Evaluation of SMOS Soil Moisture products over continental US using the SCAN/SNOTEL network

Al Bitar Ahmad, Delphine Leroux, Yann H. Kerr, Olivier Merlin, Philippe Richaume, A. Sahoo, E. F. Wood

Evaluation of SMOS Soil Moisture products over continental US using the

SCAN/SNOTEL network

Ahmad Al Bitar[1]

DelphineLeroux[1]

YannH. Kerr[1]

Olivier Merlin[1]

Philippe Richaume[1]

AlokSahoo[2]

Eric F. Wood[2]

[1]Centre d'Etudes Spatiales de la Biosphère, France

[2]Princeton University, Princeton, NJ 08544, United States

Corresponding author details:

email: ahmad.albitar@cesbio.cnes.fr

ph: +33561558537

fax: +335 61558599

Abstract—The Soil Moisture and Ocean Salinity (SMOS) satellite has opened the era of soil moisture products from passive L-Band observations. In this study validation of SMOS products over continental US is done by using the SCAN/SNOTEL soil moisture monitoring stations. The SMOS operational product and the SMOS reprocessing product are both usedand compared over year 2010. First, a direct node to site comparison is performed by taking advantage of the oversampling of the SMOS product grid. The comparison is performed over several adjacent nodes to site, and several representative couples of site-node are identified. The impact of forest fraction is shown through the analysis of different cases across the US. Also the impact of water fraction is shown through two examples in Florida and in Utah close to Great Salt Lake. A radiometric aggregation approach based on the antenna footprint and spatial description is used. A global comparison of the SCAN/SNOTEL versus SMOS is made.Statistics show an under estimation of the soil moisture from SMOS compared to the SCAN/SNOTEL local measurements. The results suggest that SMOS meets the mission requirement of 0.04 m^3/m^3 over specific nominal cases, but differences are observed over many sites and need to be addressed.

Index Terms—microwave, soil moisture,L-Band, validation, SMOS, SCAN, SNOTEL.

I. INTRODUCTION

Surface soil moisture is a key variable in hydrological modeling and weather forecast because it controls the fluxes of water and energy at the interface between the soil and the atmosphere. Soil moisture is used in hydrological models to determine infiltration and runoff rates at local scale. In land surface models soil moisture is a crucial parameter to determine the evaporative fraction at the surface and the infiltration rate into the root zone. Several studies showed the importance of soil moisture for climate change studies [1], surface atmosphere interactions [2], weather forecast [3] and agriculture applications [4].

Soil moisture can be obtained at different scales using local instruments, airborne or space borne sensors at microwave frequencies. Many algorithms and products have been developed to retrieve global soil moisture from AMSR-E(Advanced Microwave Scanning Radiometer – Earth Observing System): NASA AMSR-E official product [5], USDA [6], VUA [7], SWI [8] and from ASCAT [9]and ERS[10]. Based on the same satellite observations, those algorithms can give very different results [11]. Many studies have showed the utility of passive L-Band observation for soil moisture ([12], [13]). These observations are available since January 2010 from the Soil Moisture and Ocean Salinity (SMOS) satellite([14], [15]). SMOS mission has been delivering valuable data since its launch in November 2009. This is the second Earth Explorer Opportunity mission to be developed as part of ESA's Living Planet Program. The mission is led by ESA with contributions fromthe Centre National d'EtudesSpatiales (CNES) in France and the Centro para el DesarrolloTecnologico Industrial (CDTI) in Spain. SMOS presents a unique opportunity to probe the Earth with its L-Band 1.4Ghz 2D interferometer [16], it is also the first satellite operating in passive L-Band microwave. Since the atmosphere is invisible at microwave frequencies, SMOS is an all-weather system with a revisit period of less than 3 days. Crossing times are around 6 am and 6 pmlocal time for ascending and descending orbits respectively. SMOS has a 1000km swath width and a nominal resolution of about 43 km.

Soil moisture from passive microwave has the advantage to cover large areas and to spot large-scale events such as large scale precipitation patterns or floods,but it aggregates heterogeneities from local to regional scale ([17], [18]) which renders validation difficult. So taking into account the spatial and temporal variability is important when validating SMOS satellite data over ground measurements ([19],[20]). This can be done either by up-scaling the local observations or downscaling the satellite observations into an intermediate resolution. Many studies have been made to compare and validate satellite data from airborne and ground measurements by using up-scaling technics. Monitoring stations have been mainly used to

validate AMSR-E products over the US ([21],[22],[23], [24]),over Australia[25] and over Sahelian regions[26]. Up-scaling relations can be derived for local monitoring stations, as in [27]for SCAN sites and [28] for AMMA sites, based on temporal stability concept proposed by [18].Validation of AMSR-E soil moisture products is done over four densely instrumented soil moisture networks located in four different climate regions in[23]. This paper points out the difficulty in matching the footprint of passive microwave sensors with the network of local instruments while considering the local variability in topography, vegetation or soil type. On the other hand, airborne measurements can be presented also as an alternative[29] but generally they have limited temporal coverage and are not capable of representing the temporal dynamics of soil moisture. Another alternative is to upscalevia modeling. In this way a validation approach for SMOS based on distributed land surface modeling is presented in[30]. An assimilation approach for validation of satellite data is presented in [11]. Downscaling,or disaggregating,the observation is also a mean to bridge the gap of spatial resolution. A physical based disaggregation approach is proposed in ([31],[32)and a statistical approach in [33].

The goal of this study is to validate the SMOS data over continental US using local measurements from the SCAN[34] and SNOTEL network. This network has the advantage of representing a variety of conditions across the United States where radio frequency interference in L-Band is very low. On the other hand the SCAN and SNOTEL networks have a sparse density. In this paper we make comparisons using several approaches. The first approach is a direct site to observations comparison. The goal of this approach is to detect combinations of site and satellite nodes with good global statistics and representative dynamics. The relatively high temporal resolution of SMOS observations (3 days), the over sampling of the SMOS products (15km) and the size of the SCAN/SNOTEL network (more than 444where (0-5) cm soil moisture measurements are available) make this goal achievable. Even though no scale change is operated, the impact of heterogeneity is analyzed through the description of the

surface in the retrieval algorithm. Careful analysis of the results enhances our understanding of the SMOS retrieval algorithm and presents the basic aspects to look for in a future scale change approach. The second approach consists of an up-scaling of local observations by a weighted average when many stations are available. This approach makes use of the antenna footprint and the surface heterogeneity. The last approach is to make a global comparison over the SCAN/SNOTEL dataset. The simplification of a site to node comparison is here compensated by the number of considered observations. Aninter-comparison between two SMOS products versions is also examined.

II. Datasets

A. SMOS soil moisture products

The main products used in this paper are the SMOS Soil Moisture Level 2 User Data Products (SML2UDP) as delivered through ESA. Theseproducts contain the retrieved geophysical parameters (soil moisture, optical thickness of the low vegetation, etc…), complementary parameters (number of TB records used, surface level modeled TB at 42.5° incidence angle, etc…) and flags. A Data Quality Index (DQX) is associated to each retrieved geophysical parameter and can be associated to the uncertainty of the retrieved parameter; it is equal to -999 in case of no retrieval. The product is provided over the ISEA-4H9 grid (Icosahedral Snyder Equal Area Earth fixed) with equally spaced nodes at 14.989 km. Products for descending and ascending half-orbits are separated. A detailed description of the level 2 algorithm used for the retrievals is provided in [35] and[36].Theradiative model used in the algorithm is the LMEB model provided in [38]. The resolution associated with the TBs varies with the observation angle as the ellipsoidal footprint changes in size and shape with the viewing angle. An analytical formulation for a mean (angular independent) footprint is derived in [35].Fig. 1(a)shows the 3db mean footprint versus distance from node center. At 20km radius

for a 3db attenuationthe contribution is 0.5. By computing the normalized cumulative sum of the mean footprint one can determine the contribution of a circular area to a SMOS observation over a homogeneous surface. As an example 80% of the signal can be associated to a circular area of 20km radius and 90% of the signal to a circular area with a radius of 25 km. The SMOS product's grid is over sampled at 15 km compared to the signal resolution. The over sampling rate is lower than the nominal Nyquist rate of 21.5km and induces high correlation in the SMOS product. Fig. 1(b)shows the spatial-correlation function of the mean footprint in the case of a homogeneous surface. This function is obtained numerically by computing the autocorrelation of the convolution of a randomly generated TB field by the mean antenna footprint. FromFig. 1(b),one can associate a 0.746 correlation coefficient with 25km separation distance. As indicated above this correlation is for a randomly distributed surface cover.

Emissions from microwave radars in the protected L-Band are affecting SMOS acquisitions. This is common to many microwave missions but has an important impact on SMOS acquisitions because of the low emissions in the L-Band and the size of the swath. Over the US the probability of radio frequency interference (RFI) is low and efforts to maintain this situation are made. Nevertheless in some cases emissions in the SMOS protected band are observed and need to be accounted for. The percentage of infected TB records is given by the ratio of potentially infected TB record over the total number of records Equation(1).

$$PRFI = \frac{(NRFIx + NRFIy - NCLEANED) \times 100}{Ntot} \tag{1}$$

withTBx and TBy the brightness temperatures in x and y polarization respectively at antenna reference plane. $NRFIx$ and $NRFIy$ the number of potentially infected TBx and TBy and $NCLEANED$ the number of outlier TBthat are detected during level 2 soil moisture retrieval and added to $NRFIx$ and $NRFIy$. $Ntot$is the total number of TB records at node during a visit. The

number of potentially infected *TBx* and *TBy* are provided from level 1C processing. This percentage of RFI focuses on the hard RFI detection based on L1C TB.

The level 2 soil moisture processor provides also a complementary product the SMOS Soil Moisture Level 2 Data Analysis Product (SML2DAP) that enables data check in expert scientific laboratories. In this paper we consider the geometric fractions and the radiometric fractions extracted from this product. The geometric fractions correspond to the fraction of total surface covered by one of the 11cover types (soil, forest, barren, snow …) considered in the SMOS processor. The radiometric fraction is the geometric fraction convoluted with the mean footprint at a 4km resolution. It is more representative of the surface as observed by the instrument.Bare soil and/or low vegetation are considered as nominal surfaces. SMOS retrievals are expected to be of better quality over those surfaces. The retrieval is considered over forest cover when the radiometric forest fraction is more than 60%. Taking these fractions into consideration is important as the processor optimizes the soil moisture inversion over the dominant surface fraction and all other fraction contributions are computed using fixed parameters (soil moisture, vegetation optical thickness, surface roughness…). So the retrieved soil moisture is expected to be representative of a surface fraction on which inversion has been operated.Table 1 shows the surface fractions for selected nodes analyzed in the results section.

Two SMOS datasets are considered. The first is processed at DPGS (Data Processing Ground Segment) and corresponds to the operational SMOS product delivered by the European Space Agency (ESA). We refer to it as DPGS in this paper. The start period for the used data is the 14th ofJanuary 2010 and spans for a year. This dataset has heterogeneous processing configurations and processor versions. The Level1 processor that constructs the TBs and the Level2 that retrieves Soil Moisture have been updated several times since the beginning of the operational processing. Fig. 2 and Fig. 3 show the processing versions and configurations of the DPGS dataset respectively.

The second SMOS dataset uses the reprocessing campaign products. We refer it as REPv4. The level1C product has been reprocessed at DPGS center and L2 reprocessed at CNES center for ESA. The main advantage of this dataset is that it has a uniform processing configuration using latest processor versions. The level1 processor uses uniform calibration over the whole year: processor version is 3.4.6 and configuration file version is 3. In this study, products from the version 4 of the soil moisture level2 processor with configuration file version 5 are used. This version uses only the dual polarization TBs to retrieve soil moisture due to high noise values in the mixed polarization acquisitions.2958 soil moisture products are used for the DPGS and REPv4 datasets.

B. ECMWF SMOS products

The SMOS Level2 processor uses a custom made climate data product (AUX_ECMWF) from the European Centre for Medium-range Weather Forecasting (ECMWF)as ancillarydata. The product is obtained by a pre-processor that computes spatiotemporal averaging of the ECMWF forecast products on the ISEA-4H9 grid. The ECMWF product is considered as an internal product.

The skin soil moisture in the ECMWF product is used as a fixed (default) soil moisture value for the (non-nominal) fractions and to initialize the soil moisture overthe nominal fraction if needed. The skin term refers to the surface between the soil and the atmosphere. The soil moisture retrieval is made over the nominal fraction only. The product also contains air temperature, skin temperature, root zone soil temperature, precipitations, etc...These are used as parameters to compute the contributions and check threshold values for processing configuration.

C. SCAN/SNOTEL datasets

The Soil Climate Analysis Network (SCAN) from the National Soil Survey Center (NRCS) of the U.S. Department of Agriculture. (USDA) gives free near real time (NRT) with hourly sampling access to climatic stations data across the U.S. (http://www.wcc.nrcs.usda.gov)[34]. The stations are equipped with a multitude of sensors (air temperature, relative humidity, soil moisture at different depths, soil temperature at different depths, solar radiation, wind speed, precipitation, etc.). The most important sensors for this study are the soil moisture at 2in (~5cm), soil temperature at 2in and precipitation. The soil moisture instrument is the Hydra Probe (Stevens). The original objective of the SCAN network was to improve decision-making in agriculture, but the network has been extensively used in research activities. ([22], [23], [27] and [37])used the SCAN sites data combined with other data sources for validation of soil moisture products. The network has a low density compared to the heterogeneity of soil moisture over the SMOS footprint but covers a wide variety of soil types and climates over continental US (Fig. 4). The highest density is in the great southern plains near the Lower Mississippi River Basin with an average distance of 32km. The NRT data is provided after screened sensor limits and no additional processing or quality check is provided.

The NRCS gives also NRT access to the SNOwpackTELemetry (SNOTEL) network. This network covers the Western U.S. and Alaska. SNOTEL stations are in majority installed over mountainous regions (Mountains Rocky, Colorado) with forests. The network has been widely used by the research snow community. In many SNOTEL sites a soil moisture Hydraprobe is installed. In this study, only stations with an installed hydraprobe at 2in are considered. Periods with no snow (null Snow Water Equivalence) and no soil freezing occurrence (Tsoil> 0°C) are considered.

After excluding stations where soil moisture at 5 cm wasnot available, 444 of the 979 available SCAN and SNOTEL sitesremained(Fig. 4). Many sites are located in mountainous

regions with forest surface cover. In the considered SMOS products,soil moisture is not retrieved for nodes flagged with strong topographic index [35]. Also for forest covers when forest LAI is very high the contribution from the soil is very low and soil moisture is not retrieved. After removing the sites where the corresponding SMOS nodes do not have soil moisture retrievals we get 335 sites where soil moisture can be compared. For 2010, hourly data have been downloaded for the SCAN and SNOTEL sites of interest over continental US. Stations without soil moisture instruments at 2in are excluded. Also records where soil moisture is zero and constant for a period of more than 3 days are ignored. No further data processing is done. Table 2 shows the characteristics of some selected sites analyzed in the results sections. The forest cover description over the sites is obtained from the USDA Forest Inventory Data Online (FIDOII). The dynamics of the vegetation cover is obtained from the Leaf Area index (LAI) acquired from Moderate-Resolution Imaging Spectroradiometer (MODIS) 10 days products.

III. DATA EXTRACTION

The data extraction has been performed over the neighboring nodes located less than 20km from the SCAN/SNOTEL sites and not over the closest node because it might not be the most representative.In fact, the monitoring station position can be over the representative surface fraction, or not, depending on the chosen node. For example an increase in the water fraction can degrade the soil moisture retrieval. As an indicative value a 5% change in water fraction can induce a $0.04m^3/m^3$ change in soil moisture[15].The first comparison consists in performing a node-to-site comparison. Data over the nodes surrounding a SCAN/SNOTEL site are extracted. The extraction is done over the orbits covering continental US. The soil moisture value is considered if the inversion is successful and if the percentage of radio frequency interference (RFI) is lower than 10%. The inversion is successful if the DQX associated with the retrieved

soil moisture is different from -999 and if the unsuccessful retrieval flag is not activated

(NO_PROD_FLAG = 0).

Alternatively to node to site comparison, average soil moisture from close sitesare computed. All sites at a distance of less than 40 km are used to compute a radiometric average. Averaging is operated if more than three sites data are available using the radiometric fractions as in Equation(2). This averaging method follows the same philosophy as in the SMOS soil moisture retrieval algorithm [35].

$$\overline{SM} = \frac{\sum FRAC_i Mean\_WEF(\rho)_i SM_i}{\sum FRAC_i Mean\_WEF(\rho)_i} \tag{2}$$

with $FRAC_i$ the geometric fraction at the node corresponding to the local surface cover of site i, $Mean\_WEF(\rho)_i$ the mean footprint value at distance ρ from the node center, SM_i local soil moisture and $\overline{SM}$ average soil moisture over node.The mean footprint (*Mean_WEF*) is an approximate analytical function of the average,over all observation angles, of the weighting function used to convolute the brightness temperatures from finer resolution. It is obtained from the synthetic antenna pattern[35]. The spatial averaging used here assumes that the spatial up scaling relation between soil moisture and TB is linear.So the Mean_WEF is used in Equation (2)to compute the soil moisture averaging.The Mean_WEFis available in [35] and is given here in Equation(3).

$$Mean\_WEF(\rho) = C_{MWEF\ 2} + WEF\left[\frac{\rho}{C_{MWEF\ 1}} \cdot \frac{\pi}{C_{WEF\ 1}}\right] \text{ for } \rho \in [0, C_{MWEF\ 1}] \tag{3}$$

with ρ the distance on the earth coordinates from node center, C_{MWEF1}=40km,C_{MWEF2}= 0.027 and C_{WEF1}=73.3 are prescribed parameters. The WEF weighting function is the angle dependant it is used here to compute the Mean_WEF.

$$WEF(\rho_{DC}) = \frac{[\text{sinc}(C_{WEF\ 1} \cdot \rho_{DC})]^{C_{WEF\ 2}}}{1 + C_{WEF\ 3} \cdot \rho_{DC}^{C_{WEF\ 4}}} \text{ for } C_{WEF\ 1} \cdot \rho_{DC} \leq \pi \tag{4}$$

withρ_{DC} the distance in the director cosines reference, C_{WEF1}= 73.3, C_{WEF2} = 1.4936,C_{WEF3} = 524.5 and C_{WEF4} = 2.103.

The fractions are provided in the SML2DAP. The fractions are considered if no freezing or snow is detected. Each SCAN/SNOTEL site is classified as nominal or forest depending on local site configuration.

IV. RESULTS AND DISCUSSIONS

The soil moisture products described in section (II)are compared to in-situ data. First The objective is to present comparison over some selected sites presenting large variability in terms of soil properties, climatic conditions (dry weather, freezing), and surface cover (low vegetation, bare soil, agricultural area). Weshow the impact of forest fraction and water fraction on the retrieval of soil moisture over nominal surfaces (surfaces with more than 60% bare soil and low vegetation surfaces). Time series and scatterplots are presented over those sites. Emphasis is given on the surface description and its impact on the retrieval algorithm. The global results are presented afterwards.

A. Node to site comparison

Comparision over nominal surfaces

In this section we consider direct node to site comparison over selected sites where the nominal (bare soil and low vegetation) fraction is above90%.Fig. 5shows the retrieved soil moisture and vegetation optical thickness for DPGS dataset and REPv4 dataset over node 165158 and SNOTEL site 581. The SNOTEL 581 site is located in Dawson, Montana on a relatively flat area at an altitude of 817m. Primary land use is cropping with annual rainfalls of 320 mm. The surface fraction is 100% nominal (bare soil and low vegetation). This region is subject to harsh winters with soil freezing. Soil texture is 63% sand and 13% clay. At the beginning of the year no soil moisture retrieval are performed as the retrieval model associated with the surface description (snow and freezing soil) doesnot enable any soil moisture retrieval. But several soil moistureretrievalsare attempted at the end of January(four for DPGS and one

for REPv4). The soil moisture retrieval algorithm will retrieve selected physical parameters depending on surface description and dynamic ancillary data. The soil moisture retrieval is attempted even though there are harsh weather conditionsbecause the surface description from ancillary data does not indicate freezing or snow. This soil moisture should be considered here as an indicator for frozen soil related to the dielectric constant of the frozen soil and its temperature. Retrievals givea low soil moisture value. This result is consistent with the expected results for a frozen soil covered with dry snow([39], [40], [41]). Note that the retrieved soil moisture can be considered here as an indicator and not a physical parameter because the soil is frozen. The site data associated to those retrievalsshow also a low soil moisture value of 0.05 m^3/m^3.After March the surface fraction is nominal and soil moisture retrieval is attempted more frequently. The soil moisture retrieved by SMOS follow the dynamic of the site data very well. The values are globally respected with a bias of 0.021 m^3/m^3 and a RMSE of 0.057 for REPv4, and a bias of 0.044 m^3/m^3 and armse of 0.069 m^3/m^3 for the DPGS dataset. The SMOS results have a range of 0.01to 0.32 m^3/m^3, the site data show a wider range of 0.05 to 0.4 m^3/m^3. From the time series it is evident that SMOS underestimates the lower soil moisture values. Probable explanationsare:

- the penetration depth at L-Band is lower than 5 cm[42].

- the spatial heterogeneity is not well represented at 40km : this explanation can be dismissed as it cannot explain the negative bias we are obtaining. As we are comparing many revisits over many sites, we have many samples. Since the number of samples is high we can consider that all cases are represented. So the lake of low spatial resolution is compensated with the number of samples.

- In-situ soil measurements at 5 cm are overestimating soil moisture.One reason for the high values of soil moisture (0.2 m^3/m^3) in dry periods is soil compaction[43].

- the soil moisture retrieval algorithm is making a systematic bias.

- erroneous ancillary data like soil texture and climate data

To answer those questions a thorough analysis is needed with an extensive dataset: several soil moisture instruments for redundancy and gravimetric measurements.During rain events SMOS soil moisture overestimates soil moisture with an overshoot.The most probable reason is the presence of a water lens during rain events as the infiltration capacity of the soil is exceeded.The uncertainty associated with the soil moisture is very high (larger than 0.1 m^3/m^3), when we have an overshoot. So uncertainty can be used to filter the data.ECMWF soil moisture product overestimates the soil moisture values with a bias of 0.081 m^3/m^3. The vegetation optical thickness is relatively stable across the time serieswith a mean of 0.12 and a standard deviation of 0.078. There is noclear correlation between Leaf Area index (LAI) from MODIS and the optical thickness. The highest optical thickness values do not coincide systematically with the rain events so theycannot be associated with rain interception when low vegetation is present.Finally, there is no correlation between distance from nadir and optical thickness values.

Fig. 6 shows the time series over 8 selected sites around the US for the REPv4 dataset. The SCAN sites are selected in order to have a high dynamic range in soil moisture (minimum soil moisture < 0.2 and maximum soil moisture > 0.3). The majority of SCAN/SNOTEL sites, like site 2084 (Fig. 6 (h)), have a high minimum soil moisture value (0.2 m^3/m^3) throughout theyear. These values are not realistic for the majority of the sites. They may be due to the choice of the instrument calibration curves, soil compaction and ponding effects.The high valuesover the Mississippi can be associated with the stagnant water in low elevation regions with high clay percentage[37]. This is a reasonable explanation for the Mississippi sites but doesnot apply to many sites covering a wide variety of climates. The results over the different sites bring the same conclusion as for site 581: SMOS data at 40km have a higher dynamics than the local site measurements. An overshoot is observed in SMOS soil moisture retrievals during rain events but it is not systematic. During dry periods SMOS retrievals tend to be dryer than observations.

As explained earlier this can be associated with a shorter penetration depth. As for site 581, the SMOS-ECMWF soil moisture systematically overestimates soil moisture. The SMOS retrievals match well with the dynamics of the site measurements but local disparities can be observed. Note that the SMOS-ECMWF is different from the ECMWF operational soil moisture products.

Fig. 7 shows the scatter plot corresponding to the time series shown in Fig. 6. The scatters show the high range of SMOS data,a limited range for some SCAN/SNOTEL sites, and the underestimation of SMOS retrievals mentioned earlier.Table 3 shows the bias, the correlation coefficientand the rmseof the selected sites. The rmse ranges between 0.032 m^3/m^3 and 0.11 m^3/m^3. The highest rmse value (0.11 m^3/m^3) corresponds to site 2084 where minimum soil moisture in dry season is 0.18 m^3/m^3. The use of a higher density of site stations should enhance the results on those same nodes which suggests that the mission objective of rmse = 0.04 m^3/m^3 over nominal surfaces is attainable. SMOS retrievals follow the dynamics of the site soil moisture. This result, combined with the comparison of other sites across continental US, shows that in some conditions a direct comparison between remote sensed soil moisture at 40 km and site data is feasible.

Effect of forest cover on soil moisture retrieval

In this section the impact of forest cover over the retrieval of soil moisture is studied.The objective is not to make an in depth analysis of the radiometric model behind the soil moisture retrievals under forests but to give an insight on the quality of inversions over nominal surfaces with forest fraction. Forest fractions decrease the accuracy of soil moisture retrieval because:

- they are generally located in mountainous regions for the SNOTEL sites (increasing the topographic index and reducing the quality of the retrievals).

- the presence of wet snow.

- the presence of mulch makes it difficult to determine the depth to which soil moisture is associated[44].

- forest covers can be sparse and non-uniform[45].

Fig. 9(a) shows the retrieved soil moisture over node near SCAN site 2002. This site is located at 30km North of Minneapolis. Forest cover in this area is very sparse and represents 12% of SMOS fraction cover. The retrieved soil moisture underestimates the soil moisture of 5%. Many rain events are registered by the site observations that are not captured by the satellite observation.Fig. 9(b) shows the retrieved soil moisture over SNOTEL site 774. The site is located in Idaho at an altitude of 2034m over a mountainous region with mild-topography. The forest cover is mainly evergreen coniferous trees (Douglas pin and Pinyon-Juniper)sparsely distributed over 10% of the surface. The retrieval results from SMOS follow the seasonal variation with a rmse of 0.069 m^3/m^3. The site data are not available for a big part of the year. The third site (393) is in Utah in the Ashley national forest at 1800 m.The forest is mainly evergreen coniferous treesrepresenting 40% of the radiometric surface fraction. Soil moisture retrievals are only activated from June to November due to the presence of snow (Fig. 9(c)). The soil moisture fits well to the observation dynamics (bias=0.001, r=0.61 and rmse = 0.05).Fig. 9(d) shows retrievals for SCAN site 2024in Talahatchi county, Mississippi at 83m altitude. The site has a low slopewith heterogeneous sparse distribution of forest cover.An underestimation of $0.08m^3/m^3$is observed but the dynamics are well represented. Fig. 9(e) shows soil moisture retrievals for SCAN site 2053. The forest areas are characterized by oak pine on a mild-slope region.A bias of $0.11m^3/m^3$ is obtained on this site. Site 2079in Fig. 9(f) is located at Mammoth cave national park in the Kentucky. It is characterized by oak hickory in a region highly modified by anthropicactivities. The results show a correlation of 0.7 and a bias of 0.062 m^3/m^3.

The results show a better agreement with the SNOTEL sites concentrated on Western US than the SCAN sites located mainly in central and eastern US,independent of forest type and topography. The results suggest that the SCAN network calibration gives an overestimation of

soil moisture. From the retrieval point of view, a possible explanation could be that the retrieved soil moisture over nominal surfaces will decrease with the inversed surface fraction to compensate the ECMWF wetsoil moisture used over non inversed surface fractions. As explained earlier, the soil moisture retrieval will consider the dominant surface fraction for the retrieval and the forest fraction as a default contribution. The contribution over default fractions are computed based on the ECMWF soil moisture thatoverestimates systematically the soil moisture. This applies to sites where forest biomass density is low enough to see through forests. This explanation is less probable because no increase in the underestimation is noticed as the forest cover fraction increases.

Effect of water fraction

SMOS retrieval is highly impacted by the water fraction. TBover water fractions is close to 150 Kin L-Band which is much lower than the average TB on ground. So a small fraction of water can have a large impact on the retrieval in Level 2. When large water bodies are present the reconstructed TBs can be highly impacted at Level1 processing by Gibbs effect. In fact the high contrast occurring at the coast creates a discontinuity that produces an overshoot in the Fourier series known as the Gibbs phenomena[46]. Water fraction determination can be also erroneous because of scale effects and temporal change of the water fraction by evaporation, flooding, and tides. Two sites have been selected to show the effect of water bodies on the retrievals. The first site is located in Miami-Dade near Florida. Fig. 10(a) shows the retrieval results over node 5023541 at 19km from this site.Soil texture at 10 cm near the station is 31.2% sand and 35.3% clay. Surface distribution is 40% bare soil or low vegetation, 1.1% forest, 43.3% open fresh water, 8.8 % of sea water and 7% of urban. All surrounding nodes to the site except the considered node have no soil moisture retrievalsbecause the surface is classified as heterogeneous and no soil moisture retrieval is associated with this class.The soil moisture retrievals underestimate the reference soil moisture with a bias of 0.064 m^3/m^3. SMOS soil

moisture exhibit less dynamicscompared to the site data but seasonal dynamics are respected.
The second site (SNOTEL 332) is located in Utah at 40km east of Salt Lake in the Wasatch
Mountains. The impact of water bodies is shown here by taking advantage of the oversampling
of the grid. As shown in section II-A the correlation coefficient for two adjacent nodes is 0.746
for homogeneous surfaces. In Fig. 10we show the soil moisture timeseries for nodes176383
located at 10km from the site 332 on the eastern side.The node has the following surface
fractions: 74.7% bare soil and low vegetation,5.9% forest, 17.03% water and the comparison
with site data gives a correlationof 0.66 and rmse of 0.086. The second one is node177407.The
surface fractions for this node are48.3% bare soil and low vegetation,4.7% forest,42.7% water.
Notice the increase in water fraction of about 25.67 % between the two nodes. On this node the
retrievals are noisy (r = 0.27)with an overestimation of soil moisture (bias). Thisover estimation
can be due to a wrong estimation of the open water surface, a weak sensitivity to soil moistureas
water fraction increases, or a weak quality of the reconstruction due to the Gibbs effect. The fact
that the water is salty increasesthe over estimation. A high concentration of salt would increase
the dielectric constant, reduce the emissivity andincrease soil moisture.

B. Aggregated soil moisture

Over some areas, several SCAN and SNOTEL stations are available inside SMOS antenna
footprint.In this case the aggregation method proposed in Equation (2) is used to obtain average
site soil moisture. Fig. 8 shows the aggregated soil moisture over node 241087 using 5 SCAN
sites (2053, 2057, 2059, 2075, 2076 and 2078) located at less than 20 km from the node center.
The SMOS observation presents a very good match with the observations in the dryer site but
not with the aggregated soil moisture. The wetter sites are increasing the aggregated mean
artificially as there is no reason to have minimum soil moisture of 0.11 m^3/m^3over the year. It
seems that a recalibration of station data may be needed. In the present status, the comparison of
node to site is recommended compared to aggregated soil moisture.

C. Global analysis

In this section the global comparison over the SCAN and SNOTEL sites is presented. From the list of original sites available, 238 sites have soil moisture observations at 5cm with associated soil moisture retrievals from SMOS, and 98of those are associated with a 90% nominal surface fraction. Table4shows the number of records used for each dataset.Fig. 11(a-b) shows the histogram of the correlation coefficient for the 238 sites with the nearest node. The average correlation coefficient is enhanced by the reprocessing campaign. The two datasets have sites with negative correlation coefficients. The negative correlation can be explained mainly by the lake of spatial sampling in the direct site to node comparison.. The number of records for the REPv4 version is 13% lower than the DPGS version. REPv4 has fewer retrievals but of better quality since only dual polarization are used in the retrieval. In fact mixed polarizations are not used in version 4, so the total number of TBs is reduced and the retrieval is less attempted but with better quality. Fig. 11 (c-h) shows the results correlation, bias and RMSE for the 98 sites with 90% of nominal fraction.The mean bias is about 0.08 m^3/m^3. As mentioned earlier it is unlikely that the spatial representative in the direct site to node comparison be the reason for this bias. One reason could be the use of inadequate ancillary data in the SMOS retrieval algorithms. The ancillary data impacting the retrieval could be the static soil properties or the dynamic climate data from ECMWF. Another reason could be the presence of a systematic overestimation of TB due to emissions in the protect L-band that are affecting the signal at global scale. In fact SMOS reconstruction is based on interferometry between its radiometers that are each observing a large part of the globe so RFI sources could eventually have large impacts. Verifying this hypothesis needs to go back to the SMOS reconstruction algorithm. The most probable reason is the sensing depth. If this is overestimated than reducing it will give fast dynamics in the soil moisture and dryer soil moisture values. Even though the two results go in the wright direction the answer could be more complex as

sensing depth is also dependent on the soil properties and soil moisture itself. So sensing depth can be dynamic. The way soil properties upscale at 40km for soil moisture observation at L-band observations is also an open question. Finally the in-situ soil moisture measurements can worsen the SMOS observed bias because some sites show high soil moisture values all through the year.

V. CONCLUSIONS

Soil moisture is a crucialvariable for a variety of hydrological applications. Continuous effortsare needed to validate satellites based microwave derived soil moisture products. With the launch of SMOS new soil moisture products in L-Band microwave became available. This paper gives a first insighton the validation of SMOS data over continental US. Time series from SCAN/SNOTEL datasets are compared to SMOS soil moisture retrievals.The results show that at over a variety of sites SMOS retrievals are of good quality but globally SMOS observations show a negative bias. Results also suggest that if a better spatial representation is taken into account, SMOS should be able to meet the mission objective of 0.04 m^3/m^3 rmse over nominal surfaces. Some direct node - site comparison results have a rmse lower than 0.04 m^3/m^3 while representing the dynamics of soil moisture. Aradiometric aggregation approach is used to take the spatial heterogeneity of soil moisture into account. The results show that the aggregation doesnot always enhance the comparison. This emphasises the need for a pre-processing of the SCAN and SNOTEL data based on the previous works of [27]. The effect of water fraction is presented through two sites. Retrieval over nominal surfaces with a forested cover is also inspected through some site examples. One way to accomplish that on semi-arid sites is by using physically based disaggregation like the DisPATchmodel (Disaggregation based on Physical And Theoretical scale Change)[47]. A combination of the DisPATch model with an

assimilation scheme would enable us to take into account the temporal dynamic in the validation.

ACKNOWLEDGMENT

REFERENCES

[1] H. DouvilleandF. Chauvin,"Relevance of soil moisture for seasonal climate predictions : a preliminary study",*Climate Dynamics*, 16:719–736. 13, 2000.

[2] R. D. Koster, P. A. Dirmeyer, Z. C. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C. H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, O. Taikan, K. Oleson, A. Pitman, Y. C. Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. K. Xue, and T. Yamada, "Regions of strong coupling between soil moisture and precipitation", *Science*, vol. 305 (5687), pp. 1138–1140, DOI: 10.1126/science.1100217, 2004.

[3] M. Drusch, "Initializing numerical weather prediction models with satellitederived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set". *Journal of Geophysical*, 2007

[4] D. W. Shin, J. G.Bellow, T. E.LaRow, S.Cocke, and J. J., O'Brien, "The Role of an Advanced Land Model in Seasonal Dynamical Downscaling for Crop Model Application", *Journal of Applied Meteorology and Climatology*, 45(5):686–701. 14, 2006

[5] E.G.Njoku, "AMSR-E/Aqua daily L3 surface soil moisture, interpretive parameters, and QC EASE-grids", *Boulder, CO: National Snow Ice Data Center*, 2008, July 2002 to December 2007, daily updated digital media.

[6] T.J.Jackson, "Measuring surface soil moisture using passive microwave remote sensing", *Hydrological Processes,* 7 (2), 1993.

[7] M.Owe, R.A.M.de Jeu, and Holmes, T.H.R., "Multisensor historical climatology of satellite-derived global land surface soil moisture", *Journal of Geophysical Research*, 113 (F1), 2008.

[8] R.Scofield, J.G.LaDue, N.Grody and R.Ferraro, "A soil wetness index for monitoring the great flood of 1993", *Proceedings of the 7th Conference on Satellite Meteorology and Oceanography*, 1994.

[9] Naeimi V., K. Scipal, Z. Bartalis, S. Hasenauer and W. Wagner (2009), An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 47, pp. 1999-2013., DOI: 10.1109/TGRS.2008.2011617.

[10] Wagner,W., G. Lemoine, and H. Rott, "A method for estimating soil moisture from ERS scatterometer and soil data", *Remote Sens. Environ.*, 70, 191– 207, 1999.

[11] W.T.Crow, D.G.Miralles and M.H.Cosh, "A quasi-global evaluation system for satellite-based surface soil moisture retrievals", *IEEE Transactions on Geoscience and Remote Sensing*, 48 (6), 2010.

[12] Y. H.Kerr, "Soil moisture from space: where are we?",*Hydrogeology Journal*, 15, 117−120,2007.

[13] W.Wagner, G.Bloschl, P.Pampaloni, J. -C.Calvet, B.Bizzari, J. -P.Wigneron, andY.Kerr, "Operational readiness of microwave remote sensing of soil moisture forhydrologic applications",*Nordic hydrology*, 38(1), 1−20.

[14] Y. H.Kerr, P.Waldteufel, , J. -P.Wigneron, , J. -M.Martinuzzi, , Font, J., & Berger, M.,"Soil moisture retrieval from space: the soil moisture and ocean salinity (SMOS) mission",*IEEE Transactions on Geoscience and Remote Sensing*, 39, 1729−1735, 2001.

[15] Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M. J., Font, J., Reul, N., Gruhier, C., Juglea S. E., Drinkwater, M. R., Hahne, A., Martin-Neira, M. and Mecklenburg, S, "The SMOS mission: new tool for monitoring key elements of the global water cycle",*Proc. IEEE 98 (5)*, 666-687, doi:10.1109/JPROC.2010.2043032, 2010.

[16] McMullan, K. D., Brown, M. A., Martin-Neira, M., Rits, W., Ekholm, S., Marti, J. and Lemanczyk J., "SMOS: the payload", *IEEE Transactions on Geoscience and Remote Sensing*, vol 46, n°3, March 2008.

[17] Vinnikov, Konstantin Y., Alan Robock, Shuang Qiu, Jared K. Entin, Manfred Owe, Bhaskar J. Choudhury, Steven E. Hollinger and Eni G. Njoku, "Satellite remote sensing of soil moisture in Illinois, USA", *J. Geophys. Res.*, 104, 4145-4168, 1999.

[18] Vachaud, G., Passerat De Silane, A., Balabanis, P., Vauclin, M., "Temporal stability of spatially mesured soil water probability density function",*Soil Sci. Soc. Am. J.*, 49, 822–827, 1985.

[19] McCabe, M.F., Wood, E.F. and Gao, H., "Initial soil moisture retrievals from AMSR-E: multiscale comparison using in situ data and rainfall patterns over Iowa", *Geophysical Reasearch Letters*, 32, 2005.

[20] Brocca, L., Melone, F., Moramarco, T. and Morbidelli, R., "Spatial-temporal variability of soil moisture and its estimation across scales", *Water Resources Research*, 46, 2010.

[21] E.G.Njoku, T.J.Jackson, V.Lakshimi, Y.K.Chan and S.V.Nghiem, "Soil moisture retrieval from AMSR-E", *IEEE Transactions on Geoscience and Remote Sensing*, 41 (2), 2003.

[22] A.K.Sahoo, P.R.Houser, C.Ferguson, E.F.Wood, P.A.Dirmeyer, and M.Kafatos, "Evaluation of AMSR-E soil moisture results using the in-situ data over the Little River Experimental Watershed, Georgia", *Remote Sensing of Environment*, 112, 2008.

[23] T.J.Jackson, M.H.Cosh, R.Bindlish, P.J.Starks, D.D.Bosch, M.Seyfried, D.C.Goodrich, M.S.Moran and J.Du, "Validation of Advanced Microwave Scanning Radiometer soil moisture products", *IEEE Transactions on Geoscience and Remote Sensing*, 48(12), 2010.

[24] T. J. Jackson, M. H. Cosh, R. Bindlish, P. J. Starks, D. D. Bosch, M. Seyfried, D. C. Goodrich, M. S. Moran, and J. Du, "Validation of Advanced Microwave Scanning Radiometer soil moisture products," IEEE Trans. Geosci. Remote Sens., vol. 48, no. 12, pp. 4256–4272, 2010, doi:10.1109/TGRS.2010.2051035.[25] C.S.Draper, J.P.Walker, P.J.Steinle, R.A.M.de Jeu and T.R.H.Holmes, "An evaluation of AMSR-E derived soil moisture over Australia", *Remote Sensing of Environment*, 113, 2009.

[26] C. Gruhier, P. de Rosnay, S. Hasenauer, T. Holmes, R. de Jeu, Y. Kerr, E. Mougin, E. Njoku, F. Timouk, W. Wagner, and M. Zribi, "Soil moisture active and passive microwave products: intercomparison and evaluation over a Sahelian site", *Hydrology and Earth System Sciences*, 14, 141-156, 2010.

[27]De Lannoy, G.J.M., P.R. Houser, N.E.C. Verhoest, V.R.N. Pauwels, and T.J. Gish, "Upscaling of point soil moisture measurements to field averages at the OPE3 test site", *Journal of Hydrology*, 343(1-2), 1-11, 2007.

[28] C. Gruhier, P. de Rosnay, Y. Kerr, E. Mougin, E. Ceschia, J.-C. Calvet, and P. Richaume, "Evaluation of AMSR-E soil moisture product based on ground measurements over a temperate and semi-arid regions", *Geophys. Res. Lett.*, 35, L10405, doi:10.1029/2008GL0333330, 2008.

[29] R.Bindlish, T.J.Jackson, A.J.Gasiewski, M.Klein andE.G. Njoku, "Soil moisture mapping and AMSR-E validation using PSR in SMEX02", *Remote Sensing of Environment*, 103, 2006.

[30] S.Juglea, Y.Kerr, A. Mialon, J.-P.Wigneron, E. Lopez-Baeza, A.Cano, A. Albitar, C.Millan-Scheiding, M.Carmen Antolin, and S.Delwart, "Modelling soil moisture at SMOS scale by use of a SVAT model over the Valencia Anchor Station", *Hydrol. Earth Syst. Sci. Discuss.*, 649–686, 2010, http://www.hydrol-earth-syst-sci-discuss.net/7/649/2010/. 1147, 1149, 1154

[31] O. Merlin, J. P. Walker, A. Chehbouni, and Y. Kerr, "Towards deterministic downscaling of SMOS soil moisture using MODIS derivedsoil evaporative efficiency," *Remote Sens. Environ.*, vol. 112, pp. 3935–3946, doi:10.1016/j.rse.2008.06.012, 2008.

[32] O. Merlin, A. Al Bitar, J. P. Walker, and Y. Kerr, "An improved algorithm for disaggregating microwave-derived soil moisture based on red, near-infrared and thermal-infrared data," *Remote Sens. Environ.*, vol. 114, no. 10, pp. 2305–2316, 2010,doi:10.1016/j.rse.2010.05.007.

[33] M. Piles, A. Camps, M. Vall-llossera, I. Corbella, R. Panciera, C. Rüdiger, Y. Kerr and Jeff Walker, *"*Downscaling SMOS-derived soil moisture using MODIS visible/infrared data", *IEEE Trans. on Geosci. and Remote Sens., (in press)*

[34] Schaefer, Garry L., Michael H. Cosh, Thomas J. Jackson, "The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN)",*J. Atmos. Oceanic Technol.*, 24, 2073–2077,2007.

[35] Y. H. Kerr, P. Waldteufel, P. Richaume, P. Ferrazzoli, and J.-P. Wigneron, "SMOS level 2 processor soil moisture Algorithm TheoreticalBasis Document (ATBD)",Toulouse, France: CESBIO, vol. SO-TN-ESL-SM-GS-0001, V3.a, 2008.

[36] Y. H.Kerr, P.Waldteufel, P.Richaume, A.Mahmoodi, J.-P.Wigneron, P.Ferrazzoli, A.Al Bitar, F.Cabot, D.Leroux, A.Mialon, S.Delwart, "The SMOS soil moisture retrieval algorithm", *IEEE Transactions on Geoscience and Remote Sensing,* (submitted)

[37] De Lannoy, G. J. M.; Ufford, J.; Sahoo, A. K.; Dirmeyer, P. & Houser, P. R., "Observed and simulated water and energy budget components at SCAN sites in the lower Mississippi Basin", *Hydrological Processes*25(4), 634—649,2011.

[38] Wigneron, J.-P., Kerr, Y.H., Waldteufel, P., Saleh, H., Escorihuela, M., Richaume, P., Ferrazzoli, P., De Rosnay, P., Gurney, R., Calvet, J.-C., Grant, J., Guglielmetti, M., Hornbuckle, B., Mätzler, C., Pellarin, T. and Schwank M., "L-band Microwave Emission of the Biosphere (L-MEB) model: description and calibration against experimental data sets over crop fields", *Remote Sensing of Environment*, 107, 2007.

[39] Mätzler C., "Thermal microwave radiation: applications for remote sensing", *IET Electromagnetic Waves Series*, 2006.

[40] Schwank M., M. Stähli, H. Wydler, J. Leuenberger, C. Mätzler and F. Hannes, "Microwave L-Band Emission of Freezing Soil", *IEEE Transactions on Geoscience and Remote Sensing*, 42(6): 1252-1261, 2004.

[41] Al BitarA., JacquetteE., KerrY., MialonA., CabotF., QuesneyA., MerlinO., and RichaumeP., "Event detection of hydrological processes with passive L-band data from SMOS", *Proc. SPIE 7824*, 78240J 2010, doi:10.1117/12.865074.

[42] Escorihuela M.J., Chanzy A., Wigneron J-P. and Kerr Y. ,"Effective soil moisture sampling depth of L-band radiometry: a case study" *Remote Sensing of Environment*, 114, 995-1001, doi: 10.1016/j.rse.2009.12.011, 2010.

[43] MillerC. J., YesillerN., YaldoK., andMerayyanS., "Impact of Soil Type and Compaction Conditions on Soil Water Characteristic", J. Geotech. Ge*oenviron. Eng.* 128, 733, doi:10.1061/(ASCE)1090-0241(2002)128:9(733), 2002.

[44] Saleh, K., J.-P. Wigneron, P. de Rosnay, J.-C. Calvet, M.-J. Escorihuela, Y. Kerr, P. Waldteufel, "Impact of rain interception by vegetation and mulch on the L-band emission of natural grass", *Remote Sens. Env.*, 2006, 101: 127-139.

[45] Grant, J.P., Van de Griend, A.,A., Wigneron, J,-P., Saleh, K., Panciera, R., and Walker, J. P.," Influence of Forest Cover Fraction on L-band Soil Moisture Retrievals from Heterogeneous Pixels using Multi-Angular Observations", *Remote Sensing of Environment*, 114: 1026-1037 doi:10.1016/j.rse.2009.12.016, 2010.

[46] Gibbs, J. W., "Fourier's Series". Nature 59, 200 (1898) and 606 (1899).

[47] O. Merlin, C. Rudiger, A. Al Bitar, P. Richaume, J. P. Walker,and Y. H. Kerr, "Disaggregation of SMOS soil moisture overthe AACES area with DisPATCh",*IEEE Transactions on Geoscience and Remote Sensing*, (*submitted)*

Table 1

Description of selected DGG nodes (FNO = nomial fraction, FFO = forest fraction, FWO = water fraction)

Node	Site	Fractions			
		FNO	FFO	FWO	Other
165158	SNOTEL 581	100	0	0	0
172276	SCAN 2160	100	0	0	0
186675	SCAN 2018	100	0	0	0
203609	SCAN 2093	98	0	0	2
203626	SCAN 2001	98	0	0	2
218480	SCAN 2092	97	1	0	2
219434	SCAN 2168	97	1	0	2
235420	SCAN 2030	93	2	3	2
237475	SCAN 2084	95	4	0	1
187758	SCAN 2002	81	12	4	3
163040	SNOTEL 774	90	10	0	0
180486	SNOTEL 393	60	40	0	0
240043	SCAN 2024	65	31	2	1
241088	SCAN 2053	53	43	2	2
231355	SCAN 2079	50	49	0	1
5023541	SCAN 2051	40	1	42	7
176383	SNOTEL 332	75	6	17	2
177407	SNOTEL 332	48	5	43	4

Table 2
Description of representative SCAN/SNOTEL sites from[34]

Site	Stat	Cove type	Texture	
			Sand (%)	Clay (%)
SNOTEL 581	MT	Crops	63.0	13.0
SCAN 2160	UT	Grassland (mountain)	-	-
SCAN 2018	WY	Grassland	80.3	5.5
SCAN 2093	KS	Crops	5.8	22.4
SCAN 2001	NE	Crops	7.2	35.5
SCAN 2092	KS	Crops	72.4	7.5
SCAN 2168	NM	Grassland (mountain)	-	-
SCAN 2030	AR	Crops	9.7	12.4
SCAN 2084	MS	Crops	6.1	15.6
SCAN 2002	MN	Grassland -	74.4	9.7
SNOTEL 774	ID	Douglas pin and Pinyon-Juniper	-	-
SNOTEl 393	UT	Forest - evergreen coniferous trees	-	-
SCAN 2024	MS	Grassland	2.2	16.1
SCAN 2053	AR	Crops - oak pine	7.0	31.5
SCAN 2079	KY	Grassland -	40.5	9.7
SCAN 2051	FL	Wetlands	35.3	31.2
SNOTEL 332	UT	Forest- evergreen coniferous trees	-	-

Table 3
Statistical skills of presented SCAN/SNOTEL sites

Node	Site	Stats		
		Bias	R	RMSE
165158	SNOTEL 581	-0.019	0.71	0.053
172276	SCAN 2160	0.032	0.57	0.054
186675	SCAN 2018	0.023	0.68	0.056
203609	SCAN 2093	-0.035	0.77	0.063
203626	SCAN 2001	-0.038	0.71	0.07
218480	SCAN 2092	0.034	0.62	0.073
219434	SCAN 2168	0.007	0.61	0.032
235420	SCAN 2030	0.026	0.82	0.067
237475	SCAN 2084	-0.08	0.7	0.109
187758	SCAN 2002	-0.049	0.42	0.082
163040	SNOTEL 774	0.022	0.56	0.069
180486	SNOTEL 393	-0.001	0.61	0.05
240043	SCAN 2024	-0.087	0.65	0.119
241088	SCAN 2053	-0.115	0.62	0.128
231355	SCAN 2079	0.062	0.7	0.091
5023541	SCAN 2051	-0.064	0.35	0.092
176383	SNOTEL 332	0.021	0.66	0.086
177407	SNOTEL 332	0.08	0.28	0.13

Table4
Number of SCAN and SNOTEL sites and soil moisture records used in the global comparison

	Nb sites	Nb. of records	
		DPGS	REPv4
ALL	235	17348	16420
Nominal	98	8500	8317
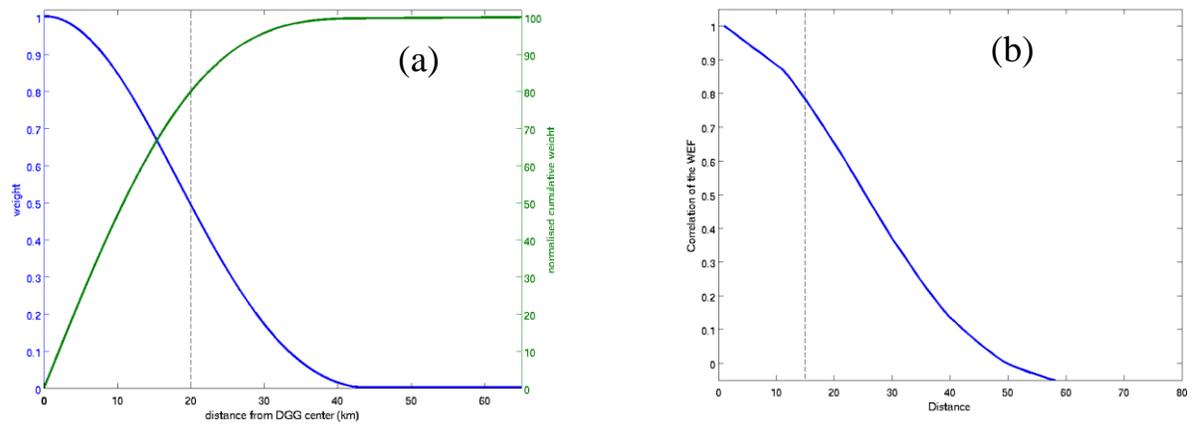
Fig. 1 Mean weighting function of SMOS footprint (a), spatial-correlation function of the mean weighting function (b)
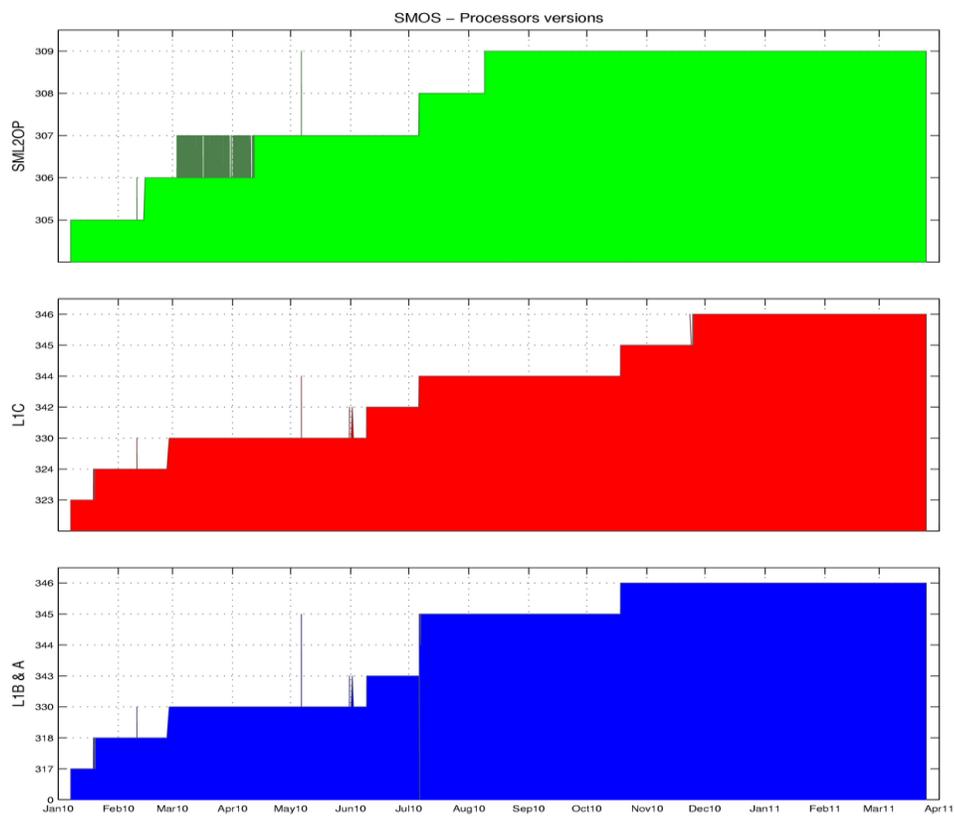
Fig. 2 Level 1 and Level 2 soil moisture operational processors (L1A,B,C and L2) versions for the SMOS DPGS soil moisture dataset.
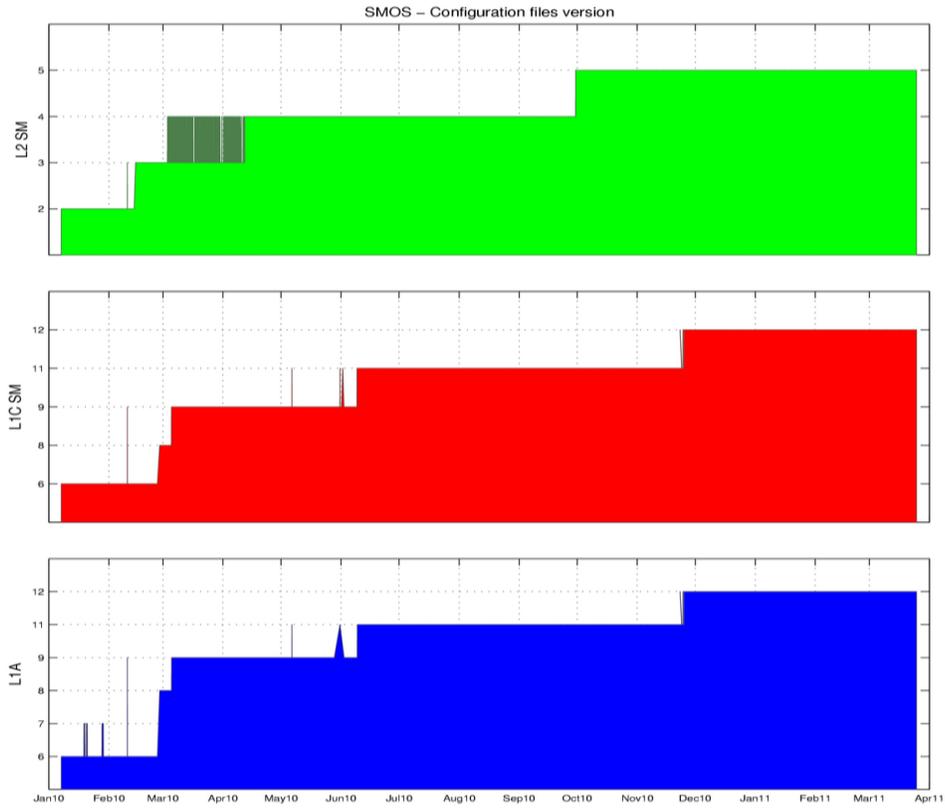
Fig. 3 Level 1 and Level 2 soil moisture operational processors (L1A,B,C and L2) configuration files versions for the SMOS DPGS soil moisture dataset
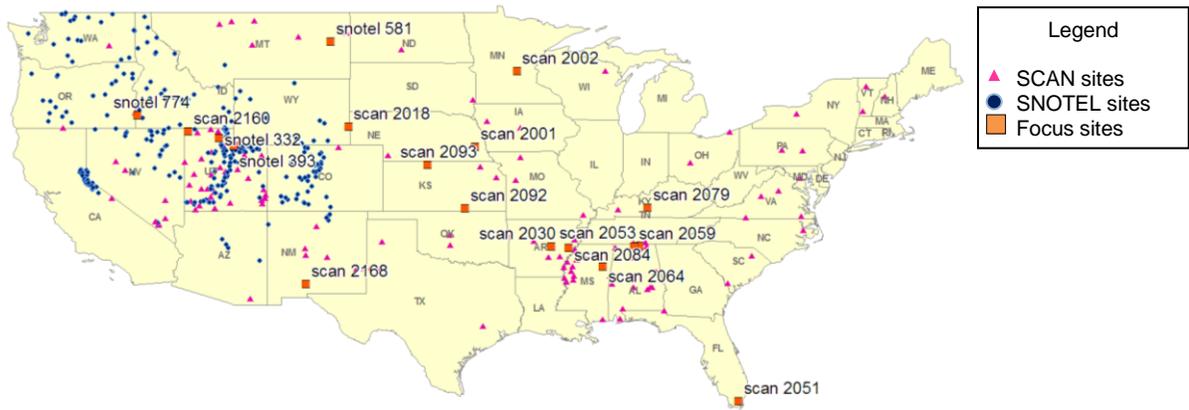

Fig. 4 Location of selected SCAN/SNOTEL sites
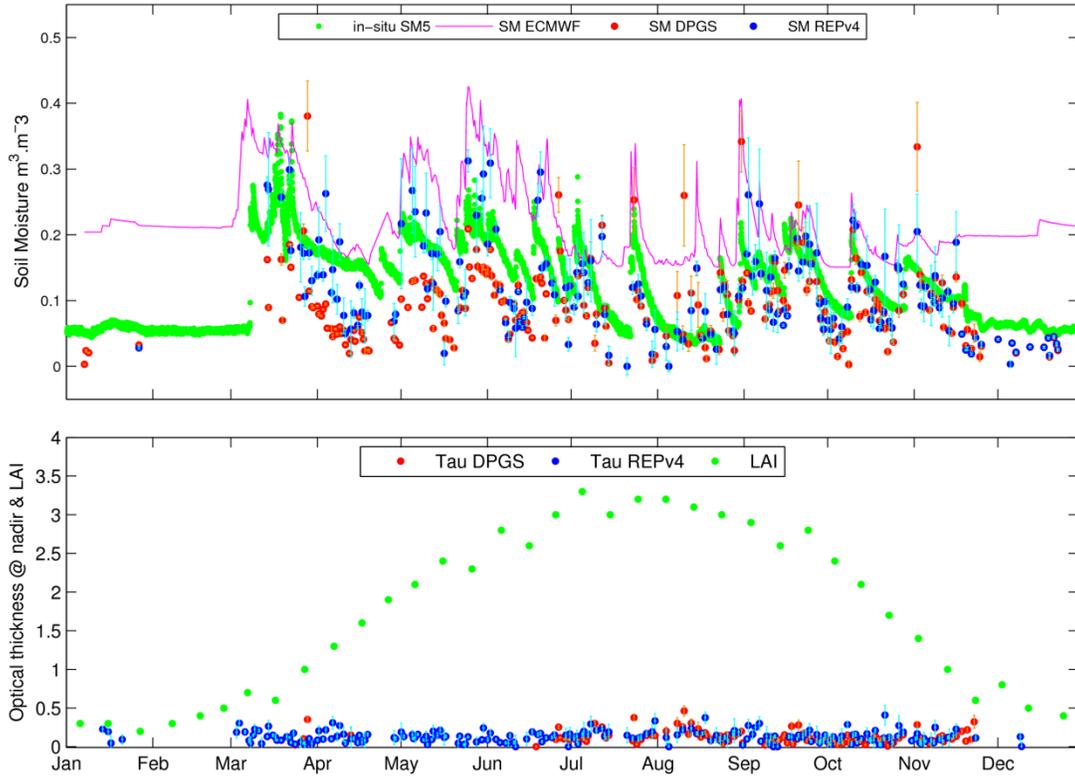
Fig. 5 Comparison of SMOS retrievals vs SNOTEL site 581(upper) soil moisture time resies from SMOS ECMWF and in-situ data,,(lower) figure the Leaf area index from Modis and optical thickness at Nadi from SMOS
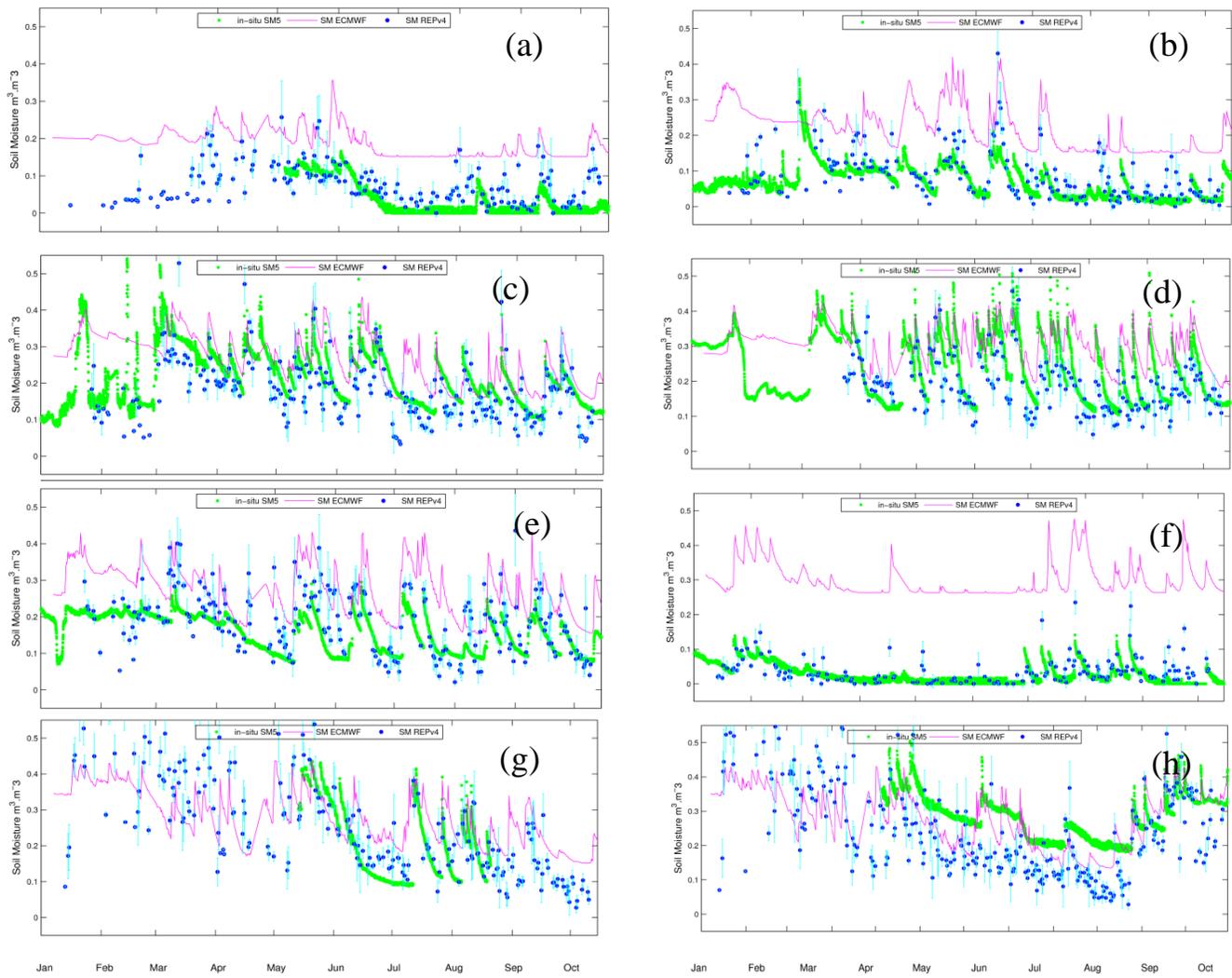
Fig. 6Time series plots of soil moisture retrievals from SMOS vs SCAN site 2160 (a), 2018 (b), 2093 (c), 2001 (d), 2092 (e), 2168 (f), 2030 (g), 2084 (h).
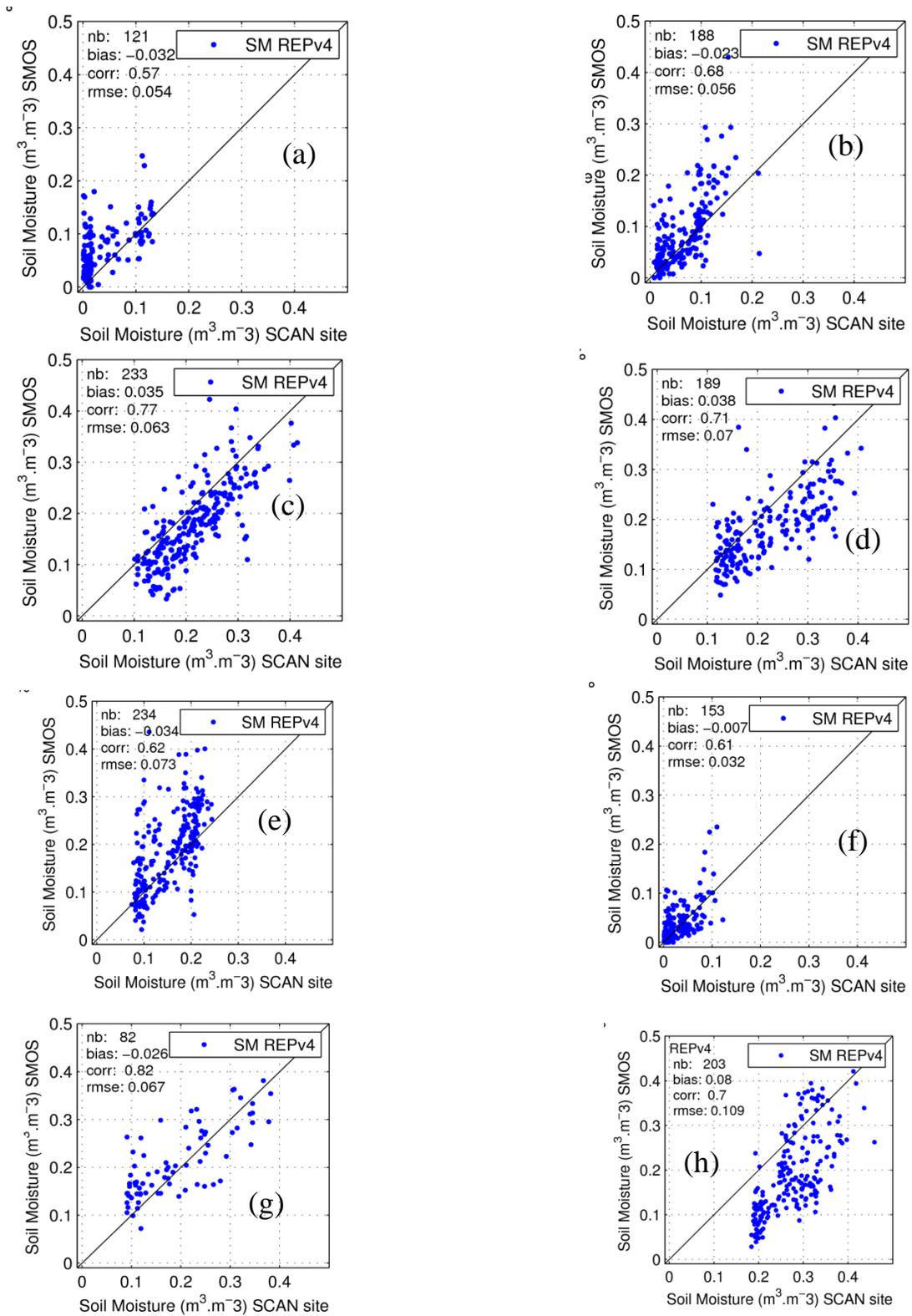
Fig. 7Scatter series plots of soil moisture retrievals from SMOS vs SCAN site 2160 (a), 2018 (b), 2093 (c), 2001 (d), 2092 (e), 2168 (f), 2030 (g), 2084 (h).
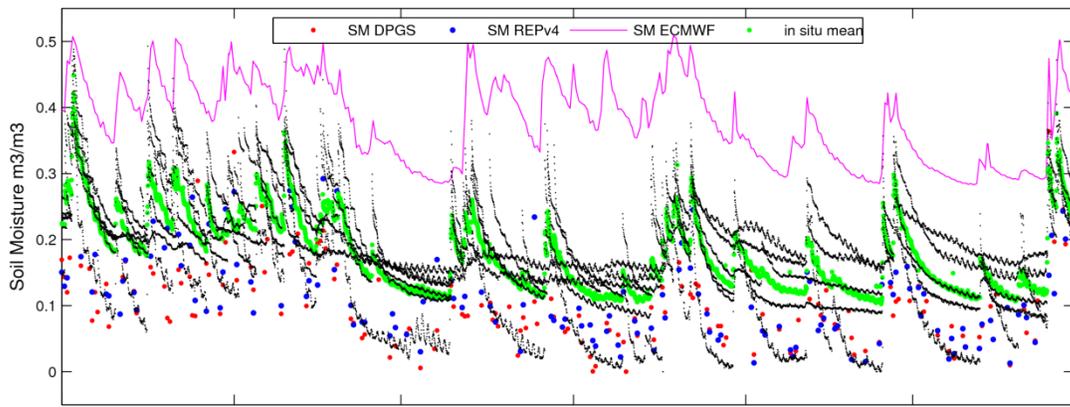
Fig. 8Soil moisture time seriesover node 241087 in the lower Mississippi. In-situ Average using 6 SCAN sites (green), SMOS retrieval (red for DPGS, Blue for REPv4), SMOS-ECMWF product.
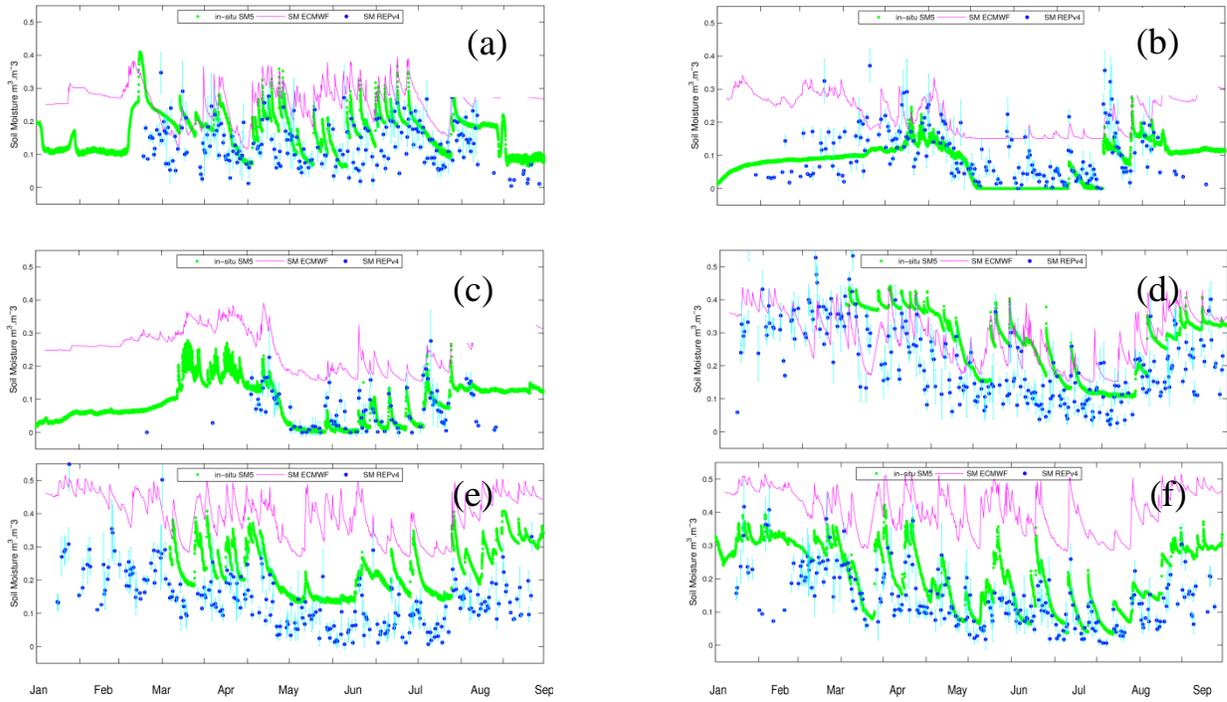
Fig. 9time series plots of soil moisture retrievals from SMOS vs site 2002 (a), 774 (b), 393 (c), 2024 (d), 2053 (e), 2079 (f).
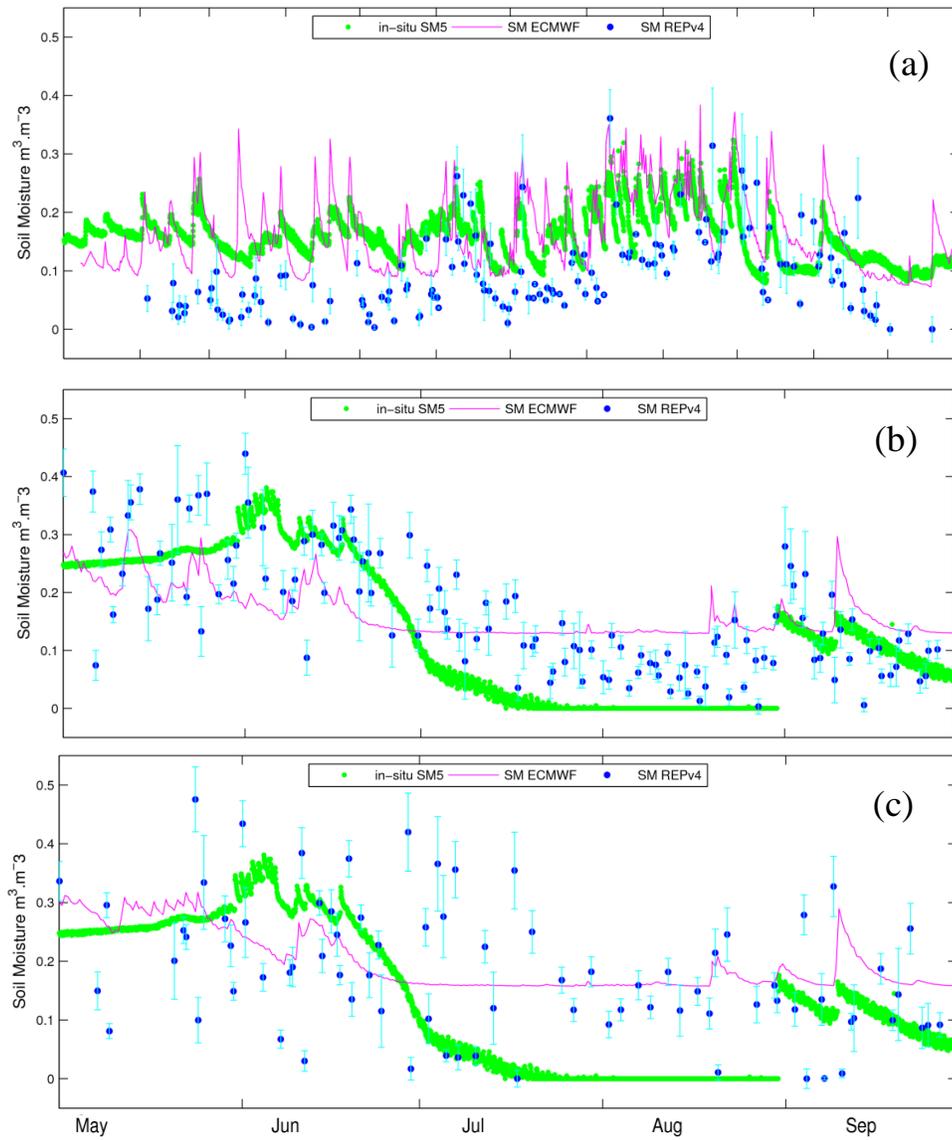
Fig. 10time series plots of soil moisture retrievals from SMOS vs SCAN site 2051 (a), 332 (b), 332 (c).
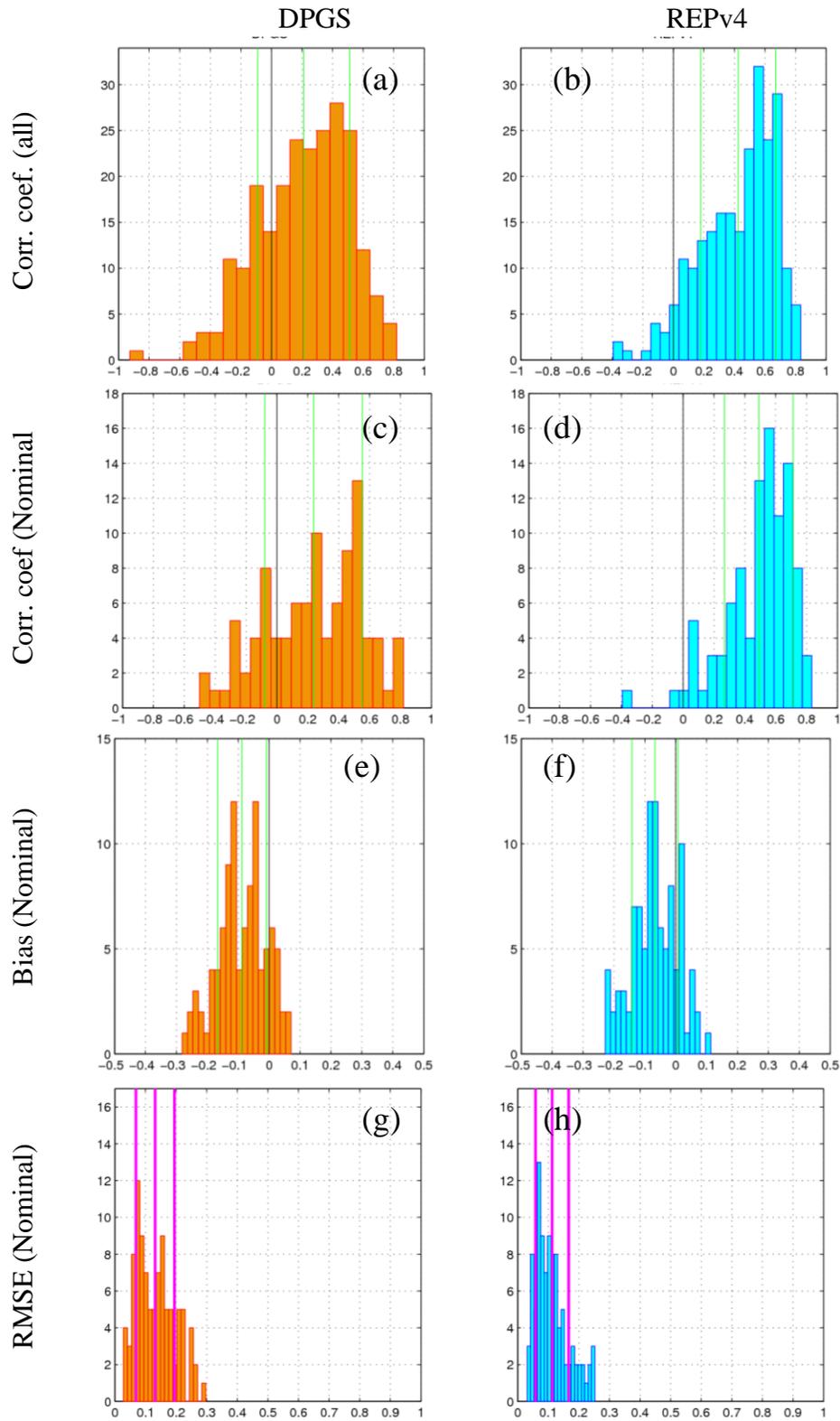
Fig. 11 Histograms of global stats (a-b) all available sites, (c-h) sites associated with nominal nodes