Atmos. Chem. Phys. Discuss., 2, S426–S436, 2002
www.atmos-chem-phys.org/acpd/2/S426/


Atmospheric Chemistry and Physics Discussions



# *Interactive comment on* "Application of the Spectral Structure Parameterization technique: retrieval of total water vapor columns from GOME" *by* R. Lang et al.

**Anonymous Referee #2**

Received and published: 25 September 2002

## 1. General Comments

Lang *et al.* present an method for water vapor column (WVC) retrieval based on a Spectral Structure Parameterization (SSP) technique. SSP constitutes a faster (by a factor of 4–5), but less accurate implementation of the Optical Absorption Coefficient Spectroscopy (OACS) method by the same authors.

The advantage in retrieval speed gained by SSP over other types of WVC retrieval makes SSP an interesting alternative, and it is well worth being published in a scientific journal. ACP is an appropriate forum for this paper. In its current form, however, the manuscript cannot be accepted for publication. Substantial revisions, followed by another round of reviews, are required before the manuscript can be considered for publication.

On the positive side, The theoretical basis of the SSP method is well presented and concisely written. Besides some clarifications on the dimension of the $w$ parameter and the effect of off-nadir viewing geometries, this part only requires minor revisions that are mostly technical in nature.

The sections dealing with bias adjustment and application to GOME data, on the other hand, require substantial additions and rewriting for both, textual and scientific clarification. A particularly important omission is the determination of fitting uncertainties. These should not be confused with the difference between model input and retrieval results, which are given in the manuscript. An further important addition are the correlations between the parameters derived in the fitting – most notably those between the WVC, the albedo, and the multiple scattering correction. Finally, no quality assessment of the retrieved WVC is provided. Comparisons to ECMWF observations are limited to hand-waving "chi by eye" comparisons, and a (by itself) not very meaningful linear fit to pairs of (ECMWF, SSP) data tuples.

A number of scientific and technical issues are detailed in the two sections below. After the authors have addressed these questions, the manuscript should be subjected for a second round of reviews.

## 2. Specific Comments

1. **Fitting uncertainties:** My major point of criticism is that no uncertainties for the retrieved WVC and associated parameters – most notably the surface albedo and the multiple scattering correction – have been given. SSP uses a MATLAB implementation of a non-linear, large-scale trust region method for the fitting of reflectivity spectra to GOME measurements. This software package should provide estimates of parameter uncertainties, or at least provide the covariance matrix of the fitting problem, from which the uncertainties can be inferred. Furthermore, the off-diagonal elements of the covariance matrix contain information on the correlations of the different parameters determined in the fit. This should provide valuable insights into the possible interference of surface albedo and multiple scattering correction retrieval – a problem identified by the authors as a source of uncertainty in WVC retrieval with SSP.

   If fitting uncertainties from the MATLAB routines are unavailable, a sensitivity analysis can be performed through small variations in the input parameters that are used in the



ACPD
2, S426–S436, 2002

LBL computations, followed by finite differencing of the retrieval results. WVC, surface albedo and multiple scattering corrections are of particular importance here.

2. **Bias correction (Section 6):** The derivation of the bias correction needs to be clarified. Exactly how are the variables $A$–$D$ determined: Simultaneously with the WVC? Off line? The manuscript states (p 1108, l 17 ff)

   > In order to make a direct comparison, realistic values fro the variables $A$ to $D$ were obtained from the best fit of the forward-lbl modeling results of GOME measurements taken at corresponding geolocations.

   From this it is impossible to infer how the parameters $A$–$D$ influence the WVC retrieval, and, ultimately, how the bias correction has been derived.
3. **SSP–OACS comparison:** More emphasis needs to be placed on the numerical comparison of WVC retrieval results from the two methods. Currently, this comparison is limited to a very busy plot

   (Figure 6), from which no real understanding of the relative performance of the two methods can be gained. A scatterplot of the type of Figure 9 and/or a table including a summary of both retrievals should be added, with uncertainties and backed up by a proper assessment in the main text.
4. **Dimension of $w$ parameter (Section 2)** What is the physical dimension of $w$? From its definition in Equation (4), its use in Equation (2) and the definition of $< \tau >$ it must be dimensionless. But Figure 2(a) lists it with dimension $\Delta\lambda^{-1}$, and it is introduced in the text as "$w$ (...) in units of $\delta\lambda^{-1}$".
5. **Path-length factor (Section 3, p 1103):** The geometric path length factor $\tilde{\mu}$ is defined as $(\mu_0^{-1} + 1)$, which holds strictly only for direct nadir viewing. This poses the following questions:

   - Has SSP retrieval only be performed for GOME center pixels?
   - Can the method be applied to non-center pixels? What are the required modifications?
   - Due to the large size of a GOME footprint, the radiances from the edges of the ground pixel are observed under line of sight angles of $\pm 10.3°$. Has this been accounted for in the path length correction? How large is the effect?
   - Later in the manuscript, SSP retrievals are said to include "all GOME pixels", and the number of GOME pixels considered for a 3-day period is larger than can be explained by GOME center pixels alone. Is there a contradiction?



**ACPD**

2, S426–S436, 2002



Full Screen / Esc

Print Version

Interactive Discussion

Original Paper

6. **Surface reflectivity (p 1102, l 26):** Clarify what is meant by "$R_{surf,j}$ is the reflectivity of the direct light path reflected at the earth surface". How does $R_{surf,j}$ relate to the surface albedo $\Lambda$, which appears as a multiplicative factor to $R_{surf,j}$ in Equation (5)?
7. **Use of HITRAN '96:** Why has HITRAN '96 been chosen over the more recent HITRAN 2000, which contains updates to the water vapor line parameters? The comparison between Learner *et al.* (2000) and HITRAN should be made with extreme care: The Learner *et al.* measurements are far from undisputed, and surely not the final answer in water vapor line measurements. Thus, any difference between Learner *et al.* and HITRAN '96/2000 does not automatically mean a problem in HITRAN. This is not to say that HITRAN is right – as Dr. Rothman is probably one of the first to admit.
8. **GOME $F_0$ measurements:** The frequency of the $F_0$ measurements is closer to "once per day" than "once per orbit".
9. **Profile constraints (p 1107, l 17 ff):** Rather than just saying "low", "medium" and "high" values for the profile constraints, the oder of magnitude of the WVC subcolumn should be specified. This will provide a better idea of what the constraints are.
10. **Section 8:** This section requires substantial rewriting. Here are some points to note.

    Paragraph 1:

    - None of the symbols stated in the text correspond to the ones plotted in Figure 6.
    - Stating the corner lat/lon coordinates for a GOME orbit provides little information on where the orbit occurred on the globe; a schematic map with the outline of the orbit serves this purpose much better. Either include a map or delete all references to the corner coordinates.
    - Are the 476 GOME ground pixel only "center" pixels?

    Paragraph 2:

    - A cloud cover that "significantly exceeds 10%" is not a useful quantification; when does the cloud cover become "significant"?
    - How does SSP, since it is a retrieval code, "model" enhanced photon path lengths?
    - High cloud cover fraction (from ICFA or otherwise) should not be confused with "thick clouds".
    - How often occurs the "occasional" cancellation of the two competing effects of clouds on WVC retrieval?



**ACPD**

2, S426–S436, 2002

Paragraph 3:

- Figure 6 is a very busy figure. In particular, it is virtually impossible to infer from it whether the bias correction improves the retrieval results (the main text is no help here, either). More details on the SSP–OACS comparison should be provided in form of a table, or with a scatter plot of the type of Figure 9.
- "datasets" should be replaced by "cases".

Paragraph 4:

- If the subcolumn profile retrieved by SSP influences the retrieval of WVC, how many retrievals in total are performed? First a subcolumn profile, followed by WVC? Clarify!
- Rephrase the sentence starting "The retrieval of an unrealistic shape of the retrieved subcolumn profile. . . "
- Unless the effect of albedo and multiple scattering on WVC retrieval are mutually exclusive, change "or" in p 1112, l 10 to "and/or".
- "very well correlated" should be replaced by "strongly correlated".
- Are the "upper fit constraints" those for the upper atmospheric layer mentioned above? If so, state clearly.
- Delete "treatment of the" in line 15.

Figure 7:

- The figure caption states "An underestimation of the surface albedo leads to an overestimation in the retrieved WVC". I don't see any evidence for this in the figure - clarify or delete.
- Also in the figure caption: Does "including all ground pixels" mean east, center, west, and back-scan pixels?

11. **Section 9:**

- How does the number "$47,000$" for the used GOME ground pixels compute? At 3 days of 14 orbits each, and approx. 2000 ground pixels between $\pm 70^{\circ}$ latitude, I would expect either $\sim 84,000$ (if all 4 types are included), or $\sim 21,000$ (if only center pixels are included).
- What are "globally corrected results"?



**ACPD**

2, S426–S436, 2002

- Cloud fractions of less than 10% are not “cloud free”!
- Koelemeier *et al.* (1999) is not a proper reference for either ICFA or the ISCCP. Instead, for ICFA use

  A. Kuze and K.V. Chance, Analysis of Cloud Top Height and Cloud Coverage from Satellites Using the $O_2$ $A$ and $B$ Bands, Journal of Geophysical Research, 99, 14,481–14,491 (1994);

  and for ISCCP use, for example,

  W.B. Rossow and R.A. Schiffer, ISCCP Cloud Data Products, Bulletin of the American Meteorological Society, 72, 2–20 (1991).

12. **Figure 8:**

- Data gaps in panel (c), “WVC for cloud fractions $<$ 10%”, are supposed to originate either from “calibration periods”or “processing failures”. No definition for the latter are be given. I would expect the constraint of “$<$10% cloud fraction” to account for most of the gaps; (rejected) cases of cloud fraction $\geq$ 10% should not be classified as “processing failures”.
- How exactly is the “interpolation of narrow swath orbits to the standard grid” performed? I would expect that extrapolation, rather than interpolation, is required to convert a narrow swath GOME pixel of size 70$\times$40 km$^2$ (lon$\times$lat) to the standard footprint of 320$\times$40 km$^2$. If global coverage over a 3-day period is desired, why was this observation period chosen in the first place?
- What is the correlation coefficient (*Pearson's r*) between the ECMWF and SSP data (panels (a) and (b))? This should be computed for all longitudes and latitudes, as well as separately for the different latitude regions that are differentiated by the “upper fitting constraints”.

13. **Figure 9:**

- As stated in the figure caption , the “gradient of 0.98” has been be determined from a linear fit to the (ECMWF, SSP) data tuples. By itself, this contains virtually no information, especially since not even the uncertainties of the gradient are given.
- What is the correlation coefficient (*Pearson's r*) of the two data sets, what is the fitting uncertainty for the gradient, and what does that mean for the quality of the SSP retrieval?



**ACPD**

2, S426–S436, 2002

14. **Section 10, last paragraph:**

    - A systematic error can be present for all cases, not just for those with cloud fractions less than 10%.
    - The 18% difference between ICFA and ATSR-2 are the MEAN difference, NOT the absolute one. Absolute difference in cloud fraction can be up to 100%, naturally.

15. **Section 11, fourth paragraph:**

    - The larger spatial coverage (NOT "higher spatial resolution") of the continuous GOME measurements compared to ground-based ECMWF is an advantage of the GOME instrument. It thus pertains to any WVC retrieval, not just SSP.
    - How can the ground-based ECMWF observations that are, as the authors state, "provided for a specific time and day", and results from a 3-day period of GOME observations, be "correlated in time and geolocation (space?)"?
    - What exactly does a "good correlation between the two data sets of correlated cloud-free ocean pixel" mean?

# 3. Technical Corrections

## 3.1. General comments:

1. In many places the understanding of the manuscript in hindered by long and intricate sentences; those should be rephrased and/or broken up into shorter ones
2. The authors should check the proper usage of "utilized" and "respectively". While "utilized" should be be replaced with "used" in almost all occurrences in the manuscript, "respectively" would best be abandoned for a direct reference. For example, a sentence like (p 1111, l 1)

   > After application of the empirical correction (Table 2), the error due to multiple scattering for the maritime and rural case is less than 6 and 2%, respectively (clear sky cases).

   becomes



ACPD
2, S426–S436, 2002

After application of the empirical correction (Table 2), the error due to multiple scattering for the clear sky cases is less than 6% (maritime) and 2% (rural).

It is left to the authors to check for occurrences of “utilize” and “respectively”.

3. The usage of “optical thickness”, “optical depth” and “optical density” should be clarified:
   - **optical thickness** is a dimensionless quantity, and it is used whenever reference to the total optical property of a finite layer of a medium (aerosol, cloud, *etc.*) is made. For example: An aerosol layer of (total) optical thickness $\tau$.
   - **optical density** is also a dimensionless quantity; it often used to replace altitude as the independent variable in radiative transfer problems. It is a measure of “how deep” one has propagated into an optical medium.
   - **optical density** is a local quantity of a medium and usually carries a dimension; it is used synonymously to “extinction coefficient”.

   Again, it is left to the authors to make the necessary corrections.

## 3.2. Typographical errors, textual changes and clarifications

1. **p 1099, l 10** SCIAMACHY expands to “SCanning Imaging Absorption spectroMeter for Atmospheric CHartorgraphY”.
2. **p 1100, l 9** “profile” becomes “profiles”.
3. **p 1100, l 12** Change “nadir satellite” to “satellite-based, nadir viewing”.
4. **p 1100, l 14** “. . . nonhomogenous atmospheres, including light paths where photons undergo a single scattering event”. The sentence reads as if single scattering is not included in the treatment of homogeneous atmospheres. Is this the intended meaning?
5. **p 1100, l 21** Delete “both”.
6. **p 1100, l 24** If possible, a reference for the WVC from ECMWF should be included.
7. **p 1101, l 8** “, therefore,” becomes “and”.
8. **p 1102, l 3** “18 **homogeneous** levels”.
9. **p 1102, l 4** What is meant by “For real measurements”?
10. **p 1102, l 13** What kind of pixel is “pixel-averaged”?



**ACPD**

2, S426–S436, 2002



Full Screen / Esc

Print Version

Interactive Discussion

Original Paper

11. **p 1103, l 24** Does “low surface albedo of 0.03 and 0.1” constitute part of the “worst case scenario” described in this sentence?
12. **p 1104, l 4** The parameters $A, B, C, D$ have not been introduced at this point – they appear first in Equation (10). Either move Eq. (10) or include a reference to it at this point.
13. **p 1104, l 15** What is meant by “a realization of the cross section $\sigma$”?
14. **p 1104, l 16** “HITRAN ’96 database **(Rothman *et al*, 1998)**”.
15. **p 1104, l 19** “, which” becomes “that”.
16. **p 1105, l 22** “these” becomes “the”.
17. **p 1105, l 26** Delete “rather”.
18. **p 1106, l 23** “which is” becomes “, which are”.
19. **p 1107, l 7** “consists out of” becomes “consists of”.
20. **p 1107, l 14** What is meant by “we scale the upper profile constraint differently”? If this means that subcolumn limits for the higher altitudes are chosen differently for different latitude regions, then say so.
21. **p 1107, l 19** “densities lower than 0.1**, where analytical derivation for $w$ and $S$ are a good approximation of the real value, and the non-linear nature of the absorption of individual lines per layer becomes weakened; and . . .** ”
22. **p 1107, l 24** “limit.” Delete rest of sentence.
23. **p 1108, l 5** “below 0.1, even for a high WVC of . . . ”
24. **p 1108, l 7** I don’t see the consequence between the sentence starting with “Consequently” and the one preceding it.
25. **p 1108, l 9** “with respect to” becomes “than”.
26. **p 1108, l 10** Why “real $w$ parameter **per layer**” when the sentence discusses $w_{max}$?
27. **p 1108, l 16** Delete “significantly different”; “by” becomes “from”.
28. **p 1108, l 22** Delete “respectively”.
29. **p 1108, l 25** “which” becomes “that”.
30. **p 1108, l 27** “corresponds roughly to latitudes $> 70^{\circ}$” becomes “is representative of solar zenith angles in GOME measurements taken at latitudes above $70^{\circ}$.”
31. **p 1109, l 1** Delete “instances of”.
32. **p 1109, l 3** “found” becomes “determined”.



**ACPD**

2, S426–S436, 2002

33. **p 1109, l 9** “Even in the absence of clouds . . . ” reads as if the aerosol induced error is larger if a cloud is present. Why?
34. **p 1109, l 11** Aerosols will not “affect the assumptions” made, but may invalidate them.
35. **p 1109, l 13** Delete “full”.
36. **p 1109, l 15** Does the DAM model compute all components of the Stokes vector but has been run in “intensity only” mode? If so, state clearly.
37. **p 1109, l 21** Either delete “solely” or clarify what is meant by it.
38. **p 1110, l 2** “pressure; aerosol optical properties are taken from . . . ”
39. **p 1110, l 8** “may be neglected” becomes “is negligible (see Figure 5(b)).”
40. **p 1110, l 15** Delete “down”.
41. **p 1110, l 16** “The first-order polynomial in $\lambda$. . . ”
42. **p 1113, l 9** “North-west Carolina” becomes “western North Carolina”.
43. **p 1113, l 25** “during instances” becomes “for cases”.
44. **p 1113, l 26** “uncertain contribution” becomes “uncertainty in the contribution”.
45. **p 1113, l 27** Specify the type of “aerosol profile” which is the critical parameter: vertical(?) distribution of the extinction coefficient? the particle size distribution? other?
46. **p 1114, l 9** “occasionally” becomes “from”.
47. **p 1114, l 16+19** “1110 and 685 nm” becomes “685 and 1110 nm”.
48. **p 1117, l 1** “nadir” becomes “nadir viewing”.
49. **p 1117, l 5+6** “surface albedo from global databases once available and tested”. At least two possible candidates come to mind: GOME derived surface albedo (Koelemeier?) and MODIS land products.
50. **p 1117, l 13** “in principle”. Either SSP it is suitable for profile retrieval or it isn’t.

### 3.3. Figures

1. Figure 2: Check dimensions of $w$ parameter in subpanel (a).
2. Figure 3: Change “in dependence” to “as a function of”; change “but than” to “the same quantities for”.



**ACPD**

2, S426–S436, 2002

3. Figure 5:Panels (a) and (b) should be interchanged to reflect the order of their reference in the caption; the y-axis caption becomes "Residual$\times 10^{22}$ $[molec/cm^2]$"; change the caption to "... differences between the WVC retrieved by SSP and as used in the lbl model for 26 sample cases ...".
4. Figure 6: The figure needs to be simplified or broken up into subfigures; change the caption to "(a) SSP ... (filled circles) ... (b) The cloud coverage ... GDP level-2 data.".
5. Figure 7: The y-axis label becomes "(OACS–SSP)/OACS $[\%]$"; the legend should be moved to avoid overlap with the plot lines.
6. Figure 8: "vapour" becomes "vapor" throughout; change "cloud free" in the title of panel (c) to "cloud fraction $<$ 10%".
7. Figure 9: In the caption, change "At 24 October" to "On 24 October".

### 3.4. References

1. The following references are out of alphabetical order: Buchwitz *et al.* (2000); Kato *et al.* (1999).
2. Kato *et al.* (2000): "... approximation for **a** shortwave ...".
3. Lacis *et al.* (1991): "mutliple" becomes "multiple". shortwave ...".
4. Maurellis *et al.* (2000b): "Retrievalof" becomes "Retrieval of"; "Highly-Structure" becomes "Highly-Structured".
5. Add the proper references for ICFA (*Kuze and Chance*, 1994) and ISCCP (*Rossow and Schiffer*, 1991) as given above.



**ACPD**

2, S426–S436, 2002