Improving the SMAP Level-4 Soil Moisture Product

Rolf Reichle*, Qing Liu, and Randy Koster

NASA Global Modeling and Assimilation Office, NASA/GSFC, Greenbelt, MD

*Rolf.Reichle@nasa.gov, 301-614-5693
1. Motivation and Overview
2. In Situ Validation
3. Assimilation Diagnostics
4. Summary and Outlook
Motivation

Key Objectives of the Level-4 Soil Moisture (L4_SM) product:

1. “Root-zone” soil moisture (0-100 cm)
2. Spatially & temporally complete
Algorithm Overview

Precipitation observations → Catchment Model 9-km → NWP surface meteorology → Ensemble KF → L4_SM Product: 9-km, 3-hourly, global, 2.5-day latency → SMAP observations 36-km brightness temperatures

Surface & root-zone soil moisture, soil temperature, snow, surface fluxes, surface met. forcing. 
Brightness temp. (obs & modeled), assimilation diagnostics, uncertainty estimates. 
Land model constants.
### Key Changes in Version 4

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ancillary data</strong></td>
<td>Improved land cover (MODIS+Geoland), topography (SRTM) and veg. height (Lidar). &lt;br&gt; Longer L-band &amp; forcing time series to derive model climatology and Tb scaling parameters. &lt;br&gt; New SMAP Tb calibration (3-4 K over land!).  &lt;br&gt; Rescaled background precipitation to GPCP climatology (Africa, high latitudes).</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td>Reduced upward recharge of surface soil moisture under non-equilibrium conditions.  &lt;br&gt; Revised treatment of surface energy balance (impact on surface soil temperature).</td>
</tr>
<tr>
<td><strong>Analysis</strong></td>
<td>Removed “catchment deficit” from EnKF state vector.</td>
</tr>
<tr>
<td><strong>Metadata</strong></td>
<td>Added “projection coordinates” for improved interoperability (ArcGIS, OPeNDAP).</td>
</tr>
</tbody>
</table>

---

**Note:**

- Improved land cover (MODIS+Geoland), topography (SRTM), and vegetation height (Lidar).
- Longer L-band and forcing time series to derive model climatology and Tb scaling parameters.
- New SMAP Tb calibration (3-4 K over land).
- Rescaled background precipitation to GPCP climatology (Africa, high latitudes).
- Reduced upward recharge of surface soil moisture under non-equilibrium conditions.
- Revised treatment of surface energy balance (impact on surface soil temperature).
- Removed “catchment deficit” from EnKF state vector.
- Added “projection coordinates” for improved interoperability (ArcGIS, OPeNDAP).
Version 4 minus Version 3

- Version 4 surface soil moisture slightly drier in most regions because of reduced upward recharge.
- No change in global average root zone soil moisture.
- Changes in Africa and high latitudes owing to rescaling of GEOS precip. to GPCPv2.2 climatology.
1. Motivation and Overview
2. In Situ Validation
3. Assimilation Diagnostics
4. Summary and Outlook
In Situ Validation

• Validation period: Apr 2015 – Mar 2018 (3 years) – unless noted otherwise

• Core sites provide locally dense in situ networks in 18 watersheds.
• Sparse networks provide point-scale measurements at hundreds of locations.

• Compare to model-only simulation (without assimilation of SMAP Tbs):
  • NRv4.1 is model for Version 3.
  • NRv7.2 is model for Version 4.
33 km Core Site at Little River, Georgia, USA

Porosity = 0.41 m³ m⁻³, wilting point = 0.11 m³ m⁻³, clay fraction = 0.07, sand fraction = 0.77

Vv3030: uRMSE = 0.041 m³ m⁻³, bias = 0.037 m³ m⁻³, R = 0.62
Tv4000: uRMSE = 0.037 m³ m⁻³, bias = 0.009 m³ m⁻³, R = 0.74

Vv3030: uRMSE = 0.030 m³ m⁻³, bias = 0.071 m³ m⁻³, R = 0.70
Tv4000: uRMSE = 0.028 m³ m⁻³, bias = 0.071 m³ m⁻³, R = 0.65
• Both versions meet accuracy requirement (ubRMSE < 0.04 m³/m³).
• Both versions meet accuracy requirement (ubRMSE < 0.04 m$^3$/m$^3$).
• Compared to model-only estimates (NR[x]), ubRMSE and correlation metrics in both versions are improved.
Both versions meet accuracy requirement (ubRMSE < 0.04 m$^3$/m$^3$).

Compared to model-only estimates (NR[x]), ubRMSE and correlation metrics in both versions are improved.


Surface soil moisture bias smaller in Version 4 than in Version 3, but opposite holds for root zone bias.

Correlation metrics unchanged between versions.
Soil Moisture Skill

- ubRMSE v.core sites (M09): Number of sites: Surface = 18, Root zone = 6
- Bias v.core sites (M09): Number of sites: Surface = 18, Root zone = 6
- R v.core sites (M09): Number of sites: Surface = 18, Root zone = 6
- anomR v.core sites (M09): Number of sites: Surface = 11, Root zone = 6

NRv4.1, Version 3, NRv7.2, Version 4

- ubRMSE v. 5 sparse networks soil moisture (all): N sites: Surface = 427, Root zone = 300
- Bias v. 5 sparse networks soil moisture (all): N sites: Surface = 427, Root zone = 300
- R v. 5 sparse networks soil moisture (all): N sites: Surface = 427, Root zone = 300
- anomR v. 5 sparse networks soil moisture (all): N sites: Surface = 369, Root zone = 231

Sparse Networks
Reduction of upward recharge into surface layer was calibrated using sparse network measurements. Improvements did not translate to core site metrics.
Runoff Validation

- On average, the model generates too little runoff.
- Mean runoff is better in NRv7.2 than NRv4.1.
- Model improvements do not translate into better Version 4 product.
1. Motivation and Overview
2. In Situ Validation
3. Assimilation Diagnostics
4. Summary and Outlook
Beginning with Version 3, brightness temperature scaling parameters are based on SMAP data where SMOS climatology is unavailable due to RFI.
Water balance (nearly) closes after considering analysis increments and impact of perturbations.
Version 4 has smaller “natural” imbalance and smaller profile s. m. increments than Version 3.
Version 4 is nearly bias-free in global average, but has slightly larger typical bias magnitude.
Version 4 better able to forecast Tb (possibly helped by better obs).
• Normalize O-Fs with (assumed) error std-devs supplied to the analysis.
• Version 4 better (less under-estimation) in regions where Tb analysis impacts soil moisture.
• Version 4 worse (more over-estimation) in forested regions (where Tb provides less information on soil moisture).
Std-dev Increments

Version 3

Surface Soil Moisture [mm/day] (avg=0.073)

Root-Zone Soil Moisture [mm/day] (avg=0.103)

Version 4

Surface Soil Moisture [mm/day] (avg=0.091)

Root-Zone Soil Moisture [mm/day] (avg=0.051)

Larger than in Version 3

Smaller than in Version 3
L4_SM provides uncertainty estimates (“ensemble std-dev”) for surface and root-zone soil moisture. These estimates should characterize the actual errors in the L4_SM product (“ubRMSE”).

Version 4 uncertainty estimates:

- better capture the average ubRMSE than in Version 3
- but are still not (spatially) correlated with ubRMSE.
Based on original climatology file

Original climatology file is based on NRv7.2 only.

Revised climatology file is adjusted for 3-yr seasonally varying mean difference between Version 4 and NRv7.2a (forced with MERRA-2 during SMAP period).

This corrects for:

1) the discontinuity between retrospective (MERRA-2) and current (GEOS FP) forcing, and

2) the effect of ensemble perturbations.
Summary

The L4_SM algorithm assimilates SMAP brightness temperature (Tb) observations into the NASA Catchment model using a distributed (3d) EnKF. The L4_SM product provides global, 9-km, 3-hourly estimates with ~2.5-day latency. The L4_SM algorithm now also assimilates SMAP Tbs in RFI-prone regions. The L4_SM soil moisture meets accuracy requirements (ubRMSE<0.04 m³ m⁻³).

Compared to Version 3, Version 4 has:

- slightly larger surface soil moisture ubRMSE,
- generally drier surface soil moisture and larger differences in Africa and high-lats,
- larger surface s. m. increments and smaller root-zone and profile s. m. increments,
- improved Tb model forecasts (smaller O-F std-dev),
- better representation of actual errors in North American and Eurasian plains,
- larger (more realistic) uncertainty estimates,
- still no correlation between uncertainty estimates and actual errors,
- no improvement in runoff skill.

Perturbations make up non-negligible fraction of water balance in desert regions.
Mismatch in layer depths (L4_SM: 0-5 cm; in situ: ~3-7 cm) and in situ measurement errors adversely impact the validation and result in over-estimated ubRMSE values.

Preliminary reprocessing stream (using L1C Tb test inputs) should complete in July.

Future work:
• Repeat reprocessing (by November 2018) using re-derived Tb scaling parameters based on published L1C Tb reprocessed data and microwave RTM parameters calibrated to NRv7.2.
• Explore assimilation of enhanced resolution and/or water-corrected Tbs.
EXTRA SLIDES
The L4_SM climatology changes significantly from Version 3 to Version 4. Do not mix versions.

<table>
<thead>
<tr>
<th>L4_SM Stream</th>
<th>L1C_TB Inputs</th>
<th>Data Period</th>
<th>Production Status (Wall Date)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vv4010</td>
<td>R16 ops (T15x test data for Tb scaling)</td>
<td>5 Jun 2018 – …</td>
<td>ongoing</td>
<td>Version 4 ops (initial stream)</td>
</tr>
</tbody>
</table>
Algorithm Overview

Precipitation observations → NWP surface meteorology → SMAP observations

SMAP observations: 36-km brightness temperatures

Land Model: 9-km

Multi-layer soil water and energy balance modeling

Data assimilation

Ensemble Kalman Filter (EnKF):
Optimal update of land model’s soil moisture states using error covariance statistics sampled from a 24-member ensemble.

Surface & root-zone soil moisture, soil temperature, snow, surface fluxes, surface met. forcing.
Brightness temp. (obs & modeled), assimilation diagnostics, uncertainty estimates.
Land model constants.

L4_SM Product: 9-km, 3-hourly, global, 2.5-day latency
Land Modeling System

Soil Moisture Analysis

**LAND MODEL**

- \( SM^{-}(t) \) 9 km
- \( TB(t) \) 9 km
- Aggregate to "36 km"
- Clim. Mean Adjustment

**SMAP OBSERVATIONS**

- \( SM^{+}(t) = SM^{-}(t) + \Delta SM \) 9 km
- \( \Delta TB(t) \) "36 km"
- Clim. Mean Adjustment

**EnKF**

\[
\Delta SM = G \times \Delta TB
\]

\( G = \text{optimal gain matrix} \)

\[
G = \text{Cov}(SM_{\text{ens}}, TB_{\text{ens}}) \left[ \text{Cov}(TB_{\text{ens}}) + \text{Cov}(TB_{\text{obs}}) \right]^{-1}
\]

# Changes in L4_SM Modeling System

<table>
<thead>
<tr>
<th>Parameter</th>
<th>NRv4/NRv4.1 (Version 3)</th>
<th>NRv7.2 (Version 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+ latest boundary conditions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ more numerically stable surface energy calculations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ decoupling of surface and root zone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ other model changes</td>
</tr>
<tr>
<td>Soils</td>
<td>HWSD/STATSGO2</td>
<td></td>
</tr>
<tr>
<td>LAI</td>
<td>MODIS/GEOLAND</td>
<td></td>
</tr>
<tr>
<td>Albedo</td>
<td>MODIS 8-day</td>
<td></td>
</tr>
<tr>
<td>Topography</td>
<td>Hydro1K</td>
<td>SRTM</td>
</tr>
<tr>
<td>Land cover</td>
<td>GLCCv2</td>
<td>GLOBCOVER</td>
</tr>
<tr>
<td>Canopy height</td>
<td>Look-up table</td>
<td>Lidar (Simard et al. 2011)</td>
</tr>
<tr>
<td>Precip. corrections</td>
<td>v2 (rescaled CPCU only)</td>
<td>v3 (+rescaling of GEOS-5 background precip. to GPCPv2.2)</td>
</tr>
<tr>
<td>Retrospective forcing</td>
<td>RP/FP-IT</td>
<td>MERRA-2</td>
</tr>
<tr>
<td>CSOIL_2</td>
<td>200</td>
<td>70,000</td>
</tr>
<tr>
<td>Catchment.F90</td>
<td>Dampen oscillations (except ITYP = 1)</td>
<td>Dampen oscillations for all types</td>
</tr>
<tr>
<td>In RZDRAIN</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>WEMIN</td>
<td>26 kg/m2</td>
<td>13 kg/m2</td>
</tr>
<tr>
<td>z0_formulation</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

**In RZDRAIN**

\[
\text{IF}(\text{SRFLW}<0.) \quad \text{SRFLW} = 0.04 \times \text{SRFLW}
\]

**WEMIN**

*Carry-over from NRv5, had been missing in this Table*
# Updated L4_SM Analysis

<table>
<thead>
<tr>
<th></th>
<th>Vv3030</th>
<th>Tv4000</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model version</strong></td>
<td>NRv4.1</td>
<td>NRv7.2</td>
</tr>
<tr>
<td><strong>RTM parameters</strong></td>
<td>v1 (calibrated to NRv4)</td>
<td>v2 (calibrated to NRv5)</td>
</tr>
<tr>
<td><strong>RTM soil temperature:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>broadleaf evergreen</td>
<td>0.5*(TSURF + TSOIL1)</td>
<td>TSOIL1</td>
</tr>
<tr>
<td>all other landcover</td>
<td>TSOIL1</td>
<td>TSOIL1</td>
</tr>
<tr>
<td><strong>L1C Tb observations</strong></td>
<td>Version 3</td>
<td>T15160 (Version 4)</td>
</tr>
<tr>
<td><strong>Tb scaling parameters</strong></td>
<td>SMOS (6 years)</td>
<td>SMOS (7 years)</td>
</tr>
<tr>
<td></td>
<td>SMAP (2 years)</td>
<td>SMAP (2.4 years)</td>
</tr>
<tr>
<td><strong>EnkF state vector</strong></td>
<td>catdef, rzexc, srfexc, tc1/2/4</td>
<td>rzexc, srfexc, tc1/2/4</td>
</tr>
</tbody>
</table>
Changes in L4_SM Analysis

- Assimilated SMAP Tbs generally warmer by a few K than in Version 3 owing to new L1 calibration.
- L-band RTM uses TSOIL1 for all land cover classes. (Version 3 system used ½ (TSURF+TSOIL1) for broadleaf evergreen.)
- RTM parameters calibrated using NRv5 modeling system. (New calibration using NRv7.2 is in progress and will be tested for future version upgrade.)
- Tb scaling parameters based on longer record:

<table>
<thead>
<tr>
<th>L4_SM Stream</th>
<th>SMOSv6 Period</th>
<th>SMAP Period and Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vv3030</td>
<td>Jul 2010 – Jun 2016 (6 yrs)</td>
<td>Apr 2015 – Mar 2017 (2.0 yrs; R14 ops)</td>
</tr>
<tr>
<td>Tv4000</td>
<td>Jul 2010 – Jun 2017 (7 yrs)</td>
<td>Apr 2015 – Aug 2017 (2.3 yrs; T15160 OASIS)</td>
</tr>
<tr>
<td>Vv4010</td>
<td>Jul 2010 – Jun 2017 (7 yrs)</td>
<td>Apr 2015 – Feb 2018 (2.9 yrs; T15160/570 OASIS)</td>
</tr>
</tbody>
</table>

- Removed “catdef” model prognostic variable from EnKF state vector.
- Bug fix: Fore and aft Tbs from same location and half-orbit now always used at same analysis time.
- Removed check for excessive std-dev in fore-minus-aft Tbs. (Obsolete with improved L1 processing.)
"Natural" water balance does NOT close: $\Delta TWS - P + E + R \sim 0.17 \text{ mm/d}$

- $P = 2.57 \text{ mm/d}$
- $E/P = 0.72$
- $R/P = 0.34$
**Water Balance (Tv4000)**

"Natural" water balance does NOT close: \( \Delta TWS - P + E + R ~ 0.16 \text{ mm/d} \)

- \( P = 2.44 \text{ mm/d} \)
- \( E/P = 0.71 \)
- \( R/P = 0.34 \)
Core Validation Sites

<table>
<thead>
<tr>
<th>Horizontal scale</th>
<th>Surface soil moisture</th>
<th>Root zone soil moisture</th>
<th>Surface Soil Temperature (6am)</th>
<th>Surface Soil Temperature (6pm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>36 km</td>
<td>17</td>
<td>36 km</td>
<td>36 km</td>
<td>36 km</td>
</tr>
<tr>
<td>9 km</td>
<td>17</td>
<td>9 km</td>
<td>9 km</td>
<td>9 km</td>
</tr>
</tbody>
</table>

Number of different core sites: 17
Number of reference pixels: 17
O-F auto-correlations increased in Tv4000. Assimilation is slightly less “efficient”.
O-F Auto-Correlations

global average number of O-F data pairs and fraction of global land area contributing to sample auto-correlation (per grid cell, Vv3030, all/all)

number of O-F data pairs
fraction of global land area (times 1000)

lag [days]

0 1 2 3 4 5 6 7 8 9 10

spatial average O-F auto-correlation (Vv3030, all/all)

auto-correlation [-]

0 0.5 1

lag [days]

0 1 2 3 4 5 6 7 8 9 10

spatial average O-F auto-correlation (Tv4000, all/all)

auto-correlation [-]

0 0.5 1

lag [days]

0 1 2 3 4 5 6 7 8 9 10
Slightly more runoff in Tv4000 in mid-latitudes.

Changes in Africa and at very high latitudes owing to rescaling of GEOS-5 precip to GPCPv2.2 climatology.
Climatology – Runoff

- Tv4000 root-zone slightly wetter in most regions.
- Changes in Africa and at very high latitudes owing to rescaling of GEOS-5 precip to GPCPv2.2 climatology.