

## Relevance of time-varying and time-invariant retrieval error sources on the utility of spaceborne soil moisture products

Wade T. Crow

Hydrology and Remote Sensing Laboratory, USDA/ARS, Beltsville, Maryland, USA

Randal D. Koster and Rolf H. Reichle<sup>1</sup>

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Hatim O. Sharif

Department of Civil and Environmental Engineering, University of Texas at San Antonio, San Antonio, Texas, USA

Received 11 October 2005; revised 18 November 2005; accepted 25 November 2005; published 31 December 2005.

[1] Errors in remotely-sensed soil moisture retrievals originate from a combination of time-invariant and time-varying sources. For land modeling applications such as forecast initialization, some of the impact of time-invariant sources can be removed given known differences between observed and modeled soil moisture climatologies. Nevertheless, the distinction is seldom made when evaluating remotely-sensed soil moisture products. Here we describe an Observing System Simulation Experiment (OSSE) for radiometer-only soil moisture products derived from the NASA Hydrosphere States (Hydros) mission where the impact of time-invariant errors is explicitly removed via the linear rescaling of retrievals. OSSE results for the 575,000 km<sup>2</sup> Red-Arkansas River Basin indicate that climatological rescaling may significantly reduce the perceived magnitude of Hydros soil moisture retrieval errors and expands the geographic areas over which retrievals demonstrate value for land surface modeling applications. **Citation:** Crow, W. T., R. D. Koster, R. H. Reichle, and H. O. Sharif (2005), Relevance of time-varying and time-invariant retrieval error sources on the utility of spaceborne soil moisture products, *Geophys. Res. Lett.*, 32, L24405, doi:10.1029/2005GL024889.

### 1. Introduction

[2] Despite the fact that the subsurface moisture prognostic variable in a land surface model is often labeled “soil moisture”, it is more accurately considered a model-dependent index of wetness [Koster and Milly, 1997]. Simply put, measured and modeled soil moistures do not have the same meaning and are not directly interchangeable. The assimilation of a measured soil moisture into a land surface model, for applications such as forecast initialization, requires that the measurement and model data be scaled to a common climatology [Reichle and Koster, 2004].

[3] One approach is to transform both modeled and remotely retrieved soil moisture ( $\theta_{mod}$  and  $\theta_{obs}$ ) into standard normal deviates to ensure they share common first and second moment statistics. If  $\mu_{obs}$  and  $\mu_{mod}$  represent the

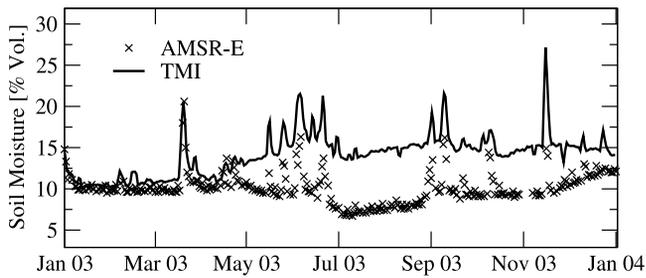
mean soil moisture from observations and the model, respectively, at a given time of year (over many years) and a given location, and if  $\sigma_{obs}$  and  $\sigma_{mod}$  represent the corresponding standard deviations, then the standard normal deviates of both soil moisture variables ( $\theta'_{mod}$  and  $\theta'_{obs}$ ) can be calculated following

$$\theta' = (\theta - \mu) / \sigma. \quad (1)$$

[4] Transformation to consistent normal deviates effectively removes the impact of time-invariant additive and multiplicative errors from comparisons between modeled and observed soil moisture and provides an objective basis for the intercomparison of anomalies. Such transformations have profound implications for the error estimates assigned to remotely-sensed soil moisture. Consider the following extreme example. A satellite retrieves volumetric soil moisture with an absolute error of 4% volumetric. The error takes the form of a constant bias resulting from the incorrect assignment of some time-invariant property such as soil type. In this example, the time series of measurement normal deviate anomalies from the sensor is identical to the time series of anomalies that would be derived from an unbiased measurement system. Furthermore, because the scaling process in (1) focuses on anomalies alone, the impact of the bias on the scaled observations is completely removed. For model applications, the biased time series of soil moisture measurements can be considered perfect after scaling. Of course, errors in sensor measurements have both time-invariant and time-varying sources. Still, if the time-invariant sources are a significant fraction of the total, the effective error of a soil moisture measurement, from a modeling standpoint, may be significantly less than the stated absolute error. Despite this fact, validation goals for spaceborne soil moisture products are typically given in terms in absolute (or unscaled) root-mean-square (RMS) error.

[5] Preliminary results with remote sensing products imply that the impact of time-invariant errors may be large. Figure 1 plots independent surface soil moisture retrievals obtained from the Tropical Rainfall Mission Microwave Imager (TMI) [Bindlish et al., 2003] and Advanced Microwave Scanning Radiometer (ASMR-E) [Njoku et al., 2003] for a 1° lat/long box in the U.S. Southern Great Plains (centered at -102.5°W and 35.5°N) during the 2003 calendar year. A seasonally-varying bias results in TMI retrievals being generally wetter than comparable AMSR-E retrievals.

<sup>1</sup>Also at Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, Maryland, USA.



**Figure 1.** Time series of 2003 AMSR-E and TMI surface soil moisture retrievals for the  $1^\circ$  grid square centered at  $-102.5^\circ\text{W}$  and  $35.5^\circ\text{N}$ .

The presence of this bias implies that each product possesses a distinct climatology - with respect to each other as well as to any modeled or in situ soil moisture product [Drusch *et al.*, 2005]. These differences will inflate remote sensing RMS error statistics without degrading the real value of retrievals for modeling applications. Of course, the single year plotted in Figure 1 is insufficient to determine whether the biases seen in Figure 1 are stable on an inter-annual basis and thus constitute a true climatological difference. Given temporal shortcomings in both remotely-sensed and ground-based soil moisture data sets, Observing System Simulation Experiments (OSSE's) provide the only currently feasible approach for studying this issue. Here we present a 20-year OSSE executed to examine the relative contribution of time-varying and time-invariant error sources to the overall uncertainty of radiometer-only soil moisture retrievals from the planned NASA Hydrosphere States (Hydros) mission [Entekhabi *et al.*, 2004]. Implications of this decomposition on the perceived value of Hydros soil moisture products for land surface modeling applications are examined.

## 2. OSSE Approach

[6] The observing system simulation experiment (OSSE) described here is based on a methodology previously developed by Crow *et al.* [2001, 2005]. It has four distinct components: (1) a land surface model to predict surface geophysical states, (2) a forward microwave emission model to convert these states into microwave brightness temperature ( $T_B$ ), (3) an orbit and sensor model to realistically degrade  $T_B$  fields, and (4) a retrieval model to invert simulated  $T_B$  observations back into soil moisture.

### 2.1. Land Surface Modeling

[7] The 1-km surface (0–5 cm) soil moisture ( $\theta$ ), 5-cm soil temperature ( $T_{5\text{ cm}}$ ), and surface skin temperature fields ( $T_0$ ) underlying the OSSE are derived from long-term (October 1980 to July 2000) TOPmodel Land Atmosphere Transfer Scheme (TOPLATS) [Peters-Lidard *et al.*, 1997] simulations over the 575,000 km<sup>2</sup> Red-Arkansas River Basin (H. O. Sharif *et al.*, Multi-decadal high-resolution hydrologic modeling of the Arkansas/Red River Basin, submitted to Journal of Hydrometeorology, 2005).

### 2.2. Forward Microwave Emission Modeling

[8] Following Crow *et al.* [2005], model-generated  $\theta$ ,  $T_{5\text{ cm}}$  and  $T_0$  are combined with ancillary data to simulate

1-km H-polarization microwave brightness temperature ( $T_B$ ) via:

$$T_B = T_s(1 - r) \exp(-\tau / \cos(\phi)) + T_0(1 - \omega) \cdot [1 - \exp\{-\tau / \cos(\phi)\}] + r \exp\{-\tau / \cos(\phi)\} \quad (2)$$

where  $T_s$  is the effective soil temperature defined as  $(T_0 + T_{5\text{ cm}})/2$ ,  $\tau$  is the vegetation nadir opacity,  $\phi$  the incidence angle,  $\omega$  the single scattering albedo, and  $r$  the soil reflectivity. In turn, vegetation nadir opacity is defined as

$$\tau = bW. \quad (3)$$

The coefficient  $b$  varies with vegetation type and  $W$  is the total columnar vegetation water content. Foliar vegetation water content is estimated for the Red-Arkansas River basin using archived (1980 to 2000) 8-km AVHRR NDVI products and the regression relationship of Jackson *et al.* [1999]. A woody-vegetation correction factor is applied to convert the foliar  $W$  value into the total columnar (i.e. trunk, branches, and foliar) value required by (3).

[9] Soil reflectivity ( $r$ ) is calculated as

$$r = r_s \exp(-h) \quad (4)$$

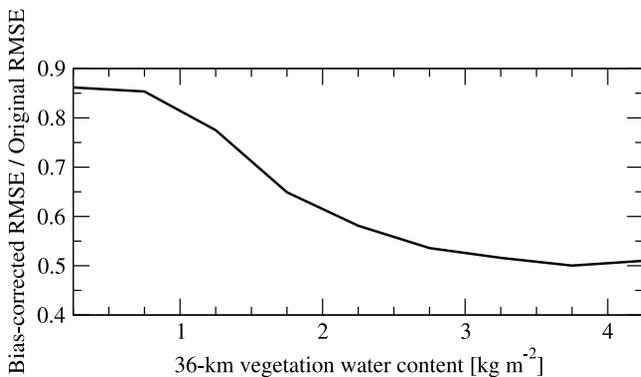
where  $h$  is assumed to be 0.1 of the surface RMS height and  $r_s$  is the Fresnel reflectivity of the equivalent smooth surface. This reflectivity depends on  $\phi$  and the dielectric constant of the soil. Dielectric information can be derived from soil moisture using the mixing model of Dobson *et al.* [1985] and known soil sand ( $S$ ) and clay ( $C$ ) percentages. For water surfaces, the forward calculation is

$$T_B = T_0(1 - r_s), \quad (5)$$

and  $r_s$  is calculated via Klein and Swift [1977]. Vegetation parameters  $b$ ,  $\omega$  and  $h$  are assigned using a 1-km land cover classification and a lookup-table populated with typical literature values [Crow *et al.*, 2005]. Soil  $S$  and  $C$  are based on a soil classification map.

### 2.3. Orbital and Scanning Model

[10] A simplified orbital simulation model of the Hydros platform and antennae [Crow *et al.*, 2001] provides acquisition times and spatial locations for all footprint centers that fall within the Red-Arkansas River Basin during an estimated two-year Hydros mission life. Hydros orbit and scan characteristics are based on preliminary mission design parameters presented by Entekhabi *et al.* [2004]. For each individual footprint over the forward half of the radiometer's scan, a two-dimensional Gauss function is used to obtain a weighted averaged of the  $T_B$  field and approximate the spatial support of the actual Hydros antennae gain function [Drusch *et al.*, 1999]. Footprint  $T_B$  antennae noise is randomly sampled from a mean-zero Gaussian distribution with a standard deviation of 1 K. Afterwards, perturbed  $T_B$  observations are binned onto a 36-km fixed earth grid. Observations from both the ascending and descending orbital scans are used and multiple footprint centers within a single bin are merged via simple averaging. Footprint-centers calculated during a



**Figure 2.** The reduction in RMS error (RMSE) associated with bias correction of OSSE results.

two-year Hydros mission life are recycled ten times in order to cover the approximately 20-year period over which land surface model output is available.

#### 2.4. Soil Moisture Retrieval Model

[11] Coarse-scale ancillary values of  $b$ ,  $h$ ,  $T_s$ ,  $W$ ,  $\omega$ ,  $S$ , and  $C$  are obtained by aggregating the 1-km fields used in the forward-modeling component of the OSSE to 36-km. After aggregation, synthetic noise is added to 36-km  $T_s$ ,  $W$  and  $b$  fields to mimic actual operational conditions. 36-km  $T_s$  and  $W$  noise is sampled from a mean-zero Gaussian distribution with standard deviations of 1.5 K and  $0.3 \text{ kg m}^{-2}$ , respectively. All  $W$  and  $T_s$  noise is modeled as both spatially and temporally independent. In addition, some misclassification of surface vegetation is expected. To represent this impact,  $b$  noise is sampled from a mean-zero Gaussian distribution with a standard deviation of 0.2. Because this error arises from uncertainty in a static classification and/or look-up table, it is assumed to be time-invariant. Using these ancillary data fields, the simulated 36-km  $T_B$  fields described in Section 2.3 are converted then into 36-km soil moisture products using the *Jackson* [1993] algorithm. The approach neglects differences between soil and canopy temperatures by assuming  $T_0 = T_s$ . Given known 36-km ancillary values of  $h$ ,  $\omega$ ,  $T_s$ ,  $b$ ,  $W$ ,  $\phi$ ,  $S$  and  $C$ , this allows (2) to be solved for  $r_s$  which, in turn, is converted into a soil moisture estimate using the Fresnel equations and the soil-mixing model of *Dobson et al.* [1985].

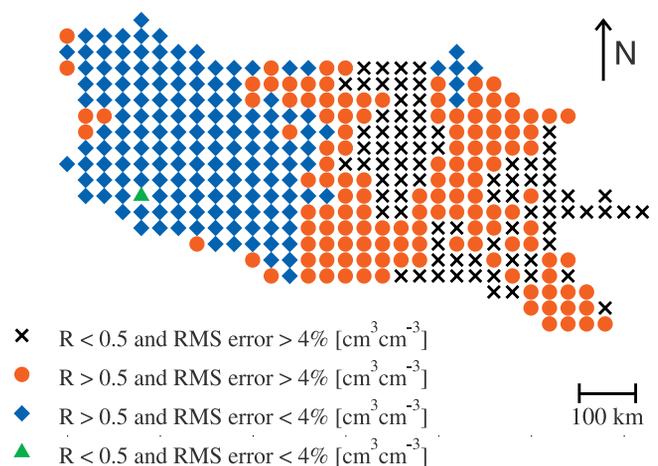
### 3. Results

[12] Using the approach described above, 20-years of OSSE-simulated Hydros soil moisture products are produced at a footprint-scale of 36-km. The fine-scale (1-km) TOPLATS soil moisture fields underlying the OSSE (see section 2.1), determined through the forcing of the land model with observational data, are directly aggregated to 36-km and are treated as benchmark “truth”. For each 36-km grid cell in the modeling domain and each day of the year, we then calculate the mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of the retrieved ( $\theta_{obs}$ , simulated via the OSSE) and the benchmark ( $\theta_{mod}$ , generated via the land surface model) soil moisture products. In order to mini-

mize uncertainty in OSSE-generated statistics due to random effects, all results are based on the pooling of five separate OSSE realizations perturbed using independent synthetic noise.

[13] Differences between the benchmark and OSSE-simulated Hydros soil moisture products are due to a discrete number of error sources: simplifying assumptions employed in the retrieval model, noise added to 36-km  $b$ ,  $W$  and  $T_s$  fields, noise added to  $T_B$  retrievals, gridding interpolation errors, and aggregation errors reflecting the neglect of sub-footprint-scale heterogeneity in the retrieval process. A portion of these errors manifest themselves as a bias in simulated soil moisture retrievals. The 20-year length of the OSSE allows  $\mu_{mod}$  and  $\mu_{obs}$  to be calculated for each day of the year and, subsequently, the correction of climatological biases on a daily time step. Figure 2 demonstrates the impact of this correction by plotting the ratio between the RMS error of the original OSSE retrievals ( $\theta_{obs}$ ) and the RMS error calculated after OSSE results have been corrected for the impact of time-invariant biases ( $\theta_{obs} + \mu_{mod} - \mu_{obs}$ ). Especially over heavily vegetated areas, the correction of long-term biases leads to a substantial reduction in RMS retrieval error.

[14] If a sufficient heritage of observations exists to calculate climatological statistics, time invariant errors in spaceborne soil moisture products can be removed by scaling raw retrievals into appropriate standard normal deviates using (1). For modeling applications, the truest reflection of retrieval value is how well these deviates correlate with actual soil moisture anomalies. Using the 20-year history of the OSSE, we calculated climatological soil moisture statistics and rescaled both benchmark and OSSE-retrieved soil moisture products into standard normal deviates using (1). Unscaled RMS error results and the temporal correlation between rescaled OSSE and benchmark soil moisture fields were then examined on



**Figure 3.** Fraction of the 575,000 km<sup>2</sup> Red-Arkansas River Basin where OSSE-simulated soil moisture retrievals satisfy RMS and R accuracy thresholds. RMS error is calculated between unscaled OSSE-simulated and benchmark soil moisture products. R is calculated between standard normal deviates of both soil moisture products.

a grid-cell by grid-cell basis within the Red-Arkansas River Basin. Blue diamonds (and the single green triangle) in Figure 3 indicate lightly vegetated grid cells where the RMS error of unscaled retrievals is below 4% volumetric (the official absolute accuracy goal for Hydros retrievals). Red circles in Figure 3 denote the fraction of the basin where this RMS threshold is not met but the correlation coefficient ( $R$ ) between the benchmark and OSSE normal deviates remains above 0.5. Past work with real spaceborne soil moisture retrievals has indicated that an anomaly correlation coefficient of 0.32 (against in situ data) leads to improved soil moisture estimates after assimilation of the retrievals into a global land surface model [Reichle and Koster, 2005]. A slightly higher correlation coefficient may be required to add value when focusing on regions where higher quality forcing data is available. Therefore, as a conservative estimate, we assume that a correlation coefficient of 0.5 represents the minimum required to add value to global model predictions. Based on this threshold, a substantial fraction of the total basin area failing the RMS threshold (red circles and black crosses) retains value for land surface modeling by virtue of a sufficiently large anomaly  $R$  (red circles only). In these areas, the value of Hydros retrievals is underestimated by the impact of time-invariant errors on RMS statistics.

[15] Red circles in Figure 3 cover approximately 36% of the Red-Arkansas River Basin and denote areas where Hydros observations retain the ability to detect soil moisture anomalies despite failing to meet a 4% volumetric RMS error threshold. The spatial extent of these areas exhibits some sensitivity to assumptions concerning the type of error sources represented by the OSSE. For instance, adding a time invariant RMS error of 0.1 to the sand ( $S$ ) and clay ( $C$ ) fractions used in the retrieval model will increase the extent of the red circles from 36% to 42% of the basin. Conversely, modeling all error sources (i.e.  $b$ ,  $W$ ,  $S$  and  $C$ ) as completely uncorrelated in time reduces the spatial fraction of red circles in Figure 3 from 36% to 30%. It is worth noting that even in this final case, where decidedly pessimistic assumptions are made concerning the temporal properties of retrieval errors, significant levels of anomaly correlation are retained in some areas of the Red-Arkansas Basin otherwise failing the 4% volumetric RMS error threshold. Despite temporally uncorrelated parameter perturbations, OSSE-simulated retrieval errors retain some temporal invariance due to land surface aggregation efforts which tend to manifest themselves as temporally persistent biases [Crow et al., 2001].

#### 4. Conclusions

[16] In this study we assess the relative contribution of time-invariant versus time-varying sources of uncertainty on soil moisture retrieval errors simulated within an Observing System Simulation Experiment (OSSE) for Hydros radiometer-only soil moisture products. While both types of error contribute directly to RMS retrieval errors, time-invariant errors can be removed given known differences between observed and modeled soil moisture climatologies and will not impact the value of Hydros retrievals for most land surface modeling applications.

[17] Results demonstrate that time-invariant biases do contribute significantly to absolute errors in OSSE-simulated Hydros retrievals (Figure 2). Consequently, retrievals over large areas of the Red-Arkansas River basin that nominally fail the Hydros RMS error goal of 4% volumetric (due to heavy vegetation) retain both an ability to detect the presence of soil moisture anomalies and their value for land surface modeling (Figure 3). Traditional remote sensing validation relies heavily on unscaled absolute RMS errors to assess retrieval value. The analysis suggests that this reliance may lead to the underestimation of the spatial extent over which Hydros soil moisture retrievals possess value for land surface model applications.

[18] Two caveats are worth noting. While our analysis assumes that climatological soil moisture statistics are accurately known, such statistics are seemingly difficult to calculate for sensors with 1- to 3-year mission lives. However, it appears possible to approximate long-term soil moisture statistics from shorter time periods by using spatial coverage as a proxy for temporal averaging [Reichle and Koster, 2004]. In addition, while the spatial and temporal dynamics of soil moisture fields and spaceborne retrieval errors has been modeled to our best knowledge, OSSE results can only approximate the magnitude and structure of errors in Hydros soil moisture products. More definitive treatment of retrieval errors will only be possible once the Hydros radiometer is in orbit.

[19] **Acknowledgments.** The high-resolution TOPLATS modeling approach used in this analysis was developed under the direction of Eric F. Wood at Princeton University through support of NASA grants NAG8-1517 and NAG5-6494 and NOAA grant NA96GP0413. Rolf Reichle's contribution to this paper was supported under NASA grant NNG05GB61G.

#### References

- Bindlish, R., T. J. Jackson, E. F. Wood, H. Gao, P. Starks, D. Bosch, and V. Lakshmi (2003), Soil moisture estimates from TRMM microwave imager observations over the southern United States, *Remote Sens. Environ.*, *85*, 507–515.
- Crow, W. T., M. Drusch, and E. F. Wood (2001), An observation system simulation experiment for the impact of land surface heterogeneity on AMSR-E soil moisture retrievals, *IEEE Trans. Geosci. Remote Sens.*, *39*, 1622–1632.
- Crow, W. T., T. Chen, D. Entekhabi, P. Houser, A. Hsu, T. Jackson, E. Njoku, P. O'Neill, J. Shi, and X. Zhan (2005), An observing system simulation experiment for Hydros radiometer-only soil moisture products, *IEEE Trans. Geosci. Remote Sens.*, *43*, 1289–1303.
- Dobson, M. C., F. T. Ulaby, M. T. Hallikainen, and M. A. El-Rayes (1985), Microwave dielectric behavior of wet soil. Part II: Dielectric mixing models, *IEEE Trans. Geosci. Remote Sens.*, *23*, 35–46.
- Drusch, M., E. F. Wood, and R. Lindau (1999), The impact of the SSM/I antennae gain function on land surface parameter retrieval, *Geophys. Res. Lett.*, *26*(23), 3481–3484.
- Drusch, M., E. F. Wood, and H. Gao (2005), Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture, *Geophys. Res. Lett.*, *32*, L15403, doi:10.1029/2005GL023623.
- Entekhabi, D., et al. (2004), The Hydrosphere State (Hydros) mission concept: An Earth system pathfinder for global mapping of soil moisture and land freeze/thaw, *IEEE Trans. Geosci. Remote Sens.*, *42*, 2184–2195.
- Klein, L. A., and C. T. Swift (1977), An improved model for the dielectric constant of sea water, *IEEE Trans. Antennas Propag.*, *25*, 104–111.
- Koster, R. D., and P. C. D. Milly (1997), The interplay between transpiration and runoff formulations in land schemes used with atmospheric models, *J. Clim.*, *10*, 1578–1591.
- Jackson, T. J. (1993), Measuring surface soil moisture using passive microwave remote sensing, *Hydrol. Processes*, *7*, 139–152.
- Jackson, T. J., D. M. Le Vine, A. Y. Hsu, A. Oldak, P. J. Starks, C. T. Swift, J. Isham, and M. Haken (1999), Soil moisture mapping at regional scales

- using microwave radiometry: The Southern Great Plains Hydrology Experiment, *IEEE Trans. Geosci. Remote Sens.*, 37, 2136–2151.
- Njoku, E. G., T. L. Jackson, V. Lakshmi, T. Chan, and S. V. Nghiem (2003), Soil moisture retrieval from AMSR-E, *IEEE Trans. Geosci. Remote Sens.*, 41, 215–229.
- Peters-Lidard, C. D., M. S. Zion, and E. F. Wood (1997), A soil-vegetation-atmosphere transfer scheme for modeling spatially variable water and energy balance processes, *J. Geophys. Res.*, 102, 4303–4324.
- Reichle, R. H., and R. D. Koster (2004), Bias reduction in short records of satellite soil moisture, *Geophys. Res. Lett.*, 31, L19501, doi:10.1029/2004GL020938.
- Reichle, R. H., and R. D. Koster (2005), Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model, *Geophys. Res. Lett.*, 32, L02404, doi:10.1029/2004GL021700.
- 
- W. T. Crow, Hydrology and Remote Sensing Laboratory, U.S. Department of Agricultural Research Service, Rm. 104, Bldg. 007, BARC-W, Beltsville, MD 20705, USA. (wcrow@hydrolab.arsusda.gov)
- R. D. Koster, and R. H. Reichle, Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.
- H. O. Sharif, Department of Civil Engineering, University of Texas at San Antonio, San Antonio, TX 78249, USA.