Estimating surface soil moisture from SMAP observations using a Neural Network technique

J. Kolassa[a,b,*], R.H. Reichle[b], Q. Liu[c,b], S.H. Alemohammad[d], P. Gentine[d], K. Aida[e], J. Asanuma[e], S. Bircher[l], T. Caldwell[m], A. Colliander[f], M. Cosh[g], C. Holifield Collins[n], T.J. Jackson[g], J. Martínez-Fernández[h], H. McNairn[i], A. Pacheco[i], M. Thibeault[j], J.P. Walker[k]

[a] *Universities Space Research Association/NPP, Columbia, MD, USA*
[b] *Global Modelling and Assimilation Office, NASA Goddard Spaceflight Center, Greenbelt, MD, USA*
[c] *Science Systems and Applications Inc., Lanham, MD, USA*
[d] *Columbia University, New York, NY, USA*
[e] *University of Tsukuba, Tsukuba, Japan*
[f] *Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA*
[g] *USDA ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD, USA*
[h] *Instituto Hispano Luso de Investigaciones Agrarias (CIALE), Universidad de Salamanca, Salamanca, Spain*
[i] *Agriculture and Agri-food Canada, Ottawa, Ontario, Canada*
[j] *Comisiòn Nacional de Actividades Espaciales (CONAE), Buenos Aires, Argentina*
[k] *Department of Civil Engineering, Monash University, Clayton, Victoria, Australia*
[l] *Centre d'Etudes Spatiales de la BIOsphère (CESBIO-CNES, CNRS, IRD, Université Toulouse III), Toulouse, France*
[m] *University of Texas at Austin, TX, USA*
[n] *USDA ARS Southwest Watershed Research Center, Tucson, AZ, USA*

Abstract

A Neural Network (NN) algorithm was developed to estimate global surface soil moisture for April 2015 to March 2017 with a 2-3 day repeat frequency using passive microwave observations from the Soil Moisture Active Passive (SMAP) satellite, surface soil temperatures from the NASA Goddard Earth Observing System Model version 5 (GEOS-5) land modeling system, and Moderate Resolution Imaging Spectroradiometer-based vegetation water content. The NN was trained on GEOS-5 soil moisture target data, making the NN estimates consistent with the GEOS-5 climatology, such that they may ultimately be assimilated into this model without further bias correction. Evaluated against in

*Corresponding author
Email address: jana.kolassa@nasa.gov (J. Kolassa)

situ soil moisture measurements, the average unbiased root mean square error (ubRMSE), correlation and anomaly correlation of the NN retrievals were 0.037 m^3m^{-3}, 0.70 and 0.66, respectively, against SMAP core validation site measurements and 0.026 m^3m^{-3}, 0.58 and 0.48, respectively, against International Soil Moisture Network (ISMN) measurements. At the core validation sites, the NN retrievals have a significantly higher skill than the GEOS-5 model estimates and a slightly lower correlation skill than the SMAP Level-2 Passive (L2P) product. The feasibility of the NN method was reflected by a lower ubRMSE compared to the L2P retrievals as well as a higher skill when ancillary parameters in physically-based retrievals were uncertain. Against ISMN measurements, the skill of the two retrieval products was more comparable. A triple collocation analysis against Advanced Microwave Scanning Radiometer 2 (AMSR2) and Advanced Scatterometer (ASCAT) soil moisture retrievals showed that the NN and L2P retrieval errors have a similar spatial distribution, but the NN retrieval errors are generally lower in densely vegetated regions and transition zones.
Keywords: soil moisture remote sensing, SMAP, data assimilation, microwave radiometer

1. Introduction

Soil moisture is a key variable for many surface and boundary layer processes, such as the coupling of the water and energy cycles (*Seneviratne et al.*, 2006; *Gentine et al.*, 2011; *Bateni and Entekhabi*, 2012) or the partitioning of precipitation into runoff and infiltration (*Philip*, 1957; *Corradini et al.*, 1998; *Assouline*, 2013). Soil moisture is also a key determinant of the carbon cycle (*McDowell*, 2011; *Sevanto et al.*, 2014; *Jung et al.*, 2017). The importance of soil moisture has been recognized by the World Meteorological Organization by naming it an Essential Climate Variable (*GCOS*, 2009) and thus encouraging efforts to obtain better soil moisture observations, which is challenging because of its high variability both in space and time.

One avenue to obtain observations of soil moisture is through satellite instruments that provide global observations with a relatively short revisit period of 2-3 days. In particular, L-band (1.4 GHz) microwave instruments exhibit a high sensitivity to soil moisture in the top ∼5 centimeters of the soil in sparsely to moderately vegetated areas. This has led to the launch of two L-band satellite missions to observe soil moisture, the European Soil Moisture and Ocean Salinity (SMOS) mission in 2009 (*Kerr et al.*, 2010) and the NASA Soil Moisture Active Passive (SMAP) mission (*Entekhabi et al.*, 2010) in 2015.

Traditionally, satellite soil moisture retrievals from L-band (and other) sensors are implemented through the inversion of Radiative Transfer Models (RTMs) (e.g. *Owe et al.* (2001); *Kerr et al.* (2012); *O'Neill et al.* (2015)), which explicitly formulate the physical relationships linking surface soil moisture to satellite brightness temperature observations. The RTM inversion technique is used to produce the official SMOS and SMAP retrieval products, and is able to provide high quality soil moisture estimates (*Al Bitar et al.*, 2012; *Chan et al.*, 2016b; *Colliander et al.*, 2017) with a typical latency of 12 to 24 hours. However, this approach requires accurate knowledge of the physical relationships between the surface state and the satellite observations as well as their associated parameters, which are often empirically estimated and thus uncertain. Moreover, RTM inversions also require explicit information on other surface states, including surface soil temperature and vegetation, and are thus typically ill-posed problems. Additionally, for time critical applications, such as near real time flood prediction or soil moisture assimilation into weather prediction models, retrieval products with a shorter latency are required.

Data assimilation provides another option to generate improved soil moisture estimates through the merging of satellite and model information, and can yield soil moisture estimates that are of higher quality than estimates from satellite observations or models alone (e.g. *Entekhabi et al.* (1994); *Walker and Houser* (2001); *Liu et al.* (2011); *Lahoz and De Lannoy* (2014)). For soil moisture assimilation, the observations and model estimates have to be unbiased with respect to each other, which is typically achieved by locally matching the mean and variability of the satellite observations to those of the model (*Reichle and Koster*, 2004). While this satisfies the requirements of the assimilation system, it has the side effect of removing some independent information in the satellite observations. Given the high quality of soil moisture observations from SMOS and SMAP this is undesirable.

As an alternative to RTM inversions, statistical Neural Network (NN) retrieval algorithms have been successfully implemented for a number of sensors in recent years (*Aires et al.*, 2005; *Chai et al.*, 2009; *Kolassa et al.*, 2013, 2016; *Rodriguez-Fernández et al.*, 2015; *Santi et al.*, 2016). Instead of explicitly formulating physical relationships, NNs are calibrated on a sample of satellite observations and corresponding soil moisture estimates (the target data) to model the global statistical relationship between the satellite observations and surface soil moisture. As a result, NN retrievals can offer several general advantages over traditional RTM inversions. First, they do not require an explicit parameterization of physical relationships and are thus not affected by errors in our knowledge of these relationships or their parameters. Second, after a one-time calibration, NNs are computationally extremely efficient and can provide soil moisture estimates almost immediately after arrival of the instrument data, thereby shortening the latency. Third, training a NN non-locally on target data from a model, yields NN retrievals that are globally unbiased with respect to the model, with spatial and temporal patterns that are driven by the satellite observations (e.g. *Jimenez et al.* (2013); *Kolassa et al.* (2016); *Alemohammad et al.* (2017)). This may reduce the need for bias correction prior to an assimilation and at the same time retain more of the independent information contained in the spatial and temporal patterns of the satellite observations.

In this study, we develop the first NN algorithm to retrieve global surface soil moisture from SMAP observations. The motivation for this work is twofold. First, we investigate statistical retrieval techniques as a possible alternative or supplement to the existing physically-based SMAP retrieval algorithms. Since statistical techniques require less ancillary data and are subject to different algorithm-related errors than physically-based retrievals, NN retrievals may provide useful information where and when RTMs are known to be uncertain. For SMOS, the NN technique has been successfully implemented (*Rodriguez-Fernández et al.*, 2015). However, it is not obvious that a NN for SMAP will work equally well, given the differences between SMOS and SMAP in the observing geometry (multiple vs. single incidence angle) and instrument error characteristics (*De Lannoy et al.*, 2015). Second, we aim to investigate the potential of statistical techniques to generate a soil moisture product with characteristics beneficial to SMAP soil moisture assimilation. The NN algorithm retrieves soil moisture in the climatology of the target model and thus may reduce the need for bias correction prior to data assimilation. In a follow-on study, we will investigate whether this results in a more efficient use of SMAP observations during data assimilation.

The NN retrieval algorithm is trained with SMAP brightness temperatures and two ancillary datasets as inputs, and with target data from the NASA Goddard Earth Observing System version 5 (GEOS-5) model (section 2). Using the trained NN, we compute global estimates of volumetric surface soil moisture for the period April 2015 to March 2017 and evaluate them using a number of different metrics and techniques (section 3). We compare the SMAP NN soil moisture estimates to the target GEOS-5 model soil moisture to identify the independent information provided by the SMAP observations that can potentially inform the model during data assimilation (section 4.1). Next, we assess the SMAP NN retrievals against independent in situ measurements and compare their skill to that of the SMAP Level-2 passive (L2P) retrieval product and the GEOS-5 model soil moisture (section 4.2). Finally, we assess the global error distributions of the SMAP NN, GEOS-5 and SMAP L2P products using a triple collocation (TC) analysis in conjunction with soil moisture retrievals based on observations from the Advanced Microwave Scanning Radiometer 2 (AMSR2) and the Advanced Scatterometer (ASCAT), which have independent errors with respect to the SMAP and GEOS-5 products (section 4.3).

2. Datasets

2.1. Neural Network Inputs and Target Datasets

2.1.1. SMAP Observations

The main input to the NN soil moisture retrieval algorithm are the SMAP brightness temperatures. SMAP was launched in January 2015 and is equipped with an L-band (1.4 GHz) radiometer observing on four different channels, horizontal and vertical polarization as well as the 3rd and 4th Stokes' parameter. SMAP is in a sun-synchronous, near-circular orbit with equator crossings at 6 AM and 6 PM local time and a revisit time of 2-3 days (*Entekhabi et al.*, 2010). Brightness temperature data have been collected since 31 March 2015.

For our NN retrieval product we use SMAP Level-1C brightness temperatures (*Chan et al.*, 2016a) for the April 2015 to March 2017 period. The data are provided on the 36-km resolution Equal-Area Scalable Earth version 2 (EASEv2) grid (*Brodzik et al.*, 2012) as daily half-orbit files. We only use observations from the 6 AM overpass, in order to minimize observation errors due to Faraday rotation and the difference between the soil and canopy temperatures (*Entekhabi et al.*, 2010; *O'Neill et al.*, 2015). A test of different input combinations indicated that using data from all four SMAP channels as inputs to the retrieval algorithm yielded the best NN retrieval performance (not shown). While the 3rd and 4th Stokes' parameters are not directly sensitive to soil moisture, including them as inputs helps the NN algorithm to distinguish between different observing conditions and thus determine the weight for a given brightness temperature observation.

2.1.2. GEOS-5 Model Surface Soil Moisture and Temperature

The model soil moisture estimates used here are generated using the GEOS-5 Catchment land surface model (*Koster et al.*, 2000; *Ducharne et al.*, 2000). The Catchment model version used in this study is very similar to that of the SMAP Level-4 Soil Moisture (L4_SM) version 2 system (*Reichle et al.*, 2015, 2016, 2017b *(*in press), but SMAP brightness temperature observations are not assimilated. The surface meteorological forcing data were provided at 0.25° resolution by the GEOS-5 Forward Processing atmospheric data assimilation system (*Lucchesi*, 2013). The GEOS-5 precipitation forcing data were corrected using global, daily, 0.5 ° resolution, gauge-based observations from the Climate Prediction Center Unified (CPCU) product, which have been scaled to the Global Precipitation Climatology Project (GPCP) v2.2 pentad precipitation product climatology (*Reichle and Liu*, 2014; *Reichle et al.*, 2017a,b). The GEOS-5 background precipitation was also scaled to the GPCP v2.2 climatology. Output fields were produced as 3-hourly time averages and provided on the 9-km EASEv2 grid.

In this study, we use two model output fields: (1) the surface soil moisture (0-5 cm soil layer) and (2) the surface soil temperature (0-10 cm soil layer). The GEOS-5 soil moisture fields served as target data in the NN training (section 3.1) and were also used in the evaluation phase to assess the skill of the NN retrieval compared to that of the target model. The surface soil temperature data were used as an input to the retrieval algorithm to account for the surface soil temperature contribution to the observed brightness temperatures (section 3.1). Using surface soil temperature estimates from the target model potentially introduces some of the GEOS-5 spatial patterns into the NN estimates and could lead to model dependency issues during a later assimilation of the NN estimates into the GEOS-5 model. The same would be true, however, for the assimilation of the SMAP L2P product, which also uses GEOS-5 ancillary soil temperatures (section 2.2.1). We assume here that the canopy temperature and surface soil temperature are in equilibrium for the 6 AM (local time) SMAP observations used here, so only a single temperature estimate is required. The surface soil temperature data were also used in the data quality control to identify frozen soil conditions (section 2.3).

2.2. Validation Datasets

2.2.1. SMAP Level-2 Passive Retrievals

The SMAP L2P soil moisture retrieval product uses SMAP radiometer Level-1C brightness temperatures to provide soil moisture estimates on the 36-km EASEv2 grid as daily half-orbit files. The retrieval algorithm is based on a physical tau-omega model (*Wigneron et al.*, 1995; *O'Neill et al.*, 2015) to isolate the soil emission from the total observed surface emission (soil and vegetation) and to subsequently convert it into a soil moisture estimate through the use of soil emission and mixing models. The surface soil temperature data required by the tau-omega model are provided by the quasi-operational GEOS-5 Forward

Processing system (*Lucchesi*, 2013) with a 0.25° resolution. The tau-omega model also requires information on the vegetation water content (VWC), which is estimated from a climatology of the Normalized Difference Vegetation Index based on Moderate Resolution Imaging Spectroradiometer (MODIS) observations using an empirical relationship established from prior investigations. No retrieval is performed for frozen soil conditions based on GEOS-5 surface soil temperature. Soil moisture retrievals are flagged as 'not recommended' when the VWC within the satellite footprint exceeds 5 kg m^{-2} (*O'Neill et al.*, 2015).

In this study, we use version 4 of the L2P 'baseline' soil moisture estimates derived from the SMAP morning (6 AM) overpass vertical polarization brightness temperatures (*O'Neill et al.*, 2016). Only data points flagged as having the 'recommended' retrieval quality were used.

2.2.2. AMSR2 and ASCAT Soil Moisture Retrievals

The Advanced Multichannel Scanning Radiometer 2 (AMSR2) is a multichannel passive microwave satellite instrument that has been collecting data since July 2012. AMSR2 measures brightness temperatures at frequencies ranging from 6.9 GHz to 89 GHz with a revisit time of approximately 2 days and equator crossings at 1.30 AM and 1.30 PM local time (*Kasahara et al.*, 2012).

Here we use the Japan Aerospace Exploration Agency AMSR2 soil moisture product computed from the 10.7 GHz and 36.5 GHz vertical and horizontal polarization brightness temperatures (*Maeda and Taniguchi*, 2013). The data are provided as daily estimates of volumetric surface soil moisture on a grid with 0.1° × 0.1° resolution spacing.

The Advanced Scatterometer (ASCAT) (*Figa-Saldaña et al.*, 2002) is an active microwave satellite instrument aboard the MetOp satellites, which have been collecting data since 2006. ASCAT measures surface backscatter at C-band (5.3 GHz) with a revisit time of 1-2 days and equator crossings at 9.30 AM and 9.30 PM.

Here we use the ASCAT surface soil moisture product developed by *Wagner et al.* (2013). The data are provided in units of surface degree of saturation with a sampling distance of 12.5 x 12.5 km and were converted into estimates of volumetric surface soil moisture using the soil porosity data of *Reynolds et al.* (2000).

Despite being posted on finer resolution grids, the spatial resolution of the AMSR2 and ASCAT observations is very similar to the SMAP 36-km resolution.

2.2.3. In Situ Measurements

SMAP Core Validation Sites. The SMAP core validation sites (referred to here as 'core sites') represent locally dense networks of in situ soil moisture measurements that are specifically designed for the calibration and validation of SMAP soil moisture products (*Colliander et al.*, 2017). Each site features an array of soil moisture sensors to represent the different spatial scales of the SMAP products (3 km, 9 km and 36 km). The measurements from each site's sensors are combined into and area-weighted average to yield one soil moisture time series per site that is representative of a 36-km satellite grid cell.

Table 1: Overview of the SMAP Cal/Val core sites. Shown are (from left to right) the site name, reference pixel ID (RPID), location, climate, land cover and the average number of sensors that contribute to the reference pixel average. Soil moisture is measured at 5 cm depth or over the top 5 cm. (*Colliander et al.*, 2017)

Site (abbreviation)	**RPID**	**location**	**climate**	**land cover**	**number of sensors**
REMEDHUS (RM)	03013602	Spain	temperate	croplands	14
Reynolds Creek (RC)	04013601	USA (Idaho)	arid	grasslands	5
Yanco (YC)	07013601	Australia	arid	croplands	26
Carman (CM)	09013601	Canada	cold	croplands	19
Twente (TW)	12043602	Holland	temperate	croplands / natural mosaic	9
Walnut Gulch (WG)	16013603	USA (Arizona)	arid	shrub open	20
Little Washita (LW)	16023602	USA (Oklahoma)	temperate	grasslands	16
Fort Cobb (FC)	16033602	USA (Oklahoma)	temperate	grasslands	12
Little River (LR)	16043602	USA (Georgia)	temperate	croplands / natural mosaic	19
South Fork (SF)	16073602	USA (Iowa)	cold	croplands	18
Monte Buey (MB)	19023601	Argentina	temperate	croplands	10
Kenaston (KN)	27013601	Canada	cold	croplands	26
TxSON (TX)	48013601	USA (Texas)	temperate	grasslands	32
Mahasri (MH)	53013601	Mongolia	cold	grasslands	5

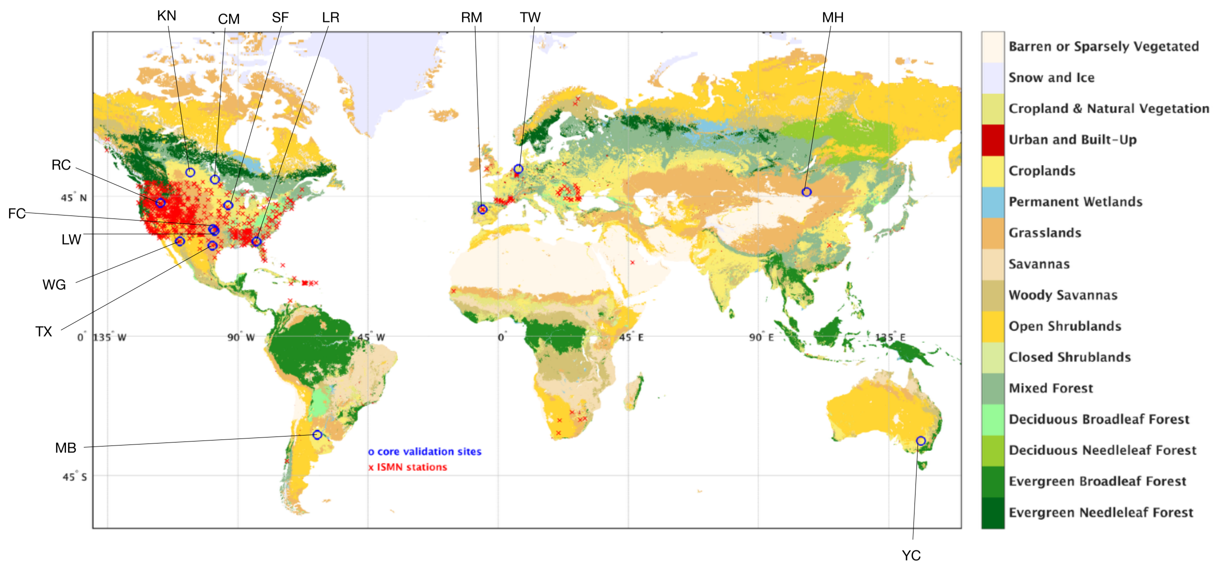

Figure 1: Location of the SMAP core validation sites (blue circles) and ISMN stations (red crosses). The background shows the dominant International Geosphere-Biosphere Program (IGBP, (*Belward et al.*, 1999)) land cover class for each location.

Table 1 summarizes the main characteristics of the 36-km core sites used here. Out of the 14 locations, nine are in North America, two in Europe, and one each in Asia, Australia and South America. The sites represent a range of different climatic conditions and land cover types, and the average number of sensors that contribute to the 36-km reference pixel data ranges between 5 and 32. Figure 1 shows the distribution of the SMAP core sites and their corresponding dominant land cover.

International Soil Moisture Network (ISMN).

Table 2: Overview of the ISMN (*Dorigo et al.*, 2011). Shown are the location, number of stations per network and the network-specific reference.

Network	location	# stations	reference
Dahra	Senegal	1	(*Tagesson et al.*, 2015)
FMI	Finland	27	(*Dorigo et al.*, 2011)
iRON	USA	6	(*Taylor et al.*, 2015)
PBO H2O	USA	161	(*Larson et al.*, 2008)
REMEDHUS	Spain	24	(*Sanchez et al.*, 2012)
RSMN	Romania	20	(*Dorigo et al.*, 2011)
SCAN	USA	181	(*Schaefer et al.*, 2007)
SMOSMANIA	France	21	(*Calvet et al.*, 2007)
SNOTEL	USA	441	(*Leavesley et al.*, 2008)
SOILSCAPE	USA	171	(*Moghaddam et al.*, 2016)
USCRN	USA	115	(*Diamond et al.*, 2013)

We further evaluate the NN retrieval product against independent in situ soil moisture measurements from the International Soil Moisture Network (ISMN), a database of soil moisture networks hosted at the Technical University (TU) of Vienna (*Dorigo et al.*, 2011) and referred to here as the 'sparse networks'. We used only ISMN networks that are not part of the SMAP core sites (Table 2). The REMEDHUS network comprises a different set of sensors for the core site and as a sparse network and thus appears for both in situ data types. The measurement depth, repeat frequency, coverage, station density and measurement method depend on the contributing network. The number of stations in each network ranges between 1 and 441 (Table 2), but - unlike for the core sites - there is typically only one sensor per 36-km grid cell. That is, the ISMN measurements are not necessarily representative of the spatial scale of the satellite observations. Figure 1 shows the spatial distribution of the ISMN stations and the dominant land cover at each location.

For two of the ISMN networks, SCAN (*Schaefer et al.*, 2007) and USCRN (*Diamond et al.*, 2013), the data were already available in-house and had been subjected to additional quality control as described in *De Lannoy et al.* (2014) and (*Reichle et al.*, 2015b) (their Appendix C). Hence, the in-house data were used for SCAN and USCRN instead of the data provided through the ISMN. As a result, more reliable metrics could be estimated for these two sparse networks.

2.3. Data Preprocessing

2.3.1. Satellite Observations and Model

We co-located all datasets spatially and temporally, using the 36-km EASEv2 grid and the SMAP morning (6 AM) overpass times as a reference. The GEOS-5, AMSR2 and ASCAT data were aggregated from their higher-resolution native grids to the 36-km EASEv2 grid using simple averaging. The temporal co-location was implemented by using the GEOS-5 3-hourly average that includes the SMAP morning overpass for a given location and day. For the AMSR2 and ASCAT retrieval products, only data from their night-time/morning overpasses for the same day - at 1.30 AM and 9.30 AM, respectively - were used since these are closest in time to the SMAP overpass at 6 AM. Likewise, for the L2P retrievals we used only the morning overpass estimates, and no regridding was required because the SMAP-based NN and L2P products are provided on the same 36-km EASEv2 grid.

We additionally applied several quality control steps to the satellite and model data sets to identify and exclude conditions in which a soil moisture retrieval was not feasible. Using the GEOS-5 surface soil temperature, we excluded times and locations with a surface soil temperature below 1°C. The MODIS-based VWC estimates provided with the L2P data were used to exclude times and locations with a VWC higher than 5 kg m^{-2}, where the SMAP radiometer is not expected to provide reliable retrievals. Finally, we excluded all pixels within 72 km of a water body - defined as a grid cell with a water fraction in excess of 5% according to the GEOS-5 land mask - to mitigate the impact of water bodies, as their low brightness temperatures cause erroneously high soil moisture retrievals (*O'Neill et al.*, 2015).

2.3.2. In Situ Data

The core site measurements are representative of the 36-km spatial resolution of the retrievals and the aggregated model, however, the geographical center of the in situ sensors for a given reference pixel does not generally coincide with the EASEv2 grid cell center of the satellite and model products. Similarly, the location of a (single point) ISMN measurement is typically offset from the center of a EASEv2 grid cell. To account for this, the retrieval and (aggregated) model soil moisture were linearly interpolated to the in situ location using data from the nearest EASEv2 grid cell and its 8 surrounding neighbors, requiring a minimum of 4 data points. Where applicable, ISMN measurements located in the same EASEv2 grid cell were averaged and their average location was used for the interpolation. For each day, the in situ measurement closest in time and within a 3 hour window of the SMAP overpass was used.

Using the GEOS-5 surface temperature for the ISMN measurements and the in situ surface soil temperature observations for the core site measurements, the in situ data were screened for (nearly) frozen soil conditions by applying the same 1°C threshold that was used for the satellite and model data.

3. Methodology

3.1. Neural Network Retrieval Algorithm

In this study we use a NN approach to retrieve global surface soil moisture with a 2-3 day repeat using SMAP brightness temperatures, GEOS-5 soil temperatures and the MODIS-based VWC climatology that is used in the generation of the SMAP L2P product. The NN retrieval algorithm is first calibrated (trained) using a subset of the available SMAP and model data to determine the statistical relationship between the satellite observations and surface soil moisture. Once calibrated, the trained NN is used to retrieve surface soil moisture from the entire set of satellite observations.

3.1.1. Neural Network Architecture

A neural network is a group of computational nodes arranged in a layered and inter-connected architecture. Figure 2 shows a schematic of a basic NN for soil moisture retrievals. The NN used here consists of 3 layers: (1) an input layer that receives the satellite observations and ancillary inputs, (2) one hidden layer, and (3) an output layer that produces the soil moisture estimates. This architecture is sufficient to approximate any continuous function (*Cybenko*, 1989). The inputs for the SMAP NN retrieval algorithm are the observations from the four SMAP channels, the GEOS-5 surface soil temperature and the MODIS-based VWC estimates. The output from the NN algorithm is an estimate of the volumetric surface soil moisture.

While the number of neurons in the input and output layers is determined by

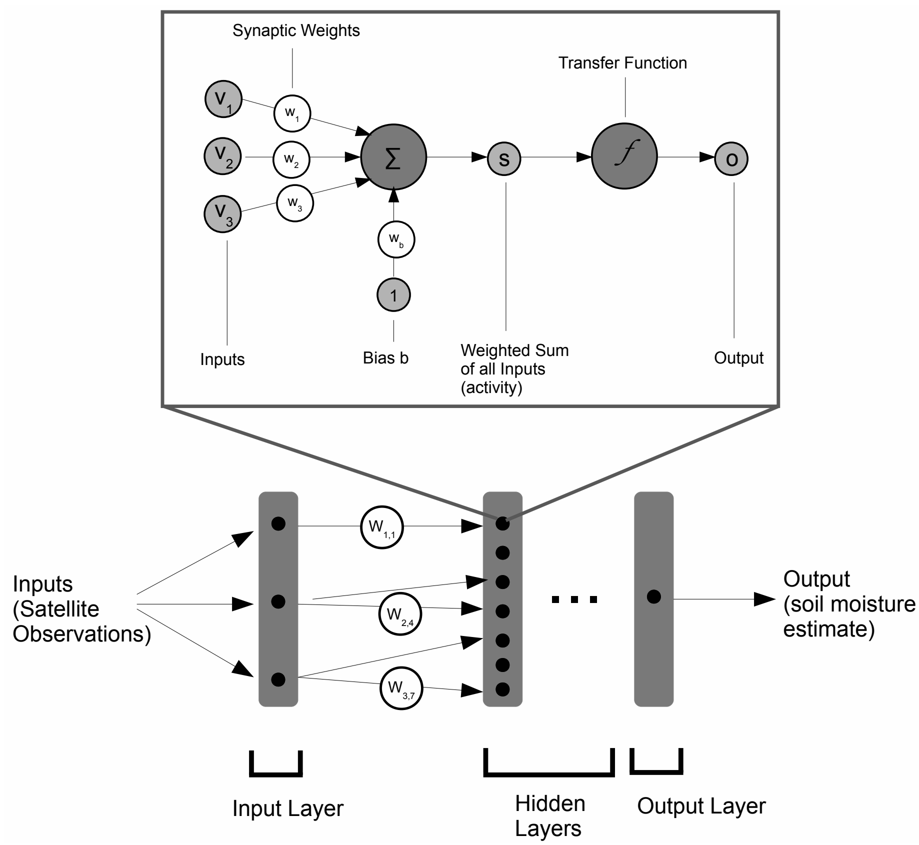

Figure 2: Schematic of a neural network with close-up of a single neuron (adapted from *Kolassa* (2013)).

the number of input and output variables (here, 6 for the input layer and 1 for the output layer), the optimal number of neurons in the hidden layer depends on the problem complexity. We found that for this study 15 hidden layer neurons constituted the lowest number of neurons that was able to converge to a solution during the NN training. We use a fully connected feed-forward network, in which all neurons from one layer are connected to all neurons in the next layer. These connections are assigned weights - the synaptic weights - used by each neuron to compute a weighted sum of all its input plus a bias before applying a transfer function. Neurons in the input and output layers use a linear transfer function, while hidden layer neurons use the typical tangent-sigmoid transfer function.

3.1.2. Neural Network Training

In order to determine the statistical function that relates the NN input data, including the satellite brightness temperature observations, to surface soil moisture, the NN is calibrated on a sample set of NN inputs and coincident soil moisture estimates (the target data), together referred to as the training data. This process is referred to as the NN training and is schematically illustrated in Figure 3 (a). To generate a training dataset representative of all expected conditions, we used the first year (April 2015 - March 2016) of our study period for NN training. The second year (April 2016 - March 2017) of the study period was used for the evaluation presented in sections 4.1 and 4.2. Model soil moisture estimates from GEOS-5 are used as the target data, because (1) the model estimates have a similar resolution as the satellite observations while providing complete spatio-temporal coverage and (2) training on a model yields

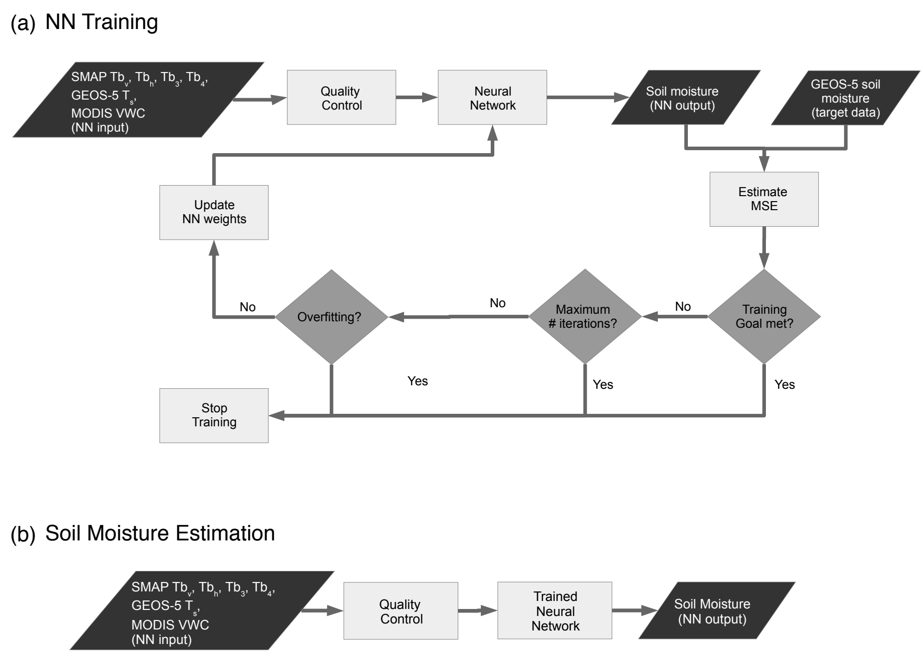

Figure 3: The two phases of the NN soil moisture retrieval approach. (a) NN training and (b) soil moisture estimation using the trained network. NN inputs include the SMAP brightness temperatures at vertical and horizontal polarization (Tb_v and Tb_h), the 3rd and 4th Stokes' parameters (Tb_3 and Tb_4), the GEOS-5 surface soil temperature (T_s), and the MODIS-based vegetation water content (VWC).

NN estimates in the global model climatology, which could be beneficial for a later assimilation of the retrieved soil moisture.

The total training dataset is split into three subsets - the calibration, validation and test data - by sampling the total dataset. The calibration data constitute 60% of the total training data and are used to optimize the NN synaptic weights (Note: In the literature these data are often referred to as 'training data'. In order to avoid confusion with the total training dataset, we have decided to use the term 'calibration data' instead). The validation data constitute 20% of the total training data and are used to detect over-fitting of the NN weights (see below). These are part of the training data and should not be confused with the independent evaluation data used in sections 4.1 and 4.2 to assess the SMAP NN retrieval quality. The test data constitute the remaining 20% of the training data and are used to assess the NN fit.

The NN training is non-localized, meaning that one NN is fitted to a global training dataset that contains data from the entire training period (April 2015 - March 2016). Furthermore, no information regarding the location and acquisition time of the training points is provided to the NN. The NN training thus essentially involves an association of the same sets of input values (that is, the same brightness temperatures, Stokes' parameters, and ancillary data) with the mean value of the corresponding target soil moisture data. If, for example, the target data in a specific region overestimate the soil moisture, they will appear as outliers in the NN training, and the NN will thus not inherit such regional errors (e.g., (*Jimenez et al.*, 2013)). As a result, the spatial and temporal patterns of the NN estimates are mostly driven by the input satellite observations. Moreover, the NN estimates match the global (single-value) mean and variability of the target data, but mean differences in the spatial patterns between the satellite observations and the model estimates are retained. These remaining local biases could represent an issue during an assimilation of the NN product. Further investigation will be needed to determine whether the disadvantage of local biases in the assimilation is outweighed by the benefit of retaining more of the independent information in the assimilated SMAP observations.

The training itself consists of an iterative optimization of the NN synaptic weights to minimize the error between the NN output and the target data (Figure 3 (a)). Three different scenarios cause the NN training to stop. First, the training is stopped when the mean squared error between the NN outputs and the target data is less than 0.001 m^3m^{-3} and the training goal is met. Second, the training is stopped when the NN training does not converge to a solution after a maximum number of iterations - set here at 1000. Third, training is stopped when over-fitting of the NN weights to the calibration data is detected. For this, the error between NN estimates computed from the validation input data and the validation model soil moistures is estimated upon each iteration. A divergence of the validation estimates from the corresponding validation model soil moisture indicates an over-fitting of the NN weights to the calibration data and a loss of the NN's generalization ability. When such a divergence is detected for six subsequent iterations, the training is stopped and the weights from the last iteration before the occurrence of the divergence are used as the final solution.

Here we use a Levenberg-Marquardt training algorithm (*Levenberg*, 1944; *Marquardt*, 1963) and apply an error back-propagation approach (*Rumelhart and Chauvin*, 1995) to update the weights. The Levenberg-Marquardt algorithm stops when a local minimum is found and thus does not permit a full exploration of the error surface. To account for this, the NN training is repeated four times, using a different random initialization for the NN weights (and thus a different starting point on the error surface) each time. This corresponds to four repetitions of the training process illustrated in Figure 3 (a). After the training is stopped, we compute the root mean square error (RMSE) between the NN estimates computed from the test data and the corresponding test model soil moistures to assess the NN fit. The NN with the lowest RMSE error out of the four repetitions is then retained as the optimal NN and used to generate the soil moisture retrieval product.

The trained NN is used to compute global estimates of volumetric soil moisture from the complete set of satellite observations and ancillary data (Figure 3 (b)). The soil moisture estimates are computed for the period April 2015 to March 2017 and include both the training data (first year) and the evaluation data (second year) that was not used in the training phase.

3.2. Evaluation Metrics

As part of the NN retrieval development, we evaluate our retrieval product against in situ soil moisture measurements and assess its fit with respect to the target model. To quantify different aspects of the retrieval product and model skill, we use the correlation R, anomaly correlation R_{anom} and unbiased root mean square error $ubRMSE$. These metrics have been chosen, because they evaluate different aspects of the retrieval products and provide complementary information on the product skill. Additionally, they are well-established for the evaluation of soil moisture retrievals (*Al Bitar et al.*, 2012; *Albergel et al.*, 2013; *Chan et al.*, 2016b; *Colliander et al.*, 2017). The evaluation metrics are computed with respect to the model soil moisture estimates (section 4.1) and in situ measurements (section 4.2).

The correlation (R) estimates the ability to capture soil moisture variations at all time scales and is computed as the Pearson correlation coefficient between the raw soil moisture and reference data time series in each location. The anomaly correlation (R_{anom}) estimates the ability to capture individual drying and wetting events and is computed similarly to the correlation, but using the anomaly time series, with the anomalies defined with respect to the 30-day moving average centered on the current day. The $ubRMSE$ measures the RMSE excluding the bias and is computed after removing the long-term mean from the soil moisture and reference data time series in each location. When assessing the fit between the NN retrieval product and its target model (section 4.1), we use the term unbiased root mean square *difference* ($ubRMSD$) to indicate that the target model is not considered the truth in this case. Rather, the $ubRMSD$ simply aims to identify differences between the observed and modeled soil moistures.

When evaluating the skill of the retrieval and model products against in situ measurements, only data points common to all four datasets (i.e., the NN and L2P retrievals, GEOS-5 model estimates, and in situ measurements) contributed to the metric calculations, with a minimum of 30 data points required. For the evaluation against ISMN data, we report the average metrics across all stations in a network. Following the approach used by *De Lannoy and Reichle* (2016), we employ a k-means clustering to avoid a dominance of areas with a high station density and to obtain realistic confidence intervals. The spatial extent of each cluster is limited to 1° around its center. Additionally, we report average metrics computed across all sites for the evaluation against core site data and across all networks for the evaluation against the ISMN data, applying the same clustering approach.

3.3. Triple Collocation Analysis

The evaluation of the NN retrieval product against in situ observations is limited by the availability of the in situ measurements and thus only covers a limited range of climate regions and land cover types. However, for most applications, and in particular for data assimilation, retrieval error estimates are required for every location. Here, we implement a triple collocation (TC) analysis (*Stoffelen*, 1998; *McColl et al.*, 2014) in order to compute a global map of error estimates for the NN soil moisture product.

Triple collocation resolves the linear relationships between three independent datasets of the same variable (here, soil moisture) in order to estimate the errors in each dataset independent of a reference. It is a localized technique that estimates the errors for all three datasets in each location independently, yielding a map of error estimates. Several studies have successfully applied TC to estimate soil moisture retrieval errors (e.g., *Scipal et al.* (2008); *Draper et al.* (2013); *Su et al.* (2014); *Chen et al.* (2016)). Here, we use TC to estimate the NN retrieval product errors and, for comparison, the errors of the GEOS-5 model and L2P soil moisture. However, one of the main assumptions of the TC analysis is an independence of the errors in the three datasets that constitute the triplet. In the case of the NN, GEOS-5 and L2P products this assumption cannot be made, since the NN uses information from the GEOS-5 model while the NN and L2P retrievals rely on the same satellite input data. We therefore use the independent soil moisture retrieval products from AMSR2 and ASCAT (section 2.2.2) to create three suitable triplets: [SMAP NN, AMSR2, ASCAT], [GEOS-5, AMSR2, ASCAT] and [SMAP L2P, AMSR2, ASCAT]. This allows us to derive error estimates for SMAP NN, GEOS-5 and SMAP L2P.

Following *McColl et al.* (2014) and *Draper et al.* (2013), we apply the extended TC to the anomaly soil moisture time series (section 3.2) and compute an error estimate in each location with at least 10 common data points in the three contributing datasets. A bootstrapping approach with 100 samples is applied to ensure a robust error estimation. To mitigate the error dependence on the (product- and location-specific) soil moisture variability, we estimate the fractional error standard deviation (*Draper et al.*, 2013; *Gruber et al.*, 2016), defined here as the error standard deviation divided by the soil moisture standard deviation of the corresponding product in each location. The fractional error standard deviation is an approximation of the noise-to-signal ratio, with values below 1 indicating that the noise is smaller than the signal and values greater than 1 indicating that the noise exceeds the signal.

4. Results and Discussion

4.1. Neural Network Fit

As a first assessment, we compare the NN soil moisture estimates to the GEOS-5 modeled soil moisture used as the target data. The purpose of this is to (1) assess the NN fit with respect to the target data over the training period, (2) evaluate the NN's ability to generalize beyond the training data and (3) identify areas of disagreement between the SMAP driven NN estimates and the model soil moisture. In such areas, an assimilation of the NN retrievals should result in the largest changes to the model.

Over the training period, the domain average ubRMSD, correlation and anomaly correlation between the NN and GEOS-5 soil moistures are 0.037 m^3m^{-3}, 0.60 and 0.53, respectively. These fit values are typical for daily NN soil moisture retrievals (for example (*Kolassa et al.*, 2016)). For the NN training it is not desirable to obtain a perfect fit with respect to the target data, since the non-localized calibration results in spatial and temporal patterns that are driven by the satellite input observations and are thus expected to differ from patterns in the target data (*Jimenez et al.*, 2013). Nevertheless, the fairly high correlation and low ubRMSD values indicate that the SMAP based NN soil moisture corresponds with the estimates generated by the model in most regions.

To assess the NN's ability to generalize beyond the training dataset and to investigate the spatial distribution of the differences between the NN estimates and the GEOS-5 soil moisture, we also compared both datasets over the evaluation period, i.e., using only data points that were not part of the training dataset. Figure 4 shows maps of the ubRMSD, correlation, anomaly correlation and bias between the NN estimates and the model soil moisture. Averaging across these maps yields a ubRMSD, correlation and anomaly correlation of 0.037 m^3m^{-3}, 0.61 and 0.55, respectively, which are similar to the average metrics obtained for the training period and indicate that the NN is able to generalize beyond the training dataset.

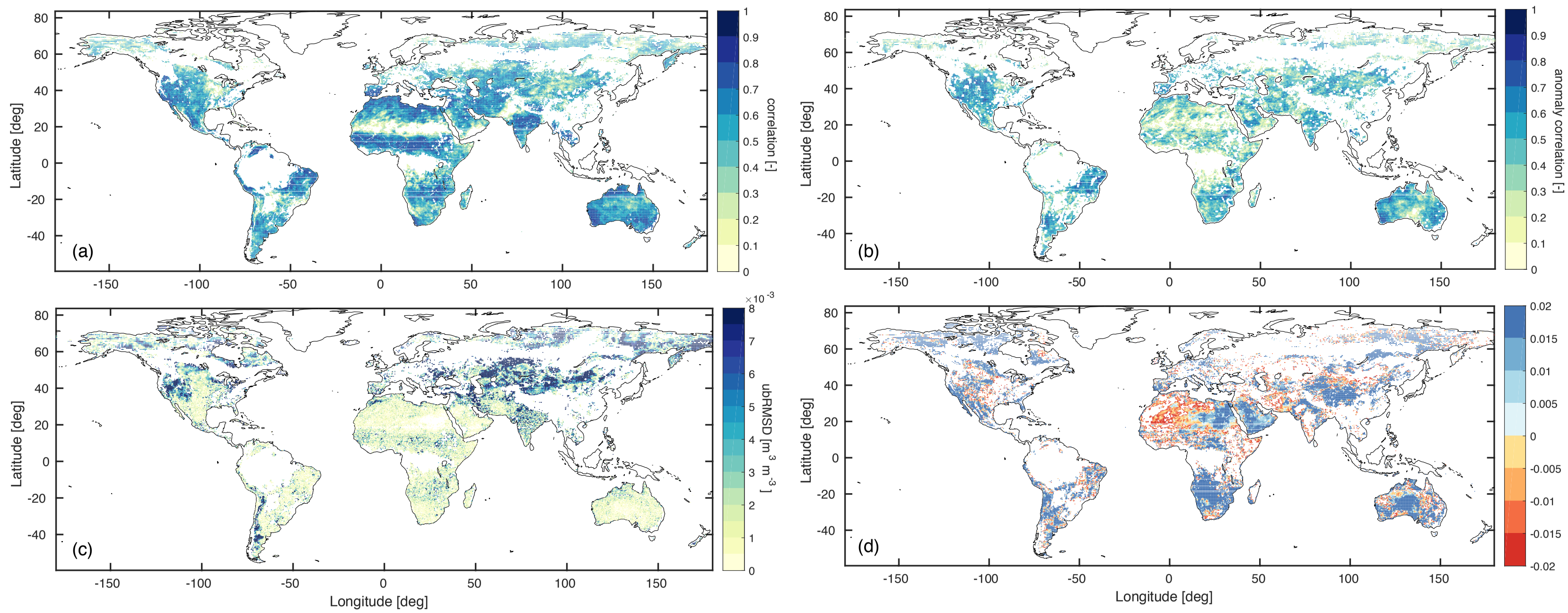

Figure 4: (a) Correlation, (b) anomaly correlation, (c) ubRMSD and (d) bias between the SMAP NN soil moisture retrievals and the GEOS-5 model soil moisture for data points from the evaluation period (April 2016 - March 2017). White spaces indicate areas with less than 30 data points, for which no robust metric could be computed.

The correlations (Figure 4 (a)) and anomaly correlations (Figure 4 (b)) exhibit similar spatial patterns, with high values in the transition zones between wet and dry climates and in regions with strong soil moisture variability, such as the Sahel, Eastern Brazil and India. However, strong correlations and anomaly correlations are also observed in semi-arid, sparsely to moderately vegetated regions, such as the Western US, the Arabian Peninsula and large parts of Australia. The (anomaly) correlations are lowest in arid regions (e.g., the Sahara and Central Australia), where the soil moisture signal tends to be small and noisy, as well as in extensive cropland regions (e.g., the US corn belt or the croplands of Argentina, Uruguay and Paraguay), where irrigation and other agricultural practices are likely to cause differences between the satellite retrieval product and the model.

The spatial patterns of the ubRMSD between the SMAP NN estimates and the GEOS-5 soil moisture (Figure 4 (c)) are different from those observed for the (anomaly) correlations, with large portions of the globe showing a ubRMSD of less than 0.001 m^3m^{-3}, including Africa, Australia and large parts of South America (excluding the Andes). Larger differences occur near mountainous regions, such as the Rocky Mountains or the Southern Andes, likely caused by higher uncertainty in the SMAP retrieval product. High-latitude boreal regions, where the data availability is low and the model precipitation forcing is less reliable (*Reichle and Liu*, 2014), also exhibit larger differences between the NN retrieval product and the model. Finally, the ubRMSD between the NN retrievals and the model estimates is large in the croplands of the US as well as Southern Russia and Kazakhstan, which is possibly a result of the missing representation of irrigation and other agricultural practices in the model that is being corrected by the NN.

The bias between the NN estimates and the GEOS-5 model over the training period (Figure 4 (d)) ranges between -0.02 $m^3\ m^{-3}$ and 0.02 $m^3\ m^{-3}$ and, by design, has a global average close to zero. In arid regions such as the Arabian Peninsula, Central Australia or the Kalahari, the NN retrievals tend to indicate wetter conditions than the GEOS-5 model. An exception is the Western Sahara, where the NN retrievals show a dry bias with respect to the GEOS-5 estimates, which might be an artifact of increased surface roughness in this region that lowers the observed soil emissivity.

In order to illustrate the behavior of the NN retrievals relative to the GEOS-5 model soil moisture in the training and evaluation periods, Figure 5 shows the anomaly time series with respect to a 30-day moving average of the NN soil moisture estimates (red squares) and the GEOS-5 model soil moisture (blue diamonds) for three SMAP core site stations - TxSON, Walnut Gulch and Carman. (The figure also shows the in situ and L2P data, which will be discussed in section 4.2.1). For better readability and to reduce the effect of seasonal differences, we only plot the months April - September for 2015 and 2016 to represent the training and evaluation periods, respectively, with the former indicated through gray background shading. There is no obvious difference between the behavior of the NN retrieval product in both periods, underlining once more the ability of the trained NN to generalize beyond the training dataset. For the TxSON

(Figure 5 (a)) and Walnut Gulch (Figure 5 (b)) sites, the time series average and dynamic range of the NN retrieval product and the GEOS-5 soil moisture are comparable, but there are differences in the response to individual events, illustrated for instance during the stronger drying in the NN soil moisture at the TxSON station in June 2015. At the Carman site (Figure 5 (c)), the NN soil moisture has a stronger variability compared to the model. This illustrates that while the NN estimates globally match the bias and variability of the target data, local biases and differences in variability between the NN estimates and the target data occur.

4.2. Evaluation against In Situ Observations

In this section, we evaluate the skill of the NN retrieval product against independent in situ soil moisture measurements from the SMAP core sites and the ISMN (section 2.2.3). The skill of the NN retrievals is compared against that of the GEOS-5 model soil moisture and the L2P retrievals. Only data from the period April 2016 - March 2017 are used in the evaluation, since these data did not contribute to the NN training.

4.2.1. Core Site Data

First, we assess the skill of the soil moisture products against core site in situ measurements. The NN retrieval product has a higher correlation than the GEOS-5 soil moisture for most core sites (Figure 6(a)), which is reflected in the higher average correlation of 0.70 for the NN retrievals compared to 0.64 for the model. The model has higher correlations than both retrieval products

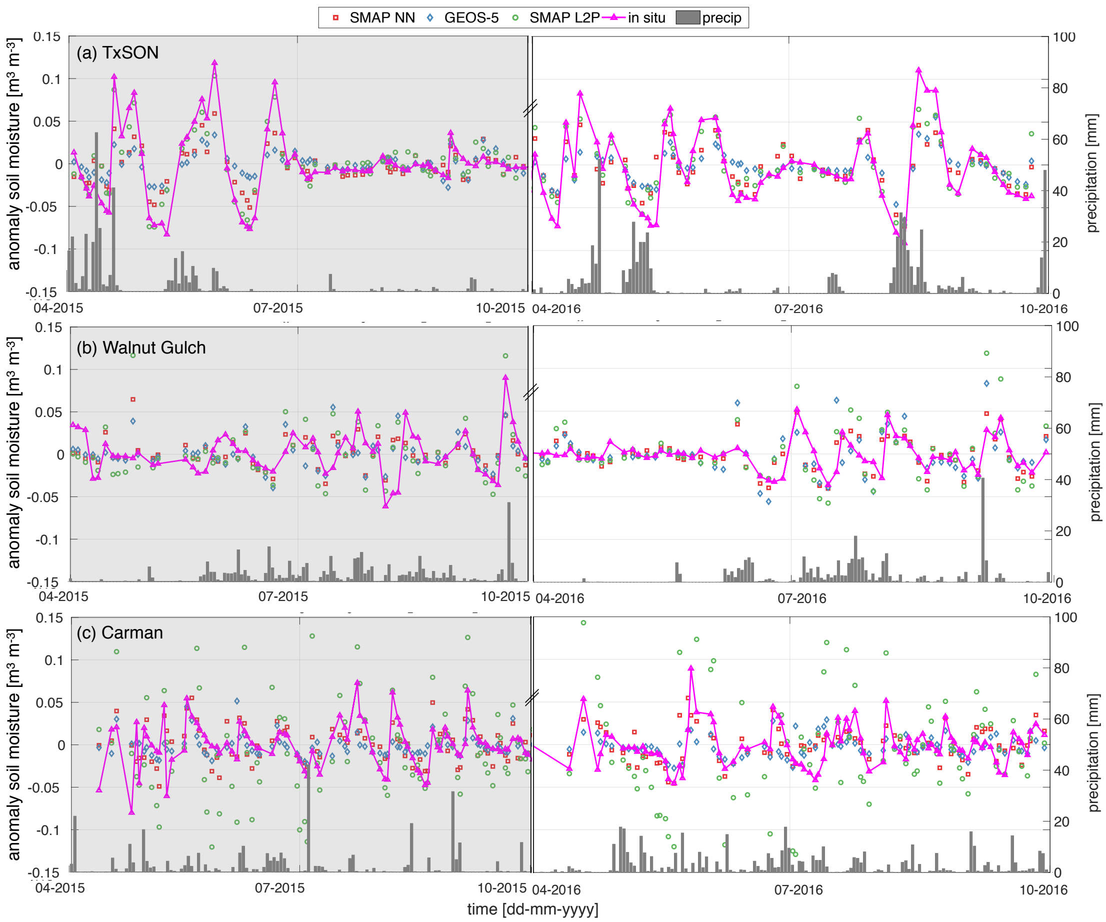

Figure 5: Soil moisture anomalies with respect to a 30-day moving average at the (a) TxSON, (b) Walnut Gulch and (c) Carman core sites for April - September of 2015 and 2016. Shown are the SMAP NN retrievals (red squares), the GEOS-5 model soil moisture (blue diamonds), the SMAP L2P retrievals (green circles) and the core site in situ soil moisture measurements (magenta triangles). Gray bars indicate the corrected GEOS-5 precipitation (section 2.1.2) interpolated to the ground station site. The gray background shading indicates data belonging to the training period.

at Reynolds Creek and a higher correlation than the NN retrievals at Carman and Kenaston. The Carman and Kenaston watersheds are both located at high latitudes where an incomplete seasonal cycle due to frozen soil filtering could prevent the NN from accurately learning the SM-Tb relationship for such conditions in the training phase. The NN retrievals tend to have a notably higher skill than the model in moderately vegetated regions, such as the shrub- and grassland sites of Little Washita or TxSON, as well as at most of the sites characterized by an arid climate (see Table 1). However, while the results appear to connect the relative performances of the NN product and model with climate and land cover characteristics, more sites would be required to draw a firm conclusion. The poor performance of the model at the South Fork site is partly due to agricultural tile drainage, which is not accounted for in the model.

The L2P retrieval product has a higher correlation skill than both of the other soil moisture products for the majority of core sites and consequently has the highest average correlation of 0.78 (Figure 6(a)). The magnitude of the skill difference between the two retrieval products is not obviously related to the climate or land cover of the in situ sites. In regions with a moderate to strong seasonal cycle, the correlation (R) primarily reflects the skill of capturing seasonal soil moisture variations. Hence, the above results indicate a better representation of the soil moisture seasonal cycle in the two retrieval products compared to the model.

In terms of the anomaly correlations (Figure 6(b)), the NN retrieval product has higher skill than the model estimates for most core sites and an average

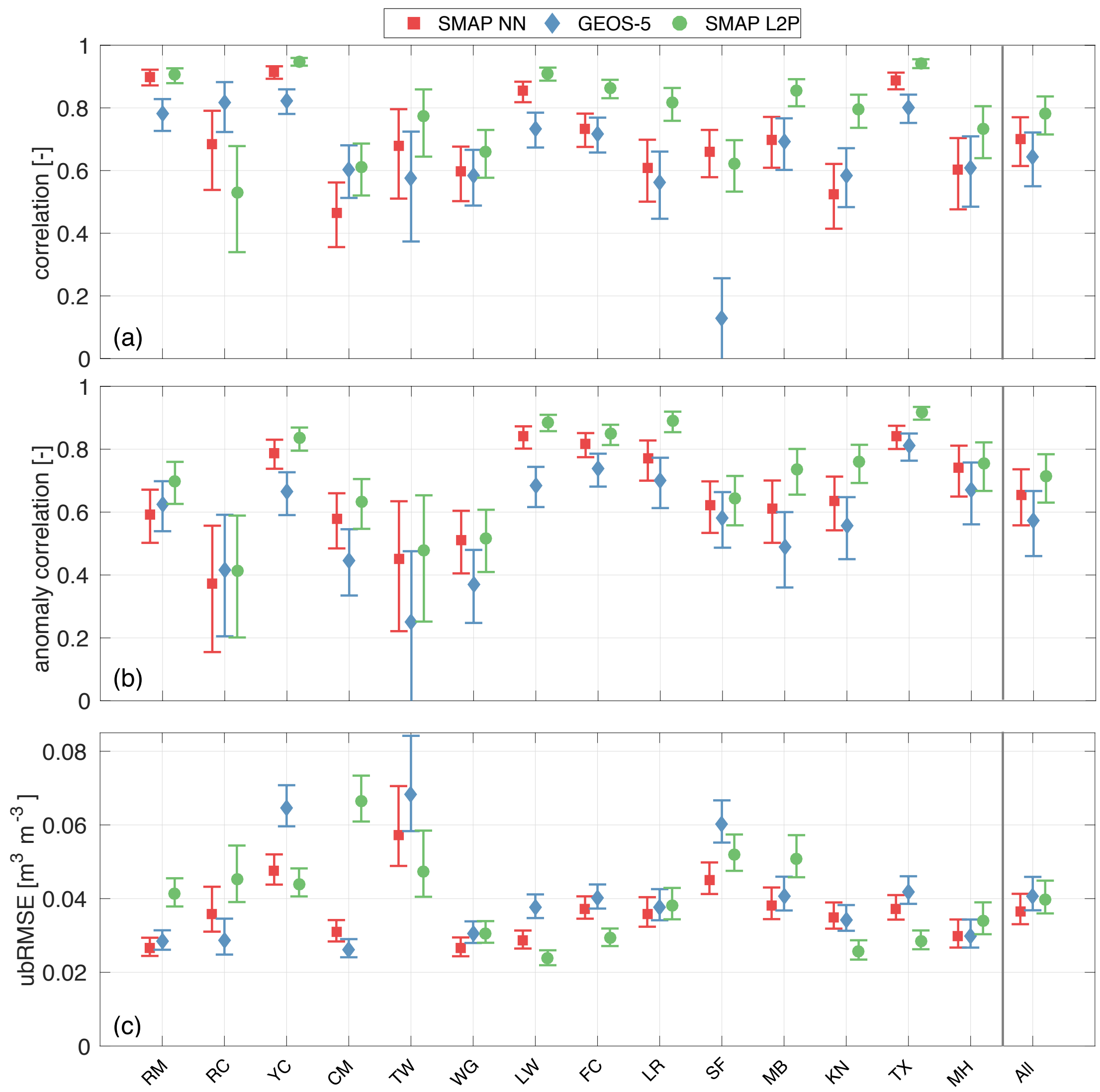

Figure 6: (a) Correlation, (b) anomaly correlation and (c) ubRMSE between the core site in situ measurements and the SMAP NN retrievals (red squares), the SMAP L2P retrievals (green circles) and the GEOS-5 model soil moisture (blue diamonds). Shown are the metrics for each site as well as the average across all sites. The error bars represent the 95% confidence interval.

skill of 0.66 compared to 0.57 for the model. The L2P retrieval product has the highest average skill overall (0.71) as well as for a majority of the core sites. In terms of the ubRMSE (Figure 6(c)), the skill of all three products is more similar. The NN product has a somewhat lower error than the L2P product at a majority of the stations and an overall lower average error of 0.037 m^3m^{-3} compared to 0.041 m^3m^{-3} for the L2P and GEOS-5 model estimates.

Our findings for the L2P skill are consistent (within error bars) with those of *Colliander et al.* (2017) (not shown). The only significant difference occurs at the Twente site, where *Colliander et al.* (2017) used a different set of sensors. Compared to *Chan et al.* (2016b), we obtain higher correlations and a slightly larger ubRMSE for the L2P product. This is in part a result of the more refined validation approach used by *Chan et al.* (2016b), who generated special L2P retrievals on custom grid cells that better match the locations of the in situ measurements and thus did not perform the spatial interpolation that was required for the published L2P retrievals used here (section 3.2). Other factors contributing to the differences in the L2P metrics are the different validation periods and L2P product versions used here and by *Chan et al.* (2016b).

To further investigate the cause for the skill differences between the retrieval products and the model at select sites, we now revisit Figure 5. At the TxSON and Walnut Gulch sites the anomalies for both retrieval products follow the in situ measurements very closely. The different average anomaly correlations obtained for these sites are mostly due to different responses to isolated events. An example is the dry down in June 2015 at the TxSON site, which is better captured by the L2P retrievals than by the NN retrievals.

At the Carman site, both retrieval products are very noisy compared to the model and in situ measurements (Figure 5(c)). The L2P product is noisier than the NN product, which is also reflected in its higher ubRMSE at this site (see Figure 6(c)). The higher ubRMSE might be caused by ancillary soil texture data in the L2P retrieval algorithm that poorly describes the highly variable conditions in the Carman watershed. This suggests that the NN retrieval approach has the potential to supplement the physically-based SMAP retrievals in regions where the ancillary data used in the RTM are uncertain. Additionally, both retrieval products suffer from using a VWC climatology that does not accurately describe the rapidly changing vegetation dynamics at Carman.

The above results show that both SMAP retrieval products have higher correlations than the model soil moisture with respect to the in situ measurements (Figure 6). This is encouraging, given that most of the core sites are located in North America, where models typically have been well tested and already have a high skill (e.g. *Albergel et al.* (2013)). Additionally, the retrievals are at a slight disadvantage in the comparison, since for most locations the SMAP emission depth will be less than the 5 cm depth represented by the in situ measurements and the model estimates. The better correlations of the retrieval products thus illustrate the high quality of the SMAP observations and their potential to provide independent information that is not captured in the models, likely related to agricultural practices, land use differences or phenology. This is corroborated by the benefit of the SMAP brightness temperature assimilation performed in the Level-4 soil moisture algorithm (*Reichle et al.*, 2017b *(*in press).

Against the core site data, the L2P retrievals generally have a higher skill than the NN retrievals in terms of the correlations and anomaly correlations, while the NN retrievals have a better average ubRMSE (Figure 6). This behavior could indicate the existence of a conditional bias in the SMAP NN retrievals, as a result of dynamic range reduction that is typical for statistical techniques (e.g. (*Kolassa et al.*, 2013)). The global average of the anomaly soil moisture temporal standard deviations for the SMAP NN retrievals, the SMAP L2P retrievals and the GEOS-5 estimates are 0.020 m^3 m^{-3}, 0.036 m^3m^{-3} and 0.015 m^3m^{-3}, respectively, suggesting that the lower dynamic range of the NN retrievals compared to the L2P retrievals is driven by the lower dynamic range of the model. At the core sites in Figure 5, the NN estimates appear to better match to dynamic range of the in situ measurements than the L2P retrievals, however, the limited number of core validation sites does not permit conclusions regarding the general suitability of the retrieval products' dynamic range.

A notable exception from the typical relative skill ranking is the Reynolds Creek site, where the NN retrievals have a significantly higher skill than the L2P retrievals in terms of the correlations and ubRMSE. Since the retrieval inputs are very similar for both products, the skill difference is likely caused by uncertainties in the ancillary data used by the L2P algorithm (for example the soil texture or roughness).

From the NN retrieval perspective, differences in the core site correlation skill between the NN and L2P retrievals can be caused by (1) errors in the target data, (2) errors in the satellite input data or (3) missing information in the NN inputs. The first two error sources affect the quality of the NN fit, whereas the latter would prevent the NN from capturing the full range of soil moisture variability. Errors in the SMAP observations would affect both of the retrieval products, such that target data errors or missing input information are more likely causes for the slightly lower NN retrieval correlations against the core site measurements. The results indicate that for the purpose of generating a 'stand-alone' soil moisture retrieval product, the L2P retrieval algorithm is slightly more suitable than the NN approach. However, our findings also demonstrated the potential of the NN retrievals to supplement the physically-based approaches in regions where the ancillary data or RTM parameterization is uncertain. The core site results also show that the NN retrievals are of sufficient quality to warrant further study into their assimilation as motivated above.

4.2.2. International Soil Moisture Network

Next, we analyze the NN retrieval skill against in situ measurements from the ISMN. While these are single point measurements and thus less suitable than the core site data for evaluating satellite retrievals, they are more numerous and are available for a greater variety of climate and land cover conditions. As before, we also estimate the skill of the L2P retrievals and the GEOS-5 model estimates against the ISMN data for comparison.

In contrast to the evaluation against the core site data, the correlation skill of the three soil moisture products against the ISMN measurements is more similar, with an average correlation of 0.52 for the GEOS-5 model, 0.58 for the

NN retrievals and 0.56 for the L2P retrievals (Figure 7(a)). This suggests that at the ISMN sites the NN retrievals slightly better capture the soil moisture seasonal variations. However, the lower correlations compared to the core site evaluation also illustrate that the single-sensor measurements of the ISMN less adequately represent the retrieval and model spatial scales.

To further interpret the correlation differences between the three products, Figure 8 maps the ranking of the three datasets, with the marker at each ISMN site indicating the dataset with the highest skill. For better readability we only plotted sites located in the contiguous US (i.e., iRON, PBO H2O, SCAN, SNOTEL, SOILSCAPE and USCRN), which constitutes the majority of sites used in this study. A large part of the ISMN stations where a skill assessment was possible are located in the Western US, as the screening for dense vegetation reduces the data availability in the Eastern US below the threshold for computing a skill metric.

The model shows the highest correlation skill at many of the stations located in or near the Rocky Mountains (Figure 8 (a)). In mountainous and rough terrain the microwave retrievals are less reliable, because of the increased surface roughness at the instrument footprint scale (*Schmugge et al.*, 1980). Furthermore, the screening for frozen soil removes a large part of the SMAP time series and reduces the retrieval algorithm's ability to correctly capture the soil moisture seasonal cycle in the training phase. In flatter regions away from the mountains, such as the Central Valley, Arizona, South East New Mexico or North Dakota, the retrievals mostly have higher correlations than the

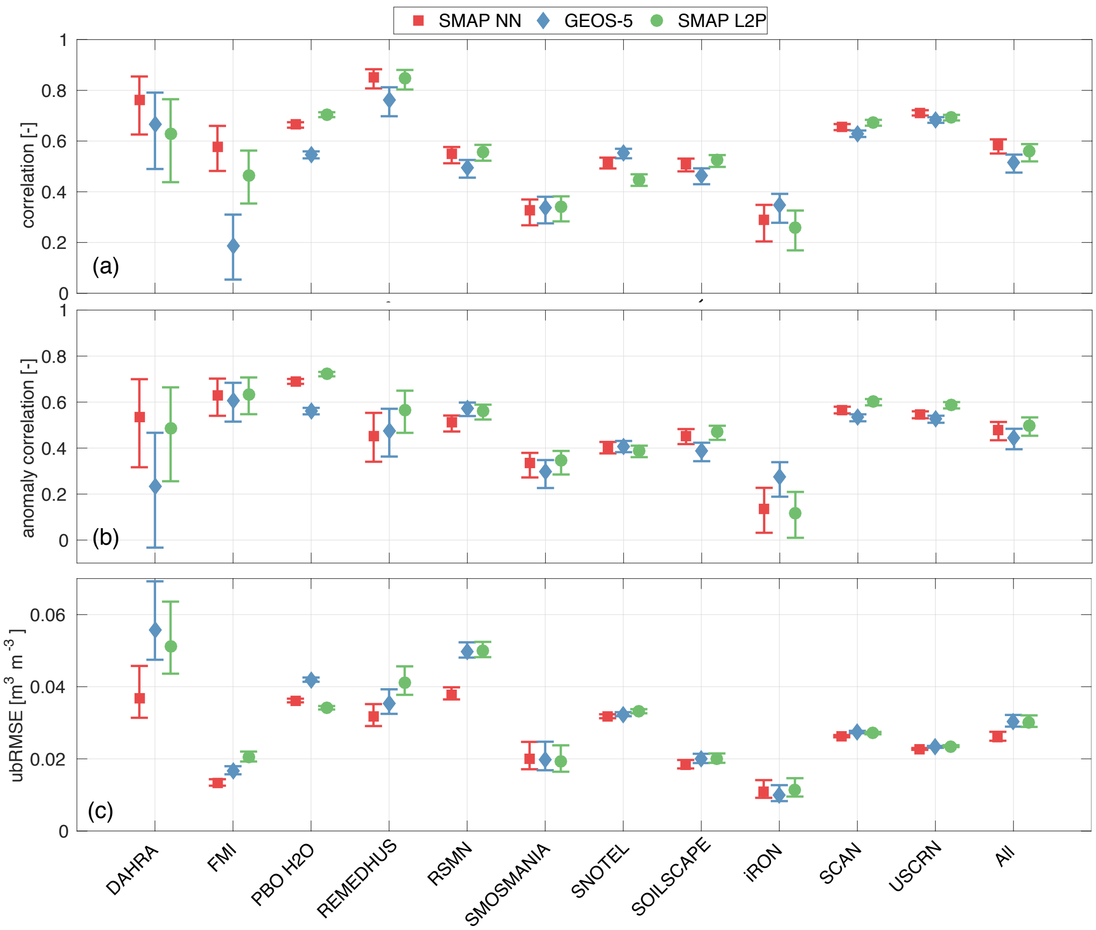

Figure 7: Network average (a) correlation, (b) anomaly correlation and (c) ubRMSE between the ISMN in situ observations and the SMAP NN retrievals (red squares), the SMAP L2P retrievals (green circles) and the GEOS-5 model soil moisture (blue diamonds). Shown are the metrics for each network as well as the average across all networks. All averages are cluster-based (section 3.2) The error bars represent the 95% confidence interval.

model (Figure 8 (a)). Thus, the high station density near the Rocky Mountains slightly skews the average correlation in favor of the model resulting in a model correlation that is comparable to those of the retrieval products. Our clustering approach (section 3.2) mitigates this skew to some extent, but with a cluster spatial extent limited to 1°, we still use a higher number of clusters in the Rocky Mountain region than in other parts of the US. A longer SMAP time series will allow for more correlations to be computed for stations in the Eastern US and would likely lead to different relative correlation skill values for the retrievals and the model estimates.

It is worth noting that our correlation value of 0.65 versus SCAN for the SMAP NN retrievals (Figure 7 (a)) is similar to the 0.61 correlation versus SCAN obtained for SMOS NN retrievals by *Rodriguez-Fernández et al.* (2015). However, it is not possible to draw firm conclusions regarding the relative quality of the SMAP and SMOS NN products, owing to the differences in the validation period and data quality control between *Rodriguez-Fernández et al.* (2015) and our study.

In terms of the network average anomaly correlations (Figure 7(b)), the L2P retrievals have the highest skill with an average anomaly correlation of 0.50 compared to 0.48 and 0.44 for the NN and GEOS-5 products, respectively. Investigating the ranking in terms of the anomaly correlations (Figure 8(b)) shows that the L2P product has the highest anomaly correlation for most of the stations leading to the highest average anomaly correlation.

Finally, the NN retrievals have the lowest average ubRMSE of 0.026 m^3

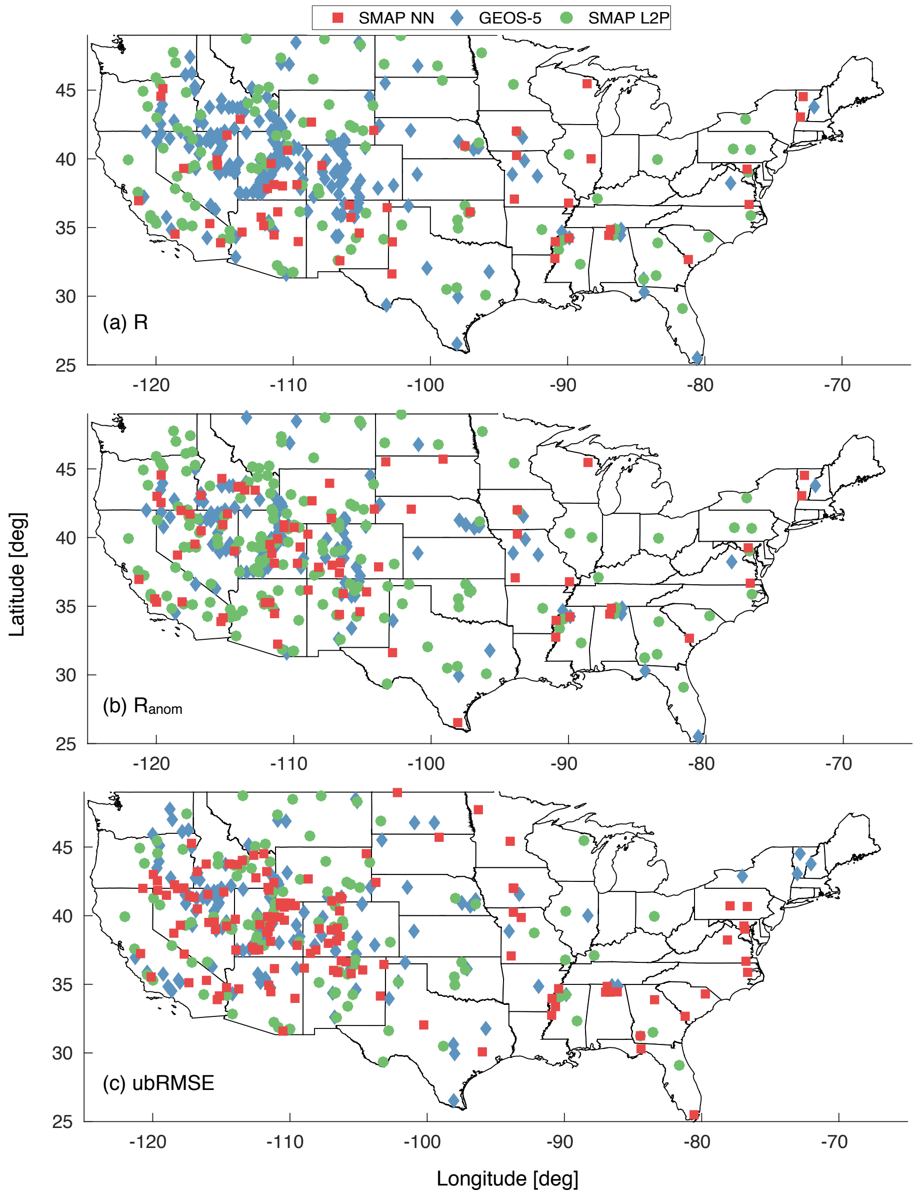

Figure 8: Skill ranking in terms of (a) correlation, (b) anomaly correlation and (c) ubRMSE of the SMAP NN retrievals (red squares), the GEOS-5 model soil moisture (blue diamonds) and the SMAP L2P retrievals (green circles) at the ISMN stations located in the US. Each marker indicates the dataset that obtained the highest skill at a given station.The contributing networks are iRON,PBO H2O, SCAN, SNOTEL, SOILSCAPE and USCRN.

m^{-3} compared to 0.030 m^3 m^{-3} for the L2P retrievals and the GEOS-5 estimates (Figure 7(c)). This relative behavior is largely driven by a significantly lower ubRMSE for the NN retrievals against the DAHRA and RSMN networks. Across all stations, the ubRMSE ranking of the three products in Figure 8(c) is fairly evenly distributed. This also indicates that the lower network average errors observed for the L2P product are not consistent, but driven by a few stations with a low L2P error.

Overall, the correlations of all three products with respect to the ISMN data are lower than for the comparison against the core site data, owing to the lower representativeness of the ISMN stations compared to the core sites.

4.3. Triple Collocation Analysis

For a global evaluation of the SMAP retrieval products and the GEOS-5 model estimates, we estimate the fractional error standard deviations using the TC analysis (section 3.3).

The fractional error spatial patterns mostly show good agreement across the three soil moisture products (Figure 9), corroborated by the very similar global mean fractional error of ~ 1.1 for all three products. All products have fractional errors higher than 1 in the arid and semi-arid regions of the Sahara, the Tibetan Plateau, Northern Mexico and the Northern Arabian Peninsula, indicating that the noise (even though it is small in absolute terms) dominates the small soil moisture signal here and limits the accuracy of all three products. Other arid and semi-arid regions, however, including most of Australia, Southwest Africa and the Southern Andes, have low fractional errors for all products. This indicates that the local fractional errors are driven by a combination of factors, likely including the mean soil moisture level, the surface roughness, land cover and soil type.

Despite a general similarity of the fractional error spatial patterns of all three soil moisture products, several differences between the retrieval and model error patterns exist. For example, the GEOS-5 estimates have higher errors than the retrieval products in the high latitude boreal regions of Alaska and Eastern Siberia, where the precipitation forcing is less reliable (*Reichle and Liu*, 2014). In contrast, both retrieval products have higher fractional errors than the model in areas surrounding the tropical forests, where a denser vegetation cover limits the canopy penetration of the microwave signal and the higher surface roughness increases the signal noise.

The NN and L2P retrieval products show a generally good agreement of the fractional error spatial patterns, but differences in the absolute values exist (Figure 10). For example, the L2P retrievals tend to have lower fractional errors (or noise-to-signal ratios) in the arid regions of Central Australia, the Kalahari or the Southern Sahara, possibly indicating that the ancillary soil data used by the L2P algorithm allows it to better account for the effect of surface roughness, which can be significant in arid regions. However, this behavior is not observed in other arid areas, such as the Central Sahara or the Arabian Peninsula. The NN retrievals have a lower fractional error in moderately to densely vegetated regions and transition zones, such as India, Central Africa, Eastern Brazil and Northern Australia. This suggests that in these regions, the NN method can

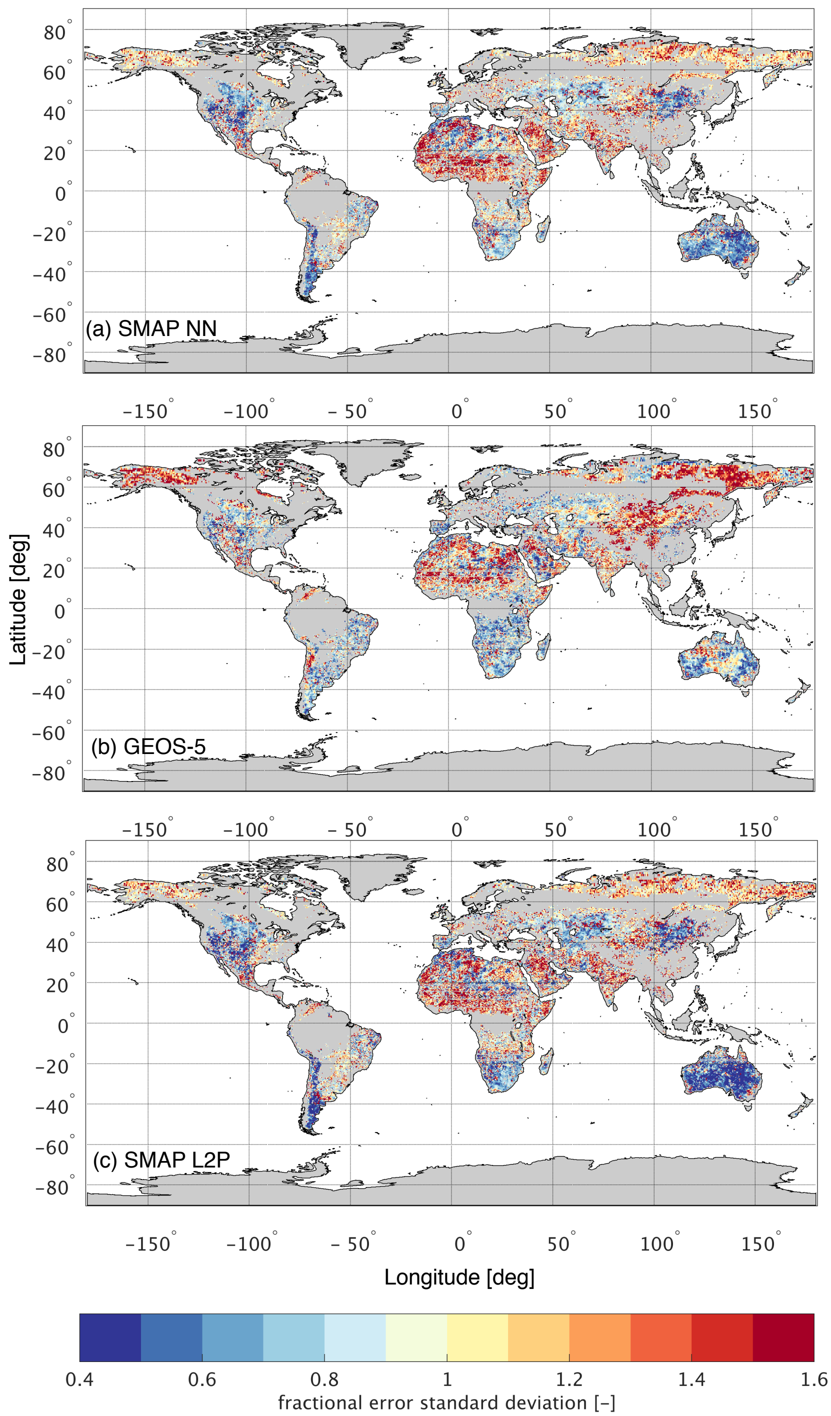

Figure 9: Fractional error standard deviations estimated from TC for the (a) SMAP NN retrieval product, (b) GEOS-5 modeled soil moisture and (c) SMAP L2P retrieval product.

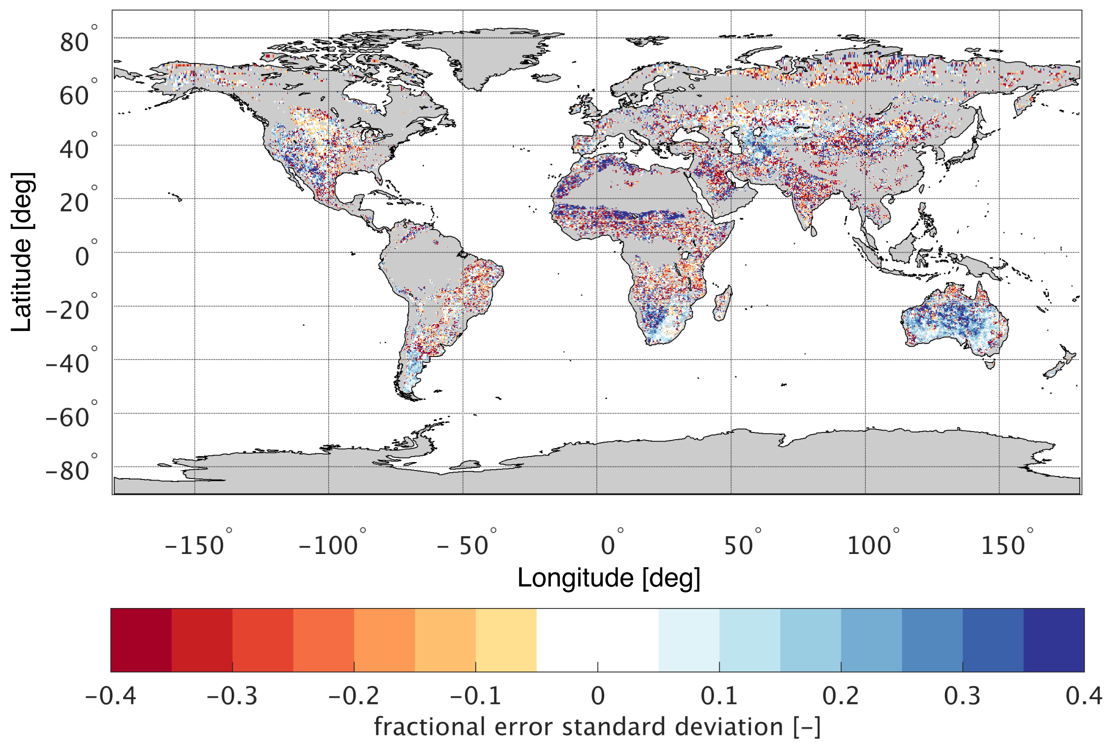

Figure 10: Difference of the fractional error standard deviations between the NN and L2P retrievals (NN - L2P). Negative values (red) indicate a lower fractional error and higher signal-to-noise ratio for the NN retrievals and positive values (blue) indicate a lower fractional error of the L2P retrievals.

produce soil moisture estimates with a higher certainty and could be used to supplement or improve the L2P retrievals. However, due to the lack of in situ stations in these areas, this finding cannot be further corroborated.

While the characterization of the global error distributions is informative, it is important to keep in mind that the error estimates derived from the TC analysis here are also subject to uncertainties. These are related to (1) differences in the overpass times between AMSR2 and ASCAT relative to SMAP and the simulation times of the model, (2) the slightly lower emission depth of the higher frequency AMSR2 and ASCAT data compared to SMAP and the depth of the model's surface layer, and (3) potential errors in the porosity data used to convert the ASCAT data into volumetric surface soil moisture estimates.

5. Summary

In this study we developed and evaluated a NN based retrieval algorithm to estimate global surface soil moisture from SMAP brightness temperatures. The SMAP NN retrieval product was trained on GEOS-5 model estimates and evaluated against in situ measurements from the SMAP core validation sites and the ISMN. The skill of the NN retrieval was compared against that of the GEOS-5 estimates and the SMAP L2P retrievals.

The comparison of the SMAP NN retrieval product against the GEOS-5 model soil moisture showed that globally the two datasets agree well. Differences occur in mountainous regions, where the microwave satellite retrievals are uncertain, and in agricultural areas, where the satellite retrieval product possibly captures the result of agricultural practices (such as irrigation, tilling and harvesting) that are not represented in the model. Combined with the generally higher skill of the SMAP retrievals against in situ measurements, the results confirm the potential for the SMAP observations to inform a model through data assimilation, as has been shown with the SMAP Level-4 products (*Reichle et al.*, 2017b *(*in press).

The SMAP NN soil moisture estimates compare favorably against the SMAP core site in situ measurements with an average correlation and anomaly correlation of 0.70 and 0.66, respectively, and an average ubRMSE of 0.037 m^3m^{-3}. Evaluated against ISMN sparse network in situ measurements, the correlation and anomaly correlation were 0.58 and 0.48, respectively, and the ubRMSE was 0.026 m^3m^{-3}. The core site data better represent the spatial scales of a satellite footprint or model grid cell, leading to the higher skill of the NN retrieval against core site data than against ISMN data.

The NN retrievals had a higher correlation (by 0.06) and a higher anomaly correlation (by 0.09) against core site in situ measurements than the GEOS-5 model estimates, which were used as the NN target data. The corresponding average ubRMSE of the NN retrievals was 0.004 m^3m^{-3} lower than that of the GEOS-5 estimates. Evaluated against ISMN data, the relative skill of the NN retrievals and model estimates was comparable to that found during the core site evaluation.

Overall, the results suggest that (1) the NN retrievals are able to use the SMAP brightness temperatures to correct potential errors in the model-based target data and (2) the NN retrievals capture soil moisture information not present in the model, resulting in better agreement with the core site and ISMN in situ measurements. The latter indicates that the NN retrievals may be beneficial in data assimilation, in particular for the short-term soil moisture variations (captured by the anomaly correlations against the cores sites) for which the skill difference between the retrievals and the model estimates is highest.

Generally, the (anomaly) correlation skill of the NN retrievals against core site measurements is lower than that of the SMAP L2P product (by 0.08 and 0.05 for the correlations and anomaly correlations, respectively). The ubRMSE of the NN retrievals, however, is lower than that of the L2P retrievals by 0.004 m^3m^{-3}. Evaluated against ISMN data, which represent a more diverse set of local conditions but only provide point-scale information, the NN and L2P retrievals have a very similar (anomaly) correlation skill, but the NN retrievals have a lower ubRMSE (by 0.04 m^3m^{-3}) than the L2P retrievals. The slightly lower (anomaly) correlation skill of the NN retrievals at the core sites is most likely related to errors in the training target data or missing information in the input data, whereas the higher ubRMSE of the L2P retrieval at the core sites is likely related to the higher time series variability of this product.

A triple collocation analysis using AMSR2 and ASCAT soil moisture retrievals as the additional two datasets showed that at the global scale all three products have comparable errors relative to their respective soil moisture dynamic range. The NN and L2P retrieval products have very similar error spatial patterns, but the NN retrievals have a better skill than the L2P product in densely vegetated regions and transition zones outside of CONUS. The GEOS-5 model has a slightly different error spatial patterns compared to the retrievals, with notable differences in high latitudes, where the model has higher errors owing to the increased uncertainty in its precipitation forcing, and in densely vegetated areas, where the retrieval products are less reliable owing to the lower soil moisture sensitivity of SMAP brightness temperatures in the presence of dense vegetation.

Overall, the skill of the SMAP NN retrievals is only slightly worse that of the SMAP L2P retrieval product, but the NN retrievals are provided in the global climatology of the GEOS-5 model, which may reduce the need for further bias correction before data assimilation. Local biases between the NN retrievals and the model, however, are retained in the NN retrievals, which would violate typical data assimilation requirements. Additionally, local discrepancies between the dynamic range of the NN retrievals and the model estimates could result in non-orthogonal errors between the observations and the model estimates, which would also violate typical data assimilation requirements. Consequently, further investigation is needed to determine the impact of such violations on the quality of the hydrological fields and surface flux estimates obtained from data assimilation, and whether the assimilation system can use NN retrievals more efficiently than standard retrievals or brightness temperatures.

The natural next step is thus to assimilate the SMAP NN retrieval product and compare the resulting analysis skill against that of assimilation experiments using traditional localized or other non-localized bias correction techniques, and against the assimilation of L2P retrievals and brightness temperatures. Another possible extension to this study would be to use the higher-resolution SMAP Enhanced Level-1C brightness temperature product (*Chaubell et al.*, 2016) to generate SMAP NN soil moisture retrievals at a higher spatial resolution.

Acknowledgments

J. Kolassa was supported by the NASA Postdoctoral Program, administered by Universities Space Research Association under contract with NASA. Additional funding was provided by the NASA Soil Moisture Active Passive mission. P. Gentine was supported by grant NNX15AB30G. Computational resources for this study were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at the Goddard Space Flight Center. We thank Aaron Berg, Tracy Rowlandson and Erica Tetlock for providing soil moisture measurements from the Kenaston network, which is funded by Environment and Climate Change Canada and the Canadian Space Agency. We also thank Mark Seyfried and Rogier van der Velde for providing soil moisture measurements from the Reynolds Creek and Twente networks, respectively.

References

Aires, F., Prigent, C., and Rossow, W.B. (2005), Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: 2. Global statistical relationships, *J. Geophys. Res., 110,* D11103, doi:10.1029/2004JD005094.

Albergel, C., Dorigo, W., Reichle, R.H., Balsamo, G., De Rosnay, P., Muoz-Sabater, J., Isaksen, L., De Jeu, R. and Wagner, W., (2013). Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing. *Journal of Hydrometeorology*, 14(4), pp.1259-1277. doi:10.1175/JHM-D-12-0161.1 Vancouver

Al Bitar, A., Leroux, D., Kerr, Y.H., Merlin, O., Richaume, P., Sahoo, A. and Wood, E.F., (2012). Evaluation of SMOS soil moisture products over continental US using the SCAN/SNOTEL network. *IEEE Transactions on Geoscience and Remote Sensing,* 50(5), pp.1572-1586. doi:10.1109/TGRS.2012.2186581

Alemohammad, S.H., Fang, B., Konings, A.G., Green, J.K., Kolassa, J., Prigent, C., Aires, F., Miralles, D. and Gentine, P., (2017) Water, Energy, and Carbon with Artificial Neural Networks (WECANN): A statistically-based estimate of global surface turbulent fluxes using solar-induced fluorescence. *Biogeosciences Discussions.* doi:10.5194/bg-2016-495

Assouline, S. (2013), Infiltration into soils: Conceptual approaches and solutions, *Water Resources Research 49(4)*, 1755-1772. doi:10.1002/wrcr.20155

Bateni, S. M., and Entekhabi, D. (2012). Relative efficiency of land surface energy balance components. *Water Resources Research, 48(4)*.doi:10.1029/2011WR011357

Bartalis, Z., Naeimi, V., Hasenauer, S. and Wagner, W., (2008). ASCAT soil moisture product handbook. *ASCAT Soil Moisture Report Series,* 15.

Belward, A.S., Estes, J.E. and Kline, K.D., (1999). The IGBP-DIS global 1-km land-cover data set DISCover: A project overview. *Photogrammetric Engineering and Remote Sensing*, 65(9), pp.1013-1020.

Brodzik, M.J., Billingsley, B., Haran, T., Raup, B. and Savoie, M.H., (2012). EASE-grid 2.0: incremental but significant improvements for Earth-gridded data sets. *ISPRS International Journal of Geo-Information*, 1(1), pp.32-45. doi:10.3390/ijgi1010032

Calvet, J.C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A. and Piguet, B., (2007), July. In situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network. *Geoscience and Remote Sensing Symposium*, 2007. IGARSS 2007. IEEE International (pp. 1196-1199). IEEE.

Chai, S.S., Walker, J.P., Makarynskyy, O., Kuhn, M., Veenendaal, B. and West, G., (2009). Use of soil moisture variability in artificial neural network retrieval of soil moisture. *Remote Sensing*, 2(1), pp.166-190. doi:10.3390/rs2010166

Chan, S., E. G. Njoku, and A. Colliander. (2016). SMAP L1C Radiometer Half-Orbit 36 km EASE-Grid Brightness Temperatures, Version 3. Boulder, Colorado USA. *NASA National Snow and Ice Data Center Distributed Active Archive Center*. doi: http://dx.doi.org/10.5067/E51BSP6V3KP7.

Chan, S.K., Bindlish, R., O'Neill, P.E., Njoku, E., Jackson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J. and Yueh, S., (2016b). Assessment of the SMAP passive soil moisture product. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8), pp.4994-5007. doi:10.1109/TGRS.2016.2561938

Chaubell, M. J., S. Chan, R. S. Dunbar, J. Peng, and S. Yueh. (2016). SMAP Enhanced L1C Radiometer Half-Orbit 9 km EASE-Grid Brightness Temperatures, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi:10.5067/2C9O9KT6JAWS.

Chen, F., Crow, W.T., Colliander, A., Cosh, M.H., Jackson, T.J., Bindlish, R., Reichle, R.H., Chan, S.K., Bosch, D.D., Starks, P.J. and Goodrich, D.C., (2016). Application of triple collocation in ground-based validation of Soil Moisture Active/Passive (SMAP) level 2 data products. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* doi:10.1109/JSTARS.2016.2569998

Colliander, A., Jackson, T.J., Bindlish, R., Chan, S., Das, N., Kim, S.B., Cosh, M.H., Dunbar, R.S., Dang, L., Pashaian, L. and Asanuma, J., (2017). Validation of SMAP surface soil moisture products with core validation sites. *Remote Sensing of Environment,* 191, pp.215-231.

Corradini, C., Morbidelli, R., and Melone, F. (1998). On the interaction between infiltration and Hortonian runoff. *Journal of Hydrology, 204(1)*, 52-67. doi:10.1016/S0022-1694(97)00100-5

Cybenko, G.V. (1989), Approximation by superpositions of a sigmoidal function, *Mathematics of Control, Signals and Systems, 2*, 303–314. doi: 10.1007/BF02551274

De Lannoy, G. J., Koster, R. D., Reichle, R. H., Mahanama, S. P., and Liu, Q. (2014). An updated treatment of soil texture and associated hydraulic properties in a global land modeling system. *Journal of Advances in Modeling Earth Systems, 6(4)*, 957-979. doi: 10.1002/2014MS000330

De Lannoy, G. J. M., R. H. Reichle, J. Peng, Y. Kerr, R. Castro, E. J. Kim, and Q. Liu (2015), Converting Between SMOS and SMAP Level-1 Brightness Temperature Observations Over Nonfrozen Land, *IEEE Geoscience and Remote Sensing Letters, 12,* 1908-1912, doi:10.1109/LGRS.2015.2437612.

De Lannoy, G.J. and Reichle, R.H., (2016). Global assimilation of multiangle and multipolarization SMOS brightness temperature observations into the GEOS-5 catchment land surface model for soil moisture estimation. *Journal of Hydrometeorology,* 17(2), pp.669-691. doi:10.1175/JHM-D-15-0037.1

Diamond, H.J., Karl, T.R., Palecki, M.A., Baker, C.B., Bell, J.E., Leeper, R.D., Easterling, D.R., Lawrimore, J.H., Meyers, T.P., Helfert, M.R. and Goodge, G., (2013). US Climate Reference Network after one decade of operations: Status and assessment. *Bulletin of the American Meteorological Society*, 94(4), pp.485-498. doi:10.1175/BAMS-D-12-00170.1

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T. (2011). The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, *Hydrology and Earth System Science*, 15, 1675-1698. doi:10.5194/hess-15-1675-2011

Draper, C.S., Reichle, R., de Jeu, R., Naeimi, V., Parinussa, R. and Wagner, W., (2013). Estimating root mean square errors in remotely sensed soil moisture over continental scale domains. *Remote Sensing of Environment*, 137, pp.288-298. doi:10.1016/j.rse.2013.06.013

Ducharne, A., Koster, R.D., Suarez, M.J., Stieglitz, M. and Kumar, P., (2000). A catchment-based approach to modeling land surface processes in a general circulation model: 2. Parameter estimation and model demonstration. *Journal of Geophysical Research*, 105(24), pp.823-24. doi:10.1029/2000JD900328

Entekhabi, D., Nakamura, H. and Njoku, E.G., (1994). Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. *IEEE Transactions on Geoscience and Remote Sensing*, 32(2), pp.438-448. doi:10.1109/36.295058

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N. (2010), The Soil Moisture Active Passive (SMAP) Mission, *Proceedings of the IEEE 98*(5), 704–716. doi:10.1109/JPROC.2010.2043918.

Figa-Saldaña, J., Wilson, J. J., Attema, E., Gelsthorpe, R., Drinkwater, M. R., and Stoffelen, A. (2002). The advanced scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for European wind scatterometers. *Canadian Journal of Remote Sensing*, 28(3), 404-412. doi:10.5589/m02-035

World Meteorological Organization - Global Climate Observing System (2009). Guideline for the Generation of Satellite-based Datasets and Products meeting GCOS Requirements. *WMO Technical Document* 1488 https://public.wmo.int/en/programmes/global-climate-observing-system/Publications/gcos-143.pdf

Gentine, P., Polcher, J., and Entekhabi, D. (2011). Harmonic propagation of variability in surface energy balance within a coupled soil–vegetation–atmosphere system. *Water Resources Research, 47(5).* doi:10.1029/2010WR009268

Gruber, A., Su, C.H., Zwieback, S., Crow, W., Dorigo, W. and Wagner, W., (2016). Recent advances in (soil moisture) triple collocation analysis. *International Journal of Applied Earth Observation and Geoinformation,* 45, pp.200-211. doi:10.1016/j.jag.2015.09.002

Jackson, T.J., Colliander, A., Kimball, J., Reichle, R., Derksen, C., Crow, W., Entekhabi, D., O'Neill, P. and Njoku, E., 2014. SMAP Science Data Calibration and Validation Plan (Revision A), *Soil Moisture Active Passive (SMAP) Mission Science Document. JPL D-52544.* Jet Propulsion Laboratory, Pasadena, CA. https://smap.jpl.nasa.gov/files/smap2/CalVal_Plan_120706_pub.pdf

Jackson, T.J., P. O'Neill, E. Njoku, S, Chan, R. Bindlish, A. Colliander, F. Chen, M. Burgin, S. Dunbar, J. Piepmeier, M. Cosh, T. Caldwell, J. Walker, X. Wu, A. Berg, T. Rowlandson, A. Pacheco, H. McNairn, M.

Thibeault, J. Martnez-Fernndez, . Gonzlez-Zamora, M. Seyfried, D. Bosch, P. Starks, D. Goodrich, J. Prueger, Z. Su, R. van der Velde, J. Asanuma, M. Palecki, E. Small, M. Zreda, J. Calvet, W. Crow, Y. Kerr, S. Yueh, and D. Entekhabi (2016). Calibration and Validation for the L23_SM_P Version 3 Data Products, *SMAP Project, JPL D-93720,* Jet Propulsion Laboratory, Pasadena, CA. http://nsidc.org/data/docs/daac/smap/sp_l2_smp/pdfs/L2SMP_validated_assess_rpt_rel2_v10a_final.pdf

Jimenez, C., Clark, D. B., Kolassa, J., Aires, F., and Prigent, C. (2013). A joint analysis of modeled soil moisture fields and satellite observations. *J. Geophys. Res, 118(12)*, 6771-6782. doi: 10.1002/jgrd.50430

Jung, M., Reichstein, M., Schwalm, C.R., Huntingford, C., Sitch, S., Ahlström, A., Arneth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P. and Gans, F., (2017). Compensatory water effects link yearly global land CO2 sink changes to temperature. *Nature,* 541(7638), pp.516-520. doi:10.1038/nature20780

Kasahara, M., Imaoka, K., Kachi, M., Fujii, H., Naoki, K., Maeda, T., Ito, N., Nakagawa, K. and Oki, T., (2012), Status of AMSR2 on GCOM-W1. *SPIE Remote Sensing* (pp. 853307-853307). International Society for Optics and Photonics. http://proceedings.spiedigitallibrary.org/pdfaccess.ashx?url=/data/conferences/spiep/69394/853307.pdf

Kerr, Y.H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J., Reul, N., Gruhier, C., Juglea, S.E., Drinkwater, M.R., Hahne, A., Martin-Neira, M., Mecklenburg, S., (2010) The SMOS

Mission: New Tool for Monitoring Key Elements of the Global Water Cycle, *Proceedings of the IEEE, 98(5)*, 666-687. doi:10.1109/JPROC.2010.2043032

Kerr, Y.H., Waldteufel, P., Richaume, P., Wigneron, J.P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S.E. and Leroux, D., (2012). The SMOS soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing,* 50(5), pp.1384-1403. doi:10.1109/TGRS.2012.2184548

Kolassa, J., (2013). Soil moisture retrieval from multi-instrument satellite observations. *Doctoral dissertation, Université Pierre et Marie Curie (Paris 6).*`http://www.theses.fr/2013PA066392`

Kolassa, J., Aires, F., Polcher, J., Prigent, C., Jimenez, C., and Pereira, J. M. (2013). Soil moisture retrieval from multi-instrument observations: Information content analysis and retrieval methodology. *J. Geophys. Res*, 118(10), 4847-4859. doi: 10.1029/2012JD018150

Kolassa, J., Gentine, P., Prigent, C., and Aires, F. (2016). Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 1: Satellite data analysis. *Remote Sensing of Environment*, 173, 1-14. doi:10.1016/j.rse.2015.11.011

Koster, R.D., Suarez, M.J., Ducharne, A., Stieglitz, M. and Kumar, P., (2000). A catchmentbased approach to modeling land surface processes in a general circulation model: 1. Model structure. *Journal of Geophysical Research: Atmospheres, 105(D20),* pp.24809-24822. doi:10.1029/2000JD900327

Lahoz, W.A. and De Lannoy, G.J., (2014). Closing the gaps in our knowledge of the hydrological cycle over land: conceptual problems. *Surveys in Geophysics*, 35(3), pp.623-660. doi:10.1007/s10712-013-9221-7

Larson, K.M., Small, E.E., Gutmann, E.D., Bilich, A.L., Braun, J.J. and Zavorotny, V.U., (2008). Use of GPS receivers as a soil moisture network for water cycle studies. *Geophysical Research Letters, 35(24).* doi:10.1029/2008GL036013

Leavesley, G.H., David, O., Garen, D.C., Lea, J., Marron, J.K., Pagano, T.C., Perkins, T.R. and Strobel, M.L., (2008), December. A modeling framework for improved agricultural water supply forecasting. *AGU Fall Meeting Abstracts.*

Levenberg, K., (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics, 2(2)*, pp.164-168.

Liu, Q., Reichle, R.H., Bindlish, R., Cosh, M.H., Crow, W.T., de Jeu, R., De Lannoy, G.J., Huffman, G.J. and Jackson, T.J., 2011. The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system. *Journal of Hydrometeorology*, 12(5), pp.750-765. doi: 10.1175/JHM-D-10-05000.1

Lucchesi, R., (2013) File Specification for GEOS-5 FP. *NASA Global Modeling and Assimilation Office (GMAO) Office Note* No. 4 (Version 1.0), 63 pp, available from `http://gmao.gsfc.nasa.gov/pubs/office_notes`.

Maeda, T. and Taniguchi, Y., (2013). Descriptions of GCOM-W1 AMSR2 Level

1R and Level 2 Algorithms. *Japan Aerospace Exploration Agency Earth Observation Research Center: Ibaraki, Japan.* `suzaku.eorc.jaxa.jp/GCOM_W/data/doc/NDX-120015A.pdf`

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics, 11(2),* 431-441. doi:10.1137/0111030

McColl, K.A., Vogelzang, J., Konings, A.G., Entekhabi, D., Piles, M. and Stoffelen, A., (2014). Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. *Geophysical Research Letters*, 41(17), pp.6229-6236. doi:10.1002/2014GL061322

McDowell, N. G. (2011). Mechanisms linking drought, hydraulics, carbon metabolism, and vegetation mortality. *Plant physiology, 155(3)*, 1051-1059. doi:10.1104/pp.110.170704

Moghaddam, M., Silva, A.R., Clewley, D., Akbar, R., Hussaini, S.A., Whitcomb, J., Devarakonda, R., Shrestha, R., Cook, R.B., Prakash, G. and Vannan, S.S., (2016). Soil Moisture Profiles and Temperature Data from SoilSCAPE Sites, USA. *ORNL DAAC*, Oak Ridge, Tennessee, USA. doi:10.3334/ORNLDAAC/1339

O'Neill, P., Chan, S., Njoku, E., Jackson, T., and Bindlish, R. (2015) SMAP Algorithm Theoretical Basis Document: L2 & L3 Radiometer Soil Moisture (Passive) Products. *SMAP Project, JPL D-66480*, Jet Propulsion

Laboratory, Pasadena, CA. https://nsidc.org/sites/nsidc.org/files/technical-references/L2_SM_P_ATBD_v7_Sep2015-po-en.pdf

O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, and R. Bindlish. 2016. SMAP L2 Radiometer Half-Orbit 36 km EASE-Grid Soil Moisture, Version 3. Boulder, Colorado USA. *National Snow and Ice Data Center Distributed Active Archive Center*. doi:10.5067/PLRS64IU03IT. Accessed: June 2016.

Owe, M., R. deJeu, and J. Walker(2001), A Methodology for Surface Soil Moisture and Vegetation optical Depth Retrieval Using The Microwave Polarization Difference Index, *IEEE Transactions on Geoscience and Remote Sensing, 39*, 1643–1654. doi: 10.1109/36.942542

Philip, J. R. (1957). The theory of infiltration: 5. The influence of the initial moisture content. *Soil Science, 84(4)*, 329-340.

Reichle, R. H., and R. D. Koster (2004), Bias reduction in short records of satellite soil moisture, *Geophysical Research Letters,* 31,L19501. doi: 10.1029/2004GL020938

Reichle, R.H., Koster, R.D., De Lannoy, G.J., Forman, B.A., Liu, Q., Mahanama, S.P. and Touré, A., 2011. Assessment and enhancement of MERRA land surface hydrology estimates. *Journal of Climate*, 24(24), pp.6322-6338. doi: 10.1175/JCLI-D-10-05033.1

Reichle, R. H., and Q. Liu, (2014). Observation-Corrected Precipitation Estimates in GEOS-5. NASA/TM 2014-104606, Vol. 35. doi: 20150000725

Reichle, R.H., De Lannoy, G.J., Liu, Q., Colliander, A., Conaty, A., Jackson, T., Kimball, J. and Koster, R.D., (2015b). Soil Moisture Active Passive (SMAP) Project Assessment Report for the Beta-Release L4_SM Data Product. *NASA Technical Report Series on Global Modeling and Data Assimilation*, NASA/TM-2015, 104606, pp.40-63. https://nsidc.org/sites/nsidc.org/files/technical-references/Reichle788.pdf

Reichle, R. H., G. J. M. De Lannoy, Q. Liu, A. Colliander, A. Conaty, T. Jackson, J. Kimball, and R. D. Koster (2015), Soil Moisture Active Passive (SMAP) Project Assessment Report for the Beta-Release L4_SM Data Product, *NASA Technical Report Series on Global Modeling and Data Assimilation, NASA/TM-2015-104606,* Vol. 40, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA, 63pp. Available at: https://gmao.gsfc.nasa.gov/pubs/

Reichle, R.H., De Lannoy, G.J., Liu, Q., Ardizzone, J.V., Chen, F., Colliander, A., Conaty, A., Crow, W., Jackson, T., Kimball, J. and Koster, R.D., (2016). Soil Moisture Active Passive Mission L4_SM Data Product Assessment (Version 2 Validated Release). *NASA Global Modelling and Assimilation Office (GMAO) Office Note* No. 12 https://nsidc.org/sites/nsidc.org/files/technical-references/Reichle851.pdf

Reichle, R.H., Liu, Q., Koster, R.D., Draper, C.S., Mahanama, S.P. and Partyka, G.S., (2017). Land surface precipitation in MERRA-2. *Journal of Climate*, 30, pp. 1643-1664. doi:10.1175/JCLI-D-16-0570.1

Reichle, R.H., De Lannoy, G.J., Liu, Q., Ardizzone, J.V., Colliander, A., Conaty, A., Crow, W., Jackson, T.J., Jones, L.A., Kimball, J.S. and Koster, R.D., (2017). Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using In Situ Measurements. *Journal of Hydrometeorology (in press)* , doi:10.1175/JHM-D-17-0063.1.

Reynolds, C.A., Jackson, T.J. and Rawls, W.J., (2000). Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions. *Water Resources Research*, 36(12), pp.3653-3662. doi:10.1029/2000WR900130

Rodrguez-Fernández, N.J., Aires, F., Richaume, P., Kerr, Y.H., Prigent, C., Kolassa, J., Cabot, F., Jimnez, C., Mahmoodi, A. and Drusch, M., (2015). Soil moisture retrieval using neural networks: Application to SMOS. *IEEE Transactions on Geoscience and Remote Sensing*, 53(11), pp.5991-6007. doi:10.1109/TGRS.2015.2430845

Rumelhart, D.E., and Chauvin, Y. (1995), Backpropagation: theory, architectures, and applications, *Lawrence Erlbaum Associates.* ISBN: 0-8058-1258-X

Sanchez, N., Martínez-Fernández, J., Scaini, A. and Perez-Gutierrez, C., (2012). Validation of the SMOS L2 soil moisture data in the REMEDHUS network (Spain). *IEEE Transactions on Geoscience and Remote Sensing, 50(5),* pp.1602-1611. doi:10.1109/TGRS.2012.2186971

Santi, E., Paloscia, S., Pettinato, S. and Fontanelli, G., (2016). Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors. *International Journal of Applied Earth Observation and Geoinformation*, 48, pp.61-73. doi:10.1016/j.jag.2015.08.002

Schaefer, G.L., Cosh, M.H. and Jackson, T.J., (2007). The USDA natural resources conservation service soil climate analysis network (SCAN). *Journal of Atmospheric and Oceanic Technology,* 24(12), pp.2073-2077. doi:10.1175/2007JTECHA930.1

Schmugge, T.J., Jackson, T.J. and McKim, H.L., (1980). Survey of methods for soil moisture determination. *Water Resources Research,* 16(6), pp.961-979.

Scipal, K., Holmes, T., De Jeu, R., Naeimi, V. and Wagner, W., (2008). A possible solution for the problem of estimating the error structure of global soil moisture data sets. *Geophysical Research Letters*, 35(24). doi:10.1029/2008GL035599

Seneviratne, S. I., Lüthi, D., Litschi, M., and Schär, C. (2006). Land-atmosphere coupling and climate change in Europe. *Nature*, 443(7108), 205-209.doi:10.1038/nature05095

Sevanto, S., McDowell, N. G., Dickman, L. T., Pangle, R., and Pockman, W. T. (2014). How do trees die? A test of the hydraulic failure and carbon starvation hypotheses. *Plant, cell & environment, 37(1)*, 153-161. doi:10.1111/pce.12141

Stoffelen, A., (1998). Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *Journal of Geophysical Research*, 103(C4), pp.7755-7766. doi:10.1029/97JC03180

Su, C.H., Ryu, D., Crow, W.T. and Western, A.W., (2014). Beyond triple collocation: Applications to soil moisture monitoring. *Journal of Geophysical Research: Atmospheres*, 119(11), pp.6419-6439. doi:10.1002/2013JD021043

Tagesson, T., Fensholt, R., Guiro, I., Rasmussen, M.O., Huber, S., Mbow, C., Garcia, M., Horion, S., Sandholt, I., Holm?Rasmussen, B. and Göttsche, F.M., (2015). Ecosystem properties of semiarid savanna grassland in West Africa and its relationship with environmental variability. *Global change biology*, 21(1), pp.250-264. doi:10.1111/gcb.12734

Taylor, J.R., Osenga, E.C., Jack-Scott, E., Arnott, J.C. and Katzenberger, J., (2015), December. Establishing a Long Term High-Altitude Soil Moisture Monitoring Network at the Watershed Scale. *AGU Fall Meeting Abstracts*.

Wagner, W., Lemoine, G., and Rott, H. (1999). A method for estimating soil moisture from ERS scatterometer and soil data. *Remote Sensing of Environment*, 70(2), 191-207. doi:10.1016/S0034-4257(99)00036-X

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., ... and Rubel, F. (2013). The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications. *Meteorologische Zeitschrift,* 22(1), 5-33. doi: 10.1127/0941-2948/2013/0399

Walker, J.P. and Houser, P.R., (2001). A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations. *Journal of Geophysical Research*, 106(D11), pp.11761-11774. doi:10.1029/2001JD900149

Wigneron, J.P., Chanzy, A., Calvet, J.C. and Bruguier, N., (1995). A simple algorithm to retrieve soil moisture and vegetation biomass using passive microwave measurements over crop fields. *Remote Sensing of Environment*, 51(3), pp.331-341. doi:10.1016/0034-4257(94)00081-W

Zreda, M., Desilets, D., Ferrè, T.P.A. and Scott, R.L., (2008). Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. *Geophysical Research Letters*, 35(21). doi:10.1029/2008GL035655