number of data points correspond to the number of samples present in the final data post merge with remote sensed data as described in Section 2c.

b. Input Data

Input data for our models were downloaded from Google Earth Engine (Gorelick et al., 2017), which provides a number of different satellite products and geophysical variables that can be combined and exported in a format suitable for machine learning. Earth Engine also facilitates the processing of imagery such as scaling it to a given spatial resolution, performing temporal joins, projecting the imagery to a specific projection etc.

As mentioned in the introduction there are a number of types of remote sensing sources that can be useful for soil moisture estimation (Ahmed et al., 2011). We select sources that have a significant correlation with soil moisture and/or have potential to help with the disaggregation of low resolution soil moisture estimates.

1) HIGH RESOLUTION SOURCES

(i) Sentinel-1 (S1)

The Copernicus Sentinel-1 (Torres et al., 2012) mission (2014-present) by the European Space Agency (ESA) provides global Synthetic Aperture Radar (SAR) readings (Rosen et al., 2000) at regular intervals. It has a revisit time of 6 days at the equator (Torres et al., 2012).

We use the Sentinel-1 GRD (Ground Range Detected) product from Earth Engine which consists of VV (Vertical Transmit - Vertical Receive Polarisation), VH (Vertical Transmit - Horizontal Receive Polarisation) and angle imagery corresponding to dual-band cross-polarized data at a 10 meter resolution. The scenes undergo thermal noise removal, radiometric calibration and terrain correction with the Sentinel-1 toolbox (Veci et al., 2014) to despeckle and denoise.

(ii) Sentinel-2 (S2)

The Copernicus Sentinel-2 (Drusch et al., 2012) mission (2015-present) by the ESA provides high-resolution multispectral imagery (Table 3 lists the Sentinel-2 bands used in this study). It has a revisit time of 5 days at the equator.

Unlike SAR, optical, NIR, and SWIR imagery are dependent on cloud cover and the time of the day of acquisition. A large fraction of the Sentinel-2 scenes have significant cloud cover. We filter the data to retain only scenes containing less than 30% cloud cover (we use the QA60 cloud_mask band from Sentinel-2 to do this).

We use the L1C top of atmosphere product, which consists of multiple bands with varying resolutions ranging from 10-60m. All the bands were upscaled (using the nearest neighbour method) or maintained at a 10m resolution to use as inputs for our models.

Band	Resolution	Wavelength	Description
B2	10 meters	496.6nm (S2A) / 492.1nm (S2B)	Blue
B3	10 meters	560nm (S2A) / 559nm (S2B)	Green
B4	10 meters	664.5nm (S2A) / 665nm (S2B)	Red
B8	10 meters	835.1nm (S2A) / 833nm (S2B)	NIR
B11	20 meters	1613.7nm (S2A) / 1610.4nm (S2B)	SWIR 1
B12	20 meters	2202.4nm (S2A) / 2185.7nm (S2B)	SWIR 2

Table 3. Sentinel-2 bands we use and their wavelengths/properties.

(iii) *NASA Digital Elevation Model (DEM)*

Digital elevation models capture the topography of bare ground which helps estimate the amount of moisture the surface can hold. The ‘elevation’ band from NASA Digital Elevation Model (DEM) (Jpl, 2020) on Earth Engine provides 30 meter resolution estimates and is a reprocessing of the widely used SRTM product (Farr et al., 2000) which improves the model globally. Data from ASTER GDEM, ICESat GLAS and PRISM are incorporated into the SRTM product to produce the refined NASA DEM. This is a single time acquisition product (acquired in the year 2000) but the variation over time is small, hence the product is still relevant for the task at hand.

2) LOW RESOLUTION SOURCES

(i) *Soil Grids*

SoilGrids (Hengl et al., 2017) provides us with various environmental and soil profile layers with global coverage. We specifically use soil texture – i.e., sand, silt, and clay fractions - and bulk density. All of these mapped products are present at a resolution of 250m.

We do not use pedotransfer functions explicitly since our models can learn the required mapping. Soil information is essential because it is related to maximum water holding capacity (porosity), infiltration and to some extent, evaporation rates, which are critical controls on water retention and storage (Beale et al., 2019; Pan et al., 2012). Soil information obtained from soil grids has been used as a proxy for the disaggregation of soil moisture in prior work (Leenaars et al., 2018; Montzka et al., 2018).

3) COARSE SOIL MOISTURE PRODUCTS

(i) *SMAP-USDA Soil Moisture*

The 'NASA USDA enhanced Soil Moisture Active Passive (SMAP) soil moisture product (Mladenova et al., 2020) provides surface level (0-5 cm) soil moisture estimates at a 10km resolution. The product is produced by applying a Palmer model (*Meteorological Drought*, 1965) followed by 1-D Ensemble Kalman Filter (EnKF) (Evensen, 2003) to assimilate the Level 3 SMAP product. It has a temporal revisit period of ~3 days.

The ‘ssm’ band provided in this product corresponds to surface soil moisture and is in units mm which is equivalent to kg / m². We convert this to volumetric soil moisture content by using 1000 kg / m³ as the density of water and a measurement depth of 5cm¹.

Ideally, we would have liked to use the SMAP L4 9 km product here, but we chose SMAP-USDA as it was available in Earth Engine, which enabled us to build large scale data pipelines for this task easily. SMAP L4 data was unfortunately not available in Earth Engine.

(ii) *GLDAS Soil Moisture*

Global Land Data Assimilation System (GLDAS) 2.1 (Rodell et al., 2004) uses land surface modelling and data assimilation techniques to model various land surface states and fluxes. They provide soil moisture products at various depths, of which 0-10 cm (‘SoilMoi0_10cm_inst’ band) is closest to the depth that we want to estimate at (0-5 cm) at a 25 km resolution. Although their soil moisture product does not map exactly to the top level surface soil moisture, it still correlates well and is a useful input. Being a modelling product, GLDAS provides information indirectly from recent rainfall and meteorological inputs. Estimates are produced at a ~3 hr interval.

Similar to the SMAP-USDA data, we convert GLDAS SM provided in mm (kg/m²) to volumetric soil moisture using a measurement depth of 10 cm.

4) INPUT NORMALIZATION

All inputs from Earth Engine data sources are normalized into a consistent range to prepare inputs for our machine learning model. A linear min-max scaling ($x' = (x - x_{min}) / (x_{max} - x_{min})$) is used for most of the data sources as shown in Table 4.

Source	Band	Min	Max
--------	------	-----	-----

¹ *Additional information on the volumetric soil moisture conversion at <https://ldas.gsfc.nasa.gov/faq/>*

Sentinel-1	VV	-25 dB	5 dB
	VH	-25 dB	5 dB
	Angle	0 Degrees	90 Degrees
Soil Grids	All bands	0 cg/cm ³	800 cg/cm ³
NASA DEM	elevation	0 m	3000 m
GLDAS	SoilMoi0_10cm_inst	0 kg/m ²	100 kg/m ²
SMAP-USDA	ssm	50 mm	255 mm

Table 4. Source statistics for each of the Earth Engine sources we use. Each of these sources is normalized via linear min-max scaling.

Coming to Sentinel-2, cloudy pixels in Sentinel-2 scenes have large reflectance values compared to non cloudy pixels. Due to this, a min-max linear scaling across the entire data results in a small dynamic range for non-cloudy reflectance values. To account for this, we use a logarithm-based nonlinear scaling method (Brown et al., 2022) that results in better dynamic ranges for non-cloudy pixels. Additional details are specified in the Appendix.

c. Creating the dataset

Data from all the sources specified above are joined with a distributed Earth Engine export pipeline to create datasets for training and testing.

This pipeline consists of the following steps:

1. Filter ISMN dataset based on the quality flags. We retain data that has the following ISMN flags - "G" (Good), "C02" (Soil moisture > 0.6), "C03" (Soil moisture > saturation point), "C02,C03" (Soil moisture > 0.6 and soil moisture > saturation point).
2. For each remaining ISMN data point consisting of latitude, longitude, timestamp, and soil moisture reading:
 - a. For each earth engine source
 - i. [Sentinel-2 source only] Filter out images with >30% cloud percentage.

- ii. Perform a spatial join to find matching images where the (latitude, longitude) of ISMN data point lies within the image bounds.
- iii. Perform a temporal join to retain images within a specified time bound in the past, where the bound depends upon the input source as specified in Table 5.
- iv. Pick the temporally closest image to the ISMN data point from the filtered images. If none are available, we discard this data point.
- v. Reproject the image to 10 m resolution and the corresponding UTM projection based on the UTM zone of the data point.
- vi. Crop the image to extract a 512x512 sized region centered around the (latitude, longitude) of the ISMN data point.
- vii. Normalize the image following Table 4.
- viii. Pair this image with the data point.

Since each of the input sources have different revisit intervals that don't align perfectly with each other - eg.: Sentinel-1 (revisit of 6 days²) and Sentinel-2 (revisit of 5 days), we temporally anchor the data to Sentinel-1 (our primary high-res input) whilst setting temporal bounds for the other sources that allow us to acquire meaningful information. The next section describes the dataset created.

1) DATASETS CREATED

(i) *Sentinel 1 Anchored Dataset - Full*

- For creating the dataset, we used Sentinel-1 data as the primary input. We chose to anchor upon Sentinel-1 since it is resilient to atmospheric conditions and is directly sensitive to soil moisture, unlike Sentinel-2 which primarily provides information for scene understanding. Data from other sources was used to enrich the Sentinel-1 inputs, so we allowed for a greater time slack there.
- The dataset is created using the Sentinel-1 anchored temporal bounds specified in Table 5. It is then split into train, validation, and test splits with a 60:20:20 ratio in an IID

² Sentinel-1 had a revisit period of 6 days before the Sentinel-1 B malfunction.

(independent and identically distributed) manner at a sensor level ie. 20% of the sensors are put under validation, 20% under test and the remaining are used for training. Data points belonging to a single sensor are always together.

- The dataset comprises a total of 124,207 data points of which 80,867 (65.1%) data points are in the train split, 22,543 (18.1%) data points are in the validation split and 20,797 (16.8%) data points in the test split. Note that since all the sensors don't have the same number of datapoints, the 60:20:20 ratio is not strictly present across data points although it is present across sensors.
- Sample input imagery and labels are present in the Appendix.

Source	Temporal bounds for Sentinel-1 anchored dataset
Sentinel-1	1 hr
Sentinel-2	14 days
SMAP-USDA	3 days
GLDAS	6 hrs
NASA DEM	Static
Soil Grids	Static

Table 5. Temporal bounds for each of the sources used during the join while creating the dataset. Note that in-situ readings are present on an hourly basis. The bold cell indicates the strictest bound (1hr).

In Table 5, the 1 hr bound ensures that each image in Sentinel-1 is paired with a maximum of 1 in-situ sensor readings, as the sensor readings are never less than 1 hour apart. This ensures that we don't create duplicate data points in the dataset where the exact same set of imagery is paired with multiple in-situ readings.

(ii) Sentinel 1 Anchored Dataset - Hard

- The dataset created above (1) is filtered further to create a harder dataset. The train split remains the same as above. The validation and test splits, however, are filtered to ensure that for each datapoint, the distance to the closest training sensor is $> 25\text{km}$. This results in much stronger validation/test splits that showcase model generalizability.
- This also results in a smaller dataset with fewer sensor networks present in validation/test splits but provides us with a stronger dataset for us to validate our approach against allowing us to test model generalizability and applicability in the real world.
- The dataset comprises a total of 99,411 data points of which 80,867 data points are in the train split, 8,794 data points are in the validation split and 9,750 data points in the test split.

Note that we use the hard dataset primarily throughout the paper with an exception of Sections 5d (Benchmarking) and 6 (Discussion on metrics vs distance from nearest train sensor) where the full dataset is used since it is better suited. Additional details are provided in the respective sections.

2) DATASET STATISTICS

All the statistics provided in this section correspond to the full Sentinel-1 anchored dataset.

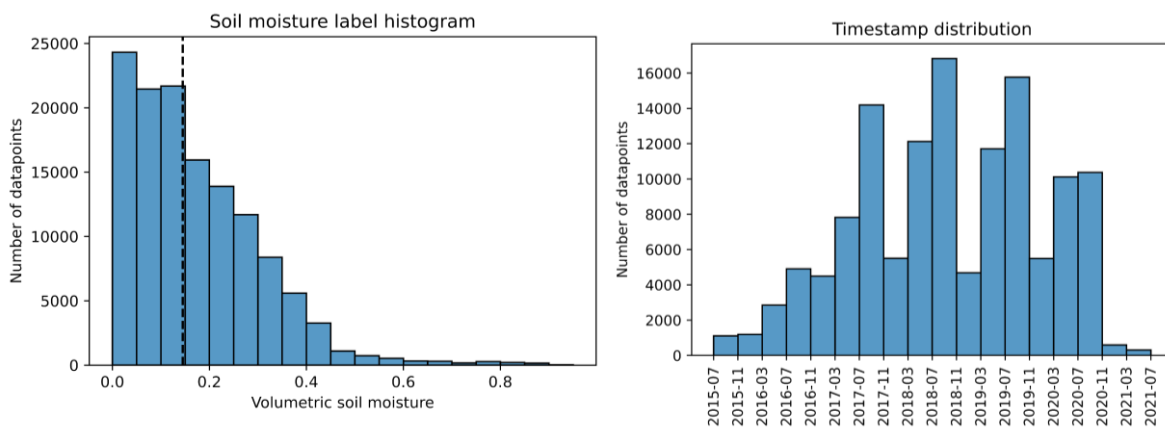


Fig. 1. *Left*: Volumetric soil moisture m^3/m^3 label distribution. The dashed line indicates the median soil moisture. *Right*: Timestamp distribution.

Volumetric soil moisture label distribution in Figure 1 shows that the labels are skewed towards the lower end, this can be attributed to the fact that most soils are not usually saturated.

The timestamp distribution in the same figure shows that there's a fair spread of data across time, however, winter seasons (seasonality considering the northern hemisphere since most of our data is from the northern hemisphere) have fewer data points noticeable as spiky dips, which can be attributed to filtering out soil moisture readings where the temperature is below 0°C via ISMN flags as well as cloud filtering on Sentinel-2.

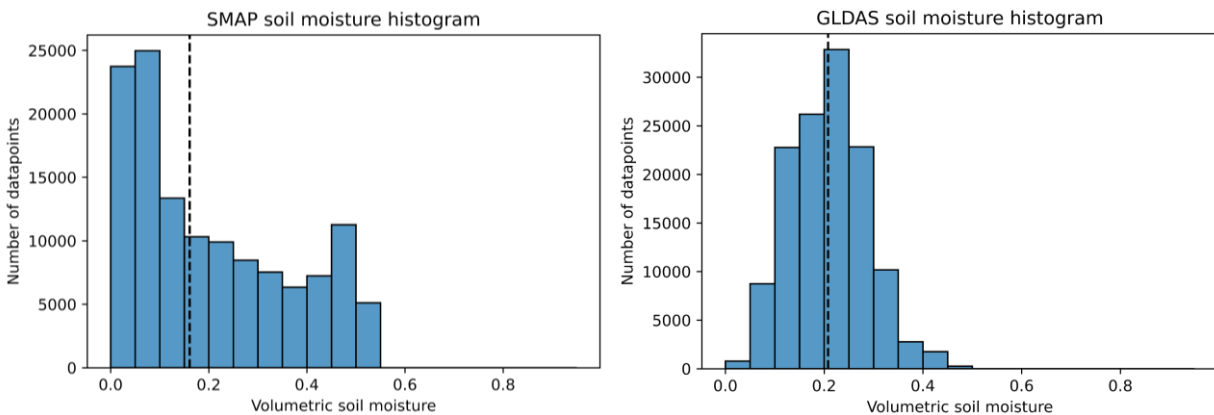


Fig. 2. Distribution of coarse soil moisture products. *Left*: SMAP-USDA volumetric soil moisture m^3/m^3 . *Right*: GLDAS volumetric soil moisture m^3/m^3 . The dashed lines indicate the corresponding median soil moisture values. Note that both the distributions are computed from the same set of datapoints.

Distribution of coarse soil moisture products in Figure 2 shows that the SMAP-USDA soil moisture histogram resembles the label histogram in Figure 1 more closely when compared to the GLDAS histogram, likely because GLDAS estimates soil moisture at a depth of 10cm.



Fig. 3. A heatmap of in-situ sensor locations present in the dataset (obtained using *Heatmaps*, n.d.)).

Figure 3 shows a heatmap of sensor locations present in the dataset. Around 83% of the sensors are in the US, 11% in Europe and the remaining around the globe. Note that this shows the distribution of in-situ sensors and not the actual data points so each sensor is represented only once.

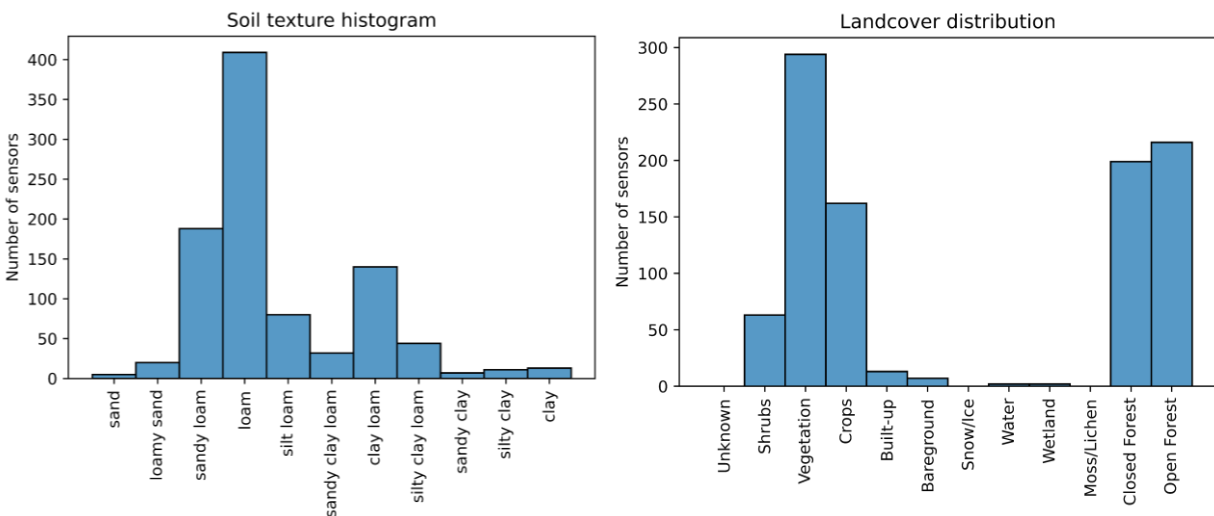


Fig. 4. *Left*: USDA based soil texture distribution. *Right*: Land cover distribution derived from the Copernicus Land Cover Map across the sensors present in the dataset.

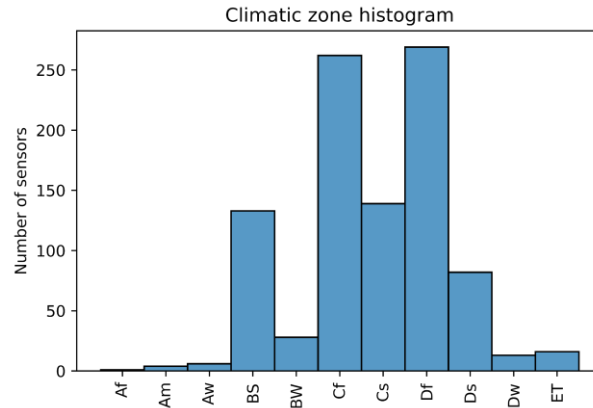


Fig. 5. Köppen climate zone distribution across the sensors present in the dataset. We use the first two identifiers of the Köppen classification only in order to group similar climatic zones together.

The soil texture distribution in Figure 4 shows the distribution of our sensors based on USDA soil texture classification (Groenendyk et al., 2015) where the input fractions of sand, silt, and clay were obtained from SoilGrids. Land cover distribution in Figure 4 shows that the majority of sensors are spread across vegetation, croplands and forests. Note that the two sensors that fall under the “Water” class correspond to sensors that are very close to rivers/water bodies and are not actually inside the water. This is a fairly broad distribution and allows us to see how our model performs across these different land cover types. We obtain land cover classes for each sensor from the Copernicus Land Cover Map (Buchhorn et al., 2020) on Earth Engine. Figure 5 shows the distribution of our sensors across the different climate zones (Kottek et al., 2006). The land cover, soil texture and climate zone distributions are similar for training, validation, and test (these distributions are shown in the Appendix).

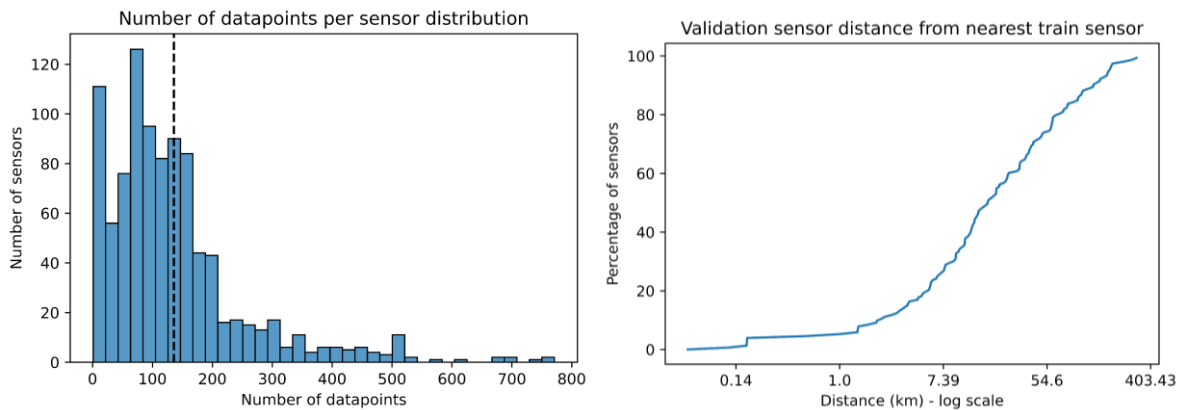


Fig. 6. *Left*: The distribution of the number of data points per sensor. The dashed line indicates the median. *Right*: Cumulative histogram of validation sensors with respect to the distance from the nearest train sensor. The distance on the x-axis varies at a log scale.

The distribution of the number of data points per sensor in Figure 6 (left) shows that we have ~100-200 data points per sensor on average. This ensures that the data spans multiple seasons and years for a majority of the sensors. For each validation sensor that we evaluate on, we compute the geodesic distance to the nearest train sensor and show these as a density curve in Figure 6 (right). This provides an understanding of how validation sensors are distributed with respect to train sensors.

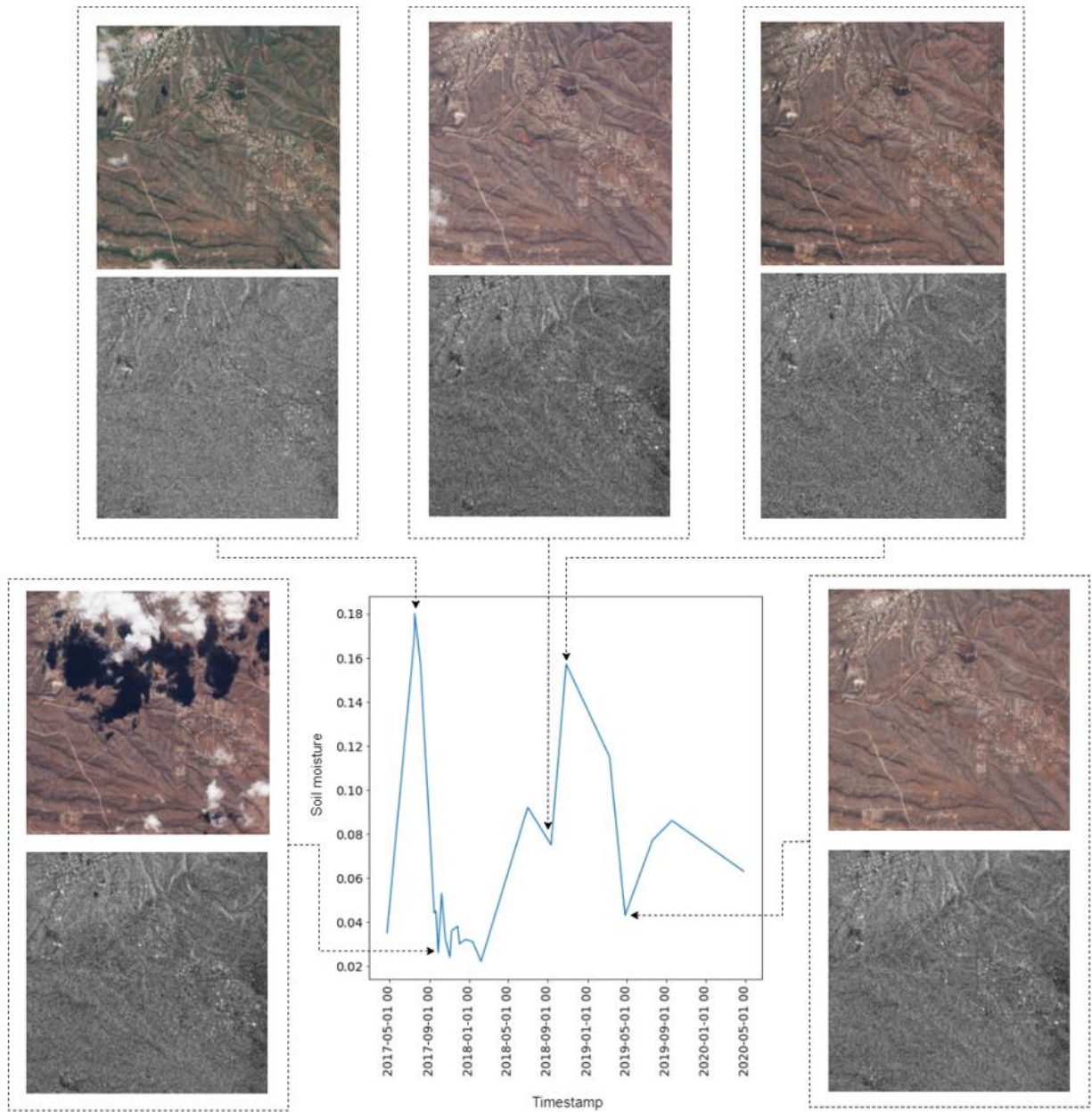


Fig. 7. Variation of soil moisture at a randomly selected sensor across time and corresponding input Sentinel-2 RGB (top in each dotted image pair) and Sentinel-1 VV (bottom in each dotted image pair) imagery.

Figure 7 captures the temporal variation of soil moisture at a single sensor. We observe that Sentinel-2 inputs can be quite visually indicative of the vegetation growth and dryness of a region which correlate with soil moisture.

3. Methods

The problem of soil moisture estimation is framed as an image based regression task (Babu et al., 2016; Fu et al., 2018; Rogez et al., 2017) where remote sensed sources and geophysical variables are used as inputs. These input sources provide the spatial and spectral covariates to estimate surface level soil moisture. We employ various deep learning techniques that have proven to work well for image based regression. These techniques allow our models to be site-agnostic, which results in better generalization compared to site-specific/calibrated models, providing the capability to scale globally.

a. Model

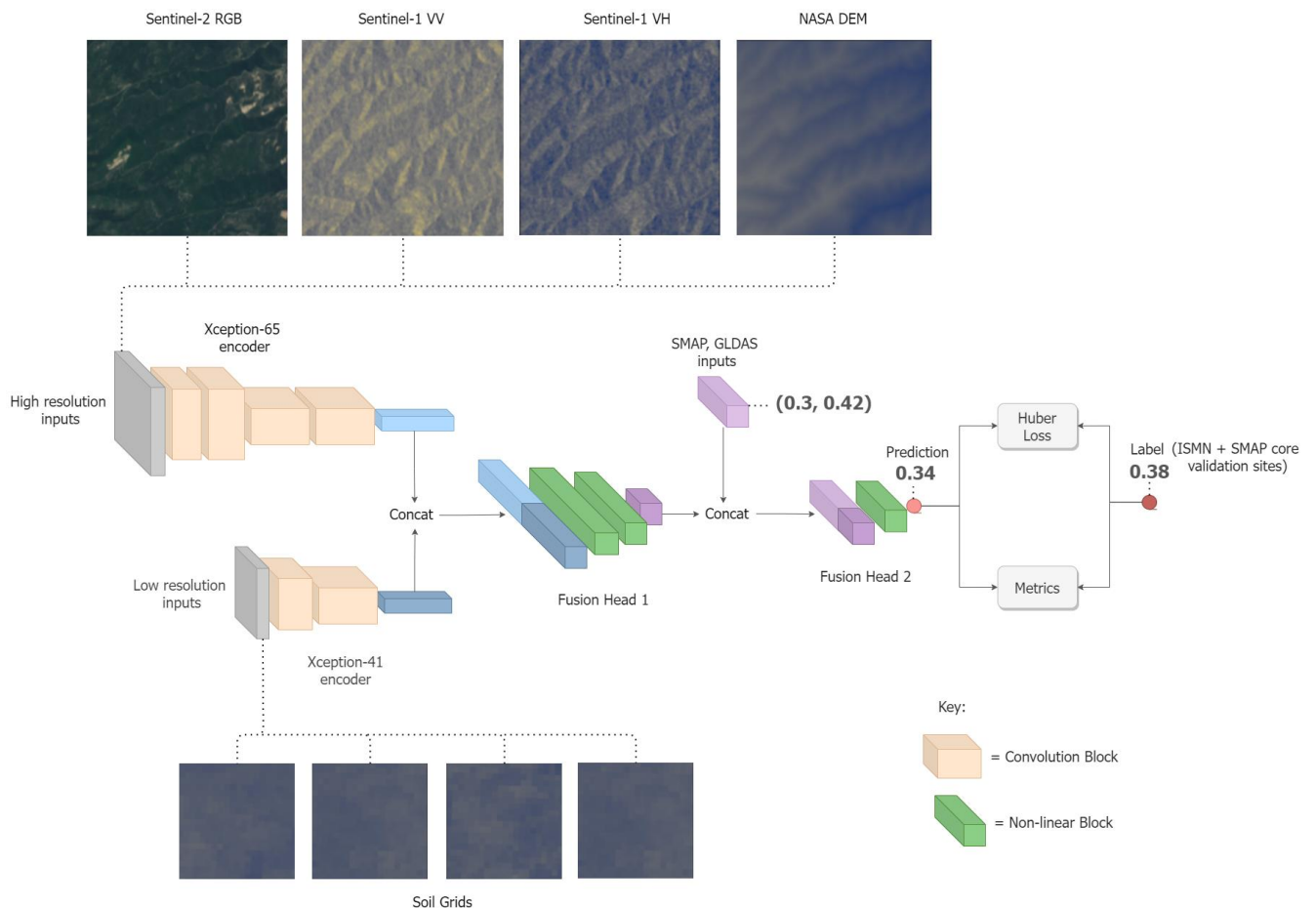


Fig. 8. Our model architecture for the task of soil moisture estimation.

Starting with the two sets of inputs (high-resolution and low-resolution), we encode each set of the inputs with Xception based feature encoders generating a single embedding (a compact

feature representation with no spatial dimension - detailed in the next section) for each set of inputs. The hypothesis is that embeddings generated from the high resolution inputs capture a combination of the scene embedding (Main-Knorn et al., 2017; Raiyani et al., 2021) from Sentinel-2 and DEM sources and a potential rough soil moisture estimate (Paloscia et al., 2013) from the Sentinel-1 and Sentinel-2 sources. Likewise, the embeddings generated from the low resolution inputs would capture soil properties such as the water holding capacity, soil type, density and texture. These embeddings are then fused together by first concatenating both embeddings and then passing the concatenated vector through fusion head 1, which is a stack of [Dropout (Srivastava et al., 2014), Fully Connected (FC), Batchnorm (Ioffe & Szegedy, 2015), Activation] layers followed by a [Dropout, FC] block at the end, described in detail in Section 3a2, resulting in a compact representation of features relevant for soil moisture.

We then fuse the coarse soil moisture inputs (SMAP and GLDAS SM estimates) with the output of the first fusion head, by passing the concatenated vector through another set of fully connected layers, called fusion head 2. This produces the final soil moisture estimate that is passed to the loss function to compare with training labels.

We also considered early-fusion approaches where the low resolution inputs are concatenated with the high-resolution inputs. However, these require the low resolution inputs to be scaled to the size of the high-resolution inputs which involves a large amount of duplication and wasted compute.

1) INPUT ENCODERS

We use the Xception (Chollet, 2017) encoder for the feature encoders. We have also experimented with ResNet (He et al., 2016) and MobileNet-V2 (Sandler et al., 2018) encoders – empirically Xception performed the best out of these choices of encoders.

In particular, we use the Xception-65 encoder to encode high-resolution imagery and an Xception-41 to encode low resolution imagery. Low resolution imagery has a much smaller amount of input pixels which contain more semantically meaningful features and hence a smaller

encoder with a fractional depth multiplier (Chollet, 2017) works well and allows us to utilize model capacity better - faster training and fewer parameters.

The input to the high-res encoder consists of a 256 x 256 (we take a center crop from the 512 x 512 image in the dataset for computational reasons - ablation study in section 5d1) pixel image centered at the location of the sensor. High resolution imagery is sampled at a 10m resolution resulting in a 2.56 x 2.56 km region being used as the input. The low-res encoder uses a 16 x 16 input size where low resolution imagery (at 250m original resolution) is bilinearly resampled at 160m to ensure that we cover the same 2.56 x 2.56 km region. The large input region allows the model to understand the context around the center of the image to estimate the soil moisture at the center. The impact of varying the region size on the model performance is explored in the experiments section.

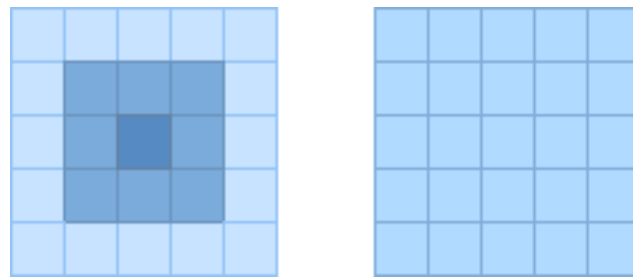


Fig. 9. A representative comparison of the features generated by the Xception encoder. *Left:* Center weighted pooling. *Right:* The standard average pooling. Darker colors indicate a higher weight placed on a particular pixel. The sum of weights in both cases equals 1.

The high resolution encoder produces 8 x 8 x 2048 sized features. We then apply a center weighted global pooling as shown in Figure 9 on them to ensure that the embeddings generated at the center are given the most importance since we are estimating the soil moisture at the center of the image. This produces a 1 x 1 x 2048 sized embedding. The low resolution encoder produces 1 x 1 x 2048 sized features directly. These embeddings are flattened and then stacked to produce a 4096 sized embedding which is passed on to the fusion head.

2) FUSION HEAD

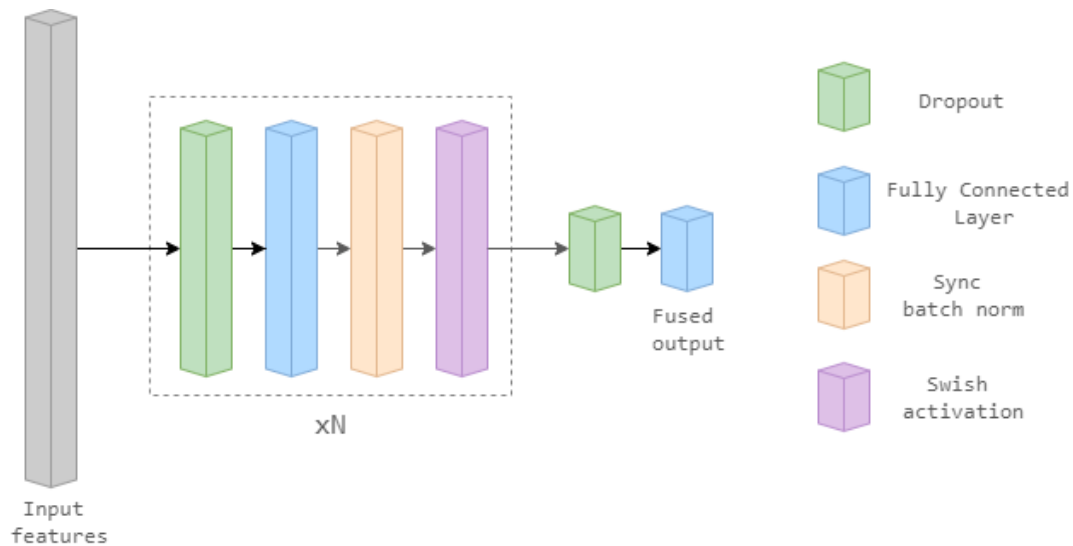


Fig. 10. Our fusion head architecture consists of fully connected, batch norm, activation and dropout layers. The dotted region shows one block of the fusion head which is repeated N times.

The fusion head consists of stacked blocks of [Dropout, Fully Connected (FC), Batchnorm, Activation] with an additional set of dropout and FC layers towards the end without an activation or batch normalization since we don't want to restrict our predicted output range. This head allows for nonlinear fusion of concatenated input features. Each fusion head can be configured by specifying the number of output channels per block and the number of blocks present.

We use two fusion heads in our model. The first one (with 3 blocks and output channels per block [128, 64, 1]) fuses information from the two input encoders encoding high-resolution and low resolution imagery respectively. The second one (with 2 blocks and output channels per block [8, 1]) fuses information from the coarse soil moisture inputs with the fused embedding generated by the first fusion head producing the final soil moisture estimate. Additional hyperparameter details are presented in Section 3c.

b. Training

We train our model using Tensorflow in a synchronized distributed TPU (v1) (Cheng et al., 2017) setting where each model is trained on a 2x2 pod (8 TPU cores). We use Stochastic Gradient Descent (SGD) with momentum (Rumelhart et al., 1986) as the optimizer. Training

takes around 8-10 hrs on average to complete. The learning rate transitions from the initial value to 0 via a polynomial decay schedule with a power of 0.9 across each minibatch during training. Additional hyperparameter details are presented in Section 3c.

In addition to the distributed training strategy, we also utilize the Tensorflow data service (Murray et al., 2021) allowing us to distribute dataset pre-processing across a cluster of CPUs.

1) LOSS

We use the Huber loss (Huber, 1992) to train our models. The Huber loss is a combination of mean squared error (MSE) and mean absolute error (MAE) which allows our model to efficiently optimize inliers whilst paying lesser attention to outliers, delta specifying the threshold on determining what constitutes as an outlier.

$$loss = \begin{cases} \frac{1}{2} * (x - y)^2 & \text{if } (|x - y| \leq \delta) \\ \delta * |x - y| - \frac{1}{2} * \delta^2 & \text{otherwise} \end{cases}$$

(Huber, 1992)

We tried using MSE and MAE loss functions, but the Huber loss performed better empirically.

2) REGULARIZATION

Although the dataset consists of a large number of data points, it contains only 958 distinct locations since the number of sensors is limited. Due to this, our models could overfit to the small set of train locations. Regularization helps prevent this and improve generalization. We employ several regularization techniques to reduce overfitting such as Weight Decay (Loshchilov & Hutter, 2019), Dropout (Srivastava et al., 2014) and Drop path (Huang et al., 2016)

(i) *Augmentations*

We also apply various augmentations on our inputs as an additional form of regularization (Shorten & Khoshgoftaar, 2019). We only use center preserving augmentations since we are estimating the soil moisture at the center of the image (see discussion about center-weighted pooling in Section 3a1). We don't use any color transform augmentations since our remote sensing inputs have a much larger spectral range (i.e., not just RGB imagery) and we do not want to lose the signal present in the absolute value of pixels.

Augmentation	Probability	Extent
Rotate	0.8	+ - 180 degrees
Flip Horizontal	0.5	None
Flip Vertical	0.5	None
Shear X	0.3	+ -9 degrees
Shear Y	0.3	+ - 9 degrees

Table 6. Details on the exact set of augmentations used. The probability here indicates the probability with which a specific augmentation is applied on an input.

For every datapoint, each of the augmentations are applied based on their probabilities. This results in a large number of combinations of these augmentations allowing us to produce a wide variety of augmented images, similar to the augmentation policies described in autoaugment (Cubuk et al., 2019).

c. *Hyperparameters*

The high-resolution encoder uses an Xception-65 backbone with an input crop size of 256 (selected from a sweep of [512, 256, 128, 64] discussed in the ablation in Table 14). The backbone has an output stride of 32 and uses a drop-path probability of 0.6 (selected from a sweep of [0.8, 0.6, 0.4, 0.2]). The low-resolution encoder uses an Xception-41 backbone with an input crop size of 16 and an output stride of 16 with the same drop-path probability as the high-resolution encoder.

The first fusion head has a shape of (128, 64, 1) with drop-keep probabilities of (0.5, 0.5, 0.6) and uses Swish (Ramachandran et al., 2017) activations (selected from a sweep of [Swish, ReLU, Leaky ReLU]). The second fusion head is relatively simple and has a shape of (8, 1) with drop-keep probabilities of (0.9, 0.9) and uses the same Swish activations similar to the first fusion head.

We use a global batch size of 64, which is split across the 8 TPUs for training (data parallel distribution) with synchronized batch norm. We trained each model for a total of 600k steps to ensure training curves plateau (not reported).. We use a huber-loss delta of 0.4 (selected from a sweep across [0, 0.2, 0.4, 0.6, 0.8, 1.0]). We use the momentum optimizer with momentum set to 0.9. We perform a joint sweep on the learning rate across [0.1, 0.05, 0.03] and weight-decay across [0.0001, 0.00005, 0.00003] and pick the best performing model on our validation set.

We tried using ImageNet checkpoints to initialize the Xception encoders since ImageNet based initializations have shown to perform well in various transfer learning scenarios (Kornblith et al., 2019), but we did not see any significant improvement in performance for the task at hand. Hence, we just use random initialization for the weights of our models.

d. Evaluation

We evaluate our models both quantitatively and qualitatively. As a part of the qualitative evaluation, we plot multiple time series of model estimates and labels to visually assess temporal variation at randomly selected sensors. This is useful to identify whether any biases are present at different sensor locations, and to visually assess how biases might differ based on properties of the location such as the soil type, land cover etc. For quantitative evaluation, we look at a set of metrics on the best model checkpoint during training as described below. During evaluation, we clamp model predictions to a range of [0., 1.] since that is the valid range for soil moisture predictions.

1) METRICS

To measure the performance of our models, we compute various metrics for each sensor and then average them over all our sensors where at least 13 data points are present at each sensor.

We wanted to evaluate only on sensors where there is data for at least 3 months or more. Considering a Sentinel-1 revisit period of around 6-7 days, 3 months would have at least 13 data points.

$$\text{final_metric} = \frac{\sum_{s=1}^M \text{metric}_s}{M}$$

Where,

M = number of sensors having ≥ 13 data points

metric_s = per sensor metric

The metrics that we report are: (i) ubRMSE (m^3/m^3), (ii) RMSE (m^3/m^3), and (iii) correlation (r), which are described as follows:

Notation used in the metrics below,

x_i = Model estimate

y_i = Label

\bar{x} = Mean model estimate

\bar{y} = Mean of labels at a particular sensor location

N = Number of data points

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}}$$

$$\text{ubRMSE} = \sqrt{\frac{\sum_{i=1}^N ((x_i - \bar{x}) - (y_i - \bar{y}))^2}{N}}$$

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}$$

We focus on the ubRMSE since it is often the case that soil moisture estimates have large climatological differences with in-situ sensors, however, we are often concerned with per-location dynamics instead of absolute values. RMSE and ubRMSE are unbounded metrics, and

therefore difficult to contextualise outside of inter-model comparisons. We therefore also report the standard Pearson product-moment correlation coefficient.

e. Model resolution

Our models produce soil moisture estimates at a nominal resolution of 320m. This is because our encoders reduce the high resolution input (256 x 256 pixels @ 10m per pixel) by a factor of 32 to create a feature representation that is 8 x 8 pixels @ 320m per pixel. We use the term “nominal” because, although the model produces representations at a 320m resolution, nearby predictions are not completely independent and even when we stride out inputs at 320m to get estimates for adjacent locations, the inputs still overlap.

Note that we use point-estimates from in-situ sensors at the center of the image as labels to train the model rather than an average of soil moisture values across a 320x320m pixel which is the resolution at which the model produces soil moisture estimates. Ideally, we would have liked to use dense sensor grids within 320x320m regions and take the average across in-situ sensors in the grid to produce the labels. However, due to data limitations, we stick to using just the readings from single in-situ sensors at the center of each datapoint as the label.

4. Experiments

We adopt the following model naming convention: the names of all the sources that are used as inputs to the model are concatenated to identify the model. Please refer to Figure 8 for details on how each of the inputs are used.

Hereafter, we refer to the Sentinel-1 + DEM + Sentinel-2 + SoilGrids + SMAP-USDA + GLDAS model as “our model” and will be the model used in experiments, results and other analyses unless specified otherwise.

a. Overall model performance

The following set of baselines and ablations are used to measure the performance of our model quantitatively. Results are reported on the test set of the Sentinel-1 anchored hard dataset.

- **Baseline:**
 - **SMAP-USDA + GLDAS NN (Neural Network)** - The coarse soil moisture inputs passed through a fusion head to produce a soil moisture estimate.
- **Sentinel-1 + DEM + Sentinel-2 + SoilGrids + SMAP-USDA + GLDAS** - Our model as described in Figure 8.
- **Ablations:**
 - **Sentinel-1 + DEM** - A Sentinel-1 only model (DEM allows the model to factor in the terrain)
 - **Sentinel-1 + DEM + Sentinel-2** - A Sentinel-1 and Sentinel-2 combined model where both of the inputs are passed together to the high resolution encoder whose embeddings are then passed through a soil moisture head.
 - **Sentinel-1 + DEM + Sentinel-2 + SoilGrids** - Similar to the previous model, except that SoilGrids is also used.

b. Spatial and Temporal Analysis

1) SPATIAL STRATIFICATION OF MODEL PERFORMANCE

To understand model fidelity, multiple stratified analyses are performed. We look at the performance variation across land cover classes, soil texture types and climate zones. The Copernicus Global Land Cover map is used for the land cover class, SoilGrids for the soil texture type and the Köppen climate map for the climate zone.

To understand variation at a finer scale, quantitative analysis at an individual sensor level is also performed. This is done for each of the validation sensors present in the US (since a majority of the sensors are located in the US). We use the validation split of the Sentinel-1 anchored hard dataset for all these analyses.

2) TIME SERIES ANALYSIS

To understand temporal variations in the performance of our model, we look at the time series of model estimates vs in-situ labels for a few randomly selected validation sensors from the Sentinel-1 anchored hard dataset. This helps understand the temporal stability of model estimates and visualize the bias present (if any).

3) LARGE SCALE SPATIAL ESTIMATION

Lastly, large scale estimation is performed using our model where we move the model pixel by pixel and estimate the soil moisture at each location. This provides insight into spatially coherency of the model estimates and variation with respect to the inputs.

c. Model exploration studies

Ablation studies and sensitivity analyses are performed to see how important various model inputs/parameters are. All the studies here were performed on the validation set of the Sentinel-1 anchored hard dataset.

1) INPUT SIZE SENSITIVITY ANALYSIS

Each of the high-resolution sources correspond to 10x10m per pixel. The sensitivity study here uses the Sentinel-1 + Sentinel-2 + DEM model. We only use high resolution sources here since they are the most sensitive to a change in input size. This helps us understand the amount of context required by the model better and how it ties to performance.

2) INPUT FEATURE ABLATION STUDY

To identify which features are the most important for our models, we perform the following experiments. Starting with our model with all the input bands (Sentinel-1, DEM, Sentinel-2, SoilGrids and the coarse soil moisture inputs - SMAP-USDA, GLDAS) as the reference model, we remove the source for which we want to calculate the feature importance and then see how much of a drop in performance is observed.

d. Benchmarking

We pick some of the best existing global methods along with a few top regional models to compare and evaluate our model performance against. Information about these works is presented in Table 7.

Work	Temporal range of in-situ	Validation sensor networks	Validation information	Additional details
-------------	----------------------------------	-----------------------------------	-------------------------------	---------------------------

	data			
(Y. Kerr et al., 2016)	2010-2013	AMMA-CATCH, SCAN, SNOTEL, USCRN	All sensors used for validation.	SMOS validation. Only provide median correlation metrics instead of mean across sensors in each of their networks. Still a useful comparison though since the median is often better calibrated and robust to outliers.
(Bi et al., 2016)	2010-2012 for NAQU, 2008-2011 for MAQU	MAQU, NAQU	All sensors used for validation. In-situ data averaged in each LSM pixel.	Validation of 4 different Land Surface Models (LSMs) - VIC (Best in MAQU), Noah (GLDAS), Mosaic, CLM (Best in NAQU). We compare against the best for each network.
(Albergel et al., 2012)	2008-2010	AMMA-CATCH, SCAN, SMOSMANIA, REMEDHUS	All sensors used for validation.	ECMWF soil moisture validation. The model estimates soil moisture at 0-7 cm but validation labels come from 0-5cm depth. No ubRMSE metrics reported. However, the RMSE and bias are reported per sensor for some of the sensor networks, so we compute the ubRMSE as $\sqrt{\text{RMSE}^2 - \text{bias}^2}$ per sensor and then take an average to get the ubRMSE for the network.
(Wang et al., 2021)	2017-2018	FR_Aqui, RSMN, SCAN, SMOSMANIA, SNOTEL, USCRN, REMEDHUS	All sensors used for validation. Sensors within a single CCI/SMOPS pixel are averaged. All readings are averaged over 24 hrs.	They study two products CCI and SMOPS, CCI performs better overall so we compare against that. Only correlation and RMSE metrics present, since bias metrics aren't provided, we can't compute the ubRMSE.
(Fang et	2018-2019	SCAN, USCRN	All sensors used	They produce soil moisture

al., 2021)			for validation.	estimates at 400m resolution using a statistical approach.
(Balenzano et al., 2021)	2015-2018/2020 (for some sites)	LITTLEWASHITA, REMEDHUS, TXSON, YANCO	All sensors used for validation.	They obtain data from ISMN along with a couple of SMAP validation sites. They use temporal Sentinel-1, land cover and soil texture inputs in a physics based method.
(Dente et al., 2012)	2010	MAQU, TWENTE	All sensors used for validation. In-situ data averaged in each SMOS pixel.	Validation of SMOS retrievals in MAQU and Twente.
(Collow et al., 2012)	2010	FORTCOBB, LITTLEWASHITA	All sensors used for validation. In-situ data averaged in each SMOS pixel.	Validation of SMOS retrievals in FortCobb and LittleWashita.
(Jing et al., 2018)	2008-2011	KYEAMBA, YANCO	All sensors used for validation.	They validate ECMWF based multilayer temporal LSMs.
(Karthikeyan & Mishra, 2021)	2015-2019	FORTCOBB, LITTLEWASHITA, TXSON	Sensor level splits.	Build and use multiple region specific XGBoost models. Only provide median metrics for each of their networks.
(Abbaszadeh et al., 2019)	2015	SCAN, USCRN, LITTLEWASHITA	80:20 train/val split over entire data (not at a sensor level).	They use a set of 12 random forest models where each model is used for a different soil texture type.
(Vergopolan et al., 2021)	2015-2019	SCAN, SNOTEL, LITTLEWASHITA, FORTCOBB, TXSON	80:20 train/test split at a sensor level. They ensure that no training site is within 25km from a test site.	They use random forests based models. They use daily aggregates of in-situ data to evaluate against.

Table 7. Information on the works we compare against.

Not all methods in Table 7 use the same kind of train/validation/test splits and acquisition time ranges for in-situ data, so it's not possible to make rigorous comparisons. However, the sensor networks used are the same in all comparisons and sampling is always performed over a large time range which provides a meaningful comparison. A lot of works also perform K-fold cross validation while reporting their results, but we don't do that on our end due to practical constraints - deep learning models take a long time to train. We instead do a rigorous train/validation/test split, where we only looked at the numbers on the test set exactly once (at the time of writing this paper), and after all model parameters were finalized on the validation split.

Note that in order to benchmark our results, we use the Sentinel-1 anchored full dataset since the Sentinel-1 anchored hard dataset contains only a small fraction of sensor networks that we can benchmark against. Additionally, most of the works we compare against use IID splits and we do the same to remain consistent while comparing our results. In addition to benchmarking, we also provide overall model performance results on the Sentinel-1 anchored full dataset here for completeness.

5. Results

a. Overall model performance

	Test		
Experiment	ubRMSE	RMSE	Correlation
Baselines (Low-resolution)			
SMAP-USDA	0.096	0.133	0.626
GLDAS	0.064	0.099	0.597
SMAP-USDA + GLDAS NN	0.059	0.091	0.671
Ours (High-resolution)			
Sentinel-1 + DEM	0.078 (-32.2%)	0.116 (-27.5%)	0.3 (-55.3%)

Sentinel-1 + DEM + Sentinel-2	0.071 (-20.3%)	0.099 (-8.8%)	0.481 (-28.3%)
Sentinel-1 + DEM + Sentinel-2 + SoilGrids	0.062 (-5.1%)	0.096 (-5.5%)	0.566 (-15.6%)
Sentinel-1 + DEM + Sentinel-2 + SoilGrids + SMAP-USDA + GLDAS	0.055 (+6.8%)	0.089 (+2.2%)	0.707 (+5.4%)

Table 8. Test set results on the Sentinel-1 anchored hard dataset. All results presented are in volumetric m^3/m^3 units. The number in brackets denotes the percentage change with respect to the SMAP-USDA + GLDAS NN baseline. A +ve change in percentage denotes an increase in Pearson correlation and decrease in ubRMSE/RMSE. The test split consists of 9,750 data points.

A quantitative evaluation of our models is presented in Table 8 along with comparisons to baselines (validation set results are present in Table A2). Our model performs better than the baseline SMAP-USDA + GLDAS NN in all the metrics. These results together show that the high resolution sources/geophysical variables and coarse soil moisture sources provide complementary information and combining them gives us the best performance.

b. Spatial and Temporal Analysis

1) SPATIAL STRATIFICATION OF MODEL PERFORMANCE

	Validation		
Land cover	ubRMSE	RMSE	Correlation
Shrubs*	-	-	-
Vegetation	0.053	0.09	0.656
Agriculture	0.046	0.073	0.67
Built-up*	-	-	-
Bareground*	0.026	0.028	0.764
Snow/Ice*	-	-	-
Water*	-	-	-
Wetland*	-	-	-

Moss/Lichen*	-	-	-
Closed forest	0.058	0.082	0.76
Open forest	0.062	0.089	0.718

Table 9. Validation results on the Sentinel-1 anchored hard dataset for our model stratified by land cover type. All results presented are in volumetric m^3/m^3 units. An asterisk (*) indicates that there are < 5 sensors available for the specific land cover class in the test data.

	Validation		
Soil texture type	ubRMSE	RMSE	Correlation
Sand*	0.023	0.068	0.25
Loamy sand*	0.123	0.124	0.73
Sandy loam	0.042	0.066	0.7
Loam	0.053	0.087	0.72
Silt loam	0.058	0.087	0.81
Silt*	-	-	-
Sandy clay loam*	-	-	-
Clay loam	0.053	0.083	0.67
Silty clay loam	0.055	0.072	0.52
Sandy clay*			-
Silty clay*	-	-	-
Clay*	0.032	0.033	0.54

Table 10. Validation results on the Sentinel-1 anchored hard dataset for our model stratified by soil texture type. All results presented are in volumetric m^3/m^3 units. An asterisk (*) indicates that there are < 5 sensors available for the specific soil texture class in the test data.

	Validation

Climatic zone	ubRMSE	RMSE	Correlation
Tropical rainforest (Af)*	-	-	-
Tropical monsoon (Am)*	-	-	-
Tropical wet and dry (Aw)*	-	-	-
Semi arid (BS)	0.049	0.074	0.61
Dry desert (BW)	0.034	0.046	0.744
Humid subtropical (Cf)	0.056	0.084	0.674
Summer Mediterranean (Cs)	0.054	0.088	0.76
Humid continental (Df)	0.055	0.087	0.72
Mediterranean-influenced (Ds)	0.058	0.069	0.817
Monsoon influenced (Dw)*	-	-	-
Tundra (ET)*	-	-	-

Table 11. Validation results on the Sentinel-1 anchored hard dataset for our model stratified by the climatic zone. All results presented are in volumetric m^3/m^3 units. An asterisk (*) indicates that there are < 5 sensors available for the specific climate zone in the validation data.

Our model performs well across different kinds of land cover classes as shown in Table 9, soil texture types as shown in Table 10 and climatic zones as shown in Table 11. Metrics on classes containing less than 5 sensors should be disregarded since the sample size is too small to get a reliable aggregate. In all of these stratifications, the model performs fairly consistently (correlation within ± 0.2) across all the classes showing adaptability and robustness of the model.

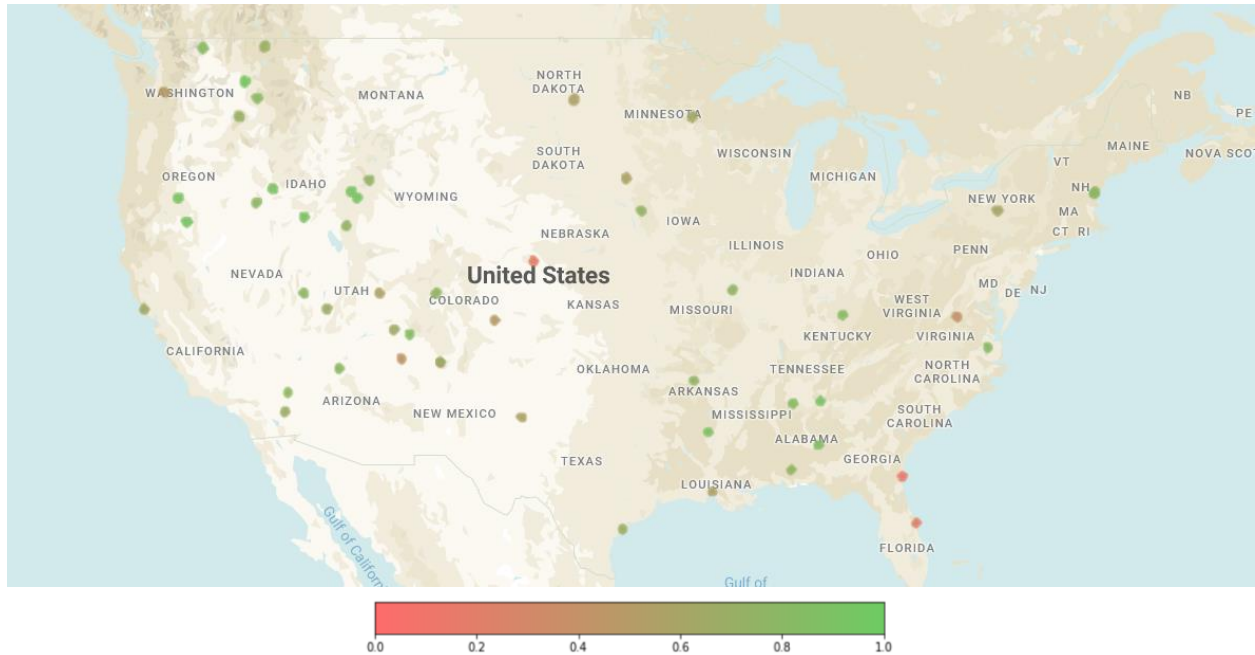


Fig. 11. Correlation between model estimates and ground truth labels at each validation sensor from the Sentinel-1 anchored hard dataset located in the US. Red-Green circles indicate in-situ sensor locations at which we report the correlation.

Figure 11 shows the correlation at a sensor level for each of our validation sensors across the US. Sensors along the coastlines show a slight drop in performance overall. Similar results for ubRMSE are provided in Figure A16.

2) TIME SERIES ANALYSIS

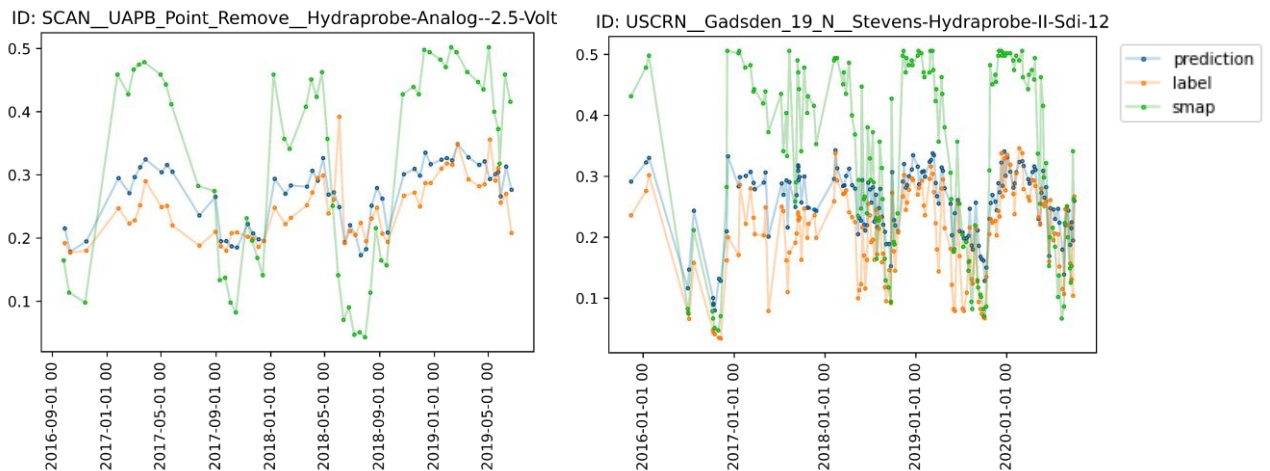


Fig. 12. Sample time-series plots showing how the estimates of the model m^3/m^3 vary over time at a few randomly picked validation sensors from the Sentinel-1 anchored hard dataset compared to the labels/SMAP estimates. The sensor corresponding to the figure on the left has the following properties: land cover - agriculture, soil type - silt loam, climate zone - Cfa. The

sensor on the right has the same properties as the one on the left except the soil type which is loam.

Figure 12 shows time series plots for a few randomly selected sensors. In a majority of cases, our model estimates follow the variation in the sensors well. They often have a small amount of bias based on the location but the overall trend is captured well by the model. This reflects what we see in our quantitative metrics - correlation and ubRMSE. Additional time series plots are presented in Figure A14.

3) LARGE SCALE SPATIAL ESTIMATION

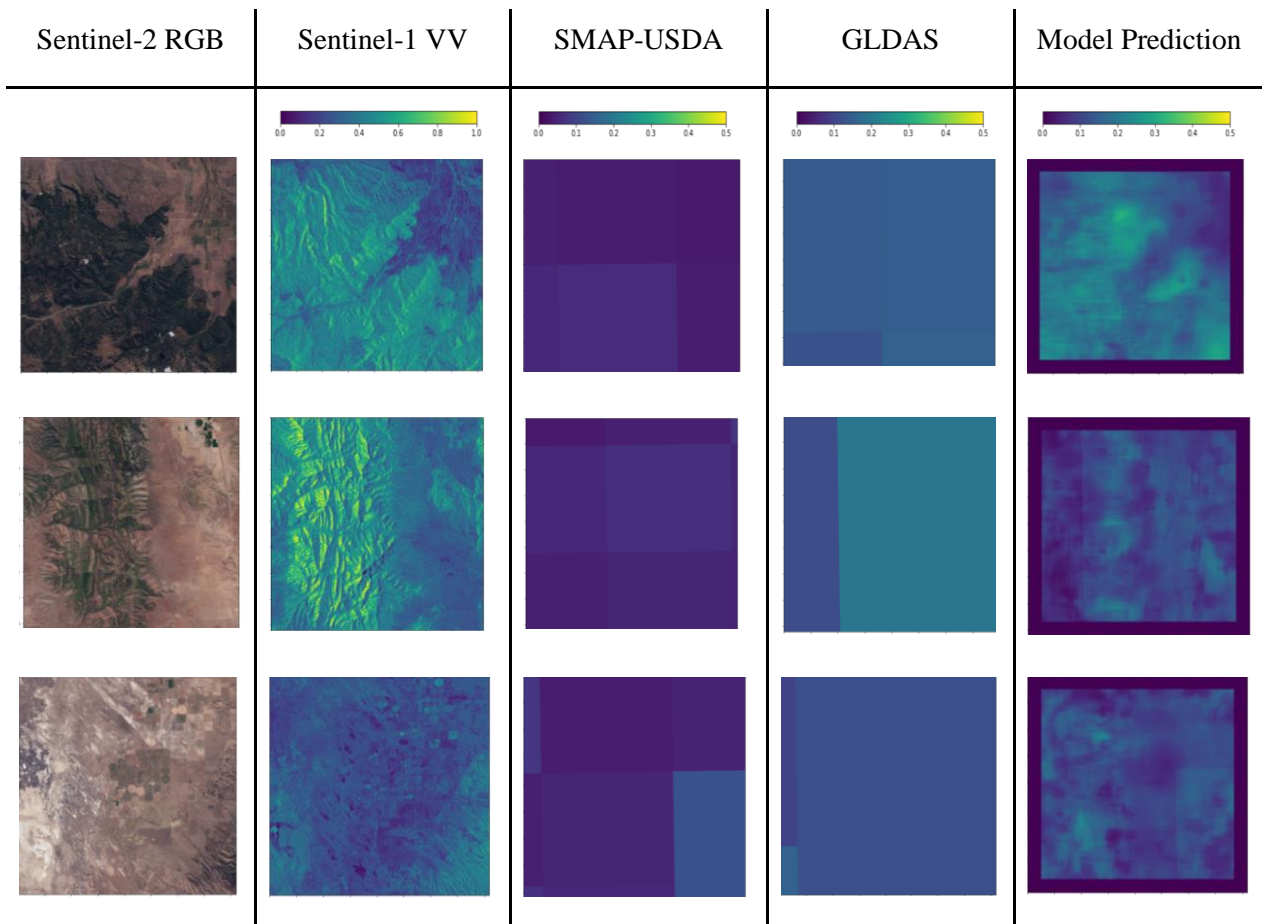


Fig. 13. Sample volumetric soil moisture m^3/m^3 prediction results with our model on $\sim 20 \times 20$ km regions selected at random. The colorbar shows the normalized input/output range for each source/prediction.

In Figure 13, we notice a fair amount of variation at a sub-SMAP-USDA pixel (10×10 km) level as well. The local variation in the estimates can be explained to a fair extent by looking at

the input imagery. These results along with the metrics we observe in Table 8 show that the model produces accurate high resolution estimates of soil moisture.

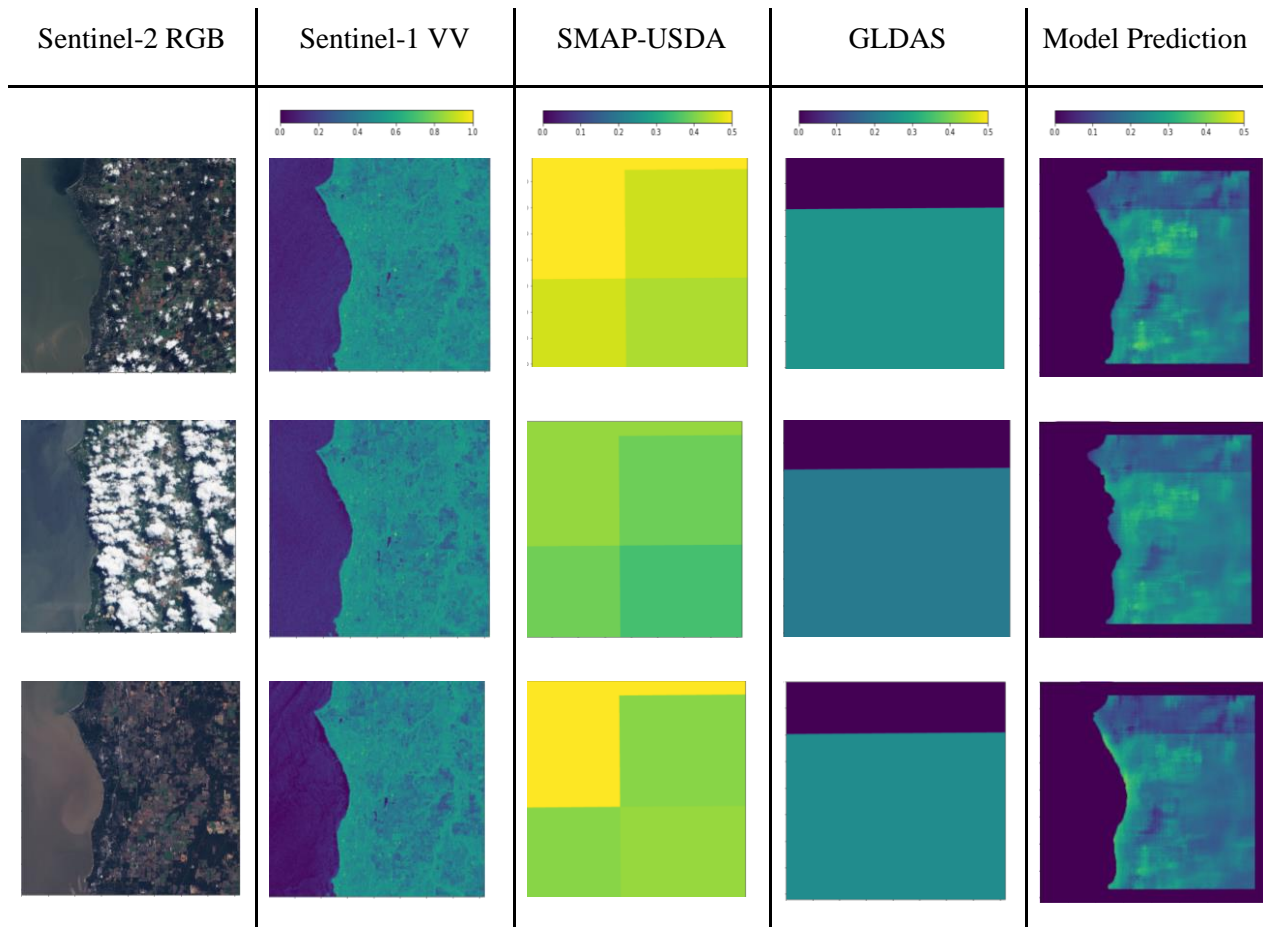


Fig. 14. Model volumetric soil moisture m^3/m^3 prediction results across time (each row separated by a year) on a single $\sim 20 \times 20 \text{ km}$ region selected at random.

In Figure 14, we notice that although the actual estimated values of soil moisture change, the patterns visible remain fairly consistent which shows that the model is sensitive to spatial changes. At times however, tiling artifacts at SMAP-USDA/GLDAS boundaries appear when there is a large variation in adjacent SMAP-USDA/GLDAS tiles.

c. Model exploration studies

1) INPUT SIZE SENSITIVITY ANALYSIS

	Validation		
Input size	ubRMSE	RMSE	Correlation
64x64	0.068	0.097	0.427
128x128	0.066	0.092	0.464
256x256	0.068	0.093	0.496
512x512	0.066	0.089	0.503

Table 12. Validation results on the Sentinel-1 anchored hard dataset for the input size sensitivity analysis. All results presented are in volumetric m^3/m^3 units.

Looking at the correlation, larger input sizes provide better results since the model has more spatial context and a larger number of model parameters to work with leading to greater representation power, however the improvements decline as we continue to scale up. ubRMSE however, remains fairly consistent throughout. Considering practical constraints, we went ahead with 256x256 for the input size.

2) INPUT FEATURE ABLATION STUDY

	Validation		
Feature/Source removed	ubRMSE	RMSE	Correlation
None	0.053	0.081	0.701
DEM	0.054 (-1.9%)	0.083 (-2.5%)	0.688 (-1.9%)
Sentinel-1	0.055 (-3.8%)	0.093 (-14.8%)	0.671 (-4.3%)
Sentinel-2	0.055 (-3.8%)	0.083 (-2.5%)	0.69 (-1.6%)
SoilGrids	0.058 (-9.4%)	0.086 (-6.2%)	0.636 (-9.3%)
SMAP-USDA	0.056 (-5.7%)	0.085 (-4.9%)	0.65 (-7.3%)
GLDAS	0.053 (+0%)	0.095 (-17.3%)	0.681 (-2.9%)
Sentinel-1 + Sentinel-2 (High resolution inputs)	0.057 (-7.5%)	0.093 (-14.8%)	0.646 (-7.8%)

SoilGrids + DEM (Static inputs)	0.058 (-9.4%)	0.083 (-2.5%)	0.634 (-9.6%)
SMAP-USDA + GLDAS (SM estimates)	0.059 (-11.3%)	0.088 (-8.6%)	0.614 (-12.4%)

Table 13. Validation results on the Sentinel-1 anchored hard dataset for the feature importance ablation study. Shows the importance of each feature set/source. All results presented are in volumetric m^3/m^3 units. The number in brackets denotes the percentage change with respect to the first row. A +ve change in percentage denotes an increase in Pearson correlation and decrease in ubRMSE/RMSE

SoilGrids is the most important source (largest drop from the baseline in terms of performance, looking at the correlation) for the model followed by SMAP-USDA and Sentinel-1 (the rest following afterwards) in Table 13. We notice that SMAP-USDA has a higher importance than GLDAS - this could be because GLDAS has a coarser resolution and a different sensing depth of 10cm instead of 5cm.

Some of these inputs are related and in order to understand their importance together, we drop features in groups. Dropping SMAP-USDA + GLDAS (SM estimates) together results in the largest drop in performance, followed by SoilGrids + DEM (static inputs) and finally Sentinel-1 + Sentinel-2 (High resolution inputs). All three sets of inputs result in a significant drop in metrics which shows that the model uses all these different types of sources in order to estimate soil moisture accurately.

d. Benchmarking

	Correlation														
Network	Ours	Kerr et al ³	Bi et al ^{4*}	Albergel et al ^{5*}	Wang et al ⁶	Fang et al ⁷	Balenzano et al ⁸	Dent et	Collozet	Jing et	Karthikeyan et	Abbaszadeh et	Vergopolan et	SMAP-USDA	GLDAS

³ (Y. Kerr et al., 2016) (SMOS)

⁴ (Bi et al., 2016)

⁵ (Albergel et al., 2012) (ECMWF)

⁶ (Wang et al., 2021)

⁷ (Fang et al., 2021)

⁸ (Balenzano et al., 2021)

								al ^{9*}	al ^{10*}	al ^{11*}	al ^{12**}	al ^{13**} *	al ¹⁴		
AMMA-CATCH	0.93	0.77		0.53										0.886	0.889
BIEBRZA_S-1	0.847													0.574	0.434
FR_Aqui	0.65							0.78						0.595	0.638
MAQU	0.504		0.623					0.63						0.453	0.542
RSMN	0.654				0.62									0.555	0.606
SCAN	0.695	0.6		0.63	0.55	0.65 6						0.65	0.66	0.618	0.612
SMOSMANIA	0.881			0.8	0.67									0.839	0.872
SNOTEL	0.735	0.48			0.48								0.58	0.586	0.477
USCRN	0.735	0.7			0.58	0.68 8						0.7		0.674	0.599
KYEAMBA	0.731									0.84				0.727	0.768
LITTLEWASHIT A	0.754						0.28		0.76		0.77	0.82	0.74	0.665	0.626
REMEDHUS	0.625			0.77	0.65		0.2							0.608	0.598
TWENTE	0.749							0.68						0.712	0.554
TXSON	0.812						0.33				0.83		0.82	0.722	0.644
YANCO	0.427						0.62				0.84			0.388	0.473

Table 14. A comparison of Pearson correlation (for our best model) vs other works spread across stratified by the sensor network. Bold metrics on each row indicate the best performing model for the network in question. All results presented are in volumetric m^3/m^3 units. An asterisk (*) indicates that the time range during which in-situ data was obtained is different from

⁹ (Dente et al., 2012)

¹⁰ (Collow et al., 2012)

¹¹ (Jing et al., 2018)

¹² (Karthikeyan & Mishra, 2021)

¹³ (Abbaszadeh et al., 2019)

¹⁴ (Vergopolan et al., 2021)

the one we use for training our models. A double asterisk (**) indicates that the work uses region-specific models. A triple asterisk (***) indicates that the work does not ensure a sensor level train/validation/test split (data points from the same sensor can be present in train and validation).

Network	ubRMSE										GLDAS
	Ours	Bi et al ^{15*}	Albergel et al ^{16*}	Fang et al ¹⁷	Balenzano et al ¹⁸	Collow et al ^{19*}	Jing et al ^{20*}	Karthikeyan et al ^{21**}	Abbaszadeh et al ^{22***}	SMAP-USDA	
AMMA-CATCH	0.022		0.036							0.078	0.036
BIEBRZA_S-1	0.077									0.13	0.132
FR_Aqui	0.06									0.132	0.061
MAQU	0.067	0.046								0.067	0.063
RSMN	0.036									0.117	0.049
SCAN	0.06			0.068					0.047	0.098	0.066
SMOSMANIA	0.045		0.053							0.111	0.053
SNOTEL	0.063									0.109	0.085
USCRN	0.04			0.08					0.04	0.086	0.055
KYEAMBA	0.068						0.06			0.104	0.062
LITTLEWASHITA	0.037				0.095	0.05		0.041	0.035	0.105	0.046
REMEDHUS	0.049				0.093					0.09	0.054
TWENTE	0.106									0.121	0.135

¹⁵ (Bi et al., 2016)

¹⁶ (Albergel et al., 2012) (ECMWF)

¹⁷ (Fang et al., 2021)

¹⁸ (Balenzano et al., 2021)

¹⁹ (Collow et al., 2012)

²⁰ (Jing et al., 2018)

²¹ (Karthikeyan & Mishra, 2021)

²² (Abbaszadeh et al., 2019)

TXSON	0.034				0.085			0.039		0.093	0.045
YANCO	0.094				0.08		0.05			0.159	0.084

Table 15. A comparison of ubRMSE (for our best model) vs other works stratified by the sensor network. Bold metrics on each row indicate the best performing model for the network in question. All results presented are in volumetric m^3/m^3 units. An asterisk (*) indicates that the time range during which in-situ data was obtained is different from the one we use for training our models. A double asterisk (**) indicates that the work uses region-specific models. A triple asterisk (***) indicates that the work does not ensure a sensor level train/validation/test split (data points from the same sensor can be present in train and validation)

Benchmark comparisons for correlation and ubRMSE in Table 14 and Table 15 respectively show that our model performs well across various sensor networks. Even in cases where existing benchmark models perform better, our model trails closely behind. Our model however, performs significantly weaker than most benchmarks on MAQU and RSMN networks in terms of correlation and YANCO in terms of ubRMSE.

Do note that these are not an exact 1:1 set of comparisons since the time ranges, sampling strategies etc differ for some benchmarks and some of these methods perform an aggregation of in-situ sensor readings falling within a single pixel of their model inputs or an aggregation of in-situ sensor readings across a period of time which simplifies the task to some extent.

To provide a comprehensive set of results to compare against other works, test set results on the Sentinel-1 anchored full dataset are presented in Table 16. These results can be used to compare against existing works that use IID based splits.

Experiment	Test		
	ubRMSE	RMSE	Correlation
Baselines (Low-resolution)			
SMAP-USDA	0.1	0.135	0.634
GLDAS	0.07	0.11	0.579
SMAP-USDA + GLDAS NN	0.063	0.098	0.666

Ours (High-resolution)			
Sentinel-1 + DEM	0.076 (-20.6%)	0.109 (-11.2%)	0.432 (-35.1%)
Sentinel-1 + DEM + Sentinel-2	0.069 (-9.5%)	0.097 (+1%)	0.57 (-14.4%)
Sentinel-1 + DEM + Sentinel-2 + SoilGrids	0.061 (+3.2%)	0.095 (+3.1%)	0.647 (-2.9%)
Sentinel-1 + DEM + Sentinel-2 + SoilGrids + SMAP-USDA + GLDAS	0.056 (+11.1%)	0.087 (+11.2%)	0.729 (+9.5%)

Table 16. Test set results on the Sentinel-1 anchored full dataset. All results presented are in volumetric m^3/m^3 units. The number in brackets denotes the percentage change with respect to the SMAP-USDA + GLDAS NN baseline. A +ve change in percentage denotes an increase in Pearson correlation and decrease in ubRMSE/RMSE. The test split consists of 20,797 data points.

6. Discussion

We developed machine learning based models that fuse information from different remote sensing, geophysical and meteorological data sources at varying resolution to produce soil moisture estimates at a nominal resolution of 320m. The result is a trained model applicable in a large variety of settings across the world, and outperforms SMAP-USDA/GLDAS baselines and most of the other methods that we compare against in most of the sensor networks. We perform various input sensitivity and ablation studies which provide useful insight into each of the remote sensing sources used in an empirical setting. We haven't used additional DEM based topographic position indices or wetness indices as model inputs currently and plan to add them as inputs in the future.. We have also tried using vegetation indices such as NDVI and leaf area indices as inputs to the model but didn't see any improvement (shown in Table A3). To understand possible biases that arise during cloud filtering, we perform a few analyses in Figures A12 and A13. We observed that cloud filtering does lead to a slight temporal bias (some months lose a higher fraction of data compared to other months) as well as a bias towards lower soil moisture values. Although this is not very significant, we wanted to ensure that we are aware of this.

In order to further understand model generalizability, we looked at how model performance varies across validation sensors as we move farther away from the train set of sensors in Figure

15. Specifically, for each validation sensor, we computed the distance from the nearest train sensor and plot the aggregate correlation across all validation sensors greater than a certain distance x from the nearest train sensor, for all x . This gave us insight into how model performance varies as validation sensors move further away from train sensors. In general we expect this graph to trend downwards because we expect more out of domain examples to show up. Models that pick up relevant and general features will expect a smaller drop compared to models that pick up location specific features. While we expect this graph to trend downward, it doesn't always strictly do so because the geography changes as distance changes, and it's possible we run into easy/hard regions as we vary this threshold. We report results on both the "hard" and "full" splits to understand this generalizability behaviour better, and provide results for the full range of distance thresholds in Figure 15 for the validation set. The results also highlight the importance of ensuring that data splits are strong to validate performance and direct IID splits might not be representative of true performance when models are deployed in the real world.

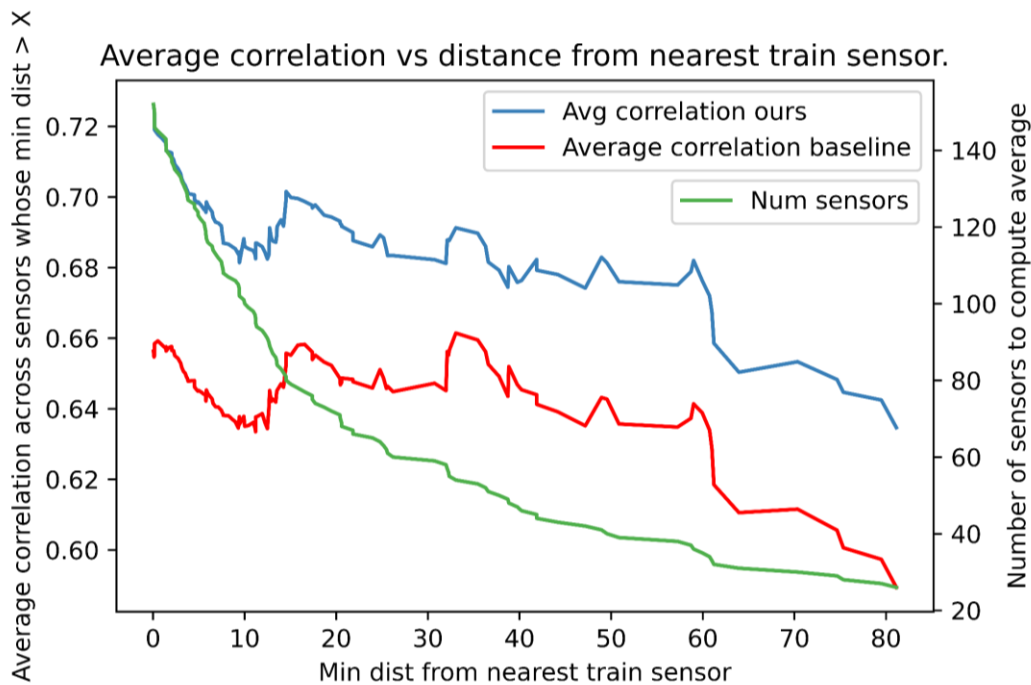


Fig. 15. Aggregate correlation of validation sensors vs distance from the nearest train sensor. The x-axis is in units of distance, the y-axis (on the left) is the metric value (m^3/m^3) and the y-axis (on the right) shows the number of validation sensors used at any given point to compute the metric average. Each point in the plot represents the average metric value for all validation sensors whose nearest train sensor is $\geq x$ distance away.

Our models are able to capture soil moisture well even in cases where there is high input variability in terms of the terrain and topography as shown in Figure A11. They are not restricted to being applied on homogenous regions only and can be applied in a wide range of settings. This is important since most agricultural farmlands often have a high amount of variability in a given spatial context (Addis et al., 2015). We do notice some unexpected model behaviour in the last row of Figure 13 where irrigated/farmland regions (towards the center of the image) have lower soil moisture than the neighbouring regions. Soilgrids data for the region showed that the sand fraction and bulk density for the farmland regions seem to be higher than surrounding areas which could be the reason for a lower SM estimate in these regions.

An important learning from this study is that tasks in hydrology such as soil moisture estimation can benefit from modern advances in deep learning. Through this study, we observed that a larger context around the location of interest can be leveraged by deep learning models to use their scene understanding capability, and learn location agnostic models (this is seen from the significant improvement in performance as the amount of context is increased). Image based CNNs, while being more computationally expensive than traditional ML techniques, are still fairly efficient on images and allow us to utilize additional input context with ease. Traditional methods like random forests etc require significant feature engineering (especially with image inputs) and don't match deep learning methods from ML literature. Our models also make it very simple to experiment with additional input sources and fuse them together in a multi stage manner which has shown large performance improvements in our experiments.

Lastly, we also observed that different remote sensing/geophysical sources are complementary and the combination of them leads to better performance. Looking forward, we would like to add additional inputs such as thermal bands, precipitation time series and extend our model capacity to process time series inputs. This would enable the model to leverage auto-correlations in time, and enable forecasting of soil moisture. We are also interested in applying more advanced self/semi supervised methods (Patel et al., 2021) to perform the task at hand to improve model generalization to unseen regions and scale better.

We release our models to be used by anyone interested. Additionally, we curate and release a large scale soil moisture dataset that can be used by others to train and evaluate remote sensing based soil moisture models with ease.

7. Acknowledgements

We would like to thank Prof Muddu Sekhar, Christopher H Van Arsdale, Kevin James McCloskey and Rob von Behren for helping us review the paper in whole and Nikhilesh Kumar for providing invaluable suggestions throughout the experimental process.

a. Data acknowledgements

- 1) The USDA Agricultural Research Service, Grazinglands Research Laboratory, El Reno, Oklahoma provided FORTCOBB and LITTLEWASHITA data.
- 2) The OzNet hydrological monitoring network provided YANCO and KYEAMBA data.
- 3) The University of Twente via DANS provided TWENTE data.
- 4) The University of Texas, Austin provided TXSON data via the Texas Data Repository.

8. Data Availability Statement

We release the dataset we created and used in the paper along with our trained models at https://github.com/google-research/google-research/tree/master/soil_moisture_retrieval.

APPENDIX

Sentinel-2 normalization scheme

We use the following non-linear normalization scheme where normalization is applied in the log space which ensures that we have a broader dynamic range for non-cloudy imagery post normalization.

Formula:

$$\text{scaled_band} = e^{\left(\frac{(\log(\text{original_band} \cdot 0.005 + 1.0) - \log_mean)}{\log_std} \right)} \cdot 5.0 - 1.0$$

$$\text{normalized_band} = \left(\frac{\text{scaled_band}}{\text{scaled_band} + 1.0} \right) * 255.0$$

Where the log_mean and log_std for each of the bands are as specified in Table A1. These were obtained by computing statistics over the United States.

Source	Band	log_mean	log_std
Sentinel-2	B2	1.7417268007636313	2.023298706048351
	B3	1.7261204997060209	2.038905204308012
	B4	1.6798346251414997	2.179592821212937
	B5	1.7734969472909623	2.2890068333026603
	B6	2.289154079164943	2.6171674549378166
	B7	2.382939712192371	2.773418590375327
	B8	2.3828939530384052	2.7578332604178284
	B11	2.1952484264967844	2.789092484314204
	B12	1.554812948247501	2.4140534947492487

Table A1. Normalization statistics for Sentinel-2 bands.

Sentinel-1 anchored dataset validation/test statistics

We provided statistics for the Sentinel-1 anchored full dataset in Section 2c2. But to ensure that our validation/test splits are truly IID, we compute the same set of statistics on each of the splits from the full dataset.

1) VALIDATION SPLIT

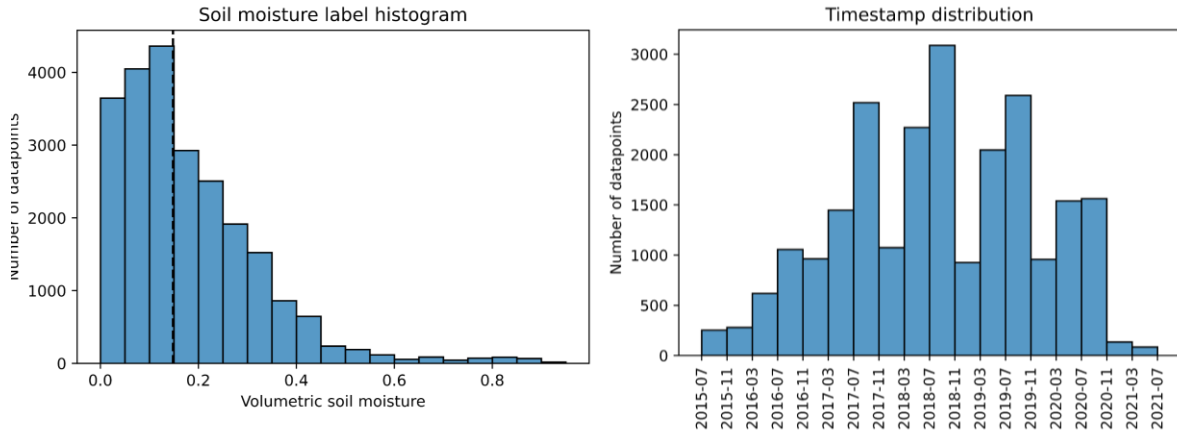


Fig. A1. *Left*: Volumetric soil moisture m^3/m^3 label distribution. The dashed line indicates the median soil moisture. *Right*: Timestamp distribution on the validation split of the dataset.

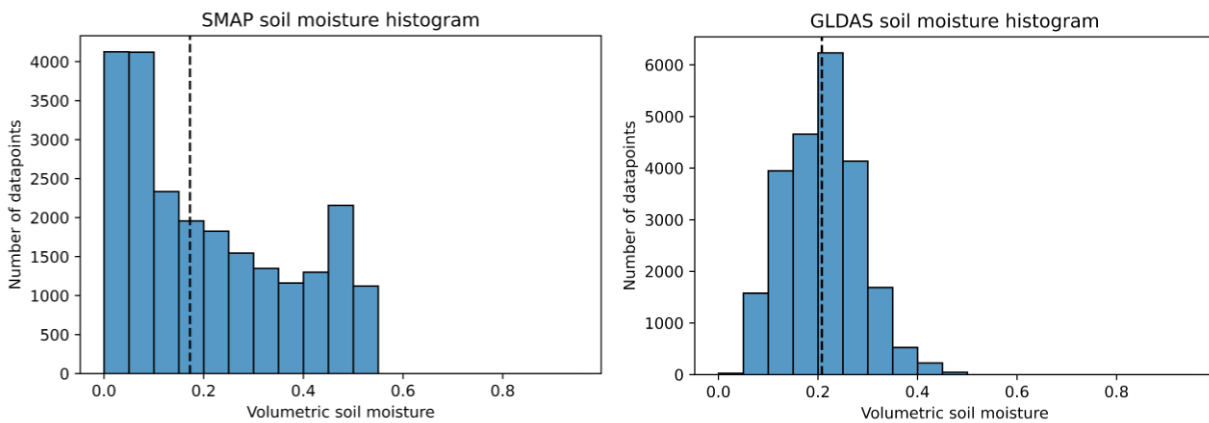


Fig. A2. Distribution of coarse soil moisture products. *Left*: SMAP-USA volumetric soil moisture m^3/m^3 . *Right*: GLDAS volumetric soil moisture m^3/m^3 on the validation split of the dataset. The dashed lines indicate the corresponding median soil moisture values.



Fig. A3. A heatmap of in-situ sensor locations present in the validation split of the dataset.

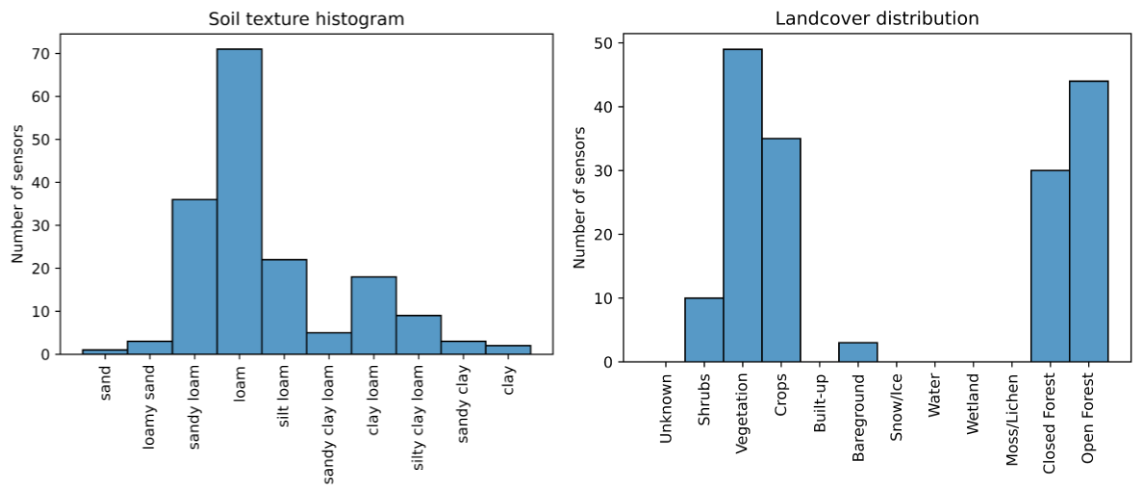


Fig. A4. *Left*: USDA based soil texture distribution. *Right*: Land cover distribution derived from the Copernicus Land Cover Map on the validation split of the dataset.

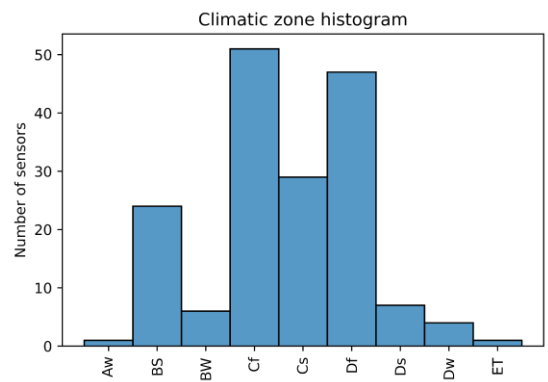


Fig. A5. Köppen climate zone distribution on the validation split of the dataset. We use the first two identifiers of the Köppen classification only in order to group similar climate zones together.

2) TEST SPLIT

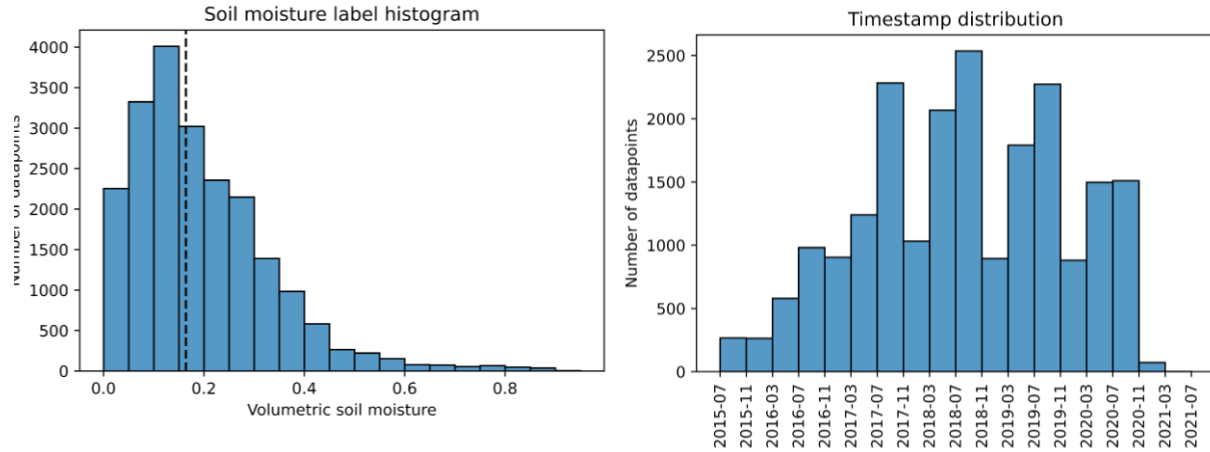


Fig. A6. *Left*: Volumetric soil moisture m^3/m^3 label distribution. The dashed line indicates the median soil moisture. *Right*: Timestamp distribution on the test split of the dataset.

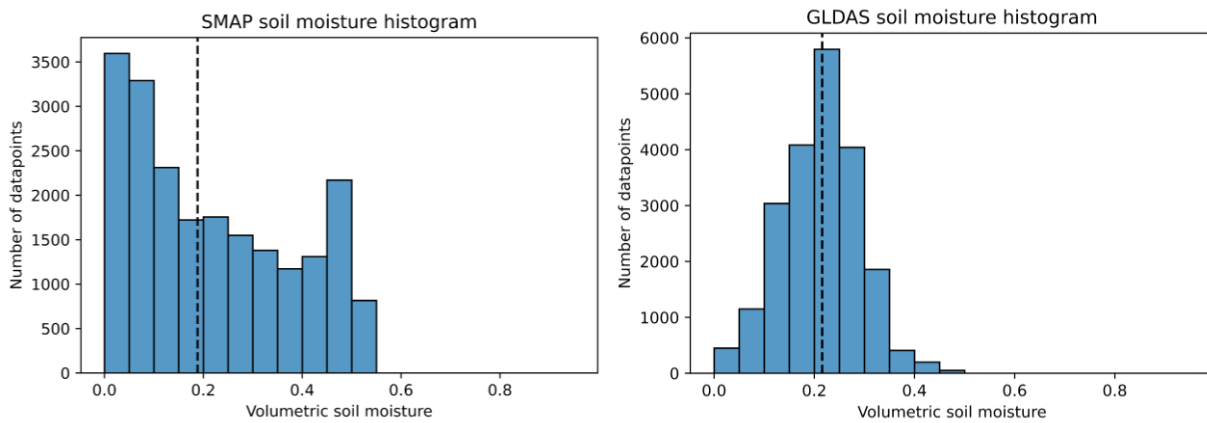


Fig. A7. Distribution of coarse soil moisture products. *Left*: SMAP volumetric soil moisture m^3/m^3 . *Right*: GLDAS volumetric soil moisture m^3/m^3 on the test split of the dataset. The dashed lines indicate the corresponding median soil moisture values.

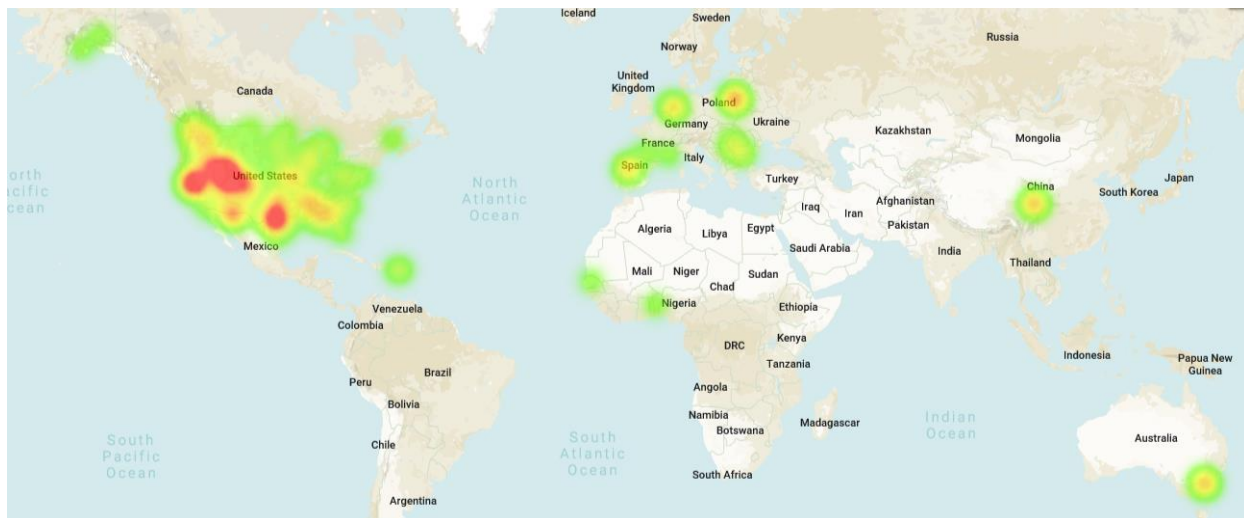


Fig. A8. A heatmap of in-situ sensor locations present in the test split of the dataset.

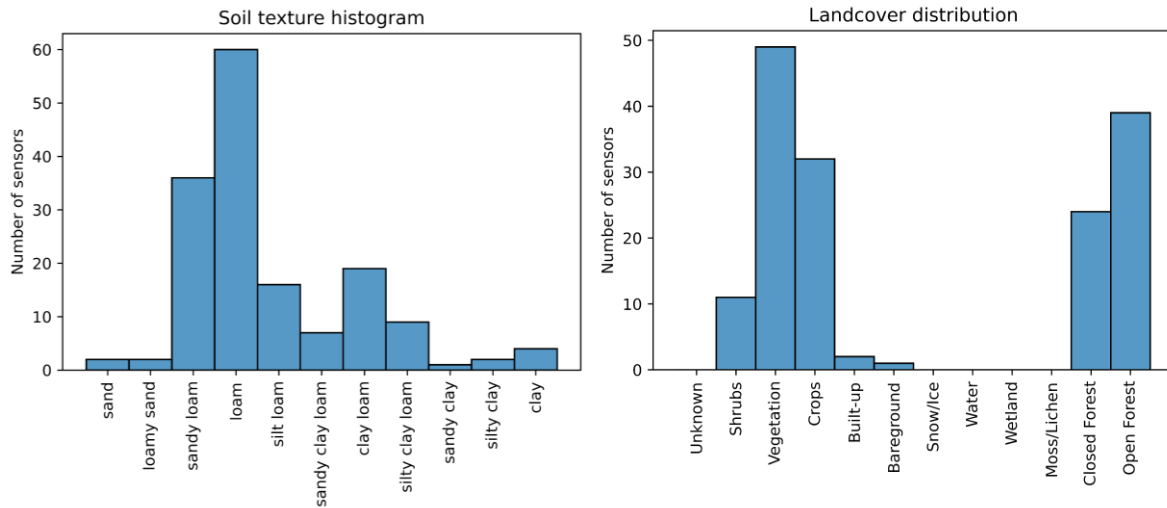


Fig. A9. *Left*: USDA based soil texture distribution. *Right*: Land cover distribution derived from the Copernicus Land Cover Map on the test split of the dataset.

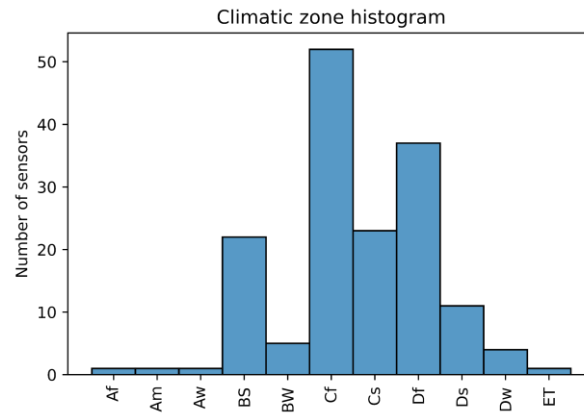


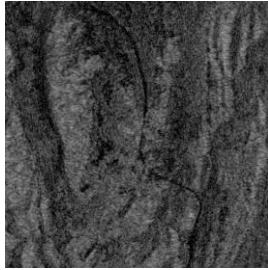
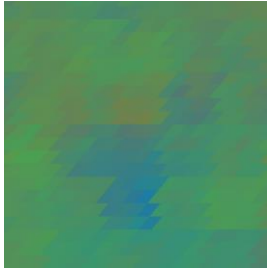

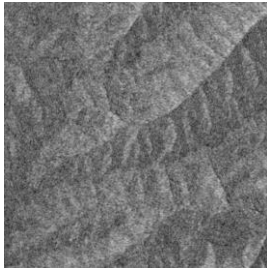
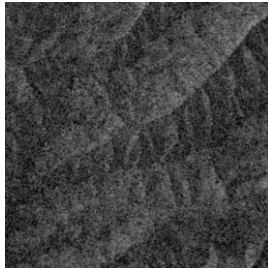


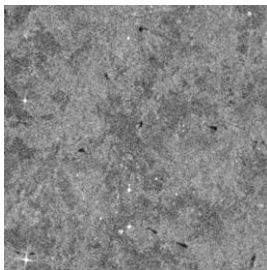
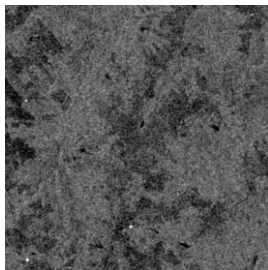

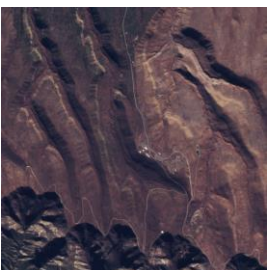
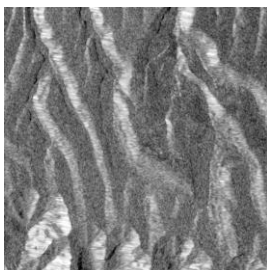
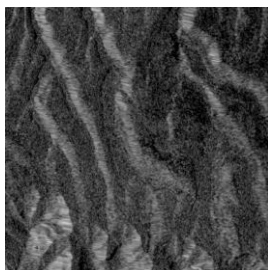

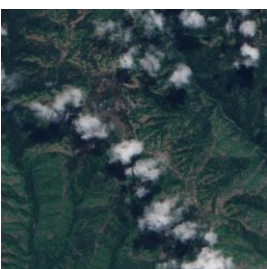
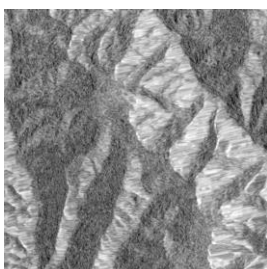
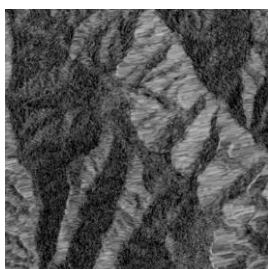



Fig. A10. Köppen climate zone distribution on the test split of the dataset. We use the first two identifiers of the Köppen classification only in order to group similar climate zones together.

From the plots above in Figures A1, A2, A3, A4, A5 and Figures A6, A7, A8, A9, A10, we observe that the distribution of data in our validation and test is similar to our train distribution. This validates that our splits are indeed IID.

Samples from the Sentinel-1 anchored full dataset

Sentinel-2 RGB	Sentinel-1 VV	Sentinel-1 VH	SoilGrids	Soil moisture
----------------	---------------	---------------	-----------	---------------

				label
				0.264
				0.244
				0.203
				0.111
				0.141

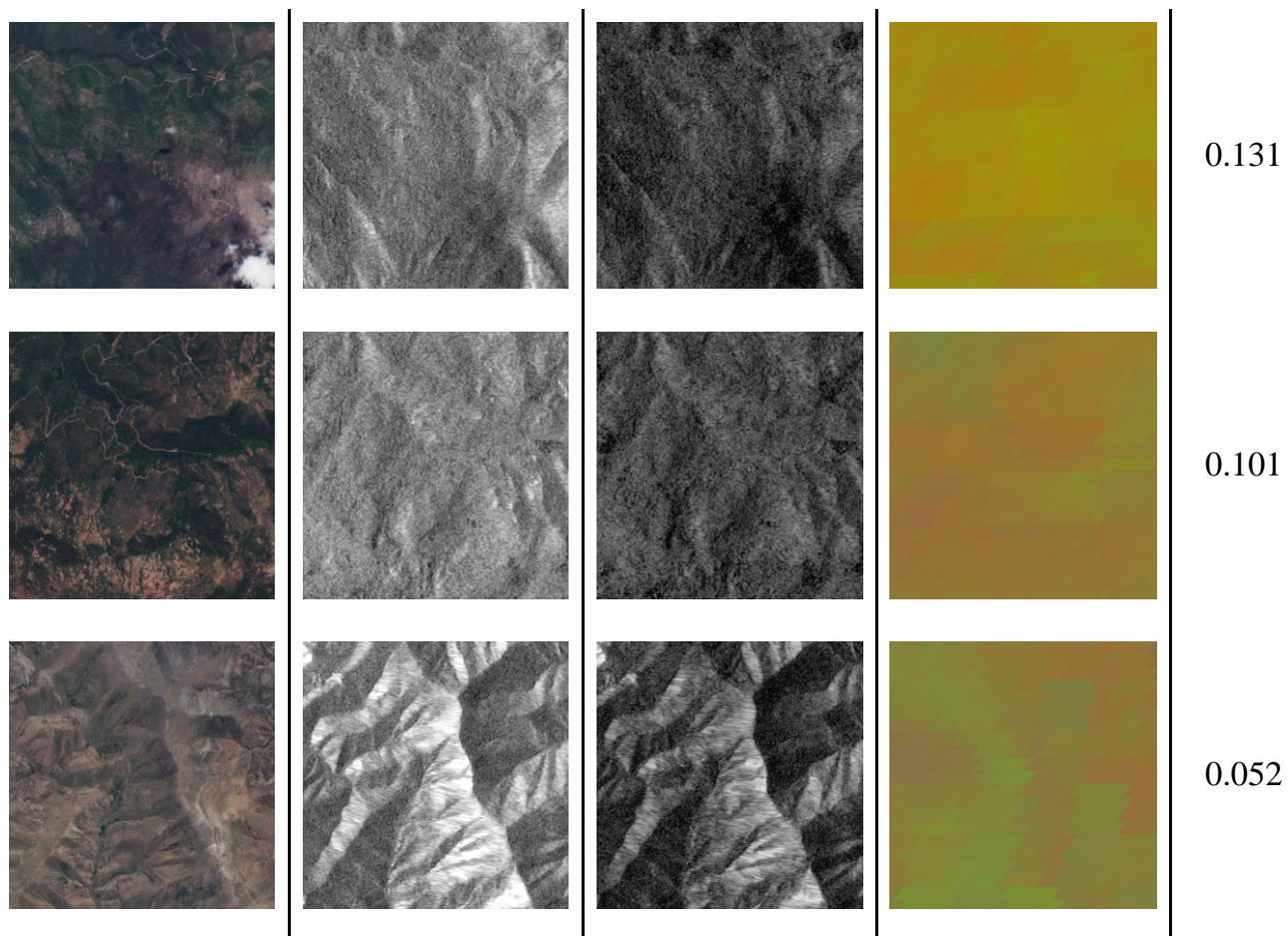


Fig. A11. Randomly selected inputs and labels from the dataset. The SoilGrids image is a false RGB image generated from the [sand, silt, clay] bands for R,G,B respectively. The labels are in volumetric m^3/m^3 units.

Overall model performance on the validation set

Experiment	Validation		
	ubRMSE	RMSE	Correlation
Baselines (Low-resolution)			
SMAP-USDA	0.097	0.146	0.612
GLDAS	0.064	0.098	0.569
SMAP-USDA + GLDAS NN	0.057	0.084	0.645
Ours (High-resolution)			

Sentinel-1 + DEM	0.07	0.095	0.414
Sentinel-1 + DEM + Sentinel-2	0.068	0.093	0.496
Sentinel-1 + DEM + Sentinel-2 + SoilGrids	0.059	0.088	0.614
Sentinel-1 + DEM + Sentinel-2 + SoilGrids + SMAP-USDA + GLDAS	0.053	0.081	0.701

Table A2. Validation results for our models on the Sentinel-1 anchored hard dataset. The number in brackets denotes the percentage change with respect to the SMAP-USDA + GLDAS NN baseline. All results presented are in volumetric m^3/m^3 units. A +ve change in percentage denotes an increase in Pearson correlation and decrease in ubRMSE/RMSE. The validation split consists of 8,794 data points.

Experiment	Validation		
	ubRMSE	RMSE	Correlation
Baselines (Low-resolution)			
SMAP-USDA	0.098	0.144	0.634
GLDAS	0.07	0.114	0.568
SMAP-USDA + GLDAS NN	0.062	0.099	0.658
Ours (High-resolution)			
Sentinel-1 + DEM	0.073	0.099	0.474
Sentinel-1 + DEM + Sentinel-2	0.069	0.094	0.587
Sentinel-1 + DEM + Sentinel-2 + SoilGrids	0.06	0.095	0.674
Sentinel-1 + DEM + Sentinel-2 + SoilGrids + SMAP-USDA + GLDAS	0.055	0.085	0.729

Table A3. Validation results for our models on the Sentinel-1 anchored full dataset. The number in brackets denotes the percentage change with respect to the SMAP-USDA + GLDAS NN baseline. All results presented are in volumetric m^3/m^3 units. A +ve change in percentage

denotes an increase in Pearson correlation and decrease in ubRMSE/RMSE. The validation split consists of 22,543 data points.

Tables A2 and A3 provide the validation set results for all the models specified in Tables 8 and 16. We notice similar sets of results on both the validation and the test sets on both the datasets.

Experiments with vegetation/leaf area indices

The vegetation indices - Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) are commonly used in soil moisture estimation methods. Similarly, the MODIS Leaf Area Index (LAI) is also a widely used input.

We tried to incorporate these into our models - computed vegetation indices from Sentinel-2 added to the high resolution inputs and MODIS LAI (inputs ranging from [0, 100] normalized to [0, 255]) added to the low resolution inputs. However, we noticed that they didn't improve the performance of our models and even dropped performance slightly as shown in Table A4.

	Validation	
Feature/Source added	ubRMSE	Correlation
None (Baseline)	0.055	0.729
EVI/NDVI	0.0566	0.714
MODIS LAI	0.0567	0.714

Table A4. Validation results on the Sentinel-1 anchored full dataset for the feature importance study on additional model inputs. All results presented are in volumetric m^3/m^3 units.

Experiments and results on an Sentinel-1 + SMAP-USDA (Active-Passive) model

We also explore an active-passive microwave only model, we use Sentinel-1 + DEM (active microwave source) and SMAP-USDA (passive microwave source) as inputs to a model and compare the performance to the rest of our models/baselines. We notice that this mostly performs better than either of SMAP-USDA or Sentinel-1 + DEM individually, however significantly falls short in comparison to our best model which uses optical imagery and soil

properties and coarse GLDAS soil moisture estimates along with these inputs as shown in Table A5.

Experiment	Validation		
	ubRMSE	RMSE	Correlation
SMAP-USDA	0.098	0.144	0.634
Sentinel-1 + DEM	0.073	0.099	0.474
Sentinel-1 + DEM + SMAP-USDA (Active-passive)	0.064	0.093	0.617
Full model (For comparison)			
Sentinel-1 + DEM + Sentinel-2 + SoilGrids + SMAP-USDA + GLDAS	0.055	0.085	0.729

Table A5. Validation results on the Sentinel-1 anchored full dataset for the active-passive model compared with SMAP-USDA baselines and the best performing model. All results presented are in volumetric m^3/m^3 units.

Experiments and results on an SMAP-USDA anchored dataset

We create an SMAP-USDA anchored full dataset similar to the Sentinel-1 anchored full dataset used in the paper. The purpose of this SMAP-USDA anchored full dataset is to have a dataset to evaluate the SMAP-USDA baseline against, given that there are large time differences (~3 days) between the Sentinel-1 anchored full dataset and SMAP-USDA overpass times. This SMAP-USDA-anchored full dataset was created using the temporal bounds specified in Table A6. This dataset is created to evaluate if anchoring the dataset closer to the SMAP-USDA overpass improves the SMAP-USDA+GLDAS baseline number. The dataset consists of ~400k data points, this is larger than the Sentinel-1 anchored dataset since SMAP-USDA has a shorter revisit interval compared to Sentinel-1 which results in more matches with the in-situ data.

Source	Temporal bounds for SMAP-USDA anchored dataset
Sentinel-1	7 days
Sentinel-2	14 days
SMAP-USDA	1 hr
GLDAS	6 hrs
NASA DEM	50 years (One time)
Soil Grids	50 years (One time)

Table A6. Temporal bounds for the SMAP anchored dataset.

We perform a few experiments on this SMAP anchored full dataset in addition to our experiments on the Sentinel-1 anchored full dataset to see how the SMAP + GLDAS NN baseline performance changes depending on the data anchoring we choose.

Experiment	Validation			
	Sentinel-1 anchored data		SMAP-USDA anchored data	
	ubRMSE	Correlation	ubRMSE	Correlation
Baseline				
SMAP-USDA + GLDAS NN	0.062	0.658	0.062	0.642
Ours				
Sentinel-1 + DEM + Sentinel-2 + SoilGrids	0.06	0.674	0.061 (+1.5%)	0.632 (-1.6%)
Sentinel-1 + DEM + Sentinel-2 + SoilGrids + SMAP-USDA + GLDAS	0.055	0.729	0.056 (+9.5%)	0.703 (+9.5%)

Table A7. Validation results on the Sentinel-1 and SMAP-USDA anchored full datasets. The number in brackets denotes the percentage change with respect to the SMAP-USDA + GLDAS NN baseline. All results presented are in volumetric m^3/m^3 units. A +ve change in percentage denotes an increase in Pearson correlation and decrease in ubRMSE/RMSE.

Looking at the results in Table A7, there are two sets of comparisons we'd like to make here,

- 1) Considering the Sentinel-1 anchored and SMAP-USDA anchored datasets individually
 - a) Our best models outperform their corresponding baselines by a significant margin in both cases.
- 2) Comparing the SMAP-USDA anchored dataset results vs the Sentinel-1 anchored dataset results
 - a) On the SMAP-USDA + GLDAS NN baseline, there isn't much of a change, in fact the correlation drops a little. This indicates that using an Sentinel-1 anchored dataset for our baseline results in our experiments, ablations and sensitivity analyses does not undermine the performance of the baseline model.
 - b) On the model using Sentinel-1 + DEM + Sentinel-2 + SoilGrids (no coarse soil moisture inputs), we notice a decrease in performance - both ubRMSE and correlation. This indicates that anchoring data being used by the model is important (Sentinel-1 anchoring in this case).
 - c) On our best model which uses all the inputs, we still notice a slight decrease in performance. When we anchor on Sentinel-1, SMAP-USDA has a 3 day time bound whereas when we anchor on SMAP-USDA, Sentinel-1 has a 7 day time bound owing to the smaller revisit interval of SMAP-USDA when compared to Sentinel-1. The results indicate that anchoring on Sentinel-1 is more meaningful for our models than anchoring on SMAP-USDA (although both Sentinel-1 and SMAP-USDA are used as inputs in this model). The hypothesis could be either that Sentinel-1 is a more significant input for our models compared to SMAP-USDA, the asymmetry in time bounds between Sentinel-1 and SMAP-USDA based on the anchoring or the sampling of labels could be slightly different in both cases.

Cloud filtering during dataset creation

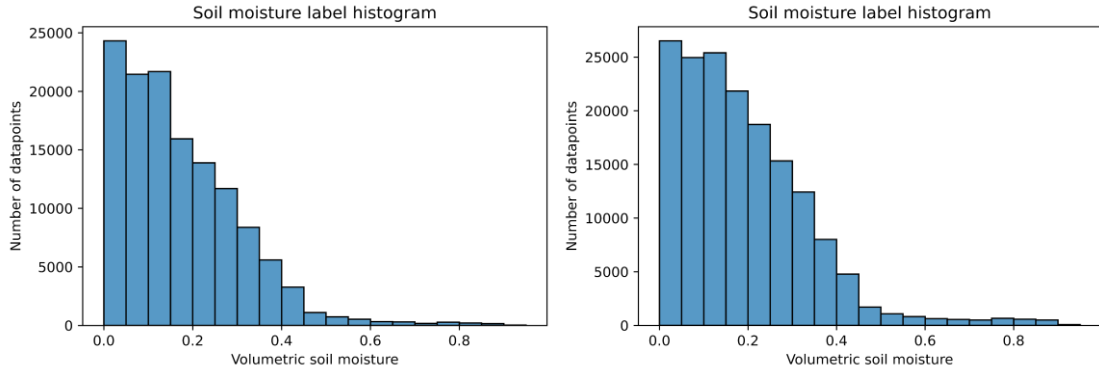


Fig A12. Left: Volumetric soil moisture m^3/m^3 label histogram with cloud filtering. Right: Soil moisture label histogram without cloud filtering.

From Figure A12, we notice that cloud filtering does end up dropping a larger percentage of higher soil moisture labels compared to lower ones. However, the ratio of higher to lower soil moisture readings remains fairly similar.

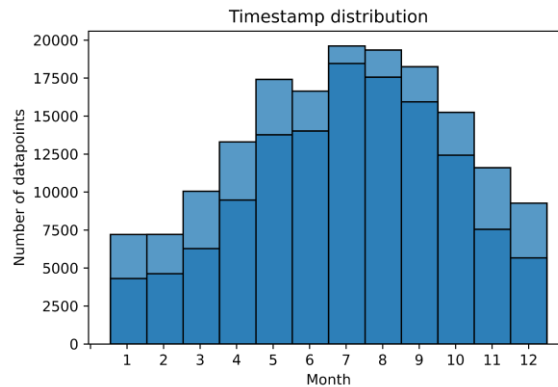


Fig A13. Soil moisture datapoints monthwise distribution. Dark blue lines correspond to the cloud filtered dataset. Light blue bars correspond to data without cloud filtering. Note that light blue bars extend all the way from the x axis and are just placed behind the dark blue bars for easier comparison.

From Figure A13, we notice that during the months of July/August/September (summer months), lesser data is dropped compared to the other months since the amount of cloud cover is relatively lower during these months.

Additional results

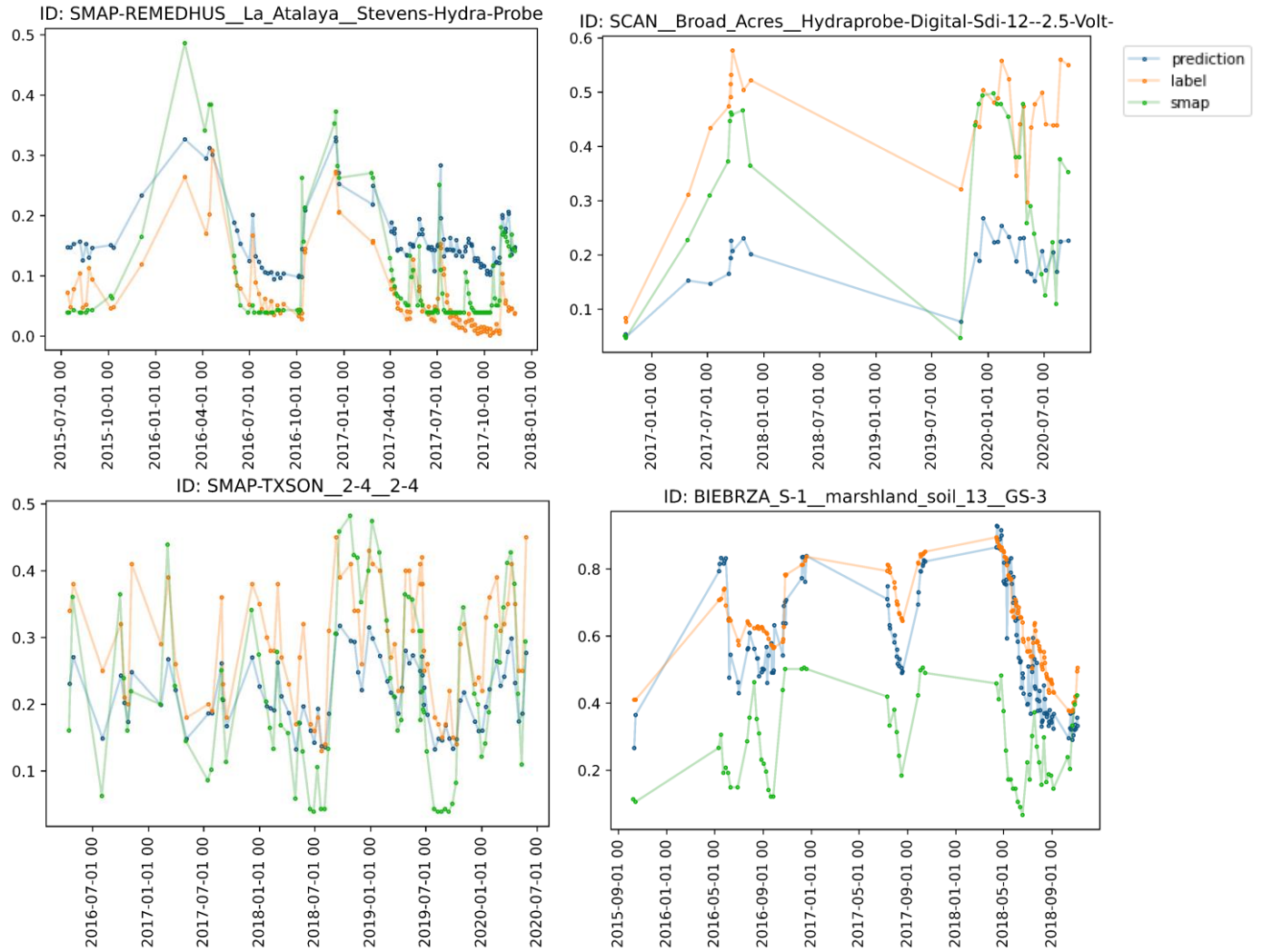


Fig. A14. Additional randomly selected time series plots showing how volumetric soil moisture model predictions m^3/m^3 vary against volumetric soil moisture m^3/m^3 labels/SMAP-USA predictions across time on our best performing model.

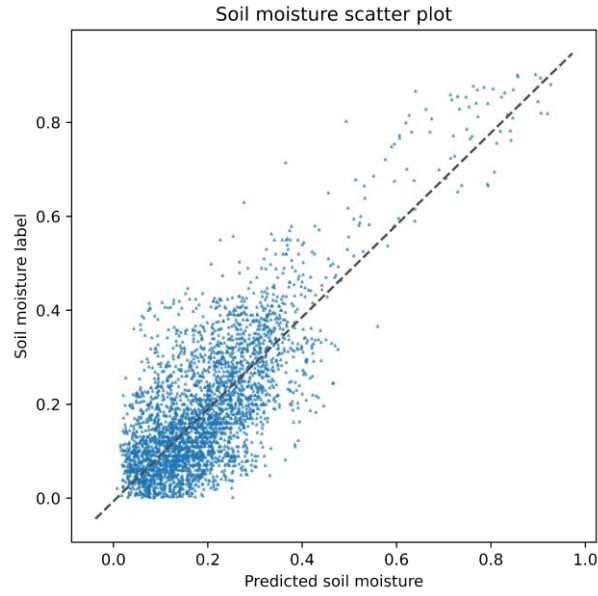


Fig. A15. Scatter plot showing predicted volumetric soil moisture m^3/m^3 vs labels m^3/m^3 across the validation set (subsamped at 15% and plotted with an opacity of 0.6 to see trends clearly) on our best performing model.

In Fig A15, we notice that the model has no significant bias at the lower side of soil moisture values. However, at higher soil moisture values, we notice that the model has a slight negative bias.

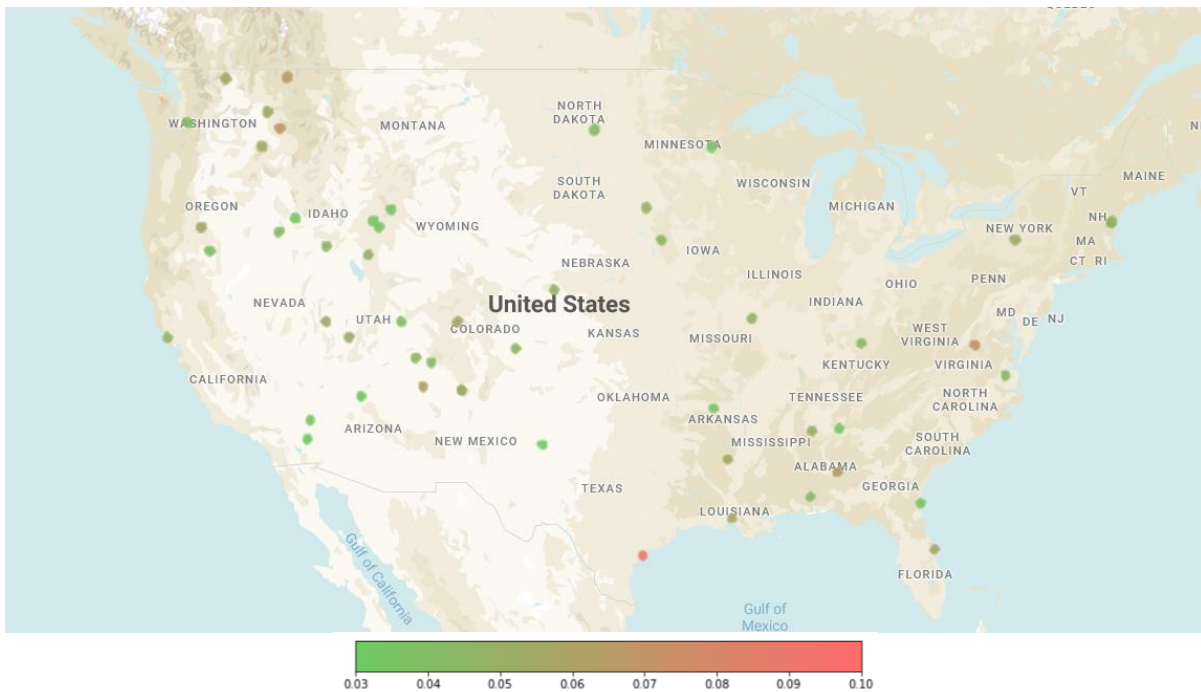


Fig. A16. ubRMSE m^3/m^3 between model estimates and ground truth labels at each validation sensor from the Sentinel-1 anchored hard dataset located in the US. Red-Green circles indicate in-situ sensor locations at which we report the ubRMSE.

REFERENCES

- Abbaszadeh, P., Moradkhani, H., & Zhan, X. (2019). *Downscaling SMAP Radiometer Soil Moisture Over the CONUS Using an Ensemble Learning Method*.
<https://doi.org/10.1029/2018WR023354>
- Addis, H., Klik, A., & Strohmeier, S. (2015). *Spatial Variability of Selected Soil Attributes under Agricultural Land Use System in a Mountainous Watershed, Ethiopia*.
<https://pdfs.semanticscholar.org/8ce1/7fe12502a942f7717c137049d9278451cf35.pdf>
- Ahmed, A., Zhang, Y., & Nichols, S. (2011). *Review and evaluation of remote sensing methods for soil-moisture estimation*. <https://doi.org/10.1117/1.3534910>
- Albergel, C., de Rosnay, P., Balsamo, G., Isaksen, L., & Muñoz-Sabater, J. (2012). Soil Moisture Analyses at ECMWF: Evaluation Using Global Ground-Based In Situ Observations. *Journal of Hydrometeorology*, 13(5), 1442–1460. <https://doi.org/10.1175/JHM-D-11-0107.1>
- Alyaari, A., Dayau, S., Chipeaux, C., Aluome, C., Kruszewski, A., Loustau, D., & Wigneron, J. (2018). *The AQUI Soil Moisture Network for Satellite Microwave Remote Sensing Validation in South-Western France*.
<https://pdfs.semanticscholar.org/185b/768b43b92cb0c765d2604d897310031b789b.pdf>
- Al-Yaari, A., Wigneron, J.-P., Kerr, Y., Rodriguez-Fernandez, N., O'Neill, P. E., Jackson, T. J., De Lannoy, G. J. M., Al Bitar, A., Mialon, A., Richaume, P., Walker, J. P., Mahmoodi, A., & Yueh, S. (2017). Evaluating soil moisture retrievals from ESA's SMOS and NASA's

- SMAP brightness temperature datasets. *Remote Sensing of the Environment*, 193, 257–273.
<https://doi.org/10.1016/j.rse.2017.03.010>
- Babu, G. S., Zhao, P., & Li, X. (2016). *Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life*. https://doi.org/10.1007/978-3-319-32025-0_14
- Baghdadi, N., Gherboudj, I., Zribi, M., Sahebi, M., King, C., & Bonn, F. (2004). *Semi-empirical calibration of the IEM backscattering model using radar images and moisture and roughness field measurements*. <https://doi.org/10.1080/01431160310001654392>
- Balenzano, A., Mattia, F., Satalino, G., Lovergine, F. P., Palmisano, D., Peng, J., Marzahn, P., Wegmüller, U., Cartus, O., Dąbrowska-Zielińska, K., Musial, J. P., Davidson, M. W. J., Pauwels, V. R. N., Cosh, M. H., McNairn, H., Johnson, J. T., Walker, J. P., Yueh, S. H., Entekhabi, D., ... Jackson, T. J. (2021). Sentinel-1 soil moisture at 1 km resolution: a validation study. *Remote Sensing of the Environment*, 263(112554), 112554.
<https://doi.org/10.1016/j.rse.2021.112554>
- Beale, J., Snapir, B., Waine, T., Evans, J., & Corstanje, R. (2019). *The significance of soil properties to the estimation of soil moisture from C-band synthetic aperture radar*. <https://pdfs.semanticscholar.org/c406/f5359fc2b2082dbc0f26adec8cb6f98e34f5.pdf>
- Bell, J., Palecki, M., Baker, C., Collins, W., Lawrimore, J., Leeper, R., Hall, M., Kochendorfer, J., Meyers, T., Wilson, T., & Diamond, H. (2013). *U.S. Climate Reference Network Soil Moisture and Temperature Observations*. <https://doi.org/10.1175/JHM-D-12-0146.1>
- Bi, H., Ma, J., Zheng, W., & Zeng, J. (2016). Comparison of soil moisture in GLDAS model simulations and in situ observations over the Tibetan Plateau. *Journal of Geophysical Research*, 121(6), 2658–2678. <https://doi.org/10.1002/2015jd024131>

- Bittelli, M. (2011). *Measuring Soil Water Content: A Review*.
<https://doi.org/10.21273/HORTTECH.21.3.293>
- Brocca, L., Ciabatta, L., Massari, C., Camici, S., & Tarpanelli, A. (2017). Soil moisture for hydrological applications: Open questions and new opportunities. *WATER*, 9(2), 140.
<https://doi.org/10.3390/w9020140>
- Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., & Fernández-Prieto, D. (2018). *How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products*. <https://doi.org/10.1016/j.jag.2018.08.023>
- Brown, C. F., Brumby, S., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J. C., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., & Tait, A. M. (2022). *Dynamic World, Near real-time global 10 m land use land cover mapping*.
<https://www.ncbi.nlm.nih.gov/pubmed/35680939>
- Buchhorn, M., Lesiv, M., Tsendbazar, N., Herold, M., Bertels, L., & Smets, B. (2020). *Copernicus Global Land Cover Layers - Collection 2*. <https://doi.org/10.3390/rs12061044>
- Cai, X., Pan, M., Chaney, N., Colliander, A., Misra, S., Cosh, M., Crow, W., Jackson, T., & Wood, E. (2017). *Validation of SMAP soil moisture for the SMAPVEX15 field campaign using a hyper-resolution model*. <https://doi.org/10.1002/2016WR019967>
- Caldwell, T. G., Bongiovanni, T., Cosh, M. H., Jackson, T. J., Colliander, A., Abolt, C. J., Casteel, R., Larson, T., Scanlon, B. R., & Young, M. H. (2019). *The Texas Soil Observation Network: A Comprehensive Soil Moisture Dataset for Remote Sensing and Land Surface Model Validation*. <https://pubag.nal.usda.gov/catalog/6773476>
- Calvet, J., Fritz, N., Berne, C., Piguet, B., Maurel, W., & Meurey, C. (2016). *Deriving*

- pedotransfer functions for soil quartz fraction in southern France from reverse modeling.*
<https://pdfs.semanticscholar.org/705f/b20a93f11a11609ab714710da5535129f853.pdf>
- Chan, S., Njoku, E., & Colliander, A. (2012). *Soil Moisture Active Passive (SMAP) Algorithm Theoretical Basis Document (ATBD)*. [https://www.semanticscholar.org/paper/Soil-Moisture-Active-Passive-\(SMAP\)-Algorithm-Basis-Chan-Njoku/0acf97080b0aa53b04bd726bc172ce3972c52027](https://www.semanticscholar.org/paper/Soil-Moisture-Active-Passive-(SMAP)-Algorithm-Basis-Chan-Njoku/0acf97080b0aa53b04bd726bc172ce3972c52027)
- Cheng, H.-T., Haque, Z., Hong, L., Ispir, M., Mewald, C., Polosukhin, I., Roumpos, G., Sculley, D., Smith, J., Soergel, D., Tang, Y., Tucker, P., Wicke, M., Xia, C., & Xie, J. (2017). *TensorFlow Estimators: Managing Simplicity vs. Flexibility in High-Level Machine Learning Frameworks*. <http://dl.acm.org/citation.cfm?id=3098171>
- Chen, K. S., Wu, T.-D., Tsang, L., Li, Q., Shi, J., & Fung, A. (2003). *Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations*. <https://doi.org/10.1109/TGRS.2002.807587>
- Chen, Q., Zeng, J., Cui, C., Li, Z., Chen, K. S., Bai, X., & Xu, J. (2018). *Soil Moisture Retrieval From SMAP: A Validation and Error Analysis Study Using Ground-Based Observations Over the Little Washita Watershed*. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8101554>
- Choker, M., Baghdadi, N., Zribi, M., Hajj, M., Paloscia, S., Verhoest, N., Lievens, H., & Mattia, F. (2017). *Evaluation of the Oh, Dubois and IEM Backscatter Models Using a Large Dataset of SAR Data and Experimental Soil Measurements*. <https://pdfs.semanticscholar.org/32fe/df4dbeb2c65f7168f9d8148da6f6dffe4a2d.pdf>
- Chollet, F. (2017). *Xception: Deep Learning with Depthwise Separable Convolutions*. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8099678>

- Colliander, A., Jackson, T., Bindlish, R., Chan, S., Das, N., Kim, S., Cosh, M., Dunbar, R., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K. K., Goodrich, D., Jassar, H. K. A., ... Yueh, S. (2017). *Validation of SMAP surface soil moisture products with core validation sites*.
<https://doi.org/10.1016/J.RSE.2017.01.021>
- Collow, T. W., Robock, A., Basara, J., & Illston, B. G. (2012). *Evaluation of SMOS retrievals of soil moisture over the central United States with currently available in situ observations*.
<https://doi.org/10.1029/2011JD017095>
- Cubuk, E. D., Zoph, B., Mané, D., Vasudevan, V., & Le, Q. V. (2019). *AutoAugment: Learning Augmentation Strategies From Data*.
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8953317>
- Dąbrowska-Zielińska, K., Musiał, J., Malinska, A., Budzynska, M., Gurdak, R., Kiryla, W., Bartold, M., & Grzybowski, P. (2018). *Soil Moisture in the Biebrza Wetlands Retrieved from Sentinel-1 Imagery*.
<https://pdfs.semanticscholar.org/36cc/0014895ff1dfe7eba5acf5122174791a72bb.pdf>
- Das, N., Entekhabi, D., Dunbar, R., Chaubell, M., Colliander, A., Yueh, S., Jagdhuber, T., Chen, F., Crow, W., O'neill, P., Walker, J. P., Berg, A., Bosch, D., Caldwell, T., Cosh, M., Collins, C. H., López-Baeza, E., & Thibeault, M. (2019). *The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product*.
<https://doi.org/10.1016/j.rse.2019.111380>
- Das, N., Entekhabi, D., Dunbar, R., Colliander, A., Chen, F., Crow, W., Jackson, T., Berg, A., Bosch, D., Caldwell, T., Cosh, M., Collins, C. H., López-Baeza, E., Moghaddam, M., Rowlandson, T., Starks, P., Thibeault, M., Walker, J. P., Wu, X., ... Njoku, E. (2018). *The*

- SMAP mission combined active-passive soil moisture product at 9 km and 3 km spatial resolutions.* <https://doi.org/10.1016/J.RSE.2018.04.011>
- Davenport, I., Sandells, M., & Gurney, R. (2008). *The effects of scene heterogeneity on soil moisture retrieval from passive microwave data.* <https://doi.org/10.1016/J.ADVWATRES.2008.06.002>
- Dente, L., Su, Z., & Wen, J. (2012). *Validation of SMOS Soil Moisture Products over the Maqu and Twente Regions.* <https://www.ncbi.nlm.nih.gov/pubmed/23112582>
- Dobriyal, P., Qureshi, A., Badola, R., & Hussain, S. A. (2012). *A review of the methods available for estimating soil moisture and its implications for water resource management.* <https://doi.org/10.1016/J.JHYDROL.2012.06.021>
- Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A. D., Zamojski, D., Cordes, C., Wagner, W., & Drusch, M. (2013). Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network. *Vadose Zone Journal*, 12(3). <https://doi.org/10.2136/vzj2012.0097>
- Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., Preimesberger, W., Xaver, A., Annor, F., Ardö, J., Baldocchi, D., Blöschl, G., Bogaen, H., Brocca, L., Calvet, J., Camarero, J., Capello, G., Choi, M., Cosh, M. C., ... Sabia, R. (2021). *The International Soil Moisture Network: serving Earth system science for over a decade.* <https://pdfs.semanticscholar.org/f348/9cf3a90905b4a5830bc558f4b0e87a94ce61.pdf>
- Dorigo, W., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Oevelen, P. V., Robock, A., & Jackson, T. (2011). *The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements.*

<https://pdfs.semanticscholar.org/811e/bab424de0997c5df1c042ab6484439ed8afe.pdf>

- Drusch, M., Bello, U. D., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). *Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services*. <https://doi.org/10.1016/J.RSE.2011.11.026>
- Dubois, P. C., van Zyl, J., & Engman, T. (1995). Measuring soil moisture with imaging radars. *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society*, 33(4), 915–926. <https://doi.org/10.1109/36.406677>
- Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., ... Van Zyl, J. (2010). The Soil Moisture Active Passive (SMAP) Mission. *Entekhabi, D. et Al. "The Soil Moisture Active Passive (SMAP) Mission." Proceedings of the IEEE*, 9, 8.5. <http://hdl.handle.net/1721.1/60043>
- Evensen, G. (2003). *The Ensemble Kalman Filter: theoretical formulation and practical implementation*. <https://doi.org/10.1007/S10236-003-0036-9>
- Falloon, P., Jones, C., Ades, M., & Paul, K. (2011). *Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty*. <https://doi.org/10.1029/2010GB003938>
- Fang, B., Lakshmi, V., Cosh, M., & Hain, C. (2021). *Very High Spatial Resolution Downscaled SMAP Radiometer Soil Moisture in the CONUS Using VIIRS/MODIS Data*. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9416730>
- Farr, T., Rosen, P., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M.,

- Rodríguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2000). *The Shuttle Radar Topography Mission*.
<https://doi.org/10.1029/2005RG000183>
- Foster, T., Mieno, T., & Brozović, N. (2020). *Satellite- Based Monitoring of Irrigation Water Use: Assessing Measurement Errors and Their Implications for Agricultural Water Management Policy*. <https://doi.org/10.1029/2020WR028378>
- Fu, H., Gong, M., Wang, C., Batmanghelich, K., & Tao, D. (2018). *Deep Ordinal Regression Network for Monocular Depth Estimation*.
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8578312>
- Galle, S., Grippa, M., Peugeot, C., Moussa, I., Cappelaere, B., Demarty, J., Mougín, E., Panthou, G., Adjomayi, P., Agbossou, E., Ba, A., Boucher, M., Cohard, J., Descloîtres, M., Descroix, L., Diawara, M., Dossou, M., Favreau, G., Gangneron, F., ... Wilcox, C. (2018). *AMMA-CATCH, a Critical Zone Observatory in West Africa Monitoring a Region in Transition*.
<https://doi.org/10.2136/vzj2018.03.0062>
- Gao, Q., Zribi, M., Escorihuela, M., & Baghdadi, N. (2017). *Synergetic Use of Sentinel-1 and Sentinel-2 Data for Soil Moisture Mapping at 100 m Resolution*.
<https://pdfs.semanticscholar.org/9a01/4048f616dd9df4be363a7934526fab07657.pdf>
- González-Zamora, Á., Sánchez, N., Pablos, M., & Martínez-Fernández, J. (2019). *CCI soil moisture assessment with SMOS soil moisture and in situ data under different environmental conditions and spatial scales in Spain*.
<https://doi.org/10.1016/J.RSE.2018.02.010>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). *Google Earth Engine: Planetary-scale geospatial analysis for everyone*.

<https://doi.org/10.1016/J.RSE.2017.06.031>

Greifeneder, F., Notarnicola, C., & Wagner, W. (2021). A Machine Learning-Based Approach for Surface Soil Moisture Estimations with Google Earth Engine. *Remote Sensing*, 13(11), 2099. <https://doi.org/10.3390/rs13112099>

Groenendyk, D. G., Ferré, T., Thorp, K., & Rice, A. (2015). *Hydrologic-Process-Based Soil Texture Classifications for Improved Visualization of Landscape Function*. <https://pdfs.semanticscholar.org/2ea3/f960d72f116f47766147394d8c05345cdc16.pdf>

Haider, S., Said, S., Kothiyari, U. C., & Arora, M. K. (2004). *Soil moisture estimation using ERS 2 SAR data: a case study in the Solani River catchment/Estimation de l'humidité du sol grâce à des données ERS-2 SAR: étude de cas dans le bassin de la rivière Solani*. <https://doi.org/10.1623/hysj.49.2.323.34832>

Heatmaps. (n.d.). Google Developers. Retrieved October 7, 2022, from

<https://developers.google.com/maps/documentation/javascript/examples/layer-heatmap>

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep Residual Learning for Image Recognition*. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7780459>

Hengl, T., de Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PloS One*, 12(2), e0169748. <https://doi.org/10.1371/journal.pone.0169748>

Huang, G., Sun, Y., Liu, Z., Sedra, D., & Weinberger, K. Q. (2016). *Deep Networks with Stochastic Depth*. <https://arxiv.org/pdf/1603.09382.pdf>

- Huber, P. J. (1992). Robust Estimation of a Location Parameter. *Breakthroughs in Statistics*, 492–518. https://doi.org/10.1007/978-1-4612-4380-9_35
- Ioffe, S., & Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. <https://arxiv.org/pdf/1502.03167.pdf>
- Jing, W., Song, J., & Zhao, X. (2018). Validation of ECMWF multi-layer reanalysis soil moisture based on the OzNet hydrology network. *WATER*, 10(9), 1123. <https://doi.org/10.3390/w10091123>
- Jpl, N. (2020). *NASADEM Merged DEM Global 1 arc second V001* [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MEASURES/NASADEM/NASADEM_HGT.001
- Karthikeyan, L., & Mishra, A. (2021). *Multi-layer high-resolution soil moisture estimation using machine learning over the United States*. <https://doi.org/10.1016/j.rse.2021.112706>
- Kerr, Y., Al-Yaari, A., Rodríguez-Fernández, N., Parrens, M., Molero, B., Leroux, D., Bircher, S., Mahmoodi, A., Mialon, A., Richaume, P., Delwart, S., Al Bitar, A., Pellarin, T., Bindlish, R., Jackson, T., Rudiger, C., Waldteufel, P., Mecklenburg, S., & Wigneron, J. (2016). *Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation*. <https://doi.org/10.1016/J.RSE.2016.02.042>
- Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, 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: A Publication of the IEEE Geoscience and Remote Sensing Society*, 39(8), 1729–1735. <https://doi.org/10.1109/36.942551>
- Kerr, Y., Waldteufel, P., Richaume, P., Wigneron, J., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S., Leroux, D., Mialon, A., & Delwart, S. (2012). *The SMOS*

Soil Moisture Retrieval Algorithm.

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6161633>

Kim, S., Zhang, R., Pham, H., & Sharma, A. (2019). *A Review of Satellite-Derived Soil Moisture and Its Usage for Flood Estimation.* <https://doi.org/10.1007/s41976-019-00025-7>

Klute, A. (1986). *Methods of soil analysis. Part 1. Physical and mineralogical methods.* <https://doi.org/10.1097/00010694-198808000-00014>

Kolassa, J., Reichle, R., Liu, Q., Alemohammad, S. H., Gentine, P., Aida, K., Asanuma, J., Bircher, S., Caldwell, T., Colliander, A., Cosh, M., Collins, C. H., Jackson, T., Martínez-Fernández, J., Mcnairn, H., Pacheco, A., Thibeault, M., & Walker, J. (2018). *Estimating surface soil moisture from SMAP observations using a Neural Network technique.* <https://www.ncbi.nlm.nih.gov/pubmed/29290638>

Kornblith, S., Shlens, J., & Le, Q. V. (2019). *Do Better ImageNet Models Transfer Better?* <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8954384>

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). *World Map of the Köppen-Geiger climate classification updated.* <https://doi.org/10.1127/0941-2948/2006/0130>

Kumar, A., Kamal, K., Arshad, M. O., Mathavan, S., & Vadamala, T. (2014). *Smart irrigation using low-cost moisture sensors and XBee-based communication.*

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6970301>

Laiolo, P., Gabellani, S., Campo, L., Silvestro, F., Delogu, F., Rudari, R., Pulvirenti, L., Boni, G., Fascetti, F., Pierdicca, N., Crapolicchio, R., Hasenauer, S., & Puca, S. (2016). Impact of different satellite soil moisture products on the predictions of a continuous distributed hydrological model. *International Journal of Applied Earth Observation and Geoinformation*, 48, 131–145. <https://doi.org/10.1016/j.jag.2015.06.002>

- Lary, D. J., Alavi, A., Gandomi, A., & Walker, A. (2016). *Machine learning in geosciences and remote sensing*. <https://doi.org/10.1016/J.GSF.2015.07.003>
- Leavesley, G., David, O., Garen, D., Nrcs-USda, N., Goodbody, A., Lea, J., Marron, J., & Strobel, M. (2010). *A MODELING FRAMEWORK FOR IMPROVED AGRICULTURAL WATER-SUPPLY FORECASTING*. <https://www.semanticscholar.org/paper/A-MODELING-FRAMEWORK-FOR-IMPROVED-AGRICULTURAL-Leavesley-David/c5405fe8403fb6aa198c1af35ae7245dcd916e09>
- Lee, C.-S., Sohn, E., Park, J. D., & Jang, J.-D. (2018). *Estimation of soil moisture using deep learning based on satellite data: a case study of South Korea*. <https://doi.org/10.1080/15481603.2018.1489943>
- Leenaars, J., Claessens, L., Heuvelink, G., Hengl, T., González, M. R., Van Bussel, L. V., Guilpart, N., Yang, H., & Cassman, K. (2018). *Mapping rootable depth and root zone plant-available water holding capacity of the soil of sub-Saharan Africa*. <https://www.ncbi.nlm.nih.gov/pubmed/30122789>
- Legates, D. R., Mahmood, R., Levia, D. F., DeLiberty, T. L., Quiring, S. M., Houser, C., & Nelson, F. E. (2011). Soil moisture: A central and unifying theme in physical geography. *Progress in Physical Geography*, 35(1), 65–86. <https://doi.org/10.1177/0309133310386514>
- Liu, Y., Jing, W., Wang, Q., & Xia, X. (2020). *Generating high-resolution daily soil moisture by using spatial downscaling techniques: a comparison of six machine learning algorithms*. <https://doi.org/10.1016/j.advwatres.2020.103601>
- Loshchilov, I., & Hutter, F. (2019). *Decoupled Weight Decay Regularization*. <https://openreview.net/forum?id=Bkg6RiCqY7>
- Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., & Gascon, F. (2017).

Sen2Cor for Sentinel-2. <https://doi.org/10.1117/12.2278218>

Mao, H., Kathuria, D., Duffield, N., & Mohanty, B. (2019). *Gap Filling of High- Resolution Soil Moisture for SMAP/Sentinel- 1: A Two- Layer Machine Learning- Based Framework*.

<https://doi.org/10.1029/2019WR024902>

Ma, T., Han, L., & Liu, Q. (2021). Retrieving the soil moisture in bare farmland areas using a modified Dubois model. *Frontiers of Earth Science*, 9.

<https://doi.org/10.3389/feart.2021.735958>

Mattia, F., Toan, T., Souyris, J., Carolis, G., Floury, N., Posa, F., & Pasquariello, G. (1997). *The effect of surface roughness on multifrequency polarimetric SAR data*.

<https://doi.org/10.1109/36.602537>

Meteorological Drought. (1965). Google Books.

https://books.google.com/books/about/Meteorological_Drought.html?id=kyYZgnEk-L8C

MirMazloumi, S. M., & Sahebi, M. R. (2016). Assessment of different backscattering models for bare soil surface parameters estimation from SAR data in band C, L and P. *European Journal of Remote Sensing*, 49(1), 261–278. <https://doi.org/10.5721/eujrs20164915>

Mladenova, I. E., Bolten, J. D., Crow, W., Sazib, N., & Reynolds, C. (2020). Agricultural Drought Monitoring via the Assimilation of SMAP Soil Moisture Retrievals Into a Global Soil Water Balance Model. *Frontiers in Big Data*, 0.

<https://doi.org/10.3389/fdata.2020.00010>

Montzka, C., Cosh, M., Bayat, B., Al Bitar, A., Berg, A., Bindlish, R., Bogen, H., Bolton, J. D., Cabot, F., Caldwell, T., Chan, S., Colliander, A., Crow, W., Das, N., Lannoy, G., Dorigo, W., Evett, S., Gruber, A., Hahn, S., ... Nickeson, J. (2020). *Soil moisture product validation good practices protocol, version 1.0*.

<https://doi.org/10.5067/DOC/CEOSWGCV/LPV/SM.001>

- Montzka, C., Rötzer, K., Bogena, H. R., Sanchez, N., & Vereecken, H. (2018). A New Soil Moisture Downscaling Approach for SMAP, SMOS, and ASCAT by Predicting Sub-Grid Variability. *Remote Sensing*, *10*(3), 427. <https://doi.org/10.3390/rs10030427>
- Moran, M. S., Doorn, B., Escobar, V., & Brown, M. E. (2015). Connecting NASA science and engineering with Earth science applications. *Journal of Hydrometeorology*, *16*(1), 473–483. <https://doi.org/10.1175/jhm-d-14-0093.1>
- Murray, D., Simsa, J., Klimovic, A., & Indyk, I. (2021). *tf.data: A Machine Learning Data Processing Framework*. <https://arxiv.org/pdf/2101.12127.pdf>
- Njoku, E., Jackson, T., Lakshmi, V., Chan, S., & Nghiem, S. (2003). *Soil moisture retrieval from AMSR-E*. <https://doi.org/10.1109/TGRS.2002.808243>
- Njoku, E., Wilson, W., Yueh, S., Dinardo, S., Li, F. K., Jackson, T., Lakshmi, V., & Bolten, J. (2002). *Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99*. <https://doi.org/10.1109/TGRS.2002.807008>
- Oh, Y., Sarabandi, K., & Ulaby, F. T. (1992). An empirical model and an inversion technique for radar scattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society*, *30*(2), 370–381. <https://doi.org/10.1109/36.134086>
- Ojha, N., Merlin, O., Amazirh, A., Ouaadi, N., Rivalland, V., Jarlan, L., Er-raki, S., & Escorihuela, M. (2021). *A Calibration/Disaggregation Coupling Scheme for Retrieving Soil Moisture at High Spatio-Temporal Resolution: Synergy between SMAP Passive Microwave, MODIS/Landsat Optical/Thermal and Sentinel-1 Radar Data*. <https://pdfs.semanticscholar.org/7a17/9ab7de8ad53ed3b1ab4e026eab866c29a1f5.pdf>

- Oliver, C., & Quegan, S. (1998). *Understanding Synthetic Aperture Radar Images*.
<https://www.semanticscholar.org/paper/Understanding-Synthetic-Aperture-Radar-Images-Oliver-Quegan/6f04e7388117ce4263ab24706c3f7f8e0f4f1955>
- Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., & Reppucci, A. (2013). *Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation*.
<https://doi.org/10.1016/J.RSE.2013.02.027>
- Panciera, R., Tanase, M. A., Lowell, K., & Walker, J. P. (2014). Evaluation of IEM, Dubois, and oh radar backscatter models using airborne L-band SAR. *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society*, 52(8), 4966–4979. <https://doi.org/10.1109/tgrs.2013.2286203>
- Pan, W., Boyles, R., White, J. G., & Heitman, J. (2012). *Characterizing Soil Physical Properties for Soil Moisture Monitoring with the North Carolina Environment and Climate Observing Network*. <https://doi.org/10.1175/JTECH-D-11-00104.1>
- Patel, C., Sharma, S., & Gulshan, V. (2021). *Evaluating Self and Semi-Supervised Methods for Remote Sensing Segmentation Tasks*. <https://arxiv.org/pdf/2111.10079.pdf>
- Raiyani, K., Gonçalves, T., Rato, L., Salgueiro, P. D., & da Silva, J. R. M. (2021). *Sentinel-2 Image Scene Classification: A Comparison between Sen2Cor and a Machine Learning Approach*. <https://doi.org/10.3390/rs13020300>
- Ramachandran, P., Zoph, B., & Le, Q. V. (2017). *Searching for Activation Functions*.
<https://doi.org/10.48550/arXiv.1710.05941>
- Robinson, D. A., Campbell, C. S., Hopmans, J. W., Hornbuckle, B. K., Jones, S. B., Knight, R., Ogden, F., Selker, J., & Wendroth, O. (2008). Soil moisture measurement for ecological and hydrological watershed- scale observatories: A review. *Vadose Zone Journal: VZJ*, 7(1),

358–389. <https://doi.org/10.2136/vzj2007.0143>

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J., Walker, J. P., Lohmann, D., & Toll, D. (2004). *THE GLOBAL LAND DATA ASSIMILATION SYSTEM*.

<https://doi.org/10.1175/BAMS-85-3-381>

Rogez, G., Weinzaepfel, P., & Schmid, C. (2017). *LCR-Net: Localization-Classification-Regression for Human Pose*.

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8099617>

Rosen, P., Hensley, S., Joughin, I., Li, F. K., Madsen, S., Rodríguez, E., & Goldstein, R. (2000). *Synthetic aperture radar interferometry*.

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=838084>

Rossato, L., Dos Santos Alvala, R. C., Marengo, J. A., Zeri, M., Do Amaral Cunha, A. P. M., Pires, L. B. M., & Barbosa, H. A. (2017). Impact of Soil Moisture on Crop Yields over Brazilian Semiarid. *Frontiers of Environmental Science & Engineering in China*.

<https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=2296665X&v=2.1&it=r&id=GALE%7CA513797143&sid=googleScholar&linkaccess=abs>

Rumelhart, D., Hinton, G. E., & Williams, R. J. (1986). *Learning representations by back-propagating errors*. <https://doi.org/10.1038/323533a0>

Sandler, M., Howard, A. G., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). *MobileNetV2: Inverted Residuals and Linear Bottlenecks*.

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8578572>

Sandric, I., Diamandi, A., Oana, N., Saizu, D., Vasile, C., & Lucaschi, B. (2016). VALIDATION AND UPSCALING OF SOIL MOISTURE SATELLITE PRODUCTS IN ROMANIA.

- XXIII ISPRS Congress, Commission II (Volume XLI-B2) - 12-19 July 2016, Prague, Czech Republic, XLI-B2*, 313–317. <https://doi.org/10.5194/isprs-archives-XLI-B2-313-2016>
- Schaefer, G. L., Cosh, M., & Jackson, T. (2007). *The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN)*. <https://doi.org/10.1175/2007JTECHA930.1>
- Shi, J., Wang, J. R., Hsu, A., O’neill, P., & Engman, E. (1997). *Estimation of bare surface soil moisture and surface roughness parameter using L-band SAR image data*. <https://doi.org/10.1109/36.628792>
- Shorten, C., & Khoshgoftaar, T. (2019). *A survey on Image Data Augmentation for Deep Learning*. <https://doi.org/10.1186/s40537-019-0197-0>
- SMAP/In Situ Core Validation Site Land Surface Parameters Match-Up Data*. (n.d.). Retrieved June 1, 2022, from <https://nsidc.org/data/nsidc-0712/versions/1/documentation>
- Smith, A. B., Walker, J., Western, A., Young, R., Ellett, K., Pipunic, R., Grayson, R., Siriwardena, L., Chiew, F., & Richter, H. (2012). *The Murrumbidgee soil moisture monitoring network data set*. <https://doi.org/10.1029/2012WR011976>
- Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). *Dropout: a simple way to prevent neural networks from overfitting*. <http://dl.acm.org/citation.cfm?id=2670313>
- Starks, P., Fiebrich, C., Grimsley, D., Garbrecht, J., Steiner, J., Guzman, J., & Moriasi, D. (2014). *Upper washita river experimental watersheds: meteorologic and soil climate measurement networks*. <https://www.ncbi.nlm.nih.gov/pubmed/25603072>
- Su, Z., Wen, J., Dente, L., Velde, R. V. D., Wang, L., Ma, Y., Yang, K., & Hu, Z. (2011). *The Tibetan plateau observatory of plateau scale soil moisture and soil temperature, Tibet - Obs, for quantifying uncertainties in coarse resolution satellite and model products*.

<https://pdfs.semanticscholar.org/fb9d/7599c58353e2e354e7ab37a4e75d8d7a2dee.pdf>

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Flourey, N., Brown, M., Traver, I., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., ... Rostan, F. (2012). *GMES Sentinel-1 mission*. <https://doi.org/10.1016/J.RSE.2011.05.028>

Tramblay, Y., Bouaicha, R., Brocca, L., Dorigo, W., Bouvier, C., Camici, S., & Servat, E. (2012). Estimation of antecedent wetness conditions for flood modelling in northern Morocco. *Hydrology and Earth System Sciences*, *16*(11), 4375–4386. <https://doi.org/10.5194/hess-16-4375-2012>

Veci, L., Prats-Iraola, P., Scheiber, R., Collard, F., Fomferra, N., & Engdahl, M. (2014). *The Sentinel-1 Toolbox*. <https://www.semanticscholar.org/paper/The-Sentinel-1-Toolbox-Veci-Prats-Iraola/500a708794ce247642c0c366c4c430da1190a5a5>

Velde, R. van der, & Benninga, H.-J. F. (2021). *Ten years profile soil moisture and temperature measurements in Twente* [Data set]. Data Archiving and Networked Services (DANS). <https://doi.org/10.17026/DANS-ZRN-N8NH>

Vergopolan, N., Chaney, N. W., Pan, M., Sheffield, J., Beck, H. E., Ferguson, C. R., Torres-Rojas, L., Sadri, S., & Wood, E. F. (2021). SMAP-HydroBlocks, a 30-m satellite-based soil moisture dataset for the conterminous US. *Scientific Data*, *8*(1), 1–11. <https://doi.org/10.1038/s41597-021-01050-2>

Verhoest, N., Lievens, H., Wagner, W., Álvarez-Mozos, J., Moran, M. S., & Mattia, F. (2008). *On the Soil Roughness Parameterization Problem in Soil Moisture Retrieval of Bare Surfaces from Synthetic Aperture Radar*. <https://pdfs.semanticscholar.org/7074/068f96afe6b5f9c629983fc23bfe796c0324.pdf>

- Walker, J. P., Willgoose, G., & Kalma, J. (2004). *In situ measurement of soil moisture: a comparison of techniques* | NOVA. The University of Newcastle's Digital Repository. <https://doi.org/10.1016/J.JHYDROL.2004.01.008>
- Wang, Y., Leng, P., Peng, J., Marzahn, P., & Ludwig, R. (2021). Global assessments of two blended microwave soil moisture products CCI and SMOPS with in-situ measurements and reanalysis data. *International Journal of Applied Earth Observation and Geoinformation*, 94(102234), 102234. <https://doi.org/10.1016/j.jag.2020.102234>
- Wu, X., Walker, J. P., Rüdiger, C., Panciera, R., & Gao, Y. (2017). *Intercomparison of Alternate Soil Moisture Downscaling Algorithms Using Active–Passive Microwave Observations*. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7784765>
- Zafar, A., Prathapar, S., & Bastiaanssen, W. (2021). Optimization of Canal Management Based on Irrigation Performance Analysis Using Satellite Measurements. In *Asian Development Bank*. <https://www.adb.org/sites/default/files/publication/673096/adb-brief-165-canal-management-irrigation-performance.pdf>