CMUG Deliverable
Number: D3.1_v1B
Due date: October 2012
Submission date: 22. November 2012
Version: 0.5

Climate Modelling User Group

Deliverable 3.1_v1B

Technical note on CMUG ECV Soil Moisture Assessment Report

Centres providing input: MPI-M

Version nr.	Date	Status
0.1	12 Oct 2012	Initial document
0.2	12 Nov 2012	Revised version with full content
0.3	18 Nov 2012	Final report draft
0.4	21 Nov 2012	Final report
0.5	28. Jan 2013	Added section on ECV_SM preprocessing steps on request of ESA

Contact: alexander.loew@zmaw.de

CMUG ECV Soil moisture Assessment Report

Table of Contents

1 Introduction

Soil moisture is an important Essential Climate Variable (ECV) which has impact on regional to global terrestrial water, energy and carbon fluxes. Soil moisture controls the partitioning of the available energy into latent and sensible heat flux and conditions the amount of surface runoff. By controlling evapotranspiration, it is linking the energy, water and carbon fluxes (Koster et al., 2004; Dirmeyer et al., 2006; Seneviratne and Stöckli, 2008). Due to its long term memory, soil moisture provides important information for seasonal climate forecasts (Fischer et al., 2007).

Up to now, soil moisture observations are based on either in situ measurements (Dorigo et al., 2011) or satellite observations. Recently, the first multidecadal satellite based soil moisture record has become available (Liu et al., 2011).

The purpose of the present study is to evaluate potential applications of this novel soil moisture data set for climate modelling applications with the Max-Planck-Institute for Meteorology Earth System Model (MPI-ESM, Giorgetta et al., 2012). The overarching objectives of the analysis in the present study are:

- Evaluate the potential of using satellite soil moisture observations for climate model evaluation
- Evaluate the potential of using satellite soil moisture data for climate model development
- Analyze the use of long term soil moisture data for climate studies

After a brief introduction to the datasets and models used, the analysis methods applied are briefly described. Results of the analysis are presented and discussed afterwards and an outlook on further potential use of satellite soil moisture data is given.

2 Data and models

2.1 MPI-ESM land surface model (JSBACH)

The model used in the present study is the land surface scheme of the MPI-ESM, JSBACH (Reick et al., 2012). The model is implicitly coupled to the atmospheric component of MPI-ESM (ECHAM6) and simulates all relevant land surface water, energy and carbon fluxes in an interactive manner (Figure 1). A new soil hydrology scheme with multiple layers for the zone until the bedrock allows for the simulation of soil moisture dynamics in varying depths. The soil layers have a thickness of dz=[0.065,0.254,0.913,2.902,5.9] [m]. As satellite observations provide soil moisture information for the upper few centimeters of the soil only, the new scheme allows for a comparison with the satellite data.

The present analysis uses version 2.03 of JSBACH which is comparable to the model version which was used for the Coupled Model Intercomparison Project 5 (CMIP5) (Taylor et al., 2012). The only major difference between the CMIP5 model and the model version used in the present study is the inclusion of the new 5-layer soil hydrology scheme.

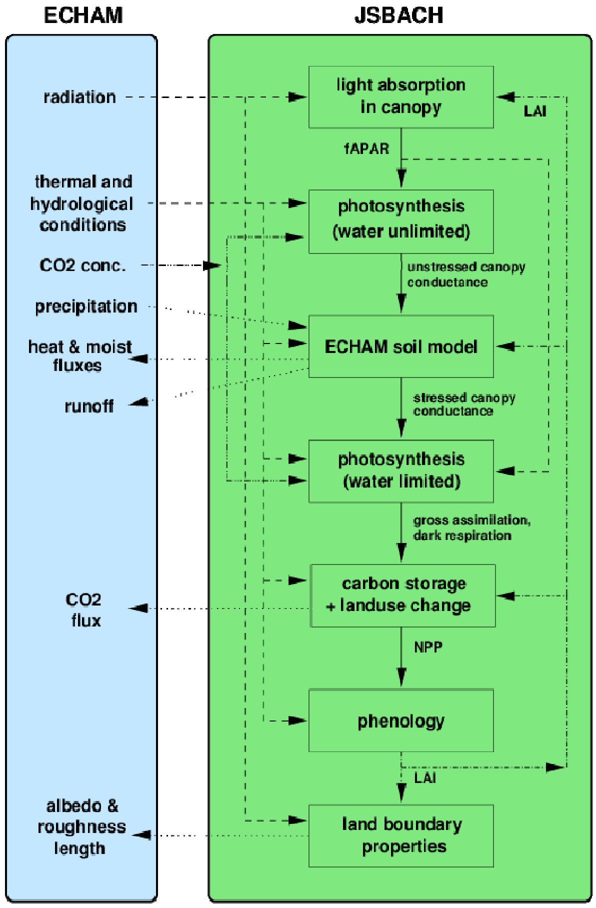

Figure 1: Components of JSBACH

2.2 *WATCH forcing data*

JSBACH simulations can be performed either by forcing the land surface scheme with any kind of meteorological forcing data (e.g. station measurements, reanalysis data) or by coupling JSBACH directly to a Global Circulation Model (GCM), like ECHAM.

For the present study, we use model simulations that are forced with observed meteorological data. This allows for a direct comparison of the satellite soil moisture observations with the model simulations and enables to quantify the accuracy of the soil moisture simulations with help of the satellite observations. The forcing data set, created in the EU WATCH project is used for that purpose.

The Watch Forcing Data (WFD) were created in the framework of the WATCH project (www.eu-watch.org) and are documented in Weedon et al. (2010, 2011). Originally, they comprise the period from 1958 to 2001 and include a wide range of meteorological forcing variables, e.g. surface temperature, precipitation, radiation fluxes and wind speed. The data set is available at 0.5° spatial resolution and sub-daily time steps.

The WFD is based on ERA40 reanalysis data (Uppala et al., 2005). An extensive post-processing was conducted in which the data were bi-linearly interpolated from the original Gaussian 1.125° grid to a rectangular 0.5° resolution grid. An elevation correction took place for most variables. Furthermore, rainfall and snowfall were subject to a wet day, bias and undercatch correction ensuring that the monthly statistics are similar to the GPCC observations while the daily variability resamples ERA40.

Recently, an extension of the WFD became available that applied the WFD methodology to the ERA-Interim data for the time period of 1979 to 2009. Similar to the WFD, the WFDEI have a spatial resolution of 0.5°. In order to serve as forcing for JSBACH they were remapped conservatively to a T63 grid (about 1.8° to 1.8°).

2.3 Soil moisture data

Different soil moisture datasets are analyzed in the present study (Table 1). The usage of different datasets allows for analysis of the consistency and differences between the different datasets. Two remote sensing based soil moisture datasets are used, including the recently released multidecadal soil moisture data record developed by partners from the ESA CCI soil moisture team (ECV_SM v0.1). The different datasets will be briefly introduced in the following.

Table 1: List of used datasets in the present study

Dataset	*Version*	*Timeperiod*	*References*
VUA AMSR-E soil moisture	V0.5	07/2002 – 10/2011	Owe et al., 2008
ESA ECV Soil Moisture (ECV_SM)	V0.1	11/1978 – 12/2010	Liu et al., 2011; Liu et al., 2012
ERA-Interim	-	1979 - 2010	Dee et al., 2011
GPCP	V2.2	1979 - 2010	Adler et al., 2003

2.3.1 AMSR-E soil moisture (VUA)

The AMSR-E soil moisture dataset from VUA is based on the Land Surface Parameter Retrieval Model (LPRM) (Owe et al., 2008) which allows for a joint retrieval of soil moisture and vegetation optical depth using an analytical approach (Meesters et al., 2005) and prior land surface temperature estimates. It has been widely used by the land surface research community.

The recent version 0.5 of the dataset comprises different retrieval results from X- and C-band. It provides individual soil moisture retrievals from LPRM for ascending and descending orbits including soil moisture errors as derived from error propagation (Parinussa et al., 2011). Further the LPRM retrieval results are normalized to the climatology of the Noah land surface model using CDF matching. The CDF matching reference is the same as used for ECV_SM v0.1. Within this report, the different data products will be analyzed.

2.3.2 ESA ECV Soil Moisture (ECV_SM) v0.1

The ESA ECV_SM v0.1 data product has been released in summer 2012 as a result of the ESA WACMOS project. This first ever multidecadal satellite-based soil moisture dataset covers the period from 1978 to 2010. It combines radiometer based products from SMMR (11/1978 – 08/1987), SSM/I (07/1987 – 2007), TMI (1998-2008) and AMSR-E (07/2002 – 2010) with scatterometer-based products from ERS-1/2 (07/1991 – 05/2006) and ASCAT (2007-2010).

In time periods, where data from different sensors is available, a binary merging technique was developed that only uses data from a single sensor in a specific time period (Liu et al., 2011). The dataset generation as well as first analyses of the dataset are reported in detail in Liu et al. (2011,2012) and Dorigo et al. (2012).

Data pre-processing

As the data pre-processing and harmonization has a major impact on some of the conclusions drawn in this report, a summary of the major preprocessing steps developed for the generation of ECV_SM is given, which is basically a summary of the more detailed description as given by Liu et al. (2012).

The ECV_SM dataset is a compilation of soil moisture observations from various instruments on different satellite platforms. These were merged into a harmonized product, whereas the observations of a particular sensor are used for specific periods in time and dedicated regions over the globe. Figure 2 shows the general temporal and spatial coverage of the datasets used for the final ECV_SM generation.

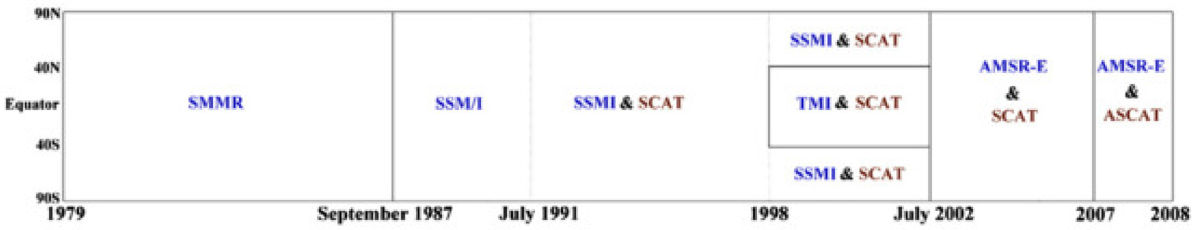

Figure 2: Spatial and temporal data coverage of soil moisture from different sensors in the final long term ECV_SM dataset (Fig. 13, Liu et al., 2012)

The passive microwave instruments are used in regions with sparse vegetation, identified by a typical vegetation optical thickness (VOD) ≤ 0.38 , while the active microwave observations were used for more denser vegetated areas (VOD> 0.38). Regions with similar skill in representing soil moisture dynamics were identified through a correlation analysis between active and passive microwave observations. In case of a correlation $r \geq 0.65$, observations from both, active and passive microwave sensors were used for the final product (for details see Liu et al., 2012). The resulting spatial and temporal field of this “binary” blending is shown in Figure 3. It shows, that e.g. over Europe, the coverage is dominated by observations from the active microwave sensors.

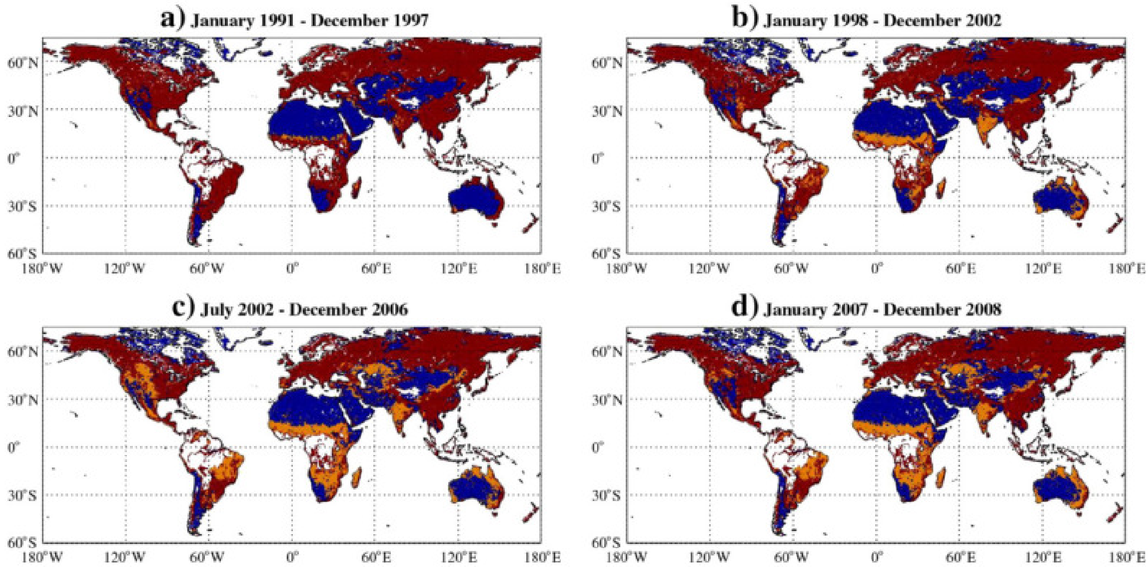

Figure 3: Spatial and temporal coverage of passive (blue) and active (red) microwave products in the final blended ECV_SM product. Transitional zones with information from both instrument types are indicated in orange colors. Areas with high vegetation density or permanents snow/ice cover are masked. (Fig.14 in Liu et al., 2012)

To generate the blended ECV_SM data product, the ECV_SM team first generated two independent time series from active and passive microwave observations, which were then blended like discussed before (Figure 4). First (step 1), the different passive microwave observations were harmonized using AMSR-E observations as a common reference and the active microwave observations were harmonized likewise (step 2).

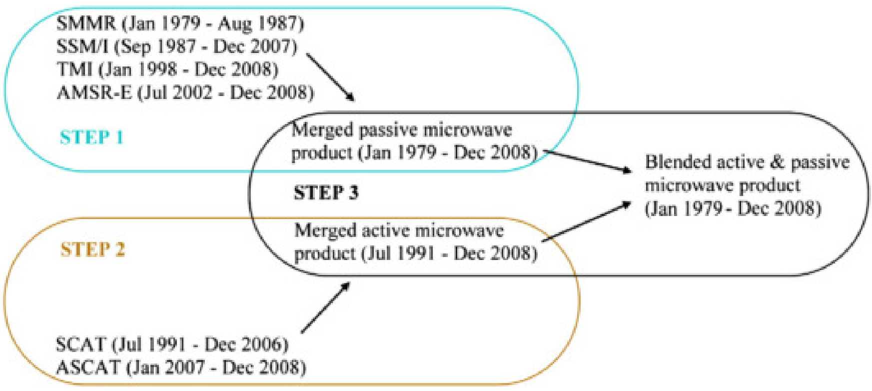

Figure 4: General workflow to generate ECV_SM data product (Fig.2, Liu et al., 2012)

The passive microwave retrieval scheme (Owe et al., 2008) provides volumetric soil moisture in units of [m³/m³], while the scatterometer based soil moisture retrievals are based on a change detection approach which provides soil moisture in terms of relative saturation degree. For the generation of a combined product, it is therefore necessary to a) translate between relative and absolute soil moisture values as well as to b) compensate for potential biases in the individual products.

The approach which was applied for ECV_SM is the so called CDF matching procedure, which is described in detail in Liu et al. (2011). Given two time series of soil moisture for a given location (Figure 5), one can calculate the cumulated distribution function (CDF) from each of the timeseries and match them using e.g. lower order polynomials. Figure 6 shows examples of CDFs calculated from the timeseries of Figure 5. To harmonize the timeseries of passive and active microwave observations over a longer time period, a common reference is needed. For ECV_SM, the common reference chosen was the soil moisture data product from the Noah land surface model. The remote sensing based soil moisture observations are therefore matched to the Noah soil moisture dynamics on a grid point basis using the CDF matching procedure. As a consequence, the soil moisture statistics (percentiles) of the blended ECV_SM soil moisture product resembles the soil moisture dynamics of the Noah land surface model. **It is however emphasized, that this rescaling just affects the overall soil moisture statistics (e.g. mean, percentiles) but not the time series information contained in the remote sensing based soil moisture observations. Thus, temporal autocorrelation structures as well as soil moisture anomaly structures are not expected to be affected by the applied blending approach.**

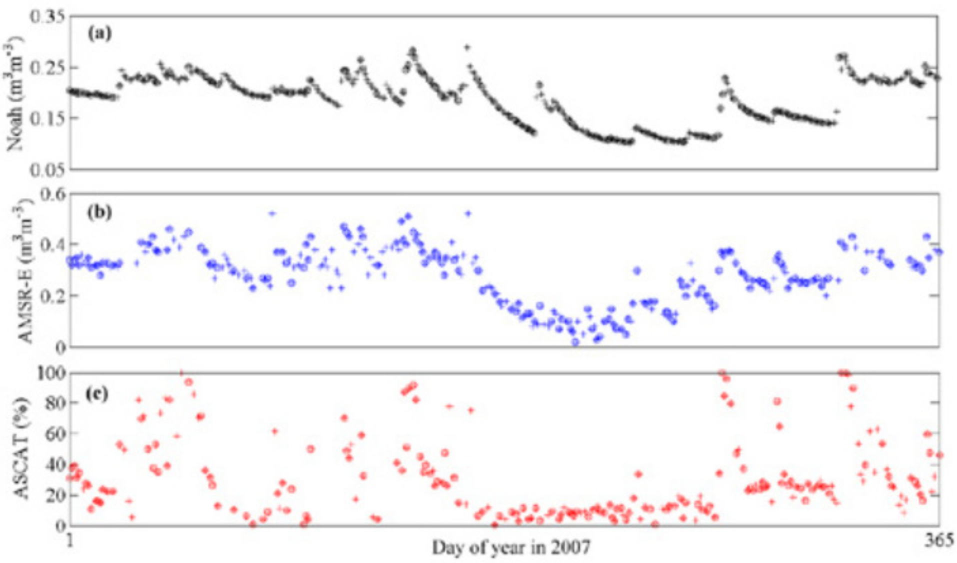

Figure 5: Example soil moisture timeseries for passive (blue) and active (red) soil moisture products as well as Noah GLDAS land surface model (black) (Fig. 2 from Liu et al., 2011)

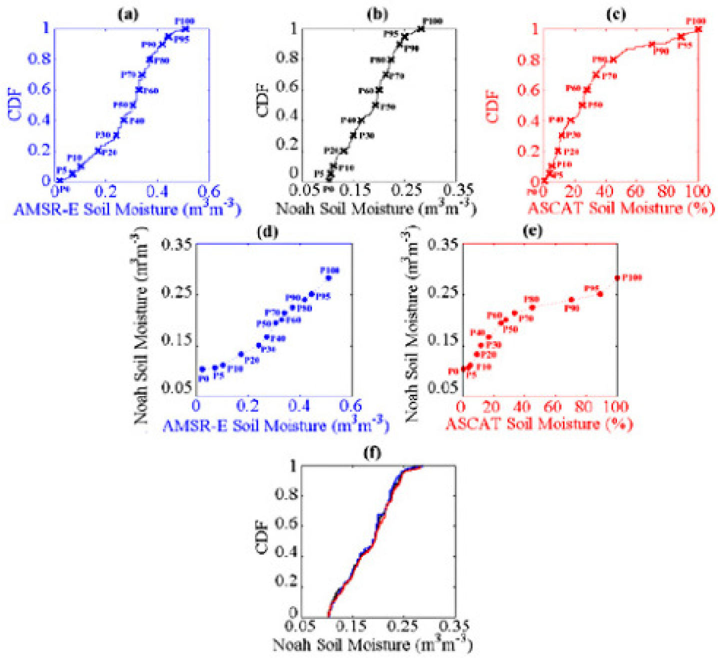

Figure 6: Example of CDF for three different soil moisture datasets (top); their relationship (middle) and final CDF statistics after matching of remote sensing datasets to Noah-LSM statistics (Fig. 3, Liu et al., 2011)

2.3.3 ERA interim soil moisture

ERA-Interim is the latest global reanalysis of the European Centre for Medium Range Weather Forecasts (ECMWF). It covers the period the period from 1979 until present and is based on a variational data assimilation system that assimilates a multitude of in situ and satellite observations in a consistent framework (Dee et al., 2011).

Soil moisture is a prognostic variable in ERA-Interim and is provided for four soil layers with thickness of 0.07, 0.21, 0.72 and 1.89 [m]. The ERA-Interim soil moisture data was extracted from the ERA-Interim data archive for the period 1979-2010.

2.4 GPCP precipitation data

The rainfall data used in this study is based on gauge corrected satellite retrievals. The Global Precipitation Climatology Project (GPCP, v2.2) (Adler et al., 2003) data product is used in the present study. It has a spatial resolution of 2.5° x 2.5° and has been available since 1979. It is based on a blended gauge-satellite product which combines precipitation retrievals from polar-orbiting passive microwave imagers (SSM/I) as well as geostationary observations (IR data).

The GPCP data is independent from the reanalysis data used as a forcing for the MPI-ESM model. It therefore allows for the analysis of precipitation and soil moisture dynamics independently from the reanalysis data which was used for the forcing of JSBACH.

3 Methods

3.1 Data preprocessing

A harmonized pre-processing of the input datasets is applied (Figure 7). These comprise:

1) **Data quality flagging**: The satellite soil moisture data sets provide a multitude of quality flags which indicate if a reliable soil moisture estimate is expected to be delivered in the data product. Snow covered areas or frozen ground is typically masked and no data is available. In addition, dense vegetated areas with high optical depth are not expected to provide reliable soil moisture estimates. The following data was excluded from the analysis:
   AMSR-E: quality flags: missing data, swath edges, water bodies, snow and glaciers/frozen ground
   ECV_SM: quality flags: regions with frozen ground or with soil temperatures below zero, region with dense vegetation, regions without valid soil moisture estimates
2) **Daily totals**: ERA-Interim comprise subdaily data. These were averaged prior to any further post-processing.
3) **Overlay creation**: The AMSR-E data sets provides daily data for ascending and descending orbits. An overlay was created that merges both datasets by using the average of both sets if both are available and the one or the other if the ascending or descending field contains missing values at a given grid cell and time step. This procedure decreases the overall amount of missing data.
4) **Regridding**: The soil moisture data is regridded from the native resolution of the original datasets to the MPI-ESM grid which has a resolution of T63 (approx. 2°) using a conservative remapping functionality of the Climate Data Operators (cdo)[1]. Some data gaps were removed by this procedure.
5) **Temporal smoothing**: The data was smoothed by calculating a five days running mean filter (via cdo) that removed very high frequency variations from the data. Additionally, it was used to fill short data gaps and therefore compensated some missing data.

[1] Climate data operators are available from: https://code.zmaw.de/projects/cdo

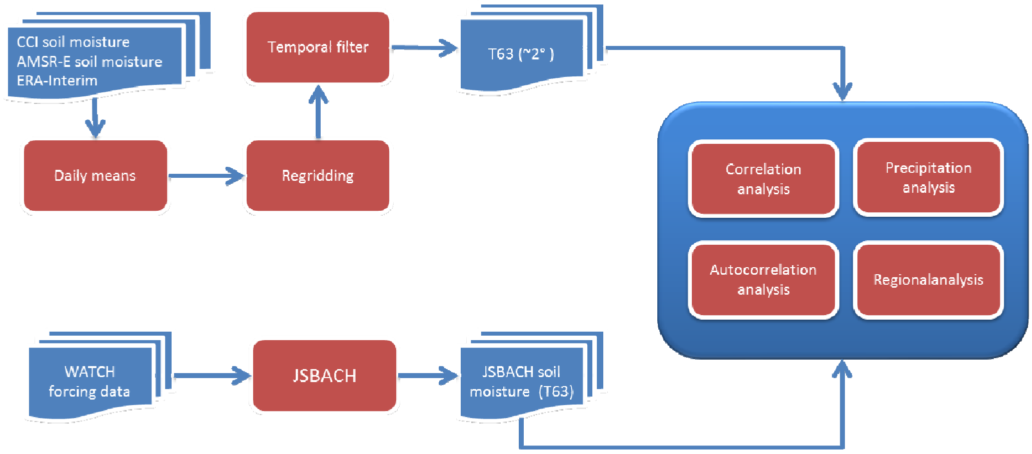

Figure 7: Analysis flowchart

3.2 Model simulations

To allow for a consistent comparison of JSBACH soil moisture simulations and satellite observations, only offline simulations of JSBACH with WATCH forcing data were conducted. Simulations were done for the period 1979-2009 on a T63 model grid. The model was initialized with PFT distributions from the LSP2 dataset (Hagemann, 2002).

3.3 Data analysis

The analysis methods are guided by the following research questions:

1) How reliable are the spatiotemporal soil moisture patterns simulated by JSBACH?
2) Does JSBACH show comparable soil moisture memory effects like the satellite observations?
3) Is the surface soil moisture a good proxy for root zone soil moisture dynamics in JSBACH?
4) How do satellite and model based soil moisture estimates relate to precipitation dynamics?
5) What can be learned about interannual and decadal soil moisture variability from the ECV_SM v0.1 dataset?

3.3.1 Soil moisture statistics

The spatiotemporal soil moisture patterns of all datasets are analyzed by comparing

a) climatological and seasonal means of global soil moisture fields
b) the percentile distribution of soil moisture for each data grid point
c) the temporal evolution of zonal mean soil moisture

The similarity between spatial fields of percentiles and global mean soil moisture fields are analyzed by calculating the correlation between two maps of soil moisture fields from the different datasets.

3.3.2 Autocorrelation analysis

The memory effect in a soil moisture timeseries $x(t)$ can be estimated from its autocorrelation function $R(\tau)$ which is given by

$$R(\tau) = \frac{E[(x_t - \mu)(x_{t+\tau} - \mu)]}{\sigma^2}$$

whereas τ is the lagtime and μ is the average of the soil moisture timeseries. The autocorrelation function is calculated for each soil layer of JSBACH and ERA-interim as well as for the satellite soil moisture datasets based on the preprocessed daily timeseries. The autocorrelation length γ, defined as the lag where $R(\tau) = e^{-1}$, is used to compare the soil moisture memory effects in the different soil moisture datasets.

3.3.3 Representativeness of the surface soil moisture information

As satellite soil moisture data provides only information on the soil moisture content of the upper few centimeters of the soil, it is crucial to investigate how the surface soil moisture dynamics relates to the deeper soil moisture dynamics. The JSBACH surface soil moisture time series are therefore correlated with the deeper soil moisture layers. The Pearson-Product Moment coefficient is used for that purpose.

3.4 Application for climate studies

The capability to capture significant climate anomalies is a crucial property of an ECV data record. Dorigo et al. (2012) have analyzed global linear soil moisture trends based on the ECV_SM v0.1 dataset. They found significant changes in the surface soil moisture in different regions of the globe.

In the present analysis we focus on the analysis of a long term and a short term climate anomaly and their respective representation in the CCI data product to evaluate the general information content of the ECV_SM data record.

3.4.1 The Sahel drought

The devastating drought in the Sahel belt in Africa in the last half of the 20th century has been the largest climate anomaly observed so far in recent times using satellite observations. The negative precipitation anomalies started in the 1960ies with a minimum around 1980 (Figure 8). Since this minimum, the rainfall recovered and it has been shown in various studies that the vegetation in the Sahel recovered subsequently (Olsson et al., 2005; Hickler et al., 2005; Fensholt et al.2012).

The ECV_SM data set is the first ever available observation-based soil moisture data product that allows for investigation into the relationship between soil moisture, precipitation, vegetation dynamics in the Sahel for over three decades. In this report, we analyze whether the soil moisture observations are able to capture the observed precipitation anomalies as well as how the observed soil moisture dynamics are related to the observed vegetation dynamics. Global observations of Normalized Vegetation Index (NDVI) from the GIMMS data set are used for that purpose (Tucker et al., 2005).

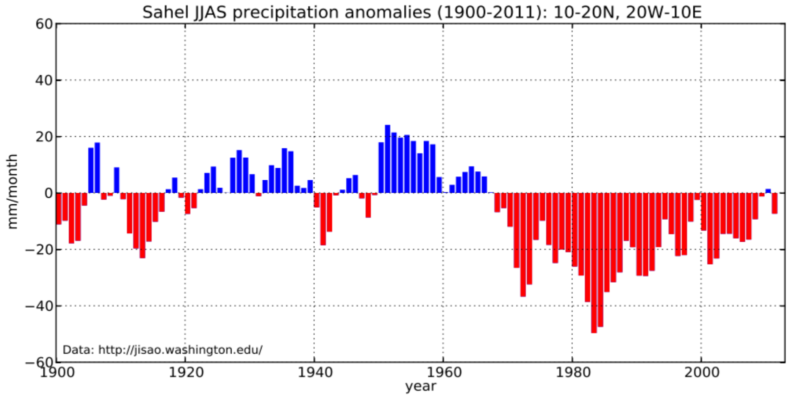

Figure 8: Sahel rainfall anomalies (1900-2011)

3.4.2 The European heat wave 2003

The European heat wave of the summer 2003 was an extreme climate anomaly that affected large parts of the European continent. The mean summertime temperatures exceeded the 1961–1990 average by about 3°C to 5°C regionally which corresponds to 5 standard deviations (Schär et al., 2004). During the first heat wave in May 2003, temperatures raised up to 30°C in Central and Southern Europe (Ferranti and Viterbo, 2006). It was very likely the hottest summer over the past 500 years (Luterbacher et al., 2004). The socioeconomic impact was disastrous. An excess above the mean mortality rate was observed across Europe, resulting in an increase of the mortality by 70 000 heat related deaths (MunichRe, 2008). Forest fires in Portugal resulted in an economic loss of US$ 1.6 billion (Heck et al., 2004) and the severe drought resulted in uninsured crop losses in Europe totaling about US$ 12.3 billion (Schär & Jendritzky, 2004). Alone in France the official statistics estimated a decrease of crop yield in the order of 15%– 28% (Zaitchik et al., 2006).

The evolution of the year 2003 heat wave has been simulated using Regional Climate Models (RCM) (Ferranti and Viterbo, 2006; Fischer et al., 2007). It was found, that the remarkable positive temperature anomalies, which resulted from an anomalous increase of sensible heat flux, are highly likely to be amplified by the spring soil water deficit, although it was not the cause of the event.

The analysis in this report will analyze, how the used soil moisture data sets capture the anomalous year 2003 soil moisture deficit in spring 2003 and its evolution throughout the year 2003.

4 Results

First, the impact of the data pre-processing is analyzed and it is investigated how representative the JSBACH surface soil moisture is for the soil moisture profile. The satellite soil moisture datasets and ERA interim and JSBACH are compared thereafter.

Finally, the detailed analyses of the two test cases (Sahelian drought, European heat wave) are presented.

4.1 Spatiotemporal data coverage

The satellite soil moisture data sets have data gaps which are due to an insufficient spatial daily coverage of the used sensors. Different pre-processing steps (section 3.1) have been performed to merge data from the same date and regrid the data to the appropriate grid of the MPI-ESM. Figure 10 shows an example of the effect of the different pre-processing steps on the completeness of the data for AMSR-E (X-band). By merging the two original (ascending, descending) data products, a daily composite image is obtained which has already nearly global coverage (except for snow covered and frozen areas). A gap free dataset is obtained by temporal smoothing of the data using a running mean 5-day average kernel.

Figure 11 shows the overall fraction of missing data in each grid cell compared to the entire time period analyzed. For the AMSR-E combined product, the mean fraction of missing data is 50% ± 19% and 57% ± 20% for the ascending and descending orbits respectively. Through the combination of the two products, the fractional coverage of the data increases significantly. Only 40% ± 28%of the data is missing on average. However, there are still large fractions of data gaps in the tropics due to dense vegetation as well as in the Northern latitudes. By aggregating the data to the coarse scale model grid (T63, 2.5 x 2.5 degree), most of the gaps in the tropics vanish due to the fact that each grid cell contains also areas where the satellite observations are sensitive to soil moisture dynamics. The temporal filtering with a 5-day average kernel reduce further the amount of time periods without data (16%± 26% of missing data). Data gaps for up to 50% of the time nevertheless remain in the Northern Latitudes.

Figure 12 shows the fraction of missing data for the ESA ECV_SM v0.1 dataset in different preprocessing steps. The raw ECV_SM data contains on average data gap fraction of 73% (± 17%). The remapping and smoothing of the timeseries decreases here also the fraction of data gaps to 60% (± 16%) after remapping and 30% (± 23%) after temporal smoothing. However the fraction of data gaps is in general higher for ECV_SM v0.1 than for AMSR-E. This is in particular the case for the high latitudes in the Northern Hemisphere. The ECV_SM dataset is based on data from a multitude of sensors. Especially in the first decade of the dataset, a large number of data gaps occurred due to the poor spatial coverage of the instruments (Figure 9). In the first pentad of the 1980ies, nearly 50% of the year was without data coverage. This explains a large portion of the data gaps still observed in the finally pre-processed dataset. Further, the merging technique, developed by Liu et al. (2011) is based on a binary blending of different soil moisture products (see Fig. 11 in Liu et al., 2011). This binary blending technique is applied for different time periods. In case of a failure of a particular instrument, no soil moisture data is reported in the ECV_SM dataset.

- The ECV_SM v0.1 data provides unique multidecadal soil moisture information from satellite observations
- The spatial and temporal data coverage is very heterogeneous and several preprocessing steps are required to generate a dataset which is compareable to climate model simulations.
- Major limitations in the data coverage are before 1988.

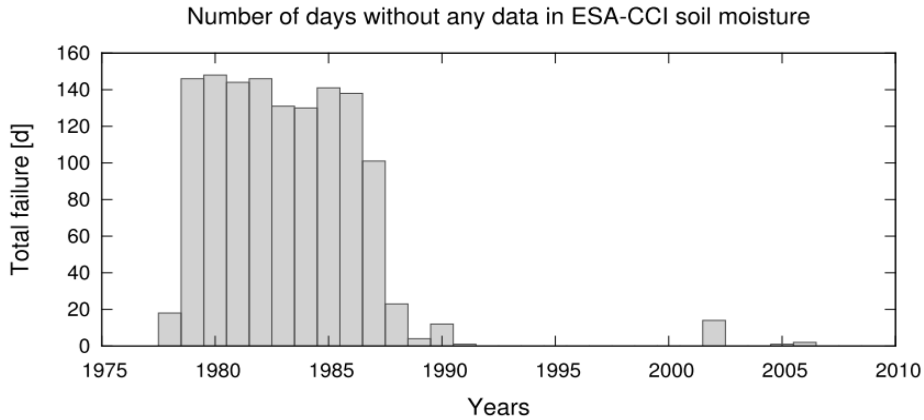

Figure 9: Number of days per year without any soil moisture observations in ECV_SM v0.1

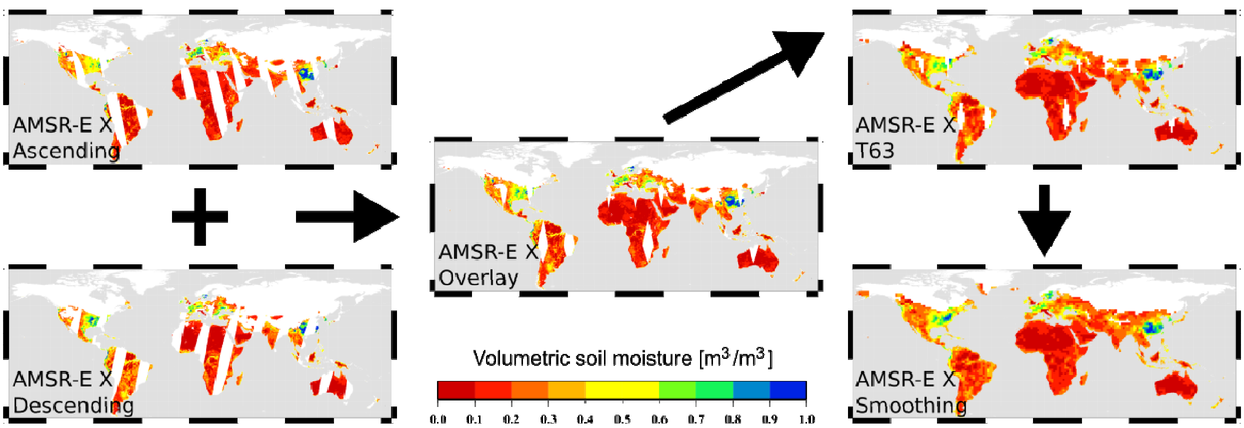

Figure 10: Example of effect of data preprocessing on data coverage for AMSR-E X-band for 15th of December 2005: original ascending/descending data coverage (left), daily merged mosaic (middle), resampled data on T63 grid and 5-day average (right)

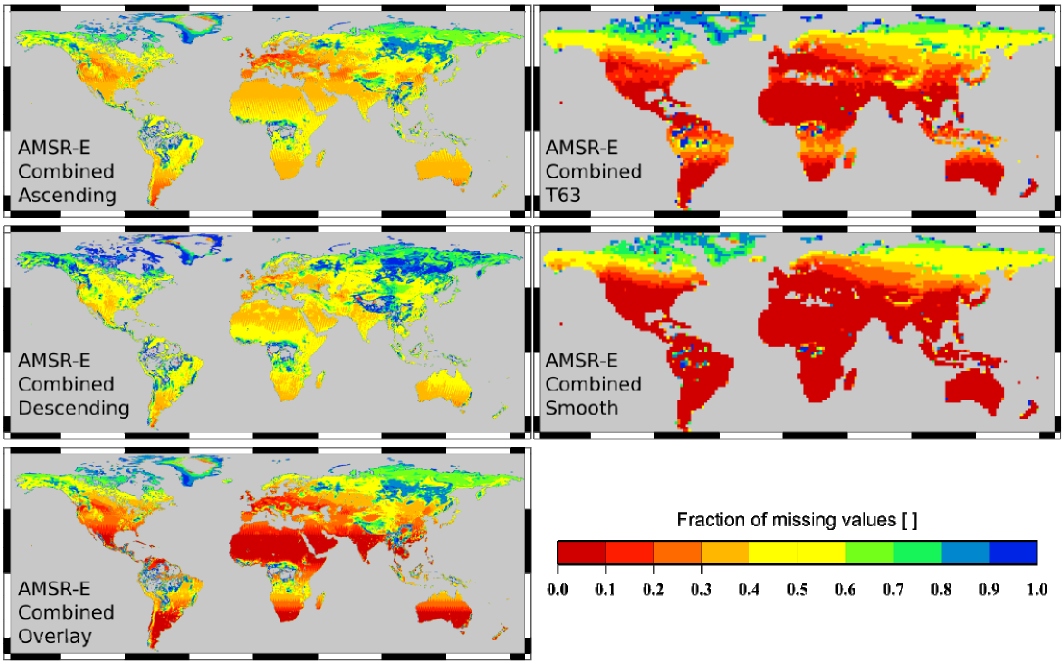

Figure 11: Effect of data pre-processing on fractional coverage of missing data for AMSR-E combined product.

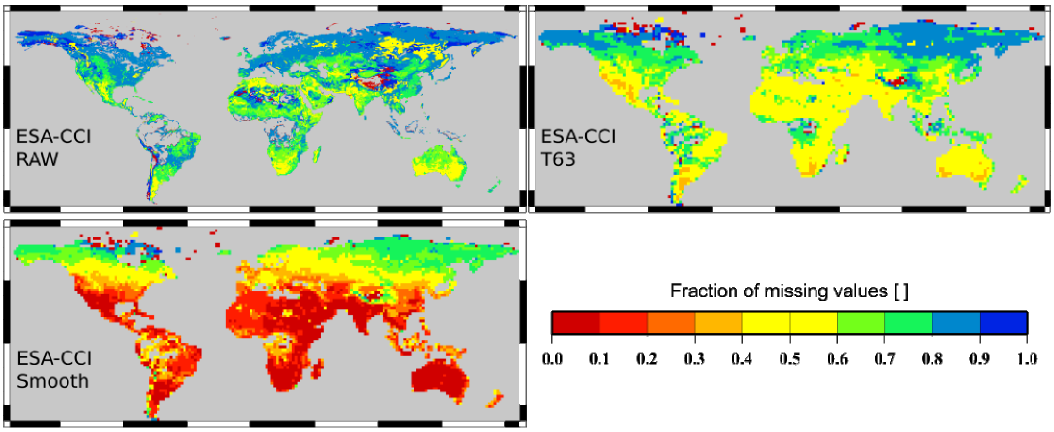

Figure 12: Fractional coverage of missing data for ECM_SM v0.1 in different pre-processing steps.

4.2 How representative is surface soil moisture?

As satellite observations provide only information on surface soil moisture dynamics, it is of vital interest to investigate how the surface soil moisture dynamics in JSBACH relates to the root zone soil moisture dynamics. The correlation between the surface soil moisture and the lower soil layers in JSBACH were therefore calculated for a) daily soil moisture data and b) daily soil moisture anomalies. Figure 13 shows the correlation coefficient between the uppermost soil moisture layer of JSBACH and the layers below while Figure 14 shows the correlation of surface soil moisture with the root zone soil moisture. Very high significant correlations are obtained for the second soil layer (0.19 m) and significant correlations are obtained for the deeper soil layers for most parts of the globe. The last soil layer (6.98) shows significant correlations for a much smaller number of grid boxes. The reason for the different behavior of the last soil layer is, that this soil layer is not active in all grid boxes for the JSBACH simulations as for large parts of the world the depth of the bedrock is already reached after the first four soil layers. Figure 14 shows in general high correlation of the surface soil moisture with the root zone soil moisture content.

In general, there is a very good agreement of the first JSBACH soil moisture layer with at least the upper first meter of the soil and even deeper soil layers. This linkage between the soil layers is likely to be even further enhanced when analyzing lagged correlations between the surface soil moisture and the root zone soil moisture, as the soil moisture dynamics in the deeper soil is typically damped and temporally lagged compared to the surface soil moisture dynamics.

These initial results indicate that surface soil moisture observations might be a reasonable proxy for deeper soil moisture dynamics, which is more relevant for the memory effects of soil moisture in JSBACH.

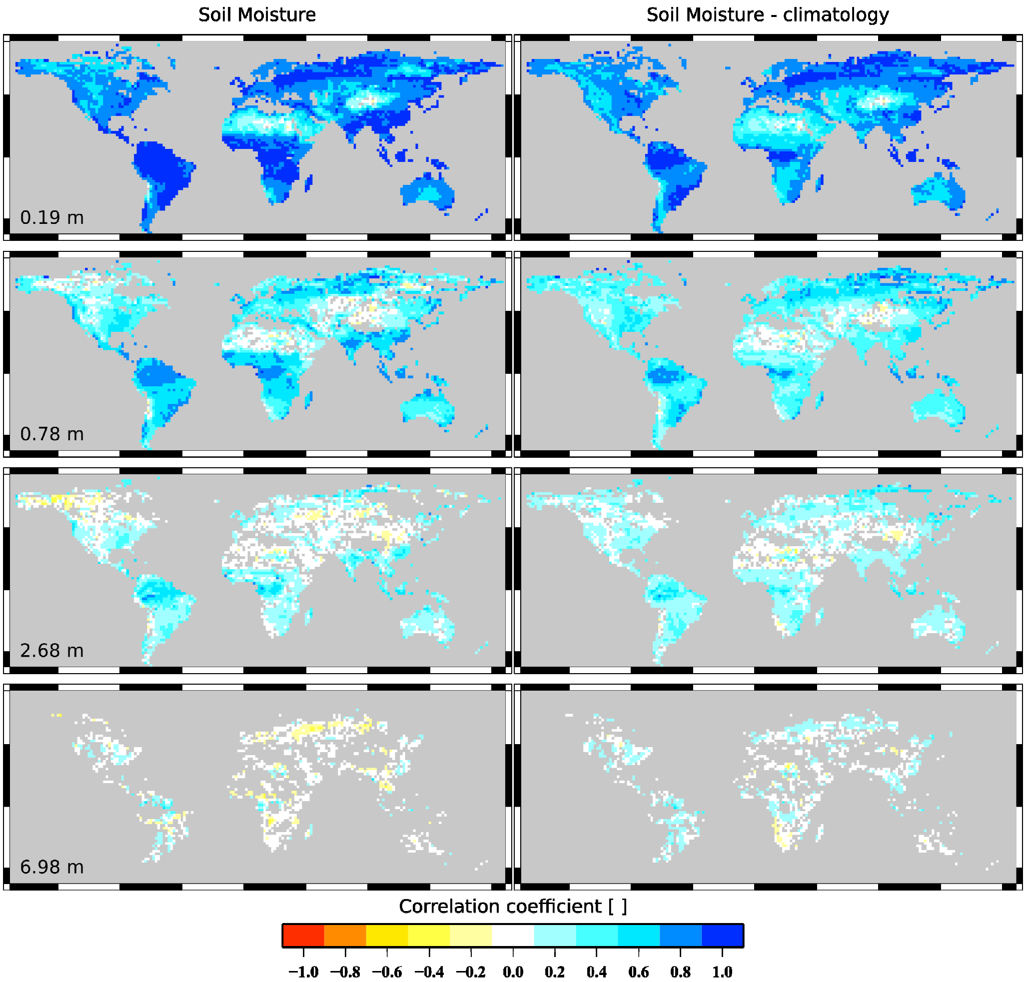

Figure 13: Correlation coefficient for unlagged correlation between surface and deeper soil layers (top-bottom) in JSBACH. Left: daily soil moisture data, right: daily soil moisture anomalies. Only significant (p<0.05) correlations are shown.

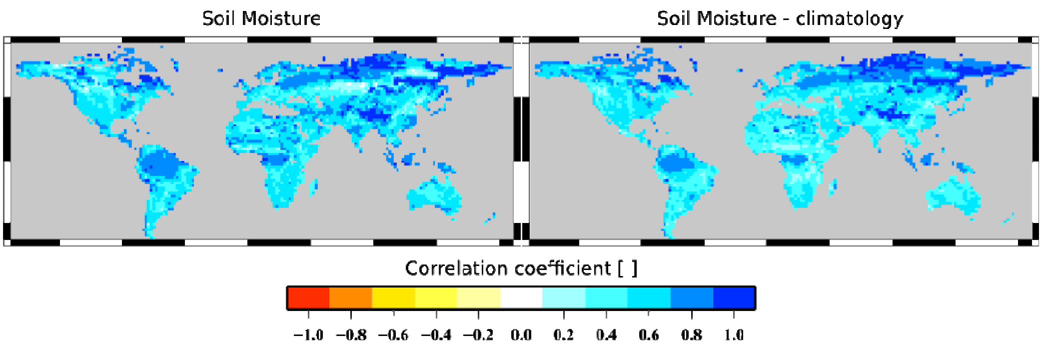

Figure 14: same as Figure 13, but for correlation between surface soil moisture and total root zone soil moisture.

4.3 Global analysis

4.3.1 Do model simulations and observations show similar soil moisture dynamics?

4.3.1.1 Global mean fields

The global mean soil moisture field for the years 2003-2009 and its coefficient of variation (CV) is shown in Figure 15. The time-latitude diagrams of zonal mean soil moisture are shown in Figure 16 for the whole period.

The ECV_SM v0.1 and AMSR-E CDF-matched data set show very similar mean fields. Contrary, the combined AMSR-E soil moisture product differs largely from the other data sets. Especially in the Northern Latitudes, the soil moisture values exceed physically meaningful values ($\theta > 0.6$). This is a well known problem in the existing dataset and is under investigation by the data providers (R. de Jeu, pers. comm.). The reason is still unknown and might be related to ponding in wetland areas (Gouweleeuw et al., 2012) or instrument related artefacts as it was observed that high soil moisture values occur much more often for AMSR-E observations than for e.g. Windsat observations.

The JSBACH simulations show a reasonable good agreement with the ECV_SM data set. ERA-interim is in general wetter than ECV_SM and JSBACH. The spatial pattern of the CV differs largely among the different datasets. JSBACH shows a much larger (relative) soil moisture dynamics compared to the other datasets. Especially in semiarid areas in Africa and Australia, the CV exceeds 50%. Contrary, ERA-interim shows very small temporal variability. On global scale, the CV rarely exceeds 15%.

The time-latitude diagrams in Figure 16 show distinct differences in the spatiotemporal mean soil moisture fields. The biases between the different datasets are clearly observable. Inconsistencies in the ECV_SM data record are clearly detectable in the figure. A drying of the zonal mean soil moisture is for instance observed in 2002 (integration of AMSR-E data). Further discontinuities can be observed in 1991 (ERS scatterometer), 1978 and 2006/2007 (METOP). Between 1988 and 2006, the latitudes between 30°S and 18°S are wetter than in the other years, while a the data coverage in latitudes below 30°S is becoming more sparse. As discontinuities in ECV time series can easily introduce artificial trends in timeseries that could be misinterpreted as a climate signal, a careful analysis of such kind of timeseries is important (Loew and Govaerts, 2010). Dorigo et al. (2012) have evaluated global trends in surface soil moisture based on the ECV_SM v0.1 dataset. To minimize the effect of outlieres and discontinuities they used a non-parameteric rank correlation technique.

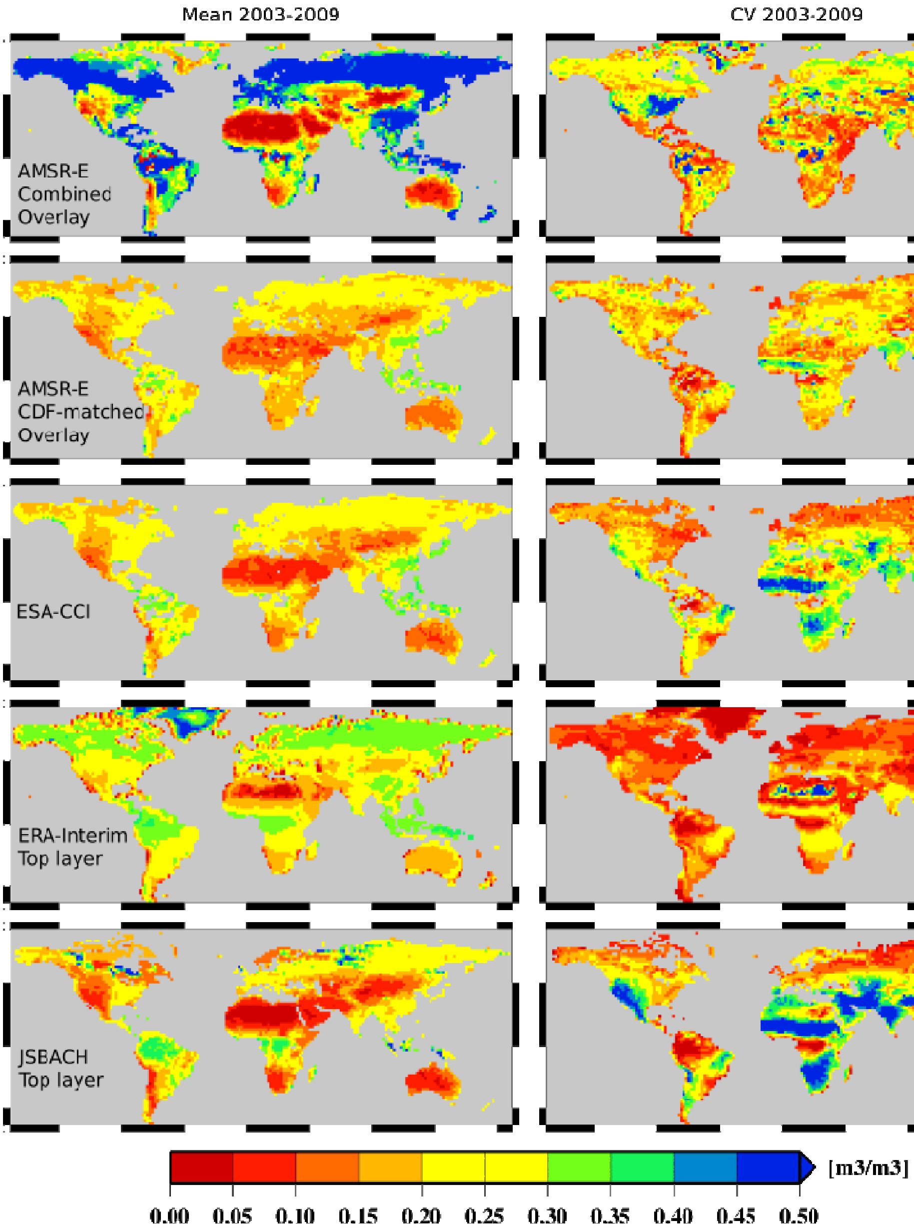

Figure 15: Mean surface soil moisture of different data sets (left) and coefficient of variation (CV) of soil moisture (right) for the period 2003-2009

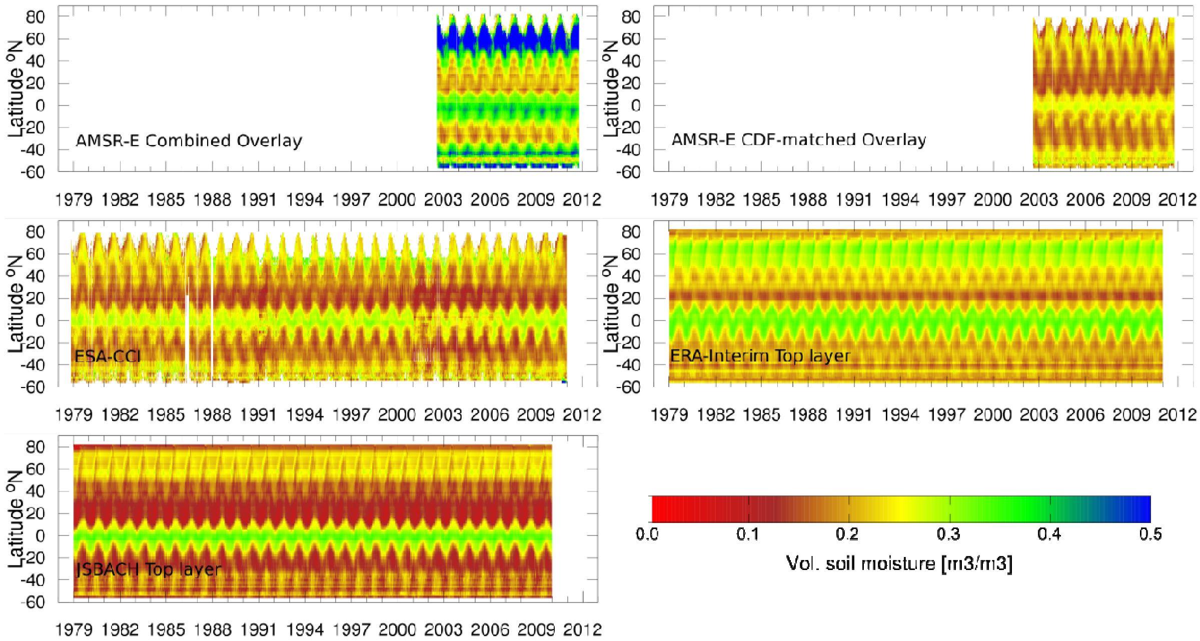

Figure 16: Time-latitude diagrams for the 5-daily smoothed surface soil moisture data sets. Colors represent zonal mean soil moisture. Be aware that the AMSR-E data is only available since 2002. Scaling on x-axis therefore differs from the other plots

4.3.1.2 Percentile distribution

The percentiles of the soil moisture were calculated from the time series of each datasets for each grid cell. Figure 17 and Figure 18 show the 5% and 95% percentiles for the different surface soil moisture data sets which are considered as the lower (dry) and upper (wet) limits of the soil moisture dynamics. The AMSR-E data shows again unrealistically high (>0.6) soil moisture for the 95% percentiles in the Northern Latitudes. ECV_SM and CDF matched AMSR-E show very similar patterns for the 5% as well as the 95% percentile respectively which would be expected as both datasets were statistically matched on the same GLDAS reference soil moisture data set. The JSBACH soil moisture fields show very good agreement with the ECV_SM 5% and 95% percentiles, while ERA-interim shows in general wetter values for both, the dry and wet case.

The similarity between the spatial patterns of each percentile was compared by calculating the correlation coefficient between the percentile maps of two datasets. Results of this correlation analysis are shown in Figure 19. Overall the JSBACH model shows very good agreement in the percentile distribution with the ECV_SM dataset. The correlation over all percentiles is in the order of r=0.8. The original AMSR-E data shows smallest values in the percentile correlation, while the AMSR-E CDF matched data product is in reasonable agreement with JSBACH and ECV_SM.

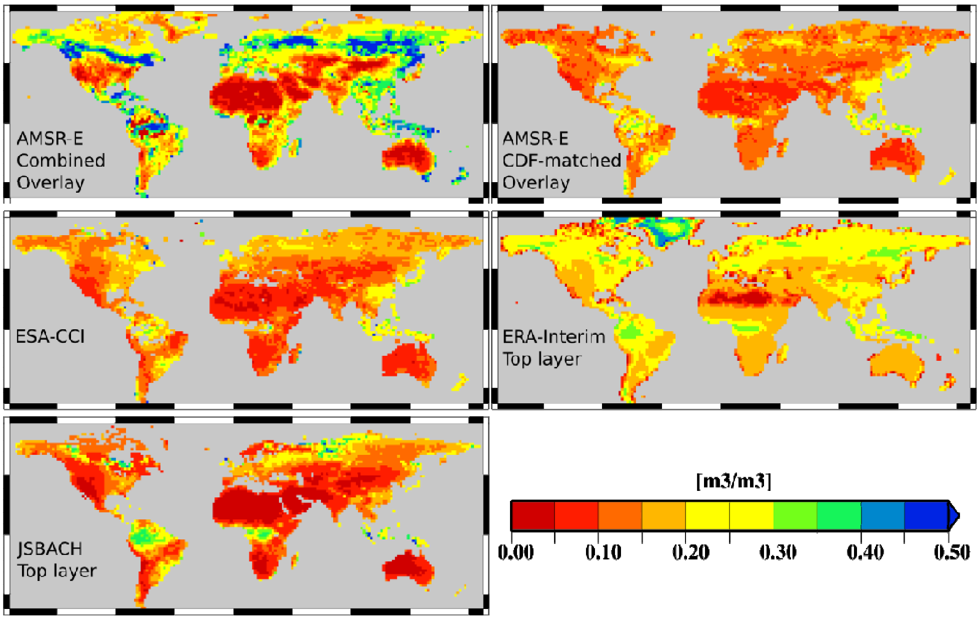

Figure 17: Maps of 5% percentiles for different soil moisture datasets

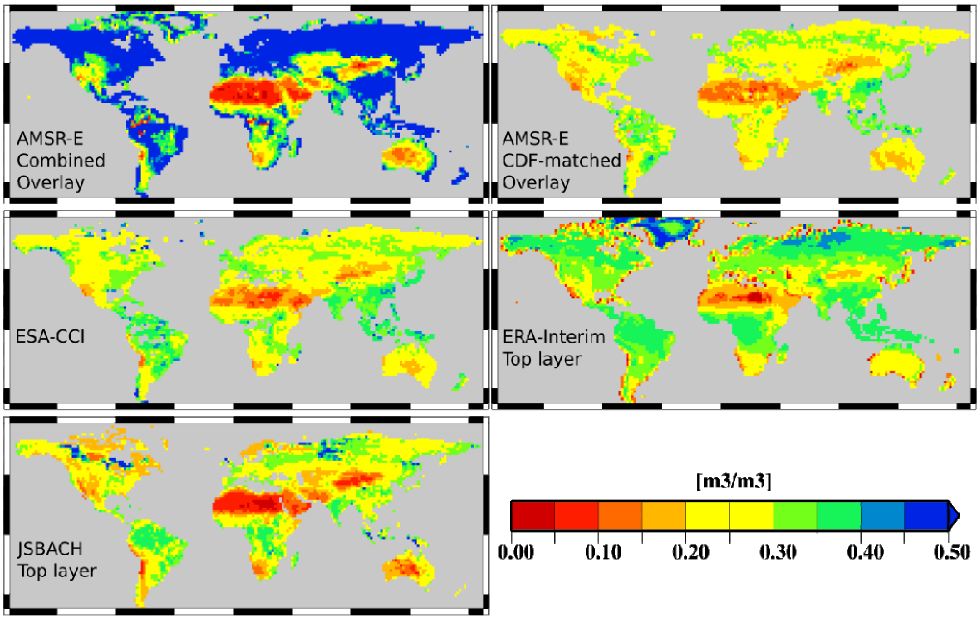

Figure 18: Maps of 95% percentiles for different soil moisture datasets

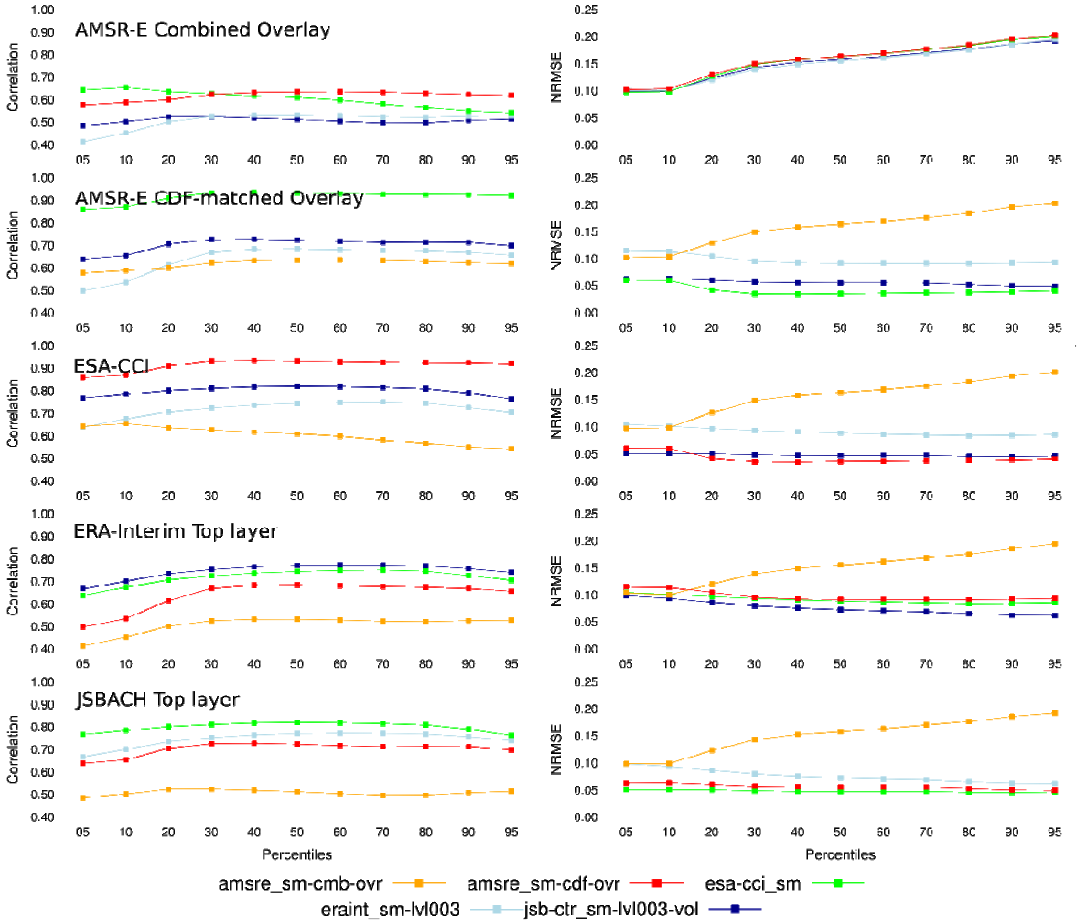

Figure 19: Correlation coefficients (left) and NRMSE (right) for relationship between soil moisture percentiles of the dataset specified in the plot title and the other soil moisture datasets. All analysis is based on data smoothed with a 5-day running mean filter. NRMSE corresponds to the root mean square error, normalized by the data range.

- JSBACH percentile distribution shows very good agreement with ECV_SM and AMSR-E CDF matched data. It remains however unclear, if this is a good argument for a skillful model or an artifact of the CDF matching procedure that simply reveals similarities between the Noah land surface model and the other land surface models investigated here.

4.3.1.3 Correlation between soil moisture fields

Figure 20 shows the correlation between the surface soil moisture and surface soil moisture anomalies of all datasets investigated. Very high significant correlations are observed for the 5-daily mean timeseries between all datasets. This indicates in general a very good skill of either the observational datasets as well as the land surface models to reproduce soil moisture dynamics at the model grid scale. Larger differences are observed in the Northern Latitudes due to the poor data coverage due to snow and frozen soil conditions. Negative correlations are here especially observed between the observations and both models.

- Both, models and observations show reasonable agreement in the observed and simulated surface soil moisture fields.
- No useful comparison seems to be possible in high latitude areas.

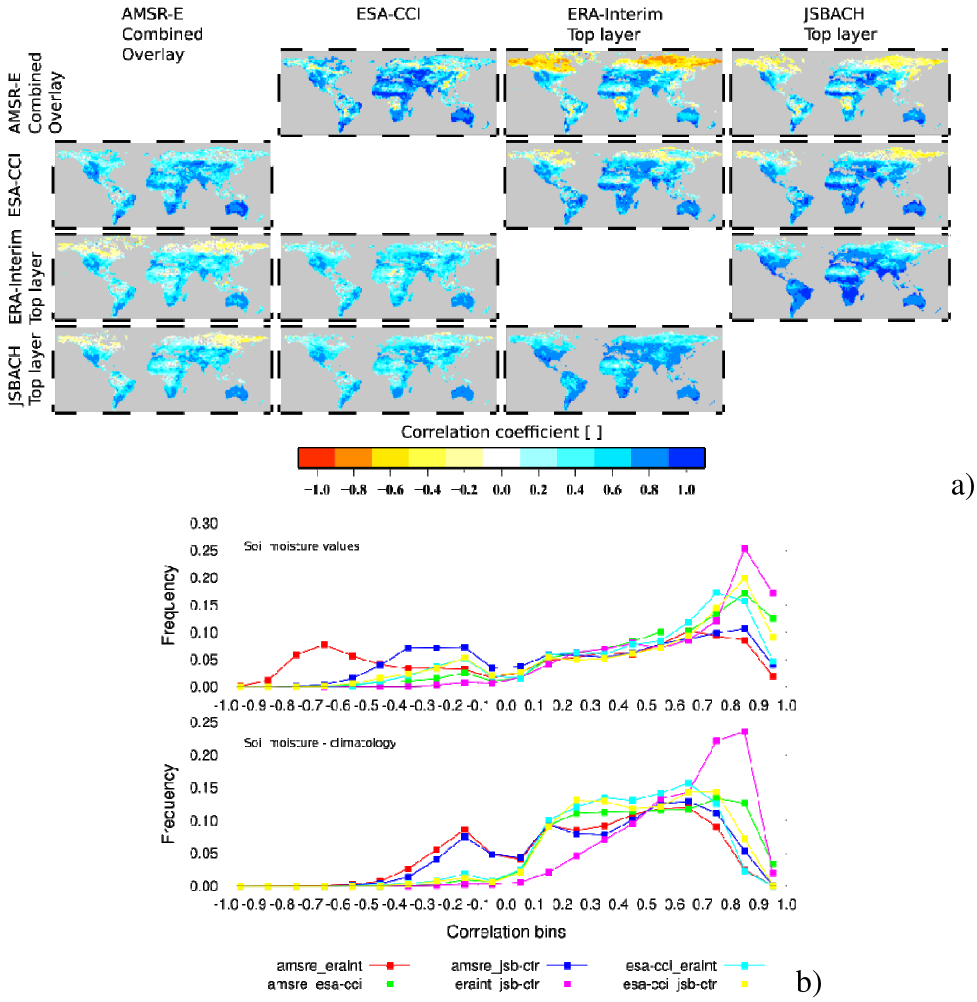

Figure 20: Correlation coefficient for the linear correlation between the different surface soil moisture datasets: upper triangle is based on 5-daily smoothed soil moisture values for the common period 2003-2009 while lower triangle is based on soil moisture anomalies

4.3.1.4 Discussion

The previous analyses have shown consistencies and inconsistencies between the various soil moisture datasets in terms of a) mean soil moisture fields, b) the temporal and spatial variability, c) correlation between soil moisture datasets and d) comparison of the statistical distribution of the datasets on a grid cell basis.

The purpose of this analysis was to use the satellite soil moisture observations as an independent dataset to evaluate the reliability of the soil moisture dynamics in JSBACH. It was assumed that JSBACH would have a reliable soil moisture dynamics and statistics of the spatial pattern of the soil moisture percentiles between JSBACH and ECV_SM would be in agreement. The latter was shown in the previous analysis.

However, the soil moisture statistics (percentiles) in ECV_SM v0.1 are not based on observational data, as the whole data record was constructed using a CDF matching technique, which used the Noah land surface model as a common reference (Liu et al., 2011). By doing so, the soil moisture dynamics in each grid cell is, by definition, bounded to the soil moisture dynamics in the Noah model and thus highly dependent on the parameterization of the Noah soil hydraulic properties. The latter are based on the FAO soil map, which is also an input to the soil map of JSBACH. Thus comparing JSBACH soil moisture statistics against ECV_SM v0.1 soil moisture statistics corresponds to indirectly compare the soil moisture statistics of two different offline land surface models. The major conclusion that can be drawn from the analysis is therefore that the Noah and JSBACH soil moisture statistics are in reasonable agreement.

- **The ECV_SM v0.1 dataset is therefore unfortunately not providing a model independent reference that can be used to evaluate the general soil moisture statistics of a climate model or offline land surface scheme. The major problem is the use of a model based soil moisture climatology as a common reference.**

What one would expect from a soil moisture Climate Data Record (CDR) is, that it provides information on the spatial distribution of soil moisture percentiles from observational evidence. This information is available in the original soil moisture datasets which have been merged in the ECV_SM v0.1 product, but it is lost due to the CDF matching technique normalizing the data to the Noah land surface soil moisture statistics. Errors in the soil moisture parameterization of the Noah model directly translate into errors in the ECV_SM statistics. Contrary, if one believes in the Noah model statistics, there would be no need to use ECV_SM for climate model evaluation studies if one could directly compare with the Noah GLDAS output.

Thus, the ECV_SM is expected to loose information on the spatial pattern of soil moisture statistics. It is however emphasized that the temporal dynamic in the dataset is not affected by the normalization procedure. Thus, comparing temporal soil moisture anomalies and compare these against observed precipitation anomalies might provide additional insight in the added value provided by ECV_SM. This will be done in the next section.

4.3.2 Relationship of soil moisture and precipitation dynamics

The relationship between precipitation and soil moisture dynamics was analyzed by comparing GPCP monthly rainfall data with the surface soil moisture data. Figure 21 shows the correlation coefficients between monthly GPCP and the surface soil moisture fields as well as between detrended GPCP precipitation anomalies and detrended soil moisture anomalies. The anomalies were obtained by first removing any linear trend in the dataset for each grid cell individually and then removing the mean seasonal cycle from timeseries of each grid cell.

The ECV_SM shows data set shows significant correlation $r = 0.41 \pm 0.22$ with GPCP. ERA-interim shows worst correlation with GPCP ($r = 0.31 \pm 0.41$) with significant negative correlations in the Northern latitudes, while JSBACH shows highest correlation values with GPCP ($r = 0.56 \pm 0.28$). The frequency distribution for the correlation with GPCP shows a clear bimodal distribution for ERA-interim. While the distribution of correlation coefficients is rather similar to that of JSBACH, ERA-interim shows much more negative correlation values. This might be related to snow effects in the Northern Latitudes. The ECV_SM data set shows significant positive correlations for most part of the globe. The correlations with GPCP are however much lower than for ERA-interim or JSBACH. The modal value for ECV_SM is $r \approx 0.8$.

The soil moisture timeseries still contain the typical precipitation and soil moisture seasonality as well as potential long term trends in precipitation and soil moisture. To analyze the skill of the different soil moisture datasets to reproduce precipitation anomalies, a correlation analysis was performed also on the detrended soil moisture anomaly timeseries. JSBACH shows the highest correlation $r = 0.56 \pm 0.16$ between GPCP anomalies and soil moisture anomalies. The ERA-interim shows lower skill in simulating the soil moisture anomalies ($r = 0.38 \pm 0.25$) and ECV_SM shows smallest anomaly correlations with monthly precipitation anomalies ($r = 0.33 \pm 0.16$). The frequency distributions in Figure 21 show a shift of the correlation coefficients to smaller values for all datasets. However, only a small shift is observed for JSBACH, which indicates still a strong dependence of JSBACH surface soil moisture to precipitation dynamics at monthly timescales. This relationship is smaller for ERA-interim, where the correlations drop faster than for JSBACH. The histogram of the ECV_SM dataset is skewed to lower correlation values, which indicates that the high correlation values observed between GPCP and ECV_SM were mainly due to common seasonality between precipitation and soil moisture.

We conclude from this analysis that JSBACH, forced with the WATCH forcing data provides reasonable estimates of soil moisture dynamics. One might argue, that the results indicate that satellite based soil moisture information do not provided additional information for land surface modeling as long as they are not superior to the precipitation data. However, the used

GPCP precipitation data has still large uncertainties and the opposite might be actually true, namely that the deviation of the ECV_SM data to the precipitation data reveals additional information about evaporative loss of surface soil moisture or different timescales of soil moisture and rainfall dynamics. Several studies have proven the general applicability of surface soil moisture information to reduce uncertainties in remote sensing based precipitation and evaporation data sets (Crow & Zhan, 2007, Liu et al., 2011a, Miralles et al., 2011).

Information on the soil moisture temporal dynamic is contained in the soil moisture autocorrelation function and provides insight into similarities and differences of temporal soil moisture dynamics of the different datasets. Details of the results of the autocorrelation analysis will be discussed in the following.

- **Further more detailed analysis of the soil-moisture precipitation interrelationship in different regions would be needed with additional in situ reference data to answer analyze the actual absolute performance of the different soil moisture datasets.**

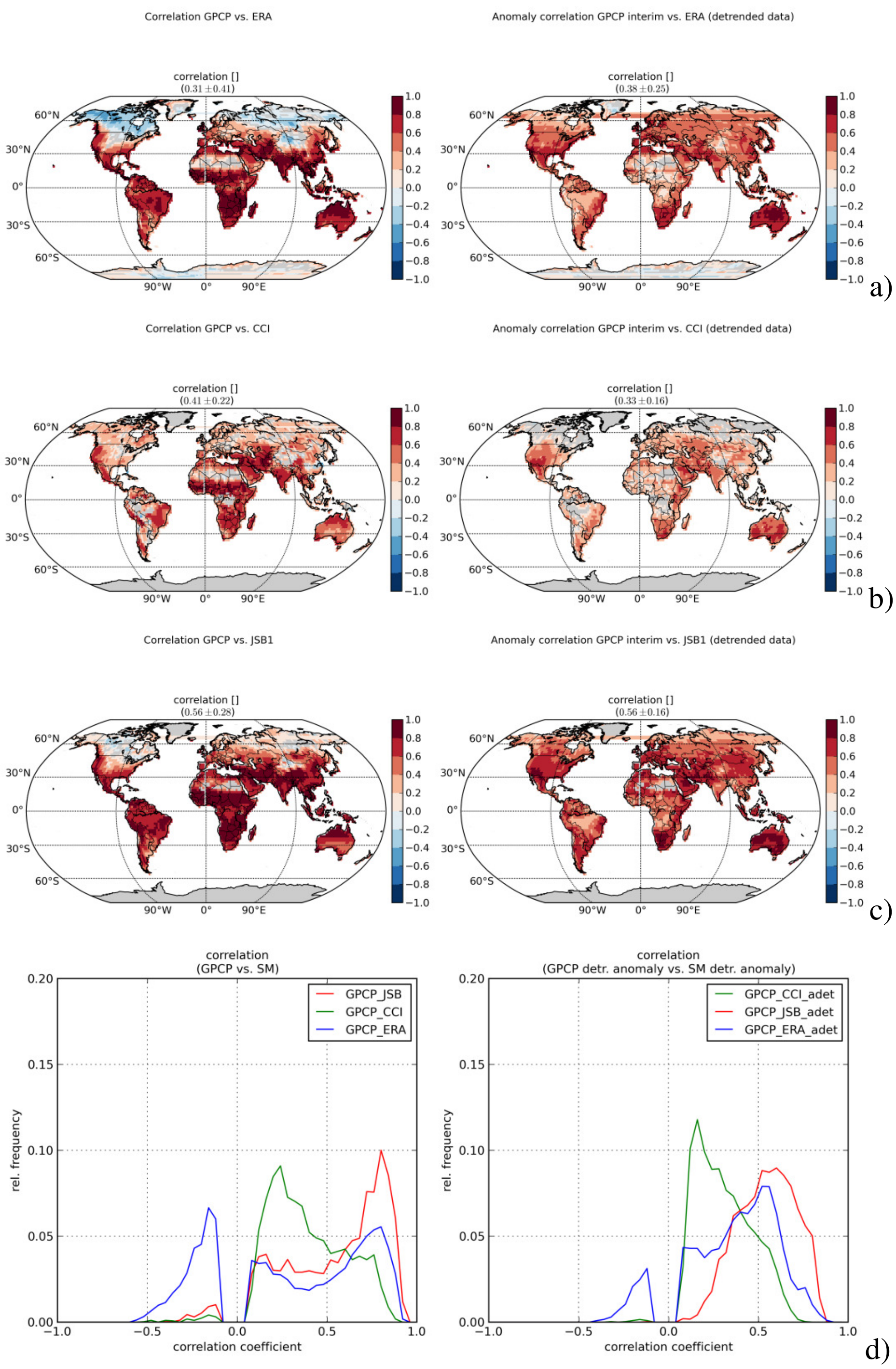

Figure 21: Significant correlation of monthly GPCP precipitation data with soil moisture (left) and GPCP anomalies with soil moisture anomalies (right) for ERA-interim (a), ECV_SM v0.1 (b) and JSBACH (c). Frequency distribution of correlation coefficient for the different datasets (d).

4.3.3 Soil moisture memory effect

The temporal autocorrelation was calculated based on the 5-daily average timeseries. Figure 22 shows examples of the estimated autocorrelation functions (ACFs) for the different soil moisture datasets and different regions. The ACFs were derived from either the original soil moisture values or the soil moisture anomalies, where the mean seasonality has already been removed. For the grid cell in North America, the JSBACH simulations show a very similar shape of the autocorrelation function and autocorrelation length compared to the ECV_SM v0.1 dataset for both, the soil moisture values as well as the soil moisture anomalies. AMSR-E and especially ERA-interim differ significantly from JSBACH and ECV_SM v0.1 for the absolute soil moisture ACF. For the soil moisture anomalies, AMSR-E shows a very similar autocorrelation pattern than ECV_SM, which indicates that both data products observe soil moisture dynamics in a coherent way. In the other example (Central Asia), the JSBACH model simulations are in between ECV_SM and ERA-interim for the ACF based on soil moisture values, while it is getting closer to the observations for the anomalies.

The results indicate that there might be good potential to exploit the surface soil moisture memory to evaluate whether a land surface model is reproducing realistically the temporal autocorrelation structures of the soil moisture dynamics. Deviations between model simulations and observations might be used to re-calibrate the land surface model's hydraulic properties on a regional to global scale.

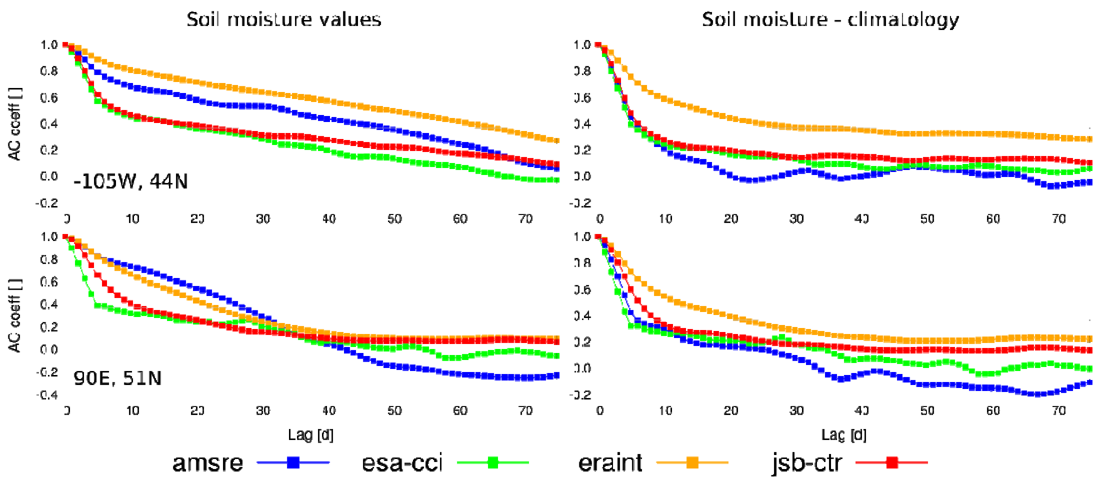

Figure 22: Autocorrelation functions for surface soil moisture datasets based on either original soil moisture values (left) and soil moisture anomalies (right). Two different regions are shown: North America (top), Central Asia (bottom)

The autocorrelation length was calculated for each of the model grid cells and all datasets. Figure 23 shows the autocorrelation length for JSBACH root zone soil moisture. The autocorrelation length is between one and two months for large parts of the globe. The autocorrelation length for soil moisture anomalies shows larger correlation lengths than for the soil moisture values in the order of six months to a year. Figure 24 shows the autocorrelation maps for surface soil moisture for the different datasets. The ACF length is in the order of one month which is consistent with previous findings (Rebel et al., 2011) and here the anomalies show a smaller autocorrelation length than the surface soil moisture data itself. The different datasets show distinct differences in the spatial pattern of the autocorrelation in some areas, while they show large agreement (e.g. in the Sahelian belt) in other areas. These differences are linked to different soil moisture dynamics as observed in the different satellite based data products as well as the different land surface models. The observed differences might offer the potential to gain further insight into the soil moisture dynamics in particular regions and might be used to recalibrate model specific parameters that control the soil moisture dynamic in JSBACH.

The good thing is, I also see this behavior in ERA-Interim. The bad thing is it cannot be validated using surface satellite observations but perhaps we can get our hands on some observations (or the behavior is already explained in the literature, I still have to look it up). In any case it is an interesting feature to discuss more deeply in a publication later on

While these preliminary results are promising, a more thorough investigation is needed, as it as shown that the autocorrelation functions of active and passive microwave observations are different in both space and time (Wanders et al., 2012; Dente et al., 2012). As the ECV_SM dataset is based on both, active and passive microwave instruments, the reason for the difference autocorrelation functions needs to be further investigated and its impact needs to be assessed.

- Using the satellite observations might therefore allow for a re-calibration of some model specific parameters.

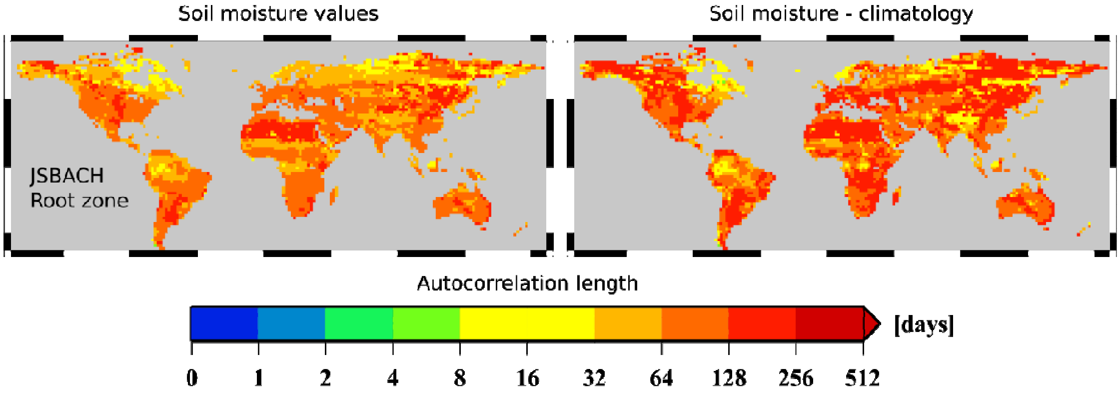

Figure 23: Autocorrelation length for JSBACH root zone soil moisture based on soil moisture timeseries (left) and soil moisture anomaly timeseries (right)

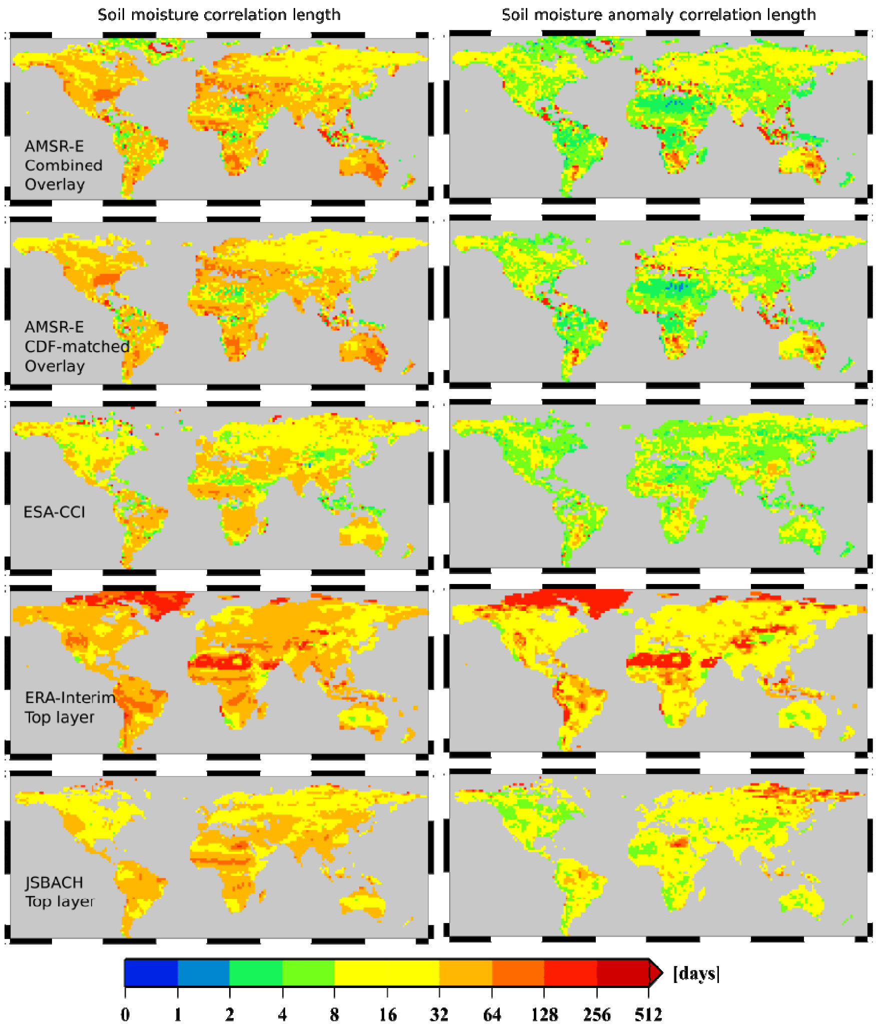

Figure 24: Temporal soil moisture autocorrelation length for soil moisture (left) and soil moisture anomaly (right) for the different soil moisture data sets

4.4 Regional test cases

4.4.1 European heat wave 2003

The severe European heat wave 2003 is not contained in the ECV_SM record. While it has been shown that remote sensing soil moisture data can provide very meaningful information on the soil moisture anomalies in the year 2003 (Loew et al., 2009), the ECV_SM data record is lacking this information. The reason for this data gap is the binary merging of different data products into the ECV_SM dataset, like introduced in Liu et al. (2011, see Fig. 11). The problem is well known and will be tackled in the frame of the ESA CCI_soil moisture project (pers. communication, W. Dorigo).

A further detailed analysis of the representation of the year 2003 European heat wave is therefore not possible in this study.

4.4.2 Sahelian drought and soil moisture variability

The precipitation in the Sahelian Region recovered since the minimum in the mid of the 1980ies. An increase in vegetation and precipitation was observed since then (e.g. Olsson et al., 2005; Huber et al., 2011; Fensholt et al., 2012). Figure 25 shows the linear trends over time for precipitation (GPCP). A clear increase in precipitation (JJAS) is observed in the Sahelian area throughout the time period covered by these datasets (1984-2005). Dorigo et al. (2012) found in the same region a significant negative trend of surface soil moisture, as derived from the ECV_SM v0.1 dataset. The positive trend precipitation and negative trend in surface soil moisture seem to be contradictory, but could be related to an increase in evapotranspiration by the increased abundance of vegetation in the area.

Figure 25 shows the linear trends of precipitation and soil moisture for the period 1979-2009 as derived from the datasets investigated in this study. ECV_SM shows significant decrease of surface soil moisture, like ERA-interim, while JSB shows no significant trends in most regions. Contrary, GPCP shows a significant increase during the same period.

We will investigate in the following the interannual and decadal variability of surface soil moisture as observed by ECV_SM v0.1 and as simulated by ERA-interim and JSBACH. As it has been shown already in section 4.3 that the different datasets show considerable biases between each other, the analysis will focus exclusively on soil moisture and precipitation anomalies.

4.4.2.1 Interannual pattern of soil moisture and precipitation anomalies

Figure 26 - Figure 29 show time-latitude diagrams for soil moisture and precipitation anomalies for the Sahelian region for all datasets. The figures differ in their input data, as they are either based on monthly mean anomalies or data from only the Sahelian rainfall season (JAS). To compare the anomalies also without a potential linear trend, the existing linear trends in the data were removed prior to the calculation of the anomalies (Figure 28,Figure 29).

ERA-interim and ECV_SM v0.1 show both a significant negative trend in soil moisture in the Sahel, while the JSBACH doesn't show such a decrease in the moisture content. The GPCP precipitation data shows a positive trend in the precipitation. This positive trend is however, mainly caused by the strong negative precipitation anomalies in the 1980ies and not significant if these years are not included in the data analysis (Loew, 2012). These negative trends in surface soil moisture for ERA-interim and ECV_SM v0.1 are even more pronounced when only the JAS season is analyzed (Figure 27). If all linear trends of the data are removed, the anomaly patterns in the different dataset evolve more clearly. The ERA-interim timeseries shows some discontinuities. A drier period (1983-1995) is followed abruptly by a wetter period which lasts until approx. 2001. A further drier period follows after 2006. It is unclear where these discontinuities are coming from, but they might be caused by a change in the observation system used in the reanalysis system.

Contrary, the ECV_SM v0.1 and JSBACH surface soil moisture shows consistent anomaly patterns with precipitation if the linear trend is removed from the data. The ECV_SM v0.1 data reproduces very well the dry and wet anomalies that are observed in the GPCP record.

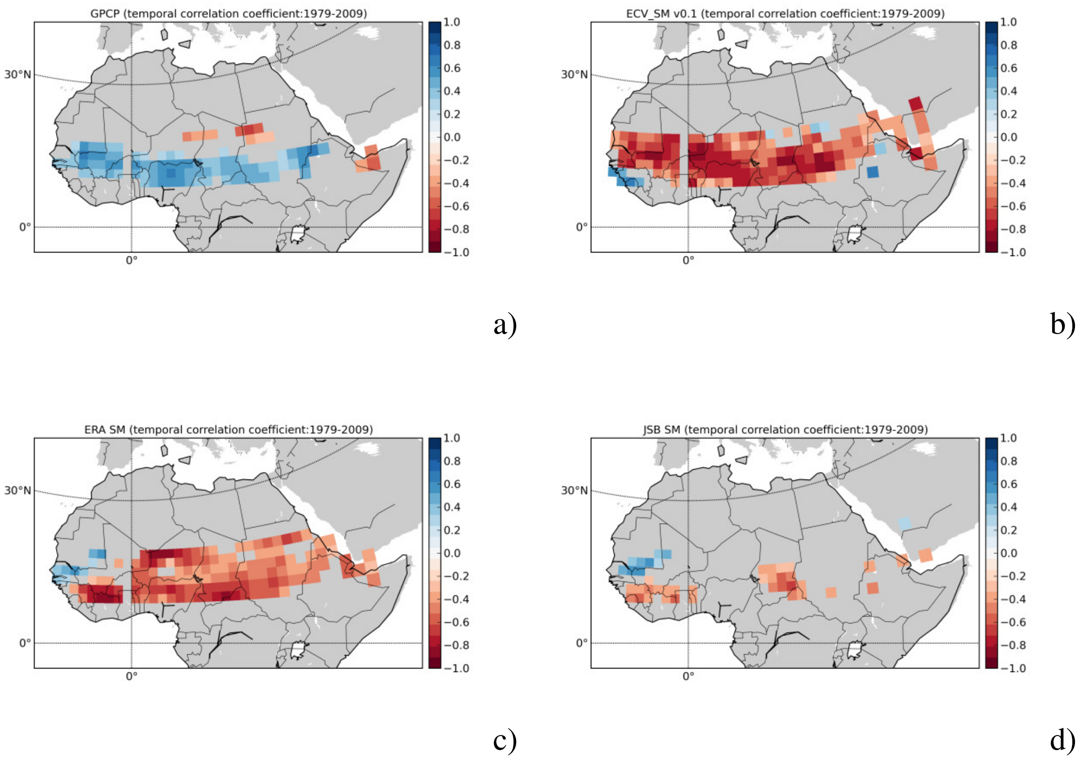

Figure 25: Pearson product moment coefficient for temporal trend of precipitation (a), ECV_SM v0.1 (b), ERA-interim soil moisture (c) and JSBACH soil moisture (d) for the Sahelian rainfall season (June, July, August, September). Only grid cells with significant correlations ($p < 0.1$) are shown.

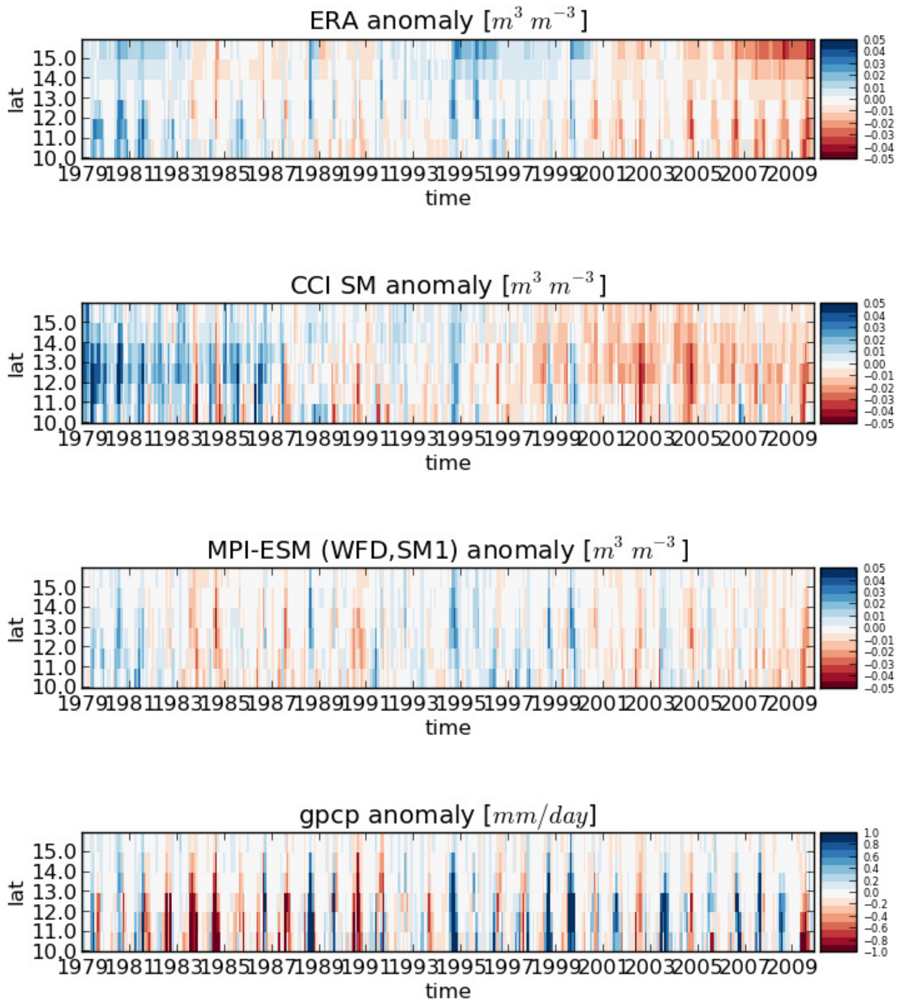

Figure 26: Time-latitude diagrams for monthly anomalies of surface soil moisture and precipitation for ERA-interim, ECV_SM v0.1, JSBACH, GPCP (top-down) in the Sahel region (20W-45E,10N-20N). Anomalies are calculated by removing the mean seasonal climatology derived from the whole timeseries (1979-2009).

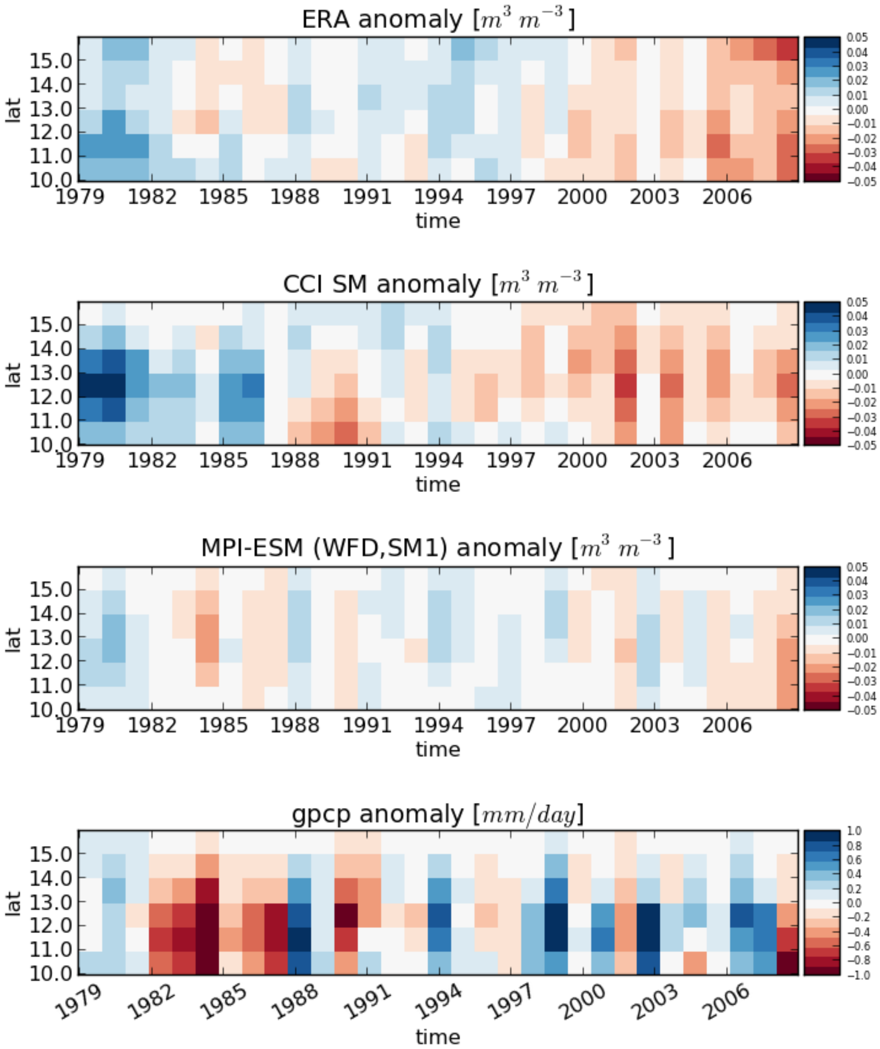

Figure 27: same as Figure 26, but for rainfall season (June,, July, August) only.

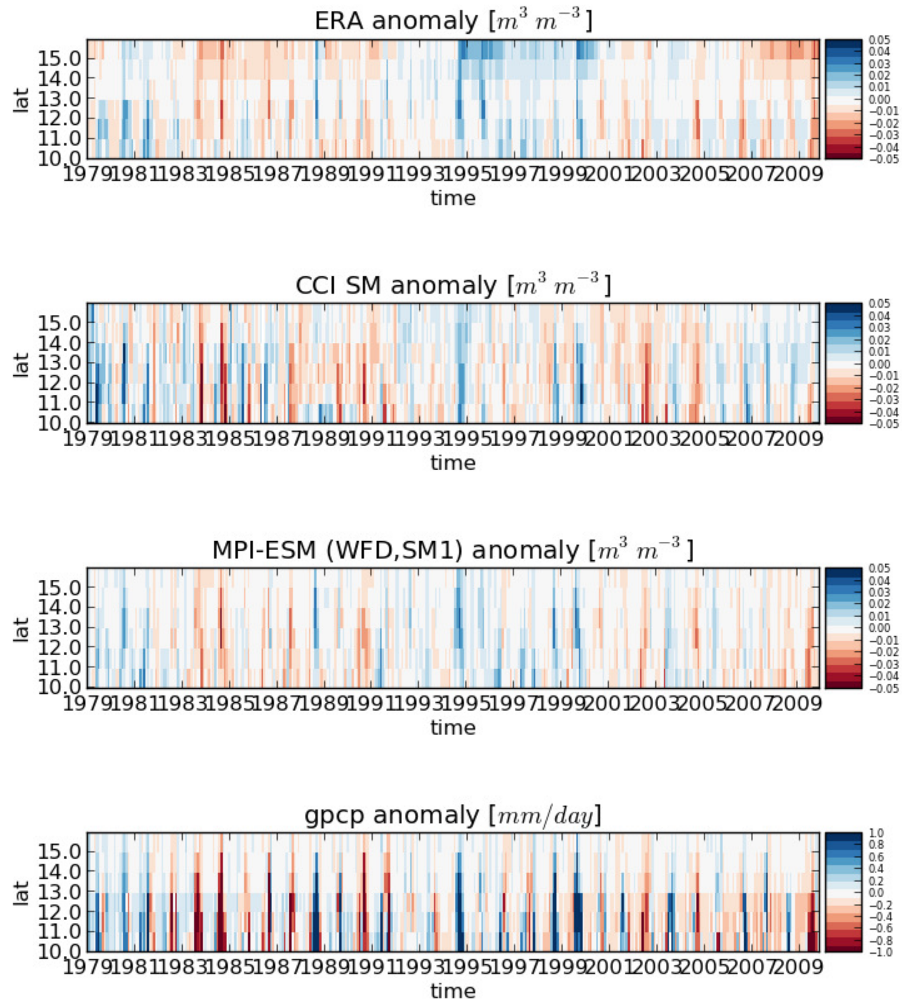

Figure 28: same as Figure 26, but for all datasets, the longterm linear trend was removed.

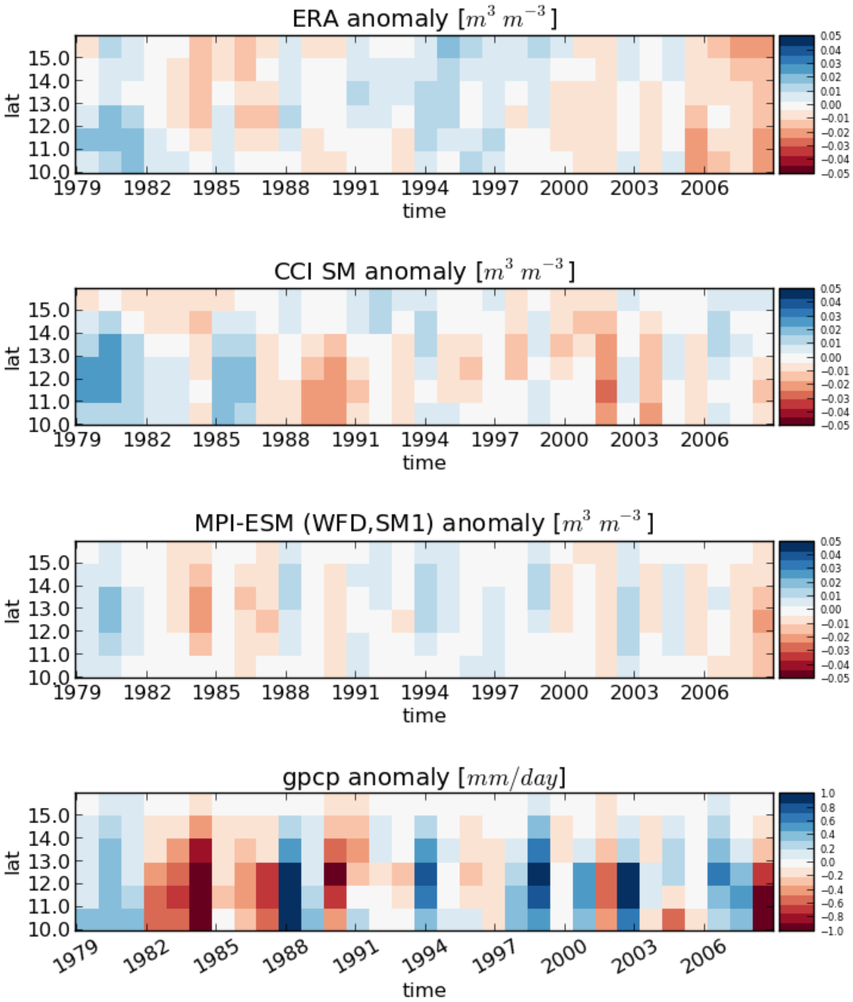

Figure 29: same as Figure 28, but for rainfall season (June,, July, August) only.

- A negative temporal trend in soil moisture is observed while a positive precipitation trend was observed during the same time
- The use of soil moisture data from re-analysis does not ensure a consistent homogeneous timeseries, as changes in the observation system might affect also the soil moisture estimates in reanalysis data.

4.4.2.2 Regional differences of precipitation-soil moisture relationship

Precipitation dynamics in the Sahel differs largely between the western, central and eastern parts. It was therefore analyzed in more detail how the relationship between surface soil moisture and precipitation differs.

A regional mean was calculated for each of the datasets and the regions specified in Table 2 (Figure 25). Figure 30 shows the timeseries of monthly soil moisture and precipitation data, their anomalies and the anomalies where the linear trend has been removed. Correlations between the ECV_SM v0.1 data and the surface soil moisture of ERA-interim and JSBACH are shown. Further, the correlation between the GPCP data (anomalies) and the soil moisture data (anomalies) is shown. One can observe clearly the negative trend in surface soil moisture for ERA-interim and ECV_SM v0.1. In general all datasets show a good significant correlation between each other, either for the original data or the anomalies. The ECV_SM v0.1 soil moisture anomalies show good agreement (r=0.3) with the GPCP precipitation anomalies. Figures for all sub-regions are given in the Annex, but show a similar coherent picture.

The anomaly correlation coefficients between precipitation and soil moisture are summarized in Figure 31. The ECV_SM correlates best with the precipitation anomalies in the western and eastern central part of the Sahel. ERA-interim shows higher correlations in the Western Sahel, while ECV_SM outperforms the ERA-interim data in the central and eastern parts. The maximum anomaly correlation for ERA interim is at the Guinea west coast (W1). JSBACH shows overall the best correlations with GPCP anomalies in all regions with maximum anomaly correlations of r=0.79 (E1).

It has been discussed that the soil moisture and precipitation show temporal trends with reversed sign. These reversed linear trends might mask some of the anomaly correlations between the different datasets. Figure 31 therefore shows also the correlations between GPCP anomalies and soil moisture anomalies for the case where all longterm linear trends have been removed from the data. One can observe an increase of the correlation for all soil moisture datasets while the relative ranking between the datasets and regions remains the same.

Table 2: Regions used for regional analysis

Region	*Bounding coordinates [°]*
W1	20W-10W / 10N-20N
W2	10W-00W / 10N-20N
C1	00W-15E / 10N-20N
C2	15W-25E / 10N-20N
E1	25W-40E / 10N-20N
Sahel	20W-40E / 10N-20N

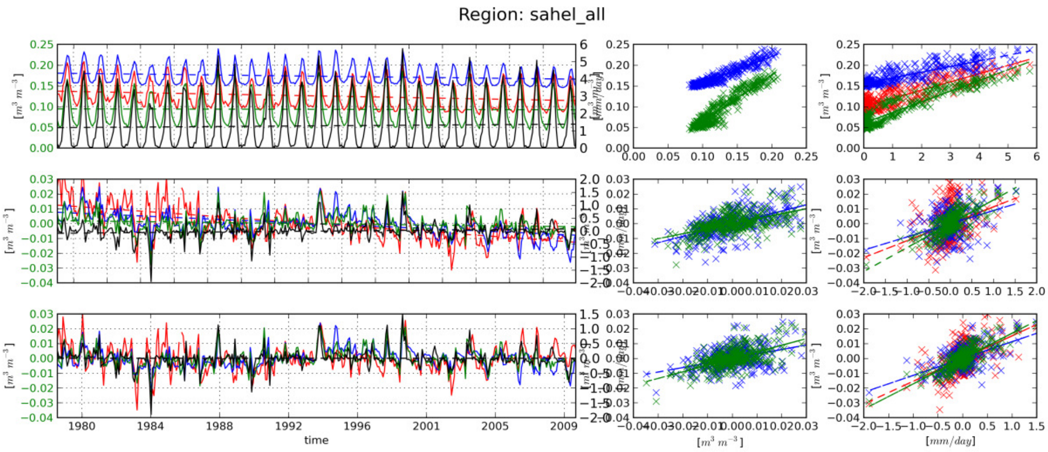

Figure 30: Analysis of surface soil moisture and precipitation for specific regions in the Sahel (Table 2): left: timeseries of soil moisture and precipitation, middle: scatterplot between ECV_SM v0.1 (x-axis) and ERA-interim, JSBACH soil moisture; right: scatterplot between GPCP and the soil moisture datasets; top: original data; middle: anomalies; bottom: detrended anomalies; colors correspond to: JSBACH=green, ERA-interim=blue, ECV_SM v0.1=red, GPCP=black

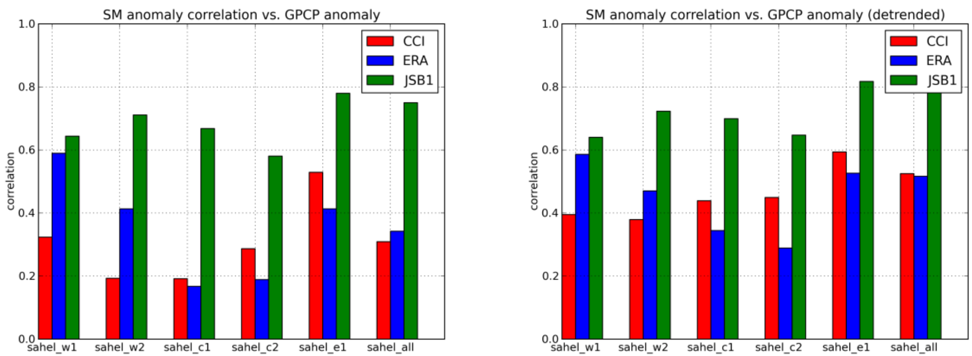

Figure 31: Correlation coefficient between GPCP precipitation anomalies and surface soil moisture data for anomalies (left) and detrended anomalies (right)

- ECV_SM shows good skills in reproducing the Sahelian interannual rainfall anomalies
- JSBACH offline simulations however show significantly higher anomaly correlations with precipitation data
- Surface soil moisture trends in ERA-interim and ECV_SM are contrary to the observed increase in precipitation in the Sahel. Potential interpretation could be an increase in evapotranspiration by the vegetation

5 Conclusions

The objectives of the present report were to assess various aspects of potential applications of multidecadal soil moisture data derived from satellite observations. Different soil moisture and precipitation datasets were used for that purpose. The research focused on the assessment of the potential of the ECV_SM v0.1 dataset for

- Climate model evaluation studies
- Climate model process development and
- Regional climate studies.

It is emphasized that all results relate exclusively to the JSBACH land surface scheme used as part of MPI-ESM. Especially the relationship between surface soil moisture and root zone soil moisture as well as the investigated memory effects are highly dependent on the parameterization of the soil hydraulic model and might therefore differ in other model setups and parameterizations.

The ECV_SM dataset is unique, as it is the first and only existing soil moisture data record for multiple decades. The analysis has shown that the present dataset shows in general good agreement with other soil moisture datasets from modeling studies as well as rainfall data. In particular the **following key advantages of the dataset** have been identified:

- **Assessing soil moisture memory effect**: It has been shown that the dataset is very useful to assess the soil moisture memory effect. Differences between the memory signal in the observed and simulated soil moisture timeseries might be used to improve model specific parameters.
- **Sahelian soil moisture dynamics**: The ECV_SM dataset has been proven to show good agreement with Sahelian soil moisture anomalies which makes it a very useful observational dataset that can be used in combination with other observational records to get a better understanding on interannual to decadal land surface dynamics in the Sahel.

However, the present study has identified several deficits of the ECV_SM dataset for its application in climate models. These are:

- **Limited coverage before 1987**: The date coverage before 1985 is limited to approximately 50% of the days of a year. The reason is the poor spatial and temporal coverage of the input data. Users of the dataset therefore need to take into account this temporal and spatial varying information in their analysis methods using appropriate statistical techniques.
- **Monthly data**: A unique feature of the ECV_SM dataset is that provides information for 30-years. In climate applications, users are often interested in monthly mean values or even seasonal climatologies. It is recommended to deliver a monthly and climatological data product for the CCI SM dataset. This data product should reflect appropriately the effect of different data density on the uncertainties of the monthly/seasonal values.

- **CDF matching**: The ECV_SM statistics is based on a CDF matching approach like was discussed in section 4.3.1. This is considered as a major disadvantage of the dataset, as the soil moisture dynamics is not based on the evidence from the satellite observations any more, but is dependent on the dynamics of the Noah land surface model used as a reference for the CDF matching procedure. The dataset can therefore not be used to evaluate whether a climate model is reproducing realistically the temporal soil moisture dynamics. Such a comparison would only reveal if a model is consistent with the Noah land surface model dynamics and nothing else. It is recommended to develop different data harmonization methods that preserve the soil moisture dynamics of the original observations. This would provide a model independent soil moisture statistics which could be better used for climate model evaluation studies. It is however emphasized, that the dataset might still provide useful insight into the temporal dynamics of soil moisture which could be exploited e.g. by correlation based analysis.
- **Temporal inconsistencies**: The analysis in section 4.3 has also revealed inconsistencies in the timeseries which is likely to be an artifact of the usage of different observation systems. Any kind of trend analysis based on the ECV_SM dataset needs to carefully take into account these potential inhomogeneities. It is recommended to develop appropriate screening techniques to detect abrupt changes in soil moisture timeseries and properly correct and document for these changes in the envisaged CCI SM dataset.
- **Varying autocorrelation lengths**: As shown by Wanders et al. (2012) and Dente et al. (2012), the data from different sensors used to compile the ECV_SM dataset show different autocorrelation lengths. It would be expected that a soil moisture climate data record is best harmonized in a way that differences in autocorrelation structure are a function of changes in the geophysical variable. In a best case, the user could use the uncertainty information in the data product to weight the different samples appropriately to estimate the autocorrelation function in a consistent manner.
- **High latitude problem**: comparisons between models and observations resulted in reasonable agreement for large parts of the globe. In Northern Latitudes however, even negative correlation between soil moisture and model output was observed. This high latitude problem is likely to be caused by inconsistent information on snow cover and frozen ground between model and observations. It is unclear, how reliable the soil moisture estimates are in these high latitude areas.
- **Data gaps due to blending**: The data blending procedure applied for ECV_SM is not optimal, as it results in the lack of observations in some areas. The European heat wave 2003 is such an example. It is suggested to develop other blending techniques for the ESA CCI soil moisture data product.

- **Information or noise?** The present study has shown that the ECV_SM soil moisture data shows good agreement with observed precipitation anomalies. It has however also be shown that the MPI-M land surface model JSBACH, forced with WATCH forcing data shows even better correlation and anomaly correlation between soil moisture and precipitation. It is evident that the model soil moisture is related to its precipitation forcing. The smaller agreement of the ECV_SM with the observed precipitation dynamics could be interpreted in two different ways:
    - The observed relationship between ECV_SM and precipitation is really weaker. This would mean that the noise in the soil moisture observations is reducing the correlation.
    - The difference and weaker correlation is due to the effect of other relevant processes. Soil moisture dynamics is not solely influenced by precipitation, but also by vegetation dynamics and evapotranspiration. In that case, the differences between soil moisture dynamics and precipitation dynamics might reveal additional insight into land surface dynamics which could be used to improve the process understanding and description in climate model land surface schemes. A further analysis is beyond the scope of the present study but will be subject of further investigations.

Overall, the ECV_SM dataset provides a first unique dataset that provides relevant information for climate studies. The possibility to retrieve relevant patterns information on the memory characteristics of soil moisture is important for further applications like climate model initialization and improvement of seasonal to multiannual climate forecasting applications.

6 References

Adler, R.F. et al., 2003. The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). Journal of Hydrometeorology, 4(6), pp.1147-1167.

Crow, W.T. & Zhan, X., 2007. Continental-Scale Evaluation of Remotely Sensed Soil Moisture Products. IEEE Geoscience and Remote Sensing Letters, 4(3), pp.451-455.

Dee, D.P. et al., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), pp.553-597.

Dente, L. Z. Verkerdy, R. de Jeu & Z. Su (2012) Seasonality and autocorrelation of satellite derived soil moisture products, International Journal of Remote Sensing, DOI: 10.1080/01431161.2012.716923

Dorigo, W. A., R. A.M. de Jeu, D. Chung, R. M. Parinussa, Y. Y. Liu, W. Wagner, and D. Fernandez-Prieto (2012), Evaluating global trends (1988-2010) in harmonized multi-satellite surface soil moisture, Geophys. Res. Lett., VOL. 39, L18405, 7 PP., 2012

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Drusch, M., et al. (2011). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. Hydrology and Earth System Sciences, 15, 1675–1698, doi:10.5194/hess-15-1675-2011.

Fensholt, R. et al., 2012. Greenness in semi-arid areas across the globe 1981-2007 - an Earth Observing Satellite based analysis of trends and drivers. Remote Sensing of Environment, 121, pp.144-158.

Fensholt, R. et al., 2012. Greenness in semi-arid areas across the globe 1981-2007 - an Earth Observing Satellite based analysis of trends and drivers. Remote Sensing of Environment, 121, pp.144-158.

Ferranti, L. & Viterbo, P., 2006. The European Summer of 2003: Sensitivity to Soil Water Initial Conditions. Journal of Climate, 19(15), pp.3659-3680.

Fischer, E.M. et al., 2007. Soil Moisture–Atmosphere Interactions during the 2003 European Summer Heat Wave. Journal of Climate, 20(20), pp.5081-5099.

Giorgetta, M. A., Jungclaus, J. H., Reick, C. H., Legutke, S., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W. A., Notz, D., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K., Stockhause, M., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., & Stevens, B. (2012). Climate change from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project 5. Journal of Advances in Modeling Earth Systems.submitted.

Gouweleeuw, B.T. et al., 2012. Space-based passive microwave soil moisture retrievals and the correction for a dynamic open water fraction. Hydrology and Earth System Sciences, 16(6), pp.1635-1645.

Hagemann, S. An improved land surface parameter dataset for global and regional climate models Max Planck Institute for Meteorology, MPI-Report 336, 2002

Hickler, T. et al., 2005. Precipitation controls Sahel greening trend. Geophysical Research Letters, 32(21).

Liu, Q. et al., 2011a. The Contributions of Precipitation and Soil Moisture Observations to the Skill of Soil Moisture Estimates in a Land Data Assimilation System. Journal of Hydrometeorology, 12(5), pp.750-765.

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., Evans, J. P. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrology and Earth System Sciences, 15, 425-436, doi:10.5194/hess-15-425-2011

Liu, Y.Y., Dorigo, W.A., Parinussa, R.M., de Jeu, R.A.M. , Wagner, W., McCabe, M.F., Evans, J.P., van Dijk, A.I.J.M. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sensing of Environment, 123, 280-297, doi: 10.1016/j.rse.2012.03.014.

Loew, A. & Govaerts, Y., 2010. Towards Multidecadal Consistent Meteosat Surface Albedo Time Series. Remote Sensing, 2(4), pp.957-967.

Loew, A. (2012): Terrestrial satellite climate data records for climate studies: How long is long enough? – a testcase for the Sahel. Theor. Appl. Clim. submitted.

Loew, A., Holmes, T. & de Jeu, R., 2009. The European heat wave 2003: Early indicators from multisensoral microwave remote sensing? Journal of Geophysical Research, 114(D5), pp.1-14.

Luterbacher, J. et al., 2004. European seasonal and annual temperature variability, trends, and extremes since 1500. Science (New York, N.Y.), 303(5663), pp.1499-503.

Meesters, A. & Jeu, R.D., 2005. Analytical derivation of the vegetation optical depth from the microwave polarization difference index. Geoscience and Remote, 2(2), pp.121-123.

Miralles, D.G. et al., 2011. (Liu et al.)lobal land-surface evaporation estimated from satellite-based observations. Hydrology and Earth System Sciences, 15(2), pp.453-469.

Olsson, L., Eklundh, L. & Ardö, J., 2005. A recent greening of the Sahel—trends, patterns and potential causes. Journal of Arid Environments, 63(3), pp.556-566.

Olsson, L., Eklundh, L. & Ardö, J., 2005. A recent greening of the Sahel—trends, patterns and potential causes. Journal of Arid Environments, 63(3), pp.556-566. Huber, S., Fensholt, R. & Rasmussen, K., 2011. Water availability as the driver of vegetation dynamics in the African Sahel from 1982 to 2007. Global and Planetary Change, 76(3-4), pp.186-195.

Owe, M., de Jeu, R. & Holmes, T., 2008. Multisensor historical climatology of satellite-derived global land surface moisture. Journal of Geophysical Research, 113(F1), p.F01002.

Parinussa R., A.G.C.A. Meesters, Y. Liu, W. Dorigo, W. Wagner & R.A.M. de Jeu An Analytical Solution to Estimate the Error Structure of a Global Soil Moisture Dataset, *IEEE Geoscience and remote sensing letters*, 8, 779-783, doi: 10.1109/LGRS.2011.2114872

Rebel, K.T. et al., 2012. A global analysis of soil moisture derived from satellite observations and a land surface model. Hydrology and Earth System Sciences, 16(3), pp.833-847.

Reick, C., Raddatz, T., Brovkin, V., & Gayler, V. (2012). The representation of natural and anthropogenic land cover change in MPI-ESM. Journal of Advances in Modeling Earth Systems. Submitted

Schär, C. & Jendritzky, G., 2004. Climate change: hot news from summer 2003. Nature, 432(7017), pp.559-60.

Schär, C. et al., 2004. The role of increasing temperature variability in European summer heatwaves. Nature, 427(6972), pp.332-6.

Taylor, K.E., Stouffer, R.J. & Meehl, G. a., 2012. An Overview of CMIP5 and the Experiment Design. Bulletin of the American Meteorological Society, 93(4), pp.485-498.

Tucker, C.J. et al., 2005. An extended AVHRR 8-km NDVI data set compatible with MODIS and SPOT vegetation NDVI data. International Journal of Remote Sensing, 26(20), pp.4485-4498.

Wanders, N. et al., 2012. Observation uncertainty of satellite soil moisture products determined with physically-based modeling. Remote Sensing of Environment, 127(October 2011), pp.341-356.

Weedon, G. P.; Gomes, S.; Viterbo, P.; Shuttleworth, W. J.; Blyth, E.; Österle, H.; Adam, J. C.; Bellouin, N.; Boucher, O. & Best, M. Creation of the WATCH Forcing Data and its use to assess global and regional reference crop evaporation over land during the twentieth century J. Hydrometeorol., 2011, 12, 823—848.

Zaitchik, B. F., A. K. Macaldy, L. R. Bonneau, and R. B. Smith (2006), Europe’s 2003 heat wave:A satellite view of impacts and land-atmosphere feedbacks, Int. J. Climatol., 26, 743–769

7 ANNEX

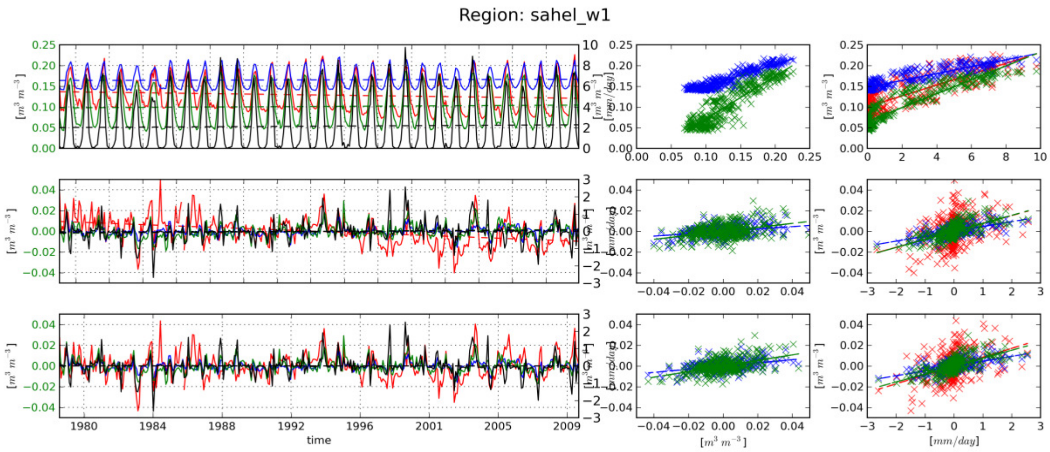

Figure 32: same as Figure 30, but for region W1

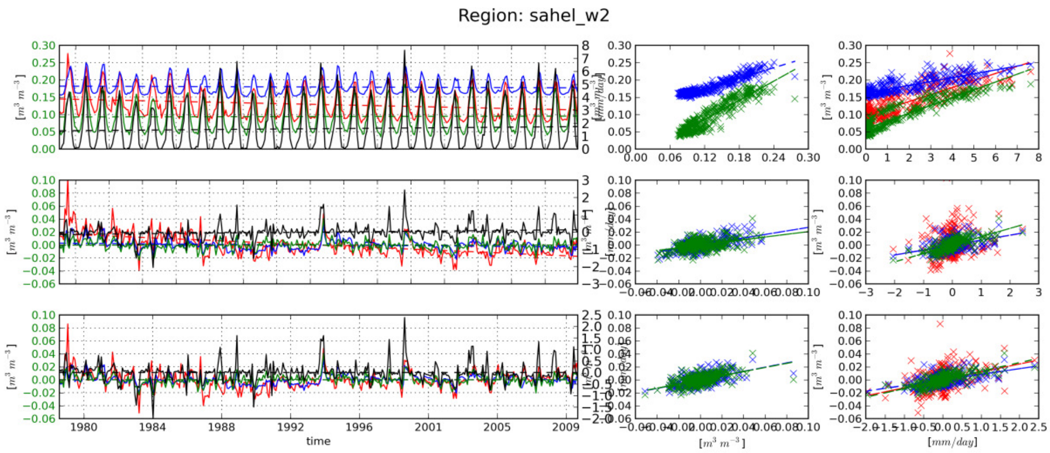

Figure 33: same as Figure 30, but for region W2

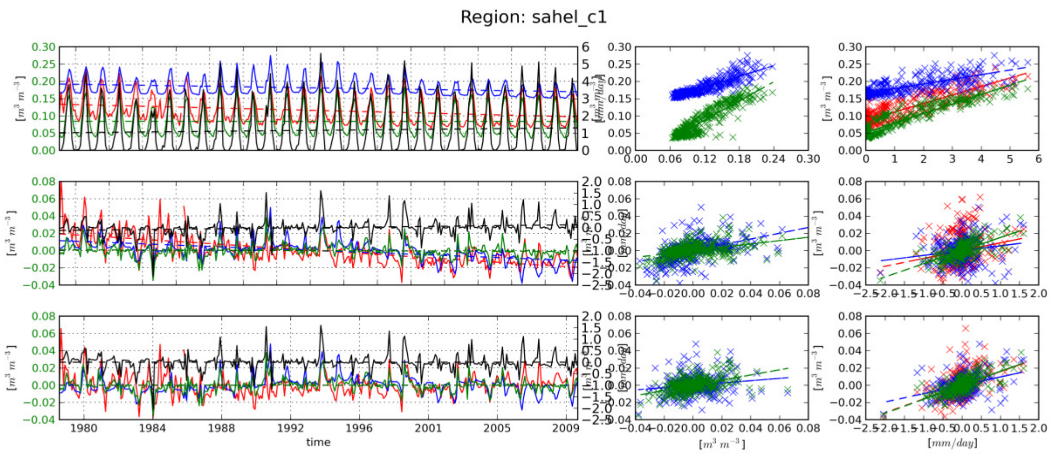

Figure 34: same as Figure 30, but for region C1

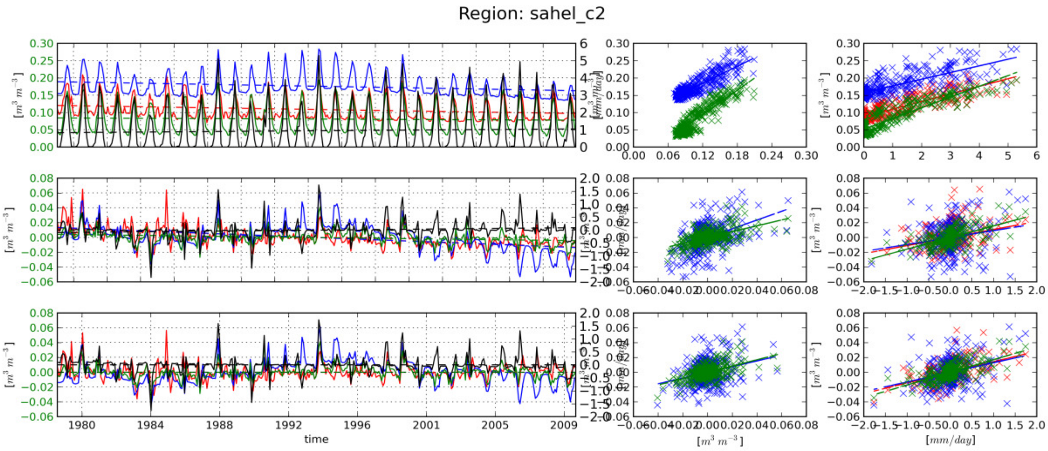

Figure 35: same as Figure 30, but for region C2

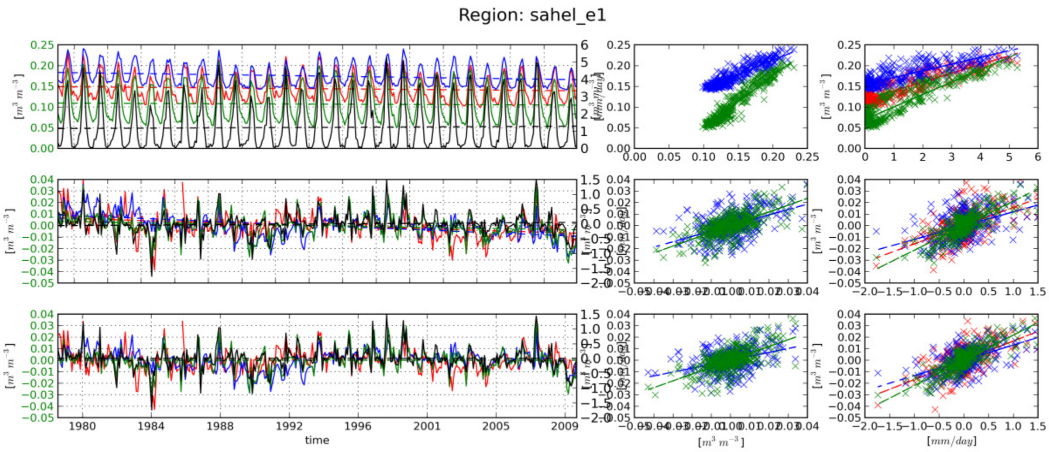

Figure 36: same as Figure 30, but for region E1

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, Hydrology and Earth System Sciences, 15, 425-436, 10.5194/hess-15-425-2011, 2011.

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P., and van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sensing of Environment, 123, 280-297, 10.1016/j.rse.2012.03.014, 2012.

Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, Journal of Geophysical Research, 113, F01002, 10.1029/2007JF000769, 2008.