AMTD-2015-219

Interactive comment on “Aerosol Optical Depth retrievals at the Izaña Atmospheric Observatory from 1941 to 2013 by using artificial neural networks” by R. D. García et al.

<u>Anonymous Referee #1 (RC, C3203-C3205, 2015):</u>

This manuscript by Garcia et al., "Aerosol optical depth retrievals at the Izaña Atmospheric Observatory from 1941 to 2013 by using artificial neural networks", presents a reconstruction of aerosol optical depth for 73 year time period. There are not very many long-term time series available for aerosol optical depth and, therefore, the topic of this study is of great interest and importance. The scope of the paper is both concise and specific, and my minor comments are mainly related to the need to clarify some of the issues. Before publication, the following points should be addressed.

<u>Authors:</u> We appreciate the positive and constructive comments of the Referee. Here we discuss and respond to his/her general and specific comments.

<u>General Comments:</u>

1. **I was missing some more information and details in the section 3. For instance, the very meaning of Eq 1. did not become clear, without any other explanations. Could you open the procedure and algorithm somewhat more. Also, in the section 3.1, it did not become clear what is the difference between 15% for validation and 15% for testing, both being independent from training, naturally. So more background about the algorithm would be welcome.**

   <u>Authors R1:</u> Following the Referee’s recommendation, the sections 3 and 3.1 (Artificial neural networks (ANNs) and Training process) have been described in greater detail. The read as follow in the final manuscript (the text included has been highlighted in bold):

   <u>Section 3. Artificial neural networks (ANNs)</u>

   *ANNs are statistical data modeling tools, inspired by the human brain, capable of simulating highly non-linear and complex relationships between inputs and outputs by a learning process, **the so-called training process**. This tool mainly consists of three layers of neurons: the input layer groups the input data in the input vector p and connects them with the hidden layer. **In this layer the input***

vector is transformed into a net input vector, a', by using adaptive weights, W^h, biases, b^h, and a transfer function, TF^h, such as a' = TF^h(n), where n=(W^hp+ b^h). Then, the hidden layer is connected with the output layer, in which the outputs obtained in the previous step, a', are transformed into the net input for the output layer, n' = (W^{out}a'+ b^{out}). Finally, the output transfer function, TF^{out}, is applied to n' to obtain the final output of the ANN, a (Jain et al., 1996, and references therein). The weights and biases used both in the hidden (W^h and b^h) and in the output layer (W^{out} and b^{out}) were previously computed in the training process.

In this work, the ANNs have been implemented by using the Matlab Neural Network Toolbox (Demuth and Beale, 1993) with the architecture shown in Fig. 3: ***the input parameters of the input layer are different meteorological observations taken at IZO (Sect. 3.2 details the selection of these inputs), and the hidden layer is made up of 30 neurons*** *with a transfer function defined by the hyperbolic tangent function of* $\boldsymbol{n}$ *(Eq. 1).*

$$\varphi = \tanh(n) = \frac{e^{2n} - 1}{e^{2n} + 1} \qquad (1)$$

where $\boldsymbol{n}$ *is the corresponding net input.* ***The hyperbolic tangent is one of the most used transfer function in ANNs, since it successfully combines a fast learning rate with reliable results (Zhang et al., 1998; Özkan and Erbek, 2013). Finally, the output layer has one neuron with the linear transfer function, which is often used in forecasting and approximation tasks (Zhang et al., 1998).***

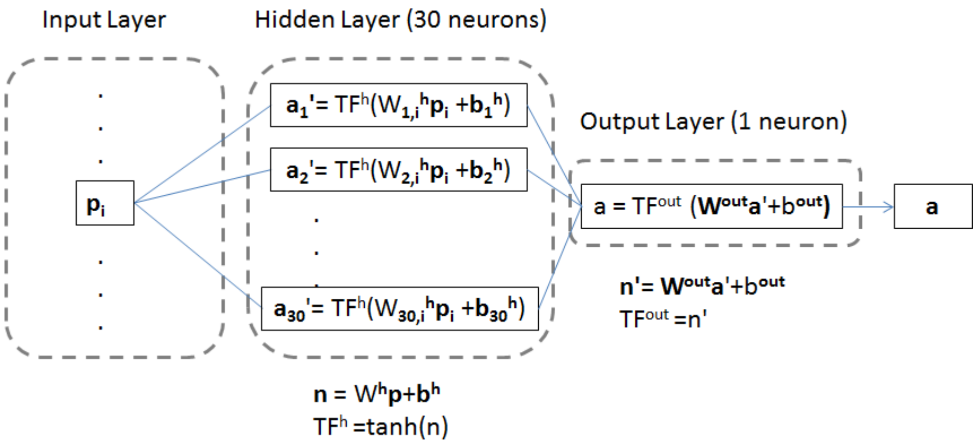

***Figure 3**. Schematic representation of the artificial neural network used in this study, where p_i = p_i(N_{di}, VIS_i, FCS_i, RH_i, $Temp_i$) with i=1, …, N, being N the total number of observations and a = ANN AOD.*

Section 3.1. Training process

The learning or training procedure plays a key role in the ANNs design and setting. ***In this process a set of inputs with known outputs (targets) are used***

to calculate the weights (W^h, W^{out}) and biases (b^h, b^{out}) to be applied in the neural network, as explained in the previous section.

The first step on this process is to randomly divide the set of known inputs and target values in three different subsets: training (70% of the data), validation (15 % of the data) and test (15 % of the data). The weights (W^h, W^{out}) and biases (b^h, b^{out}) are computed for each neuron. Then the validation subset is used to estimate the error by comparing the obtained outputs with the targets of the validation subset. The computation of weights and biases and the subsequent error estimation is iteratively repeated until the error is lower than a required value or if the assignation of new weights and biases does not decrease the error. In this work the estimation of the error is supervised by the Levenberg-Marquardt optimization algorithm, which has proved to be efficient and fast for small and medium sized networks, as the architecture used here (Foresee and Hagan, 1997; Hao and Wilamowski, 2011). The mentioned error is computed by the mean square error (MSE) defined by the following equation:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \qquad (2)$$

where N is the dimension of the validation subset, t_i the targets in the validation subset and a_i the ANN outputs obtained from the validation subset inputs. Finally, the test subset, not used in the training process, is used to check the quality of the obtained ANN by applying it to "clean" inputs, that is, inputs and targets not used in the training process (Beale et al., 2014).

Given that the division of known data in training, validation and test subsets is random, we have repeated the training process 1000 times. *Then, the best ANN is selected as the one showing the highest Pearson correlation coefficient (R), slope closer to one and lowest intercept with respect to the known outputs* ***(Hao and Wilamowski***,***2011)****.*

The AOD measurements used to train the ANNs were performed with one of the most accurate and stable instruments dedicated for atmospheric aerosol monitoring, a Precision Filter Radiometer (hereafter PFR AOD), developed at the World Radiation Center Physikalish Meteorologisches Observatorium (www.pmodwrc.ch). It was installed at IZO in the framework of a high precision world network for AOD characterization and watching (GAW/PFR) in June 2001, but continuous observations are only available since 2003. The PFR measures direct solar radiation, with a field of view of 2.5°, at 862, 500, 412 and 368 nm. Then, the AOD is estimated at all these wavelengths with an expected uncertainty of ±0.01 (Wehrli, 2000). In this study, we have used Level 3.0 of Version 3.0 AOD at 500 nm.

2. ***The idea to use FCS should have been discussed in more depth as well. For what effect, related to AOD, FCS is accounting for? Is it cloud contamination in***

AOD, as I was thinking? However, in that case, would it have been perhaps more justified approach to use FCS to exclude cases of presence of clouds and train the algorithm for those cases? Would there be some illustrative cases to demonstrate the role of FCS during the training period? At least, please provide more discussion about the role of FCS.

***Authors_R2:** The introduction of FCS in the ANN training allows the ANN to discriminate the patterns associated with possible residual cloud-cover for the days with oktas=0. Please note that the cloud-free days for the study were selected by considering a median number of oktas egual to cero, but this values was calculated from only three observations per day. Thereby, some episodes with cloud-contamination are likely and, as pointed out by the referee, this residual cloud-contamination could give artificial ANN AOD values. Figure 1 clearly illustrates the added value of including the FCS as an input parameter in the ANN training. The agreement between ANN AOD estimates and AERONET AOD values significantly improves: higher Pearson coefficient and scale factor close to one. Thus, the combination of both parameters, oktas and FCS, will assure the absence of clouds during most of the day and explains almost the 95% (R=0.97) of the observed variability in the AOD measurements.*

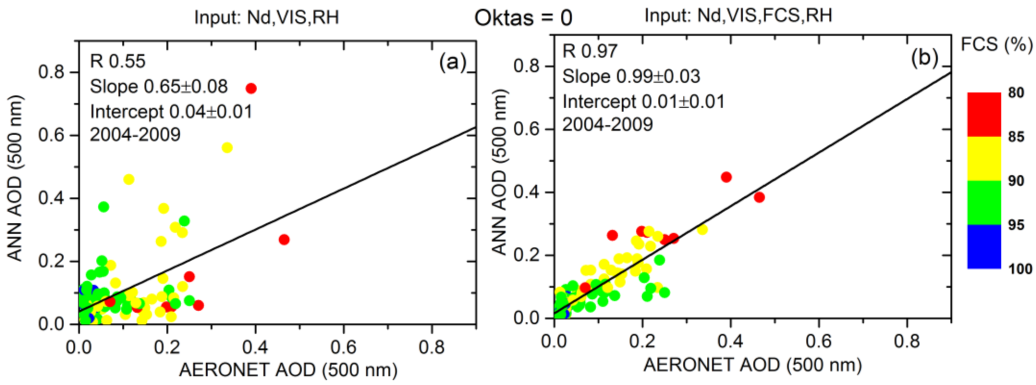

Figure 1.- Scatterplot of ANN AOD estimates using different input parameters (a) Nd, VIS and RH (b) Nd, VIS, FCS and RH vs. daily AERONET AOD at 500 nm for the periods 2004-2009 . The black solid lines indicate the least-square fits. The least-square fit parameters are shown in legend (Pearson correlation coefficient, R, slope and intercept). The colour scale indicates the fraction of clear sky values (FCS, %).

When the FCS is uniquely used to discriminate cloud-free days but not included as an input parameter, the ANN AOD estimates are not able to properly capture the real AOD variability, as observed in Figure 2. The ANN AOD values shown in this figure were obtained considering Nd, VIS and RH as input parameters without taking into account the oktas and for days with FCS≥75%.

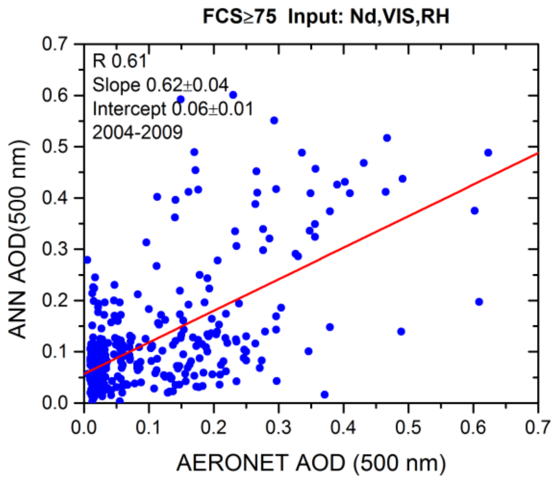

Figure 2.- Scatterplot of ANN AOD estimates using Nd, VIS and RH as inputs parameters vs. daily AERONET AOD at 500 nm for the period 2004-2009 considering cloud-free days with FCS≥75%. The red solid line is the least-square fit. The least-square fit parameters are shown in the legend (Pearson correlation coefficient, R, slope and intercept).

Following the Referee's recommendation, this discussion has been introduced in the final manuscript.

Specific Comments:

1. **Page 9077, lines 5-10, here I was thinking that perhaps Lachat and Wehrli 2013 could be cited here, since they analyzed a very nice and long time series for dimming and brightening trends.**

   Authors_R1: ***The reference has been included in the Introduction.***

2. **Page 9077, line 11-12, I can understand the meaning, however somehow the sentence is not complete. Should it continue, e.g. "... have significant role."**

   Authors_R2: ***This sentence has been modified as follows:***

   > ***"The causes of these phenomena are not fully understood currently, but it has been pointed out that changes in the transmissivity of the Earth's atmosphere play a significant role."***

3. **Page 9083, line 2, "fraction clear sky", should it be "fraction of clear sky"?**

   Authors_R3: ***Done.***

4. **Page 9083, line 3, should the ratio be other way around? Ratio between measured and SDmax?**

***Authors_R4:* Yes, the referee is right. The sentence has been modified in the final manuscript.**

> ***"…fraction of clear sky (FCS) defined as the ratio between SD performed with Campbell Stokes sunshine recorder (García et al., 2014a) and the maximum daily sunshine duration SD_{max}...."***

5. **Page 9083, line 10, "range from 1916 and 1921", the latter number is wrong? Otherwise the meteorological data are only for a very limited period.**

***Authors_R5:* This sentence has been clarified as follows:**

> ***"…The time series at IZO are from 1916 up to now for T and RH, from 1921 to present for FCS, and from 1941 to 2009 for VIS…"***

6. **Last paragraph of the section 4 remained somewhat unclear. Could you please provide some more details about the analysis to detect change points and so on.**

***Authors_R6:* We have included more information. This paragraph reads as follows in the revised manuscript:**

> *"The long-term Mark-I AOD time series also allows us to analyse the temporal consistency of the ANN AOD estimations by examining possible drifts and discontinuities in the monthly time series of the differences between ANN AOD and Mark-I AOD for July, August and September. A drift is defined as the linear trend of monthly median bias (measurements-estimations), while the change-points (changes in the median of the bias time series) are analyzed by using a robust rank order change-point test (Lanzante, 1996).* ***The Lanzante's procedure is an iterative method that applies a (single) change-point test, based on summing the ranks of the values from the beginning to each point in the series, and followed by an adjustment step (the median computed for the segments enclosed by the identified change points is used to adjust the series). In the subsequent iteration the change-point test is applied to the adjusted series and the iterative process finishes when the significance of each new change-point is less than an a priori specified level.***
>
> ***By applying this change-point test we identify 1997 as the change-point in the monthly median bias time series (see Fig. 5c)****, caused by the horizontal visibility records. Although this discontinuity is significant at 99 % confidence level, the difference of median bias is rather small (-0.013±0.001 for 1984–1997 period and +0.006±0.003 for 1998–2009 period) and within the ANN AOD and Mark-I AOD expected uncertainties. Furthermore, we observe that there are no significant drifts in the bias time series either before or after this systematic change point at 99 % of confidence level. For the rest of months, August and September, the monthly median bias time series have shown neither significant systematic change points nor temporal drifts. These findings indicate that the*

ANN AOD estimates are consistent over time and, thus, valid to reconstruct the AOD time series at IZO."

7. Table 1: HR should be likely RH?

Authors_R7: *Thank you, it was a typo corrected*

References

Lachat, D., and Wehrli, C. : Dimming and brightening trends in direct solar irradiance from 1909 to 2010 over Davos, Switzerland: Proportions of aerosol and gaseous transmission, J. Geophys. Res. Atmos., 118, 3285–3291, doi:10.1002/jgrd.50344, 2013

Beale, M. H., Hagan, M.T., and Demuth, H.B.: Neural Network Toolbox, User's Guide, The MathWorks, Inc., 2014

Özkan, C., andErbek, F. S.:The Comparison of Activation Functions forMultispectral Landsat TM Image Classification, Photogrammetric Engineering & Remote Sensing,Vol. 69(11),pp. 1225–1234, 2013

Zhang, G., Patuwo, B. E., and Hu, M. Y.: Forecasting with artificial neural networks: The state of the art. International journal of forecasting, 14(1), 35-62, 1998.