CAPÍTULO 2:

Long-term SMOS soil moisture products: a comprehensive evaluation across scales and methods in the Duero Basin (Spain)

ABSTRACT

The European Space Agency's Soil Moisture and Ocean Salinity (SMOS) Level 2 soil moisture and the new L3 product from the Barcelona Expert Center (BEC) were validated from January 2010 to June 2014 using two *in situ* networks in Spain. The first network is the Soil Moisture Measurement Stations Network of the University of Salamanca (REMEDHUS), which has been extensively used for validating remotely sensed observations of soil moisture. REMEDHUS can be considered a small-scale network that covers a 1300 km^2 region. The second network is a large-scale network that covers the main part of the Duero Basin (65000 km^2). At an existing meteorological network in the Castilla y Leon region (Inforiego), soil moisture probes were installed in 2012 to provide data until 2014. Comparisons of the temporal series using different strategies (total average, land use, and soil type) as well as using the collocated data at each location were performed. Additionally, spatial correlations on each date were computed for specific days. Finally, an improved version of the Triple Collocation (TC) method, i.e., the Extended Triple Collocation (ETC), was used to compare satellite and *in situ* soil moisture estimates with outputs of the Soil Water Balance Model Green-Ampt (SWBM-GA). The results of this work showed that SMOS estimates were consistent with *in situ* measurements in the time series comparisons, with Pearson correlation coefficients (R) and an Agreement Index (AI) higher than 0.8 for the total average and the land-use averages and higher than 0.85 for the soil-texture averages. The results obtained at the Inforiego network showed slightly better results than REMEDHUS, which may be related to the larger scale of the former network. Moreover, the best results were obtained when all networks were jointly considered. In contrast, the spatial matching produced worse results for all the cases studied.

These results showed that the recent reprocessing of the L2 products (v5.51) improved the accuracy of soil moisture retrievals such that they are now suitable for developing new L3 products, such as the presented in this work. Additionally, the validation based on comparisons between dense/sparse networks and satellite retrievals at a coarse resolution showed that temporal patterns in the soil moisture are better reproduced than spatial patterns.

RESUMEN

El producto L2 de humedad del suelo proporcionado por el satélite *Soil Moisture Ocean Salinity* (SMOS) de la Agencia Espacial Europea (ESA) y el nuevo producto L3 del *Barcelona Expert Centre* (BEC) fueron validados para el periodo comprendido entre enero de 2010 y junio de 2014. Para ello se utilizaron dos redes de estaciones de medición *in situ* de humedad del suelo localizadas en España. La primera red que se utilizó fue la Red de Estaciones de Medición de Humedad del Suelo de la Universidad de Salamanca (REMEDHUS), la cual ha sido ampliamente utilizada para validar estimaciones de humedad del suelo de forma remota. REMEDHUS puede ser considerada como una red de pequeña escala, ya que cubre un área aproximada de 1300 km^2. La segunda red utilizada fue una red meteorológica (Inforiego) que cubre la parte principal de la Cuenca del Duero (65000 km^2). En esta red, existente en la región de Castilla y León, se instalaron sondas de humedad del suelo en 2012, proporcionando datos hasta 2014. Para validar ambos productos se realizaron comparaciones de las series temporales de humedad proporcionadas por el satélite y las estaciones, utilizando diferentes estrategias (promedio total, uso y tipo de suelo), así como el uso de los datos geolocalizados en cada ubicación. Además, se calcularon las correlaciones espaciales en cada fecha para algunos días específicos. Por último, se utilizó una versión mejorada del método denominado *Triple Colocationn* (TC), es decir, la *Extended Triple Colocation* (ETC), para comparar las estimaciones de la humedad del suelo por satélite e *in situ* con las estimaciones de humedad del suelo del modelo *Soil Water Balance Model-Green-Ampt* (SWBM-GA). Los resultados de este trabajo mostraron que las estimaciones de SMOS eran consistentes con las mediciones *in situ* entre las comparaciones de las series temporales, con coeficientes de correlación de Pearson (R) y un *Agreement Index* (AI) mayores de 0.8 para el promedio total y los promedios de uso de suelo, y mayor de 0.85 para los promedios de la textura del suelo. Los resultados obtenidos en la red Inforiego fueron ligeramente mejores que en REMEDHUS, lo que puede estar relacionado con el mayor tamaño de la red. Además, los resultados fueron mejores cuando se consideraron conjuntamente las dos redes *in situ*. Por el contrario, la validación espacial obtuvo malos resultados para todos los casos estudiados.

Este estudio mostró que la versión 5.51 del producto L2 mejoró la precisión de las estimaciones anteriores de la humedad del suelo de tal manera que ahora son adecuados para desarrollar nuevos productos L3, tales como los presentados en este trabajo. Además, la validación basada en comparaciones entre redes densas y redes dispersas con las estimaciones satelitales a una baja resolución espacial mostró que los patrones temporales de la humedad del suelo se reproducen mejor que los patrones espaciales.

Palabras clave: SMOS; Humedad del suelo; Validación; Triple Colocación Extendida; Modelización hidrológica

2.1. Introduction

Although soil moisture represents a small portion of the water volume of the planet, it is a key parameter in several hydrologic and atmospheric processes. Soil moisture controls the hydrological interactions between soil, vegetation and climate forcing, and it affects the balance of water and energy between the land surface-atmosphere interface. Moreover, soil moisture is crucial in agriculture because it represents the actual reservoir of the plant available water. In 2010, soil moisture was introduced as one of the Essential Climate Variables established by the World Meteorological Organization (WMO), the Global Climate Observing System and the Committee on Earth Observation Satellites, among others, considering it as "technically and economically feasible for systematic observation" (WMO, 2010). Monitoring soil moisture is critical for improving agricultural productivity, forestry, and ecosystem health. Much effort has been focused on measuring soil moisture using diverse approaches, ranging from *in situ* measurement networks and hydrological models to satellite sensors, which are the only practical means of providing global mapping. Data from individual *in situ* networks across the globe are disseminated though the International Soil Moisture Network (Dorigo *et al.*, 2011). A variety of sensors and systems have been used for measuring soil moisture, and several satellite soil moisture data services are currently operational. The microwave spectra has been identified as the most suitable for soil moisture sensing based on the contrast between the dielectric properties of liquid water and soil material. Passive and active techniques have been satisfactorily used (Schmugge *et al.*, 2002; Wagner *et al.*, 2013). Active sensors (radars) have a high spatial resolution (~1 km), whereas passive sensors (radiometers) typically have a low resolution (~40 km). Although synthetic aperture radar can achieve spatial resolutions of meters, its temporal resolution needs considerable improvement; furthermore, its signal is significantly affected by soil roughness and vegetation, which hinders the accuracy of soil moisture retrievals (Jackson *et al.*, 1996). The most common active sensors are the European Remote Sensing satellites, ERS-1 and ERS-2, with a spatial resolution of 50 km for the vertical polarization in the C-band (5.3 GHz), and its successor, the Advanced Scatterometer (ASCAT), which has a spatial resolution of 25 km and a temporal resolution of 1-2 days (Wagner *et al.*, 2007a). In terms of passive sensors, the first microwave radiometer was the Scanning Multichannel Microwave Radiometer (SMMR), which achieved a spatial resolution of 148 km x 95 km in the 6.6 GHz channel (Gloersen and Barath, 1977). Later, the Advanced Microwave Scanning Radiometer, AMSR-E (with six bands ranging from 6.9 to 89 GHz at the HH–VV polarization) operated on board the Aqua satellite from 2002 to 2011. AMSR-E was the first satellite sensor to incorporate soil moisture as a standard product with an accuracy goal of 0.06 m^3/m^3 and a spatial resolution of 74 km x 43 km (Njoku *et al.*, 2003).

Recent proposals for dedicated soil-moisture missions have chiefly relied on passive microwave techniques in the frequency band from 1 to 2 GHz (L-band) (Jackson *et al.*, 2012) due to its high soil-penetration depth, considering that vegetation is semi-transparent up to moderate densities (Jackson and Schmugge, 1991). Thus, the Soil Moisture and Ocean Salinity (SMOS) mission launched in 2009 by the European Space Agency (ESA). The baseline SMOS payload is an L-Band (1.413 GHz) Y-shaped two-dimensional interferometric radiometer, with multi-angular and full-polarimetric capabilities. SMOS is in a sun-synchronous polar orbit, and it provides global measurements of the Earth's brightness temperature (TB) with a spatial resolution of 43 km (Kerr *et al.*, 2010). Another mission is scheduled to launch in January 2015, the National Aeronautics and Space Administration (NASA)'s Soil Moisture Active Passive (SMAP) mission, which aims to retrieve soil moisture and the freeze/thaw state. The satellite includes an L-band radiometer and a synthetic aperture radar to improve the spatial resolution of soil moisture estimates at 36, 9 and 3 km spatial resolutions (Entekhabi *et al.*, 2010). Another instrument launched on board the international Aquarius/SAC-D mission by NASA and Argentina's space agency is primarily designed to retrieve ocean salinity. Aquarius uses an

L-band radiometer and a real aperture radar (Lagerloef *et al.*, 2008) that acquires measurements at a very coarse resolution (~100 km^2). Recently, the National Snow & Ice Data Center (NSIDC) released the Aquarius soil moisture products (http://nsidc.org/data/aquarius/).

Remotely sensed soil moisture has the advantage of covering large areas and identifying large-scale events, but it aggregates heterogeneities from local to regional scales, which renders validation difficult (Ochsner *et al.*, 2013). The point scale of *in situ*-based measurements contrasts with the footprint-scale of remote soil moisture estimates. Establishing credible ground validation approaches for soil moisture requires bridging the gap between the two resolutions (Crow *et al.*, 2012). An increasing number of permanent *in situ* soil moisture measurement networks have provided data for potential validation activities from the large scale to the regional/local scale (Crow *et al.*, 2012; Ochsner *et al.*, 2013). As the data increases (from both satellite and ground-based networks), the need to optimize and standardize the comparisons to reference measurements are more critical. Additionally, upscaling strategies from point measurements to the satellite footprint must be developed.

Following the classification of Crow *et al.*, (2012), two types of networks used for the validation of remote sensing products can be found: the large scale (>10000 km^2 extent) and the small scale (between 100 km^2 and 10000 km^2). The first type has the advantage of covering large areas and a larger range of land cover soil types, but it typically lacks sampling densities that provide multiple measurements per footprint. Conversely, small-scale networks have the advantage of higher spatial densities that provide multiple measurements within a single footprint and that allow for the examination of the sub-footprint scale. A large number of dedicated studies have validated products of soil moisture from different remote sensors worldwide, i.e., AMSR-E products (Njoku *et al.*, 2003; Sahoo *et al.*, 2008; Draper *et al.*, 2009; Gruhier *et al.*, 2010; Jackson *et al.*, 2010; Xie *et al.*, 2014), ERS soil moisture products (Wang *et al.*, 2009; Reimer *et al.*, 2012), ASCAT (Wagner *et al.*, 2013; Paulik *et al.*, 2014) and SMOS (Albergel *et al.*, 2012a; Albergel *et al.*, 2012b; Bircher *et al.*, 2012; Dall'Amico *et al.*, 2012; Dente *et al.*, 2012a; Jackson *et al.*, 2012; Sánchez *et al.*, 2012a; Petropoulos *et al.*, 2014; Rötzer *et al.*, 2014), among others. The validations based on ground-based networks show that the temporal soil moisture patterns of ground observations are well reproduced by satellite data in terms of the error estimates. However, the instantaneous spatial matching at a particular time step is still an unresolved issue, particularly for passive observations. The coarse resolution of these observations hinders a proper spatial comparison with scattered ground data, including for large networks in which the low-density sampling results in less than one observation within each grid cell. This issue is also apparent for small networks in which a small number of cells should be compared.

This work aims to thoroughly test the latest version of SMOS L2 (v.5.51) and a new SMOS BEC L3 v.001 data, obtained by quality-filtering and re-gridding of SMOS L2 v.551 data from ISEA to the 25 km EASE grid (details are given in Section 2.2.2) over the Duero Basin in Spain from January 2010 to June 2014. ESA L2 and BEC L3 series were compared with *in situ* soil moisture series from two networks in Spain. Considering the different spatial scales of each network, another objective was to study the influence of using different sets of stations with very different spatial distributions. Owing the very different observations compared, the validation methods were discussed. Then, a newly extended version of the Triple Collocation (TC) methodology was applied. Spatio-temporal correlations of the L2 and L3 soil moisture series were analyzed through a comparison with point-scale measurements at the two networks and area-averaged, land-use-averaged and soil-texture-averaged data.

2.2. Data sets

2.2.1. *In situ* soil moisture data set

For the remotely sensed data set validation, two networks were used, so-called "small-scale" and "large-scale" networks. The first network provided multiple measurements within a single footprint, and the latter network provided a single point observation overlapping a unique footprint over a wider area (Fig. 2.1).
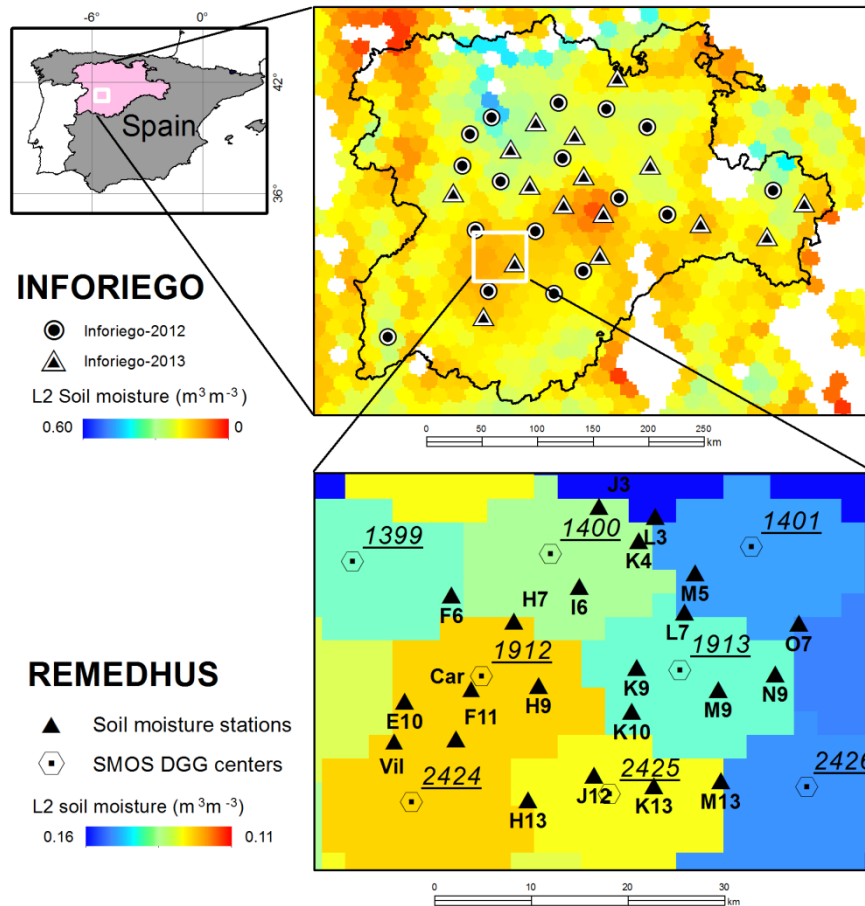

Figure 2.1. Network locations and distributions overlaid with the SMOS L2 product in the Duero Basin (Spain). The soil moisture map corresponds to 04/24/2011. DGG cells where soil moisture was missing are depicted in white.

2.2.1.1. REMEDHUS Network

The Soil Moisture Measurement Stations Network of the University of Salamanca (REMEDHUS) is located in the central Duero Basin (Fig. 2.1) over an area of 1300 km^2 (41.1° to 41.5°N and 5.1° to 5.7°W in Spain). This network is equipped with 23 automated stations that include capacitance probes (Hydra Probes, Stevens Water Monitoring System, Inc.), which measure hourly soil moisture in the top 5 cm of the soil. The homogeneity of the area makes it ideal for validating soil moisture products, as shown in many validation exercises of remotely sensed soil moisture data (Ceballos *et al.*, 2005; Wagner *et al.*, 2007b; Wagner *et al.*, 2008), including SMOS products (Sánchez *et al.*, 2012b; Petropoulos *et al.*, 2014; Piles *et al.*, 2014). The main land use is agriculture, with rainfed cereals in most cases, irrigated crops, vineyards

and forest-pasture areas. This area is nearly flat, ranging from 700 to 900 m.a.s.l. The climate of this region is continental semi-arid Mediterranean, with an average annual precipitation of 385 mm and a mean temperature of 12°C. A more detailed description of the equipment and the area of REMEDHUS can be found in previous works (Sánchez *et al.*, 2010; Sánchez *et al.*, 2012a; Sánchez-Ruiz *et al.*, 2014). Soil moisture at the satellite overpass times was obtained for each date for a total of 1642 data records for the entire period.

2.2.1.2.   Inforiego Network

The Inforiego network is located in the northwestern Iberian Peninsula over an area of approximately 65000 km^2 (Fig. 2.1). The network is deployed over the Duero River Basin. The area has a gentle topography that is surrounded by mountains to the north, south and east. The average height is 800 m.a.s.l. Similarly to the REMEDHUS area, the climate of the region is continental semi-arid Mediterranean, with an average annual precipitation of 450 mm and a mean temperature of 11.7°C. The main land uses in the Inforiego area are rainfed crops in winter and spring and irrigated crops in summer.

The soil moisture network was installed within a meteorological network that is dedicated to irrigation assessment (http://www.inforiego.org/). Out of the 56 original stations, 17 stations were selected between July 2012 and July 2013 (hereafter Inforiego-2012), and 16 stations were selected between August 2013 and August 2014 (hereafter Inforiego-2013). The criteria applied for the selection of the 33 stations were the distribution and localization of the main sub-basins of the Duero River Basin, the quality of the data collected by the weather stations, and the representativeness with respect to the environmental characteristics of the entire area, such as soil type and land use. At each of these selected stations, a Hydra probe was installed at a 5 cm depth alongside the meteorological equipment. As in the REMEDHUS network, an analysis of soil samples was conducted to validate the Hydra probes and to retrieve the soil properties at each station. As for the REMEDHUS network, the soil moisture values nearest in time to SMOS overpasses were considered. In total, 363 observations were available for Inforiego-2012 and 334 were available for Inforiego-2013.

2.2.2.   Satellite imagery

2.2.2.1.   SMOS L2 soil moisture product

SMOS is the first mission that is specifically dedicated to globally measuring the Earth's surface soil moisture (Kerr *et al.*, 2010) with a high target accuracy (0.04 m^{3}m^{-3}) and a revisit of 3 days. The SMOS Soil Moisture Level 2 User Data Product (SMUDP2 file), which is delivered through ESA, contains the retrieved geophysical parameters (i.e., soil moisture and optical thickness) and complementary parameters (number of T_B records used, atmospheric water vapor content, radio frequency interference and other flags) for the users consideration. The product is provided over the Icosahedral Snyder Equal Area Earth (ISEA-4H9) grid with equally spaced nodes at ~15 km, known as the Discrete Global Grid (DGG). A detailed description of the L2 algorithm used for the retrievals is provided in Kerr *et al.*, (2012). Since the launch of SMOS, the L2 algorithms have been modified and refined, resulting in more precise soil moisture retrievals. The Level 2 Soil Moisture Processor (L2SM) has been in operation with a stable configuration since November 2011. In June 2012, the L2SM v5.51 processor was distributed with improvements in the Radio Frequency Interference (RFI) detection and the dielectric constant model, which was modified from the Dobson model to the Mironov

formulation (Mironov *et al.*, 2004). Hence, the entire soil moisture data archives were aligned with the currently used L2SM processor baseline v5.51. A complete data set from January 2010 to the present is available using this improved version. Data of L2 soil moisture v5.51 from January 2010 to June 2014 over the Iberian Peninsula were used in this research. Two filter criteria were used. First, because each soil moisture retrieval is associated with a Data Quality Index (DQX) that represents the uncertainty of the retrieval, a threshold DQX of 0.04 (volumetric soil moisture units) was used to select the best-quality SMOS retrievals. This threshold restricts the retrieval to the theoretical accuracy of the SMOS targeted accuracy and fits the more demanding validation experiments. The second filter is based on the probability of occurrence of RFIs, which is updated in a new field (RFI_Prob flag) in the SMUDP2 product v5.51 for each grid node. Users can use this indicator as a predictor of the RFI corruption, ranging from 0 (no RFI probability) to 1 (expected corrupted data). The user-given threshold of RFI_Prob was variable among the literature, depending on the study site and data availability. In areas with a high occurrence of RFIs, the threshold could be negligible to maximize the spatio-temporal window of the data (Merlin *et al.*, 2013), but the usual thresholds range from 0.1 to 0.4 (Bircher *et al.*, 2013), meaning a probability of RFIs<40%. For this study, the use of several thresholds was explored. When a filter of 0.4 was applied, the number of data was reduced by 3.8% for ascending orbits and 3.1% for descending orbits. For a more restrictive filter of 0.3, the data reduction was 5.8% for ascending and 6.4% for descending orbits. Thus, prior to further analysis, three constraint levels of the SMOS L2 data set were tested regarding RFI filtering: applying a threshold of 0.4 or 0.3 and no filtering. Because there were no meaningful differences in the validation results, except in the RMSD (the average was reduced by 0.01%), the data were not filtered in an attempt to not overly restrict the availability of dates containing good data. Thus, only DQX filtering was applied, leading to a total of 835 (881) dates for ascending (descending) orbits for the REMEDHUS network, 1041 (1084) dates for ascending (descending) orbits of Inforiego-2012 and 1021 (1116) dates for ascending (descending) orbits of Inforiego-2013. Considering only the period when the stations were installed, 229 ascending and 248 descending dates for Inforiego-2012 remained, while 234 ascending and 233 descending dates remained for Inforiego-2013. Remarkably, the amount of data increases as the mission strengthens and consolidates over time. The REMEDHUS area is covered by 11 DGGs, and Inforiego area is covered by approximately 300 DGGs.

2.2.2.2.   SMOS BEC L3 Product

The L3 global product is generated and freely distributed by the SMOS Barcelona Expert Center (BEC) web service (http://cp34-bec.cmima.csic.es/). The BEC L3 product is developed from the operational ESA SMUDP2 product, and it includes geophysical parameters, a theoretical estimate of their accuracy, and a set of product flags and descriptors. The original L2 products were previously filtered such that grid points with negative soil moisture values and/or DQX values greater than 0.07 were discarded. Also, the failed retrievals or outliers are discarded. A weighted average is applied to bin the data to a 25 km EASE-ML regular grid. The products are provided in NetCDF format.

Soil moisture maps are calculated by BEC L3 product as follows:

$$\langle SM \rangle_k = \sum_{i=1}^{N} w_i SM_i \tag{2.1}$$

where $<SM_k>$ is the soil moisture value of each pixel, N is the number of the ISEA DGGs contained in the pixel, SM_i is the soil moisture value of the DGG, and w_i is a weighting factor determined by:

$$w_i = \frac{\frac{1}{DQX_i^2}}{\sum_{j=1}^{N} \frac{1}{DQX_j^2}} \tag{2.2}$$

w_i gives a specific weight to each DGG within the pixel, N is the number of the DGGs within the pixel, depending on the accuracy of the estimation of the original L2.

The pixel <DQX_k> is computed as:

$$\frac{1}{\langle DQX \rangle_k^2} = \sum_{i=1}^{N} \frac{1}{DQX_i^2} \tag{2.3}$$

where DQX_i is the DQX value of the DGGs within the pixel and N is the number of the DGGs.

The resulting L3 data were collected by applying the same filter to the L2 data (DQX<0.04 $m^3 m^{-3}$). For the period of study, 804, 1040, and 1012 dates were selected for the ascending orbit data and 851, 1101, and 1113 dates were selected for the descending orbit data for REMEDHUS, Inforiego-2012, and Inforiego-2013, respectively. For L2, taking into account only the usable period, 226 (233) dates of ascending data and 251 (235) dates of descending data were selected for Inforiego-2012 (Inforiego-2013). The REMEDHUS area is covered by 4 L3 pixels, and Inforiego is covered by approximately 100 pixels.

2.2.3. Modeled soil moisture

The fourth data set comes from the Soil Water Balance Model Green-Ampt (SWBM-GA) modeled soil moisture, which was previously validated in the REMEDHUS area (Gumuzzio *et al.*, 2015) and in different test sites across Europe (Lacava *et al.*, 2012; Brocca *et al.*, 2013) with satisfactory results. The modeled soil moisture time series were estimated with the parsimonious and lumped soil water balance model SWBM-GA, which was developed by Brocca *et al.*, (2008). The model employs the Green-Ampt equation for infiltration, a gravity-driven non-linear relationship for percolation (Famiglietti and Wood, 1994) and a linear relation between the actual and potential evapotranspiration (Doorenbos and Pruitt, 1977). Model simulations were performed over all the REMEDHUS and Inforiego stations using surface *in situ* measurements as a benchmark with an hourly temporal resolution. The model was previously calibrated with *in situ* measurements from both networks. The calibrated periods were 2008 to 2009 for REMEDHUS, July to December 2012 for Inforiego-2012, and August 2013 to January 2014 for Inforiego-2013; these periods are different from the period of validation for both networks in order to make the *in situ* and modeled data independent. The model inputs are rainfall and air temperature, which are computed each hour based on the weather station nearest to the soil moisture stations. Although the model was not applied as a spatially distributed model here and was not strictly an upscaling strategy, it embedded spatially averaged forcing data (e.g., precipitation and temperature). The simulations were set to represent the surface soil moisture, and outputs were used to validate satellite measurements through the Triple Collocation technique, as described in section 3.

2.3. Validation strategies

The L2 and L3 data sets were compared with the *in situ* data at the two spatial scales provided by the two networks. Spatial and temporal strategies of validation were adhered to. The first strategy followed in the validation is the temporal validation. For the temporal comparisons, the entire time series of the satellite soil moisture were compared with ground measurements that were previously collocated in time with the SMOS overpasses. Four comparisons were performed for both networks in the temporal comparisons: area-averaged, land-use-averaged, soil-texture-averaged, and point-scale measurements at each station with its collocated DGG (L2) and pixel (L3). The results are analyzed for the SMOS ascending and descending passes in order to assess their individual performances.

Soil texture and land use are closely related to soil moisture content (Crow *et al.*, 2012); hence, clustered *in situ* soil moisture time series per category (texture and land use) were compared with the remotely sensed estimates, and the consistency of the results at the point scale and the area-averaged comparisons was evaluated. Regarding the land use, four categories were taken into account (rainfed cereals, vineyard, forest-pasture and irrigated). For texture, four soil categories were considered. The relationship between the soil texture category and the textural classes is shown in Table 2.1. The distribution of the soil texture compositions of the 56 stations showed that, even though a slightly higher sand content is present at most of the stations, a noticeable texture variability can be used to assess the texture impact on the soil moisture retrieval accuracy.

Soil Texture Category	Textural Class	# Stations
Fine	Silty Clay Loam	4
	Clay Loam	4
	Clay	2
	Loam	10
Medium	Sandy Loam	17
	Sandy Clay Loam	7
Coarse	Loamy Sand	7
Very Coarse	Sand	5

Table 2.1. Number of stations by textural class and soil texture category.

The second strategy is the so-called spatial validation, even though there is no consensus regarding the terminology. Because of the different scales of the data sets, the term "spatial" may not be appropriate; thus, "daily-validation" is preferred because it characterizes a comparison of the data for each individual day. The meaning behind these terms relies on testing the soil moisture at all of the spatial locations at a given time. Spatially collocated SMOS and *in situ* observations for each time series day were compared. Each station with its corresponding (geolocated) DGG or pixel was compared for each day of data. A threshold of 16 concurrent observations was established as the minimum to compute statistics; this value is a compromise between the number of the stations in each network and preserving a robust comparison.

Both for the spatial/daily and temporal comparisons, a set of statistical metrics should be chosen. A correlation coefficient (R of Pearson) was selected for measuring the relationship between the two series. The relationship is defined by equation 2.4, where x_i and y_i are vectors of collocated (spatially/daily or temporally) *in situ* and satellite observations, respectively. The average soil moisture for those data sets is indicated by a bar.

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{2.4}$$

A measure of the differences between the data sets is provided by the root mean square difference, RMSD (Eq. 2.5), and the centered root mean square difference, cRMSD (Eq. 2.6), which avoid possible biases introduced by the RMSD between *in situ* data and satellite products (Taylor, 2001). Bias (Eq. 2.7, as the difference *in situ* minus SMOS), and standard deviation, std (Eq. 2.8), were also added to the metrics.

$$RMSD = \sqrt{\sum_{i=1}^{n}\frac{(x_i - y_i)^2}{n}} \tag{2.5}$$

$$cRMSD = \sqrt{\frac{\sum_{i=1}^{n}[(x_i - \bar{x}) - (y_i - \bar{y})]^2}{n}} \tag{2.6}$$

$$Bias = \frac{\sum_{i=1}^{n}(x_i - y_i)}{n} \tag{2.7}$$

$$std = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{2.8}$$

Finally, the Agreement Index, AI (Eq. 2.9), typically used in hydrological modeling (Willmott, 1982) was selected for assessing the accuracy of remotely sensed products. The AI represents an alternative way of computing differences; it is not a measure of correlation in a formal sense but rather a measure of the degree to which a model's predictions (remote sensing estimates, in this case) are free of error. The AI varies between 0 (total disagreement between predicted and observed values) and 1 (perfect agreement).

$$AI = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(|x_i - \bar{x}| + |y_i - \bar{x}|)^2} \tag{2.9}$$

R, std and cRMSD are complementary but not independent. These values are related by eq. 2.10 (Taylor, 2001). This relationship allows cRMSD, std and R to be displayed on two-dimensional plots known as Taylor diagrams. Std is displayed as a radial distance, and the correlation with *in situ* data is an angle in a polar plot. *In situ* data are represented by a point located on the x axis at R=1 and by the given std.

$$cRMSD^2 = std_x^2 + std_y^2 - 2 * std_x * std_y * R \tag{2.10}$$

The third strategy corresponds to the TC method, which combines three data sets to determine the errors in each. TC is generally used to resolve approximated linear relationships between different measurements (or representations) of a geophysical variable that are subject to errors. The method has been utilized in the context of calibration, validation, bias correction, and error characterization to allow comparisons of diverse data records from various direct and indirect measurement techniques, including *in situ*, remote sensing and model-based approaches (Su *et al.*, 2014). This method is useful for estimating the relative amount of error that affects different sources of geophysical measurements.

TC uses three temporally collocated time series of *N* observations to jointly provide sufficient constraints for determining the error variance. In this case, if errors in all three soil moisture products are mutually independent (uncorrelated) and if all three are degraded by random errors (Crow *et al.*, 2012), then this random error can be estimated from equations 2.11 to 2.13 (Scipal *et al.*, 2008).

$$e_{SMOS} = \sqrt{\left| \frac{\sum_{i=1}^{n}(SMOS_i - In\,Situ_i)(SMOS_i - Model_i)}{n} \right|} \tag{2.11}$$

$$e_{In\,Situ} = \sqrt{\left| \frac{\sum_{i=1}^{n}(In\,Situ_i - SMOS_i)(In\,Situ_i - Model_i)}{n} \right|} \tag{2.12}$$

$$e_{Model} = \sqrt{\left| \frac{\sum_{i=1}^{n}(Model_i - SMOS_i)(Model_i - In\,Situ_i)}{n} \right|} \tag{2.13}$$

This approach requires rescaling the data before applying the equations 2.11, 2.12, and 2.13; otherwise, the error estimates will be biased. Scipal *et al.*, (2008) proposed using linear least squares rescaling. Miralles *et al.*, (2010) showed that the approach is more effective for calculating error in soil moisture anomalies after removing the seasonal climatology effect. Moreover, the resulting errors should not be interpreted directly but instead normalized with the signal variability to express errors in terms of the signal-to-noise ratio, as recently proposed by Draper *et al.*, (2013) and McColl *et al.*, (2014).

The Extended Triple Collocation (ETC) is a new approach for TC calculations that was recently developed and initially applied to wind data sets (McColl *et al.*, 2014). This version of TC eliminates the need for de-biasing or rescaling. The choice to base the ETC calculation on raw soil moisture series differs from the recommended anomalies in the mean seasonal cycle (Draper *et al.*, 2009; Miralles *et al.*, 2010) or scaled values through standard deviation scaling of the input data (Dorigo *et al.*, 2010). Rescaling the measurement systems (e.g., by matching their temporal variances) will deliver biased RMSD estimates, since error estimation and calibration are fundamentally intertwined (Stoffelen, 1998). In addition, ETC is simpler to implement and adds no assumptions or computational cost to TC (McColl *et al.*, 2014). This method enables a direct estimation of the errors and correlations of each data set. The errors are estimated through the covariance between every two data sets, $Q_{ij}=Cov(X_i,X_j)$, using equation 2.14 if the data sets are not normalized or using equation 2.15 if the data sets are normalized. Correlations are calculated with equation 2.16, requiring no additional assumptions or computational costs. Using exactly the same assumptions as TC, ETC derives an additional performance metric, the correlation coefficient (ρ) with respect to the unknown target, to complement the RMSD (σ) (McColl *et al.*, 2014). In this regard, ETC provides new useful information that complements present data-validation strategies.

$$\sigma = \begin{bmatrix} \sqrt{Q_{11} - \dfrac{Q_{12}Q_{13}}{Q_{23}}} \\[12pt] \sqrt{Q_{22} - \dfrac{Q_{12}Q_{23}}{Q_{13}}} \\[12pt] \sqrt{Q_{33} - \dfrac{Q_{13}Q_{23}}{Q_{12}}} \end{bmatrix} \tag{2.14}$$

$$\sigma = \begin{bmatrix} \sqrt{Q_{11} - Q_{12}} \\ \sqrt{Q_{22} - Q_{12}} \\ \sqrt{Q_{33} - Q_{12}} \end{bmatrix} \tag{2.15}$$

$$\rho = \pm \begin{bmatrix} \sqrt{\dfrac{Q_{12}Q_{13}}{Q_{11}Q_{23}}} \\[12pt] sign(Q_{13}Q_{23})\sqrt{\dfrac{Q_{12}Q_{23}}{Q_{22}Q_{13}}} \\[12pt] sign(Q_{12}Q_{23})\sqrt{\dfrac{Q_{13}Q_{23}}{Q_{33}Q_{12}}} \end{bmatrix} \tag{2.16}$$

Triplets of *in situ*/modeled/SMOS data (L2 and L3 separately) were selected for the ETC approach. Both the spatial average and the spatially concurrent series were introduced into the ETC model, and σ and ρ were obtained for each triplet.

2.4. Results and discussion

2.4.1. Time series comparisons

The time series of soil moisture (Fig. 2.2; only the average REMEDHUS network is shown) reveal that the SMOS series have a larger dynamic range than the *in situ* data. *In situ* measurements have a maximum of 0.280 in winter to a minimum of 0.052 in summer with a std of 0.054 and a mean value of 0.141. However, the SMOS data ranged from 0.356 to 0.003, with a std of 0.075, for the ascending orbits. The maximum value (0.439) is higher for the descending orbits than for the ascending orbits; the minimum value is 0.009, and the std values are similar for both orbits. The SMOS data set followed the temporal dynamics of the REMEDHUS data set, although a particular underestimation can be observed, particularly in the dry periods. Additionally, a slight time lag between the SMOS response and the *in situ* measurements is noticeable, particularly in 2011 and 2012. The *in situ* observations response lags behind SMOS by 5 day. SMOS shows a quicker reactivity to rainfall events and dry-downs than the *in situ* probes. Both effects, underestimation and reactivity, were detected in previous experiments on validations in the same area (Sánchez *et al.*, 2012a) in a more general manner. The dry bias is also consistent with recent and ongoing calibration/validation studies on SMOS at other sites, which tend to indicate a general underestimation of the SMOS previous versions data with respect to 0–5 cm soil moisture measurements (Al Bitar *et al.*, 2012; Dall'Amico *et al.*, 2012; Gherboudj *et al.*, 2012; Sánchez *et al.*, 2012a; Merlin *et al.*, 2013). However, the

underestimation observed in this study is considerably smaller than that obtained with previous SMOS L2 versions. The reprocessed version 5.51 reduces outliers and leads to a more accurate estimation. The temporal patterns between SMOS L2 and L3 products matched the *in situ* soil moisture series of the two networks (area-averaged) and of the ascending and descending overpasses, with values of R~0.80 and AI~0.75; the correlation is higher for the Inforiego networks (Fig. 2.3). This result is very satisfactory, surpassing the standards of previous validations of SMOS with former reprocessing versions, e.g., for the v4.00: R of 0.43 in Albergel *et al.*, (2012b), 0.6 in Albergel *et al.*, (2012a), 0.73 in Bircher *et al.*, (2012), 0.3 in Dall'Amico *et al.*, (2012), 0.79 in Jackson *et al.*, (2012) and 0.7 in; for v5.01: 0.73 in Dente *et al.*, (2012a); and 0.28 for v5.51 in Rötzer *et al.*, (2014). The SMOS cRSMD reached the expected accuracy of 0.04 m^3m^{-3}, and the bias was always positive (i.e., indicating underestimation); the value is higher for both Inforiego networks (Fig. 2.3) than for REMEDHUS. This effect can be explained by the length of the time series, which is shorter for Inforiego, resulting in more accentuated seasonal effects that are smoothed in the longer time series of REMEDHUS. From the time series comparison, it can be stated that the SMOS soil moisture products were generally underestimated; this underestimation appears to be larger for the ascending products than for the descending (Fig. 2.3c-d) and for shorter time series (i.e., the Inforiego networks). Regarding the products and orbits, there was no difference between the L2 and L3 performances, and the ascending and descending series were similar in terms of R and cRMSD, with the bias noticeably smaller for the descending series (Fig. 2.3a-b).
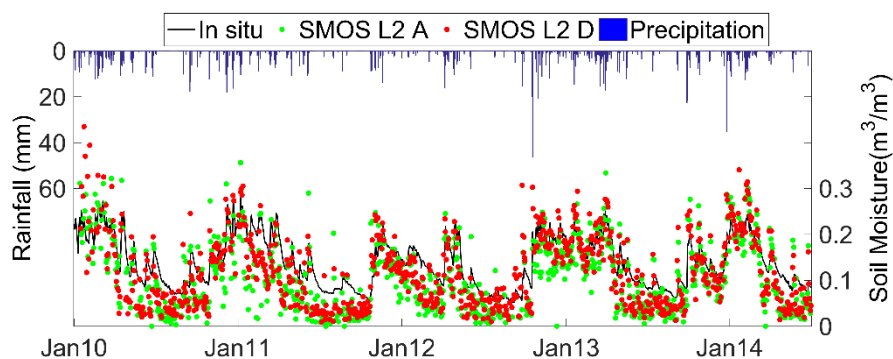

Figure 2.2. Soil moisture evolution of the *in situ* measurements (average of stations) and SMOS products (average of DGGs/Pixels), along with rainfall, during the study period (January 2010-June 2014) for the REMEDHUS network.

When computing the comparison of all 56 stations (REMEDHUS + two Inforiego), the statistics slightly improved. Increasing the number of stations and integrating the scales led to a smaller bias. It can be seen that the second period in the Inforiego network (Inforiego-2013) performed better than the first period, with a lower bias and cRMSD (Fig. 2.3a-d).

A spatially concurrent time series comparison of the two networks (each station with the corresponding L2 DGG or L3 pixel) using Taylor diagrams (Fig. 2.4, only L3 data was shown for convenience) and the REMEDHUS network revealed an R range between 0.41 and 0.8 for the L2/L3 and ascending/descending data; cRMSD is between 0.045 and 0.145. Similar metrics were obtained for all stations. The worst results were obtained for K9, H7 and I6. These poor results in H7 and I6 can be attributed to the very high sand content at these stations (85% and 90%, respectively) (Fig. 2.4). This result is consistent with a particular behavior of these soils that cannot be captured by the average values of SMOS. Panciera, (2009) and Crow *et al.*, (2012) found that soil moisture variability within a satellite footprint could be related to spatial patterns of the land use type and soil texture. Soils with higher sand content exhibited

persistently drier soil moisture conditions than soils with finer textures. The K9 station is located in an irrigated area, where the soil moisture pattern is controlled by repetitive irrigations events.

In 13 out of 23 REMEDHUS stations and in 26 out of 33 Inforiego stations, the bias was positive, indicating underestimations (not shown). Overall, the bias for each station in both networks ranged between -0.084 and 0.215.
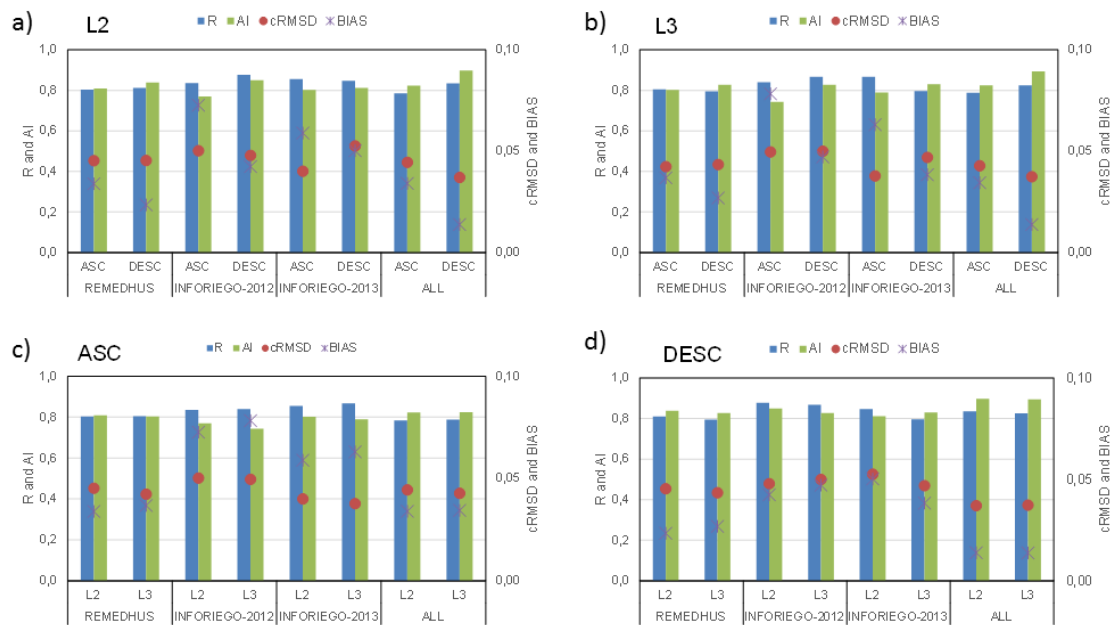

Figure 2.3. Results of the comparison between the spatially averaged series of soil moisture for the three networks in each usable time series. Figures a) and b) represent the L2 and L3 series for each network, respectively. Figures c) and d) represent the ascending and descending overpasses, respectively.
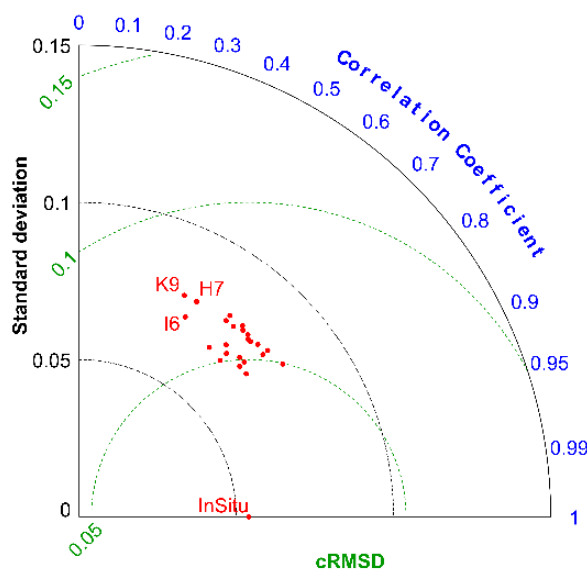

Figure 2.4. Results of the comparison between each station in the REMEDHUS network and its corresponding L3 descending pixel. Only stations with results discussed in the text are labeled.
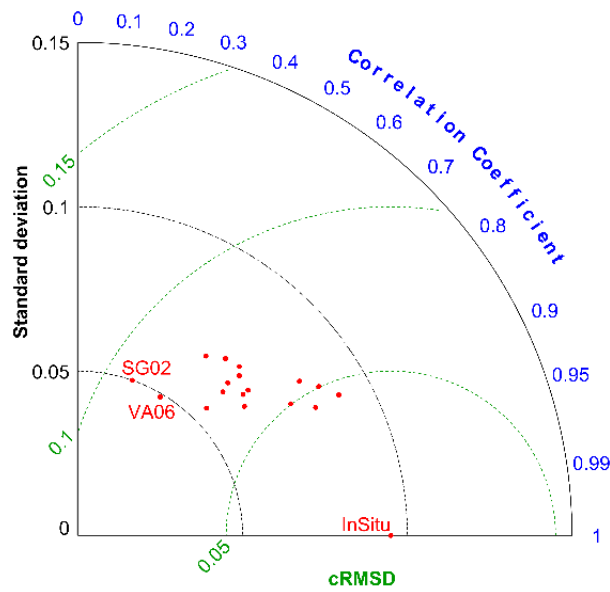
Figure 2.5. Results of the comparison between each station in the Inforiego-2012 network and its corresponding L3 descending pixel. Only stations with results discussed in the text are labeled.
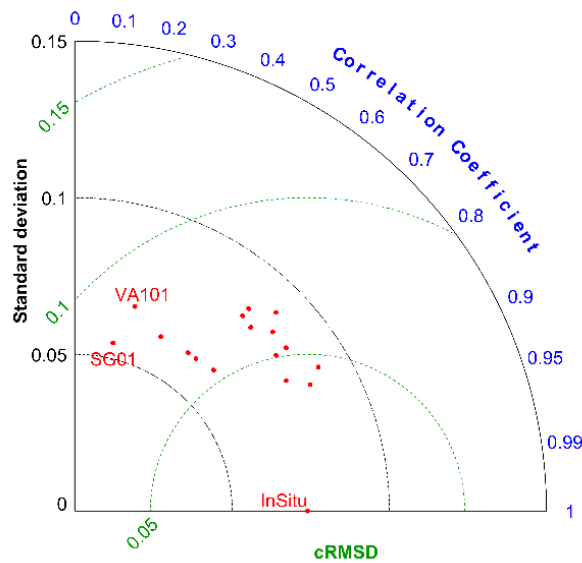

Figure 2.6. Results of the comparison between each station in the Inforiego-2013 network and its corresponding L3 descending pixel. Only stations with results discussed in the text are labeled.

The results of the comparisons, which account for soil use and texture (Fig. 2.7), showed a generally better match for the Inforiego network than for REMEDHUS in agreement with the results obtained in the area-averaged and point-scale analyses. Additionally, similar results for the ascending and descending orbits were detected (Fig. 2.7a-c). In general, values of R>0.80 and AI>0.60 were obtained, except for the irrigated land use and very coarse textures. Poor results for irrigated soils were also found in previous validation exercises in the area due to the high temporal rate of the water cycle in these soils. Fluctuating values of the actual water

content of the crops due to the high irrigation frequency could not be captured using the remotely sensed observations at a low temporal resolution (Piles *et al.*, 2014). Similarly, inconsistency between station data and SMOS was previously detected for coarse soil textures. The high sand content (coinciding with vineyard land use) led to high temporal variability in the soil moisture and to rapid soil moisture decrease in the top layer as a result of the quick infiltration and evaporation processes. More sandy soils have more pronounced effects, hindering a suitable match with SMOS readings. Nevertheless, a lower bias occurred for both very-coarse-textured soils and vineyard land use, which can be explained by the lower soil moisture content. In contrast, for the higher soil moisture contents (forest-pasture and fine textures), the bias is the highest. Rainfed land use and medium soil textures, which are common in the area, exhibited the best results in terms of R, AI, bias and cRMSD.
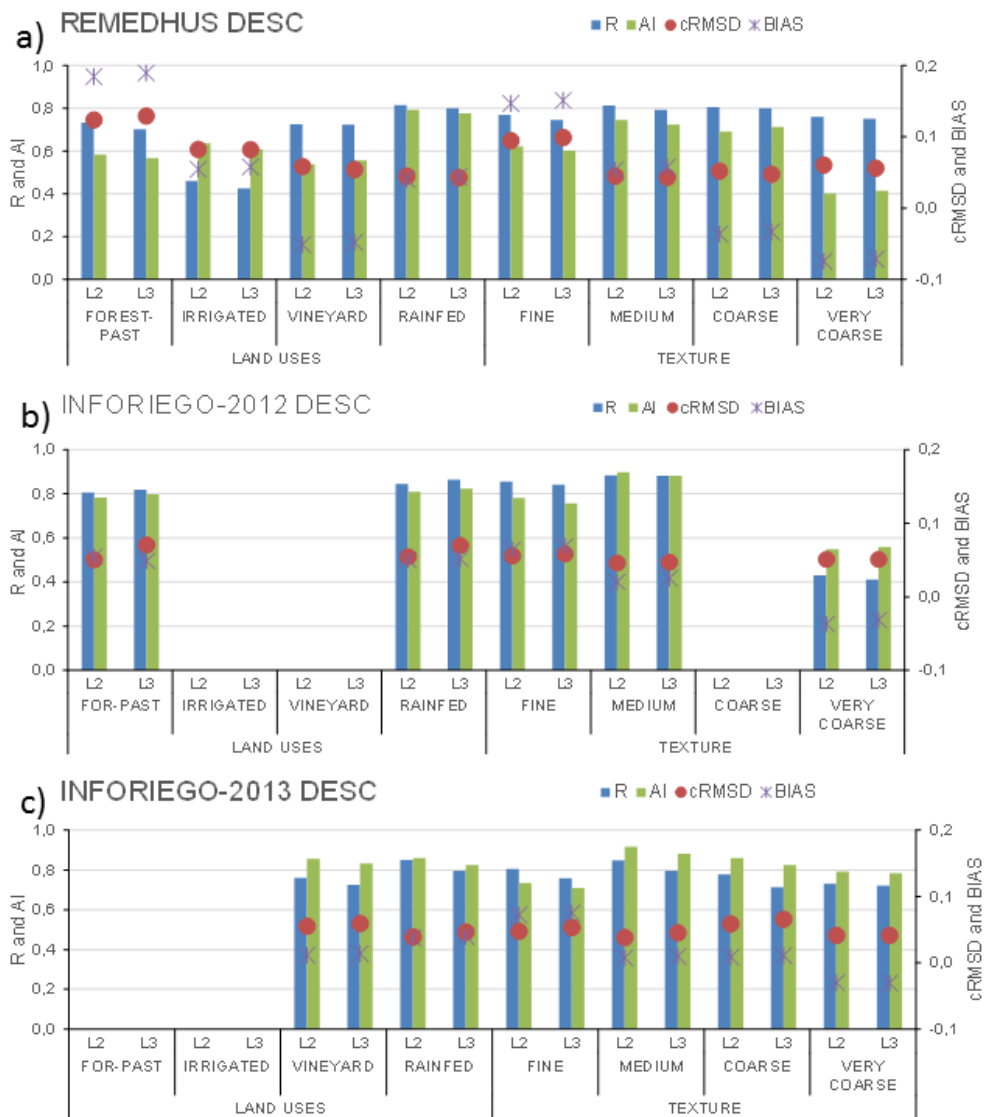

Figure 2.7. Results of the comparison between the temporal series of the soil moisture of the three networks after land-use and land-type clustering. a) REMEDHUS network, b) Inforiego-2012, and c) Inforiego-2013.

2.4.2. Spatial series comparison

Spatially collocated SMOS L2, L3 with *in situ* observations (REMEDHUS, Inforiego-2012 and Inforiego-2013) for each day were compared. Given a threshold of 16 concurrent pairwise observations for a calculation, the number of dates was reduced by one-fourth. For REMEDHUS, the L2 ascending product contained 463 days, the L2 descending product contained 461 days, and the L3 ascending and L3 descending products contained 584 and 598 days, respectively. For Inforiego, the number of days that meet the criteria also diminished: 23 days for L2 ascending, 28 days for L2 descending, 22 days for L3 ascending and 37 days for L3 descending for Inforiego-2012; and 5, 3, 1 and 3 days for Inforiego-2013. Of all the dates studied, only 13.9% of the comparisons have a p-value<0.05; negative and positive correlations existed for the series, and a very low AI was found for all products and networks.

Among the results with statistical significance (Fig. 2.8), the L2 and L3 products over REMEDHUS had correlation coefficients between R~-0.67 and R~0.72 and an agreement index between AI~0.00 and AI~0.63. The difference between the L2 and L3 products is negligible. The results are similar for the Inforiego networks, with a minimum R of -0.51, a maximum of 0.66 and an agreement index between AI~0.08 and AI~0.71. RMSD and cRMSD are similar in both networks, with a minimum of 0.045, a RMSD maximum of 0.214 and a cRMSD maximum of 0.201. The bias is negative in 85.21% of the cases, resulting in underestimation similar to the temporal validation (not shown). A representative example of the lack of significance of the spatial agreement is the BEC L3 ascending product compared to *in situ* measurements from Inforiego-2012. Out of the all dataset only two days have a correlation with statistical significance, where one day has a positive R and the other day negative.
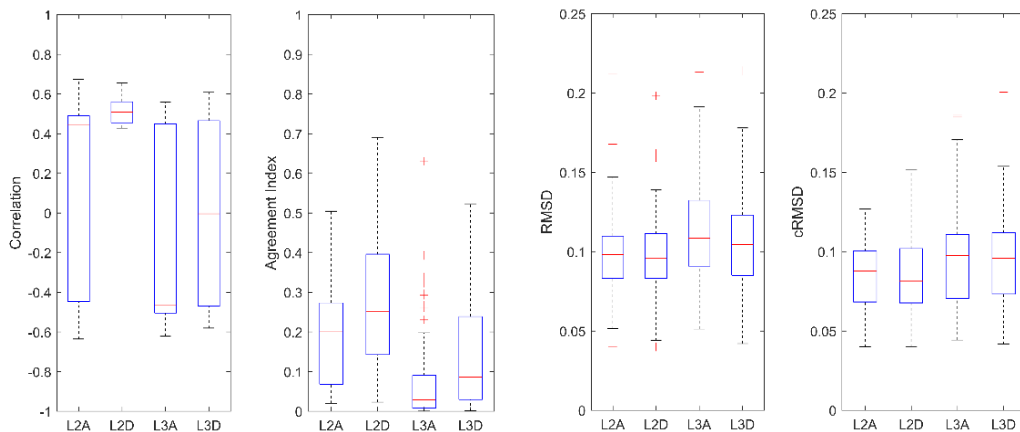

Figure 2.8. Box plot of the representing statistic of the spatial validation in REMEDHUS. For the L2 ascending (L2A) the number of significant days is 60, for L2 descending (L2D) is 64, for L3 ascending (L3A) is 99 and for L3 descending (L3D) is 82. Inforiego-2012 and Inforiego-2013 not shown because of the few significant days.

By merging the 56 stations, as conducted for the temporal validation, the significant results increase to 26.87%, but the other statistical metrics produce similar results to those studied in each individual network (not shown).

The daily validation showed poor results for both REMEDHUS (small-scale network) and Inforiego (large-scale network). It seems clear that given the spatial variability in the soil moisture fields, the mismatch between the ground sampling and the footprint scale of the remote estimates makes their direct comparison unsatisfactory. The variability between the

satellite products (15 km equally spaced grid for L2 and pixels of 25 km for L3) and the *in situ* measurements (point scale) is extremely large. Thus, a new daily correlation was performed (not shown) using a SMOS-derived product at a 1 km spatial resolution (Piles *et al.*, 2014). One may assume that if the spatial resolution of the remote estimations was enhanced, the results of the comparison may be improved; however, poor results were obtained, even at 1 km. No significant correlation was found at the 95% confidence level for either network or for the daily correlation at a 1 km resolution.

Validations based on comparisons between sparse networks and satellite retrievals at a coarse resolution showed that the temporal patterns of the soil moisture were better reproduced at both coarse and high spatial resolutions compared with the spatial patterns. As seen in most comparisons of satellite soil moisture, temporal soil moisture is well reproduced by satellite retrievals, in contrast to the spatial patterns. The lack of spatial matching should not be misattributed to retrieval uncertainty or to a degraded accuracy of remote soil moisture products; instead, the validation exercise is questionable based on such drastically different observations. In addition, the instantaneous value of several measurements of soil moisture over a large area could not be representative of the general behavior of this variable. It seems more efficient to compare the temporal evolution than a particular pattern. This issue may be overcome with dedicated and intensive field measurements, upscaling strategies or using spatially distributed surface models.

2.4.3. Extended Triple Collocation

Prior to the application of ETC, the modeled data set was validated against *in situ* observations at each station to examine the reliability of the model in simulating near-surface soil moisture in the REMEDHUS area and for the two Inforiego networks over the entire periods used in this study. In total, 1462 REMEDHUS, 213 Inforiego-2012 and 179 Inforiego-2013 daily data points were simulated. The results showed a good agreement between the modeled and *in situ* data sets, with average correlation coefficients of 0.76 for REMEDHUS and 0.89 for both Inforiego networks; acceptable RMSDs of 0.042, 0.051 and 0.040 m^3m^{-3} were obtained for REMEDHUS, Inforiego-2012 and Inforiego-2013, respectively.

The errors obtained with ETC are shown in Table 2.2. The σ is remarkably higher for SMOS (between 0.012 and 0.04 m^3m^{-3}) than for the *in situ* (between 0.003 and 0.007 m^3m^{-3}) and modeled (0.007 and 0.01 m^3m^{-3}) series. The correlations showed the same behavior, with the poorest ρ (between 0.818 and 0.902) for SMOS compared with *in situ* (between 0.942 and 0.993) and modeled (between 0.923 and 0.994) data. SMOS L2 and L3 series exhibited the highest errors and lowest correlations. These ETC results synthetize the mismatch between the spatial scales of the input data and clearly identify the complexity of comparing point-based *in situ* measurements and model simulations *versus* SMOS-averaged DGG soil moisture data.

No appreciable differences were found between the two satellite-derived products and the two networks studied. However, when all the stations were analyzed together, the correlations did not change, but the errors associated with the data set decreased by approximately one order of magnitude (reaching near-zero values). This result agreed with that obtained in the temporal comparisons, which showed that when all networks were jointly computed, the best results were obtained.

			σ (m³m⁻³)			ρ		
			SMOS	*In Situ*	**Model**	**SMOS**	*In Situ*	**Model**
L2	REMEDHUS	Ascending	0.0399	0.0166	0.0073	0.846	0.949	0.993
		Descending	0.0390	0.0170	0.0075	0.858	0.946	0.992
	Inforiego-2012	Ascending	0.0391	0.0098	0.0223	0.840	0.994	0.954
		Descending	0.0296	0.0182	0.0177	0.894	0.981	0.974
	Inforiego-2013	Ascending	0.0360	0.0159	0.0079	0.875	0.977	0.995
		Descending	0.0346	0.0171	0.0042	0.874	0.973	0.998
	All	Ascending	0.0093	0.0031	0.0018	0.816	0.961	0.990
		Descending	0.0066	0.0031	0.0016	0.867	0.961	0.992
L3	REMEDHUS	Ascending	0.0375	0.0165	0.0077	0.849	0.948	0.992
		Descending	0.0383	0.0175	0.0065	0.843	0.942	0.994
	Inforiego-2012	Ascending	0.0360	0.0102	0.0225	0.844	0.994	0.952
		Descending	0.0292	0.0201	0.0164	0.888	0.977	0.978
	Inforiego-2013	Ascending	0.0331	0.0154	0.0090	0.886	0.978	0.993
		Descending	0.0427	0.0159	0.0077	0.815	0.976	0.995
	All	Ascending	0.0090	0.0031	0.0018	0.818	0.962	0.989
		Descending	0.0067	0.0032	0.0013	0.860	0.959	0.994

Table 2.2. Errors (σ) and correlations (ρ) for the averaged *in situ*, model and satellite products resulting from the ETC method.
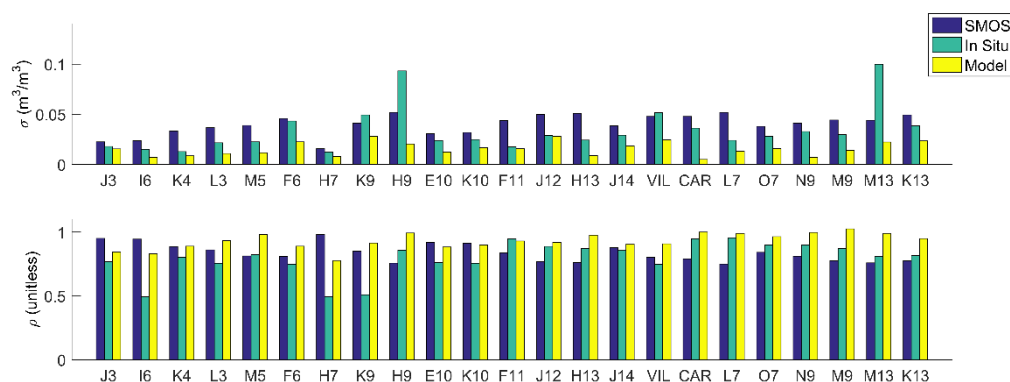

Figure 2.9. Errors (top) and correlations (bottom) obtained using the ETC method for the concurrent *in situ* REMEDHUS measurements, modeled data and SMOS L3 descending product.
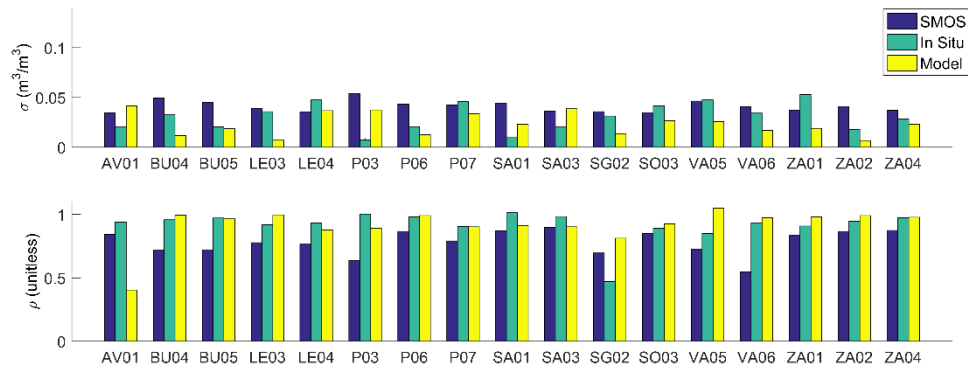
Figure 2.10. Errors (top) and correlations (bottom) obtained using the ETC method for the concurrent *in situ* Inforiego-2012 measurements, modeled data and SMOS L3 descending product.
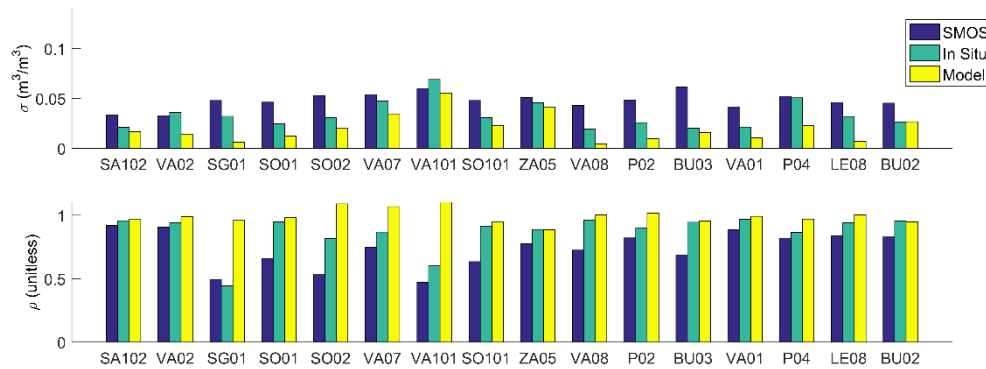

Figure 2.11. Errors (top) and correlations (bottom) obtained using the ETC method for the concurrent *in situ* Inforiego-2013 measurements, modeled data and SMOS L3 descending product. In SG01, the correlation of the model is outside of the expected range.

When the analysis was conducted for every station separately, although the errors are low (between 0.049 and 0.331 for SMOS, 0.001 and 0.264 for *in situ* and 0.001 and 0.301 m³m⁻³ for the model), the SMOS-associated errors are higher than the *in situ* and model errors (Fig. 2.9-2.11), similarly seen in the previous analysis using network-averaged soil moisture. When observing each station, two particular conditions occurred for which the ETC detected the highest errors. One case is over H9 and M13 (Fig. 2.9). These are the stations with the highest water content in REMEDHUS. For this case, the *in situ* data set had the highest errors because the ETC compared the high *in situ* water content with SMOS and the modeled soil moisture, which was remarkably drier. The other case is SG01 (Fig. 2.11), which is an extremely sand area. The ETC error was very high, and the ρ estimation resulted in a non-physical value, meaning that there is no relationship. The highly variable water content in the crops due to the hourly irrigation frequency could not be captured using the remotely sensed observations, as found in previous studies (Piles *et al.*, 2014).

A correlation that lies outside the expected range of -1 to 1 at some stations may be related to size problems in the sample. In cases where the signal-to-noise ratio is relatively low, which is common for remotely sensed soil moisture data sets, 1642 data points are not sufficient to obtain precise estimates of ρ. Note that this is a problem for classical Triple Collocation, not only ETC. For soil moisture, non-negligible cross-correlations between the errors and/or

representatively differences are common between global soil moisture data sets, potentially leading to violations of these assumptions (Draper *et al.*, 2013). Because this study is, to our knowledge, the first to apply ETC to soil moisture data sets, further research is needed for a thorough analysis.

2.5.   Conclusions

Comparisons between satellite, *in situ* and modeled data are critical in the assessment of new soil moisture products and in the release of new versions of reprocessed data. In the context of the SMOS mission, continuous validation activities using *in situ* networks have been performed over upgraded processing versions in several locations worldwide. The comparisons provide error estimates and a basis for modifying algorithms and/or parameters. For the existing and future missions devoted to soil-moisture sensing, there is a need to validate the agreement between the *in situ* data at each site (at various temporal and spatial scales) and the satellite estimates. The main challenge when comparing remote sensing data and *in situ* network observations is adequately addressing the mismatch between sparse ground-based observations (both for large- and small-scale networks) and footprint-scale soil moisture estimates.

This study conducted a validation of SMOS L2 and L3 products of soil moisture using two spatial scales provided by two *in situ* networks, along with modeled data at *in situ* stations. The products were quantitatively compared with data collected at two *in situ* networks: the REMEDHUS network, which is approximately 1300 km^2, and the Inforiego network, which is approximately 65000 km^2. Compared with previous SMOS validation exercises, this research took advantage of the latest reprocessing of the L2 product, as well as the release of new L3 products and the availability of 4-year SMOS data. Additionally, the new ETC methodology was applied with data simulated by the SWBM-GA model for both networks. A detailed analysis of soil types and land coverages was performed. Upscaling techniques and a new statistical comparison using the hydrological AI were performed.

The time series comparison results showed that the SMOS soil moisture products are generally underestimating the soil moisture observed at the studied sites in Spain, effect that appears to be larger for the ascending products and for the shorter time series. Regarding the products and orbits, there was no difference between the L2 and L3 performances; the ascending and descending series were similar. Comparisons at the point scale confirmed these favorable results, except at stations with particular conditions (very sandy soil and irrigated plots). A general dry bias is observed in the SMOS L2 and L3 series. Both for spatial/daily and temporal correlations and for ETC, the results obtained for the large-scale network were slightly better than those for the small-scale site, indicating that enlarging the spatial extent of the network resulted in a better match to averaged SMOS values. Thus, when all networks were jointly computed, the best results were obtained. However, the second period in the Inforiego network (Inforiego-2013) performed better than the first period. Thus, the improvement in the results for the Inforiego networks may be due to the increasing number of SMOS data over time, rather than an augmented spatial definition of the *in situ* references.

When the soil characteristics (soil texture and land use) were considered in the upscaling, better results were achieved, except for particular conditions (irrigated land use and very-coarse-textured soils). High sand content (coinciding with vineyard use) led to a quick soil moisture decrease in the top soil layer, hindering a good estimation. Conversely, a lower bias in very coarse soils and vineyard cover was found, probably due to their typically lower soil moisture content. The prevalent type of soils in the area (medium texture) and land use

(rainfed crops) exhibited the best results for all the statistics used. In this case, the data representativeness and favorable results are correlated.

The *in situ* soil moisture temporal patterns were correctly captured by the SMOS estimates in all cases and spatial domains, reaching the expected accuracy of the mission (4% vol.); however, the spatial patterns did not reach the expected accuracy. No significant correlation was found at the 95% confidence level for either network in terms of the daily correlations. From these results, it seems impossible (and likely useless) to depict the soil moisture spatial distribution at all stations at a single time. This result has been observed in all studies on soil-moisture retrievals from space using passive coarse-resolution radiometers, mostly due to the contrast between *in situ* measurements, which provide point-scale measurements under very different conditions at each location compared with the large footprint of the radiometer observations.

ETC is a new validation methodology that had not been applied before to soil moisture data sets. This technique proved useful for detecting particular conditions of soil moisture that are difficult to monitor from space. Although the obtained errors are relatively low, SMOS shows lightly worse results than the other data sets, which is confirmed throughout the ETC correlations.

The general results of this study showed that the newly reprocessed L2 products (5.51) improved the accuracy of soil-moisture retrievals, making them suitable for developing new L3 products as presented in this work.