AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2015GL065929

Key Points:
- Random error in measurements bias estimates of soil moisture/surface evaporation coupling
- Triple collocation provides a statistical tool to correct such estimates
- Land surface models do not overestimate the strength of this coupling

Correspondence to:
W. T. Crow,
Wade.Crow@ars.usda.gov

Citation:
Crow, W. T., F. Lei, C. Hain, M. C. Anderson, R. L. Scott, D. Billesbach, and T. Arkebauer (2015), Robust estimates of soil moisture and latent heat flux coupling strength obtained from triple collocation, *Geophys. Res. Lett.*, *42*, 8415–8423, doi:10.1002/2015GL065929.

Received 25 AUG 2015
Accepted 24 SEP 2015
Accepted article online 29 SEP 2015
Published online 23 OCT 2015

Robust estimates of soil moisture and latent heat flux coupling strength obtained from triple collocation

Wade T. Crow[1], Fangni Lei[1,2], Christopher Hain[3], Martha C. Anderson[1], Russell L. Scott[4], David Billesbach[5], and Timothy Arkebauer[6]

[1]Hydrology and Remote Sensing Laboratory, USDA ARS, Beltsville, Maryland, USA, [2]School of Resource and Environmental Sciences, Wuhan University, Wuhan, China, [3]ESSIC/NOAA-NESDIS, College Park, Maryland, USA, [4]Southwest Watershed Research Center, USDA ARS, Tucson, Arizona, USA, [5]Department of Biological Systems Engineering, University of Nebraska–Lincoln, Lincoln, Nebraska, USA, [6]Department of Agronomy and Horticulture, University of Nebraska–Lincoln, Lincoln, Nebraska, USA

Abstract Land surface models (LSMs) are often applied to predict the one-way coupling strength between surface soil moisture (SM) and latent heat (LH) flux. However, the ability of LSMs to accurately represent such coupling has not been adequately established. Likewise, the estimation of SM/LH coupling strength using ground-based observational data is potentially compromised by the impact of independent SM and LH measurements errors. Here we apply a new statistical technique to acquire estimates of one-way SM/LH coupling strength which are nonbiased in the presence of random error using a triple collocation approach based on leveraging the simultaneous availability of independent SM and LH estimates acquired from (1) LSMs, (2) satellite remote sensing, and (3) ground-based observations. Results suggest that LSMs do not generally overestimate the strength of one-way surface SM/LH coupling.

1. Introduction

Land surface models (LSMs) play an important role in diagnosing both the strength and downstream impact of water and energy feedbacks operating within the soil, vegetation, and atmospheric interface along the Earth's surface [e.g., *Van den Hurk et al.*, 2011; *Seneviratne et al.*, 2013]. A critical element of these feedbacks is the one-way coupling between soil moisture (SM) availability and surface latent heat flux (LH) [*Guo et al.*, 2006]. Such coupling is commonly quantified via the temporal correlation coefficient between collocated SM and LH estimates [*Dirmeyer et al.*, 2009]. Unfortunately, the ability of LSMs to accurately reproduce SM/LH one-way coupling strengths has not been adequately verified [*Dirmeyer et al.*, 2006].

The recent maturation of long-term, ground-based SM and LH data sets provides a tool for evaluating the accuracy of LSM coupling predictions. However, given that ground-based observations of SM and LH are known to be degraded by significant levels of random error [*Robinson et al.*, 2008; *Richardson et al.*, 2006], the possibility remains that one-way SM/LH coupling strengths estimated via comparison of independent ground-based SM and LH observations are biased low due to random measurement errors. While comparable errors surely exist in LSM-derived SM and LH predictions, they will tend to be cross correlated and therefore exert less of a degrading effect on estimated SM/LH coupling strength. Consequently, the question arises whether apparent differences between LSM-based and ground-based SM/LH coupling strengths reflect systematic errors in LSMs or the spurious impact of independent random errors in ground-based SM and LH measurements.

Recent progress on the application of triple collocation (TC) sampling strategies to land surface data sets offers an approach for addressing this question. In particular, TC provides a tool for obtaining correlation estimates which are unbiased by the presence of random error [*Draper et al.*, 2013; *McColl et al.*, 2014]. Here we apply a new variant of TC to acquire unbiased estimates of the one-way coupling strength between SM and LH based on a triplet of SM and LH estimates acquired from ground measurements, remote sensing, and LSMs.

2. Triple Collocation Approach

As noted above, our approach is based on leveraging the simultaneous availability of three SM and three LH flux estimates acquired from: ground-based (G) measurements, remote sensing (RS) retrievals, and

land surface model (LSM) predictions. To start, assume that they all can be linearly related to an unknown truth via

$$\begin{aligned} \mathrm{SM_G} &= \alpha_\mathrm{G}\mathrm{SM_{True}} + \mu_\mathrm{G} + \varepsilon_\mathrm{G} \\ \mathrm{SM_{RS}} &= \alpha_\mathrm{RS}\mathrm{SM_{True}} + \mu_\mathrm{RS} + \varepsilon_\mathrm{RS} \\ \mathrm{SM_{LSM}} &= \alpha_\mathrm{LSM}\mathrm{SM_{True}} + \mu_\mathrm{LSM} + \varepsilon_\mathrm{LSM} \\ \mathrm{LH_G} &= \beta_\mathrm{G}\mathrm{LH_{True}} + \delta_\mathrm{G} + \sigma_\mathrm{G} \\ \mathrm{LH_{RS}} &= \beta_\mathrm{RS}\mathrm{LH_{True}} + \delta_\mathrm{RS} + \sigma_\mathrm{RS} \\ \mathrm{LH_{LSM}} &= \beta_\mathrm{LSM}\mathrm{LH_{True}} + \delta_\mathrm{LSM} + \sigma_\mathrm{LSM} \end{aligned} \tag{1}$$

where μ and δ are temporally constant biases, α and β are temporally constant gains, and ε and σ are mean zero, random variables representing estimation errors in each product. These random errors are assumed to be adequately described by a temporally constant variance. All "true" variables are assumed to have a spatial support equal to a common coarse-scale grid on which both RS and LSM predictions are provided. Therefore, the ground-based error terms ε_G and σ_G also reflect upscaling errors associated with the use of a local-scale observation to characterize a (much coarser) grid-scale average.

In addition, we assume both the mutual independence of random errors among all three SM products and all three LH errors:

$$\begin{aligned} E\left[\varepsilon_i\varepsilon_j\right] &= 0 \; i \neq j \\ E[\sigma_l\sigma_m] &= 0 \; l \neq m, \end{aligned} \tag{2}$$

and the orthogonality of all errors with respect to the truth:

$$\begin{aligned} E[\mathrm{SM_{True}}\varepsilon_i] &= 0 \\ E[\mathrm{LH_{True}}\sigma_l] &= 0 \; . \end{aligned} \tag{3}$$

Note that (2) only asserts the mutual independence of SM and/or LH errors and does not address the possibility of cross correlation between SM and LH errors. For ease of notation, the general index triplets *i*, *j*, and *k* (for SM) and *l*, *m*, *n* (for LH) will be used to reflect the source of an arbitrary SM or LH product (i.e., "G," "RS," or "LSM").

If assumptions in (1)–(3) hold, and an arbitrary combination of SM product *i* and LH product *l* are defined as a reference, the system of equations in (1) can be cross multiplied, averaged, and solved to express the temporal variances of $\mathrm{SM_{True}}$ and $\mathrm{LH_{True}}$ as

$$\begin{aligned} \mathrm{Var}[\mathrm{SM_{True}}]_i &= \alpha_i^{-2}\frac{\mathrm{Cov}\left[\mathrm{SM}_i, \mathrm{SM}_j\right]\mathrm{Cov}[\mathrm{SM}_i, \mathrm{SM}_k]}{\mathrm{Cov}\left[\mathrm{SM}_j, \mathrm{SM}_k\right]} \\ \mathrm{Var}[\mathrm{LH_{True}}]_l &= \beta_l^{-2}\frac{\mathrm{Cov}[\mathrm{LH}_l, \mathrm{LH}_m)\mathrm{Cov}[\mathrm{LH}_l, \mathrm{LH}_n]}{\mathrm{Cov}[\mathrm{LH}_m, \mathrm{LH}_n]} \end{aligned} \tag{4}$$

where $i \neq j \neq k$ and $l \neq m \neq n$ [*McColl et al.*, 2014]. Likewise, for any cross combination of SM and LH products *i* and *l* where $E[\varepsilon_i\sigma_l] = 0$, the covariance of SM_i and LH_l can be written as

$$\mathrm{Cov}[\mathrm{SM}_i, \mathrm{LH}_l] = \alpha_i\beta_l\mathrm{Cov}[\mathrm{SM_{True}}, \mathrm{LH_{True}}]_{i,l}. \tag{5}$$

assuming that ε_i and σ_l are independent of both $\mathrm{SM_{True}}$ and $\mathrm{LH_{True}}$. Therefore, by combining (4) and (5), it is possible to solve for the coefficient of determination (R^2) between $\mathrm{SM_{True}}$ and $\mathrm{LH_{True}}$ as

$$R^2[\mathrm{SM_{True}}, \mathrm{LH_{True}}]_{i,l} \equiv \frac{\mathrm{Cov}[\mathrm{SM_{True}}, \mathrm{LH_{True}}]_{i,l}^2}{\mathrm{Var}[\mathrm{SM_{True}}]_i\mathrm{Var}[\mathrm{LH_{True}}]_l} = \frac{\mathrm{Cov}[\mathrm{SM}_i, \mathrm{LH}_l]^2\mathrm{Cov}\left[\mathrm{SM}_j, \mathrm{SM}_k\right]\mathrm{Cov}[\mathrm{LH}_m, \mathrm{LH}_n]}{\mathrm{Cov}\left[\mathrm{SM}_i, \mathrm{SM}_j\right]\mathrm{Cov}[\mathrm{SM}_i, \mathrm{SM}_k]\mathrm{Cov}[\mathrm{LH}_l, \mathrm{LH}_n]\mathrm{Cov}[\mathrm{LH}_l, \mathrm{LH}_m]} \tag{6}$$

Unlike SM/LH coefficients of determination sampled directly from any two SM and LH products in (1), (6) is unaffected by independent random errors in the products. In addition, partially redundant estimates of $R^2[\mathrm{SM_{True}}, \mathrm{LH_{True}}]$ can be obtained from (6) by utilizing different combinations of SM and LH products to serve as the reference data sets *i* and *l*. However, as noted above, certain combinations of SM and LH products (e.g., $\mathrm{SM_{LSM}}$ and $\mathrm{LH_{LSM}}$) are expected to possess cross-correlated errors and will therefore represent an unsuitable reference combination.

Table 1. Attributes of Ground Sites Utilized in the Analysis[a]

AmeriFlux Site	Abbreviation	Latitude/Longitude	SM Depth (cm)	*N* (Days)	*T* (Days)
Lucky Hills	WHS	31.744°/−110.052°	5	277	2
Kendall Grasslands	WKG	31.737°/−109.942°	5	494	2
Santa Rita Mesquite	SRM	31.821°/−110.866°	2.5	514	3
Santa Rita Creosote	SRC	31.908°/−110.840°	2.5–5	198	3
Tonzi Ranch	TON	38.432°/−120.966°	0–2.5[b]	523	11
Audubon Grasslands	AUD	31.591°/−110.509°	10	286	5
ARM-CART	ARM	36.606°/−97.489°	10[c]	413	34
Blodgett Forest	BLO	38.895°/−120.633°	10	186	9
Sand Hills Dry Valley	SDH	42.069°/−101.407°	10	262	3
Sand Hills Upland Dune[d]	SUH	42.066°/−101.367°	10	91	3
Duke Open Field	DK1	35.971°/−79.093°	10	304	6
Duke Hardwoods	DK2	35.974°/−79.100°	10	327	6
Duke Pine	DK3	35.978°/−79.094°	0–30	329	50
Fort Peck	FPE	48.308°/−105.102°	10	276	50
Mead Irrigated	NE2	41.164°/−96.470°	10	535	28
Mead Rainfed	NE3	41.180°/−96.440°	10	554	50
Fermi Agricultural	IB1	41.859°/−88.223°	2.5	348	3
Fermi Prairie	IB2	41.841°/−88.241°	2.5	321	3

[a] *N* relates the total number June-July-August days between 2003 and 2011 where daily AmeriFlux LH and SM, ALEXI LH, Noah LH and SM, and AMSR-E SM are all simultaneously available. *T* is the time scale parameter required to optimize the fit between filtered AMSR-E SM retrievals and ground-based SM measurements (see section 4).
[b] Described as "surface" in AmeriFlux L2 documentation.
[c] Was 5 cm prior to 04/13/2005.
[d] Not an AmeriFlux site (operated by the University of Nebraska).

3. Data

Our general approach was based on applying (6) to sites where simultaneous ground-based, RS-based, and LSM-based estimates of both SM and LH can be obtained. All data were resampled to represent daily averages (0 to 24 UTC) within June/July/August (JJA) for 2002 to 2014. This seasonal period was selected since it approximates the period of maximum SM/LH coupling. However, only JJA days in which viable SM and LH estimates are available from all six data sources in (1) were used. All daily RS and LSM estimates were spatially resampled onto a 0.25° grid prior to matching against ground-based observations. Additional data processing details are given below.

3.1. Ground-Based Observations of SM and LH

Ground-based SM and LH observations were based on nongap-filled AmeriFlux Standardized Level 2 data files (http://ameriflux.ornl.gov). Sites were selected based on the availability of long-term, simultaneous SM and LH data sets and the ability of concurrent LSM and RS predictions to adequately match ground observations (see section 4). At each AmeriFlux site, half-hourly LH and surface SM observations were aggregated up to a single daily (0 to 24 UTC) value. No values were calculated for days containing less than 10 half-hourly SM or 36 half-hourly LH observations. Since the depth of "surface" SM observations varied between sites, we used the shallowest available observation depth at each site (generally between 5 and 10 cm—see Table 1).

For more details on specific AmeriFlux sites utilized here, see *Scott* [2010], *Papuga* [2009], *Baldocchi et al.* [2010], *Krishnan et al.* [2012], *Fischer et al.* [2012], *Goldstein et al.* [2000], *Stoy et al.* [2006], *Oishi et al.* [2010], *Meyers* [2009a, 2009b], *Billesbach and Arkebauer* [2012], *Matamala et al.* [2008], and *Suyker and Verma* [2008]. In addition, a single (non-AmeriFlux) University of Nebraska Bowen Ratio flux tower site in the upland dunes portion of Nebraska's Sand Hills Ecosystem was also utilized [*Billesbach and Arkebauer*, 2012]. See Table 1 for a complete list of all sites.

3.2. LSM Estimates of SM and LH

Three separate LSM-based surface (0–10 cm) SM and LH products were obtained by individually averaging hourly Noah, Mosaic, and Variable Infiltration Capacity (VIC) LSM predictions from the North American Land Data Assimilation System-2 (NLDAS-2) [*Xia et al.*, 2012a, 2012b] into daily (0 to 24 UTC) values. All LSM simulations were forced using gauge-based daily precipitation estimates disaggregated into hourly

estimates using rain radar measurements [*Cosgrove et al.*, 2003] and topographically corrected via monthly climatological information acquired from the Parameter-elevation Regressions on Independent Slopes Model [*Daly et al.*, 1994]. Remaining nonprecipitation forcing data for NLDAS-2 simulations were obtained from National Center for Environmental Prediction North American Regional Reanalysis estimates. Noah was used as the baseline LSM, and results from other LSMs were considered to examine sensitivity to variations in LSM type.

3.3. RS Estimates of SM

Two passive and one active microwave remotely sensed surface soil moisture products were utilized to provide three separate remotely sensed estimates of SM. Baseline passive microwave soil moisture retrievals were derived from the dual-polarized C band (6.93 GHz) and X band (10.65 GHz) channels of the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E) aboard the NASA EOS Aqua satellite. The AMSR-E instrument data were acquired from the launch of Aqua in June 2002 until the failure of the instrument in October 2011. C and X band brightness temperatures measured by AMSR-E were processed into surface (1–3 cm depth) soil moisture using the Land Parameter Retrieval Model [*Owe et al.*, 2008; *Parinussa et al.*, 2011]. Based on AMSR-E soil moisture validation results [*Crow et al.*, 2010], only 1:30 A.M. (local time) descending AMSR-E overpasses were utilized in the analysis.

The Soil Moisture and Ocean Salinity (SMOS) satellite mission measures global surface soil moisture at L band (1.4 GHz) with a 3 day revisit at the equator and ascending/descending overpasses at 6:00 A.M./P.M. local solar time [*Kerr et al.*, 2010]. The SMOS mission aims at monitoring surface soil moisture at a depth of about 3 to 5 cm with 30–50 km spatial resolution. Retrievals were derived from the daily SMOS Level 3 surface soil moisture product released by the Centre Aval de Traitement des Données.

The advanced scatterometer (ASCAT) on board the ESA MetOp satellite is a real aperture radar instrument operating at C band (5.255 GHz) with an equatorial local overpass time of 9:30 P.M./A.M. for ascending and descending orbits, respectively. From January 2007 onward, ASCAT backscatter measurements were used to generate surface soil moisture estimates using the TU Wien soil moisture change detection algorithm [*Naeimi et al.*, 2009]. Soil moisture retrieval products from both overpasses were used to construct a combined daily product.

For AMSR-E surface soil moisture products, retrievals are screened out for vegetation optical depths larger than 0.8 [−] or if radio frequency interference (RFI) contamination was noted in both C and X band retrievals [*Owe et al.*, 2001]. The SMOS Level 3 retrievals with Data Quality Index values larger than 0.06 [−] or RFI probability larger than 0.3 [−] were also filtered out [*Al-Yaari et al.*, 2014]. ASCAT soil moisture retrievals were masked for frozen soil conditions.

3.4. RS Estimates of LH

Remotely sensed estimates of LH were obtained from the Atmosphere Land Exchange Inverse (ALEXI) model using thermal-infrared (TIR) remote sensing data without any precipitation input [*Anderson et al.*, 2011]. In order to estimate RS-based LH estimates, ALEXI was run at a native spatial resolution of 4 km over the period of 2003–2012 and forced with: meteorological inputs from the North American Regional Reanalysis [*Mesinger et al.*, 2006], TIR land surface temperature from Geostationary Operational Environmental Satellites, and vegetation cover fraction estimated using leaf area index (LAI) retrievals from the 8 day Terra Moderate Resolution Imaging Spectroradiometer product (MOD15A2). Instantaneous LH retrieved from ALEXI were upscaled to daytime-integrated LH estimates assuming a self-preservation of the ratio of latent heat flux and incoming shortwave radiation (f_{SUN}) during daytime hours [*Cammalleri et al.*, 2014]. Gap filling for cloudy days was based on an interpolation of clear-sky retrievals of f_{SUN} and estimates of the daily integrated incoming shortwave radiation to estimate daily ET. As described above, ALEXI LH estimates were resampled onto a 0.25° grid prior to analysis.

4. Data Preprocessing and Analysis

The application of (6) to the SM and LH data sets described in section 3 requires several additional considerations. First, the vertical measurement depth of RS-based SM retrievals is likely to be shallower than either the LSM or ground-based SM products. This inconsistency in SM measurement depth across products imperils the mutual linearity assumptions implicit in (1). Therefore, as an initial preprocessing step, the exponential

filtering approach of *Albergel et al*. [2008] was applied to all RS soil moisture sets. By removing high-frequency components of the soil moisture time series and lagging its features in time, this approach can be used to effectively increase the vertical depth of superficial soil moisture observations. Here the time scale parameter T required by the filter was assumed to be temporally constant and tuned on a site-by-site basis to maximize the temporal correlation between the filtered RS SM product and ground-based SM observations (see Table 1).

A second concern is the impact of potential nonlinearity in the mutual relationship between SM and LH products obtained from various sources. In order to minimize such effects, all SM and LH products were transformed into temporal ranks after the exponential smoothing of the RS-based SM products and prior to the application of (6). Therefore, $R^2[SM_{True}, LH_{True}]$ estimates obtained from (6) are equivalent to Spearman rank coefficients of determination. This rank transition has the benefit of expressing SM/LH one-way coupling strength in terms of a rank correlation coefficient which will not be spuriously degraded by the presence of nonlinearity in the relationship between SM and LH.

Given the relatively limited number of days in which all SM and LH estimates are simultaneously available from all data sources described in section 3 (Table 1), sampling uncertainty is likely to be large in estimates of $R^2[SM_{True}, LH_{True}]$ obtained from (6). One technique for easing such concerns is leveraging partially redundant estimates of $R^2[SM_{True}, LH_{True}]$ obtained by applying (6) to all possible combinations of SM product i and LH product l with independent errors. For the SM and LH data sets described in section 3, seven different quasi-independent estimates of $R^2[SM_{True}, LH_{True}]$ can be sampled: (1) SM_G/LH_{RS}, (2) SM_G/LH_{LSM}, (3) SM_{RS}/LH_G, (4) SM_{RS}/LH_{RS}, (5) SM_{RS}/LH_{LSM}, (6) SM_{LSM}/LH_G, and (7) SM_{LSM}/LH_{RS}. The only potential combinations not sampled were SM_G/LH_G (due to the potential for cross-correlated spatial representativeness errors in ground-based SM and LH measurements) and SM_{LSM}/LH_{LSM} (due to the potential for cross-correlated SM and LH errors in LSM-based estimates).

All seven of these viable $R^2[SM_{True}, LH_{True}]$ estimates should be unbiased but impacted by (partially independent) sampling noise. In order to leverage this redundancy, they were unified into a single, optimized estimate. Specifically, a 15,000-member boot-strapping analysis was applied to sample the mutual error covariance matrix for each of the seven quasi-independent $R^2[SM_{True}, LH_{True}]$ estimates obtained from (6). To (roughly) compensate for the potential impact of temporal autocorrelation, boot-strapped samples were based on $N/2$ samples acquired with replacement (where N is the total number of days that all SM and LH data sets are simultaneously available). In addition, in order to maximize their normality, all boot-strapped replicates of $R^2[SM_{True}, LH_{True}]$ were transformed via their third power prior to sampling of the covariance matrix.

Following *Perrons and Cooper* [1992], the resulting error covariance matrix was used to derive optimal weights for each of the (now transformed) seven separate estimates of $R^2[SM_{True}, LH_{True}]$. The error covariance matrix was also used to determine 95% confidence bounds for the single (deterministic) maximum likelihood (ML) prediction based on these weights. After all three $R^2[SM_{True}, LH_{True}]$ values (i.e., the ML estimate plus lower and upper 95% confidence bounds) were determined in this transformed space they were each inverted back to their original Spearman rank R^2 space.

5. Results

All presented results are based on daily JJA data obtained for the 0.25° pixel containing the ground sites listed in Table 1, and the availability of at least 80 separate JJA days in which all three SM and LH products are simultaneously available (at a given site). In addition, a direct Spearman rank correlation of at least $R^2 > 0.03$ [−] is required for the mutual relationship among all three SM estimates and among all three LH products. While modest, this R^2 cutoff successfully masks sites where (6) produces nonphysical results (due to any single member of the SM or LH triplet being insufficiently accurate to support the application of triple collocation).

Baseline results are based on the use of Noah SM and LH and AMSR-E SM. For each site utilized in this study, Figure 1 plots the daily SM/LH Spearman rank coefficient of determination sampled directly from ground observations ($R^2[SM_G, LH_G]$; dashed blue line) and Noah estimates ($R^2[SM_{LSM}, LH_{LSM}]$; solid black line) of SM and LH. The dashed vertical lines group sites which are colocated within a single 0.25° grid box, and sites are ordered from left to right in terms of their aridity (as measured by averaged $R^2[SM_G, LH_G]$ within a particular 0.25° grid box).

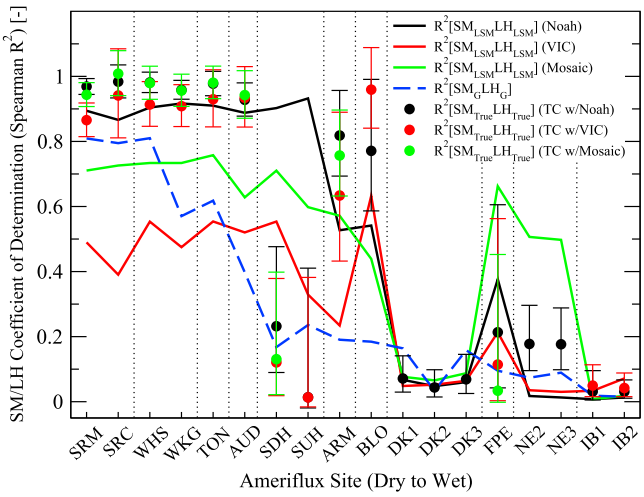

Figure 1. The Spearman rank coefficient of variation (R^2) between SM and LH estimated via the direct sampling of (various) LSM SM and LH predictions ($R^2[SM_{LSM}, LH_{LSM}]$), the direct sampling of ground-based SM and LH measurements ($R^2[SM_G, LH_G]$), and the application of the TC-based approach in (6) ($R^2[SM_{True}, LH_{True}]$) for the 0.25° grid cell over a range of ground sites. Results are shown for the application of various LSMs. Plotted error bars represent 95% confidence intervals on TC estimates. Vertical lines group sites within specific 0.25° grid cells.

Note that for the relatively highly coupled sites on the left-hand side of Figure 1, Noah generally predicts more SM/LH coupling than corresponding ground measurements. However, as discussed above, coupling estimates based on the direct sampling of independently acquired ground observations are likely biased low due to the impact of random SM and LH measurement errors.

In contrast, TC-based estimates of the true coupling strength ($R^2[SM_{True}, LH_{True}]$; circles) acquired from (6) are not biased by random error in LH or SM estimates. As a result, they tend to be larger than $R^2[SM_G, LH_G]$ for the arid sites in Figure 1. In fact, $R^2[SM_{True}, LH_{True}]$ acquired from (6) tends to be as large, or even larger than, Noah-based $R^2[SM_{LSM}, LH_{LSM}]$ (Figure 1). This implies that the high bias in $R^2[SM_{LSM}, LH_{LSM}]$ relative to $R^2[SM_G, LH_G]$ can be attributed—not to any systematic over coupling between SM and LH in Noah predictions—but rather the biasing impact of independent errors in ground-based SM and LH measurements. The single exception to this tendency is found at sites located within the Nebraska Sandy Hills region (SDH and SUH). Here values of $R^2[SM_{True}, LH_{True}]$ obtained from (6) remain larger than $R^2[SM_G, LH_G]$.

The Sand Hills region is notable for exhibiting vigorous surface-groundwater interactions and the direct uptake of ground water by vegetation within certain wet valleys [*Harvey et al.*, 2007; *Gosselin et al.*, 1999]. Since such surface/groundwater interactions are neglected in the version of Noah applied in NLDAS-2 (Noah 2.8), it is not surprising that Noah overestimates the degree to which grid-scale LH is water limited in the region. Therefore, at the SDH and SUH sites alone, there is clear evidence of significant over coupling between Noah SM and LH estimates.

An important verification strategy for TC results in Figure 1 is examining their sensitivity to variations in the products used to construct the SM and LH triplets on which they are based. This type of evaluation is especially critical for LSM-based SM and LH estimates since systematic errors in LSM SM/LH coupling may imperil the assumptions in (2) and (3) that errors in LSM-based SM and LH estimates are purely random.

In order to examine this issue, Figure 1 also plots TC-estimated $R^2[SM_{True}, LH_{True}]$ for the case of LSM-based SM and LH results derived from VIC and Mosaic. Note that VIC and Mosaic results at some sites are masked when a particular LSM fails to maintain an adequate Spearman cross correlation ($R^2 > 0.03$ [−]) versus both RS and ground-based SM and LH estimates. Prior to the application of TC, VIC and Mosaic predict substantially lower one-way SM/LH coupling ($R^2[SM_{LSM}, LH_{LSM}]$; green and red lines) than Noah.

However, with the possible exception of a modest low bias in VIC-based results, TC-corrected estimates of $R^2[SM_{True}, LH_{True}]$ obtained from (6) are generally insensitive to our choice of LSM. That is, TC-based $R^2[SM_{True}, LH_{True}]$ generated via the use of VIC, Mosaic, and Noah (as the source of LSM-based SM and LH estimates) generally fall within their mutual 95% confidence bars. This suggests that variations in the SM/LH coupling strength of LSMs do not appreciably bias TC-based estimates of $R^2[SM_{True}, LH_{True}]$ and lends credibility to coupling results based on (6).

A comparable robustness test is presented in Figure 2 where $R^2[SM_{True}, LH_{True}]$ are calculated using three different choices for the source of the RS-based SM product (i.e., AMSR-E, SMOS, and ASCAT—see section 3). The application of different RS-based SM products in (6) does not significantly alter $R^2[SM_{True}, LH_{True}]$ estimates.

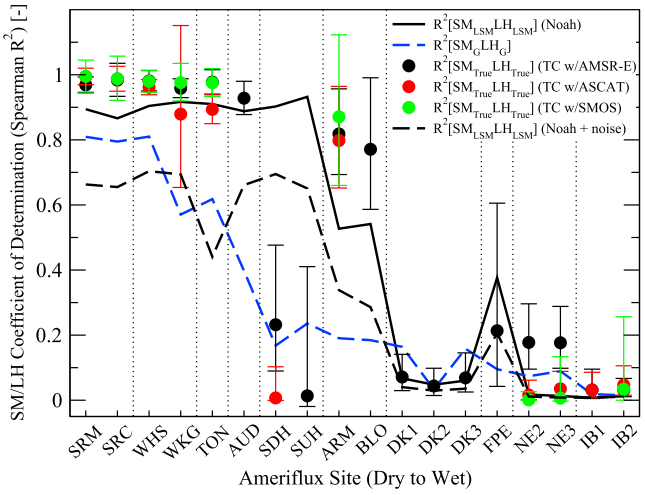

Figure 2. The Spearman rank coefficient of variation (R^2) between SM and LH estimated via the direct sampling of (various) LSM SM and LH predictions ($R^2[SM_{LSM}, LH_{LSM}]$) with and without synthetically generated noise, the direct sampling of ground-based SM and LH measurements ($R^2[SM_G, LH_G]$) and the application of the TC-based approach in (6) ($R^2[SM_{True}, LH_{True}]$) for the 0.25° grid cell covering a range of ground sites. Results are shown for the application of various RS SM products. Plotted error bars represent 95% confidence intervals on TC estimates. Vertical lines group sites within specific 0.25° grid cells.

This consistency persists despite the fact that the time period of data availability, and thus the temporal support of the $R^2[SM_{True}, LH_{True}]$ estimates, varies among the respective remotely sensed SM products.

Provided that the error independence assumptions underlying TC are respected, results should also be robust to the particular subgrid site selected for ground-based observations. Therefore, a final robustness test is examining the sensitivity of (6) to variations in the subgrid location of the ground measurement site. The dashed vertical lines in Figures 1 and 2 group ground sites by their respective 0.25° grid cells. Note that multiple ground sites located within a single grid cell (e.g., WHS/WKG, SRM/SRC, SUH/SDH, DK1/DK2/DK3, NE2/NE3, and IB1/IB2) generally yield mutually consistent TC-based $R^2[SM_{True}, LH_{True}]$ results. This suggests at least some robustness to variations in both static land cover conditions and dynamic meteorological forcing immediately surrounding each subgrid ground measurement site. For example, transitioning between ground observations obtained over shrub cover at the Lucky Hills (WHS) site and grass cover at the Kendall Grassland (WKG) site, 10 km apart, leads to comparable grid-scale coupling estimates even though these sites commonly experience different rainfall amounts during highly localized summer monsoon storm events. In Figures 1 and 2, a comparable lack of sensitivity exists between the Santa Rita Creosote (SRC, shrub cover with little grass) and Santa Rita Mesquite (SRM, tree cover with grass understory) sites and the Nebraska Sand Hill Dry Valley (SDH) and Upland Dune (SUH) sites.

6. Discussion

Relative to the other data products used in this analysis, the availability of ALEXI LH retrievals is particularly vital since all other RS-based LH products are either not sufficiently independent from LSM or RS SM products or do not provide a daily product. One consequence of this is that the robustness of results to the use of other RS-based LH products cannot be directly examined. Therefore, the most important untested assumption here is likely to be the mutual independence of LSM and RS LH estimates. Nevertheless, *Crow et al.* [2005] demonstrated that even in cases where LSMs and the Two-Source Energy Balance model (i.e., the diagnostic surface energy flux model at the core of ALEXI) are run with identical meteorological, vegetative, and radiative forcing information, they still yield LH estimates with mutually independent errors. This occurs because of the fundamentally different ways in which ALEXI and LSMs calculate surface energy fluxes [*Yilmaz et al.*, 2014].

A second issue is the realism of ground observation errors implied by this analysis. Based on (6), we attribute the observed difference between $R^2[SM_G, LH_G]$ and $R^2[SM_{True}, LH_{True}]$ to the impact of random error on ground-based SM and LH measurements. In order to further assess this claim, we assumed that LSM-based SM and LH predictions represent reality and examined how profoundly $R^2[SM_{LSM}, LH_{LSM}]$ is degraded via the introduction of artificial noise corresponding to expected levels of measurement error. To this end, Figure 2 also plots Noah-based $R^2[SM_{LSM}, LH_{LSM}]$ after independent, Gaussian random errors, with a standard deviation of 0.030 $m^3\,m^{-3}$ (volumetric) and 10 $W\,m^{-2}$ have been synthetically added to daily Noah SM and LH. These error levels are realistic expectations for the magnitude of random error present in daily ground-based SM and LH measurements [*Robinson et al.*, 2008; *Richardson et al.*, 2006]. Note that the random SM error

level is inflated slightly to capture the impact of spatial representativeness error associated with using a point-scale SM observation to represent a fetch-scale (~500 m^2) area consistent with the support of a tower-based LH measurement [*Famiglietti et al*., 2008].

The inclusion of such synthetic noise leads to a reductions in Noah-based $R^2[SM_{LSM}, LH_{LSM}]$ (compare solid and dashed black lines in Figure 2) that are generally as large as the observed difference between $R^2[SM_G, LH_G]$ and $R^2[SM_{True}, LH_{True}]$ (compare symbols to dashed blue line in Figure 2). This implies that the magnitude of random measurement error implied by (6) is consistent with our a priori expectations concerning random errors in ground-based SM and LH measurements.

7. Conclusions

Estimates of the true one-way coupling strength between SM and LH derived from the direct sampling of independent ground observations will be biased low by the presence of random measurement errors. Here we develop and apply a new triple collocation (TC) approach which provides a robust estimate of one-way SM/LH coupling strength (based on a Spearman rank coefficient of determination) which remains unbiased in the presence of random measurement and modeling errors. The TC approach is based on the simultaneous availability of SM and LH estimates acquired from ground measurements, remote sensing, and land surface modeling. Results suggest that the apparent overcoupling of LSM SM and LH predictions relative to ground measurements can be attributed to independent random errors in ground-based SM and LH measurements, and the Noah LSM provides relatively accurate predictions of SM/LH coupling strength across a range of climate and land cover types. In fact, there exists evidence of *undercoupling* within VIC and Mosaic SM and LH predictions at certain arid and semiarid sites (Figures 1 and 2). These TC-based inferences are robust to variations to LSM parameterizations, choice of RS-based SM data set and the subgrid location of ground-based SM and LH measurements. Following up on recent TC work with rainfall [*Alemohammad et al*., 2015], future work may be needed to clarify the impact of nonadditive error models for LH and SM.

Acknowledgments
Data were acquired from AmeriFlux Level 2 data (http://ameriflux.ornl.gov), SMOS Level 3 soil moisture (http://www.catds.fr/Products/Available-products-from-CPDC), ASCAT soil moisture (http://www.eumetsat.int/website/home/Data/Products/Land/index.html), and NLDAS-2 land surface model predictions (http://www.emc.ncep.noaa.gov/mmb/nldas/). Sand Hills Upland Dunes site data are available from D. Billesbach and T. Arkebauer. Daily ALEXI LH data are available from W.T. Crow.

The Editor thanks Alexandra Konings and Kaighin McColl for their assistance in evaluating this paper.

References

Albergel, C., C. Rüdiger, T. Pellarin, J. C. Calvet, N. Fritz, F. Froissard, D. Suquia, A. Petitpa, B. Piguet, and E. Martin (2008), From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations, *Hydrol. Earth Syst. Sci.*, *12*, 1323–1337.

Alemohammad, S. H., K. A. McColl, A. G. Konings, D. Entekhabi, and A. Stoffelen (2015), Characterization of precipitation product errors across the United States using multiplicative triple collocation, *Hydrol. Earth Syst. Sci.*, *19*, 3489–3503.

Al-Yaari, A., J. P. Wigneron, A. Ducharn, Y. H. Kerr, W. Wagner, and G. De Lannoy (2014), Global-scale comparison of passive (SMOS) and active (ASCAT) satellite-based microwave soil moisture retrievals with soil moisture simulations (MERRA-Land), *Remote Sens. Environ.*, *152*, 614–626.

Anderson, M. C., et al. (2011), Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery, *Hydrol. Earth Syst. Sci.*, *15*, 223–239.

Baldocchi, D. D., Q. Chen, X. Chen, S. Ma, G. R. Miller, Y. Ryu, J. Xiao, R. Wenk, and J. Battles (2010), The dynamics of energy, water and carbon fluxes in a Blue Oak (quercus douglasii) savanna in California, USA, in *Ecosystem Function in Savannas*, edited by M. J. Hill and N. P. Hanan, Taylor and Francis, Boca Raton, Fla.

Billesbach, D. P., and T. J. Arkebauer (2012), First long-term, direct measurements of evapotranspiration and surface water balance in the Nebraska Sand Hills, *Agric. For. Meteorol.*, *156*, 104–110.

Cammalleri, C., M. Anderson, and W. Kustas (2014), Upscaling of evapotranspiration fluxes from instantaneous to daytime scales for thermal remote sensing applications, *Hydrol. Earth Syst. Sci.*, *18*, 1885–1894.

Cosgrove, B. A., et al. (2003), Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, *J. Geophys. Res.*, *108*(D22), 8842, doi:10.1029/2002JD003118.

Crow, W. T., F. Li, and W. P. Kustas (2005), Intercomparison of spatially explicit distributed models for predicting surface energy flux patterns during SMACEX, *J. Hydrometeorol.*, *6*, 941–953.

Crow, W. T., D. G. Miralles, and M. H. Cosh (2010), A quasi-global evaluation system for satellite-based surface soil moisture retrievals, *IEEE Trans. Geosci. Remote Sens.*, *48*(6), 2516–2527.

Daly, C., R. P. Neilson, and D. L. Phillips (1994), A statistical-topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteorol.*, *33*, 140–158.

Dirmeyer, P. A., R. D. Koster, and Z. Guo (2006), Do global models properly represent the feedback between land and atmosphere?, *J. Hydrometeorol.*, *7*, 1177–1198.

Dirmeyer, P. A., C. A. Schlosser, and K. L. Brubaker (2009), Precipitation, recycling and land memory: An integrated analysis, *J. Hydrometeorol.*, *10*, 278–288.

Draper, C., R. Reichle, R. de Jeu, V. Naemi, R. Parinussas, and W. Wagner (2013), Estimating root mean square errors in remotely sensed soil moisture over continental scale domains, *Remote Sens. Environ.*, *137*, 288–298.

Famiglietti, J. S., D. Ryu, A. A. Berg, M. Rodell, and T. J. Jackson (2008), Field observations of soil moisture variability across scales, *Water Resour. Res.*, *44*, W01423, doi:10.1029/2006WR005804.

Fischer, M. L., M. S. Torn, D. P. Billesbach, G. Doyle, B. Northup, and S. C. Biraud (2012), Carbon, water, and heat flux responses to experimental burning and drought in a tallgrass prairie, *Agric. For. Meteorol.*, *166–167*, 169–174.

Goldstein, A. H., N. E. Hultman, J. M. Fracheboud, M. R. Bauer, J. A. Panek, M. Xu, Y. Qi, A. B. Guenther, and W. Baugh (2000), Effects of climate variability on the carbon dioxide, water, and sensible heat fluxes above a ponderosa pine plantation in the Sierra Nevada (CA), *Agric. For. Meteorol.*, *101*(2–3), 113–129.

Gosselin, D. C., S. Drda, F. E. Harvey, and J. Goeke (1999), Hydrologic setting of two inter-dunal valleys in the central Sand Hills of Nebraska, *Ground Water*, *37*, 924–933.

Guo, Z., et al. (2006), GLACE: The global land–atmosphere coupling experiment. Part II: Analysis, *J. Hydrometeorol.*, *7*, 611–625.

Harvey, F. E., J. B. Swinehart, and T. M. Kurtz (2007), Ground water sustenance of Nebraska's unique peatland fen ecosystems, *Ground Water*, *45*(2), 218–234.

Kerr, Y. H., et al. (2010), The SMOS mission: New tool for monitoring key elements of the global water cycle, *Proc. IEEE*, *98*(5), 666–687.

Krishnan, P., T. P. Meyers, R. L. Scott, L. Kennedy, and M. Heuer (2012), Energy exchange and evapotranspiration over two temperate semi-arid grasslands in North America, *Agric. For. Meteorol.*, *153*, 31–44.

Matamala, R., J. D. Jastrow, R. M. Miller, and C. T. Garten (2008), Temporal changes in C and N stocks of restored prairie: Implications for C sequestration strategies, *Ecol. Appl.*, *18*(6), 1470–1488.

McColl, K. A., J. Vogelzang, A. G. Konings, D. Entekhabi, M. Piles, and A. Stoffelen (2014), Extended Triple Collocation: Estimating errors and correlation coefficients with respect to an unknown target, *Geophys. Res. Lett.*, *41*, 6229–6236, doi:10.1002/2014GL061322.

Mesinger, F., et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*, 343–360.

Meyers, T. (2009a), Audubon research ranch FLUXNET L3 and L4 data, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., Oak Ridge, Tenn. [Available at ftp://cdiac.ornl.gov/pub/ameriflux/data/Level4/Sites_ByName/Audubon_Grasslands/.]

Meyers, T. (2009b), Fort Peck FLUXNET L3 and L4 data, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., Oak Ridge, Tenn. [Available at ftp://cdiac.ornl.gov/pub/ameriflux/data/Level4/Sites_ByName/Fort_Peck/.]

Naeimi, V., K. Scipal, Z. Bartalis, S. Hasenauer, and W. Wagner (2009), An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations, *IEEE Trans. Geosci. Remote Sens.*, *47*(7), 1999–2013.

Oishi, A. C., R. Oren, K. Novick, S. Palmroth, and G. G. Katul (2010), Interannual invariability of forest evapotranspiration and its consequence to water flow downstream, *Ecosystems*, *13*, 421–436.

Owe, M., R. de Jeu, and J. Walker (2001), A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index, *IEEE Trans. Geosci. Remote Sens.*, *39*(8), 1643–1654.

Owe, M., R. de Jeu, and T. Holmes (2008), Multisensor historical climatology of satellite-derived global land surface moisture, *J. Geophys. Res.*, *113*, F01002, doi:10.1029/2007JF000769.

Papuga, S. A. (2009), Highlight FLUXNET site Santa Rita Creosote bush, *FluxLetter: Newsletter FLUXNET*, *2*(4), 1–4.

Parinussa, R., T. Holmes, and W. T. Crow (2011), The impact of land surface temperature on soil moisture anomaly detection from passive microwave observations, *Hydrol. Earth Syst. Sci.*, *15*, 3135–3151.

Perrons, M. P., and L. N. Cooper (1992), When networks disagree: Ensemble methods for hybrid neural networks, in *Neural Networks for Speech and Image Processing*, Chapman and Hall, New York.

Richardson, A. D., et al. (2006), A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, *Agric. For. Meteorol.*, *136*, 1–18.

Robinson, D. A., C. S. Campbell, J. W. Hopmans, B. K. Hornbuckle, S. B. Jones, R. Knight, F. Ogden, J. Selker, and O. Wenderoth (2008), Soil moisture measurement for ecological and hydrological watershed-scale observatories: A review, *Vadose Zone J.*, *7*(1), 358–389.

Scott, R. L. (2010), Using watershed water balance to evaluate the accuracy of eddy covariance evaporation measurements for three semiarid ecosystems, *Agric. For. Meteorol.*, *150*(2), 219–225.

Seneviratne, S. I., et al. (2013), Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment, *Geophys. Res. Lett.*, *40*, 5212–5217, doi:10.1002/grl.50956.

Stoy, P., G. G. Katul, M. B. S. Siqueira, J. Y. Juang, K. A. Novick, J. M. Uebelherr, and R. Oren (2006), An evaluation of models for partitioning eddy covariance-measured net ecosystem exchange into photosynthesis and respiration, *Agric. For. Meteorol.*, *141*, 2–18.

Suyker, A. E., and S. B. Verma (2008), Interannual water vapor and energy exchange in an irrigated maize-based agroecosystem, *Agric. For. Meteorol.*, *148*(3), 417–427.

Van den Hurk, B., M. Best, P. Dirmeyer, A. Pitman, J. Polcher, and J. Santanello (2011), Acceleration of land surface model development over a decade of GLASS, *Bull. Am. Meteorol. Soc.*, *92*, 1593–1600.

Xia, Y., et al. (2012a), Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *J. Geophys. Res.*, *117*, D03109, doi:10.1029/2011JD016048.

Xia, Y., et al. (2012b), Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow, *J. Geophys. Res.*, *117*, D03110, doi:10.1029/2011JD016051.

Yilmaz, M. T., M. C. Anderson, B. Zaitchik, C. R. Hain, W. T. Crow, M. Ozdogan, J. A. Chun, and J. Evans (2014), Comparison of prognostic and diagnostic surface flux modeling approaches over the Nile River Basin, *Water Resour. Res.*, *50*, 386–408, doi:10.1002/2013WR014194.