

## Appendix VII - UV/Vis Instruments

Passive ultraviolet and visible (UV/Vis) spectroscopy using scattered sunlight as a source has been progressively developed since the late seventies as a powerful remote-sensing technique for the unattended long-term monitoring of stratospheric and tropospheric trace gases. One main advantage of UV/Vis spectroscopy is to allow automated daily measurements of atmospheric gases even under moderate cloud cover. This technique has been widely used in atmospheric chemistry and validated through a number of intercomparison exercises as well as through various contributions to the validation of atmospheric chemistry satellite missions like ERS-2 GOME, TOMS, ENVISAT, OMI and METOP GOME-2. While the UV/Vis technique does not require further justification for being included in NDACC observation system, individual instruments still need to be validated and the quality of long-term observations at UV/Vis sites still needs to be assessed. The present document describes the validation process for new UV/Vis instruments, as well as the criteria for maintaining data quality from existing instruments. It is written to cover measurements of the vertical column abundances of stratospheric NO<sub>2</sub> and ozone, which are the primary UV/Vis data products regularly archived in the NDACC data base. However, it is also meant to apply to measurements of any species and data products retrievable by UV/Vis spectroscopy. This includes e.g. vertical profiles of BrO, NO<sub>2</sub> and HCHO that can be derived using Multi-Axis DOAS (MAXDOAS) systems. Such data products are currently under development and still require further research to better assess their accuracy and information content. Upon definition and implementation of appropriate standards for measurements, retrieval techniques and error analysis, it is anticipated that they will progressively be added to the list of archived NDACC data sets.

### Quality criteria for the evaluation of new instruments and instrument teams

UV/Vis spectroscopic measurements can be used to answer a variety of scientific questions. In general however, the emphasis within NDACC is on long-term measurements and global studies including multi-missions satellite validation which require a long term dedicated approach to the maintenance of the quality of the measurements and the archiving of data. To determine long term trends requires stable and well calibrated instruments operated by groups that have a thorough understanding of the measurement technique. Generally speaking, the accuracy of UV/Vis data products (e.g. NO<sub>2</sub> vertical columns derived from zenith measurements) is determined by several factors:

- (1) measurement accuracy (random and systematic), which is primarily determined by instrumental factors, but also by the quality of molecular absorption cross-sections used in the retrieval process;
- (2) the accuracy of air mass factor (AMF) calculations, which depend on (a) the suitability of radiative transport models used to simulate sky radiances and (b) the choice of the atmospheric data bases used as an input (e.g. atmospheric temperature, pressure and ozone profiles);
- (3) other uncertainties due to effects not explicitly treated in the inversion process (e.g. scattering by clouds and aerosols, inelastic scattering (Ring effect) or polarisation).

For total column measurements of NO<sub>2</sub> and ozone, the limiting accuracy of the most accurate instruments operating at clean sites is determined by the accuracy of the AMF calculations and the accuracy of the calculation of the residual amount in the reference spectrum.

The process of certifying a new UV/Vis observing system of the NDACC involves two major steps: (1) an evaluation of the instrument design and of the available data analysis tools, and (2) the formal participation to a blind or semi-blind instrument intercomparison campaign. Full certification is granted to instruments and measuring groups that fulfil a set of general and specific criteria as described below.

### Evaluation of instrument design and data analysis

Before a formal intercomparison with a certified instrument is planned, the group whose instrument is being assessed is asked to supply the following to the NDACC UV/Vis working group representatives:

- A detailed technical description of the instrument including sensitivity limits and general operating parameters.
- An outline of the spectral analysis technique used with particular details of the number and source of the cross sections used, as well as the wavelength calibration procedure used.
- An example of a raw measured spectrum, a ratio spectrum of a twilight spectrum (near 90 degrees SZA) to a midday spectrum, and a spectrum demonstrating the quality of the cross sections fit to the ratio spectrum.
- Spectra or data showing the instrument resolution (slit function), an estimate of the stray light levels and the instrument polarization characteristics, in the wavelength interval expected to be used for the intercomparison. This would normally consist of:
  - (1) Slit function: one or more spectra containing lines from spectral lamps (e.g., low-pressure mercury, neon, argon, krypton or xenon) or a laser;
  - (2) Stray-light: clear sky spectra measured using fixed detector gain, with and without suitable Schott glass short I cut filters;
  - (3) Polarization: two or three spectra showing the relative transmission of the instrument for different polarization axes that have been measured using a white light source filling the field of view together with a suitable film polarizer.
- Examples of existing measurement data and (if available) results of any previous intercomparisons (including conditions and references).

Discussions and data exchange between the PI and the UV/Vis working group representatives may be required, as the instrument group must be satisfied with this part of the evaluation before proceeding to an intercomparison.

### Instrument intercomparison field campaign

Instrument intercomparison field campaigns are regularly organised according to needs, and availability of resources. So far, formal NDACC UV/Vis intercomparison campaigns have been organised in 1992 at Lauder (New-Zealand), in 1996 at OHP (South of France), in 2003 in Andoya (Norway) and in 2009 in Cabauw (The

Netherlands). The aim of such campaigns is to provide an opportunity for certification of new groups but also and more importantly to foster interactions between groups and hence promote scientific improvements for the benefit of the whole network.

The certification of a new instrument (or instrument group) relies on the successful participation in a formal instrument intercomparison exercise organised according to the “semi-blind” rules detailed below. In the general case, one or more new instruments will be evaluated by comparison with one or several already certified instrument(s) (called the reference instrument(s) in the following) under the supervision of an impartial campaign referee.

- The intercomparison will be conducted at a site selected according to current scientific priorities, i.e. a site where the probability to perform successful observation of the target data products is high enough. Generally speaking the site should be such that the probability to experience both clear and cloudy days over the intercomparison period is high enough. Measurements during both clear and cloudy conditions are important to evaluate in which way different skies impact the results.
- The intercomparison should be conducted for a period of not less than 10 days with all instruments operating correctly.
- Measurements, taken by the one or more instruments being evaluated and the reference instrument(s), should be made over the whole day with a period of high temporal sampling near midday, each day of the intercomparison irrespective of experienced weather conditions. For twilight measurements, the integration period should be less than the time taken for a 1-degree change of solar zenith angle (5 minutes at mid-latitudes) or a maximum of 5 minutes for extreme solar zenith angle values.
- Measurements taken by the instrument(s) being evaluated and the reference instrument(s) should be as much as possible coincident in time, to minimize interpolation errors when performing comparisons.
- The wavelength interval used should be the same for the instrument(s) being evaluated and the reference instrument(s). Non-standard wavelength interval might have to be selected to fulfill this important requirement.
- The cross sections used in the analysis must be from the same source, and appropriately convolved to each instrument resolution using measured instrumental slit functions.
- If the polarization characteristics are not supplied by the PI of an instrument being evaluated, they must be measured during the intercomparison according to protocols to be established
- Unless otherwise specified by the campaign referee, analysis should provide two sets of results: (1) analysis using a daily selected midday reference spectrum, to be submitted to the referee normally within 1 day of the measurements; (2) analysis using a single midday reference spectrum for the whole intercomparison data set to be submitted to the referee within maximum 3 months of the end of the intercomparison. Final “polished” results will also be submitted then. The choice of the daily reference spectrum and the single campaign reference spectrum will be made by the referee in consultation with the instrument groups.

- Blindness rules are important for formal NDACC intercomparisons. Indeed, the goal is to provide NDACC data users with evaluations that represent a “true picture” of each instrument performance. However experience has shown that campaign results strongly benefit from the adoption of a “semi-blind” intercomparison protocol where preliminary data submitted by participants can be displayed by the referee during the campaign, providing it is in a form that does not enable participants to identify individual instruments. While total “blindness” would achieve this, it limits the opportunity for groups to learn. Displays of data that do not identify groups (but enable participants to see the general form of the others measurements) were found to not compromise the integrity of the intercomparison and are therefore recommended for use in any formal UV/Vis intercomparison exercise. Note that individual results must not be exchanged between any participants being evaluated until final results are submitted by all instrument groups.

### Acceptance criteria for new instruments

The instrument group or its designated representative(s) will examine the results of the intercomparison and make a recommendation to the NDACC steering committee. While additional factors may possibly enter the considerations, the following points are considered as general criteria for acceptance:

- The instrument tests to measure resolution (slit function), polarization characteristics, and stray light levels must be acceptable.
- Good result self-consistency: this can be assessed by examining the “smoothness” of the twilight data series. At clean sites, midday result variations should remain small,  $< 1 \times 10^{15} \text{ cm}^{-2}$  for  $\text{NO}_2$  and  $< 1 \times 10^{18} \text{ cm}^{-2}$  for ozone slant columns.
- Acceptable signal to noise at high (near 95 degrees for spectra in the visible) and small (near 70 degrees) solar zenith angles. Again this can be estimated by examining residual spectra or the “smoothness” of the result series (above).
- Good consistency between the results obtained using the daily reference spectrum and the results obtained using the campaign single midday reference spectrum. This helps identify problems caused by long period (10 day) drifts in the instrument function or spectral wavelength repeatability. The reference instrument errors can be used as a guide for acceptance.

In addition measurable criteria for the certification of the primary NDACC data products ( $\text{NO}_2$  and  $\text{O}_3$  slant column measurements) have been formalised after the intercomparison campaigns of OHP in 1996 (Roscoe et al., 1999) and Andoya in 2003 (Vandaele et al., 2005).

Because no absolute calibration is possible, accuracy is determined by quantifying the consistency of each instrument to be evaluated relative to the designated reference instrument(s). Spectral measurements made during the intercomparison period are analysed by all participants using agreed criteria (wavelength interval, cross sections, etc.) to obtain “their intercomparison results”. A reliable method to determine which instruments meet a certification is a regression analysis where all

combinations of the twilight sets of measurements are intercompared. Matrices of residual error, slope and intercept are generated in order to identify the instruments that agree most closely. The results from these instruments can then be used as the reference results for comparing the results of the other participating instruments against (Roscoe et al., 1999; Vandaele et al., 2003). For example in 2003 at the Andoya intercomparison, the NO<sub>2</sub> results from 4 instruments, (two of them were designated reference instruments) agreed to within  $(1.00 \pm 0.01)$  in slope and  $(0.00 \pm 0.05) \times 10^{15} \text{ cm}^{-2}$  in intercept in one analysis (same wavelength interval and NO<sub>2</sub> cross sections). Similar or better ozone results were obtained during the OHP intercomparison in 1996. This close agreement has been the basis for choosing the following figures as acceptance criteria for NO<sub>2</sub> and O<sub>3</sub> measuring instruments:

Acceptance criteria for NO<sub>2</sub> slant column measurements

Slope =  $1.00 \pm 0.05$ ; Intercept  $\leq \pm 0.1 \times 10^{16} \text{ cm}^{-2}$ ; Residual  $\leq 0.05 \times 10^{16} \text{ cm}^{-2}$

Acceptance criteria for O<sub>3</sub> slant column measurements

Slope =  $1.00 \pm 0.03$ ; Intercept  $\leq \pm 0.15 \times 10^{19} \text{ cm}^{-2}$ ; Residual  $\leq 0.10 \times 10^{19} \text{ cm}^{-2}$

Groups with intercomparison results that meet these accuracy criteria together with the general acceptance criteria (see above) are certified for NDACC UV/Vis observations. Note that these specifications are not to be taken as rigid criteria. Some groups may have instruments that produce results that are close to these figures. Limited certification can then be used to recognise the potential of such instruments and, in this case, assistance will be offered to improve measurement accuracy and help reaching full certification at the following intercomparison.

**Quality criteria for the evaluation of continuing instruments and instrument teams**

The investigator has primary responsibility for maintaining data quality from his/her instrument on a continuous long-term basis. This should include routine procedures to check on instrument performance (e.g. stray light monitoring, instrument resolution tests, error figure monitoring and residual spectra checks).

The investigator must maintain suitable instrument operation and maintenance records. Repairs and changes to equipment must be carefully logged and calibrations made afterwards to identify any changes in accuracy.

Recommendations for standardized instrument operation and data retrieval might be agreed upon as part of activities of the UV/Vis working group. It is the responsibility of the investigator to make sure that these recommendations are properly implemented and used for the production of data to be archived in the NDACC database.

Where available, the investigator should use data from other instruments at the measurement site to compare with their UV/Vis measurements. For example: Dobson, Brewer, Sonde and Lidar data for ozone total column comparisons. The use of more than one UV/Vis instrument and the comparison of results also offer a higher level of confidence in the data.

The investigator must maintain a routine data archiving procedure. The maximum time between measurement and data submission should not exceed 12 months.

When equipment problems compromise data continuity or quality, the investigator should promptly discuss the situation with the NDACC UV/Vis Working Group. This requirement is intended to encourage the open exchange of information from which all groups can learn. Also, when instrument failure seems likely to result in an extended loss of data, other groups may be able to help (e.g. with the temporary deployment of an available instrument to fill the gap).

The investigator must be willing to participate in intercomparisons exercises. However, because of the effort and cost of these, the timing of them should be determined primarily by need, with a maximum gap between intercomparisons of 5 years. For example, when improved or new measurement or analysis techniques are proposed, the instrument group, after favorable evaluation of them, could choose to conduct a suitable intercomparison inside the 5 year window.

Instrument and analysis improvements that enhance scientific output or data quality are encouraged. The group should ensure however that data continuity and quality is maintained. Where possible an improved instrument should be operated in parallel with the existing instrument for a period of at least 6 months, and the data carefully compared. When an instrument or analysis technique improvement results in a change in the measurement results, this must be fully reported and recorded in the archive.

Approximately every 1-2 years, the NDACC UV/Vis instrument group and community should hold a workshop to discuss ways of improving measurement and analysis quality, and solving the remaining problems. Groups experience staff changes occasionally and the training and the free exchange of experience and knowledge within the UV/Vis community that workshops provide is an excellent way of providing training for new people. Workshops also foster continuing collaboration between the many groups that will help to ensure the quality of network UV/Vis measurements into the future.

Version: November 30, 2009

## References

- Hofmann, D., P. Bonasoni, M. De Maziere, F. Evangelisti, G. Giovanelli, A. Goldman, F. Goutail, J. Harder, R. Jakoubek, P. Johnston, J. Kerr, W. Matthews, T. McElroy, R. McKenzie, G. Mount, U. Platt, JP. Pommereau, A. Sarkissian, P. Simon, S. Solomon, J. Stutz, A. Thomas, M. Van Roozendael, and E. Wu, Intercomparison of UV/visible spectrometers for measurements of stratospheric NO<sub>2</sub> for the network for the detection of stratospheric change, *J. Geophys. Res.*, 100, 16,765, 1995.
- Roscoe, H.K.; Johnston, P.V.; Van Roozendael, M.; Richter, A.; Roscoe, J.; Preston, K.E.; Lambert, J.-C.; Hermans, C.; DeCuyper, W.; Dzienus, S.; Winterrath, T.; Burrows, J.; Sarkissian, A.; Goutail, F.; Pommereau, J.-P.; D'Almeida, E.; Hottier, J.; Coureul, C.; Didier, R.; Pundt, I.; Barlett, L.M.; McElroy, C.T.; Kerr, J.E.; Elokhov, A.; Giovanelli, G.; Ravegnani, F.; Premuda, M.; Kostadinov, I.; Erle, F.; Wagner, T.; Pfeilsticker, K.; Kenntner, M.; Marquard, L.C.; Gil, M.; Puentedura, O.; Arlander, W.; Kastad Hoiskar, B.A.; Tellefsen, C.W.; Heese, B.; Jones, R.L.;

Aliwell, S.R.; Freshwater, R.A. (Accepted, October 1998). Slant column measurements of O<sub>3</sub> and NO<sub>2</sub> during the NDSC intercomparison of zenith-sky UV-visible spectrometers in June 1996. *Journal of Atmospheric Chemistry*, 32: 281 - 314, 1999.

Vandaele, A.C., C. Fayt, F. Hendrick, C. Hermans, F. Humbled, M. Van Roozendael, M. Gil, M. Navarro, O. Puentedura, M. Yela, G. Braathen, K. Stebel, K. Tørnkvist, P. Johnston, K. Kreher, F. Goutail, A. Mieville, J.-P. Pommereau, S. Khaikine, A. Richter, H. Oetjen, F. Wittrock, S. Bugarski, U. Friess, K. Pfeilsticker, R. Sinreich, T. Wagner, G. Corlett, R. Leigh, An intercomparison campaign of ground-based UV-Visible measurements of NO<sub>2</sub>, BrO, and OCIO slant columns. I. Methods of analysis and results for NO<sub>2</sub>, *J. Geophys. Res.*, 110, D08305, doi:10.1029/2004JD005423, 2005.