A tool for generating fast k-distribution gas-optics models for weather and climate applications

Robin J. Hogan[1] **and Marco Matricardi**[1]

[1]European Centre for Medium-Range Weather Forecasts, Reading, UK.

Key Points:

- We describe a free software tool for generating accurate gas optics models for radiation schemes and test with line-by-line calculations.
- Extra efficiency is achieved via the use of the full-spectrum correlated-k (FSCK) method in the thermal- and near-infrared.
- The spectral properties of clouds are treated accurately via the use of sub-bands in the near-infrared and several other techniques.

Corresponding author: Robin J. Hogan, `r.j.hogan@ecmwf.int`

Abstract

One of the most important components of an atmospheric radiation scheme is its treatment of gas optical properties, which determines not only the accuracy of its radiative forcing calculations fundamental to climate prediction, but also its computational cost. This paper describes a free software tool 'ecCKD' for generating fast gas-optics models by optimally dividing the spectrum into pseudo-monochromatic spectral intervals (known as *k*-terms) according to a user-specified error tolerance and the range of greenhouse-gas concentrations that needs to be simulated. The models generated use the correlated *k*-distribution method in user-specified bands, but can also generate accurate 'full-spectrum correlated-*k*' models that operate on the entire longwave or near-infrared parts of the spectrum. In the near-infrared, the large spectral variation in cloud absorption is represented by partitioning the parts of the spectrum where gases are optically thin into three or more sub-bands, while allowing *k*-terms for the optically thicker parts of the spectrum (where clouds and surface reflectance are less important) to span the entire near-infrared spectrum. Candidate models using only 16 and 32 *k*-terms in each of the shortwave and longwave are evaluated against line-by-line calculations on clear and cloudy profiles. The 32-term models are able to accurately capture the radiative forcing of varying greenhouse gases including CO_2 concentrations spanning a factor of 12, and heating rates at pressures down to 1 Pa.

Plain Language Summary

A crucial component of atmospheric computer models used to make climate projections and weather forecasts is the 'gas optics scheme', which represents the interaction of sunlight and infrared radiation with greenhouse gases. This paper describes a free software tool 'ecCKD' that uses a number of novel techniques to generate new gas optics schemes that are computationally faster than most existing schemes while still being very accurate. For example, the schemes are able to simulate variations in carbon dioxide concentration spanning a factor of 12 and methane concentration spanning a factor of 10. Users of ecCKD can generate schemes that are optimized for specific applications, such as short-term weather forecasting. A special focus has been placed on the near-infrared part of the solar spectrum to ensure that the schemes work well when computing the interactions of sunlight simultaneously with gases and clouds, important to ensure that the impact of clouds on weather and climate are well simulated.

1 Introduction

Perhaps the most fundamental part of a climate model is the gas-optics module of its radiation scheme; in fact, one of the most influential (and indeed Nobel-prize-wining) studies of the climatic impact of increased greenhouse gases used a climate model consisting of little more than a radiation scheme coupled to a convection scheme (Manabe and Wetherald, 1967). The correlated *k*-distribution (CKD) method (Goody et al., 1989; Lacis and Oinas, 1991) has emerged as the leading technique for treating the radiative effects of gases that is fast enough to use in 3D weather and climate models. Many models still use older alternatives; DeAngelis et al. (2015) reported a large spread in the magnitude of the near-infrared water (NIR) vapor feedback amongst 14 climate models, but those using the CKD method were found to be much the most accurate compared to benchmark line-by-line (LBL) radiation calculations.

CKD models are very time consuming to develop from scratch, so when writing a new radiation scheme one must usually incorporate one of the small number of off-the-shelf models, even though it may not be optimized for ones particular application. The perceived high computational cost of radiation schemes has spurred numerous ideas to accelerate them, such as replacement of the radiation scheme (or only its gas-optics component) by a neural network (e.g. Ukkonen et al., 2020), or sub-sampling of model columns (Barker et al., 2021). The computational cost of a radiation scheme scales with the number of pseudo-monochromatic calculations (hereafter referred to as '*k*-terms' or spectral intervals) required to represent the

entire spectrum, which is determined by the gas-optics scheme. Along with several other modeling centers, the European Centre for Medium-Range Weather Forecasts (ECMWF) uses the 'RRTMG' (Rapid Radiative Transfer Model for General Circulation Models; Mlawer et al., 1997) CKD gas-optics model, which employs 112 terms in the shortwave and 140 in the longwave. However, there is a significant variation; Hogan et al. (2017) reported that the number of k-terms in seven global weather forecast models spanned more a factor of 3.7. This leads us to ask: what is the minimum number of k-terms that a CKD model needs while still being sufficiently accurate for weather and climate applications?

This paper describes the ECMWF free-software tool 'ecCKD', which allows users to generate CKD models tailored for their own applications. We aim to adopt some of the best features of CKD models reported in the literature. For example, Edwards and Slingo (1996) described a radiation scheme that is *flexible* in the sense that the spectral discretization and gas optical properties are configured at run-time by a spectral file. Their use of the concept of 'equivalent extinction' put additional constraints on the calculations required by the downstream solver such that the spectral file was, in practice, only compatible with the Edwards and Slingo (1996) radiation scheme. We use a self-describing netCDF file that consists of little more than a set of look-up tables, one per gas. While these files can be read by ECMWF's 'ecRad' radiation scheme (Hogan and Bozzo, 2018), it would in principle be straightforward to interface them to other radiation schemes.

The CKD method works by grouping non-contiguous parts of the spectrum, where the gaseous optical properties are similar, into a single pseudo-monochromatic spectral interval. One of the most interesting developments to the original CKD technique is the full-spectrum correlated-k (FSCK) approach (e.g. Modest and Zhang, 2002), based on the insight that for clear-sky radiative transfer, the parts of the spectrum grouped together in one interval can be very separated in wavelength, thus dispensing with the need for bands. Since fewer bands leads to fewer intervals overall, this is powerful way to reduce computational cost. The FSCK idea has been demonstrated in the atmosphere in the shortwave (Pawlak et al., 2004) and longwave (Hogan, 2010). The ecCKD tool is capable of producing CKD models using either traditional bands or treating the whole spectrum in a single band. It can also produce hybrid models to address the challenge posed by the large spectral variation in surface albedo and cloud optical properties in the NIR.

The classical CKD method involves reordering the absorption spectra separately at different heights and assuming perfect rank correlation between these spectra, implicitly allowing radiation to change wavelength as it traverses the atmosphere. We prefer each spectral interval of a CKD scheme to correspond to a unique set of wavelengths, independent of height, an approach taken by, for example, Bennartz and Fischer (2000), Hogan (2010) and Doppler et al. (2014). This has sometimes been referred to as the *uncorrelated* k-distribution method, although we prefer to consider this as a variant of the CKD method since its accuracy still relies on the high correlation of absorption spectra at each height, even though perfect rank correlation is not assumed. By reporting the spectral mapping in the spectral file, the optical properties of clouds and aerosols can be averaged accurately to each spectral interval. Furthermore, by allowing optical properties to be specified in individual spectral intervals, rather than only in bands, we overcome the problem identified by Lu et al. (2011) that the optical properties of clouds and water vapor are correlated within a band.

The paper is organized as follows. The steps of the method are described in section 2, and are illustrated via the production of two candidate CKD models in the shortwave and two in the longwave. Section 3 evaluates these models using independent LBL calculations for 50 independent atmospheric profiles, and in section 4 the models are evaluated in cloudy skies.

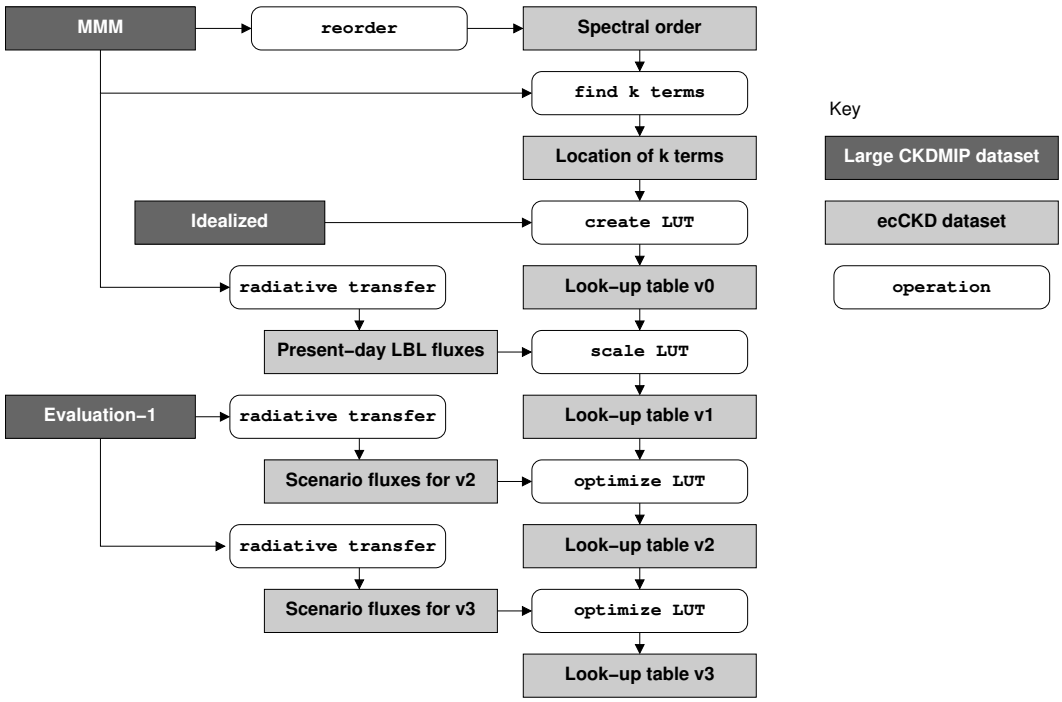

Figure 1. Flowchart illustrating the steps of the ecCKD method (white boxes), which make use of the large CKDMIP datasets of Hogan and Matricardi (2020) and store intermediate information in smaller netCDF files (light grey boxes), ultimately producing a look-up table (LUT) file for use in a radiation scheme.

2 Method

2.1 Overview

The ecCKD tool performs a sequence of tasks shown in Fig. 1, each of which is configurable by the user. Rather than computing atmospheric absorption spectra from scratch, ecCKD takes as input pre-computed absorption spectra from the Correlated K-Distribution Model Intercomparison Project (CKDMIP; Hogan and Matricardi, 2020), performing LBL radiation calculations on them as needed. The datasets are described in section 2.2, and the representation of gases in section 2.3. The user specifies the bands to be used (section 2.4), within which the spectrum is reordered separately for each gas (section 2.5). Sections 2.6 and 2.7 describe how each band is partitioned into spectral intervals, also known as *k*-terms or *g*-points, according to a user-specified error tolerance. Section 2.8 then describes how an initial set of gas-absorption look-up tables (LUTs) is created for each gas and each *k*-term, which constitutes a functioning but possibly inaccurate CKD model in the form of a file that can be used to configure the gas optical properties in a radiation scheme. The subsequent steps then refine these LUTs; in the shortwave the absorptions are scaled so as to produce an exact profile of direct irradiances for each *k*-term for one particular representative atmosphere (section 2.9). In both the longwave and shortwave, a number of optimization steps are performed to refine the LUT coefficients in order to minimize the errors in irradiances and heating rates for a set of training profiles (section 2.10).

2.2 Data

The ecCKD tool makes use of the CKDMIP LBL spectral absorption dataset described by Hogan and Matricardi (2020), appropriate for the terrestrial atmosphere. The absorbing gases considered are H_2O, O_3, CO_2, CH_4, N_2O, O_2, N_2, CFC-11 and CFC-12. Except for

H_2O, the molar absorption of all these gases can be considered independent of their concentration, so a very wide range of climate scenarios can be considered by simply scaling the absorptions. According to Meinshausen et al. (2017), this choice of gases represents 94.5% of anthropogenic greenhouse warming in terms of longwave radiative forcing since 1750, and a further 38 more minor gases (representing the remaining 5.5% of radiative forcing) can be adequately represented by using an increased 'equivalent' concentration of CFC-11. This approach is common in the various phases of the Coupled Model Intercomparison Project given that most or all climate-model radiation schemes are unable to represent all the 43 well-mixed greenhouse gases listed by Meinshausen et al. (2017). However, it should be noted that there is no reason in principle why other gases could not be added to ecCKD if LBL calculations were performed to provide additional absorption spectra in the appropriate format.

As described by Hogan and Matricardi (2020), the spectral resolution of the dataset is variable, being finest in the strong absorption bands of CO_2 because at mesospheric altitudes the CO_2 lines are only Doppler broadened and thus become very narrow. This results in a total of 7 211 999 spectral points in the longwave and 3 126 494 in the shortwave. The total volume of the dataset is around 1 TB.

2.3 Gas representation

Ultimately, the CKD model produced by ecCKD computes the molar absorption coefficient k_i (the absorption cross section per mole of *all* gases) in spectral interval i at pressure p and temperature T as the sum of the contribution from m gases, as follows:

$$k_i(p,T,\psi_1\cdots\psi_m) = k_i^0(p,T) + \sum_{j=1}^{l} k_i^j(p,T)\psi_j + \sum_{j=l+1}^{n} k_i^j(p,T,\psi_j)\psi_j + \sum_{j=n+1}^{m} k_i^j(p,T)\times(\psi_j-\psi_j^{\mathrm{ref}}), \quad (1)$$

where ψ_j is the mole fraction of gas j. The four terms on the right-hand-side represent the four different ways that gases can be represented. The first is the *background* term, a 2D look-up table representing the combined contribution from all gases with a constant, pre-defined mole fraction. For a model intended for climate simulations, this would typically include only O_2 and N_2, but for a CKD model targeting present-day NWP we could include the contribution from well-mixed greenhouse gases. The fact that the LUT includes a dependence on pressure means that a pressure dependence of these gases can be represented. The second term represents gases 1 to l, whose absorption varies *linearly* with mole fraction; in this case k_i^j is the molar absorption coefficient of gas j, i.e. the absorption cross-section per mole of the gas. The third term represents gases $l+1$ to n whose absorption varies *nonlinearly* with concentration; in this case a 3D LUT is used for molar absorption coefficient, with an additional dependence on the mole fraction of the gas in question. In the terrestrial atmosphere only H_2O is in this category, and the representation here allows the contribution from the water vapour continuum (both self and foreign) to be treated completely with no need to separate the line and continuum contributions to the absorption, as is done by many existing CKD models. The fourth term has what we refer to as a *relative-linear* dependence of absorption on mole fraction: a 'reference' mole fraction, ψ_j^{ref}, is defined for the gas, typically the mean surface present-day concentration. The absorption by present-day concentrations of the gas is then folded into the background term, while the fourth term represents the additional absorption due to perturbations (which may be negative) of concentrations from ψ_j^{ref}.

It is up to the user which of the four representations to use for each gas, and the choice depends particularly on what range of greenhouse-gas concentrations will need to be simulated by the target CKD model. The example models generated in this paper are intended to simulate the climate scenarios proposed by Hogan and Matricardi (2020). In both the shortwave and longwave, we represent O_2 and N_2 absorption by the background term, CO_2 and O_3 absorption as linear terms, H_2O as a nonlinear term, and CH_4 and N_2O as relative-linear terms. In the

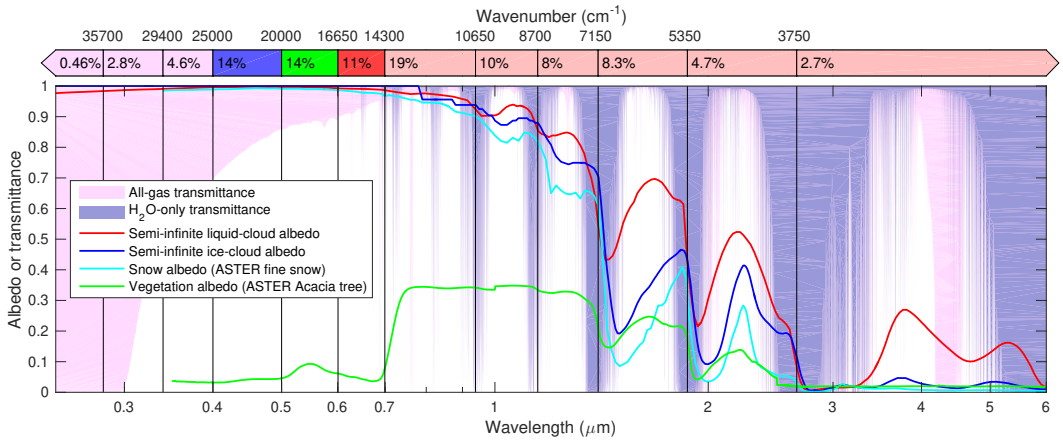

Figure 2. Spectral features of surface and atmosphere informing shortwave band selection. The shading indicates the zenith transmittance of the CKDMIP median atmosphere due to both all gases (including Rayleigh scattering) and water vapor only. The red and dark blue lines depict the albedo of semi-infinite liquid and ice clouds with effective radii of 10 and 30 μm, respectively, the latter assuming optical properties for the Baum et al. (2014) 'General Habit Mixture'. The cyan and green lines depict the albedo of fine snow and vegetation (Acacia tree) from the ASTER dataset. The bar at the top shows the fraction of incoming solar energy in a number of candidate bands.

longwave, CFC-11 and CFC-12 are represented by linear terms, while in the shortwave they are neglected (see Table 4 of Hogan and Matricardi, 2020).

2.4 Band selection

The selection of bands is entirely specified by the user, while the partitioning of each band into k-terms is automated. Many CKD models select band boundaries in order to minimize the number of absorbing gas species in a band, to ensure that the assumption of random overlap of gas absorption is valid, or to cap the error due to assuming the Planck function to be constant across each band. Since ecCKD is not subject to these limitations (any number of gases can be handled with arbitrary overlap, and the Planck function is computed exactly for each k-term), the choice of bands is driven by (1) the need to represent spectral variations of the properties of clouds, aerosols and the surface, and (2) the needs of downstream users for irradiances in specific bands.

In the longwave, the radiative effect of clouds is dominated by a fairly narrow range of wavelengths in the infrared atmospheric window. We therefore hypothesize that adequate accuracy can be achieved without the use of bands, i.e. treating the entire longwave spectrum as a single band. This is verified *a posteriori* in section 4, provided that the optical properties of clouds are computed separately per k-term.

In the shortwave there are multiple concerns that suggest the need for band boundaries, as illustrated in Fig. 2. From the point of view of gases, the important line absorption occurs almost entirely at wavelengths longer than 0.625 μm (wavenumbers less than 16 000 cm^{-1}). At shorter wavelengths, continuum absorption and Rayleigh scattering dominate, and since these tend to vary monotonically with wavelength, the k-terms selected automatically by ecCKD in this range tend to span contiguous ranges of wavelength and therefore behave exactly as bands. From the perspective of Numerical Weather Prediction (NWP) and reanalysis applications, there are several specific bands that are useful for downstream products, and it is preferable to specify them manually rather than leave ecCKD to place k-terms at arbitrary locations in

wavelength. Photosynthetically Active Radiation (PAR) is a common product consisting of the surface downwelling irradiance in the range 0.4–0.7 μm. For generating color imagery (similar to that produced by Lopez, 2020), this can be further split at 0.5 and 0.6 μm. The ecCKD tool works in wavenumber rather than wavelength, with bands specified to the nearest 50 cm^{-1}, so in the example shortwave CKD models generated in this paper we use wavenumbers of 14 300, 16 650, 20 000 and 25 000 cm^{-1} to bound the 'red', 'green' and 'blue' (RGB) bands. The solar energy in each of these is shown in Fig. 2. For more specialist applications additional bands could be considered, such as finer visible bands for modeling of marine biology (e.g. Ciavatta et al., 2014), or finer ultraviolet bands for computing photolysis rates or UV index.

In the NIR part of the shortwave spectrum, the optimum CKD model produced by ecCKD for clear skies over surfaces with spectrally constant albedo would treat the entire region (wavelengths longer than around 0.7 μm) in a single band, and indeed this is the FSCK approach taken by Pawlak et al. (2004). However, as shown in Fig. 2, there are very large variations in both cloud and surface albedo in this region, with sharp changes tending to occur in the regions of largest gas absorption. This suggests that to capture cloud and surface heating rates, band boundaries should be located between some of the NIR windows, particularly at 1.4 μm (around 7150 cm^{-1}). To reconcile these two conflicting requirements, ecCKD offers the capability to use 'sub-bands', which we demonstrate in section 2.7: only parts of the NIR spectrum that are optically thin enough for clouds and the surface to be important are partitioned into sub-bands, while the wavenumbers corresponding to optically thick parts of the spectrum are treated as a single band.

Thus, the CKD models generated in this paper to demonstrate the capabilities of ecCKD use the FSCK approach in the longwave. In the shortwave we use the 'RGB' band structure: a single ultraviolet band, red, green and blue bands, and a single NIR band. The latter employs sub-bands to allow for spectrally varying cloud and surface properties.

2.5 Reordering the spectrum

The longwave and shortwave parts of the spectrum are considered separately. The first task shown in Fig. 1 is to reorder the spectrum of each gas in order of increasing absorption within each band. We seek a unique mapping independent of height, and therefore aim to sort the high-resolution CKDMIP spectra in terms of the approximate height of the peak cooling in the longwave and height of the peak heating in the shortwave, using the median CKDMIP atmosphere with present-day (2020) greenhouse gas concentrations. In the longwave we follow a method very similar to that proposed by Hogan (2010): for each gas a LBL radiative transfer calculation is performed with all other gas concentrations set to zero and an idealized profile of temperature increasing linearly with the logarithm of pressure from -100°C at 0.01 hPa to $+15$°C at 1000 hPa. Sorting is in order of the height of peak cooling rate. This method fails for low column optical depths, τ, where heating rate peaks at the surface, so when $\tau < 0.5$ we sort by τ instead. In the shortwave a simpler approach is taken: the spectra are ordered by the height at which the optical depth from top-of-atmosphere (TOA) reaches 0.25, which is the height at which direct radiation for a solar zenith angle of 60° will have fallen to around 60% of the TOA value.

The result of the reordering is written into a spectral-order file for each gas, containing the integer rank r of each wavenumber point. Suppose the entire spectrum contains N discrete wavenumbers indexed 1 to N, and a particular band corresponds to wavenumbers indexed m to n. The ranks r_m to r_n will consist of the integers m to n but reordered. In the following sections we follow previous authors and introduce a coordinate variable for the reordered spectrum, g. In ecCKD, this simply maps the integer ranks for the bands to the range 0–1, i.e. an element of the reordered spectrum with integer rank r in a particular band would have $g(r) = (r - r_m)/(r_n - r_m)$.

2.6 Partitioning g space for individual gases

The next step shown in Fig. 1 is to partition each band into k-terms, each corresponding to a fixed set of wavenumber points. In a radiation scheme, each k-term would be treated by an independent quasi-monochromatic radiative transfer calculation, so generally more terms correspond to a more accurate but more computationally costly scheme. We use a refined version of the algorithm described by Hogan (2010): first the spectrum is partitioned separately for each gas in each band (described in this section), then the partitions for each gas are merged taking account of the spectral overlap of gases (described in section 2.7).

For each gas and band we need to find the number of intervals, n, into which g space (defined in the previous section) is to be divided, and their boundaries $g_0, g_1 \cdots g_n$, with the upper and lower bounds already defined as $g_0 = 0$ and $g_n = 1$. Most previous papers require the user to specify n and define the boundaries according to a fixed mathematical rule such as Gaussian Quadrature (e.g. Kato et al., 1999), but this is not adaptive to the spectra of individual gases. In ecCKD, the user provides a single error tolerance, and the tool attempts to partition g space such that the error associated with each g interval is approximately equal to this tolerance. The smaller the tolerance, the more g intervals will be needed and a greater overall accuracy should be achieved. The appendix describes a general algorithm for partitioning g space given a function $E(g_{i-1}, g_i)$ that returns the error associated with treating the wavenumbers corresponding to the range g_{i-1} to g_i by a single quasi-monochromatic radiation calculation. This function is similar to a cost function in estimation theory, and following Hogan (2010) is formulated as the mean squared error in heating rate over l layers, but with an additional term (weighted by f) penalizing errors in surface and TOA irradiances:

$$E(g_{i-1}, g_i) = \sum_{j=1}^{l} w_j \left(H_j^{\mathrm{CKD}} - H_j^{\mathrm{LBL}}\right)^2 + f\left[\left(F_{\uparrow\mathrm{TOA}}^{\mathrm{CKD}} - F_{\uparrow\mathrm{TOA}}^{\mathrm{LBL}}\right)^2 + \left(F_{\downarrow\mathrm{surf}}^{\mathrm{CKD}} - F_{\downarrow\mathrm{surf}}^{\mathrm{LBL}}\right)^2\right], \tag{2}$$

where H_j^{CKD} and H_j^{LBL} are the heating rates predicted by the ecCKD and line-by-line models in layer j, while $F_{\uparrow\mathrm{TOA}}$ and $F_{\downarrow\mathrm{surf}}$ are the upwelling irradiance at TOA and the downwelling irradiance at the surface, respectively. In order to weight the stratosphere and troposphere on an approximately equal basis, the vertical profile is weighted by the square-root of pressure, i.e. the weight term is given by $w_j = (p_{j+1/2}^{1/2} - p_{j-1/2}^{1/2})/p_{l+1/2}^{1/2}$, where $p_{j+1/2}$ is the pressure at the interface between layers j and $j+1$, and $p_{l+1/2}$ is the surface pressure.

Since this function is called multiple times by the algorithm described in the appendix, it cannot be too computationally costly, so partitioning is performed using only a single profile. For temperature and the concentration of the target gas, we use the median present-day profile from the CKDMIP 'MMM' dataset. The partitioning needs to account for the presence of other gases, which can dominate in some parts of the spectrum. This tends to reduce the error associated with representing the target gas and therefore reduces the number of g intervals required, but depends on the concentration of these other gases. We use the minimum concentration of these gases that the CKD model is intended to simulate. For H_2O and O_3 we use the 'minimum' concentration profile from the CKDMIP MMM dataset, which contains the minimum values in any of the 25 000 profiles of the original Eresmaa and McNally (2014) dataset. The minimum concentrations of the well-mixed greenhouse gases depends on what application the CKD model is to be used for. For a CKD model to be used solely in NWP, it is appropriate to simply use present-day concentrations for the well-mixed greenhouse gases. For simulation of past and future climate, we use the 'Glacial Maximum' values proposed by Hogan and Matricardi (2020), i.e. the minimum greenhouse gas concentrations found in the last million years.

In the longwave, the LBL radiative transfer calculation is performed with the present-day concentration of the target gas and the 'minimum' concentrations of all other gases, and when the function $E(g_{i-1}, g_i)$ is called, the LBL terms in (2) are computed simply by summing

the LBL spectral irradiances from the wavenumbers corresponding to the requested range of g. A single zenith angle is used in each hemisphere (equivalent to the two-stream method but without scattering). The 'CKD' terms in (2) are computed by first averaging the optical depths of the target gas across the wavenumbers corresponding to the requested range of g, but retaining the full spectral resolution for the other gases. This way (2) quantifies the error purely associated with approximating the target gas. The user can select the method used to average the optical depths of the target gas, the default being a linear average of the layer transmittances weighted by the Planck function at the temperature of the layer.

In the shortwave, the radiative transfer calculations are limited to the direct (unscattered) solar beam, which contains almost all of the sensitivity to gas absorption, and reduces computational cost. Thus, the upwelling terms in (2) are omitted and the heating rates consider only heating by the direct beam. Optical depths are averaged weighting by the TOA incoming solar spectral irradiance, which following Hogan and Matricardi (2020) is taken to be the 1986–2018 average of the Coddington et al. (2016) climate data record.

Shortwave partitioning includes the option to use sub-bands in the NIR, as introduced in section 2.4. Since water vapor dominates in the NIR, this can be achieved by dividing water vapor alone into sub-bands. Each of the high-resolution spectral points in the NIR band has both a wavenumber and a g value indicating the water vapor absorption strength in the NIR region. The spectral points with $g < g_{\text{crit}}$, where g_{crit} is some user-specified critical value, are deemed to be optically thin enough that sub-bands are needed, so these points are further grouped according to their wavenumber into user-specified sub-bands, although within the sub-bands the ordering by g is preserved. For $g \geq g_{\text{crit}}$, water vapor is optically thick enough that no grouping by wavenumber is needed. This is illustrated graphically at the end of the next section.

2.7 Partitioning g space for multiple gases

After each of the gases have had their reordered spectra partitioned into intervals in g space, they are combined to obtain a final set of k-terms. This is achieved using the 'hypercube partition method' of Hogan (2010): for m active gases in a particular band we consider an m-dimensional unit hypercube where dimension j represents the g space for gas j. Each wavenumber point in the band lies at a particular location in this m-dimensional space, and an individual k-term represents a subregion of the full hypercube such that the wavenumber points that lie within the subregion for a particular k-term are then treated together in a single quasi-monochromatic radiative transfer calculation.

Hogan (2010) described an optimal way to partition the hypercube into k-terms, and illustrated this in both the $m = 2$ case (two gases; his Fig. 4) and $m = 3$ case (his Fig. 5). The first k-term consists of the intersection of the first g interval for each gas, i.e. the weakest absorption. The remaining k-terms each correspond to one of the remaining g intervals for one of the gases, thereby 'specializing' in that gas. Thus, if the number of intervals required to partition the g space for gas j is n_j, then the number of k-terms required for m gases is $n_{\text{total}} = 1 + \sum_j^m (n_j - 1)$. The method supports arbitrary spectral overlap of the spectra of individual gases, and is an improvement on the approach of Ritter and Geleyn (1992) and Edwards (1996), which requires $n_{\text{total}} = 1 + \sum_j^m n_j$ terms and makes the assumption that the spectra of individual gases are randomly overlapped.

Figure 3 illustrates the resulting partitioning of the spectrum for the two longwave and two shortwave example CKD models. Vectorization and computational efficiency favors n_{total} being a power of two, so we have chosen the error tolerances to obtain 16 or 32 k-terms. Consider first Fig. 3c, showing a model in which the entire longwave spectrum has been treated as a single band and partitioned into a total of 32 k-terms, in approximate order of increasing mean optical depth. All k-terms except the first are dominated by the contribution from one particular gas, indicated on the right of the panel. The 'full-spectrum' nature of this CKD model is clear from widely separated parts of the spectrum being represented by single quasi-

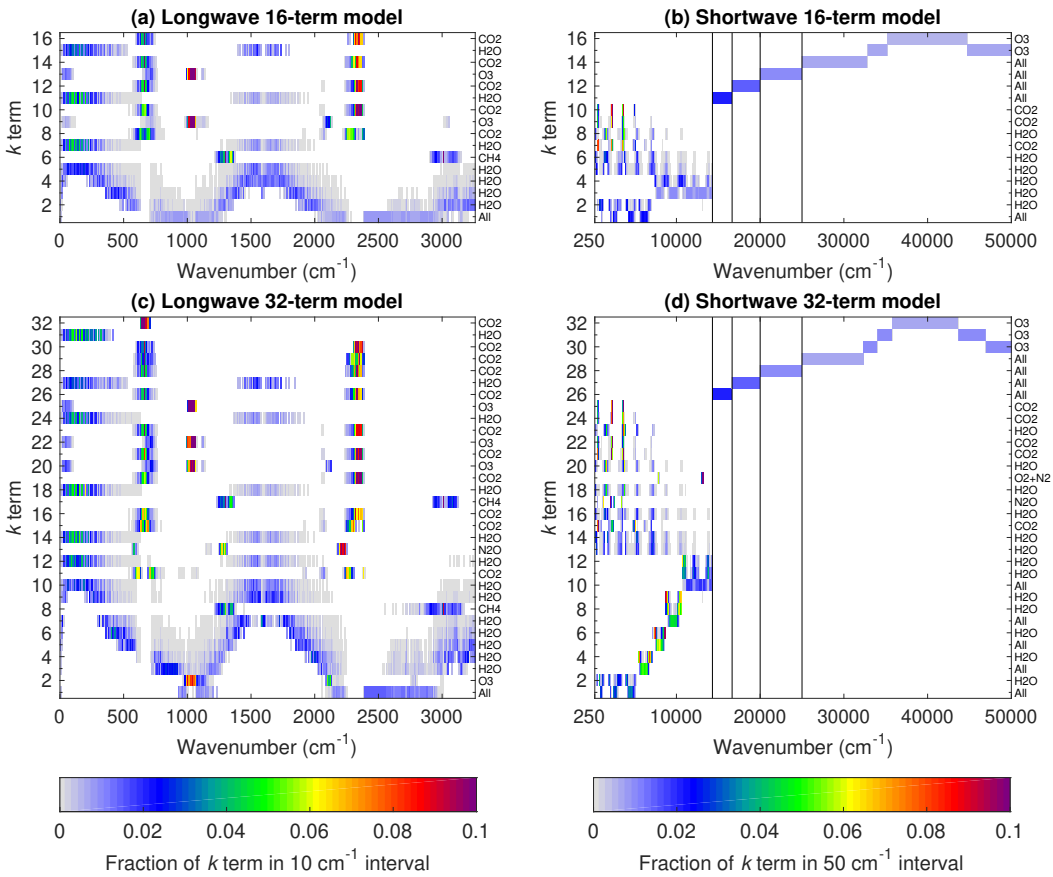

Figure 3. The contribution of each part of the spectrum to each k-term in the four CKD models considered in the text. The fractions sum to unity along each row. The main gas represented by each k term is indicated on the right-hand-side of each panel, although in principle all gases can contribute to the optical depth for all k terms. The first (least optically thick) k-term in each band or sub-band is marked 'All' since its boundaries are determined by all gases.

monochromatic k-terms; for example, term 28 represents strong CO_2 absorption from both the 4.3 μm and the 15 μm CO_2 bands.

The 32-term shortwave model in Fig. 3d is very different. The vertical black lines delineate the user-specified 'RGB' bands explained in section 2.4. The partitioning algorithm finds that only one k-term is required for each of the red, green and blue bands, and only four for the entire ultraviolet. Terms 30 and 31 each combine the effects of similar levels of O_3 absorption on opposite sides of the Hartley band, which peaks at around 40 000 cm^{-1}. In the NIR, we use sub-bands for water vapor as explained in section 2.6, with a value of $g_{\mathrm{crit}} = 0.7$ and additional wavenumber partitions at 5350, 7150, 8700 and 10650 cm^{-1} (bounding most of the NIR windows in Fig. 2). In Fig. 3d we see that k-terms 1–12 represent the intersection of regions of relatively weak water-vapor absorption ($g < g_{\mathrm{crit}}$) with weak absorption in all the other gases; these use five sub-bands in order to resolve spectral variations in clouds and the surface. The k-terms 13–25 represent regions of strong water vapor absorption ($g \geq g_{\mathrm{crit}}$) or strong absorption by one of the other gases. The 16-term shortwave model in Fig. 3b takes the same approach but with reduced k-terms via the use of $g_{\mathrm{crit}} = 0.65$ and only one additional wavenumber partition at 7150 cm^{-1}.

2.8 Creating initial look-up table

This step creates a first estimate of the LUTs in (1) using the 'Idealized' CKDMIP dataset, and indeed we use the same points (described in section 3.3 of Hogan and Matricardi, 2020): 53 logarithmically spaced points in pressure with 10 points per decade; 6 points in temperature, 20 K apart; and 12 logarithmically spaced points in water vapor mixing ratio with 2 points per decade. As shown in Fig. 1, the creation of the LUT involves reading in the location of the k-terms, i.e. a file containing the indices of the wavenumber points in the high-resolution spectrum that contribute to each term. The layer optical depths in the Idealized dataset for the relevant wavenumber points are averaged to each k-term separately for each gas, weighting by the local Planck function in the longwave and the solar spectral irradiance in the shortwave. As in section 2.6, the default averaging method is linear in layer transmittance for a zenith angle of 60°. The final step in the creation of the LUT is to convert from layer optical depth to molar absorption coefficient as used in (1).

The LUT file contains additional variables that are added at this point and remain unchanged by the subsequent steps shown in Fig. 1. The fraction of the spectrum contributing to each k-term (i.e. the information shown in Fig. 3) is provided to enable subsequent averaging of cloud, aerosol and surface properties to k-terms, with a resolution of 10 cm^{-1} in the longwave and 50 cm^{-1} in the shortwave. In the longwave we provide the Planck function for each k-term as a LUT versus temperature between 120 and 350 K at 1 K intervals, computed by simply integrating the Planck function over the wavenumber points contributing to each k-term. In the shortwave we provide the solar spectral irradiance for each k-term. Also provided is the Rayleigh molar scattering coefficient, computed for each wavenumber using the Bucholtz (1995) formula and averaged across the parts of the spectrum contributing to each shortwave k-term weighted by the solar spectral irradiance.

2.9 Scaling shortwave look-up table entries

At this point, the LUT entries have been computed only from consideration of the spectroscopy at the local pressure level, and do not necessarily perform well in radiative transfer traversing multiple levels. One of the reasons for this in the shortwave is that each k term represents the average of a range of absorption strengths that are highly correlated in the vertical. As the solar beam traverses the atmosphere, radiation in the more optically thick parts of the spectrum is attenuated more rapidly. This means that lower in the atmosphere the optically thick parts are less important and the effective average molar absorption coefficient for the k-term should be lower than the one computed weighting by the TOA solar spectral irradiance, as in section 2.8. A similar effect occurs in the longwave, so in both parts of the spectrum non-local effects need to be considered in order to derive the optimum LUT entries. This section describes the first part of this refinement in the shortwave, while section 2.10 describes a subsequent more general optimization performed in both the shortwave and longwave.

For a single atmospheric profile of temperature, pressure and gas concentrations, it is possible to derive a profile of layer optical depths for each k-term that reproduces the LBL profile of direct-beam shortwave irradiance exactly, for a particular value of the cosine of the solar zenith angle μ_0. If we define $F_{j+1/2}$ as the LBL direct irradiance at the interface between layers j and $j+1$ (counting down from TOA) integrated over the parts of the spectrum corresponding to an individual k-term, then the Beer-Lambert law states that $F_{j+1/2} = F_{j-1/2}\exp(-\tau_j/\mu_0)$. This can be inverted to obtain τ_j, the effective optical depth of layer j.

In the 'scale LUT' step in Fig. 1 we take this approach using the median present-day profile from the CKDMIP 'MMM' dataset and $\mu_0 = 1/2$, yielding a profile of τ_j values for each k-term. The same values are computed using the ecCKD v0 LUT, and the ratio of the LBL and ecCKD optical depths is calculated to provide a correction factor that varies with pressure and k term. The correction factor is then interpolated to the pressure grid of the LUT and all the molar absorption coefficients in the file are multiplied by it, producing v1 of the LUT. We have made some significant assumptions here: that the absorption of all gases should

be modified by the same proportion, and that the correction factor does not vary with the other dimensions of the LUT (temperature and water-vapor mixing ratio). Nonetheless, shortwave radiative transfer calculations using the v1 LUT are significantly more accurate than v0, and there is still the opportunity (described in the following section) for a global optimization of all the coefficients in the LUT.

2.10 Optimizing look-up table entries

The final task is to optimize the coefficients of the LUTs in order to minimize the errors in predicted irradiance and heating-rate profiles in a set of training profiles in a least-squares sense. The need for this step in terms of non-local dependencies was explained at the start of section 2.9, but it also tunes the coefficients to mitigate any errors caused by simplifications in the formulation of the CKD model. For example, (1) assumes that the average optical depths from each individual gas in a k term can be simply summed, whereas Zhang et al. (2003) argued that a more complex formulation was necessary (including much narrower bands) to treat non-random spectral overlap in the parts of the spectrum contributing to a k-term. We find that this complexity and additional cost is unnecessary if the coefficients can be optimized as described in this section.

We improve the basic method of Hogan (2010) in a number of ways. As shown in Fig. 1, the optimization may be performed in several steps, as we have found that better performance is achieved if the major gases are optimized first, with minor gases being optimized individually in subsequent steps. We define the state vector $\mathbf{x}$ to contain all the variables to be optimized in one of these steps, specifically the natural logarithm of all the non-zero entries in the look-up for each gas being optimized. With 53 points in pressure, 6 in temperature, 12 in water vapor concentration and (for example) 32 k-terms, this leads to 122 112 state variables for water vapor and 10 176 for each other gas. The cost function to be minimized is given by

$$J = (\mathbf{x} - \mathbf{x}_a)^{\mathrm{T}} \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_a) + \sum_{j=1}^{p} E_j. \tag{3}$$

The first term ensures the stability of the minimization by penalizing the squared differences between the state vector and the *a priori* LUT elements $\mathbf{x}_a$, i.e. those from the previous step in Fig. 1. The error covariance matrix $\mathbf{B}$ provides a complete description of the weighting of this term, with its diagonal elements containing the square of the user-specified root-mean-squared (RMS) error $\mathbf{x}_a$. We use a value of 8 in this paper, allowing the natural logarithm of the LUT elements to stray significantly from their prior values in the optimization, although in practice the RMS difference between the elements of $\mathbf{x}$ before and after optimization for an individual gas is around 0.25. The off-diagonal elements of $\mathbf{B}$ specify error covariances between LUT values, and have the important effect of spreading information provided by the training profiles into adjacent parts of the LUT. We model the error correlation coefficient of adjacent LUT coefficients along the pressure, temperature and water-vapor-concentration axes as ρ, and coefficients n steps apart along these axes as ρ^n. No correlation is assumed between k-terms or gases. Even though $\mathbf{B}$ is large, its inverse is very sparse and the first term in (3) is efficient to compute. Empirically we find that $\rho = 0.8$ provides the best results. The second term in (3) expresses the sum of the squared errors in heating rates and irradiances at TOA and the surface over p atmospheric profiles, where E_j has the same form as in (2) but using broadband heating rates and irradiances.

The cost function is minimized using the quasi-Newton 'L-BFGS' algorithm of Liu and Nocedal (1989), which requires the vector of gradients $\partial J/\mathbf{x}$ to be computed. This is achieved by coding the entire algorithm in C++ using the combined automatic differentiation, array and optimization library 'Adept' (Hogan, 2014), version 2.1 of which also includes an implementation of the L-BFGS algorithm. An optimization step typically takes several tens of minutes to complete.

Naturally, for a CKD model to be used in climate projections we wish to calculate not only the most accurate profiles of irradiances and heating rates, but also the radiative forcing associated with perturbations to both major and minor greenhouse gases. Unfortunately, if all gases are optimized simultaneously, the scheme tends to adjust minor gases to try to offset errors in major gases. This problem can be overcome via several separate optimization steps as shown in Fig. 1, first for the major gases and then the minor. We acknowledge that the multi-step approach is somewhat 'ad hoc' and there is scope to improve it in future versions of the software, but as will be shown in section 3, it does produce models that can accurately compute radiative forcing.

In the case of CKD models targeting climate applications (such as those depicted in Fig. 3), the first step optimizes the coefficients of H_2O, O_3, CO_2 and the background term in (1). The background term represents not only O_2 and N_2, but also the present-day 'reference' concentrations of CH_4 and N_2O, with these gases all treated at this stage as having a constant mixing ratio with pressure. The training data consist of LBL calculations performed on the 50 'Evaluation-1' CKDMIP profiles for six CKDMIP CO_2 scenarios (surface concentrations from 180 to 2240 ppmv), i.e. a total of 300 profiles.

In the second step the coefficients of CH_4 are optimized using LBL calculations on the CKDMIP scenarios in which CH_4 is perturbed from 350 to 3500 ppbv. To avoid the CH_4 coefficients being tuned to correct for remaining errors in the previous step, we train on the difference in heating-rate profiles and irradiances between perturbed and present-day CH_4 calculations, which is equivalent to bias-correcting the present-day CKD calculations from the previous step. This ensures the CH_4 coefficients are optimized to give the most accurate radiative forcing when perturbed from present-day concentrations. The third step takes exactly the same approach but optimizes the N_2O coefficients training on the Evaluation-1 profiles in which concentrations are perturbed in the range 190–540 ppbv. In the shortwave this yields the 'final' LUT ready for use in a radiation scheme. In the longwave we perform one further step to optimize the coefficients of CFC-11 and CFC-12.

3 Clear-sky evaluation

In this section we evaluate the gas-optics models generated in the previous section in clear skies. The CKDMIP Evaluation-2 dataset is used, which consists of LBL calculations on 50 independent profiles, including ones with extremes of temperature, ozone and humidity. Surface longwave emissivity and shortwave albedo are spectrally constant at 1.0 and 0.15, respectively, the latter being the approximate mean albedo of the Earth's surface. This approach is the same as that of Hogan and Matricardi (2020) to evaluate the RRTMG gas-optics model, except that they used the CKDMIP Evaluation-1 dataset. Naturally, when the ecCKD models are evaluated against the Evaluation-1 dataset (not shown) they perform better because these profiles were used in the training.

Figure 4 evaluates the performance of the FSCK-16 and FSCK-32 models, for present-day greenhouse gas concentrations. Even though relatively few *k*-terms are used compared to other CKD models, the errors are small; the root-mean-squared (RMS) error in heating rates from the surface to the upper stratosphere (4 hPa) is only 0.15 and 0.1 K d^{-1} for the 16- and 32-term models, respectively, approximately doubling in the mesosphere.

Figure 5 evaluates the instantaneous radiative forcing associated with perturbing the five main greenhouse-gas concentrations from their present-day values. Note that the CFC-11 concentrations here correspond to artificially increased values to approximately represent 38 further greenhouse gases (Meinshausen et al., 2017). In large part, both models capture the forcing associated with large perturbations to concentrations, including up to eight times preindustrial concentrations of CO_2. The 16-term model performs slightly worse in some scenarios, tending to underestimate the magnitude of the surface forcing associated with reducing CO_2 concentrations to glacial-maximum values of 180 ppmv, as well as struggling with the

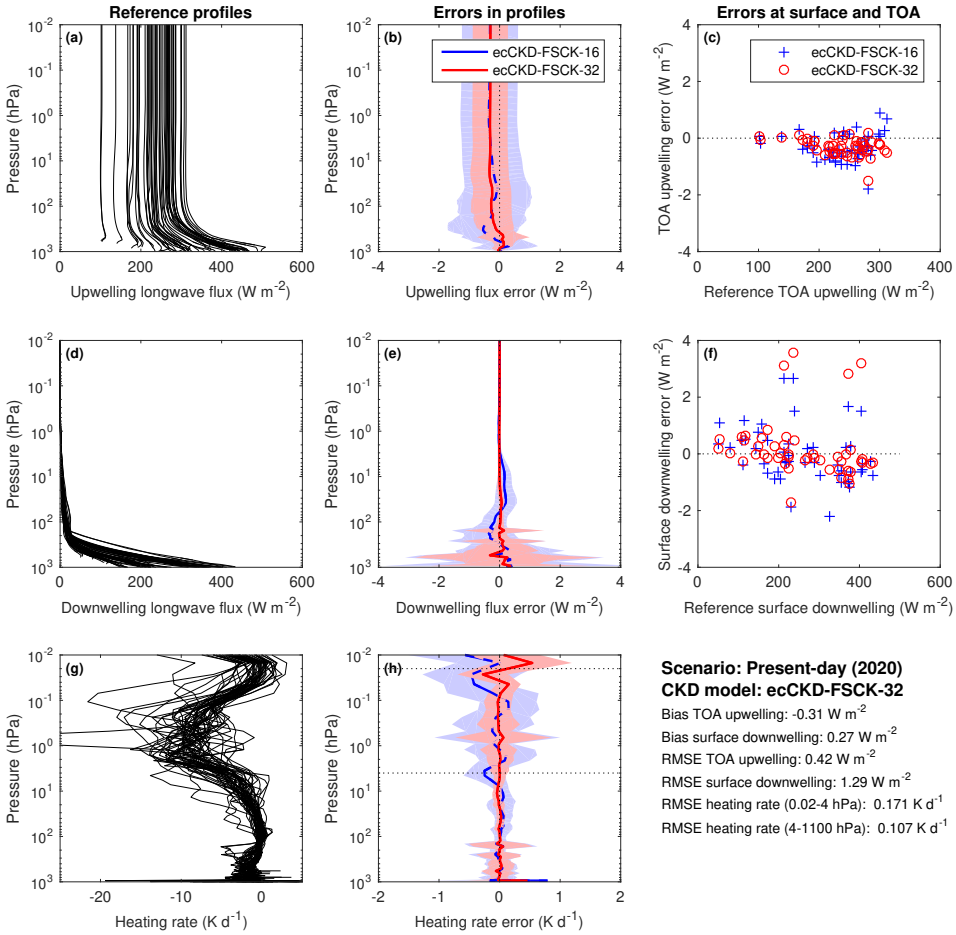

Figure 4. Evaluation of clear-sky longwave irradiances and heating rates from the two ecCKD models for the 50 independent profiles of the CKDMIP Evaluation-2 dataset with present-day concentrations of the well-mixed greenhouse gases. Panels a, d and g show quantities from the reference LBL calculations, while panels b, e and h show the corresponding biases in the ecCKD calculations using an identical radiative transfer solver with four angles per hemisphere. The shaded regions encompass 95% of the errors (estimated as 1.96 multiplied by the standard deviation of the error). Panels c and f depict instantaneous errors in upwelling TOA and downwelling surface irradiances. The statistics of the comparison are summarized in the lower right, including the root-mean-squared error (RMSE) in heating rate (weighted by the cube-root of pressure) in two ranges of pressure indicated by the horizontal dotted lines in panel h.

extreme CH_4 concentrations. As can be seen in Figs. 3a and 3c, the improvement of the 32-term model for CH_4 can be attributed to its use of two CH_4-specific k terms, rather than just one.

Figure 6 presents the corresponding present-day evaluation of irradiances and heating rates for the RGB-16 and RGB-32 shortwave models. Again, the errors are modest given the small number of k-terms, with the RMS error in heating rates from the surface to 4 hPa being 0.1 and 0.06 K d^{-1} for the 16- and 32-term models, respectively. The much larger mid-mesosphere heating-rate error for the 16-term model is associated with its poorer representation of the 4.3 μm CO_2 band; Figs. 3b and 3 show that it used only three CO_2-specific k-terms, compared to five for the 32-band model. This also explains the difference in how well the two models capture the shortwave CO_2 forcing shown in Fig. 7. This figure also indicates

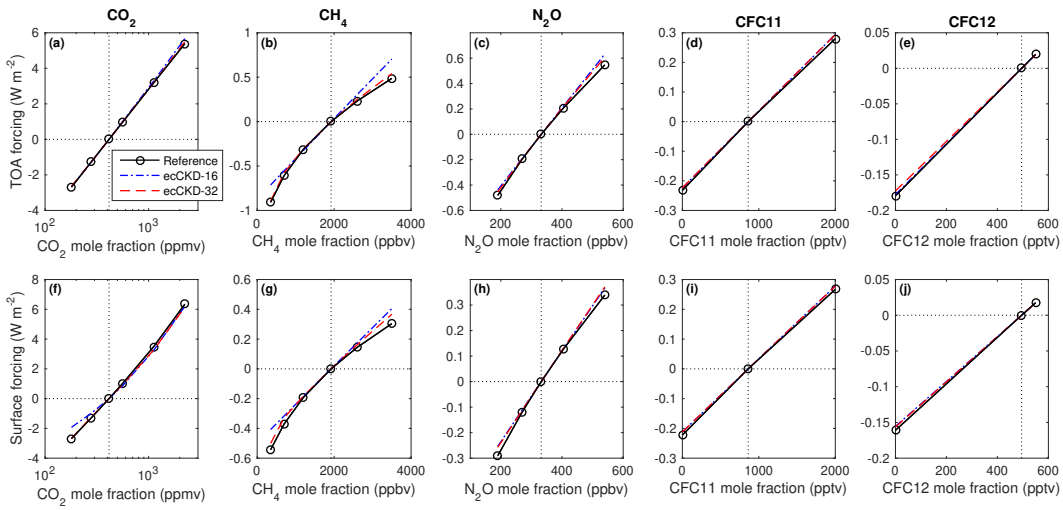

Figure 5. Comparison of reference LBL and ecCKD calculations of the instantaneous longwave clear-sky radiative forcing from perturbing each of the five well-mixed greenhouse gases from their present-day (2020) values at (a–e) top of atmosphere and (f–j) the surface, averaged over the 50 profiles of the CKDMIP Evaluation-2 dataset. The black circles correspond to scenarios 5–22 proposed by Hogan and Matricardi (2020).

Table 1. Summary of the properties of the two atmospheric profiles used to evaluate the representation of liquid and ice clouds in section 4. Both are taken from the CKDMIP dataset. The size distribution of the liquid cloud was modeled as a gamma distribution with a shape parameter of 2.

Profile	Liquid cloud	Ice cloud
Cloud effective radius	10 μm	30 μm
Cloud pressure range	726.6–907.1 hPa	184.5–457.3 hPa
Optical properties	Mie theory	Baum et al. (2014)
Origin	Evaluation-2 profile 29	Evaluation-1 profile 28
Location	31.4°S, 3.5°W	38.9°N, 25.6°W
Date and time	19 March 2014, 18 UTC	11 March 2014, 00 UTC
Surface pressure	1017.4 hPa	1021.6 hPa
Surface temperature	22.2°C	14.7°C

that the CH_4 forcing in the two models is similar; in fact neither model uses CH_4-specific *k*-terms, but rather includes the optical-depth contribution of CH_4 in all the other *k*-terms. The 16-term model also has no N_2O-specific *k*-terms and Fig. 7 shows that this leads to it tending to overestimate the N_2O forcing by around a factor of two (although the magnitude of the shortwave forcing of this gas is only a tenth of the longwave). The 32-term model introduces a single N_2O-specific *k*-term and is able to achieve a much greater accuracy.

4 Cloudy-sky evaluation

As discussed in section 2.4, clouds, aerosols and the surface can exhibit significant spectral variations in optical properties. One of the features of ecCKD is that each *k*-term has a unique mapping to specific parts of the spectrum (see Fig. 3), and this mapping is available to downstream applications (such as the ecRad radiation scheme) so that optical properties can be averaged separately for each *k* term. Nonetheless, the use of the full-spectrum correlated-*k*

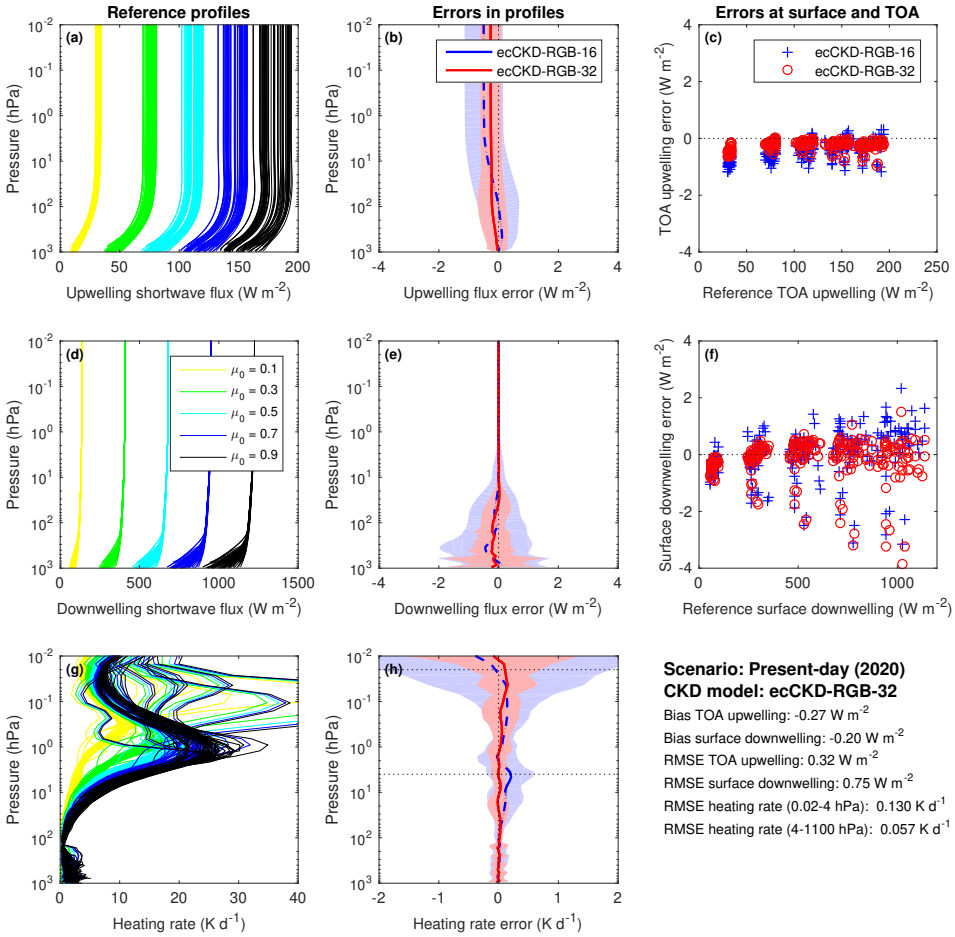

Figure 6. Similar to Fig. 4 but for the shortwave. The reference LBL calculations in panels a, d and g are for all 50 CKDMIP Evaluation-2 profiles at five values of the cosine of the solar zenith angle, μ_0 (0.1, 0.3, 0.5, 0.7 and 0.9). The subsequent evaluation considers all 250 combinations. The five clusters of points in panels c and f correspond to the five values of μ_0.

approach means that individual k-terms can represent widely separated points in the spectrum. In this section we test the impact on the accuracy of calculations of the radiative effect of liquid and ice clouds, using two real-world profiles from the CKDMIP dataset, summarized in Table 1. In each case, 26 LBL calculations have been performed for water paths ranging from 10^{-4} to 10 kg m^{-2}, plus an additional clear-sky calculation, with a vertically constant cloud mixing ratio between two pressure bounds. The optical properties of liquid clouds are computed using Mie theory at 396 wavenumbers from 5 to 50 000 cm^{-1}, while the ice properties are taken from the Baum et al. (2014) 'General Habit Mixture' available at 445 wavenumbers between 101 and 50 251 cm^{-1}. When used in LBL calculations, the mass-extinction coefficient, single scattering albedo and asymmetry factor are interpolated linearly in wavenumber space, but clamped when used at wavenumbers outside the range provided. The radiative transfer calculations use a no-scattering solver in the longwave, and the two-stream method with a solar zenith angle of 60° in the shortwave, both from the CKDMIP software package. The plane-parallel approximation is adopted, i.e. clouds are taken to be horizontally uniform with a cloud fraction of unity.

The equivalent ecCKD calculations use a version of the ecRad offline radiative transfer package that supports ecCKD gas-optics models, and a radiative transfer solver equivalent to

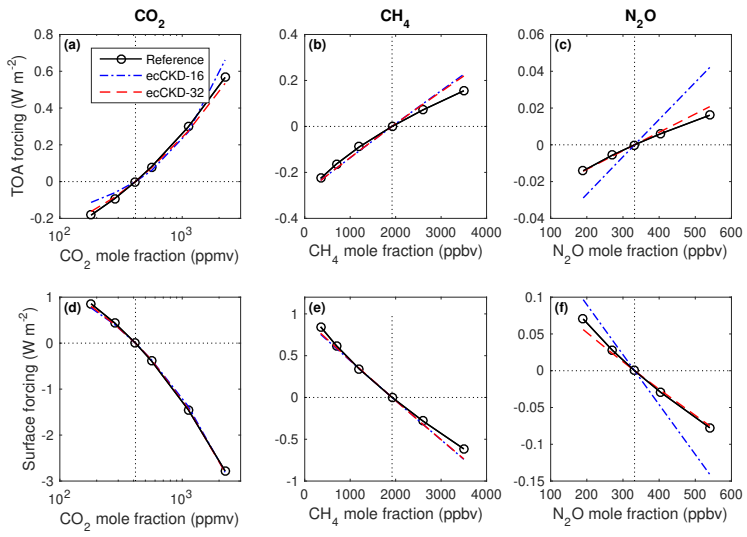

Figure 7. As Fig. 5 but for the instantaneous shortwave radiative forcing by CO_2, CH_4 and N_2O. The results for the five solar zenith angles have been averaged, so the values shown here represent a daytime average.

that used for the LBL calculations. Cloud optical properties are computed by averaging the same Mie and Baum et al. (2014) data according to the parts of the spectrum corresponding to each individual k-term (as shown in Fig. 3). Following the approach of Edwards and Slingo (1996), we consider both 'thin' and 'thick' spectral averaging. The former is appropriate in the optically thin limit and simply involves averaging of the mass-extinction coefficient, the mass-absorption coefficient, and averaging asymmetry factor weighted by scattering coefficient. The latter is more appropriate in the optically thick limit and is intended to provide the exact cloud albedo in the limit of infinite optical depth (although in practice it is not exact in this limit except in the absence of gas absorption). For a little extra accuracy, we apply delta-Eddington scaling (Joseph et al., 1976) before performing the spectral average. A further weighting is used in the averaging to approximately represent the energy at each wavenumber; in the longwave we use the Planck function at a representative atmospheric temperature of 0°C and in the shortwave at an effective solar temperature of 5777 K. In the longwave, no benefit was found from using a different reference temperature for liquid and ice clouds. For comparison we also show the results of the RRTMG gas-optics model using the same cloud optical properties, but since no information is available on the exact wavenumbers used for each of its k-terms, the optical properties are averaged to its 16 longwave and 14 shortwave bands.

Figure 8a depicts the LBL calculations of 'true' longwave cloud radiative effect at TOA and the surface for the liquid-cloud profile versus water path, with the error in these quantities shown in Figs. 8b and 8c. The errors for all models are generally less than 1 W m^{-2}. The RRTMG model performs best, although the errors associated with the two ecCKD models are still small, being up to around 5% for the FSCK-16 model and 2% for the FSCK-32 model. Naturally, the RRTMG model with its 16 longwave bands is able to capture the spectral variation of cloud optical properties, but it is nonetheless surprising how well the FSCK models perform when they consider the entire longwave spectrum in a single band. This is because (as revealed by the LBL calculations shown in Fig. S1 of the Supporting Information) over 92% of the radiative effect of this cloud at the surface and TOA is in the 8–13 μm (769–1250 cm^{-1}) longwave atmospheric window, within which the variation of cloud optical properties is quite modest, and certainly much less than in the NIR. Outside the longwave atmospheric window, clouds make a much weaker contribution to longwave cloud radiative effect either because of

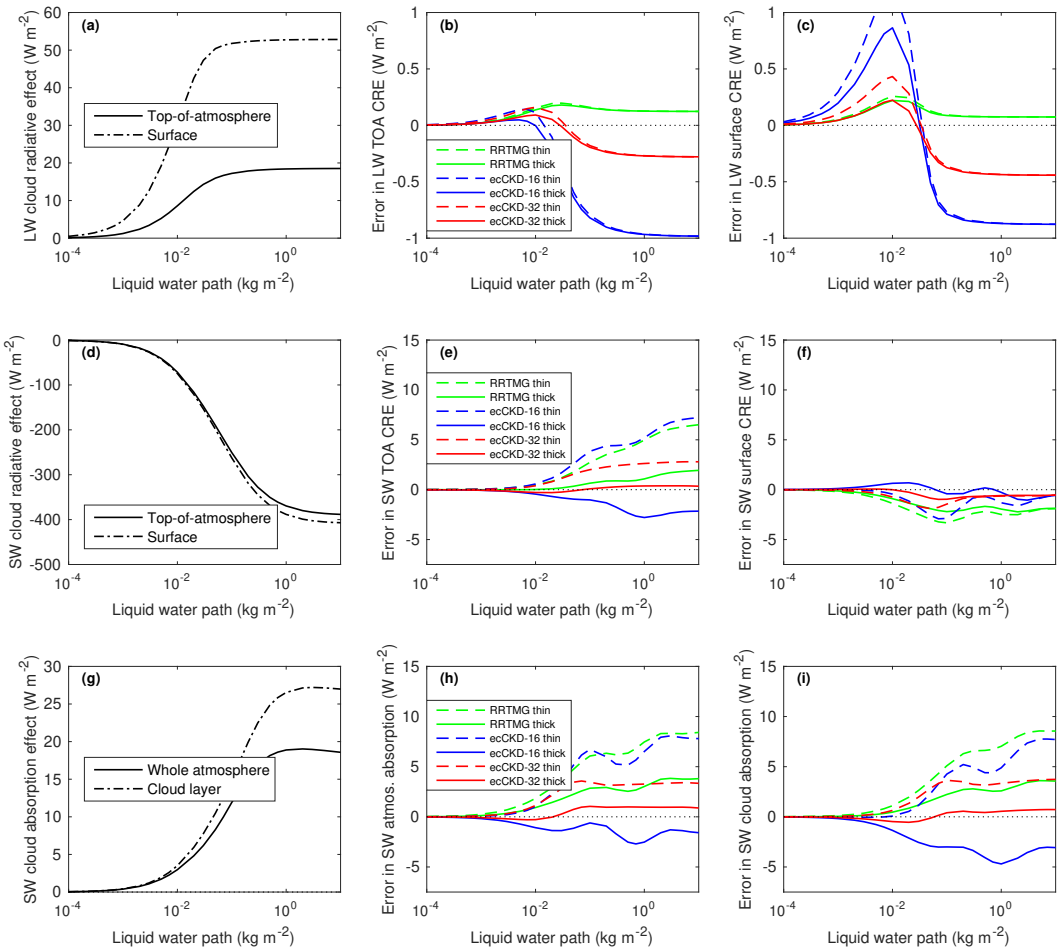

Figure 8. Error in cloud radiative effect (CRE) due to spectral discretization of RRTMG and the various ecCKD models, for a low-level liquid cloud with effective radius 10 μm in Profile 29 of the CKDMIP Evaluation-2 dataset. Panel a depicts the longwave cloud radiative effect (i.e. the change to net irradiance due to cloud) at top-of-atmosphere (TOA) and the surface as a function of LWP, for the LBL calculations. Panels b and c depict the error in these quantities for the various CKD models, and for thin and thick spectral averaging. Panels d–f show the same but in the shortwave with a solar zenith angle of 60°. Panel g shows the 'cloud absorption effect', i.e. the absorption by the entire atmosphere, and by the cloud layer alone, minus the corresponding clear-sky absorptions. Panels h and i show the error in these quantities for the CKD models.

the much stronger gas absorption or the much weaker Planck function. The top row of Fig. 9 shows the equivalent evaluation for the ice cloud profile, where the errors for all longwave gas-optics models are even less, both in an absolute and a relative sense, due to ice particles having less variation in their optical properties than liquid droplets across the longwave spectrum (see Fig. S2 in the Supporting Information). Overall, these results demonstrate that the longwave FSCK method is a viable approach for use in weather and climate models.

The middle row of Fig. 8 depicts the equivalent evaluation but in the shortwave where the magnitude of the cloud radiative effect is much larger. The best performing model is clearly RGB-32 using 'thick' averaging, with errors of only around 0.4 W m^{-1} (0.1%) for any value of liquid water path. This provides *a posteriori* justification for the use of five NIR sub-bands in Fig. 3d, bounded at the points shown in Fig. 2 where cloud optical properties tend to change most rapidly. The RGB-16 model incurs a larger error due to its employing only two

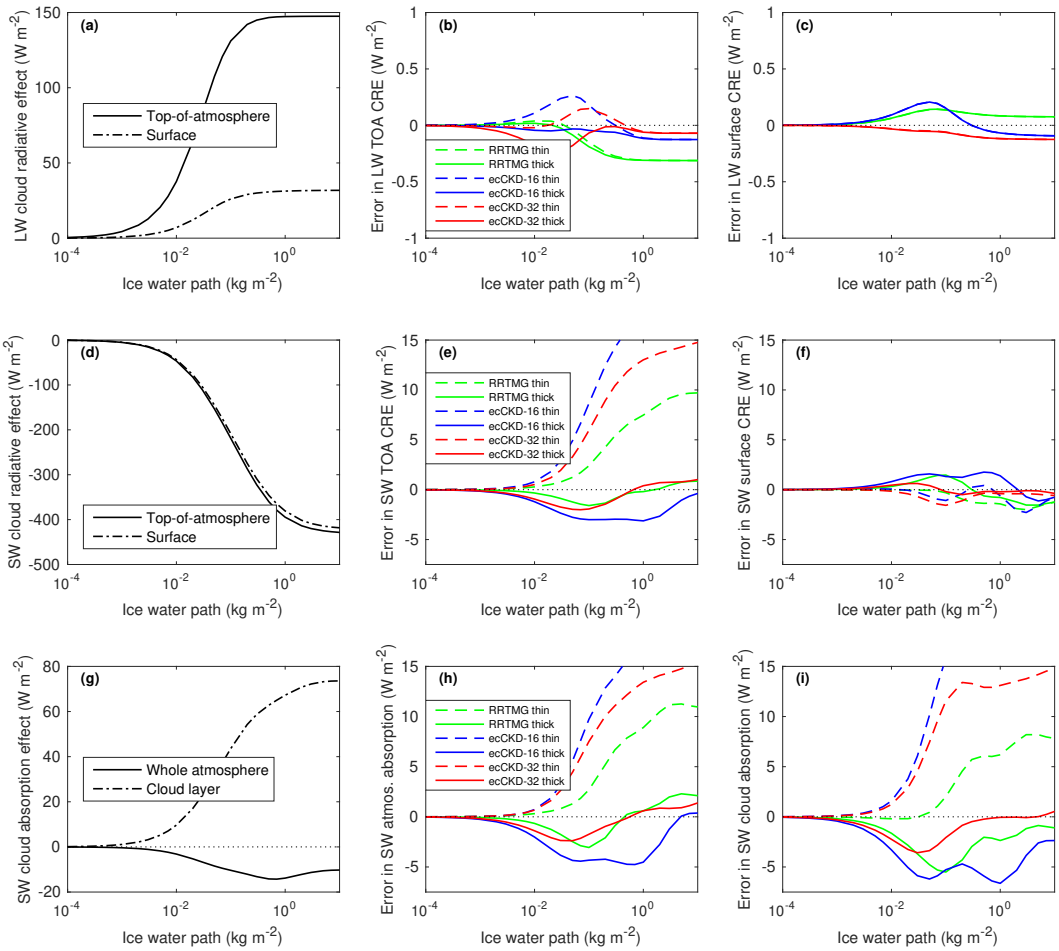

Figure 9. As Fig. 9 but for an ice cloud with an effective radius of 30 μm in Profile 28 of the CKDMIP Evaluation-1 dataset.

sub-bands. The result for ice clouds in the middle row of Fig. 9 are broadly similar. It is clear from Figs. 8e and 9e that the most accurate calculations are achieved using thick rather than thin spectral averaging, for any of the gas-optics models. This conclusion is different from that of Edwards and Slingo (1996), who advocated thick averaging for liquid clouds and thin for ice clouds.

The bottom rows of Figs. 8 and 9 consider the effect of the cloud on shortwave absorption, both by the entire atmosphere and by the cloud layer alone. Again, the RGB-32 model with thick averaging performs best. One of the interesting features of Fig. 9g is that the effect of the ice cloud is to increase shortwave absorption in the cloud layer itself, as would be expected, but to reduce absorption overall by reflecting sunlight that would otherwise have been absorbed by gases lower in the atmosphere. Thus, the sign of the impact of the cloud on whole-atmosphere absorption is dependent on two competing effects, and while the absolute magnitude of the errors shown in whole-atmosphere and cloud-layer absorption (Figs. 9h and 9i) are similar, the relative error in the latter is much larger; indeed, the RGB-16 model with thin averaging predicts that the effect of the ice cloud on atmospheric absorption is to increase rather than to decrease it.

5 Conclusions

In this paper, we have introduced a free software tool 'ecCKD' for generating fast correlated-k-distribution (CKD) gas-optics models for use in the radiation schemes of atmospheric models. The CKD models generated are both accurate and efficient, needing considerably fewer k-terms than most others in the literature. This is achieved via the use of algorithms to optimally partition the k distribution for each gas, and to optimize the look-up table coefficients for each gas in order to minimize errors against hundreds of training profiles (extending the approach of Hogan, 2010). In the shortwave, the introduction of 'sub-bands' enables the full-spectrum correlated-k (FSCK) approach to treat the entire NIR as a single band, while still enabling the large spectral differences in cloud and surface albedo to be resolved.

The tool has been demonstrated by generating and testing CKD models with only 16 and 32 k-terms in each of the shortwave and longwave, i.e. nearly a factor of 8 and 4 times fewer, respectively, than the total number used operationally at ECMWF. When evaluated against independent data, the 32-term models are shown to be very accurate, with RMS heating-rate errors of less than 0.18 K d^{-1} from the troposphere to the mid-mesosphere. The radiative forcing of the main anthropogenic greenhouse gases is captured accurately, including CO_2 varying over a factor of 12 and CH_4 over a factor of 10. The 32-term models have been found to perform well when run online in the ECMWF forecast model, to be explored in a future paper. The 16-term models are naturally somewhat less accurate, but would be suitable for short forecasts such as 12-hour forecasts performed repeatedly in a data-assimilation cycle in which efficiency is paramount.

We have used LBL calculations for profiles containing liquid and ice clouds with a large range of water contents to verify the accuracy of the FSCK approach in cloudy situations. In the longwave, provided that cloud properties are averaged over each k-term rather than per band, errors in irradiances calculated using the 32-term model are less than 0.5 W m^{-2}, an important demonstration of the viability of the longwave FSCK approach for cloudy terrestrial atmospheres. In the shortwave, the use of sub-bands in the NIR gives the 32-term ecCKD model similar accuracy to RRTMG but using only 25 rather than 78 terms in the NIR. We also find that 'thick' averaging of cloud spectral properties (Edwards and Slingo, 1996) provides the best agreement with LBL calculations for both liquid and ice clouds.

The tool described in this paper offers a number of opportunities for users of radiation schemes. Principally, it allows optimized CKD models to be generated for specific applications, from present-day NWP to palaeoclimate simulations of periods when atmospheric composition was very different. Moreover, the fact that the CKD models generated tend to be faster while of similar accuracy to existing models frees up computer time to improve the accuracy of other parts of the radiation scheme, such as the use of more than two streams (e.g. Fu et al., 1997), inclusion of 3D effects (e.g. Hogan et al., 2016), and calling the scheme more frequently in time and space (e.g. Hogan and Bozzo, 2018). Additionally, the use of a simple look-up table to compute optical depths (Eq. 1) makes it straightforward to incorporate the CKD models into different types of radiation scheme, including explicit 3D solvers (e.g. Jakub and Mayer, 2016).

Appendix A Equipartition algorithm

Section 2.6 outlined the partitioning of a reordered spectrum into intervals such that each interval was associated with around the same mean-squared error in a radiation calculation, and less than a user-specified tolerance E_{tol}. The 1D space to be partitioned is denoted g and ranges from 0 to 1. We seek the boundaries of n intervals, denoted $g_0, g_1 \cdots g_n$, such that the following two conditions are satisfied:

$$E(g_{i-1}, g_i) \leq E_{\mathrm{tol}} \qquad \text{for all } i; \tag{A1}$$

$$F \leq F_{\mathrm{tol}}, \tag{A2}$$

where the error $E(g_{i-1}, g_i)$ is a non-differentiable user-supplied function, and the second condition states that the fractional range of errors, $F = [\max(E) - \min(E)]/\bar{E}$, should be no larger than the user-supplied tolerance F_{tol}, typically 0.02. In our case, $E(g_{i-1}, g_i)$ is given by (2) and involves LBL calculations with a computational cost proportional to the width of the interval $g_i - g_{i-1}$. Therefore, a good partitioning algorithm should not require an excessive number of calculations of E, especially ones for wide g intervals. We are not aware of an off-the-shelf algorithm for performing this partitioning, so this appendix describes our solution to the problem. While it is not likely to be the fastest possible algorithm, it almost always converges to a solution that satisfies the conditions above.

The first task is to find the number of intervals required, n, in order that condition (A1) is satisfied. This is achieved by partitioning g space starting at the lower end such that for each interval (except possibly the last), $0.95E_{\text{tol}} \leq E \leq E_{\text{tol}}$. We start with a test value of $g_1 = 0.75$ and compute $E(0, g_1)$, noting that there is a lower bound on g_1 of 0 where $E(0, 0) = 0$. If the result is less than $0.95E_{\text{tol}}$ then a new lower bound for g_1 has been found, and the next test value is selected by extrapolating (but not beyond $g_1 = 1$) from the old and new lower bounds to where we would expect $E(0, g_1) = E_{\text{tol}}$ assuming a linear variation of E with g_1. On the other hand, if $E(0, g_1) > E_{\text{tol}}$ then an upper bound for g_1 has been found, and the next test value is found by linearly interpolating between the lower and upper bounds on g_1. The new test value is used to compute $E(0, g_1)$ and the process is repeated until either $0.95E_{\text{tol}} \leq E \leq E_{\text{tol}}$, or $g_1 = 1$ and $E \leq E_{\text{tol}}$. If there is remaining g space to partition then the process is repeated to compute g_2 and so on, until the process returns $g_n = 1$. We now know how many intervals are required, and have candidate values for g_i, but usually the error associated with the final interval, $E(g_{n-1}, 1)$, is significantly less than all the other errors.

The second task is to find the interior boundaries of the intervals $(g_1, g_2 \cdots g_{n-1})$ such that condition (A2) is satisfied. Note that the outer boundaries are already fixed at $g_0 = 0$ and $g_n = 1$. In the simple case of $n = 2$ we have only a single value to find, g_1; it is straightforward to progressively refine this value until (A2) is satisfied. In the more general $n > 2$ case, we use the candidate values of g_i from the first task above to compute the cumulative error as

$$C(g_i) = \sum_{j=1}^{i} E(g_{j-1}, g_j). \tag{A3}$$

A new set of candidate values g_i' is found by attempting to repartition the total error, $C(1)$, evenly amongst the n intervals. This is achieved by using linear interpolation into the C function to compute the g_i' values such that $C(g_i') = iC(1)/n$. The errors are recomputed and the process is repeated until (A2) is satisfied. If at any iteration the fractional range F increases then the iteration is not successful and instead a 'shuffle' step is performed. This consists of looping through adjacent pairs of intervals and adjusting the g point between them until their errors agree to within 2%. Thus, for intervals 1 and 2 we adjust g_1 until $E(g_0, g_1)$ and $E(g_1, g_2)$ satisfy (A2), then do the same for intervals 2 and 3 and so on up to intervals $n-1$ and n, followed by a pass back down to intervals 1 and 2. This is usually enough that subsequent partitioning iterations using (A3) lead to a reduction of F. Any further shuffle operations proceed in the opposite direction through g space as the previous one.

Acknowledgments

We thank Steve English and Richard Forbes for constructive comments on the original manuscript. This work has benefited from valuable discussions with Robert Pincus.

Data Availability Statement

The ecCKD software is available from https://github.com/ecmwf-ifs/ecckd. It makes use of data from the CKDMIP project available via the links at http://confluence.ecmwf.int/

display/CKDMIP. The netCDF files defining the 16- and 32-term gas-optics models described in this paper are available at the ecCKD web site https://confluence.ecmwf.int/x/XwU0Dw.

References

Barker, H. W., Qu, Z., Dhanraj, V., and Cole, J. N. S. (2021) Partial validation of a lossy compression approach to computing radiative transfer in cloud system-resolving models. *Q. J. R. Meteorol. Soc., 147,* 363–381, https://doi.org/10.1002/qj.3922.

Baum, B. A., Yang, P., Heymsfield, A. J., Bansemer, A., Merrelli, A., Schmitt, C., & Wang, C. (2014). Ice cloud bulk single-scattering property models with the full phase matrix at wavelengths from 0.2 to 100 μm. *J. Quant. Spectrosc. Radiat. Transfer, 146,* 123–139, https://doi.org/10.1016/j.jqsrt.2014.02.029.

Bennartz, R., & Fischer, J. (2000). A modified *k*-distribution approach applied to narrow band water vapour and oxygen absorption estimates in the near infrared. *J. Quant. Spectrosc. Radiat. Transfer, 66,* 539–553, https://doi.org/10.1016/S0022-4073(99)00184-3.

Bucholtz, A. (1995). Rayleigh-scattering calculations for the terrestrial atmosphere. *App. Opt, 34,* 2765–2773, https://doi.org/10.1364/AO.34.002765.

Ciavatta, S., Torres, R., Martinez-Vicente, V., Smyth, T., Dall'Olmo, G., Polimene, L., Icarus Allen, J. (2014). Assimilation of remotely-sensed optical properties to improve marine biochemistry modelling. *Progress in Oceanography, 127,* 74–95, https://doi.org/10.1016/j.pocean.2014.06.002.

Coddington, O., Lean, J. L., Pilewskie, P., Snow, M., & Lindholm, D. (2016). A solar irradiance climate data record. *Bull. Am. Meteorol. Soc., 97,* 1265–1282, https://doi.org/10.1175/BAMS-D-14-00265.1.

DeAngelis, A., Qu, X., Zelinka, M., & Hall, A. (2015). An observational radiative constraint on hydrologic cycle intensification. *Nature, 528,* 249–253, https://doi.org/10.1038/nature15770.

Doppler, L., Preusker, R., Bennartz, R., & Fischer, J. (2013). 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, 133,* 382–395, https://doi.org/10.1016/j.jqsrt.2013.09.001.

Edwards, J. M. (1996). Efficient calculation of infrared fluxes and cooling rates using the two-stream equations. *J. Atmos. Sci., 53,* 1921–1932, https://doi.org/10.1175/1520-0469(1996)053⟨1921:ECOIFA⟩2.0.CO;2.

Edwards, J. M., & Slingo, A. (1996). Studies with a flexible new radiation code: 1. Choosing a configuration for a large-scale model. *Q. J. R. Meteorol. Soc., 122,* 689–719. https://doi.org/10.1002/qj.49712253107.

Eresmaa, R., & McNally, A. P. (2014). Diverse profile datasets from the ECMWF 137-level short-range forecasts. NWP-SAF Document NWPSAF_EC_TR_017, available from https://www.nwpsaf.eu/site/software/atmospheric-profile-data/.

Fu, Q., & Liou, K. N. (1992). On the correlated *k*-distribution method for radiative transfer in nonhomogeneous atmospheres. *J. Atmos. Sci., 49,* 2139–2156, https://doi.org/10.1175/1520-0469(1992)049⟨2139:OTCDMF⟩2.0.CO;2.

Fu, Q., Liou, K. N., Cribb, M. C., Charlock, T. P., & Grossman, A. (1997). Multiple scattering parameterization in thermal infrared radiative transfer. *J. Atmos. Sci., 54,* 2799–2812, https://doi.org/10.1175/1520-0469(1997)054⟨2799:MSPITI⟩2.0.CO;2.

Goody, R., West, R., Chen, L., & Crisp, D. (1989). The correlated-k method for radiation calculations in nonhomogeneous atmospheres. *J. Quant. Spectrosc. Radiat. Transfer, 42,* 539–550, https://doi.org/10.1016/0022-4073(89)90044-7.

Hogan, R. J. (2010). The full-spectrum correlated-*k* method for longwave atmospheric radiation using an effective Planck function. *J. Atmos. Sci., 67,* 2086–2100, https://doi.org/10.1175/2010JAS3202.1.

Hogan, R. J. (2014). Fast reverse-mode automatic differentiation using expression templates in C++. *ACM Trans. Mathematical Softw., 40,* 26:1–26:16. https://doi.org/10.1145/

2560359.

Hogan, R. J., & Bozzo, A. (2018). A flexible and efficient radiation scheme for the ECMWF model. *J. Adv. Modeling Earth Sys., 10,* 1990–2008, https://doi.org/10.1029/2018MS001364.

Hogan, R. J., & M. Matricardi (2020). Evaluating and improving the treatment of gases in radiation schemes: the Correlated K-Distribution Model Intercomparison Project (CKD-MIP). *Geosci. Model Dev., 13,* 6501–6521, https://doi.org/10.5194/gmd-13-6501-2020.

Hogan, R. J., Ahlgrimm, M., Balsamo, G., Beljaars, A. C. M., Berrisford, P., Bozzo, A., Di Giuseppe, F., Forbes, R. M., Haiden, T., Lang, S., Mayer, M., Polichtchouk, I., Sandu, I., Vitart F., & Wedi, N. (2017). Radiation in numerical weather prediction. ECMWF Tech. Memo. No. 816, 48 pp, http://doi.org/10.21957/2bd5dkj8x.

Hogan, R. J., Schäfer, S. A. K., Klinger, C., Chiu, J.-C., & Mayer, B. (2016). Representing 3D cloud-radiation effects in two-stream schemes: 2. Matrix formulation and broadband evaluation. *J. Geophys. Res., 121,* 8583–8599, https://doi.org/10.1002/2016JD024875.

Jakub, F., & Mayer, B. (2016) 3-D radiative transfer in large-eddy simulations – experiences coupling the TenStream solver to the UCLA-LES. *Geosci. Model Dev., 9,* 1413–1422, http://doi.org/10.5194/gmd-9-1413-2016.

Joseph, J. H., Wiscombe, W. J., & Weinman, J. A. (1976) The delta-Eddington approximation for radiative flux-transfer. *J. Atmos. Sci.*, *33,* 2452–2459, http://doi.org/10.1175/1520-0469(1976)033⟨2452:TDEAFR⟩2.0.CO;2.

Kato, S., Ackerman, T. P., Mather, J. H., & Clothiaux, E. E. (1999). The *k*-distribution method and correlated-*k* approximation for a shortwave radiative transfer model. *J. Quant. Spectrosc. Radiat. Trans., 62,* 109–121, https://doi.org/10.1016/S0022-4073(98)00075-2.

Lacis, A., & Oinas, V. (1991). 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., 96,* 9027-9063, https://doi.org/10.1029/90JD01945.

Lopez, P. (2020). Forecasting the past: views of Earth from the Moon and beyond. *Bull. Am. Meteorol. Soc., 101,* E1190–E1200, https://doi.org/10.1175/BAMS-D-19-0254.1.

Lu, P., Zhang, H., & Li, J. (2011). Correlated k-distribution treatment of cloud optical properties and related radiative impact. *J. Atmos. Sci., 68,* 2671–2688, https://doi.org/10.1175/JAS-D-10-05001.1.

Manabe, S., & Wetherald, R. T. (1967). Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci., 24,* 241–259, https://doi.org/10.1175/1520-0469(1967)024⟨0241:TEOTAW⟩2.0.CO;2.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., & Weiss, R. (2017). Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geosci. Model Dev., 10,* 2057–2116, https://doi.org/10.3929/ethz-b-000191830.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res. Atmos., 102,* 16 663–16 682, https://doi.org/10.1029/97JD00237.

Modest, M. F., & Zhang, H. (2002). The full-spectrum correlated-*k* distribution for thermal radiation from molecular gas-particulate mixtures. *J. Heat Transfer, 124,* 30–38, http://doi.org/10.1115/1.1418697.

Liu, D. C., & Nocedal, J. (1989). On the limited memory method for large scale optimization. *Math. Programming B, 45,* 503–528, https://doi.org/10.1007/BF01589116.

Pawlak, D. T., Clothiaux, E. E., Modest, M. F., & Cole, J. N. S. (2004). Full-spectrum correlated-*k* distribution for shortwave atmospheric radiative transfer. *J. Atmos. Sci., 61,* 2588-2601, https://doi.org/10.1175/JAS3285.1.

Ritter, B. & Geleyn, J. F. (1992). A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Weath. Rev., 120,* 303–325, https://doi.org/10.1175/1520-0493(1992)120⟨0303:ACRSFN⟩2.0.CO;2.

Ukkonen, P., Pincus, R., Hogan, R. J., Nielsen, K. P., and Kaas, E. (2020). Accelerating radiation computations for dynamical models with targeted machine learning and code optimization. *J. Adv. Modeling Earth Sys., 12,* e2020MS002226. https://doi.org/10.1029/2020MS002226.

Zhang, H., Nakajima, T., Shi, G., Suzuki, T., & Imasu, R. (2003). An optimal approach to overlapping bands with correlated k distribution method and its application to radiative calculations. *J. Geophys. Res., 108,* 4641, https://doi.org/10.1029/2002JD003358.