# Upscaling Ground-based Soil Moisture Measurements for SMAP Validation Activities

Wade T. Crow

USDA Hydrology and Remote Sensing Laboratory

USDA-ARS HRSL Hydrology & Remote Sensing Lab Beltsville, Maryland, USA

***3rd SMAP Cal/Val Workshop Oxnard, CA, November 2012***

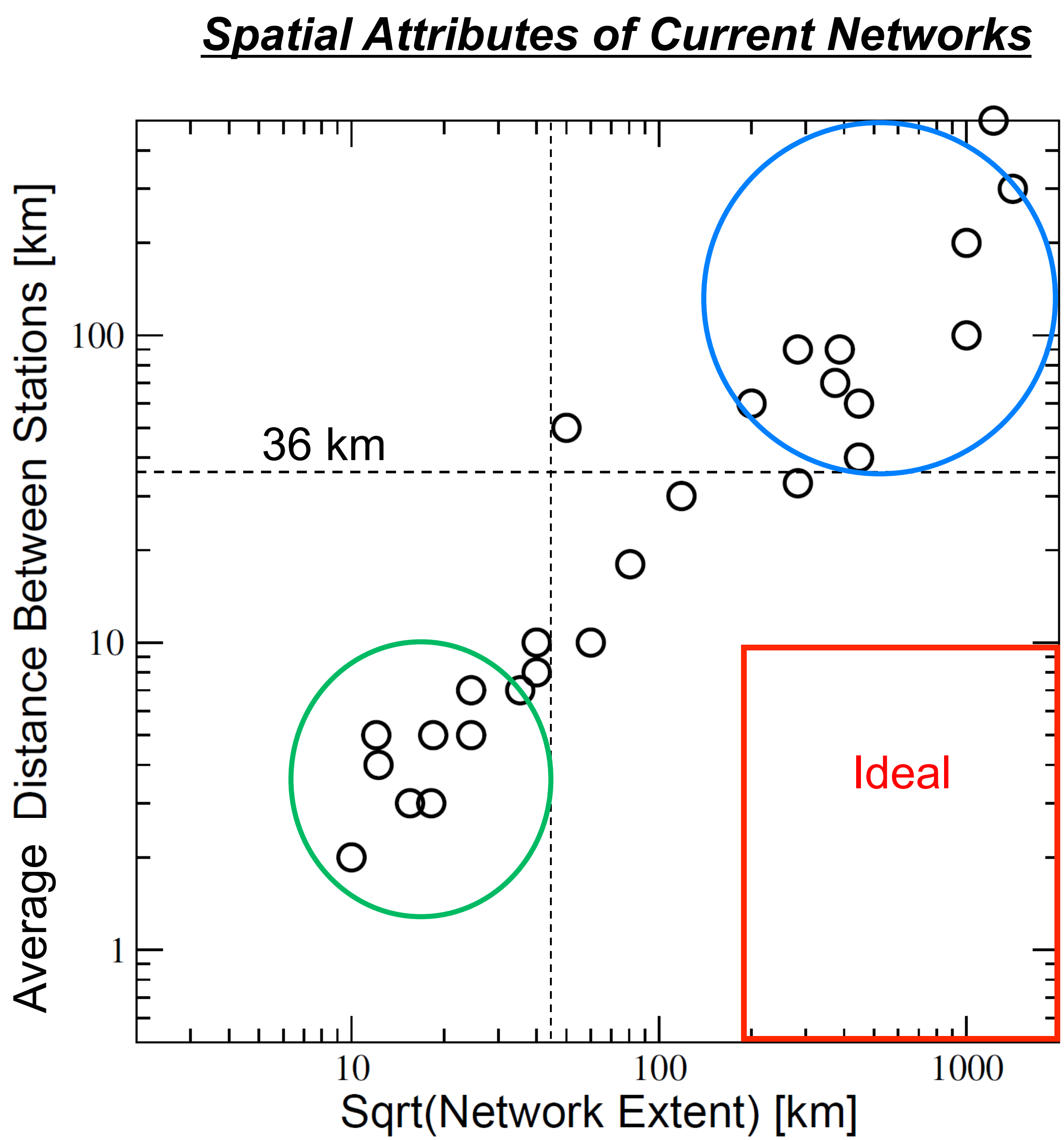
USDA Agricultural Research Service

## ***SMAP Val/Cal Working Group on Upscaling***

- Established during 1st SMAP cal/val workshop (35 members).
- Objective: Maximizing the utility of "sparse" soil moisture network observations for SMAP validation activities.
- Level 2/3 soil moisture focus (at least initially).
- Core activity: Review paper on existing soil moisture upscaling research:

Crow, W.T., A.A. Berg, M.H. Cosh, A. Loew, B.P. Mohanty, R. Panciera, P. de Rosnay, D. Ryu and J.P. Walker, "Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products," *Reviews of Geophysics*, 50, RG2002, doi: 10.1029/2011RG000372, 2012.

Presented findings ***do not*** represent a consensus view of the working group…open for discussion.




“Core” SMAP validation networks (~5-10 obs/ SMAP footprint).

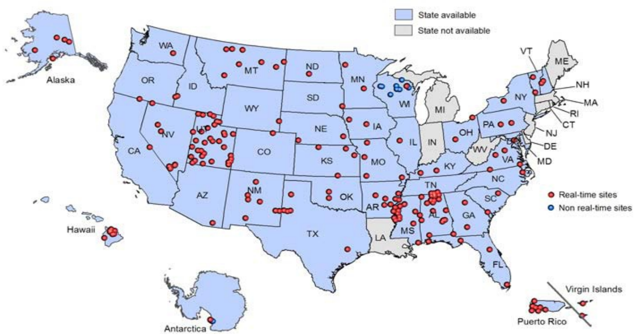
Limited extent but covering a range of biomes/climate/ land cover.

*Backbone* of SMAP validation efforts.

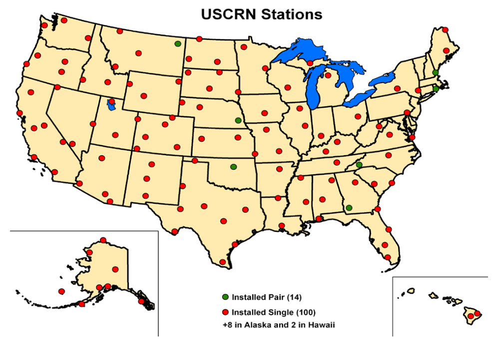
“Sparse” SMAP validation networks (~1-2 obs/ SMAP footprint).

*Supplemental* but a validation resource that cannot be neglected.

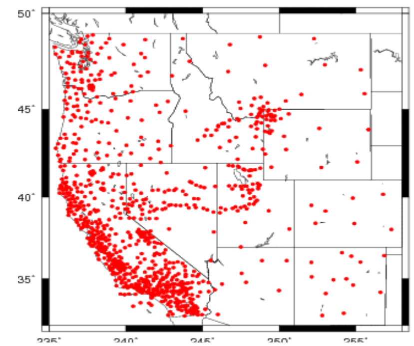
# ***Current Sparse SMAP Networks***



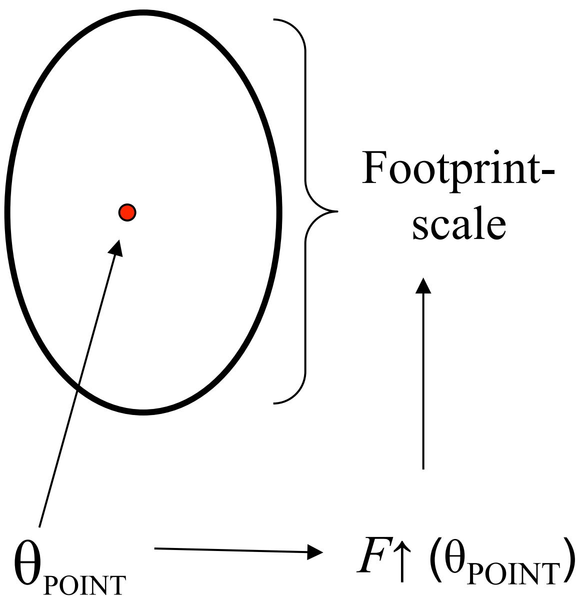

**USDA Soil Climate Analysis Network (SCAN)**



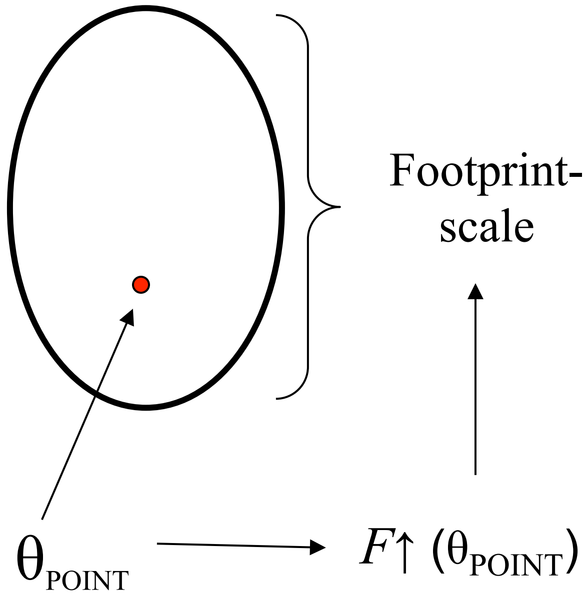

**NOAA Climate Reference Network (CRN)**




**GPS observations**




**Upscaling Challenge:** Using point-scale soil moisture observations to validate footprint-scale SMAP retrievals.




**Upscaling Challenge:**
Using point-scale soil moisture observations to validate footprint-scale SMAP retrievals.

**Proposed best practices*:**

***First***, apply temporal stability analysis (Cosh et al., 2006; 2008) to select sampling sites with temporal dynamics that best mimic footprint scale variability.

***Second***, use land surface modeling (Crow et al., 2005), kriging, and/or an intensive field campaign (De Rosnay et al. 2009) to refine understanding of the relationship between point- and footprint-scale variability (i.e., $F\uparrow$ on left).

***Third***, apply triple collocation (Miralles et al., 2010) to estimate impact of residual sampling errors on RMSE validation results.

*Based on: Crow et al., “Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products,” *Reviews of Geophysics*, 50, RG2002, doi:10.1029/2011RG000372, 2012.

# ***Triple Collocation (TC) Application to Soil Moisture:***

## ***TC using [active, passive, land surface modeling] soil moisture:***

1) Scipal et al., *GRL*, (2008)
2) Dorigo et al., *HESS*, (2010)
3) Crow et al., *IAHS Redbook* (2012)
4) Crow et al., *WRR*, (2010)
5) Parinussa et al., *HESS*, (2011)
6) Leroux et al., *IGARSS*, (2011)
7) Zieback et al., *HESS*, (2012)

## ***TC using [thermal, passive, land surface modeling] soil moisture:***

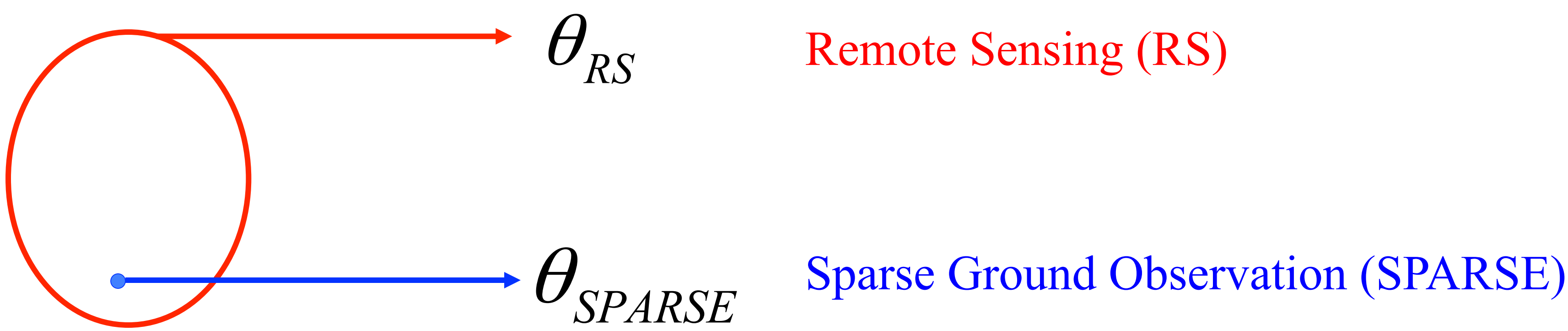
1) Hain et al., *JGR*, (2011)
2) Anderson et al, *HESS*, (2012)

## ***TC using [sparse ground, passive, land surface modeling] soil moisture:***

1) Miralles et al., *GRL*, (2011)
2) Loew et al., *HESS*, (2012)

## ***Application of Triple Co-Location To Estimate Random Sampling Error in Sparse Ground Observations:***

Mean-square difference comparisons between point-scale observations and footprint-scale retrievals will be inflated by spatial sampling errors.

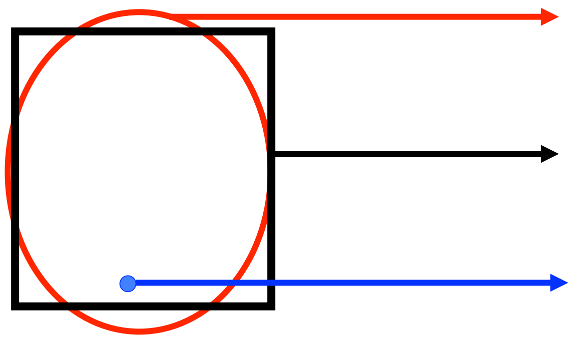



$$\mathrm{MSD}(\theta_{\mathrm{TRUE}}, \theta_{\mathrm{RS}}) = \mathrm{MSD}(\theta_{\mathrm{SPARSE}}, \theta_{\mathrm{RS}}) - \mathrm{MSD}(\theta_{\mathrm{SPARSE}}, \theta_{\mathrm{TRUE}})$$

**(TARGET)** **(AVAILABLE)** **(CORRECTION)**

## *Application of Triple Co-Location To Estimate Random Sampling Error in Sparse Ground Observations:*

1) Obtain three independent (and uncertain) estimates of *footprint-scale* soil moisture:



$\theta_{RS}$ Remote Sensing (RS)

$\theta_{LSM}$ Land Surface Model (LSM)

$\theta_{SPARSE}$ Sparse Ground Observation (SPARSE)

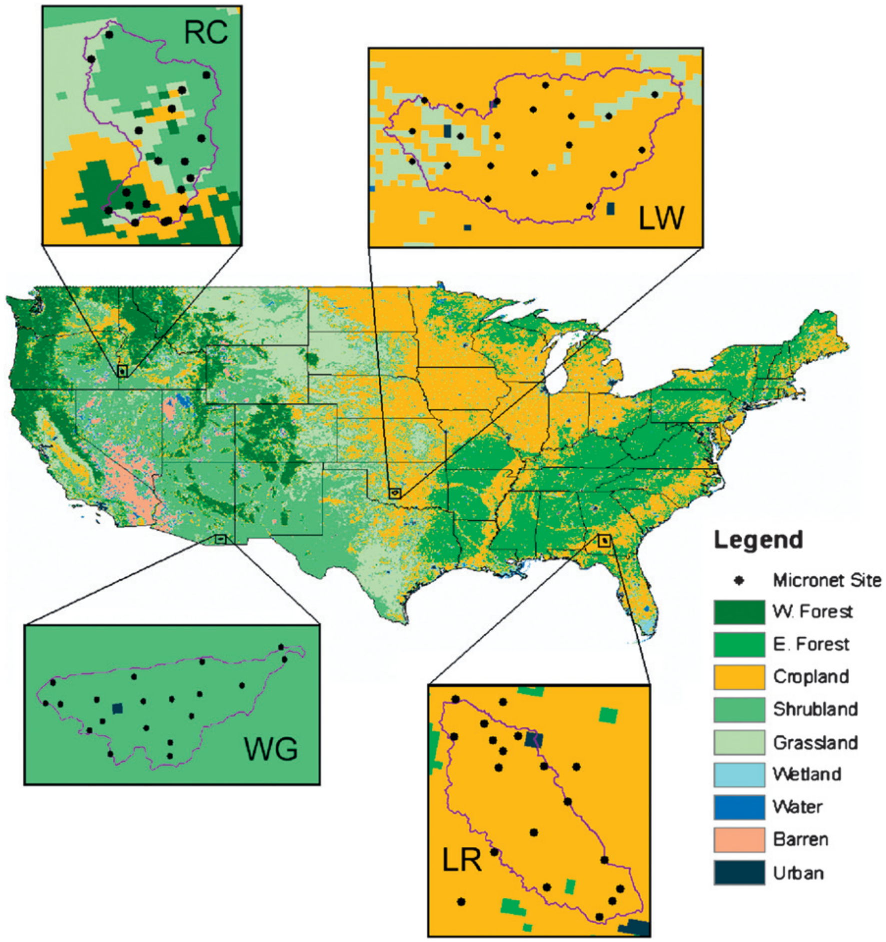
2) Assume independent errors and sample the following temporal average to estimate random sampling error in SPARSE:

$$\overline{(\theta_{SPARSE} - \theta_{RS})(\theta_{SPARSE} - \theta_{LSM})} = MSD(\theta_{SPARSE}, \theta_{TRUE})$$

3) Use this estimate to *correct* soil moisture RMSE estimates derived from RS versus SPARSE comparisons for sampling error in SPARSE.

$$\mathrm{MSD}(\theta_{\mathrm{TRUE}}, \theta_{\mathrm{RS}}) = \mathrm{MSD}(\theta_{\mathrm{SPARSE}}, \theta_{\mathrm{RS}}) - \mathrm{MSD}(\theta_{\mathrm{SPARSE}}, \theta_{\mathrm{TRUE}})$$

# ***Verification Methodology***



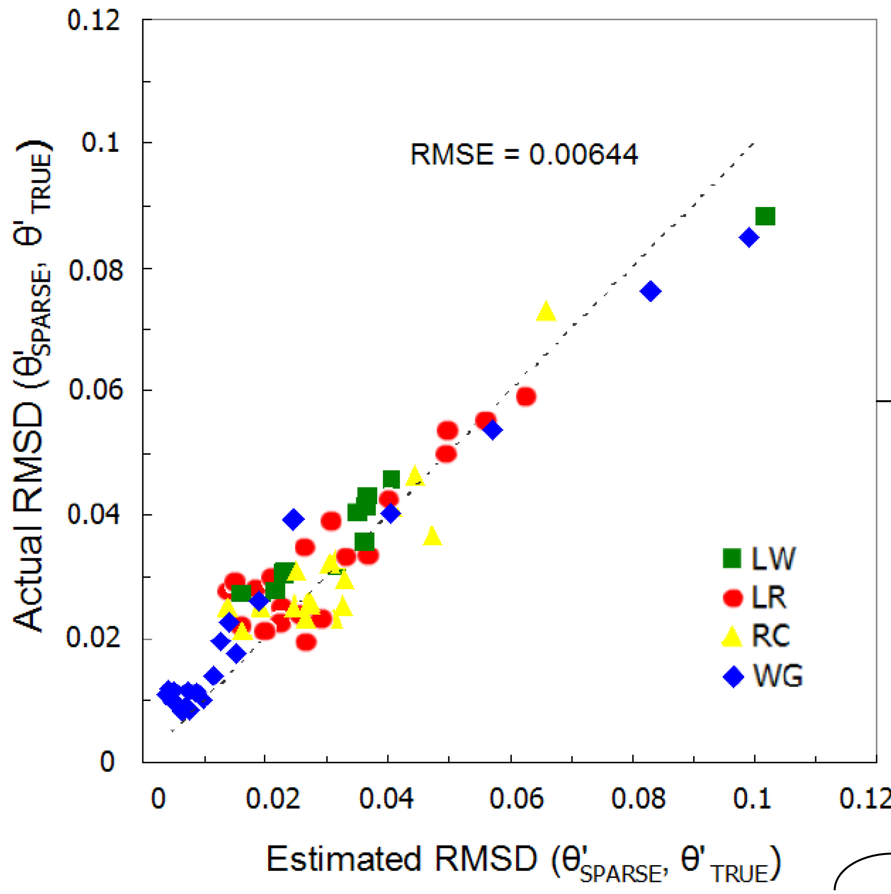

1) 1-D NOAH Land Surface Model with NLDAS forcing.
2) USDA SCA AMSR-E soil moisture product.
3) High-density soil moisture datasets from ARS watershed sites.
4) Targeting bias-corrected RMSE.

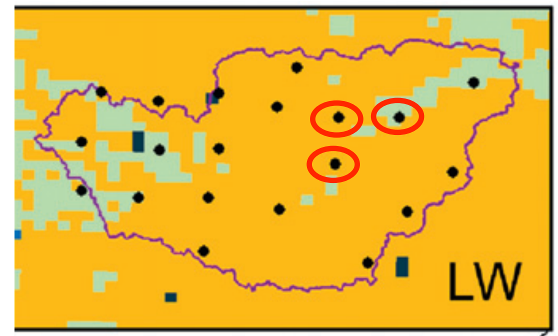
SPARSE = 1 station within each watershed.

TRUE = Comparison to average of all measurements within watershed.

Miralles, D.G., W.T. Crow and M.H. Cosh, "A technique for estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations," *Journal of Hydrometeorology,* 11(6), 1404-1410,10.1175/2010JHM1285.1, 2010.

# ***NOAH LSM + AMSRE + SPARSE TC Results***



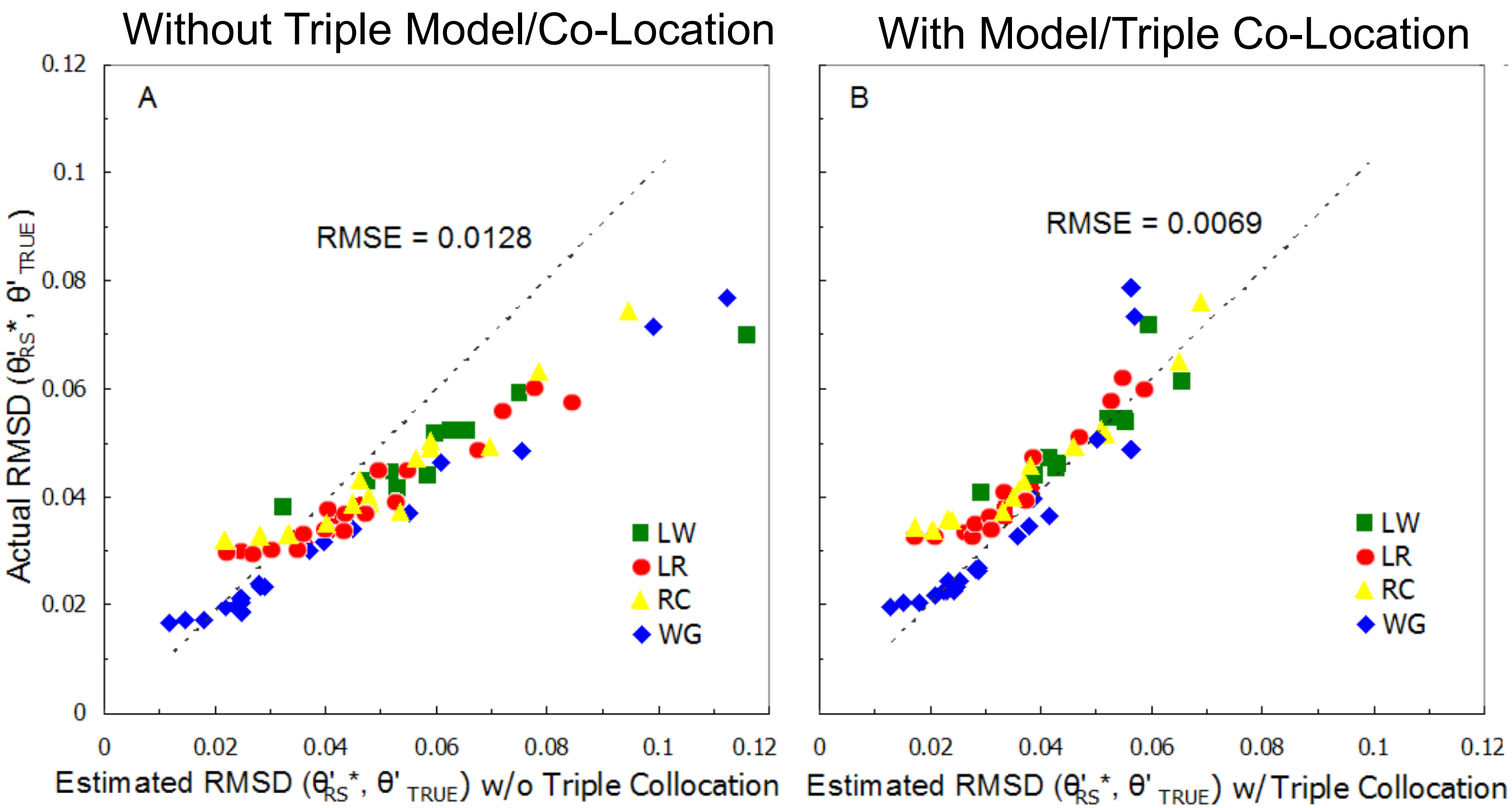

$$\underbrace{MSE(\theta_{RS}, \theta_{TRUE})}_{\text{TARGET}} = \underbrace{MSE(\theta_{RS}, \theta_{SPARSE})}_{\text{AVAILABLE}} - \overbrace{MSE(\theta_{SPARSE}, \theta_{TRUE})}^{\text{SAMPLING ERROR}}$$

**TARGET**

**AVAILABLE**

**SAMPLING ERROR**

# AMSR-E Validation Using One SPARSE station




*SPARSE ground obervations + TC can replicate (bias-corrected) RMSE results obtained from "dense" CORE sites to within 0.007 $m^3m^{-3}$*

**Triple Collocation:**

- Can also recover (and correct for) random instrument errors.

- Works with new ground-based measurement techniques with coarser spatial supports (e.g., COSMOS, GPS and fiber-optic).

**Despite promise, *not* a panacea for all scaling-related challenges….**

**Shortcomings (partial):**

- Reduced temporal sampling power

- Neglects systematic error sources

- Sensitivity to cross-correlated errors

- Uncertain application to SMAP freeze-thaw products

1) **A TC-based analysis has less temporal sampling power than an RMSD analysis against known truth.**

- Relative error of sampled MSD estimates:

For core-site validation against “known truth”: $\sqrt{\frac{2}{\mathrm{DOF}}}$

For TC-based evaluation over sparse networks: $\sqrt{\frac{5}{\mathrm{DOF}}}$

- In a twelve-month period, $40 < \mathrm{DOF} < 100$ for surface soil moisture.

- Equates to relative TC sampling errors on the order of 20%-30%.

***Impact***: May struggle to meet current SMAP timeline of completed assessment at launch+15 months.

**2) TC is only effective at measuring *random* error and will be insensitive to *systematic* error in SMAP retrievals.**

Prior to TC calculation, LSM and RS observations must be linearly rescaled (and de-biased) to "match" the SPARSE observations.

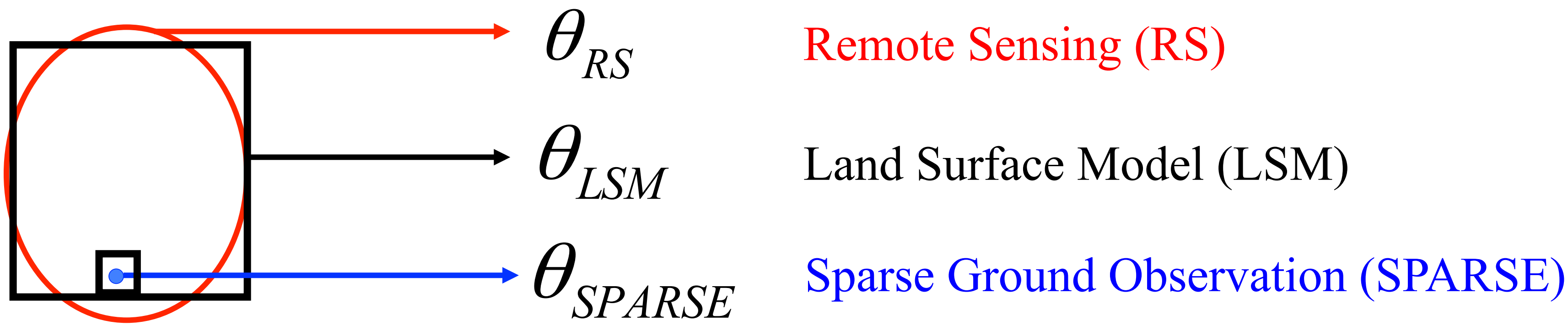
***Impact:*** TC-based RMSE estimates of point-footprint sampling errors are likely to be low (resulting is a conservative correction).

$$\mathrm{MSD}(\theta_{\mathrm{TRUE}}, \theta_{\mathrm{RS}}) = \mathrm{MSD}(\theta_{\mathrm{SPARSE}}, \theta_{\mathrm{RS}}) - \mathrm{MSD}^*(\theta_{\mathrm{SPARSE}}, \theta_{\mathrm{TRUE}})$$

↑ Biased ***high*** as a result

↑ Biased ***low*** by neglect of systematic error sources

**3) TC is sensitive to even small amounts of cross-correlated error. This cross-correlation can emerge from subtle spatial "representativeness" issues.**



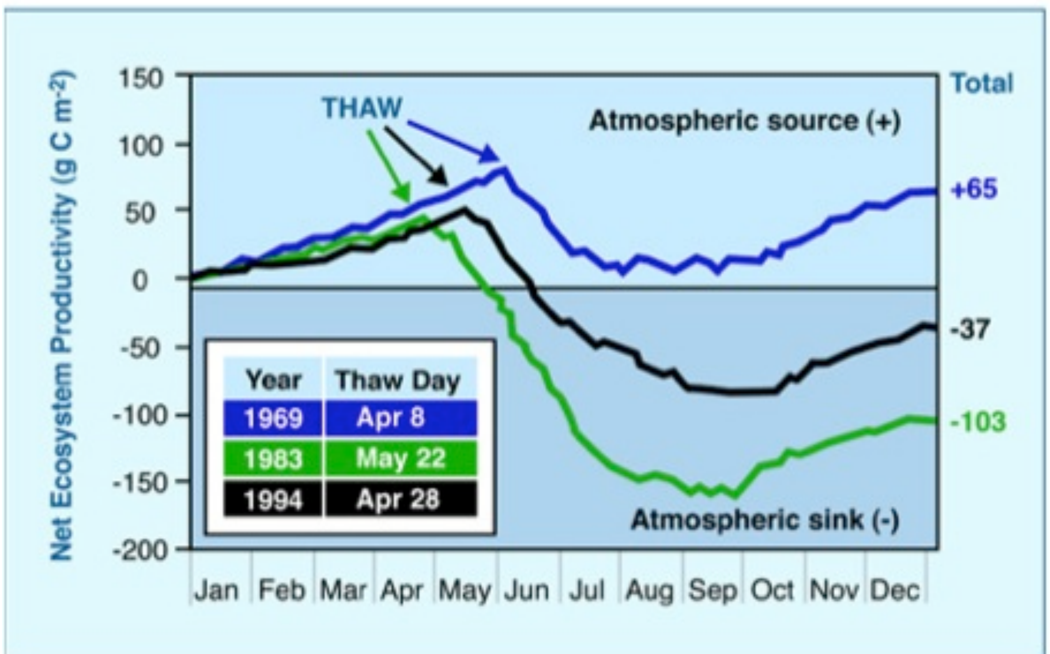

***Impact:*** Can bias TC-based MSD estimates. Potentially problematic given ambiguities in the "spatial support" of both LSM and SMAP soil moisture products.

Need to verify TC prediction additional observations (using either dense ground-based observations at core sites or independent remote sensing observations).

Potential roles for "independent" soil moisture products derived from (e.g.): ASCAT, SMAP L2/3_SM_A, and SMOS.

## 4) Uncertain application to L2/L3 freeze/thaw products.

- TC can theoretically be applied to binary variables.
- Fewer effective DOF in daily freeze/thaw time series….sampling issues will be even more severe than for soil moisture.
- However, high spatial variability...trading spatial DOF for temporal DOF?



**Use Summer 2013 rehearsal campaign to address these shortcomings:**

1) Reduced temporal sampling power

*Can sampling errors in be estimated in a short period?*

2) Neglects systematic error sources

*Can parallel approaches be developed to estimate bias?*

3) Sensitivity to cross-correlated errors

*Are errors truly independent?*

4) Uncertain application to L2/3 freeze-thaw products

*Are observed DOF in freeze-thaw time series sufficient?*

*If not, can ergodic sampling approaches be developed?*

Thanks…

Possible discussion topics:

1) Role of sparse networks in general.

2) Plans/requirements for rehearsal activities.

3) Bias-correction strategies.

4) Methodological issues with TC.

Back-up slides…..

## 2b) Uncertainties in rescaling strategies

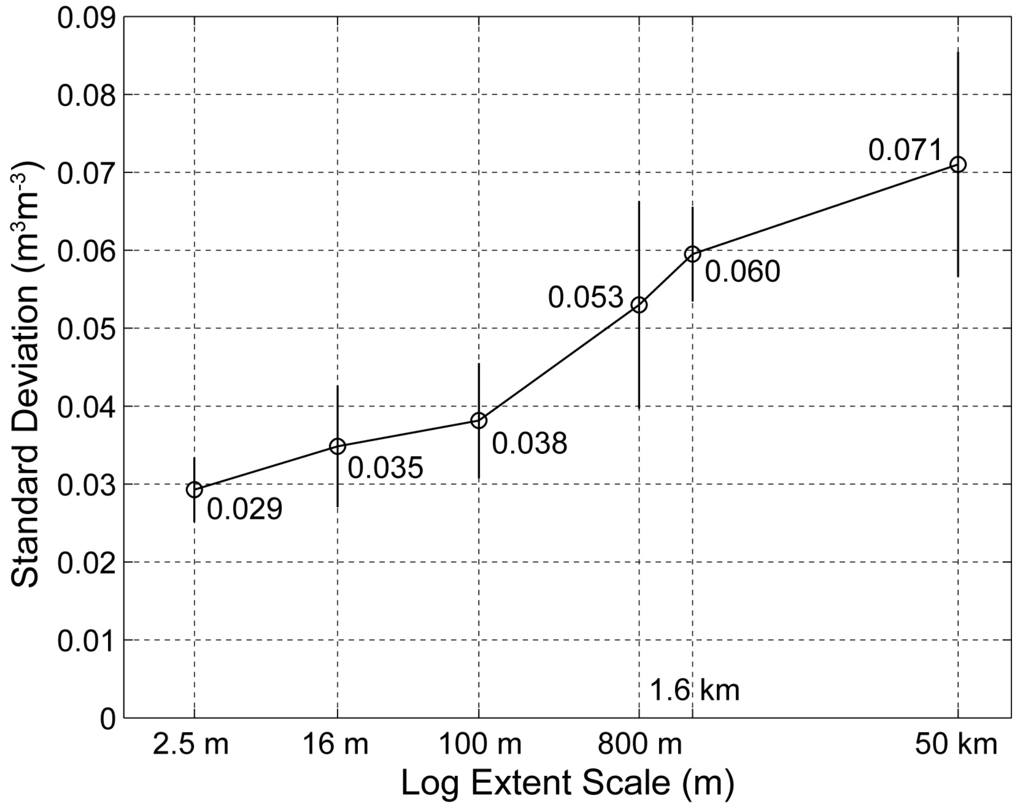
- Linear re-scaling based on equalizing means and variances is attractive but requires an implicit assumption of equal signal-to-noise ratios in all three products (Yilmaz et al., *JHM*, 2013). TC estimates will be biased if this assumption does not hold.

(easy to sample….but biased in some situations)

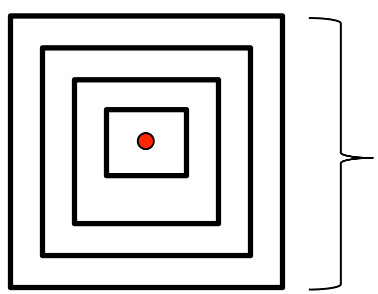
- Robust re-scaling requires sampling the ratio of cross-covariances between the three variables (Stoffelen et al., *JGR*, 1998). Prone to large sampling errors for low-skill products.

(unbiased…but difficult to sample in low-skill cases)

# ***Scope of upscaling problem/sampling density requirements***




(Famiglietti et al., 2008)



RMSE in using a single point-scale observations to characterize average soil moisture within various extent scales.

## ***Triple Collocation (TC) Application to Soil Moisture:***

***TC using [active, passive, land surface modeling] soil moisture:***

1) Scipal et al., *GRL*, (2008)
2) Dorigo et al., *HESS*, (2010)
3) Crow et al., *IAHS Redbook* (2012)
4) Crow et al., *WRR*, (2010)
5) Parinussa et al., *HESS*, (2011)
6) Leroux et al., *IGARSS*, (2011)
7) Zieback et al., *HESS*, (2012)

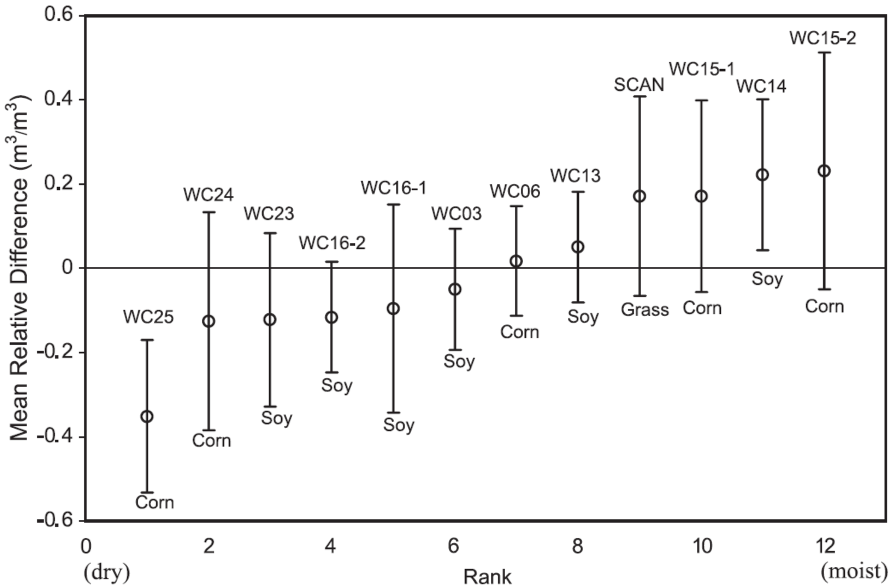
***TC using [thermal, passive, land surface modeling] soil moisture:***

1) Hain et al., *JGR*, (2011)
2) Anderson et al, *HESS*, (2012)

***TC using [sparse ground, passive, land surface modeling] soil moisture:***

1) Miralles et al., *GRL*, (2011)
2) Loew et al., *HESS*, (2012)

# ***Improve site selection based on time stability considerations***




~300-km$^2$ Walnut Creek Watershed in Central Iowa

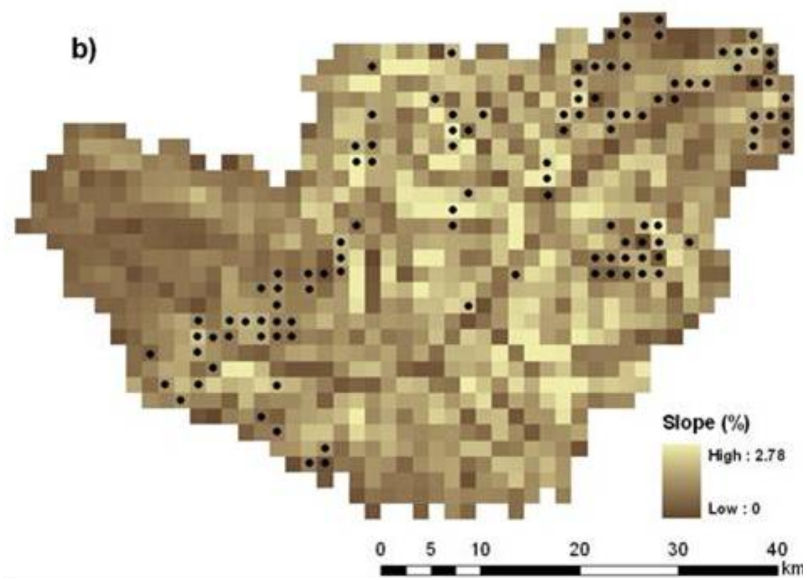
(Cosh et al. 2004)

Mean Relative Difference:

Point – Watershed / Watershed

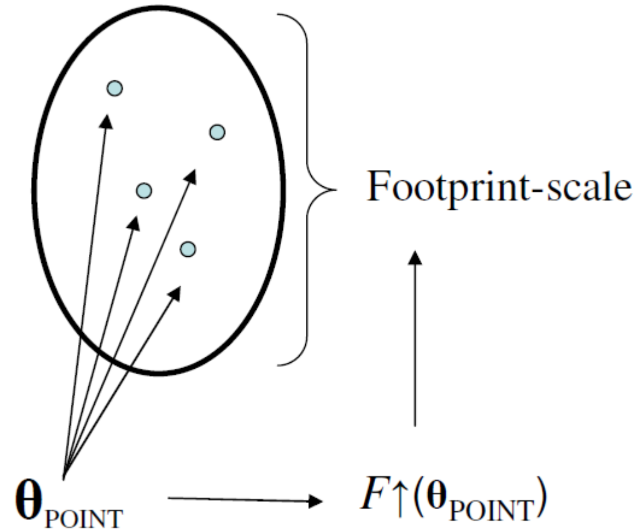
~900-km$^2$ Little Washita Watershed in Oklahoma

(Joshi et al. 2011)



## ***Improved functional forms for F↑***

Default: $\theta_{\text{UPSCALE}} = F_{\uparrow}(\boldsymbol{\theta}_{\text{POINT}}) = N^{-1}\sum_{i=1}^{N}\theta_{i,\text{POINT}}$




1. Improve using block kriging:

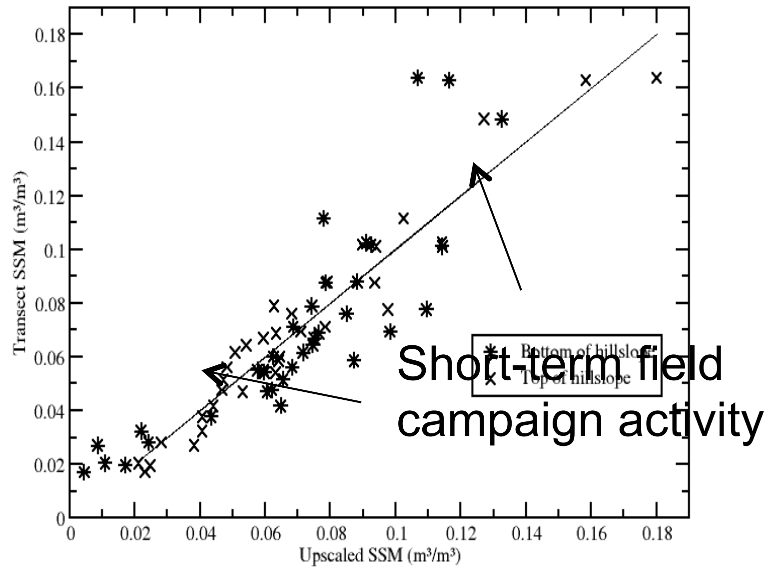
(Vinnikov et al. 1999)

$$\theta_{\text{UPSCALE}} = F_{\uparrow}(\boldsymbol{\theta}_{\text{POINT}}) = \sum_{i=1}^{N} w_i \theta_{i,\text{POINT}} = \frac{1}{\widehat{w}}\mathbf{C}^{-1}\mathbf{D}\,\boldsymbol{\theta}^{\text{T}}{}_{\text{POINT}}$$

# *Improved functional forms for F↑*

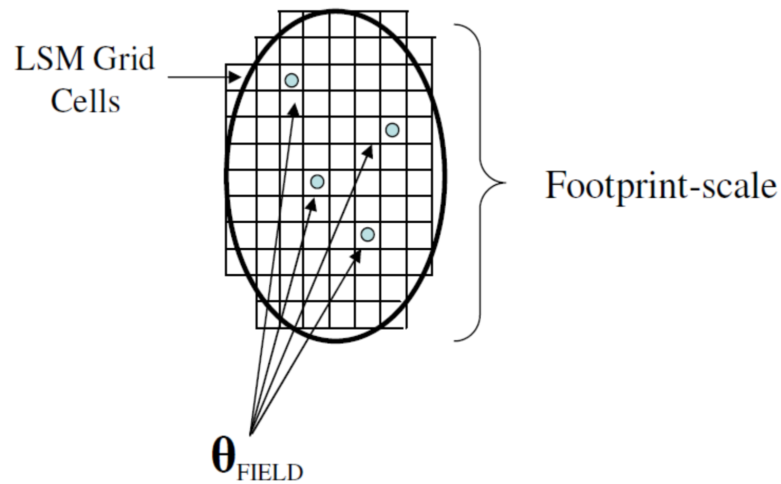
2. Improve using empirical relationships derived from field campaign data:

(de Rosnay et al. 2009)




3. Improve using Distributed Land Surface Modeling:

(Crow et al. 2004)




$$\theta_{\text{UPSCALE}} = F_{\uparrow}(\boldsymbol{\theta}_{\text{FIELD}}) = \theta_{\text{FP,LSM}} + N^{-1}\sum_{i=1}^{N}\left(\theta_{i,\text{FIELD}} - \theta_{i,\text{LSM}}\right)$$

***Different strategies can be applied simultaneously and may be complementary with regards to strengths e.g.:***

✓ Temporal stability is good for point to field upscaling, while model-based upscaling works best from field to footprint.

✓ Short-term field campaign data is good for correcting the bias component of upscaling error, Triple Collocation is good for addressing the random component.

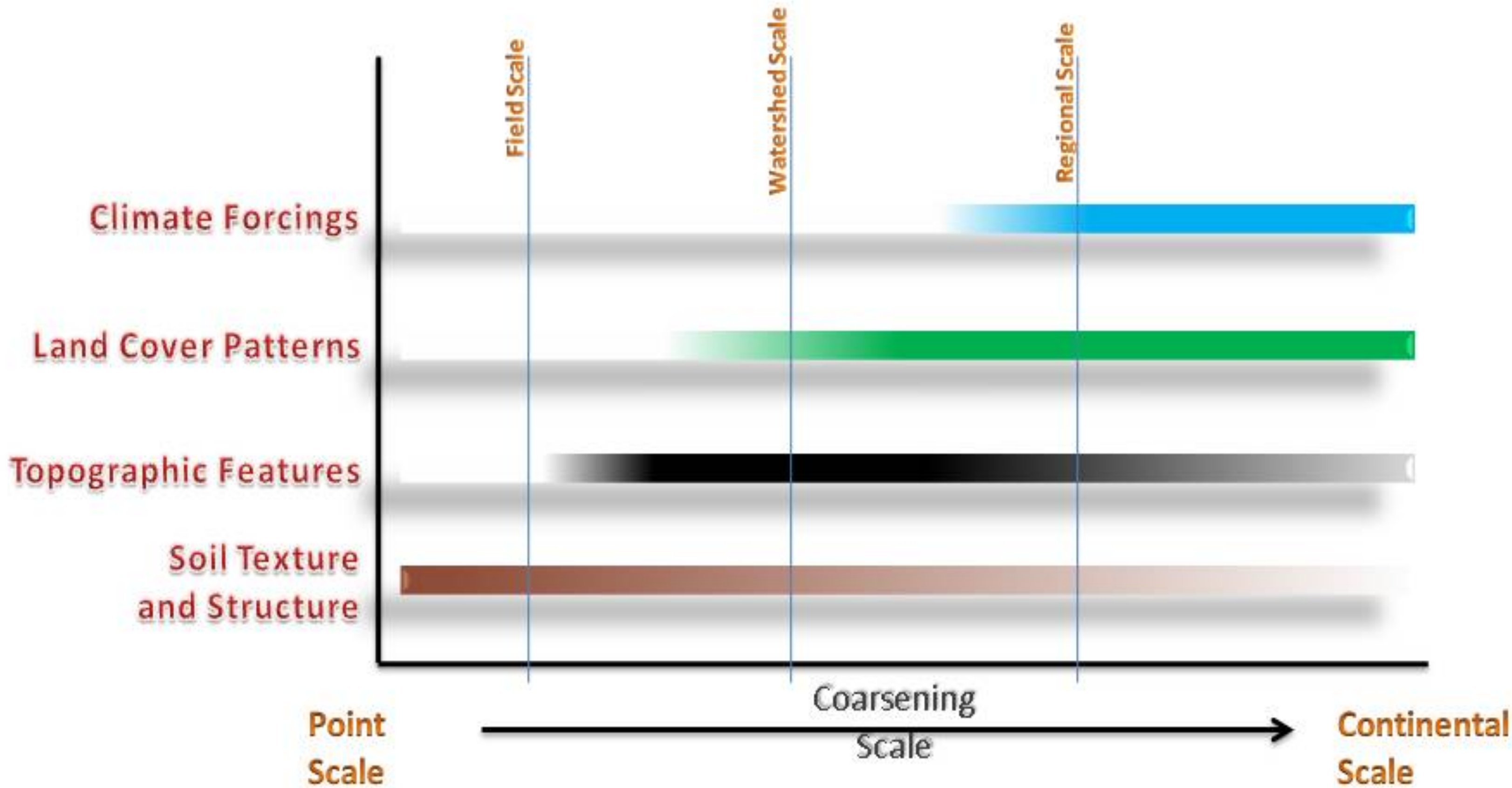
***New measurements strategies (COSMOS, GPS, fiber optic cables) may fundamentally alter the spatial support of ground-based observations.***

***Future plans:***

✓ Finalize white paper by mid-summer (still time for input!).

✓ Opportunities for evaluating upscaling scope/strategies during upcoming field campaigns.

## *Process controls on soil moisture variability*




(Jana and Mohanty 2011)