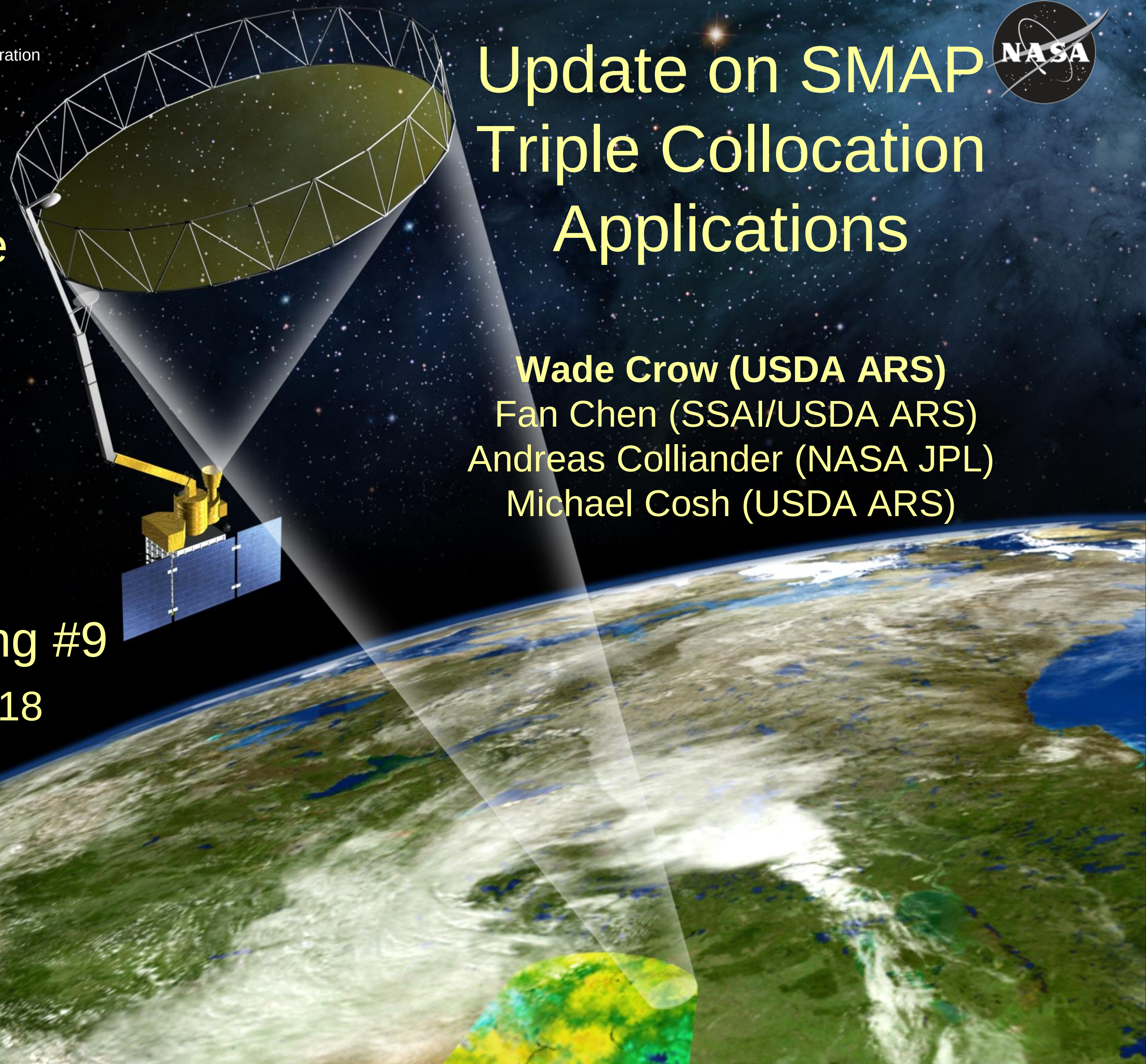
National Aeronautics and Space Administration

Update on SMAP Triple Collocation Applications

Soil Moisture Active Passive Mission

SMAP



Wade Crow (USDA ARS)
Fan Chen (SSAI/USDA ARS)
Andreas Colliander (NASA JPL)
Michael Cosh (USDA ARS)

Cal/Val Meeting #9

October, 27th 2018

Fairfax, VA

TC for SMAP Sparse Network Evaluations

- Sparse network metrics are relatively worse than those obtained at core sites (ubRMRE increased by ~0.012 m^3m^{-3} and R decreased by ~0.15 [-]).

- Original SMAP Triple Collocation (TC) approach [Miralles et al. 2013]:

  - Soil moisture triplet will consisting of [sparse, LSM, SMAP].
  - TC would be used as a effective upscaling tool to compensate sparse network validation metrics for their spatial sampling deficiencies.
  - Use core site data as a verification tool for TC assessment.

- Lessons learned during SMAP cal/val [Chen et al. 2016; 2018]:

  - TC is robust but only after seasonality has been removed to ensure a stationary time series.
  - ubRMSE correction is not possible due to lack of adequate scaling reference (i.e., one product lacking both multiplicative and additive bias).
  - “Uncorrected” comparisons to sparse network observations proved to me unexpectedly robust (in a relative sense).

Recent TC Developments/Future Directions

To date we have applied very simple statistical models to describe random soil moisture retrieval errors (i.e., additive, stationarity, white and independent).

Reality is messier and we run the risk of mis-interpreting and/or under-utilizing SMAP soil moisture products if we fail to recognize this.

Utilization of TC variants/extensions to provide a more complete picture of soil moisture retrievals errors…

Recent TC Developments/Future Directions

To date we have applied very simple statistical models to describe random soil moisture retrieval errors (i.e., additive, stationarity, white and independent).

Reality is messier and we run the risk of mis-interpreting and/or under-utilizing SMAP soil moisture products if we fail to recognize this.

<u>Variants of TC can provide:</u>

1) Temporal error auto-covariance [Dong et al. 2017].

2) Spatial structure of error cross-correlation [Gruber et al. 2018].

3) Error covariance in SM products [Pierdicca et al. 2017; Gruber et al. 2016].

4) Error covariance with ancillary variables [Zwieback et al. 2018].

5) Non-parametric mutual information [Nearing et al. 2017].

Recent TC Developments/Future Directions

To date we have applied very simple statistical models to describe random soil moisture retrieval errors (i.e., additive, stationarity, white and independent).

Reality is messier and we run the risk of mis-interpreting and/or under-utilizing SMAP soil moisture products if we fail to recognize this.

<u>Variants of TC can provide:</u>

1) Temporal error auto-covariance [Dong et al. 2017].

2) Spatial structure of error cross-correlation [Gruber et al. 2018].

3) Error covariance in SM products [Pierdicca et al. 2017; Gruber et al. 2016].

4) Error covariance with ancillary variables [Zwieback et al. 2018].

5) Non-parametric mutual information [Nearing et al. 2017].

Extension of TC to Lagged Covariances

Regular Triple Collocation

$$\sigma_x^2 = \left\langle (\mathbf{x}^* - \mathbf{y}^*)^T (\mathbf{x}^* - \mathbf{z}^*) \right\rangle$$

$$\sigma_y^2 = \left\langle (\mathbf{y}^* - \mathbf{x}^*)^T (\mathbf{y}^* - \mathbf{z}^*) \right\rangle$$

$$\sigma_z^2 = \left\langle (\mathbf{z}^* - \mathbf{x}^*)^T (\mathbf{z}^* - \mathbf{y}^*) \right\rangle$$

Yields random error variance

Modified Triple Collocation

$$L_x = \left\langle (\mathbf{x}^* - \mathbf{y}^*)^T \left(\mathbf{x}_{\mathbf{L}}^* - \mathbf{z}_{\mathbf{L}}^*\right) \right\rangle$$

$$L_y = \left\langle (\mathbf{y}^* - \mathbf{x}^*)^T \left(\mathbf{y}_{\mathbf{L}}^* - \mathbf{z}_{\mathbf{L}}^*\right) \right\rangle$$

$$L_z = \left\langle (\mathbf{z}^* - \mathbf{x}^*)^T \left(\mathbf{z}_{\mathbf{L}}^* - \mathbf{y}_{\mathbf{L}}^*\right) \right\rangle$$

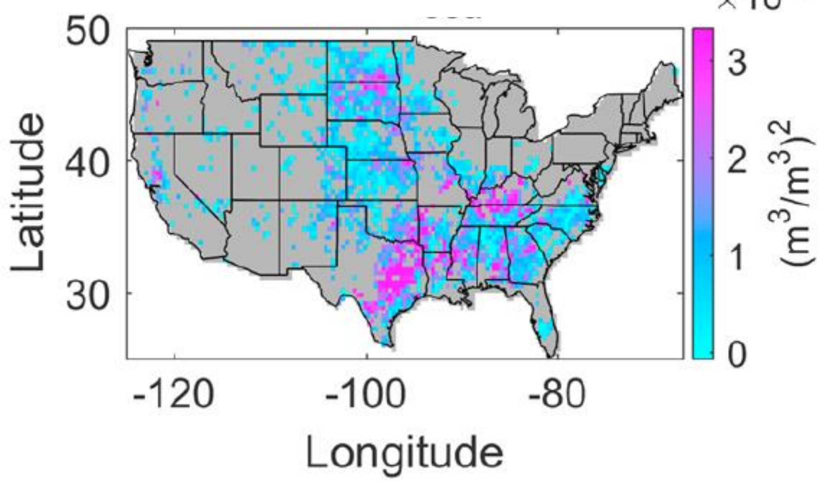
L = lag operator

Yields random error auto-covariance

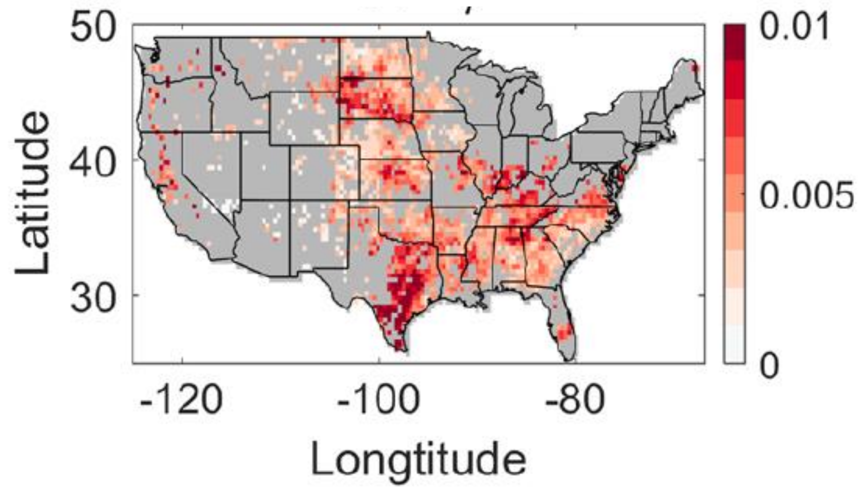
Introduced by Zwieback et al. [2013], implemented (in time) by Dong et al. [2017] and (in space) by Gruber et al. [2015; 2018]

Temporal Error Auto-Covariance

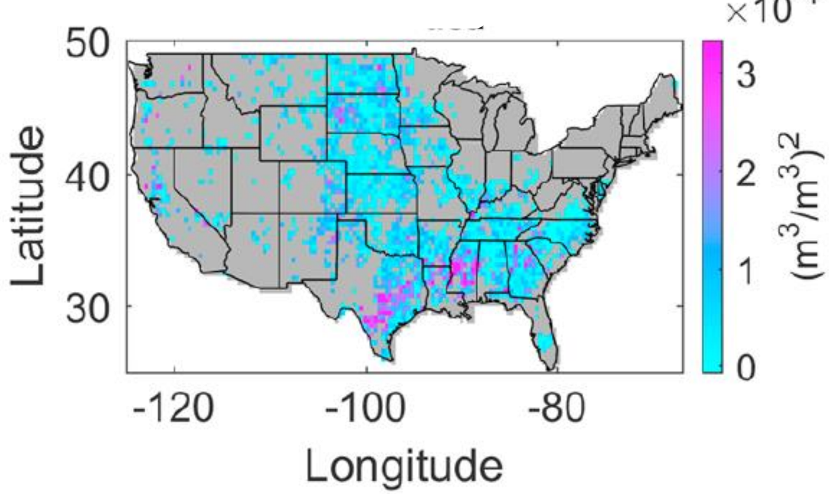
Error auto-cov. in SCA SM anomalies



Interannual NDVI variability



Error auto-cov. in DCA SM anomalies



Time scale of auto-covariance lag ~3-4 day average interval between successive SMAP AM retrievals.

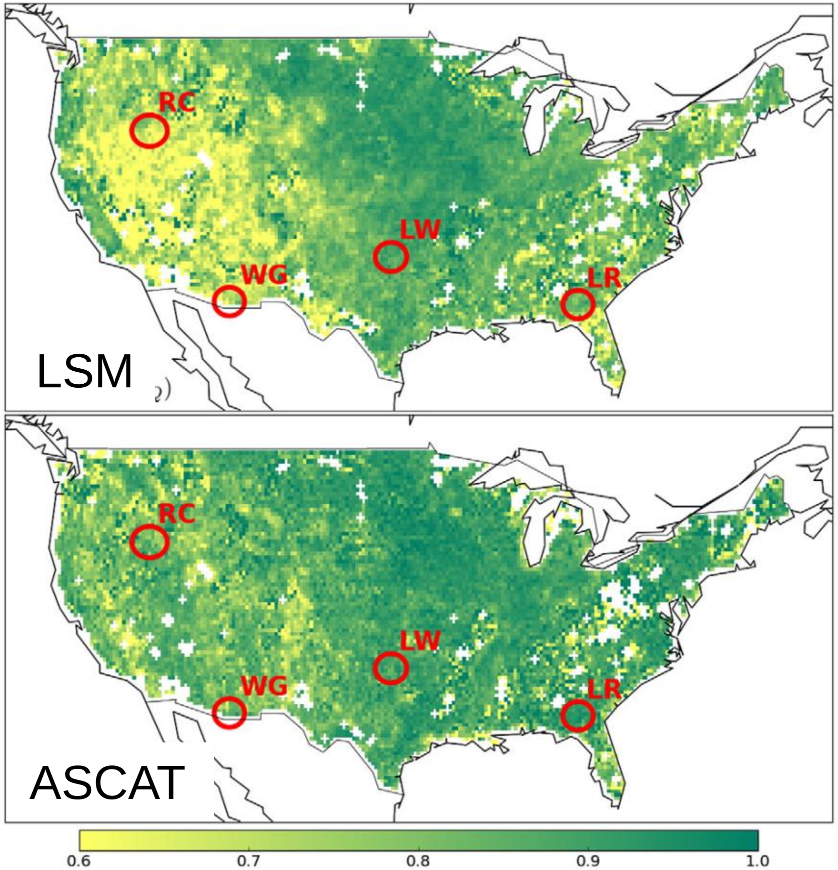
Dong et al. [2017]

Spatial Error Cross-Correlation

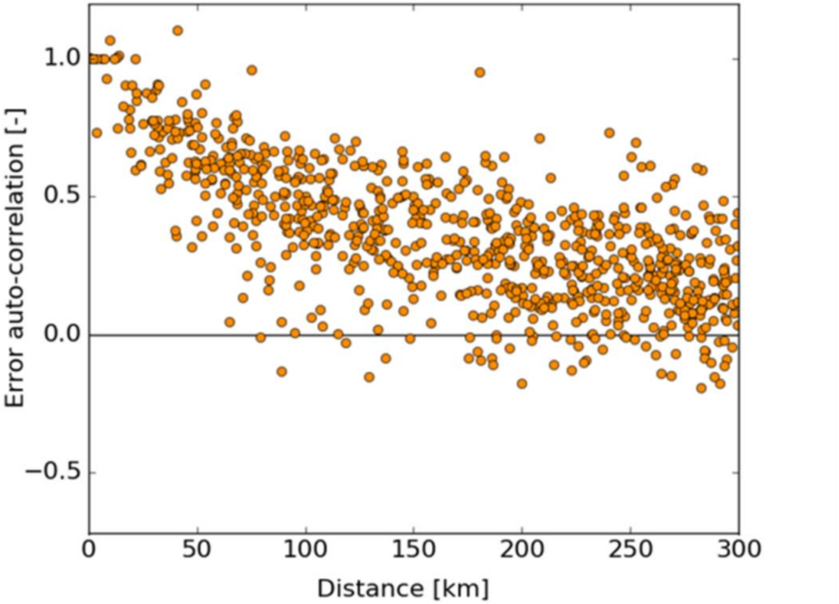
Temporal cross-correlation of error in two points separated in space.

separation = single 0.25°-pixel
(averaged in cardinal directions)

Gruber and Crow [2015]; Gruber et al. [2018]



Spatial Error Cross-Correlation Function



Recent TC Developments/Future Directions

To date we have applied very simple statistical models to describe random soil moisture retrieval errors (i.e., additive, stationarity, white and independent).

Reality is messier and we run the risk of mis-interpreting and/or under-utilizing SMAP soil moisture products if we fail to recognize this.

<u>Variants of TC can provide:</u>

1) Temporal error auto-covariance [Dong et al. 2017].

2) Spatial structure of error cross-correlation [Gruber et al. 2018].

3) Error covariance in SM products [Pierdicca et al. 2017; Gruber et al. 2016].

4) Error covariance with ancillary variables [Zwieback et al. 2018].

5) Non-parametric mutual information [Nearing et al. 2017].

SM Error Cross-Correlation

Triple collocation (TC) = **6** combination pairs to estimate **6** parameters (1 signal variance, 3 error variances, and 2 gains) [***well-posed*** if error cross-correlations are assumed zero].

Quadruple collocation (QC) = **10** combination pairs to estimate **8** parameters (1 signal variance, 4 error variances and 3 gains) [***over-constrained*** if error cross-correlations are assumed zero]

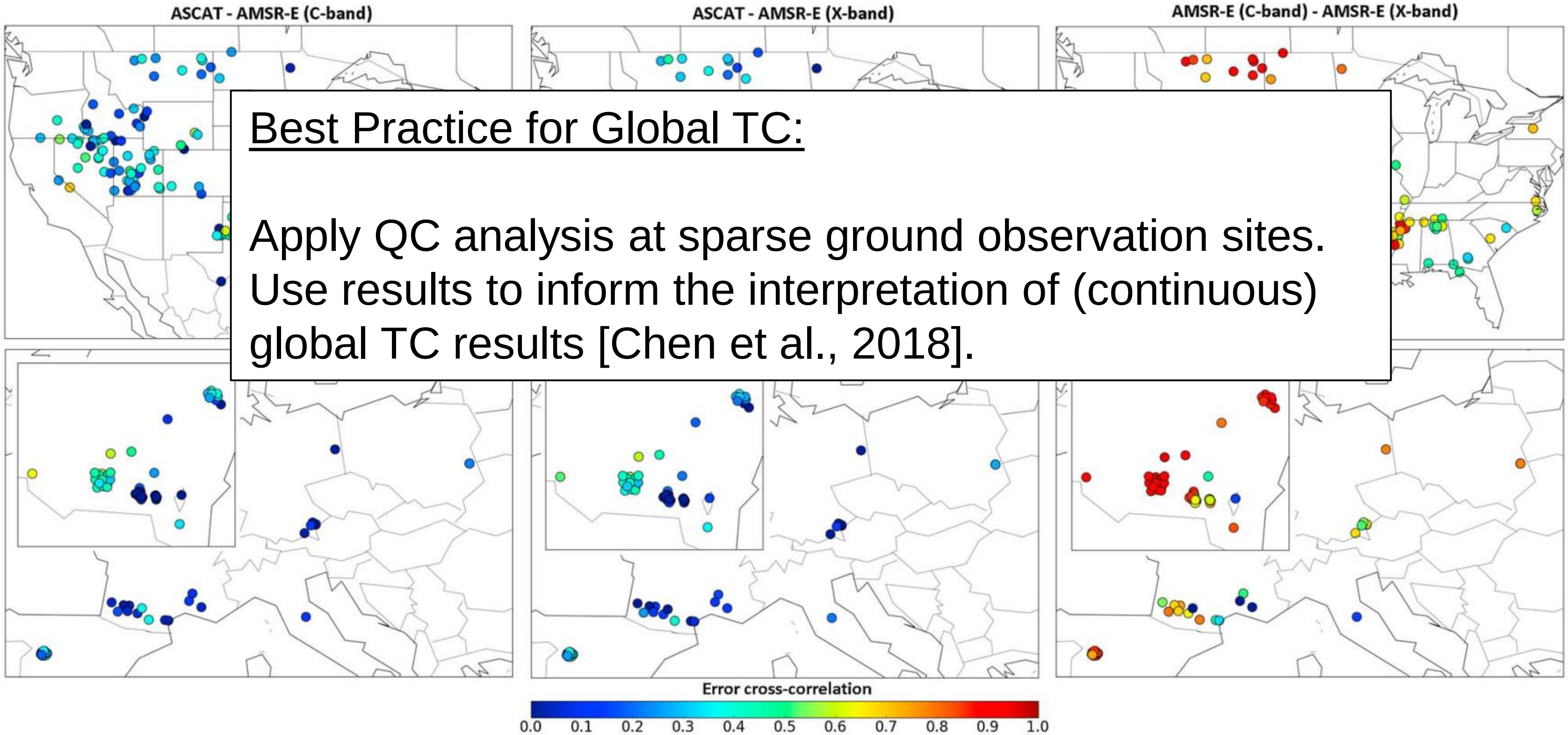
Ways in which this over-constraint can be leveraged:

1) Apply least-squares regression to reduce sampling uncertainties [Pierdicca et al. 2015].

2) Make certain assumptions about the presence of cross-correlated errors and solve for a sub-set of error cross-correlation parameters [Gruber et al. 2016; Pierdicca et al. 2017].

SM Error Cross-Correlation

[LSM, ground-based, active, passive] Gruber et al. [2016]



Best Practice for Global TC:

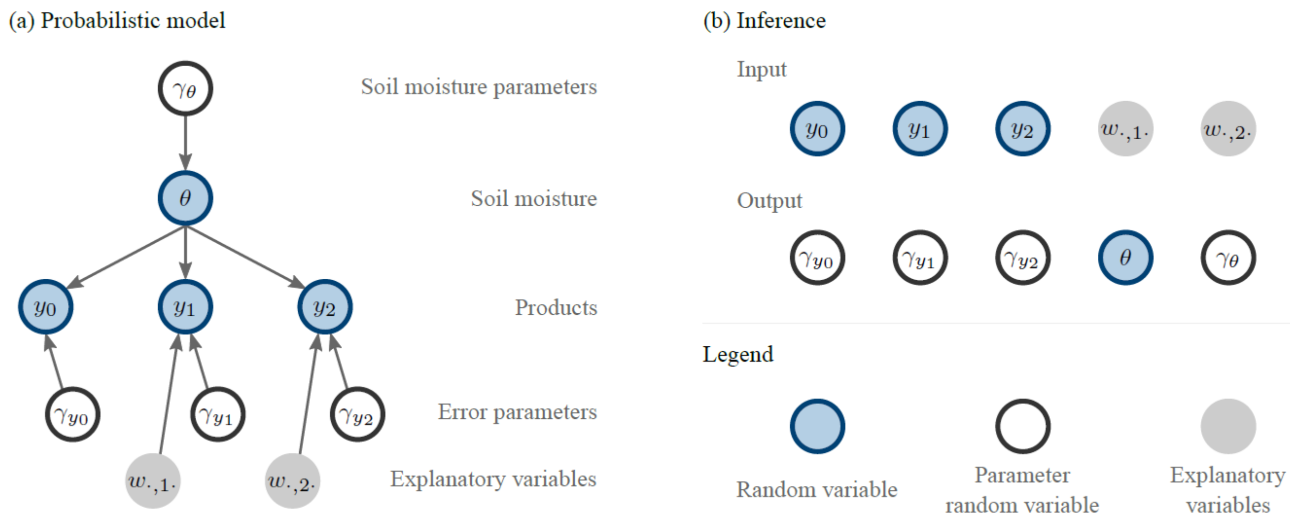
Apply QC analysis at sparse ground observation sites. Use results to inform the interpretation of (continuous) global TC results [Chen et al., 2018].

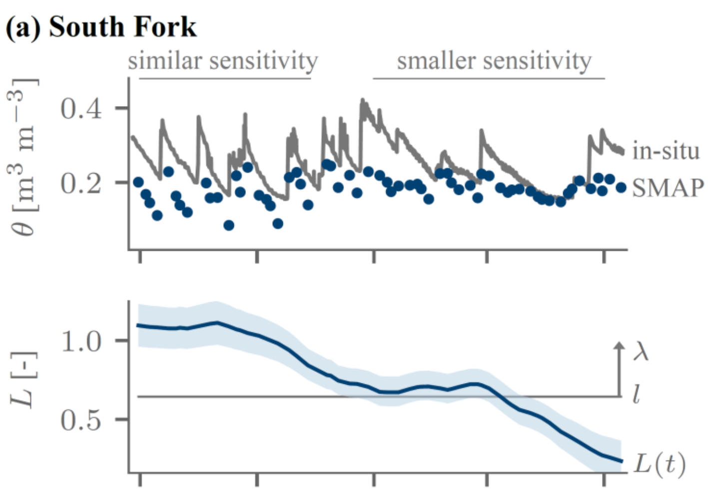
Non-negligible error cross-correlation between AMSR-E and ASCAT:

Similar results found in: Yilmaz and Crow [2014], Pierdicca et al. [2017] and Chen et al. [2018] (w/ SMAP and ASCAT).

Error Covariance with Ancillary Variables

Zwieback et al. [2018]





- SMAP retrieval errors may co-with ancillary deterministic variables (e.g., vegetation opacity).

- Large implications for SMAP-based SM coupling analysis. <u>*Is some observed coupling spurious?*</u>

Recent TC Developments/Future Directions

To date we have applied very simple statistical models to describe random soil moisture retrieval errors (i.e., additive, stationarity, white and independent).

Reality is messier and we run the risk of mis-interpreting and/or under-utilizing SMAP soil moisture products if we fail to recognize this.

Variants of TC can provide:

1) Temporal error auto-covariance [Dong et al. 2017].

2) Spatial structure of error cross-correlation [Gruber et al. 2018].

3) Error covariance in SM products [Pierdicca et al. 2017; Gruber et al. 2016].

4) Error covariance with ancillary variables [Zwieback et al. 2018].

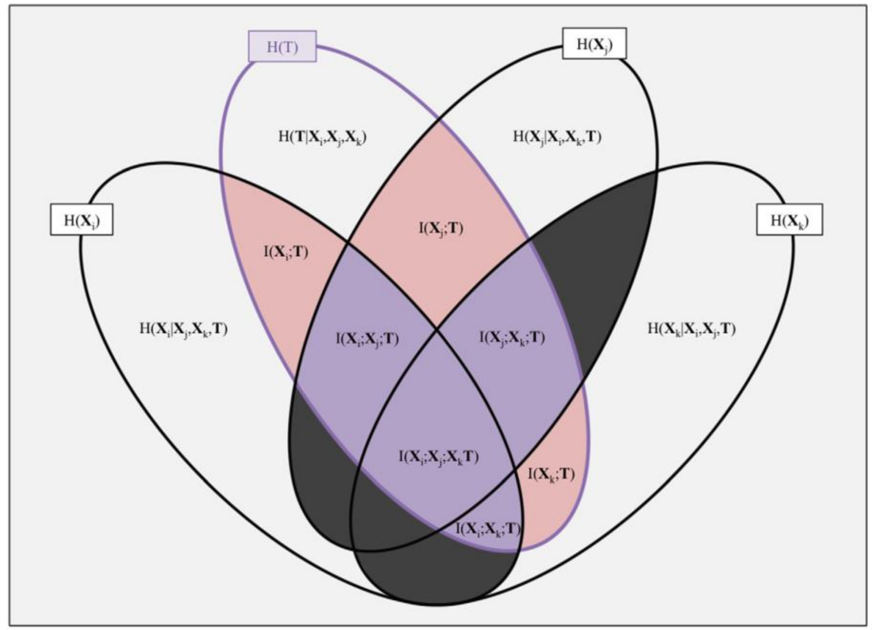
5) Non-parametric mutual information [Nearing et al. 2017].

Non-Parametric TC

Nearing et al. [2017]

$$\cancel{\begin{aligned} X &= \alpha_X T + \varepsilon_X \\ Y &= \alpha_Y T + \varepsilon_Y \\ Z &= \alpha_Z T + \varepsilon_Z \end{aligned}}$$

T = “true” soil moisture



Triple Collocation

1) Assume obs. linear w.r.t truth
2) Assume independent errors
3) Sample covariances
4) Estimate error variances

Non-Parametric TC

1) No assumed functional form.
2) Assume independent errors.
3) Cross-sample mutual information
4) Estimate residual entropies AFTER conditioning by truth.

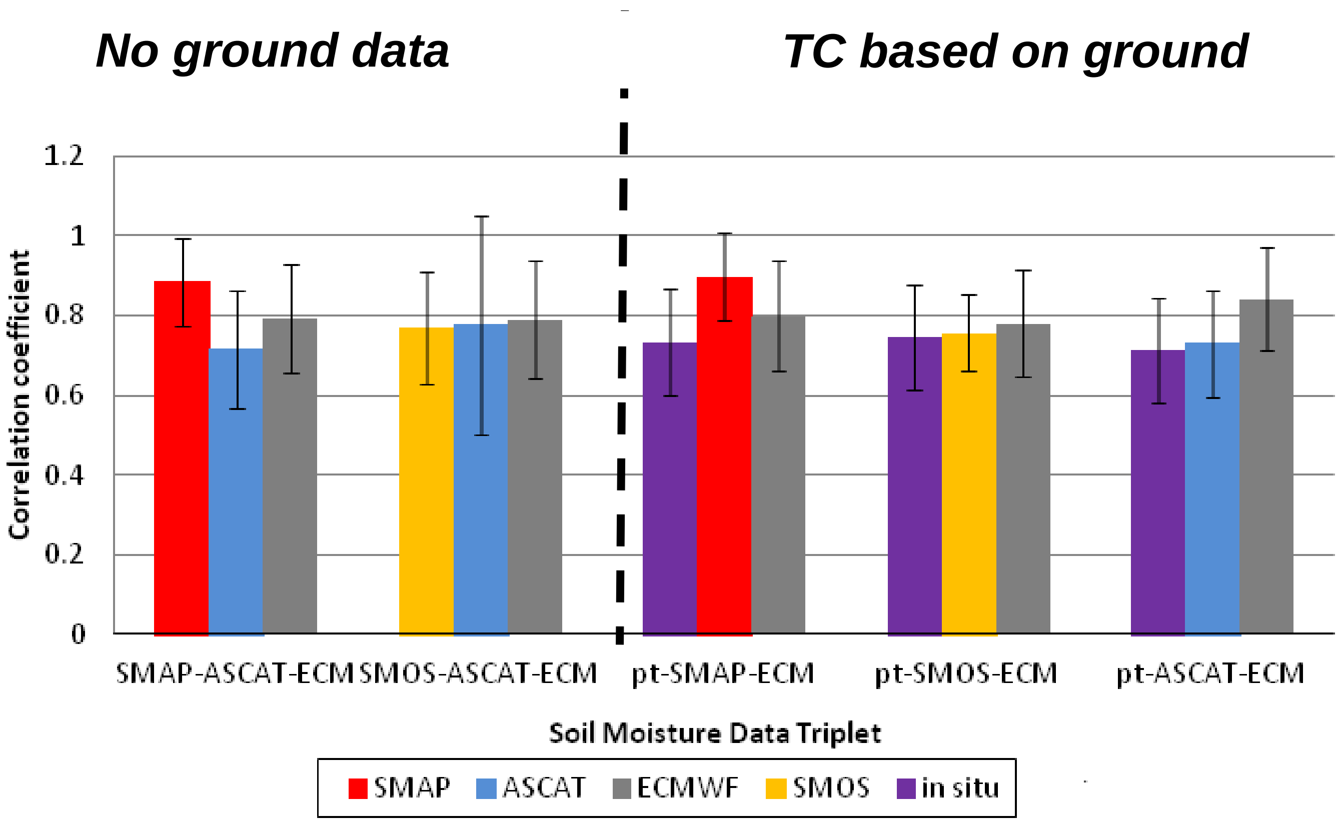
Summary

- Not a complete review (left out a number of excellent papers).
- Bad news: Purely independent, white noise error model is problematic as we push the boundaries on using SMAP data.
- Good news: Seeing steady progress towards a more sophisticated view of soil moisture remote sensing errors.
- TC and its variants are playing an important role in this progress.
- Soil moisture community is taking a lead role in the development and application of these techniques.

Work Cited

- Chen, F., Crow, W.T., Colliander, A., Cosh, M.H., Jackson, T.J., Bindlish, R., Reichle, R.H., Chan, S.K., Bosch, D.D., Starks, P.J. and Goodrich, D.C. Application of triple collocation in ground-based validation of Soil Moisture Active/Passive (SMAP) level 2 data products. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 99:1-14. 10.1109/JSTARS.2016.2569998. 2016.
- Chen, F., Crow, W.T., Bindlish, R., Colliander, A., Burgin, M.S., Asanuma, J. and Aida, K. Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation. *Remote Sensing of Environment*. 214:1-13. 10.1016/j.rse.2018.05.008. 2018.
- Dong, J., Crow, W.T. and Bindlish, R. The error structure of the SMAP single and dual channel soil moisture retrievals. *Geophysical Research Letters*. 45:758-765. 10.1002/2017GL075656. 2017.
- Gruber, A., Su, C.-H., Crow, W.T., Zwieback, S., Dorigo, W.A. and Wagner, W. Estimating error cross-correlations in soil moisture data sets using extended collocation analysis. *Journal of Geophysical Research-Atmospheres*, 121:1208-1219. 10.1002/2015JD024027. 2016.
- Gruber, A., Crow, W.T. and Dorigo, W. Assimilation of spatially sparse in situ soil moisture networks into a continuous model domain. *Water Resources Research*. 54:1353-1367. 10.1002/2017WR021277. 2018.
- Pierdicca, N., F. Fascetti, L. Pulvirenti, R. Crapolicchio and J. Muñoz-Sabater. Quadruple Collocation Analysis for Soil Moisture Product Assessment. *IEEE Geoscience and Remote Sensing Letters*, 12(8):1595-1599, 10.1109/LGRS.2015.2414654. 2015.
- Pierdicca, N., F. Fascetti, L. Pulvirenti and R. Crapolicchio, Error Characterization of Soil Moisture Satellite Products: Retrieving Error Cross-Correlation Through Extended Quadruple Collocation, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(10):4522-4530, 10.1109/JSTARS.2017.2714025. 2017.
- Zwieback, S., Colliander, A., Cosh, M. H., Martínez-Fernández, J., McNairn, H., Starks, P. J., Thibeault, M., and Berg, A.: Estimating time-dependent vegetation biases in the SMAP soil moisture product, *Hydrol. Earth Syst. Sci.*, 22:4473-4489, 10.5194/hess-22-4473-2018, 2018.

Verification of Global TC at Sparse Sites



TC(Model, RS-Passive, ASCAT) ⟵ ⟶ **TC(Model, RS, Sparse)**

RS-Passive = SMOS or SMAP SM_P

RS = SMOS, ASCAT or SMAP

Success of Sparse Network Analysis

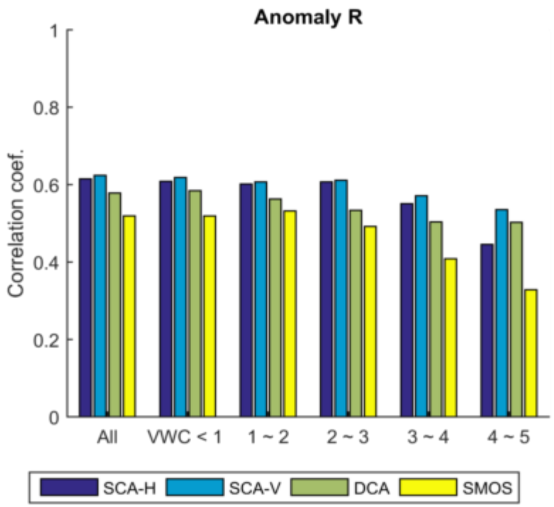
• Relative sparse network analysis consistently mirrored those of the core sites:

1) SCA-V > SCA-H > DCA

2) 6 am > 6 pm (but only slightly)

3) SMAP > SMOS

• Able to confidently extrapolate core site results into a wider range of land cover conditions:



All this was done ***without*** the application of triple collocation…..

Global TC results

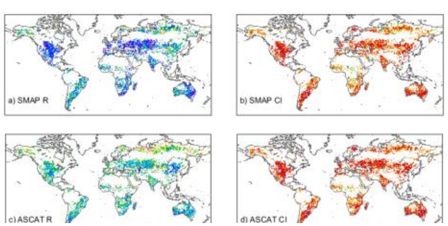
Correlation (R)

90% Confidence Interval

SMAP L3_SM_P: ver4-R14010

SMOS Level 3 v300

ASCAT 12.5 km Level 2



Summary

• Point-to-footprint scaling is daunting…however, large sample size of sparse sites make relative comparisons reliable.

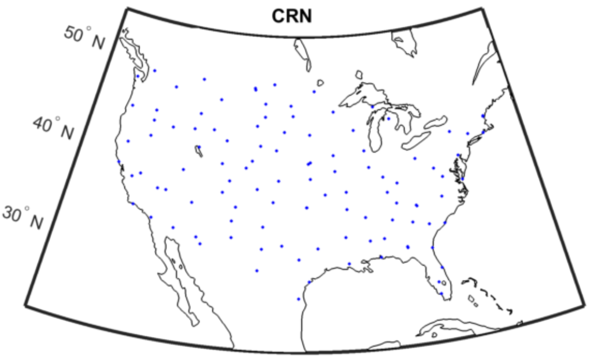
• TC (or any other up-scaling) is not required to extract useful information from sparse network locations.

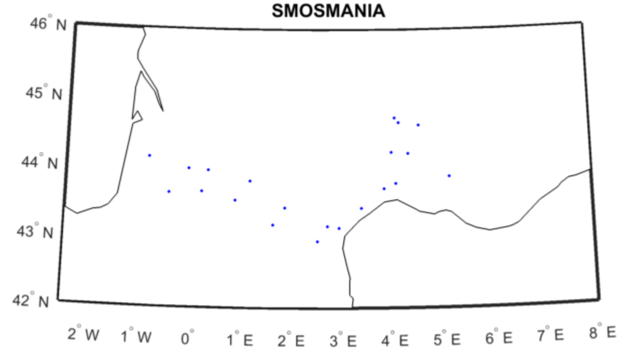
• When applied appropriately, TC is robust and is providing useful information for expanded (quasi-global domains).

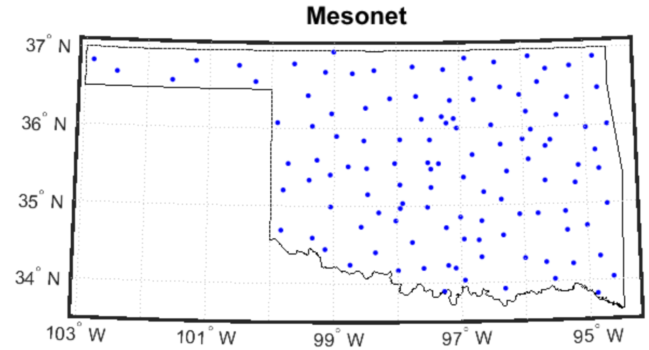
Chen, F. et al. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 99:1-14. 10.1109/JSTARS.2016.2568888. 2016.

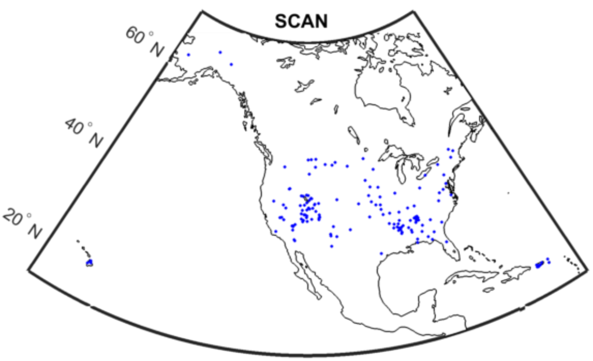
Chen, F. et al. Global-scale Evaluation of passive and active soil moisture products using Triple Collocation. *Geophysical Research Letters*. In preparation.

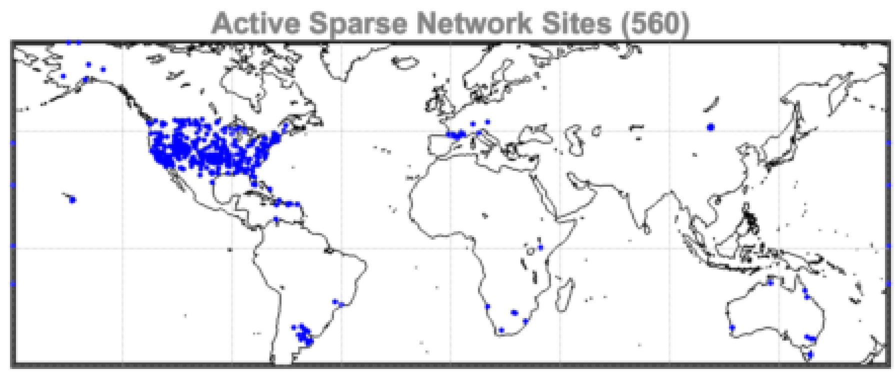
Active Sparse Networks

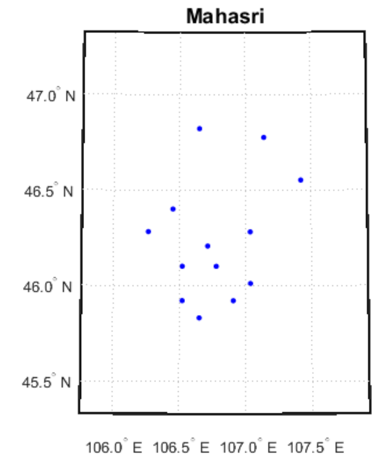


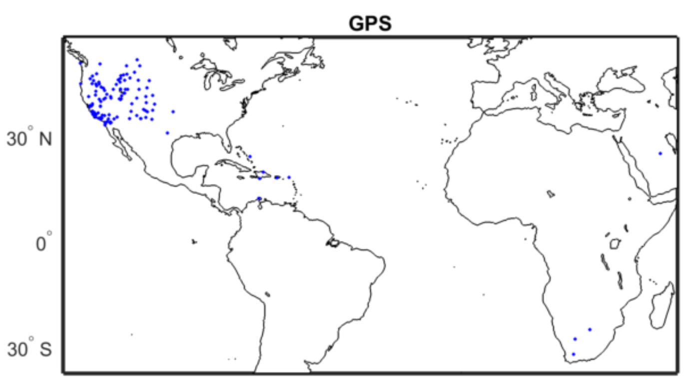


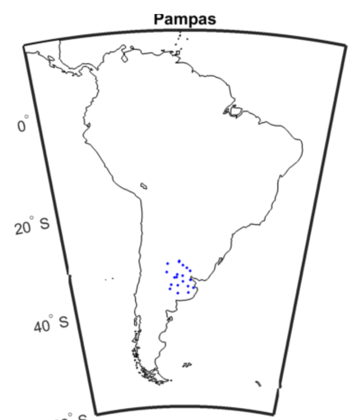


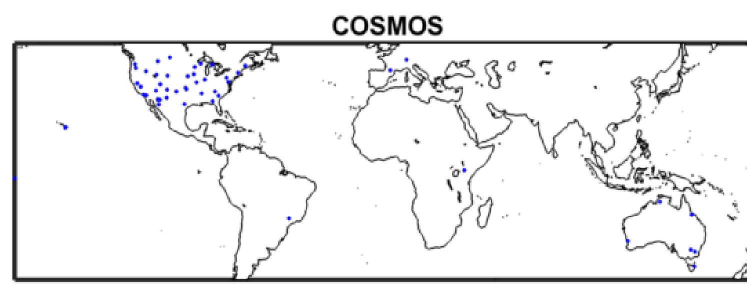












New Role of TC

To date, triple collocation (TC) results have been based on a triplet of:

TC(Model, SMAP SM_P, Sparse)

These results have been validated (Chen et al., JSTARS, 2016) but limit the analysis to sparse network cites.

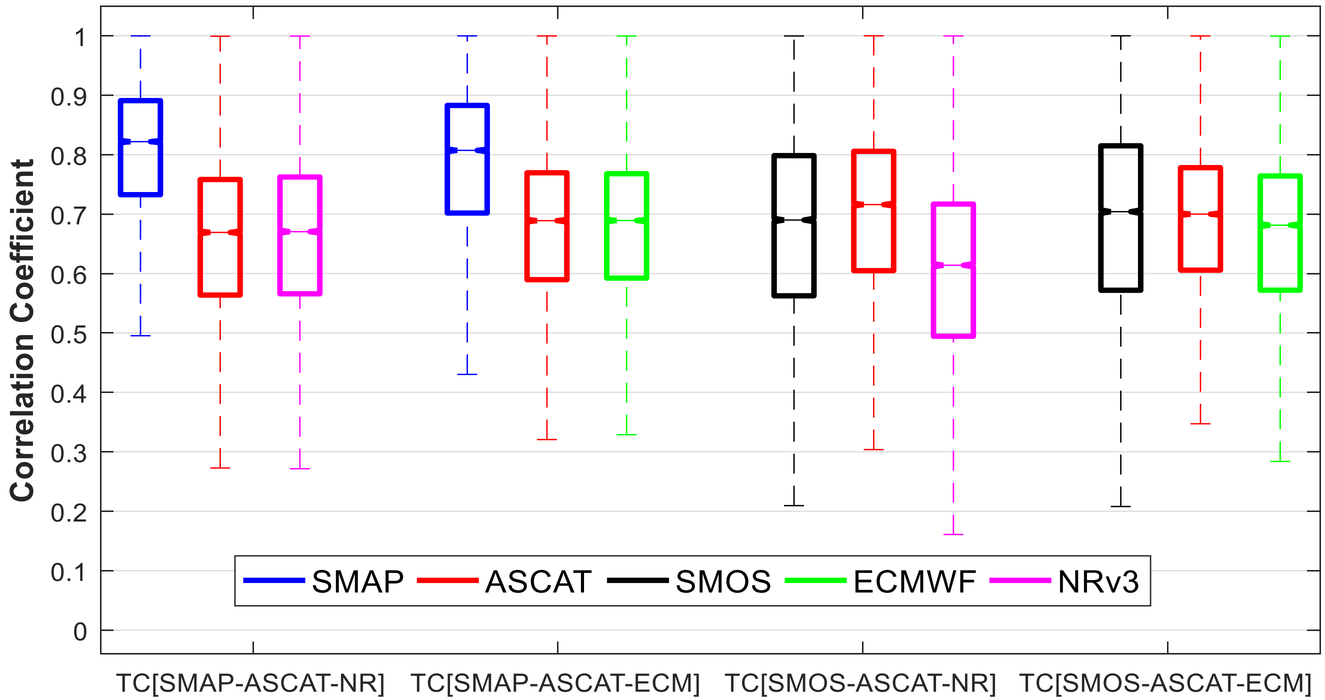
The next step is to geographically expand the scope of the TC analysis by utilizing:

TC(Model, SMAP SM_P, ASCAT)

(i.e., swap ***sparse*** site ground locations for ***ASCAT*** retrievals)

To begin, we need to verify (as predicted by TC assumptions) that this swap does induce bias into TC results….

Impact of LSM



SMAP has a relationship with NRv3 (GOES-5 Catchment)
SMOS has a relationship with ECMWF

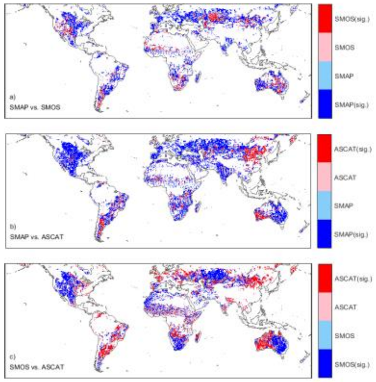
Issues: Does this skew TC results?.

Global TC results

(68%) SMAP vs (32%) SMOS

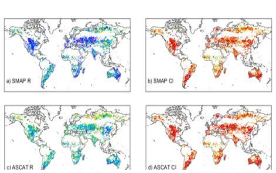
(72%) SMAP vs (28%) ASCAT

(47%) ASCAT vs (53%) SMOS



New TC Contributions to SMAP Cal/Val

1. Robust global maps of unbiased anomaly correlation (plus 95% sampling bounds) [Chen et al. 2018]:



2. ubRMSE assessment of core-site average soil moisture values [Chen et al. 2019]: