Remote Sensing of Environment 165 (2015) 109–122

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Reducing multiplicative bias of satellite soil moisture retrievals

Kurt C. Kornelsen [a,*], Paulin Coulibaly [a,b]

[a] School of Geography and Earth Science, McMaster University, Hamilton, Ontario L8S 4L8, Canada
[b] Department of Civil Engineering, McMaster University, Hamilton, Ontario L8S 4L8, Canada

ARTICLE INFO

Article history:
Received 21 October 2014
Received in revised form 17 April 2015
Accepted 27 April 2015
Available online xxxx

Keywords:
SMOS
Cumulative distribution function
Renormalization
Validation
Bias correction

ABSTRACT

Soil moisture is a principal component of the Earth's climate and hydrological systems that is difficult to monitor and model due to high variability, uncertainty in land surface characterization and uncertainty in soil moisture forcing. Satellite soil moisture retrievals and brightness temperature observations, such as those available from the Soil Moisture and Ocean Salinity (SMOS) mission, can be a valuable source of information for data assimilation and merging with other satellite retrieval datasets. To correct for biases in these data sets, bias correction methods such as cumulative distribution function (CDF) matching, linear rescaling and copulas are used to map satellite soil moisture climatology to that of in situ or model values. This study compared SMOS retrievals to soil moisture observations from the SCAN network for a calibration period of 2010–2011 and validation period of 2012–2013 before and after bias correction. The focus of the study was on the presence and removal of multiplicative bias when comparing SMOS retrievals to SCAN data. Additive bias between SMOS retrievals and SCAN observations was removed by standard bias correction techniques and a new resampling approach was found to reduce multiplicative biases. In addition, the new bias correction technique was found very competitive with the benchmark methods for both the calibration and validation periods.

1. Introduction

Soil moisture is a principal component of the Earth's climate and hydrological systems. The state of moisture in the soil controls the hydrological and energy interactions between the atmosphere, vegetation and soil at the Earth's surface, which drives the balances of water and energy. For this reason the state of soil moisture is important for both scientific and operational applications such as numerical weather prediction (Drusch, 2007), flood forecasting (Berthet, Andréassian, Perrin, & Javelle, 2009; Bronstert et al., 2012), and climate modeling (Seneviratne et al., 2013).

Monitoring and predicting soil moisture for scientific and operational purposes is a difficult task since in situ networks have poor spatial coverage and models suffer from errors in meteorological forcing, land surface characterization and simplifications of process descriptions. Recently, indirect measurements of soil moisture have become available through the use of active and passive microwave remote sensing on various platforms. Wagner, Lemoine, and Rott (1999) presented a change detection algorithm for the active microwave Advanced SCATterometer (ASCAT) on-board the MetOp (meteorological operational) satellite and several soil moisture retrieval algorithms have been implemented for the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) on-board the National Aeronautics and Space Administration's (NASA) Aqua satellite (Njoku, Jackson, Lakshmi, Chan, & Nghiem, 2003; Owe, de Jeu, & Holmes, 2008; Owe, de Jeu, & Walker, 2001). On the 2nd of November 2009, the European Space Agency (ESA) launched the Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al., 2010), the first dedicated soil moisture satellite, and in January 2015, NASA launched the Soil Moisture Active Passive (SMAP) satellite (Entekhabi et al., 2010). Both SMOS and SMAP have L-band radiometers which are considered ideal for the monitoring of soil moisture as they penetrate both the atmosphere and thin vegetation (Entekhabi et al., 2010; Kerr et al., 2010). SMOS and SMAP operate in a protected wavelength which should have minimal radio frequency interference (RFI), although in operation this has not been found to be the case (Oliva et al., 2012).

Validation of satellite soil moisture retrievals is an important step in understanding the quality of retrieval results and characterizing errors that may be present. With respect to SMOS validation, the quality of SMOS soil moisture retrievals have been evaluated using in situ networks in Spain (Sánchez, Martínez-Fernández, Scaini, & Pérez-Gutiérrez, 2012), Germany (Dall'Amico, Schlenz, Loew, & Mauser, 2012), Denmark (Bircher, Skou, Jensen, Walker, & Rasmussen, 2012), Italy and Luxembourg (Lacava et al., 2012), Canada (Gherboudj et al., 2012), the United States (Al Bitar et al., 2012; Collow, Robock, Basara, & Illsont, 2012; Jackson et al., 2012), and Australia (Su, Ryu, Young, Western, & Wagner, 2013). Large scale evaluation of SMOS data has also been conducted by comparing SMOS soil moisture retrievals to soil moisture products from ASCAT and AMSR-E as well as land surface data assimilation system (LDAS) outputs (Al-Yaari et al., 2014; Leroux, Kerr, Richaume, & Fieuzal, 2013) and by determining the impact of

* Corresponding author.
E-mail address: kornelkc@mcmaster.ca (K.C. Kornelsen).

http://dx.doi.org/10.1016/j.rse.2015.04.031
0034-4257/© 2015 Elsevier Inc. All rights reserved.

SMOS retrieved soil moisture in a simple data assimilation system (Pan et al., 2012). Similar efforts have been made for soil moisture products from ASCAT (i.e. Bartalis et al., 2007; Brocca et al., 2011) and AMSR-E (i.e. Brocca et al., 2011; Jackson et al., 2012; Pan et al., 2012). Synthesis of these various results reveals that SMOS soil moisture products have good temporal correlation to observed or modeled soil moisture (Al Bitar et al., 2012; Jackson et al., 2012; Lacava et al., 2012; Su et al., 2013) and particularly outperform other soil moisture retrievals as vegetation density increases (Al-Yaari et al., 2014; Pan et al., 2012). Under nominal conditions, SMOS retrievals are close to meeting the target root mean squared error (RMSE) of 0.04 $m^3 m^{-3}$ (Al Bitar et al., 2012; Jackson et al., 2012; Sánchez et al., 2012) although a persistent bias in SMOS, and other, soil moisture retrievals is a consistent issue that remains to be addressed (Al Bitar et al., 2012; Jackson et al., 2012; Sánchez et al., 2012; Su et al., 2013) and in some cases may be informative of algorithm performance (Jackson et al., 2012).

Biases and systematic differences between satellite retrieved soil moisture and the reference soil moisture are problematic for many applications such as LDAS and the blending of soil moisture products. For LDAS applications, several studies have shown that the assimilation of satellite data can improve the characterization of the surface states (Albergel et al., 2011; Das, Mohanty, Cosh, & Jackson, 2008; de Rosnay et al., 2013; Draper, Reichle, De Lannoy, & Liu, 2012; Reichle, Crow, & Keppenne, 2008; Reichle & Koster, 2005), however, a fundamental assumption of the Kalman filter, and many of its derivatives, is that observation noise is mean zero Gaussian with a given covariance $\mathbf{R}_k$. There is a similar requirement when blending several satellite soil moisture products to generate products representative of soil moisture climatology (Leroux et al., 2014; Liu et al., 2011, 2012; Yilmaz, Crow, Anderson, & Hain, 2012). Therefore, the correction of bias is an important step, where the temporal characteristics of the satellite data products are assumed to be more important than the absolute retrieval values. To correct for the presence of bias and variance errors, Reichle and Koster (2004) and Drusch, Wood, and Gao (2005) proposed matching the cumulative distribution function (CDF) of observed satellite data to the model climatology as observation operators for the direct assimilation of satellite soil moisture. This technique has been adopted in many studies for comparison of soil moisture retrieval performances (Brocca, Melone, Moramarco, Wagner, & Hasenauer, 2010; Lacava et al., 2012; Su et al., 2013), data assimilation (Crow & van den Berg, 2010; Draper, Mahfouf, Calvet, Martin, & Wagner, 2011) and for blending AMSR-E, ASCAT and other soil moisture products (Liu et al., 2011, 2012). Linear rescaling has been used to correct the climatology of satellite soil moisture data where the mean and standard deviation of the satellite soil moisture are rescaled to match that of the in situ reference dataset (Brocca et al., 2010; Draper, Walker, Steinle, de Jeu, & Holmes, 2009; Su et al., 2013). A third method of bias correction is the use of copulas, where a copula function is used to model the dependence between two distribution functions (Gao, Wood, Drusch, & McCabe, 2007; Leroux et al., 2014). All three bias correction techniques assume that the reference dataset does not contain noise and errors in the retrieved soil moisture can be ignored in the bias correction technique. Not accounting for the possibility of errors in the data will be shown to result in undesirable conditions such as multiplicative bias in the corrected dataset. Multiplicative bias may result in systematic under (over)-estimation of retrieved soil moisture masking the desired temporal characteristics.

The purpose of this study is to evaluate the presence of, and propose a correction for multiplicative bias in satellite retrieved and renormalized soil moisture. While the analysis will be presented using SMOS retrieved soil moisture over Soil Climate Analysis Network (SCAN) (Schaefer, Cosh, & Jackson, 2007) sites in the continental U.S., it is expected the results are indicative of errors which may be present in other applications and with other sensors (Su et al., 2013). The SCAN sites chosen have been previously selected and validated in a node-site comparison by Al Bitar et al. (2012). To expand upon the previous validation and identify the types of error present, this study will renormalize the SMOS soil moisture to that of the concurrent SCAN observations. An analysis of the difference between the two data sets will be used to demonstrate the presence of multiplicative bias before and after the climatology of the SMOS data have been matched to that of in situ observations. An assessment of the robustness of bias correction methods will also be made by temporal cross-validation, where the correction parameters will be calibrated for retrievals made during calendar years 2010–2011 and validated during 2012–2013.

2. Study areas and soil moisture measurements

The SCAN network was designed by the Natural Resources Conservation Service (NRCS) to support natural resources assessments, conservation and water resources management within the U.S. In contrast to many networks which cover a limited spatial extent, SCAN sites have been placed at selected sites distributed around the United States to collect hourly atmospheric, soil moisture and soil temperature data in different climate, physiographic and soil regions (Schaefer et al., 2007). Soil moisture at each SCAN site is collected by Stevens Hydra Probes at depths of approximately 5, 10, 20, 50 and 100 cm (Schaefer et al., 2007), of which the ~5 cm depth is analyzed herein. After filtering for SCAN sites that had sensors at the 5 cm depth and were not located in areas of strong topography or dense vegetation Al Bitar et al. (2012) focused their attention on 13 SCAN sites and 4 SNOwpack TELemetry (SNOTEL) sites for SMOS validation and analysis, of which data from the SCAN sites are re-visited herein. The use of the same study sites allows for insights gained from the previous work to be applied in this study. The distribution of the selected sites and some basic information about the sites can be found in Fig. 1, Table 1 and in Al Bitar et al. (2012).

All SCAN data were downloaded from the International Soil Moisture Network (ISMN; Dorigo et al., 2011) which is a data hosting facility for soil moisture data from various networks around the globe. In addition to quality control by the contributing network the ISMN performs basic quality control and homogenization of all its soil moisture datasets (Dorigo et al., 2011).

3. SMOS soil moisture products

The SMOS Level 2 (L2) User Data Product (SML2UDP) provided by ESA from the second data reprocessing campaign and operational product were used in this study. Data prior to April 26, 2012 were from the reprocessed dataset and following were from the operational dataset so that the entire data-series used the same soil moisture processor (ver. 5.51). Soil moisture, optical thickness and other geophysical variables are retrieved based on the inversion of the LMEB radiative transfer model (Kerr et al., 2012; Wigneron et al., 2007) by minimizing a cost function from multi-angular brightness temperature (TB) observations from the microwave imaging radiometer with aperture synthesis (MIRAS) (Kerr et al., 2012). SMOS soil moisture are provided on the (Icosahedral Snyder Equal Area Earth) ISEA 4H9 fixed grid with nodes equally spaced at 14.989 km (Kerr et al., 2010), although each SMOS grid is actually the weighted mean of an area with a radius of approximately 42 km around the node center (Kerr et al., 2012). The SMOS ascending and descending half-orbits coincide with approximately 0600 h and 1800 h local solar time (Kerr et al., 2010) where differences in the redistribution of water, heterogeneity of surface temperature caused by daytime heating and sensor orientation during the overpass all contribute to differences in accuracy of the retrieval between ascending and descending overpasses (Collow et al., 2012; Jackson et al., 2012; Rowlandson, Hornbuckle, Bramer, Patton, & Logsdon, 2012). Since the goal of this study is to assess the errors associated with bias correction and many researchers have reported a greater accuracy and stability of SMOS ascending overpasses (Jackson et al., 2012; Rowlandson et al., 2012), the SMOS data are separated and only ascending half-orbits are considered. Several indicators of quality are provided with SML2UDP