International Journal of Applied Earth Observation and Geoinformation xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors

Emanuele Santi*, Simonetta Paloscia, Simone Pettinato, Giacomo Fontanelli

Institute of Applied Physics—National Research Council (IFAC-CNR), Via Madonna del Piano 10, Firenze, Italy

ARTICLE INFO

Article history:
Received 24 March 2015
Received in revised form 29 July 2015
Accepted 6 August 2015
Available online xxx

Keywords:
Artificial neural networks
Soil moisture content
Synthetic Aperture Radar (SAR)
Scatterometer
Microwave radiometers

ABSTRACT

Among the algorithms used for the retrieval of SMC from microwave sensors (both active, such as Synthetic Aperture Radar-SAR, and passive, radiometers), the artificial neural networks (ANN) represent the best compromise between accuracy and computation speed. ANN based algorithms have been developed at IFAC, and adapted to several radar and radiometric satellite sensors, in order to generate SMC products at a resolution varying from hundreds of meters to tens of kilometers according to the spatial scale of each sensor.

These algorithms, which are based on the ANN techniques for inverting theoretical and semi-empirical models, have been adapted to the C- to Ka- band acquisitions from spaceborne radiometers (AMSR-E/AMSR2), SAR (Envisat/ASAR, Cosmo-SkyMed) and real aperture radar (MetOP ASCAT).

Large datasets of co-located satellite acquisitions and direct SMC measurements on several test sites worldwide have been used along with simulations derived from forward electromagnetic models for setting up, training and validating these algorithms. An overall quality assessment of the obtained results in terms of accuracy and computational cost was carried out, and the main advantages and limitations for an operational use of these algorithms were evaluated.

This technique allowed the retrieval of SMC from both active and passive satellite systems, with accuracy values of about 0.05 m^3/m^3 of SMC or better, thus making these applications compliant with the usual accuracy requirements for SMC products from space.

1. Introduction

The amount of water stored in the soil is an essential variable controlling many biophysical processes that impact water, energy, and carbon exchanges at the land-atmosphere interface. In-situ soil moisture (SMC) measurements are labor intensive and site-specific, moreover, frequent and spatially distributed soil moisture measurements, at different spatial scales, are advisable for the most part of the applications related to the environmental disciplines, such as climatology, meteorology, hydrology and agriculture. The possibility of observing soil moisture and its temporal evolution from space is, therefore, regarded as being extremely attractive.

Among the instruments operating from space for the observation of the Earth surface, the sensors operating in the microwave portion of the spectrum have received most attention because this frequency range has the unique ability to return information on media (atmosphere, vegetation, soil) that are opaque to the much shorter visible/near-infrared and thermal wavelength and, that is most important, because microwave scattering and emission are directly related to the water content of the observed target. In particular, remote sensing from active (Synthetic Aperture Radar—SAR and scatterometer) and passive sensors (radiometers) have demonstrated to be good and flexible tools for observing the moisture content of the first centimeter layer of soil and for detecting its spatial and temporal variations from radar (e.g., Barret et al., 2009; Notarnicola et al., 2006, 2008; Paloscia et al., 2004; Pathe et al., 2009; Pierdicca et al., 2010; Wagner et al., 1999, 2007) and radiometric microwave sensors (e.g., Jackson, 1993; Jackson et al., 2010; Mladenova et al., 2014; Paloscia et al., 2006).

These two types of sensors (radiometers and SAR) retrieve information of the Earth surface at different scales. SAR, in fact, due to antenna synthesis, reaches a very high ground resolution, in the order of a few meters; whereas, radiometers from space have a much coarser spatial resolution, in the order of tens of kilometers. These characteristics allowed the use of these sensors for different applications. Fine ground resolutions allowed in fact a more detailed investigation of the surface and can give information on small scale phenomena, useful for agricultural applications and water management at a farm or basin scale. The large scale of spaceborne radiometers is instead very useful for climatic applications

* Corresponding author. Fax: +39 055 5226434.
E-mail address: e.santi@ifac.cnr.it (E. Santi).

http://dx.doi.org/10.1016/j.jag.2015.08.002

and for detecting the trend of large scale phenomena, as global changes, e.g., deforestation and desertification processes.

The retrieval of soil parameters from active and/or passive microwave measurements is nonetheless not trivial, due to the non-linearity of the relationships between radar or radiometric acquisitions and ground parameters, and because, in general, more than one combination of soil parameters (soil moisture, roughness, vegetation cover, etc.) has the same electromagnetic response. Thus, in order to minimize the uncertainties and enhance the performance of soil parameter retrieval from remote sensing data, statistical approaches based on the Bayes theorem and learning machines are widely adopted for implementing the retrieval algorithms (e.g. Notarnicola, 2014; Pasolli et al., 2011; Pierdicca et al., 2014).

In this framework, the artificial neural networks (ANNs) represent an interesting tool for implementing accurate and flexible retrieval algorithms, which 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 they can be trained to represent arbitrary input-output relationships (Hornik, 1989; Linden and Kinderman, 1989 Linden and Kinderman, 1989). In the training phase, training patterns are sequentially presented to the network and the interconnecting weights of each neuron are adjusted according to a learning algorithm. The trained ANN can be considered as a type of non-linear least mean square interpolation formula for the discrete set of data points in the training set.

The effectiveness of ANNs in solving remote sensing problems has been well demonstrated, since they can easily merge data coming from different sources into a unique retrieval algorithm. ANNs have been successfully applied to many inverse problems in the remote sensing field, and in particular to retrieve soil moisture at local scale from SAR (e.g., Del Frate et al., 2003; Elshorbagy and Parasuraman 2008; Paloscia et al., 2008) or radiometric (e.g., Santi et al., 2012) observations. The comparison of retrieval algorithms carried out in Paloscia et al. (2008) demonstrated that ANNs, with respect to other widely adopted statistical approaches based on Bayes theorem and Nelder–Mead minimization, offer the best compromise between retrieval accuracy and computational cost.

ANNs in some cases have been used essentially as a black box, without further effort for understanding the underlying processes and the physics behind them. The strategy for minimizing these problems is mainly based on the use of both extensive datasets and model simulations for the training phase of ANN. The studies presented in Paloscia et al. (2010) and Santi et al. (2013) pointed out the potential of the ANN technique in easily and effectively ingesting information extracted from different sources for improving the retrieval process, such as NDVI derived from optical remote sensing imagery for taking into account the presence of vegetation on the ground. In Paloscia et al. (2013) nevertheless, the importance of a robust and extensive reference dataset for the training was confirmed, in order to obtain a retrieval algorithm able to work at large and global scale with a satisfactory accuracy.

An application of retrieval algorithms to MetOp-A ASCAT backscatter data can be found in Gruber et al. (2014) where the ANN techniques are compared with other retrieval algorithms, in their ability to retrieve soil moisture on a global scale. Correlation and triple collocation analysis were performed using in situ and land surface model data as a reference, pointing out again the effectiveness of the ANNs compared with other inversion approaches.

In this paper, an overview of the results obtained with ANN algorithms applied to different microwave sensors for the retrieval of soil moisture at both local and global scales is presented. We have summarized and homogenized here the results presented in other papers published in both international journals and conference proceedings, by improving and tuning the algorithms used in past research, and validating them with new datasets. In some cases, more accurate results have been obtained, with improved accuracy of the estimated parameter. In this work, a unique procedure for all the algorithms was thus implemented, by choosing the same strategy for training, test, and validation.

In Section 2 a detailed discussion of the methods and in particular of the training of ANN is carried out. In the following sections, the application of the ANN to microwave radiometers, scatterometers and SAR sensors is presented along with the main findings of the methods. Finally, some soil moisture maps obtained at different spatial scales by using the data from these sensors are shown for an overall validation of the implemented algorithms.

2. Implementing and training the artificial neural networks

The algorithms presented in this work are based on feed-forward multilayer perceptron (MLP) ANNs, with a certain number of hidden layers of neurons between the input and the output. In MLPs, successive layers of neurons are fully interconnected, with trainable connection weights that control the strength of the connections. The ANN training was based on the back-propagation learning rule, which is an iterative gradient descent algorithm that is designed to minimize the mean square error between the desired target vectors and the actual output vectors. 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^2 and RMSE, by increasing it until negligible improvements were obtained.

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. This ANN is 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 for the training data, so as to prevent overfitting problems. Overfitting is related to the oversizing of the ANN, and may cause considerable errors when testing ANN with input data that is not included in the training set (Moody et al., 1992; Tetko et al., 1995). In other words, the ANN is able to reproduce the training set with high accuracy but fails the test procedure.

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, i.e., input data out the range of the input parameters included in the training set. Logistic sigmoid (logsig) and tangental 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 is 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 non-linearity, g(a), as $1/(1+e^{-a})$. Alternatively, multilayer networks can use the tansig function, $\tanh(a) = (e^{a} - e^{-a})/(e^{a} + e^{-a})$.

Besides these problem, however, the main constraint for obtaining good accuracies with the ANN approach, as it has been demonstrated in Paloscia et al. (2013) is represented by the statis-