Dear referee #1. Thank you for the comments.

Here are answers (marked by R.) to specific comments

1. See answer to point 16

2. *"I recommend to briefly describe the strengths and the weaknesses of the Umkehr method (including the kernels) and a brief discussion how the accuracy of Umkehr profiles are determined"*

R. The information about optimization of the UMK04 algorithm for trend analysis, retrieved profile vertical resolution, AK and the accuracy of profile retrievals for UMK04 algorithm are discussed in the previously published paper by I.Petropavlovskikh et al (2005).
Petropavlovskikh, 5 I., Bhartia, P. K., and DeLuisi, J.: New Umkehr ozone profile retrieval algorithm optimized for climatological studies", Geophys. Res. Lett., 32, L16808, doi:10.1029/2005GL023323, 2005.
The current paper also indicates that effect of the stray light is insignificant in the long-term trend analysis.

Abstract

3. The Abstract should not only contain what has been done but it also needs to cover the key findings i.e. that the out-of-band stray light effect is not properly treated in the presently used Umkehr retrieval algorithms leading to falsifications of the ozone profiles of up to 7% in the upper stratosphere which has a substantially larger impact on the retrieved ozone profiles than the considered change in ozone absorption cross sections and the presently ignored effect of temperature dependence of ozone absorption cross sections.

R. The paper discusses the effect of the OOB on the Umkehr retrieved errors. The magnitude of the OOB errors depends on the instrument characteristics, quality of measurements and local ozone climatology. Therefore, there is no one number to go by. The paper concludes that further investigation is required for individual instruments.

The updated abstract will read:
"It is found that the ozone cross-section choice only minimally (within the retrieval accuracy) affects the Dobson and the Brewer Umkehr retrievals. However, significantly larger errors were found in the MLO and Boulder Umkehr ozone data (-8 and +5 % bias in stratosphere and troposphere respectively) when the out-of-band (OOB) stray light contribution to the Umkehr measurement is not taken into account (not included in the UMK04). The vertical distribution of OOB effect in the retrieved profile can be related to the local ozone climatology, instrument degradation, and optical characteristics of the instrument. However unvarying OOB errors do not contribute to the long-term ozone trends"

Introduction.

4. Line 2 ff, p. 2009: "The Dobson network has been collecting . . . since the 1930s (Dobson 1926). . . .. It has proven to be very stable ...". I think this sentence is misleading as it suggests that "the" Dobson network continuously existed since the 1930. The concept of the Dobson network design in its present form with a primary Dobson instrument goes back to the early 1970s and it was gradually extended to a global network since the second part of the 1970s. Indeed, Dobson put several instruments first in a European and thereafter in a "quasi-global" *"network" for some years, and a few stations continued operation over many decades and instruments on more sites started continuous measurements since around IGY but I don't believe that it is correct to use the term "the Dobson network" for these earlier measurements in connection with the Dobson network in its present form which started to exist only in the 1970s.*

R. The Dobson ozone network has been collecting direct-sun irradiance and zenith sky radiance data since the 1930s (Dobson, 1926). It has proven to be very stable and well maintained network capable of long-term ozone measurements that relies on regular calibration and intercomparison activities within the WMO GAW global ozone observing network.

The reviewer is correct. The network of Dobson designed instruments was sporadic prior to the 1950s, and over the years evolved into the system we have today. We change that statement to read:

"The record of stations using Dobson instruments in the WOUDC starts in 1924 with one station reporting total ozone values based on measurements of direct-sun irradiance and zenith sky radiance (Dobson, 1926). The Dobson network is defined as arrangement between stations reporting information to the common archive while using a common set of guidelines (Dobson, 1957). It has developed from a small set of mostly European sites to an approximately 60 stations worldwide by 1962. That year, a world reference Dobson instrument was defined (Komhyr, private communication 2011). It has proven to be very stable and well maintained network capable of long-term ozone measurements that relies on regular calibration and intercomparison activities within the WMO GAW global ozone observing network."

5. *Line 19 ff, p. 2009: What means: "Therefore regular calibration of the instruments . . . as is done in the Dobson network" ? I thought that Brewer instruments are supposed to be intercompared with travelling Brewer standard instruments every two years which is more often than in the Dobson network.*

R. To our knowledge, Brewer network instruments are calibrated either by comparisons with the traveling standard (by providing the reference parameter that is derived by comparing ozone derived from the traveling standard measurements and the station's to force two ozone values to be the same) or through the "Langley" method (Dobson, 1962; Grobner and Kerr, 2001. Kerr, 2002).The two methods are independent of each other in the way that we use them. The internal reference lamp produces the R6 correction, which we produce level 200 data with. The R6 correction keeps us tied to the original Brewer 017 traveling standard from IOS. Using the Langley technique we derive the ETC for ozone for each Brewer and produce the level 300 data. The Langley method is not tied to the original IOS calibration. The WMO recommended frequency of calibration is 2 year for Brewer network and 4 years for Dobson, while Dobson regional standards should be calibrated every 3 years.

We changed the text to:
"Therefore, regular calibration of the Brewer instruments in the network and post-correction of measurements is done every 2 years (Early et al., 1998; Lantz et al., 2002; Kimlin et al., 2005; GAW, 2007). The NEUBrew operational network (http://www.esrl.noaa.gov/gmd/grad/neubrew) has relied on the original calibration against the Brewer 017 traveling standard in 2006, with the continuous data corrections applied based on the internal reference lamp or Langley plots (description of procedures can be found at the website). The Dobson network carries out intercomparisons and calibration of network instruments once in 4 years, while regional standards are calibrated once in 3 years."

Dobson, G. M. B., and C. W. B. Normand, Determination of the constants etc. used in the calculation of the amount of ozone from spectrophotometer measurements and of the accuracy of the results, Annals of the International Geophysical Year., XVI, Part II, 161-191, Pergamon Press, 1962.

Grobner J. and J. B. Kerr, "Ground-based determination of the spectral ultraviolet extraterrestrial solar irradiance: Providing a link between space-based and ground-based solar UV measurements," *J. Geophys. Res. 106, D7, 7211-7217*, 2001

Kerr, J.B., "New methodology for deriving total ozone and other atmospheric variables from Brewer spectrophotometer direct sun spectra.," *J. Geophys. Res., 107, D23, 4731, doi:10.1029/2001JD001227*, 2002.

6. Line 13, p. 2010: "You might add that prior to the satellite era regular Umkehr observations from a few sites were the only source of information for ozone in the upper stratosphere."

R. Thank you for suggestion, we will include this sentence in the text

7. Line 16, p. 2010: Did BUV instrument provide reliable ozone profile information ?

R. According to the information available on the NASA/Goddard MOD web page (http://acdb-ext.gsfc.nasa.gov/Data_services/merged/mod_data.public.html) and paper published by G. Reinsel in 1982, the BUV was compared against 36 ground-based Dobson measurements for total ozone column product. While the offset of about 5 DU (10 DU by Reinsel, 1982) was determined, no calibration adjustment to Nimbus 4 measurements was done largely due to significant standard deviation of comparisons (7.8 DU), while a drift of ~ -1.53 DU per year was detected relative to the Dobson stations between 1970 and 1976. Nimbus 4 ozone profiles from 1970-1972 time period are used to extend the dataset of the BUV merged ozone profile information back in time. According to private communications with Stacey Frith (July 2011, NASA/Goddard), stratospheric ozone trends derived from the SBUV satellite data (1979-2008) and extended back to 1970s seem to intercept the Nimbus 4 data within 10 % uncertainty of trend analysis.

Reinsel, G., G. C. Tiao, R. Lewis, 1982: A Statistical Analysis of Total Ozone Data from the Nimbus-4 BUV Satellite Experiment. *J. Atmos. Sci.*, **39**, 418–430. doi: 10.1175/1520-0469(1982)039

8. Line 27, p. 2010: The sentence "The Integrated Global . . .." is not precise: The committee "ACSO" (Absorption Cross Sections of Ozone) is a joint ad hoc commission of the World Meteorological Organization (WMO) (Scientific Advisory Group (SAG) of the Global Atmosphere Watch (GAW) and IGACO-O3/UV and the International Ozone Commission IO3C) of the International Association of Meteorology and Atmospheric Sciences (IAMAS).

R. The text was changed according to suggestion.

Background

9. Line 13, p. 2012: Fig 3: The sequence of the numbers of the Figures in the text dictates the numbering of the Figures in the manuscript, i.e. Fig. number 3 should become number 1.

R. The Figure numbering is changed. Also additional description is added for new Figures 2 and 3. "Figure 2 shows the spectral dependence of the linear terms of the temperature fit for the bandpass-weighted BP and BDM ozone cross-section dataset centered at 311.2 nm Dobson short-wavelength channel. Figure 3 show spectral dependence of the linear (a) and quadratic (b) terms of the temperature fit over the Dobson's longer wavelength bandpass centered at 332.3 nm. Both figures denote spectral differences between the BP and BDM data sets."

10. Line 4, p. 2013: Some remarks concerning vertical kernel sensitivity might be appropriate here.

R.The figure representing the AK for Dobson and Brewer is added to show the sensitivity of the retrieval to ozone vertical information. AK plot shows that the information in the layer is sensed at 60-70 %, while the rest of the information comes from the adjacent layers. . It also explains how the algorithm applies vertical smoothing to the retrieved ozone profile.

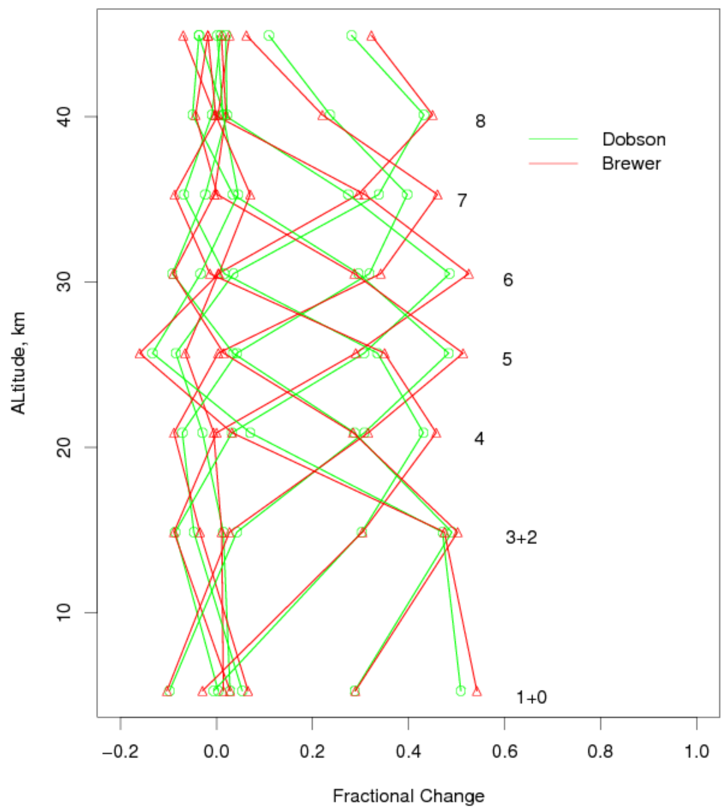


Figure 2. Averaging kernels for Dobson (green) and Brewer (red) are shown as fractional change vs. altitude. Corresponding Umkehr layers are indicated at the maximum of the AK.

Stray-light contribution in Dobson and Brewer Umkehr measurements

11. Figure 4: Legend: Could you add for clarity for which slit these measurements were performed

R. The measurements of the stray light were performed in the UV scanning mode of the instrument. Slit 1 was used for spectral range of 286.5 – 325 nm and slit 4 was used in spectral measurements between 325.5 and 363 nm. The plot in Figure 4 shows measurements done on slit 1.

12. Line 27. P. 2018: You might use the full name instead of TMTF in order to avoid confusing the reader with too many acronyms.

R. Done, Table Mountain (CO) Test Facility

Effects on Dobson and Brewer Umkehr ozone profile retrieval

13. Line 11-14, p. 2020: I suggest to quickly summarize here how the accuracy of Dobson and Brewer Umkehr ozone profiles is defined and determined and please give available numbers. The given information is very vague.

R. The errors of the retrievals are determined based on the formula provided in Rodgers (2000, Eq. 3.17 and Eq.3.19) for the maximum likelihood estimation statistical retrieval, which is based on the measurement uncertainty (known from the instrument characterization) and the profile smoothing technique (vertical resolution based on AK). Typical errors of the Dobson Umkehr ozone profile retrieval are provided in Figure 2 of Petropavlovskikh et al. (2005). It is possible to include the additional figure, see at the end of the responses to the comments) where the smoothing errors in the plot are estimated based on the climatological variability of ozone derived from the subset of the SAGE and sounding ozone profiles taken in the month of January between 1988 and 2002 at the Northern middle latitudes (McPeters, Labow and Logan, 2007).

McPeters, R. D., G. J. Labow, and J. A. Logan (2007), Ozone climatological profiles for satellite retrieval algorithms, J. Geophys. Res., 112, D05308, doi:10.1029/2005JD006823.

14. Line 14, p. 2020: To which of the two 2 references of Petropavloskikh et al., 2005 is referred to ?

R.The first one. The 2005a and 2005b are added to the reference list.

Discussion and conclusions

15. First paragraph: I don't see a clear connection of the first paragraph of p. 2022 to the other part of the conclusions. It is not clear whether the list of given publications aims to be complete. If this is not intended the papers should be declared as examples

R. The first paragraph provides examples of the previous research regarding sensitivity of ozone profile retrieval to the choice of the ozone cross-section sensitivity in the satellite retrieved ozone products. It is related to the introduction where it was indicated that the search for the better suited ozone cross-section dataset was initiated by the satellite community. Authors will be happy to include references to all missing papers.

16. Line 25, ff on page 2022: Since the paper basically only demonstrates a problem of the Umkehr method I expect more concrete information how important this problem is and how the method will be developed (compare comment 1). How large is the stray light effect in the context of the uncertainty of the method ? (Please extend the information of the accuracy of the Umkehr method given in line Line 11-14, p. 2020 and discuss it in the section Discussion and Conclusions.)

R. Figure 6 and 7 provide estimates of the cross-section, temperature and stray light effect on Dobson and Brewer Umkehr ozone profile retrievals. It is stated that the errors associated with the cross-section choice are relatively small as compared to the uncertainty of the retrieval (primary associated with measurement and vertical smoothing errors – see above).

Is it feasible to document the effect of the stray light problem by comparison with measurements of collocated instruments such as microwave or LIDAR ?

R. It takes at least a year worth of co-located and coincident measurements to develop such a correction. However, it will be an empirical approach to determine a real instrument parameter, and would depend on stability and accuracy of the referenced instrument such as microwave or lidar. Reference to the recently published paper by Nair et al. (2011) has been added to the list of references. The following is added to the text.
"Recently published paper by Nair et al (2011) compares Dobson Umkehr, lidar, ozone balloon and several satellite ozone profiles over the Haute–Provence Observatory ground station in France. Authors describe biases between co-incident Umkehr and lidar data that are similar to the above discussed errors associated with the stray light contribution to the Dobson measurements."

Or is the uncertainty in retrieved ozone profile from Dobson/Brewer instruments expected to be that large, that a refined Umkehr algorithm taking into account the stray light effects has lower priority ?

R. The work to determine stray light effects in the NOAA Dobson instruments is underway. Once the instrument is characterized, the forward model of the retrieval can be corrected by applying the measured band-pass function for individual instrument.
Although the stray light creates a bias in the derived ozone profile, it has no significant long-term change and thus does not affect the trend analysis of ozone decline and recovery, which is the main advantage of continuous long-term Umkehr ozone time series.

I furthermore miss a more clear elaboration of the further development of the retrieval; "The method is under development . . ..." (line 26, p. 2022) is not sufficient to me. Which method do you mean ? I thought the data used in this paper are sufficient to characterize stray light effects. If this is not the case this would imply that you need to declare that the used measurements are insufficient to provide enough information to describe the stray light effect in an appropriate way.

R. Further development means determining a feasible "cheap" method for optical characterization of individual instruments. It can rely on the spectrophotometer with high spectral resolution and the system built to "map" the band-pass of the Dobson, and then determine the stray light contribution based on intercomparisons with other systems that have low stray light characteristics. At this moment the resources are not available to directly measure the stray light in the Dobson instruments. Brewer instrument characterization is more easily done and the methods exist.

Or do you mean that the introduction of such measurements in a systematic way in the Dobson/Brewer network needs more work ?

R. Yes.

Should the measurement of the slit function become part of the regular Dobson intercomparisons?

R. This is under evaluation. It could be useful for the proper modeling of instrumental effects and can be used for adjustments of the "effective" (weighted by slit function) ozone cross-sections used in the total ozone calculations. It is also useful for instrument intercomparisons. However, if the slit function is unchanged - it does not affect trends, which is the main rationale for the use of the Umkehr data.

How do you plan the further development of the Umkehr retrieval algorithm for Dobson instruments? What are the steps to be planned for the retrieval developments of Brewer instruments?

R. There is a work in progress to assess similarity in the stray light response of various Dobson instruments. Once this is done, there could be an update on the UMK04 algorithm to include stray light correction.
There is also a greater need for further development, optimization and validation of the Brewer Umkehr algorithm. Although Umkehr measurements are already regularly taken by some instruments, there is no requirement for the archiving of the raw Brewer data, although there is a WMO and WOUDC organized campaign to encourage such submissions. Although the Brewer single pair ozone profile retrieval algorithm is readily available for common data processing, still the Umkehr measurements are often not processed and therefore are not methodically archived. Part of the problem is conflicting schedules of different types of measurements taken by Brewers that are dedicated to either UV or ozone programs. Under the grant from the NASA ROSES 2008 Atmospheric composition program, the author is working to develop algorithm for the Brewer Umkehr data analysis that utilizes multiple spectral channels and requires smaller range of solar zenith angles for routine observations, which removes time limitations for data collection and can be widely used for routine ozone profile retrieval in addition to UV scans.

Are further comparisons with of other measurement methods are intended (groundbased or satellite instruments)?

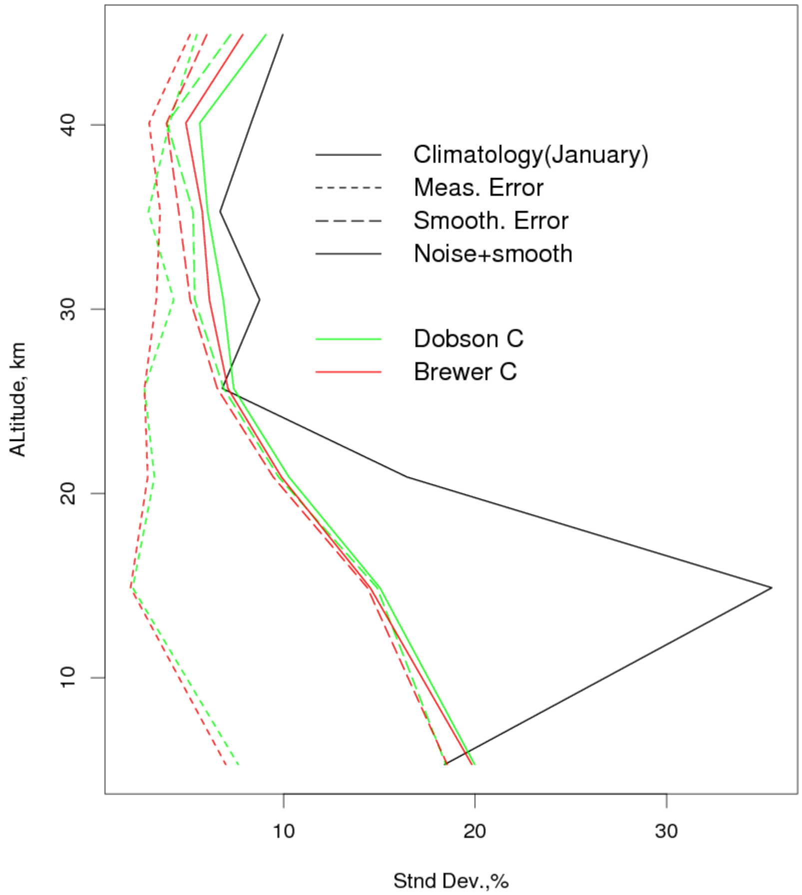
R. Yes, comparisons with other measurements (ozone sonds, LIDAR, microwave and satellite instruments) are done routinely during Dobson Umkehr intercomparisons, stations with co-incident measurements by different systems, and reports are typically published under the WMO program. There is a recent paper by Nair at al. (2011) that summarizes results of the long-term ozone measurement comparisons at the OHP, France ground-based station.

Nair, P. J., Godin-Beekmann, S., Pazmiño, A., Hauchecorne, A., Ancellet, G., Petropavlovskikh, I., Flynn, L. E., and Froidevaux, L.: Coherence of long-term stratospheric ozone vertical distribution time series used for the study of ozone recovery at a northern mid-latitude station, Atmos. Chem. Phys., 11, 4957-4975, doi:10.5194/acp-11-4957-2011, 2011.

Should one also consider to use specific temperature profiles instead of averaged temperature profiles (as used in the sensitivity analysis).

R. This could be not very practical for the historical datasets, or for the locations where the temperature profiles are not regularly available. Again, the study in this paper suggests that the temperature sensitivity is less than uncertainty of the retrieval…

Additional figure.



The errors of ozone retrieval are shown as function of altitude for Dobson (green solid line) and Brewer (red solid line). The standard deviation error is calculated relative to January middle latitude climatological ozone profile (MLL). Climatological ozone variability is shown as black line and is taken from the MLL climatology (McPeters et al, 2007) for Northern middle latitude (40-45 N) and for the month of January. The original climatological profiles and variance were converted into Umkehr layers for comparisons. Also shown separately are errors associated with the measurement (dotted) and smoothing (dashed) uncertainties that contribute to the ozone profile retrieval error. The covariance matrix used in the smoothing error calculations of the Umkehr retrieval is representative of the ozone variability over the middle Northern latitudes based on observations by SAGE II (ozone ~above 25 km) satellite instrument and ozone sonde (ozone below ~25 km) profiles collected between 1988 and 2002.

New Abstract

Remote sounding methods are used to derive ozone profile and column information from various ground-based and satellite measurements. Vertical ozone profiles measured in Dobson units (DU) are currently retrieved based on laboratory measurements the ozone absorption cross-section spectrum between 270 and 400nm published in 1985 by Bass and Paur (BP). Recently, the US National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA) proposed using the set of ozone cross-section measurements made at the Daumont laboratory in 1992 (BDM) for revising the Aura Ozone Monitoring Instrument (OMI) and Global Ozone Monitoring Experiment (GOME) satellite ozone profiles and total ozone column retrievals. Dobson Umkehr zenith sky data have been collected by NOAA ground-based stations at Boulder, CO (BDR) and Mauna Loa Observatory, HI (MLO) since the 1980s.The Umkehr retrieval algorithm is based on the BP ozone cross-section data. It is currently used for all Dobson Umkehr data processing submitted to the World Ozone and Ultraviolet radiation Data Centre (WOUDC) under the Global Atmosphere Watch (GAW) program of the World Meteorological Organization (WMO). Ozone profiles are also retrieved from measurements by the Mark IV Brewers operated by the NOAA-EPA Brewer Spectrophotometer UV and Ozone Network (NEUBrew) using a modified UMK04 algorithm (O3BUmkehr v.2.6, Martin Stanek). Records from Dobson and Brewer instruments located at MLO and BDR were used to produce Umkehr ozone retrievals using BDM ozone cross-sections and compared to profiles produced using the BP ozone cross sections. It is found that the ozone cross-section choice only minimally (within the retrieval accuracy) affects the Dobson and the Brewer Umkehr retrievals. However, significantly larger errors (-8 % in stratosphere and +5 % in troposphere) are found when the out-of-band (OOB) stray light contribution to the Umkehr measurement is not taken into account (not included in the UMK04). The vertical distribution of OOB effect in the retrieved profile can be related to the local ozone climatology, instrument replacement, and optical characteristics of the instrument. However unvarying OOB errors do not affect analysis of the long-term ozone trends.

Response to the comment by Referee # 2.
Thank you very much for reading the manuscript and finding it to be a significant contribution to understanding of Umkehr ozone profile retrievals.

Q. You have asked to "clarify the meaning of the equation found at the bottom of page 2011. Presumably the symbol F is to represent the signal detected by the spectrophotometer. But the relationship between sky brightness and the symbol ETC is not clear."

R. Equation (1) is used to explain the Umkehr measurement in generalized form for both Dobson and Brewer, although it does not actually represent the physics of the measurement.

$$N_{SZA} \equiv \log(F_{SZA}^{L_2} / F_{SZA}^{L_1}) - \log(F_0^{L_2} / F_0^{L_1})$$

The F in this generalized equation represents the signal (intensity of the zenith skylight) detected by the instrument at individual channel (Brewer) or relative difference in signals detected by Dobson in the pair of spectral channels (C-pair). The actual measurement can be a photon count (Brewer) or position of the R-dial (Dobson) that are then converted to the atmospheric radiances through calibration procedures. The equation that describes F $F_{SZA} = I_{SZA} * K * ETC$ is used to simulate the measured radiance, where I_{SZA} is zenith sky radiance normalized to the top of the atmosphere (attenuation of the scattered light through the atmosphere), K is instrumental constant or instrument transfer function, and ETC is the extra-terrestrial constant (obtained through Langley type tests). The extra terrestrial solar flux measured by other instrument (for example, Atlas-1, 2 or 3 mission used SUSIM instrument to observe sun in 1992, 1993 and 1994 respectively during the space shuttle mission, Van Hoosier, 1996) can be used to do radiance simulations.

Since neither instrumental constant K or ETC are not known for Umkehr measurements, the retrieval relies on subtraction of the nominal measurement (F_o represents Umkehr measurement at the smallest SZA) from the other Umkehr measurements (at different SZA), which effectively removes the K and ETC unknowns.

VanHoosier, M.E., Solar ultraviolet spectral irradiance with increased wavelength and irradiance accuracy, SPIE Proceedings 2831, 57-64 (1996).