SMAP CalVal Workshop

Oxnard, 3-5 May 2011

Summary of L4 Product Cal/Val Requirements

R. Reichle*  NASA/GSFC
J. Kimball** Univ. of MT

* Email: Rolf.Reichle@nasa.gov  Phone: 301-614-5693
** Email: johnk@flbs.umt.edu  Phone: 406-982-3301 x230
**SMAP Level 4 soil moisture and carbon products**

**L4_SM Product:**
Assimilating SMAP data into a land model driven with observation-based forcings yields:
– a root zone moisture product (reflecting SMAP data), and
– a complete and consistent estimate of soil moisture & related fields.

**L4_C Product:**
Combining L4_SM (SM & T), high-res L3_F/T_A & ancillary Gross Primary Productivity (GPP) inputs within a C-model framework yields:
- a Net Ecosystem Exchange (NEE) product, &
- estimates of surface soil organic carbon (SOC), component C fluxes (R) & underlying SM & T controls.
Main objectives:
• Provide estimates of **root zone** soil moisture (top 1 m) based on SMAP obs.
• Provide **global, 3-hourly, 9 km** surface and root zone soil moisture.

Baseline algorithm:
• Customized version of existing NASA/GEOS-5 Land Data Assimilation System
  – 3d ensemble Kalman filter
  – Catchment land surface model
**L4_SM inputs and outputs**

### SMAP inputs
- **Brightness temperature** (L1C_TB, 36 km)
- **Radar soil moisture** (L2_SM_A, 3 km)
- **Freeze-thaw state** (L3_F/T_A, 3 km)

### Ancillary data inputs
- Land model parameters
- Surface meteorology (incl. observation-corrected precip)
- Land assimilation parameters

### L4_SM algorithm

### L4_SM product
- 9 km, 3-hourly global output with 7-day latency
  - Surface soil moisture (≡ top 5 cm)
  - Root zone soil moisture (≡ top 1 m)
  - **Research output**
    - surface and soil temperatures (input to L4_C)
    - sensible, latent, and ground heat flux
    - runoff, baseflow, snowmelt
    - surface meteorological forcings (air temperature, precipitation, ...)
    - **error estimates** (generated by assimilation system)

In units of $m^3m^{-3}$ and percentiles
## L4_SM cal/val

### Pre-launch

- Use L4_SM system with SMOS obs (also AMSR-E, Aquarius, ...)
- Apply cal/val to the extent possible.
- Conduct OSSE’s (calibration of assimilation parameters).
Calibration within 1\textsuperscript{st} year:

Bias correction param’s ("cdf matching"), assimilation param’s (thru innovations).

Validation with in situ observations:

Surface soil moisture:
- Apply L2_SM_A/P cal/val procedures.

Root-zone soil moisture:
- In principle, cal/val is identical to surface soil moisture, but
- have fewer in situ obs. (e.g. from USDA/SCAN, NCDC/CRN)
- rarely/never have multiple in situ obs. within single grid cell

\textbf{Requirement}: Need as many root-zone soil moisture obs. as possible.
**L4_SM cal/val**

<table>
<thead>
<tr>
<th>Post-launch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Additional evaluation:</strong></td>
</tr>
<tr>
<td>– Evaluate with high-quality, independent precipitation obs (Crow 2007).</td>
</tr>
<tr>
<td>– Evaluate research product components (e.g. fluxes) to the extent possible.</td>
</tr>
</tbody>
</table>
## SMAP L4_SM validation approach

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Data</th>
<th>Importance</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Sites</td>
<td>Observed grid cell average values (time-continuous)</td>
<td>Primary</td>
<td>Anomaly correlation, RMSE, bias</td>
</tr>
<tr>
<td>Sparse Networks</td>
<td>Observed values (time-continuous)</td>
<td>Primary</td>
<td>Anomaly correlation</td>
</tr>
<tr>
<td>Satellite Products</td>
<td>Orbit-based match-ups (SMOS, ASCAT, ...)</td>
<td>Secondary: Pending continued operation</td>
<td>Anomaly correlation, RMSE, bias</td>
</tr>
<tr>
<td>Model Products</td>
<td>Global modeling and assimilation systems (ECMWF, NCEP, ...)</td>
<td>Primary</td>
<td>Anom. correlation, assim. diagnostics, RMSE, bias</td>
</tr>
<tr>
<td>Field Experiments</td>
<td>Detailed estimates for a very limited set of conditions</td>
<td>Secondary</td>
<td>Anomaly correlation, RMSE, bias</td>
</tr>
</tbody>
</table>
Synergies with L4_SM development:
Cal/val based on land modeling and assimilation

• Algorithm Testbed
  • GMAO Nature Run

• Supplemental and complementary validation approaches:
  • Include:
    • Triple collocation (L2_SM)
    • Consistency of assim. increments with independent precipitation obs. (L2_SM, L4_SM)
    • Consistency of assim. system diagnostics (e.g., statistics of “obs.-minus-forecast” residuals) (L4_SM)
  • Enable scaling from point-scale obs. to satellite-scale estimate
  • Are independent of scheduling risk associated with field campaigns
Product: Net Ecosystem CO₂ exchange (NEE = GPP – R_{eco})

- **Motivation/Objectives:** Quantify net C flux in boreal landscapes; reduce uncertainty regarding missing C sink on land (NRC Decadal Survey);

- **Approach:** Apply a soil decomposition model driven by SMAP L4_SM & ancillary (LC, GPP) inputs to compute NEE;

- **Inputs:** Daily surface (<10cm) SM & T (L4_SM), LC & GPP (MODIS, VIIRS);

- **Outputs:** NEE (primary/validated); R_{eco} & SOC (research);

- **Domain:** Vegetated areas encompassing boreal/arctic latitudes (≥45 N);

- **Resolution:** 9x9 km;

- **Temporal fidelity:** Daily (g C m⁻² d⁻¹);

- **Latency:** 14-day;

- **Accuracy:** Commensurate with tower based CO₂ Obs. (RMSE ≤ 30 g C m⁻² yr⁻¹ and 1.6 g C m⁻² d⁻¹).
Several L4_C options are being evaluated based on recommendations from an earlier ATBD peer-review; options designed to enhance product accuracy & utility include:

- Global domain encompassing all vegetated land areas;
- Internal GPP calculations using SMAP L4_SM, L3_FT & ancillary land cover (LC) & VI (e.g. NDVI from MODIS, VIIRS) inputs;
- Represent finer scale (<9km) spatial heterogeneity consistent with available LC inputs;
- Explicit representation of LC disturbance (fire) and recovery impacts;
- Algorithm calibration using available observation data (FLUXNET, soil inventories).
**L4_C cal/val**

**Pre-launch:**
- L4_C development & testing using available inputs: LC, NDVI, GPP (MODIS); SM & T (MERRA); FT (SSM/I, AMSR-E, SMOS);
- Calibration/optimization and initialization of L4_C algorithm parameters (e.g. BPLUT, SOC pools, disturbance history);

**Post-launch:**
- Re-calibration & re-initialization of L4_C parameters using SMAP L4_SM inputs;
- Verify SMAP L4_C NEE accuracy:
  - Tower site comparisons using CO₂ Obs; stand-level C-model simulations & sensitivity studies;
  - Comparisons with available soil inventories;
  - Field campaigns using nested in situ, airborne (CARVE, AirMOSS) & satellite data;
  - Atm. model inversions of L4_C outputs & comparisons of resulting C source/sink activity against available observations (CO₂ flask network, GOSAT, OCO-2).
### SMAP L4_C validation approach

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Data</th>
<th>Importance</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Sites</td>
<td>Observed grid cell average values</td>
<td>Primary</td>
<td>Correlations, RMSE, bias</td>
</tr>
<tr>
<td></td>
<td>(time-continuous)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sparse Networks</td>
<td>Observed values</td>
<td>Primary</td>
<td>Correlations, RMSE, bias</td>
</tr>
<tr>
<td></td>
<td>(time-continuous)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satellite Products</td>
<td>Orbit-based match-ups (SMOS, PALSAR, ...)</td>
<td>Secondary</td>
<td>Anomaly correlation, RMSE, bias</td>
</tr>
<tr>
<td>Model Products</td>
<td>Site &amp; Global modeling systems, model</td>
<td>Primary</td>
<td>Sensitivity diagnostics,</td>
</tr>
<tr>
<td></td>
<td>inversions (Carbontracker)</td>
<td></td>
<td>correlation, RMSE, bias</td>
</tr>
<tr>
<td>Field Experiments</td>
<td>Detailed estimates for a very limited</td>
<td>Secondary</td>
<td>correlations, RMSE, bias</td>
</tr>
<tr>
<td></td>
<td>set of conditions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Optimal L4_C validation site design

- Characterize major biomes within northern land areas (baseline)
  - Boreal ENLF, tundra, grassland, mixed forest (DBLF, ENLF), & DNLF types;
  - Disturbance history & stand succession impacts;

- Representative conditions within regional (~10x10 km) window
  - Select sites with relatively homogeneous land cover & terrain conditions;
  - Continuous measurements to characterize daily variability & cumulative annual C fluxes;

- Documented uncertainty (systematic & random error) in C flux measurements
  - Established and well defined protocols for correction & gap filling to establish complete annual C flux time series;
  - Multi-year time series to establish average conditions & year-to-year variability;

- Coincident measurements of surface meteorology, H₂O & CO₂ fluxes
  - Enable analysis of water, energy & carbon cycle linkages;
  - Measurements of component C fluxes (GPP, R_{eco}, NEE) & environmental controls (SM and soil T, surface SOC).
**L4_C Cal/Val using Tower Site Data**

**Woody Savannah (Tonzi Ranch, CA)**

Tower CO₂ flux data (FLUXNET) is used for L4_C calibration & validation (e.g. left). Baseline model performance is evaluated for expected accuracy (NEE RMSE < 30 g C m⁻² yr⁻¹ or 1.6 g C m⁻² d⁻¹). A Markov Chain Monte Carlo (MCMC) optimization is applied to minimize an objective function by adjusting biome-specific model parameters to representative tower data, including calibrating soil moisture response curves (center) for better accuracy. Uncalibrated model runs (right) using alternative remote sensing & tower inputs are also used to clarify error propagation & uncertainty sources. Available FLUXNET data includes >400 site year measurements & represent most global biome types.

1D. Baldocchi is PI of Tonzi and Vaira FLUXNET tower sites; 2R. Scott is PI of Santa Rita Site
Post-launch: L4_C model assimilation to quantify net CO₂ source-sink activity

- Apply L4_C products within carbon data assimilation system for tracking net CO₂ source/sink activity;
- Atmospheric perspective based on atmospheric transport model (TM3) constrained by satellite remote sensing and sparse surface observations;
- Accounts for fossil-fuel and fire related CO₂ emissions;
- L4_C based NEE provides land surface initial conditions;
- Provides for rigorous validation using synergistic C observations (CO₂ flask network, GOSAT, OCO-2);
- Provides means to quantify C source/sink activity (SMAP Decadal Survey objective);

1http://www.esrl.noaa.gov/gmd/ccgg/carbontracker2