Assimilation of Blended Satellite Soil Moisture Data Products to Further Improve Noah-MP Model Skills

Jifu Yin[1, 2], Xiwu Zhan[2], Michael Barlage[3], Sujay Kumar[4], Andrew Fox[4,5], Clement Albergel[6], Christopher R. Hain[7], Ralph R. Ferraro[1], Jicheng Liu[1, 2]

[1] ESSIC/CISESS, University of Maryland College Park, College Park 20740, MD, USA

[2] NOAA NESDIS Center for Satellite Applications and Research, College Park 20740, MD

[3] NOAA NCEP Environmental Modeling Center, College Park 20740, MD, USA

[4] NASA Goddard Space Flight Center, Greenbelt 20770, MD, USA

[5] GESTAR-II, Morgan State University, Baltimore 21251, MD, USA

[6] European Space Agency Climate Office, ECSAT, Harwell Campus, Oxfordshire, Didcot OX11 0FD, UK

[7] NASA Marshall Space Flight Center, Earth Science Branch, Huntsville, AL 35801, USA

*Corresponding author: Dr. Jifu Yin, Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, 5825 University Research Court suite 4001, College Park, MD 20740, USA. Email: jyin@umd.edu

Abstract:

Microwave satellite remote sensing has enabled observations of soil moisture (SM) at the global scale, and multiple SM data products have been developed in the past decades. However, single-sensor-based measurements are insufficient for continuous spatiotemporal coverage. In the context of its climate program, the Climate Change Initiative, the European Space Agency (ESA) has developed robust, long term, global scale, multi instrument satellite derived time series of climate data record for key component of the climate system, including soil moisture (CCI), while the Soil Moisture Operational Product System (SMOPS) was specifically developed by National Oceanic and Atmospheric Administration (NOAA) to offer the real time blended SM datasets through merging all available individual products. Before combining, all individual SM data ingested into both SMOPS and CCI blended products are scaled to Global Land Data Assimilation System (GLDAS) 0-10 cm SM climatology. Benefiting from land surface model evolution and the availability of high-quality forcing data, GLDAS has become more comprehensive to track SM changes and dynamic trends. The development of GLDAS and the scaling procedure in CCI and SMOPS leave an open scientific and operational question: do the blended satellite SM data products have added value comparing to the GLDAS product? This study clearly reveals that both CCI and SMOPS can provide the reliable SM observations with independent information, although their climatology matches well with GLDAS. Relative to assimilation of GLDAS 0-10 cm SM data, Noah-MP model can be further improved by assimilating the blended satellite SM observations with respect to the quality-controlled in situ measurements. The strong consistency of results presented in this paper proves that the blended satellite SM data products are more useful than the GLDAS product in terms of improving Noah-MP model performance.

1. Introduction

Through impacts on the partitioning of incoming radiation into latent and sensible heat-fluxes, soil moisture (SM) plays a crucial role in the terrestrial water cycle, energy balance and carbon exchange (Entekhabi et al., 2010a; Dorigo et al., 2017). The regional precipitation is commonly increased under surplus SM conditions, while a lack of SM triggers drought occurrences (Koster et al., 2004; Brocca et al., 2017; Dorigo et al., 2017; Yin et al., 2018). Regional gradients in SM also affect plant growth through constraining biogeochemical cycles and plant transpiration and photosynthesis (Seneviratne et al., 2010). Accurate knowledge of the SM states is thus necessary for numerical weather forecasts, as well as climate monitoring and predictions.

Ground stations may well track SM changes and variability, but they suffer from insufficient spatial coverage and inconsistencies at continental and global scales. Benefiting from decades of evolution, land surface models (LSMs) have become more comprehensive to provide spatiotemporally consistent SM estimates (Chen and Dudhia, 2001; Ek et al., 2003; Dai et al., 2003; Niu et al., 2011; Yang et al., 2011). These LSMs need vegetation parameters to represent the partitioning of available radiation into latent and sensible heat-fluxes, and soil heat exchanges (Yin et al., 2016). Similarly, the soil hydraulic properties are widely used in the current LSMs to represent moisture transport within the soil (Chen and Dudhia, 2002; Niu et al., 2011). However, the soil hydraulic properties and vegetation parameters are traditionally set as constants and assigned from lookup tables that are determined by limited field experiments (Chen and Dudhia, 2002; Ek et al., 2003; Dai et al., 2003; Niu et al., 2011). These suboptimal parameter schemes increase errors of SM estimations (Shellito et al., 2016), while the LSM simulations are subject to errors from meteorological forcing data and lack of scientific understanding in model physics (Reichle and Koster, 2004; Peters-Lidard et al., 2008; Crow et al., 2012).

Satellite land surface data products offer a great opportunity to improve LSM performance through providing global vegetation observations, soil and elevation parameters, and meteorological forcing data. In order to produce optimal fields of land surface states and fluxes, the Global Land Data Assimilation System (GLDAS) on the basis of LSMs was developed by National Aeronautics and Space Administration (NASA) through combining satellite- and ground-based observations (Rodell et al., 2004). Benefiting from the observational precipitation and downward radiation products, the GLDAS LSMs, typically including Noah, the Community Land Model (CLM), Mosaic and the Variable Infiltration Capacity model (VIC), incorporate the optimal available analyses from atmospheric data assimilation systems, and several satellite hydrological observations.

Microwave satellite observations have opened a new era for spatially distributed measurements of SM at the global scale since the 1970s. The soil dielectric constant linking SM and soil emissivity has a great impact on the microwave emission (Wang et al., 1987; Jackson and Schmugge, 1989), which allows to retrieve SM through either passive or active microwave satellite observations in a direct manner. The passive technique uses a radiometer to receive the land surface emission affected by the physical and emissivity temperature of the Earth, while an active radar senses the land surface backscatter through transmitting electromagnetic pulses. Both passive and active microwave sensors can provide SM retrievals under nearly any weather conditions. Based on X-band (8.0-12.0 GHZ) and C-band (4.0-8.0 GHZ) microwave measurements, recent satellite SM data products have included the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E; Njoku et al., 2003), AMSR-2 onboard the Global Change Observation Mission-Water (GCOM-W) satellite (Maeda et al, 2016), and WindSat (Li et al., 2010), as well as the Advanced Scatterometer (ASCAT) from Meteorological Operational

platform (MetOp)-A, MetOp-B and MetOp-C satellite series (Wagner et al., 2013). There are two more recent sensors that have been specifically designed to retrieve SM on basis of L-band (1.0-2.0 GHZ) observations, including the Soil Moisture and Ocean Salinity (SMOS; Kerr et al., 2010) and Soil Moisture Active Passive (SMAP; Entekhabi et al., 2010a).

These microwave satellite SM data products have footprints at typical resolutions from 25 to 50 km, while their daily spatial coverages depend on the sensor characteristics including swath width and revisit time. Benefiting from the development of antenna reflector, the current passive radiometers could achieve finer spatial resolution and higher coverage on the land surface brightness temperature with swath spanning from ~900 km to ~1450 km (Yin et al., 2020). Due to much more energy requirements, however, the active radar generally owns a relative narrower swath, such as ~550 km for ASCAT-A, -B and -C (Wagner et al., 2013). Higher frequencies (shorter wavelengths) microwave observations only represent SM status for shallow soil depth (Njoku et al., 2003; Wagner et al., 2013). This weakness could be compensated for by L-band frequency as used by SMOS and SMAP, which can penetrate up to the top 5 cm soil depth (Kerr et al., 2010; Entekhabi et al., 2010a). Besides lacking complete coverages in space and time and varying significantly from each other, the individual satellite SM data products have their own characteristics and are archived in different formats.

Therefore, two blended satellite SM data products were developed to overcome the aforementioned single-senor shortcomings. One is from European Space Agency (ESA)─Climate Change Initiative (CCI) combining various individual passive and active microwave satellite data products to provide the climate SM data records from 1979 in support of climate research (Dorigo et al., 2017; Gruber et al., 2019). Another one is the Soil Moisture Product system (SMOPS) developed by the National Oceanic and Atmospheric Administration (NOAA)-National

114 Environmental Satellite, Data, and Information Service (NESDS). To meet the data requirements
115 of National Weather Service (NWS), SMOPS uniquely offers a real time blended satellite data
116 product with combining all currently available individual SM observations (Liu et al., 2016; Yin
117 et al., 2015a, 2019, 2020). Before producing soil moisture CCI combined (hereafter: CCI) and
118 SMOPS blended data products, all available SM observations from the single sensors are scaled
119 to GLDAS-Noah 0-10 cm SM climatology using cumulative distribution function (CDF)-matching
120 method (Dorigo et al., 2017; Yin et al., 2015a, 2019, 2020). The scaling procedure leaves several
121 open scientific and operational questions, including:
122 1) Are the CCI and SMOPS products able to match well with GLDAS climatology?
123 2) Are there significant differences between the daily blended satellite SM data products and
124 the daily GLDAS-Noah estimations for the 0-10 cm soil layer in such a manner that CCI and
125 SMOPS can be treated as independent data sources?
126 3) Do the blended satellite SM data products have added value comparing to the GLDAS
127 product?
128 In this paper, we attempt to bridge these knowledge gaps through intercomparing the Noah-
129 Multi-parameterization (Noah-MP) model performance with benefits of assimilating GLDAS-
130 Noah 0-10 cm SM estimations and the CCI and SMOPS blended microwave satellite SM data
131 products. In the following section, the data sets used in this paper will be briefly described. Designs
132 of data assimilation (DA) strategy and evaluation methods will be provided in section 3. A
133 thorough understanding of multiyear averages is important to evaluate CCI and SMOPS
134 climatology with treating the GLDAS as a benchmark. Differences in daily changes and dynamic
135 trends are crucial to identify the quality distinctions between GLDAS estimations and the blended
136 satellite SM data products. Inter-comparisons between GLDAS and either CCI or SMOPS, as well

as Noah-MP model skill improvements from assimilating the blended satellite observations are then highlighted in Section 4. Attributes of the inter-comparisons and data assimilation results will be discussed in Section 5. Finally, all results of this study will be briefly summarized in Section 6.

2. Data Sources

2.1. GLDAS Soil Moisture

In order to produce reliable land surface fields in near real time, the GLDAS was developed jointly by NASA Goddard Space Flight Center (GSFC) and NOAA National Centers for Environmental Prediction (NCEP). It provides global hydrological estimates through combining LSM simulations with the new generation of ground- and space-based satellite observations (Rodell et al., 2004; Fang et al., 2009; Li et al., 2015). GLDAS LSMs including CLM, VIC, Noah and Mosaic are driven by observations-based precipitation and downward radiation products to reduce the uncertainties from the forcing data. Benefiting from incorporating satellite land surface data products, such as leaf area index, SM, surface temperature, snow cover and snow water equivalent, the current operational GLDAS produces optimal fields of land surface states and fluxes with uniquely unionizing the qualities of global, high resolution, near real time and offline terrestrial modeling system (Rodell et al., 2004). The GLDAS provides 3-hourly and monthly land surface state, such as SM, evaporation and soil temperature (Fang et al., 2009; Li et al., 2015).

Specifically, GLDAS V2.1-Noah 3-hourly data product at 0.25° spatial resolution is the main production stream. On basis of Noah Version 3.6 (V3.6), it provides 36 land surface fields from January 2000 to present with about 1.5-month latency. The Noah model was forced by atmospheric analysis data from NOAA Global Data Assimilation System (GDAS), daily analysis precipitation fields from the disaggregated Global Precipitation Climatology Project, and the Air Force Weather Agency's AGRicultural METeorological modeling system radiation datasets (Derber et al., 1991;

Huffman et al., 2001; Adler et al., 2003; Rodell et al., 2004). The 3-hourly GLDAS V2.1-Noah model outputs contain 0.25° SM estimations for 0-10 cm soil layer from 1 April 2015 to 31 December 2018, reprocessed to daily time-steps for this study.

2.2. ESA CCI Combined Soil Moisture Data Product

Soil moisture climate data records are fundamental for comprehensively understanding long-term changes of water and energy balances. Considering the individual satellite missions are inhibited by short mission lives and limited historical data, having a number of independent satellite SM data products does not guarantee it is straight-forward to create long-term consistent time series for climate change studies. The United Nations Framework Convention on Climate Change (UNFCCC) has thus delegated to the Global Climate Observing System (GCOS) to define what are the essential climate variables (ECVs). As a direct response to the UNFCCC needs in observations, ESA has set up the Climate Change Initiative (CCI) program that currently has 29 projects developing 27 ECVs including soil moisture (https://climate.esa.int/en/).

Combining individual satellite observations into a blended SM dataset can start from either level-1 brightness temperature and backscatter coefficients or level-2 SM retrievals. The long-term Soil Moisture CCI data product was developed by merging current and past satellite SM observation (Preimesberger et al., 2020). Three fusion products in the ESA CCI include ACTIVE, PASSIVE and COMBINED soil moisture observations (Dorigo et al., 2017). Based on a nearest neighbor search method, SM retrievals from all available single sensors are first mapped to a daily time step, and then passive and active SM data products are scaled to AMSR and ASCAT observations using the CDF-matching method, respectively (Liu et al., 2011; Dorigo et al., 2017). In ESA_CCI V04.5, the PASSIVE product combines the individual satellite SM observations from Scanning Multichannel Microwave Radiometer (SSMR), Special Sensor Microwave Imager

183 (SSM/I), Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), WindSat
184 Radiometer, AMSR-E, AMSR-2 and SMOS, while the ACTIVE data include SM observations
185 from Active Microwave Instrument (AMI), ASCAT-A and ASCAT-B. In order to generate the
186 COMBINED V04.5 datasets, both of the fusion ACTIVE and PASSIVE data products were
187 rescaled to GLDAS V2.1-Noah SM climatology for 0-10 cm soil layer to avoid the climatological
188 differences between active and passive SM retrievals (Preimesberger et al., 2020). Weights for
189 PASSIVE and ACTIVE data were derived from signal-to-noise estimates on basis of triple
190 collocation (Dorigo et al., 2017; Preimesberger et al., 2020). Specifically, individual SM retrievals
191 including ASCAT-A, ASCAT-B, AMSR-2 and SMOS were remapped to 0.25° spatial resolution
192 and then ingested into the ESA CCI COMBINED V04.5 over the 1 April 2015-31 December 2018
193 time period (Table 1) with quality control using the snow and frozen ground flags.

194 **2.3. SMOPS Blended Soil Moisture Data Product**

195 NOAA requires high-quality satellite SM observations with short latency to improve the
196 accuracy of NCEP Numeric Weather Prediction (NWP) models and National Water Model
197 (NWM) at the National Water Center (NWC). Satellite SM data products from individual satellite
198 sensors not only vary significantly from spatial resolution, data quality and archiving file format,
199 but also suffer from long-latency period and insufficient spatiotemporal coverage. To meet the
200 operational users' requirements, the SMOPS was thus developed at NOAA-NESDIS to produce a
201 blended product of SM retrievals from all available microwave satellite sensors (Liu et al., 2016;
202 Zhan et al., 2016; Yin et al., 2014, 2015a, 2019, 2020). After the first version officially released in
203 2010, the SMOPS data product has been updated twice with considerations to the users' feedbacks
204 through improving retrieval algorithms, removing old satellite platforms, and including new
205 sensors (Yin et al., 2019; 2020).

206 Specifically, SMOPS V3.0 (Table 1) provides microwave remotely sensed SM data products
207 from ASCAT-A, ASCAT-B, AMSR-2, SMOS and SMAP (Yin et al., 2019; 2020). The NWP
208 models at NOAA-NCEP need real time satellite SM retrievals arriving within the 6-hour cut-off
209 time. In order to seamlessly and conveniently provide data to the NCEP and NWC, SMOPS
210 retrieves SM data from the Level-1 brightness temperature observations to make directly available
211 with the shortest possible turn-around time (Zhan et al., 2016). The Single Channel Retrieval
212 (SCR) algorithm (Jackson, 1993) is used in SMOPS to produce AMSR-2, SMOS and SMAP near
213 real time soil moisture retrievals. Brightness temperature observations from either C- or L-band of
214 those sensors are converted to emissivity in the SCR algorithm, and in turn to be corrected by
215 vegetation information and surface roughness (Jackson, 1993). Given the dielectric constant
216 determined by the Fresnel equation, a dielectric mixing model is then used to retrieve the SM (Liu
217 et al., 2016). As ASCAT-A and ASCAT-B data products are produced and distributed in "real-
218 time" by the Vienna University of Technology via EUMETSAT (European Organisation for the
219 Exploitation of Meteorological Satellites) operationally (Wagner et al., 2013), the ASCAT Level-
220 2 relative soil moisture data are converted to the volumetric SM using a soil type map (Liu et al.,
221 2016; Zhan et al., 2016). SMOPS combines all the individual SM retrievals to a blended data layer
222 with better spatial coverage for each 6-hour time period or each day, which makes the SMOPS
223 unique (Yin et al., 2019, 2020). Before the combination, all SM observations from individual
224 sensors are remapped to SMOPS 0.25º latitude-longitude grids, and then scaled to GLDAS V2.1-
225 Noah SM climatology of the 0-10 cm soil layer using the CDF-matching method (Yin et al., 2019,
226 2020). The 6-hourly and daily SMOPS blended products are produced by merging all individual
227 SM retrievals acquired within the previous 6-hour and 24-hour windows (Yin et al., 2019). In this

228 paper, the daily 0.25° SMOPS V3.0 SM data from 1 April 2015 to 31 December 2018 are quality
229 controlled using the snow and frozen ground flags.

230 ----------

231 ***Please Insert Table 1 here.***

232 ----------

233 **2.4. SCAN Soil Moisture Observations**

234 Based on the experience of running a national SM and soil temperature pilot project, the
235 Natural Resources Conservation Service (NRCS) of the U.S. Department of Agriculture (USDA)
236 designed the Soil Climate Analysis Network (SCAN) to enable in situ observations by overcoming
237 problems including incomplete, short-term, limited areas of coverage and nonstandard sensor
238 arrays (Schaefer et al., 2007). Starting in 1999, the standard SCAN station configuration focuses
239 on the agricultural areas of the U.S in support of natural resources assessment. It consists of over
240 200 stations with hourly SM observations automatically recorded by measuring the soil dielectric
241 constant for the sensors. Hourly SCAN soil moisture and temperature data are available at National
242 Water and Climate Center home page (https://www.wcc.nrcs.usda.gov/scan/). In this paper, SCAN
243 observations from 1 April 2015 to 31 December 2018 were reprocessed to match the daily time-
244 step of Noah-MP model outputs. SM observations either outside of the physically possible range
245 or under frozen conditions on basis of SCAN soil temperature data were excluded as part of the
246 quality control strategy (Liu et al., 2011; Yin et al., 2015c). After excluding SCAN sites that
247 provide fewer than 100-day of SM observations, there are a total of 155 and 146 sites were chosen
248 in the CONUS to validate Noah-MP model 0-10 cm and 40-100 cm SM estimations.

3.0. Methods

3.1 Noah-MP Model

The Noah LSM is an important component of the NWP models for operational weather and climate predations and has been widely used by the NOAA-NCEP (Chen and Dudhia, 2002; Ek et al., 2003). However, limitations of Noah include a combined land surface layer of soil and vegetation, a bulk layer of snow and soil, using top 5 cm soil texture to represent the entire 200 cm soil column and a weak impeding effect of frozen soil on infiltration and river discharge (Niu et al., 2011; Yang et al., 2011). Noah-MP was thus developed to address the aforementioned problems through enhancing vegetation phenology, frozen soil and infiltration, the vegetation canopy energy balance, the layered snowpack, and SM-groundwater interaction and related runoff production (Niu et al., 2011; Yang et al., 2011). Compared to Noah land surface model, the new features in Noah-MP primarily include modification of two stream radiation transfer scheme to consider the 3-dimension canopy, implementation of semi-tile vegetation and bare soil, and application of interactive energy balance to estimate the canopy skin temperature (Niu et al., 2010; Yang et al., 2011). In this paper, the Noah-MP model V3.6 implemented in the NASA Land Surface System (LIS) V7.2 (Kumar et al. 2006, 2008; Peters-Lidard et al., 2011) was employed to intercompare assimilations of GLDAS 0-10 cm and the blended satellite SM estimations.

3.2. Ensemble Kalman Filter

The Ensemble Kalman Filter (EnKF) has been widely used in sequential SM data assimilation (Evensen, 1994; Kumar et al., 2009; Yin et al., 2014, 2015b). In an EnKF-based DA system, the ensemble forecast and state variable update steps are alternated by a Monte Carlo approximation of a sequential Bayesian filtering process (Evensen, 1994). The ensemble is typically created by adding Gaussian noises to forcing and/or state variables with assuming that the ensemble

perturbation cannot directly affect the mean model performance (Ryu et al., 2009). The model states ($\boldsymbol{y}$) for each ensemble member propagated forward in the forecast step is expressed as

$$\boldsymbol{y}^{t+} = y^{t-} + \mathbf{K}(\boldsymbol{x}^t - \mathbf{H}y^{t-}) \quad (1)$$

where **K**, **H**, and $\boldsymbol{x}$ vectors are the Kalman gain matrix, the observation operator and the SM observations, respectively. The observation operator primarily depends on the observation vectors, while the Kalman gain matrix is given by

$$\boldsymbol{K} = \frac{\zeta_y^t \mathrm{H}^{t^{\mathrm{T}}}}{\mathrm{H}^t \zeta_y^t \mathrm{H}^{t^{\mathrm{T}}} + \zeta_x^t} \quad (2)$$

where ζ_x^t indicates error variance for observations, while ζ_y^t is forecast error variance determined by ensemble spread. Specifically, the error variances for GLDAS 0-10 cm SM simulations, as well as SMOPS and ESA CCI blended SM data products were set as 3% as LIS examples (Kumar et al., 2009). The ensemble size was 12, which is the optimal ensemble size in a sequential SM assimilation system (Yin et al., 2015b), to update Noah-MP model states for all DA cases and the open loop run. Perturbation for meteorological forcing parameters and state variables (Table 2) were applied each individual 1-hr Noah-MP model time step.

3.3. Data Assimilation Strategies

The EnKF is an optimal DA technique if 1) satellite observation errors follow Gaussian distribution, 2) satellite observations are linearly related to the model simulations; and 3) model state errors are also jointly Gaussian-distributed. Given almost always biased to model simulations, satellite SM observations are generally bias-corrected to model simulations in order to satisfy the assumptions (Kumar et al., 2012; Nearing et al., 2013; Yin and Zhan, 2018). Thus, the Noah-MP V3.6 was initially run with the unperturbed meteorological forcing data and model state variables

to generate a 0-10 cm SM climatology for the bias-correction. In this stage, the Noah-MP V3.6 LSM was spun up by cycling 30 times through the period from 1 April 2015 to 31 December 2018, and then the simulations were conducted over the same period with 1-hour time step inputs and daily outputs. GLDAS and the both blended satellite SM estimations were then scaled to the Noah-MP model 0-10 cm SM climatology using the CDF-matching method that has a better spatiotemporal stratification (Kumar et al., 2012; Yin and Zhan, 2018).

Based on the bias-corrected GLDAS Noah V2.1, CCI and SMOPS SM datasets, the experiment structure was designed to include an open loop run and three DA cases. Specifically, 1) the open loop run (OLP) represents the Noah-MP model runs with the perturbed meteorological forcing data and model state variables (Table 2), which indicates model simulations under suboptimal forcing conditions. The OLP run is not benefited by DA. 2) The DA cases were set as assimilation of the bias-corrected GLDAS, CCI and SMOPS soil moisture estimations into the OLP run. This means they are forced by the same perturbed meteorological forcing data and model states used in the OLP run (Table 2), and in turn to highlight that the differences between DA cases and the OLP run are good metrics to evaluate the DA impacts. In this paper, DAGLD assimilates the GLDAS SM simulations for 0-10 cm soil layer, while DACCI and DASMP assimilate the ESA CCI and SMOPS blended SM data products, respectively. Given the same forcing data, model parameters, bias-correction strategy and perturbation strategies in the same EnKF-based DA system, differences among DAGLD, DACCI and DASMP are caused only by assimilating the different SM estimations.

The Noah-MP LSM with setting ensemble size as 12 was spun up by cycling 20 times through the period from 1 April 2015 to 31 December 2018. The OLP run and the three DA cases were then conducted at 0.25° spatial resolution on a gridded near-global domain (from −60°S, −180°W

316 to 90°N, 180°E) over the same period with 1-hour time step inputs and daily outputs. The GLDAS
317 and the both blended satellite SM data were assimilated into the Noah-MP LSM at each 00:00Z
318 through providing the updated 0-10 cm SM condition. All Noah-MP model simulations in this
319 study were forced by precipitation, downward longwave and shortwave radiation, near surface
320 humidity and wind, as well as surface pressure from the Global Data Assimilation System (Derber
321 et al., 1991).

322 ----------

323 ***Please Insert Table 2 here.***

324 ----------

325 **3.4. Model Performance Measurements**

326 The choice of metrics for measuring model performance mainly depends on the SM nature
327 and characteristics. A single metric is sensitive to a specific characteristic, and thus not able to
328 well obtain the SM attributes (Entekhabi et al. 2010b). In this paper, a comprehensive assessment
329 on Noah-MP model performance was thus conducted using tow widely used metrics including
330 unbiased RMSE (ubRMSE) and correlation coefficient (*r*). Based on the quality-controlled SCAN
331 observations (Θ_{SCAN}), the Noah-MP model SM simulations (Θ_{Model}) for the grid (*j*, *i*) are assessed
332 by

333 $$r(j,i) = \frac{\sum_{k=1}^{N}(\Theta_{SCAN}^{k} - \overline{\Theta_{SCAN}})(\Theta_{Model}^{k} - \overline{\Theta_{Model}})}{\sqrt{\sum_{k=1}^{N}(\Theta_{SCAN}^{k} - \overline{\Theta_{SCAN}})^2}\sqrt{\sum_{k=1}^{N}(\Theta_{Model}^{k} - \overline{\Theta_{Model}})^2}} \quad (3)$$

334 $$ubRMSE(j,i) = \sqrt{\sum_{i=1}^{N}((\Theta_{SCAN}^{i} - \overline{\Theta_{SCAN}}) - (\Theta_{Model}^{i} - \overline{\Theta_{Model}}))^2/(N-1)} \quad (4)$$

335 In this paper, the sample size *N* is 1371 for each grid, since there are 1371 days during the period
336 from 1 April 2015 to 31 December 2018. The overbar denotes the averages during the study period.

Similarly, the differences between the blended satellite and GLDAS (Θ_{GLDAS}) SM estimations are measured by root mean square error difference (RMSD)

$$RMSD = \sqrt{\sum_{k=1}^{N}(\Theta_{Satellite}^{k} - \Theta_{GLDAS}^{k})^2/(N-1)} \quad (6)$$

where $\Theta_{Satellite}$ indicates either SMOPS or ESA CCI blended SM data products that are available during the study period.

4.0 Results

4.1 Climatology of GLDAS and Blended Satellite Soil Moisture Data Products

All individual satellite SM observations are scaled to GLDAS 0-10 cm climatology before they are ingested into SMOPS and CCI blended products. The SMOPS and CCI climatologies are thus first evaluated by comparing with multi-year averaged GLDAS simulations. Figure 1 shows the temporal averaged CCI and SMOPS data products versus the GLDAS 0-10 cm SM simulations over the global domain from 1 April 2015 to 31 December 2018. The higher sample density area in red color being closer to the black line represents the blended satellite SM observations and GLDAS simulations match better, whereas the lower sample density area shading in the blue color is departure from the ideal regression curve. Compared to the GLDAS, the multiyear averaged CCI shows wetter patterns on the entire global domain (Figure 1a). Relatively, the SMOPS climatology presents drier patterns in dry areas, while exhibiting wetter patterns in the wet areas than the GLDAS (Figure 1b). With respect to the GLDAS climatology, the correlation coefficients (r) for CCI and SMOPS are 0.871 and 0.921, respectively. The good correlations imply that the climatology of both blended satellite SM data products match well with the GLDAS climatology. It also suggests that all individual satellite SM data products ingested into both CCI and SMOPS

have been successfully scaled to the GLDAS climatology, which allows to reasonably combine them together to produce the blended SM datasets.

Please Insert Figure 1 here.

4.2 Differences between GLDAS and Blended Satellite Soil Moisture Data Products

A pair of widely used metrics including correlation coefficient and RMSD are used to evaluate the differences between GLDAS and the blended satellite SM data products. Figure 2 shows temporal correlation coefficients between the daily GLDAS 0-10 SM simulations and the daily CCI and SMOPS observations from 1 April 2015 to 31 December 2018. Areas shading in red (blue) color in Figures 2a-2b indicates higher (lower) correlation coefficients. The ESA CCI shows strong consistency with the GLDAS 0-10 cm SM simulations primarily in the eastern South America, India, Australia, Sub-Sahara, and the southern Africa, while low correlations are found in the middle-high latitude areas of the north hemisphere (Figure 2a). Relatively, SMOPS presents weaker agreement with the GLDAS simulations in not only the high latitude but also middle-low latitude areas (Figure 2b). The larger correlation values for SMOPS are mainly found in the Sub-Sahara, India and the eastern South America. With respect to the GLDAS, the zonal-averaged correlation coefficients (ZACs) for CCI and SMOPS exhibit lower values (less than 0.2) from 40°N to 80°N. Both CCI and SMOPS present greater ZACs in the -40─40°N areas, where the ZACs for CCI and SMOPS are spanning 0.4-0.75 and 0.3-0.55, respectively. Compared to the SMOPS blended SM product, the ZACs ESA CCI is higher by about 0.2 in the belt areas located in -40─15°N and 25─35°N. Relative to the patterns for climatology in Figure 1, the GLDAS-based correlation coefficients for multiyear averaged CCI and SMOPS are remarkably decreased by their

daily dynamics. It suggests that the dynamic trends for the both blended satellite SM data products are hardly affected by the scaling procedure before combination.

Please Insert Figure 2 here.

Besides the correlation coefficient, the RMSD is used as another metric to measure the differences between GLDAS simulations and the blended satellite SM observations. Figure 3 shows temporal RMSDs between the daily GLDAS simulations and the daily blended satellite SM data products from 1 April 2015 to 31 December 2018. Area shading in red color in Figures 3a-3b indicates the GLDAS-based RMSD is greater, while in blue color means smaller differences. With respect to the daily GLDAS SM simulations for 0-10 cm soil layer, the CCI primarily exhibits larger RMSD values (>0.10 m^3/m^3) in the high latitude areas (Figure 3a). Relatively, the SMOPS blended SM data product presents greater GLDAS-based RMSDs on the entire global domain (Figure 3b). Besides of the high latitude areas, the greater RMSD values can also be found in the west CONUS, South America, the central Africa and the south Asia. With respect to the GLDAS 0-10 cm SM simulations, the zonal averaged RMSDs for CCI are spanning from 0.04 to 0.055 m^3/m^3 in the -50─40°N areas, while can reach to 0.09 m^3/m^3 in the high latitude areas (Figure 3c). The greater differences between the daily SMOPS and GLDAS can be found in the belt areas located in 45─65°N, which can reach to 0.1 m^3/m^3. The RMSD values for SMOPS basically range from 0.05 m^3/m^3 to 0.08 m^3/m^3 in the remaining areas. These results suggest that both CCI and SMOPS data products can be treated as independent data sources, although the individual SM observations ingested into them are scaled to GLDAS 0-10 cm SM climatology.

Please Insert Figure 3 here.

4.3 Validation on DA cases with SCAN Observations

Figures 4a-4c show the SCAN observations-based ubRMSEs for Noah-MP model 0-10 cm SM simulations from the three DA cases. SCAN sites in blue color highlight lower ubRMSE values, whereas in red color mean the DA cases perform modest. With respect to the quality-controlled SCAN measurements, Noah-MP model with benefits of data assimilation shows a reasonable performance in the west and east CONUS with the ubRMSE values below 0.05 m^3/m^3. However, large errors for the DA cases are primarily found in the Mississippi River areas and the south CONUS, which can reach to 0.1 m^3/m^3. Based on the same state variables and meteorological forcing data, inter-comparisons among the three DA cases are presented in Figures 4d-4f. Compared to DAGLD, DACCI shows better performance in the Great Plain areas, whereas the degradations caused by assimilating CCI can be found in the northwestern CONUS. With respect to the SCAN SM observations, Noah-MP model shows a better behavior with benefits of assimilating the SMOPS SM data product in the central-western CONUS in comparison with the DAGLD, while no significant degradations are found on the entire CONUS domain. Both DACCI and DASMP perform similar with the DASMP showing lower ubRMSE values in the northwest areas.

Figures 4g-4i present the differences in SCAN measurements-based ubRMSEs for the Noah-MP model SM simulations in the top soil layer between with and without DA benefits. Sites in blue color indicate improvement, whereas those in the warm color indicate degradation. Benefiting

from assimilation of GLDAS and the blended satellite SM estimations, significant improvements on Noah-MP model performance are found in the western and the central-eastern CONUS. Yet the degradations for the three DA cases can be seen in the Great Plain areas. Specifically, the CONUS domain-averaged ubRMSE for OLP is 0.0533 m^3/m^3, which can be significantly decreased by 0.0011 m^3/m^3 (2.11% reduction), 0.0012 m^3/m^3 (2.30% reduction) and 0.0018 m^3/m^3 (3.50 % reduction) by DAGLD, DACCI and DASMP, respectively (Table 3). Compared to the DAGLD, Noah-MP performance could be further improved by assimilating CCI and SMOPS blended SM data products.

Please Insert Figure 4 here.

Figures 5a-5c show the SCAN observations-based ubRMSEs for Noah-MP model 40-100 cm SM simulations from the three DA cases. The behavior of Noah-MP model SM simulations for the topsoil layer are completely mirrored in the deeper soil layer. The three DA cases present good performance in west and east CONUS areas with the ubRMSE below 0.05 m^3/m^3, whereas exhibit large uncertainties (>0.1 m^3/m^3) in the Mississippi River areas and the south CONUS. It can be found that ubRMSEs for DAGLD are tremendously reduced by DACCI in the central and west CONUS areas (Figure 5d). However, the DACCI show degradations in the south CONUS in comparison with the DAGLD, which can be significantly improved by the DASMP. Compared to DAGLD, DASMP presents significant improvement on Noah-MP model SM simulations for 40-100 cm soil layer in the central-western areas, while no significant degradations are found on the entire CONUS domain (Figure 5e). Relative to DACCI, DASMP exhibits lower ubRMSEs on the CONUS domain except the Upper Mississippi River and the south CONUS areas (Figure 5f).

The positive signals from the surface soil layer are reasonably vertically propagated to the deeper soil layer (Figures 5g-5i). With respect to the quality-controlled SCAN measurements, ubRMSEs for OLP run are significantly decreased with benefits of assimilating GLDAS SM estimations in the south and east CONUS. Both DACCI and DASMP exhibit lower ubRMSEs on the entire CONUS domain except few sites scattering in north and south areas in comparison with OLP run. Statistical results show that the CONUS domain-averaged ubRMSE for OLP run is 0.0588 m^3/m^3, which can be significantly decreased by 0.0016 m^3/m^3 (2.80% reduction), 0.0025 m^3/m^3 (4.44% reduction) and 0.0038 m^3/m^3 (6.90 % reduction) by DAGLD, DACCI and DASMP, respectively (Table 3). Compared to the DAGLD, Noah-MP model SM simulations for 40-100 cm soil layer are further improved by 1.6% and 4% by assimilating CCI and SMOPS SM data products, respectively (Table 3).

Please Insert Figure 5 here.

The second metric is the correlation coefficient (r) that is commonly used to measure the dynamic trends between model simulations and ground observations. Figure 6 shows differences in correlation coefficients among DAGLD, DACCI and DASMP cases with blue (red) indicating robust positive (negative) agreements. The DACCI is more successful to respect the SCAN SM observations in the great plain and the east CONUS areas in comparison with the DAGLD, whereas shows weaker correlations for the SCAN sites scattering in the northwest and south areas. Relatively, DASMP exhibits less degradations over the DAGLD, while presenting a relative stronger consistent with in situ observations in the central and west CONUS areas. It is very interesting to find that the DASMP presents a more robust agreement in west CONUS in

comparison with the DACCI. Compared to DASMP, however, DACCI has a better behavior in the Mississippi River areas. The CONUS domain-averaged correlation coefficient for the OLP run is 0.549, which is significantly increased by 0.015 (2.66% increase), 0.017 (3.00% increase) and 0.026 (4.52% increase) by DAGLD, DACCI and DASMP, respectively (Table 3). With respect to the in-situ observations, Noah-MP model with benefits of assimilating either CCI or SMOPS data product can more reasonably track 0-10 cm SM changes in comparison with assimilation of GLDAS estimations.

Please Insert Figure 6 here.

Compared to the top soil layer, larger differences in correlation coefficients among DAGLD, DACCI and DASMP cases are found for the SM simulations in 40-100 cm soil layer (Figure 7). Statistical results reveal that the DACCI is more successfully to track SM dynamic trends over the DAGLD in the Mississippi River and the west CONUS areas. Compared to DAGLD, DASMP is more consistent with the quality-controlled SCAN SM measurements on the entire CONUS domain except few degraded sites over the central and west areas. Relative to the DACCI, the DASMP case is more robust agreement with SCAN observations in the northwest and southeast areas, yet shows weaker correlations in the southwest and Upper Mississippi River areas. The CONUS domain-averaged correlation coefficient for the OLP run is 0.198, which is significantly increased by 0.039 (16.46% increase), 0.049 (19.84% increase) and 0.071 (26.39% increase) by DAGLD, DACCI and DASMP, respectively (Table 3). With respect to the SCAN observations, Noah-MP model with benefits of assimilating the blended satellite SM data products can more

successfully track 40-100 cm SM dynamic trends in comparison with assimilation of GLDAS estimations.

Please Insert Figure 7 here.

Please Insert Table 3 here.

5.0. Discussion

Results in Section 4.0 indicate that CCI and SMOPS blended satellite data products can well match with GLDAS 0-10 cm SM climatology, but there are significant differences in daily dynamic trends and changes between GLDAS and satellite observations. As a result, Noah-MP model can be further improved by assimilating either CCI or SMOPS in comparison with assimilation of GLDAS SM. Further considerations relevant to the investigations are discussed here associated with data characteristics, advantages of satellite observations and complementary evaluations.

5.1. Spatial Distributions

Before ingesting into the CCI and SMOPS, all individual retrievals are scaled to the GLDAS 0-10 cm SM climatology. As a result, the multiyear-averaged CCI and SMOPS SM data products match well with the GLDAS climatology (Figure 1). However, the large differences between the daily GLDAS simulations and the daily blended satellite observations can still be found (Figures 2-3), which allow to treat both of the CCI and SMOPS as the independent data sources. Interannual

variations over the 2015-2018 period indicate that GLDAS 0-10 cm SM simulations are wetter in the U.S., Canada and Europe (Figure 8) in cold season (January-April) than warm season (June-September), which are the opposite to the seasonal variations of precipitation (Adler et al., 2017). Compared to GLDAS, CCI shows a reasonable spatial distribution with narrowing the differences in season-averaged SM values between the cold and warm seasons, whereas the opposite interannual variations can still be found in the CONUS (Figure 8). Relatively, interannual variations of SMOPS present a more reasonable spatial distribution in the North Hemisphere. It exhibits a dryer pattern in cold season while the wetter SMOPS observations can be found in warm season (Figure 8). These results suggest that the blended satellite observations provide SM information in an independent manner, although they even after being re-scaled with respect to the GLDAS climatology. As a result, the problems of interannual variations in the GLDAS could be addressed by the CCI and SMOPS.

To better meet the users' requirements, the developers of soil moisture CCI and SMOPS have given a top priority to develop the "model-free" blended soil moisture data products (Madelon et al., 2022; Yin et al., 2022). Given the GLDAS model errors, the individual satellite SM retrievals could be scaled to the L-band microwave remote sensing soil moisture observations (Madelon et al., 2022; Yin et al., 2022). As a result, the "model-free" blended satellite SM observations are more successful tracking the surface soil moisture status in comparison with the GLDAS climatology-based data products (Madelon et al., 2022; Yin et al., 2022). Assimilation of the "model-free" blended satellite SM data products is expected to further improve the Noah-MP performance in the future.

Please Insert Figure 8 here.

5.2. Advantages of ESA CCI and SMOPS

Synchronous comparisons are generally used to evaluate model and satellite SM estimations with respect to the in-situ observations. Inter-comparisons in this way are more direct but subjected to the distinguished differences in sample size determined by data availability (Yin et al., 2019). GLDAS has a fully spatiotemporal coverage, which is much better than CCI and SMOPS data product. As a result, direct comparisons between GLDAS simulations and satellite observations probably tend to promote SMOPS and CCI in an unfair way. We thus designed an experiment to comprehensively intercompare Noah-MP model skills benefiting from assimilation of GLDAS and the two blended microwave SM data products. Based on the same meteorological forcing data and model structure, Noah-MP model simulations for the three DA cases under the same parameterization and perturbation conditions allow to highlight the relative advantages and disadvantages for those data products assimilated into the OLP run. Noah-MP running for DACCI and DASMP cases becomes an open loop when the blended satellite SM observations are unavailable. Given the GLDAS has a fully spatiotemporal coverage, Noah-MP is supposed to have a better performance with benefits of assimilating the GLDAS 0-10 cm SM estimations. In this paper, however, statics clearly document that both DACCI and DASMP are more successful to respect in situ observations in compassion with DAGLD. This indicates that CCI and SMOPS SM data products could provide more accurate information over the GLDAS, even though GLDAS has a higher data availability.

In an EnKF-based DA system, model simulations are corrected toward the satellite observations, and in turn to propagate model states forward in the forecast step (Yin et al., 2014, 2015b). There are still considerable uncertainties associated seasonal and interannual variations

(Figure 8) in GLDAS SM simulations, although several satellite hydrological observations are used to drive the GLDAS models (Rodell et al., 2004). This open scientific problem and challenge could be addressed by satellite SM observations. Benefiting the reasonable spatial patterns, Noah-MP model performance is further improved by DACCI and DASMP in this paper. Compared to the DAGLD, Noah-MP model SM simulations with benefits of the SMOPS and CCI blended SM observations show a better performance with respect to the in-situ observations (Figures 4-9).

5.3. Inter-comparisons of Assimilating CCI and SMOPS

Compared to CCI, SMOPS shows greater differences in GLDAS-based correlation coefficients (Figure 2) and RMSD values (Figure 3) over the global domain. According to Figure 8, however, the interannual variations for the SMOPS are more reasonable in comparison with the CCI. This means satellite observations could provide more reliable SM data in comparison with GLDAS simulations, although the latter one has a great performance. With respect to the GLDAS, more significant differences can probably guarantee that much less useful information from satellite observations is lost in the scaling procedure. Relative to CCI, SMOPS is more successful to respect the GLDAS climatology, but more independent to track SM changes and dynamic trends. As a result, the DASMP has a better behavior in comparison with the DACCI.

The spatiotemporal coverage for CCI has been significantly improved by the increasing number of available sensors in recent years (Dorigo et al., 2017), and its zonal average can thus reach to 70%-90%. The ESA CCI COMBINED V04.5 includes individual SM observations from ASCAT-A, ASCAT-B, SMOS and AMSR-2 from 1 April 2015 to 31 December 2018 (Table 1). Relatively, one more individual data product “SMAP” was ingested into the SMOPS blended SM over the same period (Table 1). As a result, many gaps remained by the CCI can be filled by the SMOPS (Wang et al., 2021), as the latter one has a 90-95% spatiotemporal coverage over the

global domain (Yin et al., 2019). With respect to the ground observations, the SMAP presents a good behavior with lower errors in comparison with the current individual SM data products (Burgin et al., 2017). Ingesting the SMAP can thus promote the SMOPS performance, and in turn to benefit the DASMP.

6.0. Conclusions

The ESA CCI and NOAA SMOPS are only two blended satellite SM products. Before the combination, all individual satellite observations ingested into CCI and SMOPS are scaled to GLDAS 0-10 cm SM climatology using the CDF-matching method. This study reveals that there is still a need to develop the blended remotely sensed SM data products, although GLDAS has a reasonable performance with benefits from LSM evolution, forcing data improvement, and parameterization promotion. Specifically, the Noah-MP model benefiting from assimilation of GLDAS simulations for the 0-10 cm soil layer was inter-compared with advances of assimilating CCI and SMOPS blended SM data products. With respect to the quality-controlled in situ observations, a comprehensive assessment using ubRMSE, correlation coefficient is conducted to measure the metrics of the three DA cases. The key results obtained in this paper include:

1) Both CCI and SMOPS can well match GLDAS climatology as the individual satellite observations are scaled to GLDAS before the combination. With respect to the GLDAS climatology, the correlation coefficients for the CCI and SMOPS are 0.871 and 0.921, respectively.

2) But both CCI and SMOPS can be treated as independent data sources, as the low correlation coefficients and large RMSD values between the daily GLDAS and the daily blended satellite SM estimations.

3) Noah-MP LSM performance can be significantly improved by assimilating GLDAS, CCI and SMOPS. The improvements can be found for not only the top soil layer but also deep soil layer with reducing unRMSE values and increasing the correlations with respect to the quality-controlled SCAN SM measurements.

4) Relative to assimilation of GLDAS 0-10 cm estimations, Noah-MP model simulations with benefits of assimilating CCI and SMOPS are more successful to track SM changes and dynamic trends with respect to in situ measurements.

Acknowledgment

This work was jointly supported by NOAA Climate Program Office (CPO)-Modeling, Analysis, Predictions and Projections (MAPP) Program, NOAA Joint Polar Satellite System (JPSS)-Proving Ground and Risk Reduction (PGRR) Program, and NOAA grant (NA19NES4320002). We thank the anonymous reviewers for helping improve the manuscript quality. NASA GLDAS outputs can be found at https://disc.gsfc.nasa.gov/datasets?keywords=GLDAS. The ESA CCI COMBINED soil moisture data are archived at https://www.esa-soilmoisture-cci.org/, while the SMOPS blended data product can be obtained from http://www.ospo.noaa.gov/Products/land/smops/smops_loops.html. The authors declare that there is no conflict of interest regarding the publication of this paper. The manuscript contents are solely the opinions of the authors and do not constitute a statement of policy, decision, or position on behalf of NOAA or the U. S. Government.

7.0. References

1) Adler R F., G. Gu, M. Sapiano, J.-J. Wang, G. J. Huffman. Global Precipitation: Means, Variations and Trends During the Satellite Era (1979–2014). Surv. Geophys., 2017, 38: 679–699
2) Adler RF, Huffman GJ, Chang A, Ferraro R, Xie PP, Janowiak J, Budolf B, Schneider U (2003) The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979-Present). J Hydrometeorol 4:1147–1167
3) Balsamo G., Agusti-Panareda, A., Albergel, C. et al., Satellite and in situ observations for advancing global Earth surface modelling: A review, Remote Sens., 2018, 10, 2038, https://doi.org/10.3390/rs10122038
4) Brocca, L., Crow, W. T., Ciabatta, L., Massari, C., De Rosnay, P., Enenkel, M., et al. (2017). A review of the applications of ASCAT soil moisture products. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10, 2285–2306
5) Burgin, M., Colliander, A., Njoku, E., Chan, S., Kerr, Y., Binglish, R., Jackson, T.J., Entekhabi, D., Yueh, S. 2017. A comparative study of the SMAP passive soil moisture product with existing satellite-based soil moisture products. IEEE Transactions on Geoscience and Remote Sensing. 55(5):2959-2971
6) Chen F, and J Dudhia. Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. Monthly Weather Review, 2001, 129: 569-585
7) Crow W. T., S. V. Kumar, and J. D. Bolten. On the utility of land surface models for agricultural drought monitoring. Hydrol. Earth Syst. Sci., 2012, 16: 3451-3460

8) Dai Y, X Zeng, R. E. Dickinson, et al. The Common Land Model (CLM), Bull. Am. Meteorol. Soc., 2003, 84: 1013–1023

9) Derber, J. C., D. F. Parrish, and S. J. Lord, The new global opera-tional analysis system at the National Meteorological Center.Wea. Forecasting, 1991, 6:538–547

10) Derber, J. C., Parrish D. F. , and Lord S. J. , 1991: The new global operational analysis system at the National Meteorological Center. Wea. Forecasting, 6 , 538–547

11) Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., et al. (2017). ESA CCI soil moisture for improved Earth system understanding: State-of-the art and future directions. Remote Sensing of Environment, 203, 186–215

12) Ek M. B., K. E. Mitchell, Y. Lin, et al. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. J. Geophys. Res., 2003, 108(D22), 8851, doi:10.1029/2002JD003296

13) Entekhabi D., R.H. Reichle, R. D. Koster, and W.T. Crow. Performance Metrics for Soil Moisture Retrievals and Application Requirements. Journal of Hydrometeorology, 2010b, 11: 832-839

14) Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010a). The Soil Moisture Active Passive (SMAP) mission. Proceedings of the IEEE, 98(5), 704–716

15) Evensen G. Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res, 1994, 99(C5): 10143-10162

16) Fang H., H. K. Beaudoing, M. Rodell, W. L. Teng, B. E. Vollemer. Global Land Data Assimilation System (GLDAS) Products, Services and Application from NASA Hydrology

674 Data and Information Services Center (HDISC). ASPRS 2009 Annual Conference Baltimore,
675 Maryland, March 8-13, 2009
676 17) Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., Dorigo, W. (2019). Evolution of
677 the ESA CCI Soil Moisture Climate Data Records and their underlying merging methodology.
678 Earth System Science Data 11, 717-739, https://doi.org/10.5194/essd-11-717-2019
679 18) Huffman, George J., and Coauthors, 2001: Global Precipitation at One-Degree Daily
680 Resolution from Multisatellite Observations. J Hydrometeorol., 2, 36–50
681 19) Jackson, T. J., "Measuring surface soil moisture using passive microwave remote sensing,"
682 Hydrological Process, 7, pp.139-152, 1993
683 20) Jackson, T.J.; Schmugge, T.J. Passive microwave remote sensing system for soil moisture:
684 Some supporting research. IEEE Trans. Geosci. Remote Sens. 1989, 27, 225–235
685 21) Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., et al. (2010).
686 The SMOS mission: New tool for monitoring key elements of the globalwater cycle.
687 Proceedings of the IEEE, 98(5), 666–687
688 22) Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., et al. (2004). Regions
689 of strong coupling between soil moisture and precipitation. Science, 305(5687), 1138–1140
690 23) Kumar S. V., R. H. Reichle, R. D. Koster, et al. Role of subsurface physics in the assimilation
691 of surface soil moisture observations. J. Hydrometeor., 2009, 10: 1534–1547
692 24) Kumar, S. V., R. H. Reichle, K. W. Harrison, C. D. Peters-Lidard, S. Yatheendradas, and J.
693 A. Santanello (2012), A comparison of methods for a priori bias correction in soil moisture
694 data assimilation, Water Resour. Res., 48, W03515, doi:10.1029/2010WR010261

695 25) Kumar, S.V., Peters-Lidard, C.D., Tian, Y., et al, 2006. Land information system: an
696 interoperable framework for high resolution land surface modeling. Environ. Model. Softw.
697 21 (10), 1402e1415. http://dx.doi.org/10.1016/j.envsoft.2005.07.004
698 26) Li L., G. B. Senay, J. P. Verdin. Evaluation of the Global Land Data Assimilation System
699 (GLDAS) Air Temperature Data Products. Journal of Hydrometeorology, 2015, 16: 2463-
700 2480
701 27) Li, L., Gaiser, P. W., Gao, B.-C., Bevilacqua, R. M., Jackson, T. J., Njoku, E. G., et al. (2010).
702 WindSat global soil moisture retrieval and validation. IEEE Transactions on Geoscience and
703 Remote Sensing, 48(5), 2224–2241
704 28) Liu, Y.Y., Parinussa, R.M., Dorigo, W.A., De Jeu, R.A.M., Wagner, W., Van Dijk, A.I.J.M.,
705 McCabe, M.F., Evans, J.P., 2011. Developing an improved soil moisture dataset by blending
706 passive and active microwave satellite-based retrievals. Hydrol. Earth Syst. Sci. 15, 425–436
707 29) Madelon R., N. J. Rodríguez-Fernández, R. van der Schalie et al., Toward the Removal of
708 Model Dependency in Soil Moisture Climate Data Records by Using an L-Band Scaling
709 Reference., 2022, 15: 831-848
710 30) Maeda, T., Taniguchi, Y., & Imaoka, K. (2016). GCOM-W1 AMSR2 level 1R product:
711 Dataset of brightness temperature modified using the antenna pattern matching technique.
712 IEEE Transactions on Geoscience and Remote Sensing, 54(2), 770–782
713 31) Nearing, G. S., Gupta, H. V., Crow, W. T. and Gong, W. (2013), An approach to quantifying
714 the efficiency of a Bayesian filter. Water Resources Research, 49(4): 2164-2173
715 32) Niu G.-Y., K. E. Mitchell, F. Chen, et al., (2011), The community Noah land surface model
716 with multiparameterization options (Noah-MP): 1. Model description and evaluation with
717 local-scale measurements, J. Geophys. Res., 116, D12109, doi:10.1029/2010JD015139

718 33) Niu G.-Y., Z-L Yang, K. E. Mitchell, et al. The community Noah land surface model with
719 multiparameterization options (Noah-MP): 1.Model description and evaluation with local-
720 scale measurements, J. Geophys. Res., 2011, 116, D12109, doi:10.1029/2010JD015139
721 34) Njoku, E. G., Jackson, T. J., Lakshmi, V., et al. (2003). Soil moisture retrieval from AMSR-
722 E. IEEE Transactions on Geoscience and Remote Sensing, 41(2), 215–229
723 35) Peters-Lidard C. D., D. M. Mocko, M. Garcia, et al. Role of precipitation uncertainty in the
724 estimation of hydrologic soil properties using remotely sensed soil moisture in a semi-arid
725 environment. Water Resour. Res., 2008, 44 (5): W05S18, doi:10.1029/2007WR005884
726 36) Peters-Lidard C.D., S.V. Kumar, D.M. Mocko, and Y. Tian. Estimating evapotranspiration
727 with land data assimilation systems. Hydrol. Processes, 2011, 25: 3979–3992
728 37) Peters-Lidard, C. D., D. M. Mocko, M. Garcia, J. A. Santanello, M. A. Tischler, M. S. Moran,
729 and Y. Wu. Role of precipitation uncertainty in the estimation of hydrologic soil properties
730 using remotely sensed soil moisture in a semiarid environment. Water Resour. Res., 2008, 44,
731 W05S18, doi:10.1029/2007WR005884
732 38) Preimesberger W., T. Scanlon, C.-H. Su, A. Gruber and W. Dorigo, "Homogenization of
733 Structural Breaks in the Global ESA CCI Soil Moisture Multisatellite Climate Data Record,"
734 in IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 4, pp. 2845-2862
735 39) Reichle R. H., and R. D. Koster, 2004: Bias reduction in short records of satellite soil moisture.
736 Geophys. Res. Lett., 31, L19501, doi:10.1029/2004GL020938
737 40) Rodell, M.; Houser, P.R.; Jambor, U.; Gottschalck, J.; Mitchell, K.; Meng, C.-J.; Arsenault,
738 K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M.; et al. The Global Land Data Assimilation
739 System. Bull. Am. Meteorol. Soc. 2004, 85, 381–394

740 41) Ryu, D., Crow, W. T., Zhan, X., Jackson, T. J. Correcting unintended perturbation biases in
741 hydrologic data assimilation. J. Hydrometeor., 2009, 10(3): 734-750
742 42) Schaefer, G.L.; Cosh, M.H.; Jackson, T.J. The USDA Natural Resources Conservation
743 Service Soil Climate Analysis Network (SCAN). J. Atmos. Ocean. Technol. 2007, 24, 2073–
744 2077
745 43) Shellito P.J., Small E.E., Cosh M.H. Calibration of Noah soil hydraulic property parameters
746 using surface soil moisture from SMOS and basinwide in situ observations. J. Hydrometeorol.
747 2016,17 (8): 2275–2292
748 44) Wagner,W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., et al. (2013). The
749 ASCAT soil moisture product: A review of its specifications, validation results, and merging
750 applications. Meteorologische Zeitschrift, 22(1), 5–33
751 45) Wang Y., P. Leng, J. Peng, P. Marzahn, R. Ludwig. Global assessments of two blended
752 microwave soil moisture products CCI and SMOPS with in-situ measurements and reanalysis
753 data. International Journal of Applied Earth Observations and Geoinformation. 2021, 94:
754 https://doi.org/10.1016/j.jag.2020.102234
755 46) Wang, J.; Engman, E.; Mo, T.; Schmugge, T.; Shiue, J. The Effects of Soil Moisture, Surface
756 Roughness, and Vegetation on L-Band Emission and Backscatter. IEEE Trans. Geosci.
757 Remote Sens. 1987, GE-25, 825–833
758 47) Yang Z.-L., G.-Y. Niu, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, K. Manning, D. Niyogi,
759 M. Tewari, and Y.-L. Xia (2011), The community Noah land surface model with
760 multiparameterization options (Noah-MP): 2. Evaluation over global river basins, J. Geophys.
761 Res., 116, D12110, doi:10.1029/2010JD015140

762 48) Yin J., X. Zhan, J. Liu, R. R. Ferraro. A New Method for Generating the SMOPS Blended
763 Satellite Soil Moisture Data Product without Relying on a Model Climatology. Remote Sens.
764 2022, 14, 1700. https://doi.org/10.3390/rs14071700
765 49) Yin J., X. Zhan, Y Zheng, C. Hain, M. EK, J Wen, L Fang, J Liu. Improving Noah Land
766 Surface Model Performance using Near Real Time Surface Albedo and Green Vegetation
767 Fraction. Agric. For Meteorol, 2016, 218-219: 171-183
768 50) Yin J., X. Zhan, Y. Zheng, C. Hain, J. Liu, L. Fang. Optimal ensemble size of Ensemble
769 Kalman Filter in sequential soil moisture data assimilation of land surface model. Geophys.
770 Res. Lett., 2015b, 16(28): 6710-6715
771 51) Yin J., X. Zhan, Y. Zheng, J. Liu, et al. Impact of quality control of satellite soil moisture data
772 on their assimilation into land surface model, Geophys. Res. Lett., 2014, 41: 7159-7166
773 52) Yin J., X. Zhan, Y. Zheng, J. Liu, L. Fang, and C. R. Hain. Enhancing Model Skill by
774 Assimilating SMOPS Blended Soil Moisture Product into Noah Land Surface Model. J. of
775 Hydrometeorol., 2015a,16(2): 917-931
776 53) Yin J., X. Zhan. Impact of Bias-Correction Methods on Effectiveness of Assimilating SMAP
777 Soil Moisture Data into NCEP Global Forecast System Using the Ensemble Kalman Filter.
778 IEEE Geosci. Remote Sens. Lett., 2018, doi: 10.1109/LGRS.2018.2806092
779 54) Yin J., Y. Zheng, X. Zhan, et al. An assessment of impacts of surface type changes on drought
780 monitoring. Int. J. Remote Sens., 2015c, 36(24): 6116-6134
781 55) Yin J., Zhan X., Liu J. NOAA Satellite Soil Moisture Operational Product System (SMOPS)
782 Version 3.0 Generates Higher Accuracy Blended Satellite Soil Moisture. Remote Sens. 2020,
783 12, 2861; doi:10.3390/rs12172861

56) Yin, J., Zhan, X., Liu, J., & Schull, M. (2019). An intercomparison of Noah model skills with benefits of assimilating SMOPS blended and individual soil moisture retrievals. Water Resources Research, 55, 2572–2592

57) Yin, J.; Zhan, X.; Hain, C.R.; Liu, J.; Anderson, M.C. A Method for Objectively Integrating Soil Moisture Satellite Observations and Model Simulations toward a Blended Drought Index. Water Resour. Res. 2018, 54, 6772–6791

Table list:

1) Table 1. Individual microwave satellite SM retrievals that are ingested into the daily SMOPS V3.0 and ESA CCI V04.5 blended data products. The acronym IFOV indicates Instantaneous Field of View.
2) Table 2. Perturbation for meteorological forcing parameters and state variables (Peters-Lidard et al. 2008; Kumar et al. 2009). The abbreviations SM1, SM2, SM3, SM4 indicate 0-10 cm, 10-40 cm, 40-100 cm and 100-200 cm SM of Noah-MP model, while SD, SW and LW are standard deviation, and short and long wave radiation, respectively.
3) Table 3. With respect to the SCAN SM observations, CONUS domain-averaged ubRMSE (m^3/m^3)and correlation coefficients for OLP, DAGLD, DACCI and DASMAP cases.

Figure list:

1) Figure 1. The temporal averaged (a) CCI and (b) SMOPS blended SM data products versus the temporal averaged GLDAS 0-10 SM estimations over the global domain from 1 April 2015 to 31 December 2018. The black diagonal line represents they are perfectly matched. The red dash line from the low-left to the upper-right is their linear regression curve. The color bar indicates sample density. The *r* is correlation coefficient.

2) Figure 2. Temporal correlation coefficients between the daily GLDAS 0-10 SM simulations and the daily (a) CCI and (b) SMOPS blended SM data products from 1 April 2015 to 31 December 2018, as well as (a) the corresponding zonal averaged correlation coefficients.

3) Figure 3. Temporal root-mean-square differences (RMSDs) between the daily GLDAS 0-10 SM simulations and the daily (a) CCI and (b) SMOPS blended SM data products from 1 April 2015 to 31 December 2018, as well as (c) the corresponding zonal averaged RMSDs (m^3/m^3).

4) Figure 4. With respect to the quality-controlled SCAN measurements, ubRMSE for 0-10 cm SM simulations during the 1 April 2015-31 December 2018 period: (a) DAGLD, (b) DACCI and (c) DASMP, as well as the corresponding differences: (d) DACCI minus DAGLD, (e) DASMP minus DAGLD, (f) DASMP minus DACCI, (g) DAGLD minus OLP, (h) DACCI minus OLP, and (g) DASMP minus OLP.

5) Figure 5. With respect to the quality-controlled SCAN measurements, ubRMSE for 40-100 cm SM simulations during the 1 April 2015-31 December 2018 period: (a) DAGLD, (b) DACCI and (c) DASMP, as well as the corresponding differences: (d) DACCI minus DAGLD, (e) DASMP minus DAGLD, (f) DASMP minus DACCI, (g) DAGLD minus OLP, (h) DACCI minus OLP, and (g) DASMP minus OLP.

6) Figure 6. With respect to the quality-controlled SCAN measurements, differences in correlation coefficients (r) for 0-10 cm SM simulations during 1 April 2015 to 31 December 2018 period: (a) DACCI minus DAGLD, (b) DASMP minus DAGLD, (c) DASMP minus DACCI.

7) Figure 7. With respect to the quality-controlled SCAN measurements, differences in correlation coefficients (r) for 40-100 cm SM simulations during 1 April 2015 to 31 December 2018 period: (a) DACCI minus DAGLD, (b) DASMP minus DAGLD, (c) DASMP minus DACCI.

8) Figure 8. Season-averaged SM (m^3/m^3) for GLDAS, CCI and SMOPS over the 2015-2018 period with left and right columns for cold (January-April) and warm season (May-September), respectively.

Table 1. Individual microwave satellite SM retrievals that are ingested into the daily SMOPS V3.0 and ESA CCI V04.5 blended data products over the time period from 2015 to 2018. The acronym IFOV indicates Instantaneous Field of View.

Sensor	IFOV (km)	Frequency (GHz)	Swath (km)	Ingested into CCI?	Ingested into SMOPS?	Reference
ASCAT-A	25–35	5.3	~550	Yes	Yes	Wagner et al., 2013
ASCAT-B	25–35	5.3	~550	Yes	Yes	Wagner et al., 2013
AMSR-2	62×35	6.925	~1450	Yes	Yes	Maeda et al., 2016
SMOS	~45	1.4	~ 900	Yes	Yes	Kerr et al., 2010
SMAP	39×47	1.41	~1000	No	Yes	Entekhabi et al., 2010a

915 Table 2. Perturbation for meteorological forcing parameters and state variables (Peters-Lidard et
916 al. 2008; Kumar *et al*. 2009). The abbreviations SM1, SM2, SM3, SM4 indicate 0-10 cm, 10-40
917 cm, 40-100 cm and 100-200 cm SM of Noah-MP model, while SD, SW and LW are standard
918 deviation, and short and long wave radiation, respectively.

		Cross correlation for forcing variable perturbations			
Perturbation type	SD	Precipitation	SW	LW	
Precipitation	0.5 (mm)	1.0	-0.8	0.5	
SW	0.3 (Wm^{-2})	-0.8	1.0	-0.5	
LW	50 (Wm^{-2})	0.5	-0.5	1.0	
		Cross correlation for state variable perturbations			
Perturbation type	SD	SM1	SM2	SM3	SM4
SM1 (0–10 cm)	6.00×10^{-3} m^3m^{-3}	1.0	0.6	0.4	0.2
SM2 (10–40 cm)	1.10×10^{-4} m^3m^{-3}	0.6	1.0	0.6	0.4
SM3 (40–100 cm)	6.00×10^{-5} m^3m^{-3}	0.4	0.6	1.0	0.6
SM4 (100–200 cm)	4.00×10^{-5} m^3m^{-3}	0.2	0.4	0.6	1.0

919
920
921
922
923
924
925
926
927
928
929
930
931
932

933 Table 3. With respect to the SCAN SM observations, CONUS domain-averaged ubRMSE
934 (m^3/m^3)and correlation coefficients for OLP, DAGLD, DACCI and DASMAP cases.

	0-10 cm Soil Moisture		40-100 cm Soil Moisture	
	ubRMSE	*r*	ubRMSE	*r*
OLP	0.0533	0.549	0.0588	0.198
DAGLD	0.0522	0.564	0.0572	0.237
DACCI	0.0521	0.566	0.0563	0.247
DASMP	0.0515	0.575	0.0550	0.269

935
936
937
938
939
940
941
942
943
944
945
946
947

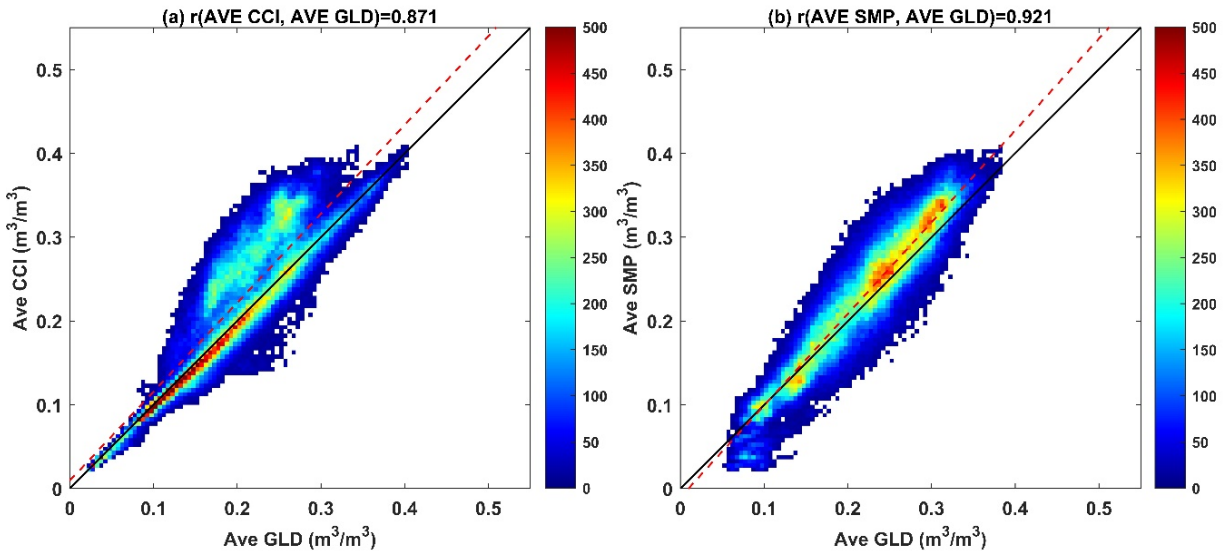

Figure 1. The temporal averaged (a) CCI and (b) SMOPS blended SM data products versus the temporal averaged GLDAS 0-10 SM estimations over the global domain from 1 April 2015 to 31 December 2018. The black diagonal line represents they are perfectly matched. The red dash line from the low-left to the upper-right is their linear regression curve. The color bar indicates sample density. The *r* is correlation coefficient.

948

949

950

951

952

953

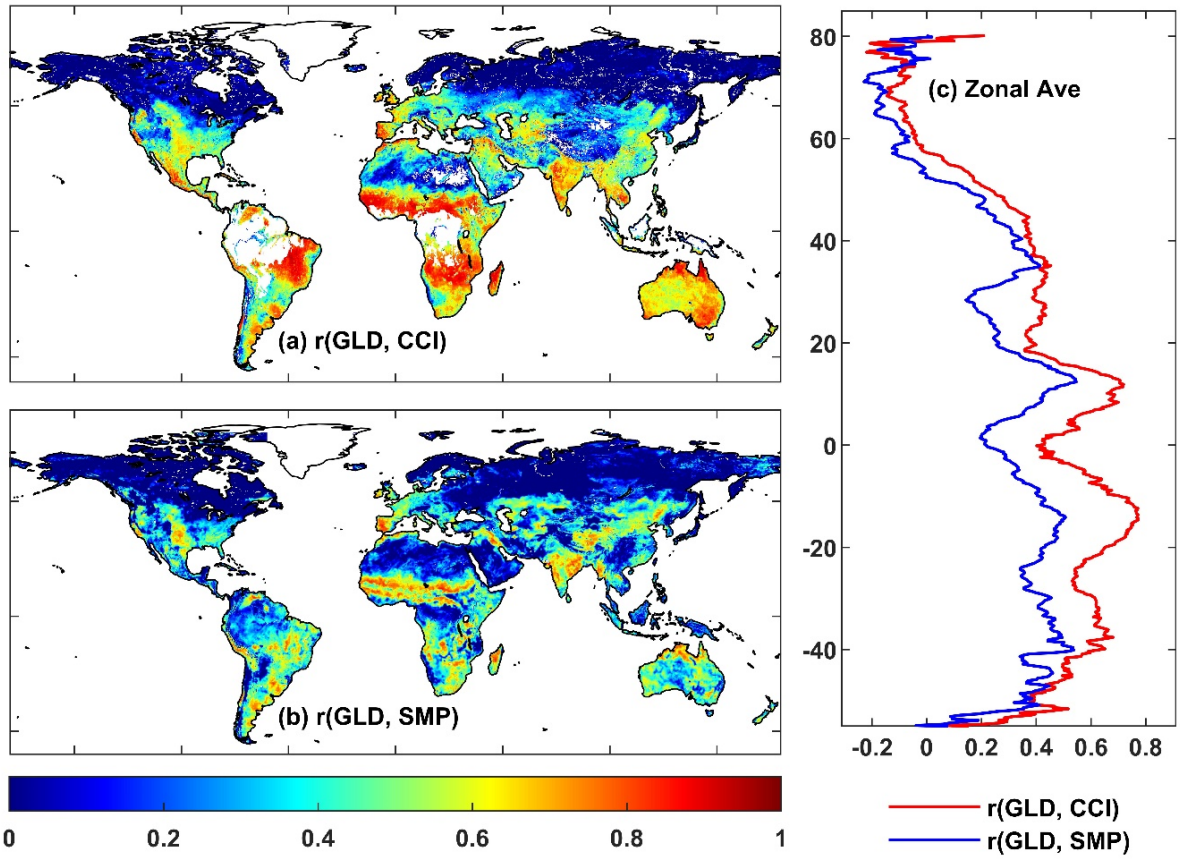

Figure 2. Temporal correlation coefficients between the daily GLDAS 0-10 SM simulations and the daily (a) CCI and (b) SMOPS blended SM data products from 1 April 2015 to 31 December 2018, as well as (a) the corresponding zonal averaged correlation coefficients.

954

955

956

957

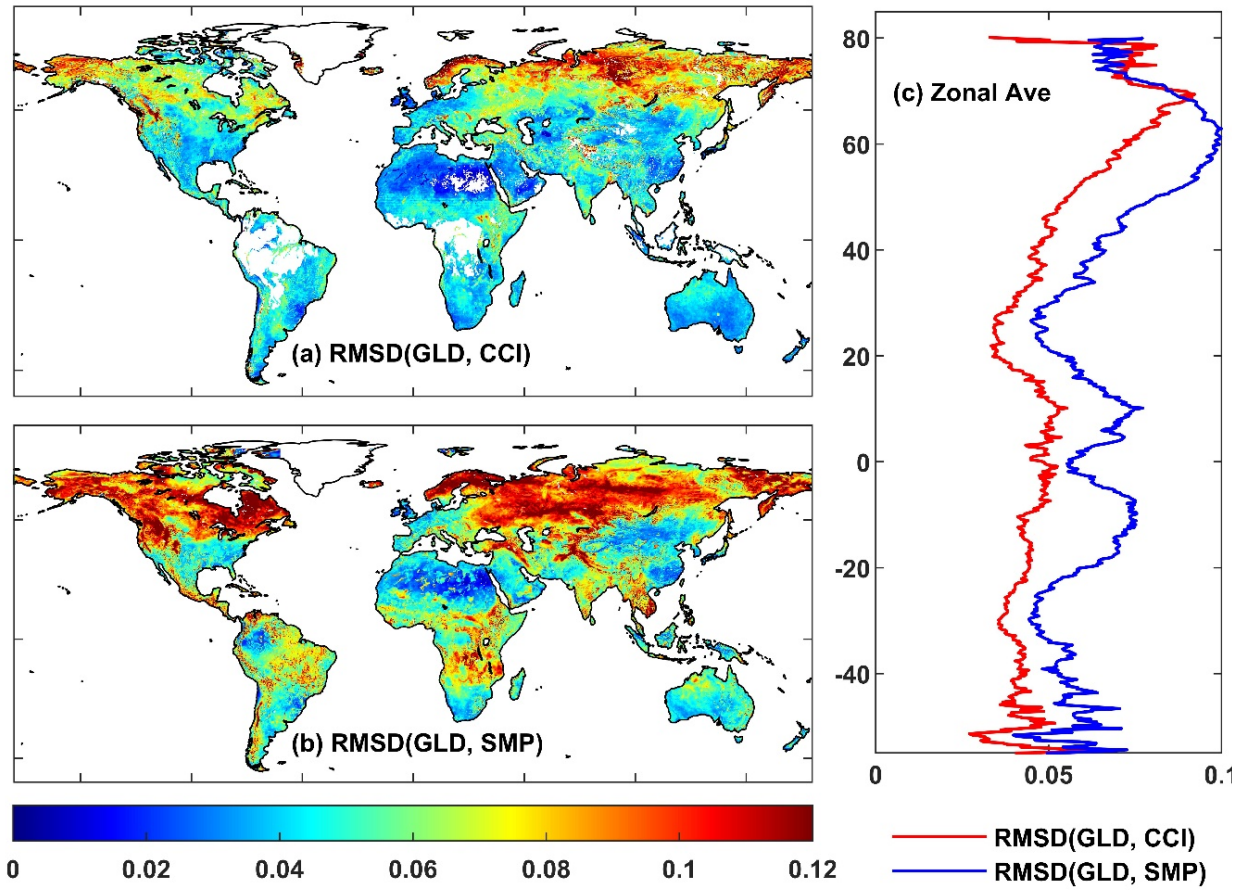

Figure 3. Temporal root-mean-square differences (RMSDs) between the daily GLDAS 0-10 SM simulations and the daily (a) CCI and (b) SMOPS blended SM data products from 1 April 2015 to 31 December 2018, as well as (c) the corresponding zonal averaged RMSDs (m^3/m^3).

958

959

960

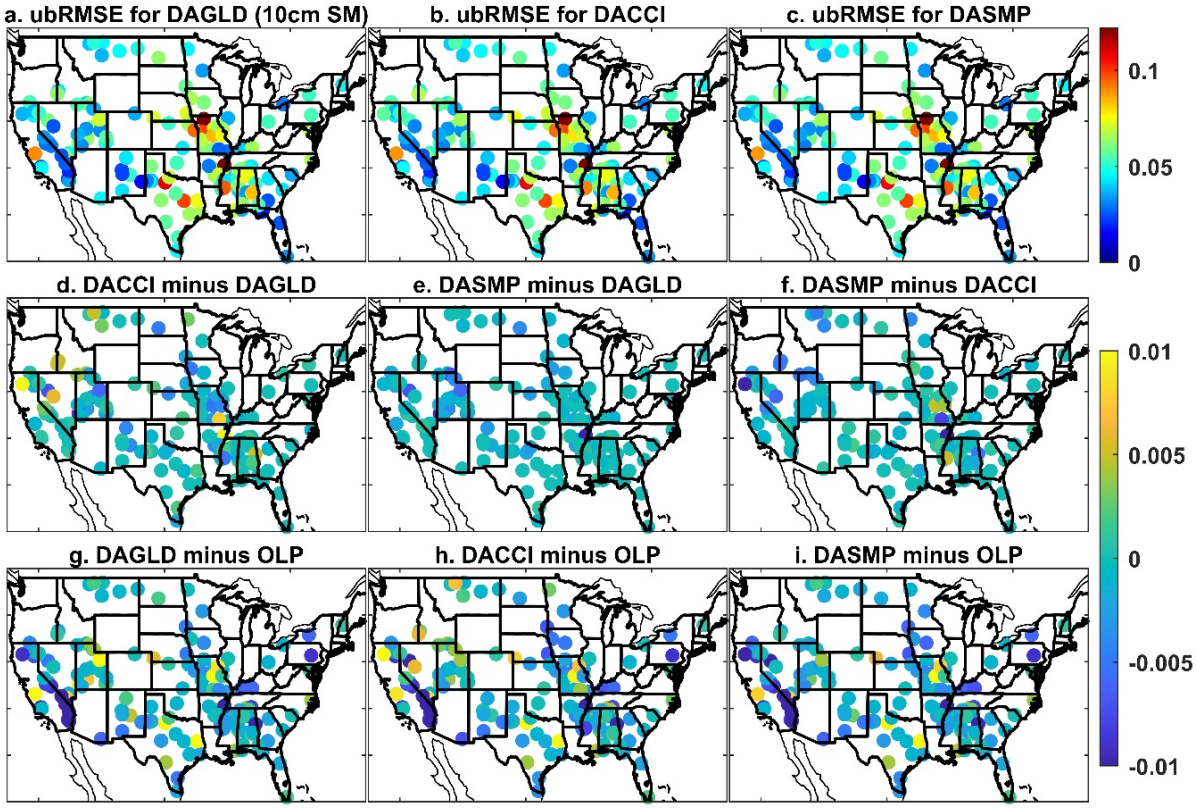

Figure 4. With respect to the quality-controlled SCAN measurements, ubRMSE for 0-10 cm SM simulations during the 1 April 2015-31 December 2018 period: (a) DAGLD, (b) DACCI and (c) DASMP, as well as the corresponding differences: (d) DACCI minus DAGLD, (e) DASMP minus DAGLD, (f) DASMP minus DACCI, (g) DAGLD minus OLP, (h) DACCI minus OLP, and (g) DASMP minus OLP.

961

962

963

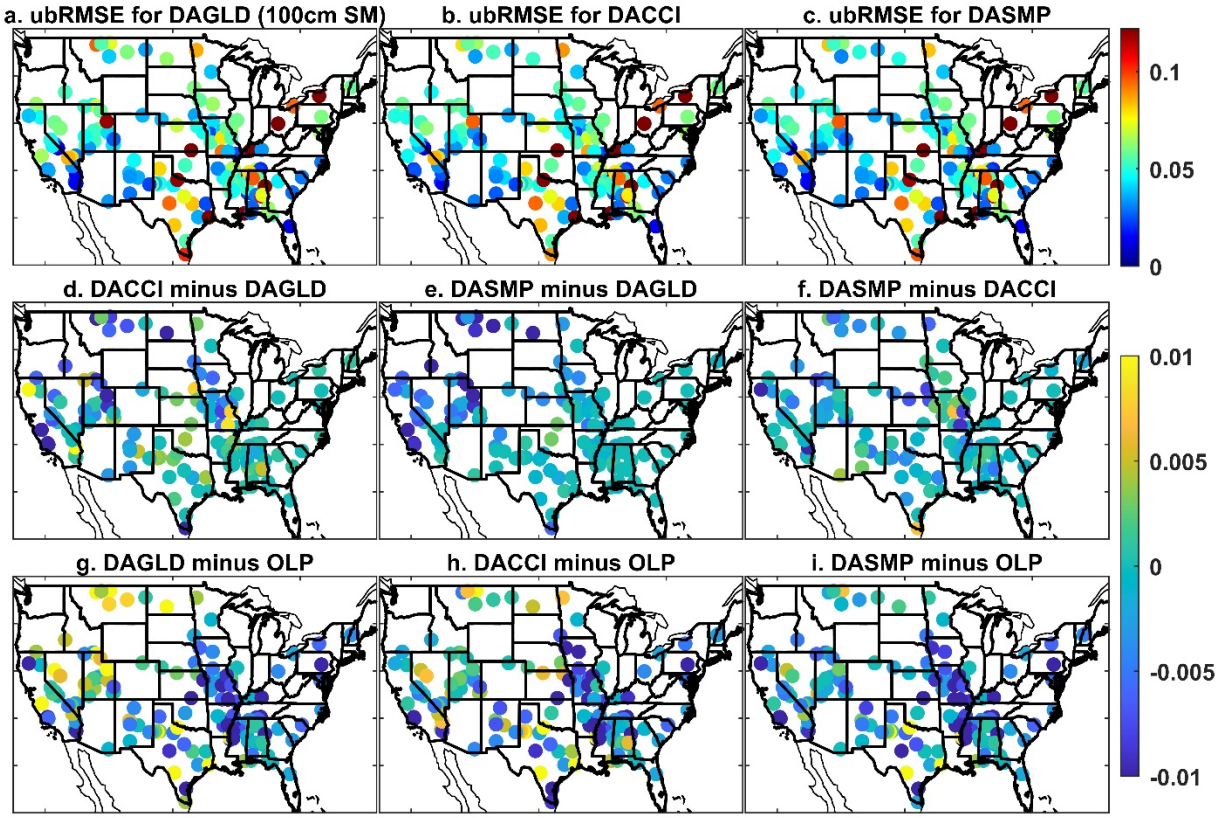

Figure 5. With respect to the quality-controlled SCAN measurements, ubRMSE for 40-100 cm SM simulations during the 1 April 2015-31 December 2018 period: (a) DAGLD, (b) DACCI and (c) DASMP, as well as the corresponding differences: (d) DACCI minus DAGLD, (e) DASMP minus DAGLD, (f) DASMP minus DACCI, (g) DAGLD minus OLP, (h) DACCI minus OLP, and (g) DASMP minus OLP.

964

965

966

967

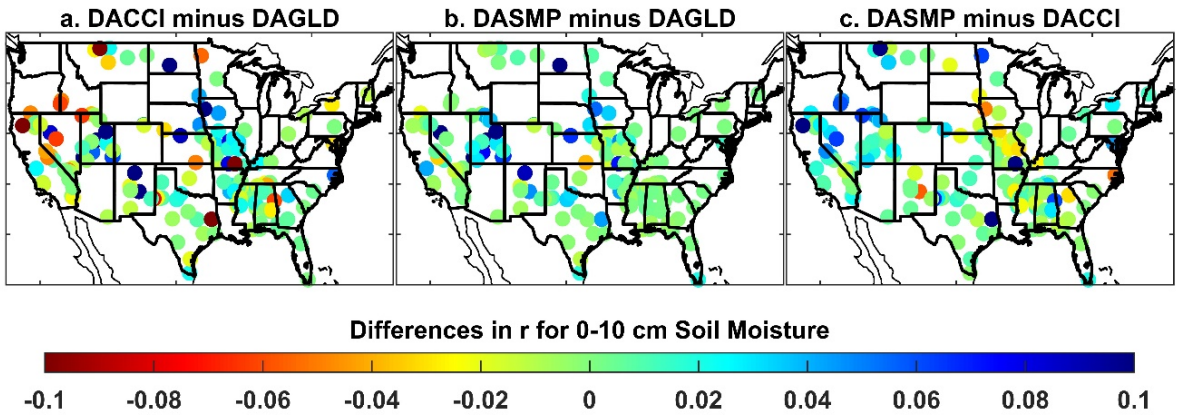

Figure 6. With respect to the quality-controlled SCAN measurements, differences in correlation coefficients (r) for 0-10 cm SM simulations during 1 April 2015 to 31 December 2018 period: (a) DACCI minus DAGLD, (b) DASMP minus DAGLD, (c) DASMP minus DACCI.

968
969
970
971
972
973
974
975
976
977

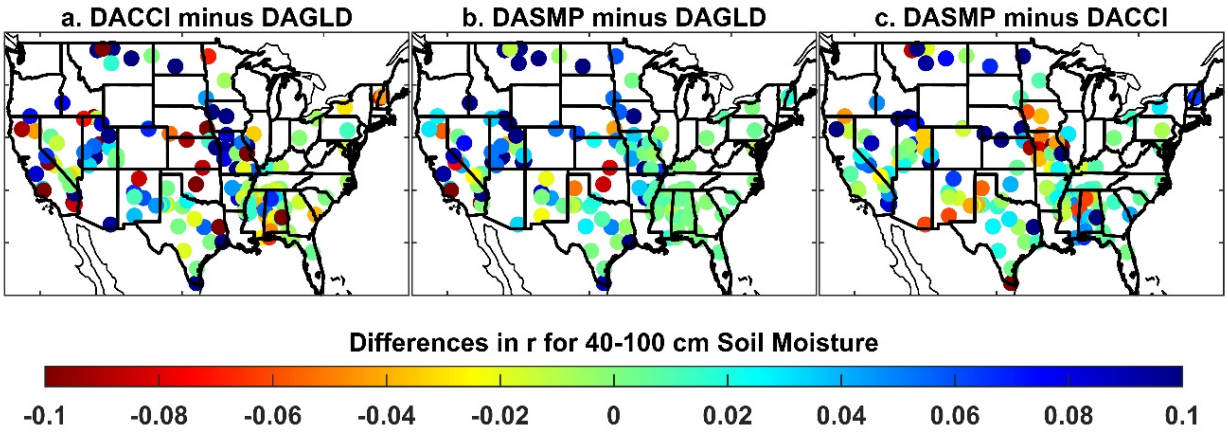

Figure 7. With respect to the quality-controlled SCAN measurements, differences in correlation coefficients (r) for 40-100 cm SM simulations during 1 April 2015 to 31 December 2018 period: (a) DACCI minus DAGLD, (b) DASMP minus DAGLD, (c) DASMP minus DACCI.

978

979

980

981

982

983

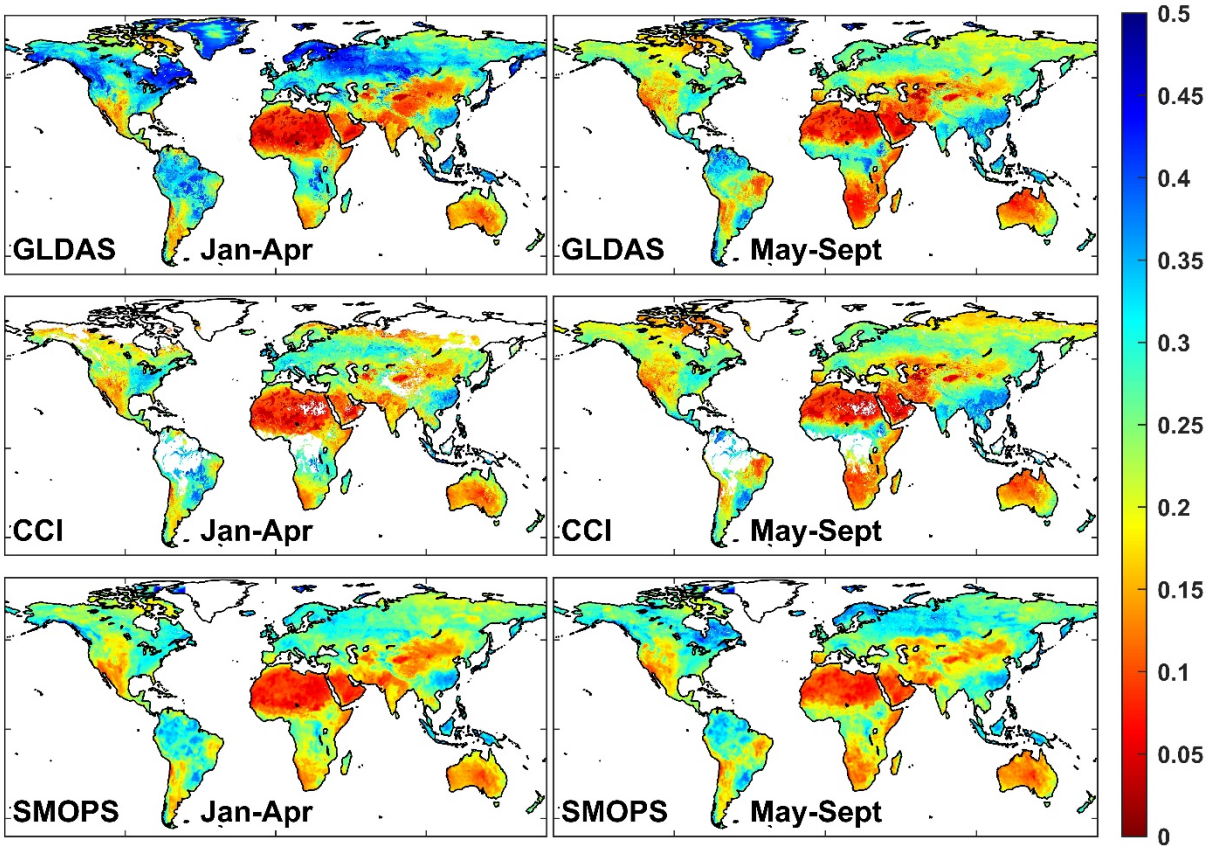

Figure 8. Season-averaged SM (m^3/m^3) for GLDAS, CCI and SMOPS over the 2015-2018 period with left and right columns for cold (January-April) and warm season (May-September), respectively.

984