Global-scale Evaluation of SMAP, SMOS and ASCAT Soil Moisture Products using Triple Collocation

Fan Chen[1,2], Wade T. Crow[1], Rajat Bindlish[3], Andreas Colliander[4], Mariko S. Burgin[4], Jun Asanuma[5], and Kentaro Aida[5]

[1] Science Systems and Applications, Inc., Greenbelt, MD, USA

[2] USDA ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD 20705, USA

[3] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

[4] NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

[5] University of Tsukuba, Tsukuba, Japan

Abstract

Global-scale surface soil moisture products are currently available from multiple remote sensing platforms. Footprint-scale assessments of these products are generally restricted to limited number of densely-instrumented validation sites. However, by taking active and passive soil moisture products together with a third independent soil moisture estimates via land surface modeling, triple collocation (TC) can be applied to estimate the correlation metric of satellite soil moisture products (versus an unknown ground truth) over a quasi-global domain. Here, an assessment of Soil Moisture Active Passive (SMAP), Soil Moisture Ocean Salinity (SMOS) and Advanced SCATterometer (ASCAT) surface soil moisture retrievals via TC is presented. Considering the potential violation of TC error assumptions, the impact of active-passive and satellite-model error cross correlations on the TC-derived inter-comparison results is examined at *in situ* sites using quadruple collocation analysis. In addition, confidence intervals for the TC-estimated correlation metric are constructed from moving-block bootstrap sampling designed to preserve the temporal persistence of the original (unevenly-sampled) soil moisture time-series. This study is the first to apply TC to obtain a robust global-scale cross-

assessment of SMAP, SMOS and ASCAT soil moisture retrieval accuracy in terms of anomaly temporal correlation. Our results confirm the overall advantage of SMAP (with a global average anomaly correlation of 0.76) over SMOS (0.66) and ASCAT (0.63) that has been established in several recent regional, ground-based studies. SMAP is also the best-performing product over the majority of applicable land pixels (52%), although SMOS and ASCAT each shows advantage in distinct geographic regions.

1. Introduction

As a key state variable in hydrological and meteorological modeling systems, the global observation of soil moisture has become a major priority. Currently, several remote sensing platforms provide continuous global surface (approximately 0-5 cm) retrievals: the National Aeronautics and Space Administration (NASA)'s Soil Moisture Active Passive (SMAP, 2015-), the European Space Agency (ESA)'s Soil Moisture Ocean Salinity (SMOS, 2009-), the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)'s Advanced SCATterometers (ASCAT, 2007-), and the Japanese Aerospace Exploration Agency (JAXA)'s Advanced Microwave Scanning Radiometer 2 (AMSR2, 2012-). The accuracy of the satellite soil moisture retrievals is typically described via their root-mean-squared-error (RMSE; e.g. Brocca *et al*. 2010; Jackson *et al*. 2010; Kerr *et al*. 2016) or de-biased/unbiased RMSE (ubRMSE; e.g. Colliander *et al*. 2017) versus ground-based observations at a footprint-scale. However, difficulty in obtaining viable estimates of ground truth soil moisture at the satellite footprint scale has limited past validation activities to a small number of locations (e.g., SMAP's core validation sites) and/or discrete time periods (e.g., field campaigns). The broader evaluation of satellite soil moisture products (across regional or continental scales) is typically based on comparisons with sparse ground soil moisture networks or modeled datasets (e.g., Paulik *et al*. 2014; González-Zamora *et al*. 2015; Piles *at al.* 2014; Al-Yaari *et al*. 2014; Polcher *et al*. 2016;

Kim *et al*. 2018). Naturally, such comparisons are unable to provide direct validation metrics relative to the ground truth, but rather metrics against a chosen reference dataset with unknown errors at the footprint-scale of satellite retrievals. For example, correlation coefficient metrics obtained from comparing with point-scale ground observations have been shown to underestimate the correlation between retrievals and true soil moisture values (Chen *et al*. 2017).

Initially designed for obtaining the calibration constants against a reference dataset in satellite wind speed products, the triple collocation (TC) (Stoffelen 1998) technique provides a solution to such challenge. In particular, TC can be applied to the estimate error variances of a geophysical measurement system and has become an important tool for satellite soil moisture assessments (e.g., Zwieback *et al*. 2012; Dorigo *et al*. 2010; Miralles *et al*. 2010; Draper *et al*. 2013). However, standard TC applications are limited to only providing relative error metrics. It requires a reference dataset to be chosen from the three collocated data products, and the resulting error variances are subject to the multiplicative and additive biases of the reference dataset (Chen *et al*. 2017). Recently developed TC-based solution – the Extended Triple Collocation, or ETC (McColl 2014) – for the Pearson's correlation coefficient metric, on the other hand, does not require a reference dataset and yields an absolute estimate of the temporal correlation between the product under evaluation and the unknown truth. Pearson's correlation coefficient is a widely reported metric for satellite soil moisture and an appropriate metric for summarizing retrieval value in a data assimilation context (Reichle *et al.*, 2008). In this analysis, we adopt the ETC solution and conduct an assessment and inter-comparison of the SMAP Level 3, SMOS Level 3 and ASCAT Level 2 soil moisture products based on the correlation metric (R). Until recently, relatively few studies have been conducted to evaluate satellite soil moisture products at a continental scale (e.g. Draper *et al*. 2013; Leroux *et al*. 2013) using TC. To the best of our knowledge, this study is the first attempt to apply TC to obtain the footprint-scale correlation metric for SMAP observations at quasi-global scale, and compare it directly with soil moisture retrievals from SMOS and ASCAT.

Our basic strategy for applying TC is to employ soil moisture data triplets comprising a passive microwave product (SMAP or SMOS), an active remote sensing product (ASCAT), and a land surface model product. TC is based on a fundamental assumption that each of these products contain uncorrelated errors. However, recent works have identified non-negligible error correlation in soil moisture products acquired from active and passive microwave sources (Gruber *et al*. 2016b; Pierdicca *et al*. 2017). This suggests that it is necessary to examine the impact of violating this assumption on SMAP-ASCAT and SMOS-ASCAT-based TC analyses. Therefore, we also apply the least-squares quadruple collocation solution (QC, Pierdicca *et al*. 2015) to estimate the error cross-correlations at over 200 sparse ground observation sites to further evaluate the robustness of our global TC analysis strategy.

This paper is organized as follows. Section 2 reviews the TC and quadruple collocation (QC) methodologies and data-processing procedures as well as the use of moving-block bootstrap re-sampling to obtain confidence intervals for TC-derived *R*. Section 3 describes the remote sensing, land surface modeling and ground observation datasets used in the analysis. Section 4 presents the QC results at sparse network sites and discusses the sensitivity of the TC analysis to both non-zero error cross-correlation between active and passive satellite soil moisture products and our choice of a particular land surface model dataset. Results and discussions of global comparison of SMAP, SMOS and ASCAT soil moisture via TC are presented in Sections 5 and 6, respectively.

2. Methodologies

2.1 **Extended Triple Collocation**

In soil moisture validation and comparison studies, TC has typically been applied to estimate the random error variance of a particular soil moisture dataset. In contrast, the extended triple collocation (ETC) approach (McColl 2014) solves for the correlation between a dataset and the unknown truth. As in TC, it requires three collocated, independent measurement systems (*X*, *Y*, *Z*, in our case representing: a passive satellite retrieval, an active satellite retrieval and a model product, respectively) that describe the same geophysical variable (in this case - average surface soil moisture of the satellite grid cell, which is approximately 40 x 40 km^2). ETC is based on the following assumptions: 1) all three datasets are linearly related to the true state (*T*); 2) zero error cross-correlation exists between *X*, *Y* and *Z*; and 3) zero correlation exists between errors and *T* and 4) the stationary of signal and error statistics (Gruber *et al*. 2016a; Draper *et al*. 2013; Zwieback *et al*. 2012). If these assumptions hold, the correlation between *X* and the *T* can be estimated as

$$R_X = \sqrt{\frac{\sigma_{XY}\sigma_{XZ}}{\sigma_X^2\sigma_{YZ}}} \qquad (1)$$

where σ_{XY} is the covariance of *X* and *Y*, and σ_X^2 is the variance of *X*. Analytical details for deriving (1) from the classic TC method (Stoffelen 1998) can be found in McColl (2014).

To ensure consistency with the assumption listed above, seasonal signals are commonly removed from the raw time-series of each product prior to the application of TC (Gruber *et al*. 2016a; Dorigo *et al*. 2010; Su and Ryu, 2015). Here, anomaly time series are generated by removing the average value of a 30-day moving window centered upon the data point being treated (i.e. from day -14 to day +15). Given the potential temporally sparse nature of satellite retrievals, a minimum of 3 observations is required in each of the first and second halves of the 30-day window, in addition to the data point being treated itself. This particular anomaly definition, versus the alternative definition of deviations from a long-term seasonal climatology, has less stringent requirements regarding the length of datasets, which is usually the limiting factor in the application of TC in satellite products. While the removal of low-frequency variability has been shown to improve the robustness of TC results (Chen *et al*. 2017), it renders our particular ETC approach insensitive to (potentially-important) error in low-frequency and/or seasonal soil moisture dynamics. The implications of this will be discussed below.

ETC-based estimates of correlation are considered viable when: 1) the collocated triple time series is comprised of at least 50 data points; 2) positive correlation is found between each of the three input anomaly time-series, and 3) ETC correlation outputs are real and positive for each of the three datasets. All other ETC correlation estimates are masked. The positive correlation requirement between input datasets (#2 above) is necessary to avoid ambiguity since ETC is unable to resolve the sign of the output R values (McColl 2014). This limitation results in the exclusion of pixels in certain regions where active and passive soil moisture retrievals are negatively correlated (see additional discussion in Section 5).

2.2 Estimation of error cross-correlation: Quadruple collocation

As noted above, a potential source of error for the TC analysis is the presence of error cross-correlation (ECC) between the soil moisture datasets, especially between active and passive remote sensing products. Non-zero ECC violates the underlying TC assumptions and can lead to biased TC results. In past studies, ECC was typically assumed to be zero between all products (e.g., Leroux *et al*. 2013). However, recent works have revealed the presence of non-zero ECC between active and passive soil moisture retrievals (Gruber *et al*. 2016b; Pierdicca *et al*. 2017). Therefore, it is prudent to re-examine ECC levels in SMAP-ASCAT and SMOS-ASCAT soil moisture data pairs utilized here.

The TC algorithm can be extended to include a fourth dataset (i.e., quadruple collocation, or QC) and the error variances can be estimated with a least squares solution (Pierdicca *et al*. 2015) with the same TC assumptions. Furthermore, the zero ECC assumption can be relaxed, and – on the condition that only one pair within of the four datasets have non-zero ECC – estimates of ECC can be obtained from the least-squares solution (Zwieback *et al*. 2012; Gruber *et al*. 2016b).

Here we adopt the formulation in Gruber *et al*. (2016b) to estimate the error cross-correlation between the active (ASCAT) and passive (SMAP, SMOS) soil moisture datasets and assess the impact of such cross-correlation on TC results. The QC analysis is conducted at sparse soil moisture network sites where ground observations can serve as the fourth soil moisture dataset. The QC formulation also provides estimates of the error variances of each dataset. In certain cases, such estimates will be more accurate than those obtained from TC since QC can account for the presence of non-zero ECC within a particular pair of collocated datasets (Yilmaz and Crow, 2014).

Given four soil moisture measurement systems *X*, *Y*, *Z*, *W*, representing a passive remote sensing, an active remote sensing, a model and point-scale ground observation, respectively, the least-squares solution for the QC problem is given by

$$y = \begin{bmatrix} \sigma_X^2 \\ \sigma_Y^2 \\ \sigma_Z^2 \\ \sigma_W^2 \\ \sigma_{XY} \\ \sigma_{XZ}\sigma_{XW}/\sigma_{ZW} \\ \sigma_{YZ}\sigma_{YW}/\sigma_{ZW} \\ \sigma_{XZ}\sigma_{ZW}/\sigma_{XW} \\ \sigma_{YZ}\sigma_{ZW}/\sigma_{YW} \\ \sigma_{XW}\sigma_{ZW}/\sigma_{XZ} \\ \sigma_{YW}\sigma_{ZW}/\sigma_{YZ} \\ \sigma_{XZ}\sigma_{YW}/\sigma_{ZW} \\ \sigma_{XW}\sigma_{YZ}/\sigma_{ZW} \end{bmatrix} A = \begin{bmatrix} 1&0&0&0&0&1&0&0&0&0 \\ 0&1&0&0&0&0&1&0&0&0 \\ 0&0&1&0&0&0&0&1&0&0 \\ 0&0&0&1&0&0&0&0&1&0 \\ 0&0&0&0&1&0&0&0&0&1 \\ 1&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0 \end{bmatrix} x = \begin{bmatrix} \beta_X^2\sigma_\Theta^2 \\ \beta_Y^2\sigma_\Theta^2 \\ \beta_Z^2\sigma_\Theta^2 \\ \beta_W^2\sigma_\Theta^2 \\ \beta_X\beta_Y\sigma_\Theta^2 \\ \sigma_{\varepsilon_X}^2 \\ \sigma_{\varepsilon_Y}^2 \\ \sigma_{\varepsilon_Z}^2 \\ \sigma_{\varepsilon_W}^2 \\ \sigma_{\varepsilon_X\varepsilon_Y} \end{bmatrix} \quad (2)$$

where Θ is the true soil moisture signal, and β is the multiplicative bias of a given dataset as in $X = \alpha_X + \beta_X\Theta + \varepsilon_X$, and ε is the zero-mean random error.

And the least squares solution for the parameters in x is given as

$$\hat{x} = (A^T A)^{-1} A^T y \quad (3)$$

Note that this solution enables the TC approach described in section 2.1 to be slightly relaxed. In particular, non-zero ECC is now allowed in one data pair (here between *X* and *Y*, where *X* is SMAP or SMOS, and *Y* is ASCAT). ECC between any other data pairs are still required to be zero (i.e., $\sigma_{\varepsilon_X\varepsilon_Y} \neq 0$, and $\sigma_{\varepsilon_X\varepsilon_Z} = \sigma_{\varepsilon_X\varepsilon_W} = \sigma_{\varepsilon_Y\varepsilon_Z} = \sigma_{\varepsilon_Y\varepsilon_W} = \sigma_{\varepsilon_Z\varepsilon_W} = 0$). As in Gruber *et al*. (2016b), we consider these conditions generally satisfied in the active-passive-LSM-*in situ* data quadruples in this study.

2.3 Confidence interval from moving block bootstrapping

Using collocated surface soil moisture retrievals from passive (SMAP or SMOS) and active (ASCAT) sensors and a land surface modeling product, the correlation metric of the three satellite products (versus an unknown truth) can be estimated via TC at a quasi-global scale. However, considerable sampling errors are expected in TC results, especially when the length of the analysis is shortened to accommodate new satellite products (e.g., the two years of SMAP considered here). Therefore, it is critical to account for sampling uncertainties when making comparisons between the satellite products.

Here, such uncertainties are quantified via bootstrap re-sampling at each pixel to construct the confidence interval (CI) of TC estimates. As noted earlier, auto-correlation in time-series will reduce the effective sample size and thus underestimate the probability that the original bootstrap confidence interval contains the true statistical property (Zwiers, 1990; von Storch and Zwiers, 1999). Since soil moisture time series typically contain large amounts of temporal auto-correlation, this effect should be considered when generating boot-strapped errors estimates for soil moisture TC results. Although mean 30-day signals have been removed from the original time-series, our analysis suggests the resulting anomaly time-series still contains significant first-order autocorrelation (not shown). This impact also applies for correlation estimated by ETC techniques since the latter is essentially an expansion upon the Pearson's correlation coefficient formula from two to three time series members (McColl, 2014). A solution is proposed in Mudelsee (2002, 2010) where a pair-wise moving block bootstrap (MBB) re-sampling technique is applied to obtain a robust estimate of the confidence intervals for Pearson's correlation coefficient in serially-correlated time-series.

Here, we have adapted the MBB method introduced in Ólafsdóttir and Mudelsee (2014) for the bi-variate correlation problem to the triple collocation problem to construct the confidence interval of the ETC correlation results. In each iteration of the re-sampling procedure, MBB is applied to draw blocks of data triplets from the original time series samples to form samples that preserve the temporal persistence of the original data. Block length is determined from the equivalent autocorrelation coefficient of the three anomaly time-series (i.e., ETC inputs) which is calculated from individual persistence time, τ, of the three time-series. Persistence times are then estimated by minimizing the sum of squares:

$$S(\tau_X) = \sum_{i=2}^{n}[x(i) - exp\{-[t(i) - t(i-1)]/\tau_X\} \cdot x(i-1)]^2 \quad (4)$$

where *n* is the length of the time-series, *x*(*i*) is the *i*th data point (i.e. soil moisture anomaly) and *t(i)* is the linear time point (in unit of day) with uneven spacing, which is typical of satellite retrievals. Note that although the land surface model time-series are evenly spaced with sub-daily frequency, only the data points that temporally matched to the satellite retrievals are considered and thereby treated as an unevenly-spaced time series. The equivalent AR(1) autocorrelation coefficient is given by $a_X = \exp(-d/\tau_X)$, where $d = [t(n) - t(1)]/(n-1)$ is the average time spacing. The autocorrelation coefficient is then bias-corrected to approximate the AR(1) process with an even time-spacing:

$$a'_X = [a_X \cdot (n-1) + 1]/(n-4). \quad (5)$$

A joint, bias-corrected equivalent autocorrelation coefficient for the triple collocation analysis is given by

$$a'_{XYZ} = (a'_X \cdot a'_Y \cdot a'_Z)^{1/3}. \quad (6)$$

The optimal block length is then estimated as

$$l_{opt} = \mathrm{NINT}\left\{\left[\sqrt{6} \cdot a'_{XYZ}/(1 - a'^{2}_{XYZ})\right]^{2/3} \cdot n^{1/3}\right\} \quad (7)$$

where NINT denotes rounding to the nearest integer. Overlapping blocks of data triplets with the length of l_{opt} are then extracted from the match-up anomaly time-series and then randomly drawn with replacement to be concatenated until the original data length is reached (see Figure 1 for an illustration of this procedure). Extra data points in the end of the newly-formed bootstrap sample are trimmed. The re-sampling procedure is repeated 1000 times in each grid pixel. Estimated 95% confidence intervals for each correlation coefficient are defined as the range between 2.5th and 97.5th percentile of the bootstrapped sampling distribution.

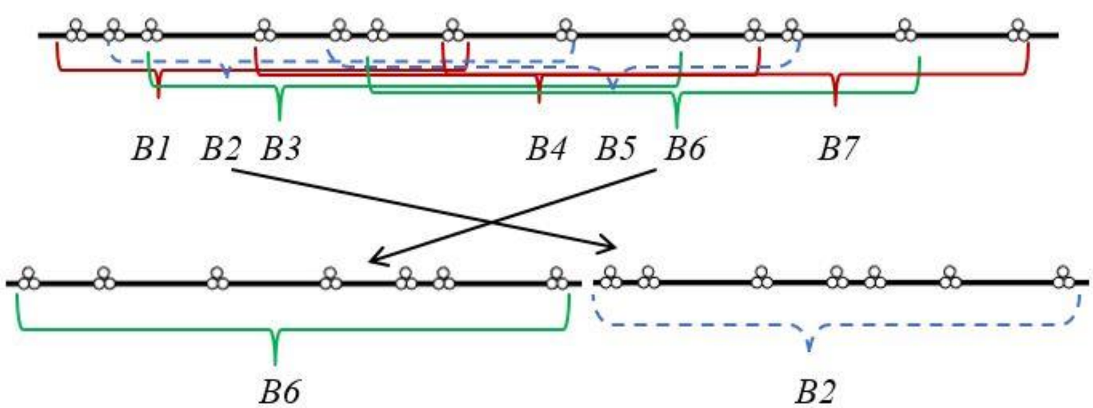

Figure 1. Schematic diagram of moving block bootstrapping for the case l_{opt}= 7 applied to a temporally sporadic time series of available soil moisture triplets. Overlapping data blocks from the original time series (top) are drawn randomly with replacement and then concatenated to generate a new bootstrap resample (bottom).

3. Data

As discussed above, three satellite surface soil moisture products (acquired between March 31, 2015 and March 31, 2017) are evaluated in this analysis: Level 3 SMAP passive radiometer retrievals (L3_SM_P, v4-R14010), Level 3 SMOS radiometer retrievals (v300), and Level 2 ASCAT scatterometer retrievals. All three retrieval time series contain retrievals obtained from both ascending and descending orbits. Details of each product are given below.

3.1 Soil Moisture Active Passive (SMAP)

Launched in January 2015, NASA’s SMAP satellite began continuous science data acquisition on March 31, 2015 with its L-band (1.41 GHz) radiometer (Entekhabi at al., 2010). The SMAP L3 data is in the format of global gridded maps of daily composites of the SMAP Level 2 Passive Soil Moisture (L2_SM_P) ascending/descending swath data, and is posted on a global cylindrical 36 km Equal-Area Scalable Earth, version 2 (EASEv2) grid. The validated SMAP L2/3 soil moisture product is based on the V-polarization single-channel (SCA-V) retrieval algorithm (Chan *et al*. 2016). Data screening is based on the soil moisture retrieval quality flag and only those flagged as “recommended for retrieval” are considered in this analysis. The retrieval quality flag is determined from a number of surface and retrieval conditions which can be found in Chan *et al*. (2016) and Chan and Dunbar (2015). Soil moisture retrievals from the ascending (6 PM LST) overpasses are now included in the SMAP Level 2/3 passive version 4 data products. Validation of the ascending (PM) retrievals indicate that it also meets the mission requirement of 0.04 m^3/m^3 unbiased root mean square error (ubRMSE), but with a small degradation compared to the descending (AM) retrievals (Jackson *et al*. 2016).

3.2 Soil Moisture Ocean Salinity (SMOS)

ESA's SMOS satellite was launched in November 2009 and measures L-band microwave emission (1.400-1.427 GHz) with equatorial ascending/descending overpasses at 6 AM/PM local solar time and a 3-day revisit period at the equator (Kerr *et al*. 2001). The SMOS soil moisture retrieval algorithm can be found in Kerr *et al*. (2013). The SMOS Level 3 (v300) soil moisture product used here is generated on a 25-km EASEv2 grid (Brodzik and Knowles, 2002) available through the Centre Aval de Traitement des Données (CATDS) (http://www.catds.fr). In this study, the SMOS L3 soil moisture data was re-gridded to the SMAP 36 km-EASEv2 grid by bilinear interpolation. Data were screened primarily by the SMOS Data Quality indeX (DQX), which takes into account the error in the retrieval parameters and the Level 1 brightness temperatures (Kerr, *et al.* 2013). DQX has been applied to screen SMOS soil moisture retrievals in several studies with thresholds varying between 0.045 and 0.07 (e.g. Polcher *et al.* 2016; Al-Yaari *et al.* 2014; Pierdicca *et al.* 2013). Here, pixels with DQX > 0.07 m^3/m^3 or covered by snow or ice were removed.. A stricter screening threshold of 0.04 m^3/m^3 for DQX is also applied to examine the impact on the overall performance SMOS relative to SMAP and ASCAT (see Section 5). The impact of varying this threshold on key results will be discussed below. However, unless otherwise noted, satellite comparison results shown below are based on the 0.07 m^3/m^3 DQX threshold to maximize the temporal and spatial coverage of the analysis.

3.3 Advanced Scatterometer

The Advanced Scatterometer (ASCAT) sensor onboard the Meteorological Operational-B (MetOp-B) satellite measures C-band (5.3 GHz) radar backscatter since September 2012, with 25-34 km spatial resolution and equatorial ascending/descending overpasses at 9:30 PM/AM local solar time and a revisit frequency of 3 days. The ASCAT Level 2 (v5) soil moisture index product utilized here is based on the change-detection algorithm developed by Vienna University of Technology (TU Wien; see Wagner *et al*. 1999; Naeimi *et al*. 2009) obtained from EUMETSAT Earth Observation Portal (EOP). As conversion to volumetric soil moisture unit is not required in calculation of correlation coefficient, potential error due to inaccurate global porosity dataset is avoided here. Pixels were masked if the probability of snow, frozen ground and estimated retrieval error are greater than 50%. The ASCAT L2 soil moisture index data are available at 12.5-km grid resolution and were re-sampled onto the SMAP 36 km-EASEv2 grid through inverse-distance-weighting interpolation.

3.4 Land surface modeling products

Two operational global land surface modeling (LSM) soil moisture datasets are used in this analysis. The first is the operational analysis layer-1 (0-7 cm) volumetric soil moisture field from the European Centre for Medium Range Weather Forecasts (ECMWF) H-TESSEL (Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land) land-surface scheme (Balsamo *et al*. 2009). The operational soil moisture analysis product data is produced by ECMWF's Land Data Assimilation System by the assimilation of 2-m air temperature and relative humidity observations (Drusch *et al*. 2009; de Rosnay *et al*. 2012). The ECMWF soil moisture analysis data is available at 00, 06, 12 and 18Z hours and in a N640 reduced Gaussian grid. Here, it was re-gridded to the nearest 36-km EASEv2 grid using a nearest neighbor approach.

The second LSM soil moisture product utilized here is the so-called SMAP Nature Run, version 3 (NRv3), available at 3-hourly interval and 9-km EASEv2 grid and were aggregated to 36-km EASEv2 grid by spatial averaging. The NRv3 data were generated with an early version of the SMAP Level 4 Surface and Root Zone Soil Moisture (L4_SM) algorithm by the NASA Goddard

Space Flight Center (GSFC) Global Modeling and Assimilation Office (Reichle *et al*. 2016), which was applied in a model-only configuration using a single ensemble member, without perturbations, and without the assimilation of SMAP observations.

The re-sampling methods for the satellite and LSM datasets were each chosen considering the features of both source and target grids (i.e. SMAP EASEv2 grid). For ECMWF, the average grid size is close to the target grid size and therefore nearest-neighbor type simple grid transformation is appropriate given that it avoids potential interpolation artifacts. For NRv3 the source grid is perfectly nested within the target grid so simple averaging is ideal. The choice of re-sampling method for SMOS and ASCAT has been made with close attention to limiting factors and after discussion with data providers. A bilinear interpolation was found to produce fewest artifacts for SMOS with its 25-km EASEv2 grid. ASCAT's grid resolution is higher (12.5 km) and the original data was provided in time-ordered format; an inverse-distance-weighting interpolation was found to be most accurate.

3.5 Sparse network ground observations

In order to verify aspects of our ETC analysis (see Section 4), two years (3/31/2015 – 3/31/2017) of ground soil moisture measurements were obtained from various sparse networks (Table 1) and applied in a QC analysis (see Section 4 below). These networks typically provide one point-scale measurement per satellite footprint at approximately 5-cm depth, except for the COsmic-ray Soil Moisture Observing System (COSMOS) and PBO H_2O/GPS networks. The cosmic-ray neutron detectors (Zreda *et al*., 2008; 2012) in the COSMOS network measure soil moisture have a footprint radius varying between ~130 to 240 meters and a dynamic penetration depth of between ~15 to 83 centimeters (Köhli *et al*. 2015). The PBO H_2O/GPS network utilizes Global Positioning System (GPS) receivers that record temporal changes in the signal-to-noise characteristic of GPS reflectometry data to estimate changes in soil moisture with a sensing depth of 2.5 cm or less (Chew *et al*. 2014) and a sensing area of approximately 120 m^2 per satellite track (Larson and Nievinski, 2013). Multiple tracks are combined to produce a daily average soil moisture value with the aggregate sensing area of approximately 1000 m^2. Except for the GPS network, hourly soil moisture measurements are generally available for all networks.

Table 1. Sparse networks providing ground measurements of soil moisture.

Network	**Instrument**	**Region**	**Number of stations**	**Reference**
Soil Climate Analysis Network (SCAN)	Hydra Probe	USA	71	Shaefer *et al*. 2007
U.S. Climate Reference Network (USCRN)	Hydra Probe II	USA	44	Bell *et al*. 2013
Oklahoma Mesonet	Campbell Scientific 229-L	Oklahoma, USA	84	Illston *et al*. 2008; Scott et al. 2013
COsmic-ray Soil Moisture Observing System (COSMOS)	cosmic-ray soil moisture probe	USA, Europe, Africa, Australia	23	Zreda *et al*., 2008, 2012
PBO H_2O (GPS)	Global Positioning System (GPS) receivers	Western USA	28	Larson *et al*. 2008
SMOSMANIA	ThetaProbe ML2X	France	8	Calvet *et al*. 2007
Pampas	ThetaProbe ML2X	Argentina	8	
Mongolia Grasslands	Time Domain Reflectometry (TDR) probes	Mongolia	5	Kaihotsu *et al*. 2009

Figure 2 shows the location of the ground observation sites used in this study. Note that some of the stations were missing in certain subsequent figures due to the limited availability of collocated satellite observations.

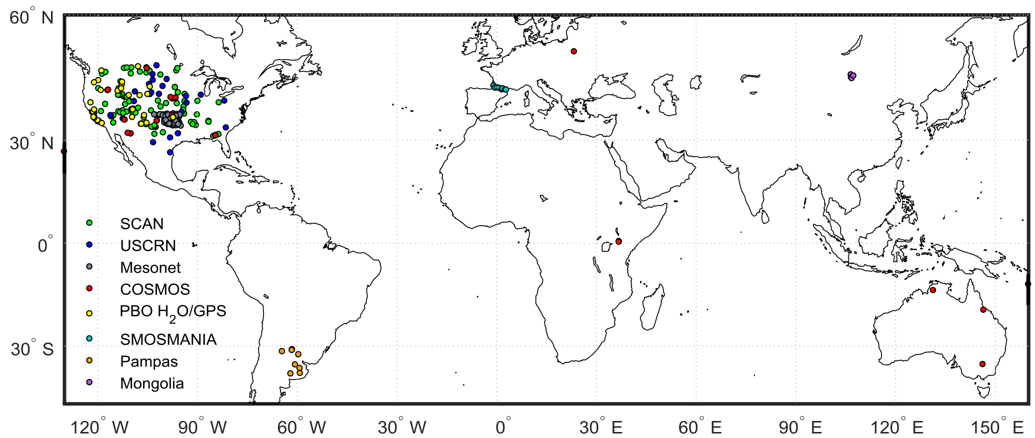

Figure 2. Location of ground observation sites (N=271) from sparse networks.

4. Validation of global TC approach

Prior to the global application of TC, we will validate aspects of the approach using ground-based observations acquired at the sparse networks shown in Figure 2. For example, it is often assumed that satellite retrievals obtained from active and passive sensors are free from error cross-correlation (ECC). As a result, the data triplets applied here consist of an active product (ASCAT), a passive product (SMOS or SMAP) and a land surface model product (ECMWF or NRv3). However, given the active-passive ECC discovered in a recent studies, it is necessary to investigate the ECC between the proposed SMAP-ASCAT and SMOS-ASCAT combinations in TC and its potential impacts on the TC-based satellite comparisons.

This investigation is made possible by adding point-scale (pts) soil moisture observations obtained at sparse networks sites (Fig. 2) into the data triplets, to obtain the data quartets [pts, SMAP, ASCAT, ECMWF] and [pts, SMOS, ASCAT, ECMWF]. Applying the least-square solution for quadruple collocation in (3) to these quartets, and assuming that non-zero ECC exists only between the active and passive soil moisture retrievals, allows us to calculate the SMAP-ASCAT and SMOS-ASCAT ECC's across the ground sites. As shown in Fig. 3 these two distributions are quite similar. That is, most sampled ECC's are positive with a median of 0.19 [-] (SMAP-ASCAT) and 0.15 [-] (SMOS-ASCAT) and an interquartile range between 0 and ~0.35 [-].

Once estimated, the impact of using of such non-zero ECC on TC results can be assessed. To this end, ASCAT R values obtained from both SMAP- and SMOS-based QC or TC analyses are averaged across all sparse sites. Since QC-generated R value takes into account the possibility of non-zero SMAP-ASCAT and SMOS-ASCAT ECC's, it is taken as a reference to evaluate the TC results. On average, TC-estimated R exhibited a slight positive bias compared with corresponding QC results, with average bias values of 0.06 and 0.05 [-] for SMAP and SMOS, respectively. Average bias for ASCAT R is 0.07 (obtained by SMAP-based TC) and 0.12 (SMOS-based TC). However, since this bias is comparable and positive for all three products, the transition from QC to TC is expected to have small net global impact on product-to-product differences. See below for additional discussion.

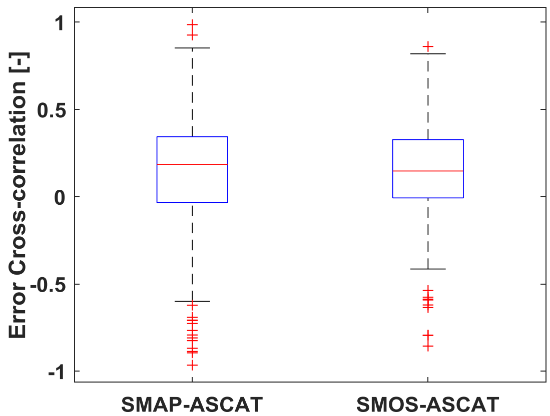

Figure 3. Distribution of ECC between SMAP-ASCAT and SMOS-ASCAT pairs estimated via the application of QC at sparse sites listed in Fig. 2. The upper and lower bounds of the boxes indicate 25th and 75th percentiles respectively and the red line in the box indicates the median. Whiskers extending from the 25th and 75th percentiles to represent 1.5 times the interquartile range.

In the TC and QC analyses above we also assume no error cross-correlation between the model and satellite products, which may not be true in all cases. For example, the SMOS soil moisture retrieval algorithm uses the ECMWF forecast temperature fields as dynamic auxiliary data input to obtain the effective soil temperature (Kerr *et al*. 2013), leading to potential ECC between the two soil moisture products. Likewise, the NASA GEOS-5 soil temperatures used in the SMAP L2_SM_P soil moisture retrieval algorithm are derived using the same GEOS-5 forward processing system that also provides the surface meteorological forcing (except precipitation) for generating NRv3. Therefore, potential ECC between SMAP and NRv3 is also of concern. An earlier study suggests small amounts of anti-correlation may exist between SMAP and NRv3 soil moisture errors that could cause slight underestimation of SMAP *R* when both datasets are used in a TC analysis (Chen *et al*. 2017). To fully address the impact of this issue on our current study, the impact of ECC on the relative evaluation of the three satellite products is examined here via both QC and TC.

Figure 4 summarizes these results. In particular, the first and second rows of Figure 4 plot the difference in correlation values (*ΔR*) between the satellite pairs obtained from TC using both ECMWF (a-c) and NRv3 (d-f) at the sparse sites. The third row (g-i) shows the ECMWF-based QC results of *ΔR*. Strong similarity in the shape of the histogram, and the values of mean *ΔR* (see dashed vertical lines) suggest that the net mean impact of potential ECC between model and passive soil moisture products is small. Furthermore, while non-zero active-passive ECC impacts absolute TC-based *R* slightly, it has very little net impact on relative *R* differences observed between SMAP, SMOS and ASCAT (compare the first and second rows against the third row in Fig. 4).

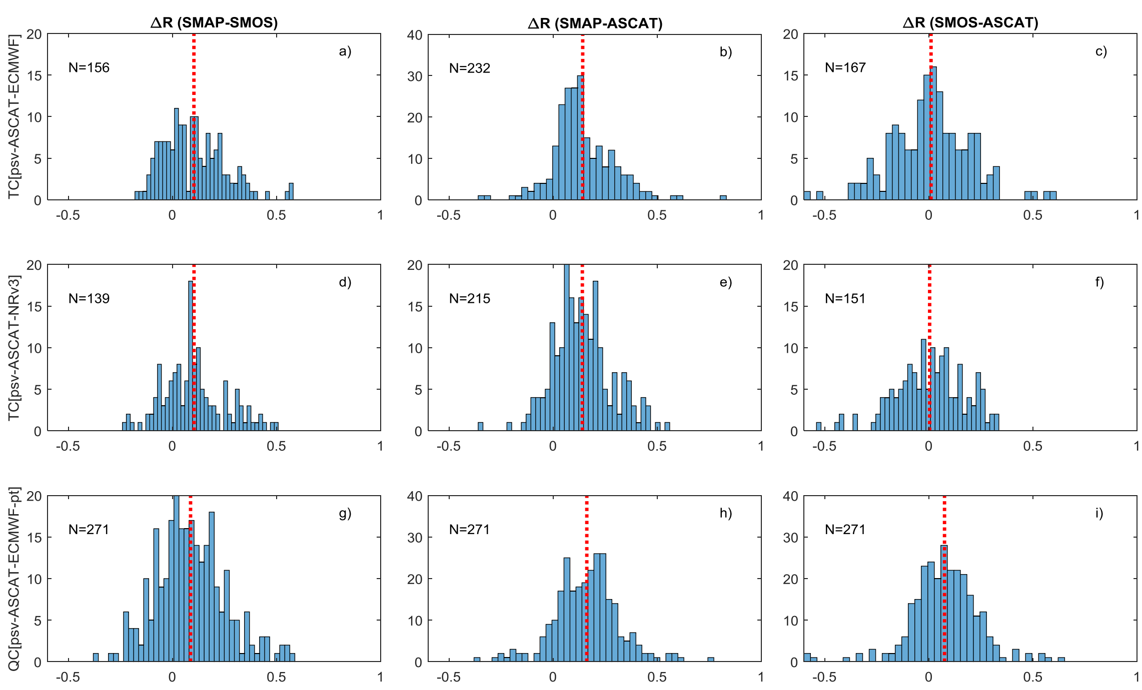

Figure 4. Comparison of differences in SMAP, SMOS and ASCAT correlation coefficients (ΔR) obtained from TC (a-f) and QC (g-i) at ground locations shown in Fig. 2. In the vertical axes, "psv" refers to passive satellite products, (SMAP or SMOS), "pt" refers to point-scale ground observations. The vertical dashed lines indicate the mean ΔR for each histogram. "N" refers to the number of stations used in each subplot.

In addition to assessing the impact of ECC on the relative global bias of TC-based R distributions (as in Fig. 3), it is useful to assess its impact on the spatial pattern of R differences observed between satellites (ΔR). Since sparse network observations are not spatially dense enough to yield continuous imagery (even after interpolation), we are restricted to the use of scatter plots when examining spatial consistency.

The spatial robustness of ΔR is examined via scatterplots comparing results obtained when utilizing different source of LSM soil moisture (Fig. 5) and QC versus TC analysis (Fig. 6). While significant sampling noise is evident, the general one-to-one correspondence suggested in Figures 5 and 6 suggest that spatially patterns present in ΔR are relatively robust to the use of competing LSM soil moisture products and the presence of ECC (accounted for in QC results but neglected in TC). While good agreement in the SMAP-SMOS ΔR and SMAP-ASCAT ΔR is observed in both cases (Fig. 5, 6), larger scatter is present in SMOS-ASCAT ΔR (Fig. 5c and 6c). This is likely due to the tendency for SMOS and ASCAT soil moisture products to exhibit relatively lower R, and thus relatively higher sampling uncertainty effects for ΔR differences, than SMAP-based results (see additional discussion in Section 5).

Therefore, across the sparse site locations in Fig. 2, relative inter-comparisons between various satellite-based soil moisture products are generally insensitive to both our choice of the collocation method (QC vs. TC) and a particular land model (ECMWF vs. NRv3).

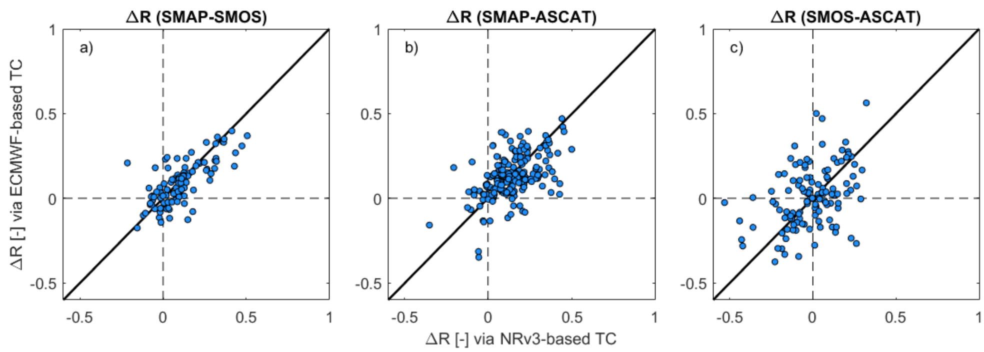

Figure 5. Comparison of *R* differences (*ΔR*) between SMAP, SMOS and ASCAT soil moisture retrievals obtained from NRv3- and ECMWF-based TC analyses. Subplots a), b) and c) include common data points shown above in Fig. 4a and 4d, Fig. 4b and 4e, and Fig. 4c and f, respectively.

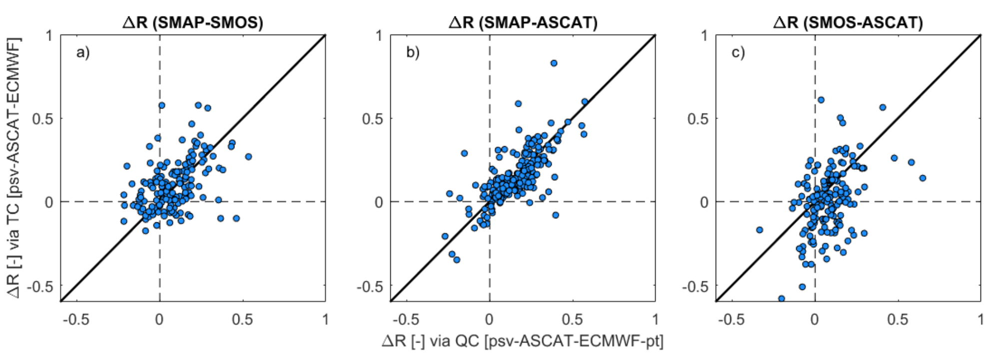

Figure 6. Comparison of *R* differences (*ΔR)* between SMAP, SMOS and ASCAT soil moisture retrievals obtained from TC and QC analyses. Subplots a), b) and c) include common data points shown above in Fig. 4a and 4g, Fig. 4b and 4h, and Fig. 4c and i, respectively.

5 Global triple collocation

QC results at ground measurement sites in Section 4 indicate that neither ECC between SMAP/SMOS and ASCAT nor ECC between the land surface model and SMAP or SMOS has a discernible impact on the inter-comparison of *R* results for SMAP, SMOS and ASCAT. Hence our strategy for a quasi-global application of TC using either a [SMOS-ASCAT-ECMWF] or [SMAP-ASCAT-ECMWF] triplet is believed to be robust. Figure 7a plots estimated *R* against true footprint surface soil moisture for SMAP, SMOS and ASCAT obtained from a TC[SMAP-ASCAT-ECMWF] (Fig.7 a, c) and TC[SMOS-ASCAT-ECMWF] (Fig. 7e) analysis. In particular, note that ASCAT results in Figure 7c are based on a TC[SMAP-ASCAT-ECMWF] analysis. Similarity of the ASCAT *R* results between the SMAP-based and SMOS-based TC analyses is shown in Fig. 8. Figures 7b, 7d and 7c show the total width of the corresponding 95% confidence interval generated from a 1,000-member moving-block bootstrap re-sampling (see Section 2.3). The global distributions of TC-based *R* results are also summarized in Fig. 8.

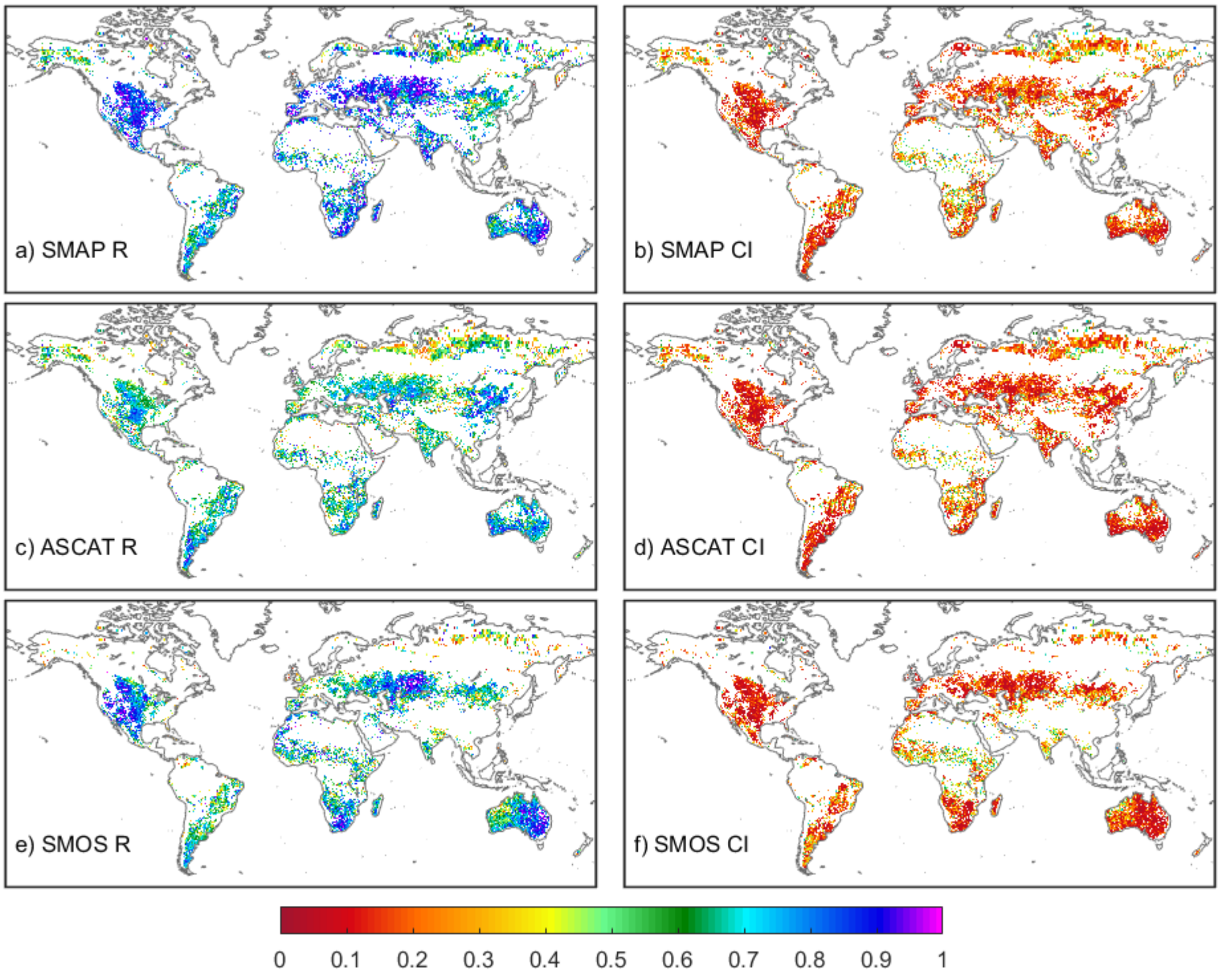

Figure 7. Quasi-global image of TC-based *R* [-] (single run, without bootstrap re-sampling) for SMAP, ASCAT and SMOS (left column: subplots a, c, e) and total width of the 95% confidence interval ('CI', right column: subplots b, d, f) derived from a 1,000-member bootstrap sampling. Subplots a) – d) are based on a [SMAP-ASCAT-ECMWF] triplet. Subplots e) - f) are based on a [SMOS-ASCAT-ECMWF] triplet.

Among the three satellite products, SMAP demonstrates the best overall performance, achieving excellent (> 0.8 [-]) *R* over the mid-latitudes of North America and Europe, as well as in southeastern Africa, India and the eastern half of Australia. Relatively good correlations (> 0.5

[-]) are found mostly elsewhere, except for parts of northern China/Mongolia and high-latitude areas of Russia where retrievals are temporally scarce due to the extended cold season.

Also retrieving from a passive radiometer, SMOS demonstrates a similar *R* pattern as SMAP, but the area of high correlation shrinks considerably in North America, Europe and Africa. SMOS also has less coverage than SMAP in the high latitudes of northern hemisphere and Asia, where correlations are relatively poor. On the other hand, SMOS has better spatial coverage and exhibits good correlations across Australia.

ASCAT presents moderate (~0.5 – 0.8 [-]) correlations in most available land pixels, and achieves higher values only in limited regions. However, higher ASCAT *R* are found in Northeastern China, where both SMAP and SMOS are out-performed by ASCAT. The 95% confidence interval (CI) (Fig. 7b, d, f) indicate relatively narrow (mostly < 0.2 [-]) ranges from Monte-Carlo simulation (i.e., small uncertainty in North America, Europe and Australia for SMAP, ASCAT and SMOS). Larger uncertainties are found in the high latitudes, tropical Africa and India, where retrieval is hindered by frequently frozen ground or high biomass. Uncertainties for SMOS are overall greater than SMAP and ASCAT over Argentina, but are smaller in South Africa.

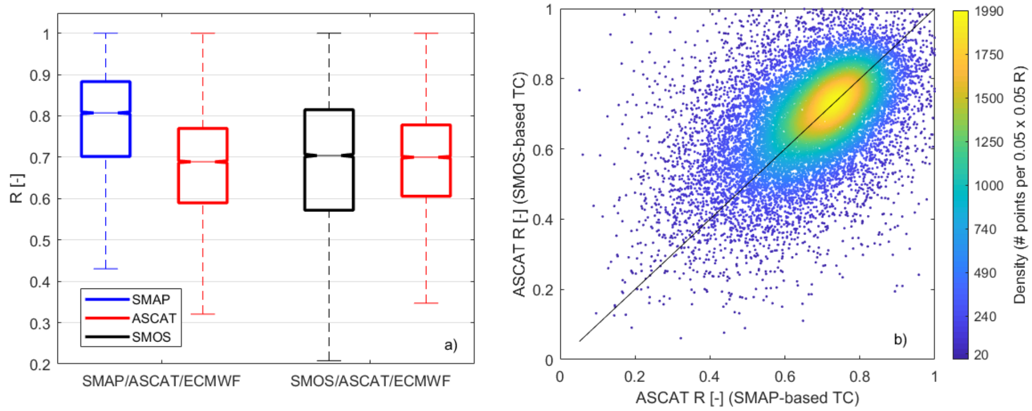

Figure 8. a) Distribution of correlation coefficients in common grid pixels (N=16,332) where both sets of single-run TC analyses (i.e., [SMAP/ASCAT/ECMWF] and [SMOS/ASCAT/ECMWF]) are available. See Fig. 3 caption for boxplot descriptions. b) Scatterplot comparison of ASCAT *R* obtained via SMAP- and SMOS-based TC.

The distribution of TC-estimated correlation values obtained globally illustrates the overall superiority of SMAP (median of ~0.8 [-]) to SMOS and ASCAT (median of ~0.7 [-]) (Fig. 8a). SMAP also presents the narrowest spread with most of its *R* values above 0.40 [-]. SMOS shows the largest spread and relatively greater number of lower values compared to SMAP and ASCAT. Note the ASCAT *R* values obtained from SMAP- and SMOS-based TC analyses are highly consistent in terms of both statistical distributions (Fig. 8a) and point-by-point comparisons (Fig. 8b). This consistency lends further support on the overall robustness of our TC approach. In particular, it suggests that the impact of non-zero ECC is nearly identical for ASCAT *R* results derived from the [SMAP-ASCAT-ECMWF] and [SMOS-ASCAT-ECMWF] triplets, and it is appropriate to simply average ASCAT *R* estimated from each triplet for comparison against SMAP and SMOS. This approach is applied later when the three remote sensing products are compared at the same time. Global-averaged *R* obtained for SMAP, SMOS and ASCAT (averaged from SMAP- and SMOS-based TC) retrievals over common pixels are 0.76, 0.66 and 0.63, respectively.

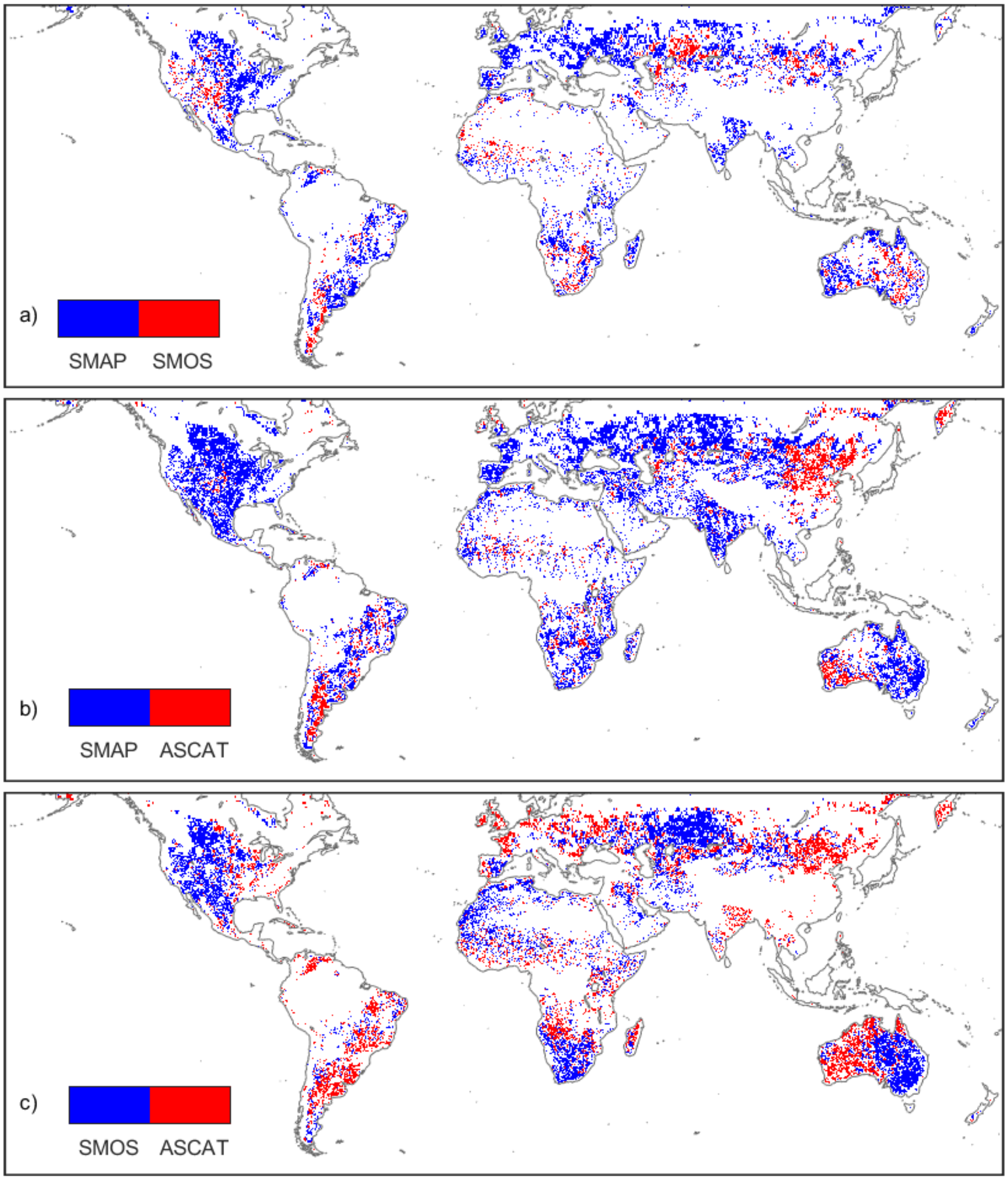

Figure 9. Comparison of TC-estimated correlation coefficients between the satellite retrieval products. Color shade indicates the product that obtains higher *R* in more than 95% of the bootstrap re-sampling

runs in a given grid cell. All areas of non-significant differences are masked. Plotted results are based on the following triplets: a) [SMAP-ASCAT-ECMWF] (for SMAP) vs. [SMOS-ASCAT-ECMWF] (for SMOS); b) [SMAP-ASCAT-ECMWF] (for SMAP and ASCAT); and c) [SMOS-ASCAT-ECMWF] (for SMOS and ASCAT).

As noted in Section 4, it is likely that *R* values in Figures 8 and 9 are uniformly biased high (on the order of 0.05 to 0.10 [-]) due to low amounts of ECC in SMAP-ASCAT and SMOS-ASCAT pairs. However, relative *R* comparisons between products are expected to be more robust. Qualitative comparisons between the satellite products are presented in Fig. 9, in which only pixels with 95% significance of comparison are shown. Superiority at 95% significance is achieved when one product has higher *R* value in more than 95% of the bootstrap re-samples. Each bootstrap replicate is treated as an independent sample and the *i*th sample TC result for SMAP is compared with the *i*th sample result for SMOS. In this way, approximately two-thirds of the pixel-wise *R* differences are identified as being significant (see Table 2).

The two L-band passive soil moisture products are compared in Fig. 9a. SMOS out-performs SMAP in areas of the Western United States, Southern Argentina, Central Asia and Eastern Australia, but 'SMAP better' pixels dominate the rest of the globe. Globally, the SMAP correlation is significantly higher than SMOS in 47% of the land pixels where comparisons are available, while SMOS is significantly higher in 14% of the pixels (Table 2). In areas of generally strong RFI pollution (e.g., Europe), the aggressive RFI mitigation efforts applied to SMAP retrievals (Mohammed *et al*. 2016; Johnson *et al*. 2016; Piepmeier *et al*. 2017) may explain their superior performance versus SMOS.

The relative performance of SMAP versus SMOS could conceivably be impacted by (somewhat arbitrary) decisions regarding data flagging and threshold for estimated quality measure. Here, the sensitivity of TC results to the SMOS data screening rules is examined by experimenting with a stricter DQX threshold (0.04 m^3/m^3). Currently a less-stringent SMOS DQX threshold (≤ 0.07 m^3/m^3) is applied in order to include more retrievals and increase the sample size for TC. As suggested in Table 2, more than 5000 pixels (or 14.4%) were removed by applying a DQX = 0.04 m^3/m^3 threshold in the TC[SMOS-ASCAT-ECMWF] analysis. Results show that the default threshold (DQX = 0.07 m^3/m^3) leads to a slight increase in the 'SMAP better' pixel classification relative to the DQX = 0.04 m^3/m^3 case (which favors SMAP in ~7% of all the commonly available pixels); however, it does not reduce the frequency of 'SMOS better' pixels as much (only ~2% pixels affected). In addition, only 0.3% of the common pixels change from a 'SMOS better' to a 'SMAP better' category when the DQX threshold is relaxed from 0.04 m^3/m^3 to 0.07 m^3/m^3. Therefore, our default DQX threshold results in only a small negative impact on SMOS performance relative to SMAP.

C-band active scatterometer retrievals from ASCAT are out-performed by SMAP in most areas except for Northeastern China, Southern Argentina and Southwestern Australia, where ASCAT retrievals demonstrate higher R (Fig. 9b). ASCAT R is significantly higher than SMAP R in only 14% of the pixels where TC results are available, while SMAP is significantly better than ASCAT at more than 50% of the available global land pixels. Note that both SMOS and ASCAT data used here were subject to processing errors due to grid transformation (to the SMAP native grid), which may cause slight under-performance and benefit SMAP in these comparisons. However, the slight global superiority of SMAP relative to SMOS is consistent with SMAP validation results at core validation sites (Chan *et al*. 2016).

The SMOS-ASCAT comparison shows a relatively even number of pixels being superior. SMOS correlation is significantly higher in most of United States, Central Asia and eastern Australia, whereas ASCAT is better in most of Northeastern China, Western Europe (areas SMOS suffers severely from RFI contamination), Argentina, and Western Australia. Considering both products being extensively validated and relatively mature, the comparison in Fig. 9c suggests that distinctive strength in each product has been firmly established in specific regions. The spatial pattern of these comparisons is largely consistent with Al-Yaari *et al*. (2014), which compared SMOSL3 and ASCAT with the Modern-Era Retrospective analysis for Research and Applications (MERRA-Land) surface soil moisture, except in Western Australia and Argentina where SMOS is found to correlate better with MERRA-Land than ASCAT.

Table 2. Pair-wise comparisons between TC-estimated correlation coefficients for various satellite products. The significance of differences is assessed using a 95% confidence threshold and the boot-strapping approach described in Section 2.3. Percentages are out of all global land pixels with viable TC estimates (see Section 2.1).

	SMAP higher		**SMOS** higher		No. pixels
	sig.	non-sig.	**sig.**	non-sig.	
SMAP vs. SMOS*	**47%**	21%	**14%**	17%	28294
SMAP vs. SMOS**	**40%**	23%	**17%**	20%	24614
	SMAP higher		**ASCAT** higher		
SMAP vs. ASCAT	**53%**	19%	**14%**	14%	39181
	SMOS higher		**ASCAT** higher		
SMOS* vs. ASCAT	**35%**	18%	**29%**	18%	36520
SMOS** vs. ASCAT	**41%**	19%	**23%**	17%	31264

* DQX $\leq 0.07\ m^3/m^3$; ** DQX $\leq 0.04\ m^3/m^3$

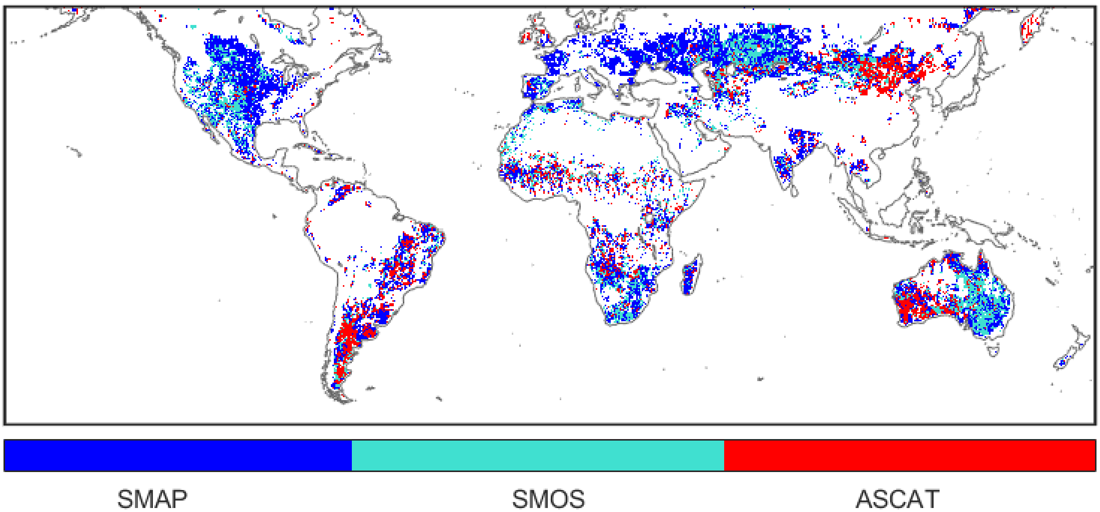

Figure 10. The satellite product (SMAP, SMOS or ASCAT) with the highest single-run TC-based correlation coefficient.

A map showing the best-performing satellite product is presented in Fig. 10. Note that regions with dense vegetation are largely masked due to a lack of successful retrievals. Likewise, in arid regions such as the Sahara Desert and Great Basin Desert, earlier studies have revealed poor or even negative correlation between active and passive products (de Jeu *et al.* 2008; Pierdicca *et al*. 2013; Burgin *et al*. 2017). This limits the area over which TC can be performed due to the masking of pixels where negative mutual correlation exists among the input datasets (see Section 2.1). As indicated above, ASCAT *R* values obtained from SMAP- and SMOS-based TC analyses are averaged for comparison. Overall, SMAP and SMOS are superior to ASCAT in most areas of North America, Europe, Southern Asia and Eastern Australia. The significant overlap of geographic regions where both passive satellites excel is generally consistent with the high level of correlation between SMAP and SMOS found earlier by Burgin *et al.* (2017). ASCAT generally performs better than SMAP and SMOS across high-latitude areas of Eastern Asia, parts of South America (mainly Argentina) and Southwestern Australia. As in Fig. 9, SMOS has higher *R* than SMAP in the Western United States, Central Asia and most inland pixels of Eastern Australia. Overall, SMAP ranks highest in 52% of the pixels with viable TC results (see Section 2.1) whereas SMOS and ASCAT each does in 24% of these pixels.

6. Summary

In this analysis, a global assessment and comparison of SMAP (L2 passive), SMOS (L3) and ASCAT (L2) surface soil moisture products is performed based on the correlation metric (*R*) obtained via triple collocation (TC). In order to produce robust TC results, *R* is estimated following removal of low-frequency variability in the soil moisture time series and therefore reflects the *R* of soil moisture anomalies relative to a 30-day moving temporal average. Given that low-frequency error sources have been previously identified in certain remotely-sensed soil moisture products (Wagner *et al*., 2014), this focus on solely high-frequency noise represents a limitation in our approach. Nevertheless, sensitivity experiments suggest that our global TC results are relatively insensitive to changing the size of the moving window from 30 to 60 days (not shown).

In addition, when comparing satellite products, it is critical to account for the sampling uncertainties due to sparse temporal availabilities or suboptimal retrieval conditions. To this end, a moving-block bootstrap re-sampling approach, with emphasis on preserving the temporal properties of the original soil moisture time series, was applied at each grid pixel to construct the confidence interval for TC estimates. The re-sampled distribution of correlation estimates is then used to obtain the significance of TC-based *R* differences between SMAP, SMOS and ASCAT soil moisture retrieval products.

Concern about the violation of TC assumption due to error cross-correlations between active-passive observations and between satellite and model products is addressed via a quadruple collocation (QC) analysis conducted within available sparse network sites (Fig. 2). Slight positive error cross-correlation is found to exist between ASCAT and both SMAP and SMOS which suggests that TC-estimated *R* for the three satellite-based products may be positively biased. However, since this bias is small and approximately equal for all three products, the relative evaluation against each other changes only slightly from QC to TC. Results also indicate limited impact associated with potential satellite-model error cross-correlations. Recent findings by Pierdicca *et al*. (2017) using a novel extended QC algorithm and 15 months of satellite and model data reveals weak SMAP-SMOS ECC that is lower than the SMAP-ASCAT ECC found. Such findings suggest the further potential of using SMAP and SMOS together in TC in future analyses. Finally, the sensitivity of SMOS TC results to the specification of the DQX threshold is shown to be low.

To the best of our knowledge, this study is the first to present a global-scale triple collocation analysis that compares the footprint-scale correlation metric of SMAP with SMOS and ASCAT soil moisture products. Results suggest that, out of these three products, SMAP has the highest global average *R* (0.76, SMOS: 0.66, ASCAT: 0.63) and is the superior product for the majority (52%) of global land pixels with a viable TC result. This finding is consistent with several recent validation studies (e.g. Kumar *et al*. 2017; Montzka *et al*. 2017; Pierdicca *et al*. 2017; Kim *et al*.

2018). For example, using information theory-based metrics, SMAP has also been found to provide higher information content than other microwave satellite soil moisture products (Kumar *et al*. 2017). Likewise, in a validation study applying both standard validation methods and triple collocation at footprint-scale soil moisture measurements from the Cosmic Ray Neutron Probes (CRNP, including some of the COSMOS stations used here) across five continents, SMAP outperformed other satellite products including AMSR2, SMOS and ASCAT (Montzka *et al*. 2017). Nevertheless, each of the three satellite retrieval products (SMAP, SMOS and ASCAT) were found to be superior (to the other two) in specific global land regions. Therefore, the global inter-comparison maps in Figures 9 and 10 provide useful information for regional-scale applications such as the choice of dataset for assimilation into rainfall-runoff models.

In closing, it should be noted that all products considered here are subject to frequent re-processing and algorithm improvements. For example, a new global daily SMOS SM product -- the SMOS-INRA-CESBIO (SMOS-IC) product was recently released and shown to yield generally higher correlations versus ground observation versus the v300 SMOS Level 3 soil moisture product considered here (Fernandez-Moran *et al*., 2017). Comparable enhanced SMAP soil moisture products are likely to arise in the foreseeable future. Therefore, the cross evaluation efforts described here are, in reality, an on-going effort requiring updating as improved products are released.

Acknowledgements

ECMWF soil moisture field was provided by European Centre for Medium-range Weather Forecasts. Nature Run soil moisture field was provided by NASA Global Modeling and Assimilation Office. Ground soil moisture measurements were contributed by individual networks as SMAP Calibration/Validation partners. Research was primarily supported by the National Aeronautics and Space Administration (NASA) Soil Moisture Active/Passive (via Wade Crow's membership on the SMAP Science Team). Addition support was provided by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

References

Al-Yaari, A., Wigneron, J.-P., Ducharne, A., Kerr, Y. H., Wagner, W., De Lannoy, G., Reichle, R., Al Bitar, A., Dorigo, W., Richaume, P., and Mialon, A. (2014). Global-scale comparison of passive (SMOS) and active (ASCAT) satellite based microwave soil moisture retrievals with soil moisture simulations (MERRA-Land), *Remote Sens. Environ.*, 152, 614-626.

Balsamo, G., Viterbo, P., Beljaars, A. C. M., van den Hurk, B. J. J. M., Hirschi, M., Betts, A. K. and Scipal, K. (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the ECMWF-IFS, *J. Hydrometeor.*, 10, 623–643.

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrinmore, J. H., Leeper, R. D., Hall, M. E., Kochendorfer, J., Meyers, T. P., Wilson, T. and Diamond, H. J. (2013). U.S. Climate Reference Network soil moisture and temperature observations, *J. Hydrometeor*, 14(3), 977–988.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., and Hasenauer, S. (2010). ASCAT soil wetness index validation through in situ and modeled soil moisture data in central Italy, *Remote Sens. Environ.*, 114(11), 2745-2755.

Brodzik, M. J. and Knowles, K. W. (2002). EASE-Grid: a versatile set of equal-area projections and grids in Goodchild, M. (Ed.) *Discrete Global Grids*. National Center for Geographic Information & Analysis. Santa Barbara, CA, USA.

Burgin, M., Colliander, A., Njoku, E. G., Chan, S. K., Francois, C., Kerr, Y. H., Bindlish, R., Jackson, T. J., Entekhabi, D., Yueh, S. H. (2017). A comparative study of the SMAP passive soil moisture product with existing satellite-based soil moisture products, *IEEE Trans. Geosci. Remote Sens.*, 55(5), 2959-2971.

Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., and Piguet, B. (2007) In situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network, *2007 IEEE Int. Geosci. Remote Sens. Symposium*, Barcelona, Spain, 1196-1199.

Chan S., and Dunbar, R. S. (2015). SMAP Level 2 passive soil moisture product specification document, JPL D-72547, Jet Propulsion Laboratory, Pasadena, CA, USA. Available: https://nsidc.org/sites/nsidc.org/files/technical-references/SMAP%20L2_SM_P%20Beta-Level%20PSD%20%28PRIMARY%29.pdf

Chan, S. K., Bindlish, R., O'Neill, P. E., Njoku, E., Jackson, T. J., Colliander, A., Chen, F., … Kerr, Y. (2016). Assessment of the SMAP Passive Soil Moisture Product, *IEEE Trans. Geosci. Remote Sens.*, 54. 1-14.

Chen, F., Crow, W. T., Colliander, A., Cosh, M., Jackson, T. J., Bindlish, R., and Reichle, R. (2017). Application of triple collocation in ground-based validation of soil moisture active/passive (SMAP) data products. *IEEE J. Sel. Topics Appl. Earth Obs. Rem. Sens*., 10 (2), 489-502.

Chew, C. C., Small, E. E., Larson, K. M., and Zavorotny, V. U. (2014). Effects of near-surface soil moisture on GPS SNR data: development of a retrieval algorithm for soil moisture, *IEEE Trans Geosci Remote Sens*, 52, 537–543.

Colliander, A., Jackson, T. J., Bindlish, R., Chan, S., Das, N., Kim, S.B., … Yueh, S. (2017). Validation of SMAP surface soil moisture products with core validation sites, *Remote Sens. Environ.*, 191, 215-231.

de Jeu, R.A.M., Wagner, W., Holmes, T.R.H., Dolman, A.J., van de Giesen, N.C., and Friesen, J. (2008) Global soil moisture patterns observed by space borne microwave radiometers and scatterometeters, *Surveys in Geophysics*, 29, 399-420. doi:10.1007/s10712-008-9044-0.

De Rosnay, P., Balsamo, G., Albergel, C., Muñoz-Sabater, J., and Isaksen, L. (2012). Initialisation of land surface variables for Numerical Weather Prediction. *Surv. Geophys.* doi:10.1007/s10712-012-9207-x.

Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., and Naeimi, V. (2010). Error characterisation of global active and passive soil moisture datasets, *Hydrol. Earth Syst. Sci.*, 14, 2605-2616.

Draper, C., Reichle, R., de Jeu, R., Naeimi, V., Parinuss, R., and Wagner W. (2013). Estimating root mean square errors in remotely sensed soil moisture over continental scale domains, *Remote Sens. Environ.*,137, 288-298.

Drusch, M., de Rosnay, P., Balsamo, G., Andersson, E., Bougeault, P., and Viterbo, P. (2009). Towards a Kalman filter based soil moisture analysis system for the operational ECMWF Integrated Forecast System, *Geophys. Res. Lett.*, 36, *L10401* doi:10.1029/2009GL037716.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., … van Zyl, J. (2010). The Soil Moisture Active Passive (SMAP) mission, *Proc. IEEE*, 98(5), 704–716.

Fernandez-Moran, R., Al-Yaari, A., Mialon, A., Mahmoodi, A., Al Bitar, A., De Lannoy, G., Rodriguez-Fernandez, N., Lopez-Baeza, E., Kerr, Y., and Wigneron, J.-P. (2017). SMOS-IC: an alternative SMOS soil moisture and vegetation optical depth product. *Remote Sens.*, 9 (5), 457, doi:10.3390/rs9050457.

González-Zamora, A., Sánchez, N., Gumuzzio, A., Piles, M., Olmedo, E., and Martínez-Fernández, J. (2015). Validation of SMOS L2 and L3 soil moisture products over the Duero basin at different spatial scales, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-7/W3, 1183-1188.

Gruber, A., Su, C.-H., Zwieback, S., Crow, W. T., Dorigo, W., and Wagner, W. (2016a). Recent advances in (soil moisture) triple collocation analysis, *Int. J. Appl. Earth Obs. Geoinf.*, 45, part B, 200-211.

Gruber, A., Su, C.-H., Crow, W. T., Zwieback, S., Dorigo, W. A., and Wagner, W. (2016b), Estimating error cross-correlation in soil moisture data sets using extended collocation analysis, *J. Geophys. Res. Atmos*., 121, 1208-1219.

Illston, R., Basara, J., Fisher, D., Elliott, R., Fiebrich, C., Crawford, K., Humes, K. and Hunt, E. (2008). Mesoscale monitoring of soil moisture across a statewide network, *J. Atmos. Oceanic Technol.*, 25: 167-182.

Jackson, T. J., Cosh, M. H., Bindlish, R., Starks, P. J., Bosch, D. D., Seyfried, M., Goodrich, D. C., Moran, M. S., and Du, J. (2010). Validation of Advanced Microwave Scanning Radiometer soil moisture products, *IEEE Trans. Geosci. Remote Sens.*,48(12), 4256-4272.

Jackson, T. J., O'Neill, P., Chan, S., Bindlish, R., Colliander, A., Chen, F., … Entekhabi, D. (2016). Calibration and Validation for the L2/3_SM_P Version 4 and L2/3_SM_P_E Version 1 Data Products, SMAP Project, JPL D-56297, Jet Propulsion Laboratory, Pasadena, CA, USA. available: https://nsidc.org/sites/nsidc.org/files/files/D56297%20SMAP%20L2_SM_P_E%20Assessment%20Report(1).pdf

Johnson, J. T., Mohammed, P. N., Piepmeier, J. R., Bringer, A., and Aksoy, M. (2016). Soil Moisture Active Passive (SMAP) microwave radiometer radio-frequency interference (RFI) mitigation: Algorithm updates and performance assessment, *2016 IEEE Int. Geosci. Remote Sens. Symposium*, Beijing, China, 123-124.

Kaihotsu, I., Koike, T., Yamanaka, T., Fujii, H., Ohta, T., Tamagawa, K., Oyunbaatar, D., and Akiyama, R. (2009). Validation of Soil Moisture Estimation by AMSR-E in the Mongolian Plateau, *J. Remote Sens. Soc. Japan*, 29. 271-281.

Kerr, Y. H., Al-Yaari, A., Rodriguez-Fernandez, N., Parrens, M., Molero, B., Leroux, D., Bircher, S., Mahmoodi, A., Mialon, A., Richaume, P., Delwart, S., Al Bitar, A., Pellarin, T., Bindlish, R., Jackson, T. J., Rüdiger, C., Waldteufel, P., Mecklenburg, S., and Wigneron J.-P. (2016). Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation, *Remote Sens. Environ.*, 180, 40-63.

Kerr, Y. H., Jacquette, E., Al Bitar, A., Cabot, F., Mialon, A., Richaume, P., and Berthon, L. (2013). In CBSA (Ed.), CATDS SMOS L3 Soil Moisture Retrieval Processor Algorithm Theoretical Baseline Document (ATBD) CBSA, Technical Note (pp. 73). Toulouse: CESBIO.

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J. M., Font, J., and Berger, M. (2001). Soil moisture retrieval from space: the soil moisture and ocean salinity (SMOS) mission, *IEEE Trans. Geosci. Remote Sens.*, 39, 1729–1735.

Kim, H., Parinussa, R., Konings, A.G., Wagner, W., Cosh, M.H., Lakshmi, V., Zohaib, M., and Choi, M. (2018). Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products, *Remote Sens. Environ*., 204, 260-275, doi:10.1016/j.rse.2017.10.026.

Köhli, M., Schrön, M., Zreda, M., Schmidt, U., Dietrich, P., and Zacharias, S. (2015). Footprint characteristics revised for field-scale soil moisture monitoring with cosmic-ray neutrons, *Water Resour. Res*., 51, 5772–5790.

Kumar, S. V., Dirmeyer, P. A., Peters-Lidard, C. D., Bindlish, R., and Bolten, J. (2017). Information theoretic evaluation of satellite soil moisture retrievals, *Remote Sens. Environ.*, in press, doi:10.1016/j.rse.2017.10.016.

Larson, K. M., and Nievinski, F. G. (2013). GPS snow sensing: results from the Earthscope Plate Boundary Observatory, *GPS Solut*., 17, 41–52.

Larson, K. M., Small, E. E., Gutmann, E., Bilich, A., Braun, J., and Zavorotny, V. (2008). Use of GPS receivers as a soil moisture network for water cycle studies, *Geophys. Res. Lett*., 35, L24405, doi:10.1029/2008GL036013.

Leroux, D. J., Kerr, Y., Richaume, P., and Fieuzal, R. (2013). Spatial distribution and possible sources of SMOS errors at the global scale, *Remote Sens. Environ.*, 133, 240-250.

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, *Geophys. Res. Lett*., 41(17), 6229-6236.

Miralles, D. G., Crow, W. T., and Cosh, M. H. (2010). Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations, *J. Hydromeorol.*, 11(6), 1423-1429.

Mohammed, P. N., Aksoy, M., Piepmeier, J. R., Johnson, J. T. and Bringer, A. (2016) SMAP L-Band Microwave Radiometer: RFI Mitigation Prelaunch Analysis and First Year On-Orbit Observations," in *IEEE Trans. Geosci. Remote Sens*., 54(10), 6035-6047.

Montzka, C., Bogena, H.R., Zreda, M., Monerris, A., Morrison, R., Muddu, S., and Vereecken, H. (2017). Validation of spaceborne and modelled surface soil moisture products with Cosmic-Ray Neutron Probes. *Remote Sens.*, 9(2), 103, doi:10.3390/rs9020103.

Mudelsee, M. (2002). TAUEST: a computer program for estimating persistence in unevenly spaced weather/climate time series, *Comput Geosci*, 28(1), 69–72.

Mudelsee, M. (2010). Climate time series analysis: classical statistical and bootstrap methods. Springer, Dordrecht Heidelberg London New York, 474pp.

Naeimi, V., Scipal, K., Bartalis, Z., and Wagner, W. (2009). An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations, *IEEE Trans. Geosci. Remote Sens*., 47(7), 1999-2013.

Ólafsdóttir, K.B., and Mudelsee, M. (2014). More accurate, calibrated bootstrap confidence intervals for estimating the correlation between two time series, *Math. Geosci.*, 46(4), 411–427.

Paulik, C., Dorigo, W., Wagner, W., and Kidd R. (2014). Validation of the ASCAT soil water index using in situ data from the international soil moisture network, *Int. J. Appl. Earth Observation Geoinf*, 30, 1-8.

Piepmeier, J. R., Focardi, P., Horgan, K. A., Knuble, J., Ehsan, N., Lucey, J., … and Njoku, E. G. (2017) SMAP L-Band Microwave Radiometer: Instrument Design and First Year on Orbit, *IEEE Trans. Geosci. Remote Sens*., 55(4), 1954-1966.

Pierdicca, N., Pulvirenti, L., Fascetti, F., Crapolicchio R., and Talone, M. (2013). Analysis of two years of ASCAT-and SMOS-derived soil moisture estimates over Europe and North Africa, European J. Remote Sens., 46:1, 759-773, doi:10.5721/EuJRS20134645.

Pierdicca, N., Fascetti, F., Pulvirenti, L., Crapolicchio, R., and Muñoz-Sabater, J. (2015). Quadruple Collocation Analysis for Soil Moisture Product Assessment, *IEEE Geosci. Remote Sens. Lett.*, 12(8), 1595-1599.

Pierdicca, N., Fascetti, F., Pulvirenti, L., and Crapolicchio, R. (2017). Error characterization of soil moisture satellite products: retrieving error cross-correlation through extended quadruple collocation, *IEEE J. Sel. Topics Appl. Earth Obs. Rem. Sens*.,10, 4552-4530, doi:10.1109/JSTARS.2017.2714025.

Piles, M., Sánchez, N., Vall-llossera, M., Camps, A., Martínez-Fernández, J., Martínez, J., and González-Gambau, V. (2014). A Downscaling Approach for SMOS Land Observations: Evaluation of High-Resolution Soil Moisture Maps Over the Iberian Peninsula, *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 7(9), 3845-3857.

Polcher, J., Piles, M., Gelati, E., Barella-Ortiz, A., and Tello, M. (2016). Comparing surface-soil moisture from the SMOS mission and the ORCHIDEE land-surface model over the Iberian Peninsula, *Remote Sens. Environ.*, 174, 69-81.

Reichle, R.H., **Crow, W.T.**, Koster, R. D., Sharif, H. and Mahanama, S. (2008). Contribution of soil moisture retrievals to land data assimilation products. *Geophys. Res. Lett.*, 35. L01404, doi:10.1029/2007GL031986.

Reichle, R. H., De Lannoy, G. J. M., Liu, Q., Ardizzone, J. V., Chen, F., Colliander, A., Conaty, A., Crow, W., Jackson, T., Kimball, J., Koster, R. D., and Smith, E. B. (2016). Soil Moisture Active Passive Mission L4_SM Data Product Assessment (Version 2 Validated Release). GMAO Office Note No. 12 (Version 1.0), 55 pp, NASA Goddard Space Flight Center, Greenbelt, MD, USA. Available: http://gmao.gsfc.nasa.gov/pubs/office_notes.

Scott, B., Ochsner, T., Illston, B., Fiebrich, C., Basara, J. and Sutherland, A. (2013). New soil property database improves Oklahoma Mesonet soil moisture estimates, *J. Atmos. Oceanic Technol.*, 30, 2585–2595.

Shaefer, G. L., Cosh, M. H., and Jackson, T. J. (2007). The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN), *J. Atmos. Oceanic Technol*., 24, 2073-2077.

Stoffelen, A. (1998). Toward the true near-surface wind speed: error modeling and calibration using triple collocation, *J. Geophys. Res*., *103*(C4), 7755-7766.

Su, C.-H. and Ryu, D. (2015). Multi-scale analysis of bias correction of soil moisture, *Hydrol. Earth Syst. Sci*., 19, no. 1,17-31.

von Storch, H., and Zwiers, F. W. (1999). Statistical analysis in climate research, Cambridge University Press, Cambridge, UK, 484pp.

Wagner, W., Lemoine, G., and Rott, H. (1999). A method for estimating soil moisture from ERS scatterometer and soil data, *Remote Sens. Environ.*, 70(2), 191-207.

Wagner, W., Brocca, L., Naeimi, V., Reichle, R., Draper, C., de Jeu, R., Ryu, D., Su, C. H., Western, A., Calvet, J. C., Kerr, Y. H., Leroux, D. J., Drusch, M., Jackson, T. J., Hahn, S., Dorigo, W., and Paulik, C. (2014). Clarifications on the "Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S.", *IEEE Trans. Geosci. Remote Sens*., 52(3), 1901-1906.

Yilmaz, M. T., and Crow, W. T. (2014). Evaluation of assumptions in soil moisture triple collocation analysis, *J. Hydrometeor*., 15(3), 1293–1302.

Zreda, M., Desilets, D., Ferré, T. P. A., and Scott, R. L. (2008). Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons, *Geophys. Res. Lett.*, 35, L21402, doi:10.1029/2008GL035655.

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., and Rosolem, R. (2012). COSMOS: The COsmic-ray Soil Moisture Observing System, *Hydrol. Earth Syst. Sci*., 16(11), 4079–4099, doi:10.5194/hess-16-4079-2012.

Zwieback, S., Scipal, K., Dorigo, W., and Wagner, W. (2012). Structural and statistical properties of the collocation technique for error characterization, *Nonlin. Processes Geophys*., 19, 69-80.

Zwieback, S., Scipal, K., Dorigo, W., and Wagner, W. (2012). Structural and statistical properties of the collocation technique for error characterization, *Nonlin. Processes Geophys*., 19, 69-80.

Zwiers, F. W. (1990). the effect of serial correlation on statistical inferences made with resampling procedures, *J. Climate,* 3, 1452–1461.

List of Figure Captions

Figure 1. Schematic diagram of moving block bootstrap sampling on collocated, temporally uneven triple-soil-moisture-product time series with an l_{opt} of 7. Overlapping data blocks from the original time series (top) are drawn randomly with replacement and then concatenated to generate a new bootstrap resample (bottom).

Figure 2. Location of ground observation sites (N=271) from sparse networks.

Figure 3. Distribution of ECC between SMAP-ASCAT and SMOS-ASCAT pairs estimated via the application of QC at sparse sites listed in Fig. 2. The upper and lower bounds of the boxes indicate 25th and 75th percentiles respectively and the red line in the box indicates the median. Whiskers extending from the 25th and 75th percentiles to represent 1.5 times the interquartile range.

Figure 4. Comparison of differences in SMAP, SMOS and ASCAT correlation coefficients (ΔR) obtained from TC (a-f) and QC (g-i) at ground locations shown in Fig. 2. In the vertical axes, "psv" refers to passive satellite products, (SMAP or SMOS), "pt" refers to point-scale ground observations. The vertical dashed lines indicate the mean ΔR for each histogram. Number of stations used in each subplot is shown as "N".

Figure 5. Comparison of ΔR (same as in Fig. 4) obtained from NRv3- and ECMWF-based TC analyses. Subplots a), b) and c) include common data points in Fig. 4a and 4d, Fig. 4b and 4e, and Fig. 4c and f, respectively.

Figure 6. Comparison of ΔR (same as in Fig. 4) obtained from TC and QC analyses. Subplots a), b) and c) include common data points in Fig. 4a and 4g, Fig. 4b and 4h, and Fig. 4c and i, respectively.

Figure 7. Quasi-global image of TC-based *R* [-] (single run, without bootstrap re-sampling) for SMAP, ASCAT and SMOS (left column: subplots a, c, e) and total width of the 95% confidence interval ('CI', right column: subplots b, d, f) derived from a 1,000-member bootstrap sampling. Subplots a) – d) are based on a [SMAP-ASCAT-ECMWF] triplet. Subplots e) - f) are based on a [SMOS-ASCAT-ECMWF] triplet.

Figure 8. a) Distribution of correlation coefficients (from single triple collocations runs) in common grid pixels (N=16,332) where both sets of TC analyses [SMAP/ASCAT/ECMWF and SMOS/ASCAT/ECMWF] are available (see Fig. 2 caption for boxplot descriptions); b) comparison of ASCAT *R* obtained via SMAP- and SMOS-based TC analyses.

Figure 9. Comparison of TC-estimated correlation coefficients between the satellite retrieval products. Color shade indicates the product that obtains higher *R* in more than 95% of the bootstrap re-sampling runs in a given grid cell. All areas of non-significant differences are masked. Plotted results are based on the following triplets: a) [SMAP-ASCAT-ECMWF] (for SMAP) vs. [SMOS-ASCAT-ECMWF] (for

SMOS); b) [SMAP-ASCAT-ECMWF] (for SMAP and ASCAT); and c) [SMOS-ASCAT-ECMWF] (for SMOS and ASCAT).

Figure 10. The satellite product (SMAP, SMOS or ASCAT) with the highest TC-based correlation coefficient ($\overline{R}$, bootstrap mean).