

A joint analysis of modeled soil moisture fields and satellite observations

Carlos Jiménez,¹ Douglas B. Clark,² Janna Kolassa,³ Filipe Aires,³ Catherine Prigent,¹

Abstract.

A methodology to conduct a joint analysis of modeled soil moisture fields from the Joint UK Land Environment Simulator (JULES) and a dataset of multi-wavelength observations is presented. It consists of building a statistical model capturing the relationships between the land surface model estimates and the satellite observations, and then using the satellite observations (mapped into soil moisture predictions by the statistical model) to evaluate the fields estimated by the land surface model. Two statistical models are tested and predict very similar soil moisture (global correlation and RMSD of ~ 0.98 and $\sim 0.02 \text{ m}^3/\text{m}^3$). A characterization of prediction uncertainty shows errors ranging between 0.01 and 0.10 m^3/m^3 , depending on biome and season. The satellite prediction and JULES soil moisture agree relatively well (global correlation and RMSD of ~ 0.92 and $\sim 0.05 \text{ m}^3/\text{m}^3$), but for some regions and periods clear differences exist. Conducted tests modifying either the predicted soil moisture or the JULES estimates show that this methodology can effectively change soil moisture towards more correct values. It can then be expected that some of the differences are the result of the satellite information modifying the modeled soil moisture fields towards more realistic values. However, proving this is difficult given the present uncertainties in modeled and observed global soil moisture products.

1. Introduction

Soil moisture accounts for a very small fraction of the total Earth water budget, but it plays an important role in the climate system as it modulates key land-surface-atmosphere processes such as the partitioning of available net energy at the surface into latent and sensible heat fluxes, soil evaporation and plant transpiration, partitioning of precipitation into infiltration and runoff, or plant growth. This means that an accurate representation of global soil moisture fields by models is crucial. However, inter-comparison exercises show that modeling this variable is still subject to a large uncertainty, with relatively large inter-model variance compared with other modeled hydrological variables [Dirmeyer *et al.*, 2006].

This situation is not helped by a lack of mature observational global soil moisture data records. The first mission to specifically measure soil moisture from space was launched in 2009 (the Soil Moisture and Ocean Salinity mission, SMOS) [Kerr *et al.*, 2010], meaning that all other observational records of soil moisture need to be derived from sensors that were not specifically designed to measure this variable. Progress is being made, mainly by exploiting existing observations from active and passive microwave missions [e.g. Wagner *et al.*, 1999; Owe *et al.*, 2001; Liu *et al.*, 2012], but product errors can still be relatively large over some regions due to the challenges related to extracting the soil moisture information from a satellite signal (non-optimized

for soil moisture) largely modulated by other processes (e.g., vegetation, roughness) [Dorigo *et al.*, 2010]. In situ observations of soil moisture also exist [e.g. Dorigo *et al.*, 2011]. They can presumably be considered as more accurate than satellite observations, and have the advantage of providing not just superficial soil moisture as the satellite observations, but also the soil moisture at deeper layers. The challenges here are to deal with the limited spatial and temporal coverage of the in situ network, and with the representativity issues associated to comparing modeled spatially-integrated estimates and in situ point measurements over heterogeneous landscapes.

Facing these difficulties, a simple comparison between existing observational and modeled soil moisture products is not sufficient, and different statistical techniques to address them are emerging. Good examples are the triple collocation method to characterize the error structure of related soil moisture datasets [e.g. Scipal *et al.*, 2008], and upscaling techniques to reduce the impact of spatial sampling errors when comparing sparse ground observation with estimates with a large ground resolution [e.g. Crow *et al.*, 2012]. For model evaluation we are interested in developing complementary methodologies focusing more on (1) exploiting the synergy between different observations in order to deal with the limitations from single observations, and (2) imposing a degree of consistency between the model and observation estimates to assure a meaningful evaluation. The latter is of special relevance when evaluating modeled soil moisture. The different model physical schemes can result in very different modeled soil water equilibrium states and simulated discharge, which can also result in large differences in the soil water content estimated by different models [e.g. Boone *et al.*, 2004].

One of these methodologies was already presented in [Aires *et al.*, 2005]. It consists of building a statistical model capturing the complex relationships between the land surface model estimates and different satellite observations, and then using the satellite observations (mapped into soil moisture estimates by the statistical model) to evaluate the

¹Laboratoire d'Etudes du Rayonnement et de la Matière en Astrophysique, Centre National de la Recherche Scientifique, Observatoire de Paris, Paris, France.

²Centre for Ecology and Hydrology, Wallingford, Oxfordshire, UK.

fields predicted by the land surface model. Here we will adopt a similar methodology to conduct a joint analysis of modeled soil moisture fields from the Joint UK Land Environment Simulator (JULES, [Best et al., 2011; Clark et al., 2011]) and a dataset of multi-wavelength observations. Although only long wave microwave observations are traditionally exploited for soil moisture products, Prigent et al. [2005] demonstrated that observations at shorter wavelengths also contain moisture-related information; they will also be used in our analysis.

The paper is organized as follows. Firstly the satellite observations and JULES model are presented in Section 2. This is followed by a description of the methodology in Section 3. The statistical model applied in [Aires et al., 2005] was based on neural networks; here we will also apply a second different statistical model to evaluate the impact of algorithm choice in the predicted soil moisture and to estimate prediction errors. The results are presented in Section 4. The JULES and predicted soil moisture from the satellite observations are compared and discrepancies between both estimates are discussed. This is complemented by some tests where the JULES soil moisture is perturbed in order to evaluate the potential of this methodology to detect inconsistencies in modeled fields. Finally a discussion of the results and the main conclusions are given in Section 5.

2. Data

2.1. Land surface model

JULES is a mechanistic model of the land surface including representations of the surface energy balance, evaporation from soil and vegetation, and snow and soil physics, including runoff generation. For the present study the global land areas were modeled on a grid of $0.5^\circ \times 0.5^\circ$ resolution. The land cover was represented by 9 land cover types: 5 plant functional Types and 4 non-vegetation types, including bare soil. In each gridbox the fractional coverage of each type was calculated by aggregating the IGBP version 2 land cover map to the model grid and mapping the IGBP land classes to the JULES land types. The IGBP data were also used to specify the Leaf Area Index of each vegetation type. Fluxes of heat and moisture in the soil are calculated by a 4-layer finite-difference model which includes representations of the effects of phase changes of soil moisture. The soil model considers a total depth of 3 m, with the surface soil layer being 10 cm thick. Soil hydraulic characteristics were also calculated from IGBP data. Snow is represented using a multi-layer approach in which the temperature, frozen and liquid water content, grain size and density of each layer are simulated. The model was run for 1951 to 2001, preceded by a multi-decadal spin-up stage, and the timestep length was one hour. Near-surface meteorology was prescribed using the WATCH Forcing Data [Weedon et al., 2011], which are based on the ERA-40 Re-Analysis product [Uppala et al., 2005] with monthly bias corrections based on observations. The simulations considered ‘near-natural’ conditions, meaning that human impacts such as irrigation were not included.

2.2. Satellite observations

The following satellite observations will be used in the analysis:

- **Visible and Near-IR.** The AVHRR instruments on board the NOAA meteorological polar orbiters provide daily observations of the Earth with a resolution as high as 1 km. The first channel is in the visible, where chlorophyll causes absorption of incoming radiation, while the second one is in the near infrared. The Normalized Difference Vegetation Index (NDVI) [Gutman, 1999] is extensively used for vegetation studies. Here the visible and near-infrared radiances are used directly instead of the NDVI product, to allow the

prediction method to find the best way of extracting the soil moisture-related information from the radiances. For this study the 10-day composite AVHRR product generated under the joint NASA and NOAA Earth Observing System Pathfinder Project [James and Kalluri, 1994], with a resolution of 8 km, is used.

- **Thermal-IR.** The ISCCP dataset of surface skin temperature is produced at 3 hour intervals since 1983 over the globe, every 30 km, combining all the infrared measurements from polar and geostationary operational weather satellites [Rossow and Schiffer, 1999]. For this study the surface skin temperatures were extracted from the ISCCP-DX product. Aires et al. [2004] developed a method to reconstruct the diurnal cycle of surface skin temperature for each location over the globe, based on a statistical analysis of the 3-hourly skin temperatures for clear scenes. The reconstructed skin temperatures are averaged to provide the monthly diurnal cycle of temperature and to derive its amplitude.

- **Active Microwave.** The European Remote-sensing Satellite ERS-1 was launched in 1991 and remained operational until 2001 [Francis et al., 1991]. Its suite of instruments included a vertically polarized radar operating at C-band (5.3 GHz). It was originally designed to measure near-surface winds over oceans, with a nominal resolution of 50 km, but it has also proved useful for soil moisture characterization [e.g., Wagner et al., 1999]. For this study, the backscattering coefficients were processed following a method similar to Frison and Mougin [1996], keeping the values at 20° and 45° .

- **Passive Microwave.** The SSM/I instruments on board the Defense Meteorological Satellite Program (DMSP) polar orbiters observe the Earth twice daily at 19.35, 22.24, 37.00, and 85.50 GHz. The fields-of-view decrease with frequency, from $43 \times 69 \text{ km}^2$ to $13 \times 15 \text{ km}^2$ [Hollinger et al., 1987]. Instead of using the raw microwave brightness temperatures, we use the estimated land surface emissivities from [Prigent et al., 2006]. These emissivities are estimated from SSM/I observations by removing contributions from the atmosphere, clouds, rain, and the surface temperature, and are related to the surface properties themselves, minimizing the contribution to the signal from the other factors.

None of these observations is optimized for characterizing surface soil moisture. The large dielectric constant at microwave frequencies of water (compared with other materials) produces a strong response to soil moisture, but this response is modulated by factors such as surface roughness, the presence of vegetation and atmospheric contributions. Notice that microwave observations at lower frequencies better suited to estimation of soil moisture also exist (e.g., the 6.6/6.9 GHz and 10.7 GHz AMSR-E observations), but their observing period falls outside the years of the land surface model runs. The surface temperature diurnal cycle is related to soil moisture through the soil thermal inertia being modulated by the soil moisture, but it also depends on other factors such as solar radiation, air temperature and humidity, and near-surface wind. The visible and infrared radiances are more closely related to vegetation than moisture, but the strong correlation between available soil moisture and vegetation growth in some regions can provide indirect information about soil moisture. A more detailed discussion about the sensitivity of these observations to soil moisture and other related processes is given in Prigent et al. [2005]; Aires et al. [2005].

2.3. Pre-processing

The satellite observations are re-gridded into a common global equal area grid ($0.25^\circ \times 0.25^\circ$ at the equator, with a pixel area of approximately 770 km^2) and averaged into

monthly means. Shorter time and space scales would be desirable, but for a first analysis we judge monthly means as a good compromise between satellite acquisition, enhancement of the signal to noise ratio of the observations, and the objective of deriving a global multi-decadal climatology. The period 1993-1999 is selected for this first product, coinciding with the period where both model outputs and our selected observations exist.

The selected satellite observations are sensitive, at best, to the first few centimetres of the surface. Therefore we select for the study the model monthly mean estimates from the top surface layer (the top 10 cm). They are given on a $0.5^\circ \times 0.5^\circ$ grid; a weighted-distance technique is used to match satellite and model resolutions. Only model estimates and observations considered to be snow-free (using the National Snow and Ice Data Center (NSIDC) monthly mean snow dataset [Armstrong and Brodzik, 2005] to filter the satellite data, and only selecting “unfrozen” soil moisture model estimates) and water-free (using the wetland dataset of Prigent *et al.* [2006]; Papa *et al.* [2010] to filter the satellite data and removing model estimates over water bodies) are used in the analysis and reported in the final product.

The regression dataset is built by selecting 200 000 matched observations and JULES soil moisture estimates from the odd months of 1993 and 1999. These estimates are selected randomly but assuring that the regression is equalized in soil moisture space, i.e. the soil moisture estimates in the regression dataset are roughly equally distributed. This avoids the regression giving greater weight to the regions of soil moisture space that have a larger number of estimates, and assures that the more extreme soil moisture values are also well considered.

The derivation of the statistical link between observations and soil moisture uses a global database, without any land cover information passed to the regression models. However, a land cover classification based on the model land types is used to report soil moisture as function of different biomes. As a given model cell can be partitioned in different land types, a cell is assigned a land type only if at least 50% of the cell surface corresponds to that land type. This means that some pixels are unclassified in terms of land cover and are not included in the statistics computed for a specific biome (although the soil moisture is always estimated for all pixels). The classification is shown in Figure 1. Notice that changing the 50% threshold to higher values, or even using a different land cover type for the analysis, does not change substantially the analysis of results based on biome types.

3. Methodology

3.1. Regression algorithms

Satellite measurements are traditionally converted to geophysical parameters with the aid of a radiative transfer model capable of simulating the observations. The retrieved values of the geophysical parameters of interest correspond to the radiative transfer inputs of the simulation satisfying some closure criteria with the observations. Here soil moisture cannot be retrieved by this approach. A joint inversion of our observations would require a multi-wavelength radiative transfer model able to reproduce all observations from a common dataset of geophysical inputs; this model does not exist. Even if this theoretical model could be put together, the existence of all the model parameters and inputs needed to allow the model to capture the complex global response of the surface at all these wavelengths could be questioned. Therefore, we propose a different approach consisting of (1) building a global database $\{\mathbf{X}^l, \mathbf{Y}^l\}_{l=1 \dots L}$ of L pairs of coincident observations (the vector \mathbf{y}) and modeled soil moisture estimates (the scalar \mathbf{x}); (2) using this database to build a statistical link between the observations

and the modeled soil moisture, and; (3) using the statistical link to predict soil moisture estimates (denoted by \mathbf{x}_{pre}) that can be compared with the modeled estimates. Notice that as in [Aires *et al.*, 2005] we prefer to use the term prediction rather than retrieval for this methodology, to make clear that this work cannot be strictly defined as a retrieval scheme per se (i.e., it is neither using a radiative transfer model to allow a physical retrieval scheme, nor a database of observations and measured soil moisture states to allow an empirical retrieval scheme, but only linking observations with soil moisture states predicted by a land surface model).

Two regression models are tested here to build the statistical link. The first regression method is based on a specific neural network topology called Multi-Layer Perceptrons (MLP). A MLP is composed of a number of neurons (the processing elements) organized in layers. If the input to the layer j is expressed as a vector \mathbf{i}^j , the weights and bias of the neurons (the adaptive parameters) are expressed respectively as a matrix \mathbf{W}^j and a vector \mathbf{b}^j , the output from the neurons is grouped as a vector \mathbf{o}^j , and the activation function of all neurons in layer j is f_j , the output of layer j is given by

$$\mathbf{o}^j = f_j(\mathbf{W}^j \mathbf{i}^j + \mathbf{b}^j). \quad (1)$$

The predicted soil moisture is given by the propagation of the observation input vector $\mathbf{y} = \mathbf{i}^1$ through the M layers of the MLP:

$$\mathbf{x}_{\text{pre}} = \mathbf{o}^M = \mathbf{f}_M(\mathbf{W}^M \mathbf{i}^M + \mathbf{b}^M) = \mathbf{f}_M[\mathbf{W}^M \mathbf{f}_{M-1}(\mathbf{W}^{M-1} \mathbf{i}^{M-1} + \mathbf{b}^{M-1}) + \mathbf{b}^M] = \dots \quad (2)$$

The number of layers, number of neurons for each layer, and type of activation functions are variables to be determined for each specific application of the MLP. Weights and biases are determined during a training phase in which the MLP is subject to adaptation to the database of examples $\{\mathbf{X}^l, \mathbf{Y}^l\}_{l=1 \dots L}$. More details about the practical implementation of the specific MLP used here can be found in Jiménez *et al.* [2009].

The second method is based on a Numerical Integration of the a Posteriori probability (NIP) of the parameters to be predicted for a given set of observations. The predicted soil moisture is given by the mean of the a posteriori probability. In practical terms the integration is solved by a summation over the L states of the database $\{\mathbf{X}^l, \mathbf{Y}^l\}_{l=1 \dots L}$ [Evans *et al.*, 2002]. If $P(\mathbf{y}/\mathbf{x})$ is the conditional probability of the observations (given a soil moisture state), the predicted soil moisture can be estimated as

$$\mathbf{x}_{\text{pre}} = \frac{\sum_{i=1}^L \mathbf{X}_i P(\mathbf{y}/\mathbf{X}_i)}{\sum_{i=1}^L P(\mathbf{y}/\mathbf{X}_i)} \quad (3)$$

If the observation error is assumed to follow a multivariate, normally distributed probability density function, and \mathbf{S}_{err} is the observation noise covariance matrix, the conditional probability of the observations (given a soil moisture state) can be expressed as

$$P(\mathbf{y}/\mathbf{X}_i) \sim \exp\left(-\frac{(\mathbf{y} - \mathbf{Y}_i) \mathbf{S}_{\text{err}}^{-1} (\mathbf{y} - \mathbf{Y}_i)^T}{2}\right) \quad (4)$$

Once \mathbf{x}_{pre} is calculated, the prediction error can be estimated as the standard deviation of the conditional a posteriori probability, which can be approximated by the summation

$$\mathbf{x}_{\text{err}} = \frac{\sum_{i=1}^L (\mathbf{X}_i - \mathbf{x}_{\text{pre}})^2 P(\mathbf{y}/\mathbf{X}_i)}{\sum_{i=1}^L P(\mathbf{y}/\mathbf{X}_i)} \quad (5)$$

This method requires specifying the observation uncertainty. We assume a diagonal \mathbf{S}_{err} (i.e. the observation errors are uncorrelated) and standard deviations of 0.012 for the emissivities, 5% of the mean value for the backscatter and the visible and near-infrared reflectances, and 5 K for the temperature diurnal cycle, based on published values. Notice that a constant observation uncertainty is an obvious simplification, as our observations are not strictly sensor radiances but derived products, the derivation of which could be more or less certain depending on the environmental conditions for each observation (e.g., saturation of NDVI at high biomass values, emissivity uncertainty propagating in the derivation of surface temperature). However, a detailed error characterization is very difficult to infer from the published work for all the observations we used, and we prefer to assume a constant uncertainty for these first tests of the algorithm. More details of this algorithm can be found in *Evans et al. [2002]*; *Rydberg et al. [2009]*.

3.2. Soil moisture prediction

In principle, both regression methods should predict similar soil moistures as it can be demonstrated that both estimate the mean state of the a posteriori probability if the uncertainty of the input (observations) and target (soil moisture) variables follows Gaussian statistics [*Bishop, 1995*; *Evans et al., 2002*]. In practice, differences are expected. For instance, the real statistics can deviate from the assumption of a perfectly normally-distributed probability. Table 1 shows the Root Mean Square Difference (RMSD, as percentage of the averaged JULES soil moisture for each land type) between the JULES soil moisture and the satellite-driven MLP and NIP prediction. The Pearson's correlation coefficient for the different soil moisture pairs is also given. Unless indicated otherwise, the statistics are calculated over the 1993-1999 period. The RMSDs and correlations for the different land types are very close for both methods, with a RMSD of $\sim 30\%$ between JULES and the predicted soil moisture and global correlations of ~ 0.90 . The RMSD between both predicted soil moistures is reduced to $\sim 13\%$, with a correlation of 0.98, indicating that the predictions agree well. Examples of the geographical distributions of soil moisture in August 1998 moisture and their differences (between JULES and the MLP prediction and between the MLP and NIP predictions) are given in Figure 2. The predicted soil moisture captures the expected soil moisture distributions related to different surface conditions and hydrological regimes, although some differences are apparent. As expected from the high correlations shown in Table 1, the soil moisture distributions from both methods look very similar. This is confirmed in the difference maps, showing much smaller differences between the MLP and NIP predictions, compared with the differences between MLP and JULES. Histograms of the predicted and JULES soil moisture values for the different land types are shown in Figure 3. The histograms for the MLP and NIP predictions are closer to one another than to the JULES histograms, again showing a relatively good agreement between both predictions. The largest differences between predicted and JULES soil moisture occur for the shrublands and C4 grass land types. These differences will be commented on in Section 4.1.

Gaps in the predicted soil moisture maps can be seen in Figure 2; these are related to missing observations. If missing one type of satellite observation does not greatly compromise the accuracy of the prediction, soil moisture could still be predicted by the same methodology but adapting the regression to consider only the existing observations. Table 2 gives the same statistics as in Table 1 for the NIP method when all observations but one are used as predictors. As expected, the RMSDs are slightly larger and correlations

slightly lower ($\sim 3\%$ and 0.03 respectively for the predictions without the temperature observations, the worst prediction). Similar figures (not given) are found for the MLP prediction. We consider these results to be acceptable and from now on predicted soil moisture values when one type of observation is missing will also be reported. More than one type of observation can be missing in some rarer cases, and the same scheme could also be applied with the remaining observations, but this is not pursued further here.

3.3. Prediction errors

The NIP method was chosen as it can be used to derive not only the mean of the a posteriori probability density function, but also its standard deviation. This can be regarded as an estimation of the uncertainty in the predicted value, and gives an idea of the prediction error. Uncertainty estimates can also be derived when using the MLP model [e.g. *Aires, 2004*], but their derivation is computationally more difficult than for the NIP algorithm. Although for some applications a general idea about the uncertainty in the predicted values is sufficient, for other applications it is necessary to always associate the predicted value with an estimate of its error. For instance, this would be the case if the satellite-driven predicted soil moisture values were to be assimilated into a numerical model [e.g. *Aires et al., 2005*].

Figure 4 shows some statistics of the prediction error. At each pixel an error distribution is produced by grouping together the prediction error for each monthly prediction at each pixel for the 7 years analyzed. The median of this distribution is reported in the upper panel of the figure. The largest absolute errors occur over areas in the boreal forest of Asia and North America, with median errors of $\sim 0.8 \text{ m}^3/\text{m}^3$. The smallest absolute errors are found over the arid and semi-arid areas. However, in relative terms these are the areas where the errors are larger; the very low soil moisture and correspondingly low signal in the satellite observations results in relatively uncertain predictions. This can be seen in Figure 3, where histograms of the prediction errors for the different land types are plotted together with the soil moisture histograms from JULES and the prediction methods. The error histograms for all land types peak at relatively similar values, while the peaks of soil moisture histograms occur over a range of values, with the driest peak being for bare soil and the wettest for broadleaf trees. To identify the areas where the prediction is more uncertain, Figure 4 also shows for each pixel the number of predictions where the predicted soil moisture value is at least three times larger than the prediction error (expressed as percentage of the total number of predictions for that pixel for the 7 years considered). The predictions are very uncertain over the dry deserts (e.g. the Sahara) and mountainous (e.g. the Andes) areas. Over Southern Asia the predicted soil moisture over India seems more uncertain than over neighbouring regions (e.g. southern China), though it should be noted that for part of the year a large part of India can be relatively dry, compared with other southern Asian regions. It is worth remembering that a constant observation uncertainty was assumed all over the globe, so this first error prediction, although of value, would require further refinement through a more realistic description of the geographical variability in observation uncertainty.

4. Results

4.1. Comparing modeled and predicted soil moisture

Figure 5 shows the 1993-1999 averaged seasonal soil moisture fields from JULES and the satellite prediction (NIP approach; the MLP fields are not displayed but are similar), and the RMSD between the two. The expected geographical distributions related to biomes and climate zones are visible in both JULES and the satellite-driven prediction

(for instance, the location of the main rainforests, deserts, mountainous regions, or transition zones such as the change from forest to grasslands in North America, arid regions and savanna in Northern Africa, or forests and desert scrub in Australia). Typical examples of seasonal variability are also captured in both estimates (e.g., the drying of the Iberian peninsula over the summer months, the wetting of India during the summer monsoon rains). However, the RMSD maps show regions and periods where the estimates differ substantially. For instance, JULES soil moisture and the satellite estimates agree relatively well over India for the SON months, while JULES is drier in the DJF and MAM months, but wetter in JJA. In North America, the satellite prediction tends to be drier for the first half of the year, with better agreement with JULES for the second half. The latitudinal location of the strong soil moisture gradients in the savanna regions of Central Africa agrees relatively well for the DJF and SON months, but for the other two seasons the shift between dry and wetter conditions occurs at higher latitudes in JULES. Along the east coast of South America, the moisture gradient between the savanna north of the Brazilian plateau and the rainforest in the south of the plateau is larger for JULES; for the same season the JULES soil moisture values for this rainforest and the rainforest in central Amazonia are closer than for the satellite-prediction, with lower soil moisture for the southern rainforest.

Differences between JULES and the satellite prediction are to be expected. On the one hand, the JULES model and associated input fields are imperfect [Blyth *et al.*, 2011, 2012]. To model soil moisture JULES is given no information other than the soil and vegetation types and the meteorological data. During periods and in regions where the main driving force of precipitation is no longer there to top up the estimated soil moisture, the internally calculated value can drift from the true soil moisture value, particularly if the soil properties are incorrectly characterised. On the other hand, there is no guarantee that the satellite observations can always capture the soil moisture signal (due to limitations of the observations and/or processing artifacts, or the impossibility of decoupling the soil moisture signal from other processes affecting the observations). There may also be issues regarding the ability of the statistical model to characterize the mapping between observations and modeled fields. Nevertheless, a large part of the differences are expected to be the consequence of the satellite information being added to the modeled fields. In that sense, we could say that the satellite observations are “correcting” potentially erroneous modeled fields towards values which are more consistent with the global relationships learned by the statistical model. This point will be illustrated in the next section with some examples.

4.2. Diagnosing model inconsistencies

A first example consists of preparing a soil moisture dataset with a large degree of consistency with the observations, followed by modifying this soil moisture estimate over specific regions and periods, and then testing whether the predicted soil moisture is able to correct the modified soil moisture field. The procedure was as follows:

- First a MLP regression model is trained with the 1993-1999 observations and JULES soil moisture fields. The resultant satellite-driven soil moisture fields will be called the “original” soil moisture and will be used as a soil moisture field that is to a large degree consistent with the satellite observations.
- These “original” soil moisture fields are then modified to simulate a soil moisture field that is not fully consistent with the satellite observations. This is called the “modified” soil moisture. The modifications consisted of (a) adding $\pm 20\%$ moisture in certain months over specific regions (shown as roughly rectangular areas in Fig. 6) (b) shifting the soil moisture values by 2 months in 4 selected African basins (colored in Fig. 6).

- A regression model (this time with the NIP technique) is trained to provide the mapping between the “modified” soil moisture fields (i.e., it is now used as if it would have been produced by a land surface model) and the observations (as was done to link JULES and the observations). The soil moisture produced by this NIP model is called the “predicted” soil moisture.

- The “predicted” soil moisture is compared with the “original” soil moisture to see whether the NIP model can correct the “modified” soil moisture and recover the “original” soil moisture at the modified regions and basins.

Figure 6 displays examples of a modified region and a basin spatially-averaged soil moisture time series. The “predicted” soil moisture is close to the “original” one, showing that the satellite-driven regression model is able to correct the modified soil moisture values toward values much closer to the original ones. The same happens for the example of a modified basin; the “predicted” soil moisture is in phase with the original soil moisture, meaning that adding information from the observations into the simulated land model field (the “modified” soil moisture) is able to correct the wrong soil moisture phasing. Figure 6 also shows histograms of the soil moisture differences to summarise the results over all modified regions and basins. The histograms of the differences between the “original” and “predicted” soil moistures are much narrower than those for the differences between the “original” and “modified” moisture.

A second test has been carried out by shifting the original JULES soil moisture time series back by one month. This modified soil moisture has been used to train a new NIP regression model. Figure 7 shows spatially-integrated soil moisture seasonal cycles for six selected basins. Due to missing observations and/or discarded pixels in the analysis (flooding or snow), for some basins the number of pixels discarded can be large, so the averages reported may differ from other published results. In all the basins analyzed here it can be seen that the soil moisture peaks one month after the JULES one-month-shifted estimates, showing that the satellite soil moisture prediction method can correct the erroneous phasing.

5. Discussion and conclusions

The methodology discussed here to jointly analyze modeled soil moisture and related observations was previously presented in Aires *et al.* [2005]. While in that study the methodology was tested with soil moisture fields from the surface modeling component of two atmospheric reanalyses (ERA-40 and NCEP/NCAR), it has been tested here with output from the JULES model. This is a land surface model dedicated to surface characterization, and for the runs considered in this analysis it was driven with one of the latest atmospheric forcing datasets for off-line hydrological and surface modeling. Greater consistency between the independent satellite observations and JULES soil moisture is observed here, with global correlations between the satellite product and model of 0.91 for JULES (compared with 0.73 and 0.83 respectively for NCEP/NCAR and ERA-40). This is likely a consequence of having a dedicated land surface model driven by state-of-the-art atmospheric forcings, but might also reflect differences in the ability of the models to simulate soil moisture.

While a neural network has been used here as in Aires *et al.* [2005], a second regression approach based on numerical integration of the soil moisture a posteriori probability (given a coincident set of observations) has also been tested

in this work. The objective was to get an approximate estimate of the errors introduced by changing the regression algorithm. Very consistent soil moisture fields were obtained by both regression methods (global correlation of 0.98), suggesting that limitations of the observations and uncertainty in the model soil moisture fields contribute more to prediction uncertainty than does the choice of regression algorithm. Although the predictions are relatively close, compared with the original JULES soil moisture, correlations for the neural network regression prediction are slightly higher (and RMSD slightly lower) for the different JULES land types, suggesting that the neural network is slightly better able to capture the JULES soil moisture features. The second regression algorithm provides not only the mean of the a posteriori probability distribution (used as the predicted soil moisture) but also the width of the distribution (reported as a standard deviation). We use this standard deviation as an estimate of the prediction error. The largest absolute prediction errors occur over some boreal regions; these estimated errors can be as large as $0.08 \text{ m}^3/\text{m}^3$. As expected, relative errors are large for those regions and periods with low soil moisture. Over these areas the prediction is more challenging as the soil moisture related signal in the observation is weak and the regression model has more difficulties to characterize the mapping between observations and modeled soil moisture.

Although the satellite prediction and JULES soil moisture agree relatively well, for some regions and periods clear differences exist. The problem is that evaluating soil moisture at large scales is very difficult. This is a problem not only for soil moisture, many other hydrological variables (e.g., precipitation, surface fluxes) are also difficult to evaluate due to the limitations of the present earth observation sensors. As mentioned in Section 1, large progress has been made exploiting the existing satellite microwave observations, but users must be very aware of the limitations of these products. There are also in situ observations, but for these the problems are related to: (1) the large mismatch between a point measurement and the spatially-integrated modeled or satellite derived soil moisture estimate; (2) sampling depth of the point measurement versus depth integration of the modeled or satellite estimate; and, (3) sparse coverage of the in situ soil moisture network, especially for earlier years such as the period analyzed here. Techniques to address these shortcomings are emerging [Miralles et al., 2010; Crow et al., 2012], but most evaluations are still done as a point-to-point comparison [e.g. Liu et al., 2012; Kolassa et al.; Albergel et al., 2012]. For instance, the JULES and satellite predicted soil moisture have been compared with the 1993-1994 in situ measurements reported in Prigent et al. [2005]. Notice that we are comparing a point measurement with a spatially integrated value for the $\sim 770 \text{ km}^2$ pixel area, and a 10 cm vertically integrated soil moisture (depth of JULES top layer) with in situ measurements for soil layers ranging from 5 to 20 cm depending on station. Only stations with at least 2 measurements per month were included. Correlations of around 0.5 were found for both estimates, in line with other reported model comparisons [e.g. Reichle et al., 2007]. Correlations were slightly larger for JULES than for the satellite prediction (0.51 and 0.48 respectively), while JULES and the satellite prediction correlated higher (0.78) over the same pixels. Notice that 95% of the matches were over C3 grass pixels, where prediction and modeled values already agree better than for other classes (see e.g. Figure 3).

Some of the discrepancies found between the satellite prediction and JULES are relatively large. This necessarily means that one or both of the estimates, either the model or the satellite prediction, is clearly different from the true soil moisture. Our tests modifying either the predicted soil moisture or the JULES estimates showed that our methodology can effectively change soil moisture towards more correct values. It can then be expected that some of these

large differences are the result of the satellite information modifying the modeled soil moisture fields towards more realistic values. In order to gain a further insight, we then compared our predictions with other reported soil moisture estimates. For instance, for the areas where we signalled model-observation discrepancies in Section 4.1 we inspected the GSWP-2 model estimates [Dirmeyer et al., 2006] and also found large differences between their individual models. Guo and Dirmeyer [2006] already reported similar findings in a detailed comparison of the GSWP-2 soil moisture predictions with ground observations, concluding that although the models could reasonably reproduce the phasing of the seasonal cycle and the interannual variation over the regions examined, the absolute values of soil moisture were poorly simulated by most models. More recently [Kolassa et al.] investigated the soil moisture predictions from the models ORCHIDEE and HTESSEL. We also compared our results with those predictions and again observed relatively large differences for some of the regions mentioned in Section 4.1. Therefore, it is difficult to reach firmer conclusions at present while global soil moisture estimates at the time and space integrations discussed here still differ considerably.

In principle our methodology could be considered not only as a mean of evaluating land surface models, but also as a way of building a satellite soil moisture data record. By construction our satellite-driven soil moisture has some dependence on the soil moisture fields used to set up the regression model. This means in practical terms that most of the a priori information used in the conversion of the observations into soil moisture estimates comes from the model. In that sense, model and predicted estimates are consistent; this should be regarded as a beneficial characteristic for certain applications (e.g., the model evaluation presented here, or assimilation of observations into models). When using external products, the degree of consistency between the model and the a priori assumptions used to derive the products could have an impact on the evaluation/assimilation [see e.g. Reichle et al., 2007]. At present the typical ill-posedness of remote inversion problems (i.e., there is not enough information in the observations to uniquely solve the inversion problem), accentuated by the limitations in the observations (not specifically dedicated to soil moisture), means that existing soil moisture products require some degree of a priori information. For instance, the merging of active and passive microwave observations reported in Liu et al. [2011] uses the absolute soil moisture values of a specific land surface model to adjust both products to a common reference. Therefore, comparison of this soil moisture product with other models can potentially be affected by this internal rescaling in the satellite product. If our satellite-driven estimates are considered as a soil moisture product, the same considerations apply, and the dependence of the estimates on the choice of land surface model needs to be assessed. In particular, similarity to satellite predictions derived from regression models set up using a different land surface model will be a strong indicator that our methodology effectively integrates the satellite information and the modeled fields to produce a more realistic soil moisture product. Indications of this were already presented in Jiménez et al. [2009] for a similar methodology applied to land surface heat fluxes, and this needs to be also confirmed for our soil moisture predictions.

Acknowledgments. We are thankful to Eleanor Blyth for providing useful insights for the evaluation of the JULES modeled fields. This research was partly funded by the European Union within the WATCH project (Contract 036946). We are grateful to three anonymous reviewers for their careful reading of the manuscript and constructive comments.

References

- Aires, F., Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 1. Network weights, *J. Geophys. Res.*, 109, D10303, 10.1029/2003JD004173, 2004.

- Aires, F., C. Prigent, and W. Rossow, Temporal interpolation of global surface skin temperature diurnal cycle over land under clear and cloudy conditions, *J. Geophys. Res.*, *109*, D04313, 10.1029/2003JD003527, 2004.
- Aires, F., C. Prigent, and W. Rossow, Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: 2. Global statistical relationships, *J. Geophys. Res.*, *110*, D11103, 10.1029/2004JD005094, 2005.
- Albergel, C., P. de Rosnay, G. Balsamo, L. Isaksen, and J. M. noz Sabeter, Soil Moisture Analyses at ECMWF: Evaluation Using Global Ground-Based In Situ Observations, *J. Hydromet.*, *13*, doi: 10.1175/JHM-D-11-0107, 2012.
- Armstrong, R., and M. Brodzik, Northern Hemisphere EASE-Grid weekly snow cover and sea ice extent version 3, *Tech. rep.*, Boulder, Colorado USA: National Snow and Ice Data Center. Digital media., 2005.
- Best, M. J., et al., The Joint UK Land Environment Simulator (JULES), model description - Part 1: Energy and water fluxes, *Geosci. Model Dev.*, *4*, doi:10.5194/gmd-4-677-2011, 2011.
- Bishop, M., *Neural networks for pattern recognition*, Oxford University Press, Inc., New York, 1995.
- Blyth, E., D. B. Clark, R. Ellis, C. H. ans S. Los, M. Pryor, M. Best, and S. Sitch, A comprehensive set of benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale, *Geosci. Model Dev.*, *4*, 255–269, doi:doi:10.5194/gmd-4-255-2011, 2011.
- Blyth, E., D. B. Clark, R. Ellis, and C. George, Using earth observation data to evaluate a land surface model in three Siberian catchments, *Boreal. Env. Res.*, *in press*, 2012.
- Boone, A., et al., The Rhône-Aggregation Land Surface Scheme intercomparison project: An overview, *J. Clim.*, *17*, 187–208, 2004.
- Clark, D. B., et al., The Joint UK Land Environment Simulator (JULES), model description - Part 2: Carbon fluxes and vegetation dynamics, *Geosci. Model Dev.*, *4*, doi:10.5194/gmd-4-701-2011, 2011.
- Crow, W. T., A. A. Berg, M. H. Cosh, A. Loew, B. P. Mohanty, R. Panciera, P. de Rosnay, D. Ryu, and J. P. Walker, Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products, *Reviews of Geophysics*, *50*(2), doi:10.1029/2011RG000372, 2012.
- Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki, GSWP-2: Multimodel analysis and implications for our perception of the land surface, *Bull. Amer. Meteor. Soc.*, *87*, 1831–1397, 2006.
- Dorigo, W., et al., The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, *Hydrol. Earth System Sci.*, *15*, doi:10.5194/hess-15-1675-2011, 2011.
- Dorigo, W. A., K. Scipal, R. M. Parinussa, Y. Y. Liu, W. Wagner, R. A. M. de Jeu, and V. Naeimi, Error characterisation of global active and passive microwave soil moisture datasets, *Hydrology and Earth System Sciences*, *14*(12), 2605–2616, doi: 10.5194/hess-14-2605-2010, 2010.
- Evans, K. F., S. J. Walter, A. J. Heymsfield, and G. M. McFarquhar, Submillimeter-wave cloud ice radiometer: Simulations of retrieval algorithm performance, *J. Geophys. Res.*, *107*, 2.1–2.21, 2002.
- Francis, R., et al., The ERS-1 spacecraft and its payload, *ESA Bulletin*, *65*, 27–48, 1991.
- Frison, P. L., and E. Mougin, Monitoring global vegetation dynamics with ERS-1 wind scatterometer data, *Int. J. Remote Sens.*, *17*, 3201–3218, 1996.
- Guo, Z., and P. A. Dirmeyer, Evaluation of the second Global Soil Wetness Project soil moisture simulations: 1. Intermodel comparison, *J. Geophys. Res.*, *111*, D22S02, 10.129/2006JD07233, 2006.
- Gutman, G. G., On the use of long-term global data of land reflectances and vegetation indices from the advanced very high resolution radiometer, *J. Geophys. Res.*, *104*, 6241–6255, 1999.
- Hollinger, J., R. Lo, G. Poe, R. Savage, and J. Pierce, Special Sensor Microwave/Imager user guide, *Tech. rep.*, Nav. Res. Lab., Washington, D.C, 1987.
- James, M., and S. V. Kalluri, The Pathfinder AVHRR land data set: An improved coarse resolution data set for terrestrial monitoring, *Int. J. Remote Sens.*, *15*, 3347–3364, 1994.
- Jiménez, C., C. Prigent, and F. Aires, Toward an estimation of global land surface heat fluxes from multisatellite observations, *J. Geophys. Res.*, *114*, D06305, doi:10.1029/2008JD011392, 2009.
- Kerr, Y. H., et al., The SMOS mission: new tool for monitoring key elements of the global water cycle, *Proceedings of the IEEE*, *98*(5), 666–687, 2010.
- Kolassa, J., F. Aires, J. Polcher, C. Prigent, C. Jimenez, and J. Pereira, Soil moisture retrieval from multi-instrument observations: information content analysis and retrieval methodology, *J. Geophys. Res.*, *in press*, doi:10.1029/2012JD018150.
- Liu, Y., W. Dorigo, R. Parinussa, R. de Jeu, W. Wagner, M. McCabe, J. Evans, and A. van Dijk, Trend-preserving blending of passive and active microwave soil moisture retrievals, *Remote Sensing of Environment*, *Remote Sens. Envir.*, *123*, doi: 10.1016/j.rse.2012.03.014, 2012.
- Liu, Y. Y., R. M. Parinussa, W. A. Dorigo, R. A. M. D. Jeu, W. Wagner, A. I. J. M. van Dijk, M. F. McCabe, and J. P. Evans, Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, *Hydrol. Earth System Sci.*, *15*, doi:10.5194/hess-15-425-2011, 2011.
- Miralles, D., C. Crow, and M. Cosh, Estimating Spatial Sampling Errors in Coarse-Scale Soil Moisture Estimates Derived from Point-Scale Observations, *J. Hydromet.*, *11*, doi: 10.1175/2010JHM1285.1, 2010.
- Owe, M., R. de Jeu, and J. Walker, A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index, *IEEE Trans. Geosci. Remote Sens.*, *39*, 1643–1654, 2001.
- Papa, F., C. Prigent, C. Jimenez, F. Aires, and W. Rossow, Interannual variability of surface water extent at global scale, 1993–004, *J. Geophys. Res.*, *115*, D12111, doi:10.1029/2009JD012674, 2010.
- Prigent, C., F. Aires, W. Rossow, and A. Robock, Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: Relationship of satellite observations to in situ soil moisture measurements, *J. Geophys. Res.*, *110*, D07110, 10.1029/2004JD005087, 2005.
- Prigent, C., F. Aires, and W. Rossow, Land surface microwave emissivities over the globe for a decade, *Bull. Amer. Meteor. Soc.*, *87*, 1573–1584, 2006.
- Reichle, R., R. Koster, J. Dong, P. Liu, S. P. P. Mahanama, E. G. Njoku, and M. Owe, Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), *J. Geophys. Res.*, *112*,D09108, doi:10.1029/2006JD008033, 2007.
- Rossov, W., and R. Schiffer, Advances in understanding clouds from ISCCP, *Bull. Amer. Meteor. Soc.*, *80*(11), 2261–2287, 1999.
- Rydberg, B., P. Eriksson, S. A. Buehler, and D. P. Murtagh, Non-gaussian bayesian retrieval of tropical upper tropospheric cloud ice and water vapour from odin-smr measurements, *Atmos. Meas. Tech.*, *2*, 621–637, 2009.
- Scipal, K., T. Holmes, R. de Jeu, V. Naeimi, and W. Wagner, A possible solution for the problem of estimating the error structure of global soil moisture data sets, *Geophys. Res. Lett.*, *35*, doi:10.1029/2008gl035599, 2008.
- Uppala, S., et al., The ERA-40 Reanalysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012, 2005.
- Wagner, W., G. Lemoine, and H. Rott, A method for estimating soil moisture from ERS scatterometer and soil data, *Remote Sens. Envir.*, *70*, 191–207, 1999.
- Weedon, R., et al., Creation of the WATCH Forcing Data and its use to assess global and regional reference crop evaporation over land during the twentieth century, *J. Hydromet.*, *q2*, 823–848, 2011.

C. Jiménez, Laboratoire d'Etudes du Rayonnement et de la Matière en Astrophysique, Centre National de la Recherche Scientifique, Observatoire de Paris, Paris, France (carlos.jimenez@obspm.fr)

Table 1. Statistics of the 1993-1999 soil moisture difference between the model (JULES) and the MLP and NIP algorithms driven by the satellite data, for different JULES land surface types. Values shown are the root means square difference (as percentage of the averaged JULES soil moisture for each land type, and in absolute values (m^3/m^3) in brackets) and the Pearson's correlation coefficient (in squared brackets).

	NIP-JULES	MLP-JULES	NIP-MLP
Broad-tree	16.6(0.054)[0.74]	16.4(0.053)[0.75]	6.0(0.019)[0.96]
Nleaf-tree	26.1(0.072)[0.48]	25.8(0.071)[0.50]	10.1(0.028)[0.78]
C3-grass	29.0(0.060)[0.82]	27.5(0.057)[0.84]	13.4(0.027)[0.95]
C4-grass	37.8(0.067)[0.85]	36.7(0.065)[0.85]	12.0(0.021)[0.97]
Shrubs	51.1(0.068)[0.75]	48.7(0.065)[0.76]	19.3(0.026)[0.94]
Bare-soil	131.5(0.046)[0.65]	128.7(0.045)[0.67]	81.5(0.028)[0.81]
All-types	30.8(0.057)[0.92]	30.0(0.055)[0.91]	13.5 (0.025)[0.98]

Table 2. Statistics of the 1993-1999 soil moisture difference between the model (JULES) and the NIP algorithm driven by different satellite data (No-XXX meaning that that input data XXX were removed) for different JULES land surface types. Values shown are the root means square difference (as percentage of the averaged JULES soil moisture for each land type) and the Pearson's correlation coefficient (in squared brackets).

	No-ERS	No-AVHRR	No-ISCCP	No-SSM/I
Broad-tree	17.5[0.70]	17.0[0.71]	17.5[0.70]	16.8[0.72]
Nleaf-tree	27.4[0.30]	26.2[0.49]	29.3[0.42]	26.9[0.43]
C3-grass	32.0[0.76]	30.7[0.80]	31.3[0.77]	29.5[0.82]
C4-grass	41.3[0.80]	37.6[0.85]	41.6[0.81]	39.4[0.83]
Shrubs	52.2[0.75]	63.7[0.71]	47.1[0.73]	51.1[0.76]
Bare-soil	129.5[0.65]	139.1[0.64]	122.3[0.63]	139.5[0.63]
All-types	33.0[0.90]	32.4[0.90]	34.0[0.88]	31.7[0.90]

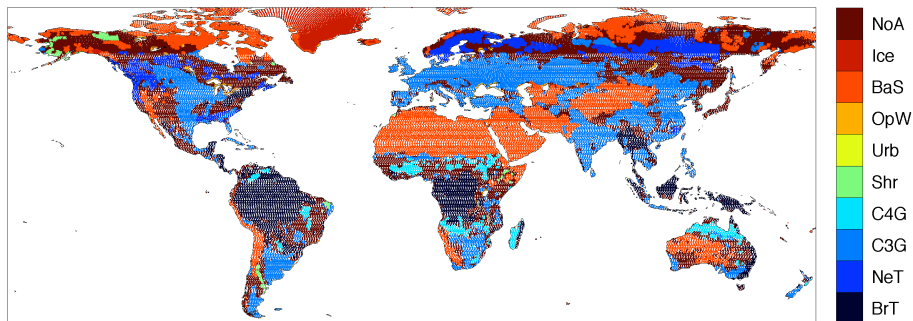


Figure 1. Geographical distribution of land types used in this study. A model cell is assigned a JULES land type if at least 50% of the cell surface corresponds to that land type. The land types are: broadleaf tree (BrT), needleleaf tree (NeT), C3 grass (C3G), C4 grass (C4G), shrubs (Shr), urban (Urb), open water (OpW), bare soil (BaS), ice (Ice), and no assigned land type (NoA).

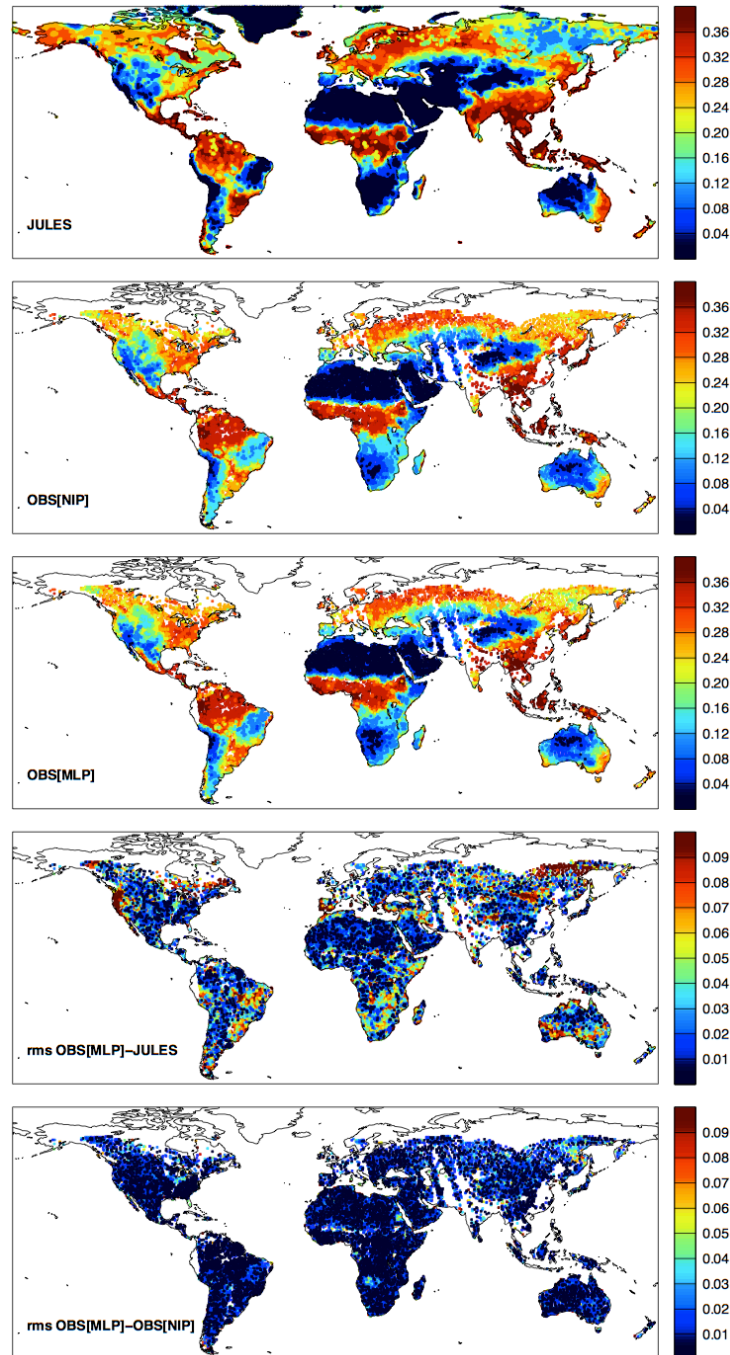


Figure 2. August 1998 soil moisture and differences between the model and predictions (m^3/m^3). From top to bottom: JULES soil moisture (first); NIP (second); MLP (third); difference between MLP and JULES soil moisture (fourth), and; difference between MLP and NIP (fifth). Missing values in the MLP and NIP predictions and differences are related to missing observational values precluding the prediction of soil moisture.

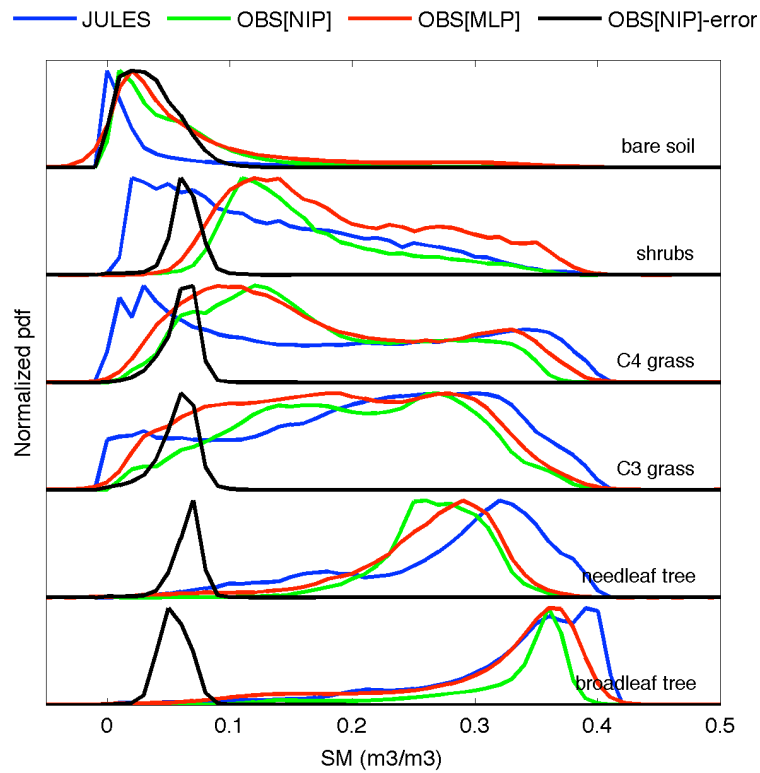


Figure 3. Soil moisture histograms for the different JULES land types. Plotted for JULES (blue), the NIP (green) and MLP (red) algorithms driven by the satellite data. The black lines display the histogram of the NIP prediction errors for each land type.

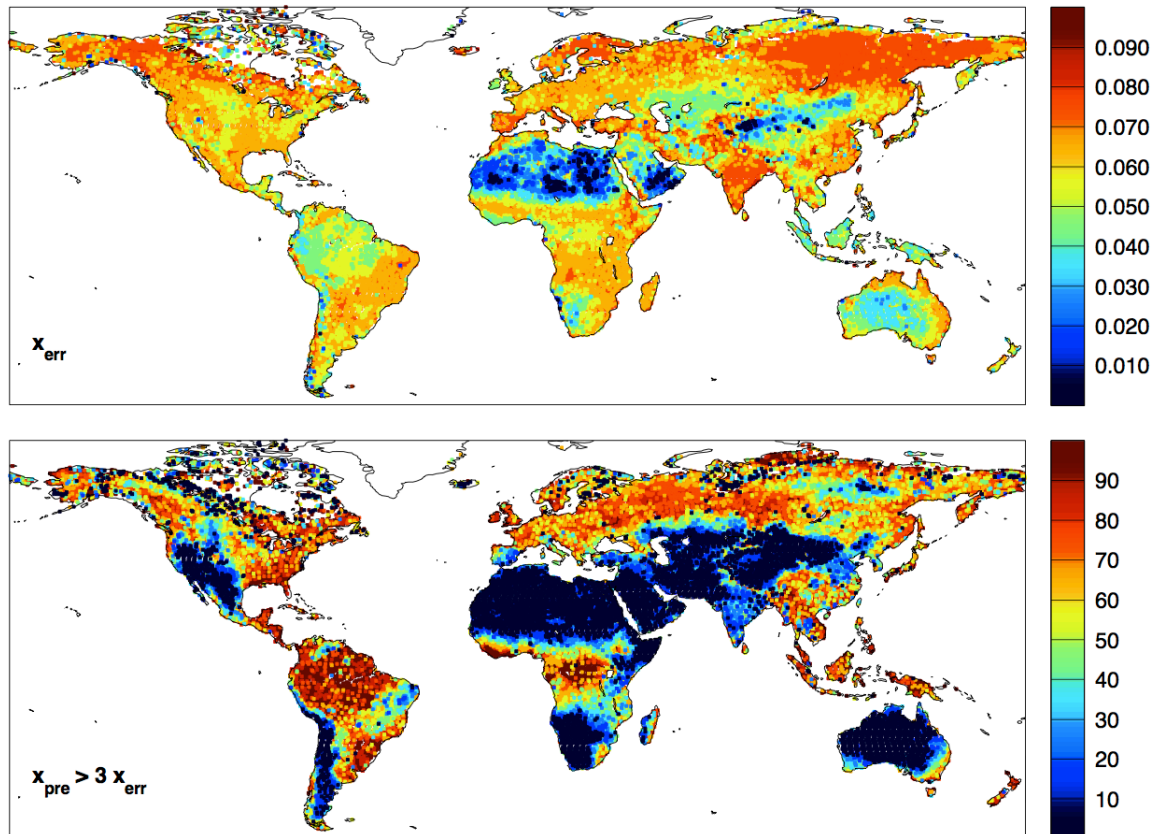


Figure 4. Statistics of prediction errors (calculated over all 1993-99 soil moisture predictions). Top: median prediction error (median value of the distribution of errors at each single pixel) (m^3/m^3). Bottom: for all predictions for a given pixel, percentage of cases where soil moisture estimates are at least 3 times larger than the prediction error.

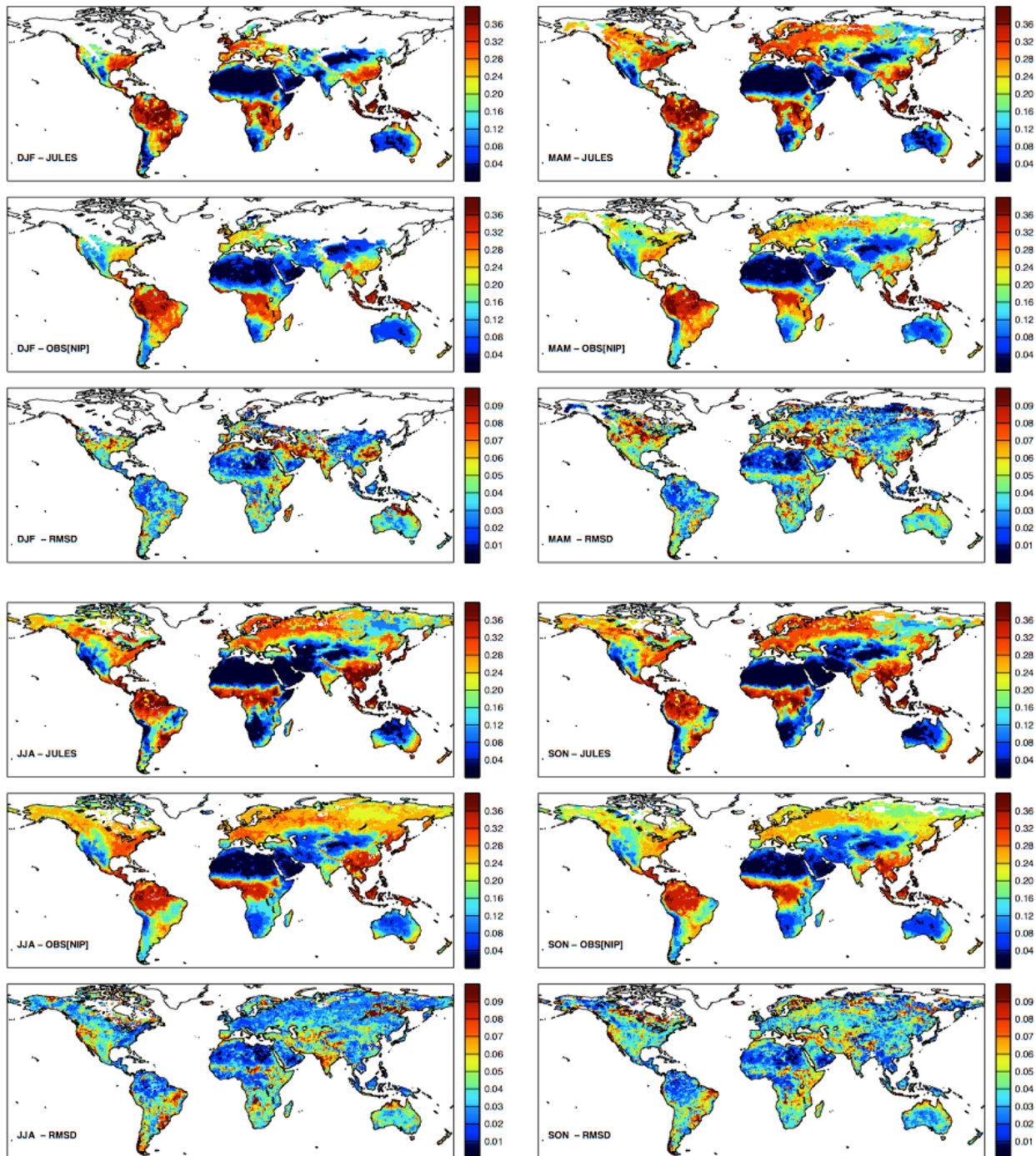


Figure 5. Seasonal JULES and satellite-driven soil moisture (top-left for Dec-Jan-Feb, top-right for Mar-Apr-May, bottom-left for Jun-Jul-Aug, and bottom-right for Sep-Oct-Nov) in m^3/m^3 and difference with the JULES soil moisture (given as the root square mean difference in m^3/m^3). Statistics computed over 1993-1999.

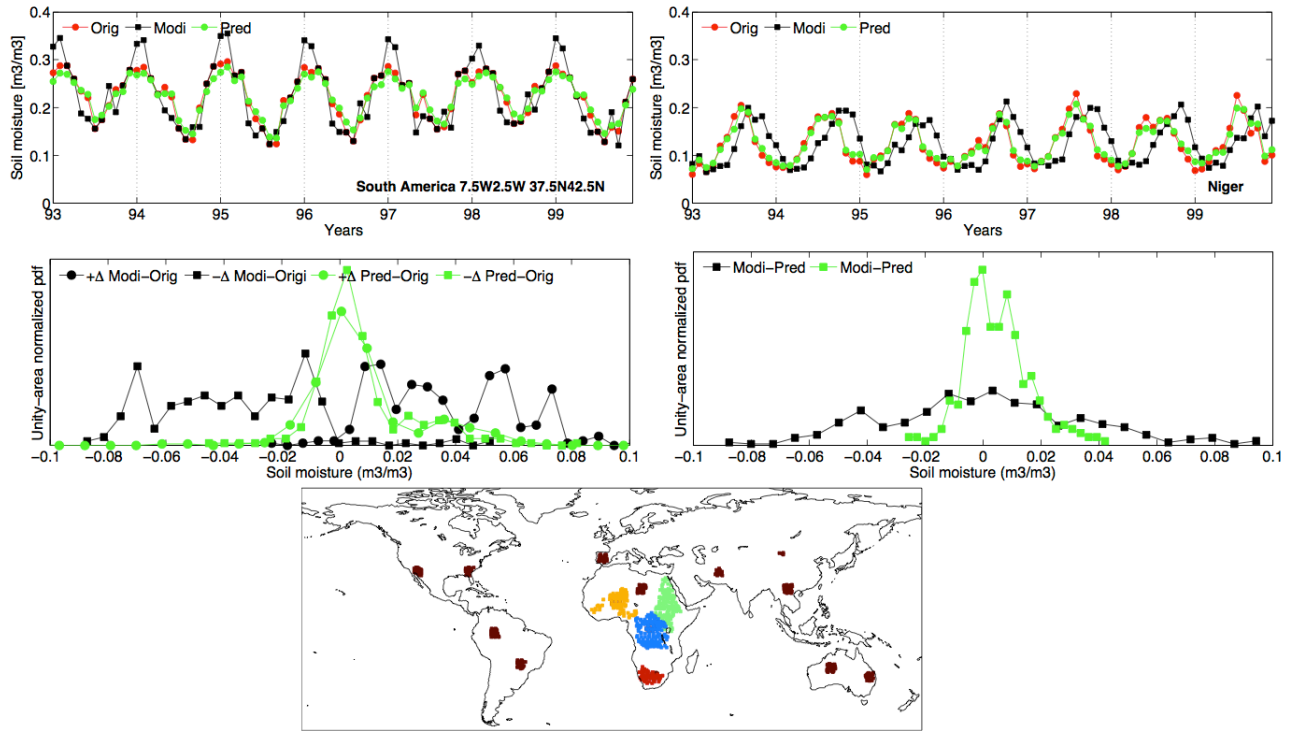


Figure 6. Summary of model diagnosing tests. Top-Left: example of original (Orig, red), modified (Modi, black), and predicted (Pred, green) soil moisture for a region in South America. Middle-left: histograms of the all-regions differences between the the predicted and the original soil moisture (green) and the modified and original (black) (two curves, the first calculated for months when a positive perturbation is added to the original soil moisture (+ Δ , circles), the second for negative perturbations ($-\Delta$, squares). Top-right: example for the Niger basin showing correction of phasing of signal; lines as in top-left panel. Middle-right: histograms of the differences; lines as in middle-left panel, but now with the histograms calculated over all months. Bottom: geographical location of disturbed regions (small areas) and basins (large areas in Africa).

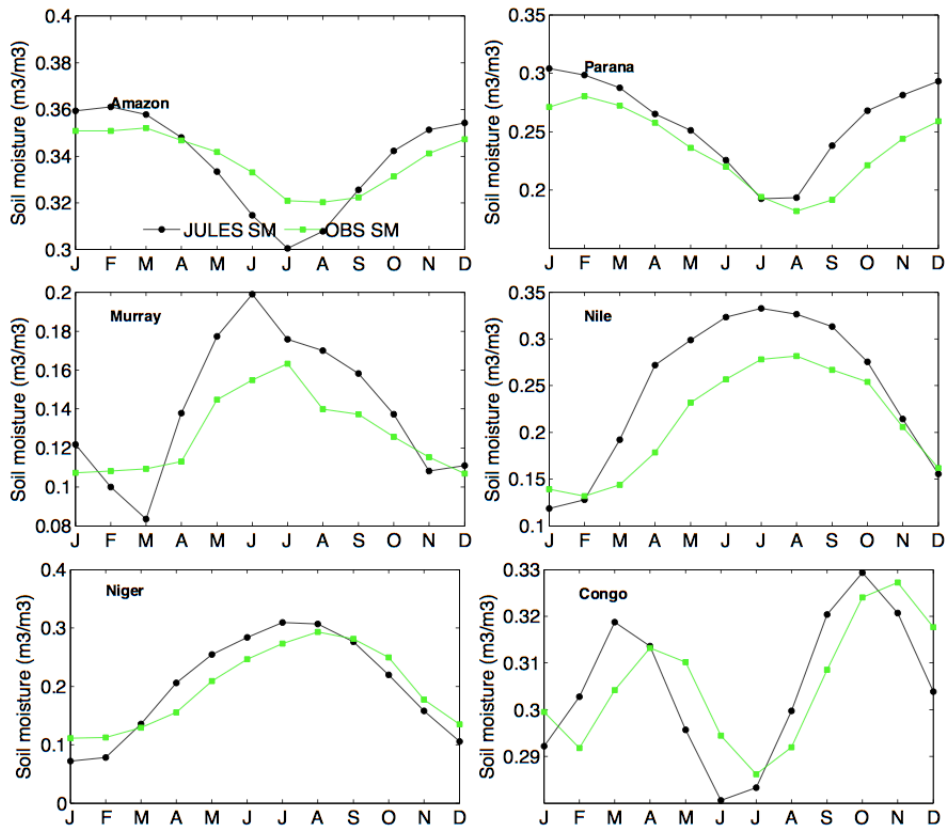


Figure 7. Seasonal cycle of soil moisture for selected basins, averaged over 1993–1999. Shown are the JULES soil moisture (shifted by one month) used to build the regression database (black), and the NIP satellite-driven prediction (green).