This paper is the accepted version of the paper published on IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, VOL. 10, NO. 9, SEPTEMBER 2017, Pages. 3861 – 3873.

Manuscript history: received December 15, 2016; revised March 13, 2017; accepted May 1, 2017. Date of publication May 28, 2017.
Digital Object Identifier 10.1109/JSTARS.2017.2703629

VEGETATION WATER CONTENT RETRIEVAL BY MEANS OF MULTIFREQUENCY MICROWAVE AQUISITIONS FROM AMSR2.

E. Santi[1], S. Paloscia[1], P. Pampaloni[1], S. Pettinato[1].
Nomaki Tomoyuki[2], Mieko Seki[2], Keiji Sekiya[2], Takashi Maeda[3].

[1]Institute of Applied Physics - National Research Council (CNR-IFAC), Florence, Italy
[2]Remote Sensing Technology Center of Japan (RESTEC), Tokyo, Japan
[3]Japan Aerospace Exploration Agency (JAXA), Tokyo, Japan

ABSTRACT. In this study, two retrieval algorithms for estimating the water content of vegetation (VWC) in the range 0-10 Kg/m^2 from multifrequency microwave radiometric data have been implemented, with the purpose of contributing to the development of an all-weather VWC product for the satellite AMSR2 radiometer of the JAXA-GCOM-W mission.

The first algorithm estimates VWC through a semi-empirical combination of the polarization index (PI) acquired at X and Ku bands, while, the second algorithm attempts to improve the retrieval accuracy adding C-band data and using an Artificial Neural Network (ANN) method. The sensitivity to vegetation biomass of multi-frequency PIs at various frequencies, which was already pointed out in previous literature, has been first further evaluated with the support of the well-known tau-omega solution of the radiative transfer model and experimental data.

Two years of AMSR2 data collected on a wide portion of Africa, which includes a large variety of vegetation types and biomasses, have been considered for implementing and testing both algorithms. VWC reference data with the same temporal and spatial coverage of AMSR2, needed for validating the algorithm outputs, have been derived from SPOT4 NDVI, downsampled to the AMSR2 ground resolution. The test results provided a correlation coefficient, R, >0.88 with root mean square error, RMSE, <1.4 Kg/m^2 for the semi-empirical algorithm, and R=0.98 with RMSE<0.5 Kg/m^2 for ANN algorithm, thus demonstrating that both approaches are able to estimate VWC.

An independent validation of the two algorithms has been then carried out on the entire Australian continent, considering 4 periods of 16 days each in different seasons of 2013.

==

E. Santi, S. Paloscia, P. Pampaloni and S. Pettinato are with the Institute of Applied Physics (CNR-IFAC), Florence - Italy (e-mail: e.santi@ifac.cnr.it, s.paloscia@ifac.cnr.it, p.pampaloni@ifac.cnr.it and s.pettinato@ifac.cnr.it).
Nomaki Tomoyuki, Mieko Seki and Keiji Sekiya are with the Remote Sensing Technology Center of Japan (RESTEC), Tokyo, Japan. (e-mail: nomaki@restec.or.jp, M.Seki@restec.or.jp, sekiya_keiji@restec.or.jp).
Takashi Maeda is with the Japan Aerospace Exploration Agency (JAXA), Tokyo, Japan (email: maeda.takashi@jaxa.jp).

In this case, the algorithm outputs have been compared with VWC derived from MODIS 16 days averaged NDVI. The independent validation resulted in R =0.82 and RMSE around 1.5 Kg/m^2 for the semi-empirical algorithm, whereas the ANN allowed obtaining R=0.88 and RMSE around 0.86 Kg/m^2, demonstrating that the microwave data from AMSR2 and polarization indices can be legitimately used to produce vegetation maps on a global scale by separating several levels of biomass, without any need of further information from other sensors.

Index Terms—*Microwave emission, Polarization Indices, optical depth, Vegetation biomass, Artificial Neural Networks, AMSR2.*

1. INTRODUCTION

The interest in monitoring vegetation biomass and its dynamical distribution on a global scale is rapidly increasing, mostly due to the impact of canopy cover on the Earth's climate evolution and particularly in view of the changes that occurred in the past years and are reasonably supposed to continue in the future [1].

Satellite measurements offer the possibility of extended and frequent observations of the vegetation spatial and temporal distribution along with its degradation, which can be attributed to various sources (e.g. increasing urbanization, air, land and water pollution, as well as wild deforestation and increasing anthropic pressure) [1-2].

Presently, most of the algorithms for monitoring vegetation are based on visible and infrared indices, such as the Normalized Difference Vegetation Index (NDVI) [3] and Enhanced Vegetation Index (EVI) [4], which are well related to the "greenness" of the vegetation canopy and then, indirectly, to the vegetation biomass. For example, a relationship between NDVI and the Vegetation Water Content (VWC in Kg/m^2), which is a parameter assumed as representative of vegetation biomass, has been established in [5].

However, optical sensors have some significant limitations. Indeed, besides being only sensitive to the surface layers of canopies, they are strongly influenced by atmospheric effects. On the contrary, microwaves can penetrate clouds and, depending on frequency, canopy cover too, thus allowing a direct investigation of the aboveground vegetation. Moreover, taking advantage of the high sensitivity of microwaves to the water of the vegetation material permittivity, they allow a direct estimate of the VWC [6].

First studies to estimate the sensitivity of the microwave brightness temperature, Tb, to vegetation biomass were carried out since 1980s on the basis of ground based experiments and model simulations [e.g. 6, 7]. Later on it was observed that, depending on frequency and dimensions of vegetation scattering elements, Tb can increase or decrease as the biomass increases. This behavior can be explained considering the prevailing effect of absorption or scattering according to the relative dimensions of the plant elements in terms of wavelength [8]. Instead, the trend of the difference between the two linear polarization components was found to be independent of the vegetation type [9] and further information on vegetation biomass can be obtained by using the polarization index (PI), i.e. the difference between the two polarization components of Tb normalized to their average value. Indeed, the polarized emission from an almost homogeneous and smooth soil is smoothed by the volumetric effect of any vegetation type. This makes it possible to establish an inversion approach to retrieve vegetation biomass independently of crop type

Experimental relationships between PI and VWC of herbaceous crops, obtained using ground-based sensors at X and Ka band, were positively compared with model simulations in [10, 11] and showed a decrease of PI as a function of increasing vegetation biomass which was rather steep at Ka band and smoother at X band. A global map of vegetation cover based on the polarization difference at 37 GHz, obtained from Nimbus 7 data, was first shown in [12]. More recently, maps of VWC retrieved from X band PI were obtained in the context of an algorithm based on an Artificial Neural Network method developed for generating soil moisture and snow depth maps from the Advanced Multifrequency Scanning Radiometer AMSRE [13].

In addition to those based on the polarization difference, other approaches for retrieving VWC from multi-frequency satellite data have been proposed. For example, in [14], Shi et al. noted that the brightness temperatures measured at a given polarization with two adjacent AMSR-E frequency channels can be described by a linear function, which includes two coefficients, both independent of the underlying soil/surface signals and dependent only on vegetation properties. One is positively correlated to NDVI and is affected by the vegetation properties and the surface physical temperature, the other is negatively correlated to NDVI and is only affected by the vegetation properties.

Other studies assumed the vegetation optical depth (VOD) as an indicator of VWC. These studies were in general carried out on the basis of experimental relationships and the well-known tau-omega solution of the radiative transfer equation. A linear dependence between these two parameters obtained on the basis of L band data, was first revealed in [15] and then confirmed in [16, 17] for several crops, including soybean, corn and grass. At higher frequencies (X and Ka bands) however, a slight deviation from the linear trend with an inverse dependence on the square root of wavelength was found in [18]. Further studies [19, 20] showed that at low frequency (L band) the dependence of VOD on VWC was close to the linear one for low vegetation, and that a change in the volume density less than about 10% for a given VWC did not influence appreciably the VOD. However, at X band and higher frequencies, the VOD depends on the vegetation volume density with a consequent deviation from the linear dependence on VWC as foreseen in [18]. Indeed, in a real situation, an increase of the water content in a crop takes place due to both changes in vegetation height and volume density.

A procedure for retrieving VOD from radiometric data using the 6.6 GHz of the Scanning Multichannel Microwave Radiometer (SMMR) was suggested in [21]. Another study, based on the VOD derived from daily time series of Ku band AMSR-E data and other ancillary products, showed that the VOD was promisingly related with the vegetation indices obtained from satellite optical-infrared (MODIS) data, and phenology cycles [22].

A comparison between VOD derived from AMSR-E, TRMM, and SSM/I, and the AVHRR derived NDVI, carried out on a global scale for more than 20 years performed in [23], showed that the correlation with NDVI was substantial over low vegetation covers, but less significant over forests. Indeed, NDVI is fundamentally sensitive to chlorophyll and leaf biomass, whilst the VOD of forests is mostly dependent on vegetation biomass. It follows the need of using more penetrating low frequency systems than those available from AMSR-E to allow a more complete investigation on dense vegetation and forests.

On this point, a significant work regarding the information achievable from L band data has been carried in the framework of the Soil Moisture Ocean Salinity (SMOS), Soil Moisture Active Passive (SMAP), and Aquarius projects. For example, the SMOS VOD product generated by the prototype L2 algorithm available in 2013 over forests was found significantly related to forest height, with low seasonal variations [24 -26].

In summary, all these studies have pointed out that the capabilities of microwave brightness temperature in estimating the VWC depend on VOD, which is a function of the observation frequency and vegetation type. Hence, multi-frequency/polarization measurements are recommended for global observations of canopy covers. Indeed, the lowest frequencies of the microwave band (1-2 GHz) are the most appropriate for investigating dense vegetation and forests, while the frequency channels available from AMSR2 and the previous AMSR-E (C to Ku bands) are well suitable in monitoring VWC of relatively low vegetated areas.

The purpose of this paper was to provide an all-weather AMSR-2 algorithm for providing a vegetation cover product for the Global Change Observation Mission of the Japan Aerospace Exploration Agency (JAXA). The need of obtaining a product related to AMSR-2 data only, and independent of other sensors, led to some compromise with respect of the accuracy achievable on forests. On the other hand, agricultural crops and grasslands represent an important carbon sink/source due to their quick growth cycle and influence significantly the carbon cycle with their seasonal evolution.

To this end, we have implemented two retrieval algorithms. The first algorithm relates VWC to a semi-empirical combination of the Polarization Index (PI) at X and Ku bands, while the second algorithm makes use of Artificial Neural Networks, fed by PI at C, X and Ku bands, for retrieving VWC. Both algorithms have been implemented and tested considering two years of AMSR2 acquisitions (from July 2012 to June 2014) on a test area in Africa, extended 20° in Latitude and 1° in Longitude, and covering a large variability of vegetation types and biomasses, from bare surfaces (Sahara desert) to dense vegetation (Equatorial forest).

An independent validation of the algorithms has been then carried out considering the entire Australian continent, and all the AMSR2 data collected in different seasons in 2013. These areas were selected as representative of the entire annual cycle of the vegetation.

Considering the unavailability of distributed ground data covering both areas, the implementation, test and validation of the algorithm were carried out by comparing the VWC derived from microwaves with the optical vegetation index (NDVI), extending the approach proposed for SMOS in [24].

After a short description of the test areas and data considered in this work, and provided in section 2, a sensitivity analysis of AMSR2 data to VWC is carried out in section 3. The implementation and test of both algorithms on the African site is described in section 4. Lastly, section 5 contains the independent validation on the Australian site and the discussion.

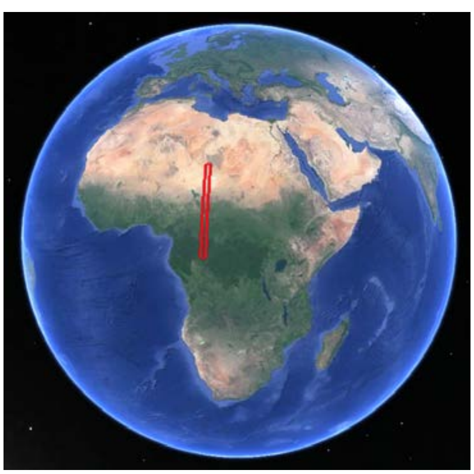
2. TEST SITES AND EXPERIMENTAL DATA

For implementing and testing the algorithm, a suitable test site was identified in a wide portion of Africa (0-20°N/16°-17°E), extended from the Sahara desert to the Equatorial forest, which includes a very high variability of vegetation types and biomasses. The area was also chosen due to the presence of large and homogeneous regions, in order to mitigate the effects related to the coarse ground resolution of AMSR2. The area extension is close to 300,000 Km^2 with a negligible percentage of rivers and open waters. Sandy and stony soils are present in the northern part, while vegetation cover increases moving to south, passing from grasslands to rainfed croplands and Savanna, which represent more than 60% of the vegetation coverage in the area. In the southern part, equatorial forests represent the dominant vegetation.

The independent validation of the algorithm was instead carried out considering the entire Australian continent, which extent is 7,617,110 km^2, with a large percentage of deserts, especially in the central part.

All the available AMSR2 Tb data collected on the African area from July 2012 to May 2014, in both ascending and descending orbits, were considered. The independent validation on Australia was instead carried out considering the AMSR2 overpasses on Australia in four different time periods of 16 days each, in January, April, July and October 2013, with the aim of representing the whole annual cycle of vegetation. In both cases, we used Level-1R AMSR2 data containing spatial-resolution matched brightness temperatures and corresponding swath data with geolocation information. With respect to the original L1B product, an antenna pattern deconvolution is applied to these data basing on the Backus-Gilbert (BG) method, in order to identify different field of views of multi frequencies [27].

All the AMSR2 data considered in this study were downloaded through JAXA GCOM data portal (http://suzaku.eorc.jaxa.jp/GCOM_W/).


(a)


(b)

Figure 1. The two areas identified from google Earth: Africa for implementing and testing the algorithm (a) and Australia for the independent validation (b)

One of the main problems encountered in this work, was in defining the reference VWC data for testing and validating the algorithm. In-situ measurements of VWC with similar spatial and temporal coverage of AMSR2 on both areas were unavailable, as in general distributed VWC measurements are barely available on a large or global scale.

An alternative strategy for validating the algorithm was therefore defined by deriving VWC values from NDVI obtained from optical satellite sensors as frequently done in similar studies [e.g. 23]. This procedure has certainly some limitations that have to be properly addressed, since optical bands are sensitive to the vegetation greenness more than to its water content, which is instead the driving parameter of microwave emission from vegetated areas. This difference can therefore introduce some uncertainties in the relationship VWC–NDVI, especially in the phase of vegetation senescence, when leaves are still green but plants are drying, or vice versa when leaves turn yellow but VWC is still high, depending on crop type, and, above all, in forests.

Nonetheless, except for these limit cases, past research pointed out that NDVI of low vegetation is in general well correlated with VWC, and linear relationships between VWC and NDVI can be found in a wide range of vegetation types [28-29]. It should be remarked that the relationship established in [29] for agricultural fields and grasslands in Italy, is very close to the one established in [28], which was developed on the USDA-ARS fields in USA, thus confirming the "robustness" of the correlation between VWC and NDVI.

In this study, we referred to the relationship presented in [29], which relates VWC to NDVI as follows:

$$VWC = 12.86 * NDVI - 2.25 \ (Kg/m^2) \qquad (1)$$

VWC was computed according to eq. (1) from SPOT4 10-days averaged NDVI on the Africa test site and 16 days averaged NDVI derived from Aqua MODIS acquisitions for the Australian site. The first set of data was obtained from http://www.vito-eodata.be/, while MODIS 16-Day L3 Global 1km data product from Aqua satellite (MYD13A2) were downloaded from the National Snow and Ice data center (NSIDC - http://nsidc.org/).

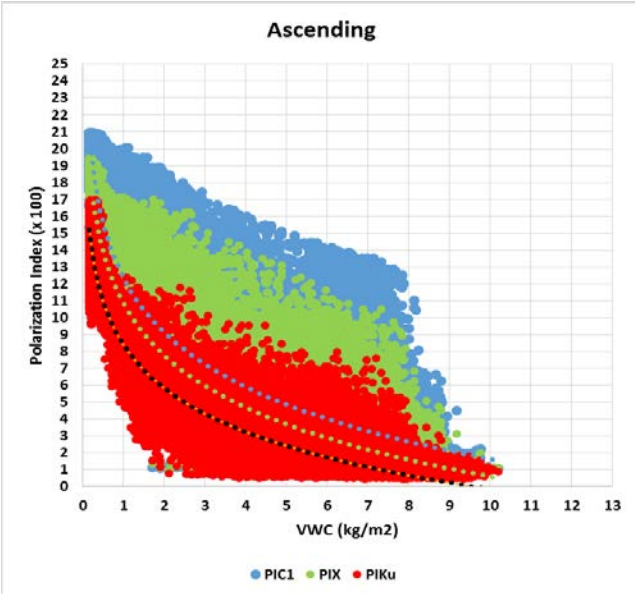
AMSR2 acquisitions were separated in ascending and descending orbits and averaged on the same time intervals of the corresponding NDVI, while NDVI data were downscaled at the AMSR2 resolution.

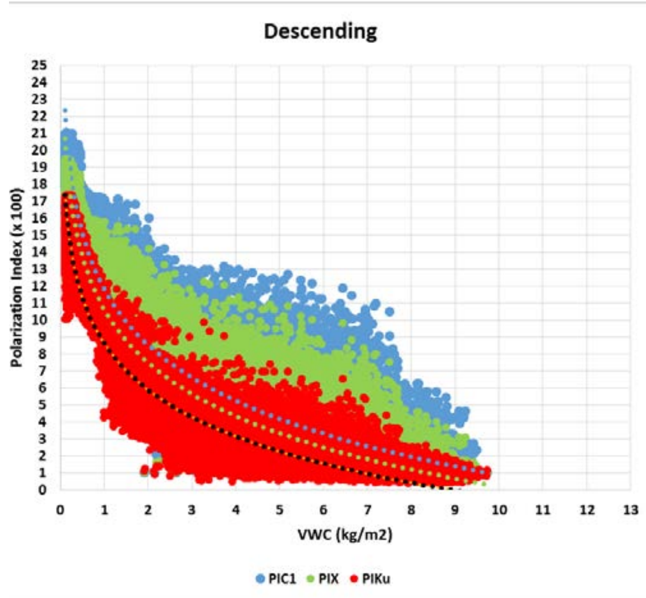
Both AMSR2 (level 3) and NDVI data were then resampled onto grids of nodes equally spaced at 0.1° x 0.1°, for a total of 225x12 nodes in Africa and 501x701 nodes in Australia, in order to co-locate Tb and NDVI for the algorithm implementation. The resulting dataset for the African site was composed by more than 132,000 co-located values of Tb (including all frequencies for the ascending and descending orbits) and corresponding NDVI. A much larger dataset was obtained for the Australian site, for which more than 3,500,000 co-located Tb and NDVI values were obtained.

Although beyond the scope of this paper, a further verification of the disaggregation technique applied to AMSR2 data was carried out by comparing L1R data considered in this work to the corresponding L1B data disaggregated using the SFIM technique proposed in [30]. The different fields of view of AMSR2 at the three frequencies, (from 62x35 Km^2 at C band to 12x7 Km^2 at Ka band), may lead indeed to large uncertainties in combining multifrequency acquisitions in the retrieval algorithm. The comparison of these data, carried out in Section 3, confirmed the effectiveness of both disaggregation procedures in mitigating the problem of different spatial resolution.

3. SENSITIVITY ANALYSIS

In the first step of this work, we evaluated the sensitivity of AMSR2 brightness temperatures to VWC derived from NDVI. The comparison between Polarization Indices (PI=(TbV-TbH)/(½(TbV+TbH))) computed at C (6.8 GHz), X (10.65 GHz), and Ku (18.7 GHz) bands and VWC derived from SPOT4 NDVI for the African site is displayed in Figure 2 (a) and (b) for ascending and descending orbits, respectively. In this figure, PI values are multiplied by 100. It should be noted that AMSR2 carries onboard two C band channels, operating at 6.8 and 7.23 GHz, respectively, to reduce the effects of radio-frequency interferences (RFI). Being the African site almost free from RFI, the Tb acquired in these two channels are well correlated (correlation coefficient, R, and slope ~ 1). The difference between the two PIs was therefore almost negligible, and thus only PI at 6.8 GHz is represented in the figures. As expected, the highest PI values correspond to desert zones and the lowest one to vegetated surfaces, with a gradual decrease according to the increase of biomass covering the surface. The scatterplots confirm the correlation between these quantities that had already been pointed out in previous works [10, 13].





(a) (b)

Figure 2. PI at C, X and Ku bands vs NDVI-derived VWC for the data collected on the African site in (a) ascending and (b) descending orbits. Blue, green and black trendlines refer to the PI_C, PI_X and PI_{Ku} band, respectively. Only the PI at the lower C band frequency (i.e. 6.8 GHz) is displayed.

Although the Tb values collected in ascending and descending orbits of the same day show marked differences, due to the surface temperature variation between the two overpasses (ascending around 01:00 am, and descending around 01:00 pm, local time), the behaviors of PI shown in Figure 2 (a, b) are similar. Indeed, the normalization for the surface temperature is intrinsic in the PI definition. Table 1 summarizes correlation, slope and intercept of the logarithmic regressions in both ascending and descending orbits.

We can note that the trends of PI at various frequencies are similar; however, the different penetration depth at each frequency causes a small shift in the curves. In both ascending and descending orbits, R values slightly increase with the frequency.

In case of extremely dense vegetation (i.e. equatorial forests), corresponding to VWC $\gg$8 kg/m^2, any further sensitivity of both microwave emission and optical data to biomass is unrealistic, due to the low penetration of radiation at these frequencies. This is evident when analyzing separately the data collected on area covered by equatorial forests, where a low correlation exists (R<0.3) between PI at all frequencies and VWC. The areas for which the sensitivity of PI to biomass is poor, and the estimate of VWC is therefore unfeasible, were identified by using a simple threshold criterion [13], which was applied to the measured PI at X band (≤ 0.5).

Lastly, negligible differences (affecting only the second digit of R) were found deriving PI from the AMSR2 L1B data disaggregated with the SFIM technique, instead of L1R. This confirmed the effectiveness of both techniques in relating each measured Tb to the nominal sampling of 0.1° x 0.1°.

Table 1. R, slope and intercept values of the logarithmic relationships between PI and VWC

	R		Slope		intercept	
	Ascending	Descending	Ascending	Descending	Ascending	Descending
C band	0.86	0.87	-4.57	-4.71	12.05	11.90
X band	0.88	0.89	-4.36	-4.52	10.65	10.66
Ku band	0.88	0.89	-3.75	-3.90	8.18	8.64

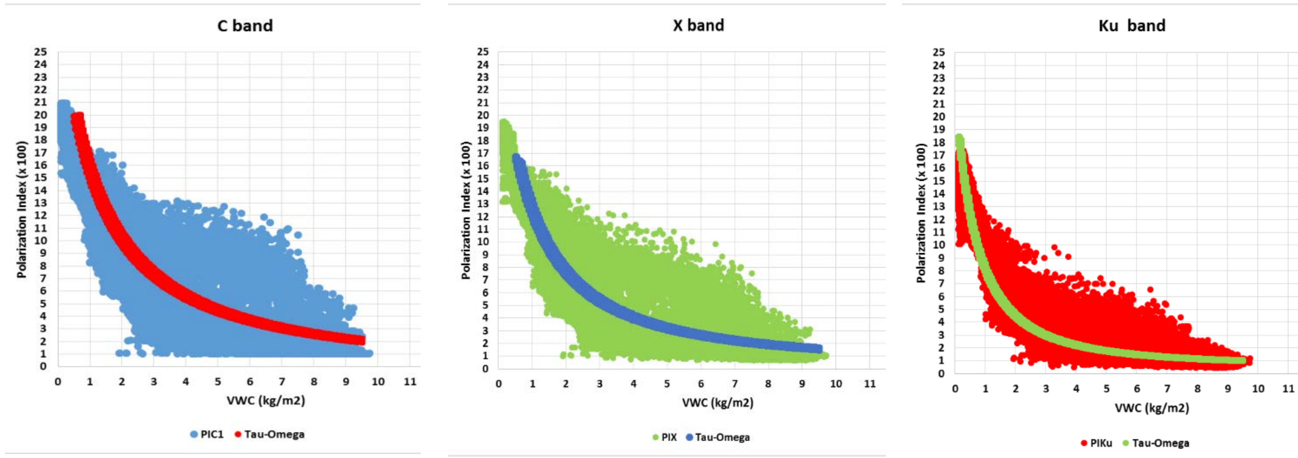


Figure 3. PI at C, X and Ku bands as a function of VWC from experimental data and tau-omega model simulations.

The obtained experimental relationships of PI as function of VOD were compared in Figure 3 with simulations carried out through the well known τ-ω solution of the radiative transfer model [17] (τ =*VOD*), by keeping constant the single scattering albedo (ω=*0.01*), the surface temperature, (*Ts*=290 K), the soil moisture (SMC= 15%), and relating VOD to VWC as in eq. 3 [18, 19]:

$$VOD = \frac{k}{\sqrt{\lambda}} * VWC \qquad (2)$$

where λ is the wavelength (m) and k is a constant depending on vegetation type, here assumed equal to 0.16 [18].

The τ- ω model was then applied for assessing the relationships between τ and VWC derived from NDVI.

For each set of AMSR2 PIs, the optimal τ and ω values that minimized the absolute error between measured and simulated PI were computed using an iterative minimization based on the Nelder-Mead algorithm [31]. The surface temperature required as model input was derived from AMSR2 Tb at Ka band, V pol., using a linear relationship [13, 32], and the soil moisture was also derived from AMSR2 through an updated implementation of HydroAlgo for AMSR2 [33].

These relationships are displayed in Figure 4, where VWC derived from NDVI is represented as a function of τ obtained from the minimization. Reasonably, the same VWC corresponds to τ values that increase with the frequency.

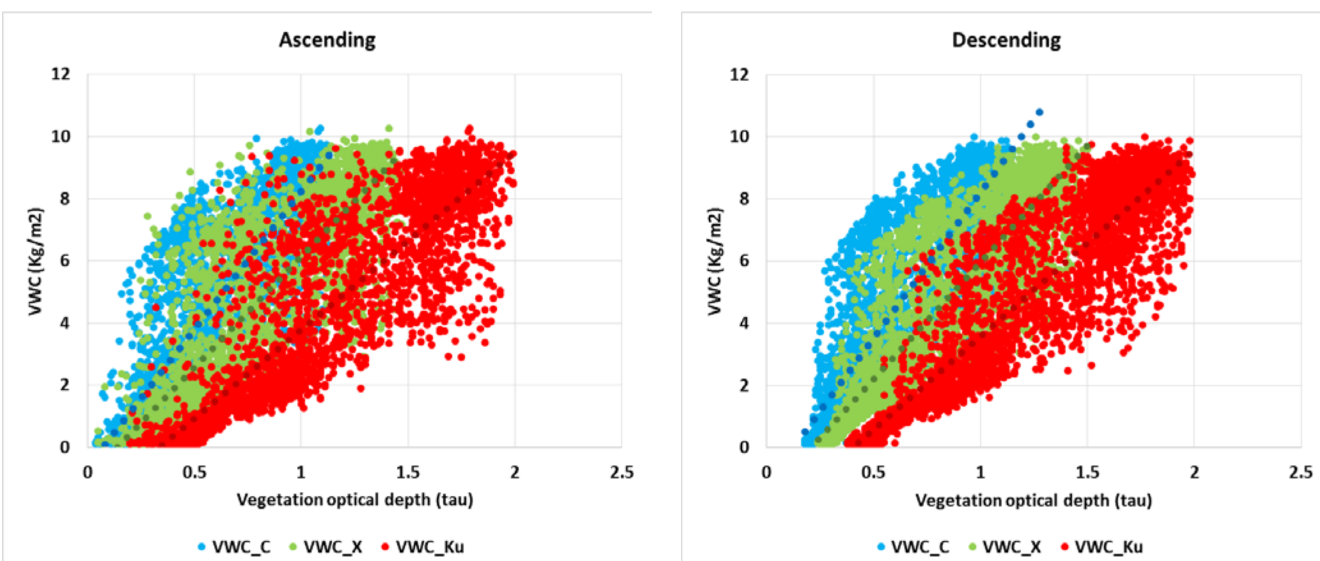


Figure 4. VWC derived from NDVI as a function of τ at C, X and Ku bands in ascending and descending orbits, respectively.

Table 2. R, slope and intercept values of the linear relationships between VWC and τ (VOD).

	R		Slope		intercept	
	Ascending	Descending	Ascending	Descending	Ascending	Descending
C band	0.81	0.81	9.24	9.37	-0.98	-1.17
X band	0.81	0.84	7.34	7.51	-1.25	-1.54
Ku band	0.82	0.87	5.80	5.99	-2.1	-2.40

The relationship presented in equation 3 and the results in Figure 4 need some more details. A linear dependence of τ on VWC was revealed by experimental results at L band in [34] and then confirmed by [16, 17] for several crops, including soybean, corn and grass. At higher frequencies (X and Ka bands), a logarithmic correlation between these two quantities was reported in [18]. However, further studies [19] demonstrated that, for relatively low values of VWC (< 6 Kg/m^2), the error in assuming a linear dependence is small also at the frequencies investigated in this study. A linear relationship for the AMSR2 frequencies was also considered in [35]. On the other hand, the intercept of the straight lines different from zero for bare soil can be interpreted as the effect of surface roughness, which affects the extinction coefficient similarly to vegetation cover [36]. The slopes, intercepts and correlation coefficients of the data presented in Figure 4 are summarized in Table 2, showing again negligible differences between ascending and descending orbits.

4. ALGORITHM IMPLEMENTATION (AFRICA)

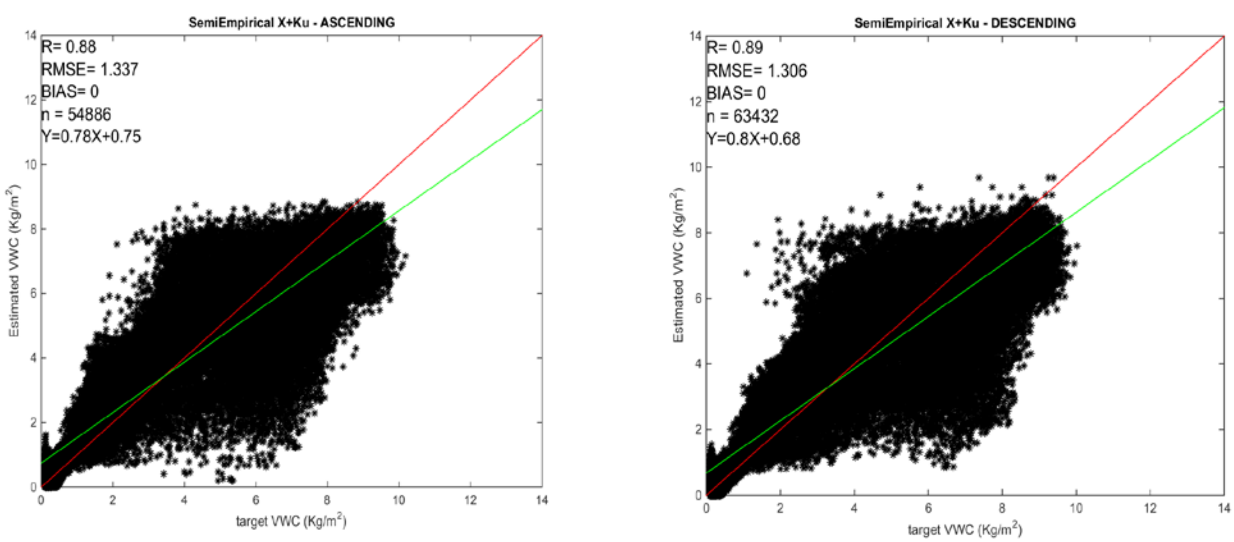
The sensitivity exhibited by PI at different frequencies to VWC suggested implementing an algorithm based on the combination of these indices for estimating biomass of different vegetation types.

4.1 Semi-empirical retrieval algorithm

The retrieval problem was first faced by developing a semi-empirical algorithm that estimates VWC as a function of the weighted combination of PI at X and Ku bands, assuming a logarithmic dependence of VWC on PI [18, 19]:

$$VWC= a*\log(PIX) +b*\log(PIKu) +c \qquad (3)$$

The values of *a*, *b* and *c* were defined by minimizing the RMSE between estimated and target VWC for a subset of about 15% of the available data from the African dataset (i.e. ~17,000 data ascending + descending), obtained through a pseudo-random sampling, in order to obtain a subset representative of the characteristics of the whole dataset. The minimization was again obtained using on the Nelder-Mead iterative algorithm and resulted in a=-0.698 (Kg/m^2), b=-1.164 (Kg/m^2) and c=5.765 (Kg/m^2).



(a) (b)

Figure 5. Test of the semi-empirical algorithm for ascending (a) and descending (b) orbits on the African site

Figure 5 represents the algorithm test conducted on the remaining part of the dataset (85% of data), which was not considered for defining the weights. The results of Figure 5 (a) and Figure 5 (b) confirm the minimum differences between ascending and descending orbits pointed out by the sensitivity analysis, as the main statistical figures listed in Table 3 demonstrate. The R values clearly reflect the correlation between PI and VWC, while the RMSE of about 1.3 Kg/m^2 depends mainly on the dispersion of PI for intermediate VWC values.

Table 3. R, RMSE, and BIAS of the correlation between estimated and target VWC using the semi-empirical algorithm

	Ascending	Descending
R	0.88	0.89
RMSE (Kg/m^2)	1.337	1.306
BIAS (Kg/m^2)	~ 0	~ 0

4.2 The ANN algorithm

A novel algorithm based on Artificial Neural Network (ANN) techniques was then defined with the aim of improving the retrieval accuracy, and especially reducing the RMSE with respect to the semi-empirical algorithm. ANN techniques demonstrated to be a good compromise between retrieval accuracy and computational cost with respect to other statistical or numerical approaches [37]. Another important advantage of ANN is in their capability of merging different input data into the same retrieval algorithm. In particular, using this technique it was possible adding easily the C band PI to the algorithm inputs, along with PI at X and Ku bands as in the semi-empirical algorithm.

The ANN implemented for this purpose was a feedforward multilayer perceptron, trained with the back propagation (BP) learning rule. The training set was composed by the same subset of 15% of the data available from the African dataset that was used for defining the weights of the semi-empirical algorithm. The test of the algorithm was instead carried out considering the remaining 85% of data, not involved in the training process.

In order to obtain a training more robust and representative of a wider range of surface conditions, the experimental dataset was increased with data simulated by the τ-ω model. According to the strategy described in [38], and deriving the range of τ and ω values as described in section 3, a dataset of 10,000 simulated PI at C, X and Ku bands and corresponding target VWC was obtained by iterating the tau omega model simulations. The relationship between frequency dependent τ values and VWC for model simulations was based again on equation 3.

Based on the training implementation of the NN Matlab ® toolbox, the training set was then divided in other three subsets composed by 60%, 20% and 20% of data for training, validation and a posteriori test, respectively.

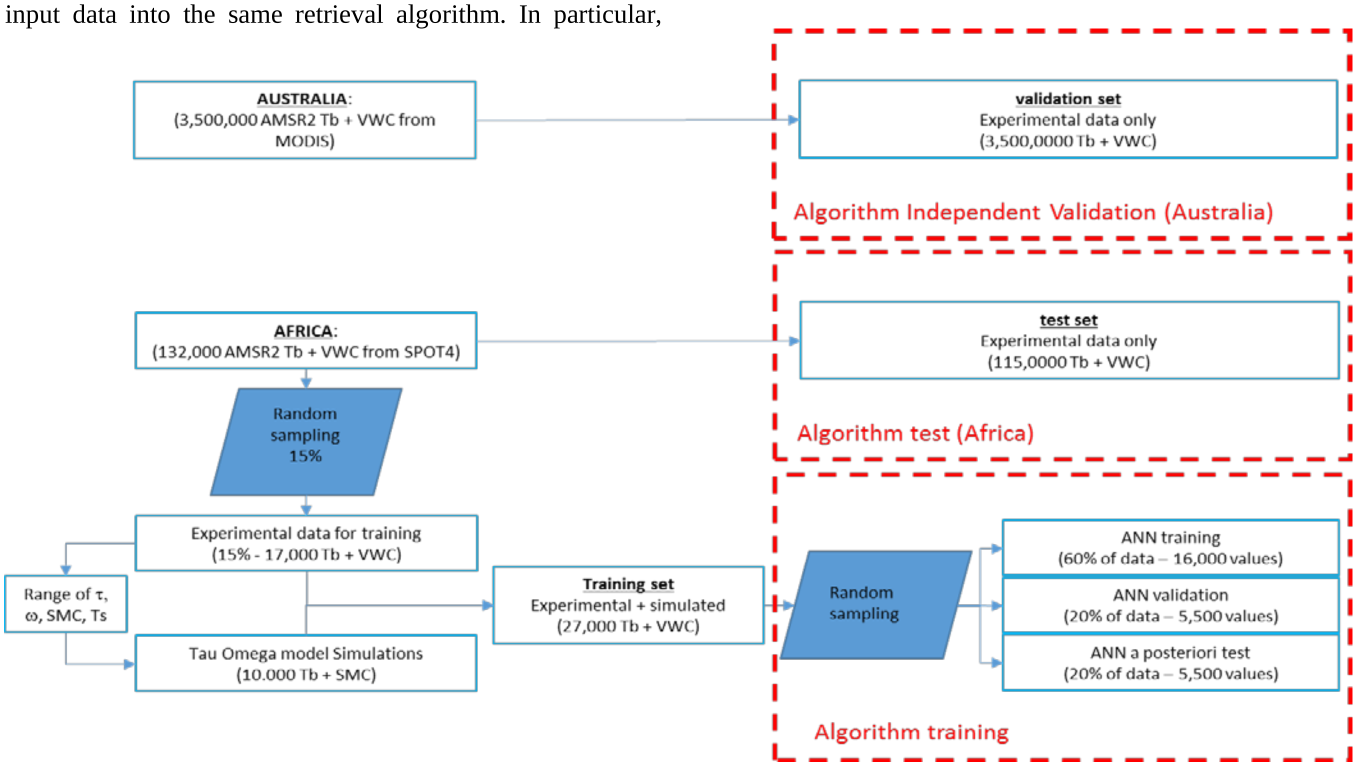


Figure 6. Strategy followed for defining the datasets considered for training, testing and validating the ANN algorithm

Following the "early stopping" rule, at each training iteration, the errors obtained on these three subsets are compared and the training is stopped when the errors on the validation and a posteriori subsets start diverging from the error on the training set. This technique allows mitigating or eliminating the problem of overfitting, which hampers the ANN capabilities in estimating correctly the target parameters on other datasets with respect to the training set. The ANN final architecture was defined applying an iterative optimization, as described in [38]: the optimization process resulted in an architecture with two hidden layers of 9+9 neurons, with a transfer function of type tansig, expressed as $\tanh(a) = (e^{a}-e^{-a})/(e^{a}+e^{-a})$. The inputs of the ANN were the PI at C, X, and Ku bands, and the output was the estimated VWC. The flowchart in Figure 6 summarizes the strategy followed for generating the training, test (Africa) and validation (Australia) datasets.

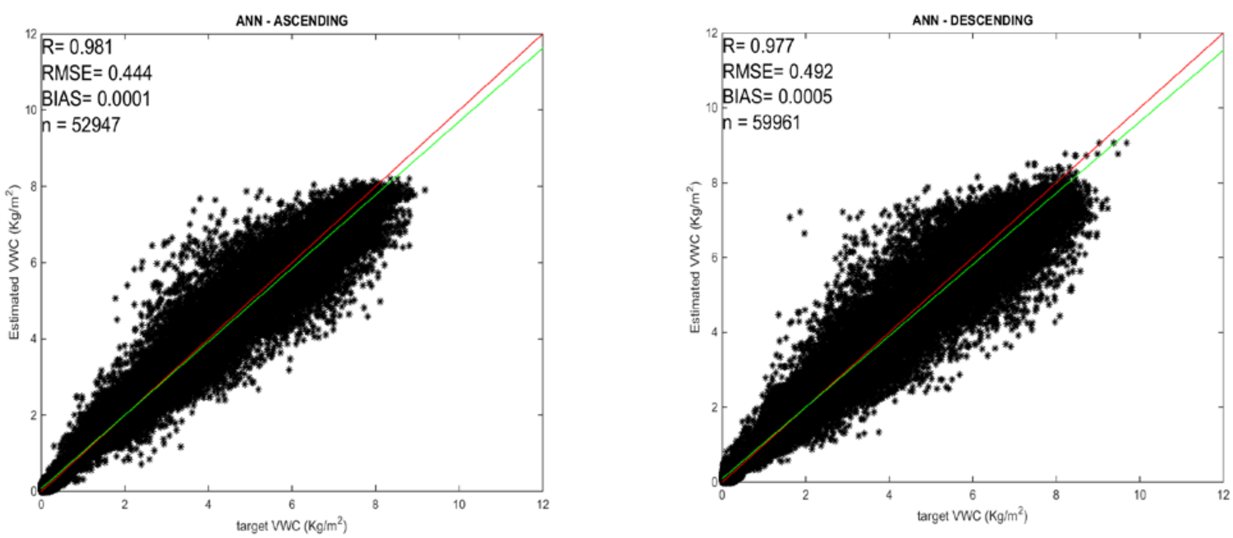
The test of the algorithm on the remaining 85% of data that were not involved in the training process is presented in Figure 7 that shows the scatterplots of retrieved vs. target VWC separated by ascending and descending orbits. The ANN algorithm appreciably improves the retrieval accuracy with respect to the semi-empirical one: the regression slope is close to 1:1, R increases up to about 0.98, and RMSE decreases down to about 0.45 Kg/m^2. The main statistical figures of the test are listed in Table 4.

Table 4. R, RMSE and BIAS of the ANN algorithm

	Ascending	Descending
R	0.98	0.98
RMSE (Kg/m^2)	0.44	0.49
BIAS (Kg/m^2)	~ 0	~ 0

The results obtained in ascending and descending orbits are very similar, but conversely to the semi-empirical algorithm, the ANN exhibits slightly higher accuracies in ascending than in descending orbits: this can be attributed to the ANN training. The results displayed in Figure 5 and Figure 7 confirm the capability of both algorithms in estimating VWC up to 8 Kg/m^2.

The proposed algorithm configuration, which adds PI at C band to the ANN inputs, slightly outperformed another implementation having as ANN inputs only PI at X and Ku bands (R increased from 0.96 to 0.98). Conversely, the inclusion of PI at C band in the inputs of the semi-empirical algorithm (eq.3) caused a small decrease of accuracy (R from 0.88 to 0.87 in ascending orbits and from 0.89 to 0.88 in descending ones). Therefore C band was not used in the semi-empirical algorithm.



(a) (b)

Figure 7. Algorithm test: estimated vs. target VWC for the African site, in ascending (a) and descending (b) orbits.

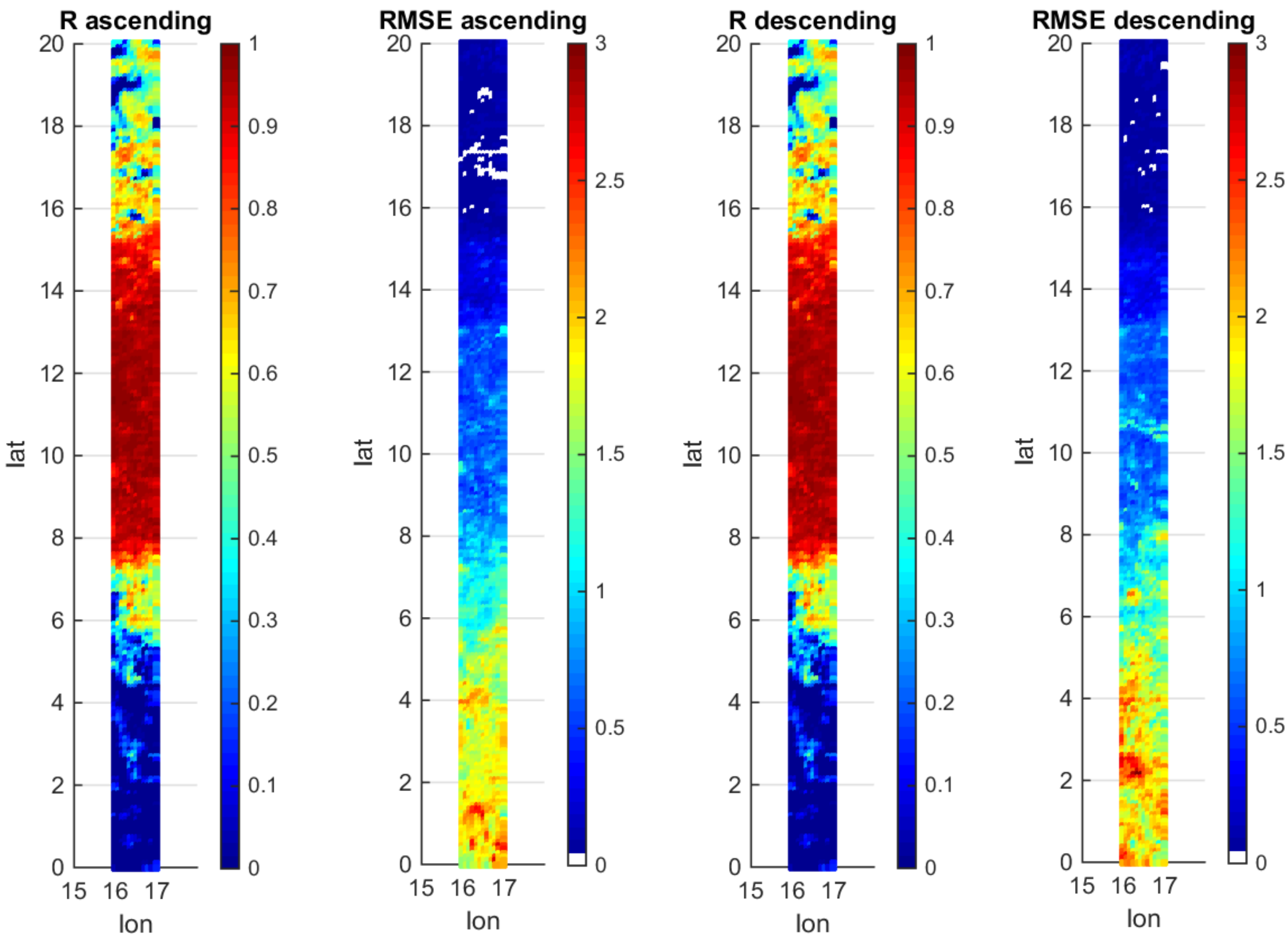


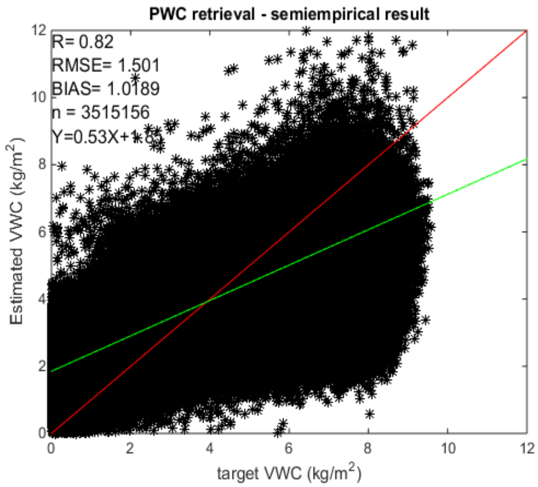
Figure 8. Algorithm test: R and RMSE maps for the African test site in ascending and descending orbits

Looking at the spatial distribution of these results, the maps of R and RMSE computed from the time series of estimated and target VWC at each of the 225x12 nodes of the fixed grid are displayed in Figure 8. The maps clearly point out the loss of correlation for the densely vegetated areas of Equatorial forest, which exhibit the lowest R values and corresponding highest RMSE - especially in descending orbits –, as it can be reasonably expected due to low penetration of radiation at these frequencies. On the other hand, the desert areas, which represent the opposite boundary of the biomass range, are characterized by the smallest RMSE and low to medium R values. Concerning vegetation covers characterized by intermediate values of biomass that are present in the central part of the area, the agreement between estimated and target values is well pointed out by the highest R and lowest RMSE values.

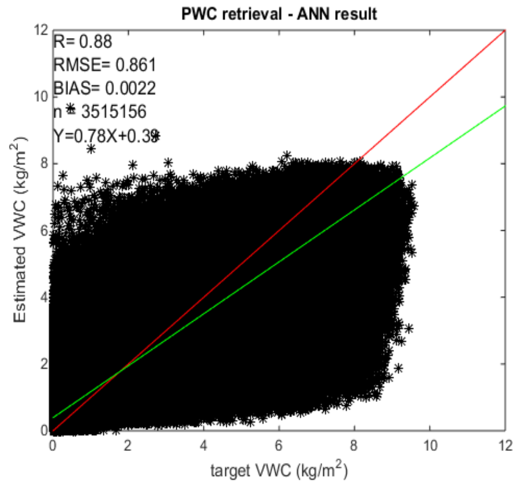
5. ALGORITHM VALIDATION (AUSTRALIA)

The validation of both algorithms was carried out considering the dataset collected on the Australian site. In this case, ascending and descending orbits were merged, and the retrieval results referred to the whole data set available. Figure 9 shows the scatter and density plots of estimated vs target VWC obtained using the semi-empirical (Figure 9 a-c) and the ANN (Figure 9 b-d) algorithms, respectively, on the >3,500,000 data of Australian site. The range of target VWC, mainly concentrated between 0 and 4 Kg/m^2, reflects the semiarid conditions of the most part of Australia. Apparently, the behavior of both scatterplots (Figure 9 a-b) belies the statistical figures of the retrieval displayed in the figures and listed in Table 5. However, looking at the corresponding density plots in Figure 9 (c-d), where a threshold was set for value occurrence < 5, it is evident that most part of data is concentrated along the regression line, while only a relatively small number of points is scattered around.

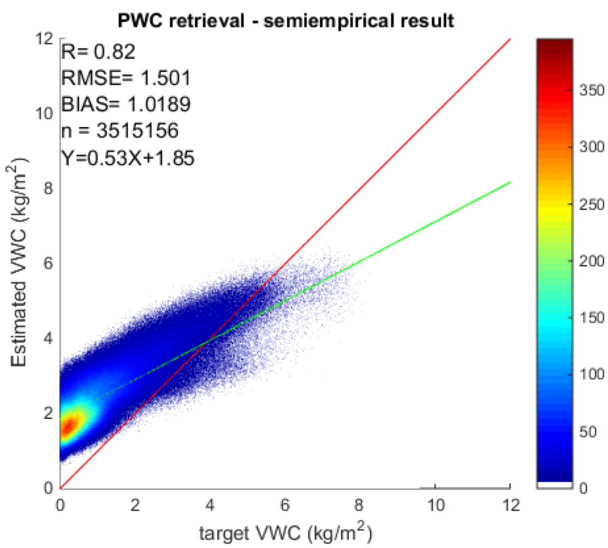
The semi-empirical algorithm tends to overestimate the low VWC and to underestimate the corresponding high values, while the ANN retrieval is closer to the 1:1 line. In general, the ANN algorithm improves the retrieval accuracy, as it is well pointed out by the comparison of R, RMSE, BIAS and regression slope in Table 5.



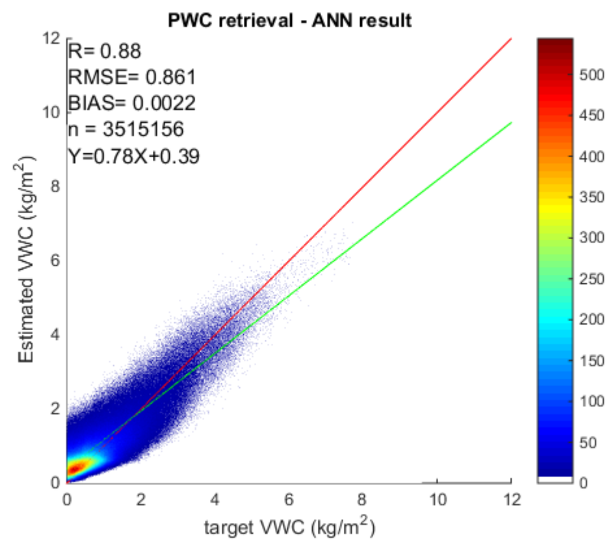
(a)



(b)



(c)



(d)

Figure 9. Scatter and density plots of estimated vs. target VWC on Australian site from semi-empirical (a-c) and ANN (b-d) algorithms, respectively.

The capability of both algorithms in detecting values of VWC up to 8 Kg/m^2 is confirmed by these results, although the retrievals are slightly worse than those obtained on African site. The reason can be attributed to the fact that the Australian dataset is completely independent of the data used for the algorithm implementation. Moreover, cross-calibration issues between NDVI derived from SPOT4 and MODIS might also affect the retrieval. Lastly, the Australian dataset is significantly larger than the African one and contains a more limited range of vegetation biomass values, thus with negative effects on the overall correlation.

Table 5. R, RMSE and BIAS obtained from the validation of both semi-empirical and ANN algorithms on Australian site

	Semi-empirical	ANN
R	0.82	0.88
RMSE (Kg/m^2)	1.501	0.861
BIAS (Kg/m^2)	1.02	0.002
slope	0.53	0.78

The R and RMSE maps obtained comparing the VWC time series estimated by both algorithms and the corresponding reference VWC at each of the 501x701 nodes of the grid are presented in Figure 10(a-b) for the semi-empirical algorithm and in Figure 10(c-d) for the ANN algorithm. The improvement obtained by using ANN is evident from the comparison of these maps. The overall correlation increases and the RMSE decreases, although some areas showing low values of R can be identified in the central-western part of Australia, which is mainly characterized by rocky deserts (Figure 10c). According to the previous findings on African site, the lowest R and highest RMSE correspond to the dense forests along the eastern coast. The retrieval accuracies of both algorithms in these areas are low and reflect the loss of correlation between PI and the VWC of dense forests derived from NDVI, as it has already been pointed out for the African site. In the RMSE maps the effect of mixed pixels along coastlines is evident for both algorithms. The disaggregation techniques applied to AMSR2 data, either Backus-Gilbert or SFIM based, are able indeed to mitigate the effects of the coarse resolution without, nevertheless, completely overcome this problem, especially in case of large Tb variations, as it happens along the coastlines.

6. CONCLUSIONS

In order to test the capability of the AMSR2 system in delivering the Vegetation Water Content (VWC) as an all-weather standard product of the JAXA GCOM-W mission, we developed two retrieval algorithms based on the measurement of the normalized polarization difference of brightness temperature, i.e. the Polarization Index (PI). Both algorithms are based on the simple τ-ω solution of the radiative transfer model and semi-empirical relationships between brightness temperature and vegetation biomass. The first one uses the inversion of a semi-empirical combination of PI at X and Ku bands, while the second method is based on an Artificial Neural Networks (ANN) and also includes data at C band.

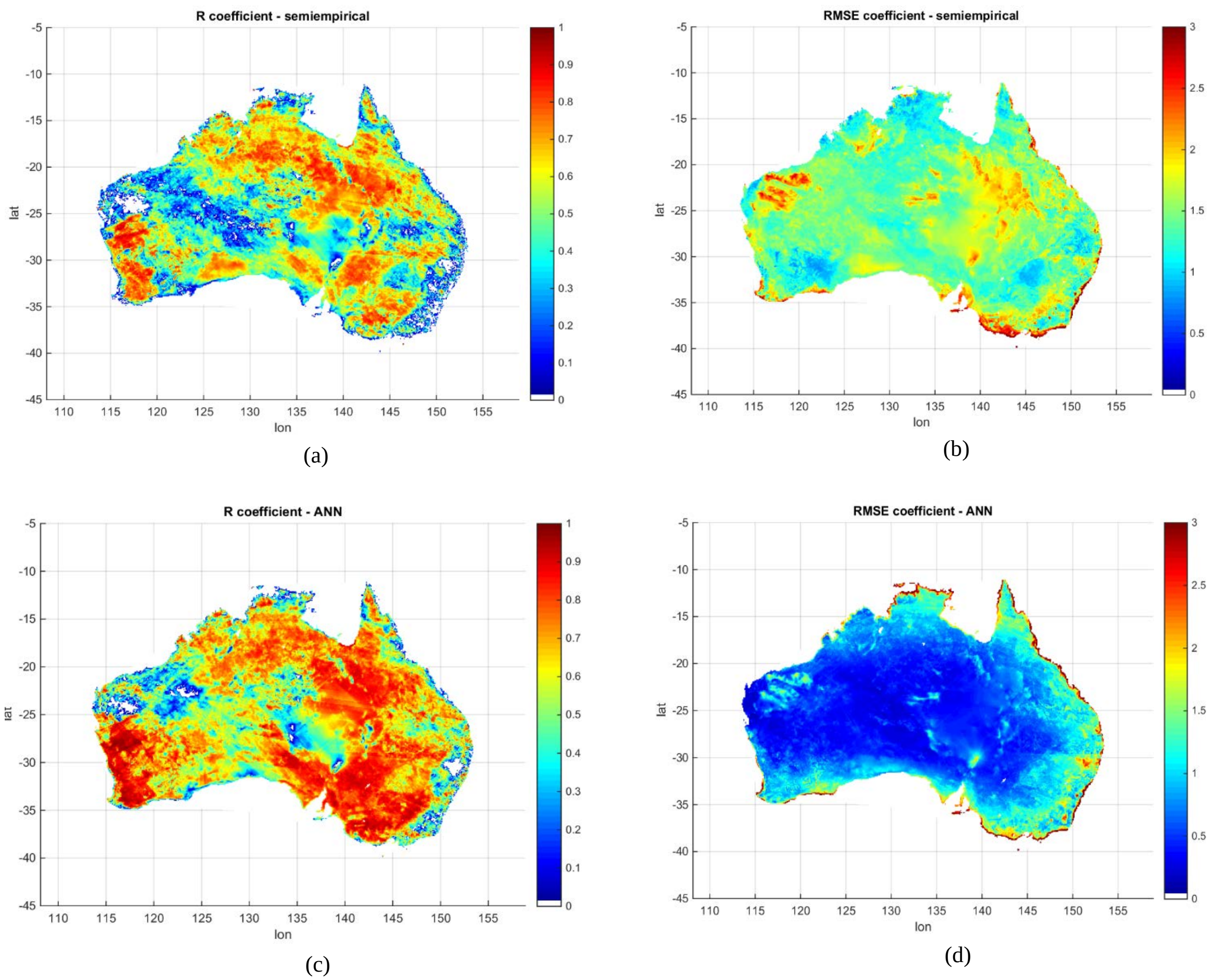


Figure 10. Maps of R and RMSE on Australia from semi-empirical (a-b) and ANN (c-d) algorithms

The two algorithms were implemented and tested considering two years of AMSR2 data collected on a wide portion of Africa that included a large variety of vegetation types and biomasses. As it has already been done in similar studies, the algorithms were validated referring to VWC derived from the Normalized Difference Vegetation Index (NDVI) computed from SPOT 4 optical data and downsampled at the AMSR2 resolution. Indeed, this approach is presently the only one capable of allow a validation on a global scale. The comparison of the two approaches showed a better performance of the ANN with a correlation coefficient, R= 0.98 and a root mean square error, RMSE, close to 0.45 Kg/m^2, while the numbers obtained from the semi-empirical algorithm were R=0.88 and RMSE = 1.3 Kg/m^2.

As expected, the best fitting was reasonably obtained on low and moderately dense vegetation (agricultural areas and prairies) where the difference in penetration between optical and microwave radiation is relatively small.

An independent validation of both algorithms was carried out on Australia, considering four periods of 16 days each, in the different seasons of 2013. In this case, the algorithm outputs were compared with VWC derived from MODIS 16 days averaged NDVI. This validation confirmed the test results on the African site, demonstrating that the microwave data from AMSR2 can be legitimately used to produce vegetation maps on a global scale by separating several levels of biomass on low and medium dense vegetation (up to 8 Kg/m^2), without any need of further information from other sensors and guaranteeing an all-weather monitoring.

7. ACKNOWLEDGEMENTS

This work has been partially carried out in the framework of the JAXA Collaborative Research Agreement for Global Change Observation Mission (GCOM).

8. REFERENCES

[1]. IPCC, 2014: climate change 2014: synthesis report. Contribution of working groups i, ii and iii to the fifth assessment report of the intergovernmental panel on climate change [core writing team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp**.** The intergovernmental panel on climate change, 2015

[2]. Piao S., Ciais P., Friedlingstein P., Peylin P., Reichstein M., Luyssaert S., et al., “Net carbon dioxide losses of northern ecosystems in response to autumn warming”. Nature, 451, 49, 2008.

[3]. Tucker, C.J. “Red and Photographic Infrared Linear Combinations for Monitoring Vegetation”, *Remote Sensing of Environment*, 8(2), 127-150 A., 1979

[4]. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, L. G. Ferreira. “Overview of the radiometric and biophysical performance of the MODIS vegetation indices”. Remote Sensing of Environment 83(2002) 195-213 doi:10.1016/S0034-4257(02)00096-2.

[5]. Jackson T. J., D.Chen, M.Cosh, Fuqin Li, M.Anderson, C.Walthall, P.Doriaswamy, E. R.Hunt, “Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans”, Remote Sensing of Environment, 92, pp.475–482, 2004.

[6]. F.T. Ulaby, R.K. Moore, A.K Fung, Microwave Remote Sensing: Active and Passive. Vol III Nordwood, MA: Artech House

[7]. Ulaby, F. T., & Long, D. G. (Eds.). Microwave radar and radiometric remote sensing. Ann Arbor (US): The University of Michigan Press, 2014.

[8]. P. Ferrazzoli, S. Paloscia, P. Pampaloni, G. Schiavon, D. Solimini, and P. Coppo, “Sensitivity of Microwave Measurements to Vegetation Biomass and Soil Moisture Content: A Case Study”. *IEEE Trans. Geosci. Remote Sensing*, vol. GE-30, pp. 750-756, 1992.

[9]. S. Paloscia. “Microwave emission from vegetation” in: Passive Microwave Remote Sensing of Land-atmosphere Interactions, pp. 358-374, B. Choudhury, E. Njoku, Y. Kerr and P. Pampaloni, Editors, VSP Press, The Netherlands, 1995.

[10]. S. Paloscia, P. Pampaloni, “Microwave Polarization Index for Monitoring Vegetation Growth”, IEEE Trans. on Geoscience and Remote Sensing, GE 26, n. 5, pp. 617-621, 1988

[11]. Paloscia S. and P. Pampaloni, “Microwave Vegetation Indexes for detecting biomass and water conditions of agricultural crops”, Remote Sens. Environ., n.40, p. 15 26, 1992

[12]. B. J Choudhury, “Monitoring global land surface using Nimbus-7 37 GHz. Theory and examples,” *International Journal of Remote Sensing*, 10, 1579−1605, 1989.

[13]. E. Santi, S. Pettinato, S. Paloscia, P. Pampaloni, G. Macelloni, and M. Brogioni: “An algorithm for generating soil moisture and snow depth maps from microwave spaceborne radiometers: Hydroalgo,” Hydrol. Earth Syst. Sci., 16, pp. 3659-3676, 2012.

[14]. J.C. Shi, T. Jackson, J. Taoa, J. Dua, R. Bindlish, L. Lud, and K.S. Chen, “Microwave vegetation indices for short vegetation covers from satellite passive microwave sensor AMSR-E,” Remote Sens. Environ., 112, pp.4285-4300, 2008

[15]. Chukhlantsev A.A. (1981). Microwave Emission from Vegetation Canopies. Ph.D. Dissertation. Moscow: The Moscow Institute of Physics and Technology (State University), 1981 (in Russian).

[16]. Jackson T.J., Schmugge T.J., and Wang J.R., “Passive microwave sensing of soil moisture under vegetation canopies”. Water Resources Research, 18:1137-1142, 1982.

[17]. Mo T., Choudhury B.J., Schmugge T.J., Wang J.R., and Jackson T.J., “A model for microwave emission from vegetation-covered fields”. J. Geophys. Res., 87:11229-11237, 1982.

[18]. P. Pampaloni and S. Paloscia, “Microwave Emission and Plant Water Content: A Comparison between Field Measurements and Theory”, IEEE Trans. Geosci. Remote Sens, vol. GE-24, no. 6, pp. 900–905, Nov. 1986.

[19]. A. Chukhlantsev, Microwave radiometry of vegetation canopies, Advances in global change research, Springer, The Netherlands, 2006,

[20]. Chukhlantsev A.A. and Golovachev S.P., “Microwave attenuation in a vegetation canopy”. Radiotekhnika i Elektronika, 34:2269-2278, 1989 (in Russian).

[21]. Owe M., R. de Jeu, and J. Walker, “A Methodology for Surface Soil Moisture and Vegetation Optical Depth Retrieval Using the Microwave Polarization Difference Index”. IEEE Transactions on Geoscience and Remote Sensing, 39-8, pp. 1643-1654, 2001

[22]. M. O. Jones, L. A. Jones, J.S. Kimball, K. C. McDonald, ‘Satellite passive microwave remote sensing for monitoring global land surface phenology’, Remote Sensing of Environment

115, pp. 1102–1114, 2011.

[23]. Yi Y. Liu, R.A M. de Jeu, M.F. McCabe, J.P. Evans, A.I. J. M. van Dijk, "Global long-term passive microwave satellite-based retrievals of vegetation optical depth", Geophysical Research Letters, 38, L18402, doi:10.1029/2011GL048684, 2011.

[24]. Lawrence H., J.P. Wigneron, P. Richaume, N. Novello, J. Grant, A. Mialon, A. Al Bitar, O. Merlin, D. Guyon, D. Leroux, S. Bircher, Y. Kerr, "Comparison between SMOS Vegetation Optical Depth products and MODIS vegetation indices over crop zones of the USA", Remote Sensing of Environment, Volume 140, pp. 396-406, ISSN 0034-4257, http://dx.doi.org/10.1016/j.rse.2013.07.021, 2014.

[25]. Rahmoune, R., Ferrazzoli, P., Singh, Y. K., Kerr, Y. H., Richaume, P., & Al Bitar, A., "SMOS retrieval results over forests: Comparisons with independent measurements". IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 3858–3866, 2014.

[26]. Vittucci C., P. Ferrazzoli, Y. Kerr, P. Richaume, L. Guerriero, R. Rahmoune, G. Vaglio Laurin, "SMOS retrieval over forests: Exploitation of optical depth and tests of soil moisture estimates", Remote Sensing of Environment, Volume 180, July 2016, Pages 115-127, ISSN 0034-4257, http://dx.doi.org/10.1016/j.rse.2016.03.004.

[27]. Maeda T., Y. Taniguchi and Keiji Imaoka, 2016. GCOM-W1 AMSR2 Level 1R Product: Dataset of Brightness Temperature Modified Using the Antenna Pattern Matching Technique. IEEE Trans. on Geoscience and Remote Sensing. 54 – 2, pp. 770 – 782

[28]. Jackson T. J., D.Chen, M.Cosh, Fuqin Li, M.Anderson, C.Walthall, P.Doriaswamy, E. R.Hunt, "Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans", Remote Sensing of Environment, 92, pp.475–482, 2004.

[29]. Paloscia, S.; Pettinato, S.; Santi, E.; Notarnicola, C.; Pasolli, L.; Reppucci, A. Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. Remote Sens. Environ., 134, 234–248, 2013.

[30]. Santi E., "An application of SFIM technique to enhance the spatial resolution of microwave radiometers". International Journal of Remote Sensing, 31(9-10), 2419- 2428, 2010.

[31]. Nelder, J. A. & Mead, R., "A Simplex Method for Function Minimization". Computer Journal, 7, 308-313, 1965.

[32]. T. R. H. Holmes, R. A. M. De Jeu, M. Owe, and A. J. Dolman, "Land surface temperature from Ka band (37 GHz) passive microwave observations," J. Geophys. Res., vol. 114, no. 4, Feb. 2009, Art. ID D04113

[33]. Santi, E., S. Paloscia, S. Pettinato, L. Brocca, L. Ciabatta, C. Massari. 2016. Soil moisture and rainfall retrieval from AMSR2 data in central Italy. Proc. Of International Geoscience and remote sensing symposium, IGARSS 2016, July 2016. In press.

[34]. Kirdiashev K.P, Chuchlantsev A.A and Shutko A.M., "Microwave radiation of the Earth's surface in the presence of vegetation cover". Radio Eng. Electron. Phys., 2: 37-56, 1979.

[35]. Sawada Y., H. Tsutsui, T. Koike, M. Rasmy, R. Seto, and H. Fujii, "A Field Verification of an Algorithm for Retrieving Vegetation Water Content From Passive Microwave Observations". IEEE Trans. On Geosci and Remote Sensing. Vol. 54-4, 2082-2095, 2016.

[36]. Eni G. Njoku, Steven K. Chan, "Vegetation and surface roughness effects on AMSR-E land observations", Remote Sensing of Environment, Volume 100, Issue 2, 30 January 2006, Pages 190-199, ISSN 0034-4257, http://dx.doi.org/10.1016/j.rse.2005.10.017.

[37]. S.Paloscia, P.Pampaloni, S.Pettinato, E.Santi, "A Comparison of Algorithms for Retrieving Soil Moisture From ENVISAT/ASAR Images", IEEE Trans. Geosci. Remote Sensing vol. 46, 10, October 2008, pp. 3274-3284, 2008

[38]. Santi E., S. Paloscia, S. Pettinato, G. Fontanelli., "Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors". Int. J. Appl. Earth Observ. Geoinf., Volume 48, Pages 61–73, 2016. http://dx.doi.org/10.1016/j.jag.2015.08.002.