



# Multi-scale analysis of bias correction of soil moisture

C.-H. Su and D. Ryu

Department of Infrastructure Engineering, University of Melbourne, 3010 Victoria, Australia

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

Received: 1 July 2014 – Published in Hydrol. Earth Syst. Sci. Discuss.: 29 July 2014

Revised: 2 December 2014 – Accepted: 4 December 2014 – Published: 6 January 2015

**Abstract.** Remote sensing, in situ networks and models are now providing unprecedented information for environmental monitoring. To conjunctively use multi-source data nominally representing an identical variable, one must resolve biases existing between these disparate sources, and the characteristics of the biases can be non-trivial due to spatio-temporal variability of the target variable, inter-sensor differences with variable measurement supports. One such example is of soil moisture (SM) monitoring. Triple collocation (TC) based bias correction is a powerful statistical method that is increasingly being used to address this issue, but is only applicable to the linear regime, whereas the non-linear method of statistical moment matching is susceptible to unintended biases originating from measurement error. Since different physical processes that influence SM dynamics may be distinguishable by their characteristic spatio-temporal scales, we propose a multi-timescale linear bias model in the framework of a wavelet-based multi-resolution analysis (MRA). The joint MRA-TC analysis was applied to demonstrate scale-dependent biases between in situ, remotely sensed and modelled SM, the influence of various prospective bias correction schemes on these biases, and lastly to enable multi-scale bias correction and data-adaptive, non-linear de-noising via wavelet thresholding.

morphic and ecological processes (Rodriguez-Iturbe, 2000; GLACE Team et al., 2004; Legates et al., 2011). It also represents an archetype of the aforementioned problem, where in situ networks, remote sensing and models jointly provide extensive SM information.

In situ networks usually provide point-scale measurements; satellite retrieval of shallow SM at a mesoscale footprint of 10–50 km must resort to a homogeneity or dominant-feature assumption, whereas modelled SM depends on the simplified model parameterization, and the quality, resolution and availability of forcing data. Subsequently, the spatial (lateral and vertical) variability of SM can lead to systematically different measurements regarded as *biases*. Descriptive or predictive spatial SM statistics can be used to relate point-scale to mesoscale estimates (Western et al., 2002), but in situ data are often limited in describing the spatial heterogeneity of SM. However, without bias correction, it is not possible to conduct meaningful comparisons between in situ, satellite-retrieved and modelled SM for validation (Reichle et al., 2004) and optimal data assimilation (Yilmaz and Crow, 2013). Standard bias correction methods are now increasingly being applied to SM assimilation in land models (Reichle et al., 2007; Kumar et al., 2012; Draper et al., 2012), numerical weather prediction (Drusch et al., 2005; Scipal et al., 2008a) and hydrologic models (Brocca et al., 2012). Reichle and Koster (2004) proposed matching statistical moments of the data, while linear methods based on simple regression and matching dynamic ranges have also been considered (e.g. Su et al., 2013a). But these methods can induce artificial biases in the signal component of the corrected data as the error statistics were ignored; this also suggests a connection that the issue of bias correction is inseparable from that of error characterisation (Su et al., 2014a).

Triple collocation (TC) (Stoffelen, 1998), which is a form of instrument-variable regression (Wright, 1928; Su et al.,

## 1 Introduction

Global environmental monitoring requires geophysical measurements from a variety of sources and sensors to close the information gap. However, different direct and remote sensing, and model simulation can yield different estimates due to different measurement supports and errors. Soil moisture (SM) is one such variable that has garnered increasing interest due to its influences on atmospheric, hydrologic, geo-

2014a), is increasingly being used to address these issues in oceanography (Caires and Sterl, 2003; Janssen et al., 2007) and hydrometeorology (Scipal et al., 2008b; Roebeling et al., 2013). In particular, it was used to estimate spatial point-to-footprint sampling errors (Miralles et al., 2011; Gruber et al., 2013), and correct biases in SM (Yilmaz and Crow, 2013). Based on an affine signal model and additive orthogonal error model, it assumes that representativity differences are manifested as additive and multiplicative biases. But these assumptions may have limited validity, as the temporal behaviour of SM may vary across different spatial scales, driven by a continuum of localised and mesoscale influences (e.g. Entin et al., 2000; Mittelbach and Seneviratne, 2012). Specifically, the coupling of SM with precipitation and evaporative losses (controlled by temperature, humidity, wind speed) varies across spatial scales. This can be more pronounced at places where surface hydrological features (e.g. topography, infiltration rate and storage capacity) are highly heterogeneous. Thus, the biases are likely to be non-systematic across short and long timescales on different spatial scales and errors are non-white, undermining the utility of the affine model. One possible remedy is to apply bias correction, either TC or statistical-moment matching, only to anomaly time series (Miralles et al., 2011; Liu et al., 2012; Su et al., 2014a), but it remains unclear how these methods affect the signal and noise components in the corrected data. Alternatively a moving time window can be used to examine the time-varying statistics of time series (Loew and Schlenz, 2011; Zwieback et al., 2013; Su et al., 2014a).

Given the possible (time)scale dependency in biases and errors, we propose an extension to TC analyses to include wavelet-based multi-resolution analysis (MRA) (Mallat, 1989) as a framework to (1) provide a fuller description of the temporal scale-by-scale relationships between coincident data sets; (2) study the influence of various prospective bias correction schemes; and (3) achieve multi-scale bias correction. To avoid excessive changes in the noise characteristics upon correction, TC can be further combined with the wavelet thresholding (Donoho and Johnstone, 1994) to (4) achieve non-linear, data-adaptive de-noising, with contrast to existing linear schemes (Su et al., 2013b). The techniques were applied to SM data from an in situ probe, satellite radiometry and land-surface model, but the proposed methods are generally enough to be applied to other geophysical variables.

The paper is organised as follows. Section 2 presents the study area over Australia and the SM data sets used in our pilot studies. Section 3 explains the theoretics behind MRA and applies it to SM, following by examination of scale-by-scale statistics in Sect. 4. Section 5 presents a new joint MRA-TC analysis framework, which is then applied to examine the influence of different bias correction schemes in Sect. 6. Importantly, both Sects. 4 and 6, using wavelet correlation, wavelet variance and scale-level TC analyses, provide evidence to support the need to extend traditional bulk

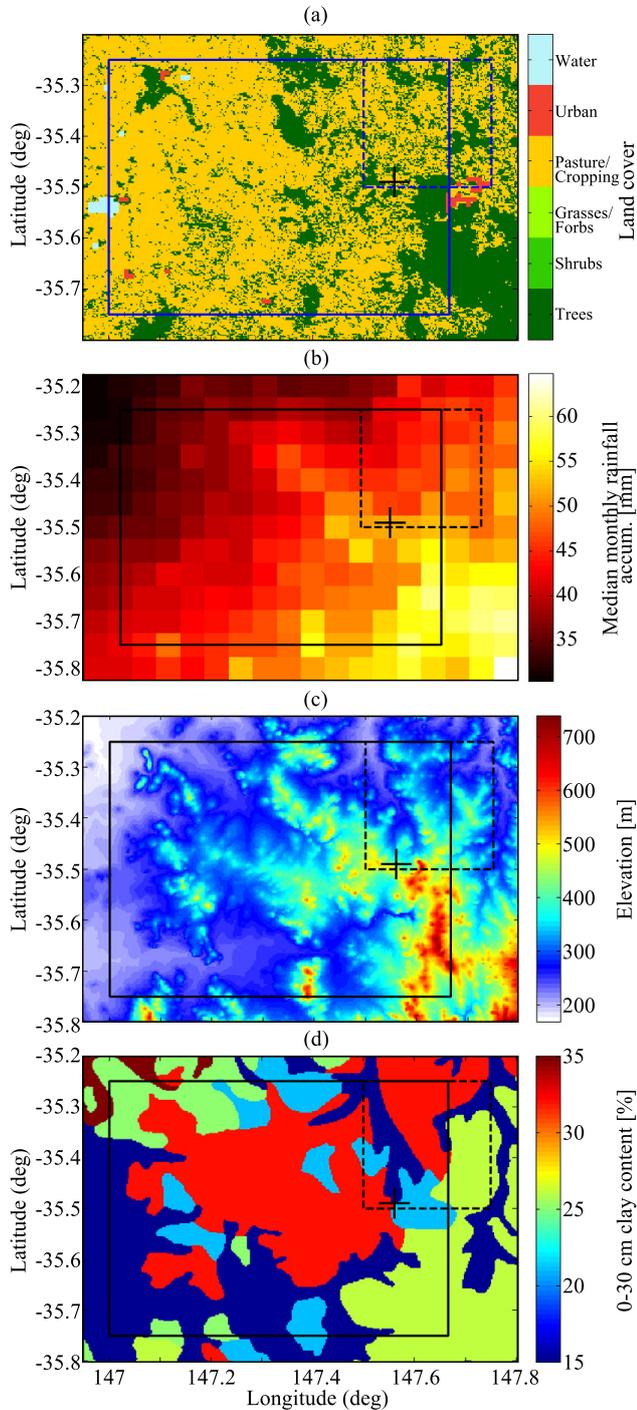
and anomaly based analyses. Section 7 demonstrates the use of wavelet thresholding to de-noise satellite SM. Section 8 offers our concluding remarks.

## 2 Study areas and data sets

We consider in situ, satellite-retrieved and modelled SM over Australia. For an in-depth study, we consider point-scale and pixel-scale SM estimates at K1 monitoring site (147.56° longitude, -35.49° latitude) situated at Kyeamba Creek catchment, southeastern Australia (Smith et al., 2012; Su et al., 2013a). The in situ SM (INS as shorthand) was sampled at 30 min intervals, 0–8 cm depth using a time-domain interferometer-based Campbell Scientific 615 probe during November 2001–April 2011. The region experiences a temperate (Cfb) climate characterised by seasonally uniform rainfall but variable evapotranspiration forcing, so that SM varies between dry in summer (December–February) to wet in winter (June–August). The creek is located on gentle slopes with rain-fed cropping and pasture, and the soil varies from sandy to loam. Figure 1 illustrates the land cover, elevation, monthly rainfall accumulation (from 2002 to 2011), and clay content over the region.

The satellite SM was retrieved by AMSR-E (Advanced Microwave Scanning Radiometer for Earth Observing System; AMS) of the AQUA satellite. The retrieval is based on an inversion of the forward radiative transfer model of a vegetation-masked soil surface, relating observed brightness temperature to soil dielectric constant estimates. A dielectric mixing model is then used to related the dielectric constant to volumetric SM. The combined C/X-band  $1/4^\circ \times 1/4^\circ$  gridded, half-daily ( $\sim 1.30$  a.m./p.m. LT – local time) version 5 product (July 2002–October 2011) is based on the Land Parameter Retrieval Model (Owe et al., 2008). C-band (X-band) has a shallow sampling depth of  $\sim 1$ –2 cm ( $\sim 5$  mm), although it is mostly C-band data over Australia due to relatively small radio frequency interference. Given the 1–2 day revisit times of the satellite, there is a significant number of missing values in the AMS data. However, we found that (not shown) over 99% (95%) of the gaps over Australia are  $\leq 1.5$  day ( $\leq 1$  day) long. For use in wavelet analysis (Sect. 4), a one-dimensional (1-D in time) interpolation algorithm (Garcia, 2010) based on discrete cosine transform (Wang et al., 2012) was applied to infill gaps of lengths  $\leq 5$  days in AMSR-E. Other interpolation methods were trialled; e.g. linear interpolated AMSR-E shows great similarities to the DCT interpolated data, while cubic spline interpolation leads to spurious peaks.

The modelled SM is taken from MERRA (Modern Era Retrospective-analysis for Research and Applications) – Land produced by the Catchment land surface model GEOS version 5.7.2. The MERRA atmospheric re-analysis is driven by a vast collection of in situ observations of atmospheric and surface winds, temperature, and humidity, and remote sens-



**Figure 1.** Spatial variability of land surface and rainfall over Kyeamba Creek. The cross denotes the location of the K1 monitoring station, and the dashed (solid) box is the pixel area of AMS (MER).

ing of precipitation and radiation (Rienecker et al., 2011). The MERRA land-only fields were post-processed by reintegrating a revised Catchment model with more realistic precipitation forcing to produce the MERRA-Land (MER as shorthand) data set (Reichle et al., 2011). The resultant SM field corresponds to the hourly averages of the uppermost layer (0–2 cm) and is gridded on a  $2/3^\circ \times 1/2^\circ$  grid.

The three data are co-located spatially via nearest neighbour and temporally at around the satellite overpass times of 1.30 a.m./p.m. LT. Their time series are plotted in blue in the first panels of Fig. 2. While co-located, the three methods observed SM dynamics over different locations and areas of the catchment (Fig. 1), due to differences in their pixel resolutions and alignments.

Continental-scale AMS and MER data over Australia are also considered. The continent has great variability in climatic and land surface characteristics. Most of the northern regions experience a tropical savannah (Aw) Köppen–Geiger climate as classified by Peel et al. (2007), central Australia is largely arid desert (BWh), and eastern mountainous areas have a temperate climate with no dry seasons (Cf). The southwestern regions similarly have a temperate climate, but with dry summers (Cs). These temperate regions have higher vegetation compared to the tropical north with moderate vegetation cover.

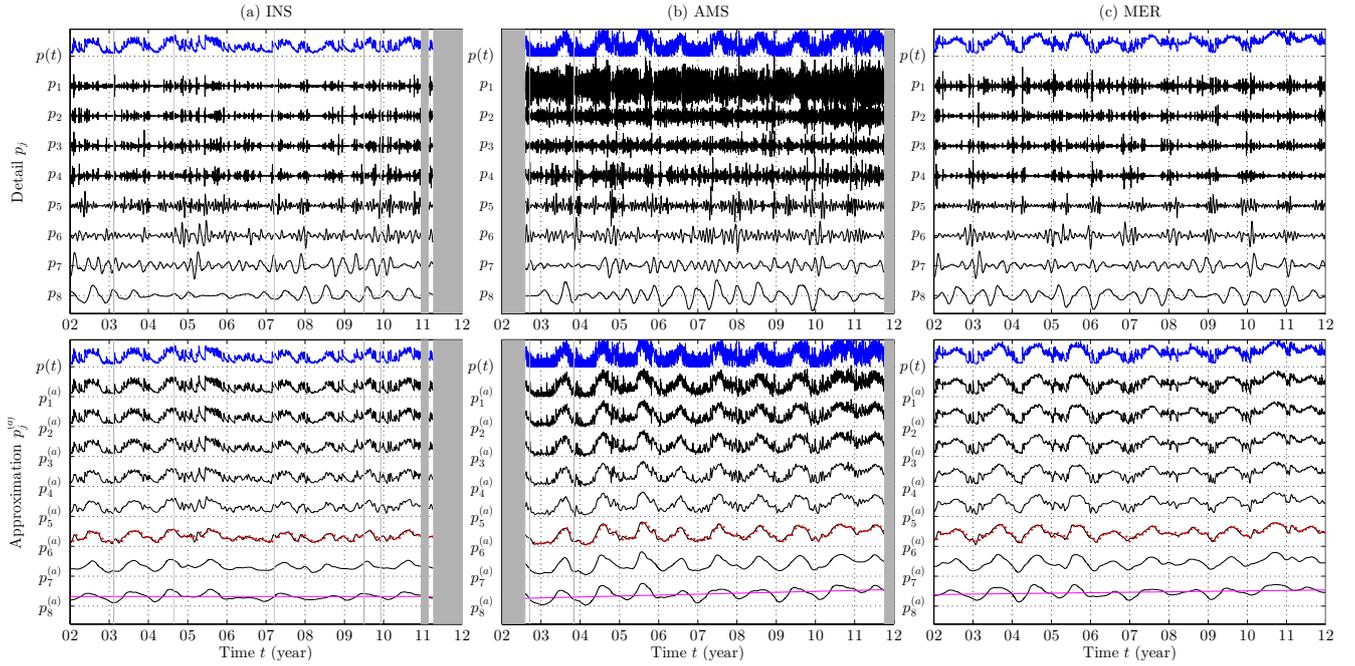
### 3 Multi-scale decomposition of soil moisture

The observed Kyeamba SM (denoted by blue curves  $p$  in Fig. 2) exhibits a long-term cycle of wet and dry years due to the El Niño–Southern Oscillation and seasonal and diurnal cycles originating from the fluctuations in vegetation and solar radiation, and experiences transient decay from various loss mechanisms, and abrupt increase from individual rainfall events. Their influences on observed SM can vary with the measurement methods. To unravel these differences, we turn to wavelets as the analysing kernels to study variability on individual broad-to-fine timescales. The *scale* under investigation is *temporal* for the rest of the paper, unless stated otherwise.

The 1-D orthogonal discrete wavelet transform (DWT) enables MRA of a time series  $p(t)$  of dyadic length  $N = 2^J$  and a regular sampling interval  $\Delta t$  by providing the mechanism to go from one resolution to another via a recursive function

$$p_{j-1}^{(a)}(t) = p_j^{(a)}(t) + p_j(t), \tag{1}$$

with an expectation value  $E(p_j^{(a)}) = E(p) = p_j^{(a)}(t) = \mu_p$  and  $E(p_j) = 0$ , where the superscript (a) labels approximated representations. The integer  $j \in [1, J]$  labels the scale of analysis, with  $j = 1$  ( $J$ ) denoting the finest (coarsest) scale, and serves to define a spectral range in a spectral analysis. The recursion therefore relates an approximation or coarse representation  $p_j^{(a)}$  of the signal at one resolution to



**Figure 2.** MRA of INS, AMS and MER SM at Kyeamba.  $p$  denotes the original time series,  $p_j$  the detail time series, and  $p_j^{(a)}$  the approximation time series. Grey shadings are  $> 5$  day data gaps, red dots superimposed in  $p_6^{(a)}$  are monthly means of  $p$ , and magenta lines are trend lines fitted to  $p_8^{(a)}$ .

that at a higher resolution  $p_{j-1}^{(a)}$  by adding some fine-scale detail denoted by  $p_j$ . The end of the recursion chain leads to reconstruction of the original time series with the equality  $p_0^{(a)}(t) = p(t)$ , and a multi-resolution decomposition of  $p$  as

$$p(t) = p_{j_0}^{(a)}(t) + \sum_{j=1}^{j_0} p_j(t) \quad (2)$$

$$= \sum_{k=1}^{n_{j_0}} p_{j_0 k}^{(a)} \phi_{jk}(t) + \sum_{j=1}^{j_0} \sum_{k=1}^{n_j} p_{jk} \psi_{jk}(t) \quad (3)$$

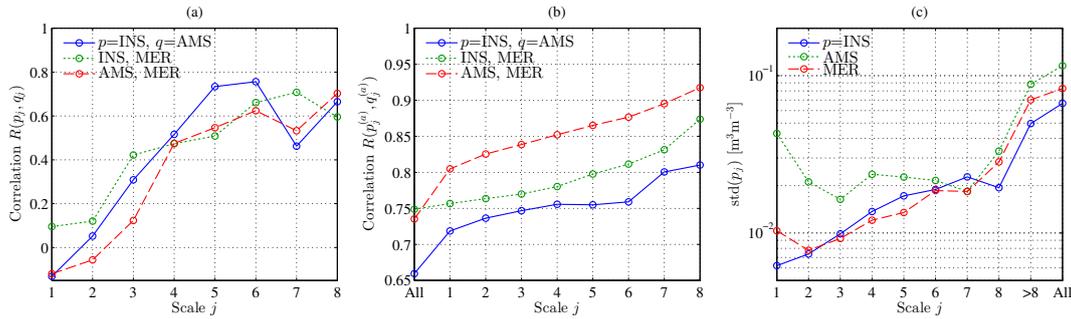
under  $j_0$  levels of decomposition. Loosely speaking, for a half-daily time series, the *detail time series*  $p_j$  for  $j = 1, 2, 3, \dots$  corresponds to (fine-scale) dynamics observed on 1 day (1 d), 2 d, 4 d, etc., timescales, while the *approximation time series*  $p_j^{(a)}$  for  $j = 1, 2, 3, \dots$  contains (broad-scale) dynamics on scales longer than 1 d, 2 d, 4 d, etc.

In Eq. (3), each of these components is further decomposed into a linear summation of  $n_j = N/2^j$  number of basis functions  $\phi_{jk}$  and  $\psi_{jk}$  with scale of variability  $2^j \Delta t$  and temporal location  $k 2^j \Delta t$ . The weighting or wavelet coefficients, determined via DWT of  $p$ , measure the similarity between  $p$  and the bases via the inner products  $p_{jk}^{(a)} \equiv \langle p, \phi_{jk} \rangle$  and  $p_{jk} \equiv \langle p, \psi_{jk} \rangle$ . Hence the coefficients indicate changes on a particular scale and location, and enable the above scale-by-scale decomposition. Note that the bases are defined in  $L^2(R)$  space and satisfy orthonormality conditions

prescribed by  $\langle \phi_{jk}, \phi_{j'k'} \rangle = \delta_{jj'} \delta_{kk'}$ ,  $\langle \psi_{jk}, \psi_{j'k'} \rangle = \delta_{jj'} \delta_{kk'}$ ,  $\langle \phi_{jk}, \psi_{j'k'} \rangle = 0$ , where  $\delta$  is the Kronecker delta function. For detailed expositions of the mathematical theory of wavelets and MRA, consult Daubechies (1992) and Mallat (1989).

The detail and approximated time series of Kyeamba's SM are illustrated in subsequent panels of Fig. 2, analysed using the Daubechies  $D(4)$  wavelet for  $j_0 = 8$ . On the finest scales  $j = 1-2$  (1–2 d), the details show variability due to rainfall wetting, and over the next set of scales  $j = 2-5$  (2–16 d) they describe transient moisture loss. The  $p_6^{(a)}$  ( $\geq 32$  d) component accounts for several scales of fluctuations over seasonal, inter-annual, and long-term timescales. For comparison, the standard monthly average analyses of the original time series  $p$  are superimposed on  $p_6^{(a)}$  (red dots).

The differences between the details of the three SM are apparent on the finest scales, with AMS and MER showing greater variability and amplitude compared to INS. However, the similarity of their temporal patterns, in both details and approximations, grows with increasing scales  $j > 3$  (see also Fig. 3). Fitting a trend line to their coarsest scale approximation series suggests  $p_8^{(a)}$  that the trends (magenta lines) in the three data show different gradients, with the trend in INS showing the smallest positive gradient. The differences in dynamic ranges of their detail and approximation time series, together with their mismatch in shape and trend, are indicative of multiplicative biases and noise.



**Figure 3.** Comparisons of correlation  $R$  and SD between INS, AMS and MER at scale levels. **(a)** compares the correlation between their detail time series  $p_j$ , and **(b)** compares between their approximation time series  $p_j^{(a)}$ . Scale  $j > 8$  corresponds to  $p_8^{(a)}$ , and “All” refers to statistics of the original time series.

#### 4 Multi-scale statistics

MRA enables direct comparisons between any two representations  $p = \{X, Y\}$  of a given variable  $f$  (e.g. SM) on various temporal scales independently, owing to the orthonormal properties of wavelet bases. It also offers an additional degree of freedom in temporal positions (using the index  $k$ ) to allow better representation of local variability. By subsetting the wavelet coefficients over certain range of  $k$  values, non-stationary statistics can also be examined. However, in this work, we consider only variability across  $j$  and assume stationarity on each scale. Pearson’s linear correlation  $R$  and variance analyses (see Appendix A) are performed on the Kyeamba’s INS, AMS and MER SM (as  $p$  in Eq. 2) detail ( $p_j$ ) and approximation ( $p_j^{(a)}$ ) time series in Fig. 3. The strength of MRA is that since the detail time series  $p_j$  on a given scale  $j$  does not contain variations on timescales greater than  $j$ , the weak-sense stationarity conditions can be better met.

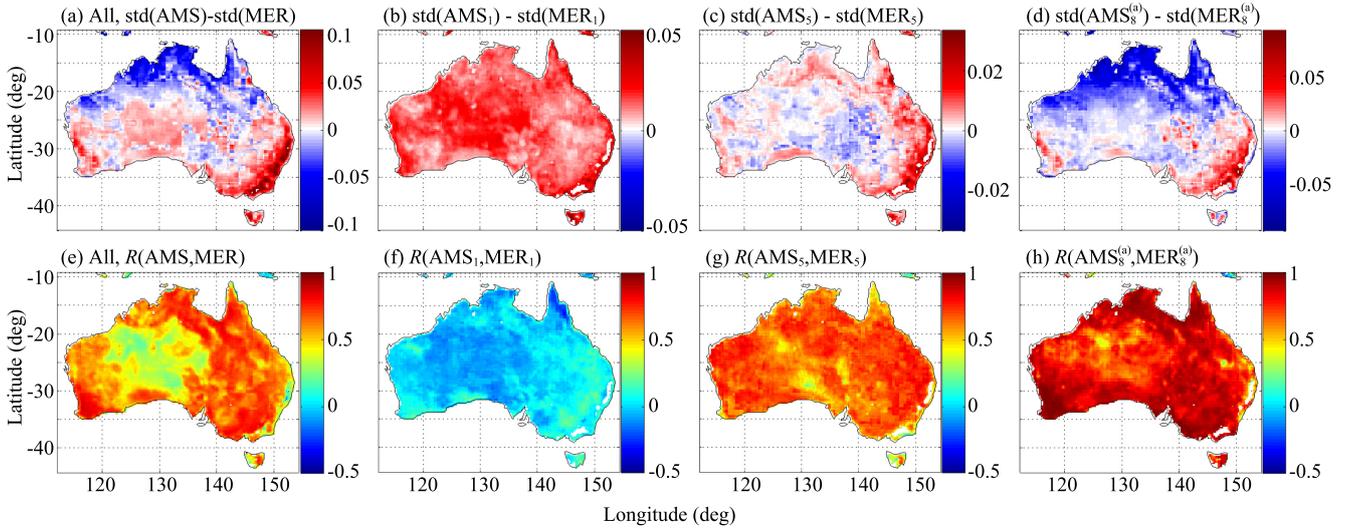
Before proceeding, we recall that weak  $R$  indicate the presence of noise and/or the presence of non-linear correlation between any pairs of the data, while differences in standard deviation can also indicate the presence of noise, but also an extraneous signal and/or multiplicative bias. Typically one invokes a linearity assumption and assumes an affine relation between the signal components of the different data and an additive noise model (more later in Sect. 5), so that the differences between the data are attributed only to an overall additive bias  $E(X) - E(Y)$ , multiplicative biases, and noise. While we adopt this simplistic viewpoint here, its limitations to properly account for variable lateral and vertical measurement supports should be noted. For instance, short-timescale SM dynamics show increasing attenuation in amplitude, but are also delayed in time in deeper soil columns (e.g. Steelman et al., 2012). Additionally, SM is physically bounded between field capacity and residual content and these thresholds can vary with soil texture, location and depths. These effects can give rise to temporal autocorrelation in errors and undermine the linearity assumption be-

tween coincident measures. Finally, the non-stationary characteristic of noise in satellite SM (Loew and Schlenz, 2011; Zwieback et al., 2013; Su et al., 2014a) due to e.g. dynamical land surface characteristics such as soil moisture (Su et al., 2014b), is not treated here.

With these considerations, we first examine the correlations between the three data. For the detail time series (Fig. 3a), their correlations are lowest on the finest scales ( $R < 0.2$ ) but generally improves with scale ( $R > 0.5$ ), as noted previously. There is however no data pair that shows consistently higher  $R$  than other pairs:  $R(INS_j, AMS_j) > R(INS_j, MER_j)$  on coarser scales  $j = 4-6, 8$ , whereas  $R(INS_j, MER_j)$  is highest on other scales. Comparing their approximation time series (Fig. 3b),  $R$  between AMS and MER are higher than the other two pairs, ranging from ( $j = 2$ ) 0.8 to 0.92 ( $j = 8$ ), largely due to the strong correlation between their respective  $p_8$  and  $p_8^{(a)}$ . In other words: on one hand, AMS and MER both show skill in representing some aspects of the in situ SM temporal variability; on the other hand, stronger AMS-MER correlations on the coarsest (temporal) scales and their mesoscale spatial resolutions would indicate lesser representativeness of in situ measurement on these spatio-temporal scales.

Furthermore, we observe that  $R(p_j^{(a)}, q_j^{(a)})$  reduces with decreasing  $j$  as more components are added to the reconstruction of  $p_j^{(a)}$  and  $q_j^{(a)}$ . The inclusion of noisy  $AMS_1$  in the makeup of AMS leads to a drop in  $R(INS, AMS)$  and  $R(AMS, MER)$ . Aside from including more noise in the approximation time series, adding components with different multiplicative biases (more later in Sect. 6) can also diminish the correlations. The scale dependence of multiplicative biases and added noise can contribute to the contrasting results of applying TC to raw versus anomaly SM time series in Draper et al. (2013). In particular, given the presence of noise in  $p_j$  for  $j \geq 7$ , error analysis of the anomaly SM (i.e. in  $p_j$  for  $j \leq 6$ ) will under-estimate the total error in the raw data  $p$ .

Next, Fig. 3c plots their wavelet spectra that decompose total variance  $\text{var}(p)$  into individual scales



**Figure 4.** Difference in SD (in units of  $\text{m}^3 \text{m}^{-3}$ ) and correlation  $R$  between AMS and MER for (a, e) all and (rest) on selected timescales.

$\text{var}(p_j) \equiv \text{std}(p_j)^2$ . The three data show clear differences in their standard deviation (SD) profile, both in the fine and coarse scales. As already noted, both noise and/or multiplicative biases are possible contributing factors such that noise can inflate the variance, while biases can cause suppression or inflation. Following the visual inspection of Fig. 2 and the noted weak correlations  $R(\text{INS}_j, \text{AMS}_j)$  and  $R(\text{INS}_j, \text{MER}_j)$  at small  $j$ , it can be argued that there is significant noise in AMS (for  $j = 1-3$ ) and MER ( $j = 1$ ). This in turn leads to their larger SD cf. INS. On coarser scales where  $R$  values are significantly higher, the differences in SD may be attributed more to multiplicative biases. For instance for their  $p_8$  and  $p_8^{(a)}$  components, AMS and MER shows larger SD and thus positively biased relative to INS.

Figure 4 extends the variance and correlation analyses between AMS and MER to the Australian continent using their coincident data from the period July 2002–October 2011. The spatial maps of SD differences ( $\Delta\text{SD}$ ) and correlations show significant variability in the statistics with timescales and spatial locations. On the finest scale  $j = 1$ , the similarity between the difference map (Fig. 4a) and the TC-derived error map of AMSR-E (see Fig. 6a in Su et al., 2014a) in terms of spatial variability and the low AMS-MER correlations (Fig. 4f) support our observation that the detail time series  $\text{AMS}_1$  is noise dominated. Weak negative correlation between  $\text{AMS}_1$  and  $\text{MER}_1$  can also be observed over arid regions. By contrast, owing to the strong correlation  $R \sim 0.6-0.9$  (Fig. 4g and h) on the coarse scales, the causes of  $\Delta\text{SD}$  (Fig. 4c and d) are related to biases. In particular, at  $j > 8$ , the  $\Delta\text{SD}$  map in Fig. 4d also suggests a possible association between biases and climatology or land cover characteristics, with negative biases dominating northern tropical (Aw) and semi-arid (BS) regions, and positive biases in temperate, vegetated regions (Cs and Cf) over southeastern and southwest-

ern Australia. The visual comparisons between scale-level  $\Delta\text{SD}$  and bulk  $\Delta\text{SD}$  enable stratification of the continent to central arid regions of higher noise identified in  $j = 1$  and 2 and temperate (tropical) regions, with a positive (negative) bias seen on coarser scales.

## 5 Joint MRA-TC analysis

In order to quantify observed differences between the data, we propose a *scale-dependent* linear model: a multi-scale (MS) model that distinguishes the signal components of the two data  $X$  and  $Y$  via an overall additive bias and a set of positive scaling coefficients  $\alpha_{p,j}$ ,  $\alpha'_p$ , and assumes an additive and zero-mean independent but non-white noise model  $\epsilon_p(t)$ . Focusing on the zero-mean signal and noise components, the “structural relationship” model reads

$$p'(t) = \alpha'_p f'(t) + \epsilon'_p(t), \quad (4)$$

$$p_j(t) = \alpha_{p,j} f_j(t) + \epsilon_{p,j}(t), \quad (5)$$

for  $p' = p_{j_0}^{(a)} - E(p)$  and  $f = f' - E(f)$ , where the signal and noise components have been decomposed into their multi-resolution forms. The standard assumptions of orthogonal and mutually uncorrelated errors are used, so that the covariance  $\text{cov}(f_j, e_{p,j}) = 0$ ,  $\text{cov}(f', e'_p) = 0$ ,  $\text{cov}(e_{p,j}, e_{q,j}) = 0$ ,  $\text{cov}(e_{p,j}, e'_{q'}) = 0$  and  $\text{cov}(e'_{p'}, e'_{q'}) = 0$  for  $p \neq q$ ,  $p, q \in \{X, Y\}$ . The differences in the values of the scaling coefficients between data, i.e.  $\alpha_{X,j} \neq \alpha_{Y,j}$ , signify multiplicative biases on individual scales. To see this, we express their mean-squared deviation  $\text{MSD} \equiv E[(Y - X)^2]$  in terms of variables in Eqs. (4) and (5) to arrive at

$$\text{MSD} = (\mu_X - \mu_Y)^2 + \sum_j^J \left[ (\alpha_{Y,j} - \alpha_{X,j})^2 \text{var}(f_j) + \text{var}(\epsilon_{X,j}) + \text{var}(\epsilon_{Y,j}) \right]. \quad (6)$$

The first term is the additive bias, and the summation consists of scale-specific multiplicative biases proportional to  $(\alpha_{X,j} - \alpha_{Y,j})^2$  and noise contributions from each datum. The interpretation of the discrepancies between  $X$  and  $Y$  can vary depending on the time period of the data and the analysis, and the adopted signal/noise model. By using the entire 9 year record of INS, AMS and MER data in MRA, the MS model does not observe a time-varying additive bias (e.g. from using the moving-window approach of Su et al., 2014a) or autocorrelated errors (from using the lagged covariance in Zwieback et al., 2013). Rather, MRA and the MS model enable a description of the systematic differences based wholly in terms of multiplicative biases at individual timescales, and the random differences in terms of additive noise. Specifically, this contrasts with the short time-window approach (e.g.  $\leq 32$  d), where multiplicative biases existing on coarse scales ( $p_6^{(a)}$ ) will manifest as both time-varying additive and multiplicative biases.

Importantly, the model allows for different scaling coefficients between scales, i.e.  $\alpha_{p,j} \neq \alpha_{p,j'}$  for  $j \neq j'$ , as a form of non-linearity with  $f$ . The equality  $\alpha_{p,j} = \alpha'_p = \alpha_p$  is therefore a special case of (bulk) linearity. As our focus of the above model is the multiplicative biases and noise, for convenience of notation, we remove the mean of the  $X$  and  $Y$  prior to MRA and bias correction. Furthermore, without the loss of generality, we choose  $X$  as the reference henceforth and let  $\alpha_{X,j}, \alpha'_X = 1$ .

By using a third independently derived representation ( $Z$ ) of  $f$ , TC enables estimation of the required scaling coefficients and noise std( $\epsilon_{p,j}$ ) (Appendix B). As we will see later, these estimates are needed for bias correction and de-noising. Within the operating assumptions of TC, TC estimates are unbiased and consistent; that is, the estimated  $\hat{\alpha}_{Y,j} = \alpha_{Y,j}$  as the asymptotic limit. However, TC's superiority is dependent on the availability of a strong instrument and a large sample for statistical analyses (Zwieback et al., 2012; Su et al., 2014a). Standard linear estimators, namely ordinary least-square (OLS) regression and variance matching (VAR), can be considered as substitutes, although they are biased estimators of  $\alpha$  when  $X$  and  $Y$  are both noisy (Yilmaz and Crow, 2013; Su et al., 2014a), e.g. OLS yields  $\hat{\alpha}_{Y,j} < \alpha_{Y,j}$ . In summary, we propose that combining these estimators with MRA via the MS linear model enables investigation into the distribution of the multiplicative biases and additive noise over  $j$ , and their response to various bias correction schemes.

## 6 Multi-scale analysis of bias correction

Consider now the bias correction of  $Y$  to produce a corrected datum  $Y^*$  that “matches”  $X$ . Different interpretations of a “match” and assumptions about signal and noise statistics lead to different bias correction schemes. To describe matching, there are different choices of optimality criterion. The first is based on matching the statistics of the signal-only component of  $Y^*$  to that of  $X$ . This approach requires consistent estimation of slope parameters  $\alpha$ 's and the resultant statistics of  $X$  and  $Y^*$  may differ due to different noise statistics. The second is based on the matching of the statistical moments between  $Y^*$  and  $X$  (e.g. VAR matching), although the statistics of their constitutive signal components may differ for the same reason. The third is based on the minimum-variance principle of minimizing the least-square difference between  $Y^*$  and  $X$  (i.e. the OLS estimation), but as already noted, the estimator becomes inconsistent when there are measurement errors in  $X$  and  $Y$ .

Following our theoretical model in Sect. 5, we define our optimality criterion based on the first criterion of matching the first two moments of the signal components in  $X$  and  $Y$  so that  $Y^*$  is suitable for bias-free data assimilation. In particular, Yilmaz and Crow (2013) have shown that residual multiplicative biases due to a sub-optimal bias correction scheme will cause filter innovations to contain residual signal and sub-optimal filter performance. Thus, within the paradigm of the MS model, our goal of bias correction is to minimise the difference  $|\alpha_{Y^*,j} - 1|$  for  $\alpha_{X,j} = 1$ , so that the multiplicative bias terms in Eq. (6) are eliminated.

- *Bulk linear rescaling* assumes bulk linearity between  $X$  and  $Y$  so that the correction equation is

$$Y^* = \frac{Y}{\hat{\alpha}_Y}, \quad (7)$$

where  $\hat{\alpha}_Y$  is given by TC for our objective. When the bulk linearity is satisfied, this approach ensures that the statistical properties (SD and higher moments) of the signal components in  $X$  and  $Y^*$  are identical. Linear rescaling using  $\hat{\alpha}_Y$  values estimated by OLS and VAR matching have previously been considered by e.g. Su et al. (2013a); but due to error-in-variable biases, they can induce artificial biases in the signal component of  $Y^*$  even if the bulk linearity condition is valid.

- *Bulk cumulative distribution function (CDF) matching* assumes non-linearity between  $X$  and  $Y$  and transforms  $Y^*$  so that (Reichle and Koster, 2004)

$$\text{cdf}(Y^*) = \text{cdf}(X), \quad (8)$$

where  $\text{cdf}(\circ)$  computes the CDF. This ensures that the mean, SD, and higher statistical moments of  $X$  and  $Y^*$  are identical, but the statistical properties of their signal and noise components that make up  $X$  and  $Y^*$  are

not necessarily identical. In particular, when the relative signal and noise statistics in the two data are different, CDF matching leads to artificial biases between the signal components in  $X$  and  $Y^*$ . As with VAR matching of first two moments, the CDF counterpart is expected to contain extraneous contribution of the noise variances in the mapping of the second moment, as well as at higher moments (Su et al., 2014a). The issue can be exacerbated by variable signal and noise statistics on different scales.

- *Anomaly/seasonal (A/S) linear rescaling* allows biases between  $X$  and  $Y$  to be different on two scales of variation. In practice, the useful information content in observations is primarily based on their representation of anomalies, where observations are assumed in a particular land surface model's unique climatology (Koster et al., 2009). The correction is therefore limited to the anomalies, although other components (e.g. seasonal fluctuation and long-term trend) may be preserved to validate model prediction. Here the linear correction using TC estimator is applied to match the characteristics of each component – anomaly ( $i = A$ ) and seasonal ( $S$ ) – separately, so that the corrected  $Y$  has the form

$$Y^* = Y_S^* + Y_A^*, \quad (9)$$

with  $Y_i^* = Y_i / \hat{\alpha}_{Y_i}$  for  $i \in \{S, A\}$ . In one approach,  $p_S$  is computed using moving window averaging of multi-year data within a window size of 31 days centered on a given day of year (Miralles et al., 2011; Su et al., 2014a), so that inter-annual cycles and long-term trends are retained in  $p_A$ . In an alternative approach (Albergel et al., 2012), a sliding 31 day window is used such that  $p_A \approx \sum_{j=1}^6 p_j$  for half-daily time series. In this work, the former, more conventional approach was taken.

- *A/S CDF matching* applies CDF matching to anomaly and seasonal components separately as per Eq. (9) but with  $\text{cdf}(Y_i^*) = \text{cdf}(X_i)$ . The application of CDF matching to the anomaly component of soil moisture data was considered by Liu et al. (2012).
- *Multi-scale (MS) rescaling* is the direct consequence of the MS model where information in  $Y$  is rescaled at individual scales,

$$Y^* = \frac{Y'}{\hat{\alpha}'_Y} + \sum_{j=1}^{j_0} \frac{Y_j}{\hat{\alpha}_{Y,j}}. \quad (10)$$

In relation to Eq. (6), this approach obviously eliminates that the multiplicative terms in the summation. The bulk and A/S linear correction schemes can be considered as special cases of MS rescaling where information from multiple scales are aggregated and corrected

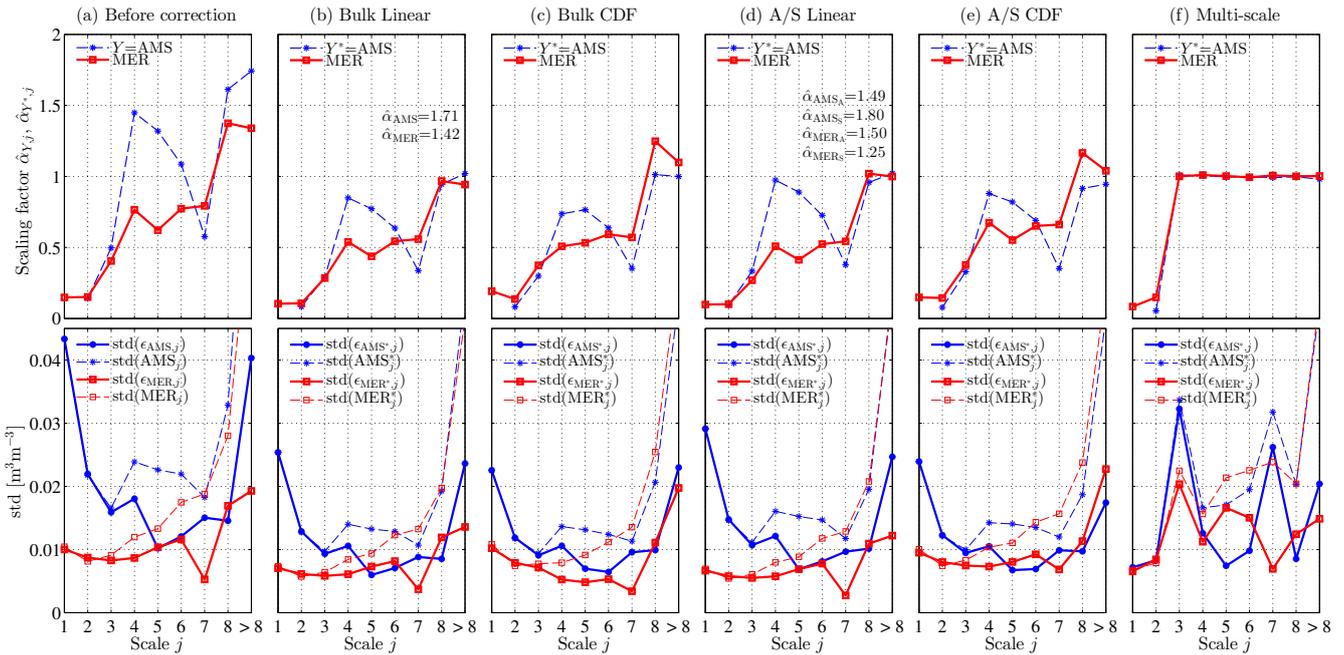
jointly. Other aggregations of the information from different subsets of scales are also possible, but they will similarly be conceived based on one's understanding or assumptions of the underlying specific processes driving SM dynamics. Investigations into suitable aggregations are beyond the scope of this work, hence we implemented the most elaborate decomposition. If joint linearity exists between two or more scales, their  $\alpha_{Y,j}$  values will be similarly valued for use in Eq. (10).

For illustrations, we correct the biases in AMS and MER SM with respect to INS SM at Kyeamba using the above five schemes. Using the above notations, AMS and MER are treated as  $Y$ , the corrected AMS\* and MER\* as  $Y^*$ , and INS as  $X$ . MRA-TC was applied to observe their consequences in Fig. 5. In the upper panel, estimated  $\hat{\alpha}_{Y,j}$  and  $\hat{\alpha}_{Y^*,j}$  values provide diagnostics for detecting the presence of multiplicative biases before and after application of the correction schemes. The lower panel plots the SD of  $Y_j$  and  $Y_j^*$  and their associated noises  $\epsilon_{Y,j}$  and  $\epsilon_{Y^*,j}$ . The values of the scaling coefficients  $\alpha_{Y,j}$  (before correction) and  $\alpha_{Y^*,j}$  (after), and the noise  $\text{std}(\epsilon_{Y,j})$  and  $\text{std}(\epsilon_{Y^*,j})$  were estimated using TC. But where TC estimates could not be retrieved (for  $j = 1-2$ ) due to negative correlation amongst the data triplet (e.g. resulting from significant noise and weak instrument), OLS-derived (under) estimates serve as a guide for the above diagnostic purposes. Similarly, the total SD is a guide for noise SD in these cases.

Figure 5a shows the MRA of the biases and noise in the pre-corrected data  $Y$ . There is considerable variability in  $\hat{\alpha}_{Y,j}$  across the scales, ranging from 0.5 to 1.8 for AMS, and from 0.5 to 1.4 for MER. In particular, their  $\hat{\alpha}'_Y$  and  $\hat{\alpha}_{Y,8}$  deviate significantly from 1, and are responsible for the larger SD (cf. INS) observed in Fig. 3c. Biases also exist on almost all other scales of AMS and MER. In the lower panel, the values of  $\text{std}(\epsilon_{Y,j})$  relative to  $\text{std}(Y_j)$  indicate the dominance of noise in the small scales  $j = 1-3$ . This explains the low  $R$  values between AMS (and MER) and INS in Fig. 3a. Furthermore, the signal-to-noise ratios are variable with scales and data sets, highlighting the importance of using a correction scheme that takes the signal-vs.-noise statistics into considerations. The TC-based scheme is limited to the linear case, and the CDF scheme ignores such a variability.

The MRA of the corrected data  $Y^*$  are shown in Fig. 5b–f. In addition we assess the level of agreement between corrected AMS\* and INS time series in Table 1 using their root-mean-squared deviation (RMSD) and correlation  $R$ . The time series plots are shown in Fig. 6 to support interpretations. These additional results focus on the AMS-INS pair that best illustrates the influence of noise in AMS.

The results of bulk, A/S and MS linear rescaling can be readily interpreted. For bulk (Fig. 5b) and A/S linear (Fig. 5d) rescaling, the values of  $\hat{\alpha}'_Y$  and  $\hat{\alpha}_{Y_j}$  used for their implementation (Eqs. 7 and 9) are listed in the figure. As these values are greater than unity for both AMS and MER, this leads to the



**Figure 5.** Bias correction of AMS and MER (as  $Y$ ) with respect to INS (as  $X$ ), showing the impact of 5 correction schemes on the scaling coefficients, noise and total SD on individual scales. Estimated  $\hat{\alpha}_{Y,j} \neq 1$  or  $\hat{\alpha}_{Y^*,j} \neq 1$  suggests multiplicative bias in  $Y_j$  or  $Y_j^*$  as per Eq. (6). (a) is the diagnosis of  $Y$  before correction, and (b–f) are that of  $Y^*$  after correction. The estimated  $\hat{\alpha}_{Y,j}$  and  $\hat{\alpha}_{Y^*,j}$  for the diagnoses are derived using OLS (for  $j = 1, 2$ ) and TC ( $j > 2$ ). The additional  $\hat{\alpha}_{Y,j}$  values listed in (b, d) are the scaling coefficients used in the implementations of bulk and A/S linear rescaling. Scale  $j > 8$  corresponds to  $Y_8^{(a)}$ .

**Table 1.** RMSD (in units of  $\text{m}^3 \text{m}^{-3}$ ) and correlation between INS and AMS SM at Kyeamba treated by various methods. The square brackets contain the 95 % confidence interval.

| Methods              | RMSD  | Correlation |
|----------------------|-------|-------------|
| None                 | 0.088 | 0.659 [14]  |
| Bulk linear          | 0.055 | 0.659 [14]  |
| Bulk CDF             | 0.053 | 0.679 [14]  |
| A/S linear           | 0.059 | 0.635 [15]  |
| A/S CDF              | 0.054 | 0.671 [14]  |
| Multi-scale (MS)     | 0.062 | 0.650 [15]  |
| Wavelet thresh. (WT) | 0.069 | 0.709 [13]  |
| WT + MS              | 0.048 | 0.711 [12]  |

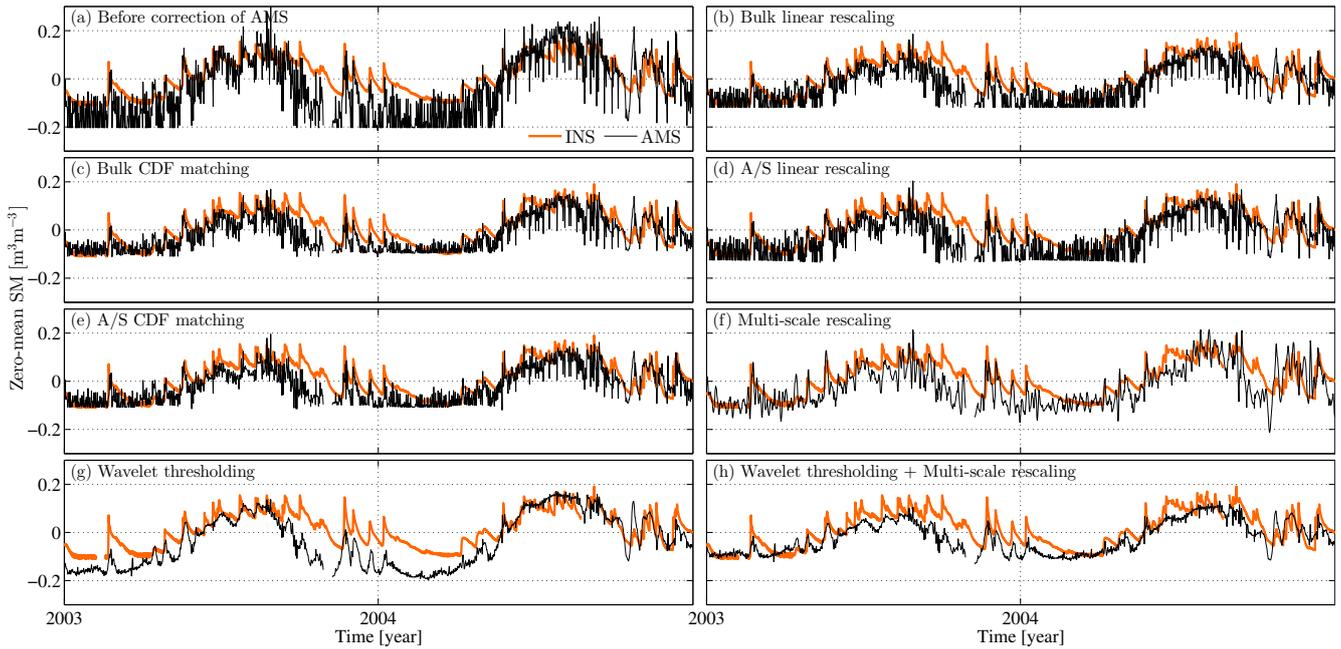
suppression of the associated signal, as well as noise, components:  $\text{std}(Y_j^*) < \text{std}(Y_j)$ , and  $\text{std}(\epsilon_{Y^*,j}) < \text{std}(\epsilon_{Y,j})$ . For AMS, the bulk linear scheme corrects the coarse-scale bias in  $Y_8^{(a)}$  component and rescales the noise variance, reducing RMSD from 0.09 to  $0.06 \text{ m}^3 \text{ m}^{-3}$ . However, the fine-scale biases in  $Y_j^*$  are still present, and increased on some scales, e.g. at  $j = 4, 7$  for  $\text{AMS}^*$ . Additionally for A/S linear rescaling,  $R(\text{AMS}^*, \text{INS})$  value does not change significantly and the noise are still clearly visible in Fig. 6b and d.

By construction, the MS rescaling uses the estimated  $\hat{\alpha}_{Y,j}$  values from Fig. 5a to correct bias on all the scales. Fig. 5f

shows the analysis of MS-corrected  $Y^*$ . The equivalence  $\hat{\alpha}_{Y^*,j} = 1$  indicates that the multiplicative biases are eliminated at  $j > 2$ . At  $j = 1-2$ , as the scaling coefficients cannot be estimated by TC, CDF matching was applied to these scales such that the biases are still present on these scales. Amid the reduction of biases, we also observed noise amplification (i.e.  $\text{std}(\epsilon_{Y^*,j}) > \text{std}(\epsilon_{Y,j})$ ) in  $\text{AMS}^*$  at  $j = 3, 7$  and in  $\text{MER}^*$  at  $j = 3-7$ , because of rescaling with less-than-unity  $\hat{\alpha}_{Y,j}$  values in Eq. (10). Indeed it is evident from Eq. (6) that it is possible to increase the noise variance and MSE when reducing the bias component of the MSE. This in turn leads to larger disagreement between INS and AMS in terms of RMSD and  $R$ , and the increased amplitudes of the noise observed in AMS in Fig. 6f.

The bulk and A/S CDF methods produced very similar results with each other, and also with their linear counterparts. There is signal and noise suppression, but the scale-level biases are retained. The signal components of  $Y^*$  are negatively biased at  $j = 3-7$  and positively biased at  $j = 8$ . The CDF-corrected  $\text{AMS}^*$  shows slightly better RMSD and  $R$  with INS, owing to the reduced noise variance and a reduced bias at  $\text{AMS}_8^{(a)*}$ .

In summary, the MRA of the bulk and A/S schemes highlights the deficiency of using a correction scheme that does not take into account the scale variability of bias and the differences in noise statistics between the two data. The im-



**Figure 6.** Time series of AMS SM at Kyeamba treated by various bias correction schemes. The use of WT-based de-noising has also been demonstrated in (g, h).

improvements in RMSD and correlation between the corrected  $Y^*$  and the reference  $X$  are somewhat superficial, masking the fact that the bias correction is limited to the coarsest scales. On the other hand, the A/S-based and MS methods can modify the original noise profiles in the data across the scales, by amplifying (or suppressing) noise in individual components (either  $Y_j$ ,  $Y_S$ , or  $Y_A$ ) with less-than (greater-than) unity pre-correction  $\alpha$ . This may be considered undesirable for an objective to produce more physically representative data with a simple error structure on the whole. Therefore, arguably, none of these methods is entirely satisfactory, in manners of not removing the multiplicative biases completely and/or changing error characteristics. From this viewpoint, the task of bias correction is seen as inseparable from that of noise reduction when considering MS (or A/S) bias correction, unless certain components in MRA were explicitly ignored.

## 7 Combining bias correction with wavelet de-noising

The last example presents an impetus to consider noise removal prior to bias correction and produce a simpler error structure in the bias-corrected data  $Y^*$ . Critically, TC provides noise and signal estimates that can be used for de-noising through thresholding of wavelet coefficients  $p_{jk}$ . The basic rationale for wavelet thresholding (WT) is that insignificant detail coefficients are likely due to noise, while significant ones are related to the signal component. Thus, a coefficient is eliminated if its magnitude is less than a given

threshold  $\lambda_p$ ; otherwise, it is modified according to a transformation function  $\Gamma(p_{jk})$  to remove the influence of the noise (Donoho and Johnstone, 1994).

One commonly used transformation is soft thresholding (Donoho, 1995), where the coefficients are modified according to

$$\Gamma_{\lambda_p}(p_{jk}) = \text{sign}(p_{jk}) \max(|p_{jk} - \lambda_p|, 0). \quad (11)$$

Such de-noising filters have near-optimal properties in the minmax sense. We follow the *BayesShrink* rule of Chang et al. (2000) to define a set of scale-dependent threshold values using

$$\lambda_{p,j} = \frac{\text{var}(\epsilon_{p,j})}{\alpha_{p,j} \text{std}(f_j)}, \quad (12)$$

where the variances are provided by TC (Appendix B). This choice of threshold is near optimal under the assumption that the signal is generalised Gaussian distributed and the noise is Gaussian. When the threshold value for  $j = 1-2$  could not be estimated using TC, CDF matching was applied. While TC is an ideal error estimator, alternative estimators for the threshold values are also available to make the de-noising a stand-alone process (Donoho and Johnstone, 1994; Donoho, 1995). After WT, the de-noised time series is constructed via inverse DWT of the modified coefficients, and can be subsequently corrected for biases. Combining with the MS bias correction scheme, biased-corrected, de-noised data are generated via

$$Y^* = \frac{Y'}{\hat{\alpha}'_Y} + \sum_{j=1}^{j_0} \sum_{k=1}^{n_j} \frac{\Gamma_{\lambda_{p,j}}(Y_{jk})}{\hat{\alpha}_{Y,j}} \psi_{jk}. \quad (13)$$

The prescription, which is essentially a two-stage operation, was applied to AMS for comparisons with the previous results. The first stage of de-noising leads to smoothing of the time series, improved  $R$  with INS by 0.05, and reduced RMSD by  $0.02 \text{ m}^3 \text{ m}^{-3}$ . The actual SM variability has become more apparent in Fig. 6g. Over-smoothing can occur due to our inability to properly distinguish signal from noise in  $\text{AMS}_1$  and  $\text{AMS}_2$  where the signal-to-noise ratio is very low. However, without the second stage of bias correction, the dynamic ranges of de-noised AMS and INS are visibly different, such that the improvement in RMSD with INS is limited. Combining WT and MS leads to improvement in both metrics of  $\text{RMSD} = 0.048 \text{ m}^3 \text{ m}^{-3}$  and  $R = 0.711$ , with Fig. 6h confirming that the reduced noise was not amplified by the MS rescaling.

## 8 Conclusions

This work combines MRA and TC in a new analysis framework with increased capacity to provide a more comprehensive view of the inter-data relations on short and long timescales. TC (or CDF) rescaling can be exploited on individual scales to reduce scale-specific multiplicative biases, and provide “prior” knowledge of noise for calibrating a WT-based de-noising filter. As a demonstration of principle, these methods are applied to SM data from in situ and satellite sensors and a land surface model. Using MRA, we found that the three data exhibit significantly different wavelet spectra and variable degrees of agreement on different timescales. On fine scales, the contribution of noise is most prominent, undermining the correlation between the data sets. By contrast, the biases are most apparent on coarse scales. Furthermore, these biases are non-systematic across timescales in the study region and across spatial locations over Australia, and the signal-to-noise ratios vary with scales and between the various data, pointing to the need to use correction schemes that are capable of handling such complexities.

These observations raised concerns about the possible inadequate treatment of SM data in the linear regime, even with anomaly/seasonal decomposition. Scale-by-scale linear rescaling based on a MRA-TC analysis framework offers a more comprehensive treatment of different biases on different scales, but error characteristics are found to be modified by variable rescaling, and can lead to undesirable noise amplification. The method of removing biases and noise on individual scales offers a remedy, although a few caveats should be noted. First, TC analysis requires a strong instrument and large sample, and in cases where these prerequisites are not met, we resort to sub-optimal estimation and rescaling methods. Second, the issue of non-stationarity in errors and scaling has not been addressed so far, and this can lead to biased estimates of the correction parameters for rescaling and de-noising. Despite this, DWT offers additional degree of freedom in translation parameter  $k$  to accommodate non-stationarity. Third, given the theoretic viewpoint presented in this work, further evaluations based on assimilation of data treated by different schemes are still warranted to assess their practical impacts. Notwithstanding these factors, MRA-TC analysis can be an important tool to allow better characterisation of the inter-sensor differences and to develop more effective strategies in harmonising a broad range of observational data records in oceanography and hydrometeorology.

### Appendix A: Wavelet statistical analysis

MRA enables the (bulk) variance  $\text{var}(p)$  of a time series  $p$  to be decomposed into wavelet variances  $\text{var}(p_j)$  on different scales  $j$ . Analogous to a Fourier spectrum, the expansion of  $\text{var}(p)$  yields a wavelet spectrum and is given by

$$\text{var}(p) = \sum_{j=1}^J \text{var}(p_j) \quad (\text{A1})$$

$$= \text{var}\left(p_{j_0}^{(a)}\right) + \sum_{j=1}^{j_0} \text{var}(p_j) \quad (\text{A2})$$

where the variance of the approximation time series  $p_{j_0}^{(a)}$  can be expressed in terms of that of the detail time series  $p_j$ .

Similarly, wavelet covariance  $\text{cov}(X_j, Y_j)$  at a given  $j$  indicates the contribution of covariance between two time series  $(X, Y)$  on that scale. Specifically, the wavelet covariance on scale  $j$  can be expressed as

$$\text{cov}(X_j, Y_j) = \frac{1}{n_j} \sum_{k=1}^{n_j} X_{jk} Y_{jk}, \quad (\text{A3})$$

noting that there is an equivalence of computing (co)variance in the wavelet and time domains. To exclude the boundary influence of a finite-length time series and missing values in the time series, an estimator of the wavelet covariance can be constructed by excluding the coefficients affected by the boundaries and gaps, followed by renormalisation. In the paper, we find it more intuitive to report the wavelet correlation, namely

$$R(X_j, Y_j) = \frac{\text{cov}(X_j, Y_j)}{\sqrt{\text{var}(X_j) \text{var}(Y_j)}}. \quad (\text{A4})$$

### Appendix B: Multi-scale triple collocation

Starting with the scale-level affine model of Eqs. (4) and (5), the associated scaling coefficients ( $\alpha'_p, \alpha_{p,j}$ ) and error variances ( $\text{var}(\epsilon'_p), \text{var}(\epsilon_{p,j})$ ) for each scale can be estimated using TC. We use solutions of Su et al. (2014a) for the data triplet  $p = \{X, Y, Z\}$  on each scale separately: with  $X$  as the reference by setting  $\alpha_{X,j}, \alpha'_X = 1$ ,

$$\hat{\alpha}_{Y,j} = \frac{\text{cov}(Y_j, Z_j)}{\text{cov}(X_j, Z_j)}, \quad (\text{B1})$$

$$\hat{\alpha}_{Z,j} = \frac{\text{cov}(Y_j, Z_j)}{\text{cov}(X_j, Y_j)}, \quad (\text{B2})$$

$$\hat{\text{var}}(\epsilon_{p,j}) = \text{var}(p_j) - \frac{\text{cov}(p_j, q_j) \text{cov}(p_j, r_j)}{\text{cov}(q_j, r_j)}, \quad (\text{B3})$$

$$\hat{\text{var}}(f_j) = \text{var}(X_j) - \text{var}(\epsilon_{p,j}) \quad (\text{B4})$$

where  $q$  and  $r$  are also data labels, but  $p \neq q \neq r$ . The hat notation is used throughout the paper to distinguish estimates from true values. It can be shown that, in probability, TC yields unbiased estimates whereby  $\hat{\alpha}_{p,j} = \alpha_{p,j}$ ,  $\hat{\text{var}}(\epsilon_{p,j}) = \text{var}(\epsilon_{p,j})$ , and  $\hat{\text{var}}(f_j) = \text{var}(f_j)$ . These expressions were used to compute the results in Fig. 5 and the threshold values for wavelet de-noising. When TC does not produce physically meaningful estimates from negative or small covariance due to weak instruments and possible inadequacy of the considered signal and noise model, the OLS estimator was used,

$$\hat{\alpha}_{Y,j}^{\text{OLS}} = \frac{\text{cov}(X_j, Y_j)}{\text{var}(X_j)}, \quad (\text{B5})$$

although its estimates are biased ( $\hat{\alpha}_{Y,j}^{\text{OLS}} < \alpha_{Y,j}$ ) for our purpose, due to the extraneous contribution of noise variance in the denominator. Similarly the VAR estimator can be used,  $\hat{\alpha}_{Y,j}^{\text{VAR}} = \sqrt{\text{var}(Y_j)/\text{var}(X_j)}$ , but it is also biased.

*Acknowledgements.* We thank Wade Crow for valuable discussions and Clara Draper for her critiques of the early drafts. We acknowledge gratefully the feedback of Simon Zwieback, two anonymous reviewers, Wolfgang Wagner, and Editor Niko Verhoest in the refinement of our manuscript. We also thank all who contributed to the data sets used in this study. Kyeamba in situ data were produced by colleagues at Monash University and the University of Melbourne who have been involved in the OzNet programme. AMSR-E data were produced by Richard de Jeu and colleagues at Vrije University Amsterdam and NASA. The MERRA-Land data set was provided by NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). The land cover/use map was produced by merging land cover (Lymburner et al., 2010) and land use (Australian Bureau of Rural Science, 2010) data sets. The recalibrated precipitation data of the Australian Water Availability Project (AWAP) (Jones et al., 2009) were obtained from the Australian Bureau of Meteorology. National soil data (McKenzie et al., 2005) were provided by the Australian Collaborative Land Evaluation Program ACLEP, endorsed through the National Committee on Soil and Terrain NCST (<http://www.clw.csiro.au/aclep>). The 9 s digital elevation map is obtained from Geoscience Australia (2008). This research was conducted with financial support from the Australian Research Council (ARC Linkage Project No. LP110200520) and the Bureau of Meteorology, Australia.

Edited by: N. Verhoest

## References

- Albergel, C., De Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., and Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, *Remote Sens. Environ.*, 118, 215–226, doi:10.1016/j.rse.2011.11.017, 2012.
- Australian Bureau of Rural Science: Land Use of Australia, version 4, 2005/2006, available at: [http://data.daff.gov.au/anrdl/metadata\\_files/pa\\_luav4g9abl07811a00.xml](http://data.daff.gov.au/anrdl/metadata_files/pa_luav4g9abl07811a00.xml) (last access: 24 July 2014), 2010.
- Brocca, L., Moramarco, T., Melone, F., Wagner, W., Hasenauer, S., and Hahn, S.: Assimilation of surface- and root-zone ASCAT soil moisture products into rainfall–runoff modeling, *IEEE T. Geosci. Remote*, 50, 2542–2555, doi:10.1109/TGRS.2011.2177468, 2012.
- Caires, S. and Sterl, A.: Validation of ocean wind and wave data using triple collocation, *J. Geophys. Res.*, 103, 3098, doi:10.1029/2002JC001491, 2003.
- Chang, S. G., Yu, B., and Yeterli, M.: Adaptive wavelet thresholding for imaging denoising and compression, *IEEE T. Image Process.*, 9, 1532–1546, doi:10.1109/83.862633, 2000.
- Daubechies, I.: *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, doi:10.1137/1.9781611970104, 1992.
- Donoho, D. L.: Denoising via soft thresholding, *IEEE T. Inform. Theory*, 41, 613–627, doi:10.1109/18.382009, 1995.
- Donoho, D. L. and Johnstone, I. M.: Ideal spatial adaptation via wavelet shrinkage, *Biometrika*, 81, 425–455, doi:10.1093/biomet/81.3.425, 1994.
- Draper, C., Reichle, R., De Lannoy, G., and Liu, Q.: Assimilation of passive and active microwave soil moisture retrievals, *Geophys. Res. Lett.*, 39, L04401, doi:10.1029/2011GL050655, 2012.
- Draper, C. S., Reichle, R. R., de Jeu, R. A., Naeimi, V., Parinussa, R. M., and Wagner, W.: Estimating root mean square errors in remotely sensed soilmoisture over continental scale domains, *Remote Sens. Environ.*, 137, 288–298, doi:10.1175/JHM-D-12-052.1, 2013.
- Drusch, M., Wood, E. F., and Gao, H.: Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture, *Geophys. Res. Lett.*, 32, L15403, doi:10.1029/2005GL023623, 2005.
- Entin, J. K., Robock, A., Vinnikov, K. Y., Hollinger, S. E., Liu, S., and Namkhai, A.: Temporal and spatial scales of observed soil moisture variations in the extratropics, *J. Geophys. Res.*, 105, 11865–11877, doi:10.1029/2000JD900051, 2000.
- Garcia, D.: Robust smoothing of gridded data in one and higher dimensions with missing values, *Comput. Stat. Data Anal.*, 54, 1167–1178, doi:10.1016/j.csda.2009.09.020, 2010.
- Geoscience Australia: GEODATA 9-Second DEM Version 3, available at: <http://www.ga.gov.au/metadata-gateway/metadata/record/66006/> (last access: 24 July 2014), 2008.
- GLACE Team, Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., and Chan, E.: Regions of strong coupling between soil moisture and precipitation, *Science*, 305, 1138–1140, doi:10.1126/science.1100217, 2004.
- Gruber, A., Dorigo, W. A., Zwieback, S., Xaver, A., and Wagner, W.: Characterizing Coarse-Scale Representativeness of in situ Soil Moisture Measurements from the International Soil Moisture Network, *Vadose Zone J.*, 12, doi:10.2136/vzj2012.0170, 2013.
- Janssen, P. A. E. M., Abdalla, S., Hersbach, H., and Bidlot, J.-R.: Error estimation of buoy, satellite, and model wave height data, *J. Atmos. Ocean. Tech.*, 24, 1665–1677, doi:10.1175/JTECH2069.1, 2007.
- Jones, D. A., Wang, W., and Fawcett, R.: High-quality spatial climate data-sets for Australia, *Aust. Meteorol. Oceanogr.*, 58, 233–248, 2009.
- Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K., and Puma, M. J.: On the nature of soil moisture in land surface models, *J. Climate*, 22, 4322–4335, doi:10.1175/2009JCLI2832.1, 2009.
- Kumar, S. V., Reichle, R. H., Harrison, K. W., Peters-Lidard, C. D., Yatheendradas, S., and Santanello, J. A.: A comparison of methods for a priori bias correction in soil moisture data assimilation, *Water Resour. Res.*, 48, W03515, doi:10.1029/2010WR010261, 2012.
- Legates, D. R., Mahmood, R., Levia, D. F., DeLiberty, T. L., Quiring, S. M., Houser, C., and Nelson, F. E.: Soil moisture: a central and unifying theme in physical geography, *Prog. Phys. Geogr.*, 35, 65–86, doi:10.1177/0309133310386514, 2011.
- Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., De Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P., and Van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, *Remote Sens. Environ.*, 123, 280–297, doi:10.1016/j.rse.2012.03.014, 2012.
- Loew, A. and Schlenz, F.: A dynamic approach for evaluating coarse scale satellite soil moisture products, *Hydrol. Earth Syst. Sci.*, 15, 75–90, doi:10.5194/hess-15-75-2011, 2011.

- Lymburner, L., Tan, P., Mueller, N., Thackway, R., Lewis, A., Thankappan, M., Randall, L., Islam, A., and Senarath, U.: 250 metre Dynamic Land Cover Dataset of Australia, 1st Edn., Geoscience Australia, Canberra, 2010.
- Mallat, S. G.: A theory for multiresolution signal decomposition: the wavelet representation, *IEEE T. Pattern Anal.*, 11, 674–693, doi:10.1109/34.192463, 1989.
- McKenzie, N. J., Jacquier, D. W., Maschmedt, D. J., Griffin, E. A., and Brough, D. M.: The Australian Soil Resource Information System: Technical Specifications, National Committee on Soil and Terrain Information/Australian Collaborative Land Evaluation Program, Canberra, 2005.
- Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., and Dolman, A. J.: Magnitude and variability of land evaporation and its components at the global scale, *Hydrol. Earth Syst. Sci.*, 15, 967–981, doi:10.5194/hess-15-967-2011, 2011.
- Mittelbach, H. and Seneviratne, S. I.: A new perspective on the spatio-temporal variability of soil moisture: temporal dynamics versus time-invariant contributions, *Hydrol. Earth Syst. Sci.*, 16, 2169–2179, doi:10.5194/hess-16-2169-2012, 2012.
- Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, *J. Geophys. Res.*, 113, F01002, doi:10.1029/2007JF000769, 2008.
- Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen–Geiger climate classification, *Hydrol. Earth Syst. Sci.*, 11, 1633–1644, doi:10.5194/hess-11-1633-2007, 2007.
- Reichle, R. H. and Koster, R. D.: Bias reduction in short records of satellite soil moisture, *Geophys. Res. Lett.*, 31, L19501, doi:10.1029/2004GL020938, 2004.
- Reichle, R. H., Koster, R. D., Dong, J., and Berg, A. A.: Global soil moisture from satellite observations, land surface models, and ground data: implications for data assimilation, *J. Hydrometeorol.*, 5, 430–442, doi:10.1175/1525-7541(2004)005<0430:GSMFSO>2.0.CO;2, 2004.
- Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P. P., Njoku, E. G., and Owe, M.: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), *J. Geophys. Res.*, 112, D09108, doi:10.1029/2006JD008033, 2007.
- Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Liu, Q., Mahanama, S. P. P., and Touré, A.: Assessment and enhancement of MERRA land surface hydrology estimates, *J. Climate*, 24, 6322–6338, doi:10.1175/JCLI-D-10-05033.1, 2011.
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA – NASA’s Modern-Era Retrospective Analysis for Research and Applications, *J. Climate*, 24, 3624–3648, doi:10.1175/JCLI-D-11-00015.1, 2011.
- Rodríguez-Iturbe, I.: Ecohydrology: a hydrologic perspective of climate-soil-vegetation dynamics, *Water Resour. Res.*, 36, 3–9, doi:10.1029/1999WR900210, 2000.
- Roebeling, R. A., Wolters, E. L. A., Meirink, J. F., and Leijnse, H.: Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data, *J. Hydrometeorol.*, 13, 1552–1566, doi:10.1175/JHM-D-11-089.1, 2013.
- Scipal, K., Drusch, M., and Wagner, W.: Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system, *Adv. Water Resour.*, 31, 1101–1112, doi:10.1016/j.advwatres.2008.04.013, 2008a.
- Scipal, K., Holmes, T., de Jeu, R., Naeimi, V., and Wagner, W.: 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, 2008b.
- Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., Grayson, R. B., Siriwardena, L., Chiew, F. H. S., and Richter, H.: The Murrumbidgee soil moisture network data set, *Water Resour. Res.*, 48, W07701, doi:10.1029/2012WR011976, 2012.
- Steelman, C. M., Endres, A. L., and Jones, J. P.: High-resolution ground penetrating radar monitoring of soil moisture dynamics: Field results, interpretation, and comparison with unsaturated flow model, *Water Resour. Res.*, 48, W09538, doi:10.1029/2011WR011414, 2012.
- Stoffelen, A.: Towards the true near-surface wind speed: Error modelling and calibration using triple collocation, *J. Geophys. Res.*, 103, 7755–7766, doi:10.1029/97JC03180, 1998.
- Su, C.-H., Ryu, D., Young, R. I., Western, A. W., and Wagner, W.: Inter-comparison of microwave satellite soil moisture retrievals over Murrumbidgee Basin, southeast Australia, *Remote Sens. Environ.*, 134, 1–11, doi:10.1016/j.rse.2013.02.016, 2013a.
- Su, C.-H., Ryu, D., Western, A. W., and W. Wagner, W.: De-noising of passive and active microwave satellite soil moisture time series, *Geophys. Res. Lett.*, 40, 3624–3630, doi:10.1002/grl.50695, 2013b.
- Su, C.-H., Ryu, D., Crow, W. T., and Western, A. W.: Beyond triple collocation: applications to satellite soil moisture, *J. Geophys. Res.-Atmos.*, 119, 6419–6439, doi:10.1002/2013JD021043, 2014a.
- Su, C.-H., Ryu, D., Crow, W. T., and Western, A. W.: Stand-alone error characterisation of microwave satellite soil moisture using a Fourier method, *Remote Sens. Environ.*, 154, 115–126, doi:10.1016/j.rse.2014.08.014, 2014b.
- Wang, G., Garcia, D., Liu, Y., de Jeu, R., and Dolman, A. J.: A three-dimensional gap filling method for large geophysical datasets: application to global satellite soil moisture observations, *Environ. Modell. Softw.*, 30, 139–142, doi:10.1016/j.envsoft.2011.10.015, 2012.
- Western, A. W., Grayson, R. B., and Blöschl, G.: Scaling of soil moisture: a hydrologic perspective, *Annu. Rev. Earth Pl. Sc.*, 30, 149–180, doi:10.1146/annurev.earth.30.091201.140434, 2002.
- Wright, P. G.: *The Tariff on Animal and Vegetable Oils*, Macmillan, New York, 1928.
- Yilmaz, M. T. and Crow, W. T.: The optimality of potential rescaling approaches in land data assimilation, *J. Hydrometeorol.*, 14, 650–660, doi:10.1175/JHM-D-12-052.1, 2013.

Zwieback, S., Scipal, K., Dorigo, W., and Wagner, W.: Structural and statistical properties of the collocation technique for error characterization, *Nonlin. Processes Geophys.*, 19, 69–80, doi:10.5194/npg-19-69-2012, 2012.

Zwieback, S., Dorigo, W., and Wagner, W.: Estimation of the temporal autocorrelation structure by the collocation technique with emphasis on soil moisture studies, *Hydrolog. Sci. J.*, 58, 1729–1747, doi:10.1080/02626667.2013.839876, 2013.