Journal of Quantitative Spectroscopy & Radiative Transfer 329 (2024) 109213

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy and Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

The k-bin tool: Fast and flexible k-distribution algorithms written in Python

Nils Madenach *, Rene Preusker, Nicole Docter, Lena Jänicke, Jürgen Fischer

Institut für Meteorologie, Freie Universität Berlin, Carl-Heinrich-Becker-Weg 6-10, Berlin, 12165, Berlin, Germany

ARTICLE INFO

Keywords:
k-distribution
Radiative transfer
Remote sensing

ABSTRACT

Radiative transfer simulations (RTS) still face significant challenges in accurately representing the highly complex gas absorption spectra of the Earth's atmosphere. Line-by-line RTS achieves high accuracy by solving radiative transfer equations for narrow spectral intervals, but at a considerable computational cost. Especially in remote sensing and climate modeling, a trade-off between efficiency and accuracy must be done. k-distribution methods are widespread in the scientific community and offer a way to make this trade-off. k-distribution methods reorder the absorption spectra k for a given spectral interval and find appropriate so-called k-bins. In the k-space much less integration points can be used, while maintaining high accuracy. The way to find optimal k-bins differs from method to method and depends on the application. In this paper, we present the flexible and fast k-bin tool. The python based lightweight k-bin tool provides a variety of different k-distribution methods and configuration options. One k-distribution method is the in-house developed k-bin approach. The different setups of the tool can be easily compared, helping to decide which method and configuration is best suited for a given application. We encourage the user of the tool to continue to optimize the k-bin tool and to extend it with new approaches and functionalities.

1. Introduction

A major challenge of radiative transfer simulations (RTS) is the treatment of the spectrally very complex gas absorption of the Earth's atmosphere. With so called line-by-line (LBL) RTS, the radiative transfer equations are solved for very narrow, quasi-monochromatic spectral intervals, allowing spectral integration to be performed with very high accuracy. However, especially in cases where scatters such as clouds or aerosols are present, the computational cost is very high and the calculations are time consuming. For both the remote sensing community and climate modellers, computational cost and time are critical factors. For each application a trade-off between efficiency and accuracy has to be made. A widely used approach to achieve a good compromise between efficiency and accuracy in the treatment of gas absorption for a given spectral range is the use of so-called k-distribution methods (see Section 2). k refers to the mass absorption coefficient, but the cross-section or as in the case of the k-bin tool, the optical thickness (τ) are also commonly used. The principle of k-distributions was first proposed in the 30s of the last century by [1] and first used for atmospheric applications by [2]. In the late 60s, the concept of "correlated spectra" was introduced [3,4] and further developed and optimized in the following years (e.g. [5,6]). Over the past decades various groups have developed k-distribution tools using different approaches (e.g. [7–16]). There are still large variations in the results of radiation schemes among climate models due to the use of different k-distribution models [17]. Thus, major efforts are ongoing to improve k-distribution models and to enhance our understanding on how different setups impact the accuracy. Recently, the *Correlated K-Distribution Model Intercomparison Project* (CKDMIP, [18]) provided open source software and datasets in order to examine these questions.

The tool presented in this paper has been developed for and used primarily by the satellite remote sensing community. Nevertheless, it is suitable and we encourage its use and optimization for climate modeling application as well. The tool is flexible in terms of the number and size of layers, the spectral range, the k-distribution method used and, in the case of the k-bin approach, the accuracy can be flexibly set for the entire atmosphere and for each layer individually. The purpose of this paper is to introduce the k-bin tool (Section 3) and to show some examples of the flexible usage of the tool by comparing different k-distribution approaches and configuration settings (Section 4). The methodology is described Section 2 and in Section 5 the paper is concluded. The tool is freely available and can be easily extended.

2. Methodology

Most RTSs are based on Beer–Bouguer–Lambert's law. The law states that the intensity I_ν at wavenumber ν of an incident beam, after

* Corresponding author.
E-mail addresses: nils.madenach@wew.fu-berlin.de (N. Madenach), rene.preusker@fu-berlin.de (R. Preusker).

https://doi.org/10.1016/j.jqsrt.2024.109213
Received 26 June 2024; Received in revised form 25 September 2024; Accepted 1 October 2024
Available online 5 October 2024

passing through an air mass described by the air mass factor (AMF) m, is reduced by the transmittance T. The transmittance for a given m is:

$$T(m) = \exp(-\tau(\nu)m) = I/I_0. \tag{1}$$

where τ is the optical thickness due to gas absorption at the wavenumber ν. For line-by-line calculation of, for example, the average transmittance of a band, the integral

$$T(m) = \frac{1}{\nu_2 - \nu_1}\int_{\nu_1}^{\nu_2} \exp(-\tau(\nu)m)d\nu. \tag{2}$$

must be evaluated. To solve this integral numerically, it can be written as a sum over N wavenumbers. Even for small spectral intervals, N must be large to ensure high accuracy, resulting in high computational cost.

$$T(m) \sim \omega \sum_{i=1}^{N} \exp(-\tau(\nu_i)m). \tag{3}$$

The weights ω depend on the integration method. For equally spaced wavenumbers ω can be written as:

$$\omega = \frac{1}{N} \tag{4}$$

In order to reduce the computational cost, the k-distribution method sorts all N values of $\tau(\nu_i)$ in an ascending order. This new smooth, monotonically increasing function depends on a new variable g ranging from 0 for τ_{min} to 1 for τ_{max}. The new function actually represents the cumulative probability function (CPF) of the optical thicknesses (see Fig. 1(b)). Solving this integral requires much fewer interpolation points, drastically reducing computational cost while maintaining a reasonably high accuracy.

$$T(m) = \int_0^1 \exp(-\tau(g)m)dg. \tag{5}$$

The package described in this paper can be used in different configurations, e.g using the correlated k-distribution method or the uncorrelated k-distribution method [12]. The main differences between these approaches are discussed below.

For the correlated k-distribution method, the gas absorption optical thicknesses are sorted as described above and separated into M bins called k-bins (see Fig. 1) where the number of bins M are given by the user. A mapping function is created that assigns each wavelength to the corresponding k-bin. [19] found that the accuracy does not always increase with increasing M. This means that the accuracy cannot always be optimized by increasing the number of bins. As the name implies, correlated k-distribution methods assume, that the absorption spectra have perfect rank correlation between different layers with varying temperature and pressure. In many situations, this assumption fulfills the requirements. However, in the case of e.g. overlapping bands, this assumption introduces errors (e.g. [20]). [21,22] describe a method where the absorption spectra can deviate from the assumption of perfect rank correlation. An example of such a case is an atmosphere with a high concentration of water vapor in the lower troposphere and a high concentration of ozone in the dry upper stratosphere. [12] referred to this method as *non-correlated k-distribution method*. [23] argues that the method should be considered as a variant of the correlated k-distribution method since it still relies on the high correlation of absorption spectra in the vertical. In the following, the method is referred to as the k-bin approach or simply k-binning.

2.1. K-bin approach

This section introduces the k-bin approach. The method is described in more detail in [12], which in turn is an evolution of the method from [22]. The idea of the k-bin approach is to find a mapping function that applies not only to the spectra of the entire atmosphere, but rather to all atmospheric layers, defined by the user. The mapping function $f(\nu) = g$ maps the τ in the wavenumber space to the τ in the g-point space where τ is sorted in ascending order (see Fig. 1(b)). $f^{-1}(g) = \nu$ is the respective inverse mapping function. The fidelity of the mapping function is defined by a threshold of allowed maximum absolute deviation of the transmission (ΔT, see Eq. (6)) derived from k-binning (T_k) compared to the transmission derived from high-resolution optical thicknesses (T_{hr}). The threshold of the maximum absolute transmission error can be specified by the user for the total atmosphere (ΔT_{toa}) and for each layer ΔT_l. To evaluate for multi-scattering, the transmissions are calculated using different air mass factors (AMF or m within equations) as defined by the user (see Table 1). Thereby, the transmission deviations for all given m (Eq. (7)) must fall below the threshold.

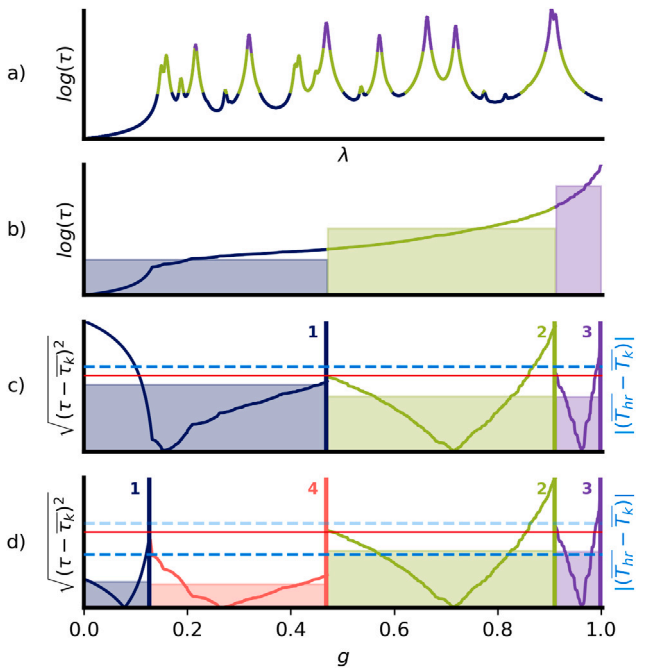

Fig. 1. Schematic illustration of the k-binning optimization procedure. Panel (a) displays an absorption spectrum with three different k-bins found by the algorithm. (b) displays the CPF of (a) in g-point space with the three k-bins. The height of the shaded areas indicates the mean optical thickness of the k-bin. In (c) the solid lines represent the spread of τ within each k-bin. The height of the shaded areas displays the standard deviation (std) of the bin and the doted line displays the absolute total transmission error. In (d) the k-bin with the largest spread/std (for k-bin 1 the shaded area is highest) has been identified and split in two. The blue dotted line shows the improved absolute total transmission error and the light blue the old error from figure (c). The red line in panel (c) and (d) represents the user defined threshold of the maximum transmission error.

$$\Delta T = T_k - T_{hr} \tag{6}$$

$$\Delta T_z(m) = Abs\left(\sum_i^M w_i \exp[-\overline{\tau}_{z,i}m] - \frac{1}{\Delta_\nu}\int_{\Delta\nu_i} \exp[-\tau_z(\nu)m]d\nu\right) \tag{7}$$

where M is the number of k-bins, z is the layer l or in the case of the total atmosphere ($z = toa$) the vertical integral of τ and $\overline{\tau}_i$ is the bin averaged optical thickness. The $\overline{\tau}_i$ is computed by taking the arithmetic mean over all τ in the bin ($average\_method = opt$) or optionally ($average\_method = trans$) over all transmissions and then transforming the k-bin mean transmission to $\overline{\tau}_i$ (see Eq. (11) and Section 3.2.3). The weight ω_i of k-bin i is calculated taking the normalized spectral response function $R(\nu)$ (Eq. (8)) into account.

$$\int_{\Delta\nu} R(\nu)d\nu = 1 \tag{8}$$

$$\omega_i = \int_{\Delta g_i} R\left(f^{-1}(g)\right)dg \tag{9}$$

Table 1
Overview of possible configurations.

Key word	Description	Example
high_resolution_optical_thickness_file	Input data (e.g. HITRAN or CKDMIP based high resolution optical thickness)	*'./indata/egasa_O2A_geo_mls.nc'*
solar_file	Input file of solar irradiance	*'.indata/IRRADTHUWL.DAT'* *'NOT_USED'*
detailed_output	T and opt are written for each gas as well	*True*
response_function_type	Relative spectral response function: either as generic description with the wavelength given in nmr or a file name	*'gaussian : center_wvl : fwhm_wvl* *'rectangular : min_wvl : max_wvl* *'by_file : ./indata/SChar_MERIS_Band_Shape.b11'*
maximum_number_of_bins	Maximum number of bins	10
maximum_error_total_atmosphere[a]	Max. allowed error in absolute units of total atm. transmission	0.01
maximum_error_per_layer[a]	A single or N numbers for each layer starting at TOA	0.1 0.1 0.1

Key word	Options					Example
kbin_method	*mean*	*correlated_toa*	*correlated_strongest_layer*	*correlated_toa_strongest_layer*	*k_binning*	*k_binning*
average_method	*opt*			*trans*		*trans*
sort_over	*lnopt*	*trans*	*rank*	*emis*		*trans*
min_amf, *max_amf*, *nnn_amf*	Defining the range of air mass factors used to evaluate the k-bins.					0.75, 10, 19
err_meas[a]	*mean_err*			*max_err*		*mean_err*

[a] Options that are only relevant for the k-bin approach.

The spectral response function (SRF) describes how sensitive an instrument's photosensor is to radiation of a particular wavelength. Using the inverse mapping function, the SRF can also be applied after the k-bins are found. This is a reasonable way to slightly modify the weights of the k-distribution after the optimal k-bins are found. For example, to account for spectral distortions of a satellite imager over the field of view.

The k-binning is initiated by subdividing the sorted absorption spectra of the total atmosphere into $M = 2$ intervals (k-bins) by separating the optical thickness in steps of $\Delta\tau = 1/M \cdot (\tau_{max} - \tau_{min})$. There are different options to sort the spectra before bin separation (see Sections 3.2.2 and 4.1). Fig. 1 displays schematically the procedure using the natural logarithm for sorting the optical thickness spectra. The scheme starts at the iteration step where three k-bins have been found. Fig. 1(a) displays the logarithmic optical thickness for a spectral range and in different colors the k-bins created by the separations. In Fig. 1(b) the optical thicknesses are displayed in the g-space or in other words the CPF. The shaded areas show here the mean optical thickness. Between each k-bin the mean optical thickness increases by $\Delta\tau$. For Fig. 1(c) the total transmission is evaluated for the different AMF m. The blue dotted line displays $\Delta T(m)$ and the red line the threshold of transmission error set by the user. If for any m, $\Delta T(m)$ (Eq. (7)) is above the threshold (blue line is above the blue line), the k-bin with the largest spread in optical thickness (quantified by the standard deviation within the k-bin, see shaded areas in 1(c) and (d)) is split into two sub-intervals (see Fig. 1(d)). With the additional k-bin (k-bin 4), the new $\Delta T(m)$ are calculated and compared to the threshold. These steps are repeated until for all m the $\Delta T(m)$ is below the threshold, as it is the case in Fig. 1(d) (blue line is below the red line). After finding the correct mapping function for the absorption spectra of the total atmosphere, the above steps are repeated for each layer l with the subsequent thresholds of each layer. After finding the correct mapping for all layers, it is evaluated if $\Delta T(m)$ for total atmosphere is still below the threshold and if not, the steps are repeated again for the spectra of the entire atmosphere. The k-bin approach allows to stack two or more profiles of temperature, pressure and gas concentration and find a mapping function that is optimized for all the given atmospheric states.

3. The k-bin tool

In this section the k-bin tool is introduced and different setups are illustrated. The k-bin tool is written in Python. Python, unlike the equations within this paper, is zero-based indexed. The tool do not uses the g-space directly, but mapping and inverse mapping functions, that maps the high-resolution spectra in the physical space and the sorted spectra to each other using their indices. The tool is flexible regarding the number and size of layers used, the input data and its configurability, e.g. in choosing different k-distribution methods. All configurables and options are listed in Table 1. The input (*high_resolution_optical_thickness_file*) must be a netCDF file containing vertically resolved spectra of high-resolution absorption optical thicknesses (e.g. based on HITRAN or CKDMIP dataset) must be provided. The spectral range can be defined in different ways using the relative spectral response function keyword (*response_function_type*). If there is no SRF that the tool should take into account for the weighting of the k-bins, a rectangular SRF can be defined by simply specifying the spectral range (*rectangular : min_wvl : max_wvl* [*nm*]). This option should be used, if no SRF is available or needed, such as in the case of the use in climate models. A Gaussian function can also be passed by specifying the center wavelength and the full width half maximum (*gaussian : center_wvl : fwhm* [*nm*]). For specific cases, such as satellite channel, a *.b11 – file* containing the relative spectral response can be provided. In this case, the spectral range is taken directly from the wavelength range within the file. The available k-distribution methods (Section 3.1) and the configuration options (Section 3.2) are described below. As a starting point, the example configuration files in the test folder can be used. The files define typical default values to start with and include some additional comments.

3.1. Available k-distribution methods

The k-bin tool has implemented four different k-distribution methods. This section gives an overview of the different options (see also Table 1). Section 4.2 discusses the advantages and disadvantages with examples. The keyword to set a specific method (*kbin_method*) will be given in parentheses in the following.

3.1.1. Correlated total atmosphere

The correlated total atmosphere option (*correlated_toa*) uses the optical thicknesses integrated over all layers and sorts them accordingly to the correlated k-distribution method explained in Section 2. The configuration parameters are the number of bins(*maximum_number_of_bins*), the sorting method (*sort_over*).

3.1.2. Correlated strongest layer

The correlated strongest layer option (*correlated_strongest_layer*) uses the layer with the strongest absorption to find optimal k-bins instead of the vertically integrated optical thickness. The configuration parameters are the number of bins (*maximum_number_of_bins*) and the sorting method (*sort_over*).

3.1.3. Correlate total atmosphere and strongest layer

This option (*correlated_toa_strongest_layer*) represents the approach developed in [22]. Both the absorption spectra of the layer with the strongest absorption and the total atmosphere integrated spectra are used for the k-binning. The configuration parameters are the number of bins (*maximum_number_of_bins*), the sorting method (*sort_over*) and the weight for the total atmosphere (*wgt_toa*). The weight for the strongest layer method is calculated internally ($1 - wgt\_toa$). The default value for *wgt_toa* is 0.7.

3.1.4. K-binning

The k-binning option (*k_binning*) uses the optimization approach described in 2.1. The configuration parameters are the maximum absolute transmission error per layer (*maximum_error_per_layer*) and for the total atmosphere (*maximum_error_total_atmosphere*), number of bins (*maximum_number_of_bins*), the sorting method (*sort_over*), the error measure (*err_meas*) (see Section 3.2.5), and the average method (*average_method*).

3.1.5. Mean

The mean option (*mean*) simply creates an arithmetic mean over the entire spectra of optical thicknesses (only 1 bin).

3.2. Available configuration options

In Table 1 all available configuration options that can be defined in the configuration file are illustrated.

3.2.1. Maximum number of k-bins

The maximum number of k-bins (*maximum_number_of_bins*) specifies how many k-bins should be used. With k-binning, the maximum number of k-bins can be exceeded if the threshold of the maximum transmission error for the total atmosphere is not reached with the given number of k-bins.

3.2.2. Sorting methods

There are four different ways to sort the spectra. The options can be set using the keyword *sort_over* in the configuration file. Besides sorting by optical thickness (*rank*), it is possible to sort the spectra using the natural logarithm of optical thickness (*lnopt*), the transmissions (*trans*) or the emissions (*emis*). Some of the options are discussed in more detail in Section 4.1 and sorting using emissions in [12].

3.2.3. Average method

This option (*average_method*) defines how the k-bin averaged optical thicknesses (τ_k) is calculated. Either the optical thicknesses (*opt*) shown in Eq. (10) or the transmissions (*trans*) displayed in Eq. (11) can be used for the averaging. N_k is the number of high resolution *opt* within the bin.

$$\overline{\tau_k} = \frac{1}{N_k} \sum_{i}^{N_k} \tau_i \tag{10}$$

$$\overline{\tau_k} = -\log\left(\frac{1}{N_k} \sum_{i}^{N_k} \exp[-\tau_i] \right) \tag{11}$$

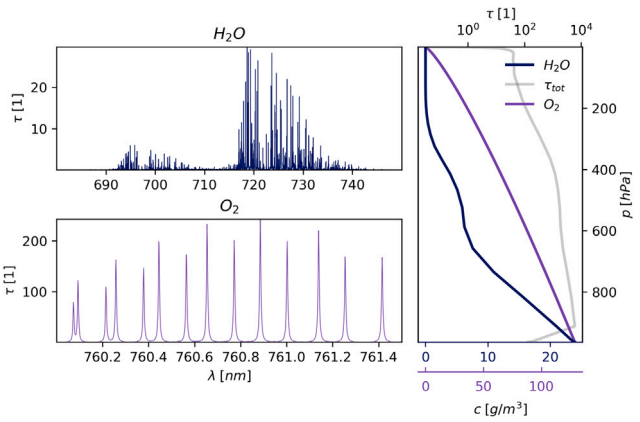

Fig. 2. Absorption spectra of the O_2 *A*-Band and the H_2O-band (left). And the vertical profile of water vapor and oxygen concentrations and the layer optical thickness (right).

3.2.4. Maximum transmission error (thresholds)

This option is only relevant for the k-bin approach (*k_binning*) and defines the threshold of the maximum transmission error for the total atmosphere (*maximum_error_total_atmosphere*) and for the maximum transmission error of each layer (*maximum_error_per_layer*) that has to be reached to stop the optimization procedure. See Table 1 for an example of typical values.

3.2.5. Error measure options

This option is only relevant for the k-bin approach (*k_binning*) and defines how the transmission error is calculated. Using the mean error (*mean_err*) the transmissions calculated from k-bin averaged optical thickness are summed up and subtracted from the total high-resolution transmission. This is done for all air mass factors and in all cases the error must fall below the given threshold (see Eq. (7)).

Using the maximum error (*max_err*) the transmissions calculated from the k-bin averaged optical thickness are subtracted from the k-bin transmissions calculated using the high-resolution optical thickness. For all M k-bins the error must be below the thresholds. See [12] for in depth explanation. The *mean_err* is used as default.

4. Comparisons

This section presents some interesting examples of configuration setups for the k-bin tool. First, different sorting options are discussed (Section 3.2.2). In Section 4.2 the correlated k-distribution approach is compared to the k-bin approach.

The comparisons are made using data from CKDMIP described in [18]. Based on the *idealized* data set, the optical thicknesses were interpolated to the temperatures, pressures and mole fractions of the given profiles.

4.1. Comparison sort over transmission vs. sort over optical thickness

In this example, we discuss the impact of the sorting method used on the correlated total atmosphere k-distribution method. The example (see below) shows, that for very high optical thicknesses, sorting the spectra based on transmission (*trans*) leads to better results. However, for cases with medium to low optical thickness, sorting based on optical thickness (*rank*) or its natural logarithm (*lnopt*, not shown) was more appropriate. Using transmissions to sort the spectra, results in larger bins for high optical thicknesses compared to sorting based on optical thickness. For very high optical thicknesses, where absorption reaches saturation, it is of advantage if the bins for high optical thicknesses are larger and therefore bins for lower optical thicknesses where no saturation occurs are smaller. An example illustrating this behavior

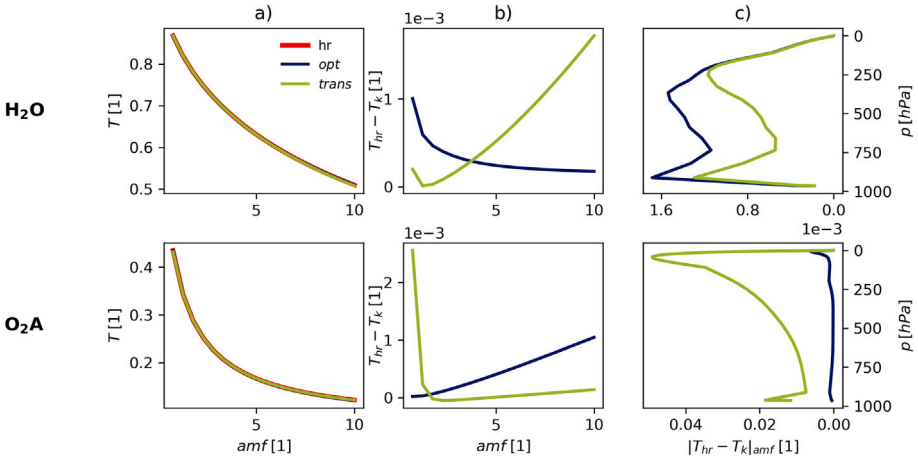

Fig. 3. Comparison of the relative transmission error for sorting the spectra using optical thickness (blue) and using the transmission (green). For the H_2O-band (upper panels) and the O_2A-band (lower panels). As k-distribution method *correlated_toa* has been used with a maximum total atmosphere error of 0.001 and a maximum layer error of 0.01. The maximum number of k-bins has been set to 35. As input the CKDMIP has been used. Column (a) displays the transmission compared to high-resolution. (b) displays the relative total error and (c) the absolute error per layer averaged over the air mass factors.

is shown in the following. Fig. 2 shows an example of a very strong absorbing spectral interval in the O_2A-band and a band in the visible region where water vapor is absorbing. The right panel displays the vertical profiles of H_2O and O_2. In Fig. 3, the results of applying the correlated k-distribution method (*correlated_toa*, $max_{bins} = 35$) with either using the transmission or the optical thickness to sort the spectra are displayed. For the spectra in the O_2A-band with very high optical thicknesses of more than 200 in the center of the absorption lines, sorting using transmission performs much better. The opposite is true for the H_2O absorbing spectra.

4.2. Comparison k-binning vs. correlated k-distribution

The k-bin approach is especially advantageous in cases where two absorbing gases are present at different heights in the atmosphere. Fig. 4 shows an illustrative example where due to an eruption of a volcano huge amounts of CO_2 have been emitted into the troposphere and lower stratosphere. At the same time, the lower troposphere contains high concentrations of CH_4. For spectral intervals where both gases absorb, e.g. between 1.6 and 1.7 μm (see Fig. 4 left panels), the k-bin approach is able to find mapping functions with much smaller transmission errors. The upper panels of Fig. 5 display the comparison of the k-binning approach and the correlated approach evaluated with different air mass factors. The relative total transmission error can be of an order of magnitude smaller for the k-binning (middle panel). This is also the case for the in-layer transmission error averaged over all air masses (right panel). The middle and lower panel illustrate why the k-bin approach performs better. The sizes are variable and smaller for segments where the optical thickness has large variance and large for regions where the variance is small (middle panel). For the correlated k-distribution approach, the k-bins are equally spaced (lower panel). K-binning also generally outperforms other k-distribution methods when applied to an atmosphere with a strongly absorbing gas in the lower troposphere (e.g. Fig. 4). In the presence of clouds, the RT does not "see" anymore the absorbing gas below the cloud. But it still uses the bins optimized with this gas absorption lines. As k-binning is optimized for all layers, the bins tend to better represent the atmosphere above the cloud without the strong absorbing gas. On the other hand, for a pure water absorbing channel, the k-bin approach does not necessarily perform better than the correlated approach for the total transmission (not shown). For the between layer transmissions errors k-binning usually produces even in such cases better results.

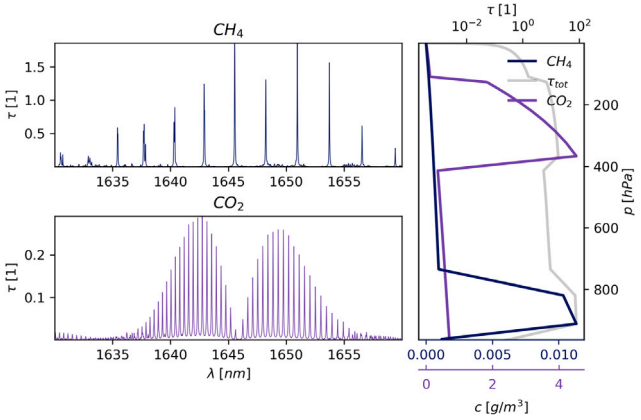

Fig. 4. As Fig. 2 but for an illustrative example of an eruption of a volcano causing a large CO_2-bloom in the upper troposphere and lower stratosphere. At the same time high methane concentrations are present near the surface. The left panels show an interval in the near infrared where both methane and carbon dioxide are strongly absorbing.

5. Conclusion

In this paper, the k-bin tool package written in Python is described and discussed. The tool is open source and can facilitate the comparison of different k-distribution methods and configuration settings to find the optimal setup for a specific application. The tool takes high-resolution optical thicknesses as input to find a mapping function for a given number of k-bins. Currently there are four different k-distribution methods to find the mapping function implemented. Configuration settings can be easily changed in the configuration file to meet the optimal needs of the user and their specific application. The advantages and disadvantages of the different methods and configurations are discussed for some selected examples. A special focus is given to the

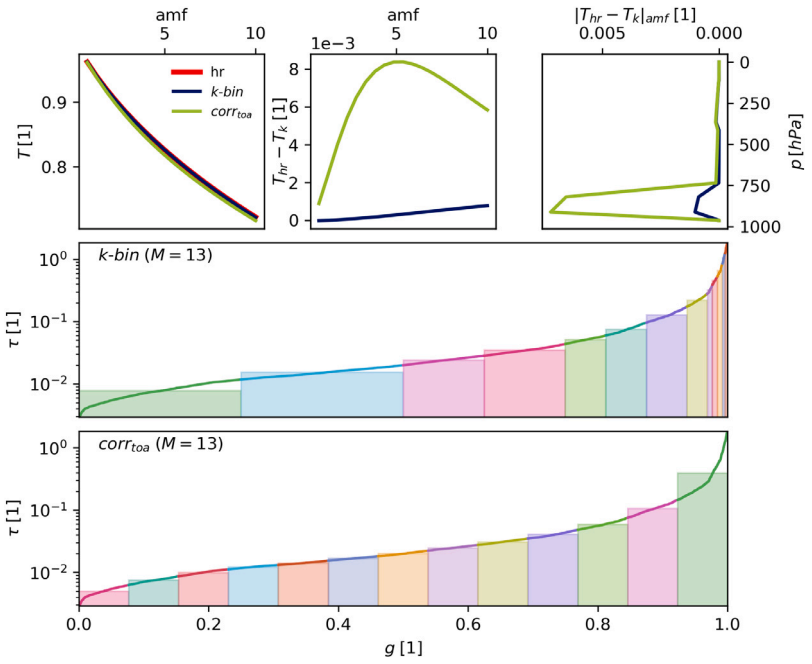

Fig. 5. Comparison of the k-distribution methods. Upper panel: transmissions using k-binning (blue), correlated total atmosphere (green) and for high resolution (red) along different air mass factors (left), total transmission error (middle) and transmission error per layer averaged over all air mass factors (right). Middle panel: locations of k-bins in g-point space using the k-bin approach. Lower panel: same as above but using the correlated total atmosphere approach.

in-house developed k-bin approach, which optimizes the k-bins until user defined thresholds for the maximum allowed error in total and in per layer transmissions are reached. This facilitates the often critical decision between accuracy and efficiency. Furthermore, for k-binning multiple atmospheres can be stacked together to find a mapping function that is optimized for all the given atmospheres. This results in better performance for situations where a single mapping function is applied to very different atmospheric states. With its flexibility and easy usage, the k-bin tool can help to better understand and improve the treatment of gas absorption using k-distribution methods. For all examples shown in the paper, the k-bin tool required less than one second, using a single core. Only the preparation (e.g. interpolation to a common wavelength's) of the data for the H2O absorption band around 720 nm took 3 s due to the high number of layers used (54). The sorting itself was always below 0.6 s.

CRediT authorship contribution statement

Nils Madenach: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Rene Preusker:** Writing – review & editing, Supervision, Software, Methodology, Conceptualization. **Nicole Docter:** Writing – review & editing, Software, Conceptualization. **Lena Jänicke:** Writing – review & editing. **Jürgen Fischer:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared a link to the code in the manuscript.

Acknowledgments

We appreciate the collaboration and discussion within CKDMIP which prompted this work. We further thank Cintia Carbajal Henken for giving valuable comments on the draft version. The two anonymous reviewers we want to thank for their thoughtful comments and efforts towards improving the manuscript.

Code availability

The code for the k-bin tool is freely available: https://gitlab.com/rene.preusker/kbin.git

References

[1] Ambartzumian V. The effect of the absorption lines on the radiative transfer equilibrium of the outer layers of stars. Publ Astron Obs Leningrad 1936;6:7–18.
[2] Lebedinsky AI. Radiative equilibrium in the Earth's atmosphere. Proc Leningrad Univ Ser Math 1939;3(31):152–75.
[3] Kondratyev KY. Radiation in the atmosphere. Academic Press; 1969, p. 912 pp..
[4] Arking A, Grossman K. The influence of line shape and band structure on temperatures in planetary atmospheres. J Atmos Sci 1972;29(6):937–49. http://dx.doi.org/10.1175/1520-0469(1972)029<0937:TIOLSA>2.0.CO;2.
[5] Goody R, West R, Chen L, Crisp D. The correlated-k method for radiation calculations in nonhomogeneous atmospheres. J Quant Spectrosc Radiat Transfer 1989;42(6):539–50. http://dx.doi.org/10.1016/0022-4073(89)90044-7.
[6] Lacis AA, Oinas V. A description of the correlated k distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres. J Geophys Res: Atmos 1991;96(D5):9027–63. http://dx.doi.org/10.1029/90jd01945.
[7] Edwards JM, Slingo A. Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. Q J R Meteorol Soc 1996;122(531):689–719.
[8] Kato S, Ackerman TP, Mather JH, Clothiaux EE. The k-distribution method and correlated-k approximation for a shortwave radiative transfer model. J Quant Spectrosc Radiat Transfer 1999;62(1):109–21. http://dx.doi.org/10.1016/s0022-4073(98)00075-2.

[9] Zhang H, Nakajima T, Shi G, Suzuki T, Imasu R. An optimal approach to overlapping bands with correlated k distribution method and its application to radiative calculations. J Geophys Res: Atmos 2003;108(D20). http://dx.doi.org/10.1029/2002jd003358.

[10] Sekiguchi M, Nakajima T. A k-distribution-based radiation code and its computational optimization for an atmospheric general circulation model. J Quant Spectrosc Radiat Transfer 2008;109(17–18):2779–93. http://dx.doi.org/10.1016/j.jqsrt.2008.07.013.

[11] Hogan RJ. The full-spectrum correlated-k method for longwave atmospheric radiative transfer using an effective Planck function. J Atmos Sci 2010;67(6):2086–100. http://dx.doi.org/10.1175/2010jas3202.1.

[12] Doppler L, Preusker R, Bennartz R, Fischer J. K-bin and k-IR: k-distribution methods without correlation approximation for non-fixed instrument response function and extension to the thermal infrared—Applications to satellite remote sensing. J Quant Spectrosc Radiat Transfer 2014;133:382–95. http://dx.doi.org/10.1016/j.jqsrt.2013.09.001.

[13] Pincus R, Mlawer EJ, Delamere JS. Balancing accuracy, efficiency, and flexibility in radiation calculations for dynamical models. J Adv Modelling Earth Syst 2019;11(10):3074–89. http://dx.doi.org/10.1029/2019ms001621.

[14] Andre F, Cornet C, Galtier M, Dubuisson P. Radiative transfer in the O_2 A-band - A fast and accurate forward model based on the L-distribution approach. J Quant Spectrosc Radiat Transfer 2021;260. http://dx.doi.org/10.1016/j.jqsrt.2020.107470.

[15] Stegmann PG, Johnson B, Moradi I, Karpowicz B, McCarty W. A deep learning approach to fast radiative transfer. J Quant Spectrosc Radiat Transfer 2022;280. http://dx.doi.org/10.1016/j.jqsrt.2022.108088.

[16] Andre F, Delage C, Guilmard L, Galtier M, Cornet C. Bridging physics and statistical learning methodologies for the accurate modeling of the radiative properties of non-uniform atmospheric paths. J Quant Spectrosc Radiat Transfer 2024;320. http://dx.doi.org/10.1016/j.jqsrt.2023.108961.

[17] Collins WD, Ramaswamy V, Schwarzkopf MD, Sun Y, Portmann RW, Fu Q, Casanova SEB, Dufresne J-L, Fillmore DW, Forster PMD, Galin VY, Gohar LK, Ingram WJ, Kratz DP, Lefebvre M-P, Li J, Marquet P, Oinas V, Tsushima Y, Uchiyama T, Zhong WY. Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). J Geophys Res: Atmos 2006;111(D14). http://dx.doi.org/10.1029/2005jd006713.

[18] Hogan RJ, Matricardi M. Evaluating and improving the treatment of gases in radiation schemes: the Correlated K-Distribution model intercomparison project (CKDMIP). Geosci Model Dev 2020;13(12):6501–21. http://dx.doi.org/10.5194/gmd-13-6501-2020.

[19] West R, Goody R, Chen L, Crisp D. The correlated-k method and related methods for broadband radiation calculations. J Quant Spectrosc Radiat Transfer 2010;111(11):1672–3. http://dx.doi.org/10.1016/j.jqsrt.2010.01.013.

[20] Goody R, Yung Y. Atmospheric radiation: Theoretical basis. USA: Oxford University Press; 1989, p. 176–7.

[21] West R, Crisp D, Chen L. Mapping transformations for broadband atmospheric radiation calculations. J Quant Spectrosc Radiat Transfer 1990;43(3):191–9. http://dx.doi.org/10.1016/0022-4073(90)90051-7.

[22] Bennartz R, Fischer J. A modified k-distribution approach applied to narrow band water vapour and oxygen absorption estimates in the near infrared. J Quant Spectrosc Radiat Transfer 2000. http://dx.doi.org/10.1016/S0022-4073(99)00184-3.

[23] Hogan R, Matricardi M. A tool for generating fast k-distribution gas-optics models for weather and climate applications. J Adv Modelling Earth Syst 2022;14. http://dx.doi.org/10.1029/2022MS003033.