remote sensing

MDPI

Article

Long Term Global Surface Soil Moisture Fields Using an SMOS-Trained Neural Network Applied to AMSR-E Data

Nemesio J. Rodríguez-Fernández [1,2,*], Yann H. Kerr [1], Robin van der Schalie [3,4], Amen Al-Yaari [5], Jean-Pierre Wigneron [5], Richard de Jeu [3,4], Philippe Richaume [1], Emanuel Dutra [2], Arnaud Mialon [1] and Matthias Drusch [6]

1 Centre d'Etudes Spatiales de la Biosphère Université de Toulouse, Centre National d'Etudes Spatiales (CNES), Centre National de la Recherche Scientifique (CNRS), Institut de Recherche pour le Dévelopement (IRD), 18 av. Edouard Belin, bpi 2801, 31401 Toulouse CEDEX 9, France; yann.kerr@cesbio.cnes.fr (Y.H.K.); phillipe.richaume@cesbio.cnes.fr (P.R.); arnaud.mialon@cesbio.cnes.fr (A.M.)

2 European Centre for Medium-Range Weather Forecasts (ECMWF), Shinfield Park, RG2 9AX Reading, UK; emanuel.dutra@ecmwf.int

3 Faculty of Earth and Life Sciences, VU University Amsterdam (VUA), 1081 HV Amsterdam, The Netherlands; vanderschalie@vandersat.com (R.v.d.S.); rdejeu@vandersat.com (R.d.J.)

4 Transmissivity B.V., Huygensstraat 34, 2201 AZ Noordwijk, The Netherlands

5 Interactions Sol Plante Atmosphére (ISPA), Unité Mixte de Recherche 1391, Institut National de la Recherche Agronomique (INRA), CS 20032, 33882 Villenave d'ornon cedex, France; amen.al-yaari@inra.fr (A.-A.Y.); jean-pierre.wigneron@inra.fr (J.-P.W.)

6 European Space Research and Technology Centre (ESTEC), European Space Agency (ESA), Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands; matthias.drusch@esa.int

* Correspondence: nemesio.rodriguez@cesbio.cnes.fr

Academic Editors: Prashant K. Srivastava, Nicolas Baghdadi and Prasad S. Thenkabail
Received: 19 July 2016; Accepted: 9 November 2016; Published: 18 November 2016

Abstract: A method to retrieve soil moisture (SM) from Advanced Scanning Microwave Radiometer—Earth Observing System Sensor (AMSR-E) observations using Soil Moisture and Ocean Salinity (SMOS) Level 3 SM as a reference is discussed. The goal is to obtain longer time series of SM with no significant bias and with a similar dynamical range to that of the SMOS SM dataset. This method consists of training a neural network (NN) to obtain a global non-linear relationship linking AMSR-E brightness temperatures (T_b) to the SMOS L3 SM dataset on the concurrent mission period of 1.5 years. Then, the NN model is used to derive soil moisture from past AMSR-E observations. It is shown that in spite of the different frequencies and sensing depths of AMSR-E and SMOS, it is possible to find such a global relationship. The sensitivity of AMSR-E T_b's to soil temperature (T_{soil}) was also evaluated using European Centre for Medium-Range Weather Forecast Interim/Land re-analysis (ERA-Land) and Modern-Era Retrospective analysis for Research and Applications-Land (MERRA-Land) model data. The best combination of AMSR-E T_b's to retrieve T_{soil} is H polarization at 23 and 36 GHz plus V polarization at 36 GHz. Regarding SM, several combinations of input data show a similar performance in retrieving SM. One NN that uses C and X bands and T_{soil} information was chosen to obtain SM in the 2003–2011 period. The new dataset shows a low bias (<0.02 m^3/m^3) and low standard deviation of the difference (<0.04 m^3/m^3) with respect to SMOS L3 SM over most of the globe's surface. The new dataset was evaluated together with other AMSR-E SM datasets and the Climate Change Initiative (CCI) SM dataset against the MERRA-Land and ERA-Land models for the 2003–2011 period. All datasets show a significant bias with respect to models for boreal regions and high correlations over regions other than the tropical and boreal forest. All of the global SM datasets including AMSR-E NN were also evaluated against a large number of in situ measurements over four continents. Over Australia, all datasets show a strong level of agreement with in situ measurements. Models perform better over Europe and mountainous regions in North America. Remote sensing

datasets (in particular NN and the Land Parameter Retrieval Model (LPRM)) perform as well as models for other North American sites and perform better than models over the Sahel region.

Keywords: soil moisture; passive radiometry; neural networks; SMOS (Soil Moisture and Ocean Salinity) mission

1. Introduction

Understanding the Earth's climate and how it is changing is one of the major challenges of the twenty-first century. Climatic patterns can only be reliably detected in time series that are sufficiently long (at least a few decades) because otherwise, natural short-scale variations can hide the longer term trends. Therefore, to understand the climate evolution, a combination of long-term models and observations is needed. Sustained observations from satellites can contribute significantly to this goal, and this is the objective of observational programs, such as the European Spatial Agency (ESA) Climate Change Initiative (CCI) [1]: to provide information on variables that were identified as "essential climate variables" (ECVs) by the Global Climate Observing System (GCOS) in the context of the United Nations Framework Convention on Climate Change (UNFCCC). This is also the goal of numerical re-analyses, such as ERA-Interim [2], which use data assimilation techniques to merge observations and models in an optimal way during long periods of time.

Soil moisture (SM) has been endorsed by the GCOS as an ECV [3], as it affects the water and energy fluxes at the land surface/atmosphere interface [4,5]. Therefore, one of the ESA CCI projects, launched in 2010, focuses on SM. As already mentioned, in order to use SM information for climate modeling, SM datasets spanning long time periods are needed. The ESA SM CCI has developed a strategy [6–9] that consists of merging a posteriori different SM datasets obtained with different algorithms applied to data from different microwave active and passive sensors. The scaling of the different datasets is done using as a reference a land surface model (such as GLDAS Noah) [6,7].

The Soil Moisture and Ocean Salinity (SMOS) satellite [10] was launched in November 2009, and it is the first mission specifically designed to retrieve SM from space. Many studies have evaluated the SMOS SM dataset in comparison to other remote sensing datasets, models and in situ measurements [11–19]. SMOS shows very good global performance although other remote sensing and model products can show better performances at some sites. In any case, datasets from the only two instruments specifically conceived to measure SM, SMOS and Soil Moisture Active Passive (SMAP), compare very well with each other [20,21].

Since the ESA SM CCI started shortly after the SMOS launch, SMOS is not included in the CCI. However, it is important to investigate ways of including SMOS data in the long-term SM datasets and of evaluating the new datasets. One of the advantages of SMOS is that it can observe at a lower frequency (1.4 GHz) than previous instruments, and this frequency is less affected by the vegetation cover [22]. A pertinent first step towards that goal is to study how to obtain a coherent dataset from SMOS and the NASA/JAXA Advanced Scanning Microwave Radiometer (AMSR-E), which operated from 2003–2011. Two approaches have been discussed in the literature: (i) adapting the Land Parameter Retrieval Model (LPRM) algorithm to SMOS data [23]; (ii) using SMOS SM as a reference to compute local linear regression equations linking AMSR-E T_b's to SMOS SM and recomputing a SM dataset from AMSR-E observations [24]. The present study is devoted to a third approach: using SMOS SM as reference to find a global non-linear regression relationship linking AMSR-E T_b's to SMOS SM and recomputing an SM dataset from the AMSR-E observations. Preliminary results were presented in [25–27].

Neural networks are a very efficient tool for computing global non-linear regressions. Recently, the ESA has released a new SMOS SM near-real-time product computed with neural networks [28]. NNs have been used successfully to retrieve SM from AMSR-E observations over Mongolia, Australia

and Italy [29,30]. NNs have also been used to test the a priori merging of active and passive microwaves [31,32], including Advanced Scatterometer (ASCAT) and AMSR-E [33] or ASCAT and SMOS [34]. In contrast, to our knowledge, this is the first time that NNs have been used to extend an SM record in time using a global remote sensing dataset as a reference to train an algorithm to retrieve SM from previous observations by another instrument. The NN training phase can be considered as a consistency check of AMSR-E T_b's and SMOS SM. It has been proposed that training an NN with a land surface model could be an efficient approach to assimilate satellite data into that same model [31,32]. However, if the goal is creating a generic long-term data record suitable for general data assimilation applications, not using a land surface model to train the NN is a potential advantage as one does not want to assimilate information from one model into another.

The rest of this paper is organized as follows. Section 2 describes the remote sensing and model datasets used. Section 3 presents the in situ measurements used for the evaluation and the comparison protocol. Section 4 discusses the training of the neural networks and the soil moisture retrieval. The new SM dataset is evaluated in Section 5 in comparison to SMOS L3 SM (Section 5.1), other AMSR-E SM datasets (Section 5.2), global models (Section 5.3) and in situ measurements (Section 5.4). Finally, the conclusions are summarized in Section 6.

2. Datasets and Pre-Processing

2.1. SMOS Soil Moisture

SMOS [10,35] was launched on 2 November 2009 to measure the thermal emission from the Earth in the 1.4-GHz protected frequency range in full-polarization and for incidence angles from 0°–∼60°. The footprint (full width at half maximum of the synthesized beam) is ∼43 km on average [35]. The Equator overpass time is 6:00 a.m./p.m. for ascending/descending orbits.

The SMOS SM retrieval algorithm is based on the iterative minimization of the difference of a forward model and the brightness temperatures (T_b's) measured by SMOS [36]. The model uses the $\tau - \omega$ (optical depth-single scattering albedo) approach to take into account the effect of vegetation [37]. In the case of forest, two contributions to the opacity are taken into account: one from the arboreous component, which is estimated from the maximum Leaf Area Index (LAI) [38], and another from the understory vegetation. Soil temperature is obtained from ECMWF Integrated Forecast System (IFS) data. For footprints with mixed land cover, the soil moisture content of the minor land cover is estimated from ECMWF IFS, and its contribution to the T_b is fixed. For such cases, the SMOS SM retrieval is only performed for the dominant land cover class within the footprint [36]. The Centre Aval de Traitement de Données SMOS (CATDS) Level 3 daily product (SMOS L3 SM, [39]) Version 3.00 was used in this study. The Level 3 algorithm uses three orbits in a seven-day window to constrain the optical depth as it is not physically expected to vary strongly in that time period [39]. CATDS products are provided on an Equal-Area Scalable Earth grid version 2 (EASEv2) with typical node spacing of 25 km [40] as two files per day, one for ascending and another of descending half-orbits (there is no temporal averaging in the L3 daily product).

2.2. AMSR-E Brightness Temperatures

The Advanced Microwave Scanning Radiometer for the Earth Observation System (AMSR-E) is a six-frequency dual-polarized total-power passive microwave radiometer dedicated to observing water-related geophysical parameters [41]. AMSR-E was launched in May 2002, on-board the Aqua satellite, which follows a near-polar Sun-synchronous orbit of 98.8-min periods, crossing the Equator at 1:30 p.m. local time for ascending orbits and at 1:30 a.m. for descending orbits. AMSR-E observed the Earth until October 2011, in six frequency bands: 6.9 GHz (C-band), 10.7 GHz (X-band), 18.7 GHz, 23.8 GHz, 36.5 GHz and 89.0 GHz, at a constant incidence angle of 55°. The footprint size depends on the frequency; for instance, it is (cross-track × along-track) 43 × 75 km^2 and 29 × 51 km^2 at 6.9 and 10.7 GHz, respectively. The sampling interval on the Earth's surface is 10.1 km along-track and

9.0 km across-track. In this study, Level 2A brightness temperatures (L2A, Version 12) have been used. The dataset was obtained from the National Snow and Ice Data Center.

2.3. AMSR-E LPRM Soil Moisture

The Land Parameter Retrieval Model (LPRM) is the baseline model for the soil moisture retrievals from the passive microwave sensors within the ESA ECV CCI soil moisture dataset [42]. The SM retrievals from LPRM have been extensively validated [12,43], and the model performance has been documented in numerous publications [6,44]. LPRM demonstrated a good soil moisture retrieval skill from AMSR-E observations over sparse to moderately-vegetated regions [15,45]. LPRM uses vertical and horizontal polarized T_b's as an input and uses a radiative transfer model to simultaneously solve for SM and vegetation optical depth [6]. An advantage of LPRM is that it can be applied on different frequency bands, including the X, C and L bands. The LPRM SM dataset used in this study was obtained with an optimized single scattering albedo, a new roughness model and using both AMSR-E and SMOS observations [23,46].

2.4. AMSR-E Reg Soil Moisture

The "AMSR-E Reg" SM dataset [24] has been obtained by computing local linear regressions linking the surface reflectivity for vertical and horizontal polarization and SMOS L3 SM using a set of three free parameters for each grid point. The local regression parameters were computed in the June 2010–October 2011 period (when both AMSR-E and SMOS were in operation), getting as many triplets of free parameters as land grid points with available AMSR-E T_b's and SMOS L3 SM data ($\sim 1.5 \times 10^5$ points). In a second step, those parameters are used to estimate SM from AMSR-E observations over the 2003–2010 period.

2.5. ESA Climate Change Initiative Soil Moisture

The European Space Agency Climate Change Initiative (CCI) has supported the generation of a long-term SM product based on multiple microwave sensors in space, such as the radiometers (passive) SMMR (Scanning Multichannel Microwave Radiometer), SSM/I (Special Sensor Microwave Imager), Tropical Rainfall Measuring Mission's (TRMM) Microwave Imager (TMI), WindSat Polarimetric Radiometer, AMSR-E and AMSR-2 and the active sensors on board the European Remote Sensing (ERS) satellites (Wind Scatterometer) and the Metop satellites (Advanced Scatterometer, ASCAT) [47]. Datasets derived from the different instruments are scaled into a common climatology by matching their cumulative distribution functions (CDF) to those of AMSR-E for the passive datasets or ASCAT for the active datasets. Since the different satellite products are provided in different units, they are scaled into a common level using GLDAS-Noah model simulations. Finally, the active and passive datasets are merged into a combined single dataset. Over areas with low vegetation, the passive one is used, while over areas with moderate vegetation density, the active one is used. In transition areas, if both passive and active data are available, the average is used, otherwise the only available data are used for the final product. The CCI SM dataset used in this study is Version 2.2 (an evaluation against in situ measurements can be found in [9]).

2.6. MERRA Land

The Modern-Era Retrospective analysis for Research and Applications-Land (MERRA-Land) [48] is an atmospheric re-analysis that makes use of the Goddard Earth Observing System model Version 5 and the data assimilation system Version 5.2.0 with an updated precipitation forcing and catchment land surface model compared to the normal MERRA dataset. It was developed by NASA and is available from the beginning of 1980 to the present day. In the current study, the top 2-cm SM is used. The data have a temporal resolution of one hour and a spatial resolution of $1/2^\circ \times 2/3^\circ$.

2.7. ECMWF ERA-Land Models

ECMWF model simulations of snow depth, soil temperature and soil moisture were used to filter the data and to compare with remote sensing soil moisture measurements. The data used in this study were extracted from a global Tiled ECMWF Scheme for Surface Exchanges over Land (TESSEL) with improved hydrology [49] simulation of surface parameters from 1979–2014. The land surface model has been forced with the atmosphere from the ERA-Interim re-analysis [2]. This dataset will subsequently be referred to as ERA-Land and differs from the ERA-Interim/Land [50] only in the precipitation forcing, which was not rescaled to match Global Precipitation Climatology Project (GPCP) monthly means since GPCP Version 2.1 was not available after 2010. A more recent version of GPCP could have been used, but it was decided to use the original ERA-Interim precipitations in this study. It has a temporal sampling of 6 h and a spatial resolution of 79 km. Only the soil moisture and temperature from the first layer of the model (0–7-cm depth) have been used in the current study.

2.8. Data Collocation and Filtering

The acquisition time difference between AMSR-E and SMOS could be significant for a dynamic variable, such as SM. Therefore, only the night orbits (descending AMSR-E and ascending SMOS) were used to avoid significant differences in the surface SM due to evaporation and convective precipitation during daytime. Possible SM differences induced by precipitation taking place in between AMSR-E and SMOS overpasses were not taken into account, but they will appear in a limited number of points at the global level, and as shown by [32], this should not affect the training. The CCI, MERRA-Land and ERA-Land datasets were spatially interpolated to the EASEv2 grid by the area-overlap weighted average. The data have also been interpolated to the time of AMSR-E acquisitions by linear interpolation. Several filters can be applied to the data:

1. Reference data uncertainty: An upper limit to the SMOS SM data uncertainty has been set for the training phase (data were used when the Dqx parameter is lower than 0.06 m^3/m^3 and there is less than 30% of forest in the footprint, the FFO parameter).
2. RFI in L-Band: SMOS SM data with a radio frequency interference (RFI) probability higher than 20% (RFI_Prob field) were filtered out.
3. RFI in C or X bands: AMSR-E brightness temperature difference from C-band to X-band for both for H and V polarizations ($T^{H,V}_{diff} = T^{H,V}_{C} - T^{H,V}_{X}$) should be in the range $-10\,\mathrm{K} \leq T^{H,V}_{diff} \leq 5\,\mathrm{K}$. Otherwise, the data are probably affected by RFI from artificial sources as discussed by [51]. Some geophysical phenomena could be identified by this criterion as RFI, for instance when the scattering is important over snow/ice or dry sand surfaces [52]. However, the effect is small, and a number of studies [6,46] have used the criterion by [51]. In addition, in the current study, snow/ice was filtered out before retrieving SM.
4. Frozen soil and snow: The soil temperature from the ERA-Land model should be higher than 274 K to avoid frozen soil (taking into account a temperature uncertainty of ~ 1 K). The snow depth from ERA-Land should be lower than 1 mm.

All filters from 1–4 have been applied to obtain a clean training subset (Section 4.1). This subset contains $\sim 4.5 \times 10^6$ samples, which allows one to study statistical retrieval techniques. For the AMSR-E NN retrieval and evaluation (Sections 5.2–5.4), only Filters 3 and 4 have been applied, except for the comparison to SMOS L3 in the common period (June 2010–October 2011; Section 5.1), for which Filter 2 was also applied.

3. In Situ Measurements and Evaluation Protocol

The AMSR-E NN SM dataset discussed in this paper, along with those presented above, was evaluated against in situ measurements for a large number of sites. All of the in situ data were obtained from the International Soil Moisture Network [53] except for the Danish Hydrological

Observatory (HOBE) and the African Monsoon Multidisciplinary Analysis (AMMA), which have been directly provided by the teams managing the experimental sites. As the C-band radiation detected from space comes from the first few centimeters of the soil, only SM measurements at the 2-, 5- and 8-cm depth range were used. The sites are located in four continents and cover a large spectrum of climate conditions. Table 1 shows a summary of the networks used, and Figure 1 shows their position. The different networks were classified in two main types:

1. Sparse networks: networks with a single site or a larger number of sites located at a distance larger than the AMSR-E footprint. Therefore, one single in situ measurement is compared to the corresponding remote sensing measurement.
2. Dense networks: networks with more than one sensor (Table 1) in a region of several hundred square kilometers (238 km^2 in Reynolds Creek or 610 km^2 in Little Washita, for instance). Since this surface is smaller than an AMSR-E footprint ($\sim$1500 km^2 at the X-band), measurements from individual sensors of the same network can be averaged to have a soil moisture value more representative of the spatial scale of the remote sensing measurement.

Table 1. In situ networks used in this study. The depths are quoted as two numbers: the first one is the upper depth, and the second one is the lower depth of the sensor. Both numbers are equal when the sensor is placed horizontally. The third column gives the number of sensors remaining for each network after applying all of the criteria used for the evaluation that are discussed in Section 3. The third column gives the average number of points in the times series for each network that satisfies the evaluation criteria.

Network	Depth (cm)	Sites	Points	Location
Sparse Networks				
ARM	0.05–0.05	38	687	USA
CARBOAFRICA	0.05–0.05	1	522	Sudan
DAHRA	0.05–0.05	1	492	Senegal
OZNET	0.00–0.05	23	630	Australia
OZNET	0.00–0.08	16	1168	Australia
REMEDHUS	0.00–0.05	25	988	Spain
MOL-RAO	0.08–0.08	1	1323	Germany
SMOSMANIA	0.05–0.05	14	576	France
SWEX-POLAND	0.05–0.05	3	380	Poland
UDC-SMOS	0.05–0.05	21	348	Germany
SCAN	0.05–0.05	88	588	USA
SNOTEL	0.05–0.05	147	459	USA
Dense Networks				
ARS Little River	0.00–0.05	29	817	USA
ARS Little Washita	0.00–0.05	20	483	USA
ARS Reynolds Creek	0.00–0.05	19	954	USA
HOBE	0.00–0.05	32	204	Denmark
AMMA Benin	0.05–0.05	9	381	Benin
AMMA Niger	0.05–0.05	5	380	Niger

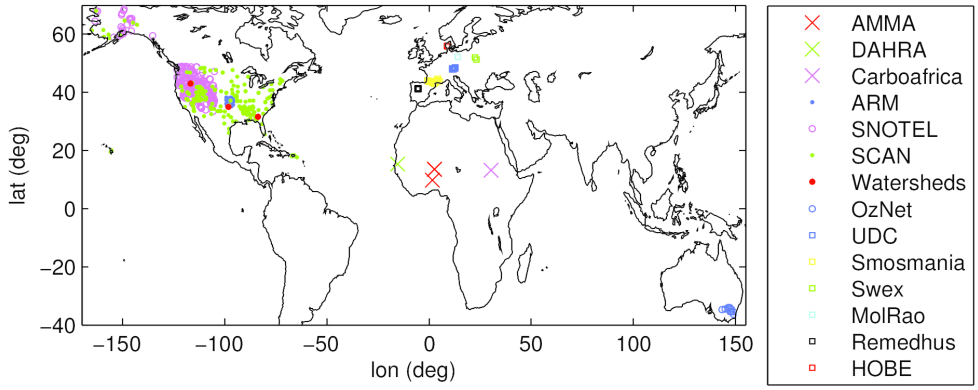

Figure 1. Locations of the in situ sensors.

3.1. North America

The SCAN (Soil Climate Analysis Network) network contains over 100 sensors/sites [54] and has been widely used to evaluate modeled and remote sensing soil moisture datasets. The sensors are located in agricultural regions with a relatively homogeneous landscape in many cases. The ARM (Atmospheric Radiation Measurement) network is run by the U.S. Department of Energy as part of the Atmospheric Radiation Measurement Climate Research Facility. SNOTEL (SNOw TELemetry) is a network [55] with a large number of stations (420) located mainly in mountainous areas. These networks are sparse networks that contain a large number of sites located at hundreds of kilometers from each other to sample soil moisture conditions at the continental scale. In addition, the United States Department of Agriculture (USDA) Agricultural Research Service (ARS) operates several dense networks. Three of these networks were used in the current study: Little River (LR) in Georgia, Little Washita (LW) in Oklahoma and Reynolds Creek (RC) in Idaho. These three watersheds cover a variety of soil types, crop types and climates. LR is the most humid site with an annual precipitation amount of 1200 mm, mostly covered by row crops and forests. With a sub-humid climate, the LW watershed is mostly used for agricultural activities with wheat crops and grass covers. Finally, RC is located in a mountainous area. Its climate is classified as semiarid, but frozen soils are very frequent in winter. These dense networks have already been used to evaluate AMSR-E and SMOS retrievals [56,57].

3.2. Europe

The Danish Hydrological OBsErvatory (HOBE) network [14,58] is situated in the Skjern River Catchment, and it is a dense network with 30 stations distributed to sample the different soil and environmental conditions on a scale of tenths of kilometers, comparable to the AMSR-E or SMOS footprints. Nearly 80% of this region is under intensive cultivation, intermixed with patches of spruce forest (∼10%), as well as heath/grassland (∼6%). The measurements of the different sensors are averaged before comparing to remote sensing or model SM datasets.

The Red de Estaciones de MEDición de la Humedad def Suelo (REMEDHUS) network [59] is located in a central sector of the Duero basin in Spain. The climate is semiarid continental Mediterranean, and the area is mainly flat covered by agricultural fields with some patchy forest. The sensors are installed horizontally at a depth of 5 cm.

The Lindenberg Meteorological Observatory-Richard Assmann Observatory (MOL-RAO) [60] site is situated close to Berlin, Germany. The Falkenberg site land cover is grassland. In this study, the horizontal probe at a depth of 8 cm has been selected.

Soil Moisture Observing System–Meteorological Automatic Network Integrated Application (SMOSMANIA) [61] is a network of Météo-France. The locations of the stations were chosen to form an Atlantic to Mediterranean transect, north of the Pyrenees mountain range, with the aim of sampling

the marked climatic gradient. In this study, the data measured by the probes at a 5-cm depth have been selected.

The Soil Water and Energy Exchange Poland (SWEX Poland) network [62] has been set up in particular to support calibration and validation of SMOS products over wetlands. In the current study, the probes placed horizontally at a 5-cm depth have been selected.

The Upper Danube catchment (UDC) SMOS network is located in southern Germany [63]. The sensors measuring soil moisture at the 5-cm depth layer have been used in the current study.

3.3. Africa

The AMMA-CATCH (African Monsoon Multidisciplinary Analysis - Coupling the Tropical Atmosphere and the Hydrological Cycle) observatory [64] sites in Niger and Benin have been used in this study. The Niger site (with five stations) is mainly tiger bush on the plateaus and fallow savannah and pearl millet crop fields on the sandy slopes (typical Sahelian rain-fed cultivated area). The annual rainfall amount ranges from 300–600 mm. The Benin site is located 400 km south of the Niger site, in a Soudanian climate. The observed annual rainfall is about 1200 mm. The vegetation is woody savannah and tropical forest. These two sites are processed as dense networks, and the soil moisture values have been averaged (five and nine stations for the Niger and Benin sites, respectively) before comparing to remote sensing and model data.

The Danish Natural Science Research Council field site near Dahra in Senegal, West Africa, (hereafter DAHRA, [65]), was established in 2002 to monitor the ecosystem properties of semiarid savannah grassland and their responses to climatic and environmental change. The site experiences a typical Sahelian climate (annual rainfall ranges from 200–600 mm), and land cover is typically Sahelian with grass species and coexisting trees (3% canopy cover). The sensors sampling horizontally at a 5-cm depth were used in this study.

The CARBOAFRICA (full project name: Quantification, understanding and prediction of carbon cycle, and other greenhouse gases, in Sub-Saharan Africa) [66] site is located in central Sudan and has collected meteorological and soil data from 2002–2012. The site is a sparse savannah in a semiarid region. The soil moisture measured by one sensor placed horizontally at a 5-cm depth is used in the current study.

3.4. Australia

Two subsets of the southeast Australia hydrological network (OzNet) have been used in this study. The Murrumbidgee Soil Moisture Monitoring Network (MSMMN) is located in southern New South Wales, Australia [67]. The Murrumbidgee Catchment shows a climate ranging from semiarid to humid. Land use includes dry land and irrigated agriculture, remnant native vegetation and urban areas. Here, the “sparse network” composed by the 18 initial stations providing soil moisture measurements integrated over the 0–8-cm soil layer has been selected because it has provided data since 2001 (hereafter, OzNet 0.00–0.08).

In addition, data from the smaller Yanco, Kyeamba and Adelong Creek catchments, which are located within the Murrumbidgee Catchment, have also been used [68]. The data started in 2006, but these second generation sites are equipped with 0–5-cm depth probes. Data from these sites have been processed as a single sparse network (hereafter, OzNet 0.00–0.05).

3.5. Local Evaluation Strategy

The AMSR-E NN SM (Section 4) and the other SM datasets (Section 2) have been evaluated against in situ measurements at the sites discussed above. RFI, snow and frozen soil filters (3 and 4 in Section 2.8) have been applied to the data. The in situ measurements have been compared to the closest point of the EASEv2 grid. To select a given in situ measurement, a time window of ±30 min around the time of AMSR-E acquisitions has been applied. Data from a given time are used to compute statistics metrics only if an SM value is available for all of the datasets, ensuring that all time series