Remote Sensing of Environment 233 (2019) 111380

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product

Narendra N. Das[a,*], Dara Entekhabi[b], R. Scott Dunbar[a], Mario J. Chaubell[a], Andreas Colliander[a], Simon Yueh[a], Thomas Jagdhuber[c], Fan Chen[d], Wade Crow[e], Peggy E. O'Neill[f], Jeffrey P. Walker[g], Aaron Berg[h], David D. Bosch[i], Todd Caldwell[j], Michael H. Cosh[k], Chandra H. Collins[l], Ernesto Lopez-Baeza[m], Marc Thibeault[n]

[a] *Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 91109, USA*
[b] *Massachusetts Institute of Technology, Cambridge, MA, 02139, USA*
[c] *German Aerospace Center (DLR), DLR Microwaves and Radar Institute, Oberpfaffenhofen, 82234 Weßling, Germany*
[d] *Science System and Application, Inc, United State Department of Agriculture, Beltsville, MD, USA*
[e] *United State Department of Agriculture, Beltsville, MD, USA*
[f] *Goddard Space Flight Center (GSFC), NASA, 8800 Greenbelt, MD, 20771, USA*
[g] *Monash University, Australia*
[h] *University of Guelph, Canada*
[i] *USDA ARS Southeast Watershed Research, Tucson, AZ, USA*
[j] *University of Texas at Austin, TX, USA*
[k] *USDA ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD, USA*
[l] *USDA ARS Southwest Watershed Research, Tucson, AZ, USA*
[m] *University of Valencia, Spain*
[n] *Comisión Nacional de Actividades Espaciales (CONAE), Argentina*

ARTICLE INFO

Edited by Jing M. Chen

Keywords:
Soil moisture active passive
SMAP
Sentinel
Soil moisture
High resolution
Active-passive
SAR
Radar
Radiometer
NASA

ABSTRACT

Soil Moisture Active Passive (SMAP) mission of NASA was launched in January 2015. Currently, SMAP has an L-band radiometer and a defunct L-band radar with a rotating 6-m mesh reflector antenna. On July 7th, 2015, the SMAP radar malfunctioned and became inoperable. Consequently, the production of high-resolution active-passive soil moisture product got hampered, and only ~2.5 months (April 15th, 2015 to July 7th, 2015) of data remain available. Therefore, during the SMAP post-radar phase, many ways were examined to restart the high-resolution soil moisture product generation of the SMAP mission. One of the feasible approaches was to substitute the SMAP radar with other available SAR data. Sentinel-1A/Sentinel-1B SAR data was found most suitable for combining with the SMAP radiometer data because of its nearly similar orbit configuration that allows overlapping of their swaths with a minimal time difference, a key feature/requirement for the SMAP active-passive algorithm. The Sentinel interferometric wide swath (IW) mode acquisition also provides the co-polarized and cross-polarized observations required for the SMAP active-passive algorithm. However, some differences do exist between the SMAP and Sentinel SAR data. They are mainly: 1) Sentinel has a C-band SAR whereas SMAP operates at L-band; 2) Sentinel has multiple incidence angles within its swath, and SMAP has one single incidence angle; and 3) Sentinel 1A/B Interferometric Wide (IW) swath width is ~250 km as compared to SMAP with 1000 km swath width. On any given day, the narrow swath width of the Sentinel observations significantly reduces the overlap spatial coverage between SMAP and Sentinel as compared to the original SMAP radar and radiometer swath coverage. Hence, the temporal resolution (revisit interval) suffers due to narrow overlapped swath width and degrades from 3 days to 12 days. One advantage of using very high-resolution resolution Sentinel-1A/Sentinel-1B data in the SMAP active-passive algorithm is the potential of obtaining the disaggregated brightness temperature and thus soil moisture at a much finer spatial resolution of 3 km and 1 km at

* Corresponding author.
E-mail addresses: Nndas@jpl.nasa.gov (N.N. Das), Darae@mit.edu (D. Entekhabi), Roy.S.Dunbar@jpl.nasa.gov (R.S. Dunbar), Mario.J.Chaubell@jpl.nasa.gov (M.J. Chaubell), Andreas.Colliander@jpl.nasa.gov (A. Colliander), Simon.Yueh@jpl.nasa.gov (S. Yueh), Thomas.Jagdhuber@dlr.de (T. Jagdhuber), Fan.Chen@ars.usda.gov (F. Chen), Wade.Crow@ars.usda.gov (W. Crow), Peggy.E.ONeill@nasa.gov (P.E. O'Neill), jeff.walker@monash.edu.au (J.P. Walker), aberg@uoguelph.ca (A. Berg), David.Bosch@ars.usda.gov (D.D. Bosch), todd.caldwell@beg.utexas.edu (T. Caldwell), Michael.Cosh@ars.usda.gov (M.H. Cosh), chandra.holifield@ars.usda.gov (C.H. Collins), Ernesto.Lopez@uv.es (E. Lopez-Baeza), mthibeault@conae.gov.ar (M. Thibeault).

https://doi.org/10.1016/j.rse.2019.111380
Received 8 February 2019; Received in revised form 9 August 2019; Accepted 20 August 2019
0034-4257/

global extent. The assessment of high-resolution product at 3 km and 1 km using the soil moisture calibration and validations sites shows reasonable accuracy of ~0.05 m^3/m^3. The SMAP-Sentinel1 active-passive high-resolution product is now available to the public (new version released in October 2018) through NSIDC (NASA DAAC). The duration of this product is from April 2015 to current date.

1. Introduction - background

On January 31st, 2015 NASA launched the Soil Moisture Active Passive (SMAP) mission. The objective of the SMAP mission is to acquire high spatiotemporal resolution surface (top ~5 cm) soil moisture and landscape freeze/thaw state at global extent (Entekhabi et al., 2010). SMAP has an L-band radiometer and a defunct L-band radar with a rotating 6-m mesh reflector antenna. On July 7th, 2015, the SMAP radar malfunctioned and became inoperable. The SMAP radiometer continues to make passive microwave measurements. Since the radar failure the SMAP project explored ways to recover the high-resolution soil moisture capability of the SMAP mission. Specifically by using other active microwave measurements (SAR) from other satellites was investigated. Characteristics and configurations of available SARs, such as global coverage, availability of data, and microwave channel wavelength were among the trade-offs considered in selecting other sources of active radar measurements. The Copernicus Project Sentinel-1A/1B synthetic aperture radar (SAR) data (*Copernicus Sentinel Data 2015. Retrieved from ASF DAAC 29,* 2015) was found suitable for this purpose due to Sentinel nearly similar orbit configuration. Hence, the Sentinel and the SMAP swaths overlap with manageable acquisition time difference, which is key to the SMAP active-passive algorithm (Das et al., 2018). The global coverage based on both Sentinel-1A and Sentinel-1B is the best among available SAR systems. The Sentinel Interferometric Wide Swath (IW) acquisition mode provides the co-pol and cross-pol backscatter observations required for the active-passive algorithm. Some differences do exist between the SMAP SAR data and Sentinel-1A/Sentinel-1B SAR data that include: 1) Sentinel has C-band and SMAP had an L-band SAR instrument; 2) Sentinel has multiple incidence angles within its swath, while SMAP had one single incidence angle at 40 degrees; and 3) Sentinel swath width is ~250 km as compared to SMAP 1000 km swath width. With regards to the last point, the SMAP and Sentinel overlap covers only ~250 km within the 1000 km swath width of the SMAP observations. Therefore, the temporal resolution (revisit interval) for the SMAP active-passive data is degraded from 3 days to 12 days when Sentinel-1A and Sentinel-1B data are used. An advantage of using Sentinel-1A and Sentinel-1B data in the SMAP active-passive algorithm is the potential of getting the disaggregated brightness temperature and soil moisture at much finer spatial resolutions (1 and 3 km). One issue with the combination of L-band radiometer and C-band SAR measurements may be that they represent different emission depths due to differences in frequency. Such differences were tested using a fine-resolution soil moisture advection-diffusion equation (Richard's solver) forced with intermittent precipitation and evaporation at the surface. At the 6:00 AM estimation times, the soil moisture profile in the absence of immediate precipitation has approached a hydrostatic balance (Montaldo and Albertson, 2001). Under these modal conditions the soil moisture profile within the top 5 and even 10 cm of the soil is uniform to within several significant digits. Comparisons with in situ soil moisture measurements (sections below) are presented below.

The active-passive algorithm disaggregates the coarse resolution SMAP radiometer-based brightness temperature (T_B) and soil moisture by using the finer spatial resolution of the Sentinel co-polarized and cross-polarized SAR data and parameters derived from a relationship between the brightness temperature and SAR data. The implementation of the active-passive algorithm is described further in a subsequent section.

The disaggregated high-resolution brightness temperatures from the SMAP-Sentinel1 active passive algorithm were then subjected to a radiative transfer model (Njoku and Entekhabi, 1996) to retrieve soil moisture. The inversion of the disaggregated brightness temperature uses the same suite of algorithms and ancillary data sources as the SMAP radiometer-only soil moisture product processing. Analyses showed that some refinements of parameters were required for the current baseline radiative transfer model (zeroth-order emission model or tau-omega) i.e., the single channel algorithm (SCA). During the initial validation the tau-omega parameters used to generate the SMAP-Sentinel1 soil moisture product (L2_SM_SP) are similar to the parameters applied in the SCA of the SMAP Level 2 Soil Moisture Passive (L2_SM_P/L2_SM_P_E) product. (Chan et al., 2016; and Chan et al., 2017). This implementation is important to maintain consistency between the SMAP-Sentinel1 L2_SM_SP and the SMAP L2_SM_P/L2_SM_P_E products.

The L2_SM_SP product uses the Sentinel-1A and Sentinel-1B SAR data to disaggregate SMAP L-band radiometer measurements from the ~40 km (half-power or − 3 [dB] definition) radiometer measurement to a 3 km and 1 km gridded products. The Sentinel C-band SAR data adds high-resolution spatial details to the radiometer product. It also adds the noise associated with the SAR observations (instrument noise, complex surface scattering, etc.). It is expected that the spatial features in the L2_SM_SP product to be at higher resolution than the SMAP Level 2 Soil Moisture Passive (L2_SM_P/L2_SM_P_E) product. The L2_SM_SP product contains disaggregated brightness temperature at 3 km and 1 km and their respective soil moisture retrievals. The purpose of providing soil moisture retrievals at 1 km is to facilitate agricultural and ecological applications that need high resolution soil moisture. However, the 3 km soil moisture retrievals are primarily used for validation using the SMAP Core Cal/Val Sites. The Core Cal/Val sites are not available at 1 km resolution because they do not satisfy the requirement of minimum number (at least 3) of in-situ sites within 1 km grid cell.

The following sections elaborates the modified active-passive algorithm, results and assessment of the L2_SM_SP product.

2. Active-passive algorithm

The originally developed SMAP Active-Passive algorithm (Das et al., 2014; Entekhabi et al., 2014; Das et al., 2018) is:

$$T_{B_p}(M_j) = T_{B_p}(C) + \beta(C) \cdot \{[\sigma_{pp}(M_j) - \sigma_{pp}(C)] + \Gamma \cdot [\sigma_{pq}(C) - \sigma_{pq}(M_j)]\} \tag{1}$$

where, $T_{B_p}(C)$ [K] is the radiometer-based brightness temperature at coarse resolution (~36 km). The radar backscatter aggregated to coarse-resolution is $\sigma_{pp}(C)$ [dB] and $\sigma_{pq}(C)$ [dB], co-pol and cross-pol, respectively. The radar backscatters $\sigma_{pp}(M)$ [dB] and $\sigma_{pq}(M)$ [dB] are at the desired high-resolution (3 km or 1 km). $\beta(C)$ [K/dB] and Γ [dB/dB] are parameters of the algorithm. The parameter $\beta(C)$ represents the covariation between $T_{B_p}(C)$ and $\sigma_{pp}(C)$ of the SMAP radiometer and radar observations, respectively, and the parameter Γ represents the vegetation-induced heterogeneity within the coarse resolution radiometer cells that is detected by the high-resolution $\sigma_{pp}(M)$ and $\sigma_{pq}(M)$ radar observations. The parameter $\beta(C)$ can be statistically estimated based on a time-series regression using pairs of SMAP radiometer $T_{B_p}(C)$ and spatially-averaged radar data $\sigma_{pp}(C)$. Subsequent repeat overpasses over the same location on the Earth grid are used in the linear time-series regression $T_{B_p}(C) = intercept + slope \cdot \sigma_{pp}(C)$. Clearly these parameters are effective across scale the coarse scale C. Γ is estimated as

$\equiv \left[\frac{\partial \sigma_{pp}(M_j)}{\partial \sigma_{pq}(M_j)}\right]_C$. The value of Γ is specific to the particular grid cell C. It is estimated based on the collection of co-polarized and cross-polarized SAR backscatter cross-section within each coarse grid cell (C). Complete description of the algorithm and parameters is available in the SMAP Active-Passive Algorithm Theoretical Basis Document (https://smap.jpl.nasa.gov/system/internal_resources/details/original/277_L2_3_SM_AP_RevA_web.pdf).

The algorithm (Eq. (1)) is based on the assumption that the linear relationship between the $T_{B_p}(C)$ and $\sigma_{pp}(C)$ holds. Therefore, it is also important to demonstrate that the similar linear relationship is found between the SMAP radiometer $T_{B_p}(C)$ and spatially-averaged Sentinel radar data $\sigma_{pp}(C)$. Fig. 1 illustrates the scatters between the SMAP $T_{B_p}(C)$ and Sentinel $\sigma_{pp}(C)$ from various regions of the world with different landcovers having varying amount of vegetation water content (VWC).

Two years (2017 and 2018) data are used to create the plots in Fig. 1 where ever there are overlap between the SMAP $T_{B_p}(C)$ and Sentinel $\sigma_{pp}(C)$. The slope of the correlation between L-band $T_{B_p}(C)$ and the C-band $\sigma_{pp}(C)$ depends on the level of VWC and the surface roughness. As expected the slope is ~0 for very highly vegetated region such as West Virginia (Fig. 1j). It is obvious from Fig. 1 that the nearly linear relationship is valid for most of the world. However, low correlation is also visible over the dry and arid Sahara desert because the dynamic range in $T_{B_p}(C)$ and $\sigma_{pp}(C)$ is not observed during the two years period. The number of samples for any given site of Fig. 1 is dependent on the availability of the Sentinel 1A/1B granules. With the current global coverage configuration from 2016 October onwards including Sentinel 1A and Sentinel1B the revisit interval is nearly 12 days over most parts of the world except Europe. In Europe, the Sentinel 1A and Sentinel 1B combination acquires observations at 6 days revisit interval. Therefore, over the European sites in Fig. 1 more samples are available. The number of samples also suffer from the SMAP and the Sentinel overlap restriction of 24 h.

The SMAP-Sentinel1 Active-Passive algorithm draws heavily from the above-mentioned algorithm but important changes in implementation of the estimation are introduced. Eq. (1) is now modified to work in emissivity space instead of brightness temperature space and the Sentinel backscatter are in linear scale [−]. Certain aspects of implementation are changed to make it more effective and applicable to accommodate the 12 days revisit interval of the Sentinel satellite. This modification is essential as with the 12 days Sentinel revisit the $T_{B_p}(C)$ and $\sigma_{pp}(C)$ time series is too sparse, and the parameter estimation through time series approach is ineffective/unfeasible. With time-series sampling for 12 days repeat cycle, accumulation of enough data pairs become sparse to allow the statistical estimation of *β(C)* may extend over periods when the vegetation or soil roughness conditions are changing with seasons.

To overcome the limitation of sparse times-series, a snapshot retrieval approach (Jagdhuber et al., 2018) is adopted to estimate the co-variation parameter from the SMAP radiometer and the Sentinel radar observations. The SMAP-Sentinel1 Active-Passive algorithm used in the L2_SM_SP product is:

$$T_{B_p}(M_j) = \left[\frac{T_{B_p}(C)}{Ts} + \beta'(C)\bullet\{[\sigma_{pp}(M_j) - \sigma_{pp}(C)] + \Gamma\bullet[\sigma_{pq}(C) - \sigma_{pq}(M_j)]\}\right] \bullet Ts \tag{2}$$

where, Ts [K] is the emission temperature of the surface soil. The parameter Γ [−] is estimated the same way as mentioned above, however, in a linear scale. The parameter $\beta'(C)$ [−] is estimated in the mentioned snapshot approach because the Sentinel revisit interval of 12 days makes the time series of the Sentinel $\sigma_{pp}(M)$ [−] and $\sigma_{pq}(M)$ [−] data very sparse. The snapshot $\beta'(C)$ is retrieved at each coarse grid cell (C) for every overlap between the SMAP and Sentinel observations,

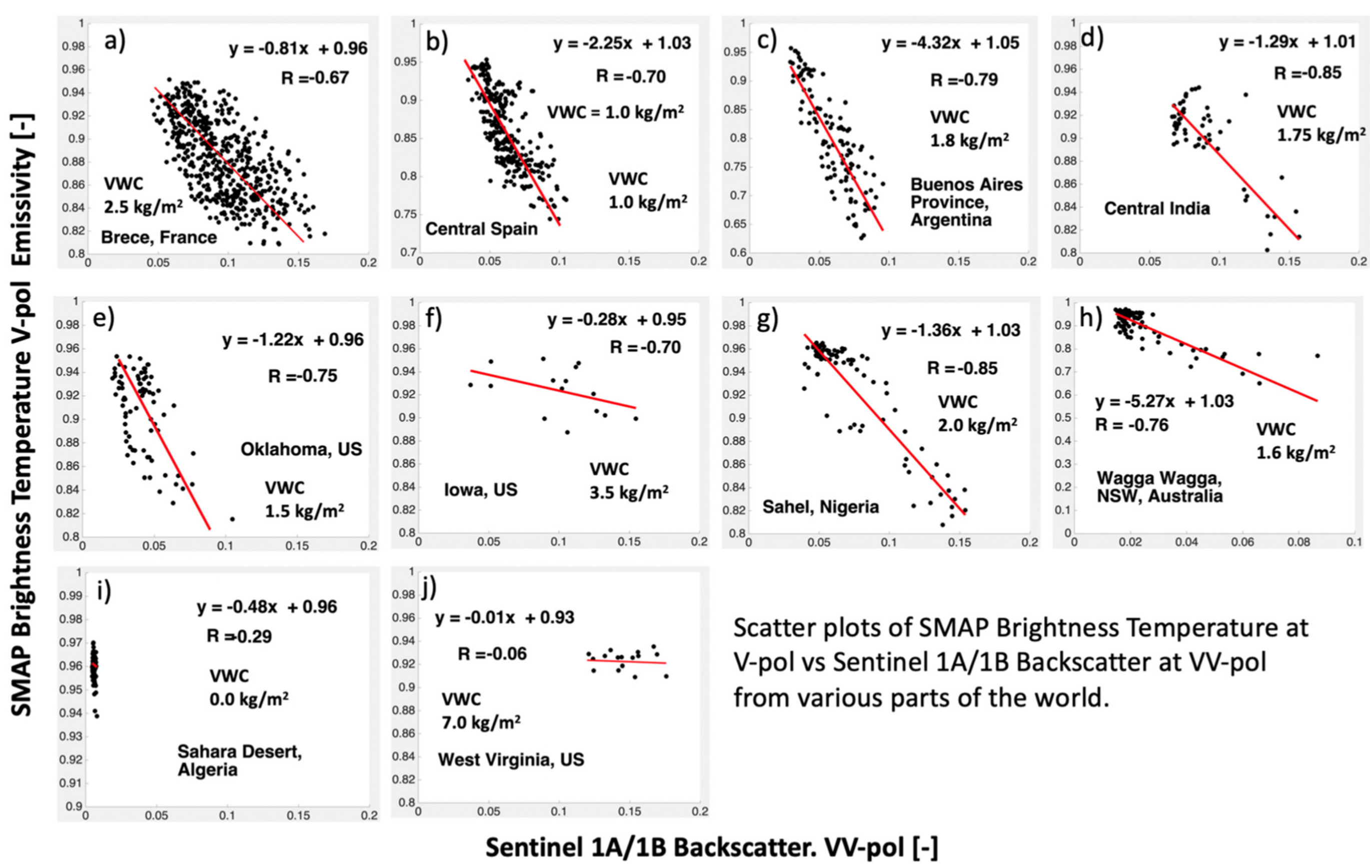

Fig. 1. Scatter plots between the SMAP L-band radiometer $T_{B_p}(C)$ and spatially-averaged Sentinel C-band radar data $\sigma_{pp}(C)$ from various regions of the world.

and is computed as (Jagdhuber et al., 2019):

$$\beta'(C) = \frac{\frac{T_{Bp}(C)}{Ts} - (\gamma + (1-\omega)(1-\gamma))}{|S_{pp}(M_j)|^2 - \mu_{pp-pq} \bullet |S_{pq}(M_j)|^2} \tag{3}$$

where, ω $[-]$ is the effective single scattering albedo, $\gamma = e^{-\tau/cos\theta}$ $[-]$ is the vegetation loss term, and θ_i [rad] is the incidence angle (Jagdhuber et al., 2018). $\beta'(C)$ in Eq. (3) results from eliminating smooth surface Fresnel reflectivity from the tau-omega model and variations in co-polarized backscatter that is due to soil moisture and not vegetation (Jagdhuber et al., 2018). The numerator is the measured surface emission minus the vegetation volume scattering and emission. The denominator is similarly the co-polarized backscatter minus the volume scattering (Jagdhuber et al., 2019). The volume scattering component in the co-polarized backscatter is the total co-polarized backscatter minus the projection of the cross-polarized backscatter onto the co-polarized backscatter. The projection is $\mu_{pp-pq} = \partial|S_{pp}(M_j)|^2/\partial|S_{pq}(M_j)|^2$.

The nadir vegetation opacity τ $[-]$ is related to the physical characteristics of the vegetation layer, such as the vegetation water content (VWC). $|S_{pp}(M_j)|^2$ is co-polarized backscatter, where $|S_{pp}(M_j)|^2 \equiv \sigma_{pp}(M_j)$, and $|S_{pp}(M_j)|^2$ is cross-polarized backscatter, where $|S_{pq}(M_j)|^2 \equiv \sigma_{pq}(M_j)$. μ_{pp-pq} is the same as Γ of Eq. (2), except using a linear regression of backscattering coefficients ($\sigma_{pp}(M_j)$ $[-]$, $\sigma_{pq}(M_j)$ in linear units) at fine scale (3 km) within each coarse-resolution TB grid cell ($T_{B_p}(C)$). These approaches to estimate $\beta'(C)$ and μ_{pp-pq} do not require time series of $T_{B_p}(C)$ and $\sigma_{pp}(C)$. The snapshot approach Eq. (3) (Jagdhuber et al., 2018; Jagdhuber et al., 2019) is capable to accommodate L-band, C-band and X-band combinations of the radiometer and SAR observations at different incident angles. At any given day, the snapshot estimate of the covariance parameter (β′) is unique and is dependent on the radiometer TB (emissivity), SAR backscatter, ω $[-]$ (the effective single scattering albedo), and $\gamma = e^{-\tau/\cos\theta}$ $[-]$ the vegetation loss term (τ is vegetation optical density and θ is incident angle of TB).

For evaluation of $\beta'(C)$ retrieved in snapshot approach, a comparison was made with $\beta(C)$ derived from the time series purely obtained from data of the SMAP mission (SMAP radar and radiometer). Both approaches converge with the $\beta'(C)$ values almost similar to $\beta(C)$ as shown in Fig. 2, except over dryland regions across the Sahara, parts of the Middle East and Central Asia. These dryland regions do not have enough soil moisture variability during the April 1 to July 7 Summer season of 2015 (when the SMAP radar data is available) to induce variations in $T_{B_p}(C)$ and $\sigma_{pp}(C)$ to allow valid time-series estimation of $\beta(C)$. Outside of these regions the magnitudes and distribution of the covariation parameter are similar between the statistical time-series and snap-shot approaches (Jagdhuber et al., 2017).

The baseline SMAP L2_SM_SP algorithm has two parameters ($\beta'(C)$ and Γ), as shown in Eq. (2). The performance of the brightness temperature disaggregation that results in the 3 km and 1 km soil moisture retrievals is heavily dependent on robust estimates of the parameters $\beta'(C)$ and Γ.

Fig. 3 shows the mean and coefficient of variation (CV) of $\beta'(C)$ at global extent using SMAP radiometer and Sentinel-1A/B backscatter data from May 1, 2015 to April 30, 2017. The global evolution of mean $\beta'(C)$ (Fig. 2) shows the typical feature of reducing magnitude (approaching zero) with increasing VWC (Jagdhuber et al., 2019). However, the CV in Fig. 2 represents high variability except over very arid regions. This is a clear indication of seasonality/variability in $\beta'(C)$ and the gradually changing values with the surface conditions, especially VWC. Some low absolute values are also observed over the Sahara desert because of local variation in roughness values leading to high backscatter even for very dry surface.

The estimation of $\beta'(C)$ through Eq. (3) does not require time series of $T_{B_p}(C)$ and $\sigma_{pp}(C)$. Therefore, space-borne radar and radiometer acquisitions with varying incidence angle can be used and the covariation parameter $\beta'(C)$ is dependent on the angle. The range of incidence

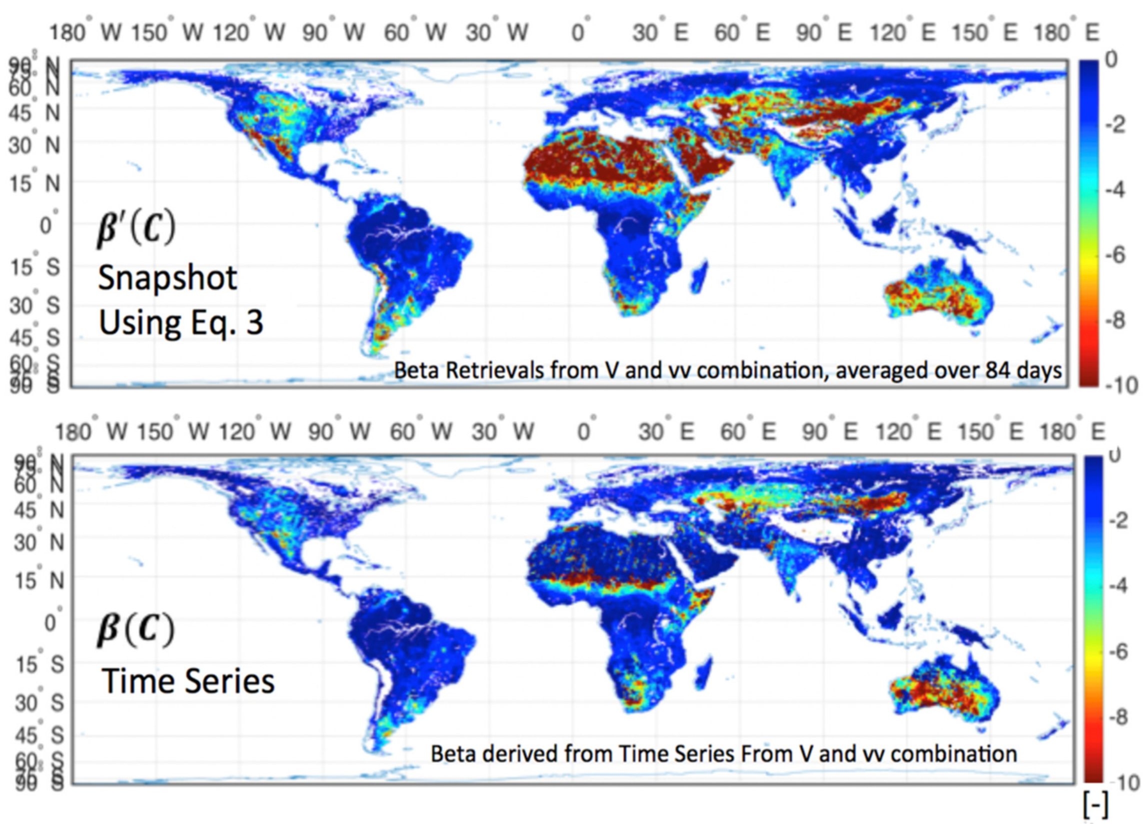

Fig. 2. Comparison of snapshot-retrieved $\beta'(C)$ and time series-retrieved $\beta(C)$ at global extent for SMAP active-passive (~2.5 months) period.

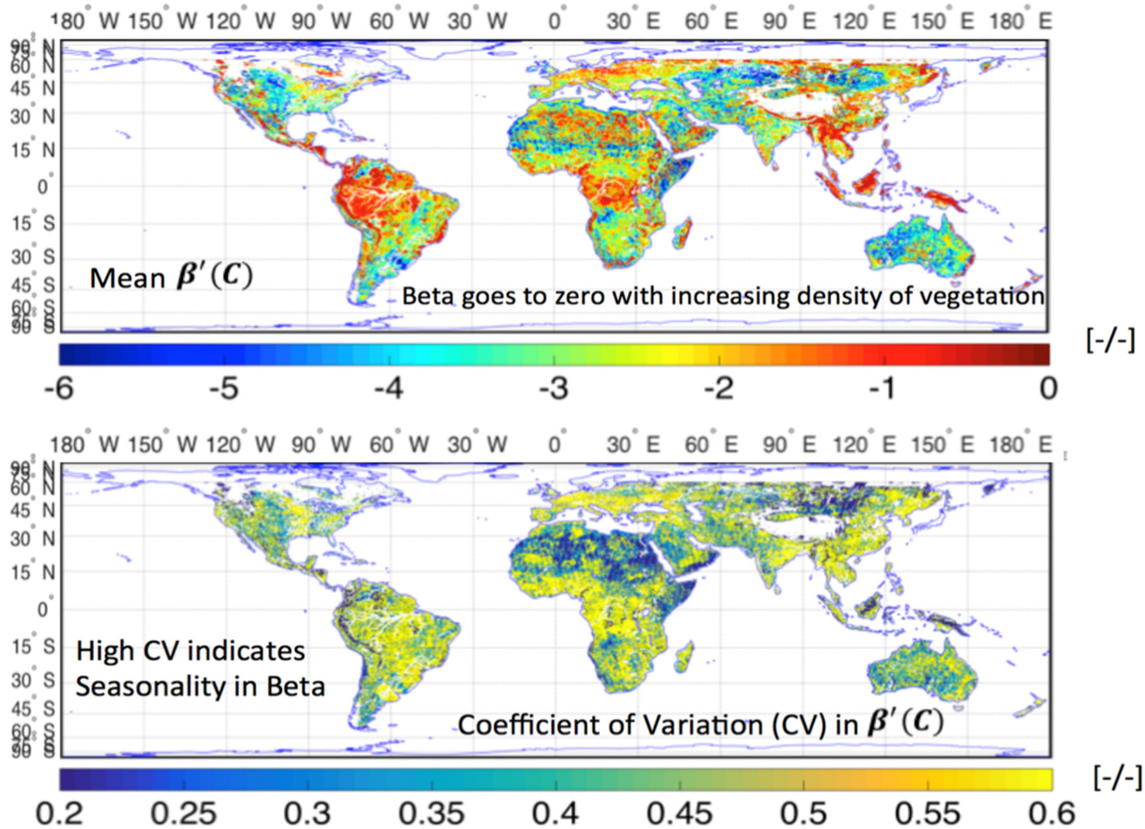

Fig. 3. $\beta'(C)$ mean and CV computed using all the available SMAP radiometer data and Sentinel-1A/Sentinel-1B σ_{pp} data from May 01, 2015 to April 30, 2017.

angles for Sentinel 1A/B observations within the (C) scale (C-scale is 33 km resolution) is ~1 deg. Therefore, linearly averaging the Sentinel 1A/B backscatter is quantitatively possible and valid. In Fig. 4 from Jagdhuber et al. (2019), the dynamics of the covariation estimation with variation of incidence angle (Sentinel-1A/B: 34 [°]–44 [°]) is presented for four different ranges of vegetation water content (VWC) within the African continent. The VWC estimates come from O'Neill et al. (2014). The covariation parameter $\beta'(C)$ from low to moderate amounts of vegetation (VWC $<$ 5 [kg/m^2]), gradually decreases in magnitude with increasing plant moisture. However, the largest change, in magnitude of $\beta'(C)$ along incidence angle (for the lowest VWC-range in Fig. 4) is around 0.5 (Jagdhuber et al., 2019). As expected, $\beta'(C)$ shows minimum sensitivity to incidence angle variations for strongly vegetated areas (VWC $>$ 6 [kg/m^2]) leveling around -1.5 [$-$]. This might be due to the insensitivity of both SMAP radiometer (L-band) and Sentinel-1A/B radar (C-band) to soil moisture variations under highly moist vegetation. One interpretation is that the incidence angle variation of active-passive microwave covariation is increasingly masked/gets absorbed by denser/thicker vegetation (Jagdhuber et al., 2019).

The parameter Γ is determined statistically for any particular overpass using the Sentinel 1A/B radar backscatters σ_{pp} and σ_{pq} at the finest available resolution (in this case at 1 km) that are encompassed within the 33 km $T_{B_p}(C)$ grid cell. The parameter Γ is projection of Sentinel 1A/B σ_{pq} space into the σ_{pp} space. It is the slope of covariance between the Sentinel 1A/B σ_{pq} and the σ_{pp} (Γ is estimated as $\equiv \left[\frac{\partial\sigma_{pp}(M_j)}{\partial\sigma_{pq}(M_j)}\right]_C$). Γ shows that the heterogeneity is captured through the spatial deviation of σ_{pq} backscatter from its mean at (C) scale. The Γ value project this spatial deviation in σ_{pq} backscatter in the σ_{pp} backscatter space. The projection of spatial deviation in σ_{pq} backscatter can be additive or negative with the σ_{pp} backscatter that depends on the vegetation and surface roughness. Therefore, one Γ value is sufficient to capture the heterogeneity of the scene within the (C) scale. In Fig. 5, the values of Γ for all arid regions of the Earth surface is between 4 and 5. This is because the range of the σ_{pq} backscatter response is much lower in the arid region than any other landcover types. Fig. 5 illustrates the mean and Coefficient of Variation (CV) of Γ values over the global extent using all data from May 1, 2015 to April 30, 2017. The Γ parameter is spatial and temporally more stable than $\beta'(C)$. At a global extent, the mean values range from 2.5 to 4.5. The CV in Γ is also very low for any given location, indicating temporal stability of this parameter.

3. Implementation of the SMAP-Sentinel1 active passive algorithm

A simplified process flow chart/processing scheme of the SMAP-

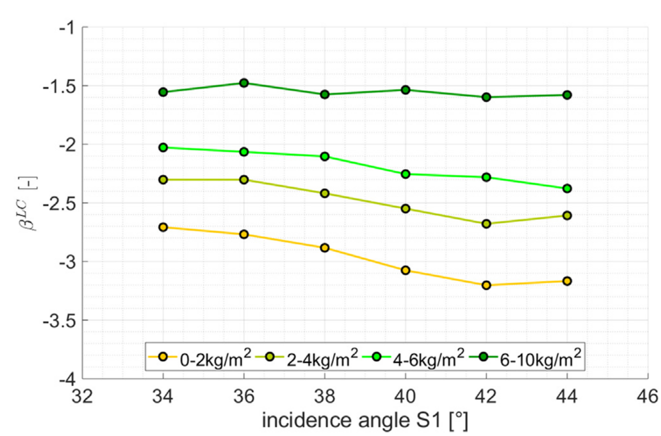

Fig. 4. Time-averaged (04/2015-04/2017) $\beta'(C)$ [$-$] along Sentinel-1A/B incidence angle [°] for four VWC-ranges in Africa; circles indicate median values for each VWC-range (Jagdhuber et al., 2019).

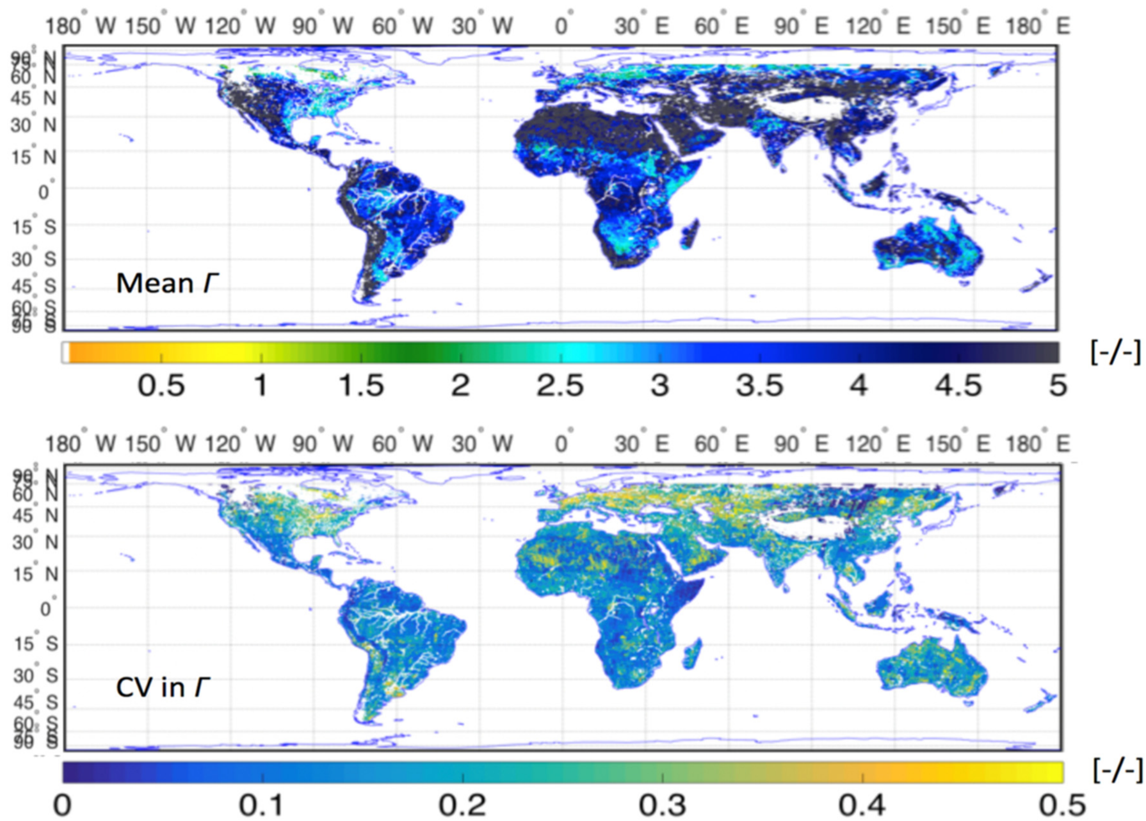

Fig. 5. Γ mean and CV computed using all the available SMAP radiometer data and Sentinel-1A/Sentinel-1B σ_{pp} data (1 km resolution) from May 01, 2015 to April 30, 2017.

Sentinel1 active-passive algorithm implementation is shown in Fig. 6. The input data are the Sentinel 1A/B Interferometric Wide (IW) Swath mode backscatter σ_{pp} (co-pol *vv*) and σ_{pq} (cross-pol *vh*) at 1 km EASE grid resolution and the brightness temperature $T_{B_p}(C)$ from the SMAP Level-2 enhanced product (L2_SM_P_E) at about 33 km spatial resolution in EASE 9 km grid.

The native resolution of Sentinel 1A/1B IW swath mode backscatter σ_{pp} (co-pol vv) and σ_{pq} (cross-pol vh) is ~25 m. The high-resolution Sentinel 1A/1B SAR backscatter data is processed for calibration, noise subtraction, terrain correction (with SRTM DEM) using the ESA Sentinel 1 toolbox (SNAP). Thereafter, the high-resolution Sentinel 1A/1B SAR backscatter data (both σ_{pp} and σ_{pq}), were subjected to filtering, and aggregation (linear averaging) to 1 km. Before aggregation of σ_{pp} and σ_{pq} from ~25 m to 1 km spatial filtering (hybrid spatial filtering

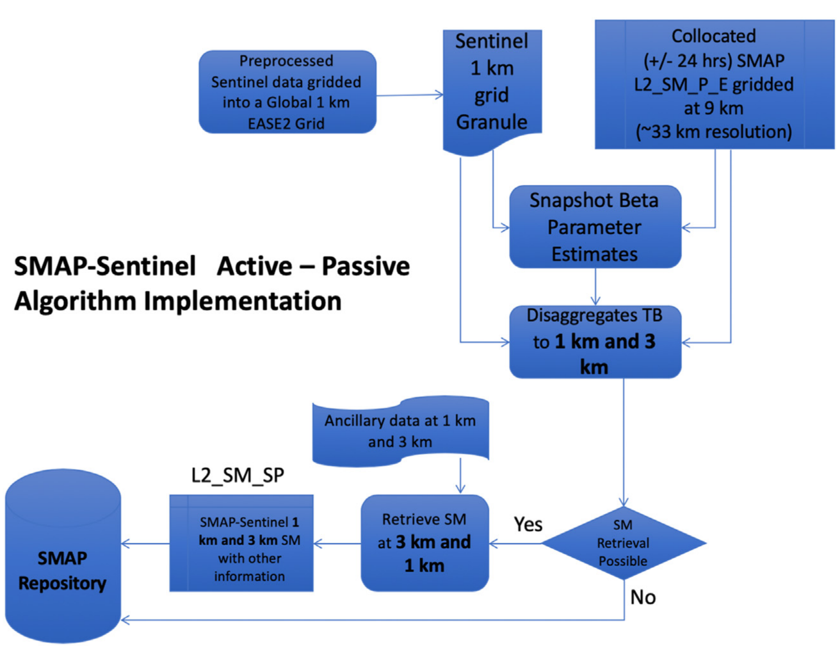

Fig. 6. Process flow/Processing scheme of the SMAP-Sentinel1 active-passive (L2_SM_SP) algorithm in the JPL Science Data System.

tool) was conducted to remove the effect of urban and manmade structures from the backscatter observations. The customized hybrid spatial filtering tool is developed at NASA JPL and is not available in the SNAP toolbox.

Several factors were addressed by the hybrid spatial filtering tool, they are: A) the tool should not affect latency; B) remove most of the unwanted measurements; C) not produce excessive averaging; and, D) preserve image details. Several techniques were studied. Techniques based on standard distribution threshold were efficient but for narrow distributions they showed that some desired features could be lost. A moving window median filter techniques were also efficient in removing undesired measurements but they were computationally expensive and produced excessive averaging when a large size window was used. To overcome all the issues mentioned above, a hybrid filter (combination of median filter and filter based on standard deviation thresholds) was implemented as follows:

1) For each 1 km^2 grid cell within a given Sentinel granule the mean (m_i) and the standard deviation (s_i) were computed, $i = 1…Nc$, where Nc is the number of 1 km^2 grid cell within the Sentinel granule.
2) The tool then computed the mean standard deviation SM over all the s_i with $i = 1…Nc$.
3) For all 1 km^2 cells with $s_i > SM$ a moving window median filter with a 9 × 9 samples window size was applied.
4) For all 1 km^2 cells with $s_i \leq SM$, we eliminated all the Sentinel samples outside the range [m_i - SM: $m_i + SM$] (Note that the threshold SM is used to avoid affecting areas with narrow distribution).

Fig. 7 illustrates the Sentinel 1A σ_{vv} data aggregated to 1 km over Southern Iowa. The high values of σ_{vv}, as highlighted in Fig. 6A, are due to non-natural scatterers (urban areas or manmade structures), these undesired high backscatter observations were filtered for the entire Sentinel granule, and then aggregated to 1 km. The filtered Sentinel 1A σ_{vv} granule is illustrated in Fig. 7.

As shown in the algorithm scheme/flow (Fig. 6), the processed Sentinel 1A and Sentinel 1B data are overlapped with the SMAP observations (descending ~6:00 AM overpasses) that is closest to the Sentinel overpass within ± 24 h time difference. The time difference between the Sentinel 1A/1B (ascending and descending) and SMAP descending is an average of ~12 h. It is expected that the spatial distribution and pattern of the soil moisture does not change significantly because of inherent memory of the soil moisture over a short period of the time difference.

The disaggregated/downscaled brightness temperature ($T_{B_p}(M_j)$) is then obtained by using the algorithm (2) on the overlapped Sentinel 1A/B (σ_{vv} and σ_{vh}) and $T_{B_v}(C)$. The implementation of Eq. (2) is conducted at 33 km resolution (C). The $T_{B_v}(C)$ values in L2_SM_P_E are gridded at/to 9 km, but keeping its inherent spatial resolution of 33 km. Therefore, the overlapped Sentinel 1A/1B data, that forms a grid of 33 rows and 33 columns at 1 km resolution, is used in the process to first compute the snapshot $\beta'(C)$ and then in Eq. (2) to obtain downscaled brightness temperature $T_{B_v}(M_j)$, as illustrated in Fig. 8.

The downscaled brightness temperature $T_{B_v}(M_j)$ is then injected into the tau-omega model (Chan et al., 2016; and Chan et al., 2017) to retrieve surface soil moisture. Various ancillary data and lookup tables are used in the tau-omega model to retrieve soil moisture (Chan et al., 2017). Prominent ancillary data are NDVI climatology from MODIS, clay fraction from global soil database, and land surface temperature (LST) from NASA GMAO, and the parameters are albedo (ω), surface roughness (h), and vegetation coefficient (b) detailed for IGBP landcover classes. These ancillary data and parameters are similar to that used in the L2_SM_P/L2_SM_P_E product (Chan et al., 2016; Chan et al., 2017), however, in a much finer resolutions (1 km and 3 km). The following section elaborates the L2_SM_SP soil moisture product and its characteristics.

4. SMAP-Sentinel1 active-passive (L2_SM_SP) product

At National Snow and Ice Data Center (NSIDC) Distributed Active Archive Center (DAAC), the L2_SM_SP data is available at URL (https://nsidc.org/data/spl2smap_s) from April 15th, 2015 through current. The coverage/overlap of SMAP and Sentinel 1 is from March 2015 onwards. Sentinel 1A is available from March 2015 to current and Sentinel 1B is available from October 2016 to current. The global 12 days global coverage is possible only when the Sentinel 1A and Sentinel 1B are composited with the present data feed from ESA. However, over Europe the coverage is ~ 6 days.

4.1. Patterns and features in the L2SMSP product

The L2_SM_SP product is available at 3 km and 1 km resolution. In this section, prior to the quantitative assessments that follow, the general features of global images are reviewed for the baseline L2_SM_SP product. With the current orbit configuration and data acquisition plan in the IW swath mode, the Sentinel-1A and Sentinel-1B spacecraft have a revisit interval of 6 days to 12 days at different regions of the world. Therefore, the composite of L2_SM_SP for 12 days should cover most parts of the Earth. Fig. 9 shows a 12-day composite of L2_SM_SP granules from 1st May 2017 to 12th May 2017 which illustrates the global coverage between +60° and −60° latitudes. Fig. 9 also provides a complete global extent of soil moisture evolution over different biomes and landcovers. Assessment of global soil moisture from the SMAP-Sentinel1 active-passive retrievals shows consistency in the soil moisture range (0.02 m^3/m^3 to 0.6 m^3/m^3), and probable

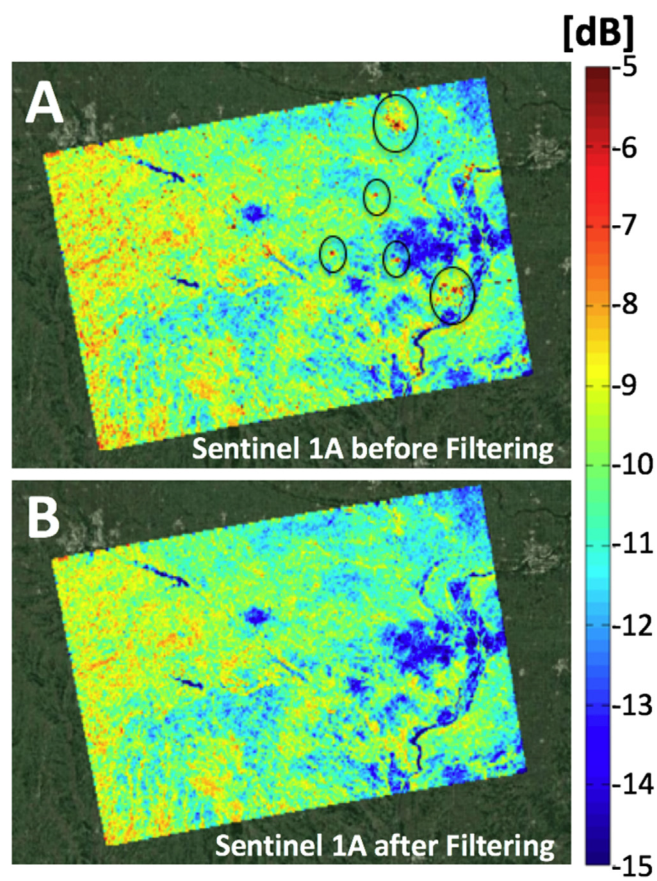

Fig. 7. Sentinel 1A σ_{vv} granule from Southern Iowa on May 05, 2018. A) σ_{vv} unprocessed data; and B) σ_{vv}data after calibration, noise subtraction, terrain correction (using SRTM DEM), filtering, and aggregation to 1 km.

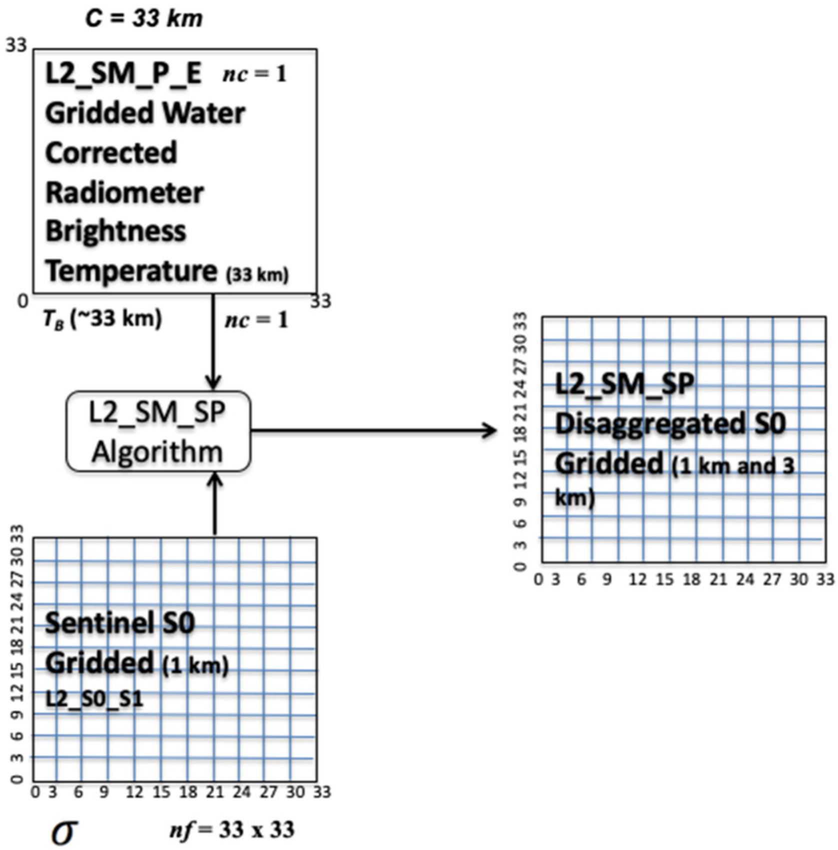

Fig. 8. Grid topology of the SMAP-Sentinel1 active-passive algorithm. C is the coarse scale (~33 km), nc is the number of radiometer grid cell, and nf is the number of Sentinel 1A/1B grid cells at 1 km.

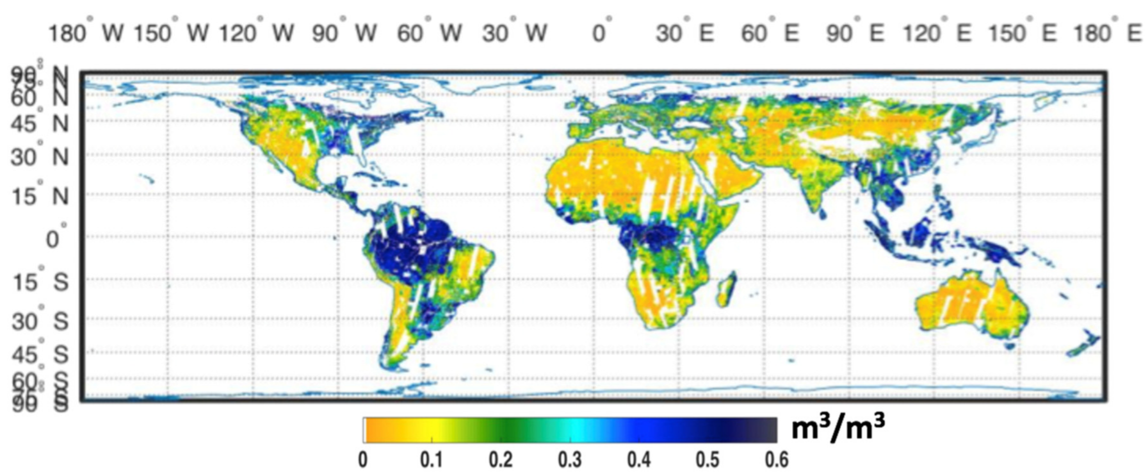

Fig. 9. Twelve Days Coverage of SMAP-Sentinel1 L2_SM_SP high-resolution (3 km) soil moisture data from 1st May 2017 to 12th May 2017.

values. For example, the regions that are very dry (i.e., the Sahara desert) and wet (i.e., the Amazon Basin) reflect the nature of the soil moisture distribution and expected variability as influenced by geophysical factors (soil types, vegetation, weather, and terrain) and landcovers.

There are a number of quality flags that are applied to L2_SM_SP products. These flags imply that the data should be used with caution while others indicate that the data should not be used in any geophysical application. A complete description of the flags and flag thresholds used in L2_SM_SP processing can be found in the Product Specification Document [L2_SM_SP Product Specification Document, available at NSIDC (https://nsidc.org/sites/nsidc.org/files/technical-references/SMAP%20L2_SM_SP%20PSD_20180531.pdf)]. The reliability of soil moisture retrieval algorithms is known to decrease when the VWC exceeds a certain threshold. For the L2_SM_SP product, a 3 kg/m^2 VWC value is used as a flag threshold to indicate areas of high vegetation where soil moisture retrievals are possibly less accurate. A quality flag value of 0 represents good quality and any value > 0 represents substandard quality due to surface flags or due to a quality flag associated with the disaggregated T_{B_p} or due to the quality of the input data ($T_{B_p}(C)$ and σ_{pp} and σ_{pq}). A surface flag is also associated with each and every soil moisture retrieval data field. The surface flags are stored in two bytes integer. There are 16 bits in two byte integer. For example, the first bit position resembles presence of waterbody. The first bit position

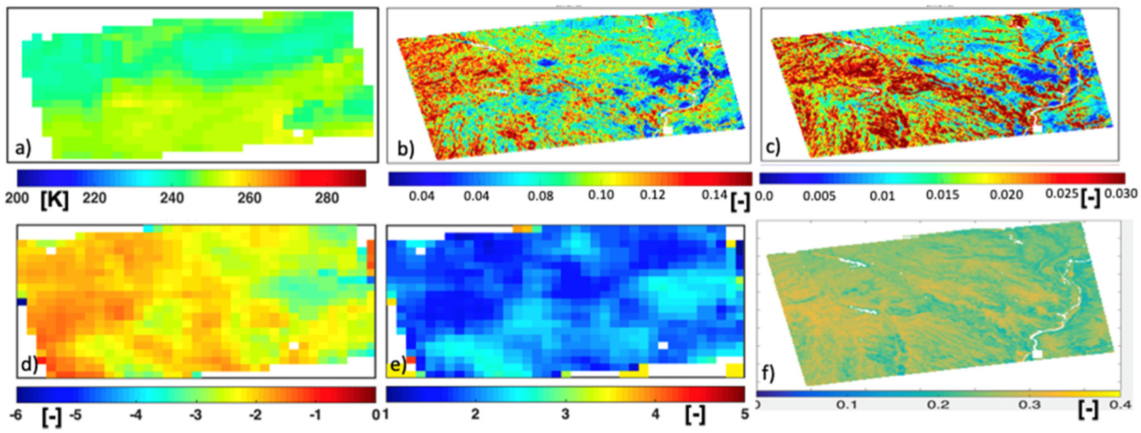

Fig. 10. An example of primary inputs to the SMAP-Sentinel active-passive algorithm. From Southern Iowa, 5th May 2018: a)SMAP radiometer brightness temperature $T_{B_v}(C)$ at about 33 km resolution but gridded at 9 km; b) Sentinel-1A/B co-polarized backscatter (σ_{vv}) at 1 km; c) Sentinel-1A/B cross-polarized backscatter (σ_{vh})at 1 km; d) Parameter $\beta'(C)$; e) Parameter Γ, and; f) Clay fraction.

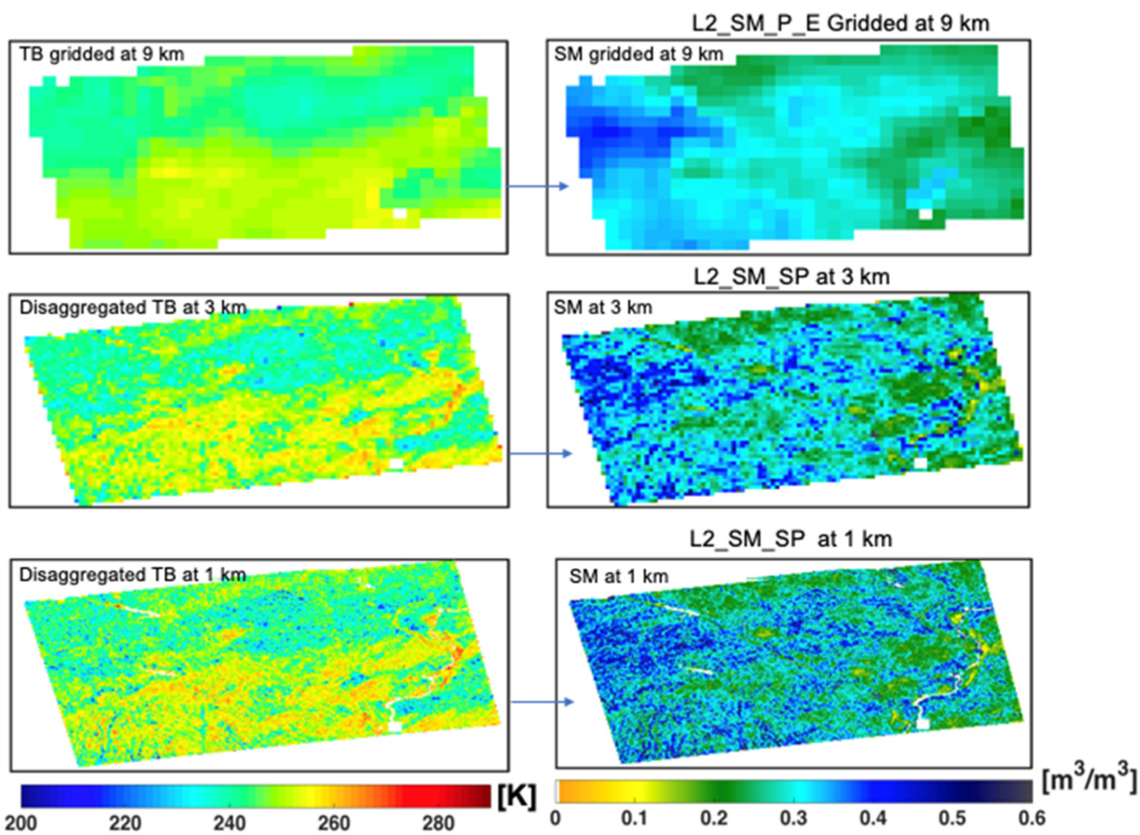

Fig. 11. Comparison of L2_SM_SP product at 1 km and 3 km resolutions with the corresponding L2_SM_P_E product gridded at 9 km. The L2_SM_SP data is from Southern Iowa, May 5th, 2018.

is set to 0 if the water fraction is less than equal to a threshold value (≤ 0.1) else the first bit position is set to 1 if the water fraction is greater than the threshold value (> 0.1). Similarly, the other bits are assigned 0 or 1 based on the threshold values of urban area, mountainous region, VWC, etc.

It is anticipated that some of the flag thresholds may be relaxed in time as the algorithms are improved for the presence of certain currently problematic surface conditions. Other areas that are flagged include regions with varied topography features (for example, mountain ranges) and near large water bodies (coastal regions and areas near large lakes).

The variability within the radiometer coarse grid cell is mostly due to soil moisture, vegetation and soil roughness (Njoku and Entekhabi, 1996; Entekhabi et al., 2010), and is captured by high-resolution Sentinel-1A/Sentinel-1B backscatter values of σ_{vv} and σ_{vh} at the finest available resolution (in this case at ~1 km). For illustration, Fig. 10 shows the primary inputs to the algorithm, the brightness temperature $T_{B_v}(C)$ values in L2_SM_P_E gridded at 9 km (~33 km resolution), and the Sentinel-1A/Sentinel-1B processed σ_{vv} and σ_{vh} backscatter data at 1 km. Fig. 11 illustrates the L2_SM_SP algorithm capability to captures high-resolution spatial features of soil moisture possible through Sentinel-1A/Sentinel-1B backscatter observations (Fig. 10b–c) that disaggregates the brightness temperature $T_{B_v}(C)$ values (Fig. 10a). Fig. 12 shows another perspective to highlight the dynamic range of brightness

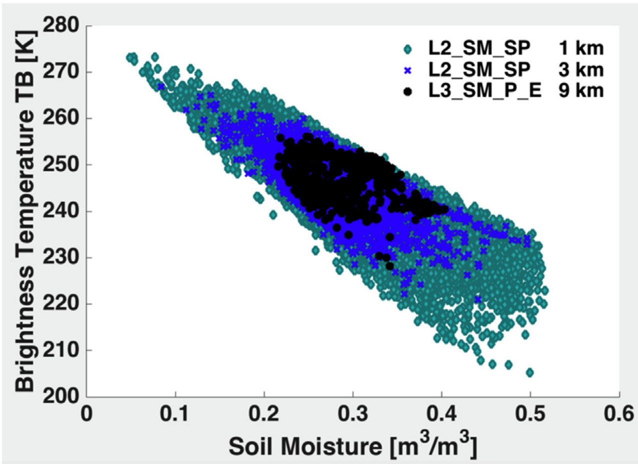

Fig. 12. Distribution of data in the soil moisture and the brightness temperature space for the L2_SM_SP product at 1 km and 3 km, and L2_SM_P_E product gridded at 9 km (~30 km effective resolution) over the Southern part of Iowa on 5th May 2018.

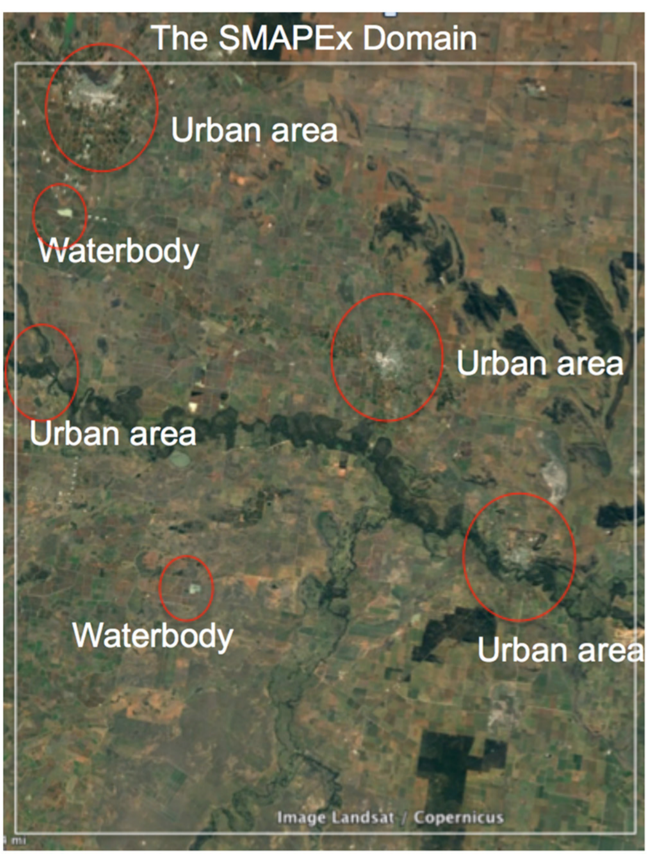

Fig. 13. Study domain (white frame) of SMAPEx Airborne campaign conducted in year 2015 (Ye et al., 2015). Red frames indicate areas not included for soil moisture estimation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

temperature and soil moisture present in the SMAP-based soil moisture products. The plot clearly shows the increase in variability and dynamic range in the L2_SM_SP product at 3 km and 1 km resolution when compared to the 9 km gridded L2_SM_P_E data that has an effective resolution of ~33 km.

5. SMAP-Sentinel1 active passive (L2_SM_SP) product validation

The assessment of the L2_SM_SP product was performed using two different approaches: 1) by comparing the disaggregated/downscaled brightness temperature with the high-resolution brightness temperature observed through an airborne platform; and 2) comparing the soil moisture retrievals from L2_SM_SP against upscaled in situ soil moisture data.

5.1. Assessment of L2_SM_SP downscaled brightness temperature

A primary part of the assessment for the L2_SM_SP algorithm is the comparison of disaggregated high-resolution brightness temperatures with L-band airborne remote sensing data. This assessment was done using airborne data from the SMAPEx 2015 campaign held/conducted in Southeastern Australia (Ye et al., 2015). The brightness temperature data from SMAPEx 2015 has a resolution of ~1 km with varying incidence angles. For better comparison with SMAP satellite data, the SMAPEx airborne data are subjected to normalization to bring all the observations to a uniform 40 deg. incidence angle (Ye et al., 2015). This process introduced an error of ~4–5 K in the SMAPEx airborne data (Ye et al., 2015). The normalized data are actually used for assessment of the L2_SM_SP disaggregated high-resolution brightness temperature. There were 2 overlapping days (May 5th, 2015 and September 13th, 2015) between SMAP-Sentinel1 L2_SM_SP product and Polarimetric L-band Microwave Radiometer (PLMR) airborne data from the SMAPEx field campaign. These concurrent acquisitions of data from different platforms provide the opportunity/possibility to validate/compare the L2_SM_SP high-resolution disaggregated brightness temperature. These specific dates of SMAPEx airborne data are also considered due to very different surface conditions in the observation domain: a) May 5th, 2015, low vegetation cover (~1 kg/m2); and b) September 13th, 2015, moderately high vegetation cover (~2.7 kg/m2). A map of the SMAPEx 2015 domain is shown in Fig. 13.

As illustrated in Fig. 13, the SMAPEx study domain contains many urban areas, small manmade structures, and waterbodies. These urban areas and waterbodies were undesirable for assessment purposes. Therefore, such data need to be flagged or masked during L2_SM_SP assessment.

Fig. 14a shows the PLMR airborne T_{B_v}data, Fig. 14b shows the Sentinel σ_{vv} data, and Fig. 14c shows the Sentinel σ_{vh} data from May 5th, 2015 over the SMAPEx study area. It is apparent that PLMR $T_{B_v}$from SMAPEx are not impacted adversely by small urban areas or manmade structures, unlike the Sentinel σ_{vv} and σ_{vh} data. (cf. Fig. 6). Fig. 13b–c also show that in the Sentinel data, the large urban areas are masked and removed but the small urban areas and manmade structures are not identified and masked. These types of undesirable outliers in the Sentinel backscatter data created anomalies in the L2_SM_SP disaggregated T_{B_v}data during a first assessment. However, the combined standard deviation and median spatial filter, as discussed in Section 3, was successfully implemented to remove the small urban areas, manmade structures and waterbodies.

Examples of disaggregated high-resolution 3 km T_{B_v} from L2_SM_SP product are shown in Fig. 15a and Fig. 15b, and compared against the SMAPEx PLMR data and the SMAP L2_SM_P_E (T_{B_v} data corrected for presence of water) product gridded at 9 km for May 5th, 2015 and Sep. 13th, 2015, respectively. The plots in Fig. 15 show the finer details captured by the L2_SM_SP active-passive algorithm due to incorporating the Sentinel backscatter observations. In addition, the finer spatial features are very similar to the PLMR T_{B_v}data. To evaluate the SMAP-Sentinel1 Active-Passive algorithm performance, the L2_SM_SP high-resolution disaggregated T_{B_v} are compared against Minimum Performance criteria to determine the value of combining Sentinel-1A/Sentinel-1B SAR data with SMAP L2_SM_P_E brightness temperature data. The Minimum Performance is the SMAP L2_SM_P_E $T_{B_v}(C)$ that is applied to all the 3 km EASE grid cells within the overlapping 9 km EASE

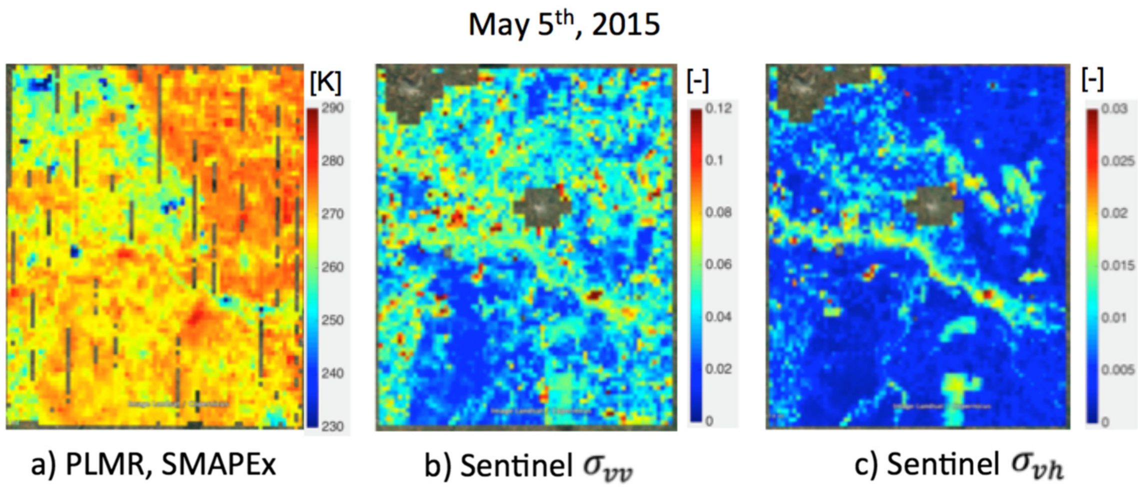

Fig. 14. PLMR and the Sentinel observations at EASE grid 1 km resolution over the SMAPEx study domain on May 5th, 2015.

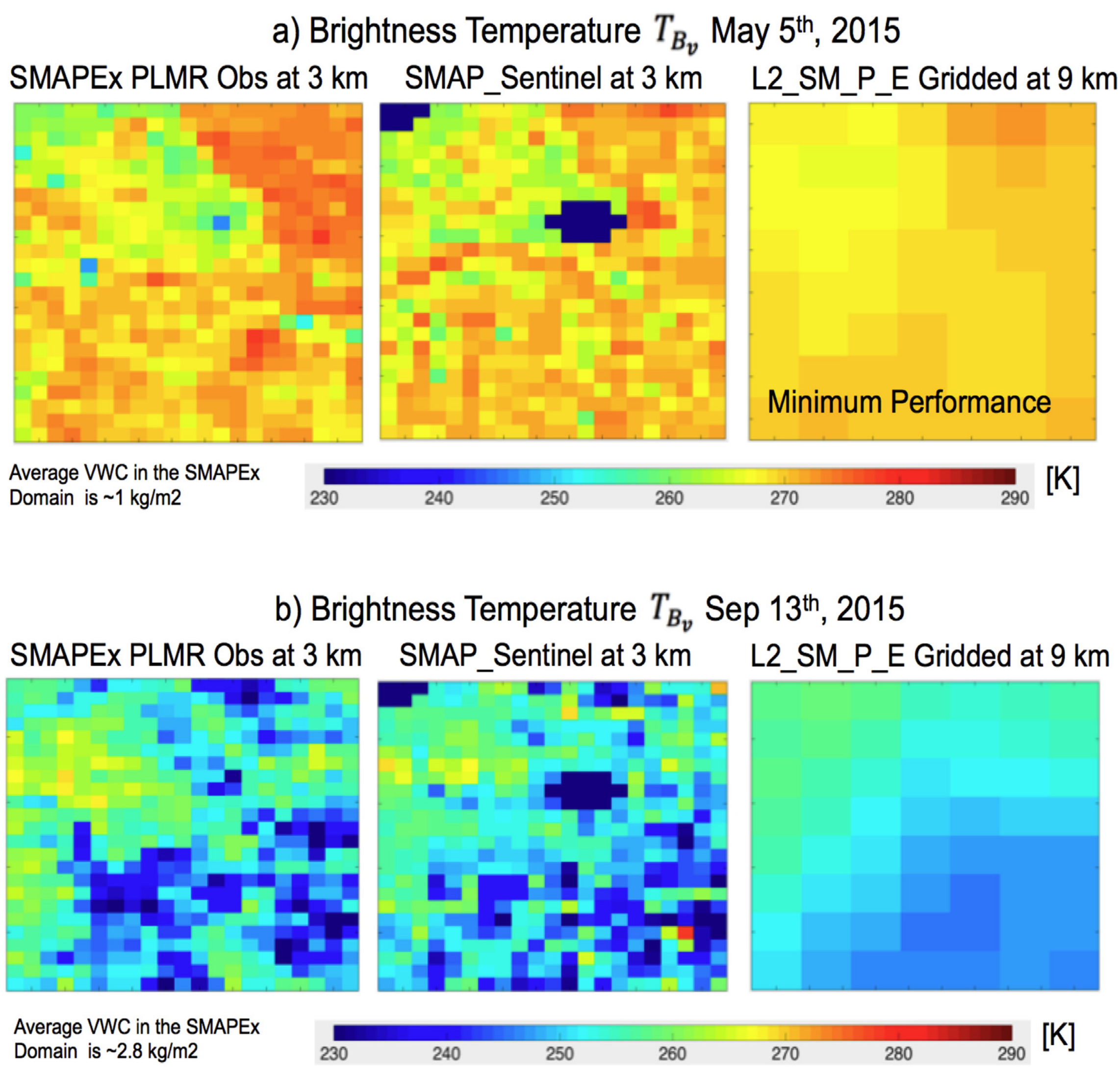

Fig. 15. Output of L2_SM_SP compared against PLMR T_{B_v}data from SMAPEx and the Minimum Performance (T_{Bv} from L2_SM_P_E at 9 km).

grid cell; it can be obtained by setting $\beta'(C) = 0$ in Eq. (2). Ideally the slope and correlation between the L2_SM_SP downscaled brightness temperature and airborne high-resolution brightness temperature should be close to one (unity). In Fig. 16, we show the slope and correlation between Minimum Performance and airborne data, between L2SMSP and airborne data and ideal performance. In the two available airborne images (May 5th, 2015and Sep 13th, 2015) the slope and correlation between L2_SM_SP downscaled brightness temperature and

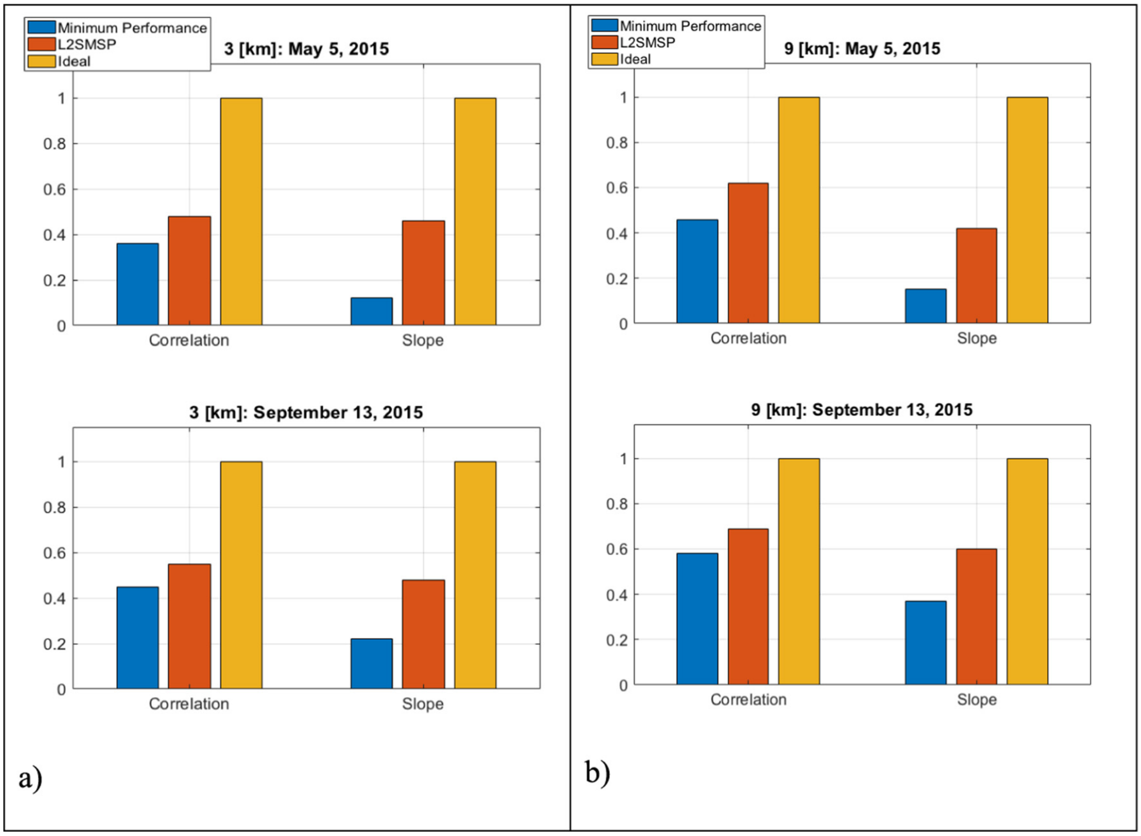

Fig. 16. a) Bar plots of SMAPEx PLMR observations against L2_SM_SP T_{B_p} at 3 km and Minimum Performance (T_{Bv} from L2_SM_P_E) at 3 km. b) Bar plots of SMAPEx PLMR observations against L2_SM_SP T_{B_p} gridded at 9 km and Minimum Performance (T_{Bv} from L2_SM_P_E) gridded at 9 km. The overall RMSE of TB is ~3.4 K for the L2_SM_SP product and ~4.6 K for the minimum performance L2_SM_P_E at 3 km resolution, and ~2.5 K for the L2_SM_SP product and ~3.3 K for the minimum performance L2_SM_P_E at gridded 9 km resolution.

Table 1
SMAP Cal/Val Partner Sites Providing Validation Data for the L2_SM_SP product.

Site name	Site PI	Area	Climate regime	IGBP land cover	Status
Walnut Gulch[b]	C. Holifield Collins	USA (Arizona)	Arid	Shrub open	Valid for 3 km and 9 km
Fort Cobb[a]	P. Starks	USA (Oklahoma)	Temperate	Grasslands	Valid for 9 km
Little Washita[a]	P. Starks	USA (Oklahoma)	Temperate	Grasslands	Valid for 9 km
South Fork[a]	M. Cosh	USA (Iowa)	Cold	Croplands	Valid for 9 km
Little River[a]	D. Bosch	USA (Georgia)	Temperate	Cropland/natural mosaic	Valid for 9 km
TxSON[b]	T. Caldwell	USA (Texas)	Temperate	Grasslands	Valid for 3 km and 9 km
Kenaston[b]	A. Berg	Canada	Cold	Croplands	Valid for 9 km
Carman[a]	H. McNairn	Canada	Cold	Croplands	Valid for 9 km
Monte Buey[b]	M. Thibeault	Argentina	Arid	Croplands	Valid for 3 km and 9 km
REMEDHUS[b]	J. Martinez	Spain	Temperate	Croplands	Valid for 3 km and 9 km
Valencia[b]	E. Lopez-Beaza	Spain	Arid	Shrub (open)	Valid for 3 km and 9 km
St Josephs[a]	M. Cosh	USA (Indiana)	Cold	Croplands	Valid for 9 km
Yanco[b]	J. Walker	Australia	Arid	Croplands	Valid for 3 km and 9 km

[a] CVS used in assessment at 9 km.
[b] CVS used for both 3 km and 9 km.

airborne data are higher than the Minimum Performance (and approaching Ideal). A similar analysis conducted at EASE grid 9 km in Fig. 16b also shows (that the L2_SM_SP $T_{B_v}(M_j)$ aggregated to 9 km has better slopes and correlations when compared against L2_SM_P_E $T_{B_v}(C)$. These results (Fig. 16a and b) clearly demonstrate that Sentinel σ_{vv} and σ_{vh} data bring/include/add valuable information to disaggregate the coarse-resolution L2_SM_P_E $T_{B_v}(C)$ to obtain L2_SM_SP $T_{B_p}(M_j)$ that matches better with the high-resolution spatial features, e.g. observed by the SMAPEx PLMR platform.

5.2. Core validation sites (CVS)

The SMAP L2_SM_SP product validation was based primarily on comparison of retrievals with in situ soil moisture measurements (Colliander et al., 2017; Chan et al., 2016; Chan et al., 2017; Das et al., 2018). The in situ measurements for the top ~5 cm from soil moisture networks with an acceptable sensor density within a 3 km EASE2 grid are the primary validation locations for the L2_SM_SP product. The SMAP project collaborated with various partners from around the world to identify such locations and established CVS (Colliander et al., 2017). These CVS have been verified as providing a spatial average of soil moisture at 3 km (with at least 3 in situ sites) and 9 km (with at least 5

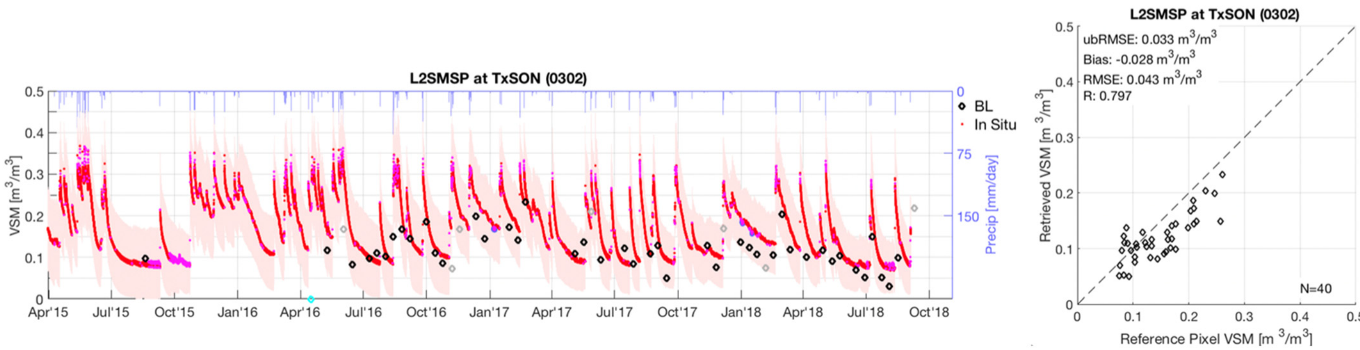

Fig. 17. L2_SM_SP assessment at 3 km (40 samples) for TxSON, Texas, USA. (BL: L2_SM_SP).

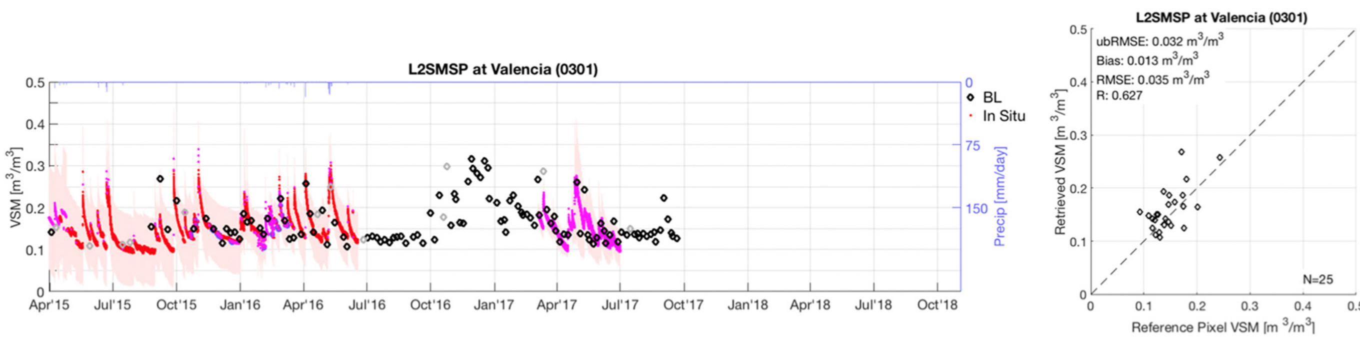

Fig. 18. L2_SM_SP assessment at 3 km (25 samples)for Valencia, Spain. (BL: L2_SM_SP).

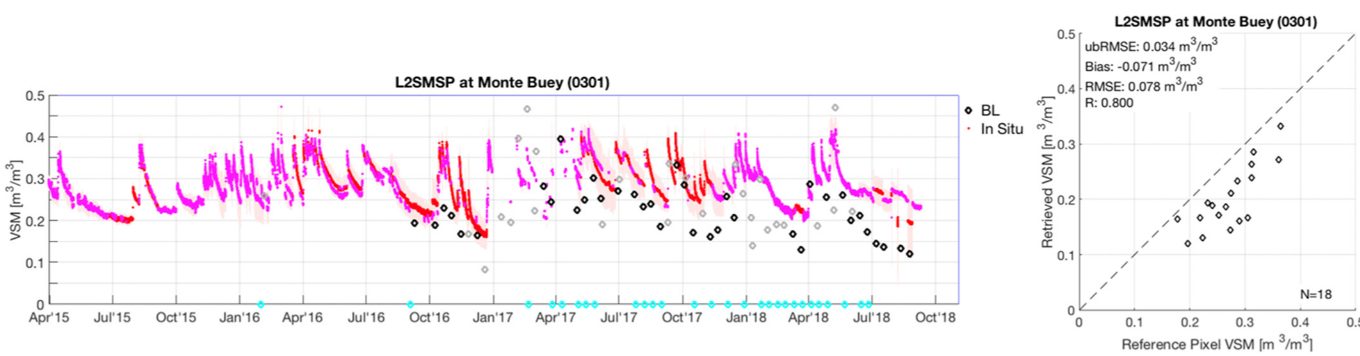

Fig. 19. L2_SM_SP assessment at 3 km (18 samples) for Monte Buey, Argentina. (BL: L2_SM_SP).

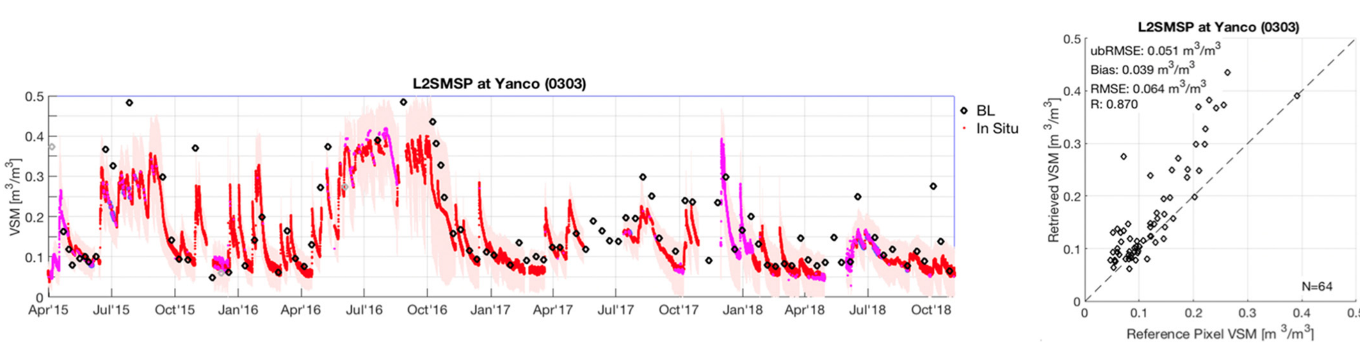

Fig. 20. L2_SM_SP Assessment at 3 km (55 samples) for Yanco, Australia. (BL: L2_SM_SP).

in situ sites) spatial resolutions. However, the spatial averages of soil moisture from CVS are not without issues because of inherent upscaling errors. Table 1 lists the CVS sites used for validation of the L2_SM_SP product. Beside the CVS, sparse networks (Chen et al., 2017) were also used as a supporting tool/option to validate the L2_SM_SP product.

The in-situ data obtained from the SMAP Cal/Val Partner Sites (Table 1) are subjected to quality control (QC) before using them to validate the SMAP products. A QC software tool was developed at JPL using the approached presented in Dorigo et al., 2013 for QC of the in-situ soil moisture data. Figs. 17–20 illustrate time series and scatter plot comparisons of L2_SM_SP product at 3 km grid cells against four CVS: TxSON, Monte Buey, Valencia, and Yanco. A total number of twelve

Table 2
SMAP L2_SM_SP assessment statistics against CVS at 3 km.

Site name	ubRMSE	Bias	RMSE	R	#Samples
Walnut Gulch	0.033	0.033	0.047	0.950	23
Walnut Gulch	0.029	0.063	0.069	0.929	21
TxSON	0.041	−0.039	0.056	0.895	24
TxSON	0.033	−0.028	0.043	0.797	40
Kenaston	0.065	−0.052	0.083	0.216	31
Kenaston	0.053	−0.047	0.071	0.603	24
Monte Buey	0.034	−0.071	0.078	0.800	20
Valencia	0.032	0.013	0.035	0.627	25
Yanco	0.082	0.013	0.085	0.489	38
Yanco	0.075	0.018	0.077	0.635	40
Yanco	0.048	0.037	0.060	0.898	64
Yanco	0.065	0.070	0.096	0.761	48
SMAP average	0.049	0.001	0.067	0.717	Total: 398

The term RMSE in the analysis is interchangeably used for root-mean-square-difference (RMSD). However, RMSD is more appropriate because the upscaled CVS value is not the truth. Nearly 30 to 50 time-matching samples are found in core sites and are used in computing he statistics.

Table 3
SMAP L2_SM_SP assessment statistics against CVS measurements at 9 km.

Site name	ubRMSE	Bias	RMSE	R	#Samples
Walnut Gulch	0.025	0.022	0.033	0.871	37
Walnut Gulch	0.028	0.049	0.056	0.863	47
TxSON	0.022	0.010	0.025	0.882	23
TxSON	0.030	0.012	0.032	0.904	42
Fort Cobb	0.030	−0.023	0.038	0.847	48
Little Washita	0.039	−0.032	0.051	0.771	93
South Fork	0.060	−0.031	0.067	0.802	39
St Josephs	0.022	−0.042	0.048	0.913	24
Little River	0.030	0.086	0.091	0.799	22
Kenaston	0.038	−0.061	0.072	0.764	28
Kenaston	0.031	−0.074	0.080	0.836	27
Carman	0.049	−0.069	0.085	0.590	20
Monte Buey	0.014	−0.049	0.051	0.967	23
REMEDHUS	0.059	0.112	0.126	0.831	63
Valencia	0.027	0.012	0.029	0.746	24
Yanco	0.055	−0.005	0.055	0.877	44
Yanco	0.049	0.037	0.061	0.831	67
SMAP average	0.036	−0.003	0.059	0.829	Total: 671

3 km grid cells from the 7 CVS were used to compute statistics for primary validation of the L2_SM_SP product. Table 2 shows the performance statistics/metrics for all the CVS used for validation. The time series plot in Figs. 17 for the TxSON site shows a good match between soil moisture trends, with some bias in soil moisture estimation compared to in situ measurements that is possible due to difference in soil texture used in the retrieval process. The performance of the L2_SM_SP product over most of the CVS with non-crop landcovers is reasonable as illustrated in Fig. 17 for TxSON and Fig. 18 for Valencia. However, the performance of the L2_SM_SP over CVS with crop cover is inferior, as shown in Fig. 19, possibly because of being out of sync with the vegetation attribute information and strong C-band interaction with vegetation might cause patterns not totally attributable to soil moisture but vegetation cover. The retrieval process uses vegetation-water-content (VWC) derived from the NDVI climatology (developed from 10 years of MODIS data), which might lead to a mismatch with the actual status of VWC. Therefore, it is likely that in Fig. 19 (Monte Buey CVS) the lack of a consistent bias and has higher errors may be caused by the mismatch. In Figs. 17–20 red color for in situ data represents good quality, and the purple color is when the in situ data quality is not satisfactory. The black dots are the data used in the scatter plot and computation of RMSE. The grey dots are the L2_SM_SP data that matches on a given day with the inferior quality in situ data and are not used in calculation of the RMSE and R values.

The validation results at 3 km resolution in Figs. 17–20 and Table 2 comes from a very limited number of CVS. Thus, another strategy was developed/followed to overcome this limitation: An upscaled L2_SM_SP product at 9 km is formed/constructed by aggregating all nine L2_SM_SP 3 km EASE grid cells within the 9 km EASE grid. The upscaled 9 km-product is then used for the CVS sites (17 sites) already established and operating for the SMAP-only Active-Passive L2_SM_AP 9 km product (Das et al., 2018). This approach optimizes the CVS usage and has potential to evaluate/validate/assess the performance of the spatially upscaled L2_SM_SP 3 km product at 9 km. The results and performance of the upscaled L2_SM_SP product at 9 km in Table 3 are encouraging. This product meets the L1 accuracy requirement of the SMAP mission (ubRMSE < 0.04 [m^3/m^3]) previously applied/established/as benchmark to the SMAP-only L2_SM_AP product. The overall ubRMSE of 0.036 m^3/m^3 for L2_SM_SP product meets the SMAP mission accuracy goal of 0.04 m^3/m^3. In Table 3 most of the R-values are relatively high (R > xx) reporting a sufficient/significant/considerable match between estimates and in situ measurements.

5.3. Sparse soil moisture networks

The intensive CVS validation performed for the SMAP L2_SM_SP product can be complemented by sparse networks as well as by new/emerging types of soil moisture networks. The important difference in interpreting these data is that they involve only 1 in situ point in a grid cell. Thus, whatever reservations might exist on the upscaling on CVS in situ measurements to resolution cells of remote sensing products. They might be of even greater concern with sparse in situ networks of soil moisture measurements. However, sparse networks do offer many sites in different environments for comparison.

The established sparse soil moisture networks utilized for the SMAP L2_SM_SP product comparison were the NOAA Climate Reference Network (CRN), the USDA NRCS Soil Climate Analysis Network (SCAN), the Oklahoma Mesonet, the MAHASRI network (in Mongolia), the SMOSMania network (in southwest Europe), the Pampas network (in Argentina), and soil moisture estimates derived from the surface reflectance at Global Position Stations (in the Western US). From these sparse soil moisture networks, ~375 sites were found to be suitable for direct comparison with the SMAP L2_SM_SP overlapping grid cells. The ~375 sites were selected based on in situ measurement data quality and continuity of the observations during the ~3 years period (April 2015 to July Oct, 2018). The defining feature of these networks were the low spatial density of in situ measurement locations that usually resulted in one point per L2_SM_SP 3 km and 1 km grid cells. This would lead to large upscaling errors due to spatial representativeness and the inability of a single in situ site location to describe/represent mean soil moisture within a 3 km or 1 km grid cell. However, despite this scaling bias, sparse networks can adequately describe relative errors.

Fig. 21A–B illustrates the L2_SM_SP product retrievals comparison with the measurements available from ~375 in situ sparse networks from many different landcovers at 3 km and 1 km, respectively. Despite the potential errors associated with spatial representativeness, the agreement between the in situ soil moisture and the L2_SM_SP is reasonably good (see Table 4). The ubRMSE and bias values obtained from these sparse networks are similar to those obtained from the CVS. These results (Fig. 21) provide further confidence in the previous conclusions based on the CVS.

6. Discussion

There is further potential for improvement in the L2_SM_SP data quality, and that is possible by reducing the errors in soil moisture retrievals. The improvements include, use of better ancillary data (e.g. optimized VWC, and better soil texture data) and optimization of the tau-omega model parameters for various landcovers at resolutions of 3 km and 1 km. Currently, the SMAP L2_SM_SP retrievals use the same tau-omega parameters as the L2_SM_P_E retrievals. Another important

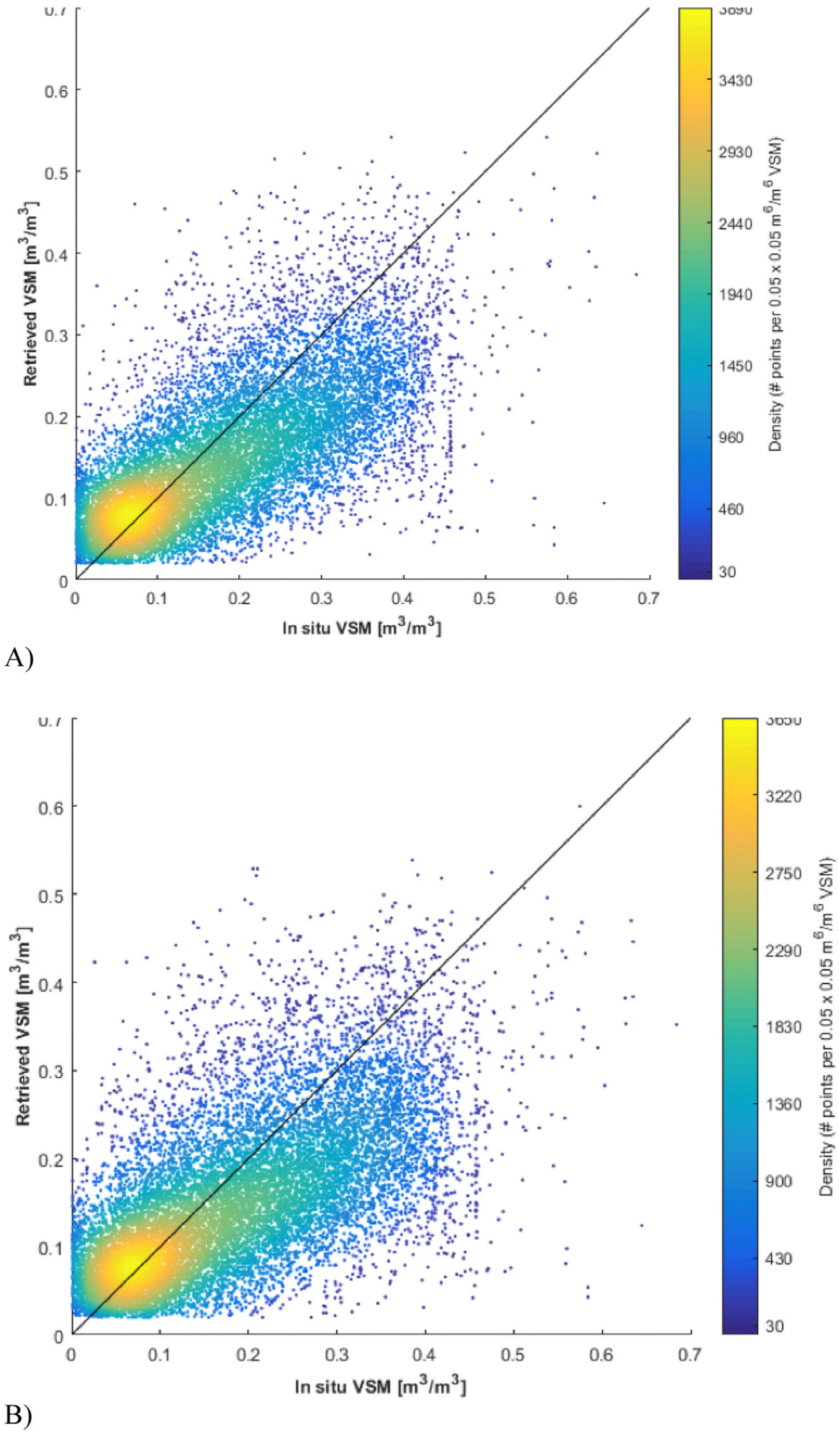

Fig. 21. Results of comparison between L2SMSP with the sparse network sites (~375 in situ sites): A) at 3 km resolution; and B) at 1 km resolution.

step to improving the L2_SM_SP data quality is the inclusion of retrieved vegetation-optical-depth (VOD), meaning tau, from dual-channel algorithms such as Konings et al. (2017). The tau values used for L2_SM_SP retrievals were derived from a 10-year (2002 − 2012) climatology of NDVI based VWC (tau = b*VWC, b is a parameter based on landcover, typically close to 0.1). The drawback of using VWC climatology for tau is prominently visible over CVS with cropland landcover. Fig. 22 illustrates one such scenario where the NDVI climatology taken from Day-of-Year (DOY) 185 is compared against the actual NDVI for DOY 185, 2017 for California. Two time series plots, one from natural landscape with shrubland cover and other with cropland are shown in Fig. 22. The actual NDVI time series over cropland (agricultural region) does not match with the climatology (2002–2012) mostly due to crop rotation and difference in planting date. However, the climatology and the actual time series over the shrubland are almost similar. Such mismatches are very possible over the CVS with crop landcover, hence leading to inferior performance of the L2_SM_SP product, as visible in Table 2 for CVS Yanco and Kenaston. Inclusion of actual NDVI in the operational L2_SM_SP process has potential to improve the overall quality (reduced RMSE), however, the 8–16 days latency of MODIS NDVI data is a constraint. As an alternative to NDVI, the cross-polarized Sentinel-1A/B measurements could also be used as a variable that is proportional to vegetation density (Vreugdenhil et al., 2018). In that case cloud-cover is no longer a constrained. We are actively pursuing

Table 4
SMAP L2_SM_SP assessment statistics against sparse network at 3 km and 1 km resolutions.

L2_SM_SP (3 km)	ubRMSE [m^3/m^3]	Bias [m^3/m^3]	RMSE [m^3/m^3]	R [−]	N
Open shrublands	0.04	0.017	0.045	0.506	34
Woody savannas	0.053	0.031	0.063	0.657	4
Savannas	0.04	−0.001	0.06	0.789	6
Grasslands	0.051	−0.032	0.064	0.647	230
Croplands	0.072	−0.033	0.087	0.531	69
Crop/natural vegetation mosaic	0.067	−0.023	0.076	0.469	14
Barren/sparse	0.026	0.031	0.04	0.514	9
Average	0.05	−0.01	0.062	0.587	370
L2_SM_SP (1 km)	**ubRMSE [m^3/m^3]**	**Bias [m^3/m^3]**	**RMSE [m^3/m^3]**	**R [−]**	**N**
Open shrublands	0.046	0.008	0.046	0.544	43
Woody savannas	0.056	−0.001	0.065	0.489	7
Savannas	0.038	0.016	0.061	0.827	4
Grasslands	0.06	−0.036	0.069	0.647	236
Croplands	0.076	−0.041	0.094	0.468	80
Crop/natural vegetation mosaic	0.068	−0.008	0.077	0.349	8
Barren/sparse	0.023	0.018	0.036	0.592	6
Average	0.052	−0.028	0.064	0.548	384

this enhancement to the SMAP-Sentinel1 product.

Another ancillary data that has potential to improve the bias and RMSE of the L2_SM_SP product is the soil texture data. The current soil texture data used in the operational processing of the L2_SM_SP product mostly comes from a blend of the Harmonized World Soil Database (HWSD) at ~10–25 km resolution, STATSGO (form Continental United States: CONUS) at 1 km resolution, and Australian Soil Resource Information System (ASRIS) at 1 km resolution. Apart from the CONUS and Australia, the rest of the world has soil data that are very coarse (~10 to 25 km) and outdated. Recent advances in the soil database such as GlobalGrid250m (Hengl et al., 2017) provides very high resolution and better accuracy. Including high-resolution and recent soil texture data in the L2_SM_SP retrieval process will definitely improve the performance at a global perspective/global scales. The impact of coarse resolution soil texture data currently used in the operational procession is not visible in the L2_SM_SP CVS because most of CVS are confined in the area of CONUS and Australia where high resolution soil texture data (STATSGO and ASRIS) is available.

7. Conclusion

The results and validations from the above sections clearly demonstrated the performance of the SMAP-Sentinel1 active-passive algorithm and the capability of this active-passive product to achieve the mission goal by producing high-resolution (1 km or 3 km) soil moisture with good accuracy (ubRMSE ≤ 0.05 m^3/m^3), however, with coarser temporal resolution of 12 days. To achieve high spatial resolution there is a tradeoff between adding spatial resolution with C-band SAR data and noise-levels. The L2_SM_SP high resolution (3 km and 1 km) comes at a cost of degradation in temporal statistics of disaggregated brightness temperature and retrieved soil moisture. Whereas the more spatially-averaged L2_SM_P_E product may have less temporal noise and temporal uncertainty when compared to L2_SM_SP, the L2_SM_SP data has more spatial resolution in term of resolving sharp and large-contrast

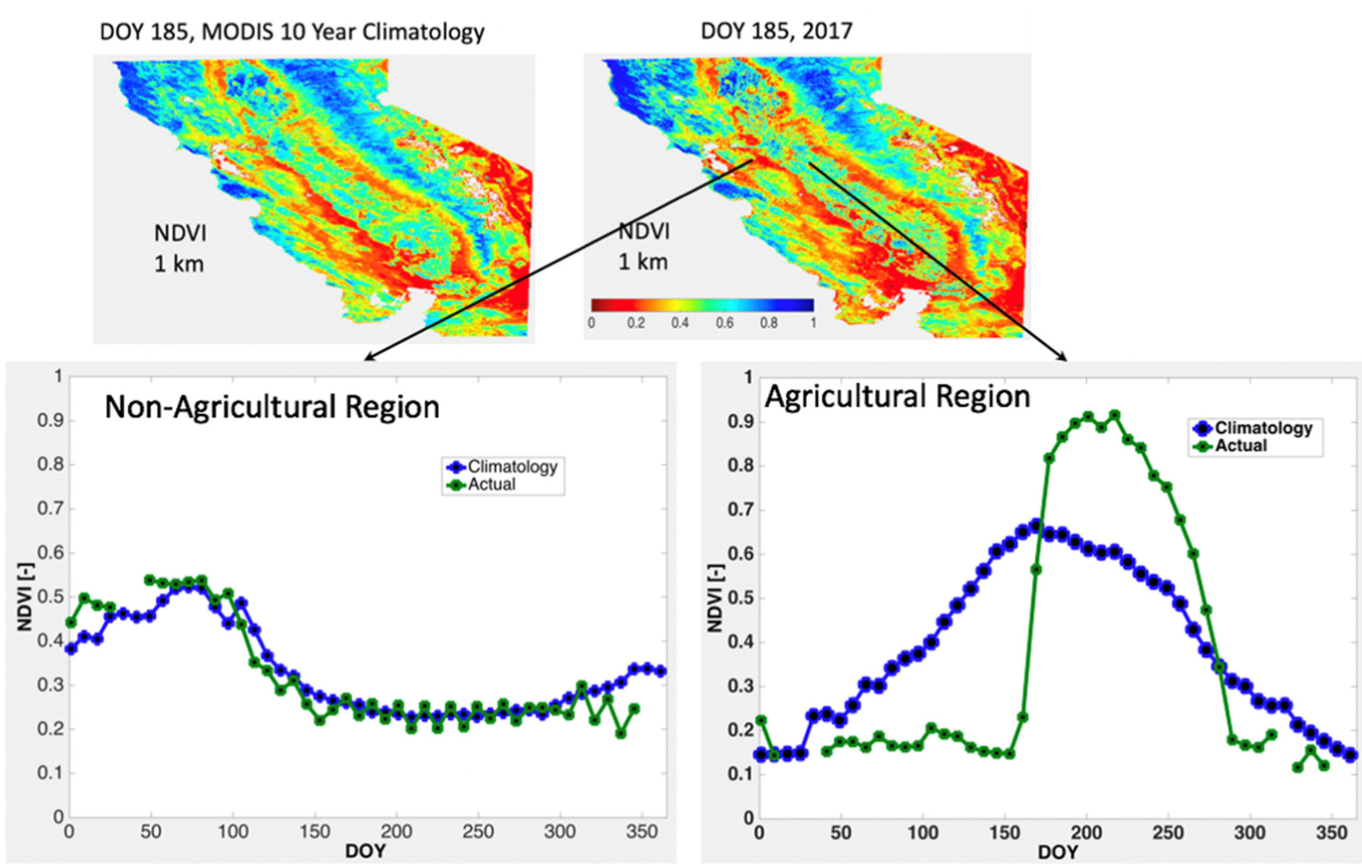

Fig. 22. Comparison of actual NDVI (green curve) and climatology of NDVI (2002–2012) (blue curve) for an agricultural region (cropland) and a non-agricultural region (shrubland) in Central Valley, California. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

features below the radiometer resolution. The degradation in accuracies is mainly due to: 1) difficulties in comprehensively characterizing the SAR signal interactions with target (land surface components), 2) the uncertainties in the parameters used in the SMAP active-passive algorithm, and 3) the random errors and biases in the static and dynamic ancillary data used for soil moisture retrievals. The high resolution L2_SM_SP product captures the spatial details and patterns of soil moisture that are not present in the SMAP radiometer-only enhanced product (L2_SM_P_E). Therefore, those users of SMAP data who require more frequent revisit and temporal accuracy can use the L2_SM_P_E product (which is posted at 9 km), and those users who need higher spatial resolution soil moisture patterns and details with slightly degraded accuracy and less frequent revisit can use L2_SM_SP data (posted at 3 km and 1 km) for their science studies and geophysical applications. The latest version of the L2_SM_SP product is made available to the public and has an ubRMSE of ~0.05 m^3/m^3 at 3 km and 1 km resolutions.

Acknowledgements

The research was carried out at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA). We acknowledge the sustained support from the SMAP Project at JPL, and the Earth Science section of the NASA HQ. We also acknowledge contributions of many staff and personnel who helped in acquiring calibration and validation data that are used for this work. We also acknowledge Massachusetts Institute of Technology (MIT) for supporting this research with the MIT-Germany Seed Fund "Global Water Cycle and Environmental Monitoring using Active and Passive Satellite-based Microwave Instruments".

References

Chan, S., Bindlish, R., O'Neill, P., Njoku, E.G., Jackson, T.J., Colliander, A., Dunbar, S., Chaubell, J., Chen, J. Piepmeier, Yueh, S., Entekhabi, D., Cosh, M., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martínez- Fernández, J., González-Zamora, Á., Seyfried, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J., Palecki, M., Small, E., Calvet, J.-C., Crow, W., Kerr, Y., 2016. Assessment of the SMAP passive soil moisture product. IEEE Trans. Geosci. Remote Sens. 54, 4994–5007.

Chan, S., Bindlish, R., O'Neill, P., Njoku, E.G., Jackson, T.J., Colliander, A., Dunbar, S., Chaubell, J., Chen, J. Piepmeier, Yueh, S., Entekhabi, D., Cosh, M., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martínez- Fernández, J., González-Zamora, Á., Seyfried, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J., Palecki, M., Small, E., Calvet, J.-C., Crow, W., Kerr, Y., 2017. Development and assessment of the enhanced passive soil moisture product. Remote Sens. Environ. 204, 931–941.

Chen, F., Crow, W.T., Colliander, A., Cosh, M., Jackson, T.J., Bindlish, R., Reichle, R., Chan, S.K., Bosch, D.D., Starks, P.S., Goodrich, D.C., Seyfried, M.S., 2017. Application of triple collocation in ground-based validation of soil moisture active/passive (SMAP) level 2 data products. IEEE J. Select. Top. Appl. Earth Observ. Rem. Sens. 10, 489–502.

Colliander, A., Jackson, T.J., Bindlish, R., Chan, S., Das, N.N., Kim, S.B., Cosh, M.H., Dunbar, R.S., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martinez-Fernandez, J., Gonzalez-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramo, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J.P., Wu, X., Monerris, A., O'Neill, P.E., Entekhabi, D., Njoku, E.G., Yueh, S., 2017. Validation of SMAP surface soil moisture products with core validation sites. Remote Sens. Environ. 191, 215–231.

April (Ed.), 2015. Copernicus Sentinel Data 2015. Retrieved from ASF DAAC 29, (processed by ESA).

Das, N.N., Entekhabi, D., Njoku, E., Shi, J., Johnson, J., Colliander, A., 2014. Tests of the SMAP combined radar and radiometer algorithm using airborne field campaign observations and simulated data. IEEE Trans. Geosci. Remote Sens. 52, 2018–2028.

Das, N.N., Entekhabi, D., Dunbar, R.S., Colliander, A., Chan, F., Crow, W., Jackson, T.J., Berg, A., Bosch, D., Caldwell, T., Cosh, M.H., Goodrich, D.C., Lopez-Baeza, E., Moghaddam, M., Rowlandson, T., Starks, P.J., Thibeault, M., Walker, J.P., Wu, X., O'Neill, P.G., Yueh, S., Njoku, E.G., 2018. The SMAP mission combined active-passive soil moisture product at 9 km and 3 km spatial resolutions. Remote Sens. Environ. 211, 204–217.

Dorigo, W., Xaver, A., Vreudgenhil, M., Gruber, A., Dostálová, A., Sanchis-Dufau, A.D., Zamojski, D., Corde, C., Wagner, W., Matthias, D., 2013. Global automated quality control of in situ soil moisture data from the international soil moisture network. Vadose Zone J. 12. https://doi.org/10.2136/vzj2012.0097.

Entekhabi, D., Njoku, E.G., O'Neill, P., Kellogg, K., Crow, W., Edelstein, W., Entin, J., Goodman, S., Jackson, T., Johnson, J., Kimball, J., Peipmeier, J., Koster, R., McDonald, K., Moghaddam, M., Moran, S., Reichle, R., Shi, J., Spencer, M., Thurman, S., 2010. The soil moisture active and passive (SMAP) mission. Proc. IEEE 98, 704–716.

Entekhabi, D., Yueh, S., O'Neill, P., Kellogg, K., et al., July, 2014. SMAP Handbook, JPL Publication JPL 400–1567. Jet Propulsion Laboratory, Pasadena, California 182 pages.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., et al., 2017. SoilGrids250m: global gridded soil information based on machine learning. PLoS One 12 (2), e0169748. https://doi.org/10.1371/journal.pone.0169748.

Jagdhuber, T., Entekhabi, D., Das, N.N., Link, M., Montzka, C., Kim, S., Yueh, S., 2017. Microwave Covariation Modelling and Retrieval for the Dual-Frequency Active-Passive Combination of Sentinel-1 and SMAP. Proc. of IGARSS 2017, July 23–28, Fort Worth, USA. pp. 2017.

Jagdhuber, T., Entekhabi, D., Das, N.N., Link, M., Montzka, C., Baur, M., Akbar, R., Kim, S., Yueh, S., Baris, I., 2018. Physics-Based Modeling of Active-Passive Microwave Covariations for Geophysical Retrievals. vol. 2018. IGARSS, pp. 250–253.

Jagdhuber, T., Baur, M., Akbar, R., Das, N.N., Link, M., He, L., Entekhabi, D., 2019. Estimation of active-passive microwave covariation using SMAP and sentinel-1 data. Remote Sens. Environ. 225, 458–468.

Konings, A.G., Piles, M., Das, N.N., Entekhabi, D., 2017. L-band vegetation optical depth and effective scattering albedo estimation from SMAP. Remote Sens. Environ. 198, 460–470.

Montaldo, N., Albertson, J.D., 2001. Robust simulation of root zone soil moisture with assimilation of surface soil moisture data. Water Resour. Res. 37 (12), 2889–2900.

Njoku, E.G., Entekhabi, D., 1996. Passive microwave remote sensing of soil moisture. J. Hydrol. 184, 101–129.

O'Neill, P.E., Chan, S., Njoku, E., Jackson, T., Bindlish, R., 2014. Algorithm theoretical basis document level 2 & 3 soil moisture (passive) data products. Jet Propuls. Lab. Calif. Inst. Technol (JPL, D-66480). https://smap.jpl.nasa.gov/system/internal_resources/details/original/484_L2_SM_P_ATBD_rev_D_Jun2018.pdf.

Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rudiger, C., Strauss, P., 2018. Sensitivity of Sentinel-1 backscatter to vegetation dynamics: an Austrian case study. Remote Sens. 10, 1–19.

Ye, N., Walker, J.P., Rudiger, C., 2015. A cumulative distribution function method for normalizing variable-angle microwave observations. IEEE Trans. Geosci. Remote Sens. 55 (7), 3906–3916.