Remote Sensing of Environment 174 (2016) 69–81

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Comparing surface-soil moisture from the SMOS mission and the ORCHIDEE land-surface model over the Iberian Peninsula

Jan Polcher [a,b,d,*], Maria Piles [c], Emiliano Gelati [a,d,e], Anaïs Barella-Ortiz [a,b], Marivi Tello [d,f]

[a] Centre National de la Recherche Scientifique (CNRS), France
[b] Laboratoire de Météorologique Dynamique du CNRS, UMR8539, CNRS, IPSL, France
[c] Dept. de Teor. del Senyal i Comunicacions, Univ. Politec. de Catalunya, Barcelona, Spain
[d] Institut Català de Ciències del Clima (IC3), Barcelona, Spain
[e] CNRM-GAME, UMR 3589 (Météo-France, CNRS), Toulouse, France
[f] Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Germany

ARTICLE INFO

Article history:
Received 4 March 2015
Received in revised form 23 November 2015
Accepted 3 December 2015
Available online 17 December 2015

Keywords:
Soil moisture
Remote sensing
SMOS
Land-surface modelling
ORCHIDEE

ABSTRACT

The aim of this study is to compare the surface soil moisture (SSM) retrieved from ESA's Soil Moisture and Ocean Salinity mission (SMOS) with the output of the ORCHIDEE (ORganising Carbon and Hydrology In Dynamic EcosystEm) land surface model forced with two distinct atmospheric data sets for the period 2010 to 2012. The comparison methodology is first established over the REMEDHUS (Red de Estaciones de MEDición de la Humedad del Suelo) soil moisture measurement network, a 30 by 40 km catchment located in the central part of the Duero basin, then extended to the whole Iberian Peninsula (IP). The temporal correlation between the in-situ, remotely sensed and modelled SSM are satisfactory ($r > 0.8$). The correlation between remotely sensed and modelled SSM also holds when computed over the IP. Still, by using spectral analysis techniques, important disagreements in the effective inertia of the corresponding moisture reservoir are found. This is reflected in the spatial correlation over the IP between SMOS and ORCHIDEE SSM estimates, which is poor ($\rho \sim 0.3$). A single value decomposition (SVD) analysis of rainfall and SSM shows that the co-varying patterns of these variables are in reasonable agreement between both products. Moreover the first three SVD soil moisture patterns explain over 80% of the SSM variance simulated by the model while the explained fraction is only 52% of the remotely sensed values. These results suggest that the rainfall-driven soil moisture variability may not account for the poor spatial correlation between SMOS and ORCHIDEE products.

1. Introduction

Moisture availability at the surface is a major constraint on the coupling between continental surface and the atmosphere at different time scales. As demonstrated for tropical regions, moisture availability can determine small scale gradients and trigger convective systems and thus atmospheric processes at short time scales (Taylor et al., 2011). For time scales up to seasonal, there is also evidence that the soil moisture memory might contribute in some regions to the predictability of climate (Koster et al., 2011). Although most of these results on the role of soil moisture in the surface atmosphere interactions are model based, and thus have to be taken with some caution (Taylor, de Jeu, Guichard, Harris, & Dorigo, 2012), it is anticipated that observing the moisture available in the soils on an operational basis will be beneficial for estimating evaporation and thus controlling the partition of energy at the surface.

The analysis of regional or even global soil moisture estimates, or more generally of land surface states, is a challenge because of the high spatial heterogeneity of surface properties and the fast time constants characterizing some surface variables. A common approach has been to run a set of land surface models under controlled conditions and use the ensemble average as best estimate (Dirmeyer, Dolman, & Sato, 1999). A related approach is to control the evolution of one model with the assimilation of all available local observations (Rodell et al., 2004) integrating the spatial variability not represented at the resolution of the model as a representativeness error of the in-situ data.

More recently, technological advancements have made possible the launch of satellite instruments designed to observe critical land surface state variables. This is the case of the ESA's Soil Moisture and Ocean Salinity (SMOS) mission, in orbit since November 2009, which provides estimates of salinity over oceans and of surface soil moisture over continental surfaces (Kerr et al., 2001). Since SMOS is mainly sensitive to near surface soil moisture (top 5 cm), it is advantageous to assimilate

* Corresponding author at: Laboratoire de Météorologie Dynamique du CNRS, Institut Pierre Simon Laplace, Ecole Polytechnique, F 91128 Palaiseau Cedex, France.
E-mail address: jan.polcher@lmd.jussieu.fr (J. Polcher).

http://dx.doi.org/10.1016/j.rse.2015.12.004

the satellite data in land surface models and thus derive more important estimates for deeper surface state variables such as root zone soil moisture, a direct indicator of the water available to plants (Draper, Reichle, De Lannoy, & Liu, 2012; Muñoz-Sabater, Jarlan, Calvet, Bouyssel, & De Rosnay, 2007). This information is very relevant to improve the characterization of surface/atmosphere interactions. The large scale nature of SMOS soil moisture estimates should ensure that they are more representative of the scales simulated by land surface models and thus perform better than in-situ observations in the assimilation process.

A first step towards assimilating satellite observed variables is to compare them with model output and ensure that both have compatible physical definitions. This has been done recently for SMOS soil moisture products, with in-situ observations and the ECMWF land surface model (Albergel et al., 2012) or the Météo-France model (Parrens et al., 2012). Nevertheless in both cases the focus of the comparison was on the temporal evolution of SSM with some spatial diagnostics presented by Parrens et al. (2012). In the present study we perform this comparison with in-situ observations and the ORCHIDEE (ORganising Carbon and Hydrology In Dynamic EcosystEm) land surface model, and extend the analysis to focus on the spatial structures that are resolved by the remotely sensed and the modelled SSM. To our knowledge it has not yet been verified that the time varying spatial structures captured by SMOS are those driven by the meteorological scales and can be compared to those simulated by land surface models.

The proposed analysis focuses on the Iberian Peninsula (IP) because it is a semi-arid region where many human activities are governed by water availability. Hence, developing tools for monitoring and predicting SSM in this region is of critical importance. Furthermore, the region is not densely vegetated and soil moisture displays a large annual cycle, thus providing an ideal setting for SMOS satellite observations. As spring time soil moisture anomalies over the IP are believed to be a pre-cursor to droughts and heat waves in Europe (Vautard et al., 2007; Zampieri et al., 2009), monitoring, assimilating and modelling surface states in this region could be key for improvements in seasonal forecasting systems.

Due to the high temporal variability of the upper 5 cm soil moisture measured by SMOS, the sampling of observations is a critical issue (De Rosnay, 2003). Thus this study starts with a presentation of the data sets and a quantification of possible biases introduced by the sampling. After these preliminaries, we compare SMOS and ORCHIDEE SSM over the REMEDHUS (Red de Estaciones de MEDición de la Humedad def Suelo) basin, where in-situ observations are available (Martínez-Fernández & Ceballos, 2005). The analysis of the temporal variability is then extended to the entire IP for the SMOS and ORCHIDEE data sets, where the spatial correlation of the surface soil moisture structures is also studied. Afterwards, we discuss the relations between the temporal and spatial variability of rainfall and SSM as retrieved by SMOS and simulated by ORCHIDEE: first the relation between the spectral characteristics of precipitation and SSM is analysed; then the coupled rainfall-SSM co-varying patterns are identified and studied by means of the singular value decomposition (SVD) method. Finally, we discuss the main findings of this study and present its general conclusions.

2. Data and methodology

2.1. The SMOS surface soil moisture product

The ESA SMOS mission was launched on November 2, 2009. Over land, the SMOS mission aims at providing surface soil moisture (originally advertised to be the top 5 cm) data at a spatial resolution better than 50 km, with a target accuracy of 0.04 m^3/m^3 and a repeat cycle of less than 3 days. The payload is a novel 2-D synthetic aperture radiometer equipped with 69 L-band receivers regularly distributed over a Y-shaped antenna array. This new design allows observing all pixels within a 1000 km wide field of view at a range of incidence angles (from 0 to 65°) and polarizations (horizontal, vertical, and mixed).

The SMOS soil moisture retrieval algorithm consists in matching measured and modelled (surface emission) brightness temperatures. The modelled values are functions of polarization, incidence angle, soil moisture (through the soil dielectric constant) and other physical parameters influencing land emissivity at L-band, namely land surface temperature, soil roughness, vegetation optical depth and volume scattering albedo. Information on land-cover and soil properties (e.g. clay fraction, bulk density, field capacity and wilting point) is also used in the forward modelling. The inversion (retrieval of geophysical variables) is performed by minimizing a cost function that accounts for the weighted squared differences between modelled and measured data (Kerr et al., 2012).

The SMOS BEC L3 v.001 soil moisture data from January 2010 to December 2012 over the IP has been selected for this research. It is obtained by quality-filtering and re-gridding of the newest version of L2 data (v.551); grid points affected by radio frequency interferences (RFI) and/or with soil moisture Data Quality Index (DQX) greater than 0.07 are discarded; and a DQX-inverse weighted average is applied to bin the data from ISEA to the 25 km EASE equal-area grid. Data and further details are available at http://cp34-bec.cmima.csic.es/. A comprehensive validation of SMOS-derived products over the IP can be found in Gonzalez-Zamora et al. (2015).

2.2. The ORCHIDEE land surface model

ORCHIDEE, the land surface model of the Institut Pierre Simon Laplace (De Rosnay & Polcher, 1998; Krinner et al., 2005), is ideally suited for analysing remotely sensed soil moisture products because of its high vertical resolution in the representation of soil moisture (De Rosnay, Bruen, & Polcher, 2000; De Rosnay, 2002; D' Orgeval, Polcher, & de Rosnay, 2008). The main motivation for developing the fine vertical resolution was to improve the interaction of the root profile with the soil moisture distribution at different depths and refine the representation of infiltration processes. With the same motivation, a sub-grid variability of soil moisture was included in the model. In each grid box three soil moisture profiles are simulated. They share the same soil texture and structure, taken from the Zobler map (Post & Zobler, 2000), but differ in their vegetation distribution. The 13 plant functional types represented by ORCHIDEE are distributed over the 3 columns: bare soil, low vegetation and high vegetation. As each is characterized by different root profiles, the soil moisture in each column will develop its own profile. It was found that this method provides the best simulation of soil moisture stress for evaporation (D' Orgeval et al., 2008). In the analysis presented here only the grid box average soil moisture profile will be considered.

From the eleven soil moisture layers simulated by ORCHIDEE, only the first 5 layers spanning the top 4.5 cm of soil is analysed here. The sum of moisture in these top layers is the closest to the SMOS theoretical penetration depth of 5 cm (Kerr et al., 2010) and is thus defined as the simulated SSM. Analyses not detailed here have shown that deeper layers are not relevant for the comparison with SMOS observations as their typical time scales are much longer than those obtained from the satellite.

ORCHIDEE is forced with two different atmospheric datasets in order to frame the uncertainty introduced by this boundary condition and test the possibility of using remote sensing observations to evaluate the quality of these simulations. The first forcing is the one extracted from the ERA-Interim (ERA-I) re-analysis (Balsamo et al., 2012; Dee et al., 2011). The corresponding simulation will be labelled OR-ERA. This type of atmospheric forcing is known to have biases that negatively affect land-surface simulations and can be corrected (Ngo-Duc, Polcher, & Laval, 2005; Weedon et al., 2011). Consequently the bias corrected product WATCH Forcing Data ERA-Interim (WFDEI) is also used

(Weedon et al., 2014). The simulation which uses this forcing will be designated OR-WFD.

The two forcing datasets share the same synoptic variability for all variables. The bias correction process affects precipitation intensity, as monthly mean precipitation and the number of wet days have been changed in order to correspond to the values given in the Climate Research Unit (CRU) compilation of observations (Harris, Jones, Osborn, & Lister, 2014). The ERA-I data are available at the horizontal resolution of the atmospheric model used in the re-analysis, i.e. 0.75° by 0.75°. On the other hand, for WFDEI the re-analysis has been extrapolated to the 0.5° by 0.5° resolution of the CRU data sets used in the bias-correction process. Thus soil moisture was simulated by ORCHIDEE at these two spatial resolutions. The overlap period between these two forcing data sets and SMOS available data is January 2010 to December 2012 and will thus be the focus for this study. Since both ERA-I and WFDEI forced simulations were started in January 1979, the analysis period does not suffer from any spin-up artefacts.

To facilitate the comparison of soil moisture daily means, the two ORCHIDEE simulations were remapped to their nearest neighbour on the 25 km EASA equal-area grid on which the SMOS BEC L3 product is provided. This preserves the spatial structures provided by the model, which is run at the coarser resolution of the atmospheric forcing. Having all surface soil moisture estimates on the same 25 km resolution grid allows to create a thinned version of the ORCHIDEE data in which all values which do not have corresponding SMOS observations are deleted. These two versions of modelled SSM are used for analysing the impact of sampling.

2.3. Sampling error analysis

The 36 months record of SMOS daily averaged SSM on the 25 km grid, not all pixels are observed daily. The month with most SSM values is November 2012 (73% of possible values are observed) and the month with the least observations is September 2012 (40%). As expected, these numbers are lower during the satellite's commissioning phase (January to June 2010). SMOS swath allows for a 3-day revisit period on the morning or evening orbit, so the IP will not be fully covered by the satellite on a daily basis. Additionally, there is an important amount of missing data due to the presence of radio frequency interferences (RFIs) (Oliva et al., 2012) overshadowing the measurements. Also, no estimates are provided when the retrieval algorithm is unable to estimate upper soil moisture within acceptable uncertainty bounds. Still, with some exceptions, the IP is well observed at the daily scale.

For a number of diagnostics presented in this study, it is helpful to average the surface soil moisture maps of SMOS over an appropriate time window to reduce measurement noise and provide more complete spatial maps. But the bias introduced by the time averaging of an incomplete sample relative to the true mean needs to be quantified. In order to evaluate this sampling error, we compare the mean SSM computed with all values simulated by ORCHIDEE with the estimated average obtained using a thinned version in which modelled values are only retained if they correspond to SMOS observations. This diagnostic is performed on the SMOS grid (i.e. a 25 km resolution). For temporal averaging windows ranging from 2 to 15 days, the true average is compared to those obtained with the thinned data using a root mean square error (RMSE) and a spatial correlation (labelled ρ here). These are the metrics most relevant for the diagnostics performed in this study. As the impact of SMOS sampling is expected to change with the season, as the frequency and intensity of rainfall events as well as the evaporative demand exhibits annual cycles, the diagnostic was performed separately for each month.

In all months, SSM exhibits a sampling error (RMSE and spatial correlation) that monotonically deteriorates from the 2 to the 5-day averaging window length and reaches a plateau for higher values. The largest errors are obtained for the month of September 2010 (where 57% of all possible points have observations), April and September 2012 (where 40 and 63%, respectively, of all possible points are observed). In these cases the RMSE reaches 10% of the true mean and the spatial correlation drops to 0.92. During the 30 months period of SMOS observations falling outside of the commissioning phase, it was not possible to determine if the annual cycle of SSM or the number of observations are responsible for the variations in the sampling error.

Given the results of the above analysis, we can assume that the SMOS sampling of SSM produces RMSE scores below 10% and spatial correlation values above 0.92 for averages up to a monthly scale. Larger discrepancies between remotely sensed and modelled soil moisture cannot be attributed to the irregular nature of satellite observations over the IP. We try to avoid this type of errors by retaining only ORCHIDEE values that correspond to valid SMOS retrievals, i.e. using a re-sampled version of the model output.

2.4. The REMEDHUS station data

To benchmark the agreement between SMOS and ORCHIDEE SSM, we select a small area in the central part of the Duero basin where a network of 19 permanent soil moisture stations is operated and serves as a SMOS validation site. This area has a typical semi-arid Mediterranean climate and its soil texture is characterized by a large sand fraction. The stations are distributed over the catchment so as to be representative of the main physiographic and pedological units (Martínez-Fernández & Ceballos, 2005; Sánchez, Martinez-Fernandez, González-Piqueras, et al., 2012; Sánchez, Martinez-Fernandez, Scaini, et al., 2012). For this analysis we select SMOS and ORCHIDEE data on the pixel which includes the REMEDHUS basin (see KML file in the Supplementary Data).

The variability across the 19 stations within the basin provides a glimpse of the small-scale spatial variability of SSM. Before computing the basin average and variance, we normalize the SSM time series at each station. In the original data the largest inter-station variability of soil moisture is observed in winter, when SSM reaches field capacity. Such variability depends on the texture and structure of the soils in which the probes are placed and is not relevant for the comparison with SMOS or ORCHIDEE, because their SSM estimates are based on slightly different assumptions about the local soil characteristics. For instance, the sand fraction is observed to be 71% in the RHEMEDUS basin (Martínez-Fernández & Ceballos, 2005), in the SMOS L2 retrieval algorithm 40.39% of sand is assumed and ORCHIDEE uses 50% of sand based on the Zobler map (Post & Zobler, 2000). To eliminate this variability each time series is linearly rescaled to the interval [0, 1] with respect to its minimum and maximum, which are used as proxies for the field capacity and wilting point. However we are aware that this normalization is likely to underestimate the maximum inter-station variability in normalized soil moisture, which generally occurs in summer. During this time of the year, because of the intra-basin water redistribution, upstream top soils are often drier than downstream ones where the wilting point may not be reached. Thus the inexact assumption about the wilting point at the downstream stations may cause an underestimation of summer inter-station variability.

2.5. Temporal variability

To study the temporal variability of the SMOS SSM, we need to take into account that the average observation frequency over the IP is 1.5 days. But there is quite a large variability in this value over the 3 years record. For the reasons previously discussed, RFIs, perturbation in the satellite operations and failure of the retrieval algorithm to converge will tend to reduce the frequency of observations.

To estimate the periodogram for irregularly distributed data, we use the Lomb–Scargle method (Press & Rybicki, 1989) and we have to determine the range of frequencies in which it can reliably be computed. Given the average frequency of observations and the record length, the spectral power could potentially be estimated from periods of

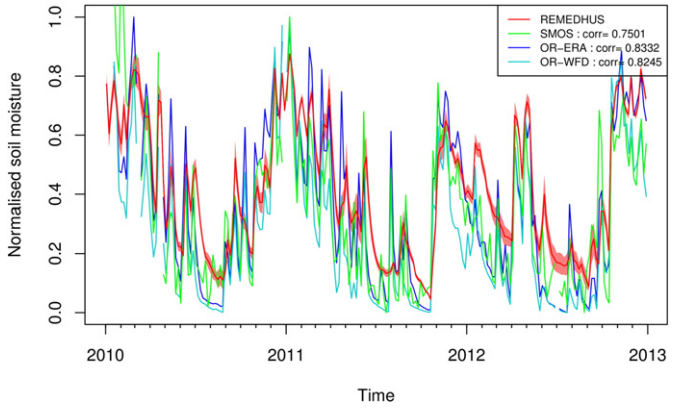

Fig. 1. Evolution of surface soil moisture from January 2010 to December 2012: observed at the REMEDHUS site (red), retrieved by SMOS (green) and simulated by ORCHIDEE (blue and cyan). The red shaded area represents the standard deviation of the 19 REMEDHUS stations. A 5 day averaging was performed before plotting. Temporal correlations with the average REMEDHUS soil moisture are reported in the legend. For placement of the stations and the grid boxes for SMOS and ORCHIDEE see the web version of this article.

3 days (the Nyquist frequency) to 1.5 years. The uneven sampling puts a limit on the estimations at the highest frequencies, as can be estimated theoretically for a Gaussian white noise process (Horne & Baliunas, 1986). As SSM is significantly auto-correlated (Entin et al., 2000), the Lomb-Scargle method will tend to overestimate the power at high frequencies. The upper limit of the frequencies which can be retrieved from the SMOS observations is determined by comparing the results of the spectral analysis of the full time series of the simulated SSM with the thinned version on a number of grid points. This shows that for periods shorter than 6 days the observations are too irregular to obtain reliable estimates of the spectral power. Thus the spectral analysis will be limited to periods between 6 and 365 days.

Another method to analyse the characteristic time scales of SSM estimates consists in decomposing the signal into a mean annual cycle and an anomaly. The first step is to compute a mean annual signal based on the 3 years of data. As this is a rather short period, the mean annual cycle is smoothed with a spline filter in order to remove sub-monthly fluctuations. The anomalies, which represent the synoptic variability, are obtained by subtracting the smoothed annual cycle from the original time series. After applying this procedure to all points of the study region, we can compare the spatial structures of the annual and of the anomaly components between the remotely sensed and the simulated SSM.

3. Results

3.1. Benchmarking SMOS and ORCHIDEE at the REMEDHUS site

Fig. 1 shows the time evolution of SMOS, ORCHIDEE and the spatially averaged REMEDHUS surface soil moisture (SSM). The drying of the surface layers during the spring and the autumn wetting are well represented. The annual cycle of the simulated SSM is very similar in phase to those observed in-situ and retrieved from SMOS.

Superposed on the annual cycle, the soil wetting events associated to rainfall episodes are well detected by the satellite and synchronous with those measured in-situ and simulated by ORCHIDEE for the top 5 cm of soil. The simulated soil moisture peaks are very similar to those observed by SMOS but generally narrower than those measured in-situ. Particularly in spring and summer, the in-situ SSM exhibits a slower drying than the one found in the SMOS retrievals and ORCHIDEE. The remotely sensed and modelled values are smaller and well outside the range of variability estimated from the 19 stations in summer. This may be caused by the normalization procedure which does not use the true wilting point for each station and thus may underestimate the spatial variability as indicated above.

The normalization of the soil moisture is problematic, but statistics such as the temporal correlation or the periodogram do not require it. As expected, the temporal correlation between the daily values for REMEDHUS, SMOS and ORCHIDEE soil moisture estimates is very high (Table 1). The variability in the correlation of each station in the basin with either the remotely sensed or modelled data is small and can be neglected. The model correlates slightly better in time to the in-situ observations than the retrieved SSM from SMOS ($r = 0.75$ vs. $r = 0.82$). These results are in the same range of the temporal correlation obtained with SMOS and HTESSEL over the same site (Albergel et al., 2012).

The high temporal correlation values are mainly driven by the strong response of SSM to rainfall events and thus the high frequency behaviour of the variable. This measure though is not very sensitive to the slower variations. In this regard, these results do not inform on the ability of the model or the remote sensing methodology to reproduce well the annual cycle of SSM.

Fig. 1 shows that the time characteristics of SSM's response to rainfall events and the annual wetting are different across the 4 data sets. This suggests that a spectral analysis might be revealing. We have estimated the spectral power using the Lomb-Scargle method (Press & Rybicki, 1989). The current SMOS data availability permits reliable estimations of the spectral power from periods of 6 days to 1.5 years as explained in Section 2.3. To account for the variability of the spectral power between the 19 stations, we use the maximum and minimum spectral power at each frequency as uncertainty estimates.

Fig. 2a provides the smoothed spectra over the REMEDHUS basin for all 4 data sets. The model is always within the variability of the 19 stations and reproduces certain remarkable spectral peaks found in the in-situ data, in particular the maximum at 9–10 days and minimum at 16–18 days. The sharpness of SSM peaks found in Fig. 1 for SMOS, relative to the other data sets, is reflected here by higher values in the high frequency part of the spectrum. For periods shorter than 8 days SMOS shows more energy than found in any of the 19 REMEDHUS stations. In contrast, at the lower frequencies all four data sets agree well.

To better understand the spectral analysis, Fig. 2b compares the soil moisture spectra of the OR-ERA simulation to the rainfall data set with which the model was forced and that of an independent gauge-based gridded rainfall product, E-OBS (also thinned according to SMOS retrieval availability) (Haylock et al., 2008). It can be noted that the spectra for rainfall are much flatter than those of soil moisture. In other words the rainfall signal is closer to white noise than the water accumulating in the upper soil. This can be explained by the autocorrelation introduced by the inertia of the soil moisture that shifts energy from the high to the lower frequencies (Entin et al., 2000). This red-shift of the

Table 1
Temporal correlation between the four surface soil moisture data sets and ratio of spectral power (R) between the 6–30 days and 90–365 day bands.

Data set	Temporal correlation (r)				Ratio R of spectral power
	REMEDHUS	SMOS	OR-ERA	OR-WFD	
REMEDHUS	1	0.75 ± 0.01	0.83 ± 0.01	0.82 ± 0.01	$0.025 \pm 1.9.10^{-4}$
SMOS		1	0.81	0.81	0.071
OR-ERA			1	0.93	0.042
OR-WFD				1	0.046

forcing rainfall spectrum is an intrinsic property of the soil. It is determined by the water holding capacity and drainage to deeper layers. Evaporative demand of the atmosphere will add further energy at lower frequencies, as its temporal evolution is slower than that of precipitation. To illustrate this phenomenon, Fig. 2b provides the spectra for several surface soil moisture layers. As layers become thicker, the spectrum shifts progressively towards red noise and the peaks are smoothed out. For instance the peak at 9–10 days, which originates from the rainfall signal (thick curves in Fig. 2b) is strong for layers up to 20 cm (1–7 layers) but loses intensity for thicker layers. In the same way the spectral peak in rainfall at 7.5 days is rapidly dampened out by the inertia of the soil and is lost at 10 cm (1–6 layers), demonstrating the time filtering operated by the vertical diffusion and evaporation.

Comparing the ERA-I rainfall spectrum with the one of the OR-ERA simulation shows that the two strong extrema in SSM (maximum at 9–10 days and minimum at 16–18 days) are partially explained by rainfall variability at this site. The strongest spectral peak in rainfall at 7.5 days corresponds to a weaker peak in SSM compared to that at 9–10 days. This can be explained by the fact that these two spectral characteristics of rainfall occur at different seasons and thus, depending on antecedent soil moisture and the atmospheric evaporative demand, their impact on SSM will be different. Analogously, the relatively weak spectral minimum of rainfall at 16–18 days corresponds to a strong surface soil moisture minimum. Detailing the seasonal dependence of the spectral characteristics of rainfall and soil moisture needs further signal processing and is beyond the scope of this study.

The shown results suggest that it would be helpful to quantify the steepness (or redness) of the spectrum. We propose here to use the ratio R of the mean energy in the sub monthly (6–30 days) and seasonal to annual (90–365 days) bands for the evaluation of the spectral characteristics of the SSM estimated by SMOS and ORCHIDEE. This allows us to evaluate the relative weight of the annual cycle and the synoptic variability in the different SSM estimates. The legend of Fig. 2b shows how R decreases as the thickness of the considered soil moisture layer increases. Table 1 gives the values of R for the REMEDHUS basin. The ratio is very stable between the 19 stations, which reflects the importance of the rainfall variability. The two ORCHIDEE simulations show very similar R values, which approximately double the R values obtained from the in-situ observations. For SMOS, the ratio is three times larger than for REMEDHUS, reflecting the high temporal variability, or low SSM memory, noted in Fig. 1.

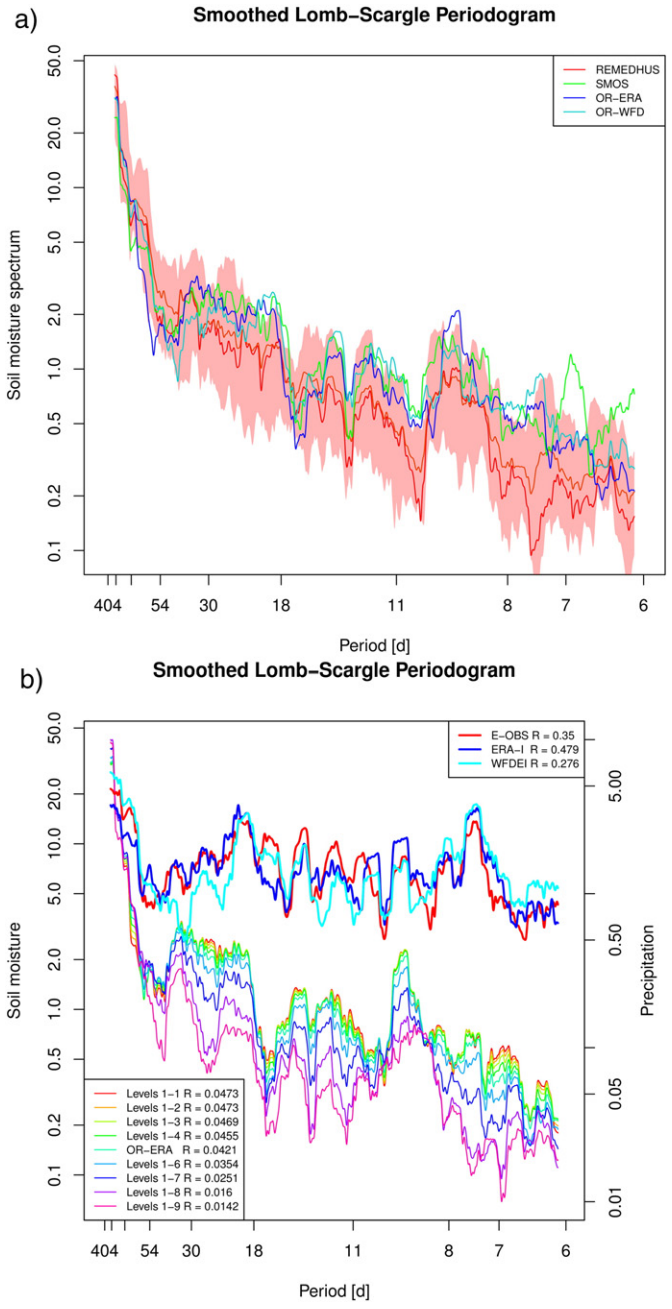

Fig. 2. a) Normalized spectral density for surface soil moisture for: the averaged REMEDHUS in-situ measurements (red), SMOS retrievals (green) and 2 ORCHIDEE simulations (blue and cyan). The shaded area around the average periodogram of all 19 observing sites (orange), represents the minimum and maximum of the spectra obtained for REMEDHUS. b) Spectral density of rainfall (thick lines) at the REMEDHUS site from E-OBS, ERA-I and WFDEI. The thin lines provide the periodogram for several soil depths of the OR-ERA simulation. The line labelled OR-ERA corresponds to the top 5 cm (integrating levels from 1 to 5). The legend displays the ratio of spectral power between the 6–30 and 90–365 day bands. The smoothing is performed with sliding windows covering 3% of the total frequency interval.

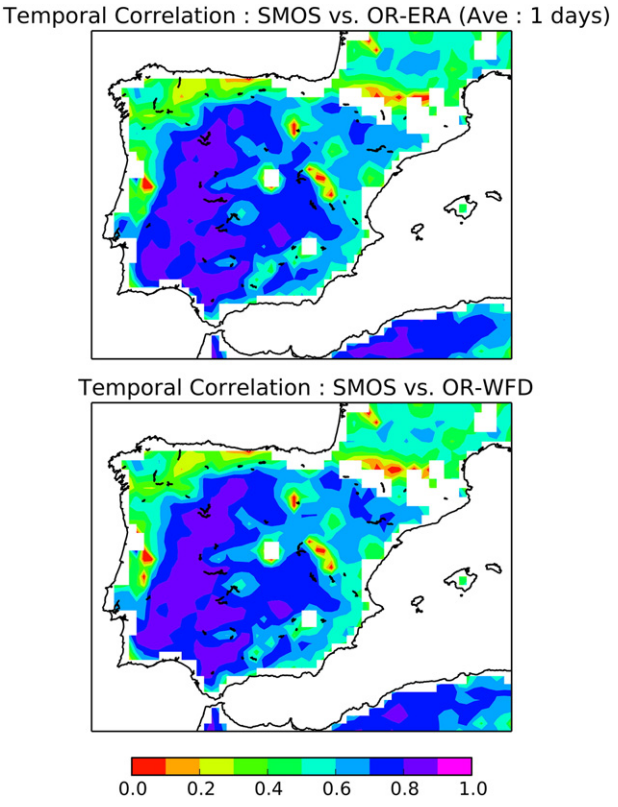

Fig. 3. Temporal correlation over the Iberian Peninsula between SSM estimates from SMOS and ORCHIDEE simulations forced by ERA-I (OR-ERA) (top) and WFDEI (OR-WFD) (bottom).

3.2. Temporal correlation

The temporal correlation of SMOS and the two ORCHIDEE simulations over the REMEDHUS site is generalized to the entire IP. Fig. 3 shows that the temporal correlation between daily time series is very good with values above 0.7 for large parts of the IP, confirming the results obtained over the REMEDHUS basin. All values are statistically significant at 95% level except for a few points, for which the correlation was set to zero in the maps. Considering the OR-ERA (upper panel of Fig. 3) or the OR-WFD (lower panel) simulation does not lead to large differences in the temporal correlation with the SMOS signal. This is as expected, since the rainfall data forcing the two simulations share the same synoptic variability, i.e. one of the re-analysis.

Most of the regions with correlations below 0.6 correspond to mountainous areas like the Iberian System (which is oriented from North-West to South-East along the 4 red dots in Fig. 3), the Central System or the Penibaetic System in the South-East. Topography is known to pose problems for SSM retrieval algorithms from passive microwave measurements (Mätzler & Standley, 2000). The steepest mountains, like the Pyrenees, are often flagged and excluded from the retrievals provided for SMOS. More intriguing are the low temporal correlations found in the north-western part of the IP. This region is dominated by an oceanic climate and a large amplitude of the rainfall annual cycle. Noteworthy as well are the relatively low temporal correlation values (between 0.6 and 0.8) in the Ebro valley, which is located between the Pyrenees and the Iberian System (see Fig. 3).

3.3. Spatial correlation

Remotely sensed products are the only means to retrieve large scale observations of water availability in the upper layers of the soil. Thus, they are a unique source of information for validating the ability of models to reproduce the spatial structures of SSM. We compute the spatial correlation (ρ) between remotely sensed and modelled data using 5-day averaged maps. The averaging is needed to obtain maps that are as complete as possible. For ORCHIDEE the average was computed based on the thinned dataset where only the values corresponding to SMOS observations are retained. From the sampling error analysis (Section 2.3) we know that without this methodology we would introduce a spatial correlation error of at most 0.08.

Fig. 4 provides the time series of spatial correlation between the remotely sensed and the simulated SSM. The average spatial correlations are rather weak: $\rho = 0.28$ for OR-ERA and $\rho = 0.26$ for OR-WFD. These values are smaller than both the spatial correlation computed between rainfall data sets ($\rho = 0.59$ for E-OBS/ERA-I and $\rho = 0.64$ for E-OBS/WFDEI) and the averaged temporal correlations, which are 0.69 for OR-ERA and 0.68 for OR-WFD. It is worth noting that the spatial correlation is slightly better for the simulation forced with ERA-I, although this forcing data-set is known to be affected by biases that are partly corrected in the WFDEI data. There appears to be some seasonal variability in these results with the highest values obtained during the spring and summer of 2012. But there is no notable increase in spatial correlation during the spring and summer of 2011 although the density of SMOS observations is about the same as the following year. The lack of observations in the first half of 2010 does not appear to worsen the spatial correlation.

To better understand the poor correlation between spatial structures, we evaluate if the differences originate in the fast or slow components of the SSM fields. For this diagnostic we decompose the SSM time series into a smooth annual cycle and an anomaly independently at each grid point (as explained in Section 2.5). The spatial correlations are then computed for the slow and fast components separately. It is important to note that the slow signal accounts for the inertia of the soil moisture, as demonstrated by the high energy at low frequencies (see Section 3.1). The physical interpretation of long term fluctuations (beyond seasonal periodicity) of SSM are also difficult because they mix very different regimes of the variable. On the other hand, the fast component is likely to integrate high frequency noise or errors.

Fig. 5 shows that the slow component exhibits low spatial correlation values, whose average is smaller than the one found for the total signals. In contrast with what is displayed in Fig. 4, the highest values are attained in autumn and the minimum in winter and spring. In other words, the spatial correlation is higher during the driest period of year for the IP, while the lowest values are found during the wettest period. The spatial correlation of the fast component is higher than the one obtained for the total signal, reaching the largest values in autumn and winter. These results indicate that the major differences in the spatial structures between SMOS and ORCHIDEE may be attributed to the slow varying component of SSM and not to its rapid response to rainfall events.

3.4. Spectral analysis

To examine in depth the poor spatial correlation and identify the regions degrading the spatial correlations, we will analyse the steepness of the periodograms and its relation to rainfall at each point. The following analysis is based on the assumption, detailed in Section 3.1, that the variability of SSM is essentially driven by rainfall and the inertia of the

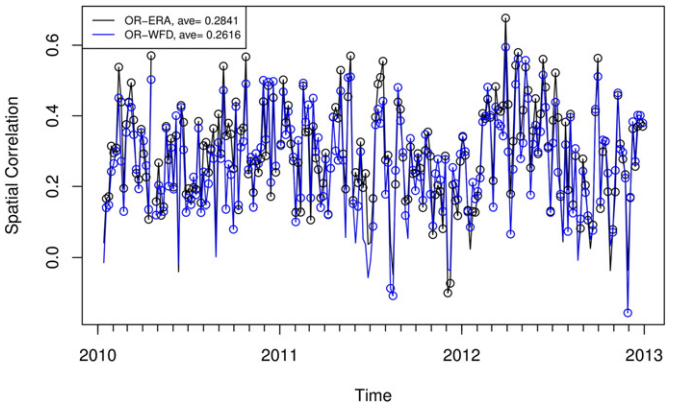

Fig. 4. Spatial correlation (ρ) of the remote sensed and simulated SSM over the Iberian Peninsula. 5 day averaged maps were correlated and significant values (p-values > 0.95) are highlighted with circles.

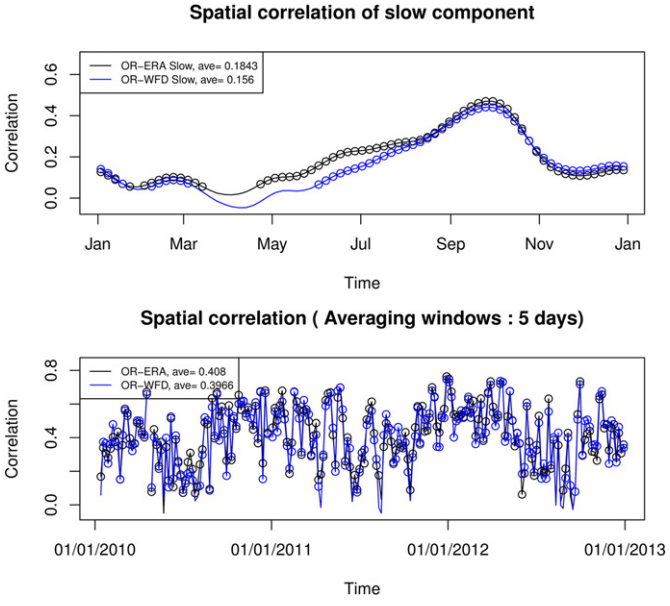

Fig. 5. Spatial correlation between the SMOS and ORCHIDEE estimates for SSM decomposed into the mean annual cycles (upper panel) and anomalies (lower panel).

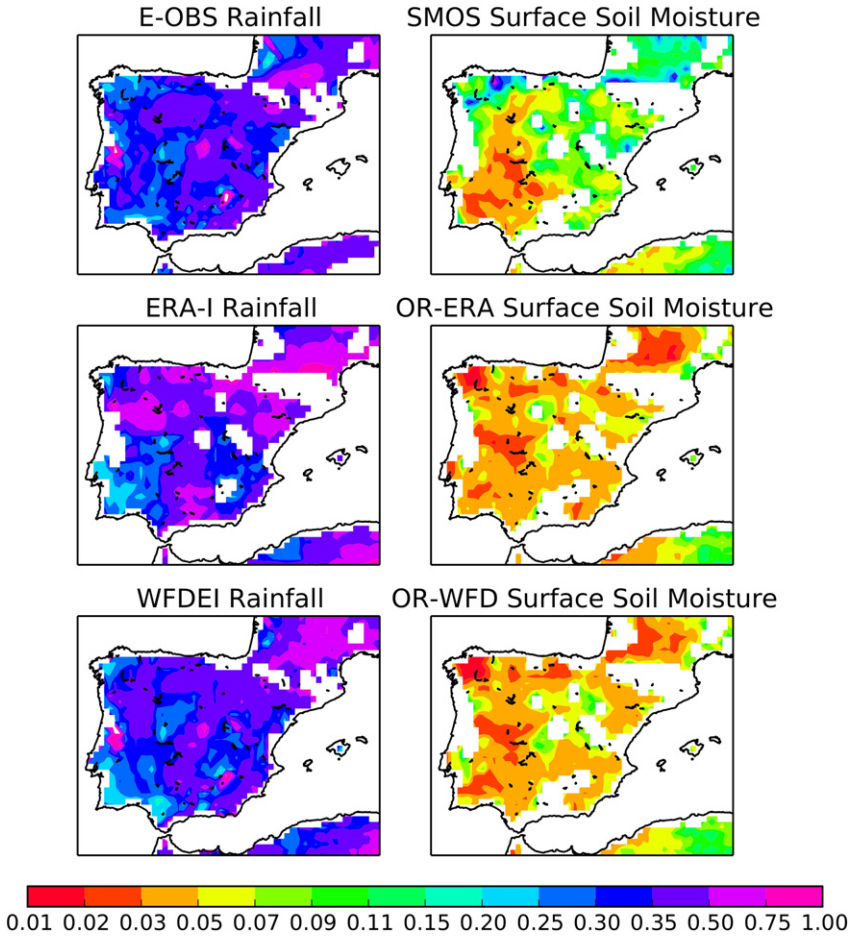

Fig. 6. Ratio of spectral power in the 6–30 days and 90–365 day bands. First row: E-OBS rainfall and SMOS surface soil moisture. Second row: ERA-I rainfall and OR-ERA SSM. Third row: WFDEI rainfall and OR-WFD SSM.

soil moisture reservoir shifts its spectrum towards lower frequencies. Thus we examine how the ratio (R) of energy between the 6–30 and 90–365 days band changes from the rainfall to SSM estimates. To avoid any biases, this diagnostic is performed on daily means sampled as the SMOS observations.

The first column of Fig. 6 provides the ratio R for the rainfall datasets. For E-OBS (an independent rain gauge based estimate) the ratio R is between 0.3 and 0.5 over most of the IP with its minima in the North-West, indicating an important role of the annual cycle in that region. This is consistent with the climatology of the IP, which is characterized by a very weak annual cycle of rainfall in the East and a strong contrast between rainy winters and relatively dry summers in the West. The ERA-I rainfall shows the largest contrasts in the spectral ratios and is characterized by high values (R above 0.5) in large parts of the North-East. The bias-correction performed to obtain the WFDEI rainfall reduces these contrasts and the R map is more similar to that of E-OBS. Despite the uncertainty in the R estimates for rainfall, the outcome of this diagnostic can be well related to the rainfall climatology of the region. Finally we note that for rainfall the ratio is always higher than 0.2.

Using the same colour scale for rainfall and SSM, the right column of Fig. 6 illustrates clearly the red-shift of the spectrum of SSM, displaying smaller values of the ratio R for SSM than for rainfall. SMOS and ORCHIDEE are in good agreement over an area in the centre of the IP to the West of border between Spain and Portugal. In this area, which includes the REMEDHUS basin, there is over 20 times more energy in the 90–365 days band than in the shorter periods (6–30 days). Over the same region, the rainfall has less than 5 times more energy in the 90–365 days band, thus showing that the red-shifts observed by SMOS and modelled by ORCHIDEE are qualitatively similar.

The Ebro basin is characterized by significant differences between the observations and the model. ORCHIDEE's SSM has more energy in the annual cycle (R below 0.05), while for SMOS R is between 0.05 and 0.09. It could well be that the model does not drain or evaporate quickly enough the water out of the top 5 cm of soil, thus overestimating the slow component of the SSM variability. It is interesting to note that the spectral differences between the two forcing rainfall datasets result in small differences between the spectra of the simulated SSM and which are significantly smaller than the discrepancies with the remotely sensed data. The possibly larger inertia of the modelled SSM could also explain the relatively low temporal correlation found over the Ebro valley in Section 3.1. However it is difficult to point at a weakness of either the model or the SMOS data without the availability of in-situ observations that could serve as an independent benchmark.

The largest discrepancies are again found in the North-West of the Peninsula, where the rainfall spectrum is characterized by relatively small R values. ORCHIDEE produces there its lowest values of the ratio for SSM, thus shifting the spectrum further towards the longer periods. This means that, according to the model, individual rainfall events impact only weakly the SSM evolution in the north-western of the IP. On the other hand the spectrum for SMOS displays in this region the highest R values, which are furthermore very close to those of rainfall. Thus, the remotely sensed SSM indicates a weak red-shift of the rainfall spectrum, i.e. only a small part of the spectral power is displaced to the lower frequencies when compared to the model. Such behaviour would be expected from ponded water, which has very little memory and thus

its variability is dominated by rainfall. The same behaviour is found for SMOS in southern France as well.

This spectral analysis of SSM reveals that the model and the remotely sensed data have very different behaviours in specific regions. The regions of the IP with the largest discrepancies in the spectral characteristics are also those with the lowest temporal correlation. The differences in SSM variability over large fractions of the IP, as shown here through discrepancies in R, can lead to different spatial covariances in the model and the remotely sensed data and thus explain the low spatial correlation. Other diagnostic tools are needed to better understand the co-variance of SSM and rainfall and perhaps identify the cause of the low spatial correlation between products.

3.5. The spatio-temporal co-variance of SSM and rainfall

The co-variance of rainfall and soil moisture can be analysed by means of the singular value decomposition (SVD) method (Bretherton, Smith, & Wallace, 1992; Wallace, Smith, & Bretherton, 1992). We apply it here to the simulated and remotely sensed daily SSM values using the independent rainfall dataset E-OBS (Haylock et al., 2008). Using these precipitation estimates as benchmarks ensures that the model is not favoured in the comparison, as it would be if we used the forcing rainfall, and that the differences found originate in SSM. Furthermore all differences in co-variant patterns can only originate in the soil moisture products. The diagnostics will be focused on the first three co-varying patterns (CPi, i = 1,2,3). They represent at least 97% of the co-variance of surface soil moisture and rainfall in all three SSM estimates (Table 2). The CPs are sorted by their fraction of explained co-variance as well as by their spatial correlation to facilitate the discussion. This is why CP2 for the modelled SSM explains less co-variance than CP3.

The SSM expansion coefficients (Fig. 7), which are the projections of the soil moisture field on the co-varying patterns identified via SVD, show that CP1 has a large annual cycle. Thus CP1 can be linked to the seasonal co-variability of rainfall and SSM. It is worth noting at this stage that the amplitude of variation of the expansion coefficients are quite similar between SMOS and ORCHIDEE. As illustrated by Fig. 7, CP2 and CP3 represent the higher frequency (synoptic) variability of soil moisture linked to rainfall events. They explain very similar fractions of the co-variance in all three decompositions, i.e. between 2% and 4%. At the REMEDHUS site (see Section 3.1), the synoptic rainfall variability is the dominant cause of the co-variance between the SSM datasets. On the other hand, at the regional scale the SVD analysis shows that the annual cycle dominates the co-variability of rainfall and SSM. This is the case for SMOS as well as for ORCHIDEE and is explained by different climatological characteristics of rainfall across the region and their impact on SSM structures. The synoptic co-variability summarized by CP2 and CP3 plays a minor role, probably because of the large spatial variability of this part of the signal. The last column of Table 2 provides the spatial correlation of the three first precipitation and SSM CPs. Generally, the fast varying CPs show higher spatial correlation values than CP1. This can be explained by the fact that the variability represented by CP1 is determined not only by rainfall, but also

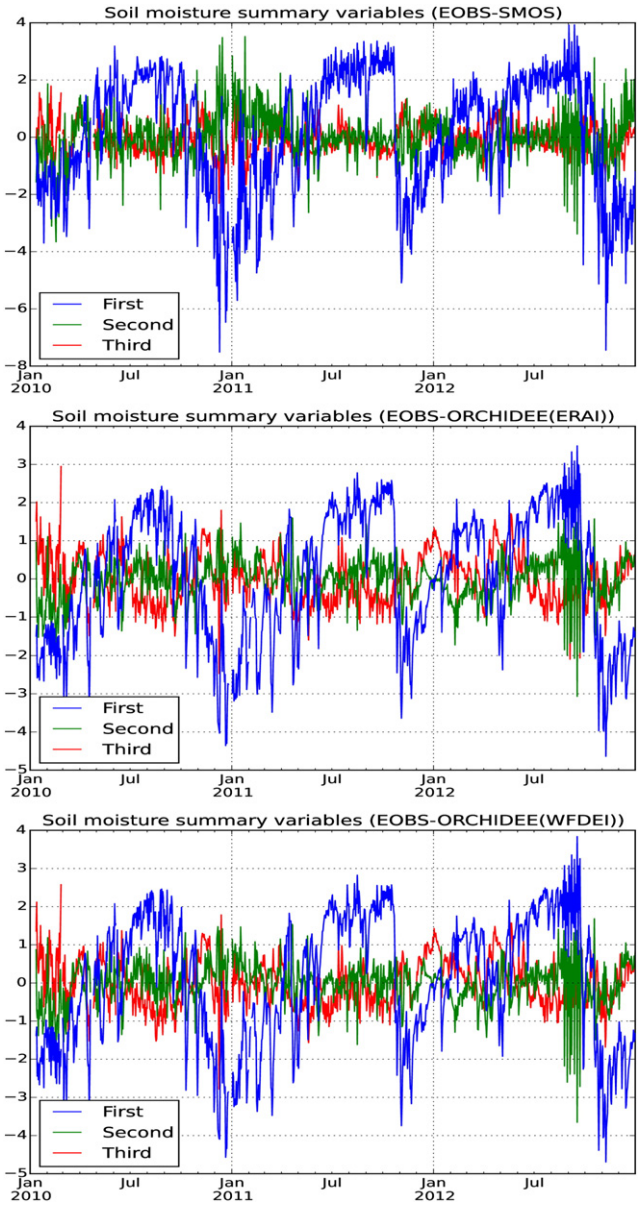

Fig. 7. Expansion coefficients for the first three pairs of co-varying patterns. From top to bottom the SVD analysis corresponds to E-OBS/SMOS, E-OBS/OR-ERA and E-OBS/OR-WFD.

Table 2
Characteristics of the first three pairs of co-varying patterns of precipitation (P) and surface soil moisture (SSM). The explained (co-)variance statistics are expressed as fractions.

Data	Co-varying Patterns	Explained co-variance	Homogeneous explained variance		Heterogeneous explained variance		Spatial correlations of P and SSM co-varying patterns.
			P	SSM	SSM exp. by P	P exp. by SSM	
E-OBS SMOS	CP1	0.91	0.34	0.39	0.08	0.07	0.47
	CP2	0.03	0.12	0.07	0.01	0.01	0.79
	CP3	0.02	0.12	0.06	0.01	0.02	0.76
E-OBS OR-ERA	CP1	0.91	0.34	0.64	0.15	0.08	0.46
	CP2	0.02	0.12	0.11	0.01	0.01	0.82
	CP3	0.04	0.13	0.13	0.02	0.02	0.91
E-OBS OR-WFD	CP1	0.91	0.34	0.63	0.16	0.08	0.45
	CP2	0.03	0.12	0.09	0.01	0.02	0.86
	CP3	0.04	0.12	0.12	0.02	0.02	0.92

by other processes such as evaporation and drainage. As for the faster varying CPs these processes are less important, we observe higher spatial correlations between rainfall and SSM patterns.

Between 15% and 16% of the ORCHIDEE SSM variance is explained by the rainfall CP1 (heterogeneous explained variance fractions in Table 2). Consistently, the precipitation variance attributable to the first CP of SSM is only 8%. The difference in these heterogeneous correlations is the consequence of the asymmetric causal link between rainfall and soil moisture that favours the SSM variance explained by rainfall. For SMOS, the heterogeneous explained variance fractions are not only lower (8% and 7%), but also the difference between the two values is very small. We may conclude that, at the seasonal scale (CP1), the link between rainfall and SSM is weaker for SMOS data than for the model. Moreover, the synoptic variability of precipitation extracted by the SVD analysis (CP2 and CP3) seems to explain a negligible part of the regional SSM variability.

The pattern maps (Figs. 8, 9 and 10) allow to locate the areas where the co-variance between rainfall and soil moisture is largest. In all three SVD analyses the spatial structure for rainfall is dominated by the western coast of the IP and the western flank of the Penibaetic System. This structure corresponds well to the dominant mode of rainfall variability identified by Rodriguez-Puebla, Encinas, Nieto, and Garmendia (1998). The spatial correlation between the three maps of rainfall CP1 is above 0.98. In the corresponding soil moisture CPs, the signal is strongest over the south-western and western IP. In ORCHIDEE the maximum extends further to the North-West than in SMOS. In the 3 pairs the spatial correlation between the CP1 for rainfall and SSM is between 0.45 and 0.47 (Table 2). Although differences between the soil moisture CP1 can be identified, the spatial correlations of SMOS CP1 with those of ORCHIDEE are above 0.63. These values of spatial correlations of the SSM CP1 are higher than the raw spatial correlations presented in Section 3.2. Moreover, the soil moisture CP1 for SMOS explains 39% of the total SSM variance, while for ORCHIDEE this fraction is between 63 and 64% (homogeneous explained variance fractions in Table 2).

The second CPs characterize the East West gradient of rainfall and are dominated by the high frequency variations as illustrated by the expansion coefficients (Fig. 7). It is consistent with the third mode of variability identified by Rodriguez-Puebla et al. (1998). Although the spatial correlation between the three rainfall CP2s is high (ρ above 0.96) some notable differences are induced in the analysis by the associated soil moisture. The CP2 of SSM reflect well the rainfall structure; spatial correlations are above 0.79, (see last column of Table 2). Some discrepancies are found in the North-West and the South-East (Murcia region) of the IP. The first region shows, as for CP1, a weaker correspondence between SMOS SSM and rainfall. While over Murcia, the modelled SSM is less correlated with rainfall. The spatial correlation of CP2 for SSM of SMOS with those of ORCHIDEE is above 0.79, but the total variance of SSM explained by the CP2 of rainfall is small in all cases (1%).

CP3 represents the North South gradient of rainfall variability and is in good agreement with the second mode identified by Rodriguez-Puebla et al. (1998). As noted above it is characterized by a strong synoptic variability. The CP3s of SSM show similar spatial structures and their correlation with the rainfall patterns is above 0.76. It is worth noting that the similarity between rainfall and soil moisture CP3s over the Murcia region is stronger for the model than for the remote sensed product. Thus, there might be a distribution of the co-variability between CP2 and CP3 which is slightly different between the three analysis. Over the Ebro basin we find that rainfall and SSM CP3s are more similar for SMOS than the model. This can be related to the result of the spectral analysis, which characterized the modelled SSM with larger inertia compared to SMOS. The spatial correlation between the CP3s of soil moisture between SMOS and the model is 0.74 for both simulations.

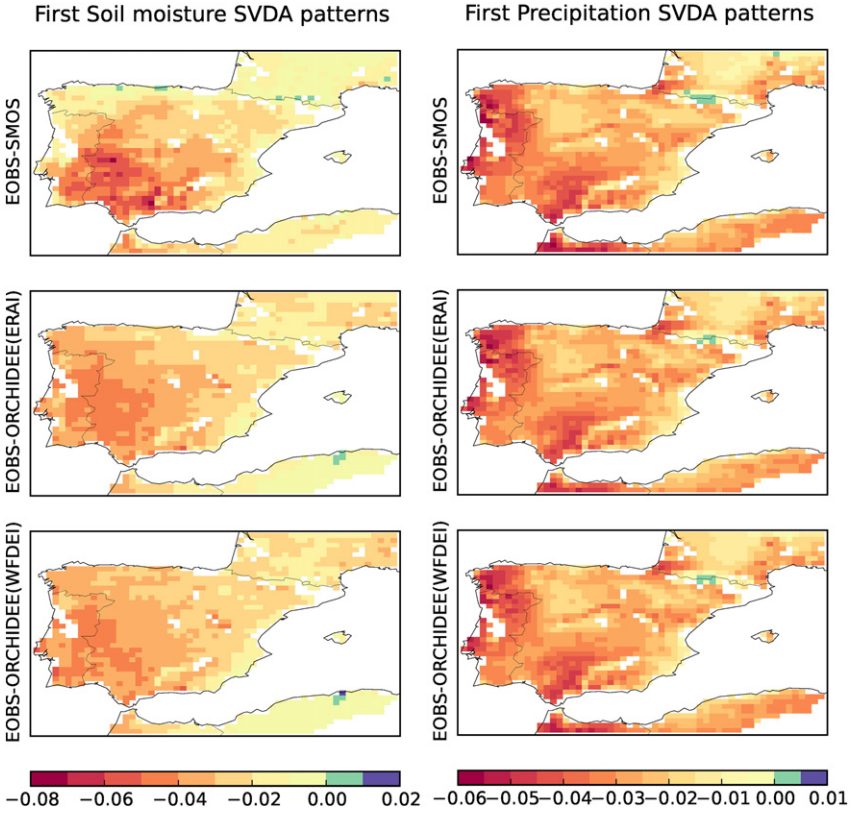

Fig. 8. First co-varying for rainfall (left column) and surface soil moisture (right column) for the three SVD analysis: First row E-OBS and SMOS. Second row E-OBS and OR-ERA. Third row E-OBS and OR-WFD.

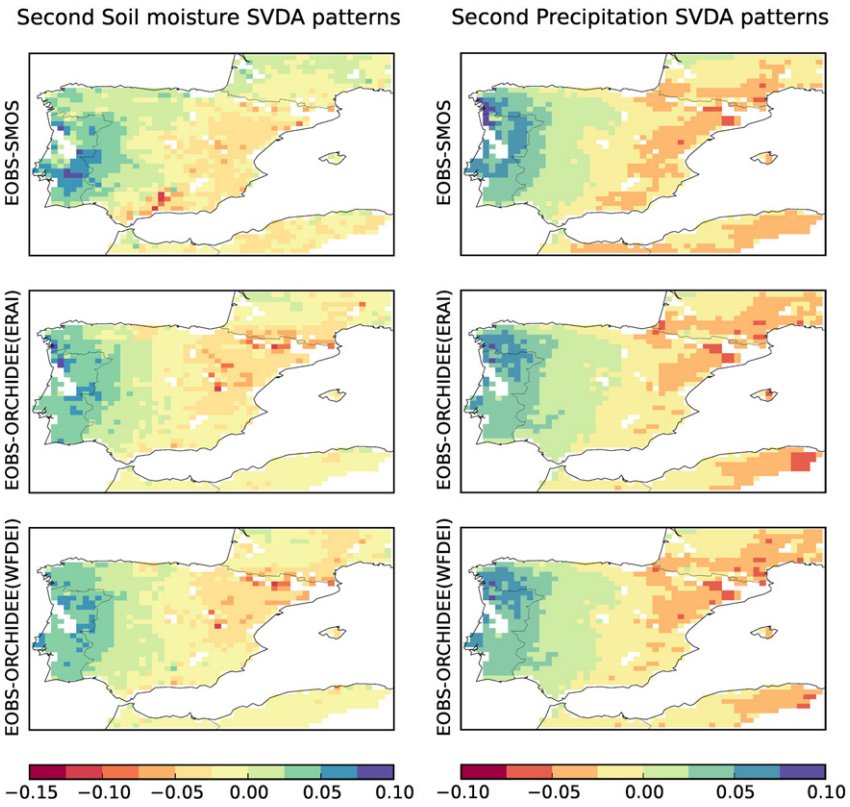

Fig. 9. As Fig. 8 but for the second pairs of co-varying patterns.

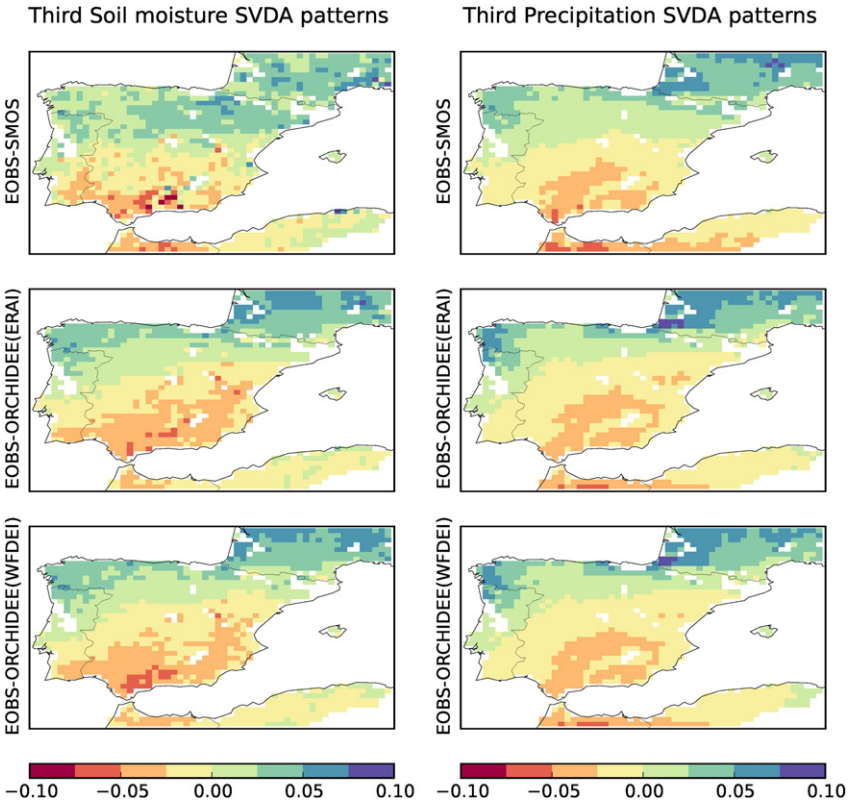

Fig. 10. As Fig. 8 but for the third pairs of co-varying patterns.

The SVD analysis between rainfall and SSM is quite revealing over this region. At this spatial scale, the co-variance of the two variables is dominated by the annual cycle. Moreover the identified co-varying patterns are similar between the model and the remotely sensed data. The largest discrepancies between SMOS and ORCHIDEE are found again in the North-West. This methodology confirms that the response of SSM to rainfall is a source of differences between the remotely sensed and modelled values but it does not seem to be major. A new aspect identified by the SVD analysis is that only 52% of SMOS SSM variability at the regional scale can be attributed to the co-variance with rainfall, while for the model this fraction is between 84% and 88%. These values are obtained by summing the homogeneous explained variance over the first three CPs.

4. Discussion and conclusions

This study presents an analysis between surface soil moisture (SSM) estimates over the Iberian Peninsula: the Soil Moisture and Ocean Salinity (SMOS) satellite retrievals are compared to the output of the ORCHIDEE land surface model. *In-situ* measurements over the REMEDHUS basin are used for an initial benchmarking. The analysis of the spatio-temporal variability of SSM is then extended at regional scale. The co-variance of SSM and precipitation is also analysed.

Our results are in good agreement with previous results showing that the time evolution of SSM measured in-situ, remotely sensed or modelled, display good correlations (Albergel et al., 2012; Parrens et al., 2012). A spectral analysis of the temporal variability of SSM brings some nuances to these results. Indeed the red-shift of the rainfall spectrum is linked to the physics of the upper soil moisture, in particular to its inertia. Both SMOS and ORCHIDEE display spectral slopes which are less steep than the ones found in the in-situ data. The excess of energy in the high frequency variations of the remotely sensed and modelled SSM can be attributed to assumptions made in SMOS and ORCHIDEE on the soil physics that are not supported by REMEDHUS data. These differences found in soil moisture dynamics and, in particular, in the effective inertia of the moisture reservoir, are not well diagnosed by time correlations but are probably reflected in the poor spatial correlations.

The temporal variability analysis is extended to the entire IP to identify regions where the SSM estimates of SMOS and ORCHIDEE may differ. The results are quite satisfactory as over a large part of the IP the temporal correlation is above 0.7. In some mountainous regions and in oceanic climate conditions, as expected the correlation falls to lower values (Mätzler & Standley, 2000). To refine our understanding of these results, we analysed the spectral slopes of the three regional SSM products. This analysis demonstrates that large temporal correlations are associated to similar spectral characteristics only over a narrow region stretching northward from the gulf of Cádiz, along the Spanish Portuguese border. Elsewhere, the SMOS spectrum is flatter than the model one. This suggests that SMOS may observe a shallower and faster soil moisture reservoir than the model does. These results support the need for more refined diagnostics than simple temporal correlations, in order to provide more in-sight in the comparisons of remotely sensed, *in-situ*, and modelled SSM.

The spectral analysis performed here is rather qualitative. Using simple models for the SSM reservoir, as for instance the one proposed by Brocca et al. (2014), the relation between the spectral characteristics of rainfall and SSM could be interpreted through some aggregated soil parameters. This way the differences between SSM estimates could be quantified. It would provide a more detailed picture of where the model and the satellite observe soils with similar characteristics.

The spatial correlation of the SSM maps obtained from SMOS and the two simulations are rather low (Fig. 4). The values are smaller than the average temporal correlations or the spatial correlation between the considered rainfall products. Thus, despite the good temporal correlations at most of the points, the small differences lead to degraded spatial co-variances between points and thus a poor match of the spatial structures. Gonzalez-Zamora et al. (2015) found a similar result when comparing a regional network of in-situ observations to the SMOS products. They attribute this result to the spatial mismatch of coarse-scale remotely sensed observations and point measurements. However, this argument does not hold when comparing a remotely sensed product with a model output, as they are both designed for similar spatial scales.

In the present study we investigate this lack of spatial correlation by trying to identify the time scale at which it is most significant. First, this was pursued by separating at each grid point the slow from the fast variations of SSM. It demonstrated that the lowest spatial correlations are obtained for the slow varying part of the signal. This methodology is simple and does not take into account that a large part of the SSM spatial structures are driven by the rainfall variability. This diagnostic was thus further refined by analysing the co-variance of rainfall and SSM via singular value decomposition (SVD). SVD identifies the patterns of co-variance between the two variables and determines the fraction of explained co-variance. This analysis demonstrates that the co-variability of SSM and rainfall is characterized by similar structures in the remotely sensed and modelled variables. The concordance is stronger than the spatial correlation found for the full soil moisture signal. Thus we may conclude that the rainfall driven structures of SSM that are captured by the model and SMOS are compatible.

The more puzzling outcome of the SVD analysis is that in ORCHIDEE over 80% of the SSM variability is explained by the identified co-varying patterns, while in SMOS this ratio is only of 52%. This leads to the hypothesis that the low spatial correlation between the SSM estimates has in large parts its origin in structures that are not driven by the synoptic and seasonal variability of rainfall. This in turn raises the question on the origins of these slow varying spatial structures, which are not rainfall-driven and which exist in SMOS but are not captured by ORCHIDEE. Answering this question might lead the community developing land surface models to add processes that are captured by the L-band measurements. Should these be artefacts of the retrieval algorithm or properties of the L-band remote sensing, then understanding these structures will lead to improvements in the retrieval methodologies.

Based on the presented results, we have some investigation avenues to propose. The spectral analysis identified the difference in steepness of the spectra of SSM as one major discrepancy between the modelled and remotely sensed SSM. It appears that the SMOS product is closer to white noise, or has less inertia, than what the model seems to indicate. This would be consistent with the findings of Escorihuela, Chanzy, Wigneron, and Kerr (2010) who indicate that the penetration depth of the L-band signal is shallower than what is assumed and that it depends on the soil moisture conditions. Thus the SSM reservoir would be more responsive to rainfall events when the climate becomes wetter, thus enhancing the higher frequency variability. The climate of the north-western corner of the IP, which exhibits a particularly poor match between SMOS and ORCHIDEE, constitutes the ideal setting for this effect to play an important role.

Analysing only SMOS-derived soil moisture estimates, it is not possible to conclude on the likely causes of the spatial mismatch with ORCHIDEE. A similar analysis performed on the observed brightness temperatures, and those which can be simulated by microwave emission model (De Rosnay et al., 2009; Drusch, Holmes, de Rosnay, & Balsamo, 2009) using ORCHIDEE moisture profiles, may indicate whether the spectral characteristics are affected by a varying penetration depth and to which extent the brightness temperature patterns should be co-varying with rainfall. The answer to this question will be an important contribution to the usage of the L-band measurements available from SMOS and Aquarius, or the recently launched SMAP mission (Entekhabi et al., 2010).

The present analysis can be repeated over any region of the world that has a sufficiently dense rainfall gauge network to establish reliable estimates of daily rainfall. Determining the characteristics of co-variance of rainfall and surface soil moisture allows to better understand the remotely sensed variable and their relation to what is represented in

land surface models. A continuous effort in this direction is an important support to all data assimilation projects that are using remote sensing products of the water cycle to build climatologies of surface state variables or to initialise forecast systems.

Acknowledgements

This work has been supported by the Spanish Ministry of Science and Education under the projects AYA2010-11062-C05-04 and AYA2010-11062-C05-05. The first author is grateful to the Institut national des sciences de l'Univers-CNRS, and in particular Jean-Marie Flaud, for allowing his secondment to IC3. This work contributes to the FP7 Earth2Observe project under grant agreement No 603608. We would like to thank N. Sánchez and J. Martínez-Fernández for making the REMEDHUS data available and G. Balsamo for providing the ERA-I forcing.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:http://dx.doi.org/10.1016/j.rse.2015.12.004. These data include a Google map of the most important areas described in this article.

References

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., ... Wagner, W. (2012). Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. *Remote Sensing of Environment*, *118*(March), 215–226. http://dx.doi.org/10.1016/j.rse.2011.11.017.

Balsamo, G., Albergel, C., Beljaars, A., Bousetta, S., Brun, E., Cloke, H., Dutra, E., et al. (2012). ERA-interim/land: a global land-surface reanalysis based on ERA-interim meteorological forcing. *ERA-Report Series*, *13*, 25.

Bretherton, C.S., Smith, C., & Wallace, J.M. (1992). An intercomparison of methods for finding coupled patterns in climate data. *Journal of Climate*, *5*(6), 541–560. http://dx.doi.org/10.1175/1520-0442(1992)0052.0.CO;2.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., ... Levizzani, V. (2014). Soil as a natural rain gauge: estimating global rainfall from satellite soil moisture data: using the soil as a natural raingauge. *Journal of Geophysical Research: Atmospheres*, *119*(9), 5128–5141. http://dx.doi.org/10.1002/2014JD021489.

D'Orgeval, T., Polcher, J., & de Rosnay, P. (2008). Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes. *Hydrology and Earth System Sciences*, *12*(6), 1387–1401. http://dx.doi.org/10.5194/hess-12-1387-2008.

De Rosnay, P. (2002). Impact of a physically based soil water flow and soil-plant interaction representation for modeling large-scale land surface processes. *Journal of Geophysical Research*, *107*(D11). http://dx.doi.org/10.1029/2001JD000634.

De Rosnay, P. (2003). A GCM experiment on time sampling for remote sensing of near-surface soil moisture. *Journal of Hydrometeorology*, *4*(2), 448–459. http://dx.doi.org/10.1175/1525-7541(2003)4<448:AGEOTS>2.0.CO;2.

De Rosnay, P., & Polcher, J. (1998). Improvements of the representation of the hydrological exchanges between the biosphere and the atmosphere in a GCM. *Hydrology and Earth System Sciences*, *2*(2–3), 239–256.

De Rosnay, P., Bruen, M., & Polcher, J. (2000). Sensitivity of surface fluxes to the number of layers in the soil model used in GCMs. *Geophysical Research Letters*, *27*(20), 3329. http://dx.doi.org/10.1029/2000GL011574.

De Rosnay, P., Drusch, M., Boone, A., Balsamo, G., Decharme, B., Harris, P., ... Wigneron, J. -P. (2009). AMMA land surface model intercomparison experiment coupled to the community microwave emission model: ALMIP-MEM. *Journal of Geophysical Research*, *114*(D5). http://dx.doi.org/10.1029/2008JD010724.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., et al. (2011). The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. http://dx.doi.org/10.1002/qj.828.

Dirmeyer, P.A., Dolman, A.J., & Sato, N. (1999). The global soil wetness project: a pilot project for global land surface modeling and validation. *Bulletin of the American Meteorological Society*, *80*, 851–878.

Draper, C.S., Reichle, R.H., De Lannoy, G.J.M., & Liu, Q. (2012). Assimilation of passive and active microwave soil moisture retrievals. *Geophysical Research Letters*, *39*(4). http://dx.doi.org/10.1029/2011GL050655.

Drusch, M., Holmes, T., de Rosnay, P., & Balsamo, G. (2009). Comparing ERA-40-based L-band brightness temperatures with Skylab observations: a calibration/validation study using the community microwave emission model. *Journal of Hydrometeorology*, *10*(1), 213–226. http://dx.doi.org/10.1175/2008JHM964.1.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., et al. (2010). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, *98*(5), 704–716. http://dx.doi.org/10.1109/JPROC.2010.2043918.

Entin, J.K., Robock, A., Vinnikov, K.Y., Hollinger, S.E., Liu, S., & Namkhai, A. (2000). Temporal and spatial scales of observed soil moisture variations in the extratropics. *Journal of Geophysical Research*, *105*(D9), 11865–11877. http://dx.doi.org/10.1029/2000JD900051.

Escorihuela, M.J., Chanzy, A., Wigneron, J.P., & Kerr, Y.H. (2010). Effective soil moisture sampling depth of L-band radiometry: a case study. *Remote Sensing of Environment*, *114*(5), 995–1001. http://dx.doi.org/10.1016/j.rse.2009.12.011.

Gonzalez-Zamora, A., Sánchez, N., Martinez-Fernandez, J., Gumuzzio, A., Piles, M., & Olmedo, E. (2015). Long-Term SMOS Soil Moisture Products: A Comprehensive Evaluation across Scales and Methods in the Duero Basin. *Physics and Chemistry of the Earth, Parts A/B/C*, *83–84*, 123–136. http://dx.doi.org/10.1016/j.pce.2015.05.009.

Harris, I., Jones, P.D., Osborn, T.J., & Lister, D.H. (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 dataset: UPDATED HIGH-RESOLUTION GRIDS OF MONTHLY CLIMATIC OBSERVATIONS. *International Journal of Climatology*, *34*(3), 623–642. http://dx.doi.org/10.1002/joc.3711.

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research*, *113*(D20). http://dx.doi.org/10.1029/2008JD010201.

Horne, J.H., & Baliunas, S.L. (1986). A prescription for period analysis of unevenly sampled time series. *Astrophysical Journal*, *302*, 757–763.

Kerr, Y.H., Waldteufel, P., Richaume, P., Wigneron, J.P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., et al. (2012). The SMOS soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1384–1403. http://dx.doi.org/10.1109/TGRS.2012.2184548.

Kerr, Y.H., Waldteufel, P., Wigneron, J.P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M. -J., et al. (2010). The SMOS mission: new tool for monitoring key elements ofthe global water cycle. *Proceedings of the IEEE*, *98*(5), 666–687. http://dx.doi.org/10.1109/JPROC.2010.2043032.

Kerr, Y.H., Waldteufel, P., Wigneron, J. -P., Martinuzzi, J., Font, J., & Berger, M. (2001). Soil moisture retrieval from space: the soil moisture and ocean salinity (SMOS) mission. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1729–1735. http://dx.doi.org/10.1109/36.942551.

Koster, R.D., Mahanama, S.P.P., Yamada, T.J., Gianpaolo, B., Berg, A.A., Boisserie, M., Dirmeyer, P.A., et al. (2011). The second phase of the global land–atmosphere coupling experiment: soil moisture contributions to subseasonal forecast skill. *Journal of Hydrometeorology*, *12*(5), 805–822. http://dx.doi.org/10.1175/2011JHM1365.1.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., ... Prentice, I.C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*(1). http://dx.doi.org/10.1029/2003GB002199.

Martínez-Fernández, J., & Ceballos, A. (2005). Mean soil moisture estimation using temporal stability analysis. *Journal of Hydrology*, *312*(1–4), 28–38. http://dx.doi.org/10.1016/j.jhydrol.2005.02.007.

Mätzler, C., & Standley, A. (2000). Technical note: relief effects for passive microwave remote sensing. *International Journal of Remote Sensing*, *21*(12), 2403–2412. http://dx.doi.org/10.1080/01431160050030538.

Muñoz-Sabater, J., Jarlan, L., Calvet, J. -C., Bouyssel, F., & De Rosnay, P. (2007). From near-surface to root-zone soil moisture using different assimilation techniques. *Journal of Hydrometeorology*, *8*(2), 194–206. http://dx.doi.org/10.1175/JHM571.1.

Ngo-Duc, T., Polcher, J., & Laval, K. (2005). A 53-year forcing data set for land surface models. *Journal of Geophysical Research*, *110*(D6). http://dx.doi.org/10.1029/2004JD005434.

Oliva, R., Daganzo, E., Kerr, Y.H., Mecklenburg, S., Nieto, S., Richaume, P., & Gruhier, C. (2012). SMOS radio frequency interference scenario: status and actions taken to improve the RFI environment in the 1400–1427-MHz passive band. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1427–1439. http://dx.doi.org/10.1109/TGRS.2012.2182775.

Parrens, M., Zakharova, E., Lafont, S., Calvet, J. -C., Kerr, Y., Wagner, W., & Wigneron, J. -P. (2012). Comparing soil moisture retrievals from SMOS and ASCAT over France. *Hydrology and Earth System Sciences*, *16*(2), 423–440. http://dx.doi.org/10.5194/hess-16-423-2012.

Post, W.M., & Zobler, L. (2000). Global Soil Types, 0.5-Degree Grid (Modified Zobler). *Available on-line [http://www.daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.*.

Press, W.H., & Rybicki, G.B. (1989). Fast algorithm for spectral analysis of unevenly sampled data. *Astrophysical Journal*, *338*(March), 277–280.

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. -J., Arsenault, K., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. http://dx.doi.org/10.1175/BAMS-85-3-381.

Rodriguez-Puebla, C., Encinas, A.H., Nieto, S., & Garmendia, J. (1998). Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula. *International Journal of Climatology*, *18*(3), 299–316. http://dx.doi.org/10.1002/(SICI)1097-0088(19980315)18:3<299::AID-JOC247>3.0.CO;2-L.

Sánchez, N., Martínez-Fernández, J., González-Piqueras, J., González-Dugo, M.P., Baroncini-Turrichia, G., Torres, E., ... Pérez-Gutiérrez, C. (2012b). Water balance at plot scale for soil moisture estimation using vegetation parameters. *Agricultural and Forest Meteorology*, *166-167*(December), 1–9. http://dx.doi.org/10.1016/j.agrformet.2012.07.005.

Sánchez, N., Martinez-Fernandez, J., Scaini, A., & Perez-Gutierrez, C. (2012a). Validation of the SMOS L2 soil moisture data in the REMEDHUS network (Spain). *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1602–1611. http://dx.doi.org/10.1109/TGRS.2012.2186971.

Taylor, C.M., de Jeu, R.A.M., Guichard, F., Harris, P.P., & Dorigo, W.A. (2012). Afternoon rain more likely over drier soils. *Nature*, *489*(7416), 423–426. http://dx.doi.org/10.1038/nature11377.

Taylor, C.M., Gounou, A., Guichard, F., Harris, P.P., Ellis, R.J., Couvreux, F., & De Kauwe, M. (2011). Frequency of sahelian storm initiation enhanced over mesoscale soil-moisture patterns. *Nature Geoscience*, *4*(7), 430–433. http://dx.doi.org/10.1038/ngeo1173.

Vautard, R., Yiou, P., D'Andrea, F., de Noblet, N., Viovy, N., Cassou, C., ... Fan, Y. (2007). Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit. *Geophysical Research Letters*, *34*(7). http://dx.doi.org/10.1029/2006GL028001.

Wallace, J.M., Smith, C., & Bretherton, C.S. (1992). Singular value decomposition of wintertime sea surface temperature and 500-Mb height anomalies. *Journal of Climate*, *5*(6), 561–576. http://dx.doi.org/10.1175/1520-0442(1992)0052.0.CO;2.

Weedon, G.P., Gomes, S., Viterbo, P., Shuttleworth, W.J., Blyth, E., Österle, H., ... Best, M. (2011). Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *Journal of Hydrometeorology*, *12*(5), 823–848. http://dx.doi.org/10.1175/2011JHM1369.1.

Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J., & Viterbo, P. (2014). The WFDEI Meteorological Forcing Data Set: WATCH Forcing Data Methodology Applied to ERA-Interim Reanalysis Data. *Water Resources Research*. http://dx.doi.org/10.1002/2014WR015638 (September, n/a – n/a).

Zampieri, M., D'Andrea, F., Vautard, R., Ciais, P., de Noblet-Ducoudré, N., & Yiou, P. (2009). Hot European summers and the role of soil moisture in the propagation of Mediterranean drought. *Journal of Climate*, *22*(18), 4747–4758. http://dx.doi.org/10.1175/2009JCLI2568.1.