The Cluster Low-Streams Regression Method for Fast Computations of Top-of-the-Atmosphere Radiances in Absorption Bands *

Ana del Águila[1][0000−0001−9006−9631], Dmitry Efremenko[1][0000−0002−7449−5072]

[1]Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany
ana.delaguilaperez@dlr.de, dmitry.efremenko@dlr.de

Abstract. Atmospheric composition sensors provide a huge amount of data. A key component of trace gas retrieval algorithms are radiative transfer models (RTMs), which are used to simulate the spectral radiances in the absorption bands. Accurate RTMs based on line-by-line techniques are time-consuming. In this paper we analyze the efficiency of the cluster low-streams regression (CLSR) technique to accelerate computations in the absorption bands. The idea of the CLRS method is to use the fast two-stream RTM model in conjunction with the line-by-line model and then to refine the results by constructing the regression model between two- and multi-stream RTMs. The CLSR method is applied to the Hartley-Huggins, O_2 A-, water vapour and CO_2 bands for the clear sky and several aerosol types. The median error of the CLSR method is below 0.001 %, the interquartile range (IQR) is below 0.1 %, while the performance enhancement is two orders of magnitude.

Keywords: Radiative transfer model, Regression model, Line-by-line model

1 Introduction

The information about the atmospheric trace gases can be retrieved from the spectral radiances measured at the top of the atmosphere. The key component of atmospheric retrieval algorithms are the radiative transfer models (RTMs). Accurate simulations in the absorption bands are based on the so called line-by-line (LBL) model [1], which requires thousands of monochromatic RTM computations per absorption band due to strong spectral variability of the absorption coefficient. Alternatives to computationally expensive LBL models are the k-distribution method [2, 3] and the principal component analysis (PCA)-based RTMs [4–9], in which the redundancies in hyper-spectral data are eliminated and the spectrum can be computed by using a small number of RTM calls. These methods are reviewed in [10].

* Supported by the German Aerospace Center (DLR) and the German Academic Exchange Service (DAAD) through the programme DLR/DAAD Research Fellowship (57424731).

In our recent work [11], the Cluster Low-Streams Regression (CLSR) method was developed to accelerate hyper-spectral computations. The idea of the CLSR method is to perform LBL computations by using a fast two-stream RTM and then to refine results by using the correlation model for the two-stream and reference multi-stream RTMs. This approach was applied to the O_2A-band and the weak CO_2 band for different atmospheric scenarios. The results were compared with the PCA-based RTMs showing an improvement over the last in terms of accuracy. Note that the idea of improving accuracy of two-stream models was exploited in numerous theoretical studies (see e.g. [12, 13] and references therein).

In this study, the CLSR method is extended to ozone Hartley-Huggins band and the water vapour band in the ultra-violet and near infrared spectral ranges, respectively. Additionally, the CLSR method is applied to several atmospheric models containing different aerosol types.

2 Methodology

2.1 Data overview

We consider the computations of the reflected spectral radiances at the top of the atmosphere (TOA) in the Hartley-Huggins, O_2A-, water vapour and CO_2 bands. Table 1 summarizes the spectral bands examined with their corresponding spectral range, spectral resolution and number of spectral points to be simulated. As a reference RTM, we use the discrete ordinates with matrix exponential (DOME) method [14, 15]. The number of discrete ordinates (streams) in the polar hemisphere N_{do} regulates the computational performance and accuracy. In the following, the model is called multi-stream (MS) when $N_{\text{do}} \geq 2$ and low-stream (LS) otherwise. Following previous analysis in [16], the multi-stream RTM with $N_{\text{do}} = 32$ discrete ordinates is used as a reference RTM.

The gaseous absorption coefficients for the O_2A-, water vapour and CO_2 bands are computed with the LBL model Py4CAtS [17], while the ozone absorption cross-sections in the Hartley-Huggins band are taken from the HITRAN 2016 database [18]. Rayleigh scattering is modeled as proposed in [19].

Table 1. Spectral ranges, spectral resolution and number of spectral points for the absorption bands used in this study.

Band	Spectral range (nm)	Spectral resolution (nm)	Number of spectral points
Hartley-Huggins	280 — 335	0.0188	2932
O_2A	755 — 775	0.0010	20000
Water vapour	770 — 1000	0.0058	40000
CO_2	1590 — 1620	0.0015	20000

The atmosphere is discretized into 35 layers with a step of 1 km between 0 and 25 km, and a step of 2.5 km between 25 km and 50 km. For all the simulations, we

assume a Lambertian surface with an albedo of 0.3. The solar zenith angle, the viewing zenith angle and the relative azimuth angle are 45°, 35° and 90°, respectively.

The atmosphere can contain one of the aerosols from the OPAC database [20], optical properties of which are summarized in Table 2 and Table 3 .

Table 2. Aerosol optical thickness (AOD) in the middle of the spectral range for the spectral bands and aerosol types considered.

Aerosol type	Hartley-Huggins (315 nm)	O_2A (760 nm)	Water vapour (885 nm)	CO_2 (1610 nm)
Tropospheric	0.133	0.018	0.015	0.010
Continental clean	0.42	0.20	0.17	0.08
Urban	4.28	0.46	0.35	0.16
Desert	0.71	0.20	0.20	0.19
Continental polluted	2.4	1.2	0.9	0.4

Table 3. Single scattering albedo (SSA) for each spectral band and aerosol type.

Aerosol type	Hartley-Huggins	O_2A	Water vapour	CO_2
Tropospheric	0.950	0.947	0.942	0.936
Continental clean	0.959	0.962	0.959	0.958
Urban	0.940	0.935	0.929	0.913
Desert	0.932	0.953	0.950	0.945
Continental polluted	0.951	0.960	0.957	0.950

Figure 1 shows the radiance spectra computed by using the multi-stream and the low-stream (i.e. two-stream) RTMs for the different absorption bands. The computations are perfomed using the LBL-framework. Note that both spectra have a similar spectral behavior. Hence, it is possible to establish a regression model between the low- and multi-stream radiances, which is a subject to the CLSR method considered further.

2.2 Cluster Low-Streams Regression (CLSR) method

The Cluster Low-Streams Regression (CLSR) method is described in detail in [11] and can be summarized as follows:

First, let us consider a LBL spectrum $\{I_{\mathrm{LS}}(\lambda_i)\}_{i=1}^N$ computed at N spectral points $\{\lambda_i\}_{i=1}^N$ by means of a low-stream RTM. We sort the radiances in ascending order and consider C clusters in the sorted radiance set (with $N_C = N/C$ radiance points). Secondly, we select n equidistant radiance points in each cluster and we compute the multi-stream radiances for the corresponding wavelengths $\{\overline{I}^c_{\mathrm{MS},q}\}_{q=1}^n$. Assuming that in each cluster c there is a linear relationship between low- and multi-stream radiances, we get

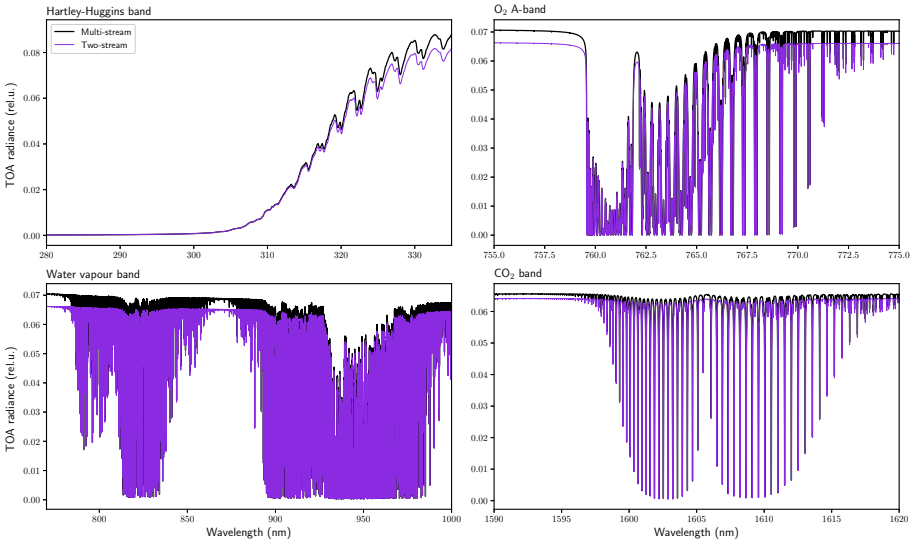

Fig. 1. Top of the atmosphere radiances computed for the absorption bands: Hartley-Huggins, O_2 A-, water vapour and CO_2 bands for one aerosol case. Black lines correspond to the multi-stream RTM while purple lines correspond to the two-stream RTM.

$$\hat{I}^c_{\mathrm{MS},i} = \alpha^c \hat{T}^c_i + \beta^c \hat{I}^c_{\mathrm{LS},i} + \gamma^c,\ i = 1, ..., N_C, \tag{1}$$

where α^c, β^c and γ^c are the regression coefficients of the c-th cluster and $\hat{T}$ is the corresponding direct transmittance. Equation (1) can be also written as follows:

$$Y = A \cdot X \tag{2}$$

with $Y = \left[\hat{I}^c_{\mathrm{MS},i}\right]$, $A = [\alpha^c, \beta^c, \gamma^c]$ and $X = \left[\hat{T}^c_i, \hat{I}^c_{\mathrm{LS},i}, 1\right]$. Finally, we find the regression coefficients as a solution to the following least square problem:

$$A = \arg\min_A \sum_{q=1}^{n} \left[\overline{I}^c_{\mathrm{MS},q} - Y\right]^2. \tag{3}$$

In this way, we can restore the spectra of the multi-stream radiances $\{\tilde{I}_{\mathrm{MS},i}\}_{i=1}^N$. Here, the "hat" notation $\hat{I}$ refers to the sorted radiances, the "bar" notation $\overline{I}$ refers to the equidistant radiances entering the regression model and the "tilde" notation $\tilde{I}$ refers to the predicted radiances. The total number of regression points, and thus the number of calls to the multi-stream RTM, is nC. Note that unlike the k-distribution method, the CLSR method provides a spectrum at the same spectral resolution as the LBL approach.

3 Simulations

3.1 Simulations of the absorption bands by using the CLSR method for the clear sky atmosphere

In this section we apply the CLSR method to simulate absorption bands. In addition to O_2A- and CO_2 bands analyzed in [11], we consider Hartley-Huggins and water vapour bands.

To estimate the accuracy of the CLSR method, we compute the residuals, the median and interquartile range (IQR). The residual for the radiance is computed at each spectral point λ_i as follows:

$$\Delta I_{\mathrm{res},i} = \frac{\tilde{I}_{\mathrm{MS},i} - I_{\mathrm{MS},i}}{I_{\mathrm{MS}}^{\mathrm{cont}}} \cdot 100, \tag{4}$$

where $\tilde{I}_{\mathrm{MS},i}$ is the radiance calculated with the CLSR method (cf. Eq. (1)), while $I_{\mathrm{MS}}^{\mathrm{cont}}$ is the radiance without absorption (i.e. the continuum radiance, which is used to avoid radiance values close to zero in the denominator of Eq. (4), when strong gas absorption is present [11, 7]).

Figure 2 shows the residuals of the CLSR method for different number of regression points per cluster for the four bands considered.

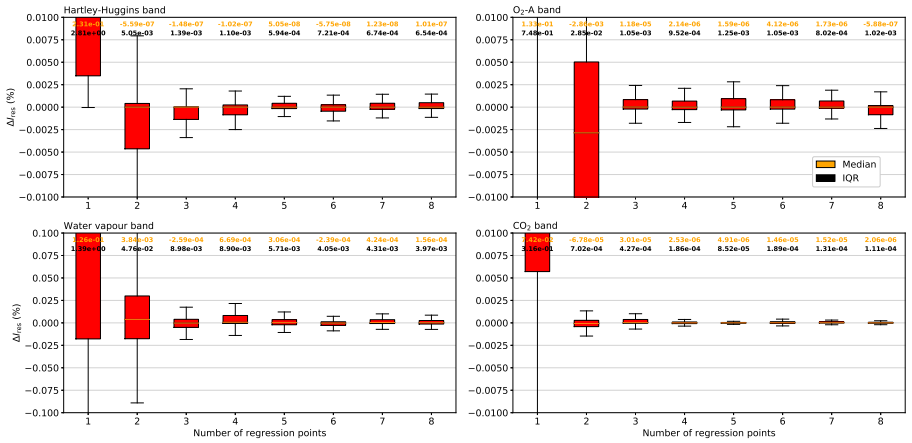

Fig. 2. Box plots of the residuals for the CLSR method for the clear sky scenario in the following absorption bands: Hartley-Huggins, O_2 A-, water vapour and CO_2 bands. The orange and black values on top of each box indicate the median values and the IQR values, respectively.

The residuals gradually decrease with the number of spectral points. In fact, they are significantly reduced when switching from 1—2 to 3 regression points. Therefore, the median values remain almost constant from 3 regression points. This trend is identical to the one found in [11] for different atmospheric scenarios. Note that the scale of residuals for the water vapour band is one order of magnitude higher than for the Hartley-Huggins, O_2A- and CO_2 bands.

3.2 Application of the CLSR method in the case of aerosols: accuracy results

In our previous work, we applied the CLSR method to the O_2A- and CO_2 bands for several atmospheric scenarios like aerosols and clouds at different heights and thicknesses. In this paper we examine the application of the CLSR method for several aerosol types and we extend the analysis to the Hartley-Huggins and water vapour bands. The computations are performed for the aerosol types outlined in Section 2.1.

Figure 3 shows the residuals for four bands and 5 aerosol models.

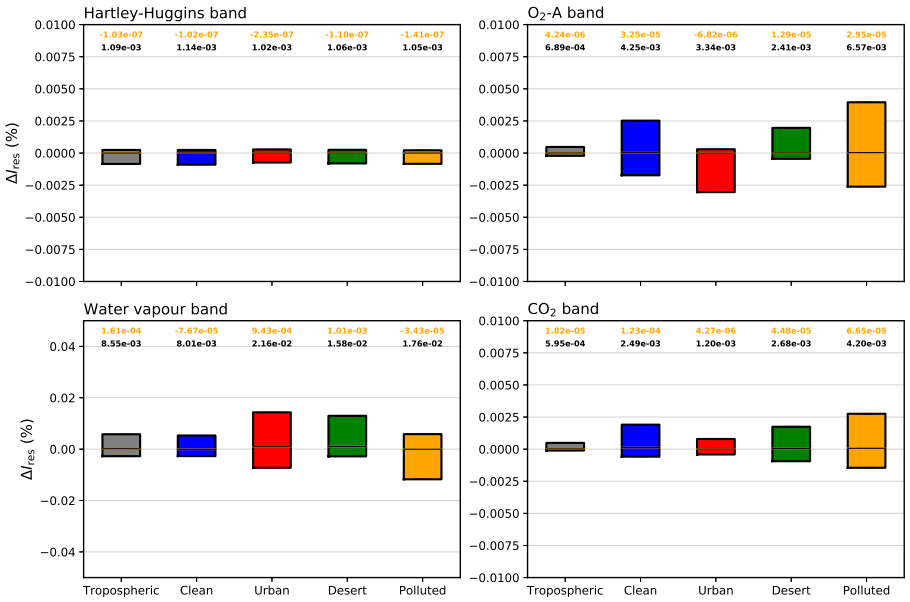

Fig. 3. Residuals for the CLSR method for the following absorption bands: Hartley-Huggins, O_2 A-, water vapour and CO_2 bands. The color of the boxes represents the type of aerosol: (grey) tropospheric; (blue) continental clean; (red): urban; (green): desert; (yellow): polluted. Note the different scale for the water vapour band compared with the other absorption bands.

Note that the residuals for the Hartley-Huggins, O_2A- and CO_2 bands are substantially smaller than those for the water vapour band. However, the median residuals are below 0.001 % and the results of the CLSR model are not biased.

In general, we conclude that the efficiency of the CLSR method is comparable to that of alternative methods like PCA-based RTMs (e.g. [7]) and our previous studies ([11]).

3.3 Assessment of the CLSR computational efficiency

In this section we analyse the computational performance of the CLSR method. Table 4 shows the number of calls to two- and multi-stream RTMs, the computational time

and the corresponding speedup factor with respect to the multi-stream LBL simulations for the O_2A-band. The computational time for monochromatic computations for the TS RTM is t_{TS} =1.6e-4 s, while for the MS RTM with $N_{\mathrm{do}} = 32$ discrete ordinates per hemisphere is t_{MS} =0.12 s, i.e., around three order of magnitude larger. However, most of the computational burden is still due to the MS RTM, and the computations of the approximate spectrum with a high spectral resolution by using the TS RTM is not a performance bottleneck in the whole CLSR processing chain. The results show that using the matrix of coefficients X with 5 clusters and 4 regression points for the CLSR method is 420 faster than using the LBL model. The speedup factor is of the same order of magnitude as in [21].

Table 4. Summary of number of calls, computational time and speedup factors for the O_2A-band with LBL and CLSR methods.

	LBL	CLSR
Number of calls to MS RTM	20000	20
Number of calls to TS RTM	0	20000
Time for MS RTM (s)	2320	2.32
Time for TS RTM (s)	0	3.20
Total computational time (s)	2320	5.52
Speedup factor	—	420

4 Conclusions

In this study, we have analysed the efficiency of the Cluster Low-Streams Regression (CLSR) method to accelerate spectral computations in several absorption bands. The CLSR method exploits the linear relationship between the low-stream and multi-stream models, where the corresponding regression coefficients are found by using the least-squares method. In our simulations several OPAC aerosol models have been considered. We reproduced the spectra with a median error below 0.001 % as compared to the reference multi-stream line-by-line model and IQR values below 0.1%. Thus, the errors present low variation and stability.

The number of calls to the multi-stream model was reduced by 3 orders of magnitude (e.g. from 20000 to 20 calls in the case of O_2A-band). The resulting performance enhancement is about 400 times. Note, that since the CLSR method is two orders of magnitude faster than the LBL model, it can be used for computations of the aerosol spectra in near-real-time applications.

In our future work, we plan to extend the CLSR method by using the asymptotic radiative transfer theory [22] and the diffuse approximation [23] instead of the two-stream RTM. Also it is of high interest to apply the CLSR method for modelling of the Stokes parameters.

References

1. Clough, S.A., Rinsland, C.P., Brown, P.D.: Retrieval of tropospheric ozone from simulations of nadir spectral radiances as observed from space. Journal of Geophysical Research **100**(D8), 16579 (1995). https://doi.org/10.1029/95jd01388
2. Fomin, B.A.: A k-distribution technique for radiative transfer simulation in inhomogeneous atmosphere: 2. FKDM, fast k-distribution model for the shortwave. Journal of Geophysical Research **110**(D2) (2005). https://doi.org/10.1029/2004jd005163
3. Fu, Q., Liou, K.: On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres. Journal of the Atmospheric Sciences **49**(22), 2139–2156 (1992). https://doi.org/10.1175/1520-0469(1992)049¡2139:OTCDMF¿2.0.CO;2
4. del Águila, A., Efremenko, D.S., Molina García, V., Xu, J.: Analysis of two dimensionality reduction techniques for fast simulation of the spectral radiances in the hartley-huggins band. Atmosphere **10**(3), 142 (3 2019). https://doi.org/10.3390/atmos10030142
5. Efremenko, D.S., Loyola, D.G., Doicu, A., Spurr, R.J.D.: Multi-core-CPU and GPU-accelerated radiative transfer models based on the discrete ordinate method. Computer Physics Communications **185**(12), 3079–3089 (2014). https://doi.org/10.1016/j.cpc.2014.07.018
6. Efremenko, D., Doicu, A., Loyola, D., Trautmann, T.: Optical property dimensionality reduction techniques for accelerated radiative transfer performance: Application to remote sensing total ozone retrievals. Journal of Quantitative Spectroscopy and Radiative Transfer **133**, 128–135 (2014). https://doi.org/10.1016/j.jqsrt.2013.07.023
7. Kopparla, P., Natraj, V., Limpasuvan, D., Spurr, R., Crisp, D., Shia, R.L., Somkuti, P., Yung, Y.L.: Pca-based radiative transfer: Improvements to aerosol scheme, vertical layering and spectral binning. Journal of Quantitative Spectroscopy and Radiative Transfer **198**, 104–111 (2017). https://doi.org/https://doi.org/10.1016/j.jqsrt.2017.05.005
8. Liu, X., Smith, W.L., Zhou, D.K., Larar, A.: Principal component-based radiative transfer model for hyperspectral sensors: Theoretical concept. Applied Optics **45**(1), 201–208 (2006). https://doi.org/10.1364/ao.45.000201
9. Natraj, V., Jiang, X., Shia, R., Huang, X., Margolis, J., Yung, Y.: Application of the principal component analysis to high spectral resolution radiative transfer: A case study of the O_2 A-band. Journal of Quantitative Spectroscopy and Radiative Transfer **95**(4), 539–556 (2005). https://doi.org/10.1016/j.jqsrt.2004.12.024
10. del Águila, A., Efremenko, D.S., Trautmann, T.: A review of dimensionality reduction techniques for processing hyper-spectral optical signal. Light & Engineering pp. 85–98 (2019). https://doi.org/10.33383/2019-017
11. del Águila, A., Efremenko, D.S., Molina García, V., Kataev, M.Y.: Cluster low-streams regression method for hyperspectral radiative transfer computations: Cases of O2 A- and CO2 bands. Remote Sensing **12**(8), 1250 (Apr 2020). https://doi.org/10.3390/rs12081250
12. Afanas'ev, V., Basov, A.Y., Budak, V., Efremenko, D., Kokhanovsky, A.: Analysis of the discrete theory of radiative transfer in the coupled "ocean–atmosphere" system: Current status, problems and development prospects. Journal of Marine Science and Engineering **8**(3), 202 (Mar 2020). https://doi.org/10.3390/jmse8030202
13. Budak, V., Efremenko, D., Shagalov, O.: Efficiency of algorithm for solution of vector radiative transfer equation in turbid medium slab. Journal of Physics: Conference Series **369**, 012021 (Jun 2012). https://doi.org/10.1088/1742-6596/369/1/012021
14. Doicu, A., Trautmann, T.: Discrete-ordinate method with matrix exponential for a pseudo-spherical atmosphere: Scalar case. Journal of Quantitative Spectroscopy and Radiative Transfer **110**(1-2), 146–158 (2009). https://doi.org/10.1016/j.jqsrt.2008.09.014

15. Efremenko, D.S., Molina García, V., Gimeno García, S., Doicu, A.: A review of the matrix-exponential formalism in radiative transfer. Journal of Quantitative Spectroscopy and Radiative Transfer **196**, 17–45 (Jul 2017). https://doi.org/10.1016/j.jqsrt.2017.02.015
16. Molina García, V., Sasi, S., Efremenko, D., Doicu, A., Loyola, D.: Radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVR measurements. Journal of Quantitative Spectroscopy and Radiative Transfer **213**, 228–240 (2018). https://doi.org/10.1016/j.jqsrt.2018.03.014
17. Schreier, F., Gimeno García, S., Hochstaffl, P., Städt, S.: Py4cats—PYthon for computational ATmospheric spectroscopy. Atmosphere **10**(5), 262 (2019). https://doi.org/10.3390/atmos10050262
18. Gordon, I., Rothman, L., Hill, C., Kochanov, R., Tan, Y., Bernath, P., Birk, M., Boudon, V., Campargue, A., Chance, K., Drouin, B., Flaud, J.M., Gamache, R., Hodges, J., Jacquemart, D., Perevalov, V., Perrin, A., Shine, K., Smith, M.A., Tennyson, J., Toon, G., Tran, H., Tyuterev, V., Barbe, A., Császár, A., Devi, V., Furtenbacher, T., Harrison, J., Hartmann, J.M., Jolly, A., Johnson, T., Karman, T., Kleiner, I., Kyuberis, A., Loos, J., Lyulin, O., Massie, S., Mikhailenko, S., Moazzen-Ahmadi, N., Müller, H., Naumenko, O., Nikitin, A., Polyansky, O., Rey, M., Rotger, M., Sharpe, S., Sung, K., Starikova, E., Tashkun, S., Auwera, J.V., Wagner, G., Wilzewski, J., Wcisło, P., Yu, S., Zak, E.: The HITRAN2016 molecular spectroscopic database. Journal of Quantitative Spectroscopy and Radiative Transfer **203**, 3–69 (2017). https://doi.org/10.1016/j.jqsrt.2017.06.038
19. Bodhaine, B., Wood, N., Dutton, E., Slusser, J.: On Rayleigh optical depth calculations. Journal of Atmospheric and Oceanic Technology **16**(11), 1854–1861 (1999). https://doi.org/10.1175/1520-0426(1999)016¡1854:orodc¿2.0.co;2
20. Hess, M., Koepke, P., Schult, I.: Optical properties of aerosols and clouds: The software package OPAC. Bulletin of the American Meteorological Society **79**(5), 831–844 (1998). https://doi.org/10.1175/1520-0477(1998)079¡0831:opoaac¿2.0.co;2
21. O'Dell, C.W.: Acceleration of multiple-scattering, hyperspectral radiative transfer calculations via low-streams interpolation. Journal of Geophysical Research **115**(D10) (2010). https://doi.org/10.1029/2009jd012803
22. Kokhanovsky, A.: Cloud Optics. Springer Netherlands (2006). https://doi.org/10.1007/1-4020-4020-2
23. Budak, V.P., Zheltov, V.S., Lubenchenko, A.V., Freidlin, K.S., Shagalov, O.V.: A fast and accurate synthetic iteration-based algorithm for numerical simulation of radiative transfer in a turbid medium. Atmospheric and Oceanic Optics **30**(1), 70–78 (Jan 2017). https://doi.org/10.1134/s1024856017010031