# Validation practices for satellite soil moisture products – What are (the) errors?

**A. Gruber**, C. Albergel, S. Alemohammad, A. Al-Yaari, L. Brocca, J.-C. Calvet, A. Colliander, M. Cosh, W. Crow, W. Dorigo, C. Draper, M. Hirschi, T. Jackson, Y. Kerr, A. Konings, W. Lahoz, G. De Lannoy, K. McColl, C. Montzka, J. Muñoz-Sabater, N. Nicolai-Shaw, J. Peng, R. Reichle, P. Richaume, C. Ruediger, T. Scanlon, R. van der Schalie, W. Wagner

9th SMAP Cal/Val Workshop

George Mason University

October 22-23, 2018

# Outline

- Towards standardized validation practices
- A little bit of theory
- Validation Good Practice Protocol

# Towards standardized practices

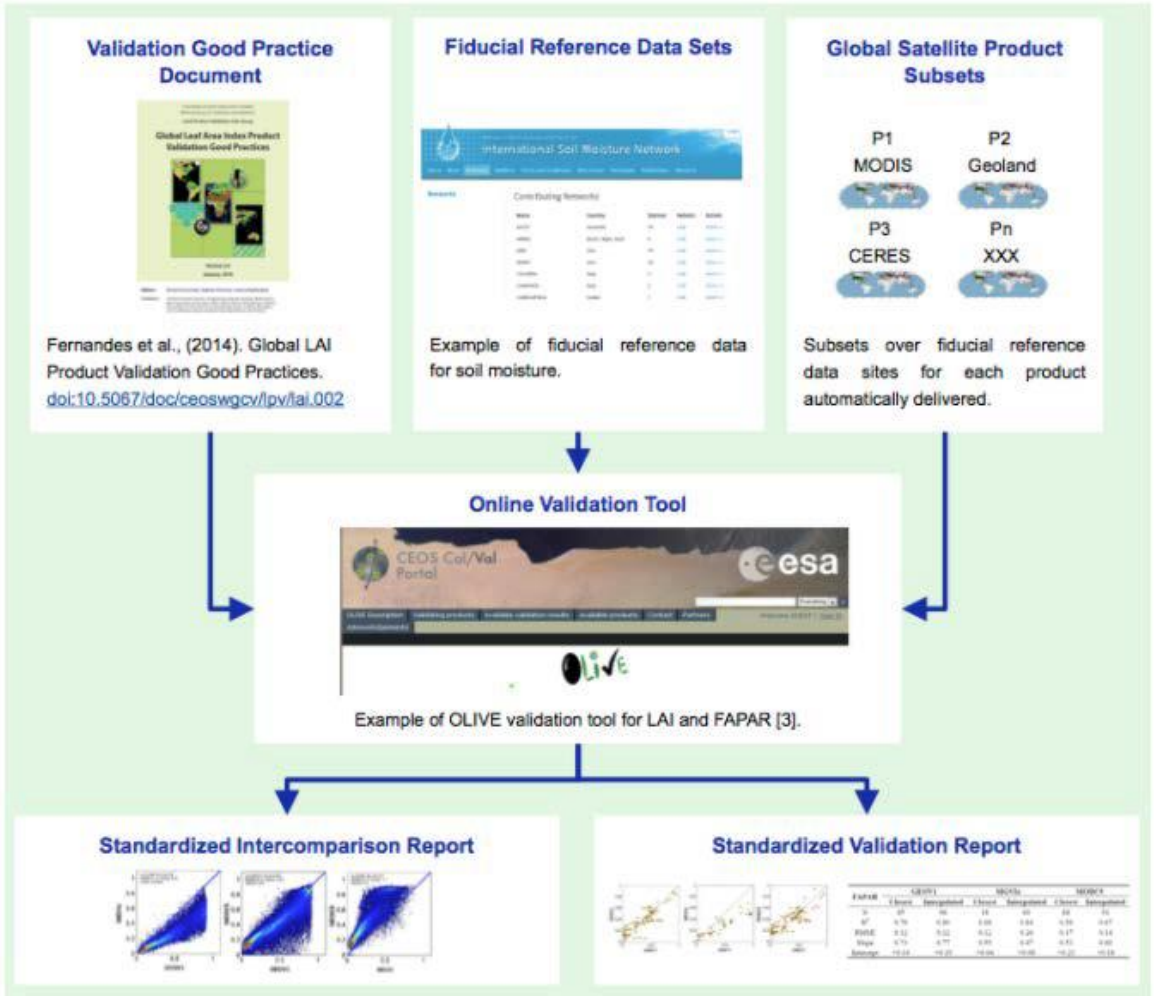
- **CEOS** – The Committee of Earth Observation Satellites
  - **WGCV** – Working Group on Calibration and Validation
    - **LPV** – Land Product Validation Subgroup
      - (Focus Area Soil Moisture)

- **QA4EO** – Quality Assurance Framework for Earth Observation
  - A suite of top-level guidelines

- **QA4ECV** – Quality Assurance for Essential Climate Variables
  - Demonstration of a QA framework on the example of some ECVs

- **QA4SM** – Quality Assurance for Soil Moisture
  - Automated Validation Framework for Soil Moisture

# Towards standardized practices




https://lpvs.gsfc.nasa.gov

# Towards standardized practices

## Validation Good Practice

Currently, most recent satellite product validation plan documents serve as best references for validation of soil moisture:

- SMAP - The SMAP Handbook includes a description of the Cal/Val plan.
- ASCAT - Brocca, L., F. Melone, T. Moramarco, W. Wagner, C. Albergel (2014) Scaling and filtering approaches for the use of satellite observations, Chapter 17 in *Remote Sensing of Engergy Fluxes and Soil Moisture Content*, G.P. Petropoulus (Ed), CRC Press, Boca Raton London New York, 411-425.
- ESA CCI Product Validation Plan
- GCOM W (AMSR2) Validation Protocol
- SMOS - Cal/Val Plan
  Y.H. Kerr, et al., Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation, *Remote Sens. of Environment*, Volume 180, July 2016, Pages 40–63.

SMAP launch, January 31, 2015

## Soil Moisture Validation Reference Data Sets

Best currently available in situ reference data for satellite-derived soil moisture product validation are provided by in situ operational networks, most of which can be found at the International Soil Moisture Network web site.

https://lpvs.gsfc.nasa.gov

# A little bit of theory

*) bias … systematic error
(additive / multiplicative)

uncertainty … random error variance

- Errors

$$e_x = x - t$$

- The standard error model for soil moisture

$$x = \alpha_x + \beta_x t + \varepsilon_x$$

- Assumptions
  - Gaussianity
  - Stationarity
  - Homoscedasticity
  - Error independence

$$\overline{x} = \alpha_x + \beta_x \overline{t}$$

$$\sigma_x^2 = \beta_x^2 \sigma_t^2 + \sigma_{\xi_x}^2$$

$$\sigma_{xy} = \beta_x \beta_y \sigma_t^2 + \sigma_{\xi_x, \xi_y}$$

All validation metrics are based on statistical moments!

# A little bit of theory

$$\overline{x} = \alpha_x + \beta_x \overline{t}$$

$$\sigma_x^2 = \beta_x^2 \sigma_t^2 + \sigma_{\xi_x}^2$$

$$\sigma_{xy} = \beta_x \beta_y \sigma_t^2 + \sigma_{\xi_x,\xi_y}$$

- How to properly interpret the metrics?

$$b_{xy} = \overline{x} - \overline{y} = \alpha_x - \alpha_y + (\beta_x - \beta_y)\overline{t}$$

$$RMSD_{xy} = \sqrt{\overline{(x-y)^2}} = \sqrt{(\overline{x} - \overline{y})^2 + \sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}}$$

$$= \sqrt{(\alpha_x - \alpha_y + (\beta_x - \beta_y)\overline{t})^2 + (\beta_x - \beta_y)^2\sigma_t^2 + \sigma_{\xi_x}^2 + \sigma_{\xi_y}^2 - 2\sigma_{\xi_x,\xi_y}}$$

$$ubRMSD_{xy} = \sqrt{RMSD^2 - b_{xy}^2} = \sqrt{\sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}}$$

$$= \sqrt{(\beta_x - \beta_y)^2\sigma_t^2 + \sigma_{\xi_x}^2 + \sigma_{\xi_y}^2 - 2\sigma_{\xi_x,\xi_y}}$$

$$R_{xy} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} = \frac{\beta_x \beta_y \sigma_t^2 + \sigma_{\xi_x,\xi_y}}{\sqrt{(\beta_x^2\sigma_t^2 + \sigma_{\xi_x}^2)(\beta_y^2\sigma_t^2 + \sigma_{\xi_y}^2)}}$$

$$\approx \mathrm{sgn}(\sigma_{xy}) \frac{1}{\sqrt{(1 + SNR_x^{-1})(1 + SNR_y^{-1})}}$$

# A little bit of theory

$$\overline{x} = \alpha_x + \beta_x \overline{t}$$

$$\sigma_x^2 = \beta_x^2 \sigma_t^2 + \sigma_{\xi_x}^2$$

$$\sigma_{xy} = \beta_x \beta_y \sigma_t^2 + \sigma_{\xi_x,\xi_y}$$

- How to properly interpret the metrics?

$$ubRMSE_x = \sqrt{\left|\overline{(x-y)(x-z)}\right|} = \sqrt{\left|\sigma_x^2 - \frac{\sigma_{xy}\sigma_{xz}}{\sigma_{yz}}\right|}$$

$$= \sqrt{\left|\beta_x^2\sigma_t^2 + \sigma_{\xi_x}^2 - \frac{(\beta_x\beta_y\sigma_t^2 + \sigma_{\xi_x,\xi_y})(\beta_x\beta_z\sigma_t^2 + \sigma_{\xi_x,\xi_z})}{\beta_y\beta_z\sigma_t^2 + \sigma_{\xi_y,\xi_z}}\right|} \approx \sigma_{\xi_x}$$

$$R_x = \sqrt{\left|\frac{\sigma_{xy}\sigma_{xz}}{\sigma_x^2\sigma_{yz}}\right|} = \sqrt{\left|\frac{(\beta_x\beta_y\sigma_t^2 + \sigma_{\xi_x,\xi_y})(\beta_x\beta_z\sigma_t^2 + \sigma_{\xi_x,\xi_z})}{(\beta_x^2\sigma_t^2 + \sigma_{\xi_x}^2)(\beta_y\beta_z\sigma_t^2 + \sigma_{\xi_y,\xi_z})}\right|}$$

$$\approx \sqrt{\left|\frac{\beta_x^2\sigma_t^2}{\beta_x^2\sigma_t^2 + \sigma_{\varepsilon_x}^2}\right|} = \frac{1}{\sqrt{1 + SNR_x^{-1}}}$$

$$SNR_x[dB] = -10\log\left(\left|\frac{\sigma_x^2\sigma_{yz}}{\sigma_{xy}\sigma_{xz}}\right| - 1\right) \approx 10\log\left(\frac{\beta_x^2\sigma_t^2}{\sigma_{\varepsilon_x}^2}\right)$$

# A little bit of theory

- All validation metrics are subject to sampling errors!

- Any validation result should be accompanied by a confidence interval!

| | | | | | | |
|---|---|---|---|---|---|---|
| ubRMSD = 0.04 | → | $CI_{95}$ = [0.035, 0.046] | for n = 100 | or | [0.038 – 0.043] | for n = 500 |
| bias = 0.1 | → | $CI_{95}$ = [0.092, 0.108] | for n = 100 | or | [0.096 – 0.104] | for n = 500 |
| R = 0.6 | → | $CI_{95}$ = [0.46, 0.71] | for n = 100 | or | [0.54 – 0.65] | for n = 500 |

# A little bit of theory

- Analytical calculation of confidence intervals (u = x – y)
  - bootstrapping as non-parametric alternative for TCA CI

$$\frac{\overline{u} - \mu_u}{\frac{s_u}{\sqrt{n}}} \sim t_{n-1}$$

$$CI_{b_{xy}} = \left[ b_{xy} + t_{n-1}^{\alpha/2} \frac{ubRMSD_{xy}}{\sqrt{n}} \, , \; b_{xy} + t_{n-1}^{1-\alpha/2} \frac{ubRMSD_{xy}}{\sqrt{n}} \right]$$

$$\frac{\sqrt{n-1}\, s_u}{\sigma_u} \sim \chi_{n-1}$$

$$CI_{ubRMSD_{xy}} = \left[ ubRMSD_{xy} \frac{\sqrt{n-1}}{\chi_{n-1}^{1-\alpha/2}} \, , \; ubRMSD_{xy} \frac{\sqrt{n-1}}{\chi_{n-1}^{\alpha/2}} \right]$$

$$z = 0.5 \ln \left( \frac{1 + R_{xy}}{1 - R_{xy}} \right) \sim \mathcal{N}_{z,(n-3)^{-0.5}}$$

$$CI_{R_{xy}} = \left[ \frac{e^{2z^{1-\alpha}} - 1}{e^{2z^{1-\alpha}} + 1} \, , \; \frac{e^{2z^{\alpha}} - 1}{e^{2z^{\alpha}} + 1} \right]$$

# A little bit of theory

- Averaging of CIs / Cis of averaged metrics?
  - Sampling errors are “uncertainties”
    - Estimation through the propagation of uncertainties
    - ”Clustering” of areas with expected sampling error correlation between samples

$$\overline{\nu} = \mathbf{w}^\intercal \mathbf{v}$$

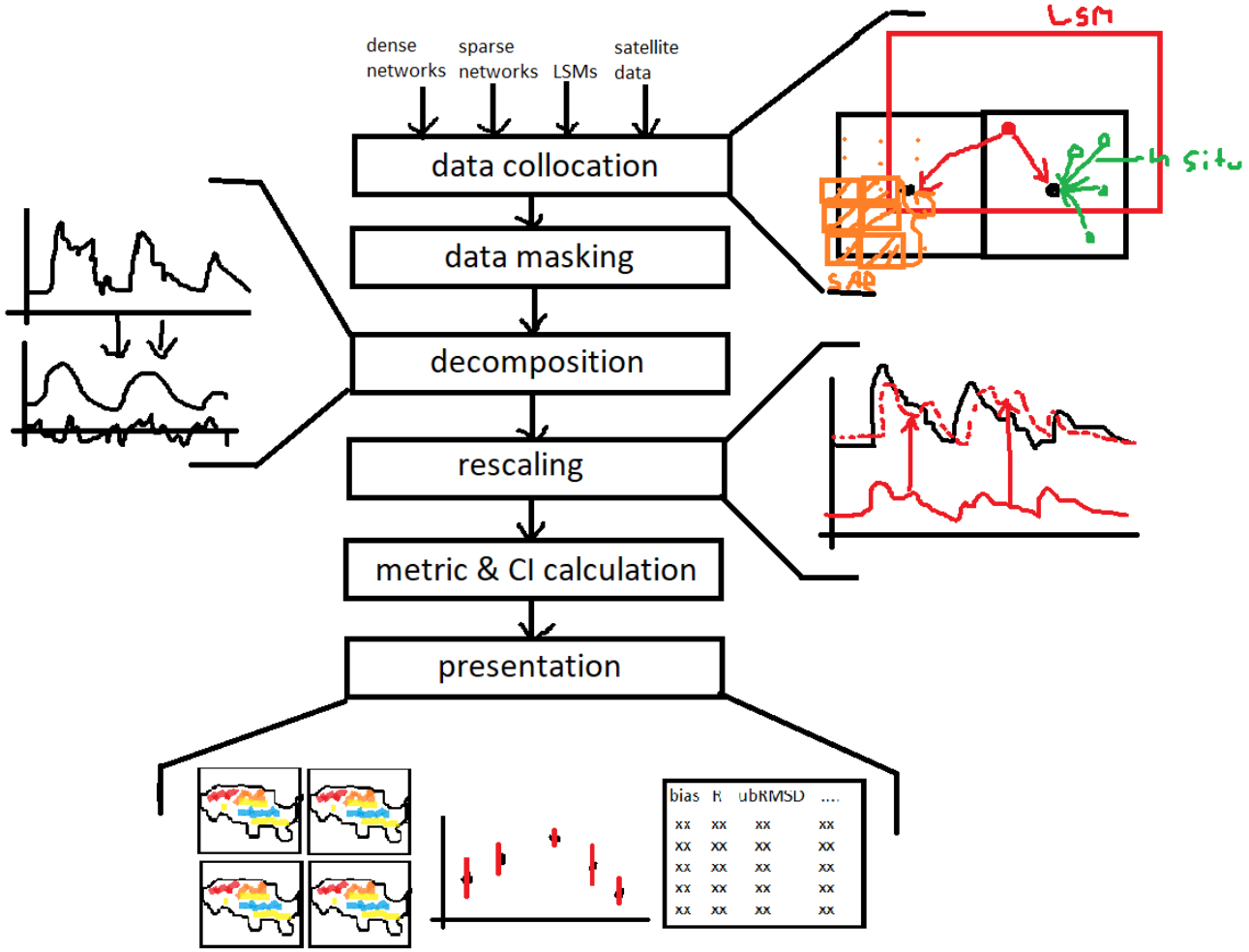
$$s^2_{\overline{\nu}} = \mathbf{w}^\intercal \mathbf{\Sigma} \mathbf{w}$$

$$\mathbf{\Sigma} = \mathbf{s}\mathbf{s}^\intercal \qquad \mathbf{\Sigma} = diag(\mathbf{s}\mathbf{s}^\intercal)$$

$$\overline{\nu} = \frac{1}{c}\sum_{i=1}^{c} \overline{\nu}_i \qquad \text{with} \qquad \overline{\nu}_i = \frac{1}{m}\sum_{j=1}^{m} \nu_j$$

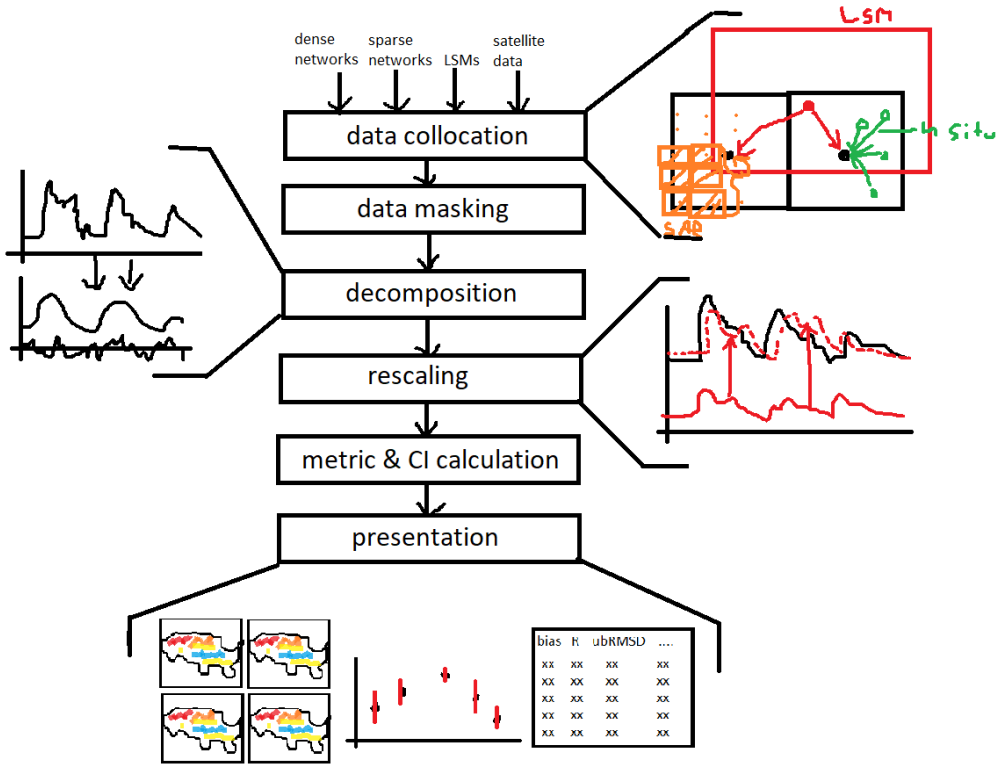
$$s^2_{\overline{\nu}} = \frac{1}{c^2}\sum_{i=1}^{c} s^2_{\overline{\nu}_i} \qquad \text{with} \qquad s^2_{\overline{\nu}_i} = \left(\frac{1}{m}\sum_{j=1}^{m} s_{\nu_j}\right)^2$$

# Validation Good Practice Protocol



# Validation Good Practice Protocol

- Collocation
  - NN resampling, 1-12 hr max distance
- Masking
  - Provider flags, SWE=0, T<4 deg, tropical forests, water bodies
- Decomposition
  - Short-term + long-term anomalies (4-8 week moving average windows)
- Rescaling
  - Mean/Std. or TCA rescaling
- Metrics:
  - Bias, ubRMSD, ubRMSE, R, SNR
- Presentation:
  - Spatial maps, average statistics, Confidence Intervals !!



# Validation Good Practice Protocol

- This protocol should provide a comprehensive, general picture of soil moisture data quality
- BUT: Current accuracy requirements (ubRMSD = 0.04 $m^3/m^3$, SNR = 0,3,6 dB) do not relate to fitness-for-purpose
- Definition of accuracy requirements (including appropriate metrics) is task of the user community
- Does it make sense to provide concrete numbers, and if so, which?
  (e.g., max. NN distance, moving average window size, …)