



A neural network algorithm for cloud fraction estimation using NASA-Aura OMI VIS radiance measurements

G. Saponaro¹, P. Kolmonen¹, J. Karhunen³, J. Tamminen², and G. de Leeuw^{1,4}

¹Finnish Meteorological Institute, Climate Change Unit, P.O. Box 503, 00101 Helsinki, Finland

²Finnish Meteorological Institute, Earth Observation Unit, P.O. Box 503, 00101 Helsinki, Finland

³Department of Information and Computer Science, Aalto University School of Science, P.O. Box 15400, 00076 Espoo, Finland

⁴Department of Physics, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland

Correspondence to: G. Saponaro (giulia.saponaro@fmi.fi)

Received: 20 October 2012 – Published in Atmos. Meas. Tech. Discuss.: 13 February 2013

Revised: 10 July 2013 – Accepted: 23 July 2013 – Published: 9 September 2013

Abstract. The discrimination of cloudy from cloud-free pixels is required in almost any estimate of a parameter retrieved from satellite data in the ultraviolet (UV), visible (VIS) or infrared (IR) parts of the electromagnetic spectrum. In this paper we report on the development of a neural network (NN) algorithm to estimate cloud fractions using radiances measured at the top of the atmosphere with the NASA-Aura Ozone Monitoring Instrument (OMI). We present and discuss the results obtained from the application of two different types of neural networks, i.e., extreme learning machine (ELM) and back propagation (BP). The NNs were trained with an OMI data sets existing of six orbits, tested with three other orbits and validated with another two orbits. The results were evaluated by comparison with cloud fractions available from the MODerate Resolution Imaging Spectrometer (MODIS) flying on Aqua in the same constellation as OMI, i.e., with minimal time difference between the OMI and MODIS observations. The results from the ELM and BP NNs are compared. They both deliver cloud fraction estimates in a fast and automated way, and they both performs generally well in the validation. However, over highly reflective surfaces, such as desert, or in the presence of dust layers in the atmosphere, the cloud fractions are not well predicted by the neural network. Over ocean the two NNs work equally well, but over land ELM performs better.

1 Introduction

The retrieval of atmospheric constituents, such as aerosols or trace gases, land or ocean surface properties from satellite data requires accurate information on the presence of clouds. Clouds strongly reflect incoming solar radiation in the ultraviolet (UV), visible (VIS) and near infrared (NIR) parts of the electromagnetic spectrum and affect the earth-emitted radiation as detected in the thermal infrared (TIR) part of the wavelength spectrum. In the UV/VIS the cloud reflectance often overwhelms the contribution of other atmospheric constituents, most land surfaces and the ocean surface, to the top-of-atmosphere (TOA) reflectance (Koelemeijer and Stammes, 1999). For instance, for the retrieval of aerosol properties all identified cloud-contaminated pixels are usually discarded (Martins et al., 2002). Cloud detection is usually performed using several tests, depending on information being available, and different algorithms have been developed to extract information on cloud microphysical properties (Ackerman et al., 1998; Kokhanovsky et al., 2011). In this paper we are concerned with cloud detection, or the determination of cloud fraction, rather than the retrieval of cloud microphysical properties. For cloud detection, the most consolidated methods are based on thresholding techniques in histograms of the measured radiance, or reflectance, at certain wavelengths using empirically estimated thresholds, or set with additional information coming from, e.g., radiative transfer models (Dybbroe et al., 2005; Loyola, 2006; Wu et al., 2006). For best results, a combination

of TIR and UV/VIS and NIR wavelength bands is required. However, such information is not always available and other methods need to be applied.

In this paper we focus on cloud detection for the Ozone Monitoring Instrument (OMI). The challenge is the coarse spatial resolution and the lack of thermal channels. The current method for the determination of the OMI cloud mask is based on two individual tests (Stammes and Noordhoek, 2002): the first one uses a radiance threshold and the UV aerosol index, while the second test considers the spatial homogeneity of the so-called small-pixel data (van den Oord, 2002). Pixels failing either of the two tests are classified as cloudy (Acarreta and de Haan, 2002). We propose an approach using neural networks (NN) for the direct determination of the cloud fraction in each OMI pixel. The approach is based on the use of the OMI radiance measurements in the VIS part of the spectrum, together with cloud information from the Aqua-MODerate Resolution Imaging Spectrometer (MODIS). OMI and MODIS both fly in the A-train constellation but on different platforms, respectively Aura and Aqua, with a time lag of about 7 min. The proposed approach is similar to that described in Preusker et al. (2008), where the cloud screening problem for the Medium Resolution Imaging Spectrometer (MERIS), suffering from a similar problem as OMI in that there are no infrared channels, was solved by applying an NN trained with a database of simulated cloudy and cloud-free spectra. In contrast, for the training we use real data obtained from MODIS, with a spatial resolution which is much higher than that of OMI, as reference data to determine the cloud fraction in an OMI pixel.

In recent years neural networks have been adopted for a wide range of applications from atmospheric sciences to electromagnetic modeling. The developed applications include, e.g., forward and inverse radiative transfer problems (Krasnopolsky, 2008), the prediction of atmospheric parameters (Grivas and Chaloulakou, 2006), the inversion and post processing of remotely sensed data (Mas and Flores, 2008; Del Frate and Schiavon, 1998), ozone retrievals (Di Noia et al., 2012; Sellitto et al., 2011, 2012), cloud classification (Christodoulou et al., 2003), land cover classification (Aitkenhead and Aalders, 2008), and feature extraction (Del Frate et al., 2005). Below, we describe the design for the cloud detection algorithm applied to OMI cloud fraction determination.

Two different learning algorithms have been used for training the neural networks, namely the back propagation (BP) and extreme learning machine (ELM). Results from the two methods are reported and their performances over land and ocean are analysed based on the comparison with an independent set of MODIS cloud fraction data. The two neural networks are trained with a training data set consisting of six randomly selected orbits. They are subsequently applied to the test data sets consisting of three other orbits, i.e., different from the training data set, and validated using the validation data set consisting of another two independent orbits.

2 Instruments

OMI is a nadir-viewing near-UV-visible spectrometer on board NASA's Earth Observing System (EOS) Aura satellite. OMI measures radiances at 751 wavelengths covering the UV/VIS wavelength range from 349–504 nm and two UV channels (UV-1: 270–314 nm, UV-2: 306–380 nm). The nominal ground footprint is $13 \times 24 \text{ km}^2$ at nadir, in the normal global operation mode. Complete global coverage was achieved daily (Levelt et al., 2006) between 2002 and 2008, while after 2008 the global coverage is achieved in two days due to the row anomaly (Yan et al., 2012) which affects the quality of OMI level 1b radiance data. Aura flies in the A-train satellite constellation, in a polar sun-synchronous orbit. MODIS, on-board Aqua, produces many cloud related products (e.g., cloud fraction, cloud top pressure, cloud optical thickness) (Hubanks, 2012; King and Bhartia, 1992; King et al., 1998). In view of the short time separation between Aura and Aqua of about 7 minutes, the MODIS products can be used together with OMI products with quite high confidence (Stammes et al., 2008; Vasilkov, 2008; Sneep et al., 2008).

3 Neural Networks

Neural network algorithms aim at identifying the relationship between input and output variables by learning either from real or simulated reference data, rather than directly from the application of a representative physical model (Haykin, 1999; Karayiannis and Venetsanopoulos, 1993).

Owing to the fact that cloud properties are highly variable and sometimes difficult to measure directly, neural networks with their adaptive learning nature offer an attractive and computationally efficient alternative for cloud screening. It has been proven that neural networks algorithms are able to approximate any continuous multivariate non-linear function, provided that the learning data set is statistically representative of the process to be modeled and an appropriate structure for the network has been selected (Hornik et al., 1989). Some applications to atmospheric sciences were referenced in the Introduction.

One important class of neural networks is the multilayer perceptron (MLP) (Werbos, 1974). Figure 1 shows the architectural graph of a multilayer perceptron with one input layer, one hidden layer and an output layer. The input signal is fed into the input layer, flows through the network on a layer-by-layer basis, and emerges at the output layer of the network as an output signal, the response of the network to the inputs.

A node receives inputs from neighbors or external sources and uses them to compute an output signal that is propagated to other units. Within the neural network there are three types of nodes: input nodes, which receive data from outside of the network, output nodes, which send data out of the network, and hidden nodes, whose input and output signals remain within the network. The behavior of the output node

depends on the activity of the hidden nodes and the weights between the hidden and output nodes. The weights between the input and the hidden nodes determine when each hidden node is active. In the MLP, the activation function of the nodes is based on a differentiable non-linear activation function, yielding a value called the unit's activation. These functions enable the network to learn complex tasks by extracting features from the input signal. Biases are simply weights from a single node whose location is outside of the main network and whose activation is always one (Bishop, 2008). A multilayer perceptron requires to be trained and several training algorithms have been developed for an MLP structure for example Levenberg–Marquardt, or Batch which are discussed in Karayiannis and Venetsanopoulos (1993). More details on artificial neural networks can be found in Bishop (2008), Ham and Kostanic (2001) and Haykin (1999).

In this work, two learning algorithms were applied to train the MLP neural network, i.e., back propagation (BP) and extreme learning machine (ELM).

3.1 Back propagation

The error back-propagation algorithm (Rumelhart et al., 1986) is a popular learning algorithm used to train neural networks by modifying the weights during the training phase in order to model a particular learning task correctly for the training examples (Haykin, 1999). The training phase updates the weights iteratively using the negative gradient of a cost function defined as the square of the norm of the error of the current training input. Basically, error back-propagation algorithms perform two passes through each layer of the network: the first pass starts with the application of the input vector to the input nodes of the network, and its effect is forwarded through the layers. This is the forward pass during which all weights of the network are fixed. Then, a set of output data is produced as the response of the network to the input signal, and is subtracted from a desired (target) response to produce an error signal. The error signal is propagated backward through the network. This process represents the backward pass. During the backward pass the weights are adjusted to move the actual response of the network closer to the desired one in a statistical sense. The model of each node is based on a non-linear activation function.

A sigmoid activation function is used for the hidden nodes and it is presented in the following equation:

$$g(x) = \tanh(ax), \quad (1)$$

where $a > 0$ is a scaling parameter, x is the activation function input, and $g(x)$ is the activation value. This function is especially advantageous for use in neural networks trained by back propagation, because it is easy to differentiate, and thus can dramatically reduce the computation burden for training. It applies to applications whose desired output values are between 0 and 1.

The back-propagation learning does not guarantee that the final solution is the best one as the convergence of the MSE is not checked. This should be taken into account when the solution is analyzed.

3.2 Extreme Learning Machine

The extreme learning machine (Huang et al., 2006) uses a special MLP network structure with one hidden layer where the weights between the input layer and the nodes of the hidden layer are chosen randomly beforehand, and similarly for the bias terms of the hidden layer nodes. The output layer is to be taken linear. The extreme learning machine method for neural networks consists of the following steps:

- Assume that we have a training set $x_i, t_i, i = 1, 2, \dots, N$, where x_i is the i th input vector of dimension n , t_i is the corresponding target vector of dimension m , and N is the number of training data pairs.
- Choose the activation function $g(t)$ and number of nodes M in the hidden layer. In our case, the selected activation function is $g(x) = \tanh(wx + \beta)$.
- Assign randomly the hidden layer scaling parameter vectors w_j and biases $\beta_j, j = 1, 2, \dots, M$.
- Calculate the $N \times M$ hidden layer output matrix \mathbf{H} . Its elements are

$$h_{ij} = g(w_j^T x_i + \beta_j). \quad (2)$$

- Calculate the $M \times m$ weight matrix \mathbf{B} of the output layer from

$$\mathbf{B} = \mathbf{H}^+ \mathbf{T}, \quad (3)$$

where $\mathbf{T} = [t_1, t_2, \dots, t_N]^T$ is the $m \times N$ target matrix and \mathbf{H}^+ is the pseudo-inverse of the matrix \mathbf{H} . The matrix \mathbf{B} is

$$\mathbf{B} = \begin{bmatrix} w_{1,m} \\ w_{2,m} \\ \cdot \\ \cdot \\ \cdot \\ w_{M,m} \end{bmatrix}. \quad (4)$$

This learning method requires an easy implementation and it runs extremely fast as compared, for example, to the standard back-propagation algorithms (see Huang et al., 2006, for more detailed information). Since this type of algorithm does not require tuning and the hidden layer parameters can be fixed, the optimal solution can be found with a system of linear equations using the least-squares method (pseudo-inverse) and avoiding problems related to gradient learning methods (Bishop, 2008), such as local minima encountered in the back propagation (Haykin, 1999).

4 Application of neural networks to cloud detection using OMI VIS radiance measurements

To investigate the potential and limits of the application of neural networks for cloud screening, a representative data set for the observed phenomena is required. The training data set needs to be as complete as possible and of sufficient quality. In the NN training phase of the cloud detection method, the input consists of OMI measurements of TOA radiances in the VIS part of the electromagnetic spectrum and collocated (in space) MODIS cloud fraction products as described below.

4.1 The OMI training input

This section describes the OMI products included in the training data set.

OMI measures radiances at a large number of wavelengths in the VIS band, but only part of these are used as described in the next section. Radiances are converted to reflectances and scaled so that the input information for the neural networks has values between 0 and 1. The conversion was done using equation:

$$\rho(\lambda) = \frac{I(\lambda)\pi}{\cos(\theta_z)L(\lambda)}, \quad (5)$$

where ρ is the calculated reflectance, I is the OMI measured radiance at wavelengths between 349 nm and 504 nm, θ_z is the solar zenith angle, and L is the solar irradiance at wavelength λ .

OMI provides, at one wavelength (388 nm in the VIS), a five times higher spatial sampling in the flight direction than normal which is called small-pixel data. This capability can be used to provide information about spatial inhomogeneity in a pixel caused by, e.g., clouds and is therefore used as one of the cloud detection criteria as mentioned in the Introduction. The small pixel radiances are included in the level-1B data set (van den Oord, 2002) and, after conversion to small-pixel reflectances using Eq. (5), were used to calculate the variance of the reflectance in each OMI pixel. This value was added to the training data set.

The solar zenith angle (SZA), providing information about the measurement geometry, and the OMI Surface Reflectance Climatology Data Product (OMLER) were also included to the training data set. OMLER is an OMI product describing the monthly climatology of the earth's surface Lambertian equivalent reflectance (LER). LER is defined as the reflectance of an isotropic surface which matches the observed top of the atmosphere (TOA) reflectance in a purely Rayleigh scattering atmosphere, i.e., in cloud- and aerosol-free conditions. The product has a spatial resolution of 0.5×0.5 degrees and has been built by using five years of OMI data, obtained between January 2005 and December 2009 (Kleipool, 2010).

4.2 Singular value decomposition for the OMI reflectance

During the training phase, the outputs of a NN come to approximate the target values given the inputs in the training set. This ability may be useful in itself, but the purpose of using a NN is to have the outputs approximating target values given inputs that are not in the training set (generalization). Selection of an appropriate number of input variables is an important issue in building a neural network with satisfactory generalization capabilities. The purpose of input variable selection is to find the smallest set of features that can avoid overfitting the NN and produce a smaller number of local minima. The OMI reflectance data consists of 751 measurements for each pixel and a dimensionality reduction is desirable to save on computation time. Dimensionality reduction is the transformation of high-dimensional data into a representation of lower dimensionality without losing valuable information. To achieve this, we have used singular value decomposition (SVD), which is a method that converts a matrix to its diagonal form (Golub and Van Loan, 1996). In the present study the SVD procedure was implemented as follows.

Consider an $N \times M$ matrix \mathbf{X} where $N \geq M$. It is possible to represent this matrix in the r -dimensional subspace where $r \leq M$. Let $\mathbf{U} = \mathbf{X}\mathbf{X}^T$ and $\mathbf{V} = \mathbf{X}^T\mathbf{X}$ be non-negative symmetric matrices with the same eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_r$, which are ordered such that $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \lambda_r$. The square roots of these eigenvalues are called the singular values of \mathbf{X} . If we form matrices Ψ and Φ from the corresponding eigenvectors of \mathbf{U} and \mathbf{V} , then \mathbf{X} can be diagonalized as

$$\mathbf{X} = \Psi \Lambda \Phi^T, \quad (6)$$

where Λ is the diagonal matrix of the eigenvalues, i.e., $\text{diag}(\Lambda) = [\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_r}]$. Basically, each singular value represents the information content of the matrix \mathbf{X} , projected into each subspace.

The reduction of the reflectance data set is achieved by using only that part of the diagonalized system where the eigenvalues are significant.

4.3 MODIS cloud fraction training data

MODIS cloud fraction is used as reference data for training the neural networks. The spatial matching between OMI and MODIS pixels was performed using the method described by Stammes et al. (2008). In this procedure OMI ground pixel latitude-longitude corner (OMPIXCOR) data are used to construct boxes representing the pixel area. OOMPIXCOR is a separate product which was used because the OMI Level 1B data product provides geodetic latitude and longitude only for the center of each ground pixel. The MODIS geolocated data is then searched for measurements falling within each box and a MODIS pixel is considered to fall within a

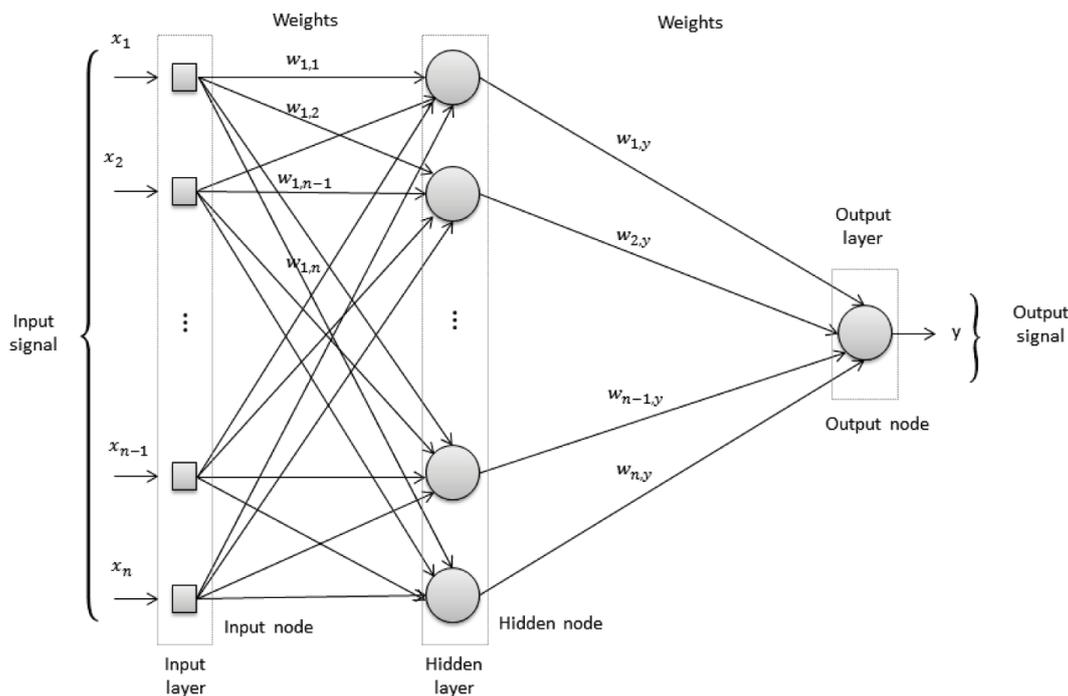


Fig. 1. Neural-network feedforward structure. x_i represents the n th input unit, and y represents the output unit.

particular OMI box if the center lies inside the OMI pixel boundary. The co-locations allow the use of MODIS data to determine the cloud fraction within an OMI pixel as the number of cloudy pixels divided by the total number of MODIS pixels falling within the considered OMI pixel. The re-gridded data is then included in the training input as the reference data.

4.4 Data set composition for training and NN structure

The data set used in this study consists of the reflectance obtained from 11 randomly chosen OMI orbits. These orbits are divided in three subsets as follows: six orbits are used for the training data set, three orbits for the test data sets and the last two are used for independent validation.

The input data set for the NN consists of the OMI SVD-reduced reflectance values, the OMLER climatological data, the solar zenith angles and the small pixel variances. The reference data set consists of the corresponding MODIS geometrical cloud fraction. The different components of the training data set which represents the input to the neural network are shown in the block diagram in Fig. 2. The neural network processes this information and provides the predicted cloud fraction.

Because of the large differences between ocean and land measurements, different models for each situation are used for each of the two neural networks used in this study (BP and ELM). To avoid overfitting of the NNs, optimization of the number of eigenvalues needed for the SVD and the

number of hidden neurons was necessary. We trained several NNs with BP and ELM and monitored the MSE and RMSE, for BP and ELM, respectively, as a function of the number of eigenvalues and hidden nodes. The combination of these parameters leading to the best performance on the test data set was used in the optimized NN. Each neural network was trained with a separate model for land and ocean. Then BP and ELM were compared as regards their performance when applied to the third, independent, validation subset.

5 Results

The back-propagation and extreme learning machine algorithms were trained with the training data set and the final weights were applied to each single orbit of the test and validation data sets. The accuracy of the cloud fraction estimates was determined. The performance of the learning algorithms in predicting cloud fraction was assessed in terms of the percentages of the pixels resulting from the test and validation data sets which were estimated to be cloudy or clear by the NN, in comparison with the same percentages as given by the MODIS re-gridded cloud fraction.

In these evaluations, two cloud thresholds were considered, i.e., 60 % and 30 %. A threshold $thr = 60\%$ implies that 60 % of an OMI pixel contains clouds. Larger values of thr imply that the pixel is cloudy while pixels with values below the thresholds indicate they are cloud-free (clear). The OMI and MODIS results are compared in Fig. 3.

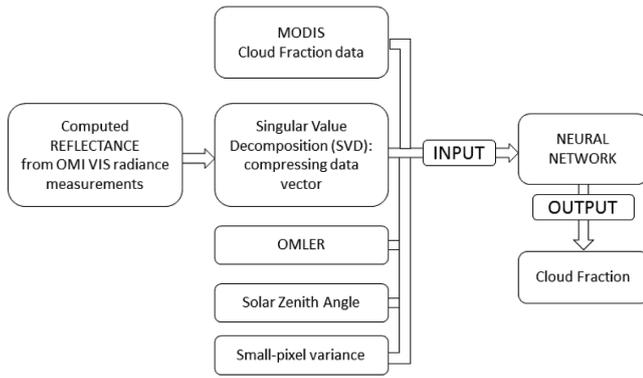


Fig. 2. Block diagram of the proposed approach for training the neural network. The training data set is composed of the target data represented by the MODIS cloud fraction re-gridded onto the OMI orbit, the compressed OMI reflectance vector data, and additional data such as climatological data (OMLER), the solar zenith angle and the computed small-pixel variance. These data form the input vector which is fed to the neural network. The neural network response is a predicted cloud fraction for the given orbit.

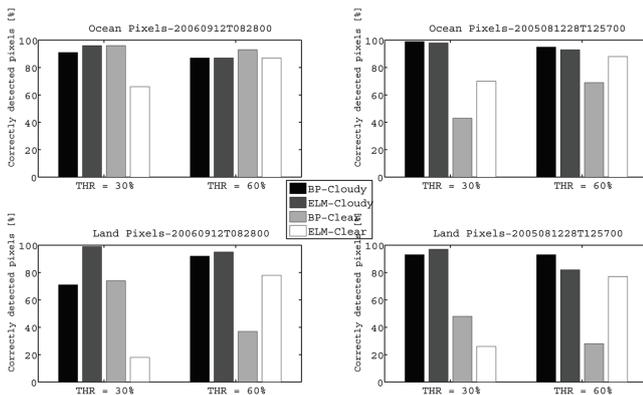


Fig. 3. Performance of backpropagation and extreme learning machine in predicting cloud fraction for divided land and ocean pixels.

The histograms in Fig. 3 represent the percentages of correctly detected cloudy and cloud-free pixels for land and ocean pixels in the validation subset for the two different thresholds and for cloudy and clear situations. The data in Fig. 3 show that both learning algorithms, BP and ELM, lead to correct estimates of cloudy pixels for both threshold values over both ocean and land in most situations. The NNs present inaccuracy when it comes to estimate small cloud fractions and on this problem the BP performs the worst.

The cloud fractions estimated from OMI data using the neural network trained by ELM or BP are compared with the MODIS re-gridded cloud fractions for each orbit of the validation data set over ocean Fig. 4 and over land Fig. 5. The color scales provide a measure for the density number of points. The images are organized in a matrix where the two rows represent the two NNs and the two columns divide the

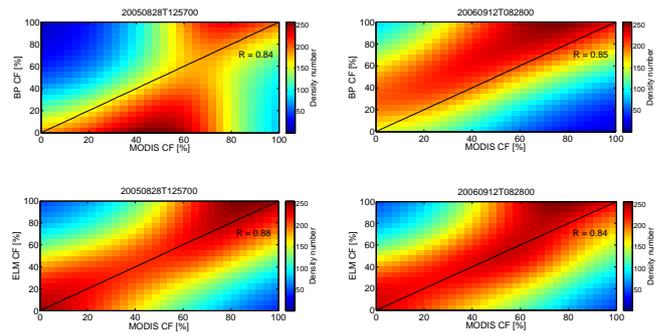


Fig. 4. Density plots showing the correlation between the cloud fraction estimated by the neural networks used in this study and MODIS data over ocean. Correlation coefficients (R) are shown in each plot.

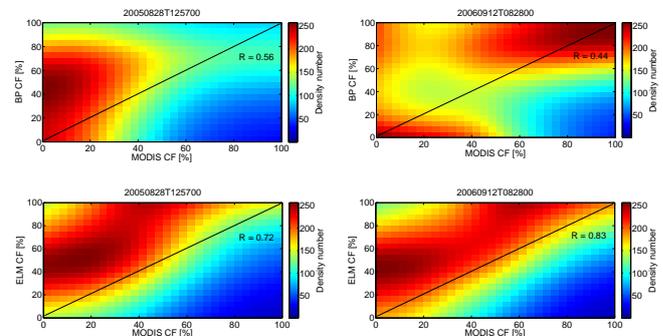


Fig. 5. Density plots showing the correlation between the cloud fraction estimated by the neural networks used in this study and MODIS data over land. Correlation coefficients (R) are shown in each plot.

orbits of the validation subset. Good correlations with R values of 0.85 and 0.88 are observed over ocean, for ELM and BP, respectively. The ocean provides a homogeneous dark surface in the UV/VIS and a good contrast between cloudy and clear pixels is expected. Over land the high reflectance measurements from bright surface represent a challenge for the NNs, although the ELM seems to be less effected than BP resulting in a R of 0.83 and 0.56, respectively, for ELM and BP.

The cloud fraction obtained with the NN using BP or ELM is compared with the MODIS geometrical CF for two validation orbits in Fig. 6 and Fig. 7. In these figures, (a) shows the MODIS geometrical cloud fraction and the grey scale indicates the cloud fraction between 0 (cloud-free) and 100 (100% cloud-covered). (b, c) show the estimated cloud fraction from BP and ELM neural networks. (d, e) show the difference between MODIS geometrical cloud fraction and the NNs estimates: the value 0 of the color scale represents a perfect agreement between the cloud fraction of the two data sets, while a value of 100 indicates a total mismatch. The results in Fig. 6 and Fig. 7 show that the cloud features are well detected, except over bright land surfaces (deserts).

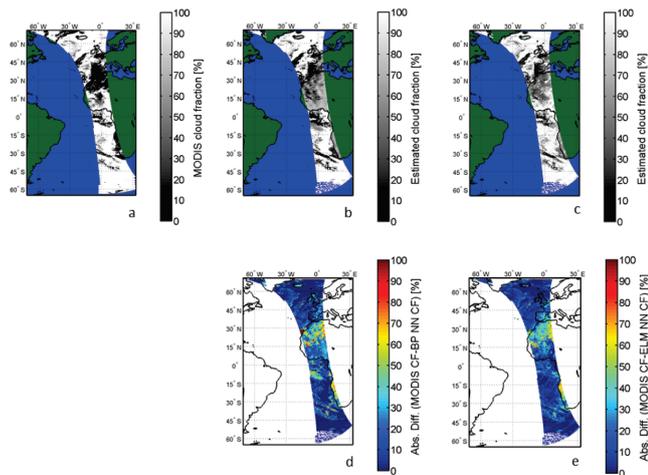


Fig. 6. Cloud fractions estimated by the BP and ELM-trained NN and comparison with MODIS CF data for the validation orbit 2005m0828t1257. **(a)** Computed MODIS geometrical cloud fraction re-located onto the OMI grid. **(b)** BP predicted cloud fraction. **(c)** ELM predicted cloud fraction. The grey-code in figures a–c ranges from 0 (cloud free) to 100 (totally cloud covered). **(d)** Absolute difference between MODIS geometrical cloud fraction and BP-predicted cloud fraction. The color-code ranges from 0 (perfect match) to 100 (complete mismatch). **(e)** Absolute difference between MODIS geometrical cloud fraction and ELM-predicted cloud fraction. The color-code in panels **(d)** and **(e)** range from 0 (perfect match) to 100 (complete mismatch)

The contribution of high ground reflectance to TOA misleads the NNs to interpret the satellite radiance measurements as if they were back reflected by clouds.

6 Conclusions

A neural networks-based solution has been explored as a contribution to detect the cloud fraction in OMI pixels using TOA radiation detected in the OMI VIS channels. This study serves as a proof of concept rather than a full study with extensive training and validation. Therefore only a limited number of OMI orbits have been used in this study, i.e., a data set of 11 OMI orbits was split in three independent data sets for training, testing and validation. In view of the vast amount of data in the VIS channels from OMI, an SVD procedure was applied to reduce the 751 channels without loss of information. This information, together with relevant auxiliary information, was used as input to two neural network learning algorithms, back propagation and ELM, and the results were compared with MODIS geometrical cloud fraction data. To this end, the selected models of the NN provide good performances during validation. The correlation coefficients between the reference MODIS cloud fraction and the estimated cloud fraction from NNs are found to be approximately 0.85 over ocean. Worst performances are observed

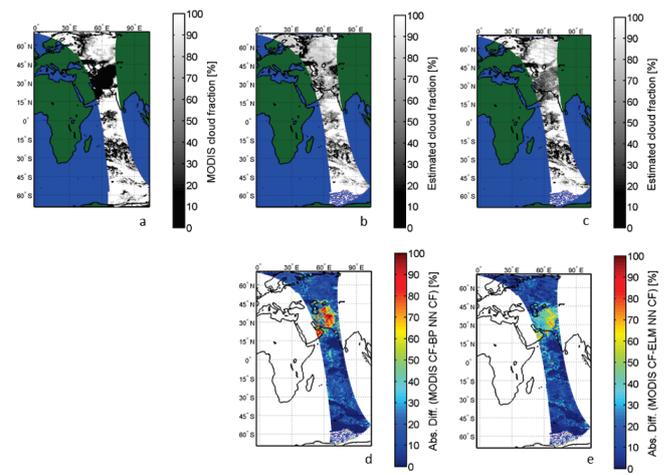


Fig. 7. Cloud fractions estimated by the BP and ELM-trained NN and comparison with MODIS CF data for the validation orbit 2006m0912t0828. **(a)** Computed MODIS geometrical cloud fraction re-located onto the OMI grid. **(b)** BP predicted cloud fraction. **(c)** ELM predicted cloud fraction. The grey-code in **(a–c)** ranges from 0 (cloud free) to 100 (totally cloud covered). **(d)** Absolute difference between MODIS geometrical cloud fraction and BP-predicted cloud fraction. The color-code ranges from 0 (perfect match) to 100 (complete mismatch). **(e)** Absolute difference between MODIS geometrical cloud fraction and ELM-predicted cloud fraction. The color-code in panels **(d)** and **(e)** range from 0 (perfect match) to 100 (complete mismatch)

over land where the ground reflectance from bright surfaces, such as deserts, misleads the NNs to interpret the high measured reflectance as if it was reflected by clouds. The spectral features alone can discriminate cloudy from clear pixels with a reasonable accuracy when proper optimization of the NNs has been performed.

Neural networks are attractive for cloud screening because of their capability of high computational speed for large data sets. Moreover, they rely on auxiliary data only during the training and they are independent from the instrument platform which makes the approach portable to other combinations of instruments such as the combination of the TROPospheric Monitoring Instrument (TROPOMI) (Veefkind et al., 2012) and the Visible Infrared Imaging Radiometer Suite (VIIRS).

Acknowledgements. All RGB MODIS granules are publicly available and are provided by the NASA Level1 and Atmosphere Archive and Distribution System (LAADS) website: <http://ladsweb.nascom.nasa.gov/data/search.html>. The authors would like to acknowledge support from and discussions with the KNMI-OMI team. Funding for this work was provided by the PP-TROPOMI project and Tekes (the Finnish funding Agency for Technology and Innovation and Ministry of Transport and Communications).

Edited by: A. Kokhanovsky

References

- Acarreta, J. R. and de Haan, J. F.: Cloud pressure algorithm based on the O₂-O₂ absorption, OMI Algorithm Theoretical Basis Document (ATBD), vol. III, Clouds, Aerosols, and Surface UV Irradiance, edited by: P. Stammes, R. Neth. Meteorol. Inst., De Bilt, 17–29, 2002.
- Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E.: Discriminating clear sky from clouds with MODIS, *J. Geophys. Res.*, 103, 32141–32157, 1998.
- Aitkenhead, M. J. and Aalders, I. H.: Classification of Landsat Thematic Mapper imagery for land covering using neural networks, *Int. J. Remote Sens.*, 29, 2075–2084, 2008.
- Bishop, C. M.: *Neural networks for pattern recognition*, Clarendon press, Oxford, 1995.
- Christodoulou, C. I., Michaelides, S. C., and Pattichis, C. S.: Multi-feature texture analysis for the classification of clouds in satellite imagery, *IEEE Trans. Geosci. Remote Sens.*, 41, 11, 2662–2668, doi:10.1109/TGRS.2003.815404, 2003.
- Del Frate, F. and Schiavon, G.: A combined natural orthogonal function/neural network technique for the radiometric estimation of atmospheric profiles, *Radio Sci.*, 33, 405–410, 1998.
- Del Frate, F., Iapaolo, M., Casadio, S., Godin-Beekman, S., and Petitdidier, M.: Neural network for the dimensionality reduction of GOME measurement vector in the estimation of ozone profiles, *J. Quant. Spectrosc. Radiat. Trans.*, 92, 275–291, 2005.
- Di Noia, A., Sellitto, P., Del Frate, F., and de Laat, J.: Global tropospheric ozone column retrievals from OMI data by means of neural networks, *Atmos. Meas. Tech. Discuss.*, 5, 7675–7727, doi:10.5194/amtd-5-7675-2012, 2012.
- Dybbroe, A., Karlsson, K.-G., and Thoss, A.: NWSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling. Part I: Algorithm description., *J. Appl. Meteor.*, 44, 39–54, 2005.
- Grivas, G. and Chaloulakou, A.: Artificial neural networks for prediction of PM10 hourly concentrations in the Greater Area of Athens, Greece, *Atmos. Environ.*, 40, 1216–1229, 2006.
- Golub, G. H. and Van Loan, C. F.: *Matrix computations*, JHU Press, 1996.
- Ham, F. and Kostanic, I.: *Principles of neurocomputing for science and engineering*, McGraw-Hill, 2001.
- Haykin, S.: *Neural Networks A Comprehensive Foundation*, Prentice Hall, 1999.
- Hornik, K., Stinchcombe, M., and White, H.: Multilayer feedforward networks are universal approximators, *Neural Netw.*, 2, 359–366, 1989.
- Huang, G. B., Zhu, Q. Y., and Siew, C. K.: Extreme learning machine: theory and applications, *Neurocomp.*, 70, 489–501, 2006.
- Hubanks, P. A.: MODIS Atmosphere QA Plan for Collection 005 and 051, Version 3.1, 2012.
- Karayannis, N. and Venetsanopoulos, A. N.: Artificial neural networks: learning algorithms, performance evaluation, and applications, *The Springer International Series in Engineering and Computer Science*, 209, 460 pp., 1993.
- King, M. D., Kaufman, Y. J., Menzel, W. P., and Tanre, D.: Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS), *IEEE Trans. Geosci. Remote Sens.*, 30, 1–27, 1992.
- King, M. D., Tsay, S. C., Platnick, S. E., Wang, M., and Liou, K. N.: Cloud retrieval algorithms for MODIS: optical thickness, effective particle radius, and thermodynamic phase, Tech. Rep. ATBD-MOD-05, NASA, Washington, DC, 1998.
- Kleipool, Q.: OMI surface reflectance climatology README, Version 2.0, data product version: 003, 2010.
- Koelemeijer, R. B. A. and Stammes, P.: Effects of clouds on the ozone column retrieval from GOME UV measurements, *J. Geophys. Res.*, 104, 8281–8294, 1999.
- Kokhanovsky, A. A., Platnick, S., and King, M. D.: Remote sensing of terrestrial clouds from space using backscattering and thermal emission techniques, The remote sensing of tropospheric composition from space, physics of earth and space environments, Springer-Verlag, Berlin, doi:10.1007/978-3-642-14791-3, 2011.
- Krasnopolsky, V. M.: *Neural network applications to solve forward and inverse problems in atmospheric and oceanic satellite remote sensing Artificial Intelligence Methods in the Environmental Sciences*, Springer-Verlag, Germany, 2008.
- Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., and Saari, H.: The Ozone Monitoring Instrument, *IEEE Trans. Geosci. and Remote Sens.*, 44, 1093–1101, 2006.
- Loyola R., D. G.: Applications of neural network methods to the processing of earth observation satellite data, *Neur. Netw.*, 19, 168–177, doi:10.1016/j.neunet.2006.01.010, 2006.
- Martins, J. V., Tanra, D., Remer, L., Kaufman, Y., Mattoo, S., and Levy, R.: MODIS Cloud screening for remote sensing of aerosols over oceans using spatial variability, *Geophys. Res. Lett.*, 29, MOD4-1–MOD4-4, doi:10.1029/2001GL013252, 2002.
- Mas, J. F. and Flores, J. J.: The application of artificial neural networks to the analysis of remotely sensed data, *Int. J. Remote Sens.*, 29, 617–663, 2008.
- Preusker, R., Hunerbein, A., Fischer, J., Brockmann, C., and Kramer, U.: MERIS Global land surface albedo maps: ATBD Cloud detection, available at: http://www.brockmann-consult.de/albedomap/pdf/atbd_cloud_detection_ama_5.pdf, 2008.
- Rumelhart, D. E., McClelland, J. L., and the PDP Research Group: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol.1 and 2, MIT Press, Cambridge, Mass., 1986.
- Sellitto, P., Bojkov, B. R., Liu, X., Chance, K., and Del Frate, F.: Tropospheric ozone column retrieval at northern mid-latitudes from the Ozone Monitoring Instrument by means of a neural network algorithm, *Atmos. Meas. Tech.*, 4, 2375–2388, doi:10.5194/amtd-4-2375-2011, 2011.
- Sellitto, P., Del Frate, F., Solimini, D., and Casadio, S.: Tropospheric ozone column retrieval from ESA-Envisat SCIAMACHY Nadir UV/VIS radiance measurements by means of a neural network algorithm, *IEEE Trans. Geosci. and Remote Sens.*, 50, 998–1011, doi:10.1109/TGRS.2011.2163198, 2012.
- Sneep, M., de Haan, J. F., Stammes, P., Wang, P., Vanbauce, C., Joiner, J., Vasilkov, A., and Levelt, P.: Three-way comparison between OMI and PARASOL cloud pressure products, *J. Geophys. Res.*, 113, D15S23, doi:10.1029/2007JD008694, 2008.
- Stammes, P. and Noordhoek, R.: OMI algorithm theoretical basis document Volume 3, Cloud, aerosols, and surface UV irradiance, ATBD-OMI-03, Version 2.0, 2002.
- Stammes, P., Sneep, M., de Haan, J. F., Veefkind, J. P., Wang, P., and Levelt, P. F.: Effective cloud fraction from Ozone Monitoring Instrument. Theoretical framework and validation, *J. Geophys.*

- Res., 113, D16S38, doi:10.1029/2007JD008820, 2008.
- van den Oord, B.: OMI small pixel data, TN-OMIE-KNMI-397, 1, 2002.
- Vasilkov, A., Joiner, J., Spurr, R., Bhartia, P. K., Levelt, P., and Stephens, G.: Evaluation of the OMI cloud pressure derived from rotational Raman scattering by comparison with other satellite data and radiative transfer simulations, *J. Geophys. Res.*, 13, D15S19, doi:10.1029/2007JD008689, 2008.
- Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruijzinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor; a GMES mission for Global Observation of the Atmospheric Composition Climate, *Air Qual. Appl. Rem. Sens. Envir.*, 120, 70–83, 2012.
- Werbos, P. J.: Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences, PhD thesis, Harvard University, 1974.
- Wu, D. L., Jiang, J. H., and Davis, C. P.: EOS MLS cloud ice measurements and cloudy-sky radiative transfer model, *IEEE Trans. Geosci. and Remote Sens.*, 44, 1156–1165, doi:10.1109/TGRS.2006.869994, 2006.
- Yan, H., Chen, L., Tao, J., Su, L., Huang, J., Han, D., and Yu, C.: Corrections for OMI SO₂ BRD retrievals influenced by row anomalies, *Atmos. Meas. Tech.*, 5, 2635–2646, doi:10.5194/amt-5-2635-2012, 2012.