Hydrol. Earth Syst. Sci. Discuss., 7, 5621–5645, 2010
www.hydrol-earth-syst-sci-discuss.net/7/5621/2010/
doi:10.5194/hessd-7-5621-2010

Hydrology and Earth System Sciences Discussions

Error characterisation of global active and passive microwave soil moisture data sets

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

Received: 23 July 2010 – Accepted: 27 July 2010 – Published: 13 August 2010

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

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Understanding the error structures of remotely sensed soil moisture products 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 data sets is often hampered by the limited availability over space and time of reliable in-situ measurements. 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 technique is a powerful tool to estimate the root mean square error while simultaneously solving for systematic differences in the climatologies of a set of three independent data sources. In addition to the scatterometer and radiometer data sets, we used the ERA-Interim and GLDAS-NOAH reanalysis soil moisture data sets 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 data set (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 data sets are uncorrelated) the errors estimated for the remote sensing products are hardly influenced by the choice of the third independent data set. 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 data sets, e.g. for improved flood forecast modelling or the generation of superior multi-mission long-term soil moisture data sets.

1 Introduction

In recent years, an increasing number of global soil moisture data sets have become available from passive and active coarse resolution satellite microwave sensors. Altogether, these data sets span a period of more than 30 years (Table 1). Knowing the quality of the different data sets 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 data sets, e.g. in relation to climate change (Liu et al., 2009). But also other applications, like the assimilation of remote sensed soil moisture conditions in flood forecasting (Brocca et al., 2009) or numerical weather prediction models (Mahfouf, 2010; Drusch, 2007; Scipal et al., 2008a) 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. Jeu et al., 2008; Wagner et al., 2007; 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 data sets. 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 simulation to estimate the uncertainty in soil moisture retrievals obtained from scatterometers

HESSD
7, 5621–5645, 2010

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 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 low 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 an independent reference data set (ERA-Interim versus 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).

HESSD
7, 5621–5645, 2010

2 Data

2.1 Scatterometer data

The Advanced Scatterometer (ASCAT) on MetOp-A operates in C-band (5.255 GHz) at VV polarization 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 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). The TU Wien model exploits the unique sensor design and the advantages of a change detection method. To correct for the effects of plant growth and decay, the model 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, which results in a relative measure of surface (<2 cm) soil moisture ranging between 0 and 1 (or 0 and 100%). The soil moisture product is available via EUMETSAT data distribution system in near real time.

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 C-band operated at 6.9 GHz and the X-band operated at 10.7 GHz. Spatial resolutions are 73x43 km and 51×30 km for the C- and X-band, respectively. In this study we only use AMSR-E night-time observations, as it was shown that these are better suited for retrieving soil moisture than day-time observations (Jeu et al., 2008).

HESSD
7, 5621–5645, 2010

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, while the last one was launched in October 2003. The SSM/I sensor operates at four frequencies but only the Ku-band (19.3 GHz) will be considered in this study. At this frequency, the footprint size is 69×43 km. Only descending mode observations from the F13 satellite (equator crossing in the early morning) 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 surface soil moisture applying the Land Parameter Retrieval Model (LPRM; Owe et al., 2008). LPRM is based on the solution of a microwave radiative transfer model and solves simultaneously for surface soil moisture, vegetation optical depth and land surface temperature without a-priori information of land surface characteristics.

2.3 Reanalysis data

2.3.1 ERA-Interim

The ERA Interim reanalysis data set contains consistent atmosphere and surface analyses for the period from 1989 to present based on the ECMWF Integrated Forecast System (IFS) model. 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^3m^{-3}, field capacity with 0.323 m^3m^{-3}, and the wilting point with 0.171 m^3m^{-3}.

HESSD
7, 5621–5645, 2010

Error characterisation of global soil moisture data sets

W. A. Dorigo et al.

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-hour time interval for a regular global grid with a spatial resolution of 0.25°. 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 kgm^{-3} which can easily be converted into volumetric soil moisture. 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. Data generated by the Noah model are publicly available from ftp://agdisc.gsfc.nasa.gov/data/s4pa/.

2.4 Spatial and temporal collocation of data sets

For the time period where observations were available for all three sources (i.e. 2007/01/01–2008/12/31) the data were binned to daily files 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

HESSD

7, 5621–5645, 2010

Error characterisation of global soil moisture data sets

W. A. Dorigo et al.

to hypothetical true soil moisture Θ in a linear fashion (Scipal et al., 2008b; Stoffelen, 1998):

$$\begin{aligned}\Theta &= \alpha_{SCAT} + \beta_{SCAT}\Theta_{SCAT} + r_{SCAT}\\ \Theta &= \alpha_{RAD} + \beta_{RAD}\Theta_{RAD} + r_{RAD}\\ \Theta &= \alpha_{MOD} + \beta_{MOD}\Theta_{MOD} + r_{MOD}\end{aligned} \tag{1}$$

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 $r^*_X = r_X/\beta_X$ (X equals SCAT, RAD, and MOD, respectively), and then eliminate the unknown truth in order to obtain Eq. (2):

$$\begin{aligned}\left(\Theta^*_{SCAT} - \Theta^*_{RAD}\right) &= r^*_{SCAT} - r^*_{RAD}\\ \left(\Theta^*_{SCAT} - \Theta^*_{MOD}\right) &= r^*_{SCAT} - r^*_{MOD}\\ \left(\Theta^*_{RAD} - \Theta^*_{MOD}\right) &= r^*_{RAD} - r^*_{MOD}\end{aligned} \tag{2}$$

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 $e^{*2}_{SCAT} = \left\langle r^{*2}_{SCAT} \right\rangle$ 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:

$$\begin{aligned}e^{*2}_{SCAT} &= \left\langle \left(\Theta^*_{SCAT} - \Theta^*_{RAD}\right)\left(\Theta^*_{SCAT} - \Theta^*_{MOD}\right) \right\rangle\\ e^{*2}_{RAD} &= \left\langle \left(\Theta^*_{SCAT} - \Theta^*_{RAD}\right)\left(\Theta^*_{RAD} - \Theta^*_{MOD}\right) \right\rangle\\ e^{*2}_{MOD} &= \left\langle \left(\Theta^*_{SCAT} - \Theta^*_{MOD}\right)\left(\Theta^*_{RAD} - \Theta^*_{MOD}\right) \right\rangle\end{aligned} \tag{3}$$

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 soil moisture data, we base our analysis on soil moisture anomalies from the long-term mean. This has the advantage that we are still able to determine the random errors of the SSM/I data set which, due to its Ku-band, has a higher sensitivity to atmospheric water vapour and vegetation and thus on many locations shows a seasonality that differs from the other microwave sensors. 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 measurements 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., 2010). Due to the short observation period of ASCAT, both ERS and ASCAT soil moisture estimates were used to obtain a reliable seasonality for ASCAT. This can be done as ERS relies on the same observation principles and retrieval concepts as ASCAT.

Finally, to estimate e^{*2}_{SCAT}, e^{*2}_{RAD} and e^{*2}_{MOD} , one has to solve for the calibration expressed in Eq. (1). Since the real truth is always unknown, one has to choose one of the data sets as a reference. This means that e^{*2}_{SCAT}, e^{*2}_{RAD}, and e^{*2}_{MOD} will be expressed in the observation space of the selected reference data set. The choice of the reference data set 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 simple linear least-squares approximation assuming errors in both variables to solve for

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

7, 5621–5645, 2010

Error characterisation of global soil moisture data sets

W. A. Dorigo et al.

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

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

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 data set $\hat{\Theta}_X(t)$ 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 data sets) are required to obtain unbiased estimates of e^{*2}_X. 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 climatology of the ERA-Interim re-analysis data set. The results of the error estimation suggest that all three data sets are characterised by a relatively low error. The mean global error is 0.017 $m^3 m^{-3}$ for the ASCAT (e^*_{SCAT}), 0.019 $m^3 m^{-3}$ for the AMSR-E C-band observations (e^*_{RAD}) and 0.018 $m^3 m^{-3}$ for ERA-Interim (e^*_{MOD}). The low error is partly due to the low dynamic

HESSD
7, 5621–5645, 2010

range of the ERS-Interim soil moisture which has been 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 climatology of ERA-Interim, it is important to realise that this data set does not profit relative to the other data sets from the selection as the reference. This means that the relative magnitude of the errors remains the same if one of the other data sets 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 ($e^*_{SCAT} = 0.028\,m^3m^{-3}$), the TMI radiometer ($e^*_{RAD} = 0.046\,m^3m^{-3}$) and ERA-Interim soil moisture ($e^*_{MOD} = 0.020\,m^3m^{-3}$). The lower average error obtained in our study for the scatterometer can be ascribed to the improved design of ASCAT with respect to the ERS series. Most remarkable is the increase in accuracy obtained for the AMSR-E radiometer compared to the TMI results by Scipal et. al. 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 lower frequency 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 consistent with the average error obtained by Scipal et al. This confirms the robustness of the triple collocation method and indicates a high spatial consistency of ERA-Interim soil moisture, as the study of Scipal et al. did not cover northern latitudes due to the limited coverage of the TMI sensor. The small difference observed for ERA-Interim is very likely a consequence of the latter. Finally, the use of soil moisture anomalies instead of original observations and the different scaling technique have a minor influence on the differences between the two studies.

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

HESSD
7, 5621–5645, 2010

than in humid regions.

Despite the similar average errors of the three data sets, several characteristic differences in the spatial distribution of the errors can be observed between the data sets (Fig. 1a–c). In very dry areas (e.g. those of central Australia) errors of soil moisture derived from AMSR-E C–band are remarkably lower than soil moisture estimates derived from ASCAT and, to a smaller degree, than the modelled soil moisture of 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 the AMSR-E observations from above the canopy in two ways. First, vegetation will absorb or scatter the radiation emanating from the soil. Secondly, the vegetation will also emit its own radiation. These two effects tend to counteract each other. The observable soil emission will decrease with increased vegetation, while emission from the vegetation 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).

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 data sets (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

HESSD
7, 5621–5645, 2010

Error characterisation of global soil moisture data sets

W. A. Dorigo et al.

be used to fill the map in Fig. 1d in order to obtain a complete global coverage.

4.2 Influence of radiometer observation frequency

Figure 2 illustrates the influence of increasing observation frequency on the error structures obtained for radiometer observations. The triple collocation was based on a combination of ERA-Interim, ASCAT and the respective radiometer data set. On average, there is a clear average increase in errors with increasing frequency, 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 do not contain a detectable soil moisture signal anymore. This is also the main reason of the reduced 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 frequency-related uncertainties of LPRM products obtained by error propagation (Parinussa et al., 2010).

Despite the fact that many areas cannot be observed due to the increased sensitivity of the signal to vegetation, observations 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. This is also illustrated by comparing the average errors of the three passive data sets over the areas where all data sets 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 $m^{-3}m^{-3}$, respectively, indicating 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 observations. However, the conclusions should be taken with caution as the trends in error structures obtained for the AMSR-E

HESSD
7, 5621–5645, 2010

and SSM/I sensors do not rely only on differences in frequency but also on differences in instrument design, radiometric accuracy, overpass times and so on.

4.3 Influence of reference data set

So far, the results presented were based on ASCAT, one of the radiometer data sets and the ERA-Interim reanalysis data set. Theoretically, the choice of the third data set should not influence the results obtained for the other data sets, given the errors of all three data set 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, data set. To be able to directly compare the error structures based on the different re-analysis data sets, errors must be expressed in the same climatology. For this reason, the AMSR-E C-band data set 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 soil moisture observations are prone to large uncertainties, like in densely vegetated areas. Nevertheless, the average global errors for both combinations are very similar (e^*_{RAD}=0.054 and 0.053 $\mathrm{m^3 m^{-3}}$ based on ERA-Interim and GLDAS-NOAH, respectively; e^*_{SCAT}=0.050

HESSD
7, 5621–5645, 2010

and 0.053 m^3m^{-3} based on ERA-Interim and GLDAS-NOAH, respectively), underlining the robustness of the triple collocation approach for different data set combinations.

5 Conclusions and outlook

The triple collocation technique is a promising method to estimate the error structures of global soil moisture data sets. The errors retrieved in this study appear reasonable and the observed patterns can be explained by known performance issues of each data set. This study underscores the conclusions drawn by Scipal et al. (2008b) concerning the differences between scatterometer, radiometer and reanalysis soil moisture data sets 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 data sets can be ascribed to the observation principle (active versus passive) and how much depends on the retrieval method itself. Including soil moisture data sets 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 decreasing frequency. 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 data sets observe the same physical phenomenon. As the measurement technique and retrieval concept of the data sets used in this study are fundamentally different, the assumption of uncorrelated errors appears justified. The second assumption is however not necessarily true. Even though all three

HESSD
7, 5621–5645, 2010

data sets represent the same physical quantity, they observe different soil layers 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 the soil moisture deviations, particularly in areas with high soil moisture dynamics and vegetation cover. 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 data sets 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, their exact accuracy 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 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 data sets. 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 frequency bands of radiometers e.g. with respect to varying land cover or climatological zones. 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 data sets.

Acknowledgements. The work was undertaken as part of ESAs STSE funded Integrated Project WAter Cycle Multimission Observation Strategy (WACMOS; http://www.wacmos.org/; ESRIN/Contract No. 22086/08/I-EC) coordinated by the International Institute of Geo-Information Science and Earth Observation (ITC). We gratefully thank Diego Miralles of the VU University Amsterdam for the fruitful discussions on his implementation of the triple collocation technique.

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, Geoscience and Remote Sensing, IEEE Transactions on, 44, 2083–2092, 2006.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Potential of ASCAT soil moisture product to improve runoff prediction, Earth Observation and Water Cycle Science Symposium, EGU Topical Conference Series on the hydrological cycle, ESA-ESRIN, Frascati, Italy, 2009,

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

Drusch, M., Wood, E. F., and Gao, H.: Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture, Geophy. 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(3), 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 prod-

ucts: 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, Geoscience and Remote Sensing, IEEE Transactions on, 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 engineering & electronic physics, 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 Transactions on Geoscience and Remote Sensing, 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., submitted, 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.

HESSD
7, 5621–5645, 2010

Error characterisation of global soil moisture data sets

W. A. Dorigo et al.

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., in press, 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 Transactions on Geoscience and Remote Sensing, 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 Transactions on Geoscience and Remote Sensing, 39, 1643–1654, 2001.
Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, Journal of Geophysical Research-Earth Surface, 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 soilution to estimate the error structure of a global soiol moisture data set, IEEE Transactions on Geoscience and Remote Sensing, In review, 2010.
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 Resou., 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.
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.
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 scat-

HESSD
7, 5621–5645, 2010

terometer 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.

HESSD

7, 5621–5645, 2010

Error characterisation of global soil moisture data sets

W. A. Dorigo et al.

Table 1. Operational products in the field of global monitoring of soil moisture using active and passive satellite microwave instruments.

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

HESSD
7, 5621–5645, 2010

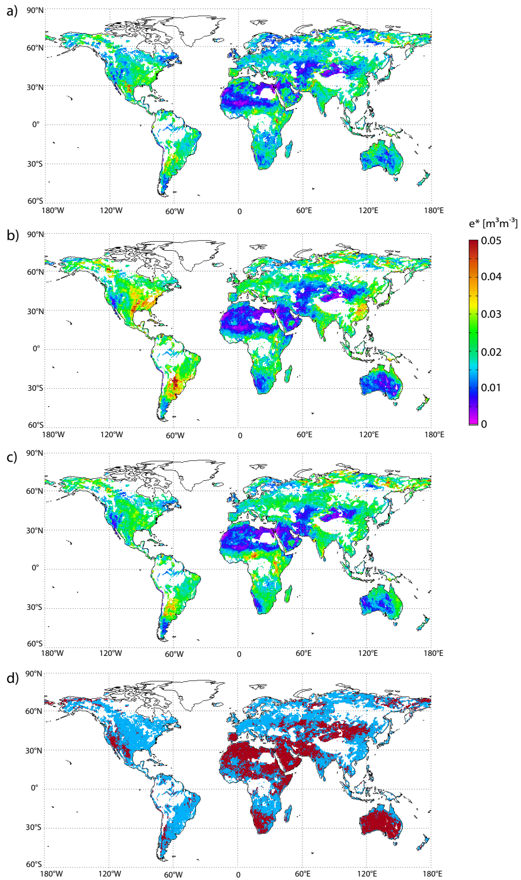

Fig. 1. Spatial errors of **(a)** ASCAT, **(b)** AMSR-E C-band, and **(c)** ERA-Interim surface soil moisture estimates. Errors are expressed in the climatology of ERA-Interim. **(d)** shows the areas in which either ASCAT (blue) or AMSR-E (red) shows the smallest error value. White areas indicate areas for which less than 100 common observations are available.

HESSD
7, 5621–5645, 2010

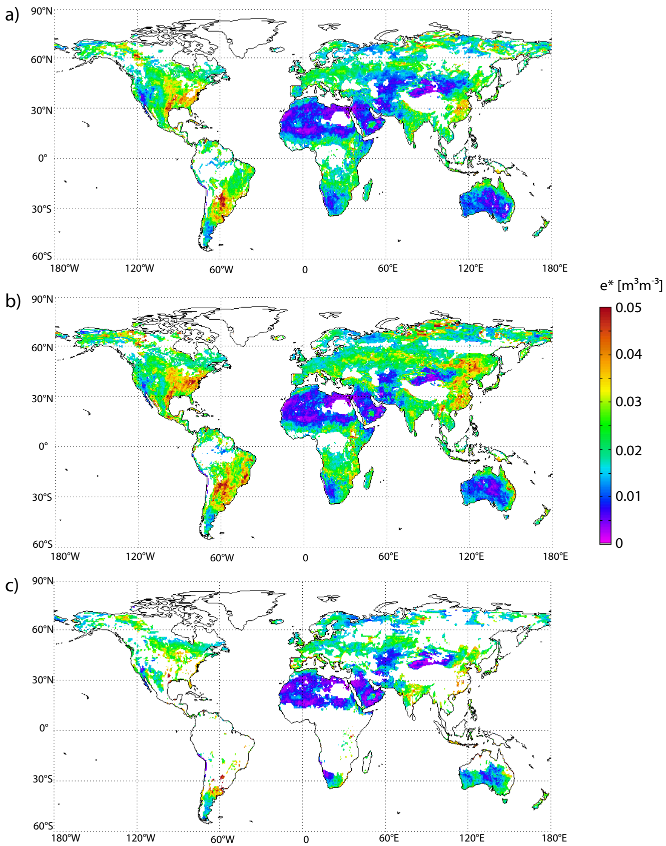

Fig. 2. Spatial errors e^*_{PAS} of **(a)** AMSR-E C-band, **(b)** AMSR-E X-band, and **(c)** SSM/I Ku-band observations obtained with triple collocation based on ERA-Interim, ASCAT, and the respective radiometer data set. Errors are expressed in the climatology of ERA-Interim.

HESSD
7, 5621–5645, 2010

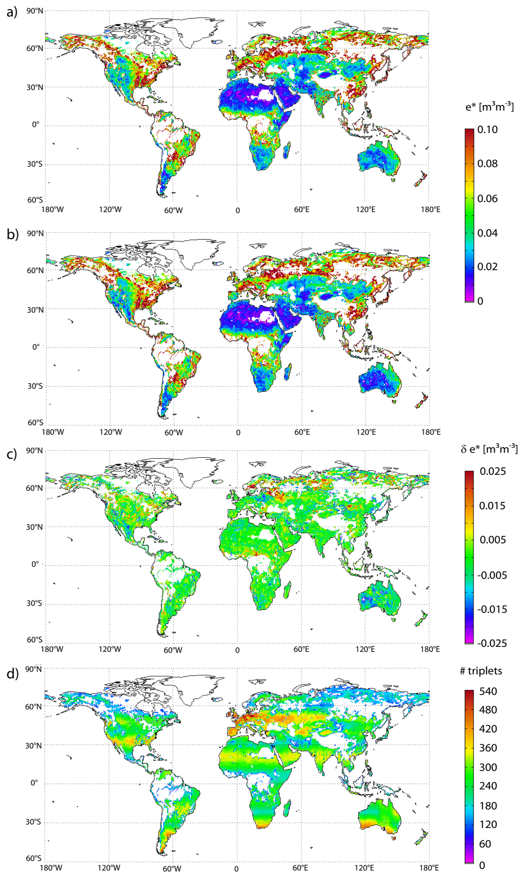

Fig. 3. Spatial errors e^*_{PAS} obtained with a combination of ASCAT, AMSR-E C-band and **(a)** ERA-Interim or **(b)** GLDAS-NOAH. Errors are expressed in the climatology of AMSR-E. **(c)** Difference between AMSR-E C-band errors obtained using GLDAS-NOAH and using ERA-Interim as third independent soil moisture data set. **(d)** Number of triplets.

HESSD

7, 5621–5645, 2010