Author's response to comments of Referee #2

AC: We thank the anonymous referee for the review of our manuscript.

RC2: I read an interesting article which aims to investigate the effect of solar zenith angle, ozone, cloud cover and surface albedo on spectral UV radiation at Marambio Base, Antarctic Peninsula. UV irradiance measurements come from a double Brewer spectrophotometer for the period 2010-2020. The effects of the different parameters on surface UV irradiance are studied using a neural network model that has been developed for this purpose. My recommendation for this article is to accept for publication after clarifying better what Figure 11 aims to show, as I had also recommended during my quick review, and after clarifying the neural networking explained in section 2.4.

AC: Thank you for your comments and suggestions. Concerning artificial neural networks (ANN), we would like to mention that dozens of cross-validated tests of randomly initialized ANN with random division of the whole dataset to training (70% of data), validation (15%) and testing (15%) subsets were carried out. The aim of these tests was to set the appropriate complexity of ANN, corresponding to the complexity of underlying relation between chosen predictors (SZA, TOC, cloudiness, albedo climatology) and predictands (spectral UV intensities). These tests revealed that about 22 neurons in the hidden layer are (quasi)optimal to avoid both over- and under-parametrisation of the relationship. The advantage of this approach is that it derives the appropriate complexity of ANN directly from real (measured) data and is not limited by an a priori assumption about the shape of the regression function. In addition, ANN are capable of modelling relationships that are difficult to describe analytically and from this point of view, they are more general than the analytical approach. Then, an ensemble of 10 neural networks, each with 22 neurons in the hidden layer were trained, again with random division of the dataset to training (70%), validation (15%) and testing (15%) subsets and with random initialization of the networks. The overtraining of the networks was tackled by adding a stopping condition: the error improvement lower than 0.0000001 in the window of 200 cycles, while the error function was defined as the sum of squares.

Based on the above-mentioned information, we will expand the explanation and interpretation of particular plots/figures in the revised version of the manuscript. Moreover, we are providing the new plots stating how we built the models and showing the differences between them (see Appendix A, which we will also upload in the manuscript).

RC2: Specific comments:
Lines 139-161: I find difficult to understand how the ten models were built and how the best neural network model was selected each time. Given that most studies use multiple regression modelling to quantify the contribution of each atmospheric parameter, lines 140-142 trigger the question how much different would the calculations from a multiple regression model be? A supplement with explanations on the neural model procedures, and comparison with estimations from a multiple regression model would help.

AC: Please note that ANN derive the relationship between predictors and predictands directly from the input data, so they are not limited by any a priori assumptions about the shape of the dependence. An example of the potential assumption based on Lambert-Bourger-Beer Law can be found e.g., in Antón et al. (2005). Moreover, with sufficient complexity of the network (the number of hidden layers and the number of neurons in them), they are able to simulate practically any dependency, even that one that is difficult to describe analytically. In this sense, they are more general than classical regression approach. A direct comparison of the classic regression approach, based on Beer's law, and the neural network is in this particular case not possible, as an independent variable (the intensity of

UV radiation at the top of the atmosphere) enters the regression relationship described in Antón et al. (2005), but it is not an input parameter of our ANN. Thus, both models are based on different sets of input data. Moreover, the neural model simulates the UV spectrum as a whole, while the regression approach, using Beer's law, calculates the intensities independently and for each wavelength separately. The neural network can therefore better simulate the interdependencies between UV radiation intensities at different wavelengths, which is especially important in a situation where the data contain inaccuracies or noisy components. This is where the two methods differ considerably, and a specific study would have to be designed to allow a proper comparison.

We agree, however, that the information we provided on the ANNs in the manuscript was limited. Therefore, based on the above-mentioned information and Appendix A, we will stress ANN modelling more in the revised version of the manuscript.

Ref.: Antón, M., Cancillo, M. L., Serrano, A., and García, J. A..: A Multiple Regression Analysis Between UV Radiation Measurements at Badajoz and Ozone, Reflectivity, and Aerosols Estimated by TOMS, Phys. Scripta, 118, 21–23, 2005.

RC2: Low albedo case (fig. 11d): the Obs. Alb. value is indeed low (0.37), but the Mod. Alb. value is 0.81, which is not low. Please check.

AC: Thank you for the comment. This was indeed an error, it should have been 0.37. However, based on the comments from Referee 1, the albedo panel as well as related calculation will be removed from the manuscript.

Appendix A

Artificial Neural Network model development and validation

Out of the ten ANN models we built, nine (ANN02 to ANN10) behaved in a similar way, while one (ANN01) was different. The differences between the models did not result from the ANN setting, which remained the same, but occurred due to the random initialization of the models and the random split of the dataset to training (70 %), testing (15 %), and validation (15 %) subsets. As seen from Fig. A1, the model ANN01 had the most data within ± 5, respective ± 10 % from observations, and it had the largest R-squared and lowest RMSE out of all ten models. However, the model was biased toward underestimation of UV irradiance throughout most of the spectrum.

For the purpose of the study, it was best to choose a model with the best precision, i.e. the lowest variability of results, highest R-squared and lowest RMSE (model ANN01). Also, it was possible to tackle the bias present within the model using a simple median correction described in the manuscript in section 2.4.

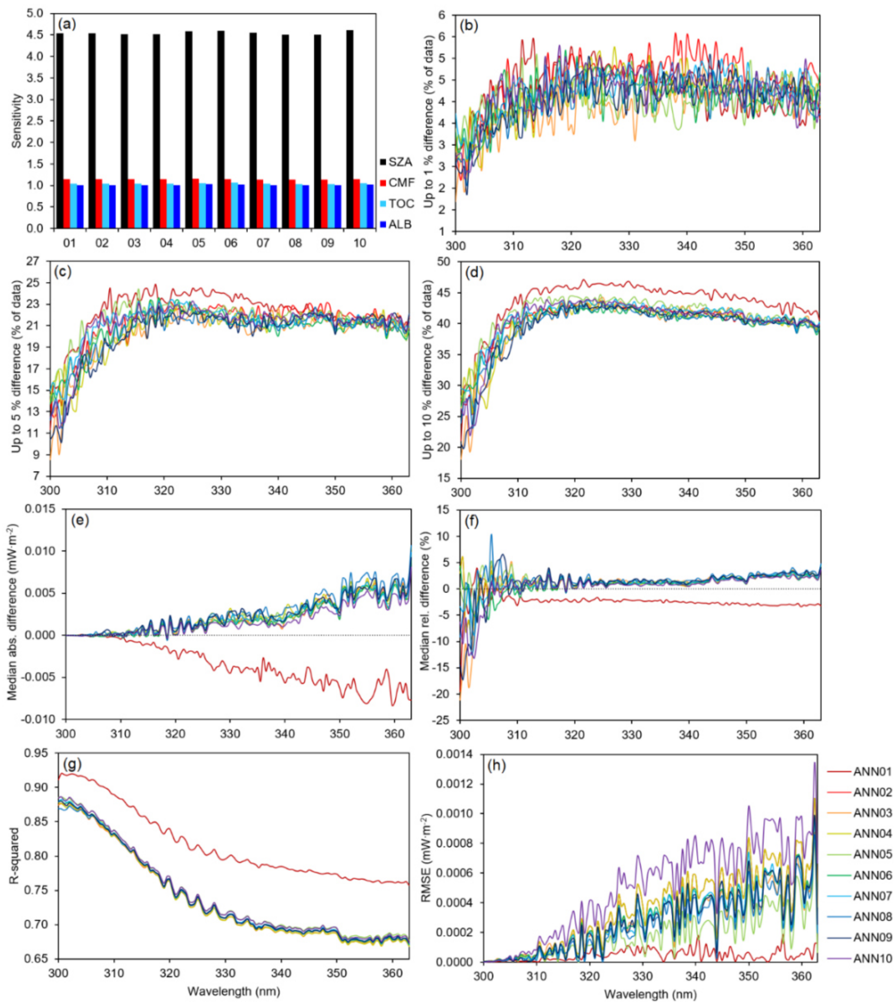

Figure A1. Development and validation of the 10 artificial neural network models, while (a) is the sensitivity of individual models to given variables; (b), (c) and (d) shows the amount of modelled data within 1, 5, and 10 % difference from observations, respectively; (e) and (f) are the absolute and relative median differences of the modelled data from the observations; (g) shows the R-squared, and (h) is the root mean square error of the individual models.