Report from the COST-726 Short-Term Scientific Missions (STSM)

Janusz W. Krzyścin (Institute of Geophysics, Polish Academy of Sciences)

Evaluation of the performance of the UV reconstruction models examined in the modeling exercise WG1 and WG2 using Taylor Diagrams tools

Hosted by Peter Koepke, Meteorological Institute, University of Munich, Germany
31 May 2006 -9 June 2006

1. Purpose of the visit.

Purpose of this STSM is a ranking of the UV reconstruction models (proposed by participants of WG2 –COST 726 with input provided by WG1) to mimic observed UV daily doses. The new statistical approach, Taylor Diagram, will be applied to evaluate the models' outcome.

The main goal of COST-726 is a reconstruction of the UV daily doses over Europe back to early 1950th. To get knowledge about performance of the reconstruction models presently being used by the participants of this action a modelling exercise has been made in spring 2006. This included 4 stations and 2 years modelled with 16 algorithms. Methodology of the Taylor Diagram will be used to quantify and visualize the overall correspondence between the modelled and observed daily sums of the erythemally weighted irradiances. The diagram provides the degree of pattern correspondence allowing one to evaluate how accurately a model simulates the natural systems. Establishing a statistical significance of the differences between various models outcome shown in Taylor diagram is the most challenging part of the STSM mission.

2. Description of the work carried out during mission

2.1 Preparation of the data

The analyzed time series of UV daily doses have a strong annual course with the maximum in late spring/early summer and minimum in winter. Thus, any model simulating such behavior will yield a high correlations coefficient and close RMS value to observed one. Thus, for better distinguishing between models' performances the annual pattern should be removed from the analyzed time series. The smoothed annual course is extracted from measured doses for each year and station using the locally weighted scatter (LOWES) smoothing techniques. Fig.1 gives examples of the measured UV daily doses and the smoothed annual patterns.

Next the deviations from the smoothed curves are calculated both for the modeled and measured time series. Further these deviations are called the absolute deviations. They are still have a seasonal course as larger deviations (in absolute units) are possible in seasons with normally high UV doses. We also examine another category of the data relative deviations that are obtained by the normalization of the absolute deviations using the smoothed annual values as the norm. The seasonality is not completely removed from the time series of the relative variations as the model's accuracy and quality of the measurements are usually poorer in winter time.

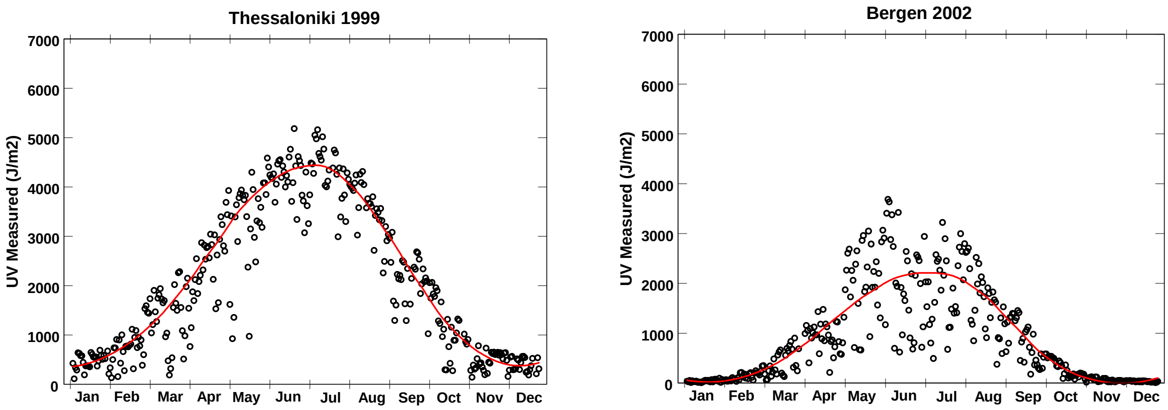

Fig. 1. The measured UV daily doses and their smoothed annual profile

The following models have been examined:

– ***auth*** Aristoteles University, Thessaloniki, Greece
– ***dwdk_day*** German Meteorological Service, Department "Climate and Environment", Freiburg, Germany. Model using global irradiance on a daily basis.
– ***dwdk_acc*** German Meteorological Service, Department "Climate and Environment", Freibur. Model using solar global radiation on a hourly basis.
– ***dwdf*** German Meteorological Service, Department "Forschung und Entwicklung", Lindenberg, Germany
– ***fmi*** Finnish Meteorological Institute, Meteorological Research Division, Helsinki, Finland.
– ***gsas*** Geophysical Institute, Slovak Academy of Sciences, Bratislava, Slovakia.
– ***igfp*** Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland.
– ***imwm*** Institute of Meteorology and Water Management, Warsaw, Poland
– ***jrc*** European Commission - Joint Research Centre, Institute for Health and Consumer Protection, Ispra, Italy.
- ***mim_cn4*** Meteorological Institute, Munich, Ludwig-Maximilians-University, Munich, Germany. Cloud neural network 4, directly taking solar irradiance.
– ***rivm*** National Institute for Public Health and the Environment, Bilthoven, The Netherlands.
– ***tobs*** Tartu Observatory, Toravere, Estonia.
– ***boku*** University of Natural Resources and Applied Life Sciences, Department „Water - Atmosphere-Environment", Vienna, Austria.
– ***mim_cn1*** Meteorological Institute, Munich, Ludwig-Maximilians-University, Munich, Germany. Cloud neural network 1, availing total cloudiness.
– ***mim_wgt*** Meteorological Institute, Munich, Ludwig-Maximilians-University, Munich, Germany; availing total cloudiness.
– ***uvwm*** University of Veterinary Medicine, Institute of Medical Physics and Biostatistics.

Last four models do include global radiation data in model's input.

2.2 Taylor Diagram

Recently a new statistical tool, Taylor diagram, has been proposed for both description and visualization of a correspondence between various simulations of a measured variable (Taylor, 2001). According to the methodology of Taylor diagram a model performance relative to measurements is visualized by a point on the polar plot. The azimuth angle φ pertaining to this point is such that $cos(\varphi)$= correlation coefficient between modeled and measured data. A radius from the origin is given as the ratio of RMS of the model values to RMS of the observed data. An ideal model (being in a full agreement with measurements) is marked by the point with coordinates φ=0 and radius=1. It means the correlation coefficient equal to 1 and the same amplitude of modeled and measured variations. Thus, in case if we have many models to be compared the best model is chosen as the model having minimum distance between its point on the Taylor diagram and the ideal model point - (0, 1). Fig.2 shows an example of the Taylor diagram illustrating the performance of all models for one selected station in 1999 and 2002.

2. 3. Results

The model performance (for each station and year) taking into account the absolute and relative deviations is visualized in Figures. Here we show example for Thessaloniki 1999 and 2002 (Fig.2). The similar pattern is seen in all plots. A group of points gathers closely to the ideal model point (0,1) and some points appear long away from this point. For selected the data category the configuration of the points remains practically unchanged in all figures. Usually the models not using global radiation as a proxy for the cloud attenuation effects stay away other model points.

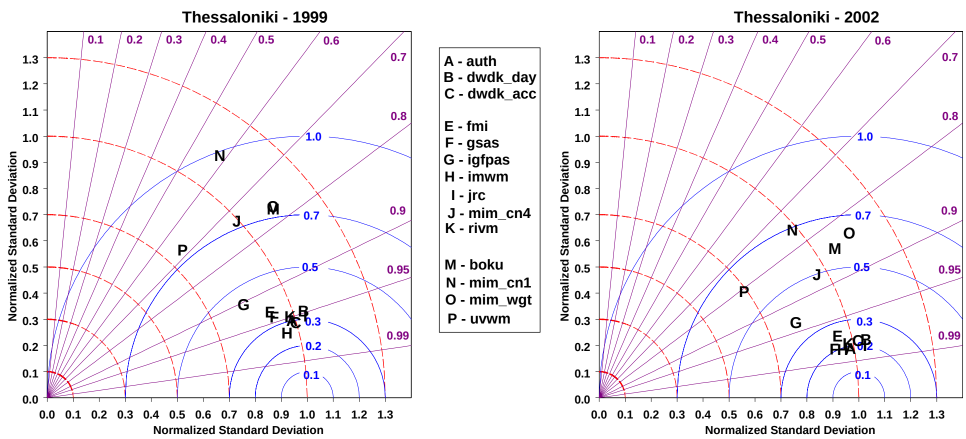

Fig.2. Taylor Diagrams for the absolute deviations; daily doses minus the smoothed annual profile calculated from of the measured doses.

Challenging problem in interpretation of the Taylor diagram is evaluation how significant are the differences between locations of various points in the diagram. During STSM mission we elaborate the method of estimation of the range of the model point distance variability to the ideal model point using the resampling methodology. The distribution of the distance is obtained from the moving-block bootstrap technique (Efron and Tibshirani, 1993). We think that it is the most important outcome of this STMS.

2.4. Resampling model

The bootstrap belongs to the category of nonparametric statistical methods. It is able to simulate the probability distribution of any statistics without making any assumptions related to the temporal or spatial covariance structure of the variables. One simply resamples, with replacement, from the original record. However, a construction of hypothetical time series must preserve the temporal structure of the original time series. The time series of absolute and relative daily deviations used here can be approximated as a simple autoregressive process or an order up 2 with small serial correlations. Thus, sequences of 5-day data blocks will be approximately independent. Resampling of blocks of data is known as the moving-blocks bootstrap first introduced by Kunsch [1989].

In our case we have a large seasonality in the data (spring/summer and winter maxima in the absolute and relative deviations data, respectively), so we assume that possible blocks for replacement are within ±1 month relative to the removed original block. It is rather an arbitrary assumption but gives ~ 10^{60} possible representatives of the original time series for each year. Both the original modeled and measured time series are bootstrapped using the same sequences of the blocks. We analyze the sample of 1000 pairs of the annual time series. For each model-measurement pair we calculate the normalized standard deviation and the correlation coefficient and finally the distance to the (0,1) point on the Taylor diagram. Sensitivity studies shown that much larger samples (10,000 and 100,000) provide similar results. The sample of the model-observation distances is sorted in ascending order and point No. 25 and No.975 define the 95% confidence range for the distance calculated from the original data.

2.5 Ranking of the models

The results are shown in Tables (joint paper is in preparation) for each station and year both for the absolute deviations and the relative deviations. To gain additional insight into the model performance we draw Taylor diagram (Fig.3) and calculate the 95% confidence ranges (Tab. 1) for each model that combines all data (together all stations and years without

Postdam 1999). The number of daily doses contributing to this time series is about 2000 for many models.

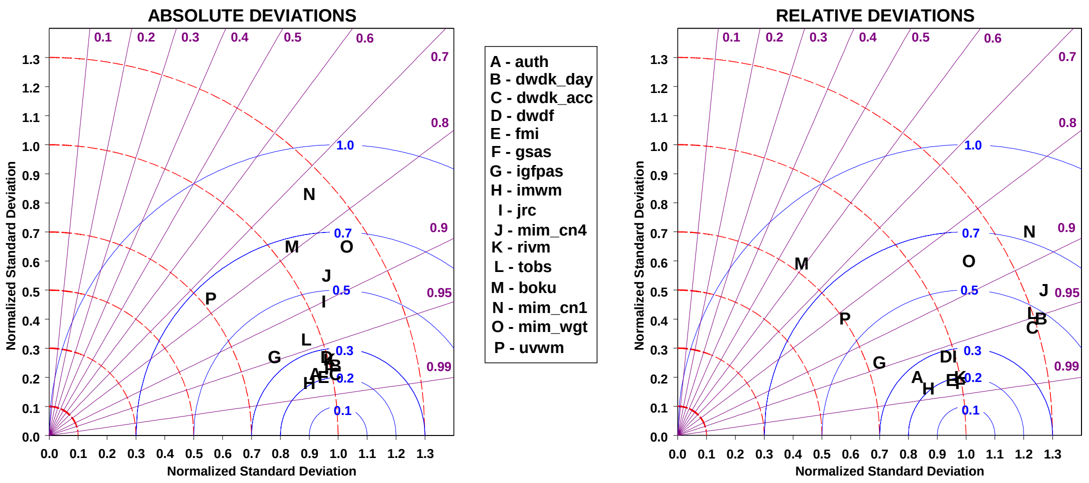

Figure 3. Taylor diagrams from all available model-measurement pairs of the deviations (together all years and stations); the absolute deviations- left, the relative deviations -right.

Model	Abs_Dev	Rel_Dev
auth	0.24(0.23,0.26)	0.27(0.25,0.29)
dwdk_day	0.26(0.24,0.28)	0.49(0.43,0.56)
dwdk_acc	0.23(0.21,0.25)	0.45(0.39,0.52)
dwdf	0.29(0.27,0.31)	0.30(0.26,0.36)
fmi	0.22(0.20,0.23)	0.21(0.18,0.25)
gsas	0.25(0.23,0.27)	0.20(0.17,0.24)
igfp	0.37(0.35,0.39)	0.40(0.39,0.43)
imwm	0.22(0.21,0.24)	0.22(0.20,0.23)
jrc	0.47(0.45,0.51)	0.29(0.25,0.34)
mim_cn4	0.56(0.53,0.59)	0.58(0.51,0.66)
rivm	0.27(0.26,0.29)	0.22(0.19,0.26)
tobs	0.37(0.31,0.42)	0.49(0.33,0.68)
boku	0.68(0.65,0.72)	0.83(0.80,0.86)
mim_cn1	0.85(0.80,0.91)	0.75(0.69,0.83)
mim_wgt	0.66(0.62,0.71)	0.61(0.51,0.72)
uvwm	0.65(0.62,0.69)	0.59(0.53,0.66)

Table 1. Taylor model-measurement distance for the selected reconstruction model derived from all available model-observation daily pairs representing: the deviations from the smoothed annual profile (derived from the measured daily doses) - **Abs_Dev**, the deviations from the smoothed annual profile expressed in percent of the smoothed values – **Rel_Dev**. 95% confidence limit is shown in the parentheses.

Results both from the combined all stations and years data and from individual station and year data lead as to the following ranking. Model ***auth***, ***fmi, gsas***, ***imgw***, and ***rivm*** form the group with the best correspondence to the measurements. It looks like that ***dwdf*** is also a candidate to this group but it requires further testing because the Tessaloniki data are not analyzed. Model ***dwdk_day, dwddk_acc,*** and ***jrc*** form another group. Their performance is similar to the first group but only for one type of data, i.e., for the absolute deviations in case of ***dwdk_day, dwddk_acc,*** and for the relative deviations for ***jrc***. The third group consists of ***igfpas*** and ***mim_cn4.*** Performance of these models is better (at least for one type of the data) than models not using global radiation but evidently not as good as models from the first group. The fourth group consists of 4 models that do not use the global radiation data: ***boku, mim_cn1***, ***mim_wgt***, and ***uvwm. tobs*** was run only for the Bergen data, so it cannot be classified at the moment.

For application the absolute doses, that are relevant for human health, are of main interest. Thus the main interest is for the results of the modelling exercise for the absolute differences. Here, the models ***auth***, ***dwdk-day***, ***dwdk-acc***, ***fmi***, ***gsas***, ***imwm***, and ***rivm*** form the group with the best correspondence to the measurements. ***dwdf*** needs additional testing with the Thessaloniki data to be included to this group. The other models show significant larger deviations. But again can be seen that the models ***igfp***, ***jrc*** and ***mim_cn4*** form another group, with better results than the models which do not take into account the global irradiance.

It should be noted that the classification of the performance of the models is based on their specific statistical properties, i.e., a correspondence to the measured UV daily doses taking into account departures from the mean annual profile derived from the measurements. If more interest is put to analyses of absolute values (daily doses) all models probably behave very similarly because of strong seasonality in the UV values. Another problem is how the model requirements limit the model use for a reconstruction of the UV doses in the past when probably not all requested input values are possible.

References:

Efron B., and R.J. Tibshirani, An Introduction to the Bootstrap, Chapman and Hall, New York, 1993.

Kunsch, H.R., The jacknife and the bootstrap for general stationary observations, Ann., Stat., 17(3),1217-1241,1989.

Taylor, K.E., Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, D7, 7183-7192, 2001.

3. Description of the main result s obtained

Establishing a statistical significance of the differences between the models outcome shown in Taylor diagram is still open problem waiting for a general solution. During STMS the original resampling methodology has been developed and successfully used to find how significant differences are between the models. The performance of the models and their ranking have been presented in appropriate graphical format, useful for a publication (in preparation).

4. Future collaboration with host institution

(see next section)

5. Project publications/articles resulting or to result from the STSM

Future collaboration with host institution is expected in area of preparation of final version of paper "Modelling solar UV radiation in the past: Comparison of algorithms and input data" by Koepke et al. to be published by the end of 2006 by the COST publishing resources.

6. Confirmation by the host institute of the successful execution of the mission

(see attached document)

7. Other comments

The participant of this STMS would like to thank Zenobia Lityńska, Bożena Łapeta, and Peter Koepke for all their efforts to organize STMS in Munich.

Munich, June 9, 2006.

doc. dr. hab. Janusz W. Krzyścin