Interactive comment on “First national intercomparison of solar ultraviolet radiometers in Italy” by H. Diémoz et al.

H. Diémoz et al.

h.diemoz@arpa.vda.it

Received and published: 26 July 2011

Answers to referee #1

The authors are very grateful to the referee for his comments and his constructive criticism.

Comment #1: The manuscript describes a measurement campaign of UV radiometers organised in Italy in June 2010. As stated by the authors, this is the first of a planned series of intercomparisons organised to assess the quality of the UV measurements performed in Italy. The participating institutions and the locations of their instruments come from different areas in Italy, and thus can serve as a preliminary core group for the eventual establishment of an Italian UV network. In that sense, the presented activity is an important first step towards achieving this objective. While the activity is of high importance for the participants to this campaign, and on a larger scale to the Italian Institutions planning to become active in the monitoring of solar UV radiation, it is not obvious that the general readership of AMT is targeted by this manuscript. Indeed, no substantial conclusions are drawn from the intercomparison but only results from individual instruments are shown without a critical appraisal of the observed deviations to the reference. Obviously, this document presents a very-well written report and serves as an excellent internal report to the participants but the outreach to external readers is very limited. I think it is therefore essential that the authors considerably expand the manuscript to demonstrate the scientific significance of this work to the AMT readership.

Answer #1: The manuscript has been updated in order to emphasize the outcomes that may be of interest to the general AMT readership. Moreover, the conclusions have been substantially expanded (a new section, 5.5, was added to the manuscript) accordingly to the suggestion of referee #1 and the observed deviations to the reference were discussed with a more critical appraisal. Additional or modified statements to the revised manuscript are reported in italic font in the order they appear in the text:

Abstract: "... on the base of their own procedures and calibration data. A radiative transfer model was successfully applied as an interpretative tool. The input parameters and output results are described in detail. The comparison was performed in terms of ...“; “An improved algorithm for comparing broadband data and spectra has been developed and is discussed in detail”; “Remarkable deviations were found for
the instruments calibrated in the manufacturers’ facilities and never involved in field intercomparison. Finally, some recommendations to the UV operators based on the campaign results are proposed”.

**Introduction:** “… health and the environment (UNEP, 2010). Well calibrated ground-based networks on a continental and national level are also essential in order to monitor the effective behaviour of the solar UV radiation in the next years on a small spatial scale (McKenzie et al., 2011). This will help to control whether the Montreal Protocol targets are fulfilled and the model predictions are correct. Moreover, quality ground-based instruments, traceable to a common reference, allow to validate satellite data mainly over polluted locations (Lee-Taylor et al., 2010) or at sites with a complex orography”: “… the campaign gives also additional information about the effectiveness of different correction procedures and the accuracy of calibration coefficients. Although this work involved instruments operated by Italian institutions, some results may be useful also for other operators. First, the variety of the radiometers participating to the campaign and the different calibration and processing procedures make the comparison representative of a wider community than the Italian one. Most radiometers (11 of 13) are commercialized worldwide. Thus, the comparison can provide a contribution to the knowledge about the general performance of narrow- and broad-band UV radiometers, their characteristics, their operation and limits. Furthermore, some radiometers participating to the campaign were calibrated few months before the comparison by their respective manufacturers. Thus, the results of the comparison may be useful to obtain some information about the effectiveness and the consistency of the calibration procedures adopted by the manufacturers. Finally, the paper presents an in-depth analysis of the algorithm used to compare broadband and spectral UV data, that can improve the theoretical basis of the methods previously reported in literature”.

**Results and discussion:** 5.5 Overall performances of the radiometers and recommendations.

Even within the uncertainty of the reference instrument, some inferences about the general performances of the radiometers can be drawn and some recommendations may be consequentially formulated.

**Figure 11** (1, in the Supplement) shows both the relative deviation from the reference during clear sky days and the IQR distance. Only results from the UV Index comparison are represented (UV-A irradiances have a similar behaviour and are omitted). Angular and spectral corrections employed by the operators are also reported in parentheses next to the instrument id. Some groups of instruments, which show similar performances, can be identified.

One of the most striking features of the graph is the importance of the traceability. Indeed, most of the instruments traceable to NTP (directly through QASUME or, with a further step, through Bentham 5541), with the only exception of instrument 12, show very low deviations with respect to the reference (blue markers). This points out the importance of a common and reliable reference scale and the effectiveness of a travelling standard such as QASUME. It is interesting to notice that the relative deviations from the reference among this group do not appreciably depend on the age of the last calibration. Broadband radiometers such as id 06 and 07 are stable, even after some years of operation after their last calibration. However, sudden changes may occur and calibrations on a more frequent basis are recommended.

Instruments which were calibrated by their respective manufacturers few months before the campaign may be included in a second group (red colour). This group is characterized by moderate to large deviations relative to the reference. Moreover, consistency among instruments calibrated by the same facility (e.g. id 05 and 04) is weirdly low. Furthermore, instrument 04 was recalibrated at the same laboratory after the comparison and is now only about 4% higher than the reference. Such large discrepancies
Finally, a third group including two radiometers designed and built by ENEA and CNR-ISAC is highlighted (violet colour). Large relative deviations with respect to the reference are found. The intercomparison helped in identifying technical problems such as deterioration of internal components and drifts in calibrations. Thus, particular attention should be taken when operating with home-made radiometers and comparisons with reference instruments should be scheduled very often.

A further classification may be done according to the processing procedures and correction algorithms. Generally, radiometers processed using a matrix (which takes into account both spectral and angular corrections) show a lower IQR. On the contrary, other kinds of corrections, such that employed for radiometer 08, are not so effective. When a fixed calibration factor is used, the IQR may be even worse, as in the case of radiometer 07. Even though low daily variations may be obtained (as with instrument 10) with a fixed calibration factor, possibly because of a good cosine response, both spectral and angular corrections should be always employed.

Generally speaking, performances of the radiometers, especially of those with the highest deviations or IQRs, could be greatly improved by using state-of-the-art correction algorithm and planning frequent intercomparison for monitoring the instrumental stability.

**Summary and conclusions:** 
"... participated to the campaign. Most radiometers (11 of 13) are commercialized worldwide. The campaign also represented a chance of comparing several data processing algorithms employed by the participating institutions."; 
"... even for clear sky days were discovered and discussed. Three radiometers, which had been calibrated few months before the campaign by their respective manufacturers, showed significant deviations to the reference. The average deviations and daily variations were very large ...".

Comment #2: Furthermore, the analysis procedure used in this manuscript is different for particular instruments (Sections 4.1 and Sections 4.2) which renders the comparison between these subsets a questionable task. I recommend the authors to use a common analysis approach for all instruments in the campaign, even if the analysis of particular radiometers could be improved (e.g. Section 4.1).

**Answer #2:** The following sentence will introduce Sect. 4 of the revised manuscript to explain the need of different analysis procedures:

Since the instruments to be compared against the reference belong to very different classes (spectral, narrow- and broad-band radiometers), a common procedure of analysis to be applied to all instruments is not suitable. Most notably, since the sampling frequencies of the instruments are considerably different, a common approach for all instruments involved in the campaign should be based on the adoption of an identical time resolution (i.e. downscaling all data to the lowest resolution among the participating instruments) which is inappropriate and can lead to ambiguous results. Indeed, the deviations to the reference would increase due to the temporal interpolation, making arduous to discriminate the dispersion originating from other factors and the dispersion from the interpolation itself.

A reliable and rigorous analysis is essential to correctly compare the processing procedures used by each participating institution, which is one of the purposes of the campaign. Therefore, an appropriate algorithm must be employed for each class of instruments for comparing the UV data recorded by the various instruments against the reference and reducing the dispersion originating from the temporal interpolation. The methods are described below.
Comment #3: The analysis procedure described in Section 4.1 for a subset of the UV Radiometers of this comparison is very elaborate and reference is made to a procedure used in a previous campaign (COST726). A significant difference between the COST726 and the present campaign is that the former was also used to calibrate the radiometers while the one described here was limited to comparing the results from the UV radiometers applying the calibration from the home institutes. Thus I expect the analysis approach to be fundamentally different. In that sense I believe that Section 4.1 contains unnecessary relict information from the COST726 intercomparison and should be considerably simplified. Lines 12-21, including equations 2-4 are unnecessary for the analysis of the present campaign and should be omitted.

Answer #3: Section 4.1 has been rewritten omitting relict information from the COST726 intercomparison, according to the referee suggestion (text from line 17 p. 2800 to line 18 p. 2801 was omitted). The text has been revised in the following way:

... an appropriate algorithm was developed.

The downscaled irradiance from broadband radiometers was calculated as

\[ I_{ds} = \frac{I_{BB}(t(\lambda))I_{0SP}(t(\lambda))CIE(\lambda)d\lambda}{\int I_{BB}(t(\lambda))CIE(\lambda)d\lambda} \]

where \( I_{BB}(t(\lambda)) \) are the CIE-weighted and cosine-corrected broadband irradiances reprocessed by the operators and \( I_{0SP}(\lambda,t(\lambda)) \) is the clear-sky irradiance simulated by the radiative transfer code at the time at which the reference instrument is measuring the wavelength \( \lambda \).

Comment #4: The main benefit of the elaborate analysis procedure in Section 4.1 is related to broken cloud conditions when the solar spectrum measured by the spectroradiometer is significantly affected. In that case however, also the broadband measurements are difficult to estimate due to the questionable way of correcting for angular response deviations (cosine corrections). As stated in 5.4, only a clear sky cosine correction was applied which is simplistic and in marked contrast to the elaborate comparison methodology. It would be interesting to determine the effect from that simplification on the performance of the radiometers during broken cloud conditions.

Answer #4: Section 5.4 has been rewritten as follows in order to better explain the benefits of the algorithm and to show that the analysis of IQRs allows to estimate the effect of the clear-sky simplification on cosine corrections, at least for analog broadband radiometers:

The effect of clouds on the ratios can be examined taking into account the series of measurements recorded in the whole campaign period (days 159 to 175), in both cloudy and clear days (Figs. 9 and 10). While median values do not change appreciably, the scatter of the ratios for each instrument and, in particular, the difference between the maximum and minimum values (i.e. the distance between the whiskers in the boxplots), increase noticeably and in many cases exceeds \( \pm 10\% \). This is to ascribe to many factors, depending on the instruments.

First, the analysis procedure which was applied to the analog broadband data allows to reduce the effect of time interpolations on the ratios. The dominant factor modulating the ratios must therefore be the influence of clouds on the cosine corrections, which
were calculated under the hypothesis of clear sky conditions. This latter approach is, of course, simplistic, but it is adopted by most of the broadband operators worldwide. Also, more rigorous corrections would require cloud cover data which are not always available. For instance, a complete spectral correction in cloudy cases would strongly depend on both cloud cover and cloud optical thickness (Mateos et al., 2011). A simpler correction is possible in the borderline case of overcast sky in presence of thick clouds, taking into account only the fraction of scattered radiation while neglecting the direct beam. Nevertheless, this diffuse correction would require a complete radiometer characterization (angular and spectral responses), which was not available for all radiometers. However, the increase of the IQRs for analog broadband radiometers during cloudy conditions with respect to clear-sky conditions is a good estimate of the error in the cosine correction originating from the clear-sky simplification. It is interesting to notice that radiometers with a good cosine response (i.e. with a smooth matrix), for example id 01, 02, 04 and 05, show slightly better performances compared to the others, e.g. id 03 and 06.

In the case of instruments with coarser time resolutions (e.g. 09 to 11 and, to a lesser degree, 12 and 13), time interpolations between measurements may produce large deviations. However, a more in-depth analysis of cloud effects on the cosine corrections for these instruments is beyond the purpose of this paper and may be studied in a future work.

Finally, it is interesting to notice that the error is amplified by the processing algorithms applied on radiometers 07 and 08. In those cases, the full range of the ratios may exceed ±20% or even ±30%.

Comment #5: line 1, page 7 I do not understand this sentence (How can the integrated clear irradiance \( I_{0BB} \) (please define it in the text) NOT change during a spectral scan? What is meant by appreciably? Please quantify?)

Answer #5: The text has been modified as follows:

We also assume that clouds act as a grey filter, i.e. the spectral transmittance of the cloud at all wavelengths can be approximated by the transmittance integrated over a wavelength band (e.g. UV-A or the erythemal range). Marking the modeled clear-sky irradiance over the spectral band of interest with \( I_{0BB}(t) \), we obtain:

\[
\frac{I_{BB}(t)}{I_{BB}(t)} \approx \frac{I_{SP}(\lambda,t)}{I_{SP}(\lambda,t)} \quad \forall \lambda
\]

As previously stated, a spectral scan of the reference instrument takes about 3 minutes. During this time, the variation of \( I_{0BB}(t) \) depends uniquely on the change of the solar zenith angle and can be considered negligible with respect to the fluctuations of the measured irradiance, \( I_{BB}(t) \), which is induced by the transit of broken clouds. Based on radiative transfer calculations, the error in considering \( I_{0BB}(t) \) as a constant is estimated to be lower than 1%. Thus, we obtain:

\[
I_{ds} \approx \int I_{SP}(\lambda,t(\lambda))CE(\lambda)d\lambda
\]

Comment #6: Section 4.2. The use of cubic splines is very handy, but it can produce nasty surprises when applied automatically. Furthermore, how do you know that the radiation changes like a cubic spline in between the missing measurements? I would recommend the use of a simple linear interpolation.

Answer #6: The authors intended just to point out the risk of an indiscriminate use of time interpolations (including cubic splines). Introduction to Sect. 5 has been rewritten as follows:
Figure 4 presents a comparison of several methods to downscale the data from broadband radiometers during three cloudy days. The aim is to draw attention to the risk of an indiscriminate use of time interpolations. First, the ratios derived by interpolating the broadband values to the most representative time of the reference spectra using a cubic spline and a simple linear interpolation are shown. Cubic splines are chosen since they were applied to depict the results of the COST726 campaign (G. Hülsen, personal communication). As can be seen from the figure, the results obtained in cloudy conditions are not optimal and present some fictitious fluctuations. The linear interpolation gives results similar to the cubic spline. The new algorithm developed in this study ...

Comment #7: Section 5.1, lines 11,12: Please provide a description of how the radiometers were calibrated relative to the Bentham, or a reference where that method is described. How does this “exercise” provide information on the reliability of the angular correction?

Answer #7: A description of how the radiometers were calibrated relative to the Bentham is already reported in Sect. 2.3.1, which has been expanded as follows:

In such a way, spectral and angular corrections could be easily introduced. More precisely, data from radiometers 01-03 were processed with a matrix calculated by the owner agency: the spectral and angular characterization of those radiometers was performed by the PMOD-WRC, then the spectral and angular corrections were calculated by ARPA Valle d’Aosta using the libRadtran model. Finally, an absolute calibration factor was determined some weeks before the comparison campaign with reference to the Bentham spectroradiometer, following the procedure described by Gröbner (2007).

Reference to Sect. 2.3.1 was added in 5.1. Also, the next paragraph has been included in Sect. 5.1 following the suggestion of the referee:

Since the IQRs of those radiometers are very low (ranging from 1.4 to 3%), this exercise also provides information on the radiometers stability and the reliability of the angular corrections.

Comment #8: Page 8, line 24. Was that statement verified by measuring the angular response of these radiometers? As far as I know this information is not supplied by the manufacturers.

Answer #8: Text in Sect. 5.2 has been rewritten in the following way:

Figures 5 and 6 clearly show that the series of measurements of every instrument have a different IQR during clear sky days. Even though the angular response of the radiometers was not measured during the campaign, most of the total variability for clear sky conditions, as we will see later, is clearly ascribable to the daily (not day-to-day) variability...

Comment #9: page 9, line 16. What does mean “the full range of ratios increases”? Could you try to rewrite this sentence?
Answer #9: The sentence has been rewritten in the following way:

The scatter of the ratios for each instrument and, in particular, the difference between
the maximum and minimum values (i.e. the distance between the whiskers in the
boxplots), increase...

Comment #10: The acronyms used in this manuscript are not consistently used
throughout the manuscript and should be spelled out at first use: Examples are:
ARPA, IBIMET, APPA, PMOD (identical to PMOD-WRC?), QASUME.

Answer #10: A table containing the acronyms (see Table 1, Supplement) of all
participating agencies has been added to the manuscript. The latter was completely
revised for a consistent use of the agency names.

Comment #11: The list of references are extensive (maybe too much for the
type of manuscript). I think some redundant references could be omitted if the
original citation is kept. Reference on page 11, line 30 seems incomplete. Is
there a web-link?

Answer #11: The following redundant references were omitted:

- Lucas et al. 2010 (p. 2791)
- McKenzie 2003 (p. 2791)
- Bernard et al. 1998 (p. 2792)
- Webb et al. 1998 (p. 2792)

Reference on line 13, p. 2810 has been rewritten as

BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2008 Evaluation of Measure-
ment Data-Guide to the Expression of Uncertainty in Measurement GUM 1995 with
minor corrections. Joint Committee for Guides in Metrology, JCGM 100

Comment #12: I am uncomfortable with the content of the reference scale
column, as I wonder how a specific instrument can represent a reference scale. I
would recommend stating the Institute or laboratory to which the measurements
are traceable to and modify the title from “reference scale” to ”Traceability”.

Answer #12: The table has been modified according to the suggestion of the referee
(see Table 2, Supplement).
Comment #13: tables 2 and 3 should state clearly the use of expanded or simple uncertainties (coverage interval etc...).

Answer #13: The tables were updated with the coverage factors.

Comment #14: Was the linear drift of the responsivity of the reference spectro-radiometer (see Figure 1) of 1% taken into account in that uncertainty estimate?

Answer #14: The text in Sect. 2.1 has been expanded following the suggestion of the referee:

Figure 1 presents the variations of the spectral responsivity, which are not strictly a drift (the responsivity measured during the first calibration, used as a reference for the following, is lower than the second and similar to the third). The changes are to ascribe to the instrumental instability (0.4% uncertainty, calculated assuming a rectangular probability distribution) and to the heating of the diffuser during the calibration (1% uncertainty). The combined uncertainty originating from the two factors is higher than the observed variability. In order to reduce the errors ...

Comment #15: Table 3: Intuitively, I would expect the wavelength uncertainty to increase with increasing SZA. This does not seem to be the case for the range 310-400 (last line of the table). Can the authors confirm these values?

Answer #15: It is a typo (as one might see from the estimate of the total uncertainty, which is correctly calculated). The wavelength uncertainty in the third column should be 0.9% (according to Gröbner 2005). This table has been corrected.

Comment #16: table 4: The radiative transfer model is used not only for the analysis, but also as an independent radiometer. Thus the use of particular parameters is worthwhile to be discussed. Indeed, can the authors comment on the following aspects.

Answer #16: The model description in Sect. 2.5 has been expanded following the suggestion of the referee:

A radiative transfer model, the libRadtran package (Mayer and Kylling, 2005), was used in the campaign for comparing different kinds of data as explained in Sect. 4 and as a further quality control. Table 4 summarizes the data set entered as input to the model.

The solar spectrum was set to the recommended value following the model documentation (Atlas-3, shifted to air wavelengths). Default summer atmospheric profiles were used. Pseudo-spherical discrete-ordinate method (DISORT) with double precision was chosen as the solver, since a simple plane-parallel DISORT solver showed relevant deviations from the reference instrument even for low zenith angles. The effective ground albedo was set to 3% (Degünther et al., 1998). Rural aerosol properties, background stratospheric aerosols and the default Shettle aerosol profile were given as inputs to the model. Since independent measurements of the aerosol single scattering albedo (SSA) were not available during the campaign, the SSA value was chosen in order to best reproduce several spectral measurements recorded with Bentham 5541 during clear-sky days in summers 2008 to 2010 at Saint-Christophe (wintertime measurements were not considered because of changes of effective ground albedo due to the snow) as explained by Ialongo et al. (2010). Therefore, the single scattering albedo (SSA) was reduced by 10% relative to the default model value (i.e. 0.90 to 0.95 depending on the wavelength). Similarly, the Molina&Molina ozone cross sections were
chosen because of their agreement with the Bentham spectral measurements in the range 295÷330 nm.

The Ångström coefficients were retrieved from the Brewer measurements in the UV and visible range, as explained in Sect. 2.2. Local atmospheric pressure was taken equal to a constant value of 950 hPa, since the measured pressure during the campaign was stable within ±5 hPa (the error introduced by using a constant value is less than 0.4% at 290 nm and even lower at higher wavelengths). The diffuse irradiance was scaled to 95% accounting for the mountain horizon under the hypothesis of isotropic diffuse radiation, as explained by Diémoz and Mayer (2007). This cosine-weighted fraction was calculated from both theodolite measurements and a digital elevation model.

The simulated spectra were then treated similarly to the instrumental data (id 14 was assigned to the model) and compared to the reference.

It should be stressed that the aim of our work was not to accurately retrieve the previously mentioned parameters, since a relatively large range of values may originate realistic spectra. The purpose was rather to restrain some relevant and free model parameters on the base of the observations and to achieve the best agreement between model and measurements. Of course, other minor factors which are not taken into account (such as aerosol and gas vertical profiles, other aerosol properties, surrounding surfaces orientation, etc.) can still influence the model.

Comment #17: a) Please state the reason for reducing the default SSA by 10% . Is there an independent reason for that change from the default values?

Answer #17: See answer #16. The spectral SSA used in the model (scaled by 10%, and integrated over the altitude) is shown in the Supplement, Fig. 3.

Comment #18: b) beta is defined at 1000 nm, while the aod from the Brewer is obtained at 320 nm (Section 2.2). How was it converted?

Answer #18: Section 2.2 has been updated as follows:

The algorithm developed by Cheymol and De Backer (2003), together with data from an in-situ Langley Plot calibration, is regularly employed to retrieve the aerosol optical depth (AOD) at 320 and 453 nm from clear-sky UV direct irradiance. The Ångström coefficients can then be estimated from measurements at these two wavelengths, according to Gröbner and Meleti (2004). The AOD and the Ångström coefficients are later included in the radiative transfer calculations.

Comment #19: c) There is an inconsistency in the use of the total ozone from the brewer and its use in the RT model because the Brewer retrieves the ozone using the Paur & Bass x-sections while the model uses the Molina&Molina X-sections. It would be interesting to see how much the model output changes if PB x-sections are used.

Answer #19: The standard Brewer algorithm estimates the ozone content from narrowband direct sun measurements in the nominal wavelength range 310-320 nm (the 306 nm slit is only employed for determining the SO$_2$ content), using Bass&Paur cross-sections as explained by the referee. The authors have decided to choose Molina&Molina cross-sections by comparing clear-sky modelled and measured spectra (obtained in summers 2008-2010) over a wider wavelength range (295-330 nm) than the Brewer's (see answer #16). Indeed, as the reference spectroradiometer is equipped with a double monochromator, irradiance measurements are reliable starting from the lowest wavelengths. The best agreement was therefore found with Molina&Molina
The following table summarizes the comparison using Molina&Molina and Bass&Paur cross-sections:

<table>
<thead>
<tr>
<th></th>
<th>Molina&amp;Molina</th>
<th>Bass&amp;Paur</th>
</tr>
</thead>
<tbody>
<tr>
<td>UV-index, mean difference</td>
<td>-0.3%</td>
<td>2.1%</td>
</tr>
<tr>
<td>(model-meas)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UV-index, IQR</td>
<td>2.6%</td>
<td>3.6%</td>
</tr>
<tr>
<td>UV-A, mean difference</td>
<td>-1.2%</td>
<td>-1.1%</td>
</tr>
<tr>
<td>(model-meas)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UV-A, IQR</td>
<td>2.9%</td>
<td>3.7%</td>
</tr>
</tbody>
</table>

Figures 4 and 5 in the Supplement show the (spectral) ratios between the model and the reference, using Bass&Paur and Molina&Molina cross-sections, respectively. The calculations take into account the different Rayleigh scattering coefficients used in the Brewer algorithm and the model, which otherwise originate a fictitious difference of about 3 Dobson Units (J. Gröbner, personal communication). As can be noticed in the figures, major differences arise at wavelengths lower than the range spanned by the Brewer (blue lines in the figure).

Comment #20: Figure 2: There are substantial obstructions compared to a clear horizon. While this will not significantly affect the instruments since they all have more or less the same angular response (did you check this assumption?) in contrast the RT Model computes its irradiance for an unobstructed horizon. Did you apply corrections to the RT model to take the true horizon into account?

Answer #20: see answer #16.

Incidentally, the cosine-weighted fraction of sky above the horizon in Saint-Christophe is similar to that obtained at the Davos station during COST726 campaign. For that comparison, the authors concluded that “The largest influence [of the horizon on the cosine correction of radiometers] is of the order of 0.8% for a total overcast sky and even less for a clear sky where this error would be reduced by the direct to diffuse radiation ratio”.

The matrices relative to the instruments belonging to ARPA Valle d’Aosta have been calculated taking into account the real horizon.

Comment #21: Figure 3: I would not call the periods before sunrise and after sunset as missing data. If really necessary, a possible quantification could be the total number of possible measurements versus the actual measurements).

Answer #21: The caption has been modified as follows:

Data for rainy periods, before sunrise, after sunset, during calibration of the reference and dome cleaning were not included in the graph.

The comparison has not been performed during these periods, since most of the matrices were calculated for clear-sky and visible sun.

Comment #22: Figures 5,6,9,10 would benefit from thicker lines. In the caption of figure 5, there is only one line inside the box (no plural).

Answer #22: The figures were adapted following the suggestion of the referee. The caption of Fig. 5 has been corrected.
Comment #23: Figure 9: Typo, the id should be 14, not 04.

Answer #23: The caption has been corrected.

Please also note the supplement to this comment:
http://www.atmos-meas-tech-discuss.net/4/C1170/2011/amtd-4-C1170-2011-supplement.pdf