NASA/TM–2015-104606/Vol. 40

Technical Report Series on Global Modeling and Data Assimilation, Volume 40

Randal D. Koster, Editor

Soil Moisture Active Passive (SMAP) Project Assessment Report for the Beta-Release L4_SM Data Product

Rolf H. Reichle, Gabrielle J. M. De Lannoy, Qing Liu, Andreas Colliander, Austin Conaty, Thomas Jackson, John Kimball, and Randal D. Koster

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771

October 2015

NASA STI Program ... in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.
- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing help desk and personal search support, and enabling data exchange services. For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov
- E-mail your question via the Internet to help@sti.nasa.gov
- Fax your question to the NASA STI Help Desk at 443-757-5803
- Phone the NASA STI Help Desk at 443-757-5802

NASA/TM–2015-104606/Vol. 40

Technical Report Series on Global Modeling and Data Assimilation, Volume 40

Randal D. Koster, Editor

Soil Moisture Active Passive (SMAP) Project Assessment Report for the Beta-Release L4_SM Data Product

Rolf H. Reichle
NASA's Goddard Space Flight Center, Greenbelt, MD

Gabrielle J. M. De Lannoy
Universities Space Research Association, Columbia, MD

Qing Liu
Science Systems and Applications, Inc., Lanham, MD

Andreas Colliander
Jet Propulsion Laboratory, Caltech, Pasadena, CA

Austin Conaty
Science Systems and Applications, Inc., Lanham, MD

Thomas Jackson
U.S. Department of Agriculture, Agricultural Research Service, Beltsville, MD

John Kimball
University of Montana, Missoula, MT

Randal D. Koster
NASA's Goddard Space Flight Center, Greenbelt, MD

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771

October 2015

Level of Review: This material has been technically reviewed by technical management

TABLE OF CONTENTS

EXECUTIVE SUMMARY

During the post-launch SMAP calibration and validation (Cal/Val) phase there are two objectives for each science data product team: 1) calibrate, verify, and improve the performance of the science algorithm, and 2) validate the accuracy of the science data product as specified in the science requirements and according to the Cal/Val schedule. This report provides an assessment of the SMAP Level 4 Surface and Root Zone Soil Moisture Passive (L4_SM) product specifically for the product's public beta release scheduled for 30 October 2015. The primary objective of the beta release is to allow users to familiarize themselves with the data product before the validated product becomes available. The beta release also allows users to conduct their own assessment of the data and to provide feedback to the L4_SM science data product team.

The assessment of the L4_SM data product includes comparisons of SMAP L4_SM soil moisture estimates with in situ soil moisture observations from core validation sites and sparse networks. The assessment further includes a global evaluation of the internal diagnostics from the ensemble-based data assimilation system that is used to generate the L4_SM product. This evaluation focuses on the statistics of the observation-minus-forecast (O-F) residuals and the analysis increments. Together, the core validation site comparisons and the statistics of the assimilation diagnostics are considered primary validation methodologies for the L4_SM product. Comparisons against in situ measurements from regional-scale sparse networks are considered a secondary validation methodology because such in situ measurements are subject to upscaling errors from the point-scale to the grid cell scale of the data product. Based on the limited set of core validation sites, the assessment presented here meets the criteria established by the Committee on Earth Observing Satellites for Stage 1 validation and supports the beta release of the data. The validation against sparse network measurements and the evaluation of the assimilation diagnostics address Stage 2 validation criteria by expanding the assessment to regional and global scales.

The beta release version of the L4_SM algorithm ingests only the SMAP L1C_TB radiometer brightness temperatures, contrary to the planned use of downscaled brightness temperatures from the L2_SM_AP product and of landscape freeze-thaw state retrievals from the L2_SM_A product. The latter two products are based on radar observations and are only available for the period from 13 April to 7 July 2015 because of an anomaly in the SMAP radar instrument. The decision to use only radiometer (L1C_TB) inputs was made to ensure homogeneity in the longer-term L4_SM beta-release data record.

An analysis of the time average surface and root zone soil moisture shows that the global pattern of arid and humid regions are captured by the L4_SM estimates. The product further reflects major events, including the extremely wet conditions in May 2015 in Texas, Oklahoma, Kansas and parts of the US Midwest. Another event that is captured well by the L4_SM product is the sharp gradient of wet versus dry conditions across Western Australia in May 2015.

Results from the core validation site comparisons indicate that the beta-release version of the L4_SM data product meets the self-imposed L4_SM accuracy requirement, which is formulated in terms of the ubRMSE: the RMSE after removal of the long-term mean difference. The overall ubRMSE of the 3-hourly L4_SM surface soil moisture at the 9 km scale is 0.036 m^3/m^3. The corresponding ubRMSE for L4_SM root zone soil moisture is 0.023 m^3/m^3. Both of these metrics are comfortably below the 0.04 m^3/m^3 requirement. The L4_SM estimates are an improvement over estimates from a model-only SMAP Nature Run version 4 (NRv4), which demonstrates the beneficial impact of the SMAP brightness temperature data. L4_SM surface soil moisture estimates are consistently more skillful than NRv4 estimates, although not by a statistically significant margin. The lack of statistical significance is not surprising given the limited data record available to date. Root zone soil moisture estimates from L4_SM and NRv4 have similar skill. Results from comparisons of the L4_SM product to in situ measurements

from 260 sparse network sites in the United States and Australia corroborated the core validation site results.

The instantaneous soil moisture and soil temperature analysis increments are within a reasonable range and result in spatially smooth soil moisture analyses. The O-F residuals exhibit only small biases on the order of 1-3 K between the (rescaled) SMAP brightness temperature observations and the L4_SM model forecast, which indicates that the assimilation system is largely unbiased. The average (RMS) magnitude of the O-F residuals is 5.8 K, which reduces to 2.6 K for the observation-minus-analysis (O-A) residuals, reflecting the impact of the SMAP observations on the L4_SM system. Averaged globally, the time series standard deviation of the normalized O-F residuals is close to unity, which would suggest that the magnitude of the modeled errors approximately reflects that of the actual errors.

The assessment report also notes several limitations of the beta-release L4_SM data product and science algorithm calibration that will be addressed prior to the release of the validated data product scheduled for summer 2016. Regionally, the time series standard deviation of the normalized O-F residuals deviates considerably from unity, which indicates that the L4_SM assimilation algorithm either over- or underestimates the actual errors that are present in the system. Planned improvements include revised land model parameters, revised error parameters for the land model and the assimilated SMAP observations, and revised surface meteorological forcing data for the operational period and underlying climatological data. Moreover, a refined analysis of the impact of SMAP observations will be facilitated by the construction of additional variants of the model-only reference data. Nevertheless, the beta-release version of the L4_SM product is sufficiently mature and of adequate quality for distribution to and use by the larger science and application communities.

1 INTRODUCTION

The NASA Soil Moisture Active Passive (SMAP) mission provides global measurements of soil moisture from a 685-km, near-polar, sun-synchronous orbit. SMAP data is used to enhance understanding of processes that link the water, energy, and carbon cycles, and to extend the capabilities of weather and climate prediction models (Entekhabi et al. 2014).

The suite of SMAP data products includes the Level 4 Surface and Root Zone Soil Moisture (L4_SM) product, which provides deeper layer soil moisture estimates that are not available in the Level 2-3 products. The L4_SM product is based on the assimilation of SMAP brightness temperatures into the NASA Catchment land surface model (Koster et al. 2000) using a customized version of the Goddard Earth Observing System, version 5 (GEOS-5) land data assimilation system (Figure 1; Reichle et al. 2014a). This system propagates the surface information from the SMAP instrument data to the deeper soil. The latency of the L4_SM product is about 2.5 days and is driven by the availability of the gauge-based global precipitation product that is used to force the land surface model (Reichle et al. 2014b).

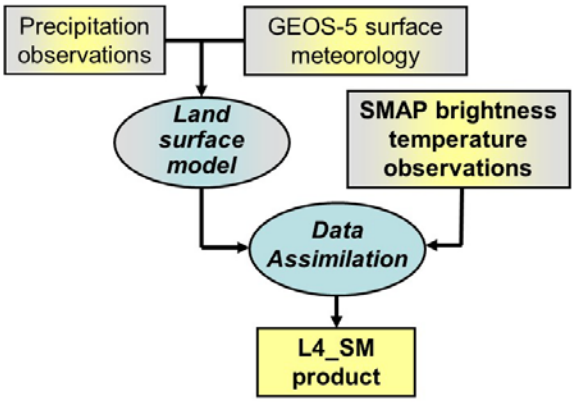

Figure 1. Schematic of the L4_SM algorithm.

The L4_SM product provides surface and root zone soil moisture (along with other geophysical fields) as 3-hourly, time-average fields on the global, cylindrical, 9 km Equal-Area Scalable Earth, version 2 (EASEv2) grid in the "geophysical" (or "gph") output Collection (Reichle et al. 2015). Moreover, instantaneous soil moisture and soil temperature fields before and after the assimilation update are provided every three hours on the 9 km global EASEv2 grid in the "analysis update" (or "aup") output Collection, along with other assimilation diagnostics and error estimates. Time-invariant land model parameters, such as soil porosity, wilting point, and microwave radiative transfer parameters, are provided in the "land-model-constants" (or "lmc") Collection (Reichle et al. 2015).

For geophysical data products that are based on the assimilation of satellite observations into numerical process models, validation is critical and must be based on quantitative estimates of uncertainty. Direct comparison with independent observations, including ground-based measurements, is a key part of the validation. This Assessment Report provides a detailed description of the L4_SM validation process and the status of the L4_SM data quality prior to the beta release of the L4_SM data product.

2 SMAP CALIBRATION AND VALIDATION OBJECTIVES

During the post-launch SMAP calibration and validation (Cal/Val) phase each science product team pursues two objectives:

1. Calibrate, verify, and improve the performance of the science algorithm.
2. Validate the accuracy of the science data product as specified in the science requirements and according to the Cal/Val schedule.

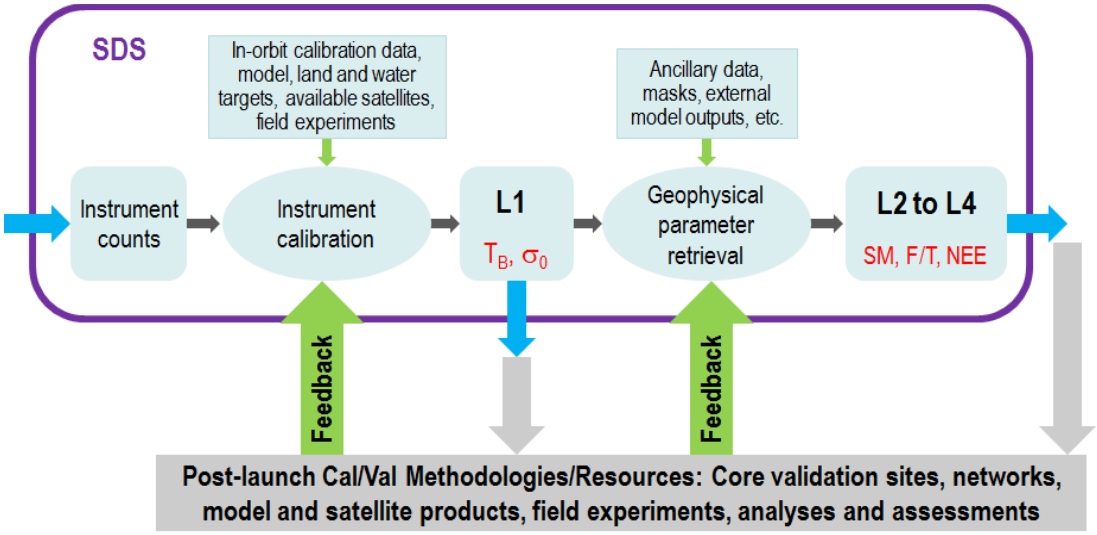

Figure 2. Overview of the SMAP calibration and validation process.

The overall SMAP Cal/Val process is illustrated in Figure 2. This Assessment Report describes how the L4_SM team addressed the above objectives prior to the beta release. The validation approach and procedures follow those described in the SMAP Science Data Cal/Val Plan (Jackson et al. 2014), the SMAP L2-L4 Data Products Cal/Val Plan (Colliander et al. 2014), and the Algorithm Theoretical Basis Document for the L4_SM data product (Reichle et al. 2014b).

SMAP established unified definitions to address the mission requirements. These are documented in the SMAP Handbook (Entekhabi et al. 2014), where calibration and validation are defined as follows:

- *Calibration*: The set of operations that establish, under specified conditions, the relationship between sets of values or quantities indicated by a measuring instrument or measuring system and the corresponding values realized by standards.
- *Validation:* The process of assessing by independent means the quality of the data products derived from the system outputs.

In order to insure the public's timely access to SMAP data, the mission is required to release beta-quality products before releasing validated products. The objectives and maturity of the beta release products are defined as follows:

- The beta release allows users to gain familiarity with data formats.
- The beta release is intended as a testbed to discover and correct errors.
- Beta-release data are minimally validated and still may contain significant errors.

- The general research community is encouraged to participate in the quality assessment and validation, but need to be aware that product validation and quality assessment are ongoing.
- Beta-release data may be used in publications as long as the authors indicate the fact that the data are beta quality. **Drawing quantitative scientific conclusions is discouraged.** Users are urged to contact of the SMAP Science Team prior to using beta-release data in publications, and to recommend members of the instrument teams as reviewers.
- The estimated uncertainties are documented.
- Beta-release data may be replaced in the archive when an upgraded (provisional or validated) product becomes available.

Due to the high quality of the SMAP L1C_TB brightness temperatures and the heritage and maturity of model-based soil moisture data products (Reichle et al. 2011), the beta release version of the L4_SM product already exceeds the above maturity requirements and is closer to a provisional release, which is defined as follows:

- Incremental improvements are ongoing. Obvious artifacts or errors observed in the provisional product have been identified and either minimized or documented.
- The general research community is encouraged to participate in the quality assessment and validation, but need to be aware that product validation and quality assessment are ongoing.
- Provisional-release data may be used in publications as long as the authors indicate the provisional quality. Users are urged to contact science team representatives prior to using the data in publications, and to recommend members of the instrument teams as reviewers.
- The estimated uncertainties are documented.
- Provisional-release data will be replaced in the archive when an upgraded (validated) product becomes available.

In assessing the maturity of the L4_SM product, the L4_SM team also considered the guidance provided by the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (CEOS 2015):

- Stage 1: Product accuracy is assessed from a small (typically < 30) set of locations and time periods by comparison with *in situ* or other suitable reference data.
- Stage 2: Product accuracy is estimated over a significant set of locations and time periods by comparison with reference *in situ* or other suitable reference data. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and time periods. Results are published in the peer-reviewed literature.
- Stage 3: Uncertainties in the product and its associated structure are well quantified from comparison with reference *in situ* or other suitable reference data. Uncertainties are characterized in a statistically robust way over multiple locations and time periods representing global conditions. Spatial and temporal consistency of the product and with similar products has been evaluated over globally representative locations and periods. Results are published in the peer-reviewed literature.
- Stage 4: Validation results for stage 3 are systematically updated when new product versions are released and as the time-series expands.

For the beta release the L4_SM team has completed Stage 1 and begun Stage 2 (global assessment). The Cal/Val program will continue through the above Stages over the SMAP mission life span.

3 L4_SM CALIBRATION AND VALIDATION APPROACH

During the mission definition and development phase, the SMAP Science Team and Cal/Val Working Group identified the metrics and methodologies that would be used for L2-L4 product assessment. These metrics and methodologies were vetted in community Cal/Val Workshops and tested in SMAP pre-launch Cal/Val rehearsal campaigns. The following validation methodologies and their general roles in the SMAP Cal/Val process were identified:

- *Core Validation Sites*: Accurate estimates at matching scales for a limited set of conditions.
- *Sparse Networks*: One point in the grid cell for a wide range of conditions.
- *Satellite Products*: Estimates over a very wide range of conditions at matching scales.
- *Model Products*: Estimates over a very wide range of conditions at matching scales.
- *Field Campaigns*: Detailed estimates for a very limited set of conditions.

With regard to the CEOS Cal/Val stages, core validation sites address Stage 1, and satellite and model products are used for Stage 2 and beyond. Sparse networks fall between these two Stages.

For the L4_SM data product, all of the above methodologies can contribute to product assessment and refinement, but there are differences in terms of the importance of each approach for the validation of the L4_SM product.

The assessment of the L4_SM data product includes comparisons of SMAP L4_SM soil moisture estimates with *in situ* soil moisture observations from core validation sites (CVS) and sparse networks. The assessment further includes a global evaluation of the internal diagnostics from the ensemble-based data assimilation system that is used to generate the L4_SM product. This evaluation focuses on the statistics of the observation-minus-forecast (O-F) residuals and the analysis increments. Together, the CVS comparisons and the statistics of the assimilation diagnostics are considered primary validation methodologies for the L4_SM product.

Comparisons against in situ measurements from regional-scale sparse networks are considered a secondary validation methodology because such in situ measurements are subject to upscaling errors from the point-scale to the grid cell scale of the data product.

Due to their very limited spatial and temporal extent, data from field campaigns play only a tertiary role in the validation of the L4_SM data product. Note, however, that field campaigns are instrumental tools in the provision of high-quality, automated observations from the core validation sites and thus play an important indirect role in the validation of the L4_SM data product.

Based on the limited set of core validation sites, the assessment presented here meets the criteria established by CEOS for Stage 1 validation and supports the beta release of the L4_SM data. The validation against sparse network measurements and the evaluation of the assimilation diagnostics address Stage 2 validation criteria by expanding the assessment to regional and global scales and suggest that the beta-release version of the L4_SM product is closer to a provisional data release (section 2).

4 L4_SM ACCURACY REQUIREMENT

There is no formal Level 1 mission requirement for the validation of the L4_SM product, but the L4_SM team self-imposed an accuracy requirement mirroring the one that applies to the L2_SM_AP product. Specifically, the L4_SM surface and root zone soil moisture estimates are required to meet the following criterion:

> **ubRMSE $\leq$ 0.04 m^3 m^{-3}** within the data masks specified in the *SMAP Level 2 Science Requirements* (that is, excluding regions of snow and ice, frozen ground, mountainous topography, open water, urban areas, and vegetation with water content greater than 5 kg m^{-2}),

where ubRMSE is the RMSE computed after removing long-term mean bias from the data (Appendix A). (The ubRMSE is also referred to as the standard deviation of the error.) This criterion applies to the L4_SM instantaneous surface and root zone soil moisture estimates at the 9 km grid-cell scale from the "aup" Collection. It is verified by comparing the L4_SM product to the grid-cell scale in situ measurements from the core validation sites (section 6.2).

L4_SM output fields other than instantaneous surface and root zone soil moisture are provided as research products (including surface meteorological forcing variables, soil temperature, evaporative fraction, net radiation, etc.) and will be evaluated against in situ observations to the extent possible given available resources.

As part of the validation process, additional metrics (including bias, RMSE, time series correlation coefficient R, and anomaly R values) will be computed for the L4_SM output fields to the fullest extent possible. This includes computation of the metrics outside of the limited geographic area for which the 0.04 $m^3 m^{-3}$ validation criterion is applied.

For the computation of the *anomaly* R metric, the seasonal cycle of the raw data (including the L4_SM product and the in situ measurements) is estimated, separately for each product and each location, by computing, for each day of the year (DOY), a climatological value of soil moisture. Anomaly time series are then computed by subtracting the mean seasonal cycle from the raw data. Lastly, the anomaly R metric is derived by computing the time series correlation coefficient of the anomaly time series. Because of the short data record of the beta-release, anomaly R metrics are not provided in this report.

The validation includes additional metrics that are based on the statistics of the observation-minus-forecast residuals and other data assimilation diagnostics (section 6.4). The appendix provides detailed definitions of all the validation metrics used here.

5 L4_SM BETA RELEASE

5.1 Process and Criteria

Since the beginning of the science data flow, the team has been conducting frequent assessments of pre-beta L4_SM data products and will continue to do so throughout the intensive Cal/Val phase and beyond. Frequent reviews of performance based upon core validation sites, sparse networks, and assimilation diagnostics were conducted for a period of 5+ months and captured a wide range of geophysical conditions. The assessment presented here is a summary of the latest status of this process.

The comparison against in situ measurements include metrics for a model-only "SMAP Nature Run," version 4 (NRv4). The NRv4 estimates are based on the same land surface model and forcing data as the L4_SM estimates, except that the NRv4 estimates do not benefit from the assimilation of the SMAP brightness temperature observations. Specifically, the NRv4 estimates are the result of a single-member, land model integration within the L4_SM system but without the ensemble perturbations and without the assimilation of the SMAP L1C_TB observations; any accuracy in the NRv4 estimates is thus derived from the imposed meteorological forcing and land model structure and parameter information. The NRv4 estimates are available for the period 1 January 2001 to present and also provide the model climatological information required by the L4_SM assimilation algorithm (Reichle et al. 2014b).

Prior to the SMAP launch, the L4_SM team formulated the following criteria for the beta release of the L4_SM data:

- The skill of the L4_SM assimilation product is no worse than that of the model-only reference (SMAP Nature Run, NRv4), where skill is measured against in situ observations from core validation sites and sparse networks.
- The statistics of observations-minus-forecast residuals are within expectation as defined by published results (AMSR-E, ASCAT, SMOS).

One key finding of this Assessment Report is that the above criteria have been met, and that the L4_SM product is sufficiently mature for a public beta release on 30 October 2015.

5.2 Processing Options and Science ID Version

The L4_SM product version used to prepare this Assessment Report has Science Version ID **Vb1004**. The data were generated in mid-September 2015 as a "forward processing parallel" (FPP) stream (ECS Version ID 199). The L4_SM algorithm slated for the beta release on 30 October 2015 is expected to have only very minor differences from Vb1004.

In anticipation of the planned L4_SM beta release on 30 October 2015, the L4_SM team defined the assessment period for this report as **11 April 2015, 0z to 19 September 2015, 0z**. The start date matches the date when the radiometer was operating under reasonably stable conditions following instrument start-up operations. The end date was selected to allow sufficient time for analysis and preparation of this Assessment Report as well as other documents required for the beta release.

The beta release version of the L4_SM algorithm ingests only the SMAP L1C_TB radiometer brightness temperatures, contrary to the planned use of downscaled brightness temperatures from the L2_SM_AP product and landscape freeze-thaw state retrievals from the L2_SM_A product. The latter two products are based on radar observations and are only available for the period from 13 April to 7 July

2015 because of an anomaly in the SMAP radar instrument. The decision to use only radiometer (L1C_TB) inputs was made to ensure homogeneity in the longer-term L4_SM beta-release data record.

The L4_SM Vb1004 algorithm used the beta-release version of the SMAP L1C_TB brightness temperature data product (CRID R11850 until 9 Sep 2015 and CRID R11920 thereafter), which meets its requirements for noise equivalent differential temperature (NEDT) and geolocation and is generally considered of good quality. A detailed report of the quality of the beta-release version of the L1C_TB data product is available from NSIDC (Piepmeier et al. 2015).

6 L4_SM DATA PRODUCT ASSESSMENT

This section provides a detailed assessment of the L4_SM data product. First, global patterns, features, and noteworthy events are discussed (section 6.1). Next, we present comparisons and metrics versus in situ measurements from core validation sites (section 6.2) and sparse networks (section 6.3). Finally, we evaluate the assimilation diagnostics (section 6.4), which includes a discussion of the observation-minus-forecast residuals, the increments, and the data product uncertainty estimates.

6.1 Global Patterns and Features

Figure 3 shows global maps of time-averaged L4_SM surface and root zone soil moisture for the validation period (11 April 2015, 0z to 19 September 2015, 0z). The global patterns are as expected – arid regions such as the southwestern US, the Sahara desert, the Arabian Peninsula, the Middle East, southern Africa, and central Australia exhibit generally dry surface and root zone soil moisture conditions, whereas the tropics (Amazon, central Africa, and Indonesia) and high-latitude regions show wetter conditions. One notable exception is that a portion of the Democratic Republic of Congo and adjacent areas appear unexpectedly dry. This is because over Africa, the beta-release version of the L4_SM algorithm uses precipitation forcing directly from the GEOS-5 Forward Processing (FP) system, which has a known dry bias in central Africa.

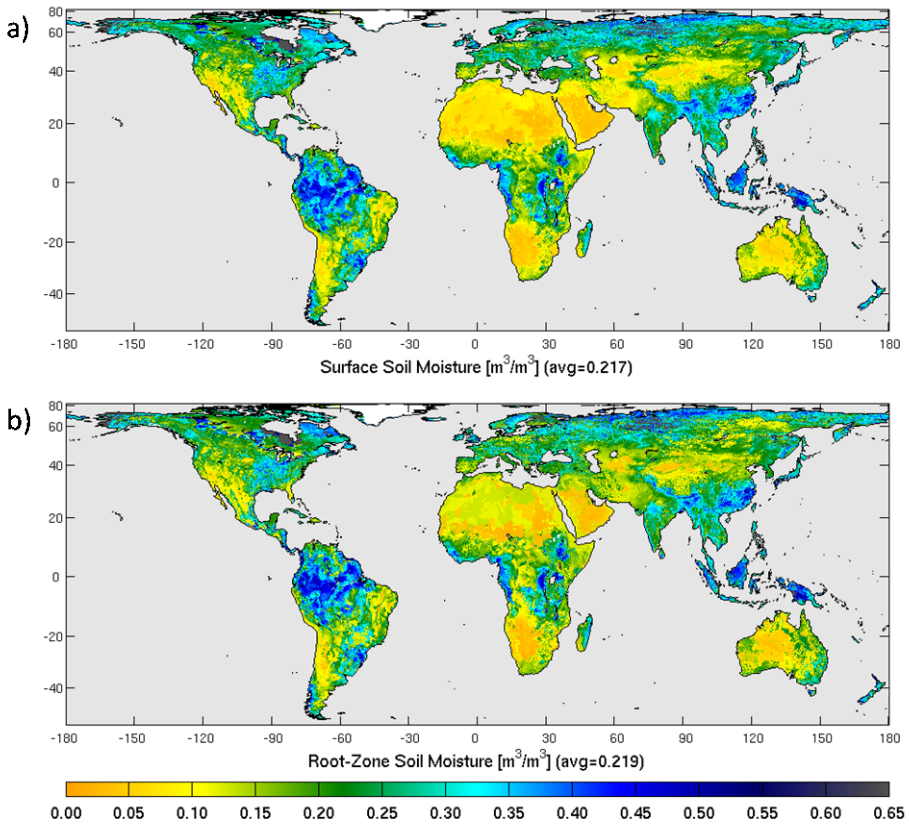

Figure 3. Time-average of L4_SM (a) surface and (b) root zone soil moisture for 11 April 2015, 0z, to 19 September 2015, 0z.

Generally, the global patterns of absolute soil moisture values are dominated by soil parameters and climatological factors. The influence of soil texture is noticeable in the coarse-scale patterns in the Sahara desert, where little is known about the spatial distribution of mineral soil fractions. Areas with peat soil include, for example, the region along the southern edge of Hudson Bay and portions of Alaska. In the land model, the soils in this region are assigned a high porosity value and show persistently wetter conditions than other areas.

The strong impact of climate on soil moisture is reflected in the overall similarity between the time-averaged fields of Figure 3 and the instantaneous fields for 1 June 2015 and 1 September 2015, both at 0z, shown in Figure 4. Some areas, however, do exhibit strong changes in soil moisture conditions between the two dates. For example, the extremely wet conditions on 1 June 2015 in Texas, Oklahoma, and Kansas and extending into the US Midwest (Figure 4a and c) resulted from well-documented extreme rainfall events throughout May 2015. The wet conditions have clearly abated by 1 September 2015 (Figure 4b and d).

Another notable feature in the global maps of Figure 4 is the strong spatial contrast in dry and wet soil moisture conditions in western Australia on 1 June 2015. This contrast resulted from parts of the region having seen unseasonably high rainfall conditions in May 2015, with a few locations recording their wettest May on record, and many locations recording their wettest May for over twenty years. In contrast, the rest of Western Australia recorded rainfall that was below to very much below average (Bureau of Meteorology 2015; http://www.bom.gov.au/climate/current).

The L4_SM product also includes a large number of output fields that are not subject to formal validation requirements. Such "research" output includes the surface meteorological forcing fields, land surface fluxes, soil temperature and snow conditions, runoff, and error estimates (derived from the ensemble). Figure 5 illustrates two of these fields for 11 April 2015: the top-layer soil temperature (at 12z) and the snow mass (3-hour average for 21z to 0z). The global patterns are again consistent with expectation. The top-layer soil temperature is hottest in the Arabian Peninsula, where the local time is around 3pm and the diurnal cycle of the soil temperature is at or near its peak. The coolest temperatures can be found at the highest northern latitudes, where the soil is still frozen. The snow mass distribution is also consistent with expectation, with nearly continuous snow cover remaining at the highest latitudes and in the northern hemisphere high mountain ranges. Snow is all but absent in the southern hemisphere at the end of the austral summer.

The temperature and snow fields shown in Figure 5 are of particular interest because they can be used to screen or flag the L4_SM soil moisture output for frozen conditions or situations where there is snow on the ground. Unlike the SMAP Level 2 and 3 retrieval products, the L4_SM product does not provide binary flags to classify the conditions at the time for which the soil moisture estimates are valid. Rather, the L4_SM product provides quantitative estimates of surface and soil temperatures, snow mass, precipitation, etc. that contain far more complete information and can readily be converted into binary flags by users should the need arise.

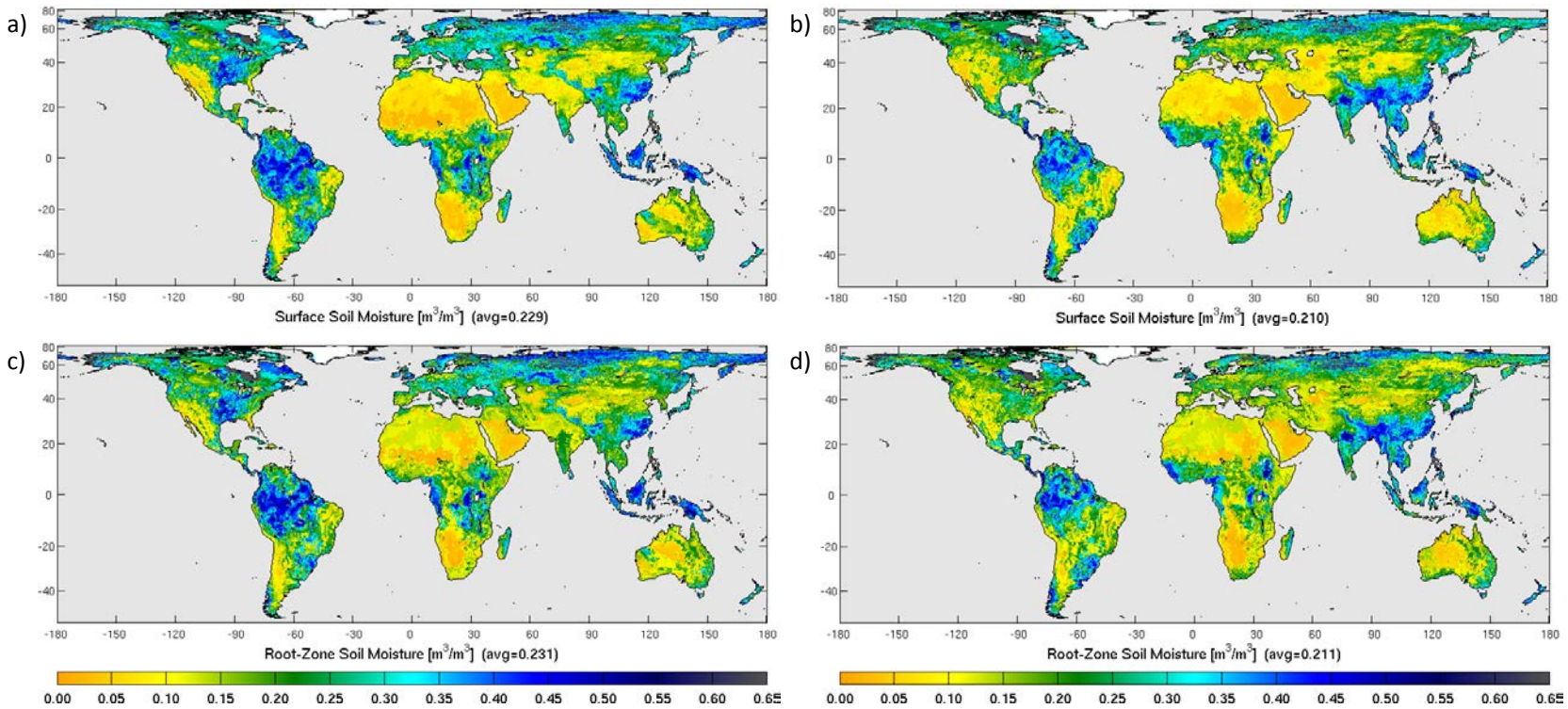

Figure 4. Snapshots of L4_SM (a,b) surface and (c,d) root zone soil moisture for (a,c) 1 June 2015, 0z, and (b,d) 1 September 2015, 0z.

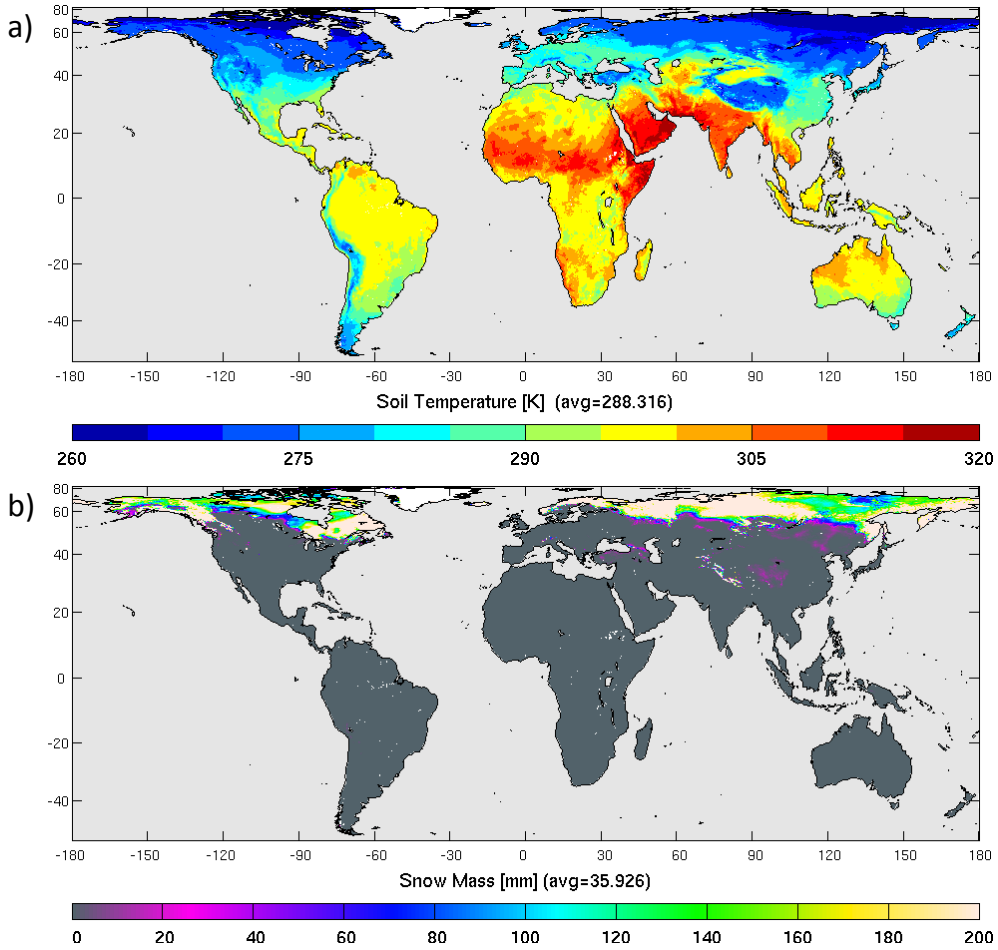

Figure 5. L4_SM (a) top layer soil temperature on 11 April 2015, 12z and (b) snow mass (three-hourly average from 11 April 2015, 21z to 12 April 2015, 0z).

6.2 Core Validation Sites

6.2.1 Method and Overview

In situ measurements are critical in the assessment of the SMAP data products. Comparisons against in situ measurements provide error estimates and a basis for modifying algorithms and/or parameters. A robust analysis requires many sites representing diverse conditions. Unfortunately, there are relatively few sites that can provide measurements of the type, quality, and quantity needed for robust validation. The core validation sites used here are the result of the Cal/Val Partner Program established by SMAP to foster cooperation with the providers of in situ soil moisture measurements and to encourage the enhancement of these resources for the support of the SMAP Cal/Val program.

For any given core validation site, the spatial distribution of the in situ sensors is typically not aligned with the grid cells of the standard EASE v2 grid of the SMAP data products. Therefore, the SMAP Cal/Val team defined custom "shifted" grid cells (or "reference pixels") that better exploit the spatial coverage of the in situ measurements at each site, but do not necessarily align with the standard EASE v2 grid. A core validation site may provide in situ measurements for one or more 9 km and/or 36 km reference pixels.

The formal L4_SM accuracy requirement is expressed in terms of the ubRMSE metric and applies to the 9 km native resolution of the data product (section 4). To maximize the use of the core validation site measurements, this section assesses the L4_SM product in terms of a variety of metrics (section 4) and for all available 9 km and 36 km core validation site pixels. The formal requirement is verified by averaging only the ubRMSE values from the 9 km reference pixels at the core validation sites.

The L4_SM assessment uses the in situ measurements at reference pixels that may not align with standard EASE v2 grid cells (Appendix C). But unlike the Level 2 algorithms, the L4_SM algorithm cannot readily be applied to reference pixels. Therefore, L4_SM estimates are co-located with the reference pixels through bilinearly interpolating the L4_SM product estimates from their standard, 9 km EASE v2 grid to the reference pixels for the purpose of the in situ comparisons.

Some sites may not offer the desired spatial distribution or quality of measurement. These sites, classified as "candidate" validation sites, have not reached the level of maturity needed for use in the verification of the formal accuracy requirement (section 4). In some cases this is simply a latency problem that will be resolved in time. Prior to launch, the L4_SM and Cal/Val teams reviewed the status of all sites to determine which sites are ready to be designated as core validation sites. The basic process was as follows:

- Define the reference pixel (or grid cell, possibly shifted w.r.t. the standard EASE v2 grid).
- Assess the site for conditions that would introduce uncertainty.
- Determine if the number of sensor locations is large enough to provide reliable grid-cell average estimates.
- Assess the geographic distribution of the sensors.
- Determine if the instrumentation has been either (1) widely used and known to be well-calibrated or (2) calibrated for the specific site in question.
- Perform a quality assessment of each sensor in the network.
- Establish a scaling function. (The default function is a straight average of all sensors.)
- Conduct a pre-launch assessment using surrogate data appropriate for the L4_SM product (e.g., the SMAP Nature Run).

- Review any supplemental studies that have been performed to verify that the site network represents the average soil moisture conditions over the reference pixel (or grid cell).

The status of candidate validation sites will be reviewed periodically to determine if they should be reclassified as core validation sites. The current set of core validation sites that provide data for the broader L4_SM assessment are listed in Table 1, along with the details of the 9 km and 36 km reference pixels. The table shows that the L4_SM validation is based on a total of 27 reference pixels from 12 different core validation sites. Surface soil moisture measurements are available for all 27 reference pixels, which include 10 reference pixels at the 36 km scale from 10 different sites and 17 reference pixels at the 9 km scale from 12 different sites. For root zone soil moisture, measurements are available for only 14 reference pixels from 6 different core sites, including 6 reference pixels at the 36 km scale from 6 different sites and 8 reference pixels at the 9 km scale from just 4 different sites. The 9 km reference pixels for root zone soil moisture belong to the core validation sites of Fort Cobb (Oklahoma), South Fork (Iowa), Kenaston (Saskatchewan), and TxSON (Texas). This very limited set obviously lacks the diversity to be representative of global conditions.

Table 1 also lists the depths of the deepest sensors that contribute to the in situ root zone soil moisture measurements. The measurements from the individual sensors are vertically averaged with weights that are proportional to the spacing of the depth of the sensors within the 0-100 cm layer depth corresponding to the L4_SM estimates. At all reference pixels except Little River, the deepest sensors are at 45 cm or 50 cm depth. At Little River, the deepest sensors are at 30 cm. The deepest sensors are therefore weighted most strongly in the computation of the vertical average. To compute the vertically averaged root zone soil moisture at a given time from a given sensor profile, all sensors within the profile must provide measurements that pass the automated quality control.

Across the reference pixels listed in Table 1, the average number of individual sensors that contribute to a given 36 km reference pixel ranges between 12.7 and 33.5, with a mean value of 21.3. At the 9 km scale, 7 of the 17 reference pixels are based on just 4 individual sensor profiles, while the rest of the 9 km reference pixels consist of about 10 sensor profiles each. The relative sampling density is therefore considerably lower for the 9 km reference pixels. For most reference pixels, individual sensor profiles tend to drop out temporarily. This leads to undesirable discontinuities in the reference pixel average soil moisture. To mitigate this effect, a minimum of 8 individual sensor profiles were required (after quality control) to compute the reference pixel average, provided at least 8 sensor profiles were in the ground. For the seven reference pixels that are based on just 4 sensor profiles, data from all 4 sensors (after quality control) was required to compute the reference pixel average.

Table 2 lists the various skill metrics for all reference pixels, both for the L4_SM product and the Nature Run v4 (NRv4; section 5.1). The table is primarily provided for reference. A detailed discussion of the skill at select core validation sites is given in sections 6.2.2-6.2.5. The results for individual reference pixels reveal many features that support the quality of the L4_SM data product and indicate potential avenues for improvement. This is followed in section 6.2.6 by a discussion of the summary metrics obtained from averaging across the skill computed at all reference pixels, along with general conclusions from the core site validation.

Table 1. Core validation sites and reference pixels for L4_SM validation. 36 km reference pixels are shown with gray shading.

Site Name	Country	Climate Regime	Land Cover	Reference Pixel							
				ID	Latitude [degree]	Longitude [degree]	Horizontal Scale [km]	Depth of Deepest Sensor [m]	Number of Sensor Profiles		
									Min	Mean	Max
REMEDHUS	Spain	Temperate	Cropland	03013602	41.28	-5.41	36	0.05	13	16.9	20
				03010908	41.32	-5.27	9	0.05	4	4.0	4
Yanco	Australia (New South Wales)	Arid	Cropland / natural mosaic	07013601	-34.85	146.17	36	0.05	9	25.5	29
				07010902	-34.72	146.13	9	0.05	8	10.7	11
Carman	Canada (Manitoba)	Cold	Cropland	09013601	49.61	-97.94	36	0.05	8	19.7	20
				09010906	49.67	-97.98	9	0.05	8	10.9	11
Walnut Gulch	USA (Arizona)	Arid	Shrub open	16013603	31.68	-110.04	36	0.05	12	20.7	24
				16010906	31.72	-110.09	9	0.05	8	8.7	9
				16010907	31.72	-109.99	9	0.05	8	10.6	11
Little Washita	USA (Oklahoma)	Temperate	Grassland	16023602	34.88	-98.09	36	0.45	8	17.0	18
				16020907	34.92	-98.04	9	0.05	4	4.0	4
Fort Cobb	USA (Oklahoma)	Temperate	Grassland	16033602	35.42	-98.62	36	0.45	10	12.7	13
				16030911	35.38	-98.57	9	0.45	4	4.0	4
				16030916	35.29	-98.48	9	0.45	4	4.0	4
Little River	USA (Georgia)	Temperate	Cropland / natural mosaic	16043602	31.60	-83.59	36	0.30	12	19.1	21
				16040901	31.72	-83.73	9	0.05	8	8.0	8
St Josephs	USA (Indiana)	Temperate	Cropland	16060907	41.45	-84.97	9	0.05	8	8.3	9
South Fork	USA (Iowa)	Cold	Cropland	16073602	42.47	-93.39	36	0.50	10	17.1	20
				16070909	42.42	-93.53	9	0.50	4	4.0	4
				16070910	42.42	-93.44	9	0.50	4	4.0	4
				16070911	42.42	-93.35	9	0.50	4	4.0	4
Kenaston	Canada (Saskatchewan)	Cold	Cropland	27013601	51.45	-106.46	36	0.50	8	30.9	35
				27010911	51.39	-106.42	9	0.50	8	12.8	14
Valencia	Spain	Cold	Savanna woody	41010906	39.57	-1.26	9	0.05	8	8.0	8
TxSON	USA (Texas)	Temperate	Grassland	48013601	30.31	-98.78	36	0.50	28	33.5	36
				48010902	30.43	-98.82	9	0.50	8	10.3	11
				48010911	30.27	-98.73	9	0.50	13	14.3	15

Table 2. Metrics at individual reference pixels. 36 km reference pixels are shown with gray shading. L4_SM metrics are shown in bold.

	Reference Pixel		Surface Soil Moisture									Root Zone Soil Moisture								
			ubRMSE [m^3m^{-3}]			Bias [m^3m^{-3}]			R [-]			ubRMSE [m^3m^{-3}]			Bias [m^3m^{-3}]			R [-]		
Site Name	ID	Horiz. Scale [km]	NRv4	**L4_SM Vb1004**	95% Conf. Interval	NRv4	**L4_SM Vb1004**	95% Conf. Interval	NRv4	**L4_SM Vb1004**	95% Conf. Interval	NRv4	**L4_SM Vb1004**	95% Conf. Interval	NRv4	**L4_SM Vb1004**	95% Conf. Interval	NRv4	**L4_SM Vb1004**	95% Conf. Interval
REMEDHUS	03013602	36	0.025	**0.020**	0.015	0.048	**0.050**	0.008	0.62	**0.73**	0.19	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	03010908	9	0.035	**0.032**	0.008	0.014	**0.017**	0.009	0.55	**0.62**	0.20	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Yanco	07013601	36	0.050	**0.034**	0.026	-0.001	**0.030**	0.044	0.87	**0.91**	0.14	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	07010902	9	0.078	**0.052**	0.029	-0.023	**0.010**	0.068	0.90	**0.93**	0.10	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Carman	09013601	36	0.022	**0.039**	0.007	-0.017	**-0.015**	0.008	0.70	**0.21**	0.17	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	09010906	9	0.025	**0.037**	0.013	0.047	**0.054**	0.009	0.63	**0.25**	0.19	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Walnut Gulch	16013603	36	0.026	**0.024**	0.015	0.053	**0.044**	0.008	0.68	**0.76**	0.14	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	16010906	9	0.031	**0.033**	0.011	0.031	**0.030**	0.007	0.59	**0.60**	0.16	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	16010907	9	0.028	**0.029**	0.012	0.037	**0.030**	0.007	0.60	**0.63**	0.15	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Little Washita	16023602	36	0.037	**0.029**	0.011	0.009	**-0.012**	0.016	0.84	**0.90**	0.07	0.033	**0.020**	0.025	-0.034	**-0.048**	0.020	0.92	**0.95**	0.24
	16020907	9	0.036	**0.030**	0.013	-0.009	**-0.029**	0.015	0.81	**0.89**	0.09	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Fort Cobb	16033602	36	0.043	**0.036**	0.016	0.035	**0.032**	0.016	0.81	**0.86**	0.09	0.025	**0.028**	0.020	0.025	**0.028**	0.024	0.90	**0.88**	0.42
	16030911	9	0.051	**0.038**	0.023	0.044	**0.045**	0.025	0.81	**0.88**	0.13	0.025	**0.029**	0.016	0.026	**0.038**	0.017	0.82	**0.94**	0.26
	16030916	9	0.034	**0.032**	0.010	-0.002	**-0.019**	0.012	0.82	**0.86**	0.08	0.020	**0.031**	0.018	-0.033	**-0.030**	0.015	0.91	**0.76**	0.26
Little River	16043602	36	0.044	**0.036**	0.019	0.086	**0.078**	0.011	0.43	**0.60**	0.23	0.028	**0.022**	0.016	0.054	**0.047**	0.009	0.72	**0.79**	0.23
	16040901	9	0.042	**0.037**	0.029	0.095	**0.092**	0.016	0.67	**0.72**	0.25	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
St Josephs	16060907	9	0.028	**0.039**	0.033	0.149	**0.131**	0.012	0.85	**0.72**	0.16	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
South Fork	16073602	36	0.037	**0.039**	0.014	0.049	**0.038**	0.013	0.65	**0.58**	0.13	0.014	**0.034**	0.008	-0.007	**-0.017**	0.010	0.72	**0.30**	0.36
	16070909	9	0.041	**0.029**	0.010	-0.014	**-0.016**	0.014	0.52	**0.80**	0.12	0.016	**0.016**	0.022	-0.074	**-0.074**	0.008	0.57	**0.83**	0.29
	16070910	9	0.044	**0.036**	0.018	0.036	**0.039**	0.018	0.54	**0.70**	0.15	0.012	**0.023**	0.011	0.018	**0.025**	0.009	0.73	**0.61**	0.35
	16070911	9	0.045	**0.038**	0.010	0.010	**0.016**	0.015	0.58	**0.71**	0.13	0.012	**0.018**	0.005	-0.010	**-0.002**	0.006	0.68	**0.78**	0.26
Kenaston	27013601	36	0.026	**0.020**	0.012	0.005	**-0.005**	0.013	0.61	**0.81**	0.11	0.017	**0.015**	0.043	-0.056	**-0.066**	0.025	0.82	**0.83**	0.65
	27010911	9	0.030	**0.032**	0.010	0.000	**-0.010**	0.011	0.52	**0.55**	0.16	0.022	**0.018**	0.033	-0.089	**-0.099**	0.013	0.543	**0.755**	0.49
Valencia	41010906	9	0.024	**0.023**	0.023	0.093	**0.085**	0.009	0.51	**0.53**	0.22	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
TxSON	48013601	36	0.043	**0.032**	0.039	0.097	**0.078**	0.024	0.82	**0.93**	0.21	0.033	**0.027**	0.142	0.062	**0.056**	0.268	0.95	**0.93**	0.50
	48010902	9	0.044	**0.044**	0.044	0.140	**0.131**	0.022	0.74	**0.82**	0.24	0.028	**0.016**	0.060	0.099	**0.100**	0.035	0.94	**0.96**	0.50
	48010911	9	0.057	**0.045**	0.046	0.135	**0.112**	0.025	0.75	**0.87**	0.25	0.031	**0.030**	0.086	0.094	**0.083**	0.075	0.90	**0.90**	0.50

6.2.2 Little Washita (Oklahoma)

The Little Washita watershed in Oklahoma has been utilized for many validation studies of microwave soil moisture retrievals. Several in situ measurement campaigns addressed both in situ sensor calibration and upscaling. Therefore, confidence in the quality of the in situ estimates is very high for this site, and performance at this site is considered to be an important factor in the L4_SM algorithm assessment.

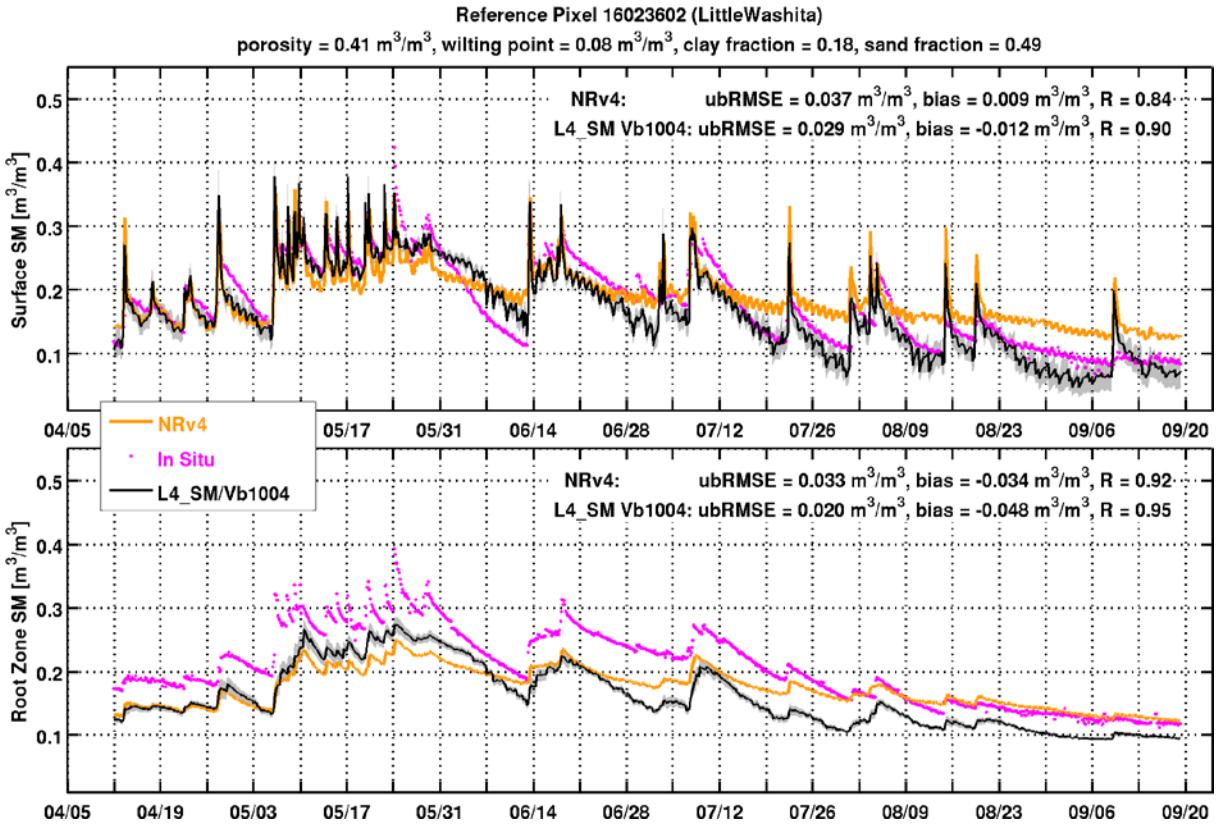

Figure 6. (Top) Soil moisture from (black solid line) L4_SM Vb1004, (orange solid line) NRv4, and (magenta dots) in situ measurements at the 36 km Little Washita reference pixel #16023602. The grey shading around the black solid line indicates the ensemble-based uncertainty estimates for surface soil moisture that are provided with the L4_SM product. (Bottom) Same as top panel but for root zone soil moisture.

The 9 km reference pixel for Little Washita (#16020907) does not have root zone soil moisture measurements, but the results for surface soil moisture at the 9 km reference pixel are similar to those at the 36 km reference pixel at that site (Table 2). Figure 6 therefore shows the L4_SM, NRv4, and in situ time series for the 36 km reference pixel (#16023602). Soil moisture varies considerably during the validation period, owing to the exceptionally wet conditions during May, which were preceded by relatively dry conditions in April and followed by a 3-month general drying trend. The L4_SM and NRv4 estimates clearly capture the overall variability, as well as the timing of the major rainstorms. The time series correlation coefficients are therefore very high, with R values of 0.90 for L4_SM surface soil

moisture and 0.95 for L4_SM root zone soil moisture, which is an improvement over the already high values of 0.84 and 0.92 for NRv4 surface and root zone soil moisture, respectively.

The improvement is also reflected in the ubRMSE metric, which decreases from 0.037 m^3m^{-3} for NRv4 surface soil moisture to 0.029 m^3m^{-3} for L4_SM, and from 0.033 m^3m^{-3} for NRv4 root zone soil moisture to 0.020 m^3m^{-3} for L4_SM. The improvements are mostly due to the increased dynamic range and the generally faster dry-downs of the L4_SM estimates that result from the assimilation of the SMAP observations and better match the in situ measurements. Bias values are very low for surface soil moisture (around 0.01 m^3m^{-3} for L4_SM and NRv4). Root zone soil moisture, however, is generally more biased, with a higher value of 0.048 m^3m^{-3} for L4_SM than the 0.034 m^3m^{-3} bias for NRv4.

6.2.3 TxSON (Texas)

While Little Washita is one of the oldest sites, TxSON (Texas) is one of the newest. The site was designed specifically to satisfy the validation of SMAP soil moisture products at the 3 km, 9 km, and 36 km spatial scales. Figure 7 shows the results for one of the 9 km TxSON reference pixels. The results are generally similar at the two other (9 km and 36 km) TxSON reference pixels. The figure shows that the precipitation pattern at TxSON was similar to that in Little Washita (section 6.2.2), with a dry April followed by an exceptionally wet May and then an extended drydown period.

As for the Little Washita reference pixels, the time series correlation coefficients for the TxSON reference pixels are very high, with R values of 0.82 for L4_SM surface soil moisture and 0.96 for L4_SM root zone soil moisture. The assimilation again results in faster drydowns and a greater dynamic range compared with NRv4, and therefore a better agreement with the in situ measurements. For the reference pixel shown in Figure 7, the improvements are reflected in the R values for surface and root zone soil moisture as well as in the ubRMSE values for root zone soil moisture. The improvements manifest themselves somewhat differently at the other TxSON reference pixels (Table 2), but overall the assimilation of SMAP observations is clearly beneficial. Unlike for Little Washita, the TxSON bias values are very high, ranging from 0.10 m^3m^{-3} to 0.14 m^3m^{-3} depending on the specific variable, data set, and reference pixel (Table 2). The bias is related to the model soil parameters for this site, including the relatively high clay fraction of 0.53 and wilting point of 0.24 m^3m^{-3}, which reflect average values across the entire 0-100 cm root zone because vertical gradients in soil texture are not represented in the soil hydrological component of the model.

One notable feature is the rapid increase in the L4_SM uncertainty estimates for surface soil moisture once the surface and root zone soil moisture values drop below the model's wilting point of 0.24 m^3m^{-3} around 15 July 2015. At that point, the model's transpiration shuts down, and modeled root zone soil moisture remains at the wilting point until a significant rain event results in sufficient infiltration to raise root zone soil moisture again. While root zone soil moisture remains stagnant at the wilting point, bare soil evaporation still taps into the surface layer soil moisture, which is no longer replenished from below and becomes highly sensitive to the perturbations in the surface meteorological forcings and the soil moisture prognostic variables, which in turn results in a dramatically increased ensemble spread in surface soil moisture.

Figure 7a also illustrates residual issues with the processing of the in situ measurements. Surface soil moisture increases slightly on July 8 and August 7, which would suggest a minor rain event. However, the increase is not due to rainfall. An analysis of the measurements from the individual sensor reveals that around the dates in question, data from several of the sensors reach extremely dry conditions and are then flagged by the quality control, which results in an average that is based on only the sensors with the somewhat wetter measurements.

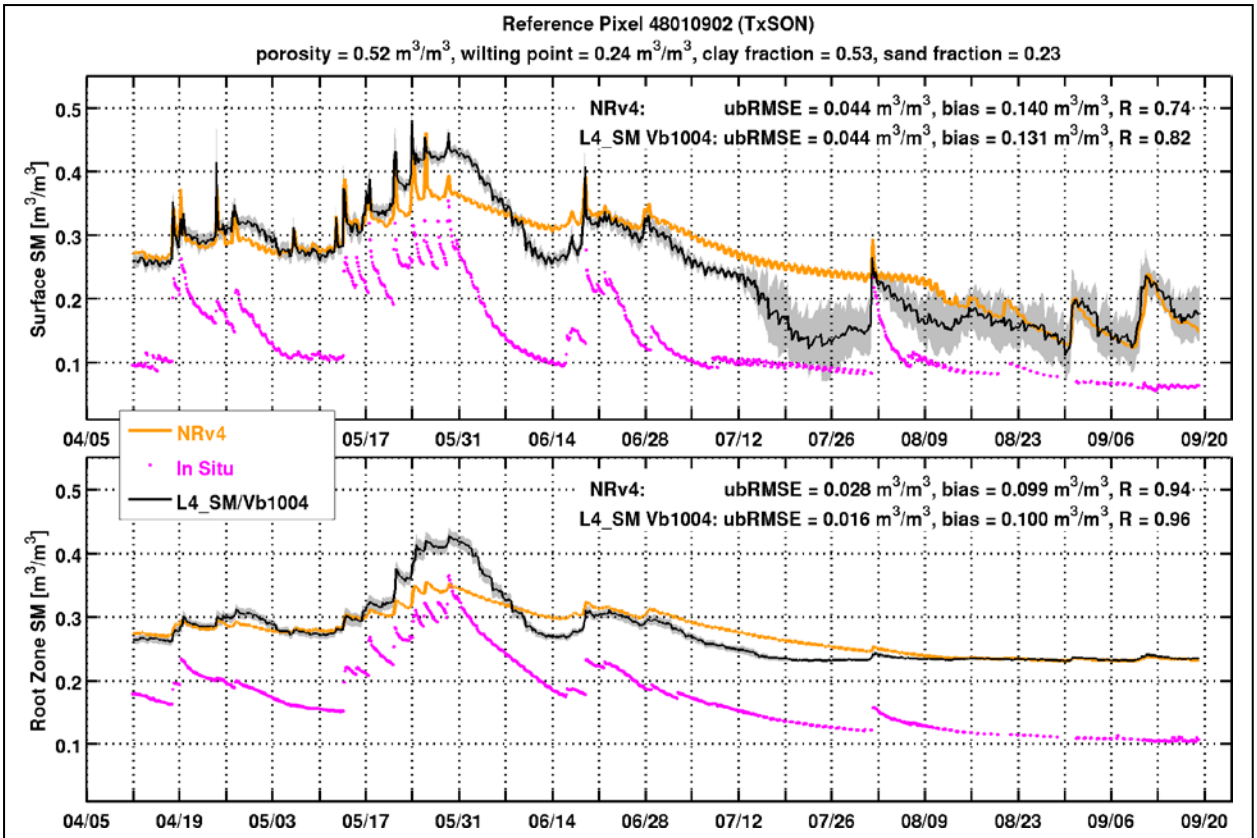

Figure 7. Same as Figure 6 but for the 9 km TxSON reference pixel #48010902.

6.2.4 Little River (Georgia)

Little River, Georgia, has been providing in situ soil moisture measurements since the launch of AMSR-E on Aqua in 2002 (Jackson et al. 2010). The site is unique in that it represents a humid agricultural environment. It also includes a substantial amount of tree cover, has very sandy soils, and is subject to irrigated agriculture. As in Little Washita, there are no in situ measurements of root zone soil moisture at the 9 km reference pixel for Little River, and the surface soil moisture results are similar for the 9 km and 36 km reference pixels. Figure 8 therefore shows the time series for the 36 km reference pixel (#16043602).

All time series reflect a drop from somewhat moister conditions in April during a long drydown in May, followed by somewhat drier conditions with frequent yet typically modest rain events during the rest of the validation period. The frequent wetting and drying events shown in the in situ measurements are reasonably captured by the L4_SM and NRv4 estimates, but the exact timing and magnitude of the storms and drydowns is less certain, with much lower R values than for the Little Washita and TxSON reference pixels. Surface soil moisture in particular has an R value of only 0.43 for NRv4, which improves to 0.60 for L4_SM with the assimilation of the SMAP observations. The correlation for root zone soil moisture is somewhat higher, with R values of 0.72 for NRv4 and 0.79 for L4_SM. The assimilation also improves the ubRMSE values of surface soil moisture estimates from 0.044 m^3m^{-3} for NRv4 to 0.036 m^3m^{-3} for L4_SM and of root zone soil moisture estimates from 0.028 m^3m^{-3} for NRv4 to 0.022 m^3m^{-3} for L4_SM. Bias values are relatively high at 0.08 m^3m^{-3} for surface soil moisture and 0.05 m^3m^{-3} for root zone soil moisture.

Figure 8b also reveals issues with the in situ measurements. Between May 17 and June 5, the reference pixel average soil moisture shows somewhat erratic behavior. In this particular case, bad data from one sensor passed the automated quality control, and sensors also drop out repeatedly during the period in question.

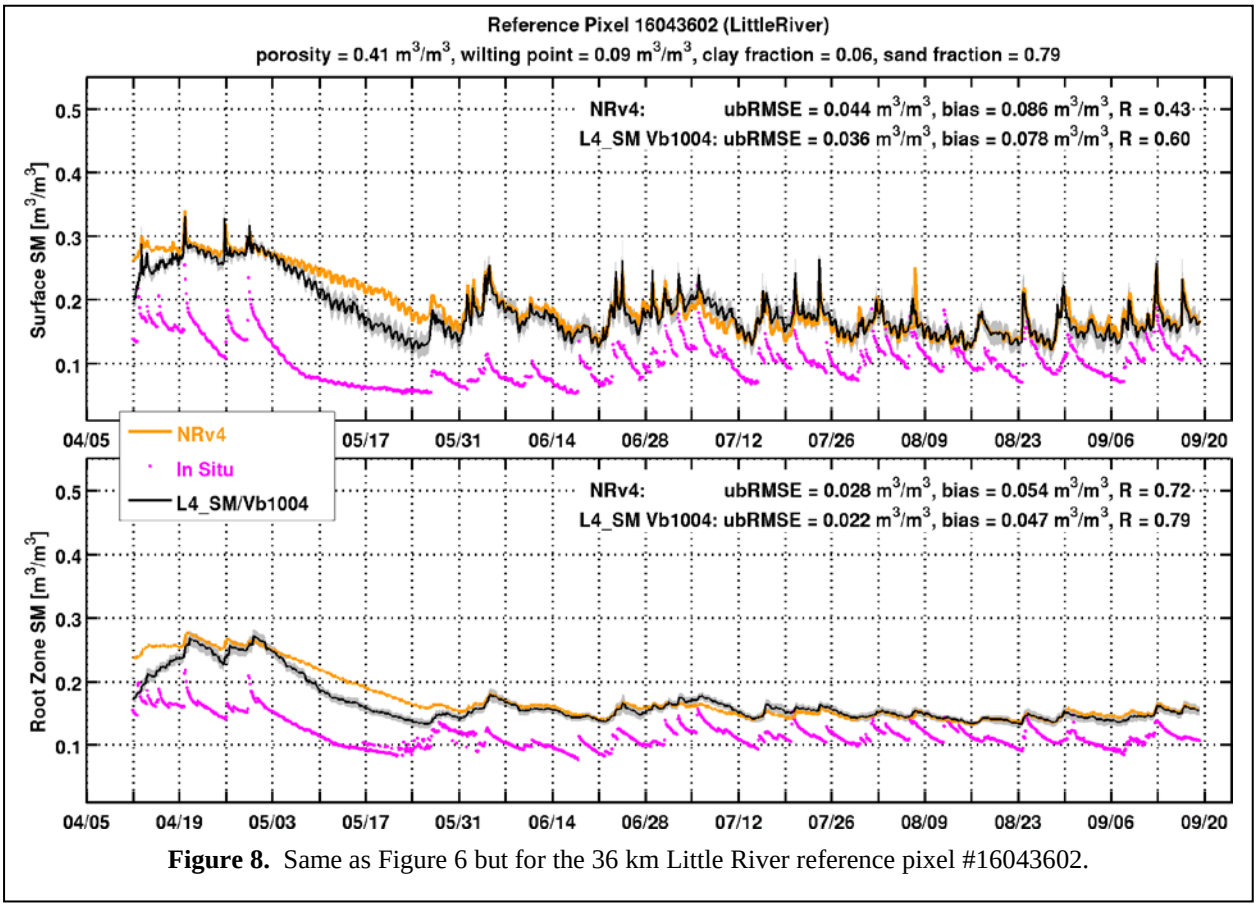

Figure 8. Same as Figure 6 but for the 36 km Little River reference pixel #16043602.

6.2.5 South Fork (Iowa)

South Fork, Iowa, is an agricultural region dominated by summer crops of corn and soybeans. Conditions in April 2015 were mostly bare soil or stubble, followed by intensive tillage that created large surface roughness. Such variations in surface roughness are difficult to capture in the microwave radiative transfer model parameters of the L4_SM algorithm. The roughness decreases again with subsequent soil treatments and rainfall, and becomes less of an issue as the growing season proceeds and crops cover the surface. By early July, corn would typically have high vegetation water content (around 3 kg m^{-2}) while that of soybeans would typically be much smaller (around 0.3 kg m^{-2}) (Jackson et al. 2004). It should also be noted that the agricultural fields are equipped with tiles to improve drainage, a process that is not captured in the global-scale Catchment land surface model of the L4_SM algorithm.

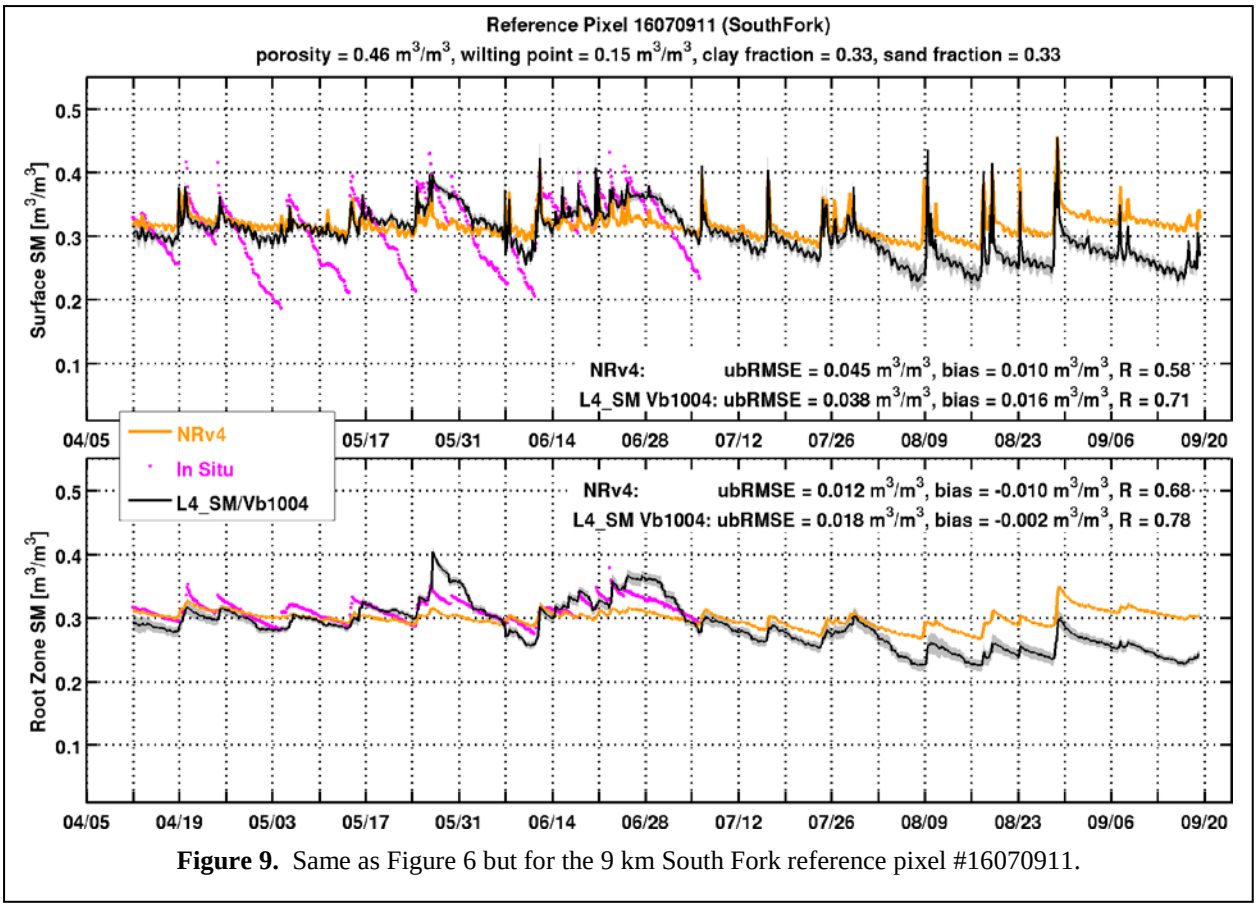

Figure 9. Same as Figure 6 but for the 9 km South Fork reference pixel #16070911.

Figure 9 shows soil moisture time series for the 9 km reference pixel #16070911, one of three 9 km reference pixels at South Fork, for which there is also a 36 km reference pixel (Table 2). Soil moisture conditions are dominated by approximately weekly rain events with subsequent drydowns, except for a generally wetter period from mid-June into early July. The L4_SM estimates capture this pattern reasonably well and present an improvement over NRv4 in terms of R values, which increase from 0.58 (NRv4) to 0.71 (L4_SM) for surface soil moisture and from 0.68 (NRv4) to 0.78 (L4_SM).

In terms of ubRMSE, L4_SM surface soil moisture estimates are also better than those of NRv4, although neither estimate captures the larger dynamic range of the in situ observations, which may be a reflection of the tile drainage. The assimilation of the SMAP observations brings the surface soil moisture ubRMSE down from 0.045 m^3m^{-3} for NRv4 to 0.038 m^3m^{-3} for L4_SM, thus ensuring that L4_SM meets the accuracy threshold of 0.04 m^3m^{-3} at this particular reference pixel. Root zone soil moisture estimates from L4_SM exhibit somewhat worse ubRMSE values than NRv4, but at 0.018 m^3m^{-3} are still well below the 0.04 m^3m^{-3} threshold. Bias values are generally low and are around 0.01 m^3m^{-3} for L4_SM and NRv4 surface and root zone soil moisture. This is encouraging because an extensive study involving sensor calibration and additional point sampling was conducted that clearly demonstrated that the network represents the average soil moisture of the 0-5 cm soil layer of the SMAP grid cell (M. Cosh 2015, USDA Hydrology and Remote Sensing Laboratory, personal communication).

6.2.6 Summary Metrics

Table 3 lists the summary metrics for surface and root zone soil moisture. The summary metrics are provided separately for the 9 km and 36 km reference pixels and are obtained by averaging across the metrics from all individual reference pixels at the given scale (Table 2). The key findings for the summary metrics (Table 3) generally match those obtained for the sample reference pixels discussed above (sections 6.2.2-6.2.5). Perhaps the most important result is that the ubRMSE values for surface and root zone soil moisture for L4_SM as well as NRv4 and at both the 9 km and the 36 km scales all meet the accuracy requirement of 0.04 m^3m^{-3}.

Table 3. Metrics averaged across core validation site reference pixels. The product (NRv4 vs L4_SM) with the better skill score is indicated by green shading.

			ubRMSE [m^3m^{-3}]			Bias [m^3m^{-3}]			R [-]		
	Horiz. Scale	Number of Reference Pixels	NRv4	L4_SM Vb1004	95% Conf. Interval	NRv4	L4_SM Vb1004	95% Conf. Interval	NRv4	L4_SM Vb1004	95% Conf. Interval
Surface Soil Moisture	9 km	17	0.039	**0.036**	0.020	0.046	**0.042**	0.017	0.67	**0.71**	0.16
	36 km	10	0.035	**0.031**	0.017	0.037	**0.032**	0.016	0.70	**0.73**	0.15
Root Zone Soil Moisture	9 km	8	0.021	**0.023**	0.032	0.004	**0.005**	0.022	0.76	**0.82**	0.36
	36 km	6	0.025	**0.024**	0.042	0.007	**0.000**	0.059	0.84	**0.78**	0.40

For a more in-depth analysis, we first compare the skill of the L4_SM and NRv4 estimates. The color-coding of the summary metrics in the table indicates whether the L4_SM or NRv4 skill is better. For all three metrics (ubRMSE, bias, and R) and for both the 9 km and the 36 km scales, the surface soil moisture skill of L4_SM exceeds that of NRv4, albeit not by a statistically significant margin. For example, at the 9 km scale the ubRMSE for L4_SM is 0.036 m^3m^{-3}, compared to 0.039 m^3m^{-3} for NRv4. The corresponding R values are 0.71 for L4_SM and 0.67 for NRv4.

The summary metrics for root zone soil moisture show a more mixed picture. At the 9 km scale, the L4_SM ubRMSE (0.023 m^3m^{-3}) is slightly higher than that of NRv4 (0.021 m^3m^{-3}), but the R value for L4_SM (0.82) is better than that of NRv4 (0.76). The numbers are reversed for the 36 km scale, where the L4_SM ubRMSE is better and the L4_SM R value is worse than the corresponding NRv4 metrics.

A closer look at the ubRMSE metric for the individual reference pixels (Table 2) reveals that the root zone soil moisture skill meets the 0.04 m^3m^{-3} threshold at all reference pixels for both L4_SM and NRv4. Surface soil moisture estimates from NRv4 fail to meet the 0.04 m^3m^{-3} threshold at 12 of the 27 reference pixels, including those at Yanco (2 out of 2), Little River (2 out of 2), TxSON (3 out of 3), Fort Cobb (2 out of 3), and South Fork (3 out of 4). By contrast, L4_SM surface soil moisture estimates fail to meet the threshold at only 3 of 27 reference pixels, the 9 km pixels at Yanco and TxSON. This result further illustrates the key role played by the assimilation of SMAP observations in improving the skill of the surface soil moisture estimates beyond the levels obtained strictly from the land surface model's integration of meteorological forcing.

Next, we compare the skill values at 9 km to those at 36 km. Here, the picture is also clearer for surface soil moisture. The L4_SM and NRv4 skill at 36 km is better for all three metrics than that at 9 km, which is consistent with the fact that the model forcing data and the assimilated SMAP brightness temperature observations are all at resolutions of about 30 km or greater. The information used to downscale the assimilated information stems from the land model parameters, which are at the finer, 9 km resolution. It is therefore not a surprise that the estimates at 36 km are more skillful than those at 9 km.

For root zone soil moisture, the results are again mixed. The R value for NRv4 at 9 km (0.76) is lower than that at 36 km (0.84). However, the R values for L4_SM as well as the ubRMSE metrics for L4_SM and NRv4 are better at 9 km than at 36 km. It should be noted, though, that the differences are not statistically significant, and that there are far fewer pixels available for evaluating root zone soil moisture than for surface soil moisture. As the time series become longer and core validation sites that are not automated (e.g., Tibet) are added, the results may become clearer.

Finally, we compare the skill of the surface estimates to that of the root zone estimates. Across all scales and metrics and for the L4_SM and NRv4 estimates, the skill of the root zone soil moisture estimates is always better than that of the surface estimates. This result makes sense because there is much more variability in surface soil moisture.

The results discussed here clearly demonstrate that the L4_SM product is of sufficient maturity and quality for beta-release to the public.

6.3 Sparse Networks

6.3.1 Method and Overview

The locally dense networks of the core validation sites are complemented by regional to continental-scale sparse networks. Although sparse networks are not ideal for soil moisture validation for the reasons discussed below, they offer in situ measurements in a larger variety of environments and provide data operationally with very short latency.

The defining feature of the sparse networks is that the there is usually just one sensor (or profile of sensors) located within a given 9 km EASE v2 grid cell. The sensor location is not necessarily representative of the grid cell, which may result in discrepancies between the L4_SM estimates and the measured soil moisture values that do not necessarily constitute errors in either data set. For example, the in situ sensor may be exposed to a locally intense rainstorm that has only a small impact on the grid cell average soil moisture. Conversely, the sensor location may not experience rainfall, but rain may be falling elsewhere the grid cell, thus impacting the L4_SM estimates. Moreover, the long-term mean and variability of soil moisture measurements from an individual sensor may not represent those of the grid cell scale soil moisture because of the heterogeneity of soil, vegetation or landscape characteristics. (In the case of some networks, such as USCRN, there are three sensors in very close proximity, which increases the robustness of the network, in particular in terms of quality control, but does not impact whether the location is representative of the grid cell.) Sparse network comparisons are therefore best interpreted in terms of time series correlation metrics. Network-average values for the mean difference (bias) may also be useful, assuming that such differences are due to a lack of representativeness for individual locations and average out when the network as a whole is considered.

The SMAP project has been evaluating methodologies for upscaling measurements from sparse networks to SMAP footprint resolutions, with a focus on using Triple Co-location approaches to correct for systematic errors in metrics derived from sparse network observations. While the Triple Co-location approaches can provide estimates of time series correlation metrics and the RMSE of random errors versus the *true* soil moisture (for which the error-prone sparse network measurements are only an approximation), it is *not* possible to compute absolute RMSE (or ubRMSE) values using Triple Co-location (Draper et al. 2013; Gruber et al. 2015). Such absolute metrics include systematic errors, such as a bias in variability, which are accounted for in the definition of the accuracy requirement (section 4).

This Assessment Report therefore focuses on metrics obtained from a direct comparison of the L4_SM product to in situ measurements, that is, metrics derived without the use of Triple Co-location. The values of the time series correlation metrics thus provided are lower than those that would be obtained with the aid of Triple Co-location and therefore provide a conservative estimate of the true skill. Moreover, the *relative* performance of the model-only and assimilation products does not depend on the use of Triple Co-location approaches.

For comparison to measurements from the single-profile sensors in the sparse networks, the L4_SM estimates are taken from the standard L4_SM 9 km EASEv2 grid cell that includes the sensor location (consistent with the approach that is used for L2 product validation). Put differently, the L4_SM estimates are not interpolated bilinearly or otherwise to the precise location of the in situ sensor locations.

Table 4. Overview of sparse networks used in L4_SM validation, with indication of the number of sites and sensor depths used here.

Network Name	Area	Number of Sites		Sensor Depths [cm]
		Surface	Root Zone	
USDA Soil Climate Analysis Network (SCAN)	USA	113	98	5, 10, 20
US Climate Reference Network (USCRN)	USA	111	85	5, 10, 20
OZNet-Murrumbidgee	Australia	36	14	4, 45
All Networks		*260*	*197*	

The current set of sparse networks used for L4_SM validation is listed in Table 4. Measurements used for L4_SM validation cover most of the United States (SCAN, USCRN) and parts of the Murrumbidgee basin in Australia (OZNet). The in situ measurements from the sparse network sites were subjected to extensive automated and manual quality control procedures by the L4_SM team following (Liu et al. 2011), which removed spikes, temporal inhomogeneities, oscillations, and other artifacts that are commonly seen in these automated measurements. Table 4 also lists the number of sites with sufficient data after quality control. A total of 260 sites provided surface soil moisture measurements, and 197 provided root zone soil moisture measurements. Most of the sites are in the continental United States, including about 100 each in the USCRN and SCAN networks. The OZNet network contributes 36 sites with surface soil moisture measurements, of which 14 sites also provide root zone measurements. Table 4 also lists the sensor depths that were used to compute the in situ root zone soil moisture. As for the core validation sites, vertical averages for SCAN and USCRN are weighted by the spacing of the sensor depths within the 0-100 cm layer corresponding to the L4_SM estimates, and the average is only computed if all sensors within a given profile provide measurements after quality control. For SCAN and USCRN sites, measurements at 50 cm (and occasionally 100 cm) depth are available, but these deeper layer measurements are not of the quality and quantity required for L4_SM validation and are therefore not used here. In future assessments, longer validation periods should facilitate the use of the measurements at 50 cm depth. For OZNet, in situ root zone soil moisture is given by the measurements at the 45 cm depth, that is, no vertical average is computed.

Because of the larger number of sparse network locations compared to the core validation site data, it is possible to examine the results stratified by general characteristics, including land cover and topographic complexity. One key distinction is whether a site is within the mask for which the formal accuracy requirement applies (section 4). A site falls outside the mask if it is in an area with mountainous topography or dense vegetation, or if it is in an urban area. The delineation used here is based on the maximum climatological LAI, the land cover class, and the variance of the elevation within the 36 km EASEv2 grid cell (a measure of topographic complexity) that contains the site. These parameters are readily available in the L4_SM modeling system. Specifically, a site is within the mask if the maximum climatological LAI less than 5, if the land cover is not forest, wetlands, or urban (that is, if the site has IGBP class 6-10, 12, or 14), and if the elevation variance around the site less than 5,000 m^2 (that is, if the standard deviation is less than 71 m).

6.3.2 Results

Figure 10 illustrates the ubRMSE values for the L4_SM estimates at the sparse network sites. The background shading in the figure also indicates whether a site is within the mask of the formal accuracy requirement (section 6.3.1). The resulting delineation (Figure 10) suggests, for example, that sites in the topographically complex western United States mountain areas and in the more densely vegetated portions of the eastern United States fall outside the mask, which is commensurate with expectation.

Overall, ubRMSE values range from 0.02 m^3m^{-3} to 0.07 m^3m^{-3}, with generally lower values for root zone soil moisture than for surface soil moisture (Figure 10). Errors are generally lowest in the dry and mountainous areas of the western United States, where the soil moisture variability is typically low, thus naturally limiting the ubRMSE values. The R values for the sparse network sites, shown in Figure 11, range from 0.3 to 0.9, with generally similar values for surface and root zone soil moisture. There is no obvious spatial pattern, except maybe that R values are almost universally high for the OZNet sites in Australia.

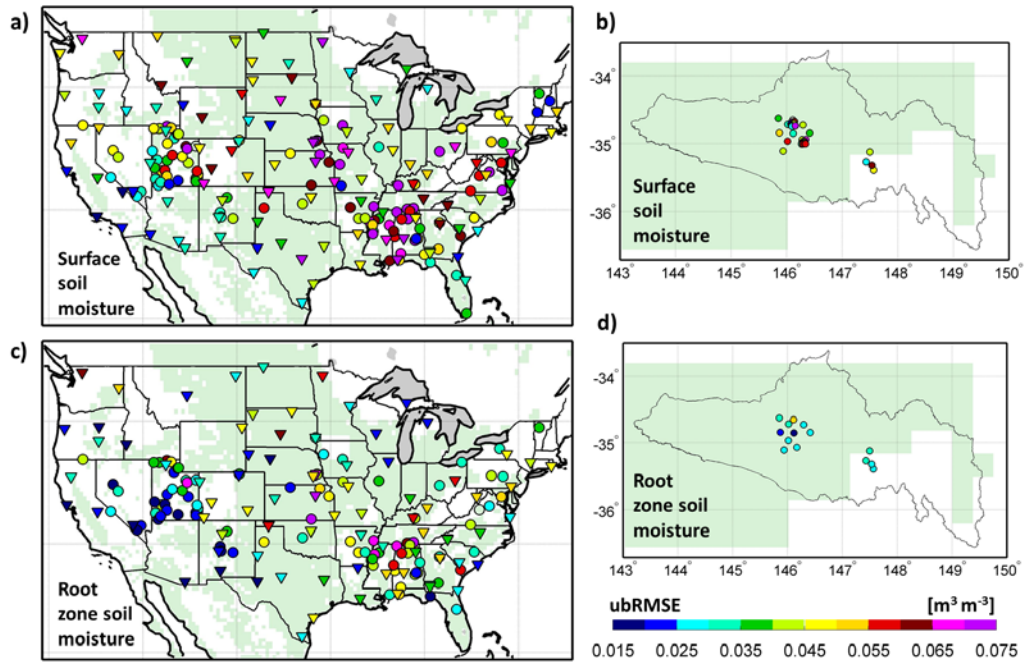

Figure 10. ubRMSE for L4_SM (a,b) surface and (c,d) root zone soil moisture. (a,c) The United States sites include (dots) SCAN and (inverted triangles) USCRN. (b,d) The Australian sites are from OZNet. Green shading indicates areas with low or modest vegetation cover and topographic complexity that are within the mask of the SMAP accuracy requirement. See text for details.

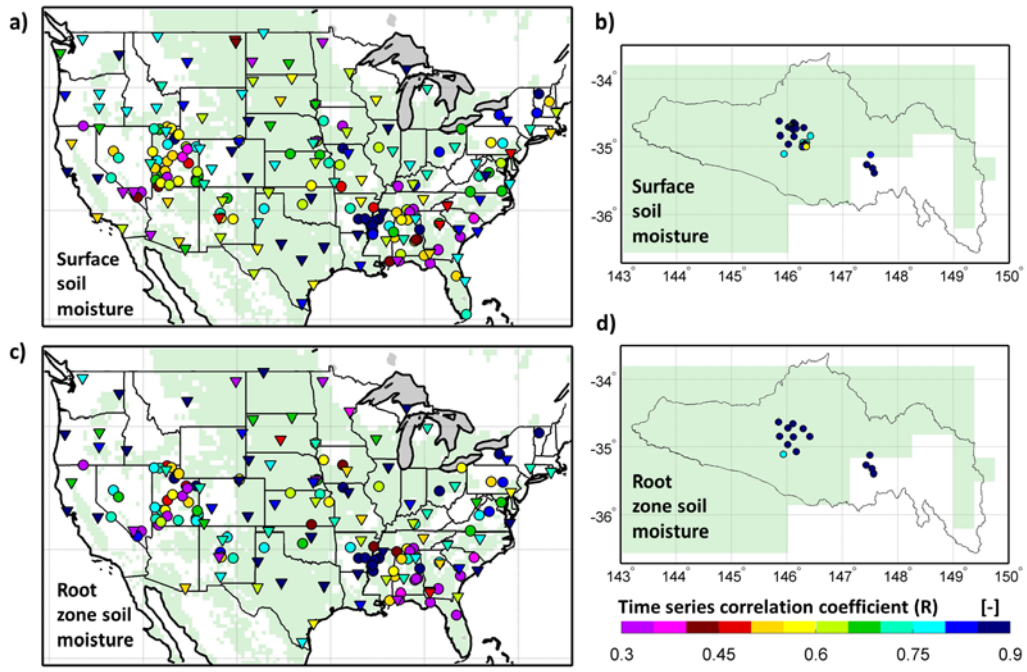

Figure 11. Same as Figure 10 but for the time series correlation coefficient (R).

Figure 12 shows the average L4_SM metrics across the sites from all networks broken down by the exclusion mask of the accuracy requirement (as indicated by the shading in Figure 10). The average metrics are computed based on a clustering algorithm that assigns the weights given to each location based on the density of sites in the surrounding region. As suggested by the map plots, Figure 12 illustrates that the ubRMSE values are considerably lower at the sites outside the mask, with values around 0.04 m^3m^{-3} for surface soil moisture and 0.03 m^3m^{-3} for root zone soil moisture, compared to around 0.05 m^3m^{-3} and 0.04 m^3m^{-3} for surface and root zone soil moisture, respectively, at sites within the mask. Again, this result is related to the much lower variability of soil moisture in the arid regions of the western United States, which also happen to lie largely in mountainous terrain. The result is reversed for the network-average absolute bias, where values are much lower within the mask (0.05-0.06 m^3m^{-3}) than outside the mask (0.08-0.09 m^3m^{-3}). The values for the time series correlation coefficients are more similar inside and outside the mask and generally range between 0.65 and 0.75.

Figure 12 also shows the skill of the NRv4 estimates. Across the board, the L4_SM skill is slightly higher than that of NRv4, reflecting the additional information contributed by the assimilation of the SMAP brightness temperature observations in the L4_SM system. The skill differences are small, though, and not statistically significant because of the relatively short data record. As for the core validation sites, the typically small differences between L4_SM and NRv4 estimates reflect the fact that the sparse network measurements are located in areas where the surface meteorological forcing takes advantage of high-quality, gauge-based precipitation measurements. Larger improvements from the assimilation of SMAP observations can be expected in areas where the precipitation forcing inputs are not as informed by gauge measurements.

Table 5 provides average skill metrics broken down by land cover as well as by the individual networks. The breakdown by network is provided for completeness, but it is difficult to interpret because of the large differences in the number and location of individual sites within each network.

The breakdown by land cover follows the IGBP classes. There are no sparse network sites in the closed shrublands, savannas, permanent wetlands, and snow/ice classes (IGBP classes 6, 9, 11, 15). Urban/built-up and barren/sparse classes include only one and two sites, respectively. We lumped the five IGBP classes for forests, including evergreen/deciduous/needleleaf/broadleaf and mixed forest (IGBP classes 1-5), into a single "forest" class. Besides this lumped forest class, there are five additional IGBP classes for which between 17 and 72 sites are available (Table 5).

As stated above, the sparse network comparisons are best interpreted in terms of time series correlation coefficients, which discounts some of the errors that arise because a given in situ site is not representative of the grid cell average soil moisture. The R values for the L4_SM product range roughly between 0.6 and 0.8, with skill on the higher end of that range in the forest, grasslands, and croplands classes, and skill at the lower end in the open shrublands and mixed crop/natural classes (Table 5). The same ranking applies to surface and root zone soil moisture estimates.

Table 5 also provides skill metrics for the NRv4 estimates. The R values of the L4_SM surface soil moisture estimates exceed those of NRv4 for all IGBP classes (except for the single site in the barren/sparse class), by an average of 0.03. For root zone soil moisture the R values of L4_SM are better than those of NRv4 except for grasslands, where NRv4 is better. Averaging across all sites, the R value for root zone soil moisture matches that of NRv4. The average ubRSME metrics across all networks similarly indicate a very small improvement in surface soil moisture (by 0.002 m^3m^{-3}) and a still smaller improvement in root zone soil moisture. These results mirror the key finding of the core validation site analysis: the assimilation of SMAP brightness temperatures primarily improves surface soil moisture estimates, and, on average, does not (yet) improve the skill of the modeled root zone soil moisture.

The bias values listed in Table 5 suggest that across the four networks, the mean soil moisture from the L4_SM and NRv4 estimates is biased high (that is, wet) by about 0.06 m^3m^{-3} for the surface and by about 0.02 m^3m^{-3} for the root zone. The root zone bias in particular is remarkably small and provides some confidence in the skill of the model-based estimates. Note, however, that the typical bias at an individual site (as measured by the mean of the absolute bias) is around 0.08 m^3m^{-3} for the surface and 0.07 m^3m^{-3} for the root zone (not shown). This is at least partly a reflection of the fact that sparse network sites are not necessarily representative of the conditions in the 9 km grid cell for which the L4_SM and NRv4 estimates are valid. This is particularly true for the forest class, because measurement sites are typically on grassy areas, regardless of the surrounding land cover. For the forest class the L4_SM and NRv4 estimates have the highest bias values, around 0.10 m^3m^{-3} for surface soil moisture and around 0.05 m^3m^{-3} for root zone soil moisture (not considering the much higher bias at the single site in the urban class).

Overall, the skill values for the sparse network sites yield results that are very similar to those obtained from the core validation sites. The beneficial impact of assimilating SMAP brightness temperature observations is greatest for surface soil moisture. Furthermore, root zone soil moisture estimates are not getting worse when SMAP brightness temperatures are assimilated. Finally, it is important to keep in mind that the skill metrics presented here underestimate the true skill because these metrics are based on a direct comparison against in situ measurements (which are subject to error). Therefore, the sparse network ubRMSE values suggest that the beta-release L4_SM estimates are at least very close to meeting the formal accuracy requirement across a very wide variety of surface conditions, beyond those that are covered by the few core validation sites that have been available to date. The sparse network results thus provide additional confidence in the conclusions drawn from the core validation site comparisons.

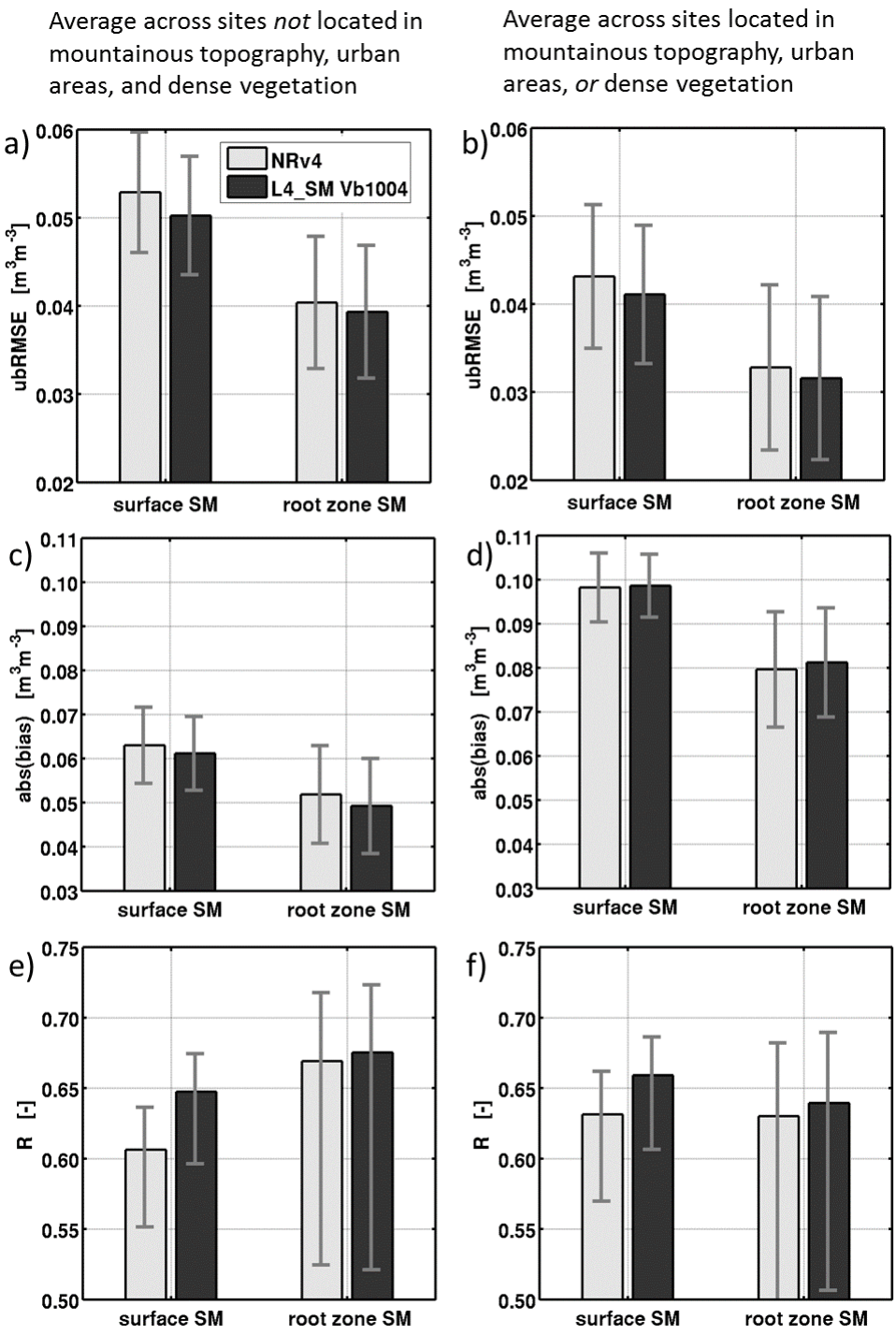

Figure 12. Skill metrics for L4_SM and NRv4 averaged over sparse network sites. (a,b) ubRMSE, (c,d) absolute bias, and (e,f) time series correlation coefficient (R). Summary metrics are averaged over sites (a,c,e) within and (b,d,f) outside the mask of the formal accuracy requirement. See text for details.

Table 5. Sparse network metrics by land cover (IGBP class), by network, and by the mask used for the (core validation site) accuracy requirement.

Sparse Network Subset	Surface Soil Moisture										Root Zone Soil Moisture									
		ubRMSE [m^3m^{-3}]			Bias [m^3m^{-3}]			R [-]				ubRMSE [m^3m^{-3}]			Bias [m^3m^{-3}]			R [-]		
	Num-ber of sites	NRv4	**L4_SM Vb1004**	95% Conf. Interval	NRv4	**L4_SM Vb1004**	95% Conf. Interval	NRv4	**L4_SM Vb1004**	95% Conf. Interval	Num-ber of sites	NRv4	**L4_SM Vb1004**	95% Conf. Interval	NRv4	**L4_SM Vb1004**	95% Conf. Interval	NRv4	**L4_SM Vb1004**	95% Conf. Interval
Forests (IGBP 1-5)	38	0.049	**0.046**	0.014	0.102	**0.095**	0.016	0.65	**0.68**	0.07	30	0.039	**0.037**	0.014	0.051	**0.045**	0.017	0.65	**0.67**	0.12
Open shrublands (IGBP 7)	25	0.035	**0.031**	0.005	0.020	**0.027**	0.005	0.50	**0.57**	0.06	17	0.029	**0.024**	0.032	0.008	**0.014**	0.052	0.60	**0.63**	0.18
Woody savannas (IGBP 8)	20	0.055	**0.047**	0.028	0.063	**0.049**	0.024	0.63	**0.71**	0.13	18	0.047	**0.041**	0.043	0.023	**0.006**	0.078	0.57	**0.65**	0.33
Grasslands (IGBP 10)	72	0.049	**0.048**	0.010	0.051	**0.057**	0.012	0.69	**0.70**	0.05	57	0.038	**0.037**	0.011	-0.001	**0.009**	0.017	0.75	**0.71**	0.12
Croplands (IGBP 12)	61	0.058	**0.053**	0.009	0.052	**0.052**	0.010	0.61	**0.67**	0.06	39	0.040	**0.040**	0.011	0.020	**0.024**	0.014	0.68	**0.69**	0.15
Urban/built-up (IGBP 13)	1	0.059	**0.057**	0.054	0.229	**0.181**	0.021	0.59	**0.62**	0.28	1	0.056	**0.052**	0.075	0.227	**0.183**	0.033	0.65	**0.68**	0.56
Crop/natural (IGBP 14)	37	0.056	**0.055**	0.006	0.041	**0.030**	0.007	0.58	**0.61**	0.06	32	0.042	**0.041**	0.007	0.012	**0.000**	0.008	0.58	**0.60**	0.15
Barren/sparse (IGBP 16)	2	0.026	**0.022**	0.007	-0.013	**-0.001**	0.005	0.40	**0.39**	0.13	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SCAN	113	0.052	**0.050**	0.005	0.039	**0.039**	0.005	0.57	**0.60**	0.05	98	0.038	**0.038**	0.007	-0.007	**-0.005**	0.008	0.60	**0.59**	0.11
USCRN	111	0.048	**0.045**	0.008	0.073	**0.070**	0.009	0.64	**0.67**	0.04	85	0.038	**0.036**	0.009	0.034	**0.030**	0.013	0.68	**0.70**	0.08
Oznet	36	0.057	**0.048**	0.039	-0.002	**0.035**	0.057	0.82	**0.84**	0.17	14	0.021	**0.030**	0.104	-0.082	**-0.046**	0.140	0.92	**0.90**	0.65
Inside mask	146	0.053	**0.050**	0.007	0.045	**0.042**	0.009	0.61	**0.65**	0.04	104	0.040	**0.039**	0.008	0.008	**0.006**	0.011	0.67	**0.68**	0.10
Outside mask	114	0.043	**0.041**	0.008	0.081	**0.080**	0.008	0.63	**0.66**	0.04	93	0.033	**0.032**	0.009	0.046	**0.046**	0.013	0.63	**0.64**	0.09
Average (all sites)	**260**	**0.048**	**0.046**	**0.007**	**0.059**	**0.058**	**0.008**	**0.61**	**0.64**	**0.04**	**197**	**0.037**	**0.036**	**0.008**	**0.023**	**0.022**	**0.011**	**0.66**	**0.66**	**0.08**

6.4 Data Assimilation Diagnostics

This section provides an evaluation of the L4_SM data assimilation diagnostics, including the statistics of the observation-minus-forecast (O-F) residuals, the observation-minus-analysis (O-A) residuals, and the analysis increments. Because the L4_SM algorithm assimilates brightness temperature observations, the O-F and O-A diagnostics are in terms of brightness temperatures (that is, in "observation space"). Strictly speaking, the analysis increments are in the space of the Catchment model prognostic variables that make up the "state vector", including the "catchment deficit", "root zone excess", "surface excess", and "top-layer ground heat content" (Reichle et al. 2014b). For the discussion below, the increments have been converted into equivalent soil moisture and soil temperature terms.

6.4.1 L4_SM Analysis

Figure 13 illustrates, for 29 May 2015, 0z, the soil moisture and temperature analysis that lies at the heart of the L4_SM algorithm. The panels only show a part of North America (the continental United States and portions of Mexico and Canada) to reveal the details of the global analysis. Figure 13a depicts the map of H-pol brightness temperature O-F residuals, which typically range between -15 K and 15 K. V-pol data exhibit similar features (Figure 13b). At the analysis time in question, brightness temperature observations were available from two ascending swaths, one crossing the eastern portion and another crossing the western portion of North America. Observations were not assimilated everywhere. For example, the quality control steps in the L4_SM processing exclude brightness temperature observations in the vicinity of open water surfaces or where model temperatures indicate surface conditions near or below freezing, for example, in the Rocky Mountains.

Note that while the O-F residuals are posted on the 9 km EASE v2 grid of the L4_SM "aup" Collection, the assimilated L1C_TB brightness temperature observations and thus the O-F residuals are effectively on the 36 km EASE v2 grid. In contrast, the increments are at the 9 km resolution of the model. One key feature of the L4_SM analysis is the downscaling of the SMAP radiometer data to the 9 km model resolution based on the modeled error characteristics, which vary dynamically and spatially.

Figures 13c, e, and g show the resulting analysis increments in surface soil moisture, root zone soil moisture, and surface soil temperature, respectively. Areas where the observed brightness temperature is warmer than the model forecast result in negative increments in soil moisture and positive increments in soil temperature. Increments in surface soil moisture typically range from -0.03 m^3m^{-3} to 0.03 m^3m^{-3}. Increments in root zone soil moisture are generally smaller and mostly range between -0.01 m^3m^{-3} and 0.01 m^3m^{-3}, reflecting the fact that the brightness temperature observations are directly sensitive only to the soil moisture and temperature in a surface layer of approximately 5 cm thickness. Root zone soil moisture increments rely on the error cross correlations between the modeled brightness temperatures and root zone soil moisture that evolve dynamically with the model ensemble.

The analysis increments are generally smooth and cover an area that is slightly larger than the coverage of the O-F residuals, owing to the spatially distributed (three-dimensional) ensemble Kalman filter used in the L4_SM analysis. Finally, Figures 13d, f, and h show the resulting analysis fields for surface soil moisture, root zone soil moisture, and surface soil temperature, respectively. The analysis fields are computed by adding the analysis increments of Figures 13c, e, and g to the model forecast (not shown). Figures 13d, f, and h illustrate that the analysis increments blend seamlessly into the model forecast fields, and that the geometrical features of the satellite swaths of assimilated observations cannot be discerned in the final analysis.

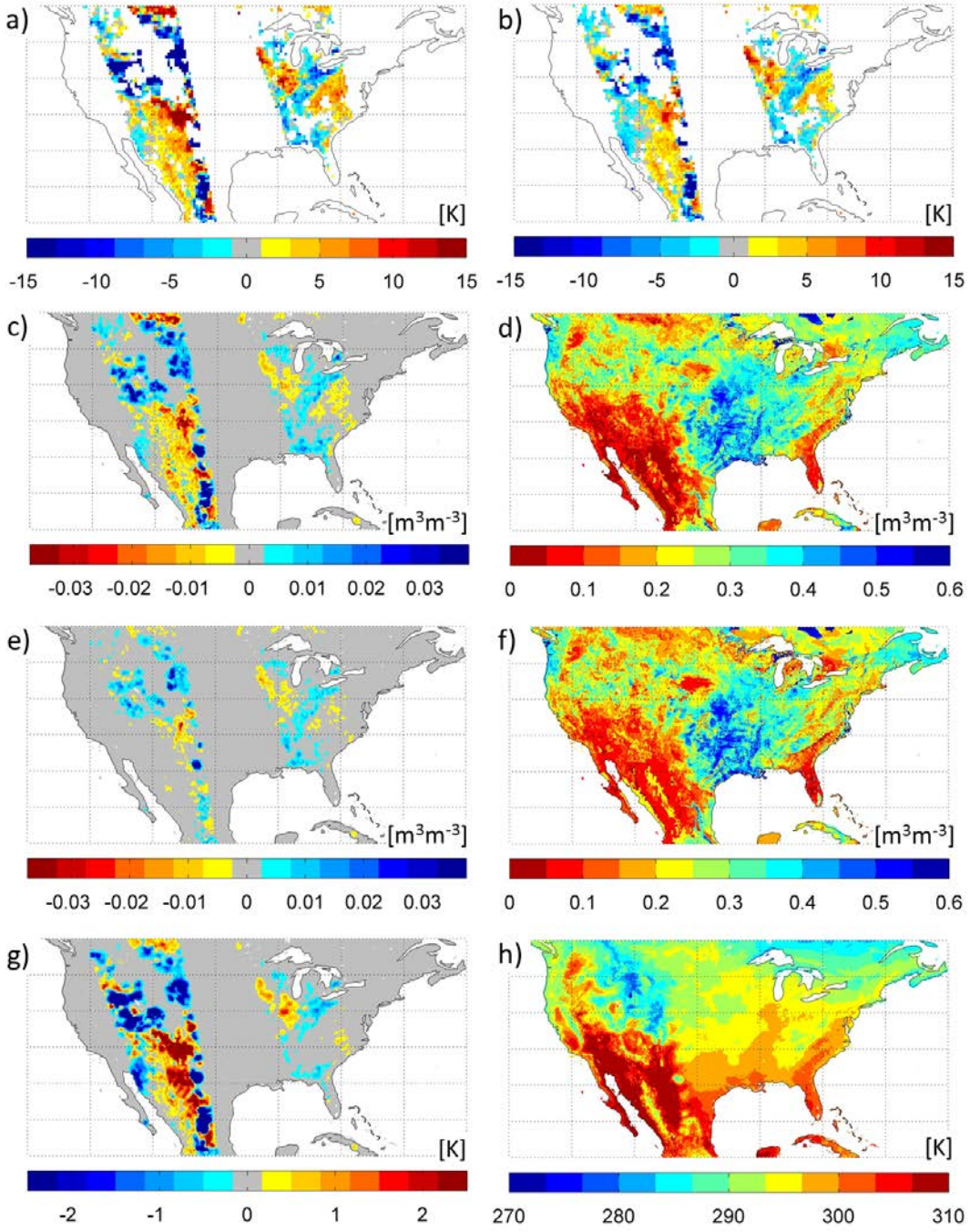

Figure 13. L4_SM analysis for 29 May 2015, 0z. Top row shows O-F residuals for (a) H-pol and (b) V-pol. Analysis increments are shown for (c) surface soil moisture, (e) root zone soil moisture and (g) surface soil temperature. The resulting analysis fields are shown for (d) surface soil moisture, (f) root zone soil moisture and (h) surface soil temperature.

6.4.2 Observation-Minus-Forecast Residuals

Figure 14 shows the global coverage of the SMAP L1C_TB observations that were used in the L4_SM analysis for 7 June 2015, 0z. The analysis window includes brightness temperature observations between 22:30z on 6 June 2015 and 01:30z on 7 June 2015. Within this window, approximately 9,000 observations were used in total, including about 2,700 H-pol and 2,700 V-pol observations from two descending half-orbits over eastern Russia and Indonesia and about 1,800 H-pol and 1,800 V-pol observations from two ascending half-orbits over the Americas.

SMAP L1C_TB observations were not used over China where L-band radio-frequency interference (RFI) is common. SMAP is equipped with a variety of hardware and software tools to deal with RFI and generally provides near-global coverage. However, the L4_SM algorithm requires knowledge of the long-term L-band brightness temperature climatology to address observation-minus-forecast bias in the system (Reichle et al. 2014b). The necessary climatological information is derived from observations provided by the Soil Moisture Ocean Salinity (SMOS) mission, which cannot provide good quality observations in the RFI-affected areas.

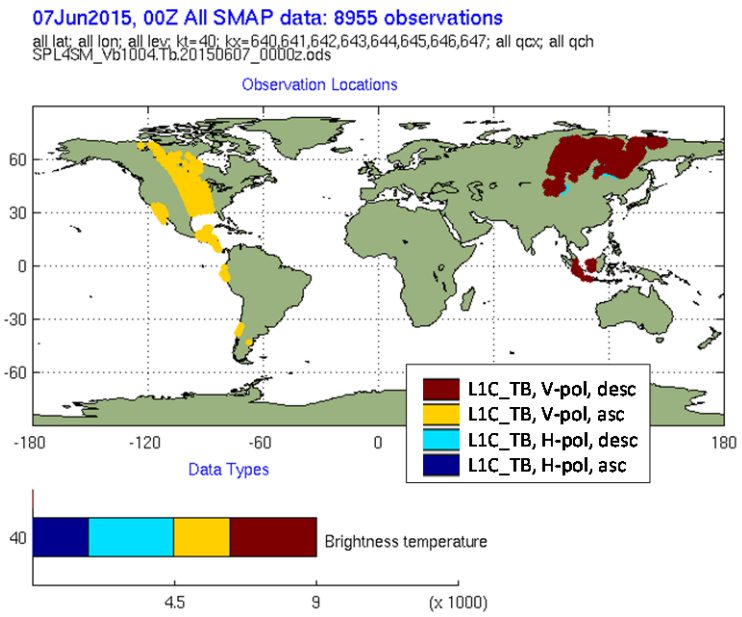

Figure 14. (Top) Coverage of the assimilated SMAP L1C_TB brightness temperature observations that were used in the L4_SM analysis for 7 June 2015, 0z (Vb1004). In the map, the coverage of V-pol observations obscures the areas where H-polarized observations are used. (Bottom) Number of brightness temperature observations assimilated in this analysis.

The gaps in spatial coverage are further illustrated in Figure 15, which shows the total number of L1C_TB observations that were assimilated during the validation period (11 Apr 2015, 0z, to 19 Sep 2015, 0z). This count includes H- and V-pol observations from ascending and descending orbits. The average data count across the globe is approximately 213 for the 161-day period, but no data are

assimilated across large areas in eastern Europe and the southern half of continental Asia due to the lack of a SMOS brightness temperature climatology because of RFI. Moreover, few or no SMAP brightness temperatures are assimilated in mountainous areas, including the Rocky Mountains and the Andes, in the vicinity of lakes, such as in northern Canada, and next to major rivers, including the Amazon and the Congo. High-latitude northern areas exhibit the highest counts of assimilated brightness temperature observations because of SMAP's polar orbit, which results in more frequent revisit times there.

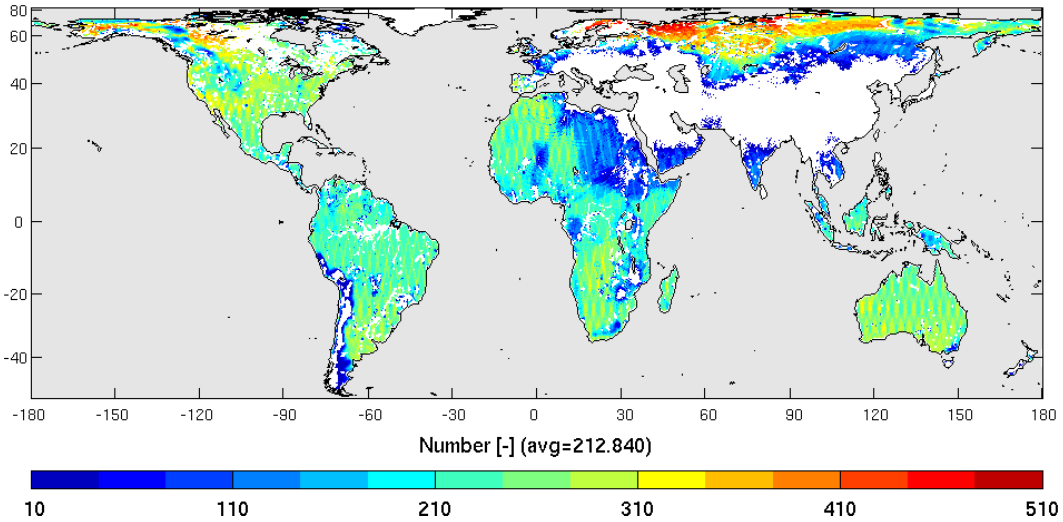

Figure 15. Number of L1C_TB observations used in the L4_SM algorithm (Vb1004) during the validation period (11 Apr 2015, 0z, to 19 Sep 2015, 0z). Data counts include H-pol and V-pol data from ascending and descending half-orbits.

Next, Figure 16 shows a time series of the global observations counts for July 2015, again including H-pol and V-pol observations from ascending and descending half-orbits. There are 8 analysis times per day (at 0z, 3z, …, 18z, and 21z), with data counts varying depending on the time of day, primarily based on the amount of land surface area where the local time is close to 6am or 6pm local time, when SMAP crosses the Equator. Each L4_SM analysis typically ingests between 5,000 and 15,000 observations, with mean and median values very close to 10,000.

The bottom panel of Figure 16 shows the time series of the O-F and O-A statistics. Global mean O-F values (cyan bars) are around 2-3 K prior to 7 July 2015, the day on which the radar anomaly occurred. Because of the changes in the observatory's thermal conditions, the calibration of the Level 1 brightness temperatures changed slightly. This is reflected in a subtle shift of the mean O-F values, which after 7 July 2015 are typically between 0 K and 2 K. Mean O-A values (orange bars) are slightly smaller than mean O-F values. Overall, the relatively small mean O-F and O-A values suggest that the assimilation system is reasonably bias-free in a global average sense.

Typical magnitudes of the O-F residuals, indicated by the RMS values (blue bars), range between 3 K and 7 K. The RMS values of the O-A residuals (red bars) are generally lower and rarely exceed 5 K, thereby reflecting the reduction in uncertainty obtained from the analysis.

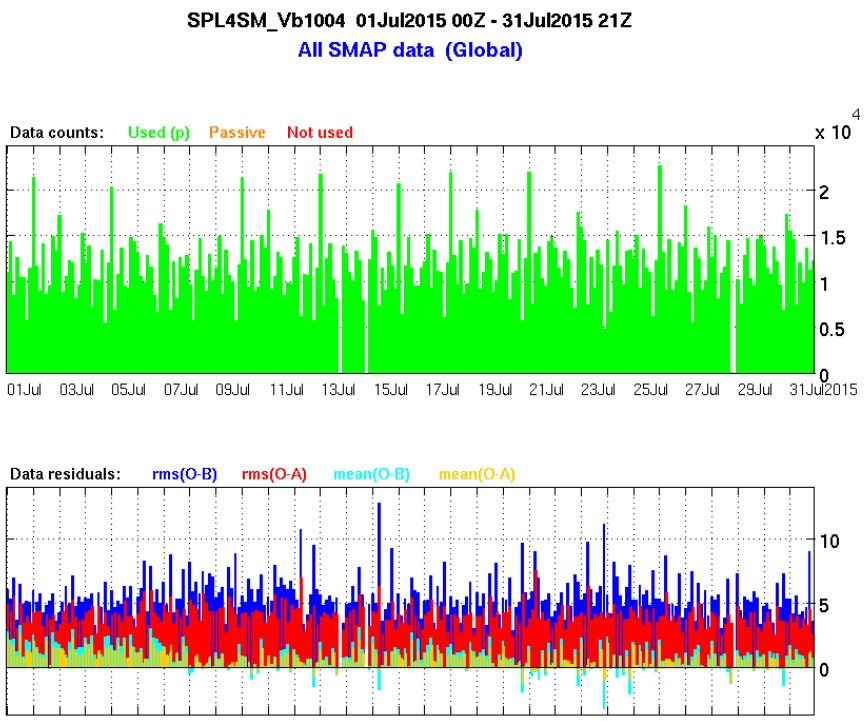

Figure 16. (Top) Counts of L1C_TB observations assimilated into L4_SM (Vb1004) during July 2015, including H-pol and V-pol data from ascending and descending orbits. (Bottom) Statistics of the corresponding O-F (or O-B) and O-A residuals, where the mean and RMS values are computed separately for each 3-hourly analysis by averaging across the global domain.

The O-F RMS values show occasional spikes exceeding 10 K. One such spike can be seen for the 15 July 2015, 6z, analysis and corresponds to a major rain event in southeastern Niger that was missed in the L4_SM precipitation forcing. This resulted in forecast brightness temperature values that were much larger than those observed by SMAP, with O-F values as low as -90 K (and thus a large O-F RMS value).

Figure 17 shows the global distributions of the time series mean and standard deviation of the O-F residuals. The time mean values of the O-F residuals are typically small and mostly range from 0 K to 2 K. Overall, there is a positive bias of 1.3 K, with very few areas exhibiting negative mean O-F values. The largest values are found in the Sahel and in central and southern Africa. Over Africa, the L4_SM precipitation forcing is not corrected to the gauge-based product. The L4_SM algorithm therefore relies heavily on the consistency of the present forcing data from the ¼ degree GEOS-5 operational forward processing (FP) system (GEOS-5.13) and the historic forcing data from the ½ degree reprocessing (RP-IT/FP-IT) system (GEOS-5.9) that was used to derive the brightness temperature rescaling factors in the calibration of the L4_SM algorithm. High values are also seen in the center of the United States, Argentina, Australia, and portions of Siberia, which indicates that the L4_SM system would benefit from further calibration.

The time series standard deviation of the O-F residuals ranges from a few Kelvin to around 15 K, with a global average of about 5.8 K. The highest values are found in central North America, the Sahel, central Asia, and the southern half of Australia. These regions have sparse or modest vegetation cover and typically exhibit strong variability in soil moisture conditions. The O-F residuals are generally smallest in more densely vegetated regions, including the eastern United States, the Amazon basin, and

tropical Africa. Small values are also found in high-latitudes, including Alaska and Siberia, and in the Sahara desert.

By construction, the analysis pulls the forecast brightness temperatures closer to the (rescaled) observations. The global average of the time mean observation-minus-analysis brightness temperature residuals is about 0.3 K, down from the 1.3 K bias of the O-F residuals (not shown). The time series standard deviation of the O-A residuals ranges from less than 1 K to around 7 K, with a spatial pattern that roughly matches that of Figure 17b and a global average of about 2.6 K (not shown).

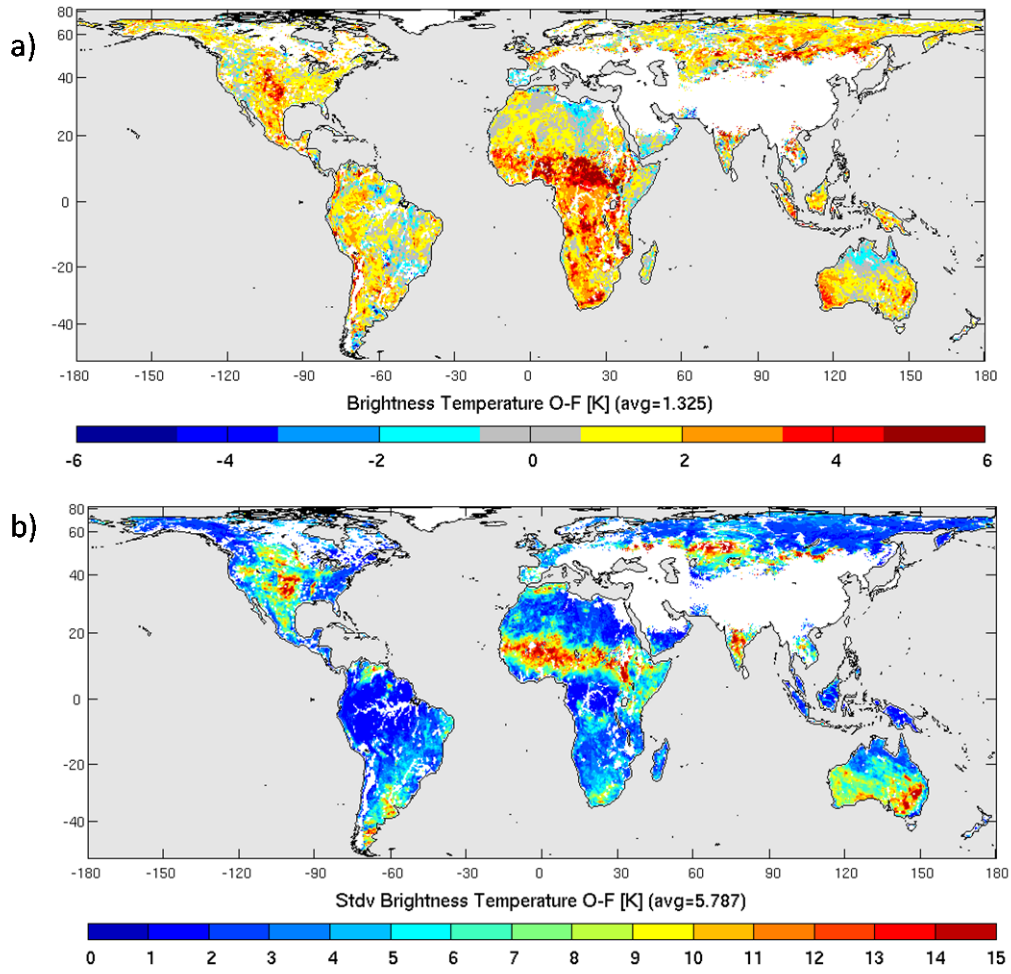

Figure 17. (a) Mean and (b) standard deviation of the O-F residuals from the L4_SM algorithm (Vb1004) for 11 Apr 2015, 0z, to 19 Sep 2015, 0z.

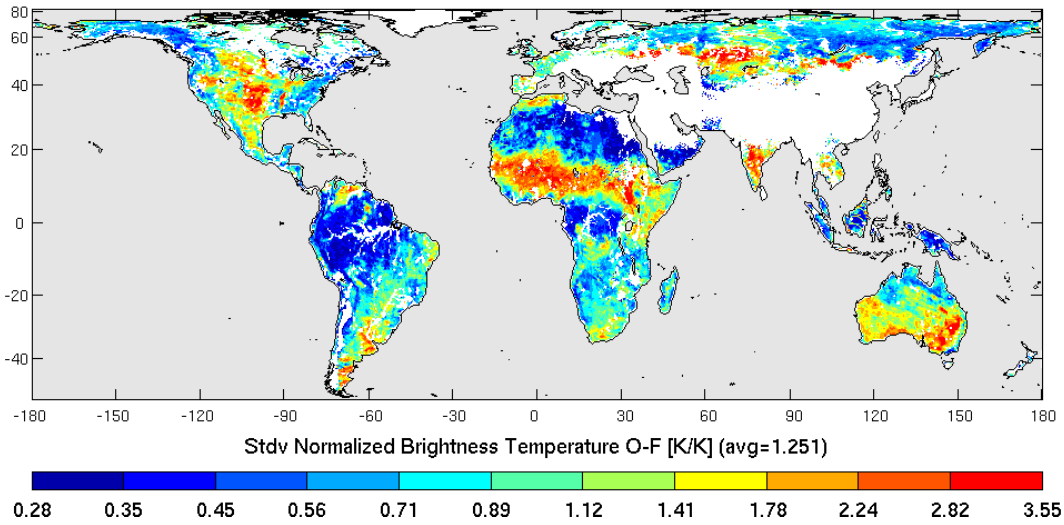

Figure 18. Standard deviation of the *normalized* O-F residuals from the L4_SM algorithm (Vb1004) for 11 Apr 2015, 0z, to 19 Sep 2015, 0z.

Finally, Figure 18 shows the standard deviation of the *normalized* O-F residuals, which measures the consistency between the expected (modeled) errors and the actual errors. Specifically, the normalization of the O-F residuals is with the standard deviation of their expected total error, which is composed of the error in the observations (including instrument errors and errors of representativeness) and errors in the brightness temperature model forecasts (Appendix B). The parameters that determine the expected error standard deviations are key inputs to the ensemble-based L4_SM assimilation algorithm. If they are chosen such that the modeled errors are fully consistent with the actual errors, the metric shown in the figure should be unity. If the metric is less than one, the actual errors are overestimated by the assimilation system, and if the metric is greater than one, the actual errors are underestimated.

The global average of the metric is 1.25 (Figure 18), which would suggest that, on average, the modeled errors are roughly consistent with the actual errors. The metric, however, varies greatly across the globe. Typical values are either too low or too high. In the Amazon basin, the eastern US, tropical and southern Africa, and the high northern latitudes, values range from 0.25 to 0.5, and thus errors there are considerably overestimated. Conversely, in central North America, the Sahel, India, and southeastern Australia, values range from 1.5 to 4, meaning that errors in these regions are considerably underestimated. Future work will focus on improving the calibration of the input parameters that determine the model and observation errors in the L4_SM system.

6.4.3 Increments

Figure 19 shows the average number of increments that the L4_SM algorithm generated per day during the validation period. The global mean is very close to unity, which means that for a given location, one increment per day is applied on average, either from an ascending or a descending overpass. The overall pattern of the increments count follows that of the count of the assimilated observations shown in Figure 15. The coverage of the increments, however, is somewhat greater than that of the observations due to the spatial interpolation and extrapolation of the observational information in the distributed analysis update of the L4_SM algorithm. The figure also reveals the diamond patterns resulting from SMAP's regular 8-day repeat orbit. Furthermore, circular features can be seen in southern Africa and northwestern Australia. It is not yet clear how these patterns arise.

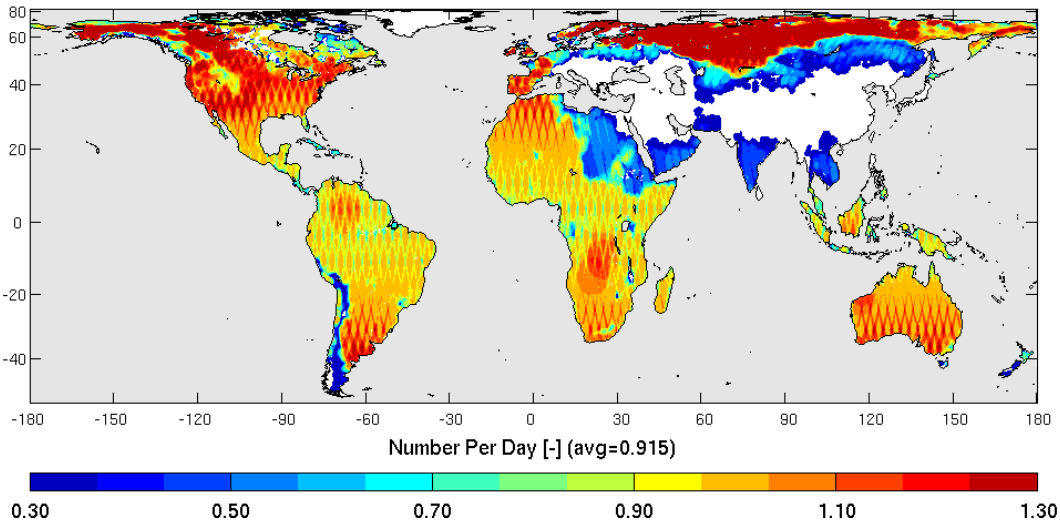

Figure 19. Average number of increments per day generated by the L4_SM algorithm (Vb1004) during the validation period (11 Apr 2015, 0z, to 19 Sep 2015, 0z). The result applies equally to all elements of the control vector, including the model prognostic variables related to surface soil moisture, root zone soil moisture, surface temperature, and top-layer soil temperature.

Figure 20 shows the time mean values of the analysis increments for surface and root zone soil moisture as well as for the surface layer soil temperature. In the long-term average, the increments for root zone soil moisture and surface soil temperature vanish nearly everywhere. Only the increments in surface soil moisture exhibit a slight bias in some regions, including the US Great Plains, the Sahel, and southern Africa, with extreme values of around -0.01 m^3m^{-3}. These drying increments are a reflection of the slight warm bias in the O-F residuals (Figure 17a). Overall, Figure 20 suggests that the analysis system is very nearly unbiased.

Figure 21 shows the time series standard deviation of the increments in surface and root zone soil moisture as well as surface soil temperature. This metric measures the typical magnitude of instantaneous increments. Typical increments in surface soil moisture are on the order of 0.015 m^3m^{-3} in the western US, central Mexico, the Sahel, eastern Africa, southern India, and most of Australia. They are around 0.005 m^3m^{-3} in the eastern US, Argentina, southern Africa, and the high northern latitudes. Over the

tropical forests, surface soil moisture increments are generally negligible, reflecting the fact that in those areas the measured SMAP brightness temperatures are mostly sensitive to the dense vegetation and are only marginally sensitive to soil moisture and soil temperature.

Typical increments in root zone soil moisture (Figure 21b) show a global pattern that is very similar to that of the surface soil moisture increments, albeit with smaller magnitudes that again reflect the lesser sensitivity of the L-band brightness temperatures to the deeper layer soil moisture. The magnitude of the average root zone soil moisture increments rarely exceeds 0.01 m^3m^{-3}, with a global average value of about 0.003 m^3m^{-3}. Finally, increments for the top layer soil temperature (Figure 21c) and surface (skin) temperature (not shown) also exhibit a pattern similar to that of the surface soil moisture increments, particularly in the northern Hemisphere for which the validation period coincides with the warmest months of the year. Surface soil temperature increments in the southern Hemisphere are smaller because the winter conditions imply a reduced diurnal cycle and thus a reduced variability of physical temperatures.

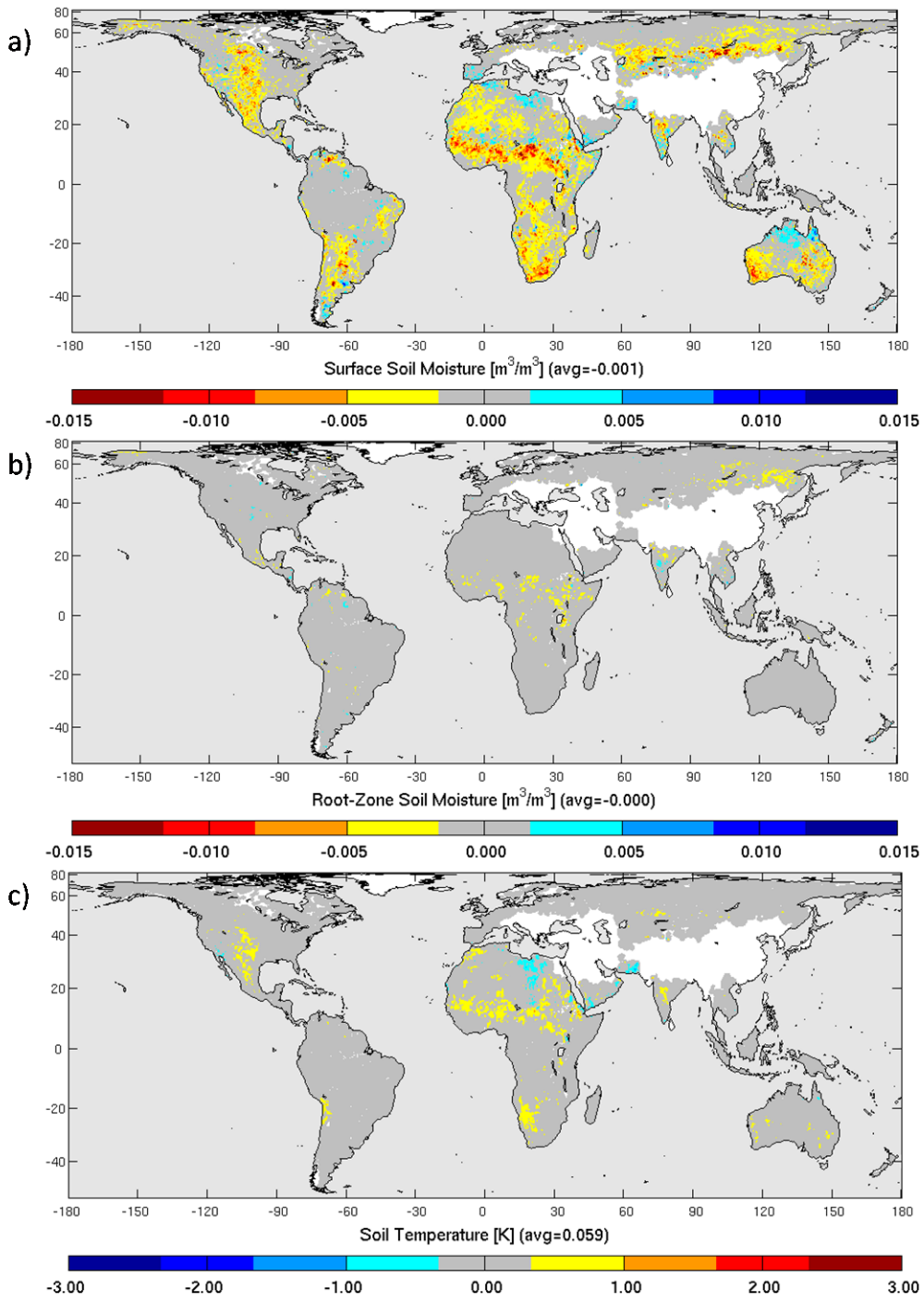

Figure 20. Time series mean of the increments for (a) surface soil moisture, (b) root zone soil moisture, and (c) top layer soil temperature from the L4_SM algorithm (Vb1004) for 11 Apr 2015, 0z, to 19 Sep 2015, 0z.

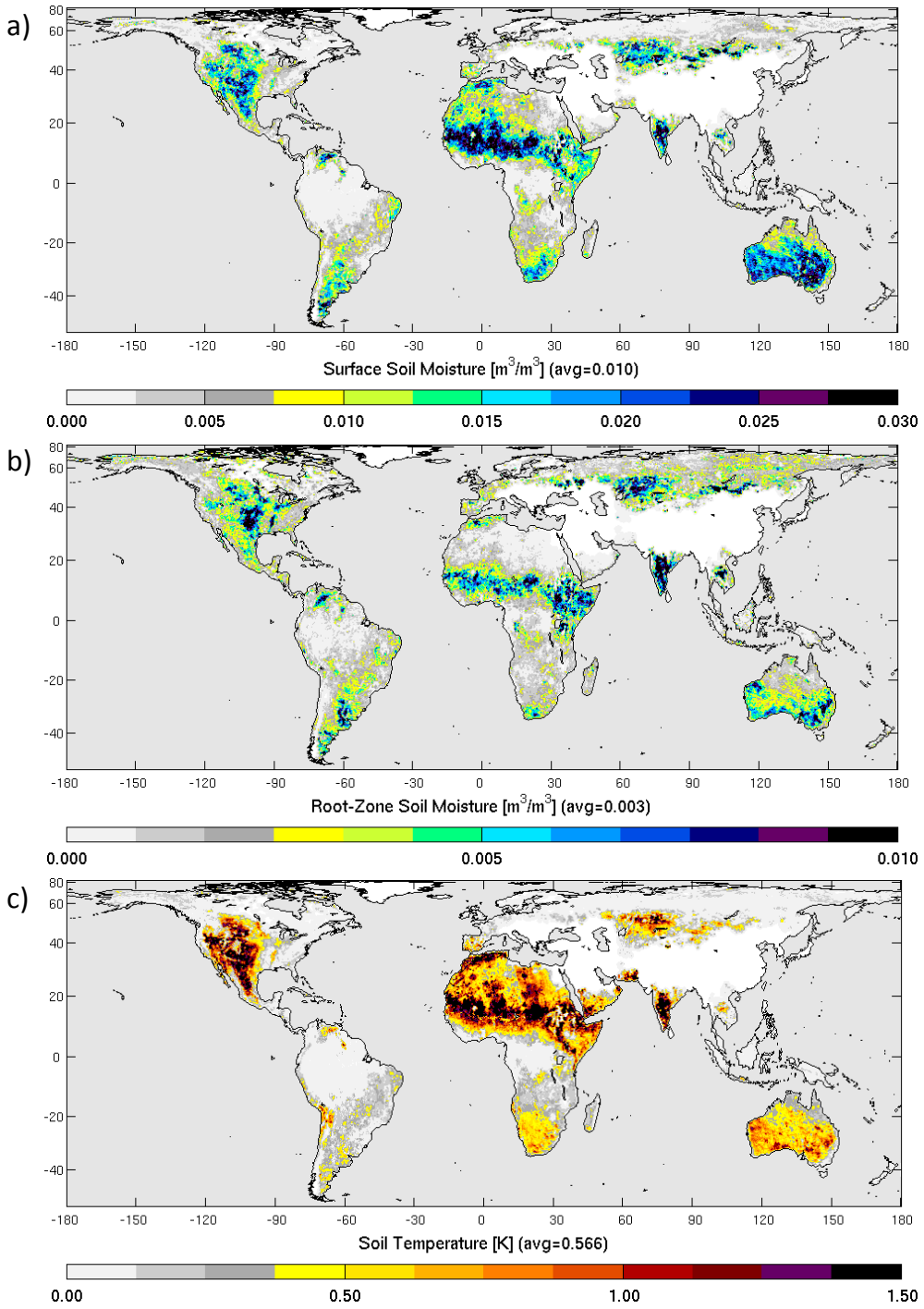

Figure 21. Same as Figure 20 but for time series standard deviation of the increments. Note the difference in colorscale for (a) and (b).

6.4.4 Uncertainty Estimates

The L4_SM data product also includes error estimates for key output variables, including surface and root zone soil moisture as well as surface soil temperature. These uncertainty estimates vary dynamically and geographically because they are computed as the standard deviation of a given output variable across the ensemble of land surface states at a given time and location. (The ensemble is an integral part of the ensemble Kalman filter employed in the L4_SM algorithm, and the ensemble mean provides the estimate of the variable under consideration.) By construction, the uncertainty estimates represent only the random component of the uncertainty. Bias and other structural errors such as errors in the dynamic range are not included.

Figure 22 shows the time mean of the uncertainty estimates for the validation period. Across the globe, surface soil moisture uncertainty typically ranges from 0.01 m^3m^{-3} to 0.03 m^3m^{-3}, with the larger uncertainties concentrated in the driest regions such as the Sahara desert. Uncertainty is also large where few or no SMAP brightness temperatures are assimilated in the L4_SM system so that the ensemble spread is never reduced through analysis updates. These regions include most of eastern Europe and the southern half of continental Asia where the lack of climatological information from SMOS prevents the assimilation of SMAP observations this early in the mission (section 6.4.2). This limitation will be removed later in the mission when a sufficient number of SMAP observations will be available to derive rescaling parameters based solely on SMAP information. Uncertainty is also high in the portions of the northern high-latitudes where few observations are assimilated, presumably because of frozen or snow-covered conditions during a considerable portion of the validation period.

Uncertainty in root zone soil moisture (Figure 22b) is generally smaller than for surface soil moisture, with typical values ranging from 0.005 m^3m^{-3} to 0.025 m^3m^{-3}. The global pattern of uncertainty in root zone soil moisture is quite different from that of surface soil moisture. The driest areas are associated with low values of uncertainty, because in arid regions the deeper layer soil moisture is mostly constant, with random errors in surface forcings (precipitation and radiation) having only a small impact. Uncertainty is highest in southern China, where root zone soil moisture is variable but SMAP observations cannot be assimilated. Finally, uncertainty in surface soil temperature (Figure 22c) ranges from 0.5 K to 2 K and is again largest in dry (and hot) regions and in regions where SMAP brightness temperatures are not assimilated.

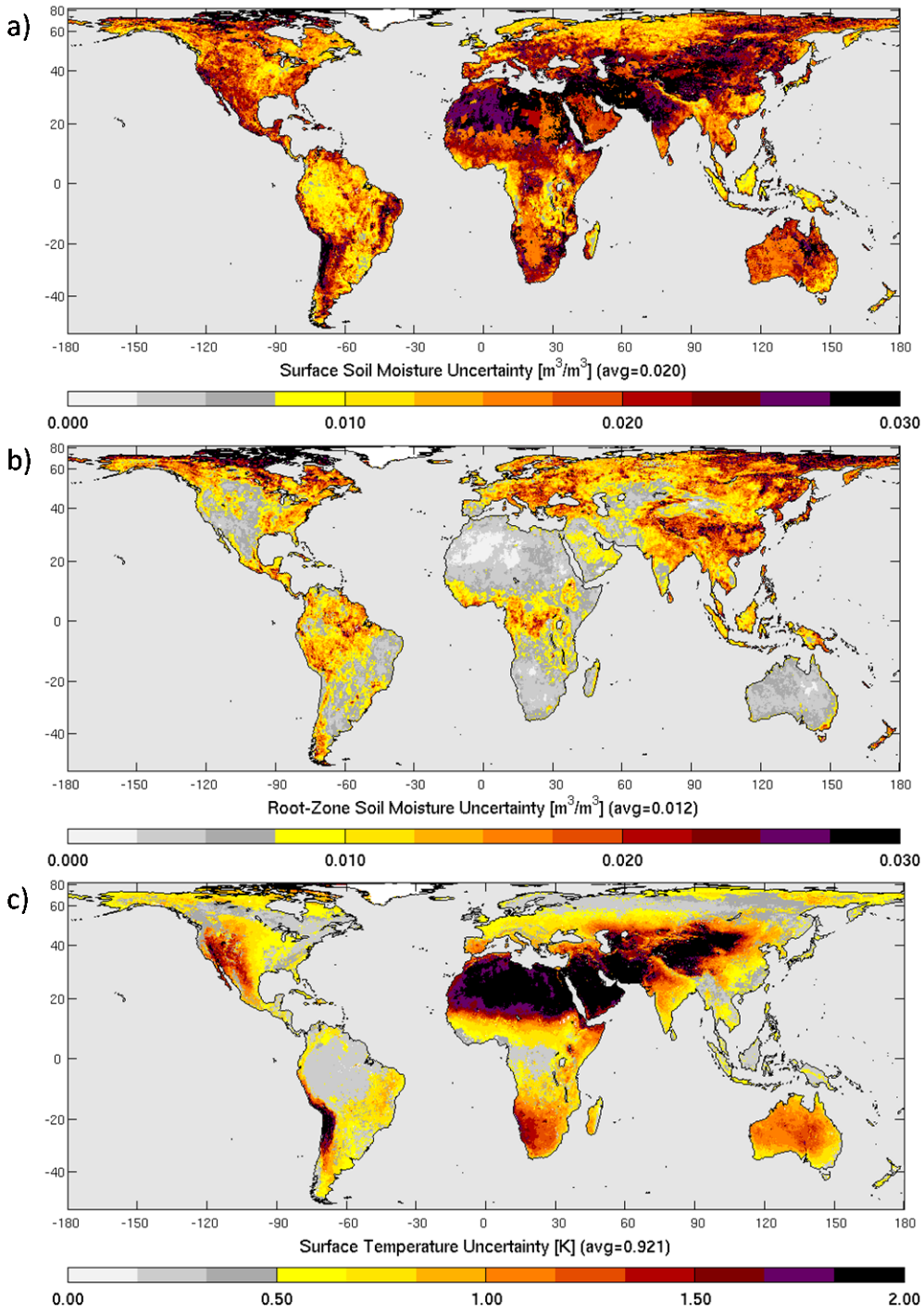

Figure 22. L4_SM uncertainty estimates for (a) surface soil moisture, (b) root zone soil moisture, and (c) top-layer soil temperature averaged across the validation period (11 April 2015, 0z to 19 September 2015, 0z). Uncertainty estimates are computed as the standard deviation across the model ensemble that is used in the L4_SM assimilation algorithm.

6.5 Summary

The SMAP L4_SM Vb1004 data product was validated using in situ soil moisture measurements from core validation sites and sparse networks. The product was further evaluated through an assessment of the data assimilation diagnostics generated by the L4_SM algorithm, such as the observation-minus-forecast residuals and the increments.

Based on the comparisons with the core validation site measurements, the L4_SM Vb1004 estimates of surface and root zone soil moisture meet the accuracy requirement (ubRMSE < 0.04 m^3m^{-3}). For surface soil moisture the ubRMSE is 0.036 m^3m^{-3} at the 9 km scale and 0.031 m^3m^{-3} at the 36 km scale. For root zone soil moisture, the ubRMSE is 0.023 m^3m^{-3} at the 9 km scale and 0.024 m^3m^{-3} at the 36 km scale.

The assimilation of SMAP brightness temperatures in the L4_SM algorithm is beneficial primarily for surface soil moisture estimates, where the improvements over the model-only SMAP Nature Run (NRv4) are consistent across the 9 km and 36 km scales and all three metrics (ubRMSE, bias, and R). Note, however, that the improvements are not statistically significant. For root zone soil moisture, L4_SM and NRv4 estimates have essentially the same skill.

The comparison with in situ measurements from sparse networks in the United States and Australia corroborate the results obtained for the core validation sites, thus moving the data product closer to compliance with Stage 2 Validation requirements.

The data assimilation diagnostics further broaden the validation to the global domain and indicate that the L4_SM system is reasonably unbiased in the global average sense. However, the observation-minus-forecast residuals of brightness temperature reveal a modest warm bias of a few Kelvin in the L4_SM product in select regions, which is reflected in a small negative (that is, dry) bias in the surface soil moisture increments of up to -0.01 m^3m^{-3} in those regions. The time mean analysis increments in root zone soil moisture and in the skin and surface soil temperatures are very small. The assimilation diagnostics further reveal that, on a regional basis, the actual errors in brightness temperature are typically over- or underestimated considerably by the L4_SM system, even though the global average of that diagnostic would suggest unbiased error estimates.

Uncertainty estimates for the analyzed surface soil moisture, root zone soil moisture, surface temperature, and top layer soil temperature are also provided with the product. These uncertainty estimates are designed to reflect the random error in key geophysical product fields. While the uncertainty estimates appear reasonable, it is not yet clear how well they reflect actual uncertainties.

It is important to keep in mind that the comparisons against the in situ measurements are impacted by the fact that the measurements themselves are prone to errors. The metrics presented here therefore underestimate the true skill of the product. In fact, the ubRMSE results presented in the report should be interpreted as the unbiased RMS *difference* between the model estimates and the in situ measurements, rather than errors with respect to actual (or true) soil moisture conditions.

Based on the results presented in this report, the public beta-release of the L4_SM data product is recommended. The results also uncovered limitations in the current version of the L4_SM data product and possible avenues for future development. These limitations and developments are addressed in the next section.

7 OUTLOOK AND PLAN FOR VALIDATED RELEASE

The assessment of the L4_SM product presented in the previous section revealed a number of current limitations as well as avenues for future development.

7.1 Bias and L4_SM Algorithm Calibration

Figure 17a showed that there are regions with a modest residual bias between the predicted brightness temperatures from the L4_SM modeling system and the SMAP observations. There are several potential causes for this. First, the surface meteorological forcing data that are used in the operational L4_SM system are based on data from the current GEOS-5 Forward Processing ("FP") system (GEOS-5.13; ¼ degree resolution). The corresponding retrospective SMAP Nature Run version 4 (NRv4) uses data from the GEOS-5 Reprocessing and Forward Processing for Instrument Teams ("RP/FP-IT") system (GEOS-5.9; ½ degree resolution). The vintage of the two systems is quite different, and the climatological parameters derived from NRv4 data do not necessarily reflect the true climatology of the L4_SM modeling component.

The second reason for the residual biases could be differences between SMAP brightness temperatures (beta-release version) and the SMOS v504 brightness temperatures that were used in the calibration of the microwave radiative transfer parameters of the L4_SM system.

To address the biases in the L4_SM algorithm, the calibration of the system will be revisited. This includes the development of a revised version of the SMAP Nature Run that is forced with the newly available MERRA-2 reanalysis data (GEOS-5.12; ½ degree resolution), which is closer to the current L4_SM forcing from the FP system. The new version of the Nature Run will also include a somewhat revised approach to precipitation corrections along with updates to the GEOS-5 land model and its ancillary data. Based on this new version of the Nature Run and the latest version of the SMOS brightness temperatures (v620), the parameters of the L4_SM microwave radiative transfer model will be recalibrated. The objective of these revisions is to provide an improved L4_SM modeling component along with retrospective data that is as consistent as possible with the present-day data in terms of its climatology.

Because SMOS observations are impacted by RFI and thus do not provide a climatology of L-band brightness temperatures in large parts of Asia and Europe, the L4_SM algorithm cannot assimilate SMAP data in those regions (Figure 15). This limitation manifests itself in two steps in the calibration of the L4_SM system. First, the L4_SM microwave radiative transfer model parameters cannot be calibrated locally. Second, the brightness temperature scaling parameters cannot be computed for the affected regions. The first limitation can be addressed by setting the microwave radiative transfer model parameters in the regions in question according to land cover (vegetation) class (De Lannoy et al. 2013, 2014). The second limitation is more difficult to overcome, as the residual seasonal model biases in brightness temperature cannot be addressed through the pentad rescaling factors used in the L4_SM algorithm. Ultimately, the solution is to calibrate the L4_SM system using a history of SMAP observations, but a robust calibration based solely on SMAP data can only be accomplished towards the end of the SMAP mission. In the meantime, we will investigate whether the L4_SM system can be provisionally calibrated in the RFI-affected regions by using land cover-based microwave radiative transfer model parameters along with a short history from SMAP to scale out the residual seasonal biases.

Another aspect of algorithm calibration involves the tuning of the L4_SM observation and model error parameters. Initial research into tuning the observation error parameters using a poor man's adaptive filter has been inconclusive. Such tuning resulted, by construction, in improved assimilation

diagnostics, but it has so far failed to consistently improve the skill metrics obtained from the comparisons against independent in situ observations.

7.2 Impact of SMAP Observations and Ensemble Perturbations

The assessment of the L4_SM product presented herein uses NRv4 estimates as the model-based reference. Both estimates use the same gauge-corrected precipitation forcing. By necessity, the assessment of the L4_SM (and NRv4) estimates versus in situ soil moisture observations is focused on regions for which the model forcing data take advantage of typically dense and reliable precipitation gauge observations. The generally good skill of the NRv4 estimates therefore leads to an underestimation of the impact of the SMAP brightness temperature observations in the L4_SM assimilation system. In regions with poor precipitation data, the impact of the SMAP observations should be larger, but because of the lack of observations of any kind in those regions, the precise impact remains unknown.

Secondly, the NRv4 estimates and the L4_SM estimates differ because the NRv4 estimates are from a single-member model run without perturbations, whereas the L4_SM estimates are based on an ensemble of model realizations that experiences perturbations to its model forcing and prognostic variables. An undesirable yet at this time unavoidable side effect of the perturbations regime is that it leads to biases between the ensemble mean estimates and the estimates from the unperturbed NRv4 model integration. This is particularly acute in very arid regions such as the Sahara desert, where the perturbations in soil moisture are, by construction, biased wet because the unperturbed, single-member model run typically remains at the lowest possible soil moisture value, thereby making negative (that is, drying) perturbations unphysical. Some of the differences between the NRv4 and L4_SM estimates will therefore partly reflect the impact of the perturbations regime rather than the use of SMAP observations.

To address these issues, the assessment will be refined by expanding the comparison of the L4_SM skill to that of additional model-only data sets. To better assess the impact of the SMAP observations, a model-only run will be conducted without the gauge-based precipitation corrections. Furthermore, an ensemble integration will be conducted in which the perturbations are applied but no SMAP data are assimilated. By comparing the L4_SM estimates to the different model-only runs, it will be possible to more clearly identify the impact of the SMAP observations.

7.3 Expanded Site Locations, Record Length, and Data Sets

The assessment of the beta release data is limited by the period of record. Only 5.5 months of data have been available for this assessment report. By the time of the validated L4_SM release in 2016, there will be one year of SMAP observations covering the full annual cycle in the Northern and Southern Hemispheres.

There are a number of in situ measurement sites that are currently classified as candidate validation sites. For several of these, the SMAP Project is only awaiting data delivery, which is not automated because of the remoteness of the locations (for example, Mongolia and Tibet). Measurements from other candidate validation sites require more processing and verification by the data providers (for example, Twente, Niger, and Benin). Several more sites are still under development and may become available within the time frame of the validated release (including Millbrook, Kuwait, and Bell Ville). It is, however, unlikely that any additional sites beyond those already known will be developed and implemented.

Sparse network validation will also be expanded. SMOSMania measurements are nearly ready for use in the L4_SM validation. Data from the COSMOS network have already been used provisionally, but further development is required before the results can be included in a formal report. Specifically, a better understanding of the variable measurement depth of the COSMOS probes is needed. Moreover, efforts are underway to complete the automated acquisition of data from additional networks, including networks in the United States, Argentina and South Africa.

Time and resources permitting, the L4_SM data product will also be evaluated by comparison with model-based estimates from other data providers, including the ERA-Land reanalysis data from the European Centre for Medium-Range Weather Forecasts and data products from Environment Canada.

7.4 L4_SM Algorithm Refinements

Despite its overall complexity, the L4_SM algorithm includes many simplifications. For example, SMAP brightness temperatures are not assimilated when the water fraction of the observed field-of-view exceeds a threshold of 5%. Rather than discarding such water-contaminated observations, the L4_SM algorithm could be refined to include the brightness temperatures of open water in its forward operator. This would require a dynamic model of the surface temperature of lakes and large rivers in the L4_SM modeling system, as well as a corresponding radiative transfer model, based, for example, on the model by Klein and Swift (1977). Due to the complexity of adding water surfaces into a modeling system that so far includes only land, it is not likely that such development is possible in time for the validated L4_SM release.

Another simplification in the current L4_SM algorithm is the use of flags in the L1 Radiometer product. Currently, the L4_SM system only checks the summary flag for the assimilated L1C_TB brightness temperatures. More specific binary flags, including dynamic surface flags, are available in the input data product and could perhaps be used in a refined L4_SM algorithm. Finally, the L4_SM algorithm uses the L1C_TB data product because of its convenient data format and posting on the 36 km EASE v2 grid. Since the L4_SM algorithm operates on the finer, 9 km EASE v2 grid, and since the SMAP radiometer greatly oversamples the brightness temperature relative to the size of its instantaneous field-of-view, the L4_SM algorithm might benefit from the direct assimilation of the time-ordered brightness temperatures from the L1B_TB data product.

As mentioned in section 5.2, the SMAP radar anomaly that occurred on 7 July 2015 means that radar-based data products from SMAP are only available for a short period at the beginning of the SMAP mission. The L4_SM algorithm originally included plans for a freeze-thaw analysis using the SMAP radar-based freeze-thaw retrievals. The SMAP Project is considering the development and generation of a freeze-thaw data product using the passive microwave observations. If an operational freeze-thaw product becomes available, the L4_SM algorithm could be expanded to assimilate it, provided adequate resources are available.

ACKNOWLEDGEMENTS

This report was made possible by the contributions of many individuals from the SMAP Project, the SMAP Science Team, the SMAP Cal/Val Partner Program, and the NASA Global Modeling and Assimilation Office. The NASA Soil Moisture Active Passive mission and the NASA Modeling, Analysis, and Prediction program supported the research. Computational resources were provided by the NASA High-End Computing Program through the NASA Center for Climate Simulation at the Goddard Space Flight Center.

REFERENCES

Colliander, A., S. Chan, N. Das, S. Kim, S. Dunbar, T. Jackson, C. Derksen, K. McDonald, J. Kimball, E. Njoku, R. Reichle, and B. Weiss (2014), SMAP L2-L4 Data Products Calibration and Validation Plan, Soil Moisture Active Passive (SMAP) Mission Science Document. JPL D-79463, Jet Propulsion Laboratory, Pasadena, CA.

CEOS (2015), Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV): http://calvalportal.ceos.org, CEOS WGCV Land Products Sub-Group: http://lpvs.gsfc.nasa.gov. Accessed 7 October 2015.

Dawdy, D., and N. Matalas (1964), Statistical and probability analysis of hydrologic data, part III: analysis of variance, covariance and time series, in *Handbook of Applied Hydrology: A Compendium of Water-Resources Technology*, pp. 8.68–8.91, McGraw-Hill, New York.

De Lannoy, G. J. M, R. H. Reichle, and V. R. N. Pauwels (2013), Global Calibration of the GEOS-5 L-band Microwave Radiative Transfer Model over Nonfrozen Land Using SMOS Observations, *Journal of Hydrometeorology, 14*, 765-785, doi:10.1175/JHM-D-12-092.1.

De Lannoy, G. J. M., R. H. Reichle, and J. A. Vrugt (2014), Uncertainty Quantification of GEOS-5 L-Band Radiative Transfer Model Parameters using Bayesian Inference and SMOS Observations, *Remote Sensing of Environment, 148*, 146-157, doi:10.1016/j.rse.2014.03.030.

Draper, C. S., R. H. Reichle, G. J. M. De Lannoy, and Q. Liu (2012), Assimilation of passive and active microwave soil moisture retrievals, *Geophysical Research Letters, 39*, L04401, doi:10.1029/2011GL050655.

Draper, C. S., R. H. Reichle, R. de Jeu, V. Naeimi, R. Parinussa, and W. Wagner (2013), Estimating root mean square errors in remotely sensed soil moisture over continental scale domains, *Remote Sensing of Environment, 137*, 288-298, doi:10.1016/j.rse.2013.06.013.

Entekhabi, D., R. H. Reichle, R. D. Koster, and W. T. Crow (2010), Performance Metrics for Soil Moisture Retrievals and Application Requirements, *Journal of Hydrometeorology, 11*, 832-840, doi:10.1175/2010JHM1223.1.

Entekhabi, D., and Coauthors (2014), SMAP Handbook, *JPL Publication, JPL 400-1567*, NASA Jet Propulsion Laboratory, Pasadena, California, USA, 182 pp.

Gruber, A., C.-H. Su, S. Zwieback, W. Crow, W. Dorigo, and W. Wagner (2015), Recent advances in (soil moisture) triple collocation analysis, *International Journal of Applied Earth Observation and Geoinformation*, in press, doi: 10.1016/j.jag.2015.09.002.

Jackson, T., D. Chen, M. Cosh, F. Li, M. Anderson, C. Walthall, P. Doraiswamy, and E. R. Hunt (2004), Vegetation water content mapping using Landsat data derived normalized difference water index (NDWI) for corn and soybean, *Remote Sensing of Environment*, *92*, 475-482.

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

Jackson, T. J., A. Colliander, J. Kimball, R. Reichle, C. Derksen, W. Crow, D. Entekhabi, P. O'Neill, and E. Njoku (2014), SMAP Science Data Calibration and Validation Plan (Revision A), Soil Moisture Active Passive (SMAP) Mission Science Document. JPL D-52544, Jet Propulsion Laboratory, Pasadena, CA.

Klein, L., and C. T. Swift (1977), An improved model for the dielectric constant of sea water at microwave frequencies, *IEEE Transactions on Antennas and Propagation*, 25, 104-111, doi:10.1109/TAP.1977.1141539.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure, *J. Geophys. Res., 105*, 24809–24822, doi:10.1029/2000JD900327.

Liu, Q., R. H. Reichle, R. Bindlish, M. H. Cosh, W. T. Crow, R. de Jeu, G. J. M. De Lannoy, G. J. Huffman, and T. J. Jackson (2011), The contributions of precipitation and soil moisture observations

to the skill of soil moisture estimates in a land data assimilation system, *Journal of Hydrometeorology*, *12*, 750-765, doi:10.1175/JHM-D-10-05000.1.

Piepmeier, J. R., and S. Chan (2015), Radiometer Brightness Temperature Calibration for the L1B_TB and L1C_TB Beta-Level Data Products, Soil Moisture Active Passive (SMAP) Mission Science Document. JPL-D93978, Jet Propulsion Laboratory, Pasadena, CA. Available at http://nsidc.org/data/docs/daac/smap/sp_l1b_tb/pdfs/L1B-L1C-Beta-Report.pdf.

Reichle, R. H., R. D. Koster, G. J. M. De Lannoy, B. A. Forman, Q. Liu, S. P. P. Mahanama, and A. Toure (2011), Assessment and enhancement of MERRA land surface hydrology estimates, *Journal of Climate*, *24*, 6322-6338, doi:10.1175/JCLI-D-10-05033.1.

Reichle, R. H., and Q. Liu (2014), Observation-Corrected Precipitation Estimates in GEOS-5, *NASA Technical Report Series on Global Modeling and Data Assimilation, NASA/TM-2014-104606, Vol. 35*, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA, 18pp.

Reichle, R. H., G. J. M. De Lannoy, B. A. Forman, C. S. Draper, and Q. Liu (2014a), Connecting Satellite Observations with Water Cycle Variables through Land Data Assimilation: Examples Using the NASA GEOS-5 LDAS, *Surveys in Geophysics*, *35*, 577-606, doi:10.1007/s10712-013-9220-8.

Reichle, R. H., R. Koster, G. De Lannoy, W. Crow, and J. Kimball (2014b), SMAP Level 4 Surface and Root Zone Soil Moisture Data Product: L4_SM Algorithm Theoretical Basis Document (Revision A), Soil Moisture Active Passive (SMAP) Mission Science Document. JPL D-66483, Jet Propulsion Laboratory, Pasadena, CA.

Reichle, R. H., R. A. Lucchesi, J. V. Ardizzone, G.-K. Kim, E. B. Smith, and B. H. Weiss (2015), Soil Moisture Active Passive (SMAP) Mission Level 4 Surface and Root Zone Soil Moisture (L4_SM) Product Specification Document. GMAO Office Note No. 11 (Version 1.4), 82 pp., NASA Goddard Space Flight Center, Greenbelt, MD, USA.

APPENDIX

A Validation Metrics Based on In Situ Measurements

This section outlines how the validation metrics are computed from the in situ measurements and the L4_SM soil moisture estimates. In situ measurements may be from SMAP core validation sites (section 6.2) or from SMAP sparse networks (section 6.3).

A.1 Long-term Average and Mean Seasonal Cycle

Let *s(i,t)* denote the SMAP product for time *t* and grid cell *i* (see Appendix C for SMAP "reference pixels"), and let *v(i,t)* denote the corresponding observation that is used for validation. Assume further that both the SMAP data product and the corresponding in situ measurements at grid cell *i* are available at $N_t(i)$ different times (excluding times and locations for which in situ measurements are eliminated in pre-processing; Appendix C). Then let $\bar{s}(i)$ denote the (long-term) average of *s(i,t)*, and let $\bar{v}(i)$ denote the corresponding (long-term) average of *v(i,t)*, that is:

$$\bar{s}(i) \equiv \frac{1}{N_t(i)} \sum_{\tau=1}^{N_t(i)} s(i,\tau) \tag{1}$$

$$\bar{v}(i) \equiv \frac{1}{N_t(i)} \sum_{\tau=1}^{N_t(i)} v(i,\tau) \tag{2}$$

The long-term average is undefined for a given grid cell *i* if $N_t(i,t)<N_{min}$, that is, if fewer than N_{min}=480 three-hourly data points are available to compute the long-term average for the grid cell in question.

Moreover, let $\bar{s}(i,t)$ and $\bar{v}(i,t)$ define the (long-term) mean seasonal cycles (denoted as periodic functions with a period of one year). These mean seasonal cycles are computed in two steps. First, 30-day moving average time series of *s(i,t)* and *v(i,t)* are computed:

$$s_{30}(i,t) \equiv \frac{1}{N_{30}(i,t)} \sum_{\tau=1}^{N_{30}(i,t)} s(i,\tau) \tag{3}$$

$$v_{30}(i,t) \equiv \frac{1}{N_{30}(i,t)} \sum_{\tau=1}^{N_{30}(i,t)} v(i,\tau) \tag{4}$$

where $N_{30}(i,t)$ is the number of data points at grid cell *i* within 15 days of time *t*. Next, the moving average time series are averaged across all available years, separately for each time-of-year:

$$\bar{s}(i,t) \equiv \frac{1}{N_a(i,t)} \sum_{\tau=1}^{N_a(i,t)} s_{30}(i,\tau) \tag{5}$$

$$(6) \qquad \bar{v}(i,t) \equiv \frac{1}{N_a(i,t)} \sum_{\tau=1}^{N_a(i,t)} v_{30}(i,\tau)$$

where $N_a(i,t)$ is the number of years for which $s_{30}(i,t)$ and $v_{30}(i,t)$ are available at grid cell i and for the time-of-year corresponding to time t. (Recall that $\bar{s}(i,t)$ and $\bar{v}(i,t)$ are denoted as periodic functions with a period of one year). The mean seasonal cycle is undefined for a given time-of-year at grid cell i if $N_{30}(i,t)<N_{min}$, that is, if fewer than N_{min}=480 three-hourly data points are available to compute the 30-day moving average for the time and grid cell in question. Furthermore, the mean seasonal cycle is also undefined for $N_a(i,t)<3$, that is, if the running mean values for a given time-of-year and grid cell are available for fewer than three different years. The mean seasonal cycle during the SMAP Cal/Val phase (approximately the first year after launch) could be computed using in situ measurements taken prior to the SMAP launch in combination with a retrospective land model-only L4_SM prototype product.

A.2 Bias and Root-Mean-Square Error

For each grid cell i (see Appendix C for SMAP validation grid cells or "reference pixels") with a sufficient number of in situ measurements, the bias and unbiased root-mean-square error (*ubRMSE*) metrics are computed as follows:

$$(7) \qquad Bias(i) \equiv \bar{s}(i) - \bar{v}(i)$$

$$(8) \qquad ubRMSE(i) \equiv \sqrt{\frac{1}{N_t(i)} \left(\sum_{t=1}^{N_t(i)} \left[(s(i,t) - \bar{s}(i)) - (v(i,t) - \bar{v}(i)) \right]^2 \right)}$$

The metrics are undefined for a given grid cell i if $N_t(i,t)<N_{min}$, that is, if fewer than N_{min}=480 three-hourly data points are available.

Note that the root-mean-square error (*RMSE*) is related to the *bias* and the *ubRMSE* as follows:

$$(9) \qquad [RMSE(i)]^2 \equiv \frac{1}{N_t(i)} \left(\sum_{t=1}^{N_t(i)} \left[s(i,t) - v(i,t) \right]^2 \right) = [Bias(i)]^2 + [ubRMSE(i)]^2$$

A.3 Time Series Correlation and Anomaly Time Series Correlation

For each grid cell i (Appendix C) with a sufficient number of in situ measurements, the time series correlation $R(i)$ and the anomaly time series correlation $R_{anom}(i)$ metrics are computed as follows:

$$(10) \qquad R(i) \equiv \frac{\left(\sum_{t=1}^{N_t(i)} \left[(s(i,t) - \bar{s}(i))(v(i,t) - \bar{v}(i)) \right] \right)}{\sqrt{\left(\sum_{t=1}^{N_t(i)} \left[(s(i,t) - \bar{s}(i)) \right]^2 \right)} \sqrt{\left(\sum_{t=1}^{N_t(i)} \left[(v(i,t) - \bar{v}(i)) \right]^2 \right)}}$$

$$(11) \qquad R_{anom}(i) \equiv \frac{\left(\sum_{t=1}^{N_{t,anom}(i)} \left[(s(i,t) - \bar{s}(i,t))(v(i,t) - \bar{v}(i,t)) \right] \right)}{\sqrt{\left(\sum_{t=1}^{N_{t,anom}(i)} \left[(s(i,t) - \bar{s}(i,t)) \right]^2 \right)} \sqrt{\left(\sum_{t=1}^{N_{t,anom}(i)} \left[(v(i,t) - \bar{v}(i,t)) \right]^2 \right)}}$$

The R and R_{anom} metrics are undefined for a given grid cell i if $N_t(i,t)<N_{min}$ or $N_{t,anom}(i,t)<N_{min}$, respectively, that is, if fewer than N_{min}=480 three-hourly data points are available. Note that $N_{t,anom}(i,t)$ is also determined by availability of the mean seasonal cycle estimates, which are themselves subject to a minimum data requirement (Appendix A.1).

Note that the *bias*, *ubRMSE*, and *R* metrics are related (Entekhabi et al. 2010; their equation (5)).

A.4 Confidence Intervals

The *bias*, *ubRMSE*, *R*, and R_{anom} metrics discussed in the previous subsections are supplemented with 95% confidence intervals, based on a Student T-distribution for the *bias*, a χ^2 distribution for the *ubMSE*, and an asymptotic normal distribution for R and R_{anom} after a Fisher transformation. The confidence intervals are first computed separately for each sparse network site or core validation site reference pixel. Since errors in soil moisture (anomaly) time series are highly auto-correlated, the number of independent data in the time series is smaller than the length of the time series. Hence, the confidence intervals are calculated using an effective sample size (Dawdy and Matalas 1964; Draper et al. 2012).

Confidence intervals for metrics that represent average skill across several sites are then computed as follows. For core validation sites, the confidence intervals across several references pixels are simply averaged. For sparse network sites, the average is computed using a k-means spatial clustering algorithm that discounts the weight of individual sensors in areas where the sensor density is large. Within each cluster, an average confidence interval is computed by taken the straight average of the confidence intervals at individual sensors within the given cluster, reflecting the fact that estimates from nearby sensors are not independent. Finally, the confidence interval for the average metric in question is computed by averaging the confidence intervals of the N_c clusters and then dividing that average by the square root of N_c, reflecting the assumption that confidence intervals for different clusters are fairly independent.

B Validation Metrics Based on Assimilation Diagnostics

The L4_SM product is the result of the assimilation of SMAP brightness temperatures into a land surface model. The assimilated brightness temperature observations along with the corresponding model forecasts are provided in the L4_SM "aup" output Collection. Differencing the two quantities yields the *observation-minus-forecast (OmF)* brightness temperature residuals. Moreover, the "aup" output Collection also provides information on the soil moisture and soil temperature analysis increments. The statistics of the *OmF* residuals and the analysis increments are computed as part of the validation of the L4_SM product.

B.1 Statistics of the Observation-Minus-Forecast Residuals

The *OmF* brightness temperature residuals are computed by differencing the assimilated brightness temperature observations and the corresponding model forecasts from the "aup" file Collection:

$$(12) \quad OmF(i,t,p,r) = tb\_[p]\_obs\_assim(i,t) - tb\_[p]\_forecast(i,t)$$

where *i* denotes the grid cell, *t* denotes the time, *p* denotes the polarization (H- or V-pol), and *r* denotes the effective resolution of the assimilated brightness temperatures (either 36 km if the observations originate from the SMAP L1C_TB product or 9 km if the observations originate from the L2_SM_AP product). The right-hand-side variable names follow the notation in the SMAP L4_SM Data Product Specification Document (Reichle et al. 2015).

The *OmF* brightness temperature residuals are normalized through division by their expected standard deviation:

$$(13) \quad OmF\_norm(i,t,p,r) = OmF(i,t,p,r) \,/\, sqrt(tb\_[p]\_errstd(i,t)^2 + tb\_[p]\_forecast\_ensstd(i,t)^2)$$

where *tb_[p]_errstd* and *tb_[p]_forecast_ensstd* are the observation and forecast error standard deviations, respectively. Data assimilation theory suggests that for an optimally configured system, the *OmF* residuals have a mean of zero, and the (normalized) *OmF_norm* residuals have a variance of one.

B.2 Statistics of the Analysis Increments

The analysis increments are computed by differencing select variables from the "aup" file Collection:

(14) *[varname]_incr(i,t) = [varname]_analysis – [varname]_forecast*

where *[varname]* is one of the five geophysical soil moisture and soil temperature variables, that is, *[varname]= {"sm_surface", "sm_rootzone", "sm_profile", "surface_temp", or "soil_temp_layer1"}* (following the notation in the SMAP L4_SM Data Product Specification Document; Reichle et al. 2015).

Note that the *spatial* mean and variability metrics of the *OmF* residuals and the analysis increments are provided in the L4_SM "quality assurance" granules (*.qa files).

C Preprocessing, Quality Control, and Upscaling of In Situ Measurements

In situ measurements from SMAP core validation sites and sparse networks undergo the following pre-processing and quality control steps:

1) Eliminate soil moisture measurements outside of the physically meaningful range defined by site-specific porosity or texture data where possible, and [0, 0.6] m^3m^{-3} otherwise.

2) Eliminate soil moisture measurements when the corresponding in situ soil temperature measurement is below 4°C.

3) Eliminate soil moisture measurements that are inconsistent with site-specific precipitation measurements (as permitted by availability of precipitation measurements and resources).

4) Eliminate soil moisture measurements that exhibit obviously unphysical characteristics. Such characteristics include spikes, sudden dry-downs (discontinuities), climatological shifts (typically related to changes in sensor calibration), and high-frequency oscillations. These checks are implemented as automatic detection tools to be run routinely, combined with visual inspection as resources permit.

5) Vertically aggregate in situ measurements to approximate the 0-100 cm layer depth of the L4_SM root zone soil moisture product. Vertical averages are weighted by the vertical spacing between the individual sensor depths and the boundaries of the root zone layer at the surface and at 100 cm depth.

6) Eliminate soil moisture measurements when the L4_SM product indicates the presence of snow (snow water equivalent > 0 kg m^{-2}) or the L4_SM soil temperature is below 4°C. This step is equivalent to the use of QC flags in the validation of the Level 2 soil moisture products.

The distributed in situ measurements from select individual core validation site sensors are aggregated into soil moisture "measurements" at the 9 km (or 36 km) scale for so-called *reference pixels*. The definition of the reference pixels is based on the characteristics of each site, including the layout of the network and land cover. This results in reference pixels that are not necessarily aligned with standard EASEv2 grid cells. The default upscaling function for the SMAP core validation site measurements is to use equal weights when averaging the sensor measurements within a reference pixel (separately for each sensor depth). If provided by the SMAP Cal/Val Partners, more sophisticated upscaling functions are used.

For core validation sites, the quality checks that are based on the in situ measurements (steps 1-4 above) are conducted prior to the aggregation using the individual sensor measurements. The vertical aggregation into "root zone" measurements (step 5) and the model-based quality checks (step 6) are applied to the horizontally aggregated in situ measurements (that is, at the scale of the reference pixels).

Previous Volumes in This Series

Volume 1 *September 1994*	Documentation of the Goddard Earth Observing System (GEOS) general circulation model - Version 1 **L.L. Takacs, A. Molod, and T. Wang**
Volume 2 *October 1994*	Direct solution of the implicit formulation of fourth order horizontal diffusion for gridpoint models on the sphere **Y. Li, S. Moorthi, and J.R. Bates**
Volume 3 *December 1994*	An efficient thermal infrared radiation parameterization for use in general circulation models **M.-D. Chou and M.J. Suarez**
Volume 4 *January 1995*	Documentation of the Goddard Earth Observing System (GEOS) Data Assimilation System - Version 1 **James Pfaendtner, Stephen Bloom, David Lamich, Michael Seablom, Meta Sienkiewicz, James Stobie, and Arlindo da Silva**
Volume 5 *April 1995*	Documentation of the Aries-GEOS dynamical core: Version 2 **Max J. Suarez and Lawrence L. Takacs**
Volume 6 *April 1995*	A Multiyear Assimilation with the GEOS-1 System: Overview and Results **Siegfried Schubert, Chung-Kyu Park, Chung-Yu Wu, Wayne Higgins, Yelena Kondratyeva, Andrea Molod, Lawrence Takacs, Michael Seablom, and Richard Rood**
Volume 7 *September 1995*	Proceedings of the Workshop on the GEOS-1 Five-Year Assimilation **Siegfried D. Schubert and Richard B. Rood**
Volume 8 *March 1996*	Documentation of the Tangent Linear Model and Its Adjoint of the Adiabatic Version of the NASA GEOS-1 C-Grid GCM: Version 5.2 **Weiyu Yang and I. Michael Navon**
Volume 9 *March 1996*	Energy and Water Balance Calculations in the Mosaic LSM **Randal D. Koster and Max J. Suarez**
Volume 10 *April 1996*	Dynamical Aspects of Climate Simulations Using the GEOS General Circulation Model **Lawrence L. Takacs and Max J. Suarez**

Volume 11 *May 1997*	Documentation of the Tangent Linear and its Adjoint Models of the Relaxed Arakawa-Schubert Moisture Parameterization Package of the NASA GEOS-1 GCM (Version 5.2) **Weiyu Yang, I. Michael Navon, and Ricardo Todling**
Volume 12 *August 1997*	Comparison of Satellite Global Rainfall Algorithms **Alfred T.C. Chang and Long S. Chiu**
Volume 13 *December 1997*	Interannual Variability and Potential Predictability in Reanalysis Products **Wie Ming and Siegfried D. Schubert**
Volume 14 *August 1998*	A Comparison of GEOS Assimilated Data with FIFE Observations **Michael G. Bosilovich and Siegfried D. Schubert**
Volume 15 *June 1999*	A Solar Radiation Parameterization for Atmospheric Studies **Ming-Dah Chou and Max J. Suarez**
Volume 16 *November 1999*	Filtering Techniques on a Stretched Grid General Circulation Model **Lawrence Takacs, William Sawyer, Max J. Suarez, and Michael S. Fox-Rabinowitz**
Volume 17 *July 2000*	Atlas of Seasonal Means Simulated by the NSIPP-1 Atmospheric GCM **Julio T. Bacmeister, Philip J. Pegion, Siegfried D. Schubert, and Max J. Suarez**
Volume 18 *December 2000*	An Assessment of the Predictability of Northern Winter Seasonal Means with the NSIPP1 AGCM **Philip J. Pegion, Siegfried D. Schubert, and Max J. Suarez**
Volume 19 *July 2001*	A Thermal Infrared Radiation Parameterization for Atmospheric Studies **Ming-Dah Chou, Max J. Suarez, Xin-Zhong, and Michael M.-H. Yan**
Volume 20 *August 2001*	The Climate of the FVCCM-3 Model **Yehui Chang, Siegfried D. Schubert, Shian-Jiann Lin, Sharon Nebuda, and Bo-Wen Shen**
Volume 21 *September 2001*	Design and Implementation of a Parallel Multivariate Ensemble Kalman Filter for the Poseidon Ocean General Circulation Model **Christian L. Keppenne and Michele M. Rienecker**
Volume 22 *August 2002*	Coupled Ocean-Atmosphere Radiative Model for Global Ocean Biogeochemical Models **Watson W. Gregg**

Volume 23 *November 2002*	Prospects for Improved Forecasts of Weather and Short-term Climate Variability on Subseasonal (2-Week to 2-Month) Time Scales **Siegfried D. Schubert, Randall Dole, Huang van den Dool, Max J. Suarez, and Duane Waliser**
Volume 24 *July 2003*	Temperature Data Assimilation with Salinity Corrections: Validation for the NSIPP Ocean Data Assimilation System in the Tropical Pacific Ocean, 1993–1998 **Alberto Troccoli, Michele M. Rienecker, Christian L. Keppenne, and Gregory C. Johnson**
Volume 25 *December 2003*	Modeling, Simulation, and Forecasting of Subseasonal Variability **Duane Waliser, Siegfried D. Schubert, Arun Kumar, Klaus Weickmann, and Randall Dole**
Volume 26 *April 2005*	Documentation and Validation of the Goddard Earth Observing System (GEOS) Data Assimilation System – Version 4 **Senior Authors: S. Bloom, A. da Silva and D. Dee** **Contributing Authors: M. Bosilovich, J-D. Chern, S. Pawson, S. Schubert, M. Sienkiewicz, I. Stajner, W-W. Tan, and M-L. Wu**
Volume 27 *December 2008*	The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. **M.M. Rienecker, M.J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H.-C. Liu, W. Gu, M. Sienkiewicz, R.D. Koster, R. Gelaro, I. Stajner, and J.E. Nielsen**
Volume 28 *April 2012*	The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna **Andrea Molod, Lawrence Takacs, Max Suarez, Julio Bacmeister, In-Sun Song, and Andrew Eichmann**
Volume 29 *May 2012*	Atmospheric Reanalyses – Recent Progress and Prospects for the Future. A Report from a Technical Workshop, April 2010 **Michele M. Rienecker, Dick Dee, Jack Woollen, Gilbert P. Compo, Kazutoshi Onogi, Ron Gelaro, Michael G. Bosilovich, Arlindo da Silva, Steven Pawson, Siegfried Schubert, Max Suarez, Dale Barker, Hirotaka Kamahori, Robert Kistler, and Suranjana Saha**
Volume 30 *September 2012*	The GEOS-ODAS, description and evaluation **Guillaume Vernieres, Michele M. Rienecker, Robin Kovach and Christian L. Keppenne**

Volume 31 *March 2013*	Global Surface Ocean Carbon Estimates in a Model Forced by MERRA **Watson W. Gregg, Nancy W. Casey and Cecile S. Rousseaux**
Volume 32 **March 2014**	Estimates of AOD Trends (2002-2012) over the World's Major Cities based on the MERRA Aerosol Reanalysis **Simon Provencal, Pavel Kishcha, Emily Elhacham, Arlindo M. da Silva, and Pinhas Alpert**
Volume 33 **August 2014**	The Effects of Chlorophyll Assimilation on Carbon Fluxes in a Global Biogeochemical Model **Cécile S. Rousseaux and Watson W. Gregg**
Volume 34 **September 2014**	Background Error Covariance Estimation using Information from a Single Model Trajectory with Application to Ocean Data Assimilation into the GEOS-5 Coupled Model **Christian L. Keppenne, Michele M. Rienecker, Robin M. Kovach, and Guillaume Vernieres**
Volume 35 **December 2014**	Observation-Corrected Precipitation Estimates in GEOS-5 **Rolf H. Reichle and Qing Liu**
Volume 36 **March 2015**	Evaluation of the 7-km GEOS-5 Nature Run **Ronald Gelaro, William M. Putman, Steven Pawson, Clara Draper, Andrea Molod, Peter M. Norris, Lesley Ott, Nikki Prive, Oreste Reale, Deepthi Achuthavarier, Michael Bosilovich, Virginie Buchard, Winston Chao, Lawrence Coy, Richard Cullather, Arlindo da Silva, Anton Darmenov, Ronald M. Errico, Marangelly Fuentes, Min-Jeong Kim, Randal Koster, Will McCarty, Jyothi Nattala, Gary Partyka, Siegfried Schubert, Guillaume Vernieres, Yuri Vikhliaev, and Krzysztof Wargan**
Volume 37 **March 2015**	Maintaining Atmospheric Mass and Water Balance within Reanalysis **Lawrence L. Takacs, Max Suarez, and Ricardo Todling**

Volume 38
September 2015

The Quick Fire Emissions Dataset (QFED) – Documentation of versions 2.1, 2.2 and 2.4

Anton S. Darmenov and Arlindo da Silva

Volume 39
September 2015

Land Boundary Conditions for the Goddard Earth Observing System Model Version 5 (GEOS-5) Climate Modeling System – Recent Updates and Data File Descriptions

Sarith P. Mahanama, Randal D. Koster, Gregory K. Walker, Lawrence L. Takacs, Rolf H. Reichle, Gabrielle De Lannoy, Qing Liu, Bin Zhao, and Max J. Suarez