Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: 2. Global statistical relationships

F. Aires
Laboratoire de Météorologie Dynamique, Institut Pierre Simon Laplace (IPSL)/CNRS, Université Pierre et Marie Curie, Paris, France

C. Prigent
Observatoire de Paris Laboratoire d'Étude du Rayonnement et de la Matière en Astrophysique (LERMA)/CNRS, Paris, France

W. B. Rossow
NASA Goddard Institute for Space Studies, New York, New York, USA

Received 4 June 2004; revised 6 October 2004; accepted 15 March 2005; published 1 June 2005.

[1] In part 1 of this study (Prigent et al., 2004), in situ measurements were used to analyze and describe the sensitivities of satellite measurements (i.e., active and passive microwave observations and surface skin temperature diurnal cycle amplitude) to the soil moisture variations to describe the complex relationships that exist between them. Soil moisture was considered in the first 10-cm layer on a 0.25° equal-area grid and a monthly timescale. In this study, the lessons from the first paper are exploited to document the sensitivity of the satellite data to the global large-scale variations of soil moisture. A statistical model based on neural networks is developed to link the satellite observations and soil moisture estimates. Given the lack of available in situ soil moisture measurements on a global basis, National Centers for Environmental Prediction (NCEP) and European Centre for Medium-Range Weather Forecasts (ECMWF) soil moisture reanalyses are used as a realistic global indicator of soil moisture. As a consequence, the statistical model cannot be considered as a retrieval scheme per se, but it shows the feasibility of such an approach. It also quantifies the information content that can be expected from the satellite observations. Applications of such a statistical model include checking the consistency of surface model, and as the basis for variational assimilation of satellite observations into a numerical surface model.

Citation: Aires, F., C. Prigent, and W. B. Rossow (2005), Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: 2. Global statistical relationships, *J. Geophys. Res.*, *110*, D11103, doi:10.1029/2004JD005094.

1. Introduction

[2] In part 1 [*Prigent et al.*, 2005], the sensitivity of active and passive microwave observations and surface skin temperature diurnal cycle amplitude to soil moisture variability is analyzed based on in situ measurements. Observations from ERS and SSM/I were used for this purpose. These two instruments were not designed to monitor soil moisture. The frequencies and the spatial/temporal resolutions have not been specifically chosen for continental studies. Future missions like SMOS or HYDROS are dedicated to soil moisture retrieval, but they will not be launched for several years. The study showed that the link between satellite observations and soil moisture is complex and can be due to an indirect relationship through the correlation between vegetation and soil moisture. Soil moisture is defined as the volumetric percent of water in the first 10 cm and a 0.25° equal-area grid, and a monthly timescale. In particular, the differences of polarization of passive microwave (SSM/I) are more sensitive to vegetation than to soil moisture itself. Depending on the location, the vegetation can be correlated or anti-correlated with soil moisture and, as a consequence, some indirect and sometimes conflicting information from satellite observations can be used to monitor soil moisture if the relationship is locally adjusted. This explains why soil moisture retrieval algorithms can work locally when a priori information about the particular location helps constrain the problem and avoid external large-scale variabilities. Although a few studies have examined the use of NDVI data in addition to infrared [e.g., *Goetz*, 1997] or passive microwave for specific regions, to our knowledge, a systematic analysis of the available observations in the various frequency domains has not been conducted on a global basis.

[3] Soil moisture retrievals from satellite observations are traditionally based on radiative transfer calculations. The

0148-0227/05/2004JD005094$09.00

radiative transfer model describes the link between the soil moisture and the satellite measured radiation, and algorithms are developed to invert this relationship, i.e., to provide an estimate of the soil moisture that corresponds to a given measured radiance. To our knowledge, no radiative transfer model is able to correctly reproduce the satellite observations over land on a global basis. Two main reasons explain this failure. First, many of the inputs for the radiative transfer model (e.g., soil texture, roughness, and moisture, vegetation characteristics) are not available on a global basis with the required accuracy. Second, the radiative transfer models are still unable to reproduce the complexity and variety of land surface mechanisms. Classic inversion schemes based on radiative transfer models can be successful in specific regions where additional information is available to "tune" out those local dependencies that are not explicitly accounted for, but extension to global applications is still questionable. In addition, using various satellite sources requires having adequate and consistent radiative transfer models for all wavelength ranges.

[4] A global soil moisture retrieval method requires: (1) multiple sources of information from various wavelengths to account for the different surface parameters involved and, in particular, to disentangle the vegetation and soil moisture effects, and (2) nonlinearity to account for the situation-dependence (the relationships between satellite observations and surface parameters can change with location) and to better exploit parameters inter dependencies. Neural network (NN) methods are a natural candidate to provide such capabilities: They are well adapted to benefit from the synergy between multiple instruments, and they are nonlinear by design.

[5] In this study, we focus on the variability of soil moisture at a global scale. The major objective of this analysis is to assess the potential of satellite observations for monitoring the large-scale variability of soil moisture. Given the scarcity of in situ measurements and the limited climate regimes they represent, National Centers for Environmental Prediction (NCEP) and European Centre for Medium-Range Weather Forecasts (ECMWF) soil moisture reanalyses are used as realistic global indicators for soil moisture. A statistical model based on NN is developed to link the satellite observations and the soil moisture estimates. This statistical model cannot be considered as a retrieval scheme per se. To develop a soil moisture retrieval scheme, we should have either a comprehensive radiative transfer model to calculate the satellite-observed radiances from a global land surface properties and soil moisture data set (i.e., physical retrieval) or a matched global set of satellites observations and soil moisture measurements (i.e., empirical retrieval). However, we lack both a comprehensive radiative transfer model and a global soil moisture data set, so to assess the potential of the available satellite observations to constrain estimates of soil moisture, we train a NN to relate the satellite observations to the soil moisture produced by the NCEP and ECMWF reanalyses. We will use in this paper the term prediction instead of retrieval, and errors refer to the prediction of the reanalyses soil moisture estimates from the satellite measurements. These soil moisture values are produced by land surface models coupled to atmospheric models with meteorological properties constrained by observations of atmospheric temperature, humidity, horizontal winds, and surface temperature and pressure.

[6] In this situation, the trained NN represents a statistical model linking satellite observations to estimated soil moisture. If the NN is able to find robust relationships between these two quantities, this link can be used to check the consistency of the reanalyses products with the satellite observations. The statistical link can also be exploited as an additional constraint in a variational assimilation scheme. This is a necessary first step toward the evaluation of existing global soil moisture data sets, the development of new land surface models, or the assimilation of satellite observations into a numerical land surface models. The results of this study should be, first, an improved understanding of both remote sensing observations and model outputs by quantifying the information content that can be expected from the various satellite observations. Second, this work should produce a methodology/data sets to be used as input and diagnostics for land surface model evaluations and comparisons.

[7] Section 2 presents the NCEP and ECMWF soil moisture reanalyses used in this study. Global linear statistics between the satellite observations and the soil moisture model outputs are examined and interpreted in section 3. The NN statistical inverse model is described in section 4 and potential applications are discussed in section 5. Conclusions and perspective are presented in section 6.

2. Soil Moisture From NCEP and ECMWF Reanalysis

[8] In this study, the standard reanalysis soil moisture products from NCEP and ECMWF are used. In numerical weather prediction (NWP) models, land surface schemes have been added, essentially to predict latent and sensible fluxes. This is valid for real time forecast and for the reanalysis mode since the same model is used in both cases. Soil moisture is a by-product of these calculations. The NWP models are fed several times daily by atmosphere observations. For NCEP, in reanalysis mode, the land surface schemes are forced with model generated precipitation and radiation, whereas today, in real time mode, observed precipitation is often added.

[9] For considerations of calculation speed, the reanalysis schemes use land surface models less sophisticated than the more advanced models available today. In the last 10 years, land surface modeling has triggered a lot of developments and significant progress has been made that has not yet benefited the reanalysis models. However, as the model intercomparisons tend to prove (Atmospheric Model Intercomparison Project (AMIP) or Global Soil Wetness Project (GSWP)), up-to-date models still have serious problems in comparison to actual soil moisture measurements, and it is very difficult to decide among them. Thirty model outputs have been compared during AMIP, more than 10 during GSWP. Which one should be selected? In this context, the NWP reanalysis have been chosen because they are easily and widely accessible products; their problems have been identified and are rather well documented.

Table 1. Linear Correlation Coefficient Between the Soil Moisture and NDVI Values and the Satellite-Derived Variables

Variable	NCEP Soil Moisture	ECMWF Soil Moisture	NDVI Vegetation
Passive MW SSM/I e19V-H	−0.26	−0.27	−0.33
Passive MW SSM/I e37V-H	−0.12	−0.12	−0.15
Active MW ERS small ang	0.58	0.55	0.63
Active MW ERS large ang	0.56	0.48	0.48
IR normalized *Ts* amplitude	−0.69	−0.58	−0.74
NVDI vegetation	0.65	0.58	1

[10] The NCEP/NCAR reanalysis covers 40 years [*Kalnay et al.*, 1996]. It uses a multi-layer soil vegetation hydrology model [*Mahrt and Pan*, 1984; *Pan and Mahrt*, 1987] that is adjusted to the *Mintz and Serafini* [1992] climatology. Soil moisture is derived from the 6-hour forecast, not from observations "and should be used with caution... but generally contains useful information" [*Chen and Mitchell*, 1999]. In numerical weather models, it is common to control the soil moisture so that it does not drift to unrealistic values: This is what is called nudging. There are different method to soil moisture nudging in the models. The most basic one consists in resetting the soil moisture to a climatological value on a regular basis. In the NCEP/NCAR reanalysis, the nudging toward the climatology leads to too large a soil moisture annual cycle amplitude and reduced interannual variability [*Roads et al.*, 1999; *Chen and Mitchell*, 1999]. Efforts are now being directed toward exploring alternatives to the climatological nudging and using observed precipitation assimilation.

[11] The ECMWF 40-year reanalysis (ERA 40) [*Simmons and Gibson*, 2000] uses a recent version of a land surface scheme [*Van den Hurk et al.*, 2000; *Viterbo and Beljaars*, 1995] and an optimal interpolation scheme [*Douville et al.*, 2000]. The emphasis of the land surface scheme is on the correct modeling of the long-time timescale of water exchanges, in contrast to most studies that focus on the short timescale. The model has four layers for moisture and temperature and has been tested against measurement campaigns in different regimes (United States, Netherlands, and Brazil) [*Viterbo and Beljaars*, 1995; *Van den Hurk et al.*, 2000; *Betts et al.*, 2003a, 2003b].

[12] The monthly mean soil moisture values are selected for the upper layer (10 cm and 7 cm in the NCEP and ECMWF reanalysis, respectively). The original NCEP and ECMWF data are mapped on Gaussian grids. To match the reanalysis data with the satellite observations, they are interpolated, in this study, onto an $0.25^\circ \times 0.25^\circ$ equal area grid, using distance-weighted averages.

[13] The satellite observations analyzed here cover a large portion of the electromagnetic spectrum: (1) the passive microwave SSM/I emissivities between 19 and 85 GHz (i.e., from 1.58 cm to 0.35 cm in wavelength); (2) ERS-1 active instrument backscattering coefficient at 5.25 GHz (wavelength = 5.71 cm); (3) the diurnal amplitude of the surface skin temperature, derived from the thermal infrared observations from both the NOAA polar orbiters and the geostationary meteorological satellites; (4) the NDVI derived from the AVHRR visible (0.58–0.68 μ m) and near-infrared (0.73–1.1 μm) reflectances. See part 1 of this study *Prigent et al.* [2005] for a detailed description of the satellite observations.

3. Global Statistical Analysis of the Monthly Data Sets

[14] The following statistics are determined for the monthly means in a 2-year data set that includes the satellite-derived variables and the NWP soil moisture estimates, all gridded on a $0.25^\circ \times 0.25^\circ$ equal area grid. The satellite data have been described in part 1 of this study [*Prigent et al.*, 2005].

[15] The linear correlations have been calculated globally between the satellite variables and the NCEP and ECMWF surface soil moistures for the 2 years, and results are reported in Table 1. Global mean linear correlations between active microwave responses, *Ts* amplitude, and the NWP soil moisture are significant and have the expected signs [see *Prigent et al.*, 2005], with rather similar values for both NWP reanalyses. In contrast, passive microwaves are not well correlated with NWP soil moisture, with the sign of the correlation opposite to the expectation. For all satellite variables, except the active microwave measurements at low incidence angles, the correlation is stronger with the vegetation index (NDVI) than with the NWP soil moistures. A strong correlation also exists between both NWP soil moistures and the NDVI.

[16] When analyzing the in situ soil moisture measurements with the satellite variables, we showed that the relationship between soil moisture and vegetation is partly responsible for the correlation between the satellite observations and soil moisture [*Prigent et al.*, 2005]. In particular, the effects of vegetation on the measurements changes their relationship with soil moisture, from region to region. In this global study, the linear correlations between the monthly-mean NCEP soil moisture and NDVI vegetation index have been calculated over the 2 years for each location and are presented in Figure 1. Strong regional patterns appear related to climate regimes. Midlatitudes are characterized by wet winters and dry summers that translate into a highly negative correlation between vegetation and soil moisture (e.g., Europe and eastern United States). In semi-arid tropical regions, vegetation growth is tightly related to the onset of precipitation, leading to a high positive correlation between soil moisture and vegetation. In desert areas in North Africa, the NWP estimates are often fixed to a nominal value, so the correlations should be considered with caution.

[17] The correlation between the ERS scatterometer observations for low incidence angles and soil moisture is generally strongly positive (Figure 1c), as expected, except in desert areas where the reanalysis values might not be representative (see remark above). The correlation between the normalized amplitude of the *Ts* diurnal cycle and the soil moisture from the model is strongly negative (Figure 1b), especially in regions where the correlation between soil moisture and vegetation is also high. At midlatitudes where the soil moisture is anti-correlated with vegetation, the linear correlation between the satellite IR derived information and the soil moisture decreases in absolute value. The amplitude of the *Ts* diurnal cycle is expected to decrease both with increasing soil moisture and with increasing

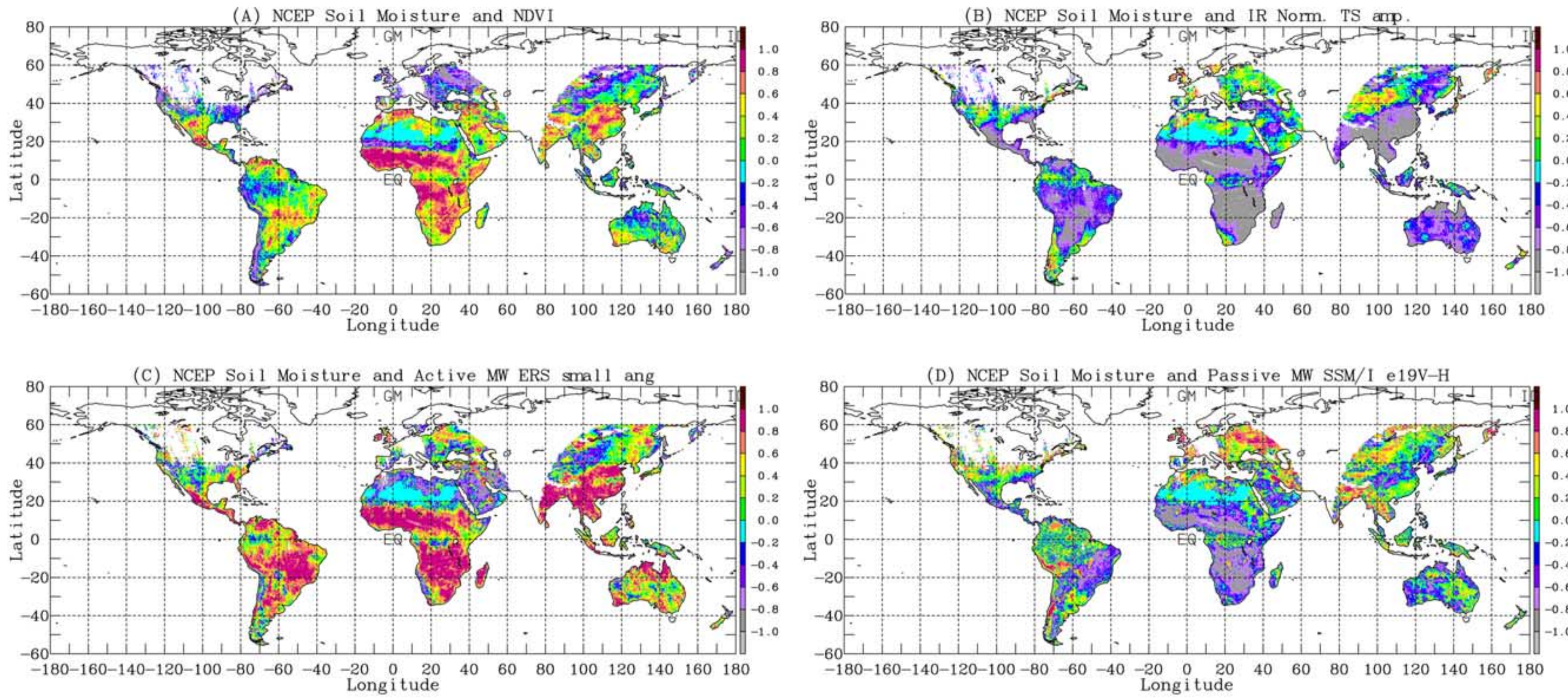

Figure 1. Correlation maps of NCEP soil moisture and (a) NDVI, (b) amplitude of the normalized diurnal cycle of the skin temperature, (c) ERS scatterometer observation, and (d) SSM/I Passive microwave emissivity at 19 GHz.

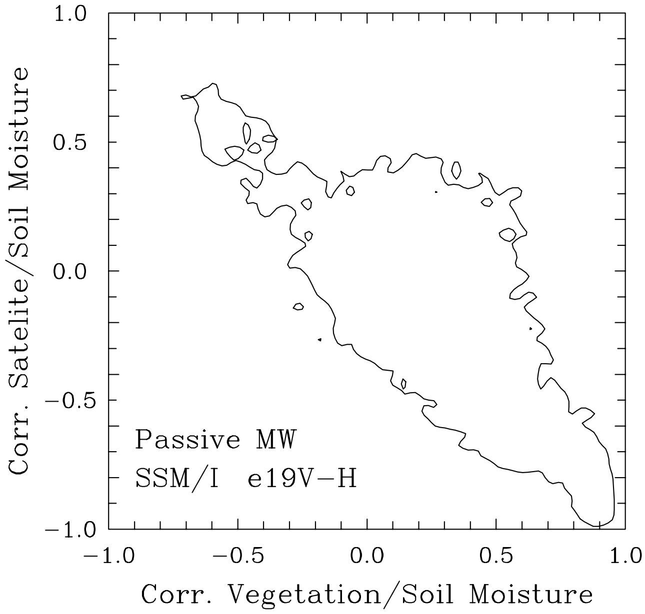

Figure 2. Linear correlation between passive MW SSM/I polarization difference at 19 GHz and the NCEP soil moisture versus the linear correlation between vegetation (NDVI) and the NCEP soil moisture.

vegetation density: When these two variables are correlated, the anti-correlation between soil moisture and the *Ts* amplitude is particularly strong, but when these two variables are anti-correlated, they have opposite effects on the *Ts* amplitude leading to a decreased correlation (in absolute value) between the *Ts* amplitude and the soil moisture. Figure 1d shows the correlation between NCEP surface soil moisture and the passive microwave emissivity polarization difference at 19 GHz. A positive correlation is expected between these two variables: With increasing soil moisture, the emissivity polarization difference should increase. In midlatitude regions (e.g., in Europe), a positive correlation is observed between soil moisture and the passive microwave observations. However, a strong negative correlation prevails in semi-arid regions, especially in Africa. This behavior is similar to what has been described with in situ soil moisture measurements in the first part of this study. The passive microwaves react primarily to the vegetation density: When the vegetation density and the soil moisture are negatively correlated, the passive microwaves vary as expected with soil moisture only because of the opposing effects of soil moisture and vegetation on the signal. Similar to *Prigent et al.*'s [2005] Figure 3, Figure 2 shows that the linear correlation between the soil moisture and the passive microwave signal depends on the linear correlation between the soil moisture and the vegetation density. Very similar results are observed with ECMWF (not shown).

[18] The above analysis confirms the interpretation of the relationships between the in situ soil moisture measurements and the satellite observations described in the first part of this study [*Prigent et al.*, 2005]. Yet the soil moisture estimates used in this analysis are very different in nature: Global model reanalysis results in contrast to local in situ measurements. Thus the derived interpretation is verified at various scales. In addition, it tends to validate the use of NWP model data that have similar fundamental behavior as the in situ measurements.

[19] Whatever the satellite observations, the relationship between the satellite-derived information and the soil moisture is complex and shows strong regional variations. Most of the time, this relation is not direct but linked to the soil moisture/vegetation relationship. For a given region, a low correlation (in absolute value) is observed between a particular satellite measurement and the soil moisture whereas other satellite measurements provide a stronger relationship to the soil moisture. The opposite might prevail in another region. Using more than one source of satellite data makes it possible to benefit from these complementarities. A neural network analysis is selected to account for these intricate relationships between the satellite observations and the soil moisture and to exploit the different regional sensitivities of the merged satellite observations to the soil moisture variations.

4. Integrated Analysis of Spectral Variations Using a Neural Network Technique

[20] In this section, we evaluate the possibility of inferring soil moisture information from satellite measurements. The method is based on a neural network (NN) analysis. Traditionally, for remote sensing applications, this statistical model requires a data set of satellite measurements (simulated or observed) collocated with a data set of the variables to be predicted (the soil moisture in this application). The role of the NN is then to reproduce the relationships inherently described by these two collocated data sets. Unfortunately, in situ measurements are scarce, and the existing measurements do not represent the large spatial and temporal diversity existing over the globe. In addition, outputs from global surface models are not sufficiently reliable and validated to be used with complete confidence. As a consequence, the goal of the present analysis cannot be the development of a definite retrieval algorithm. Instead, a sensitivity analysis of the relations of the large-scale soil moisture variability to the various sources of available satellite observations is performed at a global scale: This analysis can be considered as a feasibility study for the development of a retrieval algorithm. The NN is trained to predict the soil moisture estimates; "error" refers to the prediction errors of the reanalysis soil moisture, not error in the "true" soil moisture which is unknown. It also allows for the characterization and selection of the pertinent satellite information for soil moisture retrieval. This is an important step for the assimilation of satellite data in land surface models or for the development of a soil moisture retrieval strategy for the next generation of satellite instruments. Two direct applications of the NN model will be described in section 5.

4.1. Method

[21] The correlations between satellite observations and soil moisture content are complex and partly produced by the correlation that exists between vegetation and soil moisture. At a given location, the correlation between the vegetation and soil moisture can be positive or negative,

Table 2. Univariate Prediction Results

Satellite Observation	RMS NCEP	Correlation NCEP	RMS ECMWF	Correlation ECMWF
Passive MW SSM/I e19V-H	0.072	0.592	0.060	0.578
Passive MW SSM/I e37V-H	0.074	0.561	0.061	0.542
Active MW ERS small angle	0.062	0.721	0.057	0.633
Active MW ERS large angle	0.062	0.722	0.055	0.654
IR normalized *Ts* amplitude	0.059	0.753	0.057	0.633
NDVI	0.065	0.686	0.059	0.593

depending on the climate and the surface type (see section 3 and Figure 1). Disentangling this mixture of indirect relationships is not an easy task. Instead of avoiding or suppressing the correlation between vegetation and soil moisture, we try to use it. The idea is to integrate in the inversion scheme all the information available on the vegetation to capitalize on the indirect relationships between vegetation and soil moisture. Without a reliable radiative transfer forward model available for each satellite observation type and with only partial information content coming from each piece of information, the use of a statistical method is almost mandatory. In our case, a nonlinear statistical method is required to disentangle the difficult multivariate, indirect, and nonlinear relationships among the satellite observations, the vegetation and other surface effects, and the soil moisture. NN have proved in the past to be useful for such difficult problems.

[22] The Multi-Layered Perceptron (MLP) is adopted here [*Rumelhart et al.*, 1986]. This NN is a nonlinear mapping model. Given an input X, it provides an output Y. The inputs X are the predictors, they represent any source of information for the prediction, and the output Y represents the predicted variables. In our case, X is composed of various satellite observations. In the following, the number of inputs will vary, depending on the availability of observations (input dimension goes from 1 to 9), to test the contribution of each satellite observation. The prediction Y is the soil moisture estimate from NCEP and ECMWF reanalysis (output dimension is 2). Both NCEP and ECMWF soil moisture estimates will be predicted simultaneously in the output of the NN to provide a composite of the two models. Tests (not shown) compared their separate prediction with simultaneous predictions and show no significant difference.

[23] The neural network is trained to reproduce the behavior described by a data set of samples composed of an input X^e and its associated output Y^e ($e = 1, \ldots, N$ is the sample number in the training data set). Provided that enough samples (X^e, Y^e) are available, any continuous relationship as complex as it is, can be represented by a MLP [*Hornik et al.*, 1989; *Cybenko*, 1989]. The data set used to train the NN is composed of the satellite observations described in part 1 [*Prigent et al.*, 2005] and the soil moisture reanalyses of the NCEP and ECMWF for the 2 years of data already described (1993 and 1994).

[24] Soil moisture retrieval methods rarely consider estimating the anomaly instead of the retrieval of the absolute value. This is actually a classical approach when using a statistical model: predicting the change (second order) with respect to the mean state instead of predicting its absolute value (first order). Of course, the NN approach assumes that the mean state is known. A large number of local studies implicitly use this principle, but when working at global scales, the mean state of soil moisture content is not available. Furthermore, the mean state value is also of interest. However, if a reliable data set of soil moisture was available in the future for the training of the NN, such mean state could be used as well, and our prediction technique would greatly benefit from this information. The mean state is not used in this study.

4.2. Univariate Prediction Results

[25] In the first experiment, we use a univariate NN model: only one NN input (i.e., a particular satellite observation) and one output. All the satellite observations are individually and successively considered as inputs: the passive microwave measurements from SSM/I, the ERS active microwave observations, the IR-derived normalized *Ts* amplitude, and finally the AVHRR NDVI.

[26] Table 2 provides the RMS errors for the prediction of the NCEP and ECMWF soil moisture estimates, together with the correlation coefficient between the predicted and desired soil moisture.

[27] The RMS error is lower for ECMWF, but the correlation coefficient is larger for NCEP. This suggests that the NCEP soil moisture is more related to the satellite observations, but that its variability is larger than that of ECMWF. The correlations between the individual satellite observations and the soil moisture from NCEP or ECMWF (Table 1) have been considerably increased by transforming the satellite observation into a soil moisture estimate using the NN model. The nonlinearity of the NN applied to the satellite observation allows a better fit to the soil moisture. This indirect correlation between the observations and the soil moisture estimates (through the prediction) can be considered as a nonlinear correlation, in contrast to the direct linear correlation coefficient previously presented.

[28] The best predictors are the IR normalized *Ts* diurnal cycle amplitude and the active microwave measurements. The NDVI information is next. The passive microwave emissivity polarization differences are also acceptable predictors, even though they were not linearly correlated with the soil moisture: The NN model manages to extract nonlinearly the soil moisture information through the indirect vegetation/soil moisture relationships.

[29] The behavior of each univariate NN is illustrated in Figure 3. These curves represent how a NN with a single input (a satellite observation) changes its estimation of the soil moisture when the satellite observation is modified. Comparing all such univariate NN models provides a powerful quantification of information content of each satellite observation separately. Solid lines (dashed lines) represent the NN mapping for the NCEP (ECMWF) soil moisture estimates. The distribution of the satellite observations is represented below each mapping curve to indicate where the data are concentrated in the range of variability. Once trained, the neural networks are fixed, and by feeding each NN with the whole variation range of its particular input, it is possible to characterize the NN mapping. This makes it possible to analyze the NN, in particular the sensitivity of its outputs with respect to its inputs, and to measure the nonlinearity of the model. It is also a way of analyzing the data sets used to train the NN: Since the NN is

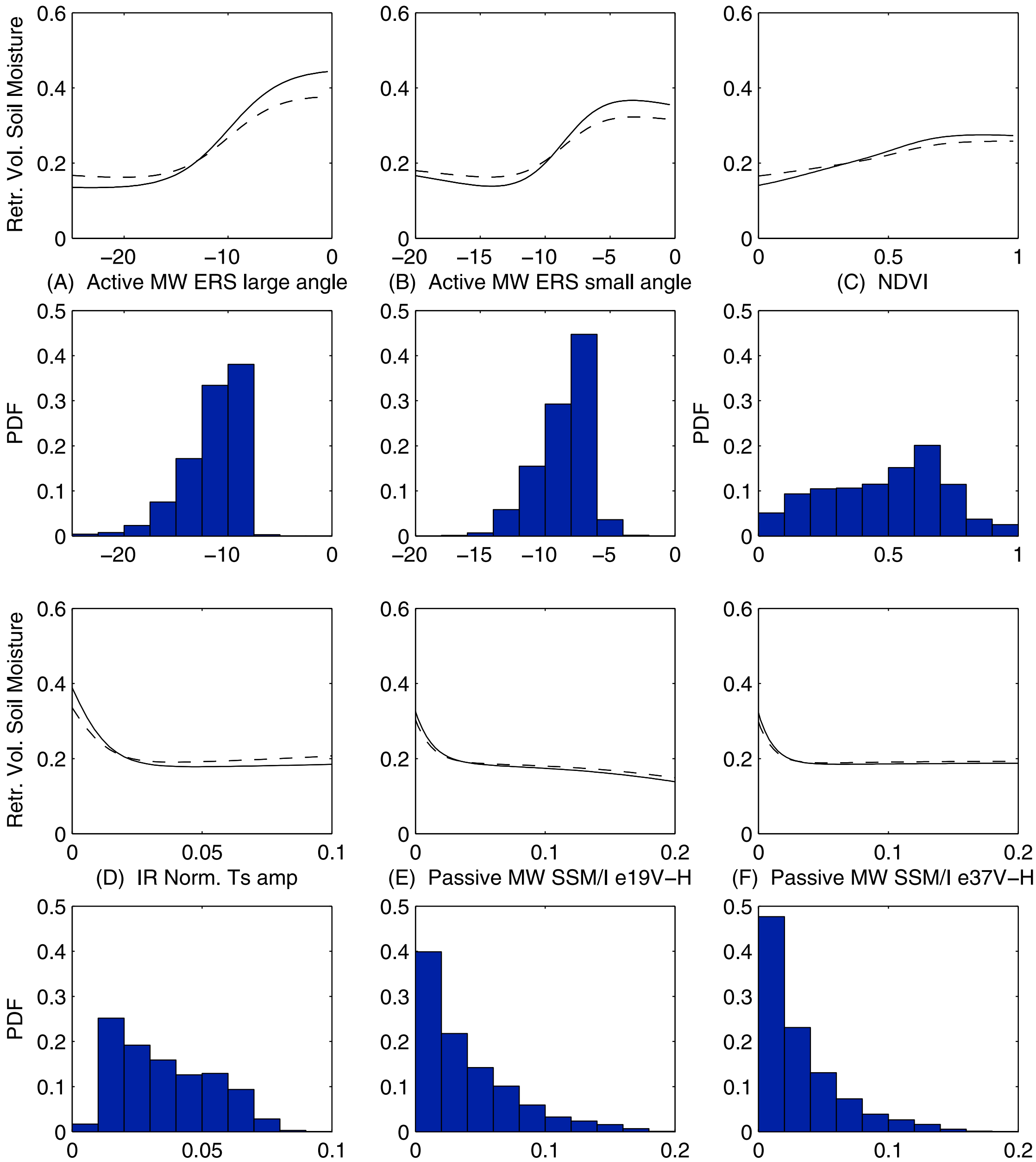

Figure 3. Satellite observations for (a) active MW ERS large angle, (b) active MW ERS small angle, (c) vegetation index from NDVI, (d) IR Normalized skin temperature diurnal cycle amplitude, (e) passive MW SSM/I Emissivity 19V-H, and (f) passive MW SSM/I Emissivity 37V-H, showing (top) monovariate NN behavior (solid lines for NCEP soil moisture estimates and dashed lines for ECMWF) and (bottom) satellite observation probability distribution function (PDF).

able to represent any continuous function with a high accuracy, the link found by the NN between the satellite observations and the soil moisture can be considered to be near optimal. Furthermore, the link is only determined by the relationships included in the data sets.

[30] The behavior of the univariate NN is coherent with the correlation and percentage of variance explained, provided in Table 2. Of course, the gradient of the curve in these figures is not the only pertinent information to consider to quantify the prediction power of a particular variable; the prediction can be more or less concentrated along this curve. The smoothness of the NN's behavior

Table 3. Multivariate Prediction Results[a]

Satellite Observation	RMS NCEP	Correlation NCEP	RMS ECMWF	Correlation ECMWF
IR normalized *Ts* amplitude	0.059	0.753	0.057	0.633
Active MW ERS small angle	0.055	0.792	0.054	0.676
NDVI	0.052	0.819	0.052	0.704
Active MW ERS large angle	0.051	0.826	0.051	0.724
Passive MW SSM/I e37V-H	0.050	0.832	0.050	0.735
Passive MW SSM/I e19V-H	0.050	0.832	0.050	0.737

[a]At each line, an additional source of satellite information is added.

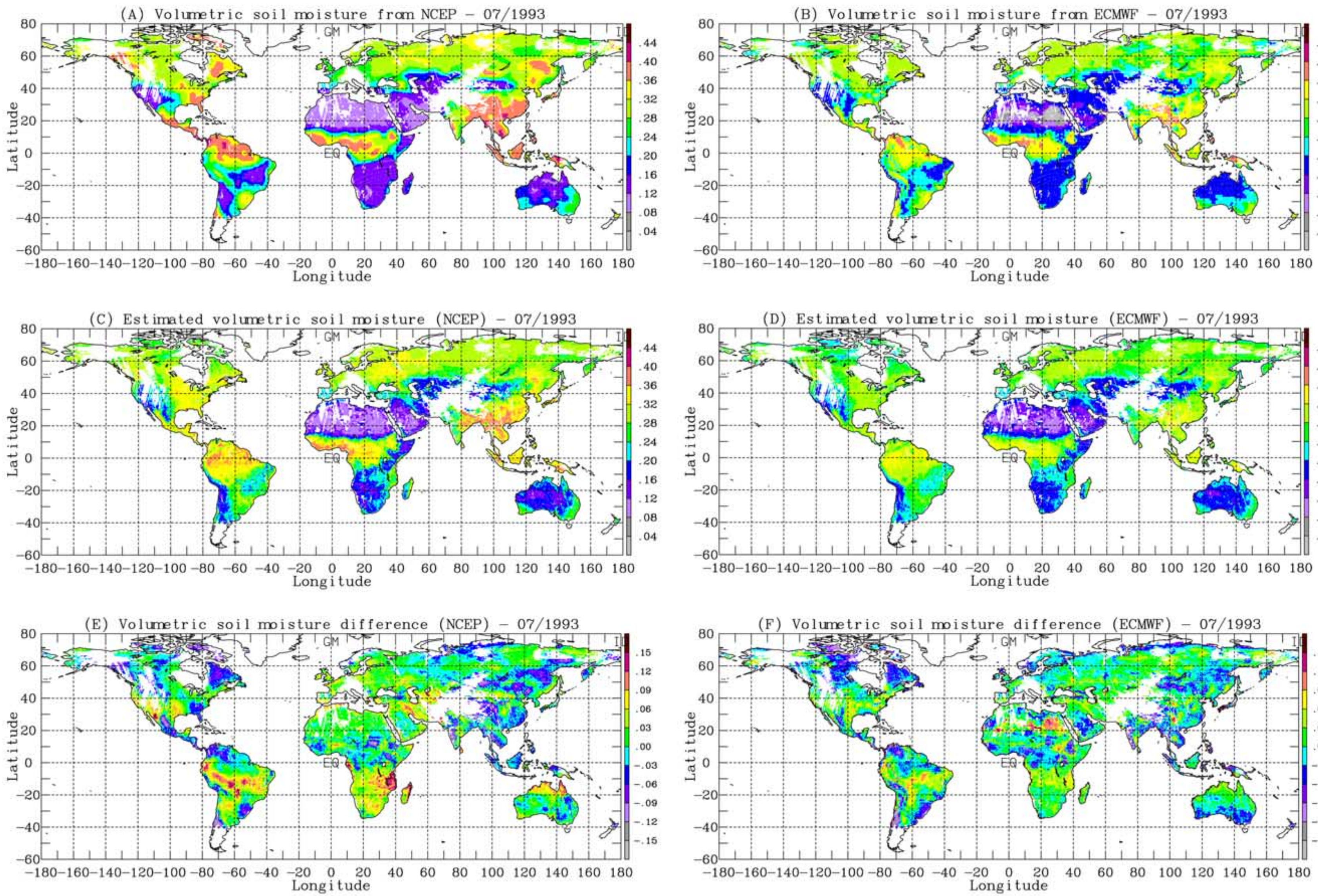

Figure 4. (a) Sample of monthly soil moisture from NCEP, (c) associated prediction from satellite observations, and (e) difference. (b, d, f) Same as Figures 4a, 4c, and 4e, respectively, but for ECMWF. Period is July 1993.

indicates that the learning phase performed satisfactorily, avoiding overtraining and overparameterization. Active microwave from ERS and IR normalized *Ts* diurnal cycle have the largest gradients, which means that these two information sources respond well to soil moisture variations from both NCEP (solid lines) and ECMWF (dashed lines). However, the gradients are stronger for NCEP than for ECMWF, which confirms the larger variability of the NCEP soil moisture values. The NDVI response is rather linear over the whole NDVI range. The passive microwave variations are again opposite to what the theory predicts, with a decrease of the emissivity polarization difference with increasing soil moisture: Variation of the vegetation density is obviously the dominant factor.

4.3. Multivariate Prediction Results

[31] Since soil moisture is related to satellite observations through a variety of intricate contributions and indirect relationships, adding together these different satellite measurements is important: Each one being sensitive to different aspects of the surface can provide different information. An ascending scheme to add up predictors in the order of their combined importance is developed here.

[32] In Table 3 we started with the best univariate NN from Table 2 (lowest RMS), i.e., with the IR normalized surface skin temperature diurnal cycle amplitude as predictor. Then, an ascendant procedure is adopted to add up, iteratively, the next best predictor. Each one of the remaining predictors is then added to the first one. A NN is tested with each of these different combinations of two inputs. The best NN is used to determine the best combination of two predictors. This scheme is repeated for determining the next best combined predictors.

[33] Using this ascending aggregation of predictors, the information content used by the NN increases: The RMS error on soil moisture is reduced from 0.059 to 0.050 for NCEP and from 0.057 to 0.050 for ECMWF, and the percentage of variance explained increases from 0.75 to 0.83 for NCEP and from 0.63 to 0.73 for ECMWF. After the addition of three or four predictors, neither the RMS errors nor the correlation coefficients are significantly improved because the information becomes redundant. However, every piece of information should be used when available: This better constrains the statistical model from satellite observations to soil moisture and therefore makes the solution more robust. In some locations, one or multiple satellite observations can be missing (for example, in India the surface skin temperature is missing). A different NN model, with different inputs, is actually defined for each one of these cases: The maximum number of pieces of information is thus exploited to constrain the model and compensate for missing data from any particular instrument, while reducing discontinuity between regions (different satellite measurements being available for different regions).

[34] An example of prediction using the NN scheme is presented in Figure 4a, for July 1993: The NCEP and ECMWF reanalyses are shown together with the prediction and the difference maps. The large-scale structures of the soil moisture fields are well reproduced with the satellite observations. The gradients from dry regions to humid ones are consistent with expectations, although large errors (differences) are present in specific regions, especially for the NCEP reanalyses. These problems will be discussed in the next section. The range of the predictions is smaller than the range in the original reanalyses. This can be explained: Much information is actually missing to describe all the soil moisture variability. When such factors are missing, the NN model, being a statistical model, describes the mean behavior in the training data set, and this tends to reduce extreme values.

[35] Figure 5 shows the RMS error statistics for the 2 years of data. Results are quite good, with a relatively uniform error over the globe for ECMWF data. For NCEP, the large structures in the errors already mentioned for the prediction sample given in Figure 4 are shown again. These regions actually correspond to known NCEP model problems (see Section 5.1).

[36] In order to check if the local-scale variability is also correctly reproduced by the NN model, time series of soil moisture are presented in Figure 6 for both NWP models and the satellite estimates. This time series is for an in situ station in India already analyzed in part 1 [*Prigent et al.*, 2005]. The soil moisture time variations predicted from the satellite data set are very similar to the corresponding NCEP and ECMWF ones. The NN method is not only capable of reproducing the large-scale spatial and temporal patterns (for which the NN scheme has been developed); it can also accurately describe the soil moisture local variations.

[37] Focusing on a smaller range of variability could improve the quality of the statistical prediction scheme. For example, various models could be set up for different types of vegetation or for different climate regimes. An alternative could also be to specialize NN models for different ranges of soil moisture content. We saw in section 3 that such filtering of data can significantly improve the correlations between the soil moisture and the satellite observations. However, thanks to its nonlinearity, the NN model has different sensitivity depending upon the situation. The NN is able to recognize the various "regimes," to change the relationships' predictors/predictands accordingly and, as a consequence, to yield good prediction statistics globally. A linear model cannot adapt itself to all the different and sometimes contradictory situations encountered over the globe. Using the spatial variability between a variety of locations, the NN is also able to address the time-variability in a particular location. This is an enormous advantage of this method. For example, since the soil moisture is being predicted indirectly from information on the vegetation, with the vegetation and soil moisture correlations negative or positive depending on the situation, some studies are only able to predict soil moisture when they are local and because the mean state of soil content is already known. The general approach developed here provides coherent results all over the globe. This is valuable for various reasons: (1) If the scheme can be tested in some locations, the model can be applied with confidence in other locations because no tuning to local conditions is performed, (2) the model can be used to check the consistency of a land surface global model outputs (see next section), (3) the sensitivity study is more general than a local study,

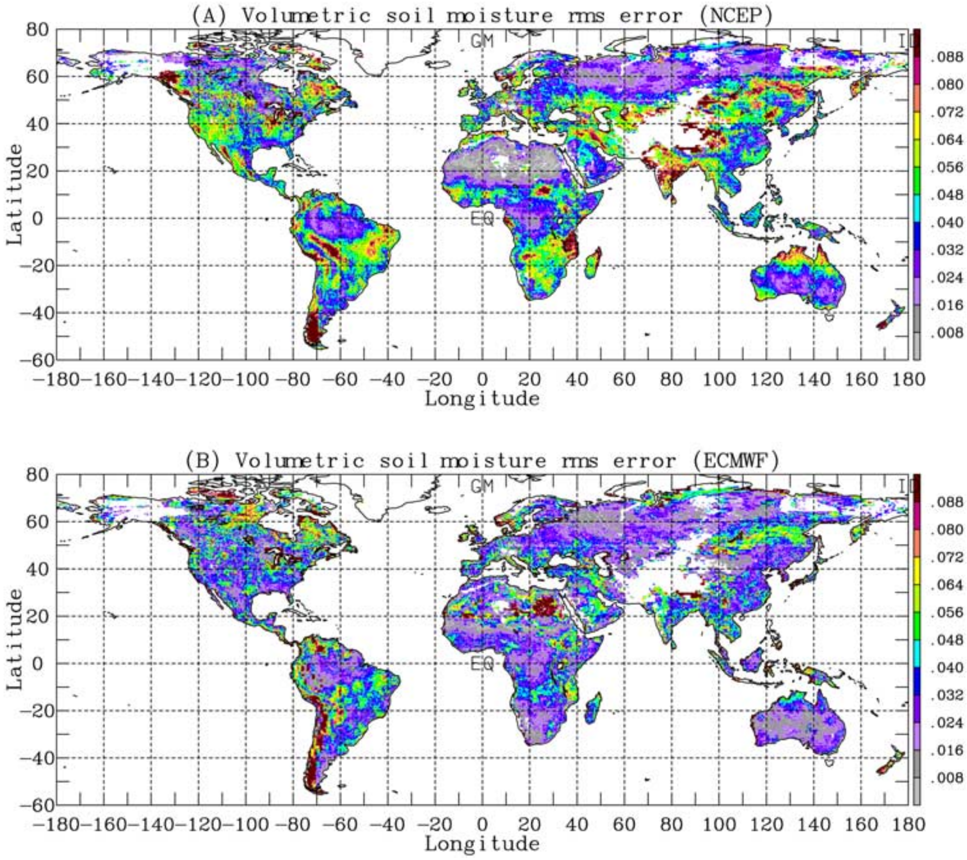

Figure 5. RMS error for the prediction of monthly (a) NCEP and (b) ECMWF soil moisture. Statistics are performed for 1993 and 1994.

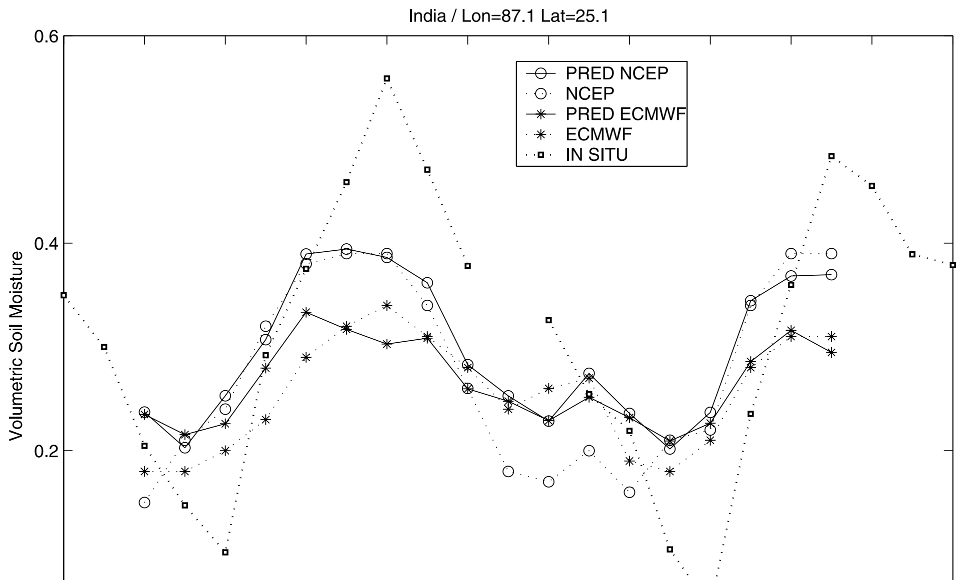

Figure 6. Time series of soil moisture from NCEP (respectively ECMWF), and predicted from satellite observations, together with in situ measurements. Station is over India.

(and 4) the conclusions of the sensitivity study can be directly used for satellite observation assimilation.

5. Applications of the NN Model

5.1. Consistency Checking

[38] The NN is a statistical model that allows for the characterization of complex relationships among variables; it is a multivariate and nonlinear model by nature. We have defined such a model to link satellite observations to soil moisture. These relationships have been examined in comparison to in situ measurements [*Prigent et al.*, 2005] and in terms of large-scale regional pattern over the globe (section 3).

[39] The NN model describes the global relationships between the satellite observations and the NWP soil moisture reanalysis. As such, in (smaller) areas where the relationships between the satellite observations and the reanalysis are anomalous (with respect to the global relations) the NN model can identify potential consistency problems in the reanalysis model: Comparisons between the NN model outputs derived from the satellite observations and the land surface model outputs should highlight the problematic regions. Any given soil moisture data set should be consistent with satellite observations. Since global relations formed with the reanalysis soil moisture values resemble the relations formed with in situ data, we can conclude that the qualitative validity of the model soil moisture treatment is high. In this way, the NN can be used as a validation/consistency check. Any given soil moisture data set can be confronted with the satellite measurements using this methodology to detect incoherencies in it. This NN consistency test is only a necessary condition for the quality of a soil moisture data set; it is not a sufficient condition. The set of satellite observations that we use is large and covers many aspects of surface characteristics: It represents a strong constraint on the data set.

[40] RMS prediction errors over the two years are presented in Figure 5 for both NWP reanalyses. As mentioned, these statistics show that the ECMWF soil moisture seems to be more coherent with the satellite observations than NCEP for which large errors appear in some areas. A region of large errors for NCEP in South America, with a north/south orientation, around 15°S west of the Andes is especially noticeable. It corresponds to an underestimate of soil moisture by the model in this area (see Figure 4). The NCEP land surface modelers were asked about this feature. It corresponds to a known problem of the NCEP model (W. Ebisuzaki, personal communication, 2002) caused by a precipitation deficit in the model in this region, especially during northern winter. The difference between observed precipitation and model precipitation exhibits a similar pattern (not shown). Soil moisture is adjusted in the model to correct for this problem: For a given pentad, the precipitation deficit is added to the top layer soil moisture during the next pentad (equal amounts every 6 hours). However, such a correction is not perfect, and some problems are still present leading to soil moisture underestimation.

[41] The soil moisture/precipitation correction should reduce the error caused by a bad model precipitation. However, it is not sufficient, especially when the problem is severe. See *Kanamitsu et al.* [2003] for a better description of the NCEP soil moisture. This example illustrates how the methodology is able to identify regions of the globe that are not consistent with the satellite observations due to difficulties in the model. Similar analysis could be applied to other land surface model outputs, such as the GSWP2 products that will soon be released.

[42] Other areas that are particularly affected by large errors are mountainous regions and the seasonal wetlands. In mountainous regions (see the Andes, for example, for the ECMWF retrieval), two main factors can contribute to the errors. First, the satellite data in the microwave domain, both active and passive, are sensitive to the topographic roughness, and this effect is likely to contaminate the retrieval. Second, in these regions of high spatial heterogeneity, the NWP models are generally less reliable. It can also be noted that for the month shown in Figure 5 (July), part of the Andes are snow covered.

[43] Extent and seasonality of the wetlands can be determined by a combination of satellite data [*Prigent et al.*, 2001], mostly driven by the sensitivity of the passive microwave observations to the presence of standing water. However, the wetland areas are not always well characterized by the NN (see, for instance, the Orinoco region in Venezuela with ECMWF). The origin of this problem comes from the model data (ECMWF or NCEP) used to train the NN: In some locations, the wetland is recognized by the model, but in other wetland locations, the model soil moisture is not affected. This introduces an incoherency in the data set to train the NN. As a consequence, the NN does not systematically detect the wetland areas, although these satellite observations already showed the required sensitivity to characterize the wetland areas [*Prigent et al.*, 2001]. This confirms that our technique is, in the current configuration, a consistency checking technique and not a definite retrieval method. As a consistency checking technique, the approach succeeded in showing that the model is not consistent in its treatment of the wetland areas.

5.2. Variational Assimilation

[44] We showed that this set of satellite observations can provide interesting information about the soil moisture. This information could be a very important constraint for land surface models. Variational assimilation is a technique designed to introduce information from observations into a numerical model. This technique requires a link between the space of the model variables (geophysical variables) and the space of the observations (satellite measurements).

[45] As already mentioned, there does not currently exist an adequate radiative transfer model to relate the satellite observations directly to the soil moisture as for other remote sensing applications (for instance, satisfactory radiative transfer models exist to reproduce the water vapor absorption in the atmosphere for water vapor profiling from satellite data). So how can this information be used in the framework of variational assimilation in land surface models? Since a direct physical relationship cannot be used, we propose to use a statistical relationship instead. Actually, each observation introduced into a variational assimilation scheme is stochastic by nature since uncertainty information needs to be associated with each assimilated observation. This is why there is no limitation in using a statistical link instead of a physical relation other than the accuracy of the

statistical link or the accuracy of the uncertainties associated to this statistical link. A neural network, as described in earlier sections, is an adequate candidate to model such a statistical link: The model is nonlinear, can describe highly complex relationships, and is a statistical model by nature [*Aires*, 2004], and uncertainty estimates can be specified in its outputs [*Aires et al.*, 2004a]. Furthermore, once this statistical model is generated, it can be analyzed in terms of its Jacobians [*Aires et al.*, 1999, 2004b] which provides extremely important additional information that insures the physical significance of the statistical constraint on the surface model.

[46] The general formalism of a 4-D variational assimilation scheme uses a cost function to combine information from observations and from the numerical model, weighted by the uncertainty estimates for both of these sources of information,

$$J_1(x_0) = \frac{1}{2}\left(x_0 - x_0^b\right)^T \boldsymbol{B}^{-1}\left(x_0 - {x_0}^b\right) + \frac{1}{2}\sum_{i=0}^{n}\left(H_i(x(t_i)) - y_i\right)^T \boldsymbol{R}_i^{-1}\left(H_i(x(t_i)) - y_i\right), \quad (1)$$

where x is the state vector, the t_i are the observation times, i is the corresponding time index, y_i are the vector of observations, H_i is the "observation operator," $\boldsymbol{R}_i$ is the observation error covariance matrix (matrices are indicated in bold) that includes measurements and representativeness errors (i.e., null-space error), and $\boldsymbol{B}$ is the background error covariance matrix for the background x_0^b (often from short-range forecast). The assimilation scheme described in this paper is 4-D, but the main comments would be valid for a simpler 1-D variational scheme too.

[47] This cost function is minimized to find the optimal state x_0. This is generally done by a gradient descent algorithm and requires the computation of the Jacobian of expression (1),

$$\bigtriangledown J_1(x_0) = \boldsymbol{B}^{-1}\left(x_0 - {x_0}^b\right) + \sum_{i=0}^{n}\boldsymbol{H}_i^T(x(t_i))\boldsymbol{R}_i^{-1}\left(H_i(x(t_i)) - y_i\right). \quad (2)$$

Estimating (2) requires the calculation of the tangent-linear operator of the forecast model and the Jacobians (or adjoint model), $\boldsymbol{H}_i$, of the observation operator H_i [*Le Dimet and Talagrand*, 1986].

[48] If the observations y_i are satellite measurements, the observation operator H_i is a radiative transfer model. Since a forward radiative transfer model linking surface parameters and satellite observations is not currently reliable enough, a statistical model, using for example a NN model, could be used instead. A statistical model such as a NN can be used for such a forward direct model, but obtaining the Jacobians of the NN forward model can be difficult [*Aires et al.*, 1999; *Chevallier and Mahfouf*, 2001], even if this is possible in some circumstances [*Aires et al.*, 2004b]. Furthermore, using such an approach would mean that we add the uncertainties of the forward model (to go from the state variables to the observations space) to the uncertainties of the Jacobians of this forward model (to go from the observations space to the state variables space).

[49] Instead of this traditional approach, we suggest using the inverse model defined in section 4. This has various benefits: (1) It avoids the estimation of the Jacobians of the neural network model, (2) it does not add up uncertainties of the forward model with the uncertainties of the Jacobians, and (3) it allows one to work directly on the state variables, which are more directly related to the numerical model. In order to assimilate the outputs of our inverse model, the cost function becomes

$$J_2(x_0) = \frac{1}{2}\left(x_0 - x_0^b\right)^T \boldsymbol{B}^{-1}\left(x_0 - {x_0}^b\right) + \frac{1}{2}\sum_{i=0}^{n}\left(x(t_i) - {x_i}^r\right)^T \boldsymbol{R}_i(x(t_i))^{-1}\left(x(t_i) - x_i^r\right), \quad (3)$$

where $x_i^r = NN(y_i)$ is the inversion of the observations, y_i, into the state space of the x_i. It should be noted that we introduce here a dependence on the situation $x(t_i)$ for the observation error covariance matrix $\boldsymbol{R}_i$. An approach to estimate the uncertainties of the NN predictions x_i^r in terms of an error covariance matrix has been developed by *Aires* [2004] and *Aires et al.* [2004a]. This is particularly interesting since it allows the variational assimilation system to give more weight to the observations when the NN inversion is reliable, and less weight when the inversion is less reliable.

[50] The Jacobian of (3) required for the minimization is simply

$$\bigtriangledown J_2(x_0) = \boldsymbol{B}^{-1}\left(x_0 - {x_0}^b\right) + \sum_{i=0}^{n}\boldsymbol{R}_i(x(t_i))^{-1}\left(x(t_i) - {x_i}^r\right). \quad (4)$$

A benefit of using the inverse statistical model instead of the forward version is that the assimilation is performed in the space of the numerical model state variables. It is known that defining a precise soil moisture quantity is difficult and this can be rather different from what would be used by a radiative transfer model. In fact, the meaning varies from model to model. By using the state variables of the particular numerical model during the training of the NN, we force coherence between the two soil moisture definitions. The assimilation scheme enforces at the same time: (1) a coherency of the state variables x_i, and (2) an external constraint for consistency with satellite observations.

[51] Another advantage of this approach is its flexibility. Any additional satellite observations can easily be used by the NN scheme: The statistical relationships linking the state variables of the numerical model would be even more complex and constraining for the solution. This variational assimilation application of the NN model defined in this paper will be the subject of a future study.

6. Conclusions and Perspectives

[52] As suggested as early as 1980 by *Schmugge et al.* [1980], an effective strategy for the detection of soil moisture should merge in situ measurements, numerical model, and remote sensing. This two-part study confirms

that each one of these components is essential: In situ measurements are used to analyze the sensitivity of satellite observations and to validate numerical models (at least locally), and remote sensing is required to initialize and constrain the surface models [*Robock et al.*, 2000].

[53] Satellite-derived quantities (active microwave backscattering, passive microwave emissivities, infrared-derived *Ts* diurnal cycle amplitude) have been systematically and objectively compared to the NWP soil moisture estimates from ECMWF and NCEP. For each satellite observation type, the optimum quantities have been previously derived and selected. For the passive microwave, the emissivities have been calculated [*Prigent et al.*, 1997, 1998], instead of using directly the brightness temperature that includes modulation by the atmosphere and by the surface temperature. The ERS scatterometer responses have been estimated for both low and high incidence angles. The infrared-derived *Ts* have been carefully analyzed to extract the diurnal amplitude [*Aires et al.*, 2004c] that is normalized by incident solar flux accounting for cloud effects.

[54] The objectives of this study were threefold: (1) to investigate the sensitivity of the satellite observations to soil moisture on a global basis, (2) to analyze the complementarity/interactions of the various satellite sources of soil moisture information, and (3) to assess the potential of the merging of such observations for soil moisture prediction.

[55] The sensitivity study conducted with the NWP soil moisture reanalysis confirms the conclusion previously derived from the analysis of satellite observations in coincidence with in situ soil moisture measurements [see *Prigent et al.*, 2005]. At the monthly timescale, the relations between the satellite measurements and the 10-cm soil moisture variations are complex and often indirect, coming through the correlation between vegetation and soil moisture. Obtaining consistent behavior with both in situ point measurements and large-scale NWP estimates gives weight to our conclusions, in addition to validating the use of the NWP soil moisture reanalysis as proxies for soil moisture at a global scale.

[56] A synergetic analysis is conducted that benefits from the optimum use of each instrument. A NN model is developed to describe the link between the satellite observations and the NWP soil moisture. No radiative transfer model today can accurately replicate this link on a global basis. The NN model can reproduce the NWP soil moisture outputs with a RMS error of 5% volumetric soil moisture; this statistical performance is remarkable, close to what is expected from the future SMOS mission retrieval (4% volumetric soil moisture). Although the NN model cannot be strictly seen as a retrieval algorithm because it is tightly related to the NWP soil moisture, the fact that the independent satellite observations can be related to model output with this level of accuracy is a positive sign for relating these observations to the real-world soil moisture. The fact that the NN model is able to work on a global scale comes from the synergetic use of observations from various wavelengths. Our study shows that a single source of information is not enough for a soil moisture monitoring generalized for all types of surfaces. We also show that the results of a particular retrieval scheme designed locally cannot be easily extrapolated to other surface conditions.

[57] Comparisons between the NN model output and the NWP soil moisture reanalysis revealed some particular problems with the NCEP land surface models that have been confirmed by the modelers. This methodology can thus help check the general consistency of output models with satellite observations and diagnose specific problems in them. We suggest using this approach to evaluate the GSWP 2 outputs that will soon be available.

[58] Finally, a method is described to use the satellite-derived NN model in the variational assimilation framework. We propose that the soil moisture estimates from our NN model should be assimilated instead of the raw satellite observations, in particular to avoid the estimation of forward model Jacobians that are even harder to obtain than the forward model itself. This is particularly useful when a good forward model does not exist. Another application in the variational assimilation context is fault detection where the soil moisture prediction by the numerical model would be monitored with the satellite observations and our statistical model.

[59] **Acknowledgments.** We would like to thank Hervé Douville from Météo-France for fruitful discussions, and Wesley Ebisuzaki for sharing some information on NCEP reanalysis. The ERS scatterometer data have been provided by IFREMER. This work was partly supported by the NASA Radiation Sciences and Hydrology Programs.

References

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

Aires, F., M. Schmitt, N. A. Scott, and A. Chédin (1999), The weight smoothing regularisation for MLP for resolving the input contribution's errors in functional interpolations, *IEEE Trans. Neural Networks*, *10*, 1502–1510.

Aires, F., C. Prigent, and W. B. Rossow (2004a), Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 2. Output errors, *J. Geophys. Res.*, *109*, D10304, doi:10.1029/2003JD004174.

Aires, F., C. Prigent, and W. B. Rossow (2004b), Neural Network Uncertainty Assessment Using Bayesian Statistics With Application To Remote Sensing: 3. Network Jacobians, *J. Geophys. Res.*, *109*, D10305, doi:10.1029/2003JD004175.

Aires, F., C. Prigent, and W. B. Rossow (2004c), Temporal interpolation of global surface skin temperature diurnal cycle over land under lear and cloudy conditions, *J. Geophys. Res.*, *109*, D04313, doi:10.1029/2003JD003527.

Betts, A. K., J. H. Ball, M. Bosilovich, P. Viterbo, Y. Zhang, and W. B. Rossow (2003a), Intercomparison of water and energy budgets for five Mississippi subbasins between ECMWF reanalysis (ERA-40) and NASA Data Assimilation Office fvGCM for 1990–1999, *J. Geophys. Res.*, *108*(D16), 8618, doi:10.1029/2002JD003127.

Betts, A. K., J. H. Ball, and P. Viterbo (2003b), Evaluation of the ERA-40 surface water budget and surface temperature for the Mackenzie River Basin, *J. Hydrometeorol.*, *4*, 1194–1211.

Chen, F., and K. Mitchell (1999), Using the GEWEX/ISLSCP forcing data to simulate global soil moisture fields and hydrological cycle for 1987–1988, *J. Meteorol. Soc. Jpn.*, *77*, 167–182.

Chevallier, F., and J.-F. Mahfouf (2001), Evaluation of the Jacobians of infrared radiation models for variational data assimilation, *J. Appl. Meteorol.*, *40*(8), 1445–1461.

Cybenko, G. (1989), Approximation by superpositions of a sigmoidal function, *Math. Control Signals Syst.*, *2*, 303–314.

Douville, H., P. Viterbo, J.-P. Mahfouf, and A. C. M. Beljaars (2000), Evaluation of the optimum interpolation and nudging techniques for soil moisture analysis using FIFE data, *Mon. Weather Rev.*, *128*, 1733–1756.

Goetz, S. J. (1997), Multi-sensor analysis of NDVI, surface temperature and biophysical variables at a mixed grassland site, *Int. J. Remote Sens.*, *8*, 71–94.

Hornik, K., M. Stinchcombe, and H. White (1989), Multilayer feedforward networks are universal approximators, *Neural Networks*, *2*, 359–366.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–470.

Kanamitsu, M., C.-H. Lu, J. Schemm, and W. Ebisuzaki (2003), The predictability of soil moisture and near surface temperature in hindcasts of NCEP Seasonal Forecast Model, *J. Clim.*, *16*, 510–521.

Le Dimet, F.-X., and O. Talagrand (1986), Variational algorithms for analysis and assimilation of meteorological observations, *Tellus, Ser. A*, *38*, 97–110.

Mahrt, L., and H. L. Pan (1984), A two-layer model of soil hydrology, *Boundary Layer Meteorol.*, *29*, 1–20.

Mintz, Y., and Y. Serafini (1992), A global monthly climatology of soil moisture and water balance, *Clim. Dyn.*, *8*, 13–27.

Pan, H. L., and L. Mahrt (1987), Interaction between soil hydrology and boundary layer developments, *Boundary Layer Meteorol.*, *38*, 185–202.

Prigent, C., W. B. Rossow, and E. Matthews (1997), Microwave land surface emissivities estimated from SSM/I observations, *J. Geophys. Res.*, *102*, 21,867–21,890.

Prigent, C., W. B. Rossow, and E. Matthews (1998), Global maps of microwave land surface emissivities: Potential for land surface characterization, *Radio Sci.*, *33*, 745–751.

Prigent, C., F. Aires, W. B. Rossow, and E. Matthews (2001), Joint characterization of the vegetation by satellite observations from visible to microwavelengths: A sensitivity analysis, *J. Geophys. Res.*, *106*, 20,665–20,685.

Prigent, C., F. Aires, W. B. Rossow, and A. Robock (2005), Sensitivity of satellite observations in the IR and microwave to soil moisture at a global scale: 1. Presentation of the satellite observations and analysis of their relations with in situ soil moisture measurements, *J. Geophys. Res.*, *110*, D07110, doi:10.1029/2004JD005087.

Roads, J. O., S. C. Chen, M. Kanamitsu, and H. Juang (1999), Surface water characteristics in NECEP's reanalysis and global spectral model, *J. Geophys. Res.*, *104*, 19,307–19,327.

Robock, A., K. Y. Vinnikov, G. Srinivasan, J. K. Entin, S. E. Hollinger, N. A. Speranskaya, S. Liu, and A. Namkhai (2000), The Global Soil Moisture Data Bank, *Bull. Am. Meteorol. Soc.*, *81*, 1281–1299.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986), Learning internal representations by error propagation, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. I, *Foundations*, edited by D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, pp. 318–362, MIT Press, Cambridge, Mass.

Schmugge, T. J., T. J. Jackson, and H. L. McKim (1980), Survey of methods for soil moisture determination, *Water Resour. Res.*, *16*, 961–979.

Simmons, A. J., and J. K. Gibson (Eds.) (2000), The ERA-40 project plan, ERA-40 Project Rep. Ser. 1, 62 pp., Eur. Cent. for Medium-Range Weather Forecasts, Reading, UK.

Van den Hurk, B., P. Viterbo, A. C. M. Beljaars, and A. K. Betts (2000), Offline validation of the ERA-40 surface scheme, *Tech. Memo. 295*, 43 pp., Eur. Cent. for Medium-Range Weather Forecasts, Reading, UK.

Viterbo, P., and A. C. M. Beljaars (1995), An improved land surface parameterization scheme in the ECMWF model and its validation, *J. Clim.*, *8*, 2716–2748.

F. Aires, Laboratoire de Météorologie Dynamique, Institut Pierre Simon Laplace (IPSL)/CNRS, Université Pierre et Marie Curie, case 99, 4 place Jussieu, F-75252 Paris, France. (filipe.aires@lmd.jussieu.fr)

C. Prigent, Observatoire de Paris, Laboratoire d'Étude du Rayonnement et de la Matiére en Astrophysique (LERMA)/CNRS, 61 avenue de l'Observatoire, Paris 75014, France. (catherine.prigent@obspm.fr)

W. B. Rossow, NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, USA. (wrossow@giss.nasa.gov)