Hydrology and Earth System Sciences

Error characterisation of global active and passive microwave soil moisture datasets

W. A. Dorigo[1], **K. Scipal**[2], **R. M. Parinussa**[3], **Y. Y. Liu**[3,4,5], **W. Wagner**[1], **R. A. M. de Jeu**[3], and **V. Naeimi**[1,6]

[1]Vienna University of Technology, Institute of Photogrammetry and Remote Sensing, Vienna, Austria
[2]ESA, ESTEC, Noordwijk, The Netherlands
[3]VU University Amsterdam, Faculty of Earth and Life Sciences, Department of Hydrology and Geo-Environmental Sciences, The Netherlands
[4]School of Civil and Environmental Engineering, University of New South Wales, Sydney, Australia
[5]CSIRO Land and Water, Black Mountain Laboratories, Canberra, Australia
[6]German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Wessling, Germany

Abstract. Understanding the error structures of remotely sensed soil moisture observations is essential for correctly interpreting observed variations and trends in the data or assimilating them in hydrological or numerical weather prediction models. Nevertheless, a spatially coherent assessment of the quality of the various globally available datasets is often hampered by the limited availability over space and time of reliable in-situ measurements. As an alternative, this study explores the triple collocation error estimation technique for assessing the relative quality of several globally available soil moisture products from active (ASCAT) and passive (AMSR-E and SSM/I) microwave sensors. The triple collocation is a powerful statistical tool to estimate the root mean square error while simultaneously solving for systematic differences in the climatologies of a set of three linearly related data sources with independent error structures. Prerequisite for this technique is the availability of a sufficiently large number of timely corresponding observations. In addition to the active and passive satellite-based datasets, we used the ERA-Interim and GLDAS-NOAH reanalysis soil moisture datasets as a third, independent reference. The prime objective is to reveal trends in uncertainty related to different observation principles (passive versus active), the use of different frequencies (C-, X-, and Ku-band) for passive microwave observations, and the choice of the independent reference dataset (ERA-Interim versus GLDAS-NOAH).

The results suggest that the triple collocation method provides realistic error estimates. Observed spatial trends agree well with the existing theory and studies on the performance of different observation principles and frequencies with respect to land cover and vegetation density. In addition, if all theoretical prerequisites are fulfilled (e.g. a sufficiently large number of common observations is available and errors of the different datasets are uncorrelated) the errors estimated for the remote sensing products are hardly influenced by the choice of the third independent dataset. The results obtained in this study can help us in developing adequate strategies for the combined use of various scatterometer and radiometer-based soil moisture datasets, e.g. for improved flood forecast modelling or the generation of superior multi-mission long-term soil moisture datasets.

1 Introduction

In recent years, an increasing number of global soil moisture products have become available from past and present passive and active coarse resolution satellite microwave sensors. Altogether, these datasets span a period of more than 30 years (Table 1). Knowing the quality of the different datasets and understanding the various error sources (sensor calibration, retrieval errors, model parameterisation, etc.) contributing to the observed soil moisture variations is indispensable if one wishes to draw conclusions on trends or anomalies in the datasets, e.g. in relation to climate change (Liu et al., 2009). But also other applications, like the assimilation of remotely

Correspondence to: W. A. Dorigo
(wd@ipf.tuwien.ac.at)

Table 1. Operational products in the field of global monitoring of soil moisture using active and passive satellite microwave instruments (sorted according to product release date).

Sensor	Producer soil moisture product	Dataset availability	First product release	Reference
ERS-1/2 scatterometer	Vienna University of Technology (TU Wien)	1991–2007	2002	Scipal et al. (2002); Wagner et al. (2007)
AMSR-E radiometer	US National Snow and Ice Data Center (NSIDC)	2002–present	2003	Njoku et al. (2003)
AMSR-E radiometer	Japanese Aerospace Exploration Agency (JAXA)	2002–present	2004	Koike et al. (2004)
AMSR-E and TRMM-TMI radiometers	United States Department of Agriculture (USDA)	2002–present	2007	Jackson (1993)
ERS-1/2 scatterometer	Centre d'Etudes des Environnements Terrestre et Planétaires	1991–present	2008	Zribi et al. (2003)
Windsat radiometer	US navy	2003–present	2008	Li et al. (2010)
SMMR, SSM/I, TRMM-TMI, AMSR-E and WindSat radiometers	Vrije Universiteit Amsterdam (VUA) and NASA	1979–present	2008	Owe et al. (2008)
Soil Moisture and Ocean Salinity mission (SMOS)	European Space Agency (ESA)	2010–present	2010	Wigneron et al. (2007)

sensed soil moisture in flood forecasting (Brocca et al., 2010) or numerical weather prediction models (Drusch, 2007; Scipal et al., 2008a; Mahfouf, 2010) require accurate estimates of the quality of the observations.

Most of the globally available microwave-based soil moisture products have been intensively validated using in-situ observations (e.g. Wagner et al., 2007; Jeu et al., 2008; Gruhier et al., 2010; Jackson et al., 2010). Even though the quality of the datasets can be established fairly accurately for the locations of the in-situ stations, available ground observations are restricted to a few locations worldwide and often cover only limited observation periods. In addition, reliable error estimation is complicated by representativeness and scaling errors, which can be larger than the actual retrieval error (Martínez-Fernández and Ceballos, 2005). Also, differences in observation times and depths, and inaccuracies of the in-situ measurements may lead to faulty interpretations of the obtained validation results (Gruhier et al., 2010).

In contrast to the locally confined in-situ validations, error propagation methods can provide a more global picture of the uncertainty of soil moisture datasets. Error propagation techniques assess the uncertainty of model estimates resulting from errors in the input variables. Naeimi et al. (2009) used a combined analytical Gaussian error propagation method and numerical propagation approach based on Monte Carlo sim-

ulation to estimate the uncertainty in soil moisture retrievals obtained from scatterometers using the TU Wien method (Wagner et al., 1999). Parinussa et al. (2010) found an analytical solution for estimating the uncertainties of soil moisture estimates from radiometers based on the LPRM model (Owe et al., 2008). A big advantage of error propagation techniques is that they allow for calculating an error estimate for each individual observation. However, the uncertainties obtained by error propagation methods only account for random errors in the model input variables but do not tell if the model itself is correct. Therefore the uncertainties obtained for different models are difficult to compare quantitatively.

Recently, Scipal et al. (2008b) introduced the triple collocation method in the field of satellite-based soil moisture research. The triple collocation method allows a simultaneous estimation of the error structure and the cross-calibration of a set of at least three linearly related datasets with uncorrelated errors (Stoffelen, 1998). By applying triple collocation to a combination of TRMM-TMI radiometer data, ERS scatterometer data, and modelled ERA-Interim reanalysis soil moisture, Scipal et al. (2008b) obtained realistic error estimates and were able to successfully distinguish spatial error trends of retrieved soil moisture. Miralles et al. (2010) successfully applied the triple collocation technique to soil moisture data products extracted from passive microwave remote

sensing, land surface modelling and high density ground-based observations with the goal of explicitly estimating the spatial sampling uncertainty of coarse-scale soil moisture estimates derived from ground observations.

This study connects to the work of Scipal et al. (2008b) and uses the triple collocation technique to establish the uncertainty of various recent passive (radiometer) and active (scatterometer) microwave soil moisture products. Herein, the prime objectives will be to reveal trends in uncertainty related to different observation principles (passive versus active), the use of different frequencies (C-, X-, and Ku-band) for passive microwave observations, and the choice of a third independent reference dataset. To address the latter we will repeat the triple collocation using two different globally available modelled reanalysis datasets (ERA-Interim and GLDAS-NOAH). Results of this study will be used for developing appropriate strategies for combining multiple satellite-based soil moisture products into a merged product (Liu et al., 2010).

2 Data

2.1 Scatterometer data

The Advanced Scatterometer (ASCAT) on MetOp-A operates in C-band (5.255 GHz) at VV polarisation and is operational since October 2006. Six radar antenna beams illuminate a continuous ground swath at six different azimuth angles (at both sides of the platform 45°, 90°, and 135° sideward from the direction of the satellite motion). Incidence angles range from 25° to 64° while the measurements used in this study have a spatial resolution of 50 km. The descending and ascending local equatorial crossing times of ASCAT are 09:30 and 21:30, respectively. The combined use of ascending and descending mode observations in this paper leads to a nearly daily revisit frequency at the Equator.

The backscatter measurements are converted to soil moisture estimates by applying the TU Wien soil moisture retrieval algorithm (Wagner et al., 1999; Naeimi et al., 2009). To correct for the effects of plant growth and decay, the method uses the multi-incidence angle measurement capability of the sensor to extract the vegetation sensitive signature from the backscatter observations. A soil moisture index is then retrieved by scaling each observation between dry and wet backscatter references representing the historically lowest and highest observed backscatter values, respectively. This results in a relative measure of surface (<2 cm) soil moisture that ranges between 0 (wilting point) and 100% (saturation). At the time of writing, the ASCAT soil moisture was available for the period January 2007 to December 2008.

2.2 Radiometer data

Since June 2002, the Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E) aboard the Aqua satellite provides a nearly daily global coverage. The instrument scans the Earth surface at an incidence angle of 55° while radiance is measured at six frequencies. The two frequencies considered in this study are the dual polarized C-band operated at 6.9 GHz and the dual polarized X-band observations operated at 10.7 GHz. Spatial resolutions are 73×43 km and 51×30 km for the C- and X-band, respectively. In this study we only use AMSR-E night-time observations (01:30 equatorial local crossing time), as it was shown that these are better suited for retrieving soil moisture than day-time observations (Jeu et al., 2008). This leads to a reduced revisit interval of approximately 2 days depending on the latitude.

The Special Sensor Microwave Imager (SSM/I) is found on board a series of Defense Meteorological Satellite Program (DMSP) platforms. The first satellite was launched in July 1987, whereas the last one was launched in October 2003. Several platforms are still operational. The SSM/I sensor operates in a dual polarization mode at four frequencies but only the Ku-band (19.3 GHz) will be considered in this study. In this band brightness temperatures are measured at a footprint size is 69×43 km. Only descending mode observations from the F13 satellite (local Equator crossing at 05:30 and a revisit frequency of about 3 days) will be used in this study for the same reason as for AMSR-E.

The brightness temperatures measured by AMSR-E and SSM/I are converted to volumetric surface soil moisture ($m^3 m^{-3}$) applying the Land Parameter Retrieval Model developed by NASA and the VU University of Amsterdam (LPRM; Owe et al., 2008). LPRM is based on the solution of a microwave radiative transfer model and uses both the horizontal and vertical polarized observations at C-, X, or Ku-band to solve for soil moisture and vegetation optical depth. The land surface temperature is calculated from the vertical polarized Ka band observations (Holmes el al., 2009). Theoretically, soil moisture values can range between 0.0–1.0 $m^3 m^{-3}$, whereas in practice, usually they do not exceed 0.6 $m^3 m^{-3}$ (De Jeu et al., 2008).

2.3 Reanalysis data

2.3.1 ERA-Interim

The ERA-Interim reanalysis dataset contains consistent atmosphere and surface analyses for the period from 1989 to present based on the ECMWF Integrated Forecast System (IFS) model (Simmons et al., 2007). The reanalysis assimilates various types of observations including satellite and ground based measurements. This system runs at T255 spectral resolution (~80 km horizontal resolution) with 91 vertical levels. In the IFS, land surface processes are described by

the Tiled ECMWF Scheme for Surface Exchanges over Land (TESSEL; Viterbo and Beljaars, 1995). In TESSEL soil processes are calculated in four layers. The lower boundary of each layer is at 0.07, 0.28, 1.0 and 2.68 m depth, respectively. To keep the land surface model simple, TESSEL uses a globally uniform soil type with fixed soil hydraulic parameters. Saturation is prescribed with a value of 0.472 m^3 m^{-3}, field capacity with 0.323 m^3 m^{-3} and the wilting point with 0.171 m^3 m^{-3}. Soil moisture estimates are provided at 00:00, 06:00, 12:00 and 18:00 UTC.

2.3.2 GLDAS-NOAH

From the year 2000 onwards, the Noah model from the Global Land Data Assimilation System (GLDAS) provides soil moisture and other atmospheric and land surface variables at a 3-h time interval for a regular global grid with a spatial resolution of 0.25° (Rodell et al., 2004). The model is forced by a combination of NOAA/GDAS atmospheric analysis fields, spatially and temporally disaggregated NOAA Climate Prediction Center Merged Analysis of Precipitation (CMAP) fields, and observation-based downward shortwave and longwave radiation fields derived using the method of the Air Force Weather Agency's Agricultural Meteorological system. The soil profile is represented by four vertical layers with a lower boundary of 0.10, 0.40, 1.00, and 2.00 m, respectively. Soil moisture is provided in kg m^{-3} which can easily be converted into volumetric water content in m^3 m^{-3}. The Noah model uses the same soil property dataset as LPRM (http://ldas.gsfc.nasa.gov/gldas/GLDASsoils.php), which is based on the Food and Agriculture Organization (FAO) Soil Map of the World linked to a global database of over 1300 soil samples. Even though theoretically possible, soil moisture values rarely drop below 0.05 m^3 m^{-3} or exceed 0.45 m^3 m^{-3}. Soil moisture and other fields are estimated 8 times per day (00:00, 03:00, 06:00, 09:00, 12:00, 15:00, 18:00, 21:00 UTC). Data generated by the GLDAS-NOAH model are publicly available from ftp://agdisc.gsfc.nasa.gov/data/s4pa/.

2.4 Spatial and temporal collocation of datasets

For the time period where observations were available for all three sources (i.e. 1 January 2007–31 December 2008) the data were binned to daily files, centred at 00:00 UTC, and collocated to a 0.25° regular grid using a nearest neighbour resampling.

3 Triple collocation

3.1 Overview of theory

Suppose three estimates Θ_{SCAT} (scatterometer-derived soil moisture), Θ_{RAD} (radiometer-derived soil moisture) and Θ_{MOD} (modelled/reanalysis soil moisture) relate to hypothetical true soil moisture Θ in a linear fashion (Stoffelen, 1998; Scipal et al., 2008b):

$$\Theta_{\text{SCAT}} = \alpha_{\text{SCAT}} + \beta_{\text{SCAT}}\Theta + r_{\text{SCAT}}$$
$$\Theta_{\text{RAD}} = \alpha_{\text{RAD}} + \beta_{\text{RAD}}\Theta + r_{\text{RAD}} \qquad (1)$$
$$\Theta_{\text{MOD}} = \alpha_{\text{MOD}} + \beta_{\text{MOD}}\Theta + r_{\text{MOD}}$$

where r_{SCAT}, r_{RAD} and r_{MOD} denote the residual errors in the estimates of Θ_{SCAT}, Θ_{RAD}, and Θ_{MOD} and α_X and β_X (with subscript X standing for SCAT, RAD, and MOD, respectively) represent the calibration constants. Goal of the triple collocation is to find an estimate of r_{SCAT}, r_{RAD} and r_{MOD}. From Eq. (1) we can first eliminate the calibration constants by introducing the new variables $\Theta_X^* = \Theta_X/\beta_X - \alpha_X/\beta_X$ and $r_X^* = r_X/\beta_X$, and then eliminate the unknown truth in order to obtain Eq. (2):

$$\Theta_{\text{SCAT}}^* - \Theta_{\text{RAD}}^* = r_{\text{SCAT}}^* - r_{\text{RAD}}^*$$
$$\Theta_{\text{SCAT}}^* - \Theta_{\text{MOD}}^* = r_{\text{SCAT}}^* - r_{\text{MOD}}^* \qquad (2)$$
$$\Theta_{\text{RAD}}^* - \Theta_{\text{MOD}}^* = r_{\text{RAD}}^* - r_{\text{MOD}}^*$$

By cross-multiplying the equations of Eq. (2) and assuming that the residual errors r_{SCAT}, r_{RAD}, and r_{MOD} are uncorrelated (i.e. the residual covariances become 0), we obtain a direct estimate of the variance of residual errors $e_X^{*2} = \langle r_X^{*2} \rangle$ if we average over a sufficiently large sample population (indicated by square brackets). The error variances are hence fully determined by three independent, calibrated soil moisture estimates:

$$e_{\text{SCAT}}^{*2} = \langle (\Theta_{\text{SCAT}}^* - \Theta_{\text{RAD}}^*)(\Theta_{\text{SCAT}}^* - \Theta_{\text{MOD}}^*) \rangle$$
$$e_{\text{RAD}}^{*2} = \langle (\Theta_{\text{SCAT}}^* - \Theta_{\text{RAD}}^*)(\Theta_{\text{RAD}}^* - \Theta_{\text{MOD}}^*) \rangle \qquad (3)$$
$$e_{\text{MOD}}^{*2} = \langle (\Theta_{\text{SCAT}}^* - \Theta_{\text{MOD}}^*)(\Theta_{\text{RAD}}^* - \Theta_{\text{MOD}}^*) \rangle$$

3.2 Implementation

In this study we follow a modification of the input data as proposed by Miralles et al. (2010). Instead of using the original absolute soil moisture values, we base our analysis on soil moisture anomalies from the long-term climatology. One should be aware that feeding the triple collocation with anomalies instead of absolute soil moisture values provides a different type of information on the dataset performance. While using absolute values provides information on the capability of the soil moisture products to estimate absolute soil moisture levels, the anomaly-based approach gives us more accurate information on the ability of the different datasets to capture single events of drying and wetting (e.g. due to rainfall). As a consequence, the anomaly-based approach tells us less about absolute deviations between datasets, e.g. like induced by a deviating seasonality. The choice of an anomaly-based approach in this study is motivated by the fact that the LPRM soil moisture product from SSM/I at many locations shows a seasonality that is different from the other

datasets considered. This is caused by the higher sensitivity of the Ku-band to atmospheric water vapour and vegetation. Whereas the use of absolute values would only tell us that the SSM/I products deviate from the other datasets (i.e. would show high errors), the use of anomalies still provides us meaningful information about the capability of this product in capturing single events.

If estimated soil moisture is assumed to be a sum of the climatology mean and an anomaly component it can be written as:

$$\Theta_X(t) = \Theta_X^{DOY}(t) + \hat{\Theta}_X(t) \tag{4}$$

where $\Theta_X^{DOY}(t)$ is the climatological expectation for soil moisture at the day-of-year (DOY) associated with time-step t, and $\hat{\Theta}_X(t)$ is the anomaly relative to this expectation. Θ_X^{DOY} values are obtained by averaging all valid soil moisture estimates in the period of observation found on the respective DOY. The resulting seasonality curves are smoothed by applying a moving window averaging with a kernel size of 31 days centred on the particular DOY (Crow et al., 2010a). Due to the short operations period of ASCAT, both ERS and ASCAT soil moisture estimates were used to obtain a reliable seasonality for ASCAT. It is feasible to this because the ERS-based soil moisture product relies on the same observation principles and retrieval concepts as the ASCAT-based product.

Finally, to estimate e_{SCAT}^{*2}, e_{RAD}^{*2} and e_{MOD}^{*2}, one has to solve for the calibration coefficients in Eq. (1). Since the real truth is always unknown, one has to choose one of the datasets as a reference. This means that e_{SCAT}^{*2}, e_{RAD}^{*2}, and e_{MOD}^{*2} will be expressed in the observation space of the selected reference dataset. The choice of the reference dataset does not influence the relative magnitude of the errors which theoretically can be scaled from one observation space into the other. Unless stated otherwise we choose Θ_{MOD} as a reference. Scipal et al. (2008b) used an iterative linear least-squares approximation to solve for the calibration constants in Eq. (1). For the linear fit they assumed errors in both the x (Θ_X) and y (Θ_Y) variables.

As in our case we are dealing with anomalies from the long-term mean, we can apply a simple rescaling of soil moisture anomaly $\hat{\Theta}_X(t)$ into the observation space of the reference dataset R based on the average and variance (VAR) of the anomalies (Miralles et al., 2010):

$$\hat{\Theta}_X^*(t) = \bar{\hat{\Theta}}_R + \sqrt{\frac{VAR(\hat{\Theta}_R)}{VAR(\hat{\Theta}_X)}} \cdot \left(\hat{\Theta}_X(t) - \bar{\hat{\Theta}}_X\right) \tag{5}$$

where $\hat{\Theta}_X^*(t)$ is the rescaled soil moisture anomaly for time step t. These rescaled values can now be inserted into Eq. (3). Notice that for the reference dataset $\hat{\Theta}_X(t)$ and $\hat{\Theta}_X^*(t)$ are equal.

Theoretically, an infinite number of common observations (i.e. at time step t observations should be available for all

three datasets) are required to obtain unbiased estimates of e_X^{*2}. Statistical tests revealed that a minimum number of 100 triplets is a good trade-off (Scipal et al., 2008b). Hence, areas with less than 100 triplets are masked in the results.

4 Results and discussion

4.1 Comparing scatterometer, radiometer and reanalysis data

Figure 1a–c show the triple collocation errors for a combination of ASCAT, AMSR-E C-band, and ERA-Interim soil moisture estimates. The errors e_X^* (i.e. the square-root of the values obtained from Eq. (3) are expressed in the dynamic range and units of the ERA-Interim reanalysis soil moisture data. The results of the error estimation suggest that all three datasets are characterised by a relatively low error. The mean global error is $0.017\,\mathrm{m^3\,m^{-3}}$ for the AS-CAT (e_{SCAT}^*), $0.019\,\mathrm{m^3\,m^{-3}}$ for the AMSR-E C-band observations (e_{RAD}^*) and $0.018\,\mathrm{m^3\,m^{-3}}$ for ERA-Interim (e_{MOD}^*). The low error is partly due to the low dynamic range between wilting point and saturation of the ERA-Interim soil moisture dataset which was used as reference. This dynamic range is known to be generally too low (Balsamo et al., 2009). Even though the errors are expressed in the dynamic range of ERA-Interim, it is important to realise that this dataset does not profit relative to the other datasets from having been selected as the reference. This means that the relative magnitude of the errors remains the same if one of the other datasets would have been used as a reference.

The average errors are lower than those obtained by Scipal et al. (2008b) for a combination of the ERS-2 scatterometer ($e_{SCAT}^* = 0.028\,\mathrm{m^3\,m^{-3}}$), the TMI radiometer ($e_{RAD}^* = 0.046\,\mathrm{m^3\,m^{-3}}$) and ERA-Interim soil moisture ($e_{MOD}^* = 0.020\,\mathrm{m^3\,m^{-3}}$). There are various reasons for this. The lower average error obtained in our study for the scatterometer can be ascribed mainly to the improved design of ASCAT with respect to the ERS series. Most remarkable are the lower errors obtained for the AMSR-E radiometer compared to the TMI results obtained by Scipal et al. (2008b). Substantial improvements can be attributed to improvements of the LPRM model and the better radiometric accuracy of AMSR-E, but large part is also explained by the larger wavelengths of AMSR-E C-band observations compared to TMI. This makes AMSR-E-based retrievals less sensitive to vegetation structure (See also Sect. 4.2). The average global error calculated for ERA-Interim is fairly consistent with the average error obtained by Scipal et al. (2008b). This confirms the robustness of the triple collocation method and indicates a high global spatial consistency of ERA-Interim soil moisture, as the study of Scipal et al. (2008b) did not cover northern latitudes due to the limited coverage of the TMI sensor. The fact that still there is a difference observed for ERA.
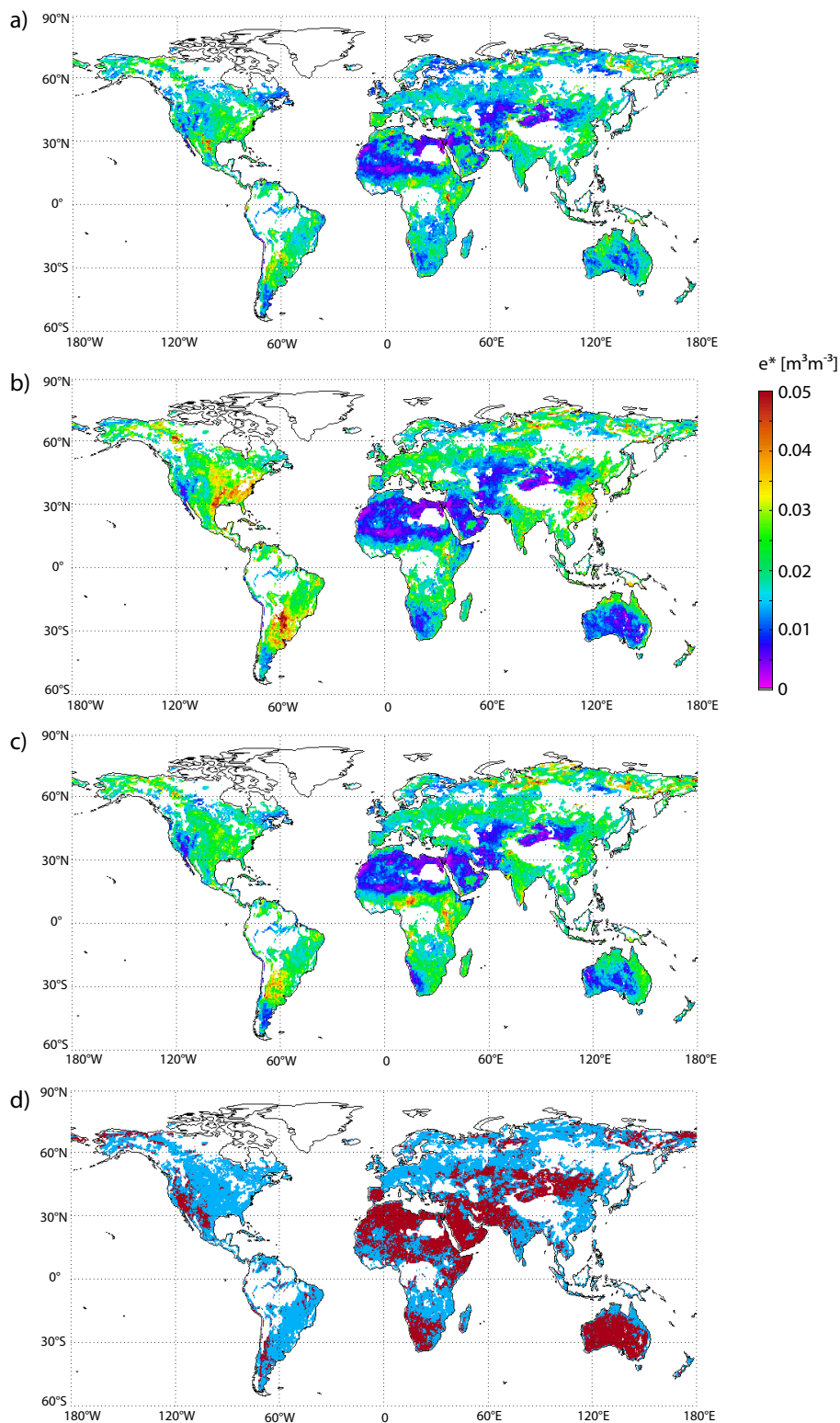
Fig. 1. Spatial errors of **(a)** ASCAT, **(b)** AMSR-E C-band, and **(c)** ERA-Interim surface soil moisture estimates. Errors are expressed in the climatology of ERA-Interim. **(d)** shows the areas in which either ASCAT (blue) or AMSR-E (red) shows the smallest error value. White areas indicate areas for which less than 100 common observations are available.

However, a quantitative interpretation and a direct comparison of the obtained errors with the results of Scipal et al. (2008b) should be taken with precaution as we based our calculations on anomalies from the long term seasonality instead of using absolute soil moisture values. Doing so, the magnitude of the computed errors may be affected in two ways. First of all, anomalies typically have a lower dynamic range than absolute soil moisture values. Nevertheless, this does not necessarily lead to smaller errors as removing the seasonal signal from the soil moisture series typically leads to a lower signal-to-noise ratio which has a negative impact on the error. Secondly, as already pointed out in Sect. 3.2, the two approaches provide us different types of error structures. The anomaly-based error estimates provide us information on the ability of the product to capture single events, but do not include any information on the systematic errors between datasets. Therefore, unless seasonalities perfectly match, errors are expected to be higher when absolute values are used. Additional research is needed to further quantify the effect of using anomalies instead of absolute values. Nonetheless, spatial patterns and relative error budgets can still give us viable information on the differences in performance between the various datasets.

Generally, error estimates are lowest in arid regions such as Southern Africa, mainland Australia, or Central Asia (Fig. 1a–c). This is explained by the very low amounts of precipitation received and hence the very low variability of soil moisture. The global picture would look different if relative instead of absolute errors were considered, as low errors in dry regions (low overall soil moisture content) have larger relative impact than in humid regions.

Despite the similar average errors of the three datasets, several characteristic differences in the spatial distribution of the errors can be observed between the datasets (Fig. 1a–c). In very dry areas (e.g. those of central Australia) error estimates derived for AMSR-E C–band are remarkably lower than error estimates derived for ASCAT and, to a smaller degree, than those obtained for ERA-Interim. In these regions the AMSR-E observations are hardly disturbed by vegetation which explains the low error estimates. The relatively high errors obtained for scatterometer data in these areas are a well-known phenomenon believed to be related to volume scattering effects in dry, loose sand and the systematic orientation of sand ripples and dunes over large areas leading to systematic influence of the azimuth viewing direction (Bartalis et al., 2006).

On the other hand, soil moisture derived from AMSR-E is prone to larger random errors in moderately to densely vegetated areas, like for instance found in south-eastern North America and northern Argentina. Vegetation affects passive microwave observations from above the canopy in two ways. First, vegetation will absorb or scatter the radiation emanating from the soil. Secondly, also the vegetation canopy itself emits radiation. These two effects tend to counteract each other. The observable soil emission will decrease with increased vegetation, while the emission from the vegetation canopy will increase. Under a sufficiently dense canopy, the emitted soil radiation will become totally masked, and the observed emissivity will be due largely to the vegetation (Owe et al., 2001). As similar vegetation interaction is expected for active microwave signals (Ulaby et al., 1982), it is suggested that the differences in errors over vegetation should be mainly attributed to the retrieval method. Recently, Crow et al. (2010b) pointed out that first order radiative transfer models are not able to accurately describe radiation attenuation in denser vegetation, especially for larger incidence angles. This finding would explain the shortcomings of LPRM, which is based on a simple linear radiative transfer model, in describing the higher order scattering that is very likely to occur in canopies with heavier vegetation cover. Hence, larger uncertainties in retrieved soil moisture would occur in these areas. In contrast, vegetation correction in the TU Wien algorithm is data-driven and therefore implicitly accounts for higher order scattering effects. And even though uncertainties in observed soil moisture increase with increasing vegetation density, effects are not as pronounced as for LPRM. However, more research and model comparison is needed to verify this hypothesis.

Figure 1d shows the areas for which either ASCAT (shown in blue) or AMSR-E (red) gives the lowest triple collocation errors. Such a map can be useful for ranking the different products in an attempt to merge the datasets (Liu et al., 2010). Nevertheless, the resulting Boolean map should be taken with precaution as, especially in transition areas, errors may be very similar and none of the products should be excluded on beforehand. In areas where less than 100 triplets are available (left blank in the image) it is expected that ASCAT would provide lower errors in moderately to densely vegetated areas while AMSR-E would show lower errors in dry areas. These assumptions could be used to fill the map in Figure 1d in order to obtain a complete global coverage.

4.2 Influence of radiometer observation wavelength

Figure 2 illustrates the influence of decreasing observation wavelength on the error structures obtained for radiometer observations. The triple collocation was based on a combination of ERA-Interim, ASCAT, and the respective radiometer dataset. On average, there is a clear increase in errors with decreasing observation wavelength, especially in areas characterised by moderate to dense vegetation cover, like in southeast Siberia. This behaviour can be explained by the fact that for the corresponding decrease in wavelength (i.e. 4.3, 2.8, and 1.6 cm for AMSR-E C-band, AMSR-E X-band and SSM/I Ku-band, respectively) the soil moisture signal emitted from the surface is increasingly absorbed by the vegetation canopy. For SSM/I Ku-band observations over moderate and dense canopies this usually implies that measured brightness temperatures no longer contain a detectable soil moisture signal. This is also the main reason of the reduced
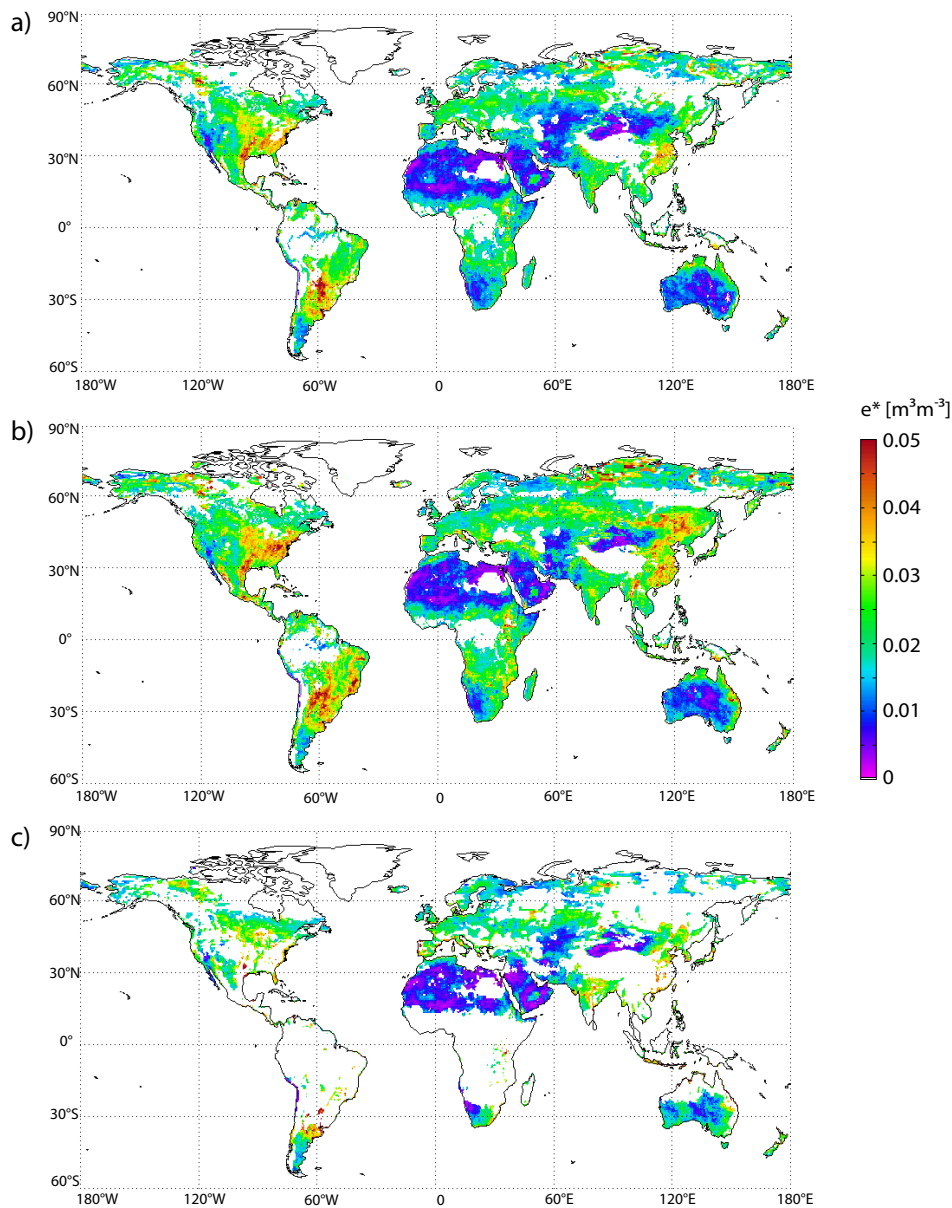
Fig. 2. Spatial errors e^*_{PAS} of **(a)** AMSR-E C-band, **(b)** AMSR-E X-band, and **(c)** SSM/I Ku-band observations obtained with triple colloca-tion based on ERA-Interim, ASCAT, and the respective radiometer dataset. Errors are expressed in the climatology of ERA-Interim.

spatial coverage seen in Fig. 2c for SSM/I as LPRM fails to converge in densely vegetated areas and hence these pixels are masked. The trends observed for the different observation bands correspond well to the trends in wavelength-related uncertainties of LPRM products obtained by error propaga-tion (Parinussa et al., 2010).

Despite the fact that many areas cannot be observed due to the increased sensitivity of the signal to vegetation, ob-servations in Ku-band can still be very valuable in arid to semi-arid areas. For example, in the desert areas of North Africa and the Middle East no significant difference can be observed between Fig. 2a–c which is an indication that in

such areas estimates from Ku-band are similarly accurate as those obtained from observations in lower frequency bands. This is also illustrated by comparing the average errors of each passive dataset over the areas where all datasets have more than 100 triplets. This area is approximately equal to the valid grid points in Fig. 2c. For this area, the average errors for AMSR-E C-band, AMSR-E X-band, and SSM/I Ku-band are 0.0158, 0.0177, and $0.0175\,\mathrm{m}^{-3}\,\mathrm{m}^{-3}$, respec-tively. This indicates that for this area there is only a slight loss in accuracy of the Ku-band compared to C-band and a nearly similar accuracy with respect to the X-band observa-tions. However, the conclusions should be taken with caution

as the trends in error structures obtained for the AMSR-E and SSM/I sensors do not rely only on differences in observation wevelength but also on differences in instrument design, radiometric accuracy, overpass times and so on.

4.3 Influence of reference dataset

So far, the results presented were based on ASCAT, one of the radiometer datasets, and the ERA-Interim reanalysis dataset. Theoretically, the choice of the third dataset should not influence the results obtained for the other datasets, given the errors of all three datasets are uncorrelated. To test this theoretical hypothesis we repeated the triple collocation with ASCAT and AMSR-E C-band while using GLDAS-NOAH instead of ERA-Interim as a third, independent, dataset. To be able to directly compare the error structures based on the different reanalysis datasets, errors must be expressed in the same units and dynamic range. For this reason, the AMSR-E C-band dataset was used as a reference against which the other observations were rescaled.

Figure 3a and b show the errors obtained for AMSR-E C-band observations, using ERA-Interim and GLDAS-NOAH as the third, independent data source, respectively. The spatial distributions of the error structures are very similar for most areas. Comparable results were obtained for ASCAT error estimates (results not shown). If we look at the areas where differences between the two combinations are largest (Fig. 3c) we see that these commonly coincide with the areas where least observations are available, like for instance around the most northern latitudes and the Sahel countries (Fig. 3d). This implies that the minimum number of 100 triplets is not in every occasion a satisfying proxy for the infinite number of common observations theoretically required, as this approximation is based on the assumption that systematic biases are absent. Even though theoretically systematic biases should have been removed by using anomalies from the long-term seasonality and the subsequent rescaling, they may still persist in areas with high soil moisture variability and in areas where the signal-to-noise ratio of soil moisture observations is low, like in densely vegetated areas and desert regions (e.g. in Southwest Australia and Southern Africa). Nevertheless, the average global errors for both combinations are very similar ($e_{RAD}^* = 0.054$ and $0.053\,\mathrm{m}^3\,\mathrm{m}^{-3}$ based on ERA-Interim and GLDAS-NOAH, respectively; $e_{SCAT}^* = 0.050$ and $0.053\,\mathrm{m}^3\,\mathrm{m}^{-3}$ based on ERA-Interim and GLDAS-NOAH, respectively), underlining the robustness of the triple collocation approach for different dataset combinations. Note, that the absolute errors shown here are considerably larger than the errors found for e_{RAD}^* and e_{SCAT}^* in Sect. 4.1. This is due to the use of the AMSR-E C-band as the reference dataset, which has a much larger dynamic range than the dynamic range of ERA-Interim.

5 Conclusions and outlook

The triple collocation technique is a promising method to estimate the error structures of global soil moisture datasets. The errors retrieved in this study appear reasonable and the observed patterns can be explained by known performance issues of each dataset. This study underscores the conclusions drawn by Scipal et al. (2008b) concerning the differences between scatterometer, radiometer and reanalysis soil moisture datasets and shows the improvements of ASCAT and AMSR-E compared to the ERS-2 and TRMM sensors, respectively. Yet unclear is to what degree the deviating trends observed for active and passive datasets can be ascribed to the observation principle (active versus passive) and how much depends on the retrieval method itself. Including soil moisture datasets based on the same sensors but obtained with different retrieval concepts could shed light on this issue and provide insight into the relative performance of the retrieval methods. Nevertheless, several studies already pointed out the limited soil moisture retrieval capability of C-band passive microwave observations over moderate to dense vegetated regions (e.g. Kirdyashev et al., 1979; Jackson et al., 1982; Parinussa et al., 2010).

In general, a decrease of random error was observed for increasing wavelength. In this prospect, extending the triple collocation analysis with SMOS observations would provide an interesting insight into the performance of soil moisture retrievals in L-band.

The results presented in this study should however be interpreted carefully. Two assumptions are central for the validity of the derived error model: (i) residual errors should be uncorrelated, and (ii) the different datasets observe the same physical phenomenon. As the measurement technique and retrieval concept of the datasets used in this study are fundamentally different, the assumption of uncorrelated errors appears justified. It may be argued that a small interdependence may exist between the reanalysis datasets and SSM/I-based soil moisture retrievals. In fact, SSM/I is one of many datasets used in the atmospheric analysis of ERA-Interim and GLDAS-NOAH, but SSM/I radiances are only assimilated over the ocean and only in the case when no rainfall is detected. The relative contribution of SSM/I to the reanalysis soil moisture estimates is therefore indirect and negligible. The second assumption, i.e. the datasets observe the same physical phenomenon, is not necessarily true. Even though all three datasets represent the same physical quantity, they observe different soil 1layers and, hence, different dynamics. Therefore, a higher order calibration might be necessary to avoid the introduction of systematic errors (Drusch et al., 2005).

The error characterisation based on different independent reanalysis datasets provides us important insight in the robustness of the triple collocation technique. The results indicate that even a minimum number of 100 joint observations in some areas are not sufficient to statistically describe
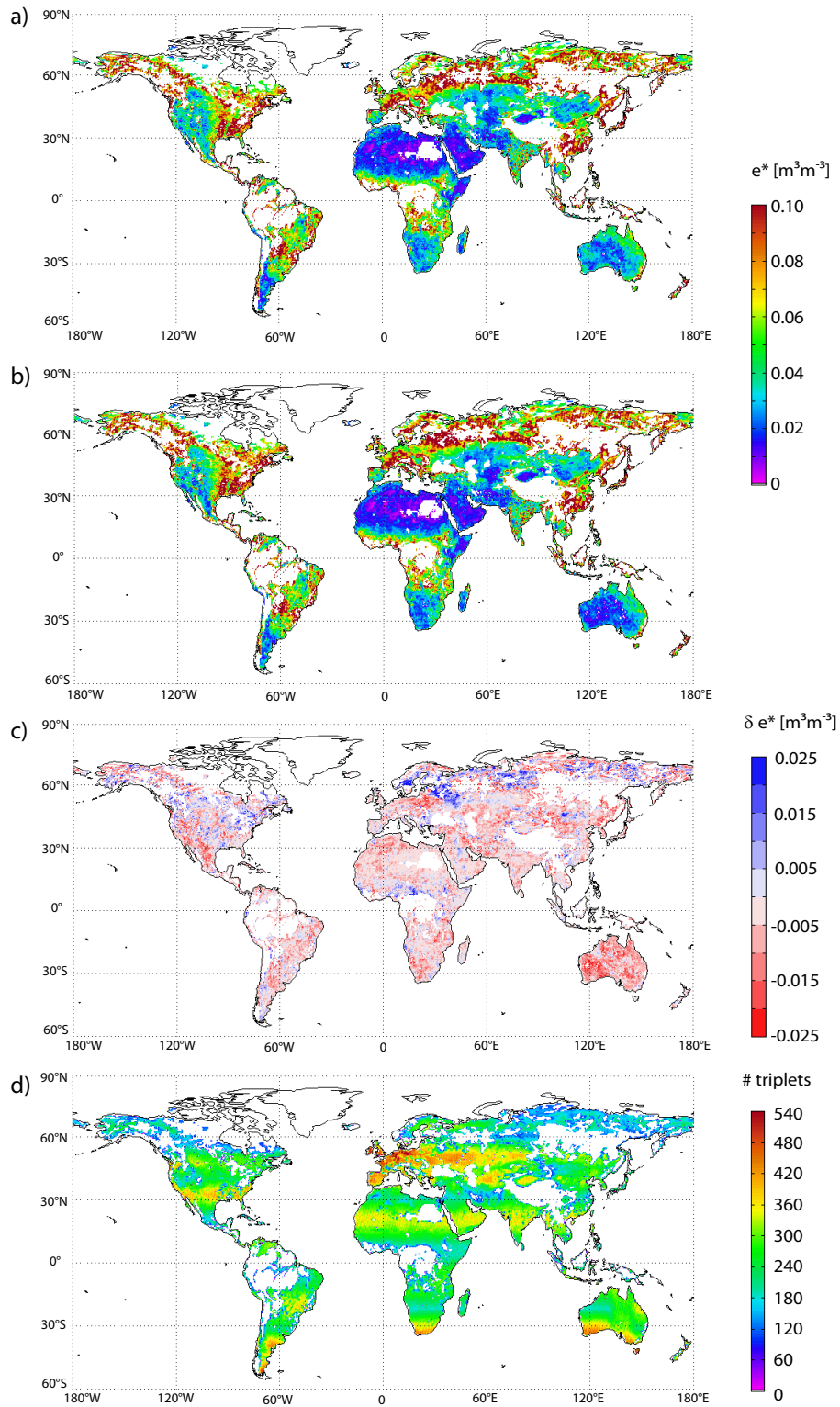
Fig. 3. (a) Spatial errors e^*_{PAS} obtained with a combination of ASCAT, AMSR-E C-band, and ERA-Interim. **(b)** Spatial errors e^*_{PAS} obtained with a combination of ASCAT, AMSR-E C-band, and GLDAS-NOAH. Both in **(a)** and **(b)** errors are expressed in the climatology of AMSR-E. **(c)** Difference between AMSR-E C-band errors obtained using GLDAS-NOAH and using ERA-Interim as third independent soil moisture dataset, respectively. **(d)** Number of triplets.

the soil moisture deviations, particularly in areas with high soil moisture dynamics and areas with a low signal-to-noise ratio of soil moisture, like in areas with dense vegetation cover or desert areas. This poses one of the major limitations of the triple collocation technique since a sufficient number of triplets can only be obtained when the overlapping time period is large enough. For the characterisation of some historic sensors (e.g. SMMR) this condition cannot be met for the combination with active soil moisture datasets and other sources need to be explored. For the most recent missions such as SMOS, this condition can only be met after a certain period of operation.

Even though the triple collocation method seems to provide plausible error estimates of soil moisture, the results presented in this paper were based on anomalies from the long-term seasonality. This means that higher order systematic deviations between the different datasets are not properly accounted for. Hence, it is expected that applying the triple collocation to absolute soil moisture values will to some extent increase the errors obtained in this study, especially for SSM/I-based soil moisture which shows a seasonality that is clearly distinctive form the other datasets. In addition, the representativeness of seasonalities based on long observation periods for present soil moisture observations should be critically evaluated in the light of climate change as some authors have reported significant trends in soil moisture over the last few decades (Liu et al., 2009; Dorigo et al., 2010).

The exact accuracy of the error estimates can only be assessed by cross-validating them with independent data sources such as in-situ soil moisture measurements. These could also provide information about the absolute biases of soil moisture estimates, something that is not accounted for by the triple collocation technique. It should also be recalled that the triple collocation technique provides only one error estimate for the entire time series and thus it can be primarily used to characterise differences between sites. Nevertheless, in combination with other uncertainty estimates, e.g. error propagation results, the triple collocation results could be used to obtain daily error budgets that are comparable between datasets. Therefore, the different accuracy assessment techniques should be seen as highly complementary.

The triple collocation results presented in this study allow us to identify systematic differences and agreements between active and passive microwave-derived soil moisture products, and between different observation bands of radiometers, e.g. with respect to varying land cover or climatological conditions. This in turn will help us in developing adequate strategies for merging active and passive observations for the generation of superior multi-mission long-term soil moisture datasets.

References

Balsamo, G., Viterbo, P., Beijaars, A., van den Hurk, B., Hirschi, M., Betts, A. K., and Scipal, K.: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system, J. Hydrometeorol., 10, 623–643, 2009.

Bartalis, Z., Scipal, K., and Wagner, W.: Azimuthal anisotropy of scatterometer measurements over land, IEEE T. Geosci. Remote., 44, 2083–2092, 2006.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff prediction through the assimilation of the ASCAT soil moisture product, Hydrol. Earth Syst. Sci., 14, 1881–1893, doi:10.5194/hess-14-1881-2010, 2010.

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

Crow, W. T., Wagner, W., and Naeimi, V.: The impact of radar incidence angle on soil-moisture-retrieval skill, IEEE Geosci. Remote. S., 7, 501–505, 2010b.

Dorigo, W., Liu, Y., Parinussa, R., Wagner, W., De Jeu, R., Hasenauer, S., and Su, Z.: Merging active and passive microwave soil moisture data to contruct long-term time series in support of climate studies, ESA Living Planet Symposium, Bergen, Norway, 2010.

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.

Drusch, M.: Initializing numerical weather prediction models with satellite-derived surface soil moisture: Data assimilation experiments with ECMWF's integrated forecast system and the TMI soil moisture data set, J. Geophys. Res.-Atmos., 112, D03102 doi:10.1029/2006JD007478, 2007.

Gruhier, C., de Rosnay, P., Hasenauer, S., Holmes, T., de Jeu, R., Kerr, Y., Mougin, E., Njoku, E., Timouk, F., Wagner, W., and Zribi, M.: Soil moisture active and passive microwave products: intercomparison and evaluation over a Sahelian site, Hydrol. Earth Syst. Sci., 14, 141–156, doi:10.5194/hess-14-141-2010, 2010.

Jackson, T. J., Schmugge, T. J., and Wang, J. R.: Passive microwave sensing of soil moisture under vegetation canopies, Water Resour. Res., 18, 1137–1142, 1982.

Jackson, T. J.: III. Measuring surface soil moisture using passive microwave remote sensing, Hydrol. Process., 7, 139–152, 1993.

Jackson, T. J., Cosh, M. H., Bindlish, R., Starks, P. J., Bosch, D. D., Seyfried, M., Goodrich, D. C., Moran, M. S., and Du, J.: Validation of Advanced Microwave Scanning Radiometer Soil Moisture Products, IEEE T. Geosci. Remote., 1–17, 2010.

Jeu, R. A. M., Wagner, W., Holmes, T. R. H., Dolman, A. J., Giesen, N. C., and Friesen, J.: Global soil moisture patterns observed

by space borne microwave radiometers and scatterometers, Surv. Geophys., 29, 399–420, 2008.

Kirdyashev, K. P., Chukhlantsev, A. A., and Shutko, A. M.: Microwave radiation of the Earth's surface in the presence of a vegetation cover, Radio Eng. Electron. P., 24, 37–44, 1979.

Koike, T., Nakamura, Y., Kaihotsu, I., Davva, N., Matsuura, N., Tamagawa, K., and Fujii, H.: Development of an Advanced Microwave Scanning Radiometer (AMSR-E) Algorithm of Soil Moisture and Vegetation Water Content, Annual Journal of Hydraulic Engineering, JSCE, 48, 217–222, 2004.

Li, L., Gaiser, P. W., Gao, B. C., Bevilacqua, R. M., Jackson, T. J., Njoku, E. G., Rüdiger, C., Calvet, J. C., and Bindlish, R.: WindSat global soil moisture retrieval and validation, IEEE T. Geosci. Remote., 48, 2224–2241, 2010.

Liu, Y. Y., Van Dijk, A. I. J. M., De Jeu, R. A. M., and Holmes, T. R. H.: An analysis of spatiotemporal variations of soil and vegetation moisture from a 29-year satellite-derived data set over mainland Australia, Water Resour. Res., 45, W07405, doi:10.1029/2008wr007187, 2009.

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., de Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., McCabe, M. F., and Evans, J. P.: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, Hydrol. Earth Syst. Sci. Discuss., 7, 6699–6724, doi:10.5194/hessd-7-6699-2010, 2010.

Mahfouf, J. F.: Assimilation of satellite-derived soil moisture from ASCAT in a limited-area NWP model, Q. J. Roy. Meteor. Soc., 136, 784–798, 2010.

Martínez-Fernández, J. and Ceballos, A.: Mean soil moisture estimation using temporal stability analysis, J. Hydrol., 312, 28–38, 2005.

Miralles, D. G., Crow, W. T., and Cosh, M. H.: Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations, J. Hydrometeorol., doi:10.1175/2010JHM1285.1, in print, 2010.

Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., and Wagner, W.: An Improved Soil Moisture Retrieval Algorithm for ERS and METOP Scatterometer Observations, IEEE T. Geosci. Remote., 47, 1999–2013, doi:10.1109/Tgrs.2009.2011617, 2009.

Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K., and Nghiem, S. V.: Soil moisture retrieval from AMSR-E, IEEE Transactions on Geoscience and Remote Sensing, 41, 215–229, doi:10.1109/tgrs.2002.808243, 2003.

Owe, M., De Jeu, R., and Walker, J.: A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index, IEEE T. Geosci. Remote., 39, 1643–1654, 2001.

Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, J Geophys. Res.-Earth, 113, F01002, doi:10.1029/2007jf000769, 2008.

Parinussa, R., Meesters, A. G. C. A., Liu, Y. Y., Dorigo, W., Wagner, W., and De Jeu, R. A. M.: An analytical solution to estimate the error structure of a global soil moisture data set, IEEE Geosci. Remote. S., in review, 2010.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, B. Am. Meteorol. Soc., 85, 381–394, 2004.

Scipal, K., Wagner, W., Trommler, M., and Naumann, K.: The global soil moisture archive 1992–2000 from ERS scatterometer data: first results, Igarss 2002: IEEE International Geoscience and Remote Sensing Symposium and 24th Canadian Symposium on Remote Sensing, Vols I-VI, Proceedings, 1399–1401, 2002.

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, 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.

Simmons, A., Uppala, S., Dee, D., and Kobayashi, S.: ERA-Interim: New ECMWF reanalysis products from 1989 onwards, ECMWF Newsletter, 110, 25–35, 2007.

Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, J. Geophys. Res., 103, 7755–7766, doi:10.1029/97jc03180, 1998.

Ulaby, F. T., Moore, M. K., and Fung, A. K.: Microwave Remote Sensing, Active and Passive: Radar Remote Sensing and Surface Scattering and Emission Theory, Vol. 2, Artech House, Norwood, MA, 1982.

Viterbo, P. and Beljaars, A. C. M.: An Improved Land-Surface Parameterization Scheme in the ECMWF Model and Its Validation, J. Climate, 8, 2716–2748, 1995.

Wagner, W., Lemoine, G., and Rott, H.: A method for estimating soil moisture from ERS scatterometer and soil data, Remote Sens. Environ., 70, 191–207, 1999.

Wagner, W., Naeimi, V., Scipal, K., de Jeu, R., and Martinez-Fernandez, J.: Soil moisture from operational meteorological satellites, Hydrogeol. J., 15, 121–131, doi:10.1007/s10040-006-0104-6, 2007.

Wigneron, J. P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M. J., Richaume, P., Ferrazzoli, P., de Rosnay, P., Gurney, R., Calvet, J. C., Grant, J. P., Guglielmetti, M., Hornbuckle, B., Mätzler, C., Pellarin, T., and Schwank, M.: L-band Microwave Emission of the Biosphere (L-MEB) Model: Description and calibration against experimental data sets over crop fields, Remote Sens. Environ., 107, 639–655, 2007.

Zribi, M., Le Hégarat-Mascle, S., Ottlé, C., Kammoun, B., and Guerin, C.: Surface soil moisture estimation from the synergistic use of the (multi-incidence and multi-resolution) active microwave ERS Wind Scatterometer and SAR data, Remote Sens. Environ., 86, 30–41, 2003.