816

TECHNICAL MEMORANDUM

Radiation in numerical weather prediction

Robin J. Hogan, Maike Ahlgrimm, Gianpaolo Balsamo, Anton Beljaars, Paul Berrisford, Alessio Bozzo, Francesca Di Giuseppe, Richard M. Forbes, Thomas Haiden, Simon Lang, Michael Mayer, Inna Polichtchouk, Irina Sandu, Frederic Vitart and Nils Wedi

Research, Forecast and Copernicus Departments

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Contents

Abstract

Radiation is a fundamental process that drives atmospheric flows at all scales, and is key to both improving short-range surface temperature forecasts, and meeting ECMWF's strategic aim of pushing the boundaries of predictability at the medium-range and longer timescales. This paper provides a detailed assessment of recent radiation developments at ECMWF and the priority areas where progress needs to be made in the coming years. A particular focus is placed on the new 'ecRad' radiation scheme that became operational in 2017 and will facilitate future developments. Five 'Grand Challenges' for radiation in NWP are then presented: the surface (particularly coastlines, forests and urban areas), clouds (particularly cloud structure, longwave scattering and 3D radiative effects), clear-sky absorption (particularly the impact of aerosols on monsoon systems, and the water vapour continuum), the middle atmosphere (particularly how reducing the large stratospheric temperature biases has the potential to improve predictive skill on monthly timescales) and efficiency (particularly optimizations in ecRad, and how future effort should reassess the balance between the spectral, temporal and spatial resolution of the radiation calculations).

1   Introduction

Radiation is a fundamental process in the earth system, providing the energy that drives both the large-scale circulation and smaller-scale processes affecting the weather experienced at the ground. While it has always been the case that a good treatment of radiative transfer is a prerequisite for reasonable near-surface temperature forecasts, ECMWF's strategic aim of pushing the boundaries of predictability at the medium-range, monthly and seasonal timescales puts additional stringent demands on the accuracy of the radiation scheme. We require not only that the Integrated Forecasting System (IFS) has an excellent climate, but also that the role of radiation in the evolution and predictability of atmospheric weather systems and regimes is well captured.

It is easy to identify specific troublesome locations and conditions where short-range near-surface temperature forecasts are poor and where improvements in the representation of radiation, in concert with the surface, boundary-layer and cloud schemes, can address them. Recent radiation changes have improved forecasts at coastlines, and an ongoing area of attention is how to prescribe the albedo of forests under snow cover. A focus in the coming years will be urban areas, which are currently ignored in the IFS (being represented as forest or grassland), yet are a WMO priority due to the impact on human health of both the urban heat-island effect and urban pollution.

By contrast, identifying what radiation changes are needed to improve the predictability of weather systems on longer timescales is much more challenging. We generally look to the radiation scheme to produce a good model climate as a pre-requisite for improving predictive skill, but the difficulty is that according to *large-scale* metrics, the climate of recent versions of the IFS is already excellent. Seasonal- and zonal-mean temperature biases are typically less than 0.5 K through most of the troposphere, which is close to our ability to measure the bias, and is of the same order as the changes that occur when switching from using prescribed sea-surface temperatures to a fully coupled ocean model. Likewise, global-mean fluxes at the surface and top-of-atmosphere are accurate to within 1 W m^{-2}. Establishing such a good climate is undoubtedly helped by some compensation of errors between the radiation, cloud and convection schemes, which means that introducing physically-based improvements to these schemes can upset the balance and increase forecast errors.

But when we look *regionally,* or above the tropopause, far larger biases emerge that are at least partly associated with radiation. In the troposphere there are persistent regional biases in the reflectance of different marine cloud systems that implicates both the clouds in the model and the way they are handled

in radiation. Errors in the distribution and optical properties of aerosols are increasingly being recognized to degrade monsoon systems in the tropics. And large, persistent temperature biases in the stratosphere, some of which are associated with errors in the treatment of ozone, are harming the potential predictive skill we would like to see through stratosphere-troposphere interactions.

This paper provides a detailed assessment of both the radiation developments at ECMWF in recent years, and the areas that need to be improved in future. It is a slightly updated version of a Special Topic Paper presented at the 46th ECMWF Science Advisory Committee (9–11 October 2017). We start in section 2 with a review of the main developments in the treatment of radiation since 2000. A particular focus is placed on the features of the new 'ecRad' radiation scheme that became operational in 2017. This scheme has been designed to be modular, flexible and extensible, and in addition to its underpinning a number of the topics explored in this paper, it is anticipated to provide the framework for the scientific developments in radiation in the coming years.

Section 3 then provides a brief assessment of the climate of the IFS, and its dependence on resolution and coupling to the ocean. We focus particularly on the surface radiation budget and tropospheric temperature biases. Sections 4–8 then consider five 'Grand Challenges' for radiation in NWP:

1. **Surface.** Two-metre temperature is an essential forecast product, yet forecast errors (both random and systematic) in some locations are unacceptably high, with radiation being at least partially to blame. We show how large temperature errors at coastlines due to the coarse radiation grid may be mitigated via the use of approximate updates to the broadband flux profiles. We then explore the potential for improving the interaction of radiation with complex underlying surfaces, such as forests and cities.

2. **Clouds.** The complex sub-grid structure of clouds presents multiple challenges to radiative transfer, and improving their interaction with radiation is confounded by the systematic errors in the location and properties of clouds predicted by the model. We show the potential for radiative biases to be revealed by careful comparison with observations, and describe improvements to cloud–radiation interactions such as a better parameterization of horizontal sub-grid heterogeneity and the ability to represent three-dimensional radiative effects in a global model for the first time.

3. **Clear-sky absorption.** While errors in clear-sky tropospheric shortwave absorption by gases and aerosols are locally smaller than those due to clouds, they can systematically affect temperatures over large areas and thereby modify regional circulation patterns. We present two examples. Firstly, a recent significant upgrade of the aerosol climatology led to a significant improvement in the Indian Summer Monsoon associated with a reduction in absorbing aerosol over Arabia. Secondly, implementing recent findings that the water vapour continuum in the near infrared could be significantly stronger than previously thought acts to reduce the cold bias of the tropical troposphere.

4. **Middle atmosphere.** The stratosphere provides a potential source of predictive skill on monthly timescales, but the downward influence of stratospheric phenomena such as sudden warmings on the troposphere is known to be too weak in the IFS. This is likely to be related to the large middle-atmosphere temperature biases in the free-running IFS. We show that monthly forecasts over Europe can be improved if the ~ 5 K polar lower stratosphere cold bias is fixed, and also demonstrate how a sequence of radiation improvements can almost eliminate the warm bias (currently up to 20 K) in the upper stratosphere and mesosphere.

5. **Efficiency.** The historically high cost of the radiation scheme has led to the need for it to be called infrequently in time and space. Despite the use of approximate updates every timestep and

gridpoint, the 3-h radiation timestep in the majority of operational model configurations is known to degrade forecast skill. The rewrite of the radiation scheme has yielded a 31–34% speed-up, and we also consider more significant optimizations in future, such as a radical overhaul of the gas optics scheme with the aim of significantly reducing the number of spectral intervals without compromising accuracy.

Finally, in section 9, we provide an outlook for radiation developments in each of these areas over the coming decade, and the likely timescale on which they will be tackled.

2   A brief history of the ECMWF radiation scheme

2.1   McRad

The radiation scheme used at ECMWF in the 1990s originated from the University of Lille; this scheme was described by Morcrette (1991) and the developments made in that decade were outlined by Morcrette et al. (2008a). Table 1 shows the changes made since 2000, a year that saw a major upgrade to the longwave scheme with the introduction of the Rapid Radiative Transfer Model for GCMs (RRTM-G; Mlawer et al., 1997). RRTM-G is a correlated-k model of gas absorption that was sponsored by the US Atmospheric Radiation Measurement programme and built on developments in both line-by-line radiative transfer and a network of high quality validation sites. The longwave component reproduced clear-sky fluxes to within a few Watts per square metre.

This was followed in 2007 by a major upgrade of the radiation package to 'McRad', which involved not only incorporation of the shortwave RRTM-G model, but also the McICA scheme of Pincus et al. (2003) to represent realistic cloud heterogeneity and overlap via a stochastic cloud generator. A detailed description was provided by ECMWF (2007), Morcrette et al. (2008a) and Ahlgrimm et al. (2016). These papers describe in detail the improvements in forecast scores and model climate forthcoming from the introduction of McRad.

The larger total number of spectral intervals used by RRTM-G made it 3.5 times slower than the Morcrette scheme, necessitating the use of a lower resolution radiation grid (Morcrette et al., 2008b). In 2016, two schemes were introduced to remedy problems associated with calling the radiation scheme infrequently in time and space. The Hogan and Bozzo (2015) scheme performs approximate updates to the fluxes at every timestep and model gridpoint to correct for errors due to sharp temperature and albedo transitions at coastlines (see section 4.1), and the improved treatment of solar zenith angle described by Hogan and Hirahara (2016) helped to reduce the stratospheric warm bias due to infrequent calls of the radiation scheme (see section 7.1). Later in 2016, the total solar irradiance was reduced by 0.4% to better match satellite observations, and a prescribed solar cycle was added (Hersbach et al., 2015), both with a very small impact.

2.2   ecRad

A difficulty with McRad from a developmental perspective is that the various components of the scheme, particularly the gas optics component (currently RRTM-G) and the solver (currently McICA), are completely intertwined, making it impossible to use it to test alternatives to these components in isolation. This motivated the development of a new radiation scheme 'ecRad' (Hogan and Bozzo, 2016) that became operational in 2017 (Cycle 43R3). As well as being 31% faster than McRad when run in the same

Table 1: Time-line of the main developments in the ECMWF radiation scheme since 2000. See section 2 for further explanation.

Cycle	22R3	23R4	25R1	26R3	28R3	32R2	35R3	41R1	41R2	43R1	43R3
Month/Year	6/2000	6/2001	4/2002	10/2003	9/2004	6/2007	9/2009	5/2015	3/2016	11/2016	7/2017
Package name	Morcrette					McRad					ecRad
SW spectral bands	4		6			14 (RRTM-G)					
LW spectral bands	16 (RRTM-G)							Updated to latest RRTM-G			
Liquid cloud optics	Fouquart (1987), Smith and Shi (1992)					Slingo (1989), Lindner and Li (2000)					SOCRATES
Ice cloud optics	Ebert and Curry (1993)					Fu (1996), Fu et al. (1998)					
Aerosol climatology	Tanré et al. (1984)			Tegen et al. (1997)							CAMS
Ozone climatology	Fortuin and Langematz (1994)						GEMS	MACC		CAMS	
Other gases	Constant mixing ratio						GEMS	MACC			
Solver	Clear/cloudy regions					McICA					Reduced noise
Surface albedo	ERBE						60-km MODIS	5-km MODIS			
Frequency	3 h		1 h in data assimilation			1 h in HRES				Approx. updates, better sun angle	
Total solar irradiance	1366 W m^{-2}									1361 W m^{-2} ($\pm$ solar cycle)	

Table 2: Comparison of efficiency-relevant parameters of the radiation schemes used at several NWP centres in their global forecast configurations, where HRES is the highest resolution deterministic system and ENS is the ensemble system. Information on non-ECMWF models was kindly provided by Yu-Tai Hou, Günther Zängl, Quentin Libois, James Manners, Paul Vaillancourt and Shoji Hirahara. 'Horizontal coarsening' is the ratio of model to radiation gridpoints. The final row shows estimates of the number of bands and g-points (pseudo-monochromatic spectral intervals) needed using the full spectrum correlated-k approach (FSCK), which can be thought of as a lower limit to the number of g-points needed in the correlated-k paradigm. See section 8 for further explanation.

Centre	Radiation timestep (h)		Horiz. coarsening		Bands		Spectral intervals		Gas optics scheme (SW; LW)
	HRES	ENS	HRES	ENS	SW	LW	SW	LW	
ECMWF	1	3	10.24	6.25	14	16	112	140	Mlawer et al. (1997)
NCEP	1	1	1	1	14	16	112	140	Mlawer et al. (1997)
DWD	0.4	0.6	4	4	14	16	112	140	Mlawer et al. (1997)
Météo France	1	1	1	1	6	16	–	140	Morcrette (1991); Mlawer et al. (1997)
Met Office	1	1	1	1	6	9	21	47	Edwards and Slingo (1996)
CMC	1	1	1	1	4	9	40	57	Li and Barker (2005)
JMA	1	1 (SW), 3 (LW)	4	4	16	11	22	156	Freidenreich and Ramaswamy (1999); Yabu (2013)
FSCK	–	–	–	–	2	1	$\sim$ 15	$\sim$ 32	Pawlak et al. (2004); Hogan (2010)

configuration (see section 8), its modular design makes testing of new ideas much easier. In principle, different configurations of ecRad could be suitable for use in a wider range of atmospheric models, and indeed it has already been implemented in Météo-France's limited-area model Meso-NH.

Since the design of ecRad has facilitated the work described in sections 4, 5 and 8, and will provide the framework for further radiation developments in the future, we spend a little time here outlining the structure of the scheme. Figure 1 depicts the five components of ecRad and the flow of data between them. We first outline the four components that work on the radiation grid (lying between the two dashed lines in Fig. 1). The *gas optics* component computes the optical properties of gases in each spectral interval, as well as the spectrally-resolved Planck function and top-of-atmosphere incoming solar radiation. Currently only RRTM-G (Mlawer et al., 1997) is available, but it would be straightforward to add alternatives in future (see section 8.2). These arrays are passed to the *aerosol optics* component, which adds the contribution from aerosols. This component allows an arbitrary combination of hydrophilic and hydrophobic aerosol types whose optical properties are specified at run-time via a configuration file. The Tegen and CAMS aerosol climatologies are supported within this framework (see section 6.1) and it would be trivial to support alternative descriptions in future. The *cloud optics* component computes cloud optical properties in each spectral band (rather than spectral interval) and in addition to supporting the same schemes used in McRad, it also supports the Baran et al. (2014) and Yi et al. (2013) schemes for ice clouds. The *solver* component then combines the clear-sky and cloudy optical properties according to the cloud fraction and appropriate assumptions about in-cloud heterogeneity and overlap, from which the flux profiles are computed. Three solvers are available: McICA (Pincus et al., 2003), Tripleclouds (Shonk and Hogan, 2008) and SPARTACUS (Hogan et al., 2016), the latter which represents 3D radiative effects for the first time in a global model (see section 5.3). The McICA solver in ecRad includes an improved cloud generator (Hogan and Bozzo, 2016) that generates less stochastic noise in atmospheric heating than the McICA solver in McRad, as demonstrated in Fig. 2. Each ecRad solver can optionally account for longwave scattering, an effect only very approximately represented in McRad but which can increase the longwave cloud radiative effect by 10% (Costa and Shine, 2006).

Regarding the treatment of the surface, ecRad in Cycle 43R3 uses a simple single-tile description as in McRad, but Fig. 1 depicts the planned introduction of a *surface optics* component. This would act on the model grid and take as input a physical description of multiple surface tiles, including ones with more complex geometry such as urban areas and vegetated surfaces (see section 4.2). It would compute the grid-mean albedo and longwave emission seen by the atmosphere above, which would be interpolated on to the radiation grid and passed to the solver. The downwelling surface fluxes output from the solver and interpolated back to the model grid would then be used to work out how the radiation penetrates down to the individual facets of the surface.

The offline version of ecRad is available for non-commercial use under the terms of the OpenIFS license.

3 Climate of the IFS

Since radiation changes tend to primarily affect the climate of the model, this section provides a brief assessment of the temperature and surface radiation biases in recent operational cycles of the free-running model, and the impact of resolution and coupling to the ocean. This puts the changes discussed in later sections in context. The 'standard' free-running model experiment in this paper consists of four 1-year simulations covering the period 2000–2004 at T_L255 resolution (around 80 km) with 137 levels and the radiation scheme called every hour. Coupled simulations are performed in experiments when the tropospheric response is important, but uncoupled simulations are sufficient for testing changes that
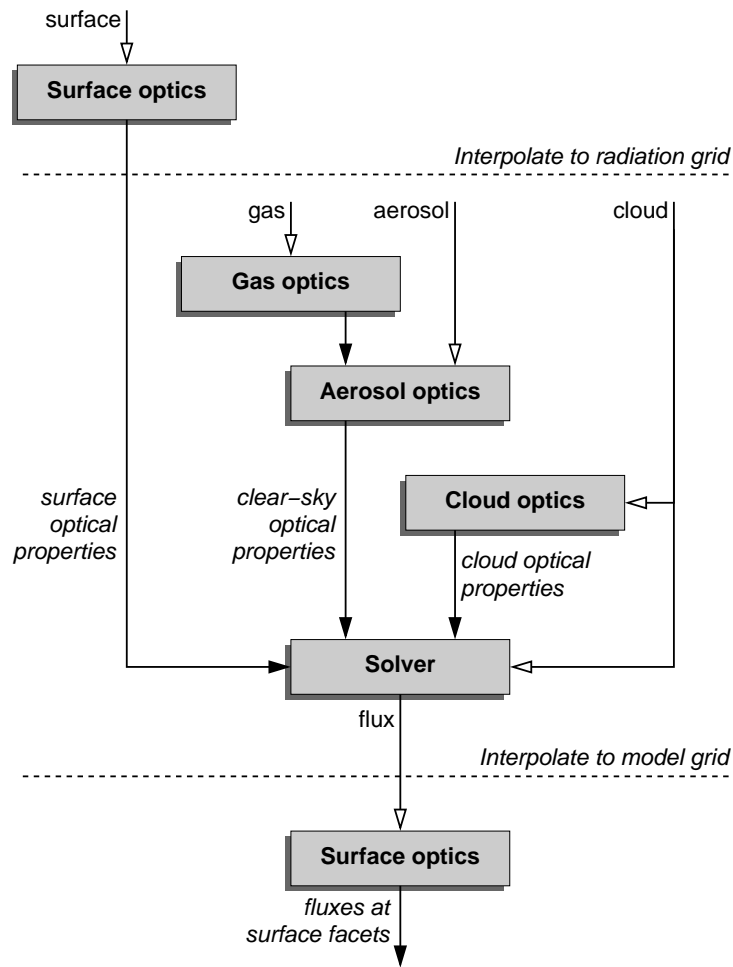
Figure 1: Schematic of the design of ecRad and the flow of data between each of the five components of the scheme. Open-headed arrows show incoming data structures containing a physical description of the system, and an outgoing data structure containing the flux profiles. Solid headed arrows denote groups of arrays containing optical properties of the atmosphere or surface in each spectral interval.

affect only the middle atmosphere. Such simulations are sufficiently long to compute global-mean fluxes to within 0.1 W m^{-2}, but not to pin down fine-scale temperature changes. Note also that 1-year coupled simulations are too short for the deep ocean to come into equilibrium with the atmosphere, but they match ECMWF's longest operational forecasting horizon, and are sufficient to capture the response of the ocean mixed layer.

We start by considering the radiation budget. Top-of-atmosphere global mean fluxes do not provide an independent test of the model as they tend to be used for the tuning of physical parameterizations, so we focus on surface fluxes. Table 3 compares the free-running IFS, ERA-Interim and ERA5 [which use the ecRad, McRad and Morcrette (1991) radiation schemes, respectively], and the average of a large number of coupled climate models, to the observational estimates of Wild et al. (2015). Over land, the climate models tend to overestimate shortwave downwelling flux by on average 8 W m^{-2} and underestimate longwave downwelling by 4 W m^{-2}, but there is a large spread amongst the models. ERA-Interim and ERA5 typically have around half the error of the mean of the climate models. When considering the free-running coupled simulations with the latest IFS cycle, it is striking that shortwave and longwave downwelling fluxes over land and globally all agree with the observations to within 1 W m^{-2}. The un-
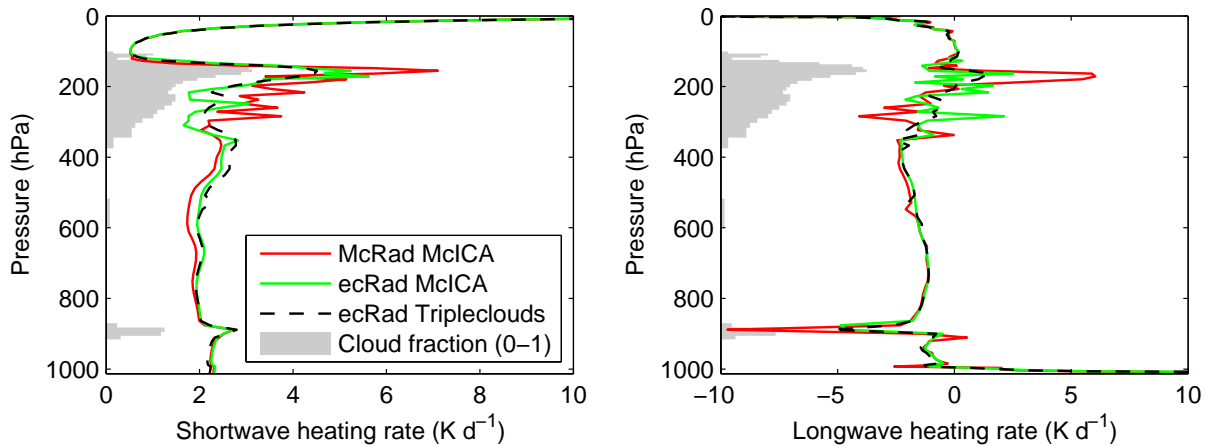
Figure 2: Comparison of the noise in instantaneous atmospheric radiative heating rates: the (left) shortwave and (right) longwave heating-rate profiles for a single 137-level IFS profile, comparing the McICA solver in the McRad radiation scheme (with noise up to 5 K d⁻¹), the reduced-noise McICA solver in ecRad, and the noise-free 'Tripleclouds' solver (available as a more expensive option in ecRad). The grey shading indicates the cloud fraction profile.
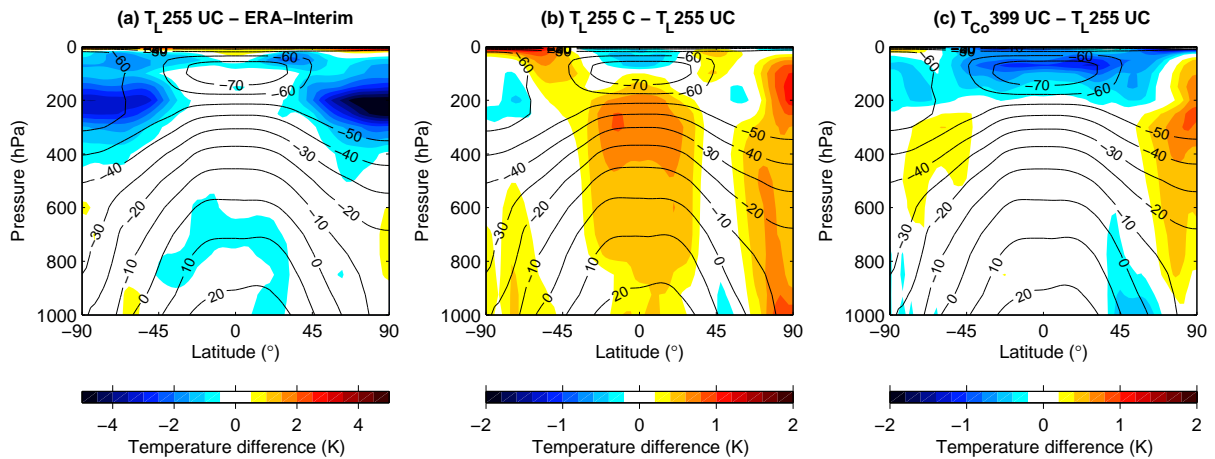

Figure 3: The black contours show the annual-mean temperature from four 1-year free-running simulations of Cycle 43R3 in the following configurations: (a) T_L255 uncoupled, (b) T_L255 coupled, and (c) $T_{Co}399$ uncoupled. The coloured contours indicate the difference with (a) ERA-Interim and (b–c) T_L255 uncoupled. Note the different colour scales.

coupled simulation is very similar. This gives us confidence that we can study regional biases associated with the surface and clouds in sections 4 and 5 in the knowledge that the global-mean radiation budget of the model is in good shape. It is perhaps surprising that the free-running models outperform both reanalysis datasets, but the radiation scheme was different in each case, and the reanalyses assimilate neither direct cloud observations nor cloud-affected solar or infrared radiances. Therefore even though temperature, pressure and humidity are strongly constrained in reanalyses, the cloud fields are entirely model generated.

Figure 3a compares the annual-mean temperature of the uncoupled IFS simulation with ERA-Interim. The tropospheric temperature bias is also very small (except for a cold bias in the tropical lower troposphere of up to around 0.5 K). Note that the tropospheric climate has not always been so good; many

Table 3: Comparison of estimates of global-mean and land-mean surface fluxes (W m^{-2}). The observational estimates are from Wild et al. (2015), who used a global network of surface flux observations that were then extrapolated to fill in the gaps using the pattern from 43 CMIP5 and IPCC AR5 climate models, the averages and standard deviations of which are shown in the second line. The climate models were run from 1850 to 2005 and averages taken from the period 2000–2004. Two ECMWF reanalysis products are shown: ERA-Interim (see Berrisford et al., 2011) and the more recent ERA5. The coupled and uncoupled IFS values are from four 1-year T_L255 Cycle 43R3 simulations. The final five rows are all 2000-2004 averages.

	Global down		Global net		Land down		Land net	
	SW	LW	SW	LW	SW	LW	SW	LW
Observations	184.7	341.5	160.1	−56.7	184	306	136	−66
43 climate models	189±5	340±4	165±4	−58±3	192±10	302±7	141±8	−69±6
ERA-Interim	188.4	341.4	164.6	−56.6	187.6	304.0	140.1	−67.0
ERA5	188.2	339.2	163.8	−57.9	189.3	303.6	137.8	−67.1
Uncoupled IFS	184.5	339.3	160.6	−57.0	183.6	305.8	134.5	−65.6
Coupled IFS	184.3	340.6	160.4	−56.9	184.4	306.7	134.9	−66.0

model cycles over the last eight years have had a troposphere 0.5 K colder. The polar lower-stratospheric cold bias of up to 5 K is a long-standing feature in the IFS and is common to almost all global models; it is discussed further in section 7.2.

Figure 3b shows that when the atmosphere is coupled to the ocean the tropical troposphere warms by up to 0.8 K, replacing a slight cold bias by a slight warm bias. This is because the coupled model has warmer sea-surface temperatures (SSTs) than the uncoupled system with prescribed bulk SSTs from OSTIA.

As mentioned in the introduction, the small tropospheric bias is likely helped by some compensation of errors between different schemes. But it presents a challenge for introducing physically-based improvements that change the bias, since even if they significantly reduce errors in a particular region or regime, their impact on forecasts is judged by metrics such as root-mean-squared error that can be strongly sensitive to bias. A solution is to recognize which parameters of the physics parameterizations are based on sound physical principles or empirical evidence, and which are under-constrained observationally and so have been set in order to minimize large-scale biases. Then, if necessary, improvements to the physics parameterizations can be accompanied by tuning of the under-constrained parameters (such as autoconversion rate or cloud-edge turbulent erosion) in order that they do not degrade the model climate.

Two further challenges related to the IFS climate should be mentioned. Firstly, Fig. 3c shows a resolution dependence of tropical lower-stratosphere temperatures; increasing resolution from 80 to 28 km leads to a cooling of up to 1 K, and shorter experiments suggest that the higher resolutions used in medium-range forecasting are a further 1 K cooler. Recent experimentation has revealed that this is because higher resolutions resolve more waves, and more stratospheric wave breaking drives faster ascent in the tropical stratosphere. This resolution dependence makes it more tricky to test improvements to the stratosphere such as those discussed in section 7.

A second challenge is that, as in many global models (Hobbs et al., 2016), the atmosphere of the IFS appears not to conserve energy: when comparing top-of-atmosphere net radiative fluxes with surface net energy fluxes, we find that it erroneously generates energy at a rate of around 2.3 W m^{-2} in the most recent IFS cycles, regardless of resolution or whether it is coupled to the ocean. Hersbach et al. (2015) reported a figure of 1.6 W m^{-2} in the ERA-20CM 20th-century atmospheric model ensemble (based on IFS Cycle 38R2). The cause of this is also not yet known, although it is unlikely to be a problem with the radiation scheme since atmospheric radiative heating rates are computed from the divergence

of the net flux profile, and consequently any net heating or cooling of the atmosphere should exactly match the net radiative fluxes at surface and top-of-atmosphere. Mayer et al. (2017) highlighted the need to account carefully for additional contributions to the surface energy budget from the moist physics, such as melting snow (either immediately when falling into the sea, or the slower melting of snow over land) and the fact that raindrops are usually colder than the surface on to which they fall. In the IFS, the heating or cooling of falling rain is currently ignored. An additional contribution to the lack of energy conservation is that the semi-Lagrangian advection scheme does not strictly conserve water vapour or thermal energy. Clearly further work is needed.

4   Challenge 1: Surface

Screen-level temperature is probably the forecast product of most interest for users, but biases can be large and have different causes in different locations. Section 4.1 summarizes the regional biases in operational 2-m temperature forecasts and explores in which situations radiation improvements can reduce these biases. Then in section 4.2 we describe the radiative changes needed to support a more radical upgrade to the surface scheme. One anticipated upgrade is the introduction of an urban tile; despite around half of the world's population living in urban areas (and presumably a similar fraction of users of ECMWF products), cities are currently represented as forests, grass or crops in the IFS. Forecast biases then arise due to errors in moisture availability, roughness and heat storage, but urban areas also present the interesting 3D radiation problem of computing the fluxes into the various facets of the urban surface. We have an efficient solution suitable for use in NWP.

4.1   Understanding 2-m temperature errors

Figure 4 presents the operational forecast biases in daily minimum and maximum 2-m temperature in winter and summer, and clear regional patterns are evident. The daily maximum temperature in many tropical land areas is underestimated by around 2 K, while the daily minimum temperature in boreal forest regions tends to be overestimated by at least 2 K. Averaging over the European and North African region[1], we find daily maximum/minimum temperatures in summer/winter all to be too low by 0.5 K, except for the summer daily minimum which is too high by 0.5 K; these biases have been consistent in operations for the last five years.

Such biases have complex causes, and at a particular location they can be down to problems with any one of the radiation, surface, boundary-layer, cloud or convection schemes (Beljaars, 2017). In practice we also need to contend with the representativity of observations, especially for stations in mountainous areas that can experience unresolved orographic flows, and for which forecast errors are introduced when interpolating or extrapolating the temperature in the model to the station height.

Although Table 3 suggests that the surface radiation budget is accurate on a global scale, this hides substantial regional errors, so we first investigate the extent to which errors in 2-m temperature are related to errors in surface downwelling fluxes. Figure 5 compares a decade of monthly-mean operational fluxes to observations at two contrasting sites, while Fig. 6 compares the diurnal cycle of 2-m temperature at the same sites, composited over the last summer and winter in the timeseries. The downwelling shortwave bias at Cabauw is very low while the downwelling longwave is underestimated by around 5 W m^{-2}. The corresponding 2-m temperature is very well predicted, on average, being within 0.5 K at all times in the diurnal cycle. This is encouraging but Cabauw is a flat rural site in a very well observed

---

[1]25°N to 70°N, −10°E to 28°E

(a) Maximum temperature, NH winter

(b) Minimum temperature, NH winter

(c) Maximum temperature, NH summer

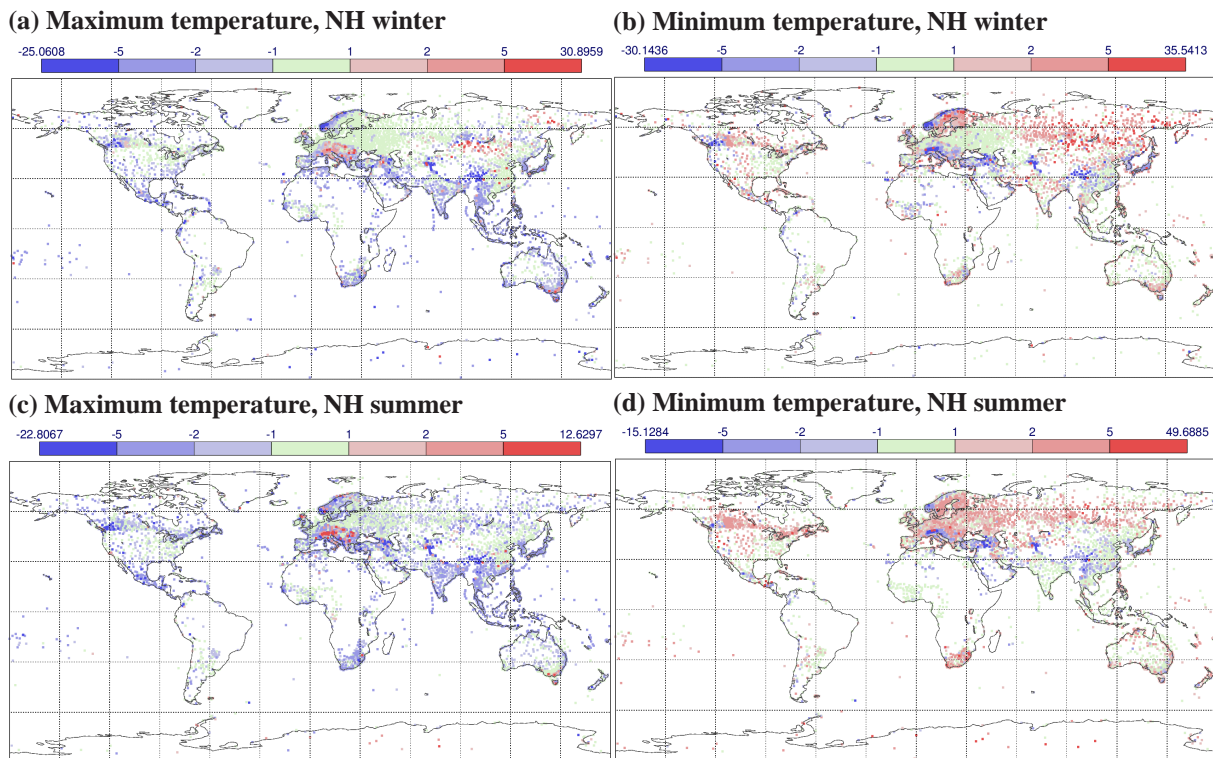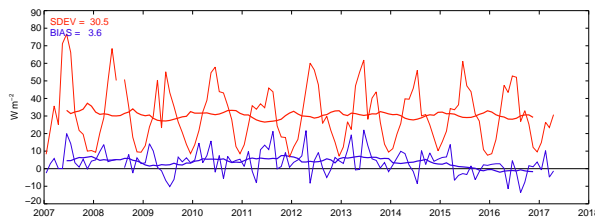**(d) Minimum temperature, NH summer**

Figure 4: Bias of daily maximum and minimum 2-m temperature from operational high-resolution uncoupled 36-h forecasts against screen observations, for (top row) December 2015 to February 2016, and (bottom row) June to August 2016.
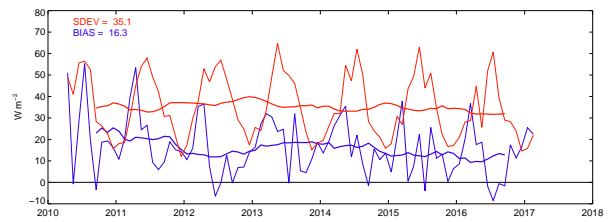
region. Sapporo is a much more complex site: it is a large coastal city surrounded by mountains with a very large annual snowfall. Shortwave is overestimated by 10 W m^{-2} in all seasons, while longwave is underestimated by around 10 W m^{-2} in summer and 25 W m^{-2} in winter. This signal is easiest explained by an underestimate in clouds; near-surface temperature is underestimated by 2–4 K, and in winter this signal could be a consequence of the large underestimate in longwave downwelling flux. However, the tendency for the largest errors to be at night is also a signature of the urban heat island effect (Oke, 1982), so the absence of urban areas in the model could also be partly to blame for this cold bias. Recently, 2-m temperature became an additional 'headline score', so more emphasis will be placed on its routine verification. In order to untangle the numerous factors that lead to biases at particular locations, it will be necessary to focus on 'super-sites' that measure not just 2-m temperature and downwelling fluxes shown here, but also sensible and latent heat fluxes, and cloud profiles.

An example of a specific 2-m temperature error entirely due to radiation was the coastline problem, whereby clear-sky nighttime forecasts at certain coastal gridpoints could be more than 10 K too cold. This was identified as being due to the use of a coarser radiation grid: warm sea points were used to compute upwelling longwave fluxes that were then applied over colder land points, leading to runaway cooling. This was solved in IFS Cycle 41R2 (see Table 1 for the dates when individual IFS cycles were operational) with the implementation of the Hogan and Bozzo (2015) scheme, which performs an approximate update of the broadband fluxes every timestep and model gridpoint to make them consistent with the skin temperature and surface albedo. Figure 7 demonstrates the improved agreement with 2-m temperature observations at a point on the coast of Norway. Note that not only did the scheme cure the cold nighttime biases, it also cured warm biases at desert coastlines due to the sharp albedo contrast, and improved the diurnal cycle at inland points in model configurations that call the radiation scheme only
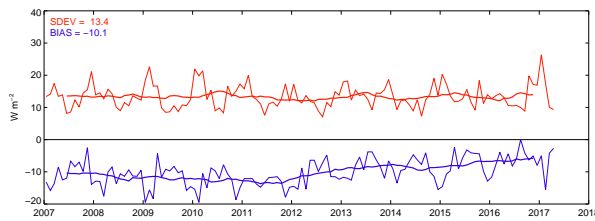
(a) Cabauw shortwave

(b) Sapporo shortwave
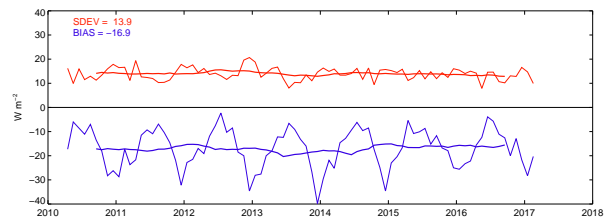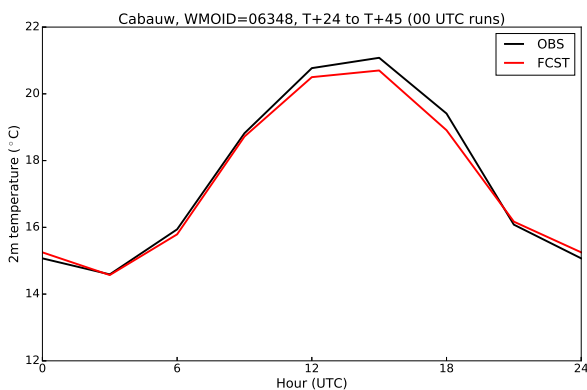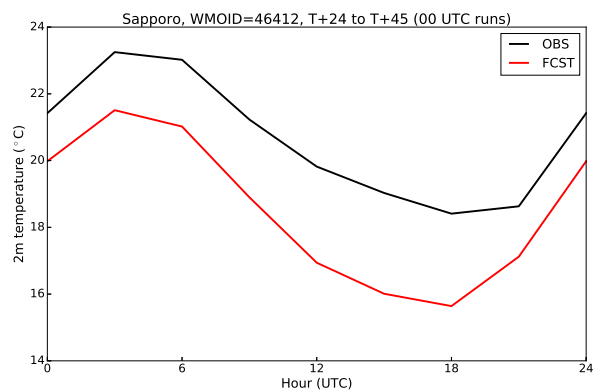
(c) Cabauw longwave

(d) Sapporo longwave

Figure 5: The bias (blue) and error standard deviation (red) of operational 72-h forecasts of surface shortwave and longwave downwelling fluxes at Cabauw, The Netherlands (52.0°N, 4.9°E) and Sapporo, Japan (43.1°N, 141.3°E). Observational data were provided by the Baseline Surface Radiation Network (BSRN) project.
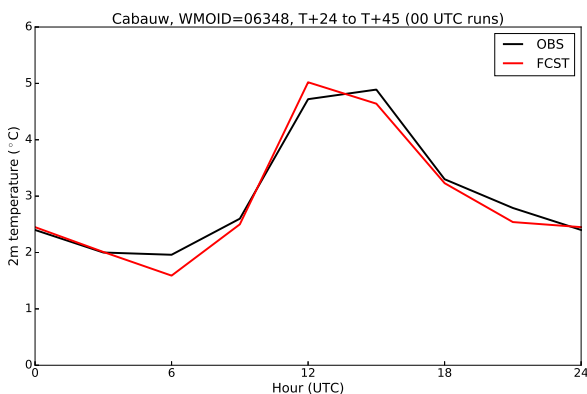
(a) Cabauw June to August 2016

(b) Sapporo June to August 2016

(c) Cabauw December 2016 to February 2017

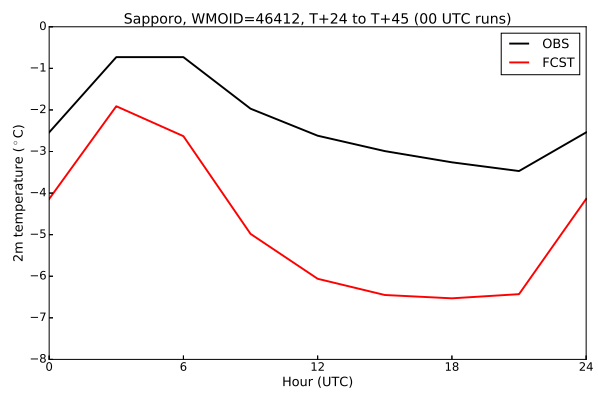**(d) Sapporo December 2016 to February 2017**

Figure 6: Diurnal composite of observed and forecast 2-m temperature at Cabauw and Sapporo in summer and winter, for comparison with Fig. 5.
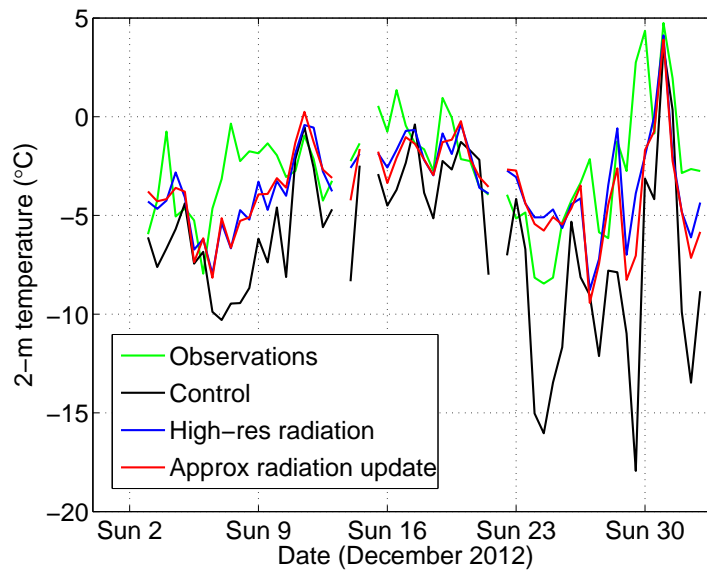
Figure 7: Comparison of T_L639 00-UTC and 12-UTC 2-m temperature forecasts (at lead times of 24 h and 36 h, respectively) against observations at Sortland, Norway (68.7°N, 15.42°E) for December 2012. The black line is for a radiation grid with 6.25 times fewer gridpoints than the model, the blue line for radiation called at every model gridpoint, while the red line is as the black but with the approximate update scheme of Hogan and Bozzo (2015) applied at every timestep and gridpoint.

every 3 h.

Morcrette et al. (2008a) reported an upgrade of the snow-free land surface albedo dataset to use a MODIS-derived climatology that prescribes separate albedos for direct and diffuse solar radiation in two spectral regions. The impact on forecasts was rather weak, increasing mean 850-hPa temperatures by only 0.08 K after 10 days. The resolution of the albedo dataset was improved in Cycle 41R1 (see Table 1), along with better snow-clearing. The next step will be to modify the scheme to account fully for the dependence of albedo on sun angle (Schaaf et al., 2002).

The computation of snow albedo is also in need of improvement; Beljaars (2017) found that the largest 2-m temperature errors in Northern Hemisphere spring occurred over exposed snow (i.e. unshaded by high vegetation). The treatment of snow in the IFS was described by Dutra et al. (2010) and includes a somewhat arbitrary treatment of the age of snow as an exponential decay in albedo with time, intended to represent the effects of particle deposition. Another area that needs revisiting is the reduced impact of snow on gridbox-mean albedo in forested areas, which Dutra et al. (2010) treated via a simple fit to MODIS satellite data . The next section describes a much more physically based method for accounting for snow in forested areas in the radiation scheme.

4.2 Representing more complex surfaces

The IFS surface scheme currently describes the surface in terms of 9 tiles within which 20 land-use types are possible, but the only information passed from this very detailed physical description to the radiation scheme is a gridbox-averaged skin temperature and a modification of the albedo to account for snow coverage. In order to represent more complex radiation processes in some of these tiles, the capability has been added to ecRad to work with the properties of individual tiles, rather than just gridbox averages. This will support the planned addition of an urban tile in the surface scheme (Balsamo et al.,
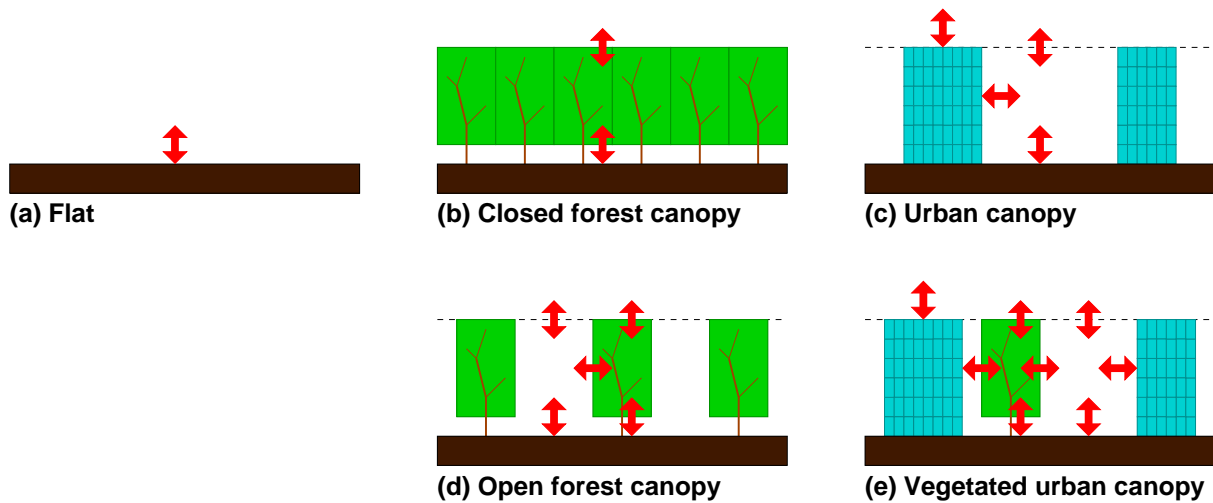
Figure 8: Schematic of the five surface geometries supported by ecRad in both the shortwave and longwave, in order of complexity. Trees are assumed to permeable to radiation while buildings are impermeable. The red arrows indicate the radiation exchanges that are handled explicitly in the scheme. The horizontal dashed lines indicate the separation of air within the canopy and air in the lowest model level above.

2014); correct treatment of the 3D effects in street canyons would then underpin predictions of the urban heat-island effect. It will also enable rigorous treatment of 3D effects in forest canopies leading to a more accurate dependence of albedo on sun angle, especially when there is snow on the ground (as demonstrated below). This will enable not only better predictions of canopy and near-surface temperatures, but also rigorous calculations of absorbed photosynthetically active radiation by the canopy, and hence evapotranspiration and (in the case of the CAMS configuration of the IFS) carbon dioxide uptake (Boussetta et al., 2013). It would also ensure radiative consistency with future improvements to the treatment of leaf-area index (LAI); experiments have demonstrated reduced forecast errors in 2-m temperature when replacing the operational LAI climatology with near-real-time satellite estimates (Boussetta et al., 2015), and testing of a prognostic LAI scheme is ongoing.

Figure 8 depicts the five surface geometries envisaged, where 'flat' is obviously what we have at the moment and 'closed forest canopy' uses the standard two-stream approach of Sellers (1985). The last three types involve horizontal fluxes into the walls of buildings and between trees and clearings. It should be stressed that adding this additional complexity in the radiation scheme would need to go hand-in-hand with an upgrade to the surface scheme to add equivalent turbulent fluxes of heat and moisture, such as via a resistance network (e.g., Best et al., 2011). The use of a single layer to represent buildings and trees can be thought of as the minimum degree of additional complexity needed to capture the main effects of vertical surfaces on both radiative fluxes and fluxes of heat and moisture, and is comparable to a recent implementation of urban areas in a global model by Li et al. (2016a). The aim is that the scheme should require a minimum of additional physiographic variables, and the extra radiation calculations through a vegetation or urban canopy would be comparable in computational cost to adding a single extra atmospheric layer.

To solve the radiative transfer problem, we have adapted the 'SPARTACUS' approach (SPeedy Algorithm for Radiative TrAnsfer through CloUd Sides) originally for computing the 3D effects of clouds (Hogan et al., 2016). The two-stream equations are solved for tree crowns and the open areas between them, but with additional terms to represent lateral transport of radiation between these two regions and, in the case of urban areas, into the walls of buildings. These terms are proportional to the area of the
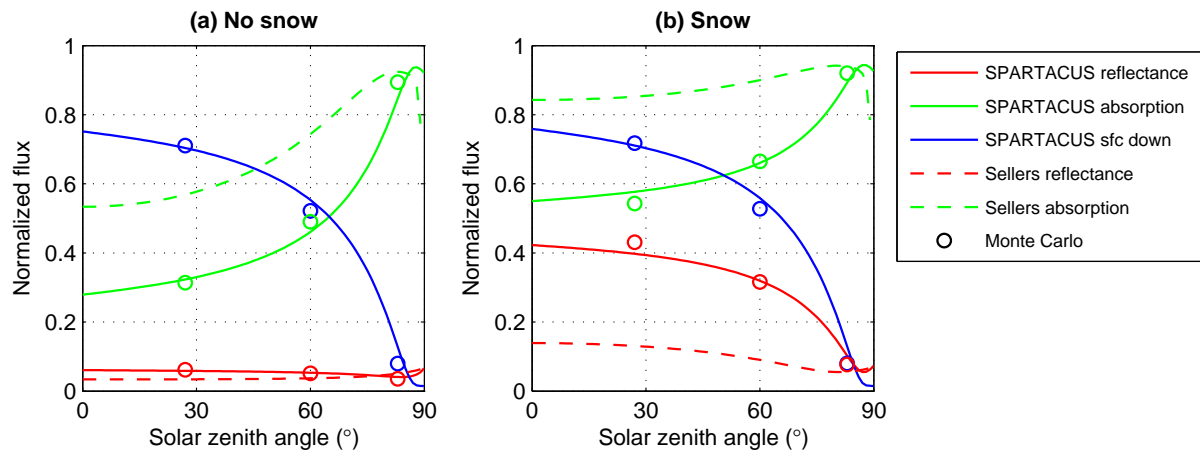
Figure 9: Evaluation of the SPARTACUS method to compute the shortwave radiative properties of an idealized open forest canopy from the RAMI4PILPS intercomparison, in which trees are represented by homogeneous 10-m spheres with an areal tree cover of 0.3 and a horizontally averaged leaf area index of 1.5. The optical properties of leaves are appropriate for the visible (photosynthetically active) region, and the surface albedos used are (a) 0.122 and (b) 0.964. The fluxes have been normalized by the incoming solar flux at the top of the canopy. The Sellers (1985) calculations were kindly provided by Tristan Quaife and Renato Braghiere. See Hogan et al. (2017) for further details.

interface between regions. This scheme requires very few geometric descriptors; for the 'urban canopy' and 'open forest canopy' types it takes just three: the canopy depth, the areal building/tree fraction, and an 'effective horizontal scale' for the buildings/trees. Trees and buildings are implicitly assumed to be randomly spaced and randomly oriented, which means there is no need for the empirical adjustments used in the vegetation radiative transfer scheme of Pinty et al. (2006), and there is no explicit assumption of a single street width (e.g., Li et al., 2016a).

In terms of sourcing the physiographic data required, forest height and tree cover are available from satellite (Simard et al., 2011; Hansen et al., 2003). Global urban data from satellite are generally limited to urban fraction (e.g., Schneider et al., 2009), but in combination with data on individual buildings (such as shown in Fig. 11a) for a limited number of cities, it should be possible to follow the example of Li et al. (2016b) and generalize properties to classes of urban area in a wide geographic region.

To test the validity of the SPARTACUS method for the 'open forest canopy', Fig. 9 shows a shortwave comparison with Monte Carlo calculations performed in the RAMI4PILPS intercomparison exercise (Widlowski et al, 2011) with and without snow on the surface. The agreement is excellent for all predictors: scene reflectance, absorption by the canopy, and the downwelling flux at the surface. This contrasts with the poor performance of the Sellers (1985) approach, which homogenizes the canopy horizontally (representing it as a 'closed forest canopy' in Fig. 8). Note that the Sellers scheme is used in state-of-the-art surface models such as JULES in the Met Office global model (Best et al., 2011). Of particular interest in Fig. 9b is the strong dependence of scene reflectance on solar zenith angle. This behaviour is not captured at all in the current albedo scheme, and yet forests with snow on the ground cover around 20% of land areas north of $20°$N in winter (Beljaars, 2017).

We next test this approach for urban areas. Figure 10 depicts the three geometric variables calculated from building data in London. We see that building fraction and mean building height are much higher in central London than in the suburbs, but effective building size[2] has very little variation; its mean value

---

[2]'Effective building size' is exactly analogous to the 'effective cloud diameter' introduced by Jensen et al. (2008): it is the

(a) Building fraction **(b) Mean building height (m)** **(c) Effective building size (m)**
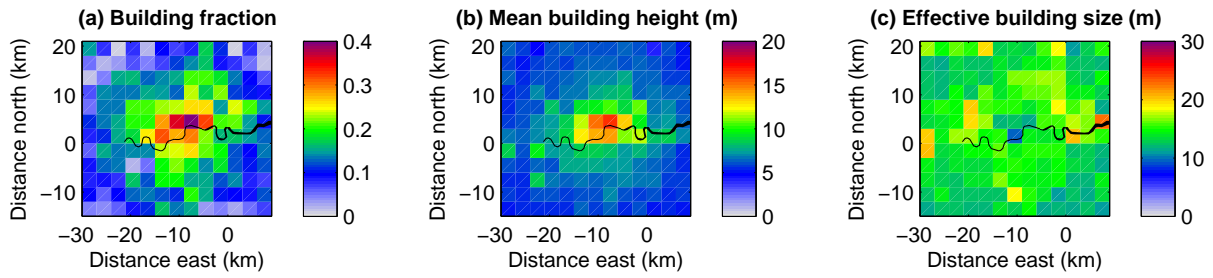
Figure 10: Building properties over London: (a) areal fraction covered by buildings, (b) mean building height, and (c) effective building size. Distances are measured relative to Greenwich at 51.48°N, 0.00°E. The River Thames is marked in black. The building database was provided by Emu Analytics.
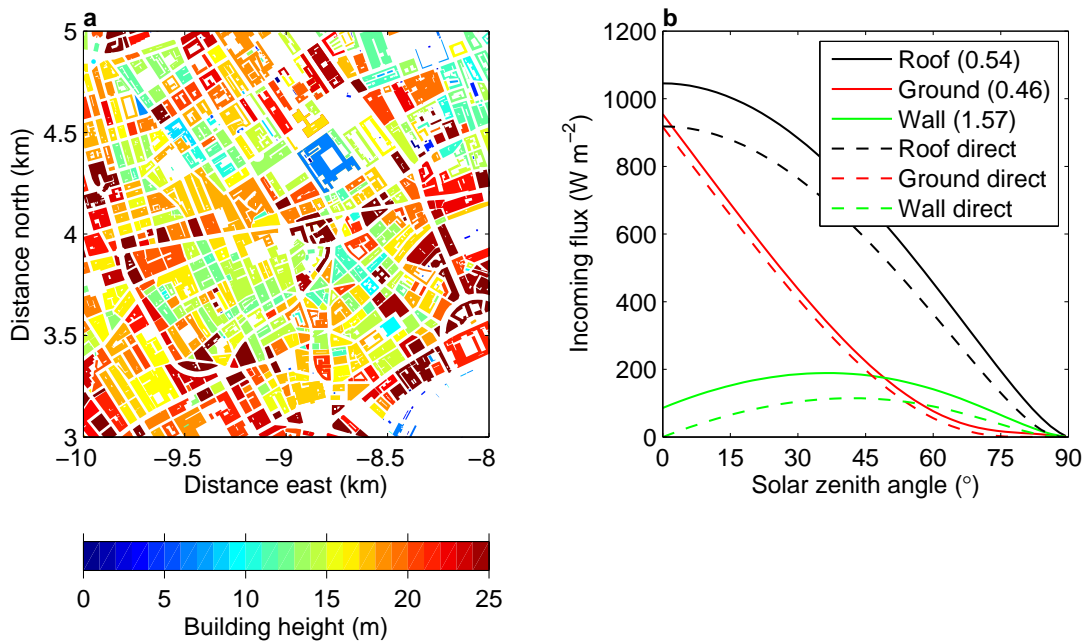

Figure 11: (a) Building geometries and heights for Central London measured relative to 51.28°N, 0.0°E (data from Emu Analytics); (b) corresponding SPARTACUS calculations of mean broadband shortwave fluxes into the roofs, ground and walls, where the numbers in brackets in the legend indicate the area of each of these surfaces divided by the horizontal area of the domain. The dashed lines provide the direct-beam solar fluxes. The calculations used a mean building height of 16.7 m and an effective horizontal building size of 10.6 m. The atmosphere above was the mid-latitude summer standard atmosphere with no clouds but a typical continental aerosol profile, and a broadband albedo of 0.2 was assumed for all surfaces.

is 14.3 m but its standard deviation is only 2.6 m. While further work is needed to compute it for other cities worldwide, the fact that it is approximately constant within a city suggests it is quite amenable to parameterization, and it might be possible to assign the same value to all the cities in a large region, as done by Li et al. (2016b). Figure 11 demonstrates the calculation of mean solar fluxes into the roofs, walls and ground for a 2×2 km region of Central London. The strong shading of street-level fluxes by high buildings is clearly evident when the sun is low in the sky. This offers the possibility for coupling

---

diameter of a cylindrical building in an equivalent idealized city composed of equally-sized cylindrical buildings with the same areal coverage and perimeter length as the actual city.

with an urban energy-balance scheme in which not only could the heat storage by the three main urban surfaces be computed, but separate forecasts of temperature could be produced at street level and roof level. While this is still simpler than many detailed urban energy-balance schemes, it is commensurate with the needs of a global model, and it was found by Grimmond et al. (2010) that simpler schemes performed just as well as complex schemes when evaluated against observations.

An additional consideration for the future, particularly as the resolution of the model increases, is the representation of 3D radiative effects in mountainous areas. These not only affect near-surface temperature forecasts, as is apparent over the Alps in Fig. 4, but also the triggering of convection. It was estimated by Manners et al. (2012) that the effect of the orientation of the terrain on the interception of the direct solar beam was around four times more important than the longwave sky-view effect in valleys, so the shortwave problem should be tackled first. The technical challenge is that to treat the problem correctly requires horizontal communication between columns, which breaks parallelism: if the surface in one gridbox receives less incoming sunlight because it is tilted away from the sun then the excess sunlight ought to be transported into the adjacent column. If this horizontal communication is ignored then strict energy conservation is violated because the downwelling shortwave flux exiting the base of the lowest atmospheric layer does not equal the downwelling flux seen by the surface. This approach was taken by Manners et al. (2012), although they commented that energy should be approximately conserved over an extended region. A simple but strictly energy-conserving alternative that could be applied in the IFS is to represent the orientation of the terrain with respect to the sun, and potentially also terrain shadowing effects, by dynamically adjusting the surface direct-beam albedo to ensure the correct shortwave absorption by the surface; this way the energy excess (or deficit) would not be ignored but would go into increased (or decreased) surface reflection. This technique could be applied at the resolution of the model grid rather than the coarser radiation grid by incorporating it into the approximate-update scheme (Hogan and Bozzo, 2015), in which the shortwave flux profiles on the radiation grid are modified to be consistent with the high-resolution albedo on the model grid. If this approach shows a positive impact then representation of the longwave sky-view effect could be considered as well (e.g., Senkova et al., 2007).

5   Challenge 2: Clouds

Accurate representation of the interaction of clouds in the radiation scheme is of foremost importance both for medium-range weather forecasts, and for establishing a good tropospheric climate needed for monthly and seasonal forecasts. The cloud scheme provides the radiation scheme with profiles of cloud fraction and the mixing ratios of liquid cloud, ice cloud, snow and rain. But this leaves the radiation scheme with a considerable number of assumptions to make, each of which can significantly change the radiative impact of the cloud.

Where possible we want the assumptions made by the cloud and radiation schemes to be the same, and the last decade has seen improvements in consistency, for example making snow active in radiation (Li et al., 2014). Morcrette et al. (2008a) anticipated the development of a PDF-based cloud scheme, in which the width of the cloud water distribution could also be fed to the McICA radiation scheme. In practice, it has not been deemed advantageous to make such a radical change to the prognostic variables of the cloud scheme, although recently we have tested the introduction of an empirical, diagnostic, regime-dependent PDF width that is consistent between the cloud and radiation schemes (Ahlgrimm and Forbes, 2016).

Also in need of attention is the fact that there is currently very little consistency between the assumptions made in the radiation scheme and in the forward operators used for assimilating cloud-affected radiances

in the data assimilation system. For example, the microwave and infrared operators assume clouds to be maximally overlapped, while the radiation scheme assumes clouds to become more randomly overlapped as their vertical separation increases. Geer et al. (2017) described the numerous other differences in assumed cloud properties, so we do not discuss the issue further in this paper.

While the discussion of cloud-radiation interactions in this paper is concerned with improving *mean* behaviour, an important task in the ensemble prediction system is to improve the description of the *uncertainty* in instantaneous cloud properties, in order that the ensemble spread provides a good estimate of root-mean-squared forecast error. At ECMWF, this is being tackled in the new Stochastically Perturbed Parameterization (SPP) scheme, in which uncertain physical parameters describing clouds and other variables are perturbed differently in each member of the forecast ensemble. Additionally, improvements are being tested to the older Stochastically Perturbed Parameterization Tendency (SPPT) scheme, by separating out the uncertainty in the heating-rate profile due to clear-sky radiative heating (which is quite accurately predicted) from that due to clouds (which is much more uncertain). Since these advances were described in detail by Leutbecher et al. (2017), we do not describe this work further here.

In terms of improving the cloud-related radiation errors in the model, section 5.1 discusses the first task: using a variety of observations to characterize the biases in the radiative impact of clouds, and trying to determine to what extent they can be attributed to errors in the model cloud fields, or in their treatment by the radiation scheme. Section 5.2 examines each of the assumptions made by the radiation scheme and recent work to improve them. Then in section 5.3 we provide a first estimate of the impact of 3D radiative transfer on the climate of a global model, made possible by ecRad being the first radiation scheme in such a model to have the option to represent 3D radiative effects.

5.1   Separating cloud errors from radiation errors

Figure 12 evaluates the top-of-atmosphere (TOA) shortwave cloud radiative effect (CRE; the net downward flux at TOA minus what would be expected in clear skies) against the best estimate from the CERES instrument[3]. The fact that the global-mean shortwave CRE agrees to within 0.5 W m^{-2} is not particularly informative since this value is routinely used to assess and tune changes to the cloud, convection and radiation schemes. Rather, it is the very substantial regional errors in shortwave CRE that highlight the priorities for model improvement. An example of the impact of these errors on local forecasts was shown in Fig. 6, where near-surface temperature errors at Sapporo were believed to be at least partially related to an underestimate in the magnitude of the CRE. The challenge from the point of view of radiation development is to diagnose which errors are due to the underlying clouds in the model and which are due to the way their interaction with radiation is treated. It is undesirable to mitigate errors in the cloud or convection schemes by tuning the radiation scheme in a particular way, so the strategy is to make all three schemes as physically and observationally based as possible, and to aim for consistency between the assumptions made by the schemes, and then if needed to tune the under-constrained parameters such as autoconversion rate or cloud-edge turbulent erosion.

Some of the regional shortwave CRE biases have a very clear source outside the radiation scheme. For example, the underestimated reflectance of marine stratocumulus off the west coasts of the major continents is known to be due to the underestimated cloud amount, a bias shared by all the climate models evaluated by Calisto et al. (2014). This bias has improved in the IFS over the last decade; Morcrette et al. (2008a) reported that just after the implementation of McRad the maximum error in the Peruvian stra-

---

[3]Longwave CRE is not compared here because the observational estimate uses nearby cloud-free profiles that are systematically drier, leading to systematic differences with the model estimate that are difficult to account for (Allan and Ringer, 2003).

Figure 12: Comparison of annual-mean top-of-atmosphere shortwave cloud radiative effect (TOA SW CRE) between the Cycle 43R3 model climate and the CERES-EBAF observations. The model runs consisted of four 1-year free-running coupled simulations at T_L255 resolution.

tocumulus sheet was 60 W m^{-2}, but Fig. 12 shows that this has now been reduced to 45 W m^{-2}. Likewise, the underestimate in reflection over the Southern Ocean is common to many climate models (Bodas-Salcedo et al., 2014), but significant improvements have recently been made in the IFS by increasing the amount of supercooled water in these clouds (Forbes and Ahlgrimm, 2014; Forbes et al., 2016), and recent improvements in Cycle 45R1 have improved the bias even further (not shown).

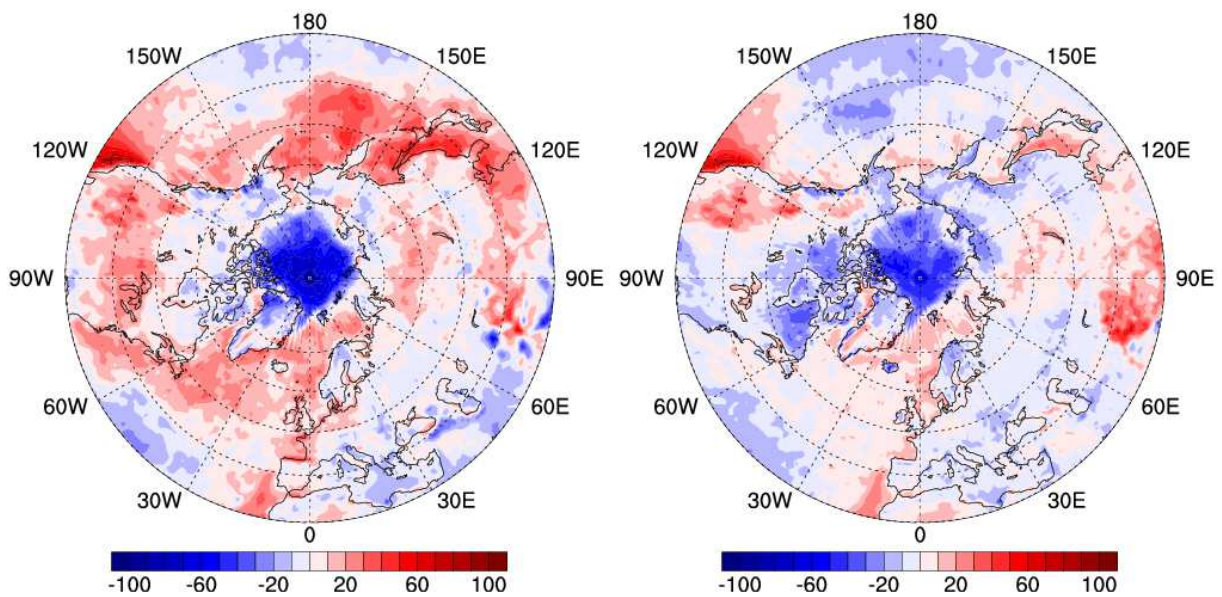*Figure 13: Bias in surface downwelling shortwave flux from (left) ERA-Interim and (right) ERA5 (W m^{-2}) as judged by the CERES satellite estimate, for the period June-August 2016.*

Another example of a radiation bias likely due to the clouds in the model is the strong underestimation of summertime surface downwelling shortwave flux over the Arctic Ocean (Fig. 13). This has important consequences for sea ice in the model, which does not melt enough in summer leading to large sea-ice cover biases in seasonal forecasts. Comparisons to the well-equipped surface stations of the International Arctic Systems for Observing the Atmosphere suggest that the bias is much weaker at the coast (Haiden, 2015). Sotiropoulou et al. (2016) attributed the bias to an overestimate in cloud occurrence in the model, which is corroborated by comparisons with the CALIPSO satellite. Summertime Arctic clouds are frequently mixed-phase, and are therefore particularly challenging to model (Forbes and Ahlgrimm, 2014), but the improvement between ERA-Interim and ERA5 shown in Fig. 13 is encouraging. It should be mentioned that the specification of surface properties is also a likely cause of forecast error, particularly in the shortwave where sea-ice albedo is strongly modulated by the presence of melt ponds. But there are issues to address in the longwave as well: Feldman et al. (2014) investigated the impact of the fact that the emissivity of the ocean is less than 0.89 at wavelengths longer than 20 μm, while most models assume much higher values (the IFS uses 0.99 for the entire longwave spectrum). This only impacts the Arctic where the atmosphere is dry enough that micro-windows open up in the far-infrared spectrum, but here the surface warming effect of lower emissivity was reported to be as much as 2 K. The Year of Polar Prediction, which started in 2017, will provide an opportunity for a more concerted effort to diagnose errors in polar forecasts and the extent to which these can be improved through changes to the treatment of clouds and radiation in the model.

There are other radiation biases whose source is less clear to diagnose, and therefore where more detailed analysis of surface measurements is required. For example, Van Weverberg et al. (2015) provided a methodology for using ground-based active and passive measurements to identify the source of 2-m temperature forecast biases over the continental United States. Another example apparent in Fig. 12 is the overestimate of reflectance in the trade cumulus regions that cover a large fraction of the ocean. If this is an example of the 'too bright, too few' error identified in many models by Nam et al. (2012), then one would expect these clouds to contain too much liquid water, but satellite estimates can give contradictory information. For example, at Hawaii the annual-mean liquid water path (LWP) in the IFS
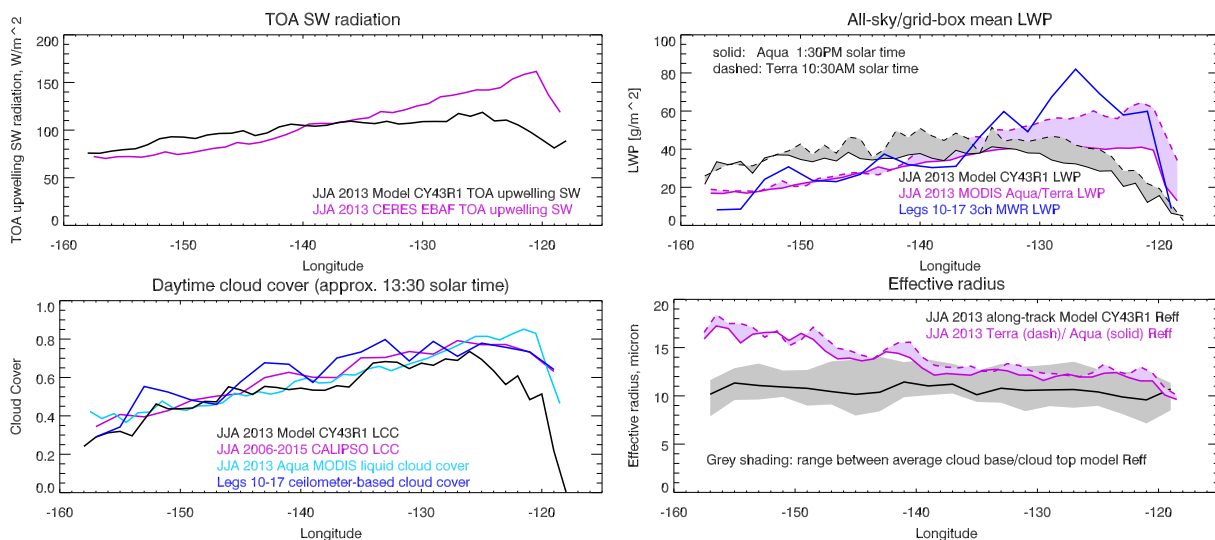
Figure 14: Transect of (black lines) model and (magenta lines) satellite data between (right site of plots) Los Angeles and (left side of plots) Hawaii averaged over JJA 2013, along with (blue lines) ship-borne estimates from eight legs of the MAGIC campaign. Note that the ship always set sail at the same time of day, so the ship-borne measurements include the diurnal cycle of cloud properties, while the MODIS instruments on Aqua and Terra each sample the clouds at just one point in the diurnal cycle.

is around 50 g m^{-2} which is 20 g m^{-2} too low according to the passive microwave estimates from SSMI but 20 g m^{-2} too high according to MODIS. Intercomparisons suggest that microwave-based LWP is overestimated in partially cloudy conditions (Seethala and Horváth, 2010), but conversely, MODIS only estimates LWP during the day and there is a significant diurnal cycle in marine boundary-layer cloud properties.

Work is in progress to unpick the marine boundary-layer cloud problem further by bringing in crucial ground-based measurements: Fig. 14 shows a summer transect of model, satellite and ship-borne observations from the Californian coast on the right, through the transition region from stratocumulus to cumulus, to Hawaii on the left. The top-left panel shows an underestimate of shortwave reflectance of up to 60 W m^{-2} in the stratocumulus and an overestimate of around 20 W m^{-2} in the cumulus. The bottom-left panel shows agreement between the daytime estimates of low cloud cover from the satellite-borne CALIPSO and MODIS instruments, and the minimum (daytime) values from the oscillating ship-borne ceilometer estimates that span the full diurnal cycle. The model significantly underestimates daytime stratocumulus cover near the Californian coast, but is close to the three observational estimates in the cumulus regime, suggesting that the model does not suffer particularly from the 'too few' bias reported by Nam et al. (2012). In terms of LWP (top-right panel), MODIS and the ship-borne measurements are in agreement that stratocumulus LWP is too low while cumulus LWP is too high, providing the most likely explanation for at least the sign of the shortwave CRE biases. However, the bottom-right panel of Fig. 14 suggests that effective radius is underestimated in the cumulus regime, which would also cause cumulus clouds to be too reflective. This ongoing study shows the need to understand the uncertainties in satellite estimates, and highlights the value of additional surface-based observations to characterize model error. The following two sections discuss a number of candidate mechanisms that could play a part in explaining radiative biases in cumulus and other cloud types.

More generally, further detailed studies using a wider range of observations are needed to target specific cloud biases in the model in different parts of the globe. The forthcoming EarthCARE satellite,
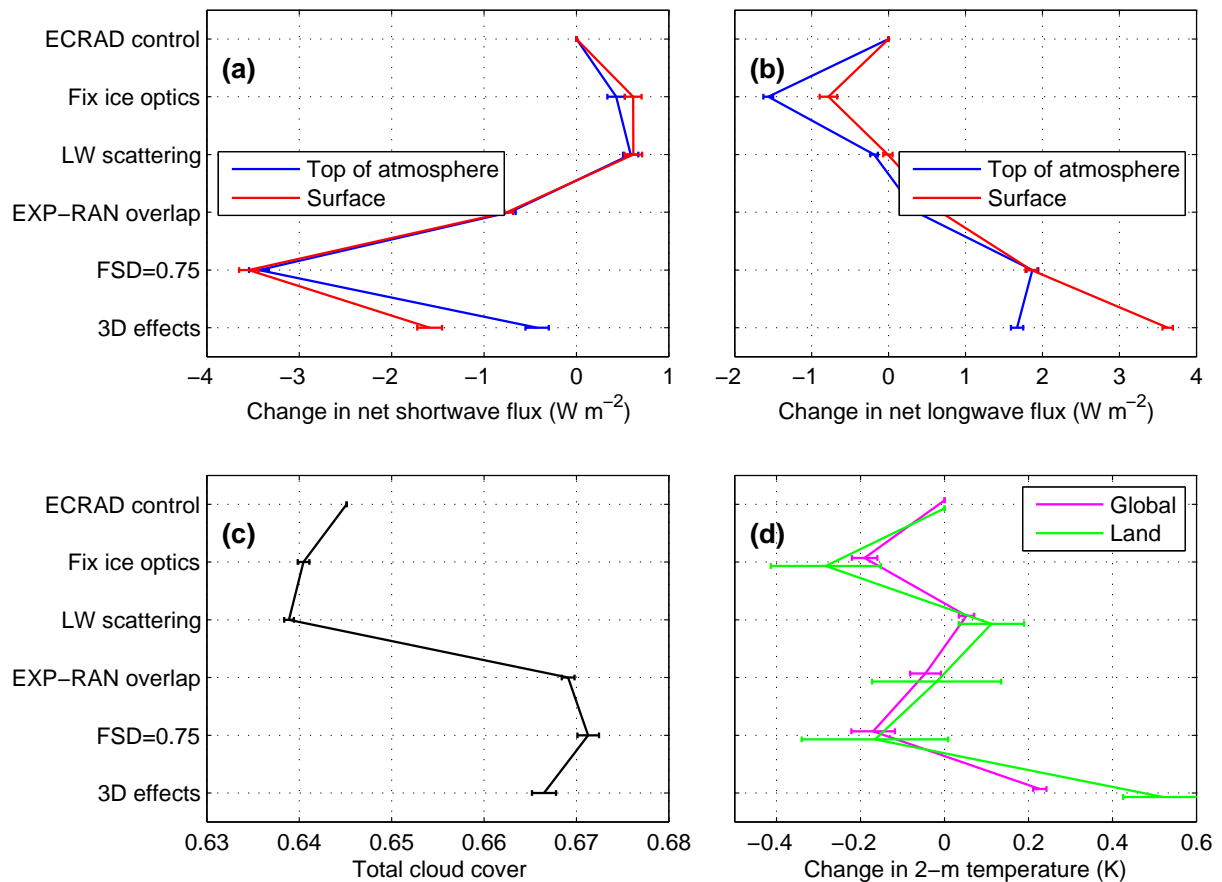
Figure 15: Impact of various modifications to the radiation scheme on (top row) surface and top-of-atmosphere net fluxes, and (bottom row) global-mean cloud cover and 2-m temperature. All modifications are cumulative in the sense that they include all the other changes made down to that point. They are computed from four 1-year free-running coupled T_L255 simulations, where the ecRad control is very close to the default configuration for Cycle 43R3.

which combines sophisticated synergistic cloud retrievals and broad-band flux estimates from the same platform, will provide a dataset that could enable this approach to be applied to cloud regimes globally. Geer et al. (2017) discussed ways to identify cloud biases in observation space using the all-sky assimilation system.

5.2   Improving treatment of clouds in the radiation scheme

In this section we discuss recent and anticipated future work on the various assumptions that need to be made to compute radiative fluxes in the presence of clouds. To put the impact of these improvements in context, Fig. 15 depicts the cumulative change that they produce in surface and top-of-atmosphere fluxes, cloud cover and mean 2-m temperature. The ecRad control refers to a configuration close to Cycle 43R3, so most of the changes presented could in principle be implemented operationally in the near future.

Longwave scattering is neglected in the ECMWF model but has been estimated by Costa and Shine (2006) to reduce outgoing longwave radiation by around 3 W m^{-2}, which is around 10% of the top-of-atmosphere longwave cloud radiative effect. It also makes the undersides of clouds partially reflective,

and their reflection of upwelling longwave leads to greater longwave downwelling, and hence warming, at the surface. In practice, there is a problem in McRad's treatment of the longwave optical properties of ice clouds that had to be reproduced in ecRad, because (as Fig. 15 shows) when it is fixed, the land surface cools by 0.3 K, exacerbating the existing cold bias. To a large extent, fixing the problem and then introducing longwave scattering returns the climate to a similar state to the control state. Nonetheless, turning on longwave scattering increases the total cost of ecRad by only around 10%, so for improved physical realism, the intention is to introduce it into a near-future operational cycle.

Cloud overlap determines the total cloud cover for a given profile of cloud fraction. It was found in observations by Hogan and Illingworth (2000) and others that cloud overlap has 'EXP-RAN' behaviour, i.e. vertically contiguous cloud layers have a correlation that decreases inverse-exponentially with their separation, while vertically separated cloud layers are randomly overlapped. At the time McRad was implemented, it was decided to use a cloud generator that implemented 'EXP-EXP' overlap (i.e. vertically separated cloud layers are correlated) coupled to an overlap decorrelation length of 2 km, which Barker (2008) found from CloudSat observations to predict about the right cloud cover using the same generator. A couple of years later, the overlap decorrelation length parameterization of Shonk et al. (2010) was adopted, which included a latitude dependence to fit cloud-radar observations at a number of ARM (Atmospheric Radiation Measurement) and European sites. However, the Shonk et al. (2010) parameterization was derived to work with an EXP-RAN scheme, leading to an inconsistency between the generator and the overlap decorrelation such that we should expect the IFS now to predict a cloud cover that is too small for a given profile of cloud fraction. To match the assumption in McRad, an EXP-EXP scheme is used in the default ecRad configuration in Cycle 43R3, but it would be desirable to switch it to an observationally supported EXP-RAN scheme. Figure 15 shows that the introduction of true EXP-RAN overlap would lead to a reduction in net surface shortwave flux of 1.4 W m^{-2}, an increase in global cloud cover of 0.03, and a cooling of the land surface by 0.1 K. A further change that could be considered for operations in future is to include the dependence of cloud overlap on wind shear (Di Giuseppe and Tompkins, 2015).

Cloud heterogeneity is another important factor controlling the impact of the cloud field on radiation. It may be characterized by the fractional standard deviation (FSD) of cloud water content or cloud optical depth, and the operational assumption in McRad, carried through to ecRad, is that FSD = 1. Shonk et al. (2010) reviewed a wide range of observational studies and found a large dependence on cloud type, but a mean value of 0.75. Figure 15 shows that the impact of reducing cloud heterogeneity to this value is to make the clouds more opaque to solar radiation, with a reduction of surface net shortwave flux of 2.7 W m^{-2} and a further surface cooling of 0.1 K, on average. Ahlgrimm and Forbes (2016) performed this experiment and reported regional reductions in annual-mean top-of-atmosphere net shortwave fluxes of up to 10 W m^{-2}. They also performed a detailed analysis of ground-based cloud observations from the ARM programme, and parameterized FSD as a function of cloud fraction and total water content; while showing that 0.75 was a reasonable average value, their parameterization captured the tendency for higher values in tropical than polar boundary-layer clouds, and the tendency for lower variability in overcast compared to partially cloudy situations. Their scheme also has the advantage of being consistent with the FSD assumed within the cloud microphysics parameterization. It turns out that it is not only the width of the water content distribution but the shape that is important: Hogan and Bozzo (2016) reported that replacing the default gamma distribution by a lognormal (keeping FSD the same) cooled the land surface by a further 0.1 K on average.

The combined effect of the introduction of more realistic cloud overlap and heterogeneity is to cool the land surface by around 0.2 K; the pattern of the change in coupled climate simulations is shown in the top-left panel of Fig. 16. In fact the entire troposphere is cooled, the top-right panel showing that this is
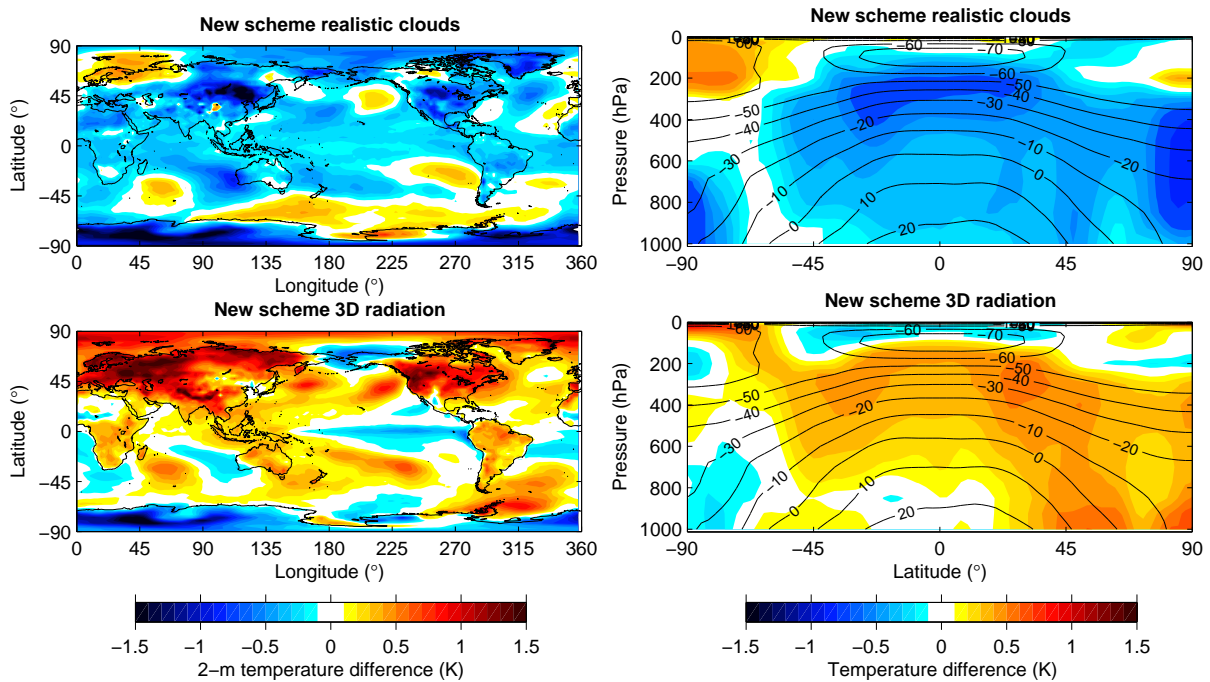
Figure 16: Impact on the ECMWF model climate of (top row) introducing realistic cloud overlap and heterogeneity, and (bottom row) introducing realistic cloud overlap and heterogeneity as well as introducing 3D radiative effects. The control against which both are compared is the ecRad configuration with longwave scattering shown in Fig. 15. The results are from 8 free-running 1-year coupled simulations at T_L255 resolution.

as much as 0.5 K in the tropical upper troposphere. In uncoupled deterministic forecasts this would be a degradation, and such a change would have to be accompanied by another modification that works in the other sense. In coupled forecasts the tropospheric cooling would act to counter the warming resulting from the coupling (Fig. 3b). The impact of 3D effects is shown in the lower panels of Fig. 16 and discussed in section 5.3.

Liquid cloud effective radius is currently computed from liquid water content and droplet number concentration, the latter which is constant with height but has a land/sea dependence, as well as a dependence on near-surface wind speed in an attempt to represent the increased lofting of aerosols from the surface. In practice, the value over sea does not exhibit realistic regional structure but rather has a value that is approximately constant except for some variability that depends on the instantaneous near-surface wind speed. The bottom-right panel of Fig. 14 illustrates the problem that this scheme fails to represent the tendency to have lower effective radius towards the coast associated with advection of continental aerosol, rising to higher values in the cleaner air of the open ocean. While recognizing that effective radius also depends on in-cloud LWP, one possible improvement would be to make use of the hydrophilic aerosol concentrations available in the aerosol climatology (see section 6.1), which exhibits a realistic variation with distance from the coast. If this resulted in a general increase in effective radius (as Fig. 14 suggests it ought to) then it would act to warm the ocean surface and hence to counter the cooling effect resulting from the improvements to cloud overlap and heterogeneity discussed above. *Ice effective radius* also needs attention since the scheme currently in use was developed before the finding that many of the aircraft observations on which parameterizations were based suffered from shattering on the probe inlets that led to an over-counting of small particles.

Cloud optical properties are being revised in ecRad. Nielsen et al. (2014) reported that the Slingo (1989)
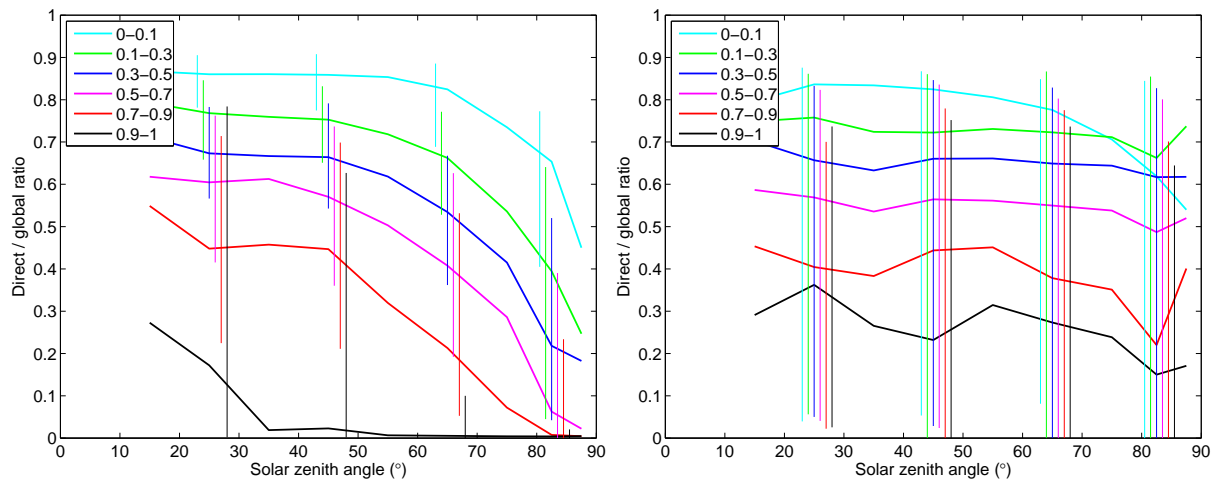
Figure 17: (Left) Ratio of direct to total downwelling surface shortwave flux from ten years of observations at the ARM Southern Great Plains site, for different ranges of low cloud cover (coloured lines) in conditions where no mid- or high-level clouds were present. (Right) The same from the ECMWF operational model for the calendar years 2011–2014 at 14 sites in the United States, Europe and Barbados.

parameterization of liquid cloud optical properties tended to overestimate optical depth, so ecRad uses by default the same parameterization as in the Met Office radiation scheme SOCRATES[4], which does not have this problem. The ice optics schemes are currently from Fu (1996) and Fu et al. (1998), but we are testing the Baran et al. (2014) and Yi et al. (2013) schemes, which account for more recent findings that ice particles are optically 'rough'.

Convective cloud and precipitation is currently ignored in the radiation scheme, but it has recently been found necessary to include it in the forward operator for all-sky microwave radiances. The main uncertainty is that the convection scheme does not predict the convective cloud and precipitation fraction, so a somewhat arbitrary value of 5% is assumed. We hope to test its inclusion in the radiation scheme in the next few years. This would also help to bring the assumptions in the radiation scheme and the radiance models used in data assimilation into better agreement.

5.3   Representing 3D radiative transfer

One process missing from the radiation schemes of all operational weather and climate models is 3D radiative transfer in the presence of clouds. This includes the flow of radiation through cloud sides as well as the horizontal transport within clear and cloudy parts of the atmosphere. For the purposes of an atmospheric model, 3D transport can be split into two parts. The first part is the lateral exchange of radiation *between* model columns, an effect that is only just beginning to be explored as cloud-resolving models with online 3D radiation schemes become available (e.g., Klinger et al., 2017). ECMWF is at least a decade from the resolution at which this effect is likely to be important. The second part concerns horizontal transport *within* a grid column, and can lead to biases in models of any resolution. In the shortwave, the direct solar beam can intercept the sides of clouds, leading to an increase in reflection when the sun is low in the atmosphere. In the longwave, emission from the sides of clouds leads to increased downwelling at the surface during both day and night.

Downwelling solar radiation in the model is split into direct (unscattered) and diffuse (scattered) compo-

---

[4]SOCRATES is the Suite of Community Radiation Codes based on Edwards and Slingo

nents, with the former being of most interest for the solar energy industry since solar panels are typically oriented towards the sun to maximize interception of the direct beam. Figure 17 presents observational evidence that the lack of 3D radiative effects in the operational ECMWF radiation scheme can lead to a significant overestimate of the fraction of surface downwelling radiation that is in the direct beam. The effect is strongest in broken low-cloud conditions when the sun is low in the sky (large solar zenith angle) since in the real world this leads to some of the direct beam intercepting cloud sides, an effect not represented in the radiation scheme.

A method to represent 3D radiative effects in a large-scale model has recently been developed by Hogan et al. (2016) in the form of the SPARTACUS solver, which is included as an option in ecRad. Using SPARTACUS instead of McICA increases the total cost of ecRad by a factor of 4.5, so it is a long way from being a candidate for an operational configuration, but it provides the capability to understand better the source of the radiative biases in the IFS. It works by solving the two-stream equations in the clear and cloudy parts of a gridbox including terms that represent the flow of radiation through cloud sides. These terms are proportional to the length of the perimeter of the cloud. Analysis of observations and cloud-resolving model output by Schäfer (2016) suggested that the cloud perimeter length could be parameterized in terms of cloud fraction and an *effective cloud scale*, and that as a first approximation a value of 1 km can be assumed for boundary-layer clouds and 10 km for all other clouds.

The final entry in Fig. 15 shows the impact on the climate of the coupled IFS model of introducing 3D radiative effects by switching the solver from McICA to SPARTACUS. We see that both net shortwave and net longwave at the surface are increased by 2 W m^{-2}, which leads to an increase in the mean 2-m temperature over land of around 0.6 K. This more than counteracts the 0.2 K cooling associated with more realistic cloud overlap and heterogeneity, and is a candidate mechanism to explain the surface cold bias in operations (Fig. 4). The bottom-right panel of Fig. 16 shows that introducing 3D effects along with more realistic cloud overlap and heterogeneity leads to a warming of up to 0.4 K in the tropical upper troposphere.

Work is ongoing to validate SPARTACUS in a wider range of cloud types than the cumulus clouds considered by Hogan et al. (2016), and to improve the appropriate values of effective cloud size to use in different parts of the globe.

6 Challenge 3: Clear-sky tropospheric shortwave absorption

While errors in the representation of the interaction of clouds and radiation are the source of the largest instantaneous errors in radiative heating rates in the troposphere, errors in the interaction with aerosols and gases can cause much smaller clear-sky biases that are nonetheless important because they act systematically over a large area. Aerosols and gases have traditionally been regarded as important primarily for climate projections, but they can also impact the predictability of weather systems through a range of mechanisms: changes to large-scale horizontal gradients of shortwave atmospheric heating lead to changes to mean wind through thermal-wind balance, changes to the vertical heating profile affects tropospheric stability and therefore convection, and atmospheric heating can act as a Rossby-wave source that impacts the global circulation (Rodwell and Jung, 2008). We consider the impact of improvements to aerosols in section 6.1 and to the shortwave water vapour continuum in section 6.2.

As shown in Table 1, the ozone climatology used in medium-range HRES and ENS configurations has been upgraded several times to make use of the ozone analyses from GEMS, MACC and CAMS. ECMWF (2007) anticipated the introduction of prognostic ozone in the radiation scheme, and indeed the linear Cariolle ozone scheme was included in the SEAS4 seasonal prediction system leading to improved

predictions of the Quasi-Biennial Oscillation. Similar results were found with the upgraded BMS ozone scheme (Monge-Sanz et al., 2011). However, the impact on tropospheric variables has been found to be slightly negative when tested in the SEAS5, monthly and medium-range systems. One factor is that linear ozone schemes are not suitable for the troposphere and indeed long integrations of linear schemes can lead to underestimates of tropospheric ozone by up to a factor of ten. The other problem is the correct initialization of stratospheric ozone: ozone initial conditions from ERA-Interim have been shown to produce worse results than climatological initial conditions. Work is ongoing to resolve these problems, but in the meantime SEAS5 and all the other operational forecast configurations use the same CAMS climatology.

6.1 Aerosols

Aerosols represent the main source of uncertainty in both clear-sky tropospheric shortwave absorption, and in the magnitude of surface solar heating in clear-sky conditions. Rodwell and Jung (2008) demonstrated the positive impacts on forecasts when the Tanré et al. (1984) climatology was replaced by Tegen et al. (1997) in the IFS. There is increasing interest in the impact of direct aerosol effects on tropical monsoon systems (e.g., Bollasina et al., 2011). The question that invariably arises is whether a *prognostic* aerosol scheme is needed, as anticipated in ECMWF (2007), or whether a climatology is sufficient for the needs of NWP. The tests performed by Morcrette et al. (2011) and Mulcahy et al. (2014) suggested that while prognostic aerosols could reduce radiation and temperature biases on a regional scale, the impact on forecast skill downstream was limited. In the case of the IFS, the computational cost of the 11 prognostic aerosol species in the CAMS configuration of the model would be prohibitive for use in NWP, which is why recent work has focused on improving the climatology. One possibility to overcome this would be to use a coarser grid for the prognostic aerosols. Another would be to use prognostic variables only for selected aerosol species and a climatology for the rest, where the likely prognostic aerosol types would be dust (e.g., Woodward, 2001) and biomass burning aerosol to capture the effect of fires. Satellite observations of fires are already used to estimate aerosol emissions to initialise CAMS forecasts (e.g., Flemming et al., 2017). However in the absence of a dynamical model to predict fire evolution, emissions from fires are kept constant during the forecast integration. In the future it may be possible to build on the existing scheme for forecasting fire danger developed for the Copernicus Emergency Management Service (Di Giuseppe et al., 2016) and develop a dynamical fire model with a stochastic fire ignition scheme that could trigger new fire events and/or extinguish ongoing fires during the forecast. Such a scheme would be expected to have a larger impact at monthly to seasonal timescales.

Cycle 43R3 saw the introduction of a new aerosol climatology based on the CAMS reanalysis (Bozzo et al., 2017). It provides a good example of how aerosols can affect tropical weather systems, and how improvements could be made in future. The new climatology consists of monthly-mean concentrations of five main aerosol species divided into 11 types, and a significant improvement over the Tegen climatology is that the radiative properties have been computed rigorously with the latest refractive indices in each band of the entire longwave and shortwave spectrum. The effect on an instantaneous visibility forecast is shown in the top two panels of Fig. 18. A further improvement is that the humidification of the hydrophilic aerosol species is now represented via a dependence of the optical properties on relative humidity. The strong influence of humidity on visibility can be seen by comparing the bottom two panels of Fig. 18.

Global aerosol distributions in models are invariably compared to each other and to observations in terms of aerosol visible optical depth, and indeed the CAMS aerosol analysis system assimilates optical depth
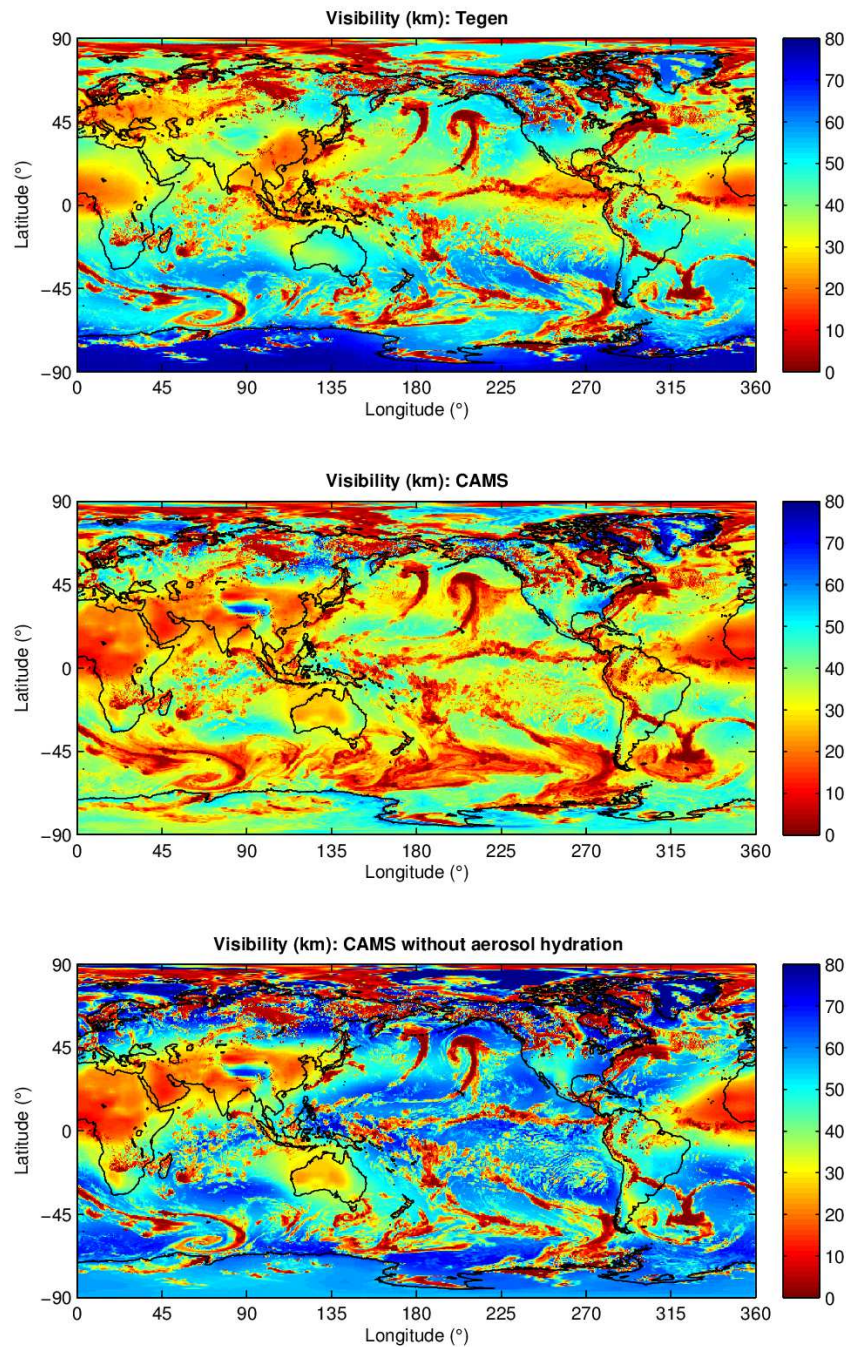
Figure 18: Comparison of instantaneous high-resolution forecasts of surface visibility computed (top) using the original Tegen climatology, (middle) with the new CAMS climatology, and (bottom) with the new CAMS climatology but setting the humidity to 0%.

retrievals from MODIS. While the optical depth determines how much solar radiation reaches the surface, the impact of aerosol on the atmosphere is often best understood by considering the change to the *absorption* optical depth, since that is what determines the rate of atmospheric heating by shortwave radiation. Unfortunately, this quantity is far less constrained by observations, and in the case of the CAMS climatology we are reliant on the optical properties assumed for the aerosol types used in the CAMS
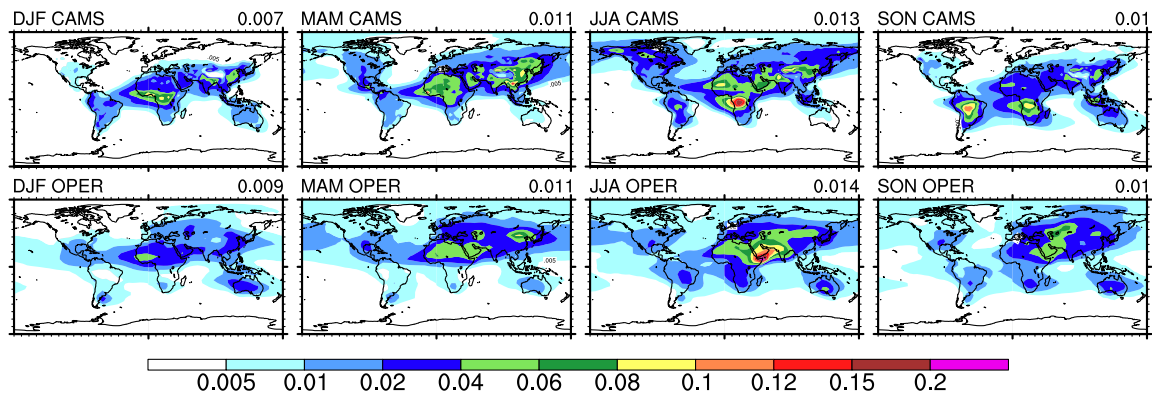
Figure 19: Comparison of aerosol absorption optical depth from (top row) the CAMS climatology introduced in Cycle 43R3 (Bozzo et al., 2017), and (bottom row) the Tegen et al. (1997) climatology operational for the previous 14 years.

system. Figure 19 compares the absorption optical depth between the old and the new climatologies. The most significant change is the reduction of absorbing aerosol over Arabia in JJA, and the increase in absorbing aerosol associated with biomass burning in central Africa.

The change over Arabia has a direct impact on the Indian Summer Monsoon flow (Vinoj et al., 2014), which Fig. 20c shows is too strong with the Tegen aerosol. By reducing the solar heating of the troposphere over Arabia, the strength of the Arabian heat low is reduced (compare Figs. 20a and 20b). This in turn reduces the latitudinal pressure gradient at low levels between Arabia and the Indian Ocean, and hence the strength of the westerly flow into India, as shown in Fig. 20d. Bozzo et al. (2017) reported that this halved the overestimate of West Indian rainfall.

The increase in absorbing aerosol over central Africa and into the Gulf of Guinea is also an interesting story: during testing of Cycle 43R3 it was found that this caused a very localized warm bias of 0.5–1 K at 850 hPa in JJA, which led to a conspicuous degradation of the root-mean-squared temperature error at this location when evaluated against analyses. The problem was solved by manually reducing the concentration of absorbing aerosol species in this region of the climatology by an annually varying amount that peaked in summer at 65%. This change was subsequently found to bring the CAMS climatology into closer agreement with both sparse AERONET sun-photometer observations and other aerosol models in the 'AeroCom' multi-model average. The reason for the overestimate is believed to be too strong a source of biomass burning aerosol in southern Africa in the CAMS system. This exercise highlights the strong sensitivity of the analysis system to radiative perturbations, something that could be exploited in future to help identify other systematic errors in the aerosol distribution and the underlying CAMS system.

6.2 Water vapour continuum

Water vapour is the strongest absorber of solar radiation in the troposphere, and much of the shortwave heating that occurs in the lower troposphere is due to absorption by the water vapour continuum, which accounts for the absorption between spectral lines. Despite its importance, the low optical depths of these spectral regions make it very difficult to characterize accurately in the laboratory. The continuum model used in RRTM-G and many other gas-optics schemes is 'MT_CKD' version 2.5 (Mlawer et al., 2012). However, various more recent laboratory measurements, notably from the 'CAVIAR' project,
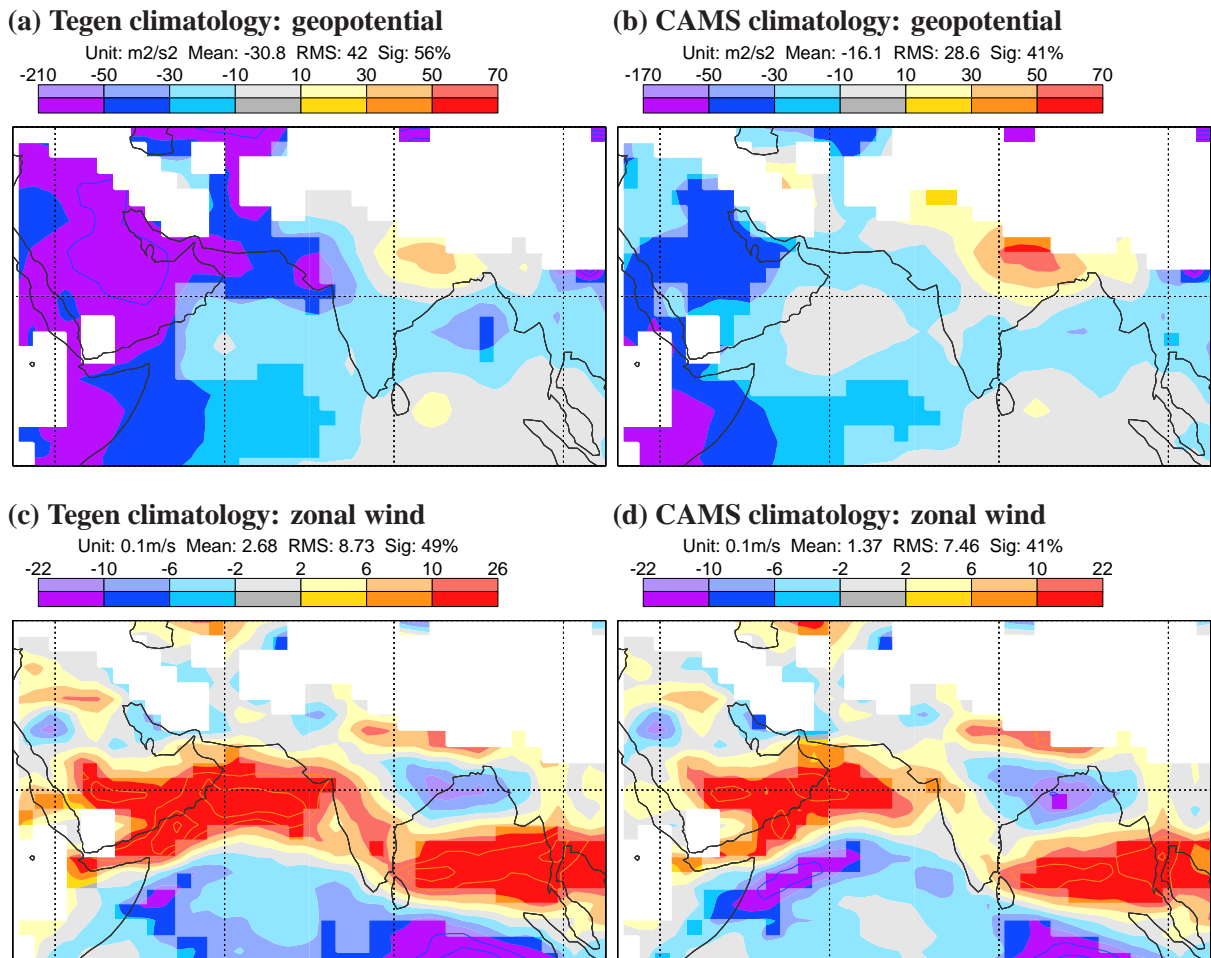
Figure 20: Geopotential and zonal-wind bias at 850-hPa for the period 1 May to 21 August 2016 over the northern Indian Ocean for forecast day 5: (left panels) operational model using Tegen aerosol climatology, and (right panels) after implementing the CAMS aerosol climatology. Bold colours indicate areas significant at the 5% level using a paired T-test with AR(1) noise. Further details available in Bozzo et al. (2017).

have estimated continuum absorption at least a factor of 10 greater than predicted by MT_CKD (see the review by Shine et al., 2016). Rädel et al. (2015) performed simulations with the Met Office climate model to show that replacing the CKD model (the predecessor to MT_CKD) with CAVIAR increased the all-sky solar absorption by 1.5 W m^{-2} in the tropics, but less in the drier extra-tropics.

To estimate the effect of this enhanced water vapour continuum in the IFS, calculations of the CAVIAR self- and foreign-continuum absorption spectra (kindly provided by Igor Ptashnik) have been used to scale the coefficients in each spectral interval of the MT_CKD model in the IFS, but retaining the same temperature dependence. The impact on coupled climate simulations is shown in Fig. 21. Panel d shows the increase in zonal-mean atmospheric shortwave absorption of up to 3 W m^{-2}, larger than found by Rädel et al. (2015) presumably due to their reference being CKD rather than MT_CKD. Panels a–c depict the widespread increase in tropospheric temperature, which in the tropics peaks in the upper troposphere at 0.3–0.4 K (similar to the combined impact of cloud changes depicted in the bottom-right panel of Fig. 16). Seasonal warmings in excess of 1 K are observed at the summer poles, due to the perpetual daylight and the shallow solar incidence angles. At the surface, net shortwave is reduced by about the same amount as the increase in atmospheric absorption, but the warmed troposphere emits more longwave
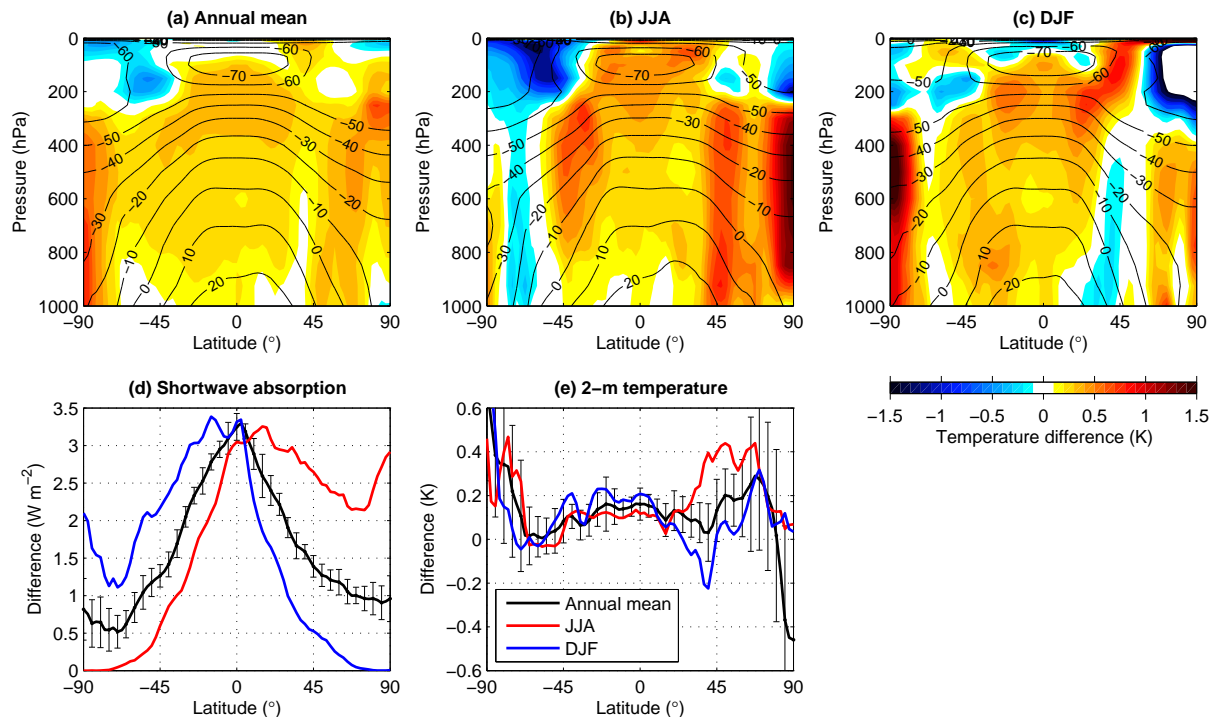
Figure 21: Impact of changing the shortwave water vapour continuum model from MT-CKD to CAVIAR on (a–c) tropospheric temperature, (d) shortwave atmospheric absorption and (e) 2-m temperature. The calculations are from eight 1-year coupled T_L255 climate simulations, and the error bars in the lower panels show the 95% confidence interval in the annual-mean change.

downward, leading to the increase in 2-m temperature shown in Fig. 21e (the global-mean increase is 0.2 K).

To investigate the impact on medium-range forecast skill, 4-month (Nov 2015 to Feb 2016) analysis experiments have been performed. Figure 22a depicts the average temperature bias of the control experiment at forecast day 8, showing the ubiquitous tropical cooling of around 0.5 K, but in this particular year a summer-pole warm bias of around 0.7 K. The 0.1–0.2 K warming of the troposphere (Fig. 22b) improves the mean and therefore leads to a statistically significant 4% reduction in tropical root-mean-squared temperature errors (Fig. 22c), with a much more mixed picture in the mid-latitudes. Figure 22d depicts the change in forecast skill as measured by the geopotential height anomaly correlation. There is a 4% improvement throughout the troposphere at 30°S, matched by a reduction in the vector wind error by around the same magnitude (not shown). One hypothesis to explain this is that the tropical warming reduces the error in the meridional temperature gradient, which in turn improves the winds at this latitude. The picture elsewhere is more mixed, with a degradation in the mid- and high-latitudes of the winter hemisphere, although this signal has not reached statistical significance.

An additional analysis experiment has been performed for a shorter period of Northern Hemisphere summer (11 May to 30 June 2016), and to a large extent the same pattern is seen in all variables shown in Fig. 22, except with the sign of latitude inverted. The water vapour continuum example illustrates well the point made in the introduction that the very good tropospheric climate of the IFS makes it difficult to show improvements in forecast skill from improvements in the physics that have a small but systematic impact on the large-scale temperature.
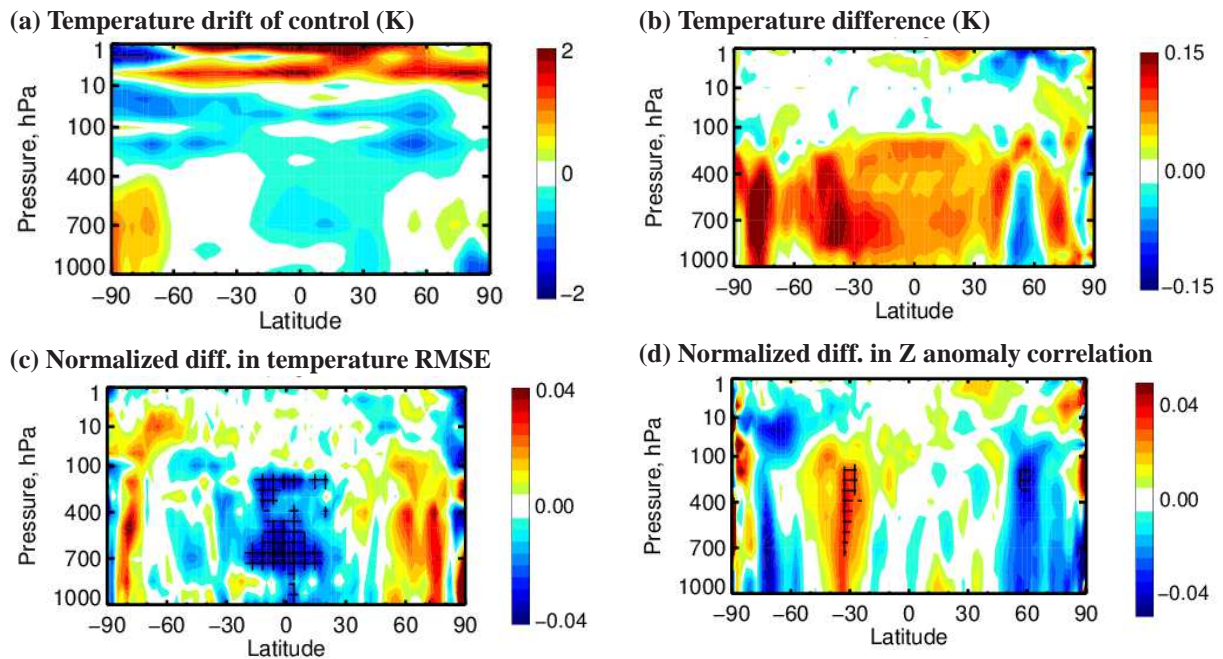
(a) Temperature drift of control (K)

(b) Temperature difference (K)

(c) Normalized diff. in temperature RMSE

(d) Normalized diff. in Z anomaly correlation

Figure 22: Impact of CAVIAR water vapour continuum on 8-day forecast skill as measured using $T_{Co}399$ analysis experiments from 2 November 2015 to 28 February 2016, with the experiment and the control each evaluated against its own analysis: (a) mean temperature drift of control forecast after 8 days, (b) CAVIAR temperature drift minus control, (c) normalized difference in root-mean-squared temperature error (negative indicates an improvement), and (d) normalized difference in geopotential height anomaly correlation (positive indicates an improvement). The hatched areas in the lower panels indicate regions where the differences are statistically significant.

7 Challenge 4: Middle atmosphere

The stratosphere offers a potential source of predictability on monthly and seasonal timescales, but has been historically rather neglected at ECMWF. Since November 2016 it has received much more attention via the activities of the Stratosphere Task Force[5] (e.g., Polichtchouk et al., 2017), which has led to some of the advances discussed in this section.

One proposed source of tropospheric predictability is from the Quasi-Biennial Oscillation (QBO, e.g. Marshall and Scaife, 2009), an oscillation of the equatorial stratospheric zonal wind with a period of approximately 28 months. After the annual and diurnal cycles, the QBO is the most predictable feature of the atmospheric circulation, and its influence on northern hemisphere winter weather was argued by Thompson et al. (2002) to be as strong as that of ENSO. It is driven by breaking gravity waves, which have some sensitivity to the radiation scheme via its role in setting the stability of the stratosphere. Section 7.1 shows that the stratosphere is too statically stable due to a large warm bias that increases with height from the mid-stratosphere to the mesosphere. A combination of improvements that reduce the shortwave heating rate are then shown to be able to almost eliminate this bias.

Another route to tropospheric predictability is through the downward influence of Sudden Stratospheric Warmings (SSWs), which shows up clearly when reanalysis data are composited around SSW events (Baldwin and Dunkerton, 2001). The evolution of SSWs in ECMWF monthly forecasts are among the

---

[5]The Stratosphere Task Force is a cross-departmental project at ECMWF running until the end of 2017, with the aim of diagnosing and improving all aspects of the stratosphere in IFS analyses and forecasts. It involves close collaboration with Ted Shepherd and Inna Polichtchouk of the University of Reading.
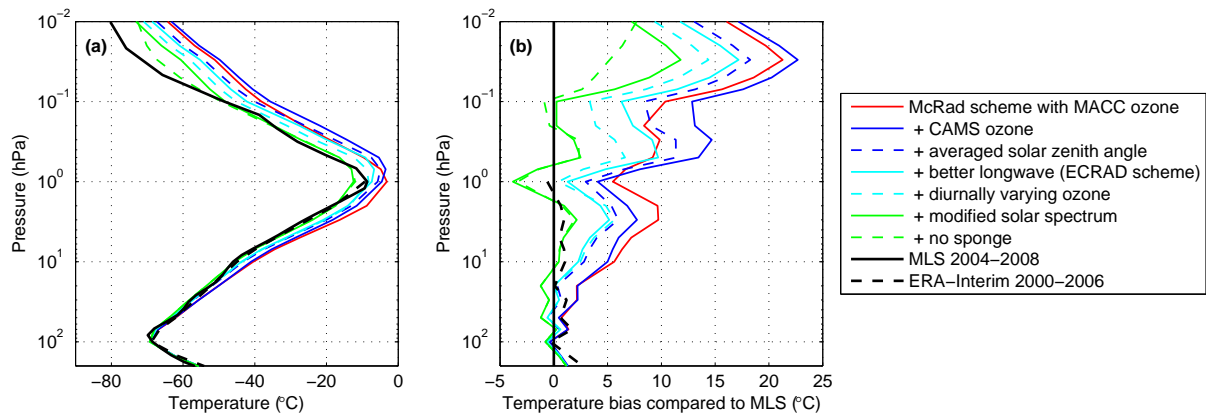
Figure 23: Annual-mean temperature from four 1-year uncoupled T_L255 137-level climate simulations of IFS Cycle 43R1 (coloured lines), the Microwave Limb Sounder (MLS; black solid line) and ERA-Interim (black dashed line). ERA-Interim has a large warm bias in the mesosphere and so is not shown above 1 hPa. The red line is approximately equivalent to Cycle 41R1 and the light blue solid line to Cycle 43R3.

best compared to other centres, but the correlation with the troposphere is significantly weaker than found in reanalyses (Vitart et al., 2014). A contributing factor could be the large polar lower-stratospheric cold bias seen in Fig. 3a; biases in the meridional temperature gradient lead to errors in the position and intensity of the subtropical upper tropospheric jet, which in turn can affect the coupling between the stratospheric polar vortex and the troposphere (e.g., Garfinkel et al., 2013). In section 7.2 we show that artificially removing the lower-stratosphere temperature bias results in an improvement to monthly forecast skill.

7.1 Upper stratosphere and mesosphere warm bias

Figure 23 compares the annual-mean temperature from free-running uncoupled simulations of the IFS with the Microwave Limb Sounder (MLS) instrument. The red line shows the climate of the model when configured to be similar to Cycle 41R1 using a 3-h radiation timestep, which matches all operational configurations except HRES. The top row of Fig. 24 shows the latitudinal structure of the errors and its dependence on season. The annual-mean warm bias is severe: up to 10 K too warm in the upper stratosphere and 20 K in the mesosphere, with larger biases in individual seasons. We next consider the sequential impact of five improvements to the physics and chemistry of the model that can reduce the warm bias, shown by the coloured lines in Fig. 23:

1. Replacing the ozone climatology derived from MACC with one derived from CAMS that is in better agreement with observed concentrations (Flemming et al., 2017) leads to a cooling of the stratosphere and a warming of the mesosphere simply because of a reduction of stratospheric ozone and an increase in mesospheric ozone. This change became operational in Cycle 43R1.

2. Hogan and Hirahara (2016) found that the 3-h radiation timestep was responsible for a warm bias of up to around 3 K, but that careful averaging of the solar zenith angle used in the radiation scheme could reduce this to 0–1 K when compared to simulations with radiation called every model timestep. This change became operational in Cycle 41R2. Hereafter, the coloured lines shown in Fig. 23 use a 1-h radiation timestep.

3. The ecRad radiation scheme introduced in Cycle 43R3 uses an exact rather than approximate
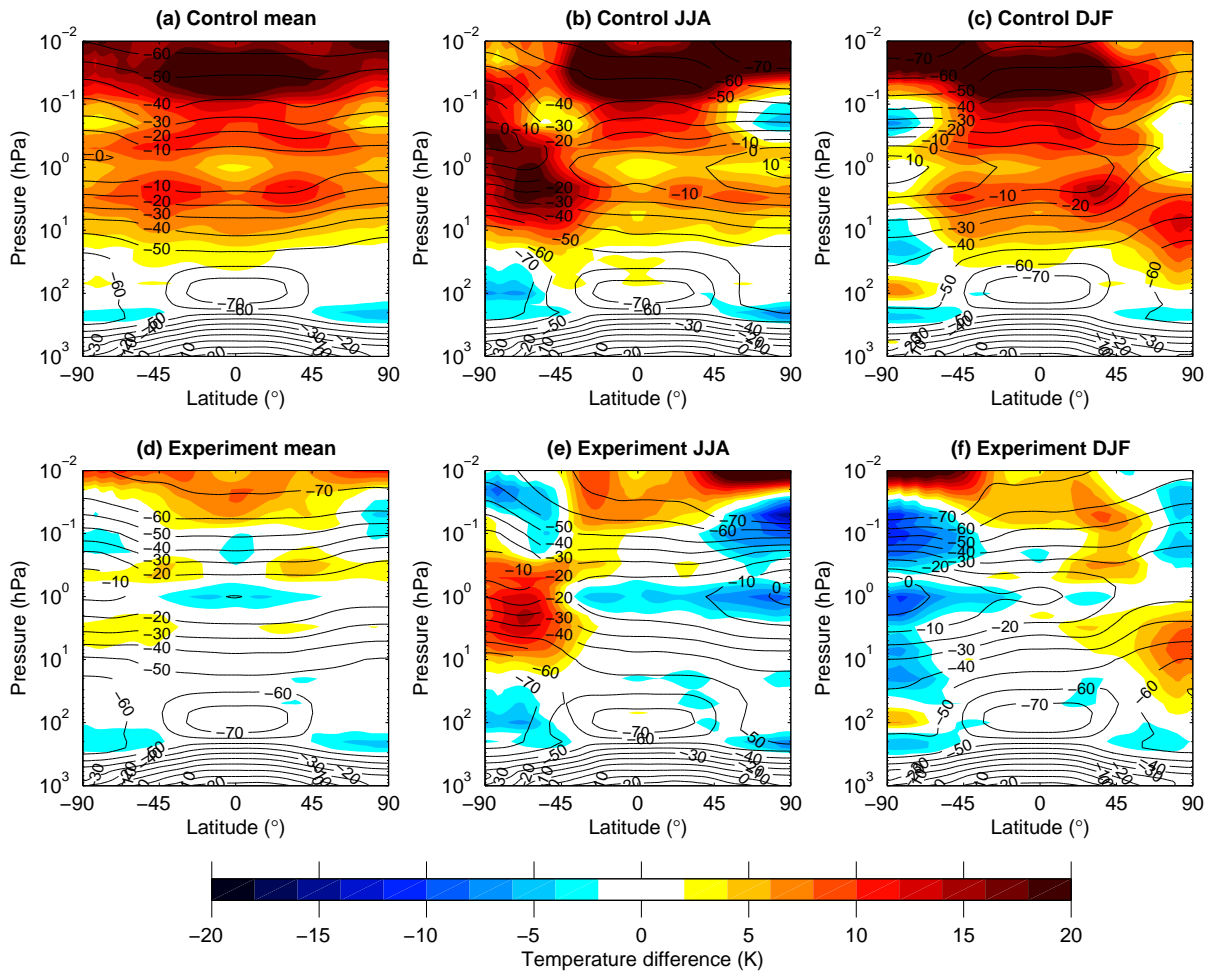
Figure 24: Mean temperature from the first and last IFS simulations shown in Fig. 23: (top row) McRad scheme with MACC ozone, and (bottom row) after multiple changes described in section 7. The black contours show temperature and the colours show the difference against a reference dataset consisting of the MLS climatology at pressures of 100 hPa and less, and ERA-Interim at pressures greater than 100 hPa. The left column shows the annual mean, the middle column the northern-hemisphere summer and the right column the northern-hemisphere winter.

solution to the longwave two-stream equations, which has the effect of reducing extrema in the temperature profile and specifically warming the tropopause and cooling the stratopause (Hogan and Bozzo, 2016).

4. An additional factor that could be considered in future is that ozone has a diurnal cycle, and the radiative impact of ozone is dominated by the shortwave for which the daytime values are important. To investigate the magnitude of this effect, an approximate diurnal cycle has been introduced by approximately fitting the results of Studer et al. (2014): at pressures lower than 0.4 hPa the daytime value is 25% lower than the midnight value, increasing to 3% more at 5 hPa and then dropping back to have no diurnal cycle at pressures greater than 20 hPa. It can be seen to have a small effect in the stratosphere but to cool the mesosphere by up to 4 K.

5. The version of RRTM-G used at ECMWF assumes the Kurucz (1994) solar spectrum. More recent measurements of the sun, particularly from the SORCE instrument, suggest a 7–8% lower ultravi-
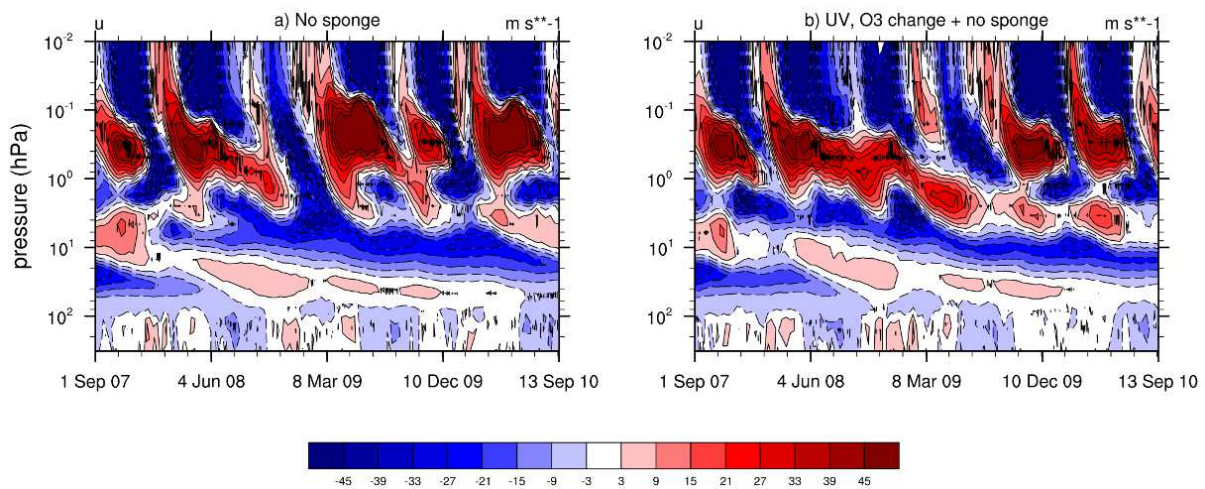
Figure 25: Zonal wind above the equator in a single 4-year uncoupled T_L255 simulation, for (a) Cycle 43R1 with ecRad but no sponge, and (b) as panel a but with the ultraviolet reduction and approximate diurnal cycle in ozone discussed in the text.

olet output for the same total solar irradiance. Fig. 23 shows that replacing the solar irradiance in each of the RRTM-G bands with the recent climate data record of Coddington et al. (2016) leads to a significant temperature reduction throughout the middle atmosphere, similar to the findings of Zhong et al. (2008) for the Met Office model.

The solid green line in Fig. 23 shows that the combination of these effects largely removes the temperature bias at pressures down to 0.1 hPa. The first three are already operational in Cycle 43R3 while the last two (or refined versions of them) will be considered for future cycles. This still leaves a warm bias of up to 10 K in the upper mesosphere. There is a strong 'sponge' in the model above the 0.1 hPa level that explicitly suppresses upwardly propagating waves so that they are not reflected by the model top at 0.01 hPa. This sponge actually has a warming effect on the upper mesosphere, as shown by the dashed green line in Fig. 23 in which it has been turned off. Work is ongoing to design a sponge that adequately suppresses waves without affecting temperature so much.

The latitudinal and seasonal structure of this final model configuration is shown by the bottom row of Fig. 24. This is clearly a large improvement in the annual mean. Figure 25 shows that this cooling has a small effect on the semi-annual oscillation in the mesosphere and the quasi-biennial oscillation in the stratosphere, which implies that the impact on predictive skill in the troposphere is likely to be weak. Further experiments exploring the impacts of various model changes on the dynamics of the middle atmosphere are described by Polichtchouk et al. (2017).

Figures 23b and 24d indicate that even without the sponge, there is still a warm bias of 7 K at the model top. This error may have a radiative component, for example due to the neglect of non-local-thermodynamic-equilibrium (non-LTE) effects. There is also still a warm bias of up to 15 K in the winter polar upper stratosphere (particularly in the southern hemisphere), presumably due to the biases in polar night being unaffected by the five changes above that reduce solar heating. This is as likely to be due to an overactive Brewer-Dobson circulation as to a radiative error. One area of future work that could bring additional insight to radiatively-driven temperature errors would be to evaluate the solar tides in the IFS, which observations show to have a diurnal amplitude of up to 3 K at 1 hPa and up to 10 K at 0.01 hPa (Huang et al., 2010).
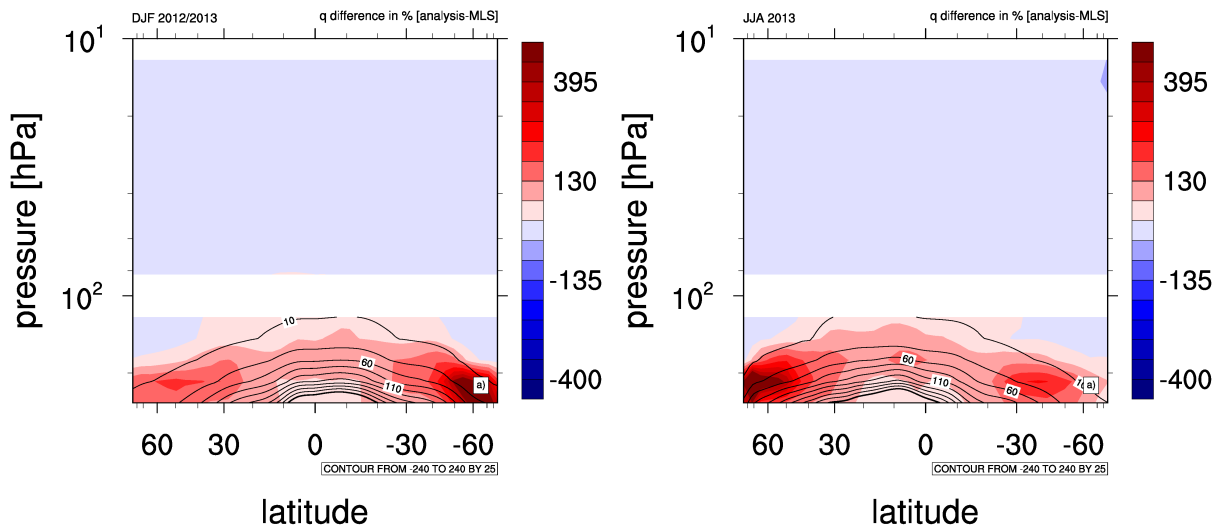
Figure 26: Percentage error of mean specific humidity in operational analyses compared to the MLS instrument for (a) December 2012 to February 2013 and (b) June to August 2013.

If the middle-atmosphere biases can be significantly reduced then the door would be opened to consider whether further benefit could be derived from higher vertical resolution or a raised model top, although it should be stressed that such a change would be driven not by the needs of radiation but by the need to better represent the dynamical evolution of phenomena such as the QBO. Raising the model top would require the radiation scheme to be upgraded to represent non-LTE effects.

7.2   Polar lower stratosphere cold bias

Figure 3a shows the annual-mean polar cold bias of around 5 K at 200 hPa. This is a very common problem in global models (Gates et al., 1999), and has been present in the IFS for at least the last 25 years. It is barely affected by horizontal resolution, and it can be seen in Fig. 24 that it is completely untouched by the improvements described in section 7.1. Its main cause is believed to be excessive transport of water vapour from the troposphere to the polar lower stratosphere (Stenke et al., 2008), presumably due to numerical diffusion from the advection scheme. The water vapour then emits too strongly in the thermal infrared. In the case of the IFS, comparison of analyses and short forecasts with both aircraft in-situ sampling (Dyroff et al., 2015) and MLS observations (Fig. 26) reveals that water vapour is indeed overestimated in the extra-tropics between 150 and 250 hPa (the latter being the lower limit of MLS) with the largest overestimate being around a factor of four over the summer pole. It appears that the analysis water-vapour biases are just as large as the forecast biases.

While it is outside the scope of this paper to solve the advection problem, we can ask what the impact would be if it were solved by artificially modifying the humidity seen by the radiation scheme. Experiments have been performed in which the humidity seen by radiation has been reduced by up to a factor of three using the pattern shown in Fig. 27, which almost completely removes the polar lower stratosphere cold bias. The impact on monthly forecast skill is shown in Fig. 28, and indicates a positive impact over Europe in all tropospheric variables, although only the signal in 500-hPa temperature has reached statistical significance. There is a need to understand the mechanisms for this improvement, but at this stage it appears not to be limited to better predictions of either SSWs (since improvements are also seen in summer) or the North Atlantic Oscillation (since the NAO index itself is not significantly improved).
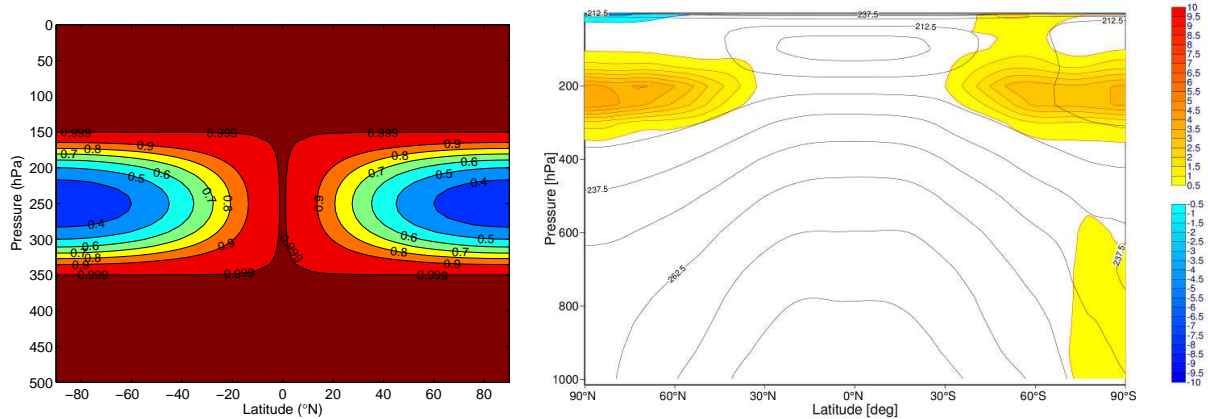
Figure 27: Impact of artificially reducing humidity seen by radiation scheme: (left) multiplier applied to specific humidity seen by radiation, and (right) change to annual-mean temperature in four 1-year T_L255 simulations (compare to the model bias in Fig. 3a).

It should also be noted that the temperature bias in the zonal mean hides larger transient biases associated with individual cyclones (not shown); it is possible that more predictive skill will emerge when the source of the humidity bias is fixed leading to more accurate simulation of the zonally asymmetric component of the radiative heating and cooling.

8 Challenge 5: Efficiency

There is a strong impetus at ECMWF to optimize the major components of the IFS, and radiation has historically been regarded as one of the most expensive. However, in the HRES configuration of current IFS Cycle 43R3 ($T_{Co}1279$ resolution), the new ecRad radiation scheme is responsible for only around 5% of the computational cost of the forecast model. This figure is much less than the 19% reported by Morcrette et al. (2008b) for the McRad radiation scheme when it became operational a decade ago in Cycle 32R2 (T_L799 resolution). Part of this difference is because ecRad is faster than McRad, as described in section 8.1. But the majority is because the increase in horizontal resolution has been accompanied both by a reduction in model timestep while the radiation timestep has been kept constant at 1 h, and an increase in the ratio of model to radiation gridpoints from 6.25 to 10.24.

Going forward, the challenge is to configure the radiation scheme to provide the maximum accuracy and hence forecast skill within an approximately fixed computational envelope. In section 8.2 it is argued that the trade-off is currently skewed too much in favour of spectral resolution rather than temporal or spatial resolution, and indeed we show that shortening the ENS radiation timestep from its current value of 3 h results in a measurable increase in forecast skill. Section 2 described a decade of progress in radiation modelling, but because of the high spectral resolution of the underlying RRTM-G gas optics scheme, neither McRad nor ecRad are fast enough to use in the simplified physics of the data assimilation system; instead the tangent-linear and adjoint calculations are based on the old Morcrette radiation scheme. If the option became available for ecRad to use a much faster gas optics scheme then it would be possible to use it also in the data assimilation system, leading to greater consistency between the full and simplified physics schemes.
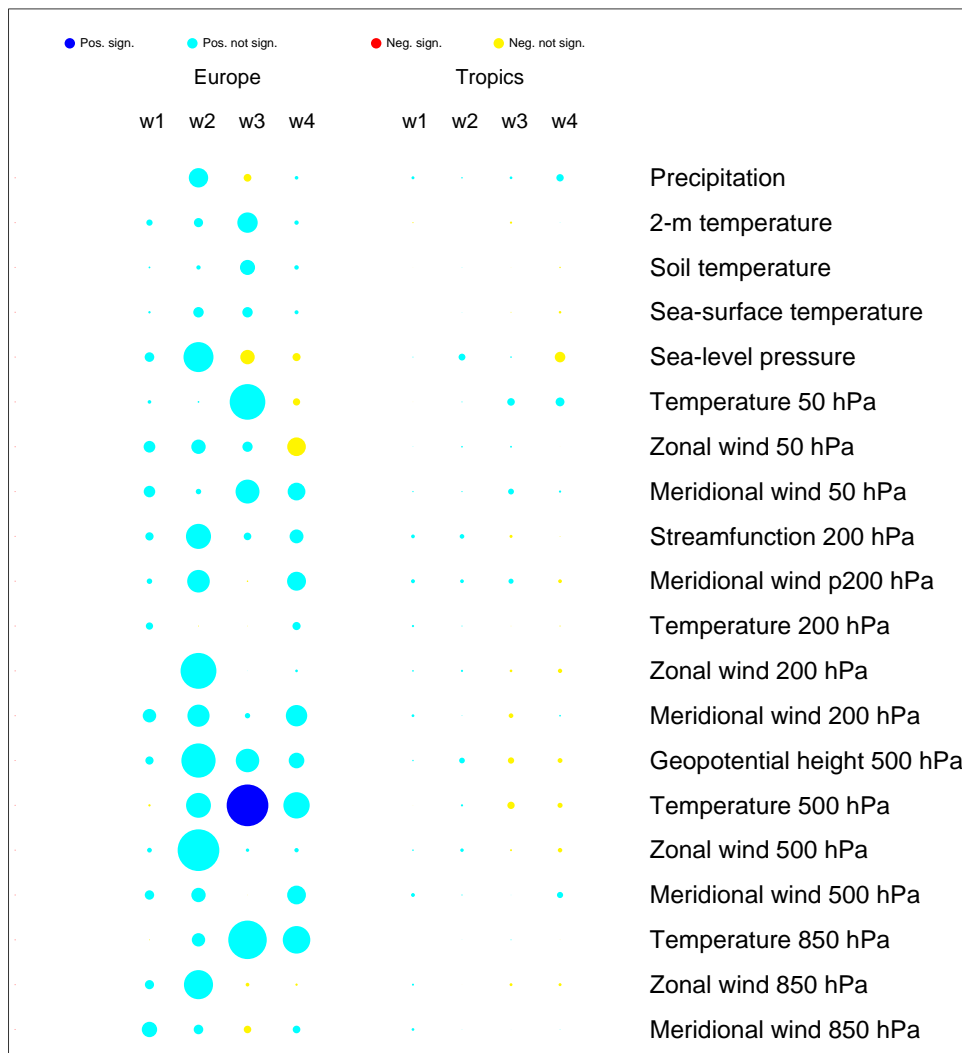
Figure 28: Scorecard indicating the change to monthly forecast skill over Europe and in the tropics (as measured by the Ranked Probability Skill Score) arising from artificially fixing the polar lower stratosphere cold bias by reducing humidity seen by radiation, between forecast weeks 1 (w1) and 4 (w4), for a range of model variables. The scores are computed from 26 years of re-forecasts for February, May, August and November start dates (a total of 104 cases).

8.1   Optimizations in ecRad

The rewriting of the radiation scheme presented the opportunity to introduce not only a much more modular structure, but also a number of significant optimizations. Figure 29 is taken from Hogan and Bozzo (2016) and compares the cost of the main parts McRad and ecRad. The main optimizations in ecRad are:

- Switching the fastest varying dimension to be spectral interval, rather than column (as in the rest of the model physics). This is because conditional operations (which inhibit vectorization) depend on the presence of cloud or whether the sun is above the horizon, and these are functions of column but not spectral interval. This change makes all parts of the scheme faster, except for gas optics since the original McRad code is still used for this part, and an additional permutation of arrays is now required at the end. In principle this could be much improved in a future implementation of
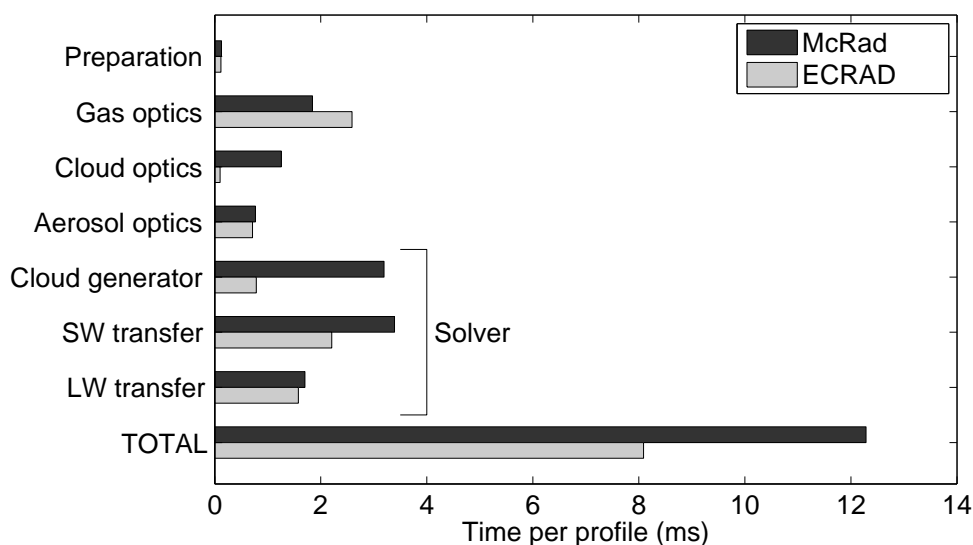
Figure 29: Comparison of the computational cost of various parts of the old 'McRad' and new 'ecRad' radiation schemes in a 1-day forecast of the IFS at $T_{Co}639$ resolution with 137 levels. All calculations performed after interpolation to the radiation grid are included. 'Preparation' consists of preparing the inputs to the radiation scheme. This is followed by the four components of the new scheme but with the 'Solver' component divided into the McICA cloud generator and then the radiative transfer calculations in the shortwave (SW) and longwave (LW). In both runs the model was configured to process atmospheric profiles in batches of eight.

the gas optics model.

- McRad recomputed the cloud optical properties using the perturbed water contents from the McICA cloud generator for each spectral interval (or g-point). A large speed-up is achieved by instead computing the optical properties for each band (there being 8–9 times fewer bands than g-points) and then simply scaling the resulting optical depths using the McICA cloud generator. This change also reduces the stack memory usage in OpenMP threads of the entire model by a factor of 2.5. Although this corresponds to a reduction of the total memory footprint of the IFS by only 3%, a reduction of memory per thread is attractive for emerging future HPC architectures.

- The stochastic cloud generator has been rewritten to use far fewer random numbers, and is around a factor of four faster (see Hogan and Bozzo, 2016, for details).

In this real-world benchmark, ecRad is around 34% faster than McRad, reducing to 31% when the computational cost of interpolating to the reduced radiation grid is included. This suggests that an immediate route to improving forecast skill for the same cost is to reinvest the saving by reducing the ENS radiation timestep from 3 to 2 h. Figures 30a and 30c quantify the increase in skill in ENS resulting from such a change. Unfortunately, when ecRad was introduced into operational Cycle 43R3, the cost saving was needed to increase the 'buffer' between the time the operational suite typically finished and the deadline for delivering forecast products to member states, so the ENS radiation timestep was kept constant at 3 h.

Another ongoing project will enable the radiation scheme (and potentially other components such as the wave model) to be run in parallel to the dynamics (Mozdzynski and Morcrette, 2014). The expected efficiency improvement of the whole model is expected to be of order 5%, although the forecast impact of the model seeing radiation fluxes lagged by 1–2 radiation timesteps needs to be assessed.
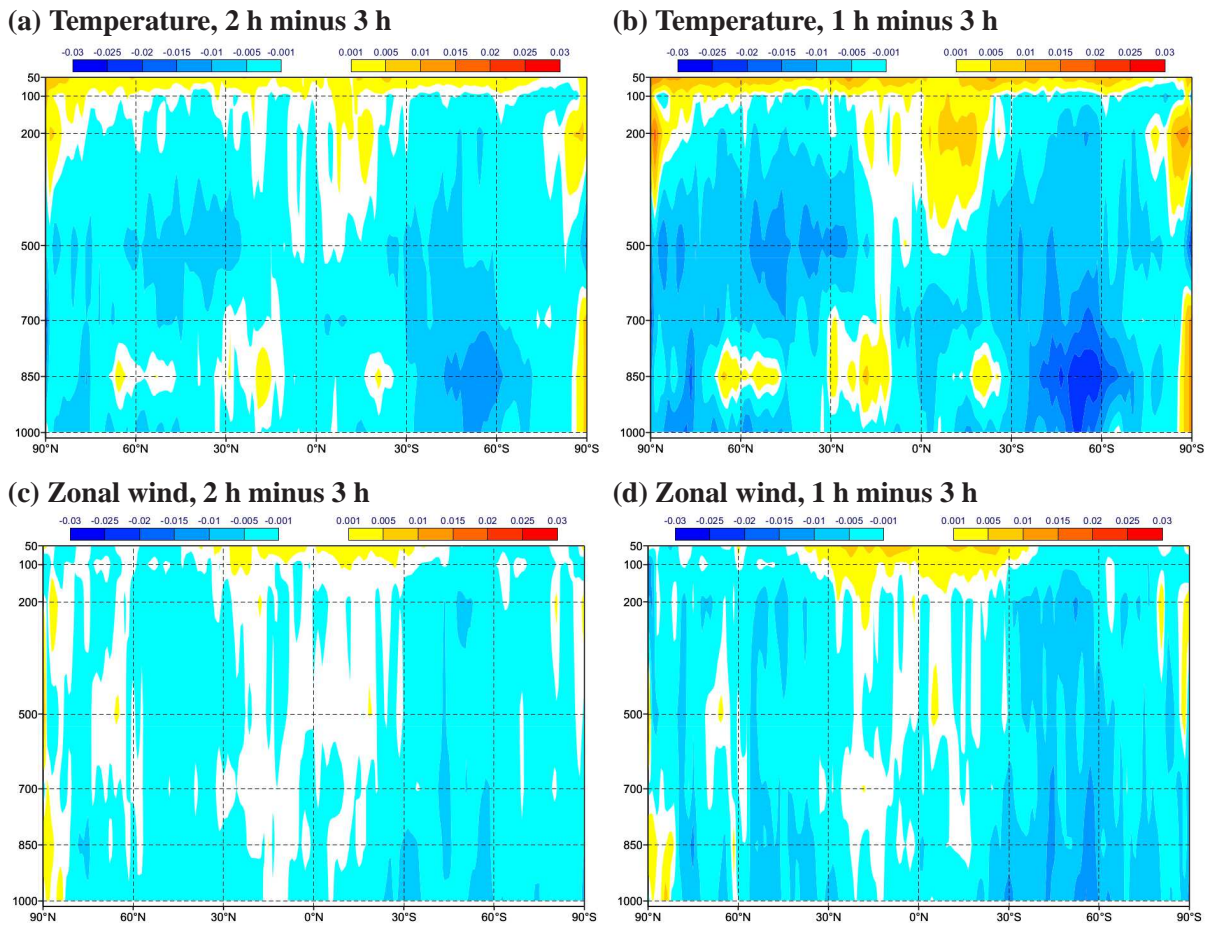
(a) Temperature, 2 h minus 3 h

(b) Temperature, 1 h minus 3 h

(c) Zonal wind, 2 h minus 3 h

(d) Zonal wind, 1 h minus 3 h

Figure 30: Change to the 48-h Continuous Ranked Probability Score (CRPS) in temperature (K) and zonal wind (m s^{-2}), versus latitude (°N) and pressure (hPa) resulting from reducing the radiation timestep from (left column) 3 h to 2 h and (right column) 3 h to 1 h. Negative values indicate an improvement of forecasting skill. To put these numbers in perspective, the root-mean-squared error in 48-h tropospheric temperature is around 0.75 K and in vector wind is around 3 m s^{-1}. Therefore, a 0.03 K reduction in temperature CPRS is an improvement of around 4% and a 0.015 m s^{-1} reduction in wind CPRS is an improvement of around 0.5%. The ensemble forecasting system consisted of 21 ensemble members at $T_{Co}399$ resolution with 91 levels and 45 start dates in 2015 using IFS Cycle 43R1 (with ecRad as the radiation scheme) evaluated against operational analyses. Note that the apparent increase in error in the stratosphere is believed to be due to a slight worsening of the temperature bias in the 91-level model, and is not seen when evaluating the impact of radiation timestep in 137-level deterministic forecasts.

8.2 Balancing spectral accuracy with other aspects of the radiation scheme

The model responds to broadband fluxes, which the radiation scheme computes by integrating over four dimensions: *time, space, angle* and *frequency*. A trade-off therefore has to be made in how finely each of these dimensions are discretized to maximize accuracy. Table 2 compares the temporal, spatial and spectral resolution of four global NWP models, and a striking difference is evident with ECMWF favouring spectral over spatial resolution compared to the other centres.

Considering first *time,* since 2004 the radiation scheme is called every 1 h in HRES and every 3 h in all other model configurations. Calling the radiation scheme every timestep gives only a marginal improvement over calling it every 1 h, but every 3 h is significantly worse, which is hardly surprising given that the path of the sun through the sky is then represented by typically only four discrete angles. Im-
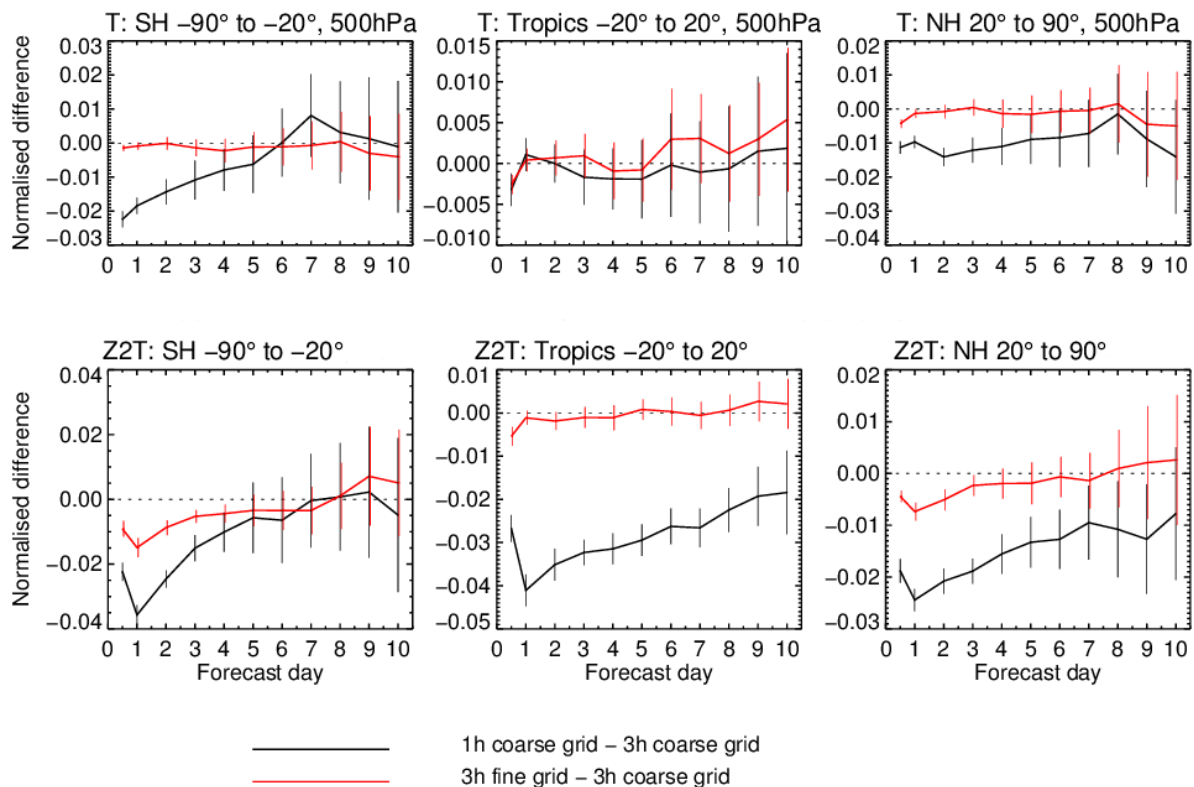
Figure 31: Comparison of the effect of improving the spatial versus temporal resolution of the radiation calls, using daily 43R3 forecasts at $T_{Co}399$ resolution from 1 June to 30 September 2016. The quantity shown is the fractional change to root-mean-squared error, so -0.01 indicates a 1% improvement. The top row shows 500-hPa temperature errors and the bottom row 2-m temperature errors. The control run calls radiation every 3 h and on a grid with 6.25 times fewer gridpoints. The black line shows the effect of reducing the radiation timestep to 1 h and the red line the effect of calling it every gridpoint.

provements were made in Cycle 41R2 to mitigate some of these problems with negligible computational cost: the Hogan and Bozzo (2015) approximate update scheme adjusts the longwave fluxes to respond to evolving surface temperature, the Manners et al. (2009) scheme corrects surface solar fluxes for the changing path length of the sun through the atmosphere, while the Hogan and Hirahara (2016) scheme averages solar zenith angle in such a way as to remove the 3 K stratosphere temperature bias due to infrequent radiation calls. Nonetheless, Hogan and Bozzo (2015) reported that only half of the forecast degradation associated with 3-h radiation was mitigated by these changes; the remainder was associated with the interaction of the changing cloud fields with radiation that could only be captured by calling the scheme more frequently. This is quantified in the right panels of Fig. 30. These results suggest that it would be highly desirable to increase the radiation frequency to once per hour in all operational model configurations.

Consider next *space;* the radiation scheme is run on a coarser grid by interpolating the model fields using a 12-point stencil (Morcrette et al., 2008b). The resulting fluxes are then interpolated back onto the model grid. The degree of coarsening can be quite severe; the $T_{Co}1279$ HRES model runs radiation on a grid equivalent to that of a $T_{Co}399$ model, which has 10.24 times fewer points. Until Cycle 41R1, this severely degraded the 2-m temperature errors at some coastal points, as described in section 4.1, but the problem was solved in Cycle 41R2 by the Hogan and Bozzo (2015) scheme. So is it any longer

as important to improve spatial resolution as temporal resolution? This is answered in Fig. 31, which considers improvements to a control model with the same coarsening as ENS: radiation called every 3 h on a grid with 6.25 times fewer points. The red lines show that if radiation were called at every gridpoint (an increase in computational cost of a factor of 6.25), random errors in temperature would be reduced by at best 1% but for only a day or two into the forecast. The black lines show that if radiation were called every hour (an increase in cost of a factor 3) then improvements of up to 3% can be expected that persist much longer into the forecast. This reveals that it would be far more useful to improve temporal rather than spatial resolution in ENS, and indeed there might be advantage to further spatial coarsening in order to be able to afford a 1-h radiation timestep.

Radiation can propagate at an arbitrary *angle,* so it is surprising that global models can get away with using the two-stream equations in which only two discrete angles are considered for the diffuse radiation field, plus the propagation of the direct beam in the case of solar radiation. More sophisticated radiation models can treat an arbitrary number of streams, N, but as a rule of thumb the computational cost of the solver is $O(N^3)$, so a four-stream solver would be eight times more expensive. Cheaper alternatives have been proposed, for example the Räisänen (2002) scheme of simply tweaking the two-stream coefficients to better predict shortwave fluxes when the sun is low in the sky. In the longwave, Fu et al. (1997) proposed an approximation to four streams that increasing the cost of the solver by 80% compared to a two-stream solver with scattering. But as discussed in section 5.2, the first step at ECMWF would be to turn on longwave scattering in a two-stream context, increasing the cost of the longwave solver by 33%, and the radiation scheme overall by 10%.

Finally we consider *frequency.* Like most state-of-the-art GCM radiation schemes, RRTM-G uses the correlated-k method for its spectral integration: within several broad *bands,* the complex gas absorption spectra are represented by a number of representative *spectral intervals* (also known as g-points) that are treated monochromatically by the rest of the scheme. The first row of Table 2 shows the total number of spectral intervals required by RRTM-G, which is also used in the global models of DWD, NCEP and (in the longwave only) Météo France. This scheme is very accurate in clear-sky conditions (e.g., Macfarlane et al., 2016), but it uses up to 3.7 times more spectral intervals than some of the other NWP centres listed in the table. Since all subsequent parts of the radiation code use the same spectral discretization, a reduction of the number of intervals scales the cost of the entire scheme. Part of the problem is that RRTM-G has such a large number of bands; Table 2 confirms the unsurprising fact that more bands leads to more spectral intervals. Hogan (2010) demonstrated that for clear-sky longwave calculations, the *full-spectrum correlated-k* method (FSCK) could be applied, and that neither the variation of the Planck function with frequency, nor the number of active gases, was a valid reason to introduce bands. With only a single band covering the entire longwave spectrum, heating-rate accuracies of better than 0.1 K d^{-1} should be achievable through most of the troposphere and stratosphere using only around 32 spectral intervals. Pawlak et al. (2004) showed that even fewer spectral intervals were required in the shortwave, although they proposed two bands to allow for the fact that cloud absorption is almost entirely limited to the near infrared. Further work is required to extend longwave FSCK to clouds. We will also investigate what can be learned from the fast single-band 'ACRANEB' radiation scheme (Mašek et al., 2015) used in the HARMONIE limited-area model.

An alternative approach to speeding up the four-dimensional integration discussed above is to employ some kind of sampling of one or more of the dimensions. Bozzo et al. (2014) tested a stochastic sampling technique in which only around 12% of the spectral intervals were computed in each call to the radiation scheme, speeding the scheme up to the extent that it could be called at every model gridpoint or more frequently in time. This was found to solve the coastline problem discussed in section 4. The downside was that it led to significant random errors in diagnosed surface fluxes, which would be a problem for users

such as the solar energy industry. In practice, the approximate update technique of Hogan and Bozzo (2015) solved the coastline problem without generating noisy fluxes, so that scheme was implemented into the operational model instead. Other sampling ideas that could be considered in future are to select the more absorbing spectral intervals that are only important in the middle atmosphere and either treat them with a cheaper solver (Li and Barker, 2005) or call them less frequently in time (Manners et al., 2009).

9 Conclusions

Radiation is a key process in the ECMWF model, and improvements in the representation of radiative transfer are needed to improve forecasts at all timescales, from short-range temperature forecasts to the predictability of regime shifts in monthly and seasonal forecasts. As has been highlighted in this paper, radiation interacts with many other parts of the model, and radiation developments need to be pushed forward on many fronts in concert with improvements to the other parts of the model. There is a particular need to improve the consistency of assumptions made in the different schemes, and to exploit new observations to improve these assumptions and to understand possible compensating errors between schemes. A unifying framework for tackling these disparate challenges is the new ecRad radiation scheme, whose flexibility and efficiency will facilitate the scientific testing and implementation of the new ideas that will enable us to meet these challenges in the coming decade.

This paper has explored a large number of potential improvements to the treatment of radiation, but naturally not all can be taken forward with the same priority. We next provide an outlook for the work highlighted in the discussion of the five challenges, and then summarize the likely order in which they will be tackled.

1. Considering first the *surface,* improving 2-m temperature forecasts in different regions requires a careful unpicking of the role of the relevant processes, including radiation. While there is always scope to improve the representation of surface albedo, such as for snow-covered surfaces, the current 'single flat tile' paradigm will never properly capture important effects in more complex terrain. We have demonstrated an efficient new way to treat 3D radiative effects in forest and urban canopies and tested it in the offline version of ecRad. This offers the potential for more accurate calculation of the reflection from forests with snow on the ground, as well as photosynthesis rates if needed. The ability to compute the shortwave and longwave fluxes into the three main facets of an urban environment will facilitate the development of an urban tile in the IFS, which would also treat turbulent fluxes of heat and other tracers between the urban surfaces and the boundary layer above. Judging from the experience of representing urban areas in some other regional NWP and climate models, this should lead to better forecasts for half the world's population that live in cities.

2. A list of desirable physically-based improvements to the treatment of *clouds* in the radiation scheme was identified. The changes with no computational cost included adopting 'EXP-RAN' overlap, regime-dependent cloud heterogeneity consistent with the cloud microphysics scheme, improved formulations of ice and liquid effective radius, and representing convective cloud in the radiation scheme. With a 10% cost increase, ecRad also allows longwave scattering to be represented. Since these changes can have opposing effects on the radiation budget, the plan is to group as many together at once before including in an IFS cycle. Although not affordable operationally, ecRad also has the option to represent 3D effects for the first time in a global model. This will

enable us to determine the importance of 3D radiative effects on the Earth's radiation budget, and if it is significant, to explore cheaper ways of approximating the impact in the operational model. It should also facilitate better forecasts of direct shortwave fluxes for the solar energy sector.

3. Two modifications with very different effects on *clear-sky shortwave absorption* were tested. An upgrade of the near-infrared water vapour continuum led to a weak but very widespread warming of the troposphere; aside from slightly countering the cold bias in the tropical troposphere, it was difficult to detect any robust improvements to weather forecasts. By contrast, an upgrade of the aerosol climatology caused stronger but more localized changes to tropospheric heating, which improved the low-level flow responsible for the Indian Summer Monsoon. It was also found that the IFS analysis system is very sensitive to errors in the aerosol *absorption* optical depth, a quantity that is very poorly observed. This opens up an intriguing research question: can we use the model's wind and temperature biases in the vicinity of other monsoon systems to diagnose locations where the aerosol absorption is wrong? This would stimulate a deeper examination of whether the aerosol sources or optical properties in the CAMS system need to be improved.

4. A concerted effort to improve the *middle atmosphere* in the IFS has been made by the Stratosphere Task Force over the last year. In this paper we have shown that the extremely large warm biases (up to 20 K) in the upper stratosphere and mesosphere can be largely removed by a combination of improvements that reduce solar heating rates, the most significant being the use of an updated solar spectrum with 7–8% less ultraviolet. However, we have also found that the stratosphere tends to cool at higher horizontal resolutions (Fig. 3c), a problem that needs to be understood and resolved before the improvements can be implemented operationally. An even longer-standing error is the polar lower stratosphere cold bias of ∼5 K in both analyses and forecasts, due to radiation acting on a very large overestimate of humidity. Artificially removing the temperature bias appears to lead to more accurate monthly forecasts, so the next step is to find a physical solution to the excessive transport of water vapour from troposphere to stratosphere, and to understand the dynamical mechanisms by which a reduction in temperature bias leads to improved predictive skill in the troposphere.

5. Finally to *efficiency,* the relative cost of radiation has dropped by a factor of four in the last decade, giving it a comparable cost to the other individual physics parameterizations. A priority is to increase the frequency of radiation calls to 1 h in all model configurations, which could be afforded by using a gas optics model with significantly fewer spectral intervals than RRTM-G. This could be achieved either by implementing an alternative existing model, or by taking on the more ambitious task of developing the full-spectrum correlated-k idea at ECMWF. The modular structure of ecRad makes it straightforward to support several gas-optics models (including RRTM-G), so that different configurations of the IFS could choose a different trade-off between spectral resolution and other priorities.

The priority to be applied to each of these items depends on their likely impact on forecasts, the time needed to do the necessary research and implementation, and parallel developments in the other parts of the IFS on which radiation improvements depend. The first items to be carried forward towards operational implementation are the package of physically-based improvements to the treatment of clouds, and the modifications that reduce the middle-atmosphere temperature bias. On a similar timescale, we envisage the aerosol climatology to be upgraded as new reanalysis datasets become available from CAMS. The development of an urban tile is envisaged to start in the next two years, at which time we would consider an overhaul of the treatment of the surface in the radiation scheme, including possible 3D radiative effects in urban and forested surfaces. Investigating faster gas-optics schemes would likely be started on

a similar timescale. The introduction of orographic radiative effects would be tested closer to the time when the next resolution upgrade will occur; a resolution of 5 km is expected in 2025.

Acknowledgments

References

Ahlgrimm, M., and R. M. Forbes, 2016: Regime dependence of cloud condensate variability observed at the Atmospheric Radiation Measurement Sites. *Q. J. R. Meteorol. Soc.,* **142,** 1605–1617.

Ahlgrimm, M., R. M. Forbes, J.-J. Morcrette and R. A. Neggers, 2016: ARM's impact on numerical weather prediction at ECMWF. *Meteorol. Monographs,* **57,** 28.1–28.13.

Allan, R. P., and M. A. Ringer, 2003: Inconsistencies between satellite estimates of longwave cloud forcing and dynamical fields from reanalyses. *Geophys. Res. Lett.,* **30,** doi:10.1029/2003GL017019.

Baldwin, M. P., and T. J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science,* **294,** 581–584.

Balsamo, G., A. Agustì-Panareda, C. Albergel, A. Beljaars, S. Boussetta, E. Dutra, T. Komori, S. Lang, J. Muñoz-Sabater, F. Pappenberger, P. de Rosnay, I. Sandu, N. Wedi, A. Weisheimer, F. Wetterhall and E. Zsoter, 2014: Representing the Earth surfaces in the Integrated Forecasting System: Recent advances and future challenges. ECMWF Technical Memorandum 729, 48 pp.

Baran, A. J., P. Hill, K. Furtado, P. Field and J. Manners, 2014: A coupled cloud physics–radiation parameterization of the bulk optical properties of cirrus and its impact on the Met Office Unified Model Global Atmosphere 5.0 configuration. *J. Clim.,* **27,** 7725–7752.

Barker, H. W., 2008: Overlap of fractional cloud for radiation calculations in GCMs: A global analysis using CloudSat and CALIPSO data. *J. Geophys. Res. Atmos.,* **113,** doi:10.1029/2007JD009677.

Beljaars, A., 2017: Towards optimal parameters for the prediction of near surface temperature and dew point. ECMWF Technical Memorandum, in preparation.

Berrisford, P., P. Kållberg, S. Kobayashi, D. Dee, S. Uppala, A. J. Simmons, P. Poli and H. Sato, 2011: Atmospheric conservation properties in ERA-Interim. *Q. J. R. Meteorol. Soc.,* **137,** 1381–1399.

Best, M. J., M. Pryor, D. B. Clark, G. G. Rooney, R. Essery, C. B. Ménard, J. M. Edwards, M. A. Hendry, A. Porson, N. Gedney and L. M. Mercado, 2011: The Joint UK Land Environment Simulator (JULES), model description: 1. Energy and water fluxes. *Geosci. Model Dev.,* **4,** 677–699.

Bodas-Salcedo, A., K. D. Williams, M. A. Ringer, I. Beau, J. N. S. Cole, J.-L. Dufresne, T. Koshiro, B. Stevens, Z. Wang and T. Yokohata, 2014: Origins of the solar radiation biases over the southern ocean in CFMIP2 models. *J. Clim.,* **27,** 41–56.

Bollasina, M., Y. Ming and V. Ramaswamy, 2011: Anthropogenic aerosols and the weakening of the South Asian summer monsoon. *Science,* **334,** 502–505.

Boussetta, S., G. Balsamo, A. Beljaars, A. Agusti Panareda, J.-C. Calvet, C. Jacobs, B. van den Hurk, P. Viterbo, S. Lafont, E. Dutra, L. Jarlan, M. Balzarolo, D. Papale and G. van der Werf, 2013: Natural land carbon dioxide exchanges in the ECMWF integrated forecasting system: Implementation and offline validation. *J. Geophys. Res. Atmos.,* **118,** 5923–5946.

Boussetta, S., G. Balsamo, E. Dutra, A. Beljaars and C. Albergel, 2015: Assimilation of surface albedo and vegetation states from satellite observations and their impact on numerical weather prediction. *Rem. Sens. Env.,* **163,** 111–126.

Bozzo, A., R. Pincus, I. Sandu and J.-J. Morcrette, 2014: Impact of a spectral sampling technique for radiation on ECMWF weather forecasts. *J. Adv. Model. Earth Syst.,* **6,** 1288–1300.

Bozzo, A., S. Remy, J. Flemming, A. Benedetti, M. J. Rodwell and J.-J. Morcrette, 2017. ECMWF Technical Memorandum 801, 33 pp.

Calisto, M., D. Folini, M. Wild and L. Bengtsson: Cloud radiative forcing intercomparison between fully coupled CMIP5 models and CERES satellite data. *Ann. Geophys.,* **32,** 793–807.

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.

Costa, S. M. S., and K. P. Shine, 2006: An estimate of the global impact of multiple scattering by clouds on outgoing long-wave radiation. *Q. J. R. Meteorol. Soc.,* **132,** 885–895.

Di Giuseppe, F., and A. M. Tompkins, 2015: Generalizing cloud overlap treatment to include the effect of wind shear. *J. Atmos. Sci.,* **72,** 2865–2876.

Di Giuseppe, F., F. Pappenberger, F. Wetterhall, B. Krzeminski, A. Camia, G. Libertá and J. San Miguel, 2016: The potential predictability of fire danger provided by Numerical Weather Prediction. *J. Appl. Meteor. Climatol.,* **55,** 2469–2491.

Dyroff, C., A. Zahn, E. Christner, R. M. Forbes, A. M. Tompkins and P. F. J. van Velthoven, 2015: Comparison of ECMWF analysis and forecast humidity data with CARIBIC upper troposphere and lower stratosphere observations. *Q. J. R. Meteorol. Soc.,* **141,** 833–844.

Dutra, E., G. Balsamo, P. Viterbo, P. M. A. Miranda, A. Beljaars, C. Schär and K. Elder, 2010: An improved snow scheme for the ECMWF land surface model: Description and offline validation. *J. Hydromet.,* **11,** 899–916.

Ebert, E. E., and J. A. Curry, 1993: An intermediate one-dimensional thermodynamic sea ice model for investigating ice–atmosphere interactions. *J. Geophys. Res. Atmos.,* **98,** 10085–10109.

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.

ECMWF, 2007: Recent advances in radiation transfer parametrizations. *Paper to 36th session of ECMWF Science Advisory Committee,* ECMWF/SAC/36(07)5, http://www.ecmwf.int/system/files/livelink/5805951/SAC36(07)5.pdf.

Flemming, J., A. Benedetti, A. Inness, R. J. Engelen, L. Jones, V. Huijnen, S. Remy, M. Parrington, M. Suttie, A. Bozzo, V.-H. Peuch, D. Akritidis and E. Katragkou: The CAMS interim reanalysis of carbon monoxide, ozone and aerosol for 2003–2015. *Atmos. Chem. Phys.,* **17,** 1945–1983.

Feldman, D. R., W. D. Collins, R. Pincus, X. Huang and X. Chen, 2014: Far-infrared surface emissivity and climate. *Proc. Nat. Acad. Sci. USA,* **111,** 16297–16302.

Forbes, R. M., and M. Ahlgrimm, 2014: On the representation of high-latitude boundary layer mixed-phase cloud in the ECMWF global model. *Mon. Weath. Rev.,* **142,** 3425–3445.

Forbes, R., A. Geer, K. Lonitz and M. Ahlgrimm, 2016: Reducing systematic errors in cold-air outbreaks. ECMWF Newsletter 146.

Fortuin, J. P. F., and U. Langematz, 1994: An update on the global ozone climatology and on concurrent ozone and temperature trends. *Proc. SPIE Atmos. Sensing and Modeling,* **2311,** 207–216.

Fouquart, Y., 1987: Radiative transfer in climate modeling. In M. E. Schlesinger (Ed.), *NATO Adv. Study Inst. on physically based modeling and simulation of climate and climate changes,* 223–283, Erice, Sicily, 11–23 May 1986.

Freidenreich, S. M., and V. Ramaswamy, 1999: A new multiple-band solar radiative parameterization for general circulation models. *J. Geophys. Res. Atmos.,* **104,** 31389–31409.

Fu, Q., 1996: An accurate parameterization of the solar radiative properties of cirrus clouds. *J. Climate,* **9,** 2058–2082.

Fu, Q., K. N. Liou, M. C. Cribb, T. P. Charlock and A. Grossman, 1997: Multiple scattering parameterization in thermal infrared radiative transfer. *J. Atmos. Sci.,* **54,** 2799–2812.

Fu, Q., P. Yang and W. B. Sun, 1998: An accurate parametrization of the infrared radiative properties of cirrus clouds of climate models. *J. Climate,* **11,** 2223-2237.

Garfinkel, C. I., D. W. Waugh and E. P. Gerber, 2013: The effect of tropospheric jet latitude on coupling between the stratospheric polar vortex and the troposphere. *J. Clim.,* **26,** 2077–2095.

Gates, W. L., J. S. Boyle, C. Covey, C.G. Dease, C. M. Doutriaux, R. S. Drach, M. Fiorino, P. J. Glecker, J. J. Hnilo, S. M. Marlais, T. J. Phillips, G. L. Potter, B. D. Santer, K. R. Sperber, K. E. Taylor and D. N. Williams, 1999: An overview of the results of the atmospheric model intercomparison project (AMIP I). *Bull. Am. Meteorol. Soc.,* **80,** 29–55.

Geer, A., M. Ahlgrimm, P. Bechtold, M. Bonavita, N. Bormann, S. English, M. Fielding, R. Forbes, R. Hogan, E. Hólm, M. Janisková, K. Lonitz, P. Lopez, M. Matricardi, I. Sandu and P. Weston, 2017: Assimilating observations sensitive to cloud and precipitation. ECMWF Technical Memorandum 815, 60 pp.

Grimmond, C. S., M. Blackett, M. J. Best, J. Barlow, J. Baik, S. E. Belcher, S. I. Bohnenstengel, I. Calmet, F. Chen, A. Dandou, K. Fortuniak, M. L. Gouvea, R. Hamdi, M. Hendry, T. Kawai, Y. Kawamoto, H. Kondo, E. S. Krayenhoff, S. Lee, T. Loridan, A. Martilli, V. Masson, S. Miao, K. Oleson, G. Pigeon, A. Porson, Y. Ryu, F. Salamanca, L. Shashua-Bar, G. Steeneveld, M. Tombrou, J. Voogt, D. Young, and N. Zhang, 2010: The international urban energy balance models comparison project: first results from Phase 1. *J. Appl. Meteor. Climatol.,* **49,** 1268–1292.

Hansen, M., R. S. DeFries, J. R. G. Townshend, M. Carroll, C. Dimiceli and R. A. Sohlberg, 2003: Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm, *Earth Interact.,* **7,** 1–15.

Haiden, T., 2015: Surface verification in the Arctic. ECMWF Newsletter 145.

Hersbach, H., C. Peubey, A. Simmons, P. Berrisford, P. Poli and D. Dee, 2015: ERA-20CM: a twentieth-century atmospheric model ensemble. *Q. J. R. Meteorol. Soc,* **141,** 2350–2375.

Hobbs, W., M. D. Palmer and D. Monselesan, 2016: An energy conservation analysis of ocean drift in the CMIP5 global coupled models. *J. Clim.,* **29,** 1639–1653.

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 A. J. Illingworth, 2000: Deriving cloud overlap statistics from radar. *Q. J. R. Meteorol. Soc.,* **126,** 2903–2909.

Hogan, R. J., and A. Bozzo, 2015: Mitigating errors in surface temperature forecasts using approximate radiation updates. *J. Adv. Model. Earth. Syst.,* **7,** 836–853.

Hogan, R. J., and A. Bozzo, 2016: ECRAD: A new radiation scheme for the IFS. ECMWF Technical Memorandum 787, 33 pp.

Hogan, R. J., and S. Hirahara, 2016: Effect of solar zenith angle specification in models on mean shortwave fluxes and stratospheric temperatures. *Geophys. Res. Lett.,* **43,** 482–488.

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

Hogan, R. J., T. Quaife and R. Braghiere, 2017: Fast matrix treatment of 3D radiative transfer in vegetation canopies: SPARTACUS-Vegetation 1.0. *Geosci. Model Dev. Discuss.,* doi:10.5194/gmd-2017-208.

Huang, F. T., R. D. McPeters, P. K. Bhartia, H. G. Mayr, S. M. Frith, J. M. Russell III and M. G. Mlynczak, 2010: Temperature diurnal variations (migrating tides) in the stratosphere and lower mesosphere based on measurements from SABER on TIMED. *J. Geophys. Res. Atmos.,* **115,** doi:10.1029/2009JD013698.

Jensen, M. P., A. M. Vogelmann, W. D. Collins, G. J. Zhang and E. P. Luke, 2008: Investigation of regional and seasonal variations in marine boundary layer cloud properties from MODIS observations. *J. Climate,* **21,** 4955–4973.

Klinger, C., B. Mayer, F. Jakub, T. Zinner, S.-B. Park and P. Gentine, 2017: Effects of 3-D thermal radiation on the development of a shallow cumulus cloud field. *Atmos. Chem. Phys.,* **17,** 5477–5500.

Kurucz, R. L., 1994: Synthetic infrared spectra. In Infrared solar physics: Proc. 154th Symp. Int. Astronomical Union, Tuscon USA, March 2–6, 1992 (Eds. D. M. Rabin, J. T. Jefferies and C. Lindsey), Kluwer Academic Publishers, Dordrecht, 523 pp.

Leutbecher, M., S. J. Lock, P. Ollinaho, S. T. K. Lang, G. Balsamo, P. Bechtold, M. Bonavita, H. M. Christensen, M. Diamantakis, E. Dutra, E., S. English, M. Fisher, R. M. Forbes, J. Goddard, T. Haiden, R. J. Hogan, S. Juricke, H. Lawrence, D. MacLeod, L. Magnusson, S. Malardel, S. Massart, I. Sandu, P. K. Smolarkiewicz, A. Subramanian, F. Vitart, N. Wedi and A. Weisheimer, 2017: Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Q. J. R. Meteorol. Soc.,* doi:10.1002/qj.3094.

Li, J. and H. W. Barker, 2005: A radiation algorithm with correlated k-distribution. 1: Local thermal equilibrium. *J. Atmos. Sci.,* **62,** 286–309.

Li, J.-L. F., R. M. Forbes, D. E. Waliser, G. Stephens and S. Lee, 2014: Characterizing the radiative impacts of precipitating snow in the ECMWF Integrated Forecast System global model. *J. Geophys. Res. Atmos.,* **119,** 9626–9637.

Li, D., S. Malyshev and E. Shevliakova, 2016a: Exploring historical and future urban climate in the Earth System Modeling framework: 1. Model development and evaluation. *J. Adv. Model. Earth Syst.,* **8,** 917–935.

Li, D., S. Malyshev and E. Shevliakova, 2016b: Exploring historical and future urban climate in the Earth System Modeling framework: 2. Impact of urban land use over the Continental United States. *J. Adv. Model. Earth Syst.,* **8,** 936–953.

Lindner, T. H., and J. Li, 2000: Parameterization of the optical properties for water clouds in the infrared. *J. Climate,* **13,** 1797–1805.

Macfarlane, S. A., J. H. Mather and E. J. Mlawer, 2016: ARM's rogress on improving atmospheric broadband radiative fluxes and heating rates. *Meteorol. Monographs,* **57,** 20.1–20.24.

Manners, J., J.-C. Thelen, J. Petch, P. Hill and J. M. Edwards, 2009: Two fast radiative transfer methods to improve the temporal sampling of clouds in numerical weather prediction and climate models. *Q. J. R. Meteorol. Soc.,* **135,** 457-468.

Manners, J., S. B. Vosper and N. Roberts, 2012: Radiative transfer over resolved topographic features for high-resolution weather prediction. *Q. J. R. Meteorol. Soc.,* **138,** 720–733.

Marshall, A. G., and A. A. Scaife, 2009: Impact of the QBO on surface winter climate, *J. Geophys. Res. Atmos.,* **114,** D18110, doi:10.1029/2009JD011737.

Mašek, J., J.-F. Geleyn, R. Brožková, O. Giot, H. O. Achom and P. Kuma, 2016: Single interval shortwave radiation scheme with parameterized optical saturation and spectral overlaps. *Q. J. R. Meteorol. Soc.,* **142,** 304–326.

Mayer, M., L. Haimberger, J. M. Edwards, and P. Hyder, 2017: Toward consistent diagnostics of the coupled atmosphere and ocean energy budgets. *J. Climate,* **30,** 9225–9246.

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.

Monge-Sanz, B. M., M. P. Chipperfield, D. Cariolle and W. Feng, 2011: Results from a new linear O3 scheme with embedded heterogeneous chemistry compared with the parent full-chemistry 3-D CTM. *Atmos. Chem. Phys.,* **11,** 1227–1242.

Morcrette, J.-J., 1991: Radiation and cloud radiative properties in the ECMWF operational weather forecast model. *J. Geophys. Res. Atmos.,* **96,** 9121–9132.

Morcrette, J.-J., H. W. Barker, J. N. S. Cole, M. J. Iacono, R. Pincus, 2008a: Impact of a new radiation package, McRad, in the ECMWF Integrated Forecasting System. *Mon. Weath. Rev.,* **136,** 4773–4798.

Morcrette, J.-J., G. Mozdzynski and M. Leutbecher, 2008b: A reduced radiation grid for the ECMWF Integrated Forecasting System. *Mon. Weath. Rev.,* **136,** 4760–4772.

Morcrette, J.-J., A. Benedetti, A. Ghelli, J. W. Kaiser and A. M. Tompkins, 2011: Aerosol-cloud-radiation interactions and their impact on ECMWF/MACC forecasts. ECMWF Technical Memorandum 660, 35 pp.

Mozdzynski, G., and J.-J. Morcrette, 2014: Reorganization of the radiation transfer calculations in the ECMWF IFS. ECMWF Technical Memorandum 721, 20 pp.

Mulcahy, J. P., D. N. Walters, N. Bellouin and S. F. Milton, 2014: Impacts of increasing the aerosol complexity in the Met Office global numerical weather prediction model. *Atmos. Chem. Phys.,* **14,** 4749–4778.

Nam, C., S. Bony, J.-L. Dufresne and H. Chepfer, 2012: The 'too few, too bright' tropical low-cloud problem in CMIP5 models. *Geophys. Res. Lett.,* **39,** doi:10.1029/2012GL053421.

Nielsen, K. P., E. Gleeson and L. Rontu, 2014: Radiation sensitivity tests of the HARMONIE 37h1 NWP model. *Geosci. Model Dev.,* **7,** 1433–1449.

Oke, T. R., 1982: The energetic basis of the urban heat island. *Q. J. R. Meteorol. Soc.,* **108,** 1–24.

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.

Pincus, R., H. W. Barker, and J.-J. Morcrette, 2003: A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous clouds. *J. Geophys. Res. Atmos.,* **108,** 4376, doi:10.1029/2002JD003322.

Pinty, B., T. Lavergne, R. E. Dickinson, J.-L. Widlowski, N. Gobron and M. M. Verstraete, 2006: Simplifying the inter-action of land surfaces with radiation for relating remote sensing products to climate models, *J. Geophys. Res.,* **111,** doi:10.1029/2005JD005952.

Polichtchouk, I., R. J. Hogan, P. Bechtold, S. Malardel, T. Stockdale, S.-J. Lock and T. G. Shepherd, 2017: How accurate is the middle-atmosphere circulation in the IFS? ECMWF Technical Memorandum, in preparation.

Rädel, G., K. P. Shine and I. V. Ptashnik, 2015: Global radiative and climate effect of the water vapour continuum at visible and near-infrared wavelengths. *Q. J. R. Meteorol. Soc.,* **141,** 727–738.

Räisänen, P., 2002: Two-stream approximations revisited: A new improvement and tests with GCM data. *Q. J. R. Meteorol. Soc.,* **128,** 2397–2416.

Rodwell, M. J., and T. Jung, 2008: Understanding the local and global impacts of model physics changes: An aerosol example. *Q. J. R. Meteorol. Soc.,* **134,** 1479–1797.

Schäfer, S. A. K., 2016: *What is the global impact of 3D cloud-radiation interactions?* PhD thesis, University of Reading, UK.

Schaaf, C. B., F. Gao, A. H. Strahler, W. Lucht, X. Li, T. Tsang, N. C. Strugnell, X. Zhang, Y. Jin, J. P. Muller and P. Lewis, 2002: First operational BRDF, albedo nadir reflectance products from MODIS. *Remote. Sens. Env.,* **83,** 135–148.

Schneider, A., M. A. Friedl and D. Potere, 2009: A new map of global urban extent from MODIS data. *Env. Res. Lett.,* **4,** article 044003.

Seethala, C., and Á. Horváth, 2010: Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds. *J. Geophys. Res. Atmos.,* **115,** doi:10.1029/2009JD012662.

Sellers, P. J., 1985: Canopy reflectance, photosynthesis and transpiration. *Int. J. Remote Sens.,* **6,** 1335–1372.

Senkova, A. V., L. Rontu and H. Savijärvi, 2007: Parametrization of orographic effects on surface radiation in HIRLAM. *Tellus A,* **59,** 279–291.

Slingo, A., 1989: A GCM parametrization for the shortwave radiative properties of water clouds. *J. Atmos. Sci.,* **46,** 1419–1427.

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.

Shonk, J. K. P., and R. J. Hogan, 2008: Tripleclouds: an efficient method for representing horizontal cloud inhomogeneity in 1D radiation schemes by using three regions at each height. *J. Climate,* **21,** 2352–2370.

Shonk, J. K. P., R. J. Hogan, J. M. Edwards and G. G. Mace, 2010: Effect of improving representation of horizontal and vertical cloud structure on the Earth's radiation budget: 1. Review and parameterisation. *Q. J. R. Meteorol. Soc.,* **136,** 1191–1204.

Simard, M., N. Pinto, J. B. Fisher and A. Baccini, 2011: Mapping forest canopy height globally with spaceborne lidar. *J. Geophys. Res. Biogeosci.,* **116,** doi:10.1029/2011JG001708.

Smith, E. A., and L. Shi, 1992: Surface forcing of the infrared cooling profile over the Tibetan plateau – 1: Influence of relative longwave radiative heating at high altitude. *J. Atmos. Sci.,* **49,** 805–822.

Sotiropoulou, G., J. Sedlar, R. Forbes and M. Tjernström, 2016: Summer Arctic clouds in the ECMWF forecast model: an evaluation of cloud parametrization schemes. *Q. J. R. Meteorol. Soc.,* **142,** 387–400.

Stenke, A., V. Grewe and M. Ponater, 2008: Lagrangian transport of water vapor and cloud water in the ECHAM4 GCM and its impact on the cold bias. *Clim. Dyn.,* **31,** 491–506.

Studer, S., K. Hocke, A. Schanz, H. Schmidt and N. Kämpfer: A climatology of the diurnal variations in stratospheric and mesospheric ozone over Bern, Switzerland. *Atmos. Chem. Phys.,* **14,** 5905–5919.

Tanré, D., J.-F. Geleyn and J. M. Slingo, 1984: First results of the introduction of an advanced aerosol-radiation interaction in the ECMWF low resolution global model. In H. E. Gerber and A. Deepak (Eds.), *Aerosols and their climatic effects,* 133–177. A. Deepak Publ., Hampton, VA.

Tegen, I., P. Hollrig, M. Chin, I. Fung, D. Jacob and J. Penner, 1997: Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results. *J. Geophys. Res. Atmos.,* **102,** 23895–23915.

Thompson, D. W. J., M. P. Baldwin and J. M. Wallace, 2002: Stratospheric connection to Northern Hemisphere wintertime weather: Implications for prediction. *J. Clim.,* **15,** 1421–1428.

Van Weverberg, K., C. J. Morcrette, H.-Y. Ma, S. A. Klein and J. C. Petch, 2015: Using regime analysis to identify the contribution of clouds to surface temperature errors in weather and climate models. *Q. J. R. Meteorol. Soc.,* **141,** 3190–3206.

Vinoj, V., P. J. Rasch, H. Wang, J. H. Yoon, P. L. Ma, K. Landu and B. Singh, 2014: Short-term modulation of Indian summer monsoon rainfall by West Asian dust. *Nature Geosci.,* **7,** 308–313.

Vitart, F., G. Balsamo, R. Buizza, L. Ferranti, S. Keeley, L. Magnusson, F. Molteni and A. Weisheimer, 2014: Sub-seasonal predictions. ECMWF Technical Memorandum 738, 45 pp.

Ward, H. C., S. Kotthaus, L. Järvi and C. S. B. Grimmond, 2016: Surface urban energy and water balance scheme (SUEWS): Development and evaluation at two UK sites. *Urban Clim.,* **18,** 1–32.

Widlowski, J.-L., B. Pinty, M. Clerici, Y. Dai, M. De Kauwe, K. de Ridder, A. Kallel, H. Kobayashi, T. Lavergne, W. Ni-Meister, A. Olchev, T. Quaife, S. Wang, W. Yang, Y. Yang and H. Yuan, 2011: RAMI4PILPS: An intercomparison of formulations for the partitioning of solar radiation in land surface models. *J. Geophys. Res. Biogeosci.,* **116,** doi:10.1029/2010JG001511.

Wild, M., D. Folini, M. Z. Hakuba, C. Schär, S. I. Seneviratne, S. Kato, D. Rutan, C. Ammann, E. F. Wood and G. König-Langlo, 2015: The energy balance over land and oceans: an assessment based on direct observations and CMIP5 climate models. *Clim. Dyn.,* **44,** 3393–3429.

Woodward, S., 2001: Modeling the atmospheric life cycle and radiative impact of mineral dust in the Hadley Centre climate model. *J. Geophys. Res. Atmos., 106,* 18155–18166.

Yabu, S., 2013: Development of longwave radiation scheme with consideration of scattering by clouds in JMA global model. *CAS/JSC WGNE Res. Activ. Atmos. Oceanic Modell.,* **43,** 04.07–04.08.

Yi, B., P. Yang, B. A. Baum, T. L'Ecuyer, L. Oreopoulos, E. J. Mlawer, A. J. Heymsfield and K.-K. Liou, 2013: Influence of ice particle surface roughening on the global cloud radiative effect. *J. Atmos. Sci.,* **70,** 2794–2807.

Zhong, W., S. M. Osprey, L. J. Gray and J. D. Haigh, 2008: Influence of the prescribed solar spectrum on calculations of atmospheric temperature. *Geophys. Res. Lett.,* **35,** L22813.