Author/s:
Su, CH;Ryu, D;Crow, WT;Western, AW

Title:
Beyond triple collocation: Applications to soil moisture monitoring

Date:
2014-06-16

Citation:
Su, C. H., Ryu, D., Crow, W. T. & Western, A. W. (2014). Beyond triple collocation: Applications to soil moisture monitoring. Journal of Geophysical Research, 119 (11), pp.6419-6439. https://doi.org/10.1002/2013JD021043.

Persistent Link:
https://hdl.handle.net/11343/297399

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2013JD021043

Key Points:
- Triple collocation is one possible implementation of instrumental variables
- Lag variables are practical instruments to lift three-data requirement
- Utility of lag-based IV is demonstrated with satellite-retrieved soil moisture

Correspondence to:
C.-H. Su,
csu@unimelb.edu.au

Citation:
Su, C.-H., D. Ryu, W. T. Crow, and A. W. Western (2014), Beyond triple collocation: Applications to soil moisture monitoring, *J. Geophys. Res. Atmos.*, *119*, 6419–6439, doi:10.1002/2013JD021043.

Received 16 OCT 2013
Accepted 6 MAY 2014
Accepted article online 11 MAY 2014
Published online 10 JUN 2014

Beyond triple collocation: Applications to soil moisture monitoring

Chun-Hsu Su[1], Dongryeol Ryu[1], Wade T. Crow[2], and Andrew W. Western[1]

[1]Department of Infrastructure Engineering, University of Melbourne, Parkville, Victoria, Australia, [2]Hydrology and Remote Sensing Laboratory, USDA Agricultural Research Service, Beltsville, Maryland, USA

Abstract Triple collocation (TC) is routinely used to resolve approximated linear relationships between different measurements (or representations) of a geophysical variable that are subject to errors. It has been utilized in the context of calibration, validation, bias correction, and error characterization to allow comparisons of diverse data records from various direct and indirect measurement techniques including in situ remote sensing and model-based approaches. However, successful applications of TC require sufficiently large numbers of coincident data points from three independent time series and, within the analysis period, homogeneity of their linear relationships and error structures. These conditions are difficult to realize in practice due to infrequent spatiotemporal sampling of satellite and ground-based sensors. TC can, however, be generalized within the framework of *instrumental variable* (IV) regression theory to address some of the conceptual constraints of TC. We review the theoretics of IV and consider one possible strategy to circumvent the three-data constraint by use of lagged variables (LV) as instruments. This particular implementation of IV is suitable for circumstances where multiple data records are limited and the geophysical variable of interest is sampled at time intervals shorter than its temporal correlation length. As a demonstration of utility, the LV method is applied to microwave satellite soil moisture data sets to recover their errors over Australia and to estimate temporal properties of their relationships with in situ and model data. These results are compared against standard two-data linear estimators and the TC estimator as benchmark.

1. Introduction

When diverse sensors and models are used to explore a geophysical target, it is important to investigate how information from multiple sources can be compared and/or integrated. This is particularly challenging in hydrometeorology where measurements are often associated with considerable measurement errors and differences in scales due to different measurement supports. There is now frequent monitoring of surface wind, wave height, terrestrial soil moisture (SM), sea surface temperature, and precipitation using a wide range of instruments such as in situ buoys and gauges, satellite altimeters, scatterometers, radiometers, and/or numerical modeling. Depending on the context, validation, calibration, bias correction, and data assimilation can only be done properly when the error statistics in the data and relationships between different data are known. Indeed, from a minimalist perspective and if the relationships are linear, this can be recast as parameter estimation of the unknown errors in the individual data sets and the multiplicative and additive biases between the data sets.

Triple collocation (TC) [*Stoffelen*, 1998] is a powerful method to achieve asymptotically unbiased estimates of the error parameters in an affine signal model under certain conditions. In particular, TC uses three colocated data sets to jointly provide sufficient constraints for determining the error variance and the affine parameters that characterize the biases. The technique is now widely used in oceanography [e.g., *Stoffelen*, 1998; *Caires and Sterl*, 2003; *Janssen et al*., 2007; *O'Carroll et al*., 2008; *Vogelzang et al*., 2011] and hydrometeorology [e.g., *Scipal et al*., 2008; *Roebeling et al*., 2012]. However, challenges remain in practice due to the scarcity of global observational data from remote sensing and in situ monitoring stations, while sufficiently large samples are required for accurate analyses. Another limitation is the assumption of stationarity of the statistics and linearity between the three estimates (of same target) across all time scales. For many geophysical variables that exhibit a mixture of short- and long-time scale dynamics, the linear model fails to account for possible seasonal differences between the climatology of different data sets, among other things. Using the anomalies from the seasonal cycle in TC [*Miralles et al*., 2010] provides a partial remedy, as changes in the physical conditions can also enter into the variance of the errors. In so far as this is true, TC will lead to inaccurate

estimates. These additional complexities may be resolved with a sufficiently narrow-moving sampling window [e.g., *Crow et al*., 2011], but this also increases sampling errors of the estimation.

The purpose of this work is twofold. First, we appraise the concept of *instrumental variable* (IV) regression [*Wright*, 1928; *Reiersøl*, 1941] to make the connection that TC is one of many possible implementations of IV estimators. From the theoretical viewpoint, the data availability of three *coincident* estimates of the *same* geophysical variables is not a necessary requirement for unbiased parameter estimation. Second, we propose an alternative implementation of IV to relax TC's three-data requirement for solving the estimation problem. We show that *a lagged variable* (LV) can serve, for practical reasons, the purpose of the third data set in TC so that only two data sets are required. This lag-based IV (LV) estimation is suitable for geophysical variables that are sampled at time intervals shorter than their autocorrelation length.

As a proof of concept, we compare the utility of the proposed and existing methods using SM as a test bed. SM is a key variable in hydrological and meteorological processes, and an Essential Climate Variable of the Global Climate Observing System program. With diverse monitoring systems ranging from satellite sensors to in situ probes and land surface models that contribute toward global mapping of SM, there is much interest in deriving a complete and consistent data set from these multiple sources dating back to 1979 [*Liu*, *et al*., 2011; *Hollmann et al*., 2013]. Error characterizations [*Scipal et al*., 2008; *Dorigo et al*., 2010; *Draper et al*., 2013; *Su et al*., 2013a; *Gruber et al*., 2013] and rescaling to correct observational biases [*Yilmaz and Crow*, 2013] are necessary steps to achieve this goal [*Yilmaz et al*., 2012]. Moreover, error estimates are needed for specifying the relative certainty of the observations in the Kalman gain for data assimilation [*Crow and van den Berg*, 2010]. To do this at mesoscales, TC is typically performed on SM estimates from passive and active microwave observations of polar-orbiting satellites. However, their orbit and swath width specifications limit the number of coincident samples for TC analyses. Therefore, SM monitoring represents a fruitful area that will benefit from the proposed LV analyses.

This paper is organized as follows. Section 2 provides the theoretical background of this work with an outline of the problem of parameter estimation in the linear regime and a review of a few standard linear estimators. This motivates the discussion of IVs in general, their association with TC, and finally the development of the LV estimator. Subsequent sections focus on LV estimation applied to SM data. Section 3 describes the in situ, satellite and modeled SM data sets used in our studies. Section 4 illustrates the sampling limitation of satellites, which motivates the use of two-data estimators. The utility and limitations of LV are illustrated with synthetic experiments in section 5 and with real data in sections 6 and 7, where various methods were applied for estimating errors in satellite SM across Australia and dynamical relationships between various SM data sets.

2. Theory

Consider a (geophysical) variable x being measured at time t by an imperfect measuring apparatus,

$$X_t = x_t + u_t, \tag{1}$$

with x_t having a (sample) mean $E(x) = \mu_x$ and variance $\mathrm{var}(x) = \sigma_x^2$. $E(\circ)$ denotes expectation value in time, and var(◦) is the variance operator. The measurement error u_t has zero mean, variance σ_u^2, and is assumed to be serially uncorrelated in time and uncorrelated with the signal component. That is, it satisfies the properties $E(u) = 0$, $\mathrm{cov}(u_t, u_{t'}) = \sigma_u^2 \delta_{tt'}$, and $\mathrm{cov}(x_t, u_{t'}) = 0$ where $\delta_{tt'}$ is Kronecker delta, cov(◦) is the covariance operator, and t, t' are time indices used to distinguish time-lag covariance. Consider a second variable y that is linearly related to x via an affine map

$$y_t = c + \alpha x_t, \tag{2}$$

with $\mathrm{var}(y) = \sigma_y^2 = \alpha^2 \sigma_x^2$ and $E(y) = \mu_y$. Depending on the context, y can be assumed to take on the same physical meaning as x. In that case, the coefficients c and α will contribute to the additive and multiplicative biases (see later in equation (4)) that are conceptually used to account for differences between x and y in their physical representations or spatiotemporal supports. More generally, x and y can take on different physical meanings (see later in section 2.2) but share an underlying affine relation where c and α represent an intercept and a scaling term, respectively. The measurements or estimates of y are also generally erroneous so that its corresponding observed value at time t is

$$Y_t = y_t + v_t \tag{3}$$

where v_t is the associated measurement error with same properties $E(v)=0$, $\mathrm{cov}(v_t, v_{t'}) = \sigma_v^2 \delta_{tt'}$ and $\mathrm{cov}(y_t, v_{t'}) = 0$. If the two data sets X and Y are independently derived, their errors are assumed to be mutually uncorrelated: $\mathrm{cov}(u_t, v_{t'}) = 0$. The model assumes that the error and signal statistics and affine structure remain unchanged during the sampling/estimation period.

It is instructive to measure the level of disagreement between X and Y with the root-mean-square difference (RMSD), under the interpretation that direct comparison between x and y is physically meaningful. Substituting equations (1)–(3) into $\mathrm{RMSD}(X,Y)^2 \equiv E[(Y-X)^2]$ leads to

$$\mathrm{RMSD}(X,Y)^2 = E_X^2 + E_Y^2 + M^2 + B^2. \tag{4}$$

This describes a decomposition into four components, namely, stochastic errors in the two data (first two terms) with respective standard deviations $E_X \equiv \sigma_u$ and $E_Y \equiv \sigma_v$, a multiplicative bias between them defined as $M = \sqrt{(\alpha-1)^2\sigma_x^2}$, and an additive bias defined as $B \equiv E(Y) - E(X) = c + (\alpha-1)\mu_x$.

Given this setup, we have, on one hand, the problem of error estimation, which involves determining the errors E_X and E_Y. On the other hand, the problem of bias correction or calibration involves determining the coefficients c and α so as to eliminate biases M and B through a linear transformation $(Y-c)/\alpha$. Regardless of these contextual differences, these two problems are essentially the same, because unbiased estimates of the errors lead to unbiased estimates of the coefficients (and vice versa). To see this, we solve five first- and second-order statistical equations $\{E(p), \mathrm{cov}(p,q); p, q = \{X, Y\}\}$ between X and Y, where there are six unknown quantities, for the error and signal variances, to arrive at the following estimation equations (distinguished by a hat):

$$\hat{\sigma}_u^2 = \mathrm{var}(X) - \frac{\mathrm{cov}(X,Y)}{\hat{\alpha}}, \tag{5a}$$

$$\hat{\sigma}_v^2 = \mathrm{var}(Y) - \hat{\alpha}\,\mathrm{cov}(X,Y), \tag{5b}$$

$$\hat{\sigma}_x^2 = \mathrm{var}(X) - \hat{\sigma}_u^2, \tag{5c}$$

which yield unbiased estimates (i.e., $\hat{\sigma}_u = \sigma_u, \hat{\sigma}_v = \sigma_v$ in probability) if and only if $\hat{\alpha} = \alpha$. This is the same for the intercept $\hat{c} = E(Y) - \hat{\alpha}E(X)$. Notably, each term of the RMSD's decomposition can be individually computed if α is known. Various methods of estimation differ only in the way they estimate α.

It is important to note that α, c, M, and B are extrinsic properties that are defined with respect to a reference data (X in this case), while the error variances of equations (5a) and (5b) are intrinsic properties of each data set and require no reference data. It is a common practice to preprocess the data at hand to correct the biases before estimation. In cases where the mean of Y is modified to match E(X), B becomes zero and $c = (1-\alpha)\mu_x$, while the observed errors remain unchanged from our definition of E_X and E_Y in equations (5a) and (5b). In other cases where Y is also rescaled to correct the multiplicative bias with some rescaling value α', bias becomes $M = \sqrt{(\alpha/\alpha' - 1)^2\sigma_x^2}$ and the observed error variance in Y is rescaled to become E_Y^2/α'^2. Of course, unless there was a priori knowledge of the true α or unbiased estimators (see later) were used, bias M is nonzero. Introducing such preprocessing would complicate the notation but would not otherwise provide new insights to the estimation problem.

In the following review of standard linear estimation methods, we simplify the discussion by assessing their properties based on their ability to retrieve an unbiased estimate of α.

2.1. Standard Linear Estimators

The first two-data estimator is the method of ordinary least squares (OLS) regression. The standard regression analysis of regressing Y onto the regressor variable X corresponds to solving for values of $\hat{\alpha}$ and $\hat{c}$ that minimize the Euclidean distance (L_2-norm) between Y and $\hat{c} + \hat{\alpha}X$. This leads to the OLS estimation equation for the scaling coefficient,

$$\hat{\alpha}_{\mathrm{OLS}} = \frac{\mathrm{cov}(X,Y)}{\mathrm{var}(X)} = \frac{\alpha}{1 + \sigma_u^2/\sigma_x^2}. \tag{6}$$

Given that $\hat{\alpha}_{\mathrm{OLS}} < \alpha$ when errors are present in X, this expression yields a biased estimate smaller than the true value. This is widely known as the *errors-in-variable* bias. Equation (6) can be interpreted to provide a lower

bound. Of course, if the error variance is known, the bias can be corrected. Otherwise, it is only in the limit of high signal-to-noise ratio (SNR) in X that the estimator becomes near optimal. In terms of $\mathrm{SNR}_X = \sigma_x^2/\sigma_u^2 > 1$, its power series expansion is

$$\hat{\alpha}_{\mathrm{OLS}} = \alpha\left(1 - \frac{1}{\mathrm{SNR}_X} + \frac{1}{\mathrm{SNR}_X^2} - \frac{1}{\mathrm{SNR}_X^3} + \ldots\right). \tag{7}$$

One may also consider a reverse OLS (rOLS) regression with Y as the regressor. In this case, overestimation occurs to provide an upper bound

$$\hat{\alpha}_{rOLS} = \frac{\mathrm{var}(Y)}{\mathrm{cov}(X,Y)} = \alpha\left(1 + \frac{\sigma_v^2}{\sigma_y^2}\right), \tag{8}$$

which depends on the SNR in Y ($SNR_Y = \sigma_y^2/\sigma_v^2$). The presence of the errors therefore leads to inconsistent estimates, with $\hat{\alpha}_{\mathrm{OLS}} \neq \hat{\alpha}_{\mathrm{rOLS}}$. Consequently, when the estimated coefficients are used for bias correction or calibration of Y with respect to X, OLS leads to overestimation of the signal component, whereas rOLS leads to underestimation. Therefore, if one seeks to use standard regression in practice, it is recommended that both forward and reverse regressions are computed to provide bounds for α, and the better choice between the two depends on the relative SNR between the two data sets.

The second two-data estimator is to match the statistical moments of the two data sets. The first-order (linear) approach is to match their first two moments. The matching of their means enables unbiased estimate of the intercept. The matching of the variance (VAR) corresponds to taking the ratio of their standard deviation

$$\hat{\alpha}_{\mathrm{VAR}} = \sqrt{\frac{\mathrm{var}(Y)}{\mathrm{var}(X)}} = \sqrt{\frac{\alpha^2 + \sigma_v^2/\sigma_x^2}{1 + \sigma_u^2/\sigma_x^2}}, \tag{9}$$

but this also contains the errors-in-variable bias, and the direction of the bias depends on the relative SNR between the two data sets. The method essentially assumes absence of errors or identical SNR in X and Y. The nonlinear method of matching all statistical moments using cumulative distribution function (CDF) matching [*Reichle and Koster*, 2004] is likely to share similar pitfalls. Heuristically, the CDF mapping will also contain the extraneous contribution of the error variances (as observed in equation (9)) in the mapping of the second moment (also, skewness and kurtosis at higher moments). This was exemplified in synthetic twin data assimilation experiments of *Yilmaz and Crow* [2013] where VAR- and CDF-matching showed almost identical performance in bias correction for assimilation. Comparing against the OLS's equation (7), the power series expansion of the variance matching (VAR) solution in equation (9) is

$$\hat{\alpha}_{\mathrm{VAR}} = \alpha\left(1 + \frac{1}{2\mathrm{SNR}_Y} - \frac{1}{2\mathrm{SNR}_X} + \ldots\right). \tag{10}$$

Thus, without prior knowledge of the relative SNR, it is more conservative to use VAR than OLS because the influence of the errors on the estimation is smaller in VAR than that in OLS.

The third method, triple collocation (TC), overcomes the shortcoming of the OLS and VAR by using three coincident and independent estimates of the same geophysical variable [*Stoffelen*, 1998]. In addition to x and y in equations (1)–(3), a third (error-free) representation (z) of the same target is considered

$$z_t = d + \beta x_t \tag{11}$$

where d and β are the associated intercept and scaling coefficients. Needless to say, the measurement record of z may also be erroneous with an unknown error component w,

$$Z_t = z_t + w_t. \tag{12}$$

As before, the error w is uncorrelated with x, u, and v. Now the means of and covariances among the three data sets (equations (1), (3), and (12)) pose sufficient constraints to resolve the relationships among the three data sets. In particular, there are nine first- and second-order statistical equations given by

$\{E(p), \mathrm{cov}(p,q); p,q = \{X,Y,Z\}\}$ for solving nine unknown parameters including the required affine coefficients. The solution for the scaling coefficient between X and Y is

$$\hat{\alpha}_{\mathrm{TC}} = \frac{\mathrm{cov}(Y,Z)}{\mathrm{cov}(X,Z)}, \tag{13}$$

By substituting equations (1), (3), (11), and (12) into equation (13), it can be shown that this estimator are unbiased in probability, i.e., $\hat{\alpha}_{\mathrm{TC}} = \alpha$, if all the assumptions considered so far hold and data sample for statistical analysis is sufficiently large. Simulations of *Zwieback et al*. [2012] showed that sample sizes of ≥500 coincident triplets are needed to limit sampling errors of the estimator within 10% uncertainty. In some cases, smaller data sets have been used for pragmatic reasons [e.g., *Scipal et al*., 2008; *Dorigo et al*., 2010]. The standard TC also assumes absence of autocorrelation in the errors [*Zwieback et al*., 2012]. While the estimate of the scaling coefficient remains unbiased if cross correlations among errors remain absent, the error estimates based on equation (5) are now influenced by the autocorrelation structure of the errors. There are also extensions to TC using lagged moments as avenues to resolve the autocorrelation structure of the errors [*Sargan*, 1959; *Zwieback et al*., 2013].

2.2. Method of Instrumental Variables

To introduce the method of instrumental variables, we first recall the theoretical reasoning for the suboptimal performance of the OLS estimator. It is equivalent to the solution to the normal equation $\mathbf{X}^T(\mathbf{Y} - \mathbf{X}\boldsymbol{\varphi}_{\mathrm{OLS}}) = \mathbf{0}$, where the column vector $\boldsymbol{\varphi}_{\mathrm{OLS}} = [c_{\mathrm{OLS}}, \alpha_{\mathrm{OLS}}]^T$ contains the required affine coefficients and superscript T denotes matrix transpose. Here $\mathbf{Y}$ is a $N \times 1$ matrix containing the second data set Y of sample size N, $\mathbf{X}$ is a $N \times 2$ matrix whose first column is $[\mathbf{1}]$, and second column contains the first data set (regressor X). The normal equation can be interpreted as a condition that the estimated regressed vector must yield a residual $\boldsymbol{\varepsilon} \equiv \mathbf{Y} - \mathbf{X}\boldsymbol{\varphi}_{\mathrm{OLS}}$, which is uncorrelated with the regressor. By inverting the normal equation, we arrive at the estimation equation $\hat{\boldsymbol{\varphi}}_{\mathrm{OLS}} = \left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{Y}$, whose second element is identical to the previous solution in equation (6). However, in the general case where the regressor X contains errors, the residual is in fact $\varepsilon = v - \alpha u$ and thus correlated with X; this violation of the normal equation condition is responsible for the observed errors-in-variable bias in the OLS estimate [*Maddala*, 2001].

This problem is circumvented by using an additional "instrumental" variable (IV) Z that is uncorrelated with this residual ε so that the above normal equation can be replaced with [*Maddala*, 2001]

$$\mathbf{Z}^T\boldsymbol{\varepsilon} = \mathbf{Z}^T(\mathbf{Y} - \mathbf{X}\boldsymbol{\varphi}_{\mathrm{IV}}) = \mathbf{0} \tag{14}$$

where $\mathbf{Z}$ is the $N \times 2$ matrix of the third data set, organized akin to $\mathbf{X}$. Solving equation (14) for $\boldsymbol{\varphi}_{\mathrm{IV}}$ leads to the following IV estimation equation for c and α

$$\hat{\boldsymbol{\varphi}}_{\mathrm{IV}} = \left(\mathbf{Z}^T\mathbf{X}\right)^{-1}\mathbf{Z}^T\mathbf{Y} \tag{15}$$

under the nonsingular assumption. In TC, it is explicit that the three data sets describe three representations of the same geophysical variable. With respect to the *X-Y* relationship, the third independently derived data Z satisfies the condition of equation (14) because it is not contained in the equation of interest and can be regarded as an IV. Indeed, the second element of equation (15) can be shown to recover the TC estimator of equation (13). More importantly, the framework of IV regression theory [*Reiersøl*, 1941; *Sargan*, 1958] generalizes TC in resolving the affine structures and highlights the fact that the above physical interpretation of the TC is superficial. Broadly speaking, an IV is any variable that is linearly correlated with the predetermined variables and independent of the residual ε from their estimated relationship. In particular, the condition of equation (14) infers that an instrument Z, defined with respect to the *X-Y* relationship, has the following properties:

$$\mathrm{cov}(Z,\varepsilon) = 0, \tag{16a}$$

$$\mathrm{cov}(X,Z) \neq 0. \tag{16b}$$

By symmetry, Y is the instrument with respect to the *X-Z* relationship, leading to a similar estimation equation $\left[\hat{d}_{\mathrm{IV}}, \hat{\beta}_{\mathrm{IV}}\right]^T = \left(\mathbf{Y}^T\mathbf{X}\right)^{-1}\mathbf{Y}^T\mathbf{Z}$. The standard TC analysis can therefore be regarded as one of the many possible implementations of IVs if there is a multiplicity of other variables that satisfy the conditions of equation (16).

The difference between TC and IV is subtle but important, where the latter is inclusive of using another geophysical or a latent variable as an IV. Though this theoretical formalism is widely known in econometrics, its relation to existing empirical work in TC has not been fully appreciated.

As with linear regression, we can consider the standard error (SE) of the estimators. For the IV estimator, SE is given by the square root of the diagonal elements of the covariance matrix of the IV estimate [Sargan, 1958]

$$\mathrm{cov}(\boldsymbol{\varphi}_{\mathrm{IV}}) = \mathrm{var}(\varepsilon)\left(\mathbf{Z}^{\mathbf{T}}\mathbf{X}\right)^{-1}\left(\mathbf{Z}^{\mathbf{T}}\mathbf{Z}\right)\left(\mathbf{X}^{\mathbf{T}}\mathbf{Z}\right)^{-1}. \tag{17}$$

under the assumption that asymptotic joint distribution between the residual and IV is normal. Expansion of equation (17) yields

$$\mathrm{var}(c_{\mathrm{IV}}) = \mathrm{var}(\varepsilon)\frac{\mathrm{var}(Z)E(X)^2 + \mathrm{cov}(X,Z)^2}{N\,\mathrm{cov}(X,Z)^2}. \tag{18a}$$

$$\mathrm{var}(\alpha_{\mathrm{IV}}) = \mathrm{var}(\varepsilon)\frac{\mathrm{var}(Z)}{N\,\mathrm{cov}(X,Z)^2}. \tag{18b}$$

Equations (17) and (18) are naturally applicable for TC and OLS. For instance, the SE for OLS estimates can be easily retrieved by replacing **Z** with **X** in equation (17). Furthermore, these uncertainties can be propagated through equation (5) using the method of partial derivatives to arrive at the following expressions for the uncertainties of the error and signal estimates

$$\mathrm{var}\left(\sigma_u^2\right) = \mathrm{var}[\,\mathrm{var}(X)] + \frac{\mathrm{cov}(X,Y)^2}{\hat{\alpha}_{\mathrm{IV}}^4}\,\mathrm{var}(\alpha_{\mathrm{IV}}) + \frac{1}{\hat{\alpha}_{\mathrm{IV}}^2}\,\mathrm{var}[\,\mathrm{cov}(X,Y)], \tag{19a}$$

$$\mathrm{var}\left(\sigma_v^2\right) = \mathrm{var}[\,\mathrm{var}(Y)] + \mathrm{cov}(X,Y)^2\,\mathrm{var}(\alpha_{\mathrm{IV}}) + \hat{\alpha}_{\mathrm{IV}}^2\,\mathrm{var}[\,\mathrm{cov}(X,Y)], \tag{19b}$$

$$\begin{aligned}\mathrm{var}\left(\sigma_x^2\right) = {} & \left[\frac{\mathrm{cov}(X,Z)}{\mathrm{cov}(Y,Z)}\right]^2\mathrm{var}[\,\mathrm{cov}(X,Y)] + \left[\frac{\mathrm{cov}(X,Y)}{\mathrm{cov}(Y,Z)}\right]^2\mathrm{var}[\,\mathrm{cov}(X,Z)] + \cdots \\ & \left[\frac{\mathrm{cov}(X,Y)\,\mathrm{cov}(X,Z)}{\mathrm{cov}(Y,Z)^2}\right]^2\mathrm{var}[\,\mathrm{cov}(Y,Z)]\end{aligned} \tag{19c}$$

where $\mathrm{var}[\mathrm{cov}(p,q)] = \{E[(p-\mu_p)^2(q-\mu_q)^2] - \mathrm{cov}(p,q)^2\}/N$ approximates the variance of the sample covariance between data p and q for $p,q = \{X,Y,Z\}$.

The IV framework also extends the first-order linear analysis to allow K-order polynomial regression analysis between x and y [*Bowden and Turkington*, 1990]

$$y_t = \Sigma_{k=1}^{K}\varphi_k x_t^k. \tag{20}$$

In which case, the solution of equation (15) remains applicable but matrices $\boldsymbol{\phi}$, **X**, and **Z** are in extended forms $\boldsymbol{\varphi}_{\mathrm{IV}}{}^{(K)} = [\,\varphi_0 \quad \varphi_1 \quad \ldots \quad \varphi_K\,]^T$, $\mathbf{X}^{(K)} = [\,\mathbf{1} \quad \mathbf{X} \quad \mathbf{X}^2 \quad \ldots \quad \mathbf{X}^K\,]$, and $\mathbf{Z}^{(L)} = [\,1 \quad \mathbf{Z}_1 \quad \mathbf{Z}_2 \quad \ldots \quad \mathbf{Z}_L\,]$, respectively. The number of regression coefficients, however, needs to be matched (at least) by the number of IVs (L) for estimability. Where the matrix $\mathbf{Z}^{\mathbf{T}}\mathbf{Y}$ is singular and noninvertible or where there are more IVs than number of regression coefficients, a consistent set of coefficients are determined via minimization of the Mahalanobis square distance, $(\mathbf{Y} - \boldsymbol{\varphi}_{\mathrm{IV}}\mathbf{X})^T\mathbf{W}\mathbf{Z}^T(\mathbf{Y} - \boldsymbol{\varphi}_{\mathrm{IV}}\mathbf{X})$, where $\mathbf{W} = (\mathbf{Z}^T\mathbf{Z}/N)^{-1}$ is a typical choice for a weighting matrix [*Sargan*, 1958; *Hansen*, 1982]. In this case, the IV estimator yields a consistent result given by

$$\hat{\boldsymbol{\varphi}}_{\mathrm{IV}} = \left(\mathbf{X}^T\mathbf{Z}\mathbf{W}\mathbf{Z}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Z}\mathbf{W}\mathbf{Z}^T\mathbf{Y}. \tag{21}$$

The asymptotic solutions presented above serve only as approximations; there are practical limitations of using IV and TC. First, the use of instruments that explain little of the variability in x can make their estimation highly sensitive to small correlation between the instrument and the residual. To see this, we recast the solutions of IV (and TC) in equations (13) and (15) in terms of linear correlation R as $\hat{\alpha}_{\mathrm{IV}} = \alpha + \sqrt{\mathrm{var}(\varepsilon)/\mathrm{var}(X)}\,R(Z,\varepsilon)/R(X,Z)$. If $R(x,z)$ is too small, then even a very small correlation between the instrument and the error can produce a large inconsistency in the IV estimate of α [*Bound et al.*, 1995]. Second and of relevance to all the estimators discussed, the standard computation of statistics, i.e., arithmetic sample mean and sample covariance, is sensitive to outliers, since they will attract the mean and inflate the covariance. To improve their robustness, it may be advisable to screen the multivariate data for outliers before estimation. Several

techniques for outlier detection are available, e.g., based on minimizing ellipsoid volume [*Rousseeuw*, 1985] or the covariance matrix determinant [*Rousseeuw and Van Driessen*, 1999], and could be implemented to obtain more robust estimates of the sample mean and covariance. Third, finite and small sample sizes can lead to biased estimates [*Bound et al*., 1995, *Nelson and Startz*, 1990].

This review of IV estimators is by no means comprehensive of their asymptotic and finite-sample properties. It is provided to prompt practitioners of TC to appreciate the theoretical and empirical advances made in IV theory and the close connections between the two areas. Readers are referred to *Bowden and Turkington* [1990] for a useful introduction to IV estimation.

2.3. Method of Lag-Based Instrumental Variables

The challenge of implementing an IV estimator is finding proper instruments. Strict observations of the conditions in equation (16) exclude the predetermined variables (*X* and *Y*) in the relationship, their lagged values, and other variables that are not constructed independently from the variables in the relationship. However, when a geophysical variable of interest had been sampled at time intervals shorter than their temporal correlation length, the lagged values of the predetermined variable can be candidate IVs under practical circumstances where simultaneous third measurements are unavailable or the number of triplet observations is inadequate. The model for the proposed lag-based IV (LV) estimator is described as follows.

Suppose the variable *x* of interest has strong lag *m* autocorrelation [*Maddala*, 2001]

$$x_t = d + \eta x_{t-m} + w_t \tag{22}$$

where η is the autocorrelation coefficient in *X* at lag *m*. Considering the lagged values of *X* as the IV by replacing *Z* with $X_{t\text{-}m}$ in equation (14), the revised IV estimation equation for LV is

$$\hat{\boldsymbol{\varphi}}_{\mathrm{LV}} = \left(\mathbf{X}_{t-m}^T \mathbf{X}_t\right)^{-1} \mathbf{X}_{t-m}^T \mathbf{Y}_t, \tag{23}$$

where $\hat{\boldsymbol{\varphi}}_{\mathrm{LV}} = [\hat{c}_{\mathrm{LV}}, \hat{\alpha}_{\mathrm{LV}}]$. This expression can be shown to be unbiased if the error in *X* is uncorrelated in time, with properties $\mathrm{cov}(e_t, e_{t'}) = \sigma_e^2 \delta_{tt'}$ and $\mathrm{cov}(e_t, w_{t'}) = 0$ for $e = \{u,v\}$. Equivalently, the lagged values of *Y* can be used as the IV such that the corresponding estimator is $\hat{\boldsymbol{\varphi}}_{\mathrm{LV},Y} = \left[\hat{c}_{\mathrm{LV},Y}, \hat{\alpha}_{\mathrm{LV},Y}\right] = \left(\mathbf{Y}_{t-m}^T \mathbf{X}_t\right)^{-1} \mathbf{Y}_{t-m}^T \mathbf{Y}_t$.

However, in the presence of autocorrelated errors in either *X* or *Y*, LV leads to biased estimates, depending of the choice of the variable as instrument. In particular, the two LVs yield inconsistent estimates of the scaling factors; notably, the second elements in vectors $\hat{\boldsymbol{\varphi}}_{\mathrm{LV}}$ and $\hat{\boldsymbol{\varphi}}_{\mathrm{LV},Y}$ are

$$\hat{\alpha}_{\mathrm{LV}} = \frac{\eta\alpha\sigma_x^2}{\eta\sigma_x^2 + \mathrm{cov}(u_t, u_{t-m})} \tag{24a}$$

$$\hat{\alpha}_{\mathrm{LV},Y} = \frac{\eta\alpha\sigma_x^2 + \mathrm{cov}(v_t, v_{t-m})}{\eta\sigma_x^2}, \tag{24b}$$

respectively. Their inconsistency can be used to detect possible presence of autocorrelated errors in either *X* or *Y*, but which estimator is closer to the true value clearly depends on the assumption of the relative autocorrelation of the errors. Possible remedies to address this may be to use a priori knowledge of temporal persistence of the model errors (through model characterization) to augment the estimator or to use lag moments in IV [*Sargan*, 1959] or TC [*Zwieback et al*., 2013] to resolve autocorrelated structure of the errors. Despite this shortfall, the influence of such errors may diminish with the length of the lag *m*. For instance, suppose the autocorrelation in the error *e* is an order-1, zero mean ($e_0 = 0$), stationary ($\rho < 1$) autoregressive (AR) process

$$e_t = \rho e_{t-1} + \sigma_e \gamma \tag{25}$$

for $\gamma \sim$ iid $N(0,1)$. Since its autocovariance diminishes exponentially with *m* as

$$\mathrm{cov}(e_t, e_{t-m}) = \frac{\rho^m \sigma_e^2}{1-\rho^2}, \tag{26}$$

the estimates of equation (24) tend to the true values when autocorrelation of the signal *x* is stronger than that of the errors $\eta\sigma_x^2 >> \mathrm{cov}(u_t, u_{t-m})$ and/or $\eta\alpha\sigma_x^2 >> \mathrm{cov}(v_t, v_{t-m})$ for larger *m* and/or smaller ρ. It is of note that for $\rho < 1$, OLS estimate $\hat{\alpha}_{\mathrm{OLS}}$ will still be more negatively biased ($< \alpha$) than $\hat{\alpha}_{\mathrm{LV}}$ in the presence of such an autocorrelated error.

It should be emphasized that if there is no practical limitation with data availability, TC with three independent data sets remains a better implementation of the IV estimator to meet the requirement of the instrument in equation (16). Modeled and remote sensing data typically contain autocorrelated errors. Temporal memory of the model can lead to serial correlations in model errors. For remote sensing, observed radiance are typically influenced by processes that operate at different time scales, and the difficulty of resolving individual processes with simplified and computationally tractable physical models leads to autocorrelated errors in retrievals. Amid these considerations, we will provide proof-of-concept demonstrations of LV estimation using actual modeled and remotely sensed SM data in sections 6 and 7.

3. Soil Moisture Data

Surface SM is used as a test variable for our examination of the various estimators. It typically shows strong autocorrelation because of the relatively slow transient behavior of drainage and evapotranspiration, although rainfall wetting decreases this autocorrelation. Thus, when SM is sampled at intervals shorter than arrival times of rainfall events, the strong signal autocorrelation is retained in the time series. The mixing of the two processes with their characteristic times (related to land surface properties and climatology), coupled with suitable sampling frequency, inform the suitability of using the LV estimator and the choice of lag m. In this case for SM, the lag-based signal model of equation (22) is also an approximation as it assumes that the amount of water loss is linearly related to SM. In reality, the loss can be nonlinear [*Famiglietti and Wood*, 1994] such that the ratio of x_t to x_{t-m} varies with time.

Our case studies use SM data sets from ground measurements, satellite remote sensing, and land surface model. Different data sets and configurations were used to illustrate the properties of various estimators using synthetic experiments and test the utility of LV with real data. The spatial domain of focus is Australia, but we also use point-scale data to bring focus on the properties of the estimators and relationships between these data sets. The data sets considered are listed as follows.

The point-scale in situ SM at 0–5/8 cm depth was measured at 14 monitoring stations at the 600 km^2 Kyeamba Creek catchment, southeast Australia. The November 2001 to December 2010 SM records from one of the stations located at −35.49° latitude, 147.56° longitude, south of the catchment, will be used in this study. The station is installed at a grassland/cropping area, with a temperate climate, Köppen Cfa [*Peel et al.*, 2007], characterized by seasonally uniform rainfall and hot summer. SM varies seasonally from dry in summer (December–February) to wet in winter (June–August) due to the seasonality of evapotranspiration and rainfall forcing. The vegetation cover is greatest during spring months around September/October and is lowest from the end of summer in February/March. Readers are referred to *Smith et al.* [2012] and *Su et al.* [2013c] for details of the monitoring network OzNet and the catchment.

Two satellite-retrieved SM products are considered. The first is a passive microwave product derived from AMSR-E (Advanced Microwave Scanning Radiometer-Earth Observing System) (June 2002 to October 2011) instrument on board the Aqua satellite. Based on observed C- and X-band emission from the surface soil, the radiometer enabled retrievals of surface SM at shallow depths; C-band has a footprint of 74×43 km^2 and topsoil sensing depth of 1–2 cm while X-band has smaller footprints but < 5 mm penetration depths. Its scans of Australia occurred around midday (1:30 P.M. local time) and midnight (1:30 A.M.). Among several retrieval approaches (see recent review by *Mladenova et al.* [2014, and references therein]), the Land Parameter Retrieval Model (LPRM) approach [*Owe et al.*, 2008] used the instrument's dual-polarization observations to estimate the vegetation optical thickness and the soil dielectric constant. A dielectric mixing model is then applied to relate the dielectric constant to volumetric SM. Here we use the version 5 LPRM, combined C/X-band SM data set, expressed on a regular 1/4° (longitude) × 1/4° (latitude) grid, but it is mainly C-band SM over Australia. The data was screened for poor-quality retrievals using the accompanying quality flag to identify swath edge effects, frozen or thawing soil, radio frequency interference, and dense vegetation.

The second is an active microwave product from ASCAT (advanced scatterometer) on board of MetOp-A satellite. It uses the backscatter of C-band emission to estimate SM. From October 2006, the satellite overpasses over Australia at around 9 A.M. and 9 P.M. The backscatter is sensitive to the moisture content of the scattering land surface, and by using a change detection algorithm [*Wagner et al.*, 1999], the SM is

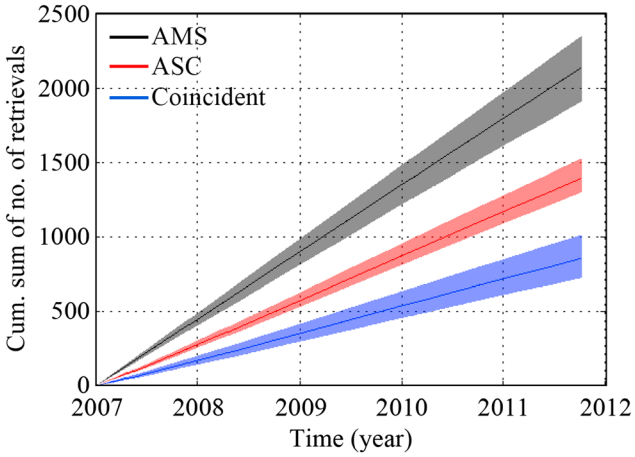

Figure 1. Cumulative sum of the number of AMS and ASC retrievals, and their coincidences at 11,130 spatial points over Australia. The solid curves correspond to medians, bounded by 10th–90th percentiles (shading).

estimated relative to historical minima and maxima as a degree of saturation (θ_{rel}). We use the ASCAT data set (January 2007 to October 2013) produced using the Water Retrieval Package [*Naeimi et al.*, 2009] version 5.5. The product is defined over a sinusoidal grid with ~12.5 km resolution. The data was also preprocessed to screen for frozen surface and temporary standing water using its retrieval quality flag and converted to volumetric SM (θ_{vol}) using ancillary information on saturable soil water content ϕ via $\theta_{vol} = \theta_{rel}\phi$ [*Wagner et al.*, 1999; *Su et al.*, 2013c]. The ancillary data is taken from the *McKenzie et al.* [2000] interpretations of Digital Atlas of Australian Soils [*Northcote et al.*, 1960-1968] for the A-horizon (top) soil layer.

The final SM data set is the modeled SM field from the MERRA (Modern Era Retrospective-Analysis for Research and Applications)-Land reanalysis. The MERRA atmospheric reanalysis is driven by a vast collection of in situ observations of atmospheric and surface winds, temperature, and humidity, and remote sensing of precipitation and radiation [*Rienecker et al.*, 2011]. The MERRA land-only fields were postprocessed by reintegrating a revised Catchment model with more realistic precipitation forcing to produce the MERRA-Land data set. [*Reichle et al.*, 2011]. We use the MERRA-Land data, produced by the catchment land surface model version Goddard Earth Observing System 5.7.2, and the hourly average uppermost land surface fields are gridded in 2/3° × 1/2° grid.

As shorthand, the four data sets are referred to as INS (in situ), AMS (AMSR-E), ASC (ASCAT), and MER (MERRA-Land).

4. Sampling Limitation of Satellite Soil Moisture

TC analyses of satellite SM typically involve a passive microwave-derived, an active microwave-derived, and a modeled SM data series [*Scipal et al.*, 2008; *Dorigo et al.*, 2010; *Draper et al.*, 2013; *Su et al.*, 2013b]. By comparison, a two-data estimator (OLS, VAR, and LV) can be applied to each satellite SM data set individually, in conjunction with a modeled SM as the second data. The applicability of TC is therefore more restricted by the sampling frequency and coincidence of the satellites in practice. To motivate the use of the two-data estimator in remote sensing of SM, we examine the sampling characteristics of AMSR-E and ASCAT by analyzing the spatiotemporal coverage of their data sets over Australia.

AMS and ASC data sets over Australia were cross referenced spatially based on nearest neighbor and temporally based on night (9 P.M./1:30 A.M.) and day (9 A.M./1:30 P.M.) overpass times of the satellites. The data at each of the 11,130 spatial locations over the regular 1/4° × 1/4° grid were examined to compute the cumulative sum of the number of retrievals (i.e., the growth of the sample size) by individual sensors and their coincidences. Figure 1 plots the cumulative frequency analysis, showing the spatial median, 10th and 90th percentiles to represent the spatial variability in the satellite coverage at different latitudes. In general, the sample size of retrieved SM grows linearly with time, with observations by AMS more frequent than ASC. This is because ASCAT has a longer revisit time of 2–3 days and a narrower effective swath width of 1100 km, c.f. 1–2 day revisit time and 1445 km swath width of AMSR-E.

The waiting period required to accumulate a certain number of coincident AMS-ASC pairs for TC is considerably longer than that for individual satellite records, ~7 months for 100 samples and more than ~2.75 years for 500 samples. For a two-data estimator that uses modeled SM as the second data, the waiting period for AMS is markedly shorter at ~3 months for 100 samples (or ~13 months for 500 samples) while ASC requires

Table 1. Setup of the Four Synthetic Experiments[a]

Ex#	N	σ_u (m^3m^{-3})	σ_v (m^3m^{-3})	α, β	m	ρ
1	366	Variable	$0.5\sigma_y$	[0.5,2]	1	0
2	Variable	0.04	0.04α	[0.5,2]	1	0
3	366	$0.5\sigma_x$	$0.5\sigma_y$	[0.5,2]	Variable	0
4	366	$0.5\sigma_x$	$0.5\sigma_y$	[0.5,2]	Variable	Variable

[a] N is the sample size, σ_u and σ_v are the standard deviation of the errors, α and β are the scaling factors, m is the delay of the lag variable, and ρ is the AR(1) coefficient of the errors in the reference data X.

~4.5 months and ~2 years, respectively. Larger samples available within shorter time windows would facilitate investigation of the temporal variability in the estimated parameters in a multiseasonal and multiannual data record.

This analysis is also relevant to other current and upcoming satellite missions for SM observations. AMSR-E has been succeeded by AMSR-2 (Advanced Microwave Scanning Radiometer 2) with similar specifications having a revisit time of 1–2 days and a swath limit of 1450 km. The current Soil Moisture and Ocean Salinity and upcoming Soil Moisture Active and Passive have lower temporal resolution with a 3 day revisit time and swath widths of ~1000 km. The next section will examine the possible impact of this infrequent sampling on estimation.

5. Synthetic Experiments

Synthetic experiments are used to illustrate the merits and limitations of the proposed LV method and to give a measure of the best possible outcomes under different scenarios. Other estimators were also implemented for comparison. Three erroneous representations $\{X,Y,Z\}$ of SM were synthetically generated from the in situ observations used as x, and the estimators were applied to determine the relationship between X and Y.

5.1. Methods

The data were prepared as follows. INS was subsampled at 0 h (local time) to produce a daily N-length time series from January 2007 to May 2011 and filtered to remove the seasonal cycle and create a time series of SM anomalies using a $W \sim 31$ day moving window

$$x_t^{(a)} = x_t - \frac{1}{(W+1)} \sum_{i=t-W/2}^{t+W/2} x_i. \tag{27}$$

This is in contrast to the previous TC studies of *Miralles et al.* [2010] and *Draper et al.* [2013] where anomalies from the climatological seasonal cycle were used. Given that Australia experienced dry (1995–2009) and wet years (2010–2011) during this data period, the latter approach is less favorable as the seasonal cycle is not entirely removed. The final synthetic data were generated by adding white noise of variances σ_e^2 to randomly rescaled SM anomalies, in accordance to equations (1)–(3), (11), and (12).

Four experiments (Ex) were performed with setups summarized in Table 1: Ex1 considers the effect of decreasing SNR_X (increasing error in X); Ex2 the effect of sample size by varying N; Ex3 the effect of using extended lag m for the LV estimator; and Ex4 the influence of autocorrelated errors on the LV. With the exception of Ex2, a 1 year (2007) daily data period is used. X_{t-m} is used as the instrument in LV estimation.

The merits of the estimators are judged based on their ability to retrieve true scaling coefficient α for each of the 1000 realizations of synthetic data sets generated with randomized c and α values. Boxplots are used to summarize the ensemble statistics of the estimation, in terms of the ratio of estimated to true values $\hat{\alpha}/\alpha$. The bias of the estimator is represented by the median of the ensemble, whereas its sampling error is related to both the interquartile range (IQR) of the ensemble and the standard error (SE) calculated using equation (18b). The SE for VAR is computed by propagation of var[var(p)] via partial derivatives (following equation (19)). However, SE is not an accurate measure of error when the estimator is biased. An ideal estimator is characterized by a unit median and, in the limit of a large sample size N, zero IQR, and SE.

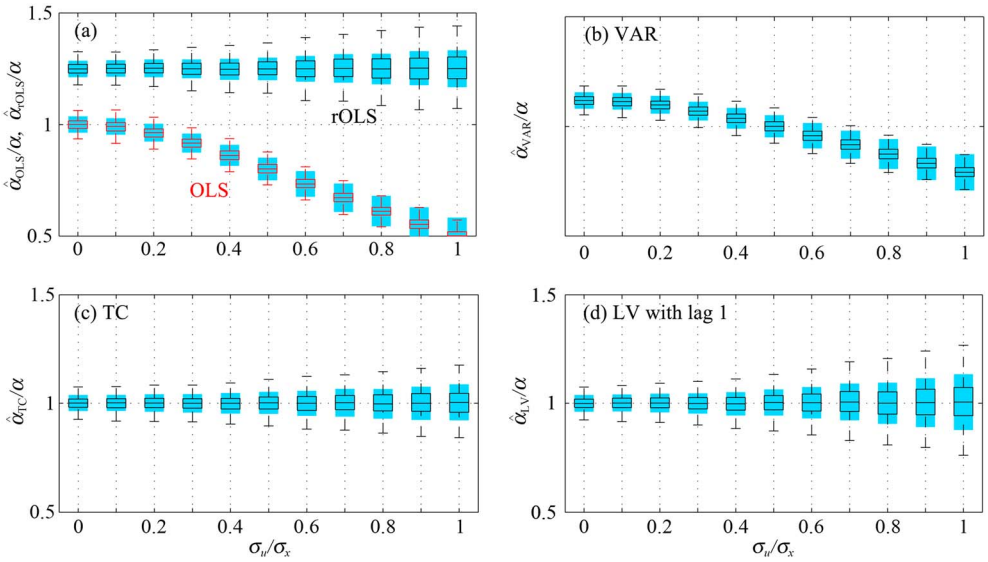

Figure 2. (a–d) Boxplots summarizing the performance of the various methods for estimating the scaling coefficient α between synthetic SM anomaly data as a function of SNR_X of X (Ex1). The central mark is the median, and the edges of the box are 25th and 75th percentiles. The standard errors are superimposed as blue shading.

5.2. Results and Discussion

5.2.1. Ex1 in Figure 2

This experiment provides an overall picture of the four linear estimators. Figure 2a illustrates the negative error-in-variables bias of the OLS estimates (red boxplots), which increases with increasing error in X. By contrast, the rOLS estimates (black boxplots) are independent of SNR_X but are positively biased because $SNR_Y = 1/2$. Figure 2b shows the interplay of the SNR of X and Y on the VAR estimator. When $SNR_X < SNR_Y$, the VAR estimates are positively biased, and the crossover to negative biases occurs at $SNR_X = SNR_Y$. In contrast, the TC and LV methods generate superior, unbiased estimates, as shown in Figures 2c and 2d. These simulation results are consistent with the discussions in section 2.1. Sampling error exists in LV and TC estimates, and they increase with measurement error in the data. The LV estimator shows marginally larger IQR and SE than TC due to the reduced correlation with the lagged variable.

5.2.2. Ex2 in Figure 3

As a consequence of the similar statistical equations, TC and LV show similar dependency on the sample size (not shown). Ex2 examines on the possible influence of infrequent remote sensing observations in practice by discarding parts of the synthetic data to mimic the temporal sampling patterns of AMSR-E and ASCAT at the

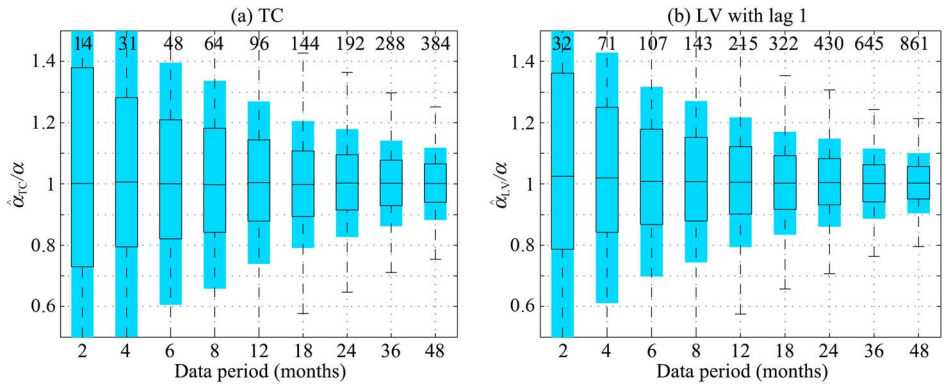

Figure 3. Boxplots summarizing the performance of TC and LV on synthetic SM anomaly data as a function of a variable data sampling period (Ex2). The notations follow Figure 2, with additional numbers (top) showing the average sample size N for each period length.

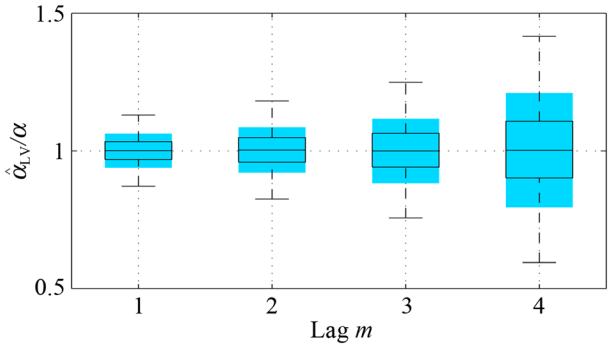

Figure 4. Boxplots summarizing the performance of LV on synthetic SM anomaly data, using variable lag m (Ex3). The notations follow Figure 2.

Kyeamba station. For TC, the data were screened based on coincidence of the two sensors; whereas for LV, its inclusion is based only on the presence of AMSR-E. The data were then randomly sampled with variable window sizes, which vary the effective sample size available for TC and LV analyses. As already noted in section 4, the sample sizes available for LV are significantly greater than that for TC. As a result, the sampling errors for LV are relatively smaller than for TC. For instance, a 6 month sampling window yields SE = 0.31 for the LV estimates, compared with 0.40 for TC. A 2 year period is needed for LV (cf. ~3 years for TC) to reduce the SE below the 0.15 threshold.

5.2.3. Ex3 in Figure 4

These results illustrate the effects of using a longer lag in LV. A longer lag does not introduce any bias, but it increases the sampling errors due to the diminishing autocorrelation of the instrument. The SE for $m = 3$ (i.e., 3 day lag) remains relatively low at 0.12, cf. 0.06 at $m = 1$. If such a level of SE at $m = 3$ is deemed acceptable, the use of multiple lags may be reasonable for irregularly sampled data sets, although this depends on the autocorrelation characteristics of the local SM anomalies. The LV estimator is degraded by the occurrence of rainfall wetting during the time period between t-m and t. To overcome this, rainfall data, e.g., 3-hourly rainfall fields from Tropical Rainfall Measuring Mission [*Huffman et al.*, 2007], can be used to identify and omit such instances from estimation, in conjunction with seasonal analysis to avoid sampling bias toward dry seasons. This is also relevant to TC when applied to noncoincident samples. For instance, the Aqua and MetOp-A satellite overpasses were 3–4 h apart. If the TC analysis includes triplets where rainfall events occur during the time lapse, the estimated error in AMSR-E could be artificially inflated.

5.2.4. Ex4 in Figure 5

This experiment considers the influence of error on LV, which is modeled using equation (25) as an AR(1) process with an autocorrelation coefficient ρ. The result illustrates an interplay between the autocorrelation in the signal of X and that in its error. Such an error leads to a negatively biased estimator with the degree of bias increasing with ρ (equation (24a)), as illustrated in Figure 5a for $m = 1$. In subsequent simulations with longer lagged variables at $m = 2$ and 3, the biases for larger ρ are visibly reduced, but the tradeoff is an increase in the sampling errors as observed in Ex3.

To conclude, in comparison to OLS and VAR estimators, the LV and TC estimators are unbiased but conditional on having serially uncorrelated errors. The experiments suggest that LV may be a suitable alternative to the TC when the latter suffers from small-sample bias or when three data sets are unavailable. While LV suffers from autocorrelated errors, this can be partly addressed by increasing the lag of the instrument.

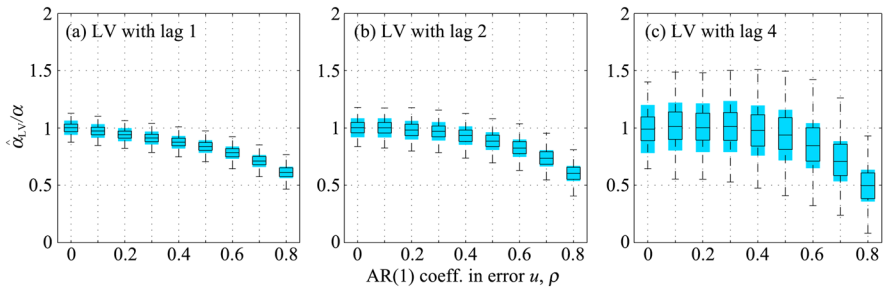

Figure 5. (a–c) Boxplots summarizing the performance of LV with variable lag m, acting on synthetic SM anomaly data in the presence of autocorrelated errors in X (Ex4). The notations follow Figure 2.

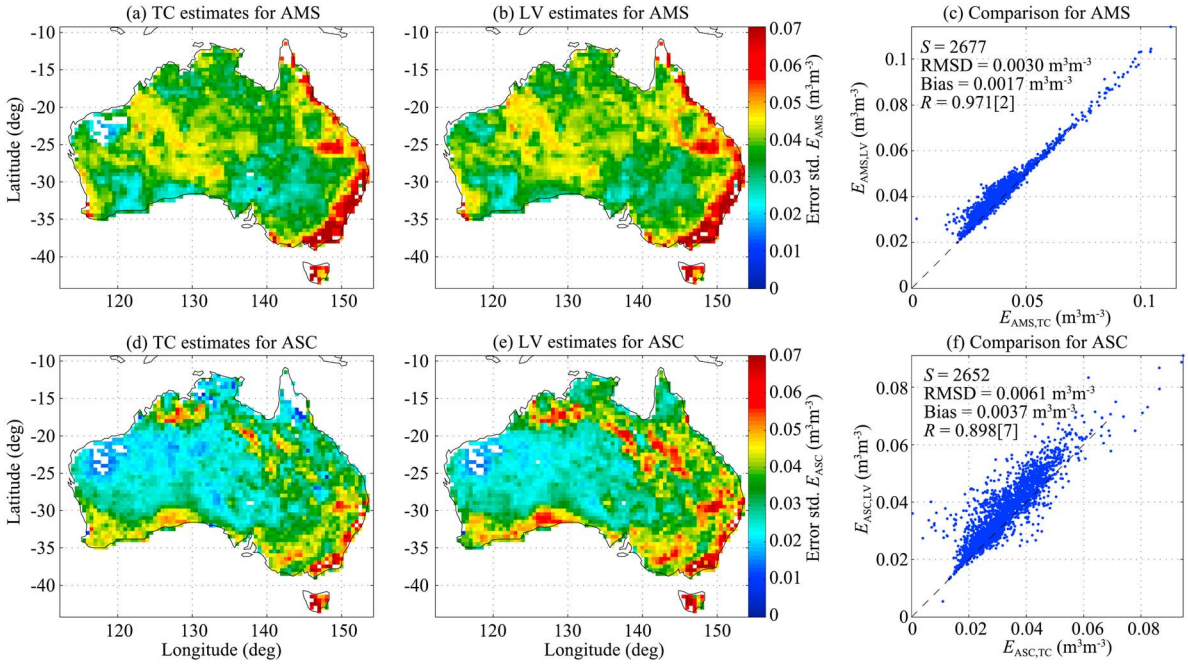

Figure 6. Comparisons of stochastic error standard deviation of (a–c) AMS and (d–f) ASC SM anomalies, derived using TC and lag-1 (half-day lag) LV estimators, over Australia. The color bars are truncated at 0.07 m^3m^{-3}.

6. Error Characterization of Satellite Soil Moisture

In this section, we demonstrate the use of the LV estimator to characterize the spatial error structures of AMS and ASC SM anomalies over Australia. TC estimates are used as the benchmark for comparison.

6.1. Methods

AMS, ASC, and MER SM were resampled spatiotemporally to a half-daily temporal grid and a regular 1/2° × 1/2° spatial grid for consistency. The estimation period is January 2007 to October 2011, corresponding to the coincident period of the three data sets. AMS and ASC were upscaled spatially using equal-weighted averaging, while MER was spatially resampled using nearest-neighbor resampling and temporally subsampled at 11.30 A.M./P.M. LST to approximately match satellite overpass times. The anomalies of each data set were computed separately using 31 day moving averages of equation (27).

In LV estimation, the error standard deviation E in AMS and ASC SM anomaly data were estimated individually, with MER SM anomalies as the second data and the lagged instrument. The use of MER as the IV takes advantage of the high temporal resolution and availability of the model data. The errors in daytime and nighttime SM data were estimated both jointly and separately, using variable lags $m \in [1, 4]$ (1/2–2 day lags) of MER. These results are validated against equivalent estimates from TC analysis of AMS, ASC, and MER triplets. To ensure direct comparison, only the triplets used in the TC analysis were used in the LV analysis, and both analyses used the error estimators of equation (5) so that their disagreement is solely due to differences in their estimates of α. Moreover, pixels were screened for nonphysical results, namely, negative covariances due to poor correlations between the data sets and negative error variances due to very low errors or possible cross correlation of errors.

For each satellite product $p = \{\text{AMS,ASC}\}$, the differences between spatial maps of the two estimates of error standard deviation E_p are measured by Pearson's linear correlation $R(\hat{E}_{p,\text{LV}}, \hat{E}_{p,\text{TC}})$, RMSD, and bias defined by

$$\text{RMSD}(\hat{E}_{p,\text{LV}}, \hat{E}_{p,\text{TC}})^2 = E_S\left[(\hat{E}_{p,\text{LV}} - \hat{E}_{p,\text{TC}})^2\right], \tag{28a}$$

$$\text{Bias}(\hat{E}_{p,\text{LV}}, \hat{E}_{p,\text{TC}}) = E_S(\hat{E}_{p,\text{LV}}) - E_S(\hat{E}_{p,\text{TC}}), \tag{28b}$$

where $E_S(\circ)$ denotes spatial averaging over S spatial points compared and $\hat{E}_{p,\text{LV}}$ and $\hat{E}_{p,\text{TC}}$ refer to estimates by LV and TC, respectively. The 95% confidence interval (CI) of the bias is given by

Table 2. Statistics Comparing TC and Lag *m* LV Estimated Errors for AMS and ASC SM Anomalies Over Australia[a]

Configurations		AMS				ASC			
SM	*m*	*S*	RMSD	Bias	*R*	*S*	RMSD	Bias	*R*
Day	1	2578	0.0055	0.0039[6]	0.935[5]	2096	0.0130	0.0105[7]	0.762[18]
	2	2564	0.0053	0.0036[6]	0.934[5]	2082	0.0128	0.0100[7]	0.747[19]
	3	2335	0.0060	0.0038[6]	0.904[7]	1953	0.0143	0.0109[8]	0.715[22]
	4	2248	0.0059	0.0035[6]	0.900[8]	1889	0.0148	0.0114[8]	0.699[23]
Night	1	2619	0.0027	−0.0005[6]	0.970[2]	2600	0.0047	0.0008[7]	0.912[7]
	2	2610	0.0072	−0.0056[6]	0.918[6]	2562	0.0066	0.0037[7]	0.906[7]
	3	2585	0.0109	−0.0087[6]	0.847[11]	2457	0.0063	0.0034[7]	0.913[7]
	4	2505	0.0115	−0.0083[7]	0.801[14]	2017	0.0103	0.0070[8]	0.859[11]
Day, Night	1	2677	0.0030	−0.0017[5]	0.971[2]	2652	0.0061	0.0037[7]	0.898[7]
	2	2676	0.0031	−0.0002[6]	0.956[3]	2629	0.0072	0.0051[7]	0.900[7]
	3	2667	0.0048	−0.0017[6]	0.911[6]	2542	0.0078	0.0054[8]	0.897[8]
	4	2655	0.0052	−0.0014[6]	0.881[8]	2393	0.0098	0.0075[7]	0.885[9]

[a] *S* is the number of spatial samples compared, RMSD and Bias are in units of m^3m^{-3}, and *R* is Pearson's correlation. The square brackets contain 95% CI, e.g., 0.0039[7] corresponds to 0.0039 ± 0.0007.

$\sqrt{(z^2/S)\left[\mathrm{var}_S\left(\hat{E}_{p,\mathrm{LV}}\right) + \mathrm{var}_S\left(\hat{E}_{p,\mathrm{TC}}\right)\right]}$ with a critical value $z = 1.96$, where $\mathrm{cov}_S(\circ)$ denotes an operator that computes the spatial variance of the error maps. The 95% CI of *R* is calculated using Fisher's transform.

6.2. Results and Discussion

Figure 6 shows the LV-versus-TC error comparison for daytime and nighttime satellite SM anomalies, where lag $m = 1$ (0.5 day) MER SM anomalies was used in LV. The pairs of TC and LV error maps show very similar spatial patterns, which is confirmed by the metrics (see also Table 2). The estimates of errors at over 2650 pixels show strong linear correlation of 0.97 (AMS) and 0.90 (ASC). The differences in values are measured by RMSD = 0.003 and 0.006 m^3m^{-3}, which are largely due to positive biases in LV compared with TC of 0.002 and 0.004 m^3m^{-3}, respectively. The results for other configurations are summarized in Table 2. TC and 0.5 day LV estimates show good agreement for individual daytime and nighttime SM anomalies, with *R* generally > 0.91 and RMSD < 0.006 m^3m^{-3}. The largest discrepancy is associated with daytime ASC SM anomalies, which have lower $R = 0.76[2]$ (i.e., 95% CI of ±0.02) and higher RMSD = 0.013 m^3m^{-3}. This suggests that jointly using day and night ASC data in LV will improve agreement with TC.

When using longer lags of up to 2 days, the LV estimator remains capable of recovering the TC-derived error maps for SM anomalies with $R = 0.70$–0.96 and RMSD = −0.0002 to 0.0114 m^3m^{-3}. Contrary to the observations in the synthetic Ex4, biases remain for these cases, perhaps due to persistence in the autocorrelations in the MER error.

It should be noted that as spatial points available for each comparison are variable, Table 2 does not allow direct comparisons. The estimates were therefore screened for common locations (not shown). With the 2054 common pixels identified for AMS, we find that with increasing lag, *R* decreases from 0.969[3] to 0.88[1] and RMSD increases from 0.0032 to 0.0050 m^3m^{-3} for combined day and night SM. For ASC, the differences at 1386 common pixels are marginal with $R = 0.90[1]$ and RMSD = 0.0060 m^3m^{-3} for $m = 1$, cf. 0.89[1] and 0.0097 m^3m^{-3} for $m = 4$.

Lastly, for completeness, the use of lagged values of the AMS and ASC SM anomalies as IVs has also been considered (not shown). In general, we found poorer agreement with TC for all metrics because the signal autocorrelation in the satellite SM anomalies is weak, leading to nonphysical error estimates (negative variances) at many locations.

In summary, the results support the utility of the LV estimators and the use of modeled SM as the instrument for estimating the errors in satellite SM anomalies. This avoids the use of three distinct SM data sets required in typical TC implementations. While the estimator with lags of 0.5–2 day delays shows very good agreement with TC, the likely presence of errors in MER led to observed (small, mainly positive) biases in the LV estimates.

7. Time-Varying Characterization of Soil Moisture

The benefit of parameter estimation with two data sets is most apparent with remote sensing and in situ data, where spatially and/or temporally coincident observations are usually limited at the global scale. The limited

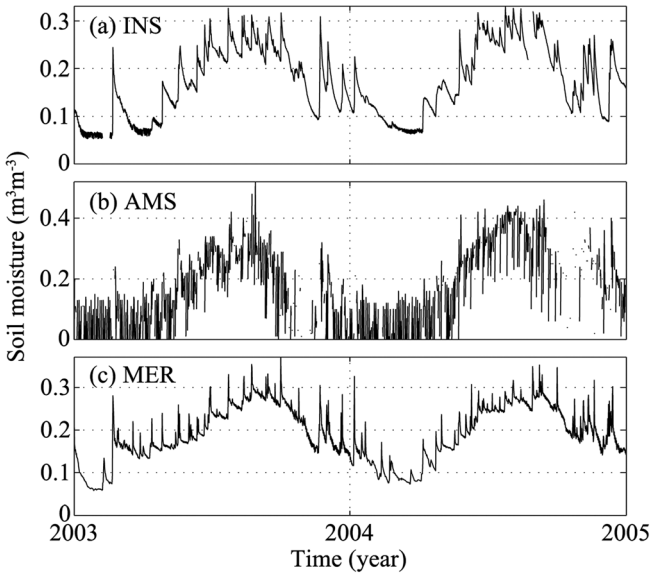

Figure 7. INS, AMS, and MER SM time series at Kyeamba.

spatiotemporal overlap of satellite sensors on different orbital tracks (section 4) can introduce small-sample bias and large sampling errors (section 5) in TC, limiting our ability to study the nonstationary relationships between data sets. LV provides an alternative that overcomes this limitation by using modeled data as both the second data and the instrument. So far, our analyses have assumed homogeneity of the errors and an affine structure between the SM data sets for the entire estimation period. This section relaxes this constraint and examines the time-varying characteristics of the relationships between different SM data sets.

7.1. Methods

This analysis uses the daytime and nighttime INS, AMS, and MER SM time series at Kyeamba from June 2002 to December 2010 (partly illustrated in Figure 7). The data with different resolutions are colocated via nearest neighbor.

We first consider an ideal scenario where ground, remote sensing, and modeled SM data are available for joint analyses. LV is used to examine the AMS-INS relationship and the MER-INS relationship separately, using lag-1 (0.5 day) INS as the instrument. The relationships are characterized by additive (B) and multiplicative (M) biases, scaling coefficient a, and stochastic error standard deviation E, which provide the makeup of RMSD in equation (4). RMSD, B, M, and α are defined with the in situ data as the reference, while stochastic error (standard deviation) E is defined independent of a reference. The absolute deviation of estimated α from unity mirrors the presence of the bias M.

For comparison, TC was implemented using the three data sets simultaneously to provide alternative estimates of the two relationships. The setup of TC between INS, AMS, and MER is similar to the work of *Miralles et al*. [2010] for estimating point-to-footprint SM sampling errors. The standard two-data estimators (OLS and VAR) were also applied to AMS-INS and MER-INS pairs separately to estimate α. To develop a complete picture of these relationships, estimations are made for both the (original) SM time series of Figure 7 and SM anomalies. The anomalies of each data were computed separately from 31 day moving averages using equation (27).

We also consider a typical scenario where only remote sensing and modeled SM data are available for analyses such as the case in section 6. We investigate the utility of LV in reproducing the TC estimates of AMS-MER SM anomaly relationships with lag-1 MER as the instrument. MER is the reference data for computing B, M, and α in this case.

To sum up, we investigate the following relationships, (A1) AMS and INS SM, (A2) MER and INS SM, (A3) AMS and INS SM anomalies, (A4) MER and INS SM anomalies, and (A5) AMS and MER SM anomalies. In all the analyses, statistics are sampled separately from the multiyear data set using a 60 day sampling window centered on a particular day of the year (DOY), thereby allowing for nonstationary statistics. The window size is chosen so that all data samples are sufficiently large (~500 to 700) to reduce sampling error to acceptable levels as observed in section 4. This approach assumes that the temporal variation of the statistics is seasonal. The prospective application of such an analysis is to produce historical DOY (or month of the year) based statistics for dynamical bias correction of observational data for data merging and validation or for specifying time-varying error variances for data assimilation. The error E estimates were

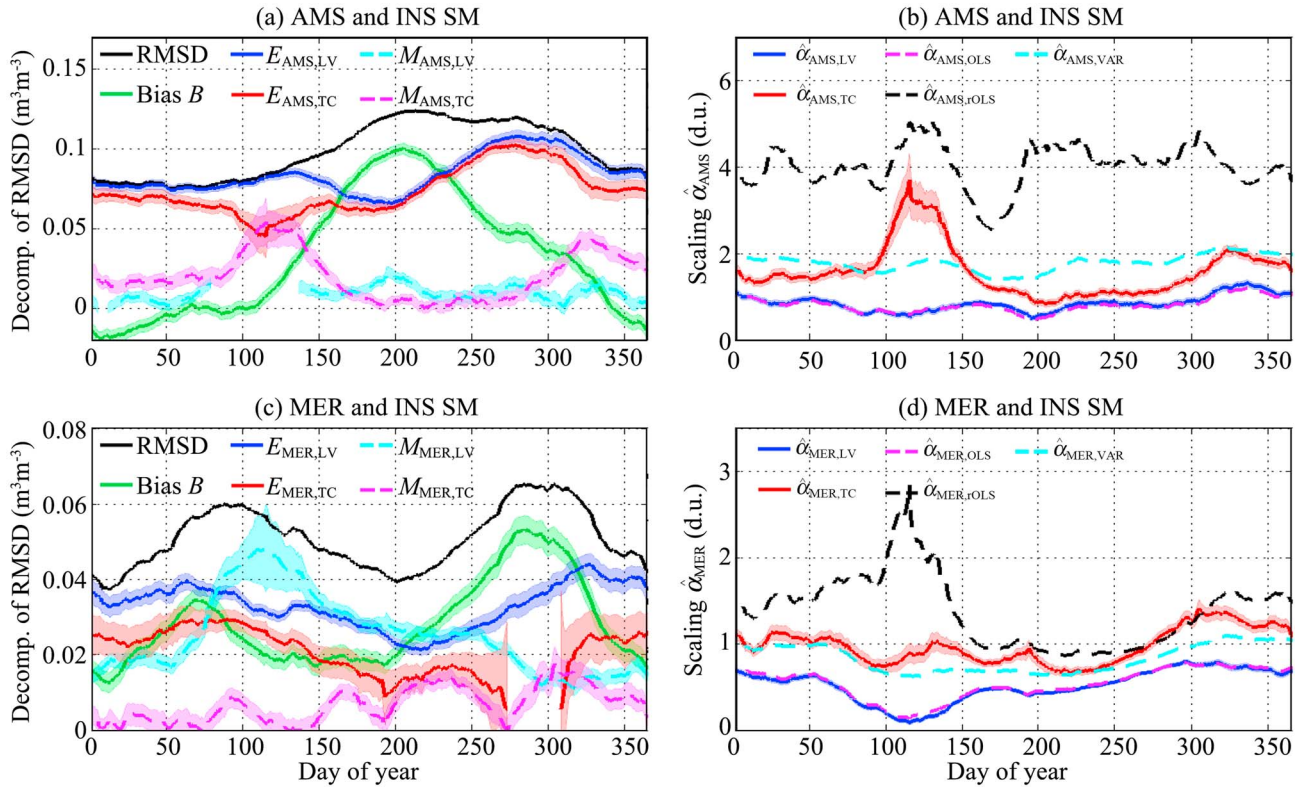

Figure 8. Estimation of relations between (a, b) AMS and INS SM (A1) and between (c, d) MER and INS SM (A2) at Kyeamba (Figure 7) with a 60 day moving window centered at DOY. Figure 8a and 8c show the respective decompositions of RMSD with respect to INS into bias estimates *B* and *M*, and stochastic error estimates *E*; Figures 8b and 8d plot scaling estimates α with respect to INS. SE for LV and TC estimates are shown by the shading around their curves. The subscripts of these terms label the associated estimator used.

also screened for negative or very low values less than 5% of the total variance of the SM or SM anomaly time series. The α and *M* estimates were also screened for weak correlations between the data sets using the lesser of a simple threshold of 0.1 and error variance > 95% of the total variance. The SEs of the TC and LV estimates were also computed to provide measures of uncertainty to aid interpretation.

7.2. Results and Discussion

7.2.1. A1 in Figures 8a and 8b

We focus first on the LV estimation of AMS-INS SM. The RMSD and its constituents show temporal variability in Figure 8a, mainly due to a seasonal cycle in bias *B* and error $\hat{E}_{AMS,LV}$. *B* is larger during the wet (high moisture) months of June–August (DOY ≈ 150–240) and the period of larger $\hat{E}_{AMS,LV}$ occurs during a later time period of September–November (spring, DOY ≈ 240–334). This is the period with maximum vegetation density, which is consistent with previous validation works that find satellite retrieval errors degrade with vegetation [*Brocca et al.*, 2011, and references therein]. While the bias *M* also shows temporal variation, its contribution to the overall difference between AMS and INS is relatively small and is accompanied by small deviations of $\hat{\alpha}_{AMS,LV}$ from unity. Moreover, the strong lag-1 autocorrelation in the in situ time series with $R(\mathrm{INS}_t, \mathrm{INS}_{t-1}) \sim 0.96$–$0.98$ leads to close correspondence between $\hat{\alpha}_{AMS,LV}$ and $\hat{\alpha}_{AMS,OLS}$ (Figure 8b) and small estimates of E_{INS} with $E\left(\hat{E}_{INS,LV}\right) = 0.011\ \mathrm{m^3m^{-3}}$ (not shown), because $\mathrm{cov}(\mathrm{AMS}_t, \mathrm{INS}_{t-1}) \approx \mathrm{cov}(\mathrm{AMS}, \mathrm{INS})$ and $\mathrm{cov}(\mathrm{INS}_t, \mathrm{INS}_{t-1}) \approx \mathrm{var}(\mathrm{INS})$.

Comparing LV and TC estimates of E_{AMS}, they are similar in trend and/or values for the majority of the year. The significant exception occurs during DOY ≈ 100–150 where TC yielded higher *M* and α estimates. Such a subseasonal and significant change in the $\hat{\alpha}_{AMS,TC}$ time series is not sensible. Based on Figure 7, the SNR in INS is better than in AMS, so the VAR estimates (as well as rOLS) serve as an upper bound for the true value of α but $\hat{\alpha}_{AMS,TC} > \hat{\alpha}_{AMS,VAR}$ is observed. We attribute this phenomenon to the weak correlation $R(\mathrm{INS,MER}) \approx 0.2$ during this period, which is at the minimum of the dynamic range in correlation of ~0.2–0.8. This instance highlights the practical difficulty of using IV techniques (TC or LV)

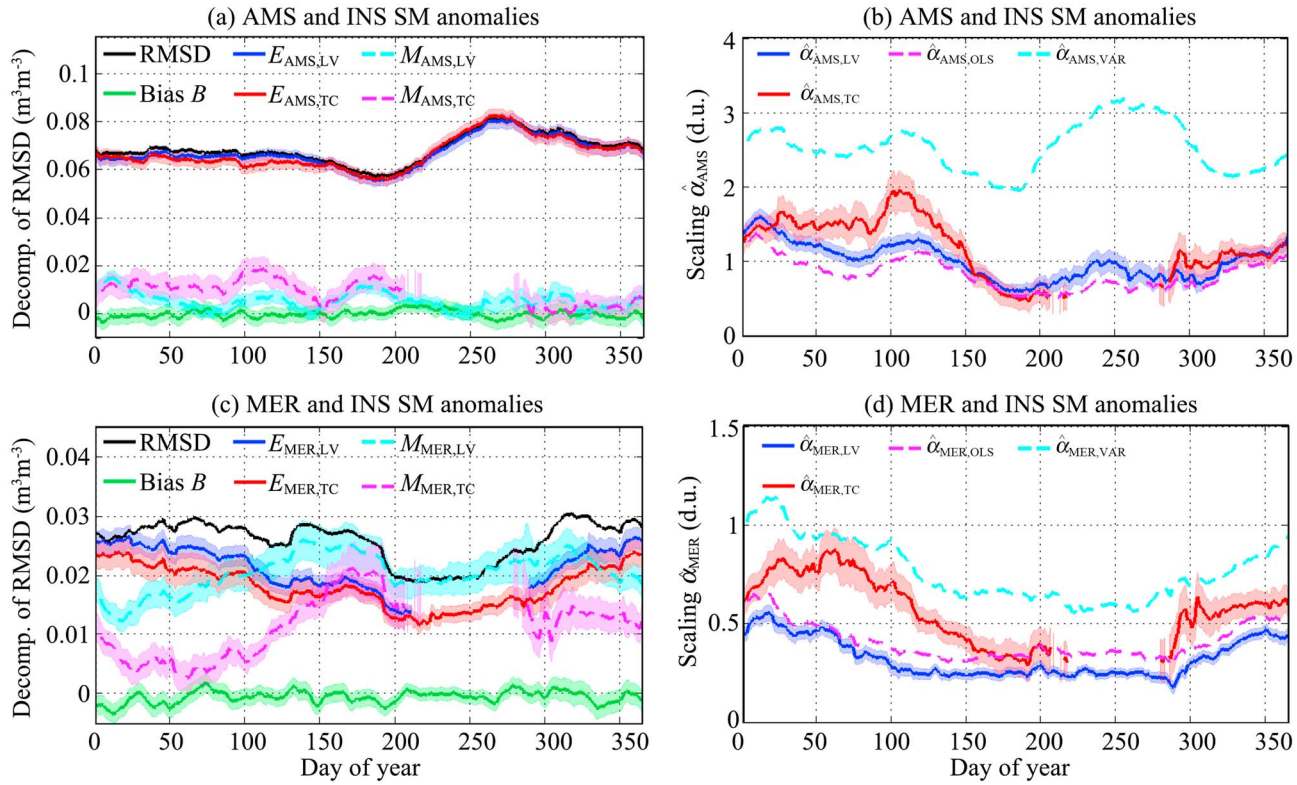

Figure 9. Same as Figure 8, but for (a and b) AMS and INS SM anomalies (A3) and (c and d) MER and INS SM anomalies (A4). rOLS estimates are excluded because they are out of the plot range.

when the chosen instrument (MER, in this case) explains little of the variation in the INS of the relationship [*Bound et al.*, 1995].

The other difference between LV and TC estimates is related to their estimates of E_{INS} (not shown), namely, $E(\hat{E}_{INS,LV}) = 0.011\ m^3m^{-3}$, cf. $E(\hat{E}_{INS,TC}) = 0.035\ m^3m^{-3}$. The quality-assured in situ data is of high fidelity with the measuring instruments having calibration accuracies RMSD of the same order of magnitude [*Smith et al.*, 2012]. However, point-to-footprint sampling errors may be larger [*Miralles et al.* 2010]. In particular, the TC analysis with three SM products with distinctive spatial supports will lead to the error estimates of the three data containing some scale-related sampling error. For instance, signal detected in AMS and MER at coarser scales but not by in situ probe is counted as error in INS and as signal of AMS and MER. Similarly, signal common to AMS and INS at finer scales will be added to error of MER. By the same argument where the resolved common signal in INS and its lagged values are regarded as errors in AMS, LV analysis is likely to cause a bias in favor of the lagged instrument leading to observed $\hat{E}_{AMS,LV} > \hat{E}_{AMS,TC}$ and $\hat{\alpha}_{AMS,LV} \approx \hat{\alpha}_{AMS,OLS}$.

7.2.2. A2 in Figures 8c and 8d

RMSD between MER and INS SM also shows a seasonal cycle that is partly related to positive bias *B* and error in MER. $\hat{E}_{MER,TC}$ was masked during DOY ≈ 270–310 due to high SNR in MER. As previously stated, TC produced higher E_{INS} estimates while LV yielded larger bias (*M*) estimates, resulting in different RMSD decompositions. Indeed, $\hat{\alpha}_{MER,TC} > \hat{\alpha}_{MER,VAR}$ indicates that TC regards the SNR in INS to be lower than MER, whereas $\hat{\alpha}_{MER,LV} \approx \hat{\alpha}_{MER,OLS}$ indicates that LV regards INS to be nearly error free. Despite these differences, their estimates of E_{MER} show a similar seasonal trend, with larger errors during summer and autumn months of DOY ≈ 300–150, and the bias *B* has two maxima around DOY = 70 and DOY = 270. Notably, the differences between AMS and MER in their physical representation manifest in different temporal dynamics of the biases and errors.

7.2.3. A3 in Figures 9a and 9b

The AMS-INS SM (A1) analysis was repeated with their SM anomalies. The additive biases for the anomaly data are essentially removed, but the errors and scaling are still variable with time. The stochastic errors in

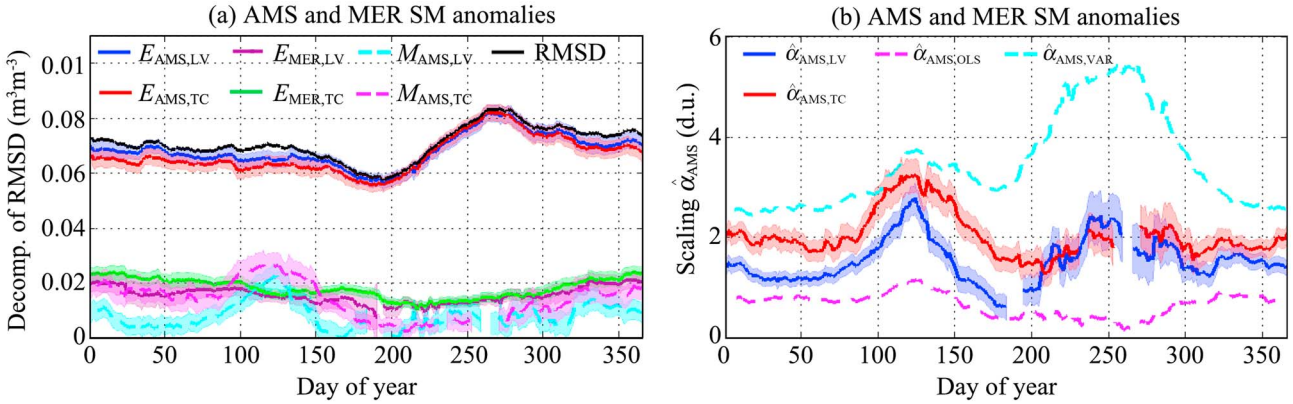

Figure 10. Estimation of relations between AMS and MER SM anomalies (A5) using various estimators with a 60 day moving window. TC estimates are taken from Figure 9. SE for LV and TC estimates are shown by the shading around their curves.

AMS SM anomalies show similar seasonal cycles to A1, suggesting that the stochastic errors reside in the high-frequency components of the SM time series that are preserved in the SM anomalies. The makeup of RMSD is primarily from E_{AMS}, with relatively little contribution from bias M. There are similarities between the TC and LV estimates, but TC yielded greater M and α estimates. TC estimates of α over DOY = 210–280 were masked due to low R(AMS,MER) ~ 0.03–0.3, which without screening will lead to biased $\hat{\alpha}_{AMS,TC} < \hat{\alpha}_{AMS,OLS}$. On the other hand, LV counterparts are marginally larger than the OLS estimates due to low measures of E_{INS} as observed in A1.

7.2.4. A4 in Figures 9c and 9d

The analysis of MER-INS SM anomalies shows that LV and TC estimates of E_{MER} are in good agreement. In contrast to AMS in A3, both the stochastic errors and bias M are equally prominent in the makeup of RMSD. Nevertheless, LV and TC differ in their estimates of M, α, and E_{INS}. Since OLS estimates serve as the lower bound for the truth, $\hat{\alpha}_{MER,LV} < \hat{\alpha}_{MER,OLS}$ appears problematic. Closer inspection of the covariances involved reveals that the large reduction from R(MER,INS) to $R(\mathrm{MER}_t,\mathrm{INS}_{t-1})$ of ~0.16 on average and relatively weak $R(\mathrm{MER}_t,\mathrm{INS}_{t-1})$ of ~0.2–0.5 are responsible for the incorrect estimates by LV.

7.2.5. A5 in Figure 10

RMSD between AMS and MER SM anomalies is largely made up of stochastic errors in AMS and MER. The LV and TC estimates of these errors are again in good agreement, but TC produced higher estimates of bias M and α. Nevertheless, the LV and TC M estimates share a similar trend, showing larger biases during April–May (DOY ≈ 80–150), while their α estimates respect the bounds provided by OLS and VAR. OLS estimates are lower than the truth because of errors in MER, while VAR are higher because SNR in MER is higher than in AMS. Hence, TC and LV estimates of α would both appear reasonable.

The decomposition of RMSD and its (estimated) constituents provides a fuller description of the difference between pairs of SM data. Regardless of the chosen estimator used in SM analyses (A1 and A2), the large variability observed in B, E_{AMS}, and E_{MER} over significant ranges are problematic if a homogeneous affine model is applied to the multiseasonal time series. One resolution for this is to apply the estimators to SM anomaly time series [*Miralles et al*. 2010; *Draper et al*., 2013]; however, our anomaly analyses (A3–A5) show that this only provides a partial remedy. All the estimators found temporal variability in errors and bias M remain in the SM anomalies. This supports the current and conservative strategy of applying estimators to multiseasonal SM data in a windowed fashion, based on DOY or month of year, although interannual variability is not taken into consideration by this approach.

Our assessments of the various two-data estimators are consistent with our theoretical discussions in section 2 and intuitions from viewing the individual time series in Figure 7. OLS provides low estimates, rOLS results in overestimates, and VAR reflects the relative SNR in INS, AMS, and MER. For the TC-versus-LV comparison, the general impression is that the estimation is sensitive to the choice of the instrument, whether it is a third measurement record or a lagged variable. The agreement between LV and TC estimates appears better in the SM anomaly analyses (A3–A5) than in the SM analyses (A1 and A2). It is, however, less conclusive about the absolute performance of LV and TC. A1–A4 show that they differ significantly over the estimation of the error in

in situ data, leading to contrasting identification of multiplicative biases. LV with ground data as the instrument has likely underestimated errors in the point-scale SM relative to the coarse-scale SM because of strong correlation between INS and its lagged values due to shared information at fine spatial scales. In turn, it overestimates errors in mesoscale SM relative to point-scale data. It can be argued that LV and TC estimates may not be necessarily compatible for direct comparison because different points of reference are adopted by TC and LV. This would be closely related to the chosen interpretation of the nature (autocorrelated error or shared signal) underlying the autocorrelation in the lagged variable. While this may be contentious, examination of the disparity between TC and LV may provide further insights into point-to-footprint sampling errors because LV has reduced complexity due to involving only two representation supports, instead of three in TC.

The A1 and A3 analyses with AMS SM and its anomalies have shown unrealistic α and E estimates by TC when correlations R(INS,MER) and R(AMS,MER) are too weak. MER is found to be a poor instrument in some instances, particularly A4 showing inconsistency in the LV-estimated α due to weak $R(\mathrm{MER}_t,\mathrm{INS}_{t-1})$. These observations highlight the need to address broader questions surrounding the relevance of the chosen instrument in TC and, when data record is short, finite sample performance.

The final caveat is that the above interpretations are based on the use of linear estimators with the affine signal model and additive error model of section 2, and thus, their correctness hinges on the validity of these assumptions. We have already observed apparent discrepancies in α estimates between SM and SM anomaly analyses, suggesting nonlinearity where the low-frequency components in SM time series are scaled differently from the high-frequency components retained in the anomaly data. Within the 60 day window, SM is influenced by various processes (rainfall, loss via evapotranspiration, and infiltration) with different characteristic times. It is conceivable that such nonlinearity also exists within 60 days. This gives cause for going beyond linear methods, possibly including polynomial regression (equation (20)), the multiresolution analysis of *Su et al*. [2013d], and extensions to methods such as CDF matching [*Reichle and Koster*, 2004].

8. Conclusions

Global environmental monitoring requires geophysical measurements from a variety of sources to close the information gap. The relationships between comparable data sets must therefore be established to create complete and consistent data records. This is nontrivial when there are measurement errors and relationships between data sets that vary with time and/or are possibly nonlinearity. Where linear approximation is invoked, the standard two-data linear estimators such as OLS and statistical moment matching are suboptimal because the errors are not resolved. TC resolves this problem but may be limited in practice by sample size requirements due to the need for triplets of coincident observations.

In the first part of this paper, we identified the conceptual association between TC and the method of IV regression, which encompasses TC as one of many possible realizations of IV estimators. Other prospective IV implementations can involve other geophysical variables or latent variables, conditional on satisfying equation (16). As a standard method in the econometric toolkit, the rich body of knowledge on IV can be leveraged to inform TC's asymptotic and, more importantly, finite-sample properties. For instance, the method provides natural extensions to cases of polynomial regression and multiple instruments.

The second part of the paper is concerned with the practical issue of data availability to propose a potentially useful two-data estimator. Under an ideal circumstance where the errors are not serially correlated, lagged variables are potential instruments that conceptually serve the role of the third data series in the TC. This lifts the three-data constraint of TC, making it easier to meet sample size requirements. Potential temporal variability in observational errors and interdata relations due to seasonal and interannual variability provide further constraints on the size of data sampling window, which also makes methods that maximize the use of measurements attractive. SM is an important variable that can benefit from the use of such a lag-based IV (LV) estimator, given limited coincident spatiotemporal coverage of remote sensing and in situ networks at present and in the foreseeable future.

Synthetic experiments based on surface SM were used to illustrate the properties of LV. With independent and uncorrelated errors, the LV estimator was shown to be unbiased. It is also resilient against weakly

autocorrelated errors when implemented with longer lagged variables. Next, the LV estimator was applied to characterize the spatial error structures of satellite SM products derived from spaceborne radiometer (AMSR-E) and scatterometer (ASCAT) observations over Australia. We found that error estimates derived from the LV method, which uses lagged values of modeled (MERRA-Land) SM as the instrument, show remarkably good agreement with the TC benchmark. However, the LV estimates were found to be slightly biased, probably due to the presence of autocorrelated errors in the modeled SM. This could perhaps be rectified if the autocorrelation structure of the model error can be characterized beforehand.

The affine signal model and additive error model provide a simple decomposition of RMSD into two stochastic error components, a multiplicative bias (i.e., related to scaling of the data), and an additive bias. By applying the LV in a moving window fashion to in situ satellite and modeled SM, we showed temporal variability in these four components and distinguished their relative contributions to their overall differences with respect to the in situ data. These results are of importance for evaluation studies of satellite and model data against ground observations. Also, in typical scenarios for data assimilation, the relationship between satellite and modeled SM needs to be estimated to correct the biases (with respect to the model) in the observations. This was demonstrated with AMSR-E and MERRA-Land. These results identified similarities between TC and LV estimates of errors, but differences in their estimates of multiplicative bias (and scaling) were found. It is clear that the choice of the instrument has a strong influence on the estimation outcomes. Some of which are attributed to shortcomings of TC and LV due to weak instruments, prompting closer examination to identify suitable instruments and poor finite sample performance in the presence of weak instruments. In general, these experiments illustrated structural changes in the affine model and errors over time, highlighting the need to rethink about the best strategies for applying linear estimators. The underlying assumptions of these estimators become invalid when sampling periods for estimation are not carefully chosen. Exploring alternative implementations of IV estimators and seeking nonlinear methods may provide other avenues to address these pressing challenges.

Acknowledgments

We thank three anonymous reviewers and Shelly Chua for their valuable comments on the manuscript. We are grateful to all who contributed to the data sets used in this study. Kyeamba in situ data were produced by Jeff Walker and colleagues at Monash University and the University of Melbourne who have been involved in the OzNet program. AMSR-E data were produced by Richard de Jeu and colleagues at Vrije University Amsterdam and NASA. ASCAT level 3 data were produced by the Vienna University of Technology within the framework of EUMETSAT's Satellite Application Facility on Support of Operational Hydrology and Water Management from MetOp-A observations. The MERRA-Land was provided by NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). National soil data were provided by the Australian Collaborative Land Evaluation Program, ACLEP, endorsed through the National Committee on Soil and Terrain NCST (www.clw.csiro.au/aclep). This research was conducted with financial support from the Australian Research Council (ARC Linkage Project LP110200520).

References

Bound, J., D. A. Jaeger, and R. M. Baker (1995), Problems with instrumental variables estimation when the correlation between instruments and the endogeneous explanatory variable is weak, *J. Am. Stat. Assoc.*, *90*(430), 443–450.

Bowden, R. J., and D. A. Turkington (1990), *Instrumental Variables*, Cambridge Univ. Press, Cambridge, New York.

Brocca, L., et al. (2011), Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe, *Remote Sens. Environ.*, *115*, 3390–3408.

Caires, S., and A. Sterl (2003), Validation of ocean wind and wave data using triple collocation, *J. Geophys. Res.*, *103*(C3), 3098, doi:10.1029/2002JC001491.

Crow, W. T., and M. J. van den Berg (2010), An improved approach for estimating observation and model error parameters in soil moisture data assimilation, *Water Resour. Res.*, *46*, W12519, doi:10.1029/2010WR009402.

Crow, W. T., M. J. Van den Berg, G. J. Huffman, and T. Pellarin (2011), Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART), *Water Resour. Res.*, *47*, W08521, doi:10.1029/2011WR010576.

Dorigo, W. A., K. Scipal, R. M. Parinussa, Y. Y. Liu, W. Wagner, R. A. M. de Jeu, and V. Naeimi (2010), Error characterisation of global active and passive microwave soil moisture datasets, *Hydrol. Earth Syst. Sci.*, *14*, 2605–2616.

Draper, C. S., R. R. Reichle, R. A. de Jeu, V. Naeimi, R. M. Parinussa, 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., and E. F. Wood (1994), Multiscale modeling of spatially variable water and energy balance processes, *Water Resour. Res.*, *11*, 3061–3078, doi:10.1029/94WR01498.

Gruber, A., W. A. Dorigo, S. Zwieback, A. Xaver, and W. Wagner (2013), Characterizing coarse-scale representativeness of in situ soil moisture measurements from the International Soil Moisture Network, *Vadose Zone J.*, *12*(2), doi:10.2136/vzj2012.0170.

Hansen, L. (1982), Large sample properties of generalized methods of moments estimators, *Econometrica*, *50*, 1029–1054.

Hollmann, R., et al. (2013), The ESA climate change initiative: Satellite data records for essential climate variables, *Bull. Am. Meteorol. Soc.*, *94*, 1541–1552.

Huffman, G. J., et al. (2007), The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *J. Hydrometeor.*, *8*, 38–55.

Janssen, P. A. E. M., S. Abdalla, H. Hersbach, and J.-R. Bidlot (2007), Error estimation of buoy, satellite and model wave height data, *J. Atmos. Oceanic Technol.*, *24*(9), 1665–1677.

Liu, Y. Y., et al. (2011), Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, *Hydrol. Earth Syst. Sci.*, *15*, 425–436.

Maddala, G. S. (2001), *Introduction to Econometrics*, John Wiley, Chichester, New York.

McKenzie, N. J., D. W. Jacquier, L. J. Ashton, and H. P. Cresswell (2000), Estimation of soil properties using the Atlas of Australian Soils. CSIRO Land and Water Technical Report 11/00. [Available at http://www.clw.csiro.au/aclep/documents/tr11-00.pdf.]

Miralles, D. G., W. T. Crow, and M. H. Cosh (2010), Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations, *J. Hydrometeor.*, *11*, 1423–1429.

Mladenova, I. E., et al. (2014), Remote monitoring of soil moisture using passive microwave-based techniques—Theoretical basis and overview of selected algorithms for AMSR-E, *Remote Sens. Environ.*, *144*, 197–213.

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*, 1999–2013.

Nelson, C. R., and R. Startz (1990), Some further results on the exact small sample properties of the instrumental variable estimator, *Econometrica*, *58*(4), 967–976.
Northcote, K. H., et al. (1960–1968), *Atlas of Australian Soils, Sheets 1 to 10. With explanatory data*, CSIRO Aust. and Melbourne Univ. Press, Melbourne.
O'Carroll, A. G., J. R. Eyre, and R. W. Saunders (2008), Three-way error analysis between AATSR, AMSR-E, and in situ sea surface temperature observations, *J. Atmos. Oceanic Technol.*, *25*(7), 1197–1207.
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.
Peel, M. C., B. L. Finlayson, and T. A. McMahon (2007), Updated world map of the Köppen-Geiger climate classification, *Hydrol. Earth Syst. Sci.*, *11*, 1633–1644.
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., et al. (2011), Assessment and enhancement of MERRA land surface hydrology estimates, *J. Clim.*, *24*(24), 6322–6338.
Reiersøl, O. (1941), Confluence analysis by means of lag moments and other methods of confluence analysis, *Econometrica*, *9*(1), 1–24.
Rienecker, M., et al. (2011), MERRA: NASA's modern-era retrospective analysis for research and applications, *J. Clim.*, *24*(14), 3624–3648.
Roebeling, R. A., E. L. A. Wolters, J. F. Meirink, and H. Leijnse (2012), Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data, *J. Hydrometeor.*, *13*(5), 1552–1566.
Rousseeuw, P. J. (1985), Multivariate estimation with high breakdown point, in *Mathematical Statistics and Applications*, vol. B, edited by W. Grossman et al., pp. 283–297, Reidel, Dordrecht.
Rousseeuw, P. J., and K. Van Driessen (1999), A fast algorithm for the minimum covariance determinant estimator, *Technometrics*, *41*(3), 212–223.
Sargan, J. D. (1958), The estimation of economic relationships using instrumental variables, *Econometrica*, *26*(3), 393–415.
Sargan, J. D. (1959), The estimation of relationships with autocorrelated residuals by the use of instrumental variables, *J. Roy. Stat. Soc. B Stat. Meth.*, *21*(1), 91–105.
Scipal, K., T. Holmes, R. de Jeu, V. Naeimi, and W. Wagner (2008), A possible solution for the problem of estimating the error structure of global soil moisture data sets, *Geophys. Res. Lett.*, *35*, L24403, doi:10.1029/2008GL035599.
Smith, A. B., et al. (2012), The Murrumbidgee soil moisture network data set, *Water Resour. Res.*, *48*, W07701, doi:10.1029/2012WR011976.
Stoffelen, A. (1998), Towards the true near-surface wind speed: Error modelling and calibration using triple collocation, *J. Geophys. Res.*, *103*, 7755–7766, doi:10.1029/97JC03180.
Su, C.-H., D. Ryu, A. W. Western, and W. Wagner (2013a), De-noising of passive and active microwave satellite soil moisture time series, *Geophys. Res. Lett.*, *40*, 3624–3630, doi:10.1002/grl.50695.
Su, C.-H., D. Ryu, A. W. Western, W. T. Crow, and W. Wagner (2013b), Error characterization of microwave satellite soil moisture data sets using Fourier analysis, *Proc. 20th International Congress of Modeling and Simulation*, 3120-3126.
Su, C.-H., D. Ryu, Y. I. Rodger, A. W. Western, and W. Wagner (2013c), Inter-comparison of microwave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia, *Remote Sens. Environ.*, *134*, 1–11.
Su, C.-H., D. Ryu, A. W. Western, W. T. Crow, and W. Wagner (2013d), Towards optimal rescaling between two noisy measurements: An application to satellite soil moisture. Satellite Soil Moisture Validation and Application Workshop, Poster presentation.
Vogelzang, J., A. Stoffelen, A. Verhoef, and J. Figa-Saldana (2011), On the quality of high-resolution scatterometer winds, *J. Geophys. Res.*, *116*, C10033, doi:10.1029/2010JC006640.
Wagner, W., G. Lemoine, and H. Rott (1999), A method for estimating soil moisture from ERS scatterometer and soil data, *Remote Sens. Environ.*, *70*, 191–207.
Wright, P. G. (1928), *The Tariff on Animal and Vegetable Oils*, Macmillan, New York.
Yilmaz, M. T., and W. T. Crow (2013), The optimality of potential rescaling approaches in land data assimilation, *J. Hydrometeor.*, *14*, 650–661.
Yilmaz, M. T., W. T. Crow, M. C. Anderson, and C. Hain (2012), An objective methodology for merging satellite- and model-based soil moisture products, *Water Resour. Res.*, *48*, W11502, doi:10.1029/2011WR011682.
Zwieback, S., K. Scipal, W. Dorigo, and W. Wagner (2012), Structural and statistical properties of the collocation technique for error characterisation, *Nonlinear Process. Geophys.*, *19*, 69–80.
Zwieback, S., W. Dorigo, and W. Wagner (2013), Estimation of the temporal autocorrelation structure by the collocation technique with emphasis on soil moisture studies, *Hydrol. Sci. J.*, *58*(8), 1729–1747.