Evaluating and improving the treatment of gases in radiation schemes: the Correlated K-Distribution Model Intercomparison Project (CKDMIP)

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

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

Correspondence: Robin J. Hogan (r.j.hogan@ecmwf.int)

Abstract. This is a draft paper, eventually to be submitted to GMD. It was last updated 28 October 2019.

1 Introduction

The correlated k-distribution (CKD) method (e.g., Lacis and Oinas, 1991) underpins the treatment of gas absorption in the radiation schemes of most current weather and climate models. It accelerates the calculation of broadband irradiance profiles by approximating the integration over hundreds of thousands of spectral lines by a much smaller number (tens to hundreds) of quasi-monochromatic irradiance calculations, which we hereafter refer to as 'g-points'. More g-points means higher accuracy but a larger computational cost, so we have a trade-off to make depending on the application. For climate modelling we require schemes that can accurately compute the radiative forcing of a number of different greenhouse gases over a wide range of concentrations. For short-range weather forecasting with present-day greenhouse gas concentrations, the priority is much more on efficiency: the radiation scheme must be called frequently to capture the local radiative impact of evolving cloud fields, and forecasts must be delivered to customers in a timely fashion. The lower model top in many limited-area weather models also means that, in principle, fewer g-points are required to compute the heating-rate profile.

Unfortunately, the tools and know-how to make this accuracy–efficiency trade-off are available to only a handful of specialists worldwide, with the result that most atmospheric models are available with only one gas-optics configuration, which is often not optimized for the application at hand. Indeed, Hogan et al. (2017) surveyed seven models used for the same application of global weather forecasting, and reported that the total number of g-points (shortwave plus longwave) ranged from 68 to 252. This raises the question as to whether some schemes can achieve the same accuracy with fewer g-points.

The purpose of the Correlated K-Distribution Model Intercomparison Project (CKDMIP) is to address these issues, and specifically:

1. To use benchmark line-by-line calculations to evaluate the accuracy of existing CKD models for applications spanning short-range weather forecasting to climate modelling, and to explore how accuracy varies with number of g-points in individual CKD schemes.

2. To understand how different choices in way that CKD models are generated affects their accuracy for the same number of g-points.

3. To provide freely available datasets and software to facilitate the development of new gas-optics models, with the ultimate aim of producing a community tool to allow users to generate their own gas-optics models targeted at specific applications.

The project has similarities to the Radiative Forcing Model Intercomparison Project (RFMIP; Pincus et al., 2016), which used line-by-line calculations to evaluate the radiation schemes of a number of climate models in terms of surface and top-of-atmosphere (TOA) irradiances for a range of atmospheric profiles and climate scenarios. However, CKDMIP goes further in that it includes weather forecasting applications, and provides the means to improve the way that CKD schemes make the trade-off between accuracy and effi-

ciency. This is possible by making available the spectral optical depth of each layer of the atmosphere due to each gas separately. The CKDMIP software package allows participants to combine the optical depths of the gases they are interested in and perform radiative transfer calculations on the result, producing their own reference profiles of spectral or broadband irradiances and heating rates. Since the per-molecule absorption is (to a very good approximation) independent of concentration for all gases except water vapour, it is simple to produce reference calculations for scaled concentrations of individual gases.

This protocol paper describes the design and generation of these datasets and software, and what comparisons will be performed. Section 2 describes the overarching design decisions of CKDMIP, including which gases to include, which weather and climate applications to target, and for climate modelling which range of gas concentrations to consider. Section 3 describes in detail how the datasets are produced, how the spectral resolution has been chosen and what radiative transfer calculations are performed. Section 4 then describes what is required of CKDMIP participants, the spectral band structures that should be used, the metrics that will be used to quantify errors in irradiances and heating rates, and how errors due to the representation of the spectral variation in cloud properties will be assessed. After the conclusions in section 5, information on obtaining the CKDMIP code and datasets is provided in section 6.

Finally a note on terminology. Throughout this paper we define a *CKD scheme* as a software component (usually embedded within the radiation scheme of an atmospheric model) that takes as input profiles of atmospheric temperature, pressure and the concentrations of a number of gases, and outputs profiles of optical depth for each of a number of g-points. It also includes a means to compute the Planck function to use for each g-point. A *CKD model* is one configuration of a CKD scheme with a particular number of g-points, which might consist of a set of look-up tables that can be used by the CKD scheme. A *CKD tool* is a method (which may be fully automated or involve some hand-tuning) for generating individual CKD models, with some means to control the trade off between accuracy and the number of g-points.

2 Design of evaluation scenarios

2.1 Which gases?

In the shortwave, the choice of which gases to include is relatively uncontroversial: H_2O, O_3, CO_2, O_2 and CH_4. The concentration of N_2 is specified as well, as it is needed (in addition to O_2) to compute the collision-induced contribution to the continuum absorption and the broadening efficiency of these molecules, where applicable.

In the longwave, there are a much larger number of greenhouse gases that could be included, many of which have a very small individual impact. However, the purpose of CKDMIP is to evaluate the techniques used by schemes for generating CKD models based on the different requirements of weather and climate modelling, rather than to produce a single optimum CKD model that explicitly represents all the greenhouse gases that anyone might want to simulate. Therefore we have chosen to follow the pragmatic approach of Meinshausen et al. (2017). They stated that 94.5% of the anthropogenic greenhouse warming (in terms of radiative forcing) between 1750 and 2014 was due to increases in CO_2, CH_4, N_2O, CFC-11 and CFC-12, with the remaining 5.5% being attributable to 38 further gases. Their 'Option 2' approximately represents the radiative forcing of these 38 gases by artificially increasing the concentration of CFC-11 (by around a factor of 3.9 in the present day); indeed, the CMIP6 (Coupled Model Intercomparison Project Phase 6) historic concentrations and future scenarios are available with these 'CFC-11-equivalent' concentrations. From Cycle 47R1, ECMWF's Integrated Forecasting System will take this approach, using concentrations from the CMIP6 SSP3-7.0 scenario (O'Neill et al., 2016). For the CKDMIP activity we therefore limit ourselves to consideration of 9 gases: H_2O, O_3, CO_2, O_2, CH_4, N_2O, CFC-11, CFC-12 and N_2. Note that the longwave absorption of O_2 and N_2 are included; their impacts on net TOA irradiance are around 0.11 and 0.17 W m^{-2}, respectively (Höpfner et al., 2012).

2.2 Numerical weather prediction

Table 1 lists the three main applications for which we envisage that CKD models could be optimized. The first two correspond to present-day Numerical Weather Prediction (NWP) at the local and global scale. Both need to represent variable water vapour and ozone, but to a good approximation can assume all other gases to have a constant mole fraction, or to vary as a function of pressure alone. (Note that since the atmosphere is an ideal gas to a good approximation, we can assume the mole fraction of a gas to be equal to its volume mixing ratio.) In principle, this allows the number of g-points to be reduced since, for example, all the well-mixed gases could be merged into a single 'hybrid' or 'composite' gas whose optical properties vary as a function of temperature and pressure alone (e.g., Ritter and Geleyn, 1992; Niemelä et al., 2001).

In terms of the present-day concentrations of the well-mixed gases, we assume that O_2 and N_2 have a constant volume mixing ratio of 0.20946 and 0.78102 mol mol^{-1}, respectively, independent of pressure (Jones and Schoonover, 2002). These concentrations are also assumed for all past and future scenarios in section 2.3. The present-day surface concentrations of the five other well-mixed gases are shown in

Table 1. The three modelling applications of radiation schemes that we envisage would need to be targeted by a different CKD model. The present-day and variable trace-gas concentrations for these scenarios are provided in Table 2. Heating rates are evaluated at pressures down to the indicated 'lowest pressure', although note that the reference line-by-line calculations are performed down to lower pressures than these.

Application	Lowest pressure	Trace gas concs.
Limited-area NWP	4 hPa	Present-day (2020)
Global NWP	0.02 hPa	Present-day (2020)
Climate	0.02 hPa	Variable

Table 2, and were taken from the CMIP6 SSP3-7.0[1] scenario for calendar year 2020. The vertical profiles of these gases are discussed in section 3.2.

The difference between the two NWP applications listed in Table 1 is in the location of the model top. The model top quoted for all current configurations of the ECMWF model and all global configurations of the Met Office model used for weather and climate is 0.01 hPa (around 80 km). In the case of the ECMWF model this actually means that the highest model layer spans the pressure range 0–0.02 hPa. Since the temperature of the highest layer of a model is typically strongly affected by the 'sponge' (Shepherd et al., 1996), we limit evaluation of heating rates to pressures greater than 0.02 hPa. For the limited-area NWP application we evaluate heating rates only for pressures greater than 4 hPa, comparable to the model top used in the Met Office UKV model.

2.3 Climate modelling

CKD models used for climate modelling should be able to simulate a wide range of greenhouse gas concentrations. The first four lines of Table 2 list individual scenarios that will be tested. They include present-day and preindustrial conditions, plus the conditions at a glacial maximum, with the values for CO_2 and CH_4 taken from Petit et al. (1999) and for N_2O from the shorter period reported by Schilt et al. (2010). The fourth row shows a 'future' scenario consisting of worst-case conditions for 2110 by extracting the maximum concentrations from any of the CMIP6 scenarios at this time. In this year, the concentration of CH_4 peaks at 3500 ppbv in the SSP3-7.0 scenario, and equivalent CFC-11 peaks at 2000 pptv in the SSP5-8.5 scenario.

The final line of Table 2 shows a range of concentrations that will be used in testing the radiative effect of individual gases, keeping all others constant. For CO_2 we match the range of 0.5–8 times preindustrial concentrations used in RFMIP (Pincus et al., 2016). For the other four well-mixed anthropogenic greenhouse gases we require the capability to simulate the minimum concentrations found in the last million years, which occured at glacial maxima, up to the maximum concentrations found in any of the CMIP6 future scenarios, which extend until 2250.

In principle, there are important applications in addition to those shown in Table 1, such as atmospheric reanalysis, which have been generated back to the mid-19th century (e.g. Compo et al., 2011). A CKD model targeted at this application would only need to span greenhouse gas concentrations from preindustrial to present-day. We decided not to include this application in CKDMIP, partly not to overload the participants, but also because of the expectation that the number of g-points required would not be very different between the reanalysis and climate modelling applications.

3 Generating datasets

Four datasets are provided in CKDMIP, listed in Table 3. Each consists of profiles of layer-wise spectral optical depth due to individual gases. The first two (*Evaluation-1* and *Evaluation-2*) each consist of 50 realistic profiles of temperature, water vapour and ozone (described in section 3.1), accompanied by vertical profiles of the well-mixed gases (described in section 3.2). *Evaluation-1* is provided to participants and may be used to train individual CKD models, while *Evaluation-2* is held back to provide independent evaluation. Section 3.3 describes the last two datasets, which could also be useful in the training of new CKD models. Section 3.4 then describes how the profiles of spectral optical depth were computed for each dataset.

3.1 Temperature, humidity and ozone

For evaluating radiation schemes in RFMIP, Pincus et al. (2016) extracted a set of 100 contrasting atmospheric profiles from the 60-layer ERA-Interim reanalysis dataset, whose highest model level spans the pressure range 0–0.1 hPa. As well as being ten times higher than the pressure of the highest model level in the current ECMWF and Met Office global models, this vertical grid not sufficient to fully resolve the strong peak in atmospheric heating and cooling rates that occurs at the stratopause, nor to test solar absorption by molecular oxygen.

Therefore, we have selected a new set of temperature, pressure, humidity and ozone profiles from the 25,000 'NWP-SAF' profiles of Eresmaa and McNally (2014), which they extracted from ECMWF operational model forecasts in 2013 and 2014. By this time the model used 137 layers with the highest layer spanning pressures 0–0.02 hPa, as in its current configuration. The 50 profiles of the '*Evaluation-1*' dataset consist of 33 randomly taken from the 5,000-profile subset that Eresmaa and McNally (2014) themselves selected to maximize variations in temperature. An additional 17 profiles are selected to contain the extreme values (both maximum and minimum) in the entire dataset of

[1] 'SSP3-7.0' is the 'regional rivalry' Shared Socioeconomic Pathway of CMIP6, with an anthropogenic radiative forcing of 7.0 W m^{-2} in 2100.

Table 2. Surface volume mixing ratios of the five main anthropogenic greenhouse gases for various scenarios considered in CKDMIP, where 'CFC-11 equivalent' is an artificially increased CFC-11 concentration to represent 38 further greenhouse gases. The numbers in brackets indicate the multiple of preindustrial concentrations. The present-day scenario will be used to test CKD models developed for the two NWP applications Table 1. All scenarios will be used to test CKD models for climate, and the final line indicates the range of concentrations over which these gases will be tested individually (keeping the others constant).

Scenario	CO_2 ppmv	CH_4 ppbv	N_2O ppbv	CFC-11 equivalent pptv	CFC-12 pptv
Glacial maximum	180 (0.64)	350 (0.5)	190 (0.7)	32	0
Preindustrial	280 (1)	700 (1)	270 (1)	32	0
Present-day (2020)	415 (1.48)	1921 (2.74)	332 (1.23)	861	495
Future (2110)	1120 (4)	3500 (5)	405 (1.5)	2000	200
Climate range	140–2240 (0.5–8)	350–3500 (0.5–5)	190–540 (0.7-2)	0–2000	0–550

Table 3. The four spectral optical depth datasets generated as part of CKDMIP, where T is temperature, p is pressure and q is specific humidity.

Name	Purpose	Layers	T profiles	Description
Evaluation-1	Training & evaluation	54	50	Realistic profiles selected from NWP-SAF dataset
Evaluation-2	Independent evaluation	54	50	Further profiles selected from NWP-SAF dataset
MMM	Training	52	3	Median, min. and max. of NWP-SAF T, H_2O & O_3 profiles
Idealized	Generating look-up tables	53	11	Idealized profiles regularly spaced in T, $\log p$ and $\log q$

(a) temperature in the layer nearest the surface, (b) temperature at 500 hPa, (c) temperature at 100 hPa, (d) temperature at 10 hPa, (e) temperature at 1 hPa, (f) specific humidity at 500 hPa, (g) specific humidity at 100-hPa (maximum only), (h) ozone concentration at 10-hPa, and (i) ozone concentration at 1 hPa.

It was apparent from inspection of the data that there was virtually no variability in stratospheric water vapour in the ECMWF model at the time the NWP-SAF profiles were generated, which is a problem for training and evaluating a gas-optics model. Therefore, additional variability has been added by multiplying the humidity profiles by the following function of pressure, p:

$$f(p,r) = \exp\left[r \times \frac{1-\mathrm{erf}\left(\frac{p-100\ \mathrm{hPa}}{50\ \mathrm{hPa}}\right)}{2}\right], \qquad (1)$$

where r is a random number drawn from a Normal distribution with mean of zero and standard deviation 0.25, and is constant for each individual profile. This function adds around 25% variability in the stratosphere and mesosphere, but leaves the troposphere virtually unchanged. Unrealistically low humidities have been removed by setting the minimum specific humidity to 10^{-7} kg kg^{-1}.

The resulting temperature, humidity and ozone mixing ratios are shown by the red and blue lines in Fig. 1. An additional '*Evaluation-2*' dataset has been generated with a different set of 33 random profiles from the original 25,000, along with the 17 profiles containing second-most extreme values of the variables listed above. These will be used to provide independent evaluation of the CKD models.

Training and evaluating a CKD model is costly both in terms of computation and storage due to the high spectral resolution required, and 137 layers is more than needed for evaluating clear-sky radiative transfer. Therefore, we interpolate the profiles on to a coarser grid with 54 layers. We use the Line-By-Line Radiative Transfer Model (LBLRTM; Clough et al., 2005), version 12.8, which takes as input the temperature, pressure and gas concentrations at the interfaces between layers. The highest two layers of the coarser grid are bounded by pressures of 0.0001, 0.01 and 0.02 hPa; the first of these represents the TOA since LBLRTM cannot compute gas properties at zero pressure. As shown in Table 1, the pressure surfaces 0.02 and 4 hPa mark the point at which evaluation of heating rates begins. We assign 15 layers between these two pressure surfaces, with the interfaces between them spaced linearly in $p^{0.15}$ space, where p is pressure. The pressures defining the remaining layers vary according to the surface pressure p_s: we assign 35 layers between 4 hPa and $p_s/1.005$, again spaced linearly in $p^{0.15}$ space. Finally, a further two layers are added very close to the surface (bounded by $p_s/1.005$, $p_s/1.002$ and p_s) in order to resolve sharp temperature gradients in the surface layer. The black dots in Fig. 1 mark the corresponding interfaces between layers for the median profiles described in section 3.3.

3.2 Well-mixed gases

Many weather and climate models assume a spatially constant mole fraction for each of the well-mixed gases, whereas for a little more realism they should decrease with height. The radiation scheme in the ECMWF model uses climatologies of these gases that vary with month, latitude and pres-

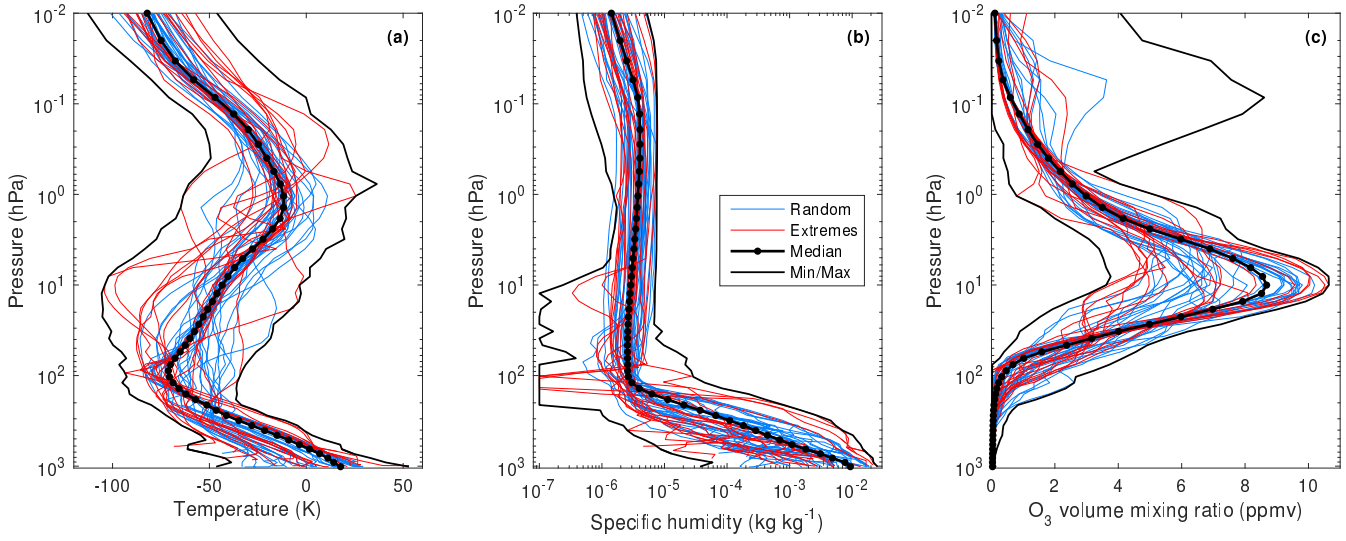

Figure 1. Vertical profiles of the temperature, specific humidity and ozone concentrations for the '*Evaluation-1*' dataset described in section 3.1.

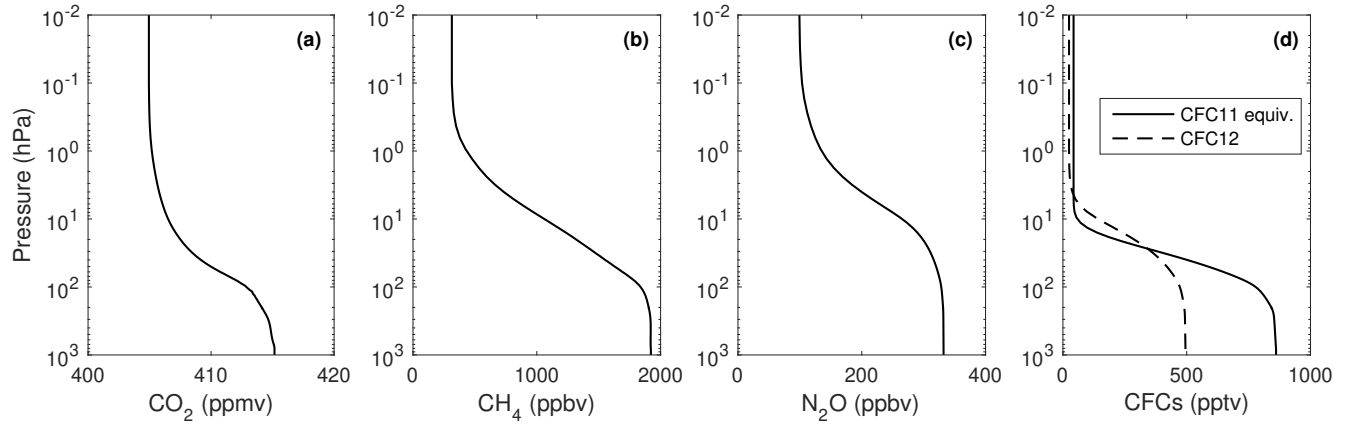

Figure 2. Vertical profiles of the five well-mixed greenhouse gases for the present-day (2020) surface concentrations listed in Table 2.

sure, with the CO_2 and CH_4 climatologies taken from the MACC analysis system (Inness et al., 2013) and the N_2O, CFC-11 and CFC-12 climatologies from the Cariolle chemistry model (Bechtold et al., 2009). Long-term changes due to anthropogenic emissions are represented by scaling these fields so that the global-mean surface values match either historic measurements (for hindcasts and reanalysis) or the CMIP6 SSP3-7.0 scenario (for operational forecasts from model cycle 47R1). We have averaged these climatologies globally and annually, and scaled them to the 2020 surface values in SSP3-7.0, to obtain the profiles shown in Fig. 2. Present-day CO_2 has a difference of 10 ppmv between the values at 1000 and 0.01 hPa. In the case of CFC-11 and CFC-12, the concentrations from the Cariolle model drop to almost zero in the upper stratosphere and mesosphere, which could be problematic for using them to train the pressure dependence in a CKD model. Therefore the profiles of these two gases have been artificially modified to fall to no less than 5% of their surface value. In order to obtain profiles with the surface concentrations shown in Table 2, we simply scale the profiles shown in Fig. 2.

We have computed that the difference in the total TOA radiative forcing of a gas with a constant mole fraction with pressure, versus the more realistic profiles in Fig. 2 but the same surface concentration, is 10% for CFC-11, 5% for CFC-12, and less than 0.2% for the other three gases.

3.3 Additional training datasets

Two additional datasets are shown at the bottom of Table 3, which are intended to facilitate the development of CKD schemes, while being consistent with the datasets that will be used to evaluate them. The '*MMM*' dataset contains the optical properties of all nine gases but using the median, min-

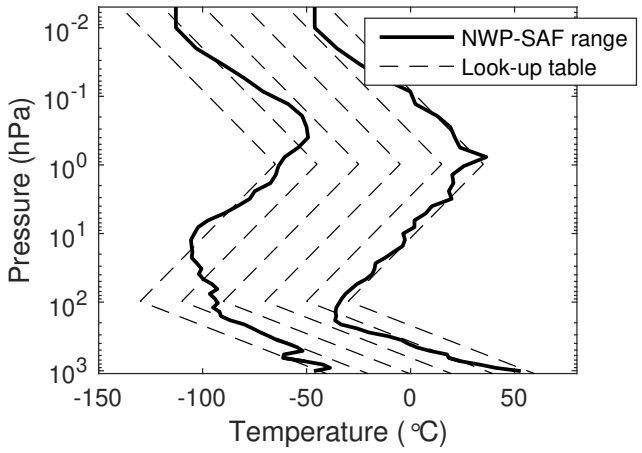

Figure 3. The solid lines show the minimum and maximum temperatures of the NWP-SAF dataset, also shown in Fig. 1a. The dashed lines show the 6 idealized temperature profiles, 20 K apart, used in the '*Idealized*' dataset in Table 3.

imum and maximum temperature profiles derived from the entire 25,000-profile NWP-SAF dataset; these temperatures are shown by the black lines in Fig. 1. In the case of H_2O and O_3 only, three concentration profiles are used for each temerature, corresponding also to the median, minimum and maximum of the NWP-SAF profiles (shown in Figs. 1b and 1c). For all other gases the present-day concentrations shown in Fig. 2 are used. The vertical grid is the same as for the *Evaluation-1* and *Evaluation-2* datasets, except that surface pressure is set to mean sea level pressure ($p_s = 1013.25$ hPa), and the two layers very close to the surface are not used so that the total number of layers is 52 rather than 54.

The final '*Idealized*' dataset contains absorption spectra for idealized temperature and concentration profiles that are intended to cover the full range of likely temperature, pressure and concentrations found in the atmospheres that any CKD model would be applied to. Therefore they can be used to populate look-up tables of molar absorption to be used by CKD models. We envisage that the maximum layer-mean pressure that needs to be accommodated by a radiation scheme is 1100 hPa, so construct a logarithmically spaced pressure profile of 53 elements, containing ten points per decade with a maximum layer-mean pressure of 1100 hPa. At each pressure, 6 temperatures are simulated spanning a 100 K range at 20 K intervals. We use idealized temperature profiles shown in Fig. 3 that are intended to encompass the maximum and minimum temperatures found in the NWP-SAF dataset. For all gases, absorption spectra are computed for mole-fraction profiles that are constant with pressure, using the present-day values for the five well-mixed gases shown in Table 2, and 5 ppmv for O_3. Since the molar absorption of these gases is very close to constant with concentration (see section 3.4), only one concentration needs to be simulated for each. In the case of water vapour whose absorption varies with concentration, we simulate 12 logarithmically spaced specific humidities from 10^{-7} to $10^{-1.5}$ kg kg^{-1}, i.e. using two values per decade.

3.4 Line-by-line modelling

TBD: Details of LBLRTM settings

An important practical consideration is to determine at what spectral resolution to produce the absorption spectra. They need to be fine enough resolution that the most narrow spectral lines are resolved and the resulting irradiance and heating-rates profiles are an accurate benchmark, but also a managable data volume for storage, processing and distribution. LBLRTM can inform the user of the spectral resolution it needs to resolve the lines at a particular pressure, and for CO_2 in the longwave at 0.01 hPa (the most important gas at the pressure where the lines are finest), it recommends a wavenumber resolution of 0.000036, 0.000156, 0.000292 and 0.000485 cm^{-1} in the wavenumber ranges 0–350, 350–1300, 1300–1700 and 1700–3260 cm^{-1}, respectively. This implies that more than 20 million spectral points are required. Using this resolution as a reference, we have experimented with degrading the spectral resolution in each of these four wavenumber ranges. Computing the heating rate error for each wavenumber range, we find that the 350–1300 cm^{-1} range is the only one of the four where anything like the LBLRTM-recommended resolution is required. Therefore, we have adopted spectral resolutions of 0.0002, 0.001 and 0.005 cm^{-1} in three spectral ranges 0–1300, 1300–1700 and 1700–3260 cm^{-1}. This leads to heating-rate errors of no more than around 0.005 K d^{-1} (all of which occur in the upper stratosphere and mesosphere) in any of the four original wavenumber ranges, even for the most challenging scenario of 8 times preindustrial concentrations of CO_2. This leads to 7,211,999 spectral points in the longwave.

A further significant reduction in data volume is possible if the absorption cross-section per molecule is independent of the concentration of that gas, so varies only as a function of temperature and pressure. In this case, for well-mixed gases, the profile of layer-wise optical depth need only be provided for a single concentration profile; if optical depths are required for concentration profiles scaled by a constant, then the optical depths themselves may simply be scaled. We have computed absorption spectra for each gas over the full range of concentrations required in Table 2, and found that to a very good approximation, molar absorption can be treated as independent of concentration for all gases except water vapour. Therefore, for the well-mixed gases, absorption spectra are provided only for present-day concentrations. The CKDMIP software accordingly allows the user to scale the optical depth of each gas before performing radiative transfer calculations on the mixture.

A similar approach is taken in the shortwave. TBD: Description of shortwave spectra. The CKD software calculates

the spectral optical depth due to Rayleigh scattering using the model of Nicolet (1984).

A realistic TOA solar irradiance spectrum was extracted from the climate data record of Coddington et al. (2016) by averaging over the last 33 years (1986–2018 inclusive), i.e. three solar cycles. It has a resolution of 1 nm at wavelengths shorter than 750 nm, and is interpolated to the spectral resolution of the shortwave gas absorption spectra.

The water vapour spectra have been computed assuming the widely used Mlawer-Tobin-Clough-Kneizys-Davies (MT_CKD) continuum model (Mlawer et al., 2012), version 2.5. However, there is still considerable uncertainty on the strength of the water vapour continuum, particularly in the near infrared (Shine et al., 2016), and indeed this could be a source of difference between individual gas optics schemes and the reference calculations produced in CKDMIP. Therefore, for each dataset, we produce an additional set of water vapour files but with no representation of the continuum. If needed, evaluation can be carried out using only the contribution from spectral lines, or alternatively different models of the continuum can be tried.

The absorption spectra are stored, one gas per file, in NetCDF4/HDF5 format with compression, so the file size depends on the spectral extent and degree of fine structure in the spectrum. In the longwave, the volume of a single file varies from 0.05 GB for CFC-11 to around 1 GB for CH_4.

TBD: Description of shortwave and total data volumes for the various datasets

3.5 Generating irradiance and heating-rate benchmarks

The CKDMIP software takes as input the spectral optical depths of each of a number of gases, optionally scales the optical depths of the well-mixed gases if a different concentration is required, and computes clear-sky aerosol-free irradiances (broadband or spectral) at layer interfaces for each of the test profiles. These can be used to compute broadband or spectral heating-rate profiles. The intention is that the radiative transfer equations are then the same as those used by large-scale atmospheric models, and the same solver will be used with the various CKD models in order that any differences to the line-by-line broadband irradiances are due to the representation of gas optics, not the details of the solver.

In the longwave we use a no-scattering solver with the following properties:

- Surface emissivity is assumed to be unity.
- The skin temperature of the surface is assumed to be equal to the air temperature at the base of the lowest atmospheric layer.
- Local thermodynamic equilibrium is assumed.
- The propagation of longwave radiation assumes a diffusivity factor of 1.66, i.e. radiation is assumed to propagate with discrete zenith angles of $\pm 53°$.
- The temperature at layer interfaces is taken as input and a linear-in-optical-depth variation of the Planck function within each layer is assumed, leading to the use of Eqs. 6–12 of Hogan and Bozzo (2018).

The shortwave scheme has the following characteristics:

- The surface is assumed to be a Lambertian reflector with an albedo of 0.06.
- It uses a direct-beam calculation plus a two-stream diffuse calculation, with the Zdunkowski et al. (1980) coefficients characterizing the rate of exchange of energy between the three streams, and the Meador and Weaver (1980) solutions to the two-stream equations in individual layers.
- Calculations are performed at a number of solar zenith angles θ_0, evenly spaced in $\cos\theta_0$ space, thereby matching the distribution of solar zenith angles striking the Earth. We do not account for the fact that individual test profiles at a particular latitude would each experience a different θ_0 distribution.
- No account is made for Earth curvature.

The atmospheric heating rate in layer i is computed from the net irradiance divergence across a layer, as:

$$\frac{dT_i}{dt} = -\frac{g_0}{C_p}\frac{F^n_{i+1/2} - F^n_{i-1/2}}{p_{i+1/2} - p_{i-1/2}}, \qquad (2)$$

where $p_{i+1/2}$ and $F^n_{i+1/2}$ are the pressure and net downward irradiance, respectively, at the interface between layers i and $i+1$, g_0 is the acceleration due to gravity (standard gravity) and C_p is the specific heat of dry air, taken to be constant at 1004 J kg^{-1} K^{-1}.

4 CKDMIP experimental protocol

Anyone with a CKD tool can take part in CKDMIP. Participants are provided with access to the *Evaluation-1*, *MMM* and *Idealized* datasets, and the software described in section 3.5 to perform line-by-line radiation calculations on them. They may use these or their own datasets as input to their CKD tool. In section 4.1 we describe the band structure that should be used by participants, if possible. Section 4.2 describes the calculations that will be performed by participants and the data they will provide. In section 4.3 we outline the how these data will be processed to quantify accuracy, and to investigate the accuracy–efficiency trade-off.

Table 4. The spectral boundaries of the (left) 'narrow' and (right) 'wide' longwave bands, in which participants will be asked to generate CKD models. The narrow bands are essentially the same as those in RRTMG, except for the final band which spans the last four bands of RRTMG.

Narrow bands			*Wide bands*	
#	Spectral interval (cm^{-1})	RRTMG g-points	#	Label
1	0–350	8	1	Far infrared
2	350–500	14		
3	500–630	16	2	Main CO_2 band
4	630–700	14		
5	700–820	16		
6	820–980	8	3	Infrared window
7	980–1080	12		
8	1080–1180	8		
9	1180–1390	12	4	Mid-infrared A
10	1390–1480	6		
11	1480–1800	8		
12	1800–2080	8	5	Mid-infrared B
13	2080–3260	10		

Table 5. As Table 4 but for the shortwave. The narrow bands are as in RRTMG, except for band 7 which spans two RRTMG bands.

Narrow bands			*Wide bands*	
#	Spectral interval (cm^{-1})	RRTMG g-points	#	Label
1	250–2600	12	1	Mid-infrared
2	2600–3250	6		
3	3250–4000	12		
4	4000–4650	8	2	Shortwave infrared
5	4650–5150	8		
6	5150–6150	10		
7	6150–8050	12		
8	8050–12850	10	3	Near infrared
9	12850–16000	8		
10	16000–22650	6	4	Visible window
11	22650–29000	6		
12	29000–38000	8	5	Ultraviolet
13	38000–50000	6		

4.1 Common band structures

Virtually all operational CKD models for weather and climate split the longwave and shortwave spectra into bands, and compute k distributions within each one. As shown in the survey of Hogan et al. (2017), the number of bands is strongly correlated to the total number of g-points, and therefore to the overall efficiency of a CKD model. The choice of bands can be dependent on the constraints of a particular CKD scheme: some schemes require the longwave bands to be narrow enough that the Planck function may be assumed constant (e.g. Fu and Liou, 1992); some need to restrict the number of active gases in a band (e.g. Mlawer et al., 1997); some assume the spectral overlap of different gases is random, invalid for wide bands (e.g. Ritter and Geleyn, 1992); while probably all assume that cloud and surface properties are constant within each band, which could lead to significant errors in the shortwave if the bands are too wide. All of these arguments deserve detailed scrutiny within CKDMIP.

We propose two band structures, shown in Table 4 for the longwave and Table 5 for the shortwave. Since RRTMG (Mlawer et al., 1997) is so widely used, our proposed 'narrow bands' are modelled on RRTMG, except that we merge a few of the very narrow or very low-energy bands that RRTMG represented with four or fewer g-points. This leads to 13 bands in both the longwave and shortwave. These bands should be narrow enough to satisfy all the needs for narrowness cited previously. To assist participants who do not wish to download all the large spectral absorption files, much smaller files are available containing benchmark irradiance profiles computed for each scenario of the *Evaluation-1* dataset, both broadband values and values averaged in each of the narrow bands.

The 'wide bands', of which there are five in both the longwave and the shortwave, consist of groupings of the narrow bands. In the longwave these are purposefully somewhat wider than in most current CKD models, in order to really test the limits of some of the restrictions cited above. Moreover, they will potentially allow the advantages of CKD schems that do not assume the Plank function to be constant across a band, or do not assume random spectral overlap, to be more clearly apparent.

Some participants may wish to use their own sub-bands within these wide bands if they think it will achieve a better accuracy–efficiency trade-off for a particular wide band. For example, Cusack et al. (1999) used two 'split bands' in the longwave, one which represented the wings of the main CO_2 band (essentially a merger of our narrow bands 3 and 5) and the other which represented the parts of the infrared window on either side of the ozone band (essentially a merger of our narrow bands 6 and 8).

Finally, CKDMIP welcomes submissions using even wider bands. Indeed, the 'full-spectrum correlated-k' (FSCK) technique has been proposed as a means to achieve good accuracy using only one band in the longwave (Hogan, 2010) and two in the shortwave (Pawlak et al., 2004). The investigation of the effect of spectral variations of cloud properties within bands and g-points described in section 4.4 will be particularly important for FSCK submissions.

4.2 Contribution of CKDMIP participants

Ideally, CKDMIP participants would use their tool to generate a CKD model for all combinations of the following:

- The longwave and shortwave.
- The three applications listed in Table 1.
- The narrow and wide band structures described in section 4.1 (and optionally even wider bands).
- A range of total number of g-points, in order that the efficiency–accuracy trade-off can be explored. Ideally at least three configurations would be provided.

This could potentially lead to a full submission involving the generation of 36 CKD models. It is recognised that this is potentially very demanding, so reduced submissions are welcome according to the scientific interests of the participant. In principle, a participant could submit just one longwave and one shortwave CKD model; if it used the narrow bands specified in section 4.1 and targeted climate modelling, then it could still be tested against other models in all scenarios.

Participants do not submit the code for their CKD models, but rather run each of them on the 50-profile *Evaluation-1* dataset with the following well-mixed gas concentration scenarios:

- For CKD models generated for the two NWP applications, the present-day scenario in Table 2.
- For CKD models generated for climate modelling, the four individual scenarios 'glacial maximum', 'preindustrial', 'present-day' and 'future'.
- For CKD models generated for climate modelling, the present-day conditions but with individual gases perturbed in the ranges shown on the last line of Table 2; CO_2 should use the values 140, 280, 1120 and 2240 ppmv. CH_4 should use the values 350, 700, 1200, 2600 and 3500 ppbv. N_2O should use the values 190, 270, 405 and 540 ppbv. CFC-11 and CFC-12 may simply use the extreme ends of the proposed range. The present-day scenario will already be available providing an additional point in each of these ranges.

For each of these scenarios, they submit a NetCDF file containing the following variables as a function of profile number in the *Evaluation-1* dataset:

- Pressure at layer interfaces, copied from the input file;
- The absorption optical depth of all gases in each layer, in each of N g-points;
- In the shortwave only, the Rayleigh scattering optical depth in each layer and g-point;
- In the shortwave only, the TOA solar irradiance integrated over the parts of the spectrum contributing to each g-point, scaled such that these numbers sum to a total solar irradiance of 1360.5 W m^{-1}.
- In the longwave only, the Planck function at each layer interface, integrated over the parts of the spectrum contributing to each g-point. At a given layer interface, these values should sum to close to σT^4, where σ is the Stefan-Boltzmann constant and T is the temperature at the layer interface (provided in the input file).

These files should be compatible with the CKDMIP software, which can then read them in and compute profiles of upwelling and downwelling irradiances, both at each g-point and as broadband values. This ensures that the radiative transfer is identical to that used in generating the line-by-line benchmarks, so that when the irradiances are compared to the benchmarks, the differences are only due to the spectral approximations made in the CKD model.

A further file is required for each CKD model generated, describing which parts of the spectrum are represented by each g-point, to be used in section 4.4 for investigating the representation of cloud optical properties. In the longwave this should be expressed at a resolution of 10 cm^{-1} and in the shortwave at a resolution of 50 cm^{-1}. This is commensurate with the spectral scale at which the optical properties of clouds vary.

After the first phase of comparisons using the *Evaluation-1* dataset in which the line-by-line benchmarks are made available to participants, a second phase of comparisons will be conducted using *Evaluation-2* dataset, in which the line-by-line benchmarks are withheld.

The protocol above assumes that participating radiation schemes have a clean separation between the generation of optical depths in each g-point and the radiative transfer perfomed on them. Allowance will need to be made for some schemes in which the separation is not so clean. For example, SOCRATES (the Suite Of Community Radiation codes based on Edwards and Slingo, 1996) uses the concept of 'equivalent extinction' to treat minor gases (Edwards, 1996). In the longwave this involves performing M no-scattering radiation calculations to work out the contribution from minor gases in a band. The net irradiance from these profiles are analyzed to work out the equivalent extinction, which is then added to the N g-points for representing the major gases in the band. A full longwave radiative transfer calculation, including scattering, is then performed on these N g-points. This approach could be accommodated in CKDMIP by the participant performing the M initial calculations themselves and providing the resulting N optical depth profiles. The CKDMIP radiative transfer software would then be run on these N g-points (verifying that it gives very similar results to the SOCRATES radiative transfer solver), but when assessing the accuracy–efficiency trade-off, the cost of the scheme would be counted as $aM + N$, the a factor being optionally

less than one to account for the fact that equivalent extinction can be computed with a cheaper solver. In the shortwave, the SOCRATES scheme uses a more sophisticated treatment of gas optics ($M+N$ g-points) for the cheap direct-beam radiative transfer calculation, and a simpler treatment of gas optics (N g-points) for the more expensive solver for scattered radiation. This could be accommodated by the participant providing CKDMIP with separate direct and a diffuse optical depths in the N g-points, and again the cost of the scheme being counted as $aM+N$, with a this time representing the cost of the direct-only versus full shortwave radiation calculation.

4.3 Error metrics

The irradiance profiles provided by participants for the relevant scenarios in Table 2 will be compared to the equivalent line-by-line benchmarks, with differences in net irradiance being characterized by the bias and root-mean-squared error (RMSE) over the 50 profiles. Atmospheric heating-rate bias and RMSE will be examined as a function of pressure. The profile of heating-rate error will be summarized by a few error metrics, such as the whole-profile RMSE, or the values for the troposphere, stratosphere and mesosphere separately. An appropriate weighting with pressure will need to be specified; rather than weighting linearly with pressure which overweights the troposphere, Hogan (2010) proposed weighting by the square-root of pressure, which increases the weighting of stratospheric errors, but other powers (e.g. the cube-root) are possible. Naturally, the heating-rate errors will only be counted down to the lowest pressure for the application in question (see Table 1). The handful of RMSE values will then be plotted as a function of number of g-points to compare how different CKD tools perform in terms of accuracy versus efficiency.

In addition, we will look at the accuracy of the CKD models for climate in terms of the TOA and surface radiative forcing they predict when the five well-mixed anthropogenic greenhouse gases in Table 2 are perturbed individually in the range given in the bottom row. This will involve simple averaging over the 50 profiles.

4.4 Errors due to the spectral variation of cloud properties

Until this point, we have considered exclusively clear-sky radiation calculations with a spectrally constant surface albedo. It is known that errors can arise in cloudy skies if cloud optical properties are assumed constant across spectral bands (Lu et al., 2011), primarily due to the spectral correlation of absorption by water vapour, liquid water and ice. In principle, this error can be ameliorated by computing cloud properties separately for each g-point, possible if we have fine-scale information on which parts of the spectrum each g-point contributes to. As described in section 4.2, this information is requested of participants for each of their CKD models.

In the final part of CKDMIP, errors in cloudy skies are estimated. This may be achieved using the clear-sky submissions of the CKDMIP participants, so requiring no additional simulations from them. Firstly, line-by-line cloudy-sky benchmarks are produced. For liquid clouds, Mie calculations have been performed for distributions of droplets at a sufficiently high spectral resolution to resolve variations of refractive index. For ice clouds we use the generalized habit mixture of Baum et al. (2014). The CKDMIP software is then used to add horizontally homogeneous clouds of varying optical depth to the gas optical depth in the *Evaluation-1* dataset for present-day conditions, and perform line-by-line calculations. Then the equivalent calculations are performed for the various CKD models, by taking their present-day optical depth files and adding the contribution from clouds. From the information they provide on the spectral contributions to each g-point, average cloud properties will be computed for each g-point using the appropriate combination of 'thick' and 'thin' averaging (Edwards and Slingo, 1996). Errors in irradiances and heating rates will then be computed.

A similar procedure would be possible for aerosols, or to quantify errors due to spectrally surface albedo, particularly over snow an vegetation where the variations are largest.

TBD: Section with some examle data and perhaps an estimate of the errors in an existing model?

5 Conclusions

6 Code and data availability

The code and technical documentation are available at the project web site http://confluence.ecmwf.int/display/CKDMIP. The username and password needed to access the FTP site containing the CKDMIP datasets are available on request from the first author.

Acknowledgments

This paper has benefited from valuable discussions with James Manners, Keith Shine, Eli Mlawer, Robert Pincus, David Paynter...

References

Baum, B. A., P. Yang, A. J. Heymsfield, A. Bansemer, A. Merrelli, C. Schmitt, and C. Wang, 2014: Ice cloud bulk single-scattering property models with the full phase matrix at wavelengths from 0.2 to 100 um. *J. Quant. Spectrosc. Radiat. Transfer,* **146,** 123–139

Bechtold, P., A. Orr, J.-J. Morcrette, R. Engelen, J. Flemming and M. Janiskova, 2009: Improvements in the stratosphere and mesosphere of the IFS. *ECMWF Newsletter No. 120,* 22–31.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara and P. D. Brown, 2005: Atmospheric radiative transfer modeling: a summary of the AER codes. *J. Quant. Spectrosc. Radiat. Transfer,* **91,** 233–244.

Coddington, O., J. L. Lean, P. Pilewskie, M. Snow and D. Lindholm, 2016: A solar irradiance climate data record. *Bull. Am. Meteorol. Soc., 97,* 1265–1282.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P. and Brönnimann, S., 2011: The twentieth century reanalysis project. *Quart. J. Roy. Meteorol. Soc.,* **137,** 1–28.

Cusack, S., J. M. Edwards and J. M. Crowther, 1999: Investigating k distribution methods for parameterizing gaseous absorption in the Hadley Centre Climate Model. *J. Geophys. Res.,* **104,** 2051–2057.

Eresmaa, R., and A. P. McNally, 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/.

Edwards, J. M., 1996: Efficient calculation of infrared fluxes and cooling rates using the two-stream equations. *J. Atmos. Sci.,* **53,** 1921–1932.

Edwards, J. M., and A. Slingo, 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.

Foster, G. L., D. L. Royer and D. J. Lunt, 2017: Future climate forcing potentially without precedent in the last 420 million years. *Nat. Commun.* **8,** 14845.

Fu, Q., and K. N. Liou, 1992: On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres. *J. Atmos. Sci.,* **49,** 2139–2156.

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.

Hogan, R. J., and Bozzo, A., 2018: A flexible and efficient radiation scheme for the ECMWF model. *J. Adv. Modeling Earth Sys.,* **10,** 1990–2008.

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

Höpfner, M., Milz, M., Buehler, S., Orphal, J., and Stiller, G., 2012: The natural greenhouse effect of atmospheric oxygen (O_2) and nitrogen (N_2), *Geophys. Res. Lett.,* **39,** L10706, doi:10.1029/2012GL051409.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team, 2013: The MACC reanalysis: an 8 yr data set of atmospheric composition. *Atmos. Chem. Phys.,* **13,** 4073–4109.

Jones, F. E., and R. M. Schoonover, 2002: Handbook of mass measurement. CRC Press, 336 pp.

Lacis, A., and V. Oinas, 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.

Lu, P., H. Zhang and J. Li, 2011: Correlated k-distribution treatment of cloud optical properties and related radiative impact. *J. Atmos. Sci.,* **68,** 2671–2688.

Meador, W. E., and W. R. Weaver, 1980: Two-stream approximations to radiative transefer in planetary atmospheres: a unified description of existing methods and a new improvement. *J. Atmos. Sci.,* **37,** 630–643.

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., and Weiss, R., 2017: Historical greenhouse gas concentrations for climate modelling (CMIP6) *Geosci. Model Dev.,* **10,** 2057–2116.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res. Atmos.,* **102,** 16 663–16 682.

Mlawer, E. J., V. H. Payne, J. L. Moncet, J. S. Delamere, M. J. Alvarado and D. C. Tobin, 2012: Development and recent evaluation of the MT_CKD model of continuum absorption. *Philos. Trans. R. Soc. A,* **370,** 2520–2556.

Nicolet, M., 1984: On the molecular scattering in the terrestrial atmosphere: An empirical formula for its calculation in the homosphere. *Planetary and Space Sci.,* **32,** 1467–1468.

Niemelä, S., P. Räisänen and H. Savijärvi, 2001. Comparison of surface radiative flux parameterizations: Part I: Longwave radiation. *Atmos. Res.,* **58,** 1–18.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M., 2016: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geosci. Model Dev.,* **9,** 3461–3482.

Pawlak, D. T., E. E. Clothiaux, M. F. Modest and J. N. S. Cole, 2004: Full-spectrum correlated-k distribution for shortwave atmospheric radiative transfer. *J. Atmos. Sci.,* **61,** 2588-2601.

Petit, J. R., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J. M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., and Delmotte, M., 1999: Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature,* **399,** 429–439.

Pincus, R., Mlawer, E. J., Oreopoulos, L., Ackerman, A. S., Baek, S., Brath, M., Buehler, S. A., Cady-Pereira, K. E., Cole, J. N., Dufresne, J. L. and Kelley, M., 2015: Radiative flux and forcing parameterization error in aerosol-free clear skies. *Geophys. Res Lett.,* **42,** 5485–5492.

Pincus, R., Forster, P. M., and Stevens, B., 2016: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6. *Geosci. Model Dev.,* **9,** 3447–3460.

Ritter, B. and 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.

Schilt, A., Baumgartner, M., Schwander, J., Buiron, D., Capron, E., Chappellaz, J., Loulergue, L., Schüpbach, S., Spahni, R., Fischer, H., and Stocker, T. F., 2010: Atmospheric nitrous oxide during the last 140,000 years. *Earth Planetary Sci. Lett.,* **300,** 33–43.

Shepherd, T. G., Semeniuk, K., and Koshyk, J. N., 1996: Sponge layer feedbacks in middle-atmosphere models. *J. Geophys. Res.,* **101,** 23447–23464.

Shine, K. P., A. Campargue, D. Mondelain, R. A. McPheat, I. V. Ptashnik and D. Weidmann, 2016: The water vapour continuum in near-infrared windows — current understanding and prospects for its inclusion in spectroscopic databases. *J. Molecular Spectroscopy,* **327,** 193–208.

Zdunkowski, W. G., R. M. Welch and G. Korb, 1980: An investigation of the structure of typical two-stream methods for the calculation of solar fluxes and heating rates in clouds. *Beitr. Phys. Atmos.,* **53,** 147–166.