We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Neural Networks Applications for the Remote Sensing of Hydrological Parameters

Emanuele Santi

Additional information is available at the end of the chapter

Abstract

The main artificial neural networks (ANN)-based retrieval algorithms developed at the Institute of Applied Physics (IFAC) are reviewed here. These algorithms aim at retrieving the main hydrological parameters, namely the soil moisture content (SMC), the plant water content (PWC) of agricultural vegetation, the woody volume of forests (WV) and the snow depth (SD) or snow water equivalent (SWE), from data collected by active (SAR/scatterometers) and passive (radiometers) microwave sensors operating from space. Taking advantage of the fast computation, ANN are able to generate output maps of the target parameter at both local and global scales, with a resolution varying from hundreds of meters to tens of kilometres, depending on the considered sensor. A peculiar strategy adopted for the training, which has been obtained by combining satellite measurements with data simulated by electromagnetic models (based on the radiative transfer theory, RTT), made these algorithms robust and site independent. The obtained results demonstrated that ANN are a powerful tool for estimating the hydrological parameters at different spatial scales, provided that they have been trained with consistent datasets, made up by both experimental and theoretical data.

Keywords: microwave satellite sensors, soil moisture content, vegetation water content, snow water equivalent, woody volume

1. Introduction

Nowadays, global phenomena such as climate warning, stratospheric ozone depletion and troposphere pollution are threatening the long-term habitability of the planet. The Earth is a complex evolving system where regional as well as global processes at all spatial and temporal scales are strongly interrelated. These include surface exchanges of water, energy, carbon and

other bio-geological processes, and exchange between land surface atmosphere, ocean and ground water. The human activity further complicated all these processes, because it transforms continuously the land surface to meet human needs associated with basic food production, population expansion and economic development.

Consequently, an ever-increasing interest in meteorological events and climate changes has led to a greater focus on the study of hydrological processes and their dynamics. The hydro-logical cycle involves the circulation of water from ocean to water vapour through the evaporation processes, the transformation of water vapour into precipitation and its return to the cycle through infiltration and evapotranspiration again. However, considering the spatial and temporal coverage needed to have reliable estimates, direct measurements of all the parameters involved in the hydrological cycle are difficult and extremely expensive. This led to an increasing interest for the observation from space, since it can meet the temporal and spatial requirements for an operational monitoring of the parameters related to the hydrolog-ical cycle. Earlier initiatives, such as the launch of a number of satellites having on-board sensors dedicated to the Earth's parameters observation, have already developed long-term applications and provided fundamental contributions in understanding global and regional ocean processes and enhancing land surface studies.

Among the instruments operating from space for the Earth surface observation, the sensors operating in the microwave portion of the electromagnetic (e.m.) spectrum have a great potential because these frequencies are capable of estimating some parameters of atmosphere, vegetation and soil that cannot be observed in the visible/near-infrared and thermal wave-length. Moreover, the scattering and emission at microwaves are related directly to the water content of the observed target. Microwave sensors, which can be classified in active (real aperture radar—RAR, and synthetic aperture radar—SAR), and passive (radiometers) are therefore particularly suitable for monitoring the key parameters of the hydrological cycle. In particular, these sensor represent a powerful tool for monitoring the soil water content or soil moisture content (SMC), the vegetation biomass, expressed as plant water content (PWC) for agricultural crops and woody volume (WV) for forests, and the snow depth (SD) or its water equivalent (SWE).

SMC is one of the driving factors in the hydrological cycle, being able to influence the runoff, the evapotranspiration, the surface heat fluxes and the biogeochemical cycles. The knowledge of SMC and its dynamics is mandatory in a wide range of activities concerning the forecasting of weather and climate, the prevention of natural disasters such as floods and landslides, the management of water resources and agriculture-related activities and many others. A huge amount of experimental and theoretical studies on the SMC retrieval from microwave acquisitions was therefore carried out since the late 1970s. The dielectric constant of soil (DC) at microwave frequencies is strongly dependent on the water content of the observed soil. At L-band, for example, the large variation in real part of the dielectric constant from dry to saturated soil results in a change of about 10 dB in radar backscatter and of 100 K in the radiometric brightness temperature. An important component required in the soil moisture inverse problem is the knowledge of the relationship between the soil dielectric constant to its

moisture content: widely adopted empirical models for assessing such relationship are given in [1, 2].

Snow is another driving factor of the hydrological cycle, since it is able to influence the Earth's climate and its response to global changes. Snow is the main component of the cryosphere, and its accumulation and extension are related to the global climate variations. The monitoring of the snow parameters, and in particular SD and SWE, is essential for the forecasting of snow–water runoffs (flash floods) and for the management of the water resources. Currently, satellite microwave radiometers are employed for generating low-resolution SD or SWE products at global scale, while the operational mapping of snow at high resolution mainly rely with optical sensors, being the microwave application still at the research stage for this application. In both cases, the currently available snow products are based on the single sensors, and thus, the temporal and spatial coverage is given by the sensor characteristics, which may not fulfil the requirements for the operational use of remote sensing data in monitoring and management of snow.

Vegetation cover on the Earth's surface is an important variable in the study of global changes, since vegetation biomass is the most influential input for carbon cycle models. The frequent and timely monitoring of vegetation parameters (such as vegetation biomass and leaf area index) is therefore of vital importance to the study of climate changes and global warming.

The retrieval of the aforementioned parameters from active and/or passive microwave measurements is nonetheless not trivial, due to the nonlinearity of the relationships between radar and radiometric acquisitions and target parameters. Moreover, in general, more than one combination of surface parameters (SMC, surface roughness—HSTD, PWC and so on) give the same electromagnetic response. Thus, in order to minimize the uncertainties and enhance the retrieval accuracy from remote sensing data, statistical approaches based on the Bayes theorem and learning machines are widely adopted for implementing the retrieval algorithms [3–5].

In this framework, the artificial neural networks (ANN) represent an interesting tool for implementing accurate and flexible retrieval algorithms, which are able to operate with radar and radiometric satellite measurements and to easily combine information coming from different sources. ANN can be considered a statistical minimum variance approach for addressing the retrieval problem, and, if properly trained, they are able to reproduce any kind of input-output relationships [6, 7].

During the training, sets of input data and corresponding target outputs are provided sequentially to the ANN, which iteratively adjusts the interconnecting weights of each neuron, in order to minimize the difference between actual outputs and corresponding targets, basing on the selected learning algorithm.

Many examples of ANN application to inverse problems in the remote-sensing field can be found in literature, in particular concerning the retrieval of soil moisture at local scale from SAR [8–10] or radiometric [11] observations. The comparison of retrieval algorithms carried out in [9] demonstrated that ANN, with respect to other widely adopted statistical approaches based on Bayes theorem and Nelder–Mead minimization [12], offer the best compromise

between retrieval accuracy and computational cost. Other comparison between ANN, Bayesian, SVM and other retrieval approaches can be found in [13–15]. All these works demonstrated that the ANN are able to provide accuracy results in line with (or better than) the other methods, with the advantages of a fast computation, that is mandatory for the online processing of high-resolution images, and the possibility of updating the training if new data are available. The cited publications refer to the retrieval of soil moisture; however, these considerations remain valid for the retrieval of the other parameters investigated here.

Basing the training on data simulated by forward electromagnetic models, namely models that are able to simulate the microwave signal emitted or scattered by the target surface, the ANN can be regarded as a method for estimating the hydrological parameters from satellite microwave acquisitions through the inversion of the given model. Following this approach, the ANN act for inverting the forward model, similarly to other physically based algorithms, but without the approximations needed for an analytical inversion. Moreover, the additional inclusion of experimental data in the training set allows retaining the advantages of the experimental-driven approaches in adapting the algorithm to the particular features of a given test site [16].

The main advantages of this technique consist of the possibility of quick updating the training with new datasets, thus adapting the algorithm to work on a given test area, but without losing the accuracy on a larger scale. Moreover, the method has the capability of easily merging data coming from different sources for improving the retrieval accuracy. The poor robustness to outliers represents instead the main disadvantage of ANN: outliers are input data out of the range of the training set. In such case indeed, the ANN may return large errors or fail completely the retrieval, requiring therefore a 'robust' training, which has to be representative of a variety of surface conditions as wide as possible.

Besides the other considerations, it should be remarked that the strategy adopted for setting up and training the ANN is fundamental for obtaining a valid retrieval algorithm. An inappropriate training can turn indeed the ANN from a powerful retrieval instrument into an inadequate approach to the given problem. Some examples can be found in literature, in which the training set is insufficient for defining all the interconnection weights of the complex architecture proposed, or in which the architecture definition and the related overfitting and underfitting have not been properly addressed. Another fundamental consideration is that ANN are able to represent any kind of input–output relationships, and therefore, a deep knowledge of the physic of the problem is mandatory for avoiding the risk of relating input and output quantities that are instead completely uncorrelated, thus generating relationships that have no physical basis.

In this work, a review of the main ANN-based algorithms developed at IFAC for estimating the soil moisture (SMC, in m^3/m^3), the water content of agricultural vegetation (PWC, in kg/m^2), the forest woody volume (WV, in t/ha) and the snow depth/water equivalent (SD, in cm, and SWE, in mm) is presented. These algorithms take advantage of an innovative training strategy, which is based on the combination of satellite measurements with data simulated by electromagnetic models, based on the radiative transfer theory (RTT).

2. Implementing and training the Artificial Neural Networks

The ANN considered in this work are feed-forward multilayer perceptron (MLP), having two or three hidden layers of nine to twelve neurons each between the input and the output. The training was based on the back-propagation learning rule, which is an iterative gradient-descent algorithm able to adjust the connection weights of each neuron, in order to minimize the mean square error between the outputs generated by the ANN at every iteration and the corresponding target values.

It should be noted that the gradient-descent method sometimes suffers from slow convergence, due to the presence of one or more local minima, which may also affect the final result of the training. This problem can be solved by repeating the training several times, with a resetting of the initial conditions and a verification that each training process led to the same convergence results in terms of R and RMSE, by increasing it until negligible improvements were obtained.

2.1. Defining architecture

In order to define the optimal ANN architecture in terms of number of neurons and hidden layers, the most suitable strategy is to start with a simple ANN architecture, generally with one hidden layer of few neurons. These ANN are trained by means of a subset of the available data, tested on the rest of the dataset, and the training and testing errors are compared. The ANN configuration is then increased by adding neurons and hidden layers; training and testing are repeated and errors compared again, until a further increase of the ANN architecture is found to have a negligible decrease of the training error and an increase in the test error. This procedure allows defining the minimal ANN architecture capable of providing an adequate fit of the training data, preventing overfitting or underfitting problems. NNs, like other flexible nonlinear estimation methods, can be affected indeed by either underfitting or overfitting. ANN configurations not sufficiently complex for the given problem can fail to reproduce complicated data set, leading to underfitting. ANN configurations too complex may fit also the noise, leading to overfitting. Overfitting is especially dangerous because it can easily lead to predictions that are far beyond the range of the training data. In other words, the ANN are able to reproduce the training set with high accuracy but fails the test and validation phases.

2.2. Selecting the transfer function

Another key issue for defining the ANN best architecture is in the selection of the most appropriate transfer function: in general, linear transfer functions give less accurate results in training and testing; however, they are less prone to overfitting and are more robust to outliers, that is input data out the range of the input parameters included in the training set. Logistic Sigmoid (logsig) and Hyperbolic Tangent Sigmoid (tansig) transfer functions are instead characterized by higher accuracies in the training and test; however, they may lead to large errors when the trained ANN are applied to new datasets. Logsig generates outputs between 0 and 1 as the neuron's net input goes from negative to positive infinity and describes the

nonlinearity, g(a), as $1/(1 + e^{-a})$. Alternatively, multilayer networks can use the tansig function, $\tanh(a) = (e^a - e^{-a})/(e^a + e^{-a})$.

2.3. Generating the training set

Nevertheless, besides these problems, the main constraint for obtaining good accuracies with the ANN approach consists of the statistical significance of the training set, which shall be representative of a variety of surface conditions as wide as possible, in order to make the algorithm able to address all the situations that can be encountered on a global scale. The datasets derived from experimental activities are in general site dependent and cannot be representative of the large variation of the surface features that can be observed on a larger scale. Therefore, a training set only based on experimental data is not sufficient for training the ANN for global monitoring applications.
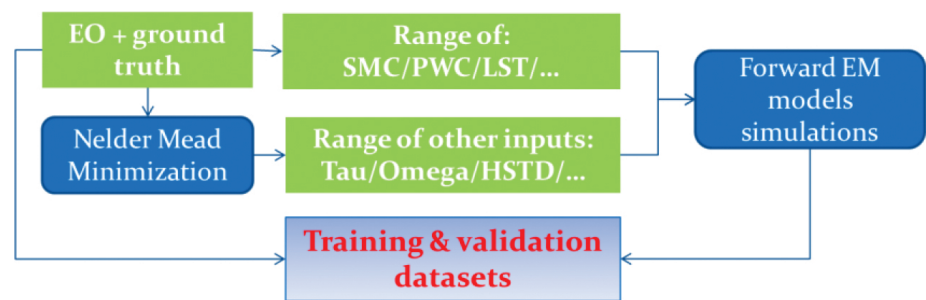

Figure 1. Strategy adopted for defining the raining and validation datasets, starting from the experimental data and using the forwards e.m. models (after [16]).

By combining the experimental in situ measurements with simulated data obtained from the e.m. models, it is possible to fill in the gaps of the experimental datasets and to better characterize the microwave signal dependence on the target parameter for a variety of surface conditions as wide as possible. The consistency between experimental data and model simulations can be obtained by deriving the range of model input parameters from the available measurements. After defining the minimum and maximum of each parameter required by the model, namely SMC, PWC, soil moisture, surface roughness (HSTD) and surface temperature (LST), the input vectors are generated by using a pseudorandom function, rescaled in order to cover the range of each parameter. Thousands of inputs vectors for running the model simulations can be generated by iterating this procedure, thus obtaining datasets of surface parameters and corresponding simulated microwave data for training and testing the ANN. The flowchart of **Figure 1** represents the main steps for generating the training from the experimental data. The same procedure allows generating the independent dataset for validating the ANN after training. In general, the available data are divided in two subsets with a random sampling; the first subset is divided again in 60–20–20% for training, test and validation phases, respectively, and the second subset is reserved for an independent test of the algorithm. The random sampling of the dataset is reiterated 5–6 times, and the training is repeated each time, in order to avoid any dependence of the obtained results on the sampling process.

This strategy has been successfully adopted for implementing and training the ANN-based algorithms that are presented in the following sections.

3. HydroAlgo ANN algorithm for AMSR-E and AMSR2

The 'HydroAlgo' algorithm [11] applies the ANN for estimating simultaneously SMC, PWC and SD from the acquisitions of the low-resolution spaceborne radiometers, like the Advanced Microwave Scanning Radiometer for the Earth observing system (AMSR-E) [17], which is no more operating, and its successor, AMSR2 [18]. We refer to [11] for a detailed algorithm description. The main characteristic of the algorithm is the exclusive use of AMSR-E/2 data, taking advantage of the multifrequency acquisitions of these sensors. It includes a disaggregation procedure, based on the smoothing filter-based intensity modulation (SFIM) technique [19], which is able to enhance the spatial resolution of the output SMC product up to the nominal sampling of AMSR-E/2 ($\sim 10 \times 10 \, \text{km}^2$). The algorithm flowchart is represented in **Figure 2**: it should be noted that the algorithm applies the already trained ANN to the input data, without repeating the training for each new set of satellite acquisitions. The trained ANN are generated once, saved and recalled for processing the available data. In particular, specific ANN have been trained for each given output product, basing on training sets composed by a combination of experimental data and simulations from e.m. models, obtained following the scheme of Section 2.
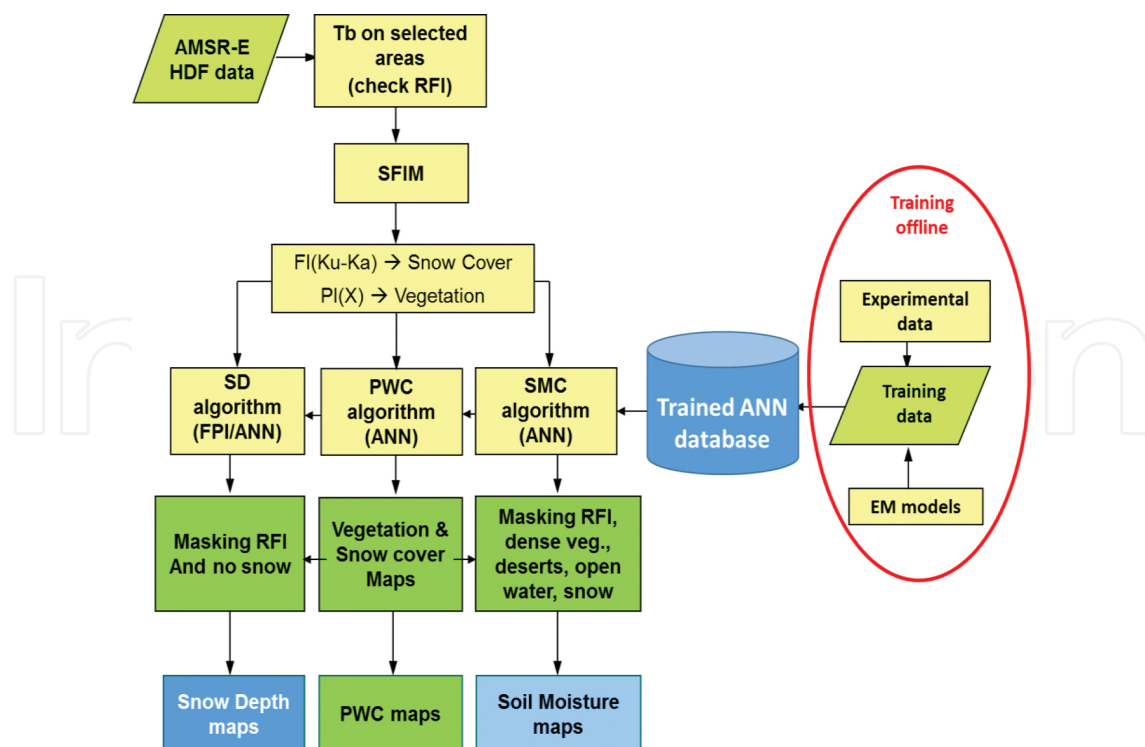

Figure 2. HydroAlgo flowchart.

3.1. SMC processor

The SMC processor was developed and tested using a set of several thousand of data, which was obtained by combining the experimental data collected in Mongolia and Australia with 10,000 values of Tb simulated by the 'tau-omega' model [20]. The experimental dataset was provided by JAXA, within the framework of the JAXA ADEOS-II/AMSR-E and GCOM/AMSR2 research programs.

The core of the algorithm is composed by two feed-forward multilayer perceptron (MLP) ANN, trained independently for the ascending and descending orbits, and using the back-propagation learning rule. Inputs of the algorithm are the brightness temperature at C-band in V-polarization, the polarization indices (PI) at 10.65 and 18.7 GHz (X- and Ku-bands), defined as $PI = 2 \times (TbV - TbH)/(TbV + TbH)$, and the brightness temperature at Ka-band (36.5 GHz) in V-polarization. C-band, that is the lowest AMSR-E frequency, was chosen for its sensitivity to the SMC, which is slightly influenced by sparse vegetation. The polarization indices at X- and Ku-bands are considered for compensating the effect of vegetation on soil emission [21], and for flagging out the densely vegetated targets, where SMC cannot be retrieved. The brightness temperature at Ka-band, V-polarization, was assumed as a proxy of the surface physical temperature, to account for the effect of diurnal and seasonal variations of the surface temperature on microwave brightness [22].
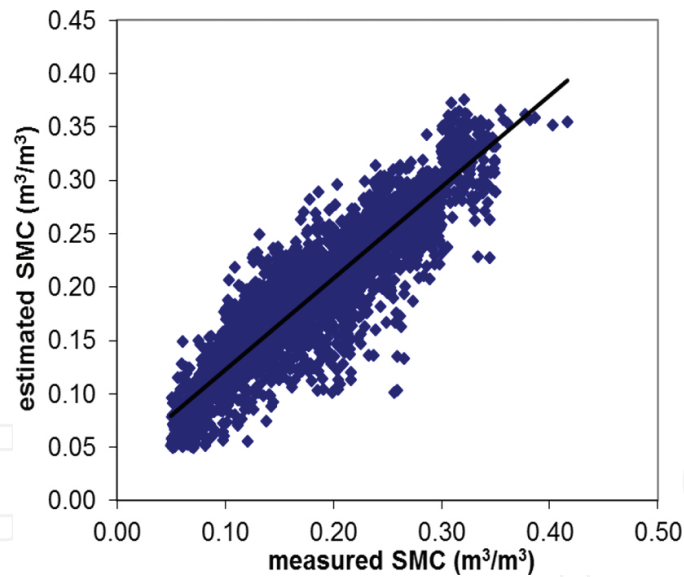

Figure 3. HydroAlgo SMC product validation (after [11]).

The SMC product validation on the Australian and Mongolian data, which were not used for the training, resulted in a determination coefficient $R^2 = 0.8$ (ANN output vs estimated SMC), root-mean-square error RMSE = 0.03 m^3/m^3, and BIAS = 0.02 m^3/m^3 (**Figure 3**).

The peculiar characteristics of ANN allowed adapting this algorithm for working on given test areas with a specific updating of the training process, which is devoted to maximize the performances of the algorithm on the area, losing something of the algorithm capabilities for

the global retrieval. Following this approach, it was possible to obtain the algorithm for the SMC retrieval in central Italy, which was presented in [23]. In that work, HydroAlgo-derived SMC has been compared with simulated SMC data obtained from the application of a well-established soil water balance model (SWBM) [24] in central Italy (Umbria region), with the aim of exploiting the potential of AMSR-E/2 for SMC monitoring on a regional scale and in heterogeneous environments too. For this application, the 10% of about 450,000 AMSR-E acquisitions collected over the test area and corresponding SMC values simulated by SWBM, obtained with a random sampling, were added to the training set of the original HydroAlgo implementation. The algorithm trained with this updated dataset was validated on the remaining 90% of the available data, allowing an appreciable improvement of the accuracy with respect to the original implementation. In detail, this 'supervised' approach allowed obtaining an overall increase of the average R from 0.71 to 0.84 and a corresponding decrease of RMSE from 0.058 to 0.052 m^3/m^3, with respect to the original implementation of HydroAlgo applied to the same dataset.

3.2. PWC processor

The PWC processor was based on the well-demonstrated sensitivity of the polarization difference, expressed as the polarization index, at various AMSR-E frequencies to PWC. Past research has shown that the microwave polarization index (PI), defined as the difference of the first two Stokes parameters (H- and V-polarization) divided by their sum, especially at X- and Ku-bands, is directly related to τ and therefore to the seasonal changes in PWC and LAI.

It is generally known indeed that microwave emission, expressed as brightness temperature (Tb), depends on canopy growth but also on plant geometry and structure. Therefore, Tb temporal trends vary according to the vegetation type in terms of scatterer dimensions and observation frequency. Tb tends to increase as the biomass of plants characterized by small leaves and thin stems increases, whereas it has an opposite behaviour for crops characterized by large leaves and thick stalks. On the other hand, the PI at the same frequency usually decreases as the biomass of different vegetation types increases, resulting rather independent on crop type [25, 26].

This allowed using PI at higher frequencies with the twofold purpose of compensating the vegetation effect on soil emission at lower frequencies and of estimating directly the PWC.

The capabilities of ANN in merging different inputs into a single retrieval algorithm allowed making synergistic use of PI at C-, X- and Ku-bands from the AMSR-E acquisitions. In order to implement, train and validate the algorithm, we identified a suitable test area in a wide portion of Africa (0–20°N/16°–17°E), which extended from the Sahara desert to Equatorial forest, and therefore included a wide range of vegetation types, biomass amount and land-scapes. The area was also chosen for the presence of large and homogeneous regions, which allowed mitigating the effects related to the coarse resolution of AMSR-E. Several AMSR-E swaths on the area have been collected and resampled on a fixed grid, in order to be repre-sentative of the entire seasonal cycle of vegetation. This process resulted in a dataset of about 10,000 radiometric acquisitions at the considered frequencies, from C- to Ka-bands. Consid-ering the difficulties in obtaining 'ground truth' data of PWC for validating the algorithm on

large or global scale, the validation was carried out referring to PWC values derived from NDVI thanks to the relationship established by [22]. Although this relationship was initially developed for corn and soybean crops, it can be considered valid for other types of vegetation too.

In detail, the 'reference' PWC was derived from NDVI data obtained from http://free.vgt.vito.be/home.php, resulting from 10 days of SPOT4 acquisitions on the African continent. These data were resampled on the fixed grid and compared with the corresponding satellite acquisitions, in both ascending and descending orbits. Two different ANN have been defined and trained independently, in order to better account for the large differences between Tb data collected in ascending and descending orbits.

A subset of 15% of the data available was considered for generating the datasets for training, testing and validating each ANN (60–20–20%, randomly sampled), and the remaining 85% of data (about 8500 samples) was considered for the independent validation of the algorithm, to which the result presented in **Figure 4** is referred.
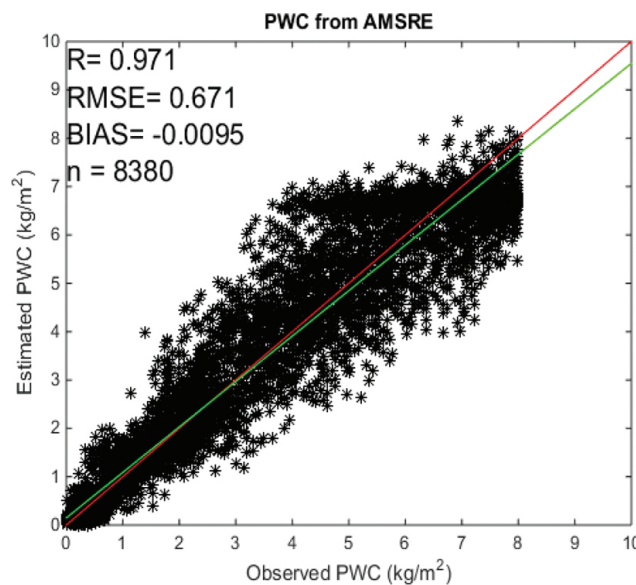

Figure 4. Validation results: PWC estimated by the algorithm as a function of the PWC derived from NDVI and considered as ground truth.

The ANN optimization process resulted in an architecture with two hidden layers of 11 + 11 neurons, with a transfer function of type 'tansig'. The validation returned encouraging results, with a RMSE error on the PWC retrieval <1 kg/m^2, and a correlation coefficient R = 0.97.

3.3. SD processor

As per the SMC processor, the implementation of the SD processor was based on a dataset provided by JAXA and composed of AMSR-E acquisitions and direct SD and air temperature measurements collected in the eastern part of Siberia. The measurements covered a flat area of about 20′ in latitude, 45′ in longitude, at an average altitude of 300 m asl, covered by low vegetation. In this region, snow was generally present from the beginning of October to the

end of May, with a depth that did not exceed 50 cm. The ground measurements were covering seven winter seasons, from October 2002 to May 2009. By combining the AMSR-E acquisitions and the related direct measurements of SD and air temperature, it was possible to obtain a dataset of 17,000 values for training and testing the ANN. As for the previously described processors, two ANN have been developed and trained separately for the ascending and descending orbits.

The validation was carried out on a different area of about 200 × 200 km located between Finland and Norway, obtaining the following statistics: R = 0.88, RMSE = 9.13 cm and BIAS = -0.95 cm.
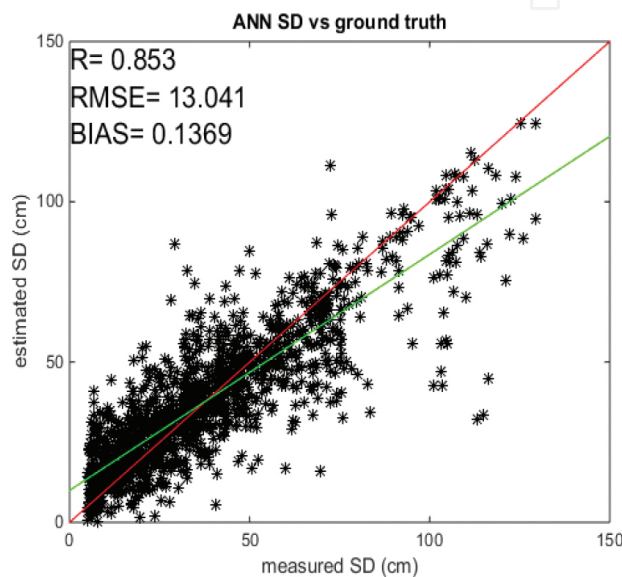
Figure 5. Validation results: SD estimated by the ANN algorithm as a function of the SD derived from ground measurements (after [26]).

The algorithm was then adapted for working on alpine areas, in which snow properties suffer dramatic spatial variations that cannot be easily reproduced by spaceborne microwave radiometers, due to their coarse spatial resolution. This limitation was overcome by setting up a method for evaluating and correcting the effects of the complex orography, of the different footprint in the different AMSR-E channels, and of the forest coverage. The detailed description of this method can be found in [27]. The test and validation were carried out on a test area of about 100 × 100 km^2 located in the eastern Italian Alps, using AMSR-E data collected during the winters between 2002 and 2011. The obtained results were encouraging: the correlation between SD estimated by the algorithm and the corresponding ground truth resulted in R = 0.85 and RMSE = 13 cm considering the descending orbits, while the retrieval accuracy worsened when considering the ascending ones (**Figure 5**).

In this case, the training of the ANN was updated by adding to the original training set a subset of the data collected on the area and the validation was carried out on the remaining part of the dataset, for a total of more than 1400 daily AMSR-E acquisitions and corresponding ground truth.

4. The SAR ANN algorithm for SMC, PWC and SWE

Similarly to HydroAlgo, a further ANN-based algorithm has been implemented for working with SAR data at C- and X-bands, aiming at generating SMC maps of bare or slightly vegetated soils, PWC maps of agricultural vegetation and SWE maps of snow covered surfaces. The algorithm takes advantage of the high resolution of the considered sensors, which can, however, provide data at local or regional scale, since SAR images cover usually areas not larger than 100×100 km^2. The other main difference in respect to HydroAlgo is that the existing SAR systems work at a single frequency and the obtainable product depends on the frequency, polarizations and ancillary information available. For instance, C-band cannot retrieve SD and it is more suitable for monitoring SMC, being less affected to the vegetation effects which drive instead the scattering mechanism at X-band. Depending on the input SAR data, the output resolution ranges between 10×10 and 100×100 m^2. **Figure 6** represents the algorithm flowchart: after a common pre-processing, the algorithm splits in three different branches, one for each output product. Details on the implementation of each processor can be found in [28, 29] for SMC processor [30], for PWC processor and [30] for SWE processor.
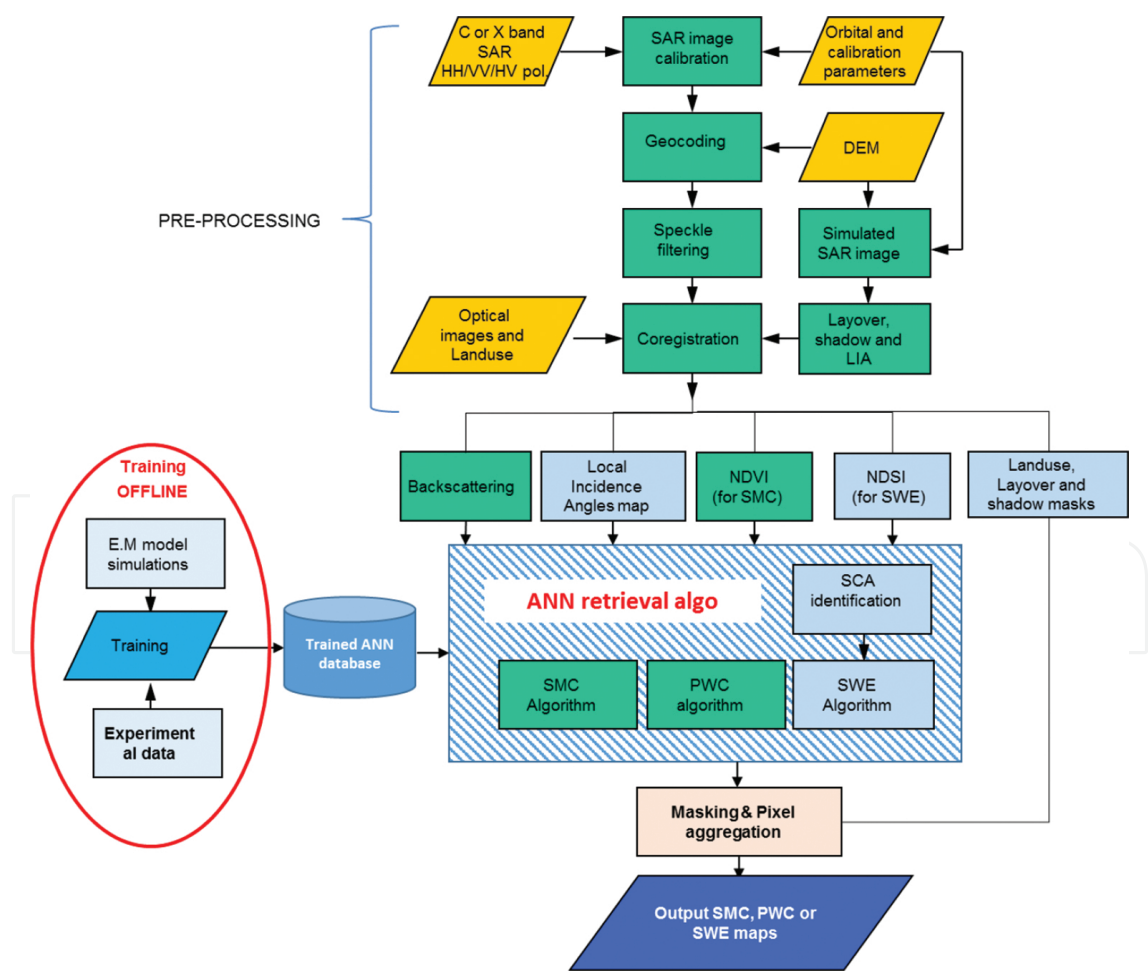

Figure 6. Flow chart of the SAR algorithm.

4.1. SMC processor

The recent generation of SAR sensors can operate in several acquisition modes and provide images at different polarizations and acquisition geometries. For enhance the retrieval accuracy, the algorithm has been implemented with a dedicated ANN for each configuration of inputs, namely the backscattering coefficients ($\sigma°$) in VV- or HH-polarization with and without the ancillary information on vegetation, represented by co-located NDVI from optical sensor, and VV + VH or HH + HV combinations. Consequently, the algorithm was composed by 6 + 6 ANN trained independently for C- and X-band, respectively. Following the strategy presented in Section 2, the dataset implemented for the ANN training was obtained by combining the available SAR images, the corresponding direct measurements of the surface parameters, and a large set of data simulated using e.m. forwards models.

Simulated backscattering values at all polarizations were obtained by coupling OH [31] and vegetation water cloud [33] models. This quite simple but widely validated combination offers several advantages with respect to more sophisticated formulations, namely the reduced set of input parameters needed for simulating the backscatter, the fast computation and the reliable accuracy. In detail, the OH model simulates the surface scattering from bare rough surfaces: with respect to IEM/AIEM, it is able to simulate both co- and cross-polarizations, accounting for the soil surface roughness by only using the height standard deviation (HSTD, in cm) parameter. The VWC model is a simplified implementation of RTT. It accounts for volume scattering of vegetation over the soil, for the attenuation effect on the soil scattering (simulated by OH model) and for the soil—vegetation interaction, requiring as inputs PWC and observation angle only. Inputs of the 'coupled' model are SMC, HSTD, PWC and the observation angle theta.

Minimum and maximum values of the soil parameters measured during the experimental campaigns (SMC, HSTD and PWC) were considered in order to define the range of variability of each soil parameter. Using a pseudorandom function drawn from the standard uniform distribution on the open interval (0, 1), rescaled in order to cover the range of each soil parameter, we generated input vectors for the e.m. model, in order to simulate the backscattering at VV, HH and HV/VH-polarizations.

This procedure was then iterated 10,000 times, thus obtaining a set of backscattering coefficients for each input vector of the soil parameters. The consistency between the experimental data and the model simulations was verified before proceeding to the training phase. The ANN training was carried out by considering the simulated $\sigma°$ at the various polarizations and the incidence angle as input of the ANN, and the soil parameters, in particular the SMC, as outputs. It should be remarked that the soil surface roughness parameter HSTD was added to the ANN outputs in order to enhance the training performances. However, an operational retrieval of surface roughness is not in the scopes of this algorithm and the roughness parameters are then disregarded in the algorithm. After training, the ANN were tested on a different dataset that was obtained by re-iterating the model simulations as described above. The use of a pseudorandom function prevented a correlation between these two datasets: this fact was particularly important in order to evaluate the capabilities of ANN to generalize the training phase and to prevent the overfitting problem. Incorrect sizing of the ANN or inadequate training could

cause the overfitting: the ANN return outputs outside the training range (outliers) when tested with input data that are not included in the training set.

The algorithm was validated using a set of experimental data collected on several test areas, mainly agricultural fields and grasslands, located worldwide. The total dataset was composed by about 700 field-averaged values of $\sigma°$ at C-band from Envisat/ASAR and about 600 at X-band from Cosmo-SkyMed (CSK), collected at various polarizations.
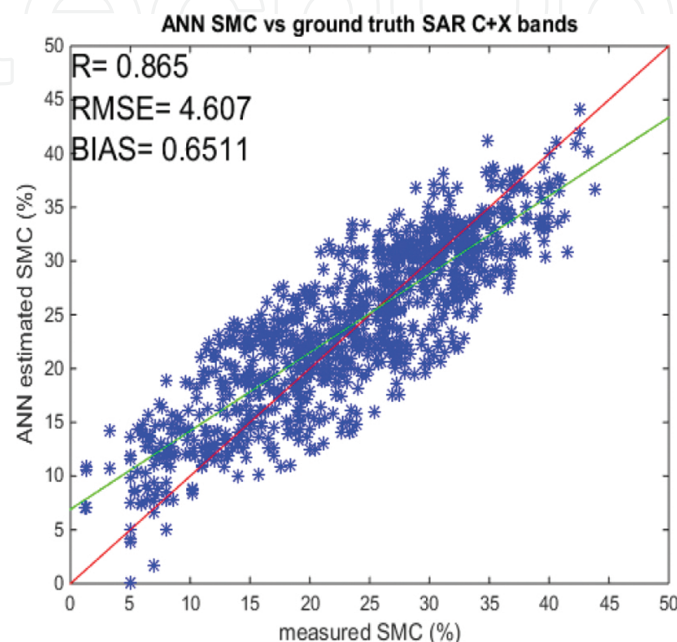
Figure 7. Validation of the SAR SMC algorithm.

Figure 7 shows the overall validation obtained by comparing the SMC values retrieved by the algorithm with the corresponding ground truth, and it corresponds to R = 0.86, RMSE = 4.6 and BIAS = 0.65. Analysing separately the two frequencies, the best results were achieved at C-band, which is more sensitive to SMC and less influenced by the vegetation than X-band. At the latter frequency, instead, the vegetation effect is dominant, although some sensitivity to SMC is detectable at least for bare and scarcely vegetated surfaces.

4.2. PWC processor

The algorithm for PWC estimate was very similar to the one for estimating SMC, and it is based on a feed-forward multilayer perceptron (MLP) ANN, trained by using the back-propagation (BP) learning rule and a RTT discrete element model, more sophisticated than the WCM [29].

The model was first validated with the experimental data collected in the 'Sesto' agricultural area located in central Italy, close to the city of Florence, mainly covered by wheat crops, and then used for generating the training set of ANN in combination with experimental data. Model simulations were iterated 10,000 times by randomly varying each input parameter in the range derived from experimental data, thus obtaining a training set able to complete the

training phase and fully define all the neurons and weights of the ANN. The dataset was split randomly in two parts, the first part for training and the second one for testing the ANN. A configuration with two hidden layers of ten perceptrons each was finally chosen as the optimal one. The validation results gave R = 0.97 and RMSE = 0.345 kg/m^2 (**Figure 8**).

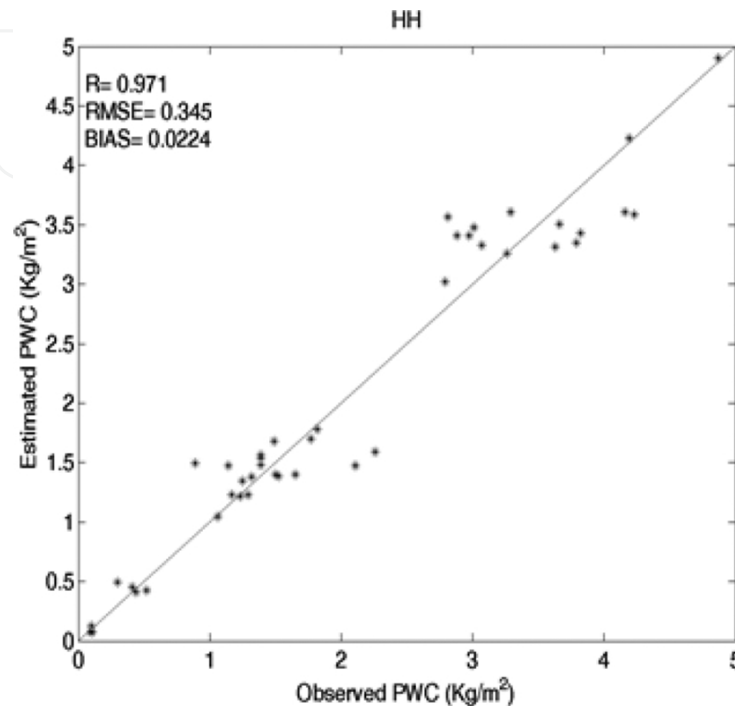**Figure 8.** PWC algorithm validation: PWC estimated vs. PWC observed.

4.3. SWE processor

In the last years, the remote-sensing community has shown a growing interest in the new generation X-band SAR satellites, such as CSK and Terra-SARX (TSX), with the aim of better understanding if at this frequency, the information on snow parameters can be retrieved and under which conditions. Although X-band is not the most suitable frequency for the retrieval of SD or SWE, since the dry snow is almost transparent at this frequency, the freezing of dedicated missions such as the ESA cold regions hydrology high-resolution observatory (CoReH$_2$O), put more interest in evaluating the potential of such a frequency for snow parameter retrieval. Basing on encouraging experimental results pointing out the relationship between σ° and SD for several winter seasons in the Italian Alps [30, 34], we have implemented an ANN-based retrieval algorithm able to estimate the SWE/SD of snow-covered surfaces from X-band SAR data.

This algorithm, which has been preliminary described in [30], is composed of two steps: first, the dry snow is identified and separated from the wet snow and from the snow-free surfaces using a well-known threshold criterion [35]. Then, the SWE retrieval by means of the ANN algorithm is attempted on the areas of the image identified as dry snow.

Inputs to the ANN are the X-band σ° measured in the available polarizations, the corresponding reference value measured in snow-free conditions, and the local incidence angle information. SWE is the ANN output.

The main problem we had to face in developing this algorithm was related to the lack of extensive sets of measurements of snow parameters, which posed some constraints in defining the training set. The available measurements are indeed sparse, and, besides being site dependent, are numerically inadequate for training the ANN and define all its neurons and weights. The training of the ANN was therefore performed by using data simulated by the dense medium radiative transfer model implementation [36, 37]. As for the other algorithms, the training set was generated by running the model simulation for the input values of snow parameters in a range derived from the direct measurements, obtaining output backscattering coefficients at the given polarizations for each input vector of snow parameters. In order to match all the acquisition modes of CSK, several ANN have been set up and trained separately, according to the combination of polarizations available from the dataset.
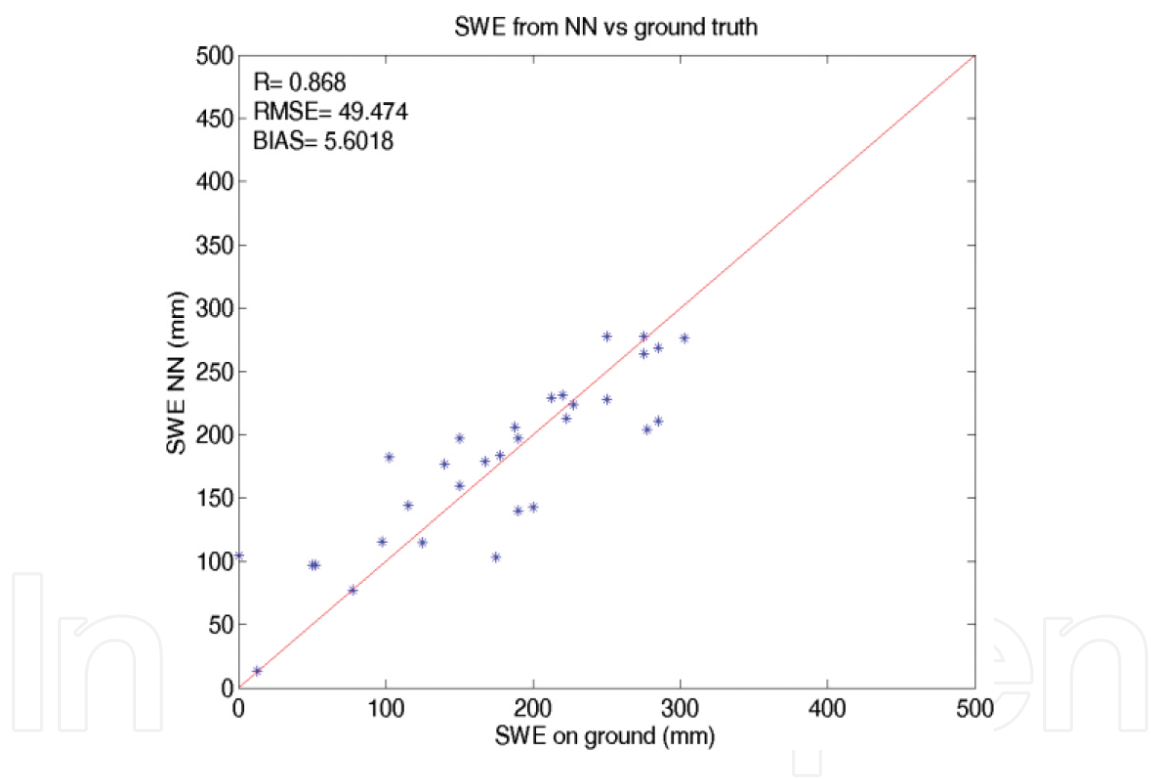

Figure 9. Validation of the SAR SWE algorithm.

Although the algorithm is not able to consider some model parameters, such the average crystal dimension, which are unavailable from in situ measurements, the training process converged successfully. In particular, for a configuration with 2 hidden layers of 13 neurons each and an activating function of 'tansig' type, assuming the availability of CSK data in two polarizations (co- and cross-polar), the validation resulted in R > 0.9, with an associated probability value (P-value) of 95% and RMSE = 50 mm of equivalent water [30]. After the test on simulated data, the algorithm was validated considering CSK images and corresponding

ground truth available on the Cordevole and Bardonecchia test areas, located in the eastern and western part of Italian Alps, respectively. The direct comparison with ground truth data resulted in an R > 0.85, RMSE = 50 mm and Bias = 5.6 mm (**Figure 9**).

5. The P + L-band SAR ANN algorithm for WV

A final example of the ANN capabilities in adapting to the retrieval of hydrological parameters from microwave remote-sensing acquisitions is represented by this ANN application to the forest WV retrieval. The algorithm takes advantage of the well-known sensitivity of low microwave frequencies such as L- and P-bands to forest biomass. However, L- and P-band SAR data available from satellite and corresponding in situ measurements were not sufficient for implementing and validating such algorithm. Therefore, we selected a dataset of airborne SAR measurements derived from the ESA project BioSAR 2010, which has been obtained through the ESA eopi portal: https://earth.esa.int/web/guest/pi-community. The dataset was composed of airborne SAR fully polarimetric images at P- and L-bands acquired in fall 2010 in Sweden by the airborne system ONERA SETHI and corresponding LiDAR measurements of forest height, which were considered as target values for training and testing the algorithm.

The WV ANN algorithm considers as inputs the co- and cross-polarized backscattering at both P- and L-bands, along with the corresponding incidence angles, without any ancillary information from other sensors. The training set was implemented by considering as ground truth the WV estimated by LiDAR. In this case, the model simulations considered for increasing the training set were based on an implementation of the water cloud model with gaps, which was initially proposed in [38, 39] and it was based on the original VWC by [32]. The model has been modified by adding a term able to account for the backscattering dependence on the observation angle. The independent validation of this algorithm was carried out on some plots for which conventional measurements of WV were available. Although the validation set was limited, the results were encouraging, with R = 0.98, RMSE = 22 t/ha and Bias = 11 t/ha.

6. Generation of maps of the target parameter at regional and global scale

The fast computation is another important advantage of the ANN-based algorithms with respect to other statistical methods. The training represents indeed the only time-consuming process; however, it is only once carried out at the beginning and it deals with the algorithm implementation, not with its application. A trained ANN are indeed able to process the input satellite data in real or near-real time. This characteristic allows an operational application of the ANN for generating maps of the target parameter at high resolution and large or global coverage. In [27], the ANN retrieval algorithm was demonstrated to be able to process 200,000 pixels/s, which correspond to about 80 s for generating a SMC map at $25 \times 25 \, m^2$ resolution from an input SAR image of $100 \times 100 \, km^2$. Besides these considerations, the output maps represent also an effective tool for verifying qualitatively the validity of the training process,

although these maps cannot be considered a real validation of the retrieval algorithms, since adequate ground truth for comparing extensively the algorithm outputs at large and global scale is barely available.
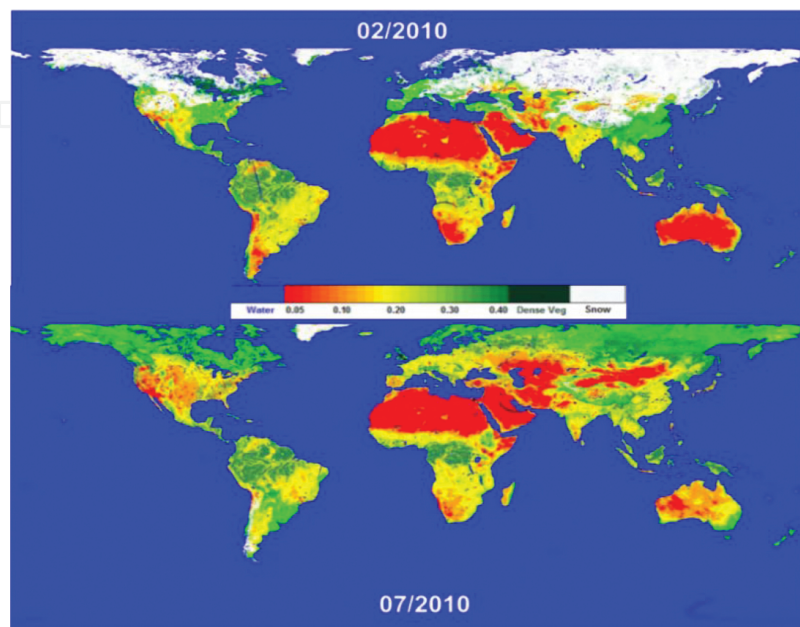

Figure 10. Global SMC and snow extent maps obtained as weekly average of AMSR-E acquisitions collected in February and June 2010 (after [11]).
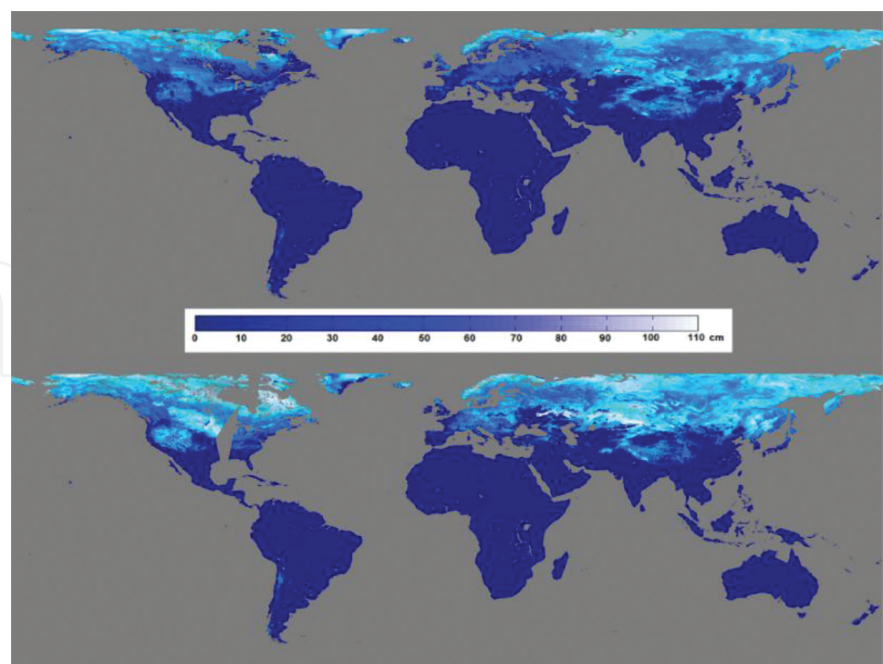

Figure 11. SD maps obtained as weekly average of AMSR-E acquisitions collected in different seasons between December 2009 and February 2010 (after [11]).
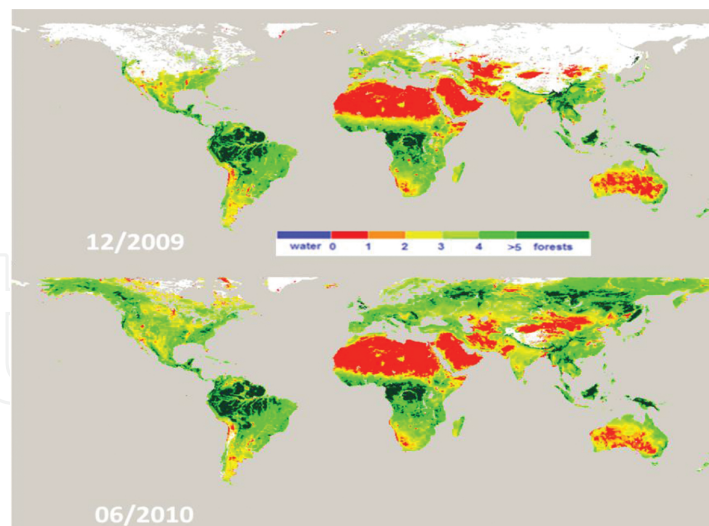
Figure 12. Global PWC maps (kg/m²) of agricultural and herbaceous vegetation from a weekly average of AMSRE ascending and descending orbits in February (top) and June (bottom) 2010.

Extreme variations of the target parameter between adjacent pixels, the presence of large percentages of outliers, and the absence of clearly detectable patterns indicate indeed that the training was not achieved successfully, although the validation in the control points resulted satisfactory. In these cases, the ANN should be retrained, by verifying that the training set is representative of the entire range of the input microwave data and output parameters considered for the specific application.
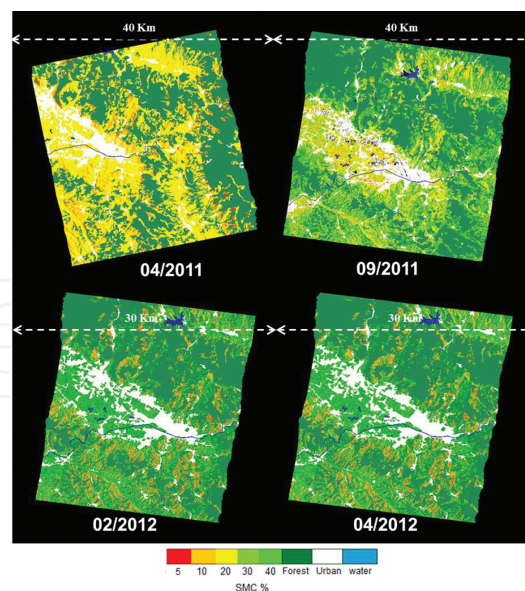

Figure 13. SMC maps from CSK images (© ASI)" for four different dates between 2011 and 2012 in the 'Sesto' test area.

As an example of the operational capabilities of the algorithms proposed here, **Figures 10** and **11** represent some examples of SMC and SD maps generated by using microwave radiometric

data through HydroAlgo and reprojected on a fixed grid spaced 0.1° × 0.1°, while examples of outputs generated by the PWC processor are represented in **Figure 12**.
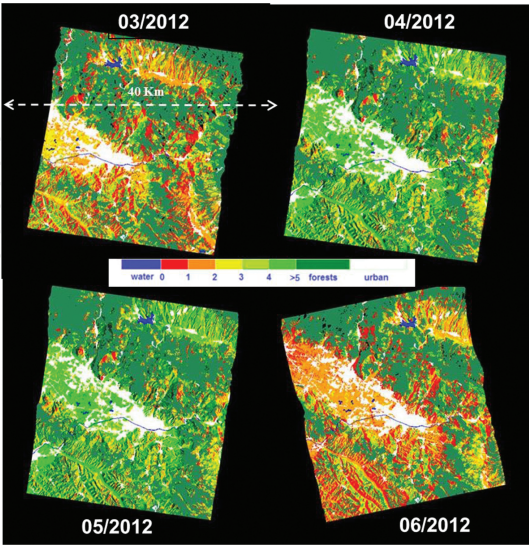

Figure 14. PWC maps from CSK images (© ASI) covering the entire seasonal cycle of wheat from march to June 2012 in the 'Sesto' test area.
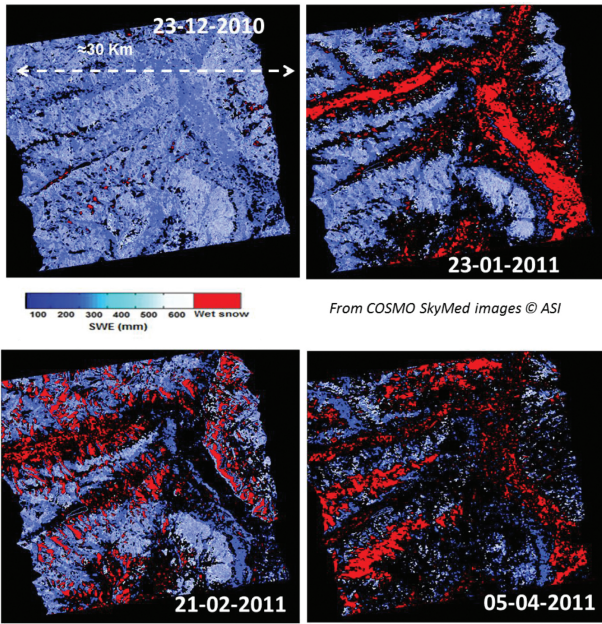

Figure 15. Maps of SWE derived from the proposed algorithm for a test area in the Italian Alps.

Maps have been obtained as weekly average of the AMSR-E acquisitions in both ascending and descending orbits for different seasons: winter and summer for SMC and PWC and two different winter periods for SD, in order to point out the sensitivity to the global spatial and temporal variations of the investigated parameter.

Examples of outputs maps generated by the SAR SMC and PWC algorithm from CSK images are represented in **Figures 13** and **14**, while maps of SWE derived from the proposed algorithm for a test area in the Italian Alps are shown in **Figure 15**. Map dimensions range between 30 × 30 and 40 × 40 km², depending on the input images. White and blue colours represent masking for urban areas and water bodies respectively.

Finally, a WV map (t/ha) has been produced for the area where SAR data at P-and L-bands have been acquired (**Figure 16**). Different colours represent different levels of forest biomass in accordance with the ground truth data collected simultaneously to the BioSAR acquisitions.

Figure 16. Forest biomass map (WV, in t/ha) from BioSAR data on the Sweden test site.

7. Conclusions

The overview of the retrieval algorithms presented here demonstrated that ANN are a powerful tool for implementing inversion algorithm, which are able to estimate the hydrological parameters from microwave satellite acquisitions, provided ANN have been trained with consistent datasets made up by both experimental and theoretical data. The flexibility of this method and the possibility of using it for both active and passive sensors with high accuracy and computational speed were confirmed. Moreover, the possibility of repeating the training with new datasets easily enables the improvement of the retrieval accuracy, making this technique flexible and adaptable to new datasets and sensors.

A further advantage of these algorithms is in their capability of merging data coming from different sources, as other sensors or ancillary information, into a unique retrieval approach. It was the case of the algorithm implemented for C- and X-band SAR, which takes advantage of the NDVI information from optical sensors (Landsat/Modis), when available, for improving the SMC retrieval accuracy.

The main constraint for accurate retrievals is due to the training process: the retrieval error may be large if the ANN are tested with data not correctly represented in the training. Large datasets are therefore needed for properly training the ANN, in order to cover the whole range of the microwave data and corresponding surface parameters. It should be noted that there is not a unique way for defining the training set. Some a priori knowledge and the support of model simulations help in setting the range of each surface parameter, in order to make the training set as representative as possible of the observed surface. Testing and validation on independent datasets (i.e. not related to the data considered for training) may indicate if the training has been achieved properly. In particular, the use of electromagnetic models for generating large training dataset is one of the best methods for avoiding the danger of 'black box' algorithms and to make sure that the results are based on physical assumptions. Since the training is performed off-line, before starting the data processing, the computational speed of ANN is not hampered by this procedure.

Acknowledgements

Author details

Emanuele Santi

Address all correspondence to: e.santi@ifac.cnr.it

Institute of Applied Physics, National Research Council (IFAC-CNR), Florence, Italy

References

[1] Dobson C., Ulaby F., Hallikainen M., El-Rayes M. 1985. Microwave dielectric behaviour of wet soil—Part II: Four component dielectric mixing models. IEEE Transactions on Geoscience and Remote Sensing. 23(4):35–46.

[2] Mironov V. L., Kosolapova L. G., Fomin S. V. 2009. Physically and mineralogically based spectroscopic dielectric model for moist soils. IEEE Transactions on Geoscience and Remote Sensing. 47(7):2059–2070.

[3] Notarnicola C. 2014. A Bayesian change detection approach for retrieval of soil moisture variations under different roughness conditions. IEEE Geoscience and Remote Sensing Letters. 11(2):414–418.

[4] Pasolli L., Notarnicola C., Bruzzone L., Bertoldi G., Della Chiesa S., Hell V., Niedrist G., Tappeiner U., Zebisch M., Del Frate F., Vaglio Laurin G. 2011. Estimation of soil moisture in an alpine catchment with RADARSAT2 images. Applied and Environmental Soil Science. Article ID 175473:12. doi:10.1155/2011/175473

[5] Pierdicca N., Pulvirenti L., Pace G. 2014. A prototype software package to retrieve soil moisture from Sentinel-1 data by using a Bayesian multitemporal algorithm. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 7(1): 153–163.

[6] Hornik K. 1989. Multilayer feed forward network are universal approximators. Neural Networks. 2(5):359–366.

[7] Linden A., Kinderman J. 1989. Inversion of multi-layer nets. Proceedings of the International Joint Conference on Neural Networks. 2:425–443.

[8] Del Frate F., Ferrazzoli P., Schiavon G. 2003. Retrieving soil moisture and agricultural variables by microwave radiometry using neural networks. Remote Sensing of Environment. 84(2):174–183. doi:10.1016/S0034-4257(02)00105-0

[9] Elshorbagy A., Parasuraman K. 2008. On the relevance of using artificial neural networks for estimating soil moisture content. Journal of Hydrology. 362:1–18. Available at: http://www.sciencedirect.com/science/article/pii/S0022169408004204

[10] Paloscia S., Pampaloni P., Pettinato S., Santi E. 2008. A comparison of algorithms for retrieving soil moisture from ENVISAT/ASAR images. IEEE Transactions on Geoscience and Remote Sensing. 46(10):3274–3284.

[11] Santi E., Pettinato S., Paloscia S., Pampaloni P., Macelloni G., Brogioni M. 2012. An algorithm for generating soil moisture and snow depth maps from microwave spaceborne radiometers: HydroAlgo. Hydrology and Earth System Sciences. 16:3659–3676. doi:10.5194/hess-16-3659-2012

[12] Nelder J. A., Mead R. 1965. A simplex method for function minimization. Computer Journal. 7:308–313.

[13] Gruber A., Paloscia S., Santi E., Notarnicola C., Pasolli L., Smolander T., Pulliainen J., Mittelbach H., Dorigo W., Wagner W. 2014. Performance inter-comparison of soil moisture retrieval models for the MetOp-A ASCAT instrument. Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 2455–2458. doi:10.1109/IGARSS.2014.6946969

[14] Mladenova I. E., Jackson T. J., Njoku E., Bindlish R., Chan S., Cosh M. H., Holmes T. R. H., de Jeu R. A. M., Jones L., Kimball J., Paloscia S., Santi E. 2014. Remote monitoring of soil moisture using passive microwave based techniques—theoretical basis and

overview of selected algorithms for AMSR-E. Remote Sensing of Environment. 144:197–213. doi:10.1016/j.rse.2014.01.013

[15] Paloscia S., Santi E., Pettinato S., Mladenova I. E., Jackson T. J., Bindlish R., Cosh M. H. 2015. A comparison between two algorithms for the retrieval of soil moisture using AMSR-E data. Frontiers in Earth Sciences. 3(16):1–10. doi:10.3389/feart.2015.00016

[16] Santi E., Paloscia S., Pettinato S., Fontanelli G. 2015. Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors. International Journal of Applied Earth Observations and Geoinformation. doi: 10.1016/j.jag.2015.08.002

[17] Lobl E. 2001. Joint advanced microwave scanning radiometer (AMSR) science team meeting. Earth Observer. 13(3):3–9.

[18] Imaoka K., Takashi M., Misako K., Marehito K., Norimasa I., Keizo N. 2012. Status of AMSR2 instrument on GCOM-W1, earth observing missions and sensors: development, implementation, and characterization II. Proceedings of SPIE 2012, 852815. doi: 10.1117/12.977774

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

[20] 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

[21] Paloscia S., Pampaloni P. 1988. Microwave polarization index for monitoring vegetation growth. IEEE Transactions on Geoscience and Remote Sensing. 26(5):617–621.

[22] Jackson T. J., Cosh M. H., Bindlish R., Starks P. J., Bosch D. D., Seyfried M. S., Goodrich D. C., Moran M. S. 2010. Validation of advanced microwave scanning radiometer soil moisture products. IEEE Transactions on Geoscience and Remote Sensing. 48(12):4256–4272.

[23] Santi E., Paloscia S., Pettinato S., Brocca L., Ciabatta L. 2015. Robust assessment of an operational algorithm for the retrieval of soil moisture from AMSR-E data in central Italy. Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS). pp. 1288–1291. doi:10.1109/IGARSS.2015.7326010

[24] Brocca L., Camici S., Melone F., Moramarco T., Martinez-Fernandez J., Didon-Lescot J.-F., Morbidelli R. 2014. Improving the representation of soil moisture by using a semi-analytical infiltration model. Hydrological Processes. 28(4):2103–2115. doi:10.1002/hyp.9766

[25] Choudhury B. J., Tucker C. J. 1987. Monitoring global vegetation using Nimbus-7 37 GHz Data Some empirical relations. International Journal of Remote Sensing. 8:1085–1090.

[26] Paloscia S., Pampaloni P. 1992. Microwave vegetation indexes for detecting biomass and water conditions of agricultural crops. Remote Sensing of Environment. 40:15–26.

[27] Santi E., Pettinato S., Paloscia S., Pampaloni P., Fontanelli G., Crepaz A., Valt M. 2014. Monitoring of Alpine snow using satellite radiometers and artificial neural networks. Remote Sensing of Environment. 144:179–186. doi:10.1016/j.rse.2014.01.012

[28] Paloscia S., Pettinato S., Santi E., Notarnicola C., Pasolli L., Reppucci A. 2013. Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. Remote Sensing of Environment. 134:234–248. doi:10.1016/j.rse.2013.02.02721

[29] Santi E., Paloscia S., Pettinato S., Notarnicola C., Pasolli L., Pistocchi A. 2013. Comparison between SAR soil moisture estimates and hydrological model simulations over the Scrivia test site. Remote Sensing. 5:4961–4976. doi:10.3390/rs5104961

[30] Paloscia S., Santi E., Fontanelli G., Montomoli F., Brogioni M., Macelloni G., Pampaloni P., Pettinato S. 2014. The sensitivity of Cosmo-SkyMed backscatter to agricultural crop type and vegetation parameters. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 7(7):2856–2868.

[31] Pettinato S., Santi E., Brogioni M., Paloscia S., Palchetti E., Xiong C. 2013. The potential of COSMO-SkyMed SAR images in monitoring snow cover characteristics. IEEE Geoscience and Remote Sensing Letters. 10(1):9–13.

[32] Oh Y., Sarabandi K., Ulaby F. T. 1992. An empirical model and an inversion technique for radar scattering from bare surfaces. IEEE Transactions on Geoscience and Remote Sensing. 30:370–381.

[33] Attema E. P. W., Ulaby F. T. 1978. Vegetation modeled as a water cloud. Radio Science. 13:357–364.

[34] Brogioni M., Cagnati A., Crepaz A., Paloscia S., Pampaloni P., Pettinato S., Santi E., Xiong C., Shi J. C. 2014. Model investigations of backscatter for snow profiles related to avalanche risk. Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS). pp. 2415–2418. doi:10.1109/IGARSS.2014.6946959

[35] Nagler T., Rott H. 2000. Retrieval of wet snow by means of multitemporal SAR data. IEEE Transactions on Geoscience and Remote Sensing. 38:754–765.

[36] Liang D., Xu X., Tsang L. 2008. The effects of layers in dry snow on its passive microwave emissions using dense media radiative transfer theory based on the quasicrystalline approximation (QCA/DMRT). IEEE Transactions on Geoscience and Remote Sensing. 46:3663–3671.

[37] Tsang L., Pan J., Liang D., Li Z. X., Cline D., Tan Y. H. 2007. Modeling active microwave remote sensing of snow using dense media radiative transfer (DMRT) theory with multiple scattering effects. IEEE Transactions on Geoscience and Remote Sensing. 45(4): 990–1004.

[38] Santoro M., Eriksson L., Askne J., Schmullius C. 2006. Assessment of stand-wise stem volume retrieval in boreal forest from JERS-1 L-band SAR backscatter. International Journal of Remote Sensing. 27:3425–3454.

[39] Fransson J. E. S., Israelsson H. 1999. Estimation of stem volume in boreal forests using ERS-1 C- and JERS-1 L-band SAR data. International Journal of Remote Sensing. 20:123–137.