Estimating uncertainty in remotely sensed soil moisture at continental to global scales

Clara Draper, NASA Global Modeling and Assimilation Office.
with
Rolf Reichle,
Richard de Jeu and Robert Parinussa,
Vahid Naeimi and Wolfgang Wagner

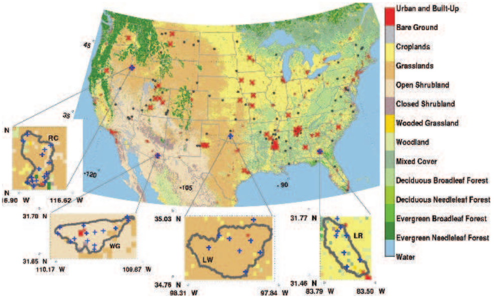
Draper et al, 2013. (RSE)

December 1, 2015

Motivation/Outline (1)

- Aim: estimate large-scale distributed Root Mean Square Errors (RMSE) in remotely sensed near-surface soil moisture
  - For data assimilation observation error specification
  - For general evaluation (comparison to target accuracies for remotely sensed products)

- Difficulty: there is no agreed upon true soil moisture across continental scale domain



SCAN and ARS in situ sensors, from Liu et al (2011)

Motivation/Outline (2)

- Demonstrate two methods for estimating distributed errors:
  - Triple colocation
  - Error propagation through remote sensing retrieval algorithms
- Apply to two remotely sensed soil moisture data sets:
  - AMSR-E X-band Land Parameter Retrieval Model (de Jeu and Owe, 2003)
  - ASCAT C-band TUWien empirical change detection (Wagner et al, 1999)
  - Jan. 2007 - Oct 2011, over a North American domain

Rescaling of continental-scale soil moisture

- At large spatial scales, systematic differences between the mean and variance of different soil moisture estimates are ubiquitous, and the true mean and variance is unknown
- When comparing different soil moisture estimates, for evaluation or data assimilation:
  - It is standard to rescale all data sets to have mean and variance consistent with a chosen reference climatology
  - Focus is then on comparing their temporal changes

Triple colocation ($RMSE^{TC}$)

- For ASCAT ($\theta(A)$), AMSR-E LPRM ($\theta(L)$), and Catchment model near-surface soil moisture ($\theta(C)$) soil moisture anomalies from the mean seasonal cycle
- Assume an additive random error model ($< \epsilon(X) >= 0$):
  $\theta(A) = \alpha(\theta + \epsilon(A))$
  $\theta(L) = \lambda(\theta + \epsilon(L))$
  $\theta(C) = \gamma(\theta + \epsilon(C))$
- $\theta =$ truth, and α, λ, γ are calibration constants
- Assume errors are not correlated with each other, or with the true state

Stoffelen (1998), Scipal et al (2008), Miralles et al (2010)

Triple colocation ($RMSE^{TC}$)

- Cannot solve for all three calibration constants and the errors, set $\theta(A)$ as the reference data set ($\alpha = 1$)
- Solve for remaning calibration constants: $\hat{\lambda}_A$, $\hat{\gamma}_A$
- Solve for $\epsilon^2 = (RMSE^{TC})^2$: $\hat{\epsilon}^2_A(A)$, $\hat{\epsilon}^2_A(L)$, $\hat{\epsilon}^2_A(C)$

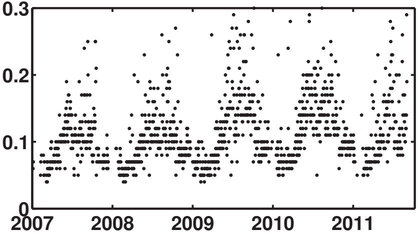
Note:

- Rescaling by variance results in biased $RMSE^{TC}$: must use triple colocation calibration
- Assumptions are violated if apply to original soil moisture time series: must use anomalies

Error propagation ($RMSE^{EP}$)

- Propagate expected errors in input observations and parameters through the soil moisture retrieval model
  - ASCAT (Naeimi et al, 2007)
  - LPRM (Parinussa et al, 2011)

LPRM RMSE^{EP} timeseries at [-103,32.5] ($m^3 m^{-3}$)



- Reported the average of the time series at each location

Comparison of methods

Triple colocation :

- Estimates RMSE of the soil moisture anomalies, relative to the selected reference data set
- Can be used to compare errors in different data sets

Error Propagation:

- Estimates soil moisture errors due to errors in retrieval model parameters and input only (assumed to be RMSE of soil moisture anomalies)
- Input uncertainties are specified somewhat arbitrarily, hence magnitude of estimated RMSE is not informative
- Designed instead to detect (temporal/spatial) variation in errors

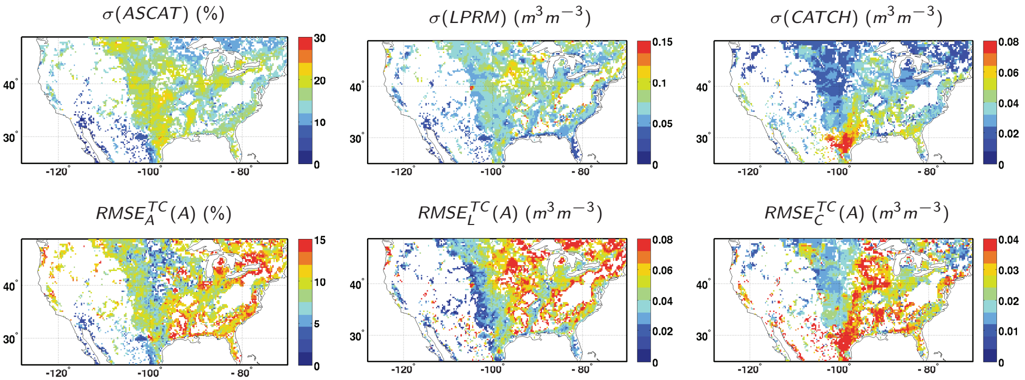
Interpretation of soil moisture RMSE over large domains

Soil moisture RMSE over large domains

- $\sqrt{<(X-Y)^2>} = \sqrt{\sigma(X)^2 + \sigma(Y)^2 - 2R\sigma(X)\sigma(Y) + b^2}$
  Because σ differs from place to place ... a single global RMSE target may not be appropriate." (Entekhabi et al, 2010)

- For error estimates based on rescaled data sets:
  - RMSE depends on the reference σ
  - Information on temporal agreement is from R

Dependence of RMSE on the reference



- Magnitude and spatial variability in RMSE strongly influenced by the reference σ

Fractional RMSE (*fRMSE*)

- Spatial differences in σ are not linear: the ratio (and ranking!) of the domain-average RMSE in $m^3 m^{-3}$ depends on the selected reference

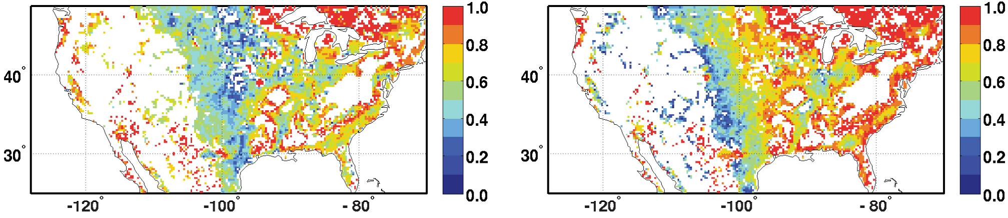
Reference	Mean $RMSE^{TC}$		
	A	L	C
ASCAT (%)	10	18	19
LPRM ($m^3 m^{-3}$)	0.07	0.06	0.09
CATCH ($m^3 m^{-3}$)	0.07	0.08	0.03

- RMSE is not well suited to continental scale soil moisture evaluation
- Here: report errors as $fRMSE = RMSE_X(X)/\sigma(X)$ (unitless, with range [0,1])

Triple Colocation Results

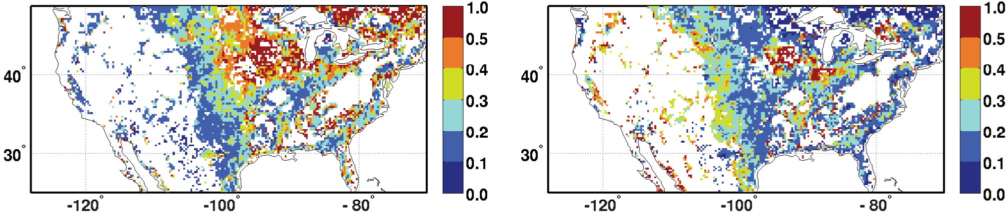
Triple colocation fRMSE (fRMSE^{TC})

ASCAT $fRMSE^{TC}$ LPRM $fRMSE^{TC}$

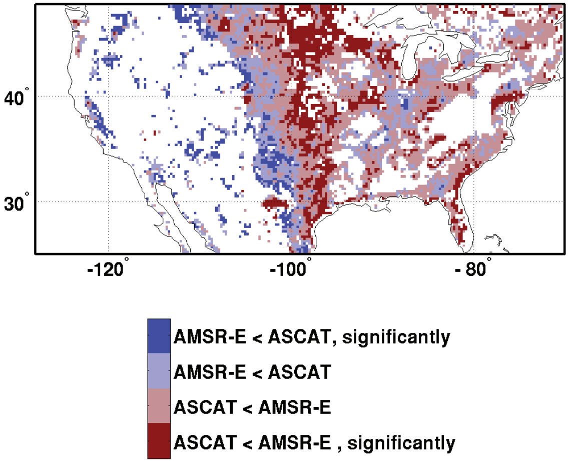


- fRMSE^{TC} uncertainty (boot-strapped)

ASCAT $fRMSE^{TC}$ 90% C.I. (full width) LPRM $fRMSE^{TC}$ 90% C.I. (full width)



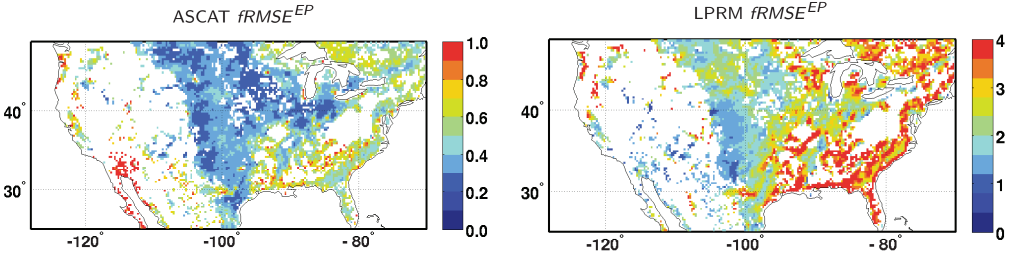
ASCAT and LPRM AMSR-E fRMSE^{TC} comparison



See also Crow and Zhan (2007), Scipal et al (2008), Dorigo et al (2010)

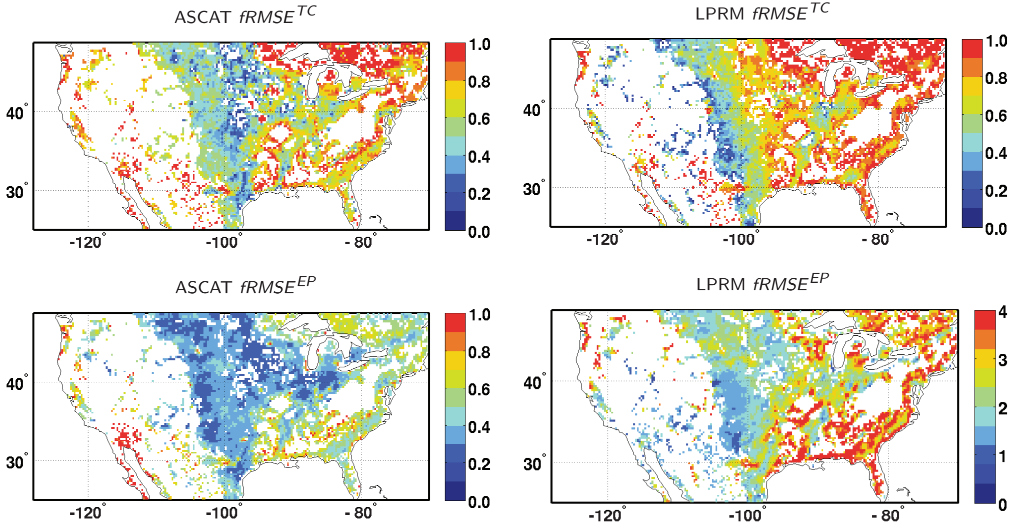
Error propagation results

Error propagation fRMSE (fRMSE^{EP})

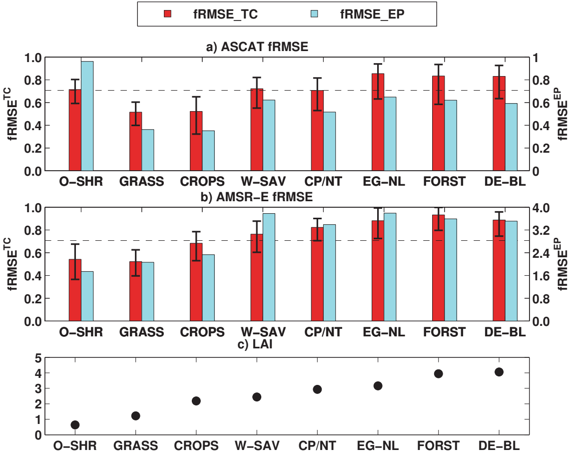


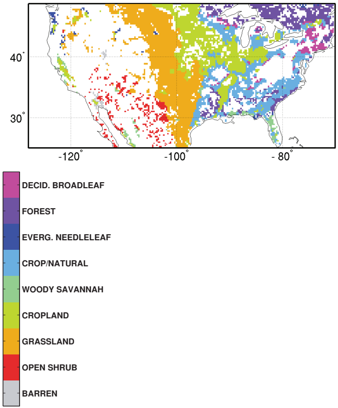
- Note: LPRM *fRMSE^{EP}* unphysically large

fRMSE^{TC} and fRMSE^{EP} maps



fRMSE^{TC} and fRMSE^{EP} by land cover





- ASCAT and AMSR-E fRMSE appear similar, although AMSR-E is more dependent on vegetation cover

Conclusions (1/3)

How should we evaluate remotely sensed soil moisture globally?

- Substantial spatial variation in RMSE and fRMSE across North America
  - Evaluation based on handful of pixels may not represent global accuracy
  - Triple colocation or error propagation both provide useful uncertainty estimates that can complement comparison to in situ observations
- RMSE not a good choice of metric for large-scale evaluation
  - Uniform RMSE target over large domain not sensible
  - Selection of reference determines RMSE magnitude (domain-average RMSE has non-linear dependence on reference)
- Alternatives: fRMSE, signal/noise, correlation

Conclusions (2/3)

How should we specify soil moisture observation errors for data assimilation?

- Specification of spatially uniform RMSE not sensible
- Use a uniform fRMSE, or even better spatial maps from triple colocation and error propagation (latter gives temporal variability)

Conclusions (3/3)

Where should we focus future research on evaluating remotely sensed soil moisture?

- Given the available data, we can estimate errors in temporal behavior with reasonable confidence
  'No amount of statistical manipulation can add any new physical information on top of what already exists in the data products in use. A "real" better error assessment can only happen when new, better and independent data sources become available'. Pan et al, 2015.
- Other types of errors need more attention
  - Systematic errors (bias in mean, variance)
  - Errors in spatial patterns
- Data assimilation would benefit from realistic spatial error covariances