

Paths to accuracy for radiation parameterizations in atmospheric models

Robert Pincus¹ and Bjorn Stevens²

Received 11 December 2012; revised 11 March 2013; accepted 15 March 2013.

[1] Radiative transfer is sufficiently well understood that its parameterization in atmospheric models is primarily an effort to balance computational cost and accuracy. The most common approach is to compute radiative transfer with the highest practical spectral accuracy but infrequently in time and/or space, though errors introduced by this approximation are difficult to quantify. An alternative is to perform spectrally sparse calculations frequently in time using randomly chosen spectral quadrature points. Here we show that purely random quadrature points, though effective in some large-eddy simulations, are not a good choice for models in which the land surface responds to radiative fluxes because surface temperature perturbations can be large enough, and persistent long enough, to affect model evolution. These errors may be mitigated by choosing teams of spectral points designed to limit the maximum surface flux error; teams, rather than individual quadrature points, are then chosen randomly. The approach is implemented in the ECHAM6 global model and the results are examined using “perfect-model” experiments on time scales ranging from a day to a month. In this application the approach introduces errors commensurate with the infrequent calculation of broadband calculations for the same computational cost. But because teams need not increase with size, and indeed may become better and more balanced with increased spectral density, improvements in radiative transfer may not need to be traded off against spatiotemporal sampling.

Citation: Pincus, R., and B. Stevens (2013), Paths to accuracy for radiation parameterizations in atmospheric models, *J. Adv. Model. Earth Syst.*, 5, 1–9, doi:10.1002/jame.20027.

1. What Does it Mean for a Parameterization to be Accurate?

[2] Unlike many closure problems faced in models of the atmosphere, the environmental factors that control the distribution of radiation in the atmosphere are very well understood, so the solution to fully specified problems is known to great accuracy. Radiation parameterizations therefore seek primarily to find an acceptable compromise between accuracy and computational cost. The accuracy of radiative transfer calculations may be measured via comparison to benchmark models [Oreopoulos *et al.*, 2012] which are themselves known to be in excellent agreement with observations [Mlawer *et al.*, 2000; Turner *et al.*, 2004]. Comparisons are normally made for clear-sky conditions, consistent with the way the parameterizations of absorption by gases are developed.

[3] State-of-the-art radiation parameterizations can reproduce benchmark calculations to within 1% for shortwave fluxes and fractions of a percent for longwave fluxes [Oreopoulos *et al.*, 2012] but this accuracy is so computationally expensive that radiation parameterizations cannot be applied at every time step of the model. Instead, radiative heating and cooling rates are normally updated less frequently than are model dynamics and, in most cases, other physical parameterizations. The choice to update radiative heating rates less frequently than other fields is an approximation made, not in the radiation parameterization, but in the coupling to the rest of the model. The simulation errors caused by this approximation may range from modest changes in temperature fields [Xu and Randall, 1995; Morcrette, 2000] to the introduction of more dramatic instabilities [Pauluis and Emanuel, 2004] but are generally difficult to quantify. To minimize simulation errors prudence dictates that the radiation time step be as close to the dynamical time step as can be afforded, although precisely how close is a subjective choice.

[4] Several approaches have been developed to accelerate the calculation of radiative fluxes to allow for more frequent calculation. One is to use physically based radiative transfer models to train fast statistical

¹University of Colorado and NOAA Earth System Research Lab, Physical Sciences Division, University of Colorado, Boulder, Colorado, USA.

²Max Planck Institute for Meteorology, Hamburg, Germany.

models (normally artificial neural networks) to emulate fluxes based on the state of the atmosphere [e.g., *Chevallier et al.*, 1998; *Krasnopolsky et al.*, 2008]. An intermediate tactic is to apply physical models sparsely in space and/or time, use simple statistical models (e.g., regression) to predict changes since the last radiation time step, and selectively update calculations based on the expected error [*Venema et al.*, 2007]. A third alternative exploits two facts, that cloud properties vary much more quickly in the atmosphere than does the concentration of gases, and that variations in clouds and gases affect fluxes in different, roughly disjoint spectral regions, to motivate updating only the cloud-affected portions of the spectrum at high frequency [*Manners et al.*, 2009]. (*Räisänen and Barker* [2004] and *Hill et al.* [2011] take a similar approach to the related problem of representing cloud variability.) The calculation of cloud-affected fluxes can be further accelerated by reducing the spectral detail used to treat absorption by gases [*Manners et al.*, 2009].

[5] Each of these methods, including infrequent radiation calculations, represent approximations which introduce errors in radiative heating rates. These errors depend on many factors including how quickly the optical properties of the atmosphere are changing. But the error characteristics of an approximation can be crucially important in determining whether the approximation affects model evolution. Since radiative fluxes at the top of the atmosphere are essentially in balance (after accounting for ocean heat storage), for example, even small ($1 < W/m^2$) biases in radiative fluxes affect multidecadal simulations and must be “tuned” away [*Mauritsen et al.*, 2012] and/or balanced by compensating errors. Random, uncorrelated noise, on the other hand, does not affect the statistical evolution of most models, whether that noise comes from parameterizations of gravity wave drag [*Eckermann*, 2012; *Lott et al.*, 2012] or radiation [*Pincus et al.*, 2003] or is externally applied in an effort to diversify ensembles of medium-range forecasts [*Buizza et al.*, 1999]. For the purposes of parameterization development this implies that unbiased algorithms, even if they introduce even quite substantial noise in heating rates, can be more accurate, in the sense of introducing smaller changes in model evolution, than other approximations including detailed algorithms used infrequently.

[6] Here we describe an approach to radiative transfer parameterization that emphasizes the accuracy relevant for hydrodynamic models, including both the radiation calculations and the ways those calculations are coupled to the rest of the model. The approach takes advantage of the local homogenization of heating rates arising from small-scale fluid dynamical processes. We have implemented these ideas in a new radiation package, PSrad (named because it is a postscript to the RRTMG package from which it descends), and initially implemented in the ECHAM climate model. PSrad is unique in that it allows only a small sample of the full broadband spectral integration to be performed, with the idea that these calculations should be performed at each time step. This spectral sampling introduces

grid-scale noise in radiative fluxes, as does the more common use of stochastic samples to represent the sub-grid-scale distribution of cloud properties [*Pincus et al.*, 2003]. Experiments show that ECHAM is insensitive to even large grid-scale perturbations to radiative heating rates within the atmosphere, but that significant perturbations in surface fluxes can introduce systematic biases in the model trajectory. Simulation bias can be limited by bounding errors in surface fluxes using carefully selected subsets of the broadband calculation. The approach is applicable to dynamical models at all scales even as significant noise is introduced into individual calculations.

[7] The next section details several strategies for coupling radiation calculations to model integrations, including infrequent radiation calculations and spectral subsets computed at higher frequency. Section 3 describes the code we have developed to implement these strategies and section 4 the impact of two classes of approximations on forecasting and quasi-climatological time scales in climate model integrations. Section 5 discusses implications for weather forecasting and climate projection applications.

2. Strategies for Spectral Integration

[8] Models of the atmosphere require broadband radiation calculations, i.e., those that account for all wavelengths of radiation emitted by the sun or the earth and its atmosphere. In parameterizations of radiative transfer this spectral integration is accomplished using weighted sums

$$F(x, y, z, t) = \sum_g^G w_g F_g(x, y, z, t), \quad (1)$$

where the individual fluxes F_g are computed using optical properties and boundary conditions appropriate to each pseudospectral interval (quadrature point). These quadrature points are frequently determined using k -distributions [*Fu and Liou*, 1992; *Lacis and Oinas*, 1991]; following this nomenclature we refer to these intervals as “ g -points”. In the shortwave the weights w_g are determined by the spectral distribution of incoming solar energy while in the longwave they depend on local temperature through the Stefan-Boltzmann relation.

[9] In most models of the atmosphere radiative fields are updated less frequently than other variables, i.e.,

$$F(x, y, z, t) \approx c(x, y, t) \sum_g^G w_g F_g(x, y, z, t_{\text{rad}}(t) | t_{\text{rad}}(t) - t \leq \Delta t_{\text{rad}}), \quad (2)$$

where c represents correction factors that may be applied to account for time-varying solar zenith angles, surface temperatures, etc. In some implementations [e.g., *Morcrette*, 2000] spatial resolution may also be reduced.

2.1. Monte Carlo Spectral Integration

[10] Infrequent broadband calculations (e.g., equation (2)) can be described as a “spectrally dense, temporally sparse” approach to computing radiation transfer. Monte Carlo Spectral Integration [MCSI; see *Pincus and Stevens, 2009*] reverses these densities:

$$F(x, y, z, t) \approx G/G' \sum_{g \in [1..G']} w_g F_g(x, y, z, t), \quad (3)$$

where each of the G' samples is chosen randomly with replacement at each location and time step.

[11] MCSI was initially introduced for large-eddy simulation where it has the advantages of being consistent (i.e., converging with increasing temporal and/or spatial resolution, as do the other approximations used in LES) and in explicitly sampling temporal variability, especially in cloud optical properties. It takes advantage of the fact that local fluid instabilities homogenize sampling noise on small scales while, on the larger spatial and temporal scales where heating rates can effect the overall evolution of the flow, sampling noise is small. In several applications, including purely radiatively driven flow, LES using MCSI is statistically indistinguishable from LES using benchmark radiation calculations (e.g., equation (1)) [*Pincus and Stevens, 2009*].

[12] MCSI implemented similarly in a global model introduces much larger and more systematic errors. The green line in Figure 1 shows the global root-mean-square (RMS) difference in 2 m air temperature as a function of forecast lead time between a reference calculation that computes broadband radiation at every time step and every grid point and one using the same radiation code (described in section 3) but applying equation (3) with $G' = 1$. The figure shows the average over 29 independent forecasts (see section 4). RMS differences with respect to the reference forecast exceed 1.5 K after the first day and grow over time. A ~ 0.5 K diurnal cycle tracks the diurnal variation in global mean 2 m temperature and occurs because the land surface is not homogeneously distributed over Earth’s surface.

[13] Why is the MCSI approximation accurate (in the sense of not disturbing the flow) in large-eddy simulations but not in a global model? There are at least two significant distinctions. First, the parameterizations used in global models, especially those for deep and shallow convection, depend more nonlinearly on the atmospheric state than do the simple subgrid-scale models for turbulence and microphysics used in LES, so even unbiased random sampling noise can bias the flow through nonlinearities. More importantly, surface properties are frequently fixed in large-eddy simulations, while global models almost always include land surfaces whose temperature changes in response to surface fluxes. Perturbations to the surface temperature caused by sampling errors are not homogenized by mixing with neighboring columns, while errors within the atmosphere are mixed by the fluid flow. The impact of surface temperature perturbations dominates: on an aquaplanet with globally specified sea surface temperatures (Figure 1, purple line),

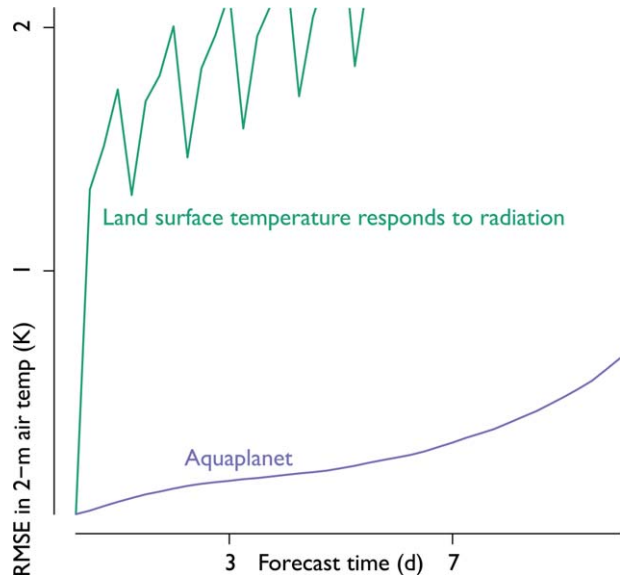


Figure 1. The impact of randomly sampling a single spectral interval at every time step (Monte Carlo Spectral Integration, equation (3)). RMS differences are computed with respect to reference forecasts (green) in which the land temperature responds to surface fluxes, and an aquaplanet with fixed sea surface temperatures (purple). The diurnal variation in RMSE for the earth-like planet tracks the diurnal variation in global 2 m air temperature caused by the asymmetric distribution of land. Sampling noise in radiative heating rates within the atmosphere drives modest differences, while noise can persist and modify the simulation when the surface temperature can change in response to surface radiation fluxes, both because heating rates are substantially larger at the surface than in the atmosphere and because sampling errors at the surface are not mixed by the flow as they are in the atmosphere. This implies that accuracy in radiation parameterizations depends in part on limiting perturbations to downwelling fluxes at the surface.

differences between simulations using equation (3) and simulations using equation (1) are small.

2.2. Bounding Errors in Surface Fluxes Using Teams of Spectral Points

[14] Figure 1 implies that the magnitude of instantaneous surface flux errors must be bounded if a radiation parameterization is to be useful in models including land surfaces. One approach would be to simply increase the numbers of samples, G' , chosen at each time step but this is painfully slow: like all Monte Carlo estimates the RMS error, for example, decreases as $1/\sqrt{G'}$ (see, e.g., section 4.2.5 of *Evans and Marshak, 2005*). A more efficient strategy is to generate sets of g -point “teams” constructed to minimize some measure of sampling error, and to sample these teams randomly.

[15] Assume a set of A representative atmospheres and define a cost function C as some measure of the error accumulated over L possible estimates F_l^a of the

true flux F^a in each atmosphere. The number of teams, M , may be chosen to be small divisor of the total number of g -points G so that each team has the same number $m = G/M$ of quadrature points. The members of these teams may then be chosen to minimize C .

[16] We have computed teams for several values of M using the radiative transfer code described in section 3. Our set of representative atmospheres is obtained from four snapshots taken over the course of a single day from a free run of ECHAM6 ($A \approx 73000$). We optimize over $L = MA$ clear-sky fluxes consistent with the way k -distributions are normally constructed. Since our goal is to minimize the possibility of very large errors in surface fluxes we use the 95% error in surface fluxes E_{srf}^{95} as our cost function C . Our minimization of C is informal: we compute fluxes for each g -point individually and choose the M g -points which make the worst proxies as that first member of each team. For each remaining team member we process teams randomly and choose the remaining g -point that minimizes the cost function for the provisional team. It is likely that the balance of teams could be modestly improved though further optimization (by simulated annealing following *Kirkpatrick et al.* [1983], for example).

[17] Fluxes can then be computed by choosing one of the teams at random:

$$F(x, y, z, t) \approx M \sum_{g \in \{1..M\}} w_g F_g(x, y, z, t). \quad (4)$$

[18] Because each g -point is included in exactly one team equation (4), like equation (3), is an unbiased estimate of the true flux given by equation (1).

[19] Using teams of g -points is effective in limiting the error for a given computational cost (Figure 2). Though the teams are chosen to minimize errors in clear-sky surface fluxes the errors for all skies are commensurate, slightly lower in the shortwave where the presence of clouds simply reduces downwelling flux and slightly larger in the longwave where clouds both increase and change the spectral distribution of downwelling flux.

[20] Teams constructed in this way are more efficient in reducing error as team size increases. Figure 3 shows the ratio of errors obtained using teams of a given size (equation (4)) to those using Monte Carlo samples (equation (3)). The ratio is small in both the shortwave and longwave for $m = 2$ but increases at four or more samples, such that random sampling using equation (3) can achieve commensurate accuracy only by increasing sample sizes by a factor of 10 or greater—in other words, by increasing computational cost to nearly that of broadband integration using equation (1).

3. PSrad/RRTMG: A New Radiation Code for Climate Models

[21] We have developed a new radiation package, PSrad/RRTMG (so named because it is a postscript to the RRTMG package), designed for use in models of the atmosphere. The longwave and shortwave components are organized along parallel lines: driver modules call routines to compute the optical properties of gases, aerosols, and clouds, then combine these to compute fluxes at the boundaries and the interfaces between model layers. The codes are modeled after the RRTMG

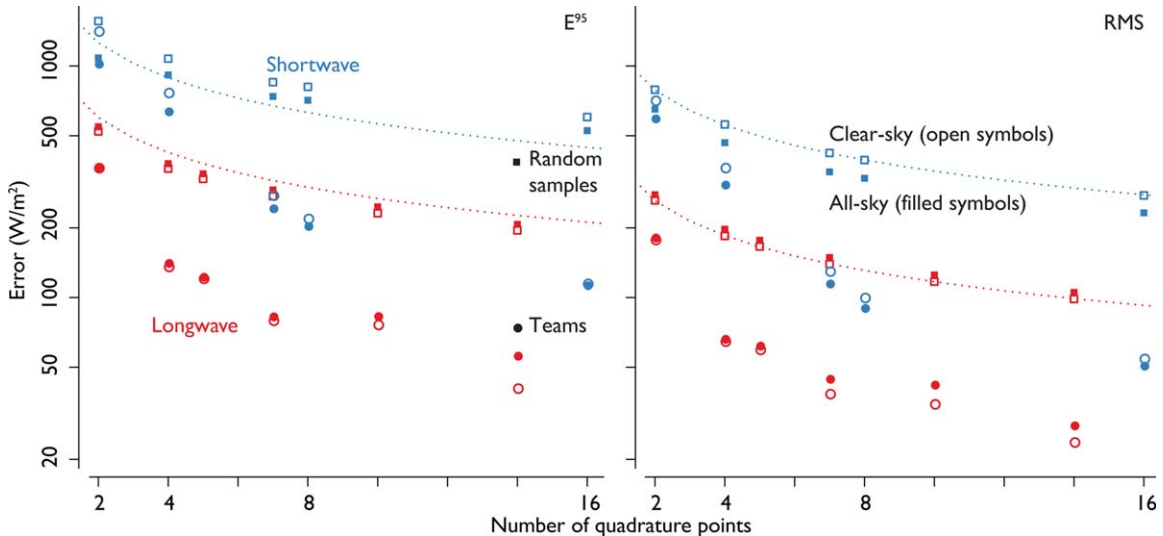


Figure 2. Errors in shortwave (blue) and longwave (red) surface fluxes for various spectral sampling strategies as a function of the number of quadrature points used. Errors are accumulated over roughly 73000 sample columns representing four snapshots from a single day of ECHAM. (right) RMS error decreases as $1/\sqrt{n}$ when samples are chosen randomly (equation (3), squares) though (left) E^{95} does not, but both decrease much more quickly using “teams” of g -points chosen to minimize E^{95} (equation (4), circles) than as compared to calculations using the same number of quadrature points, but chosen at random. Errors for all-sky fluxes (closed circles) are marginally smaller in the shortwave and larger in the longwave than for the clear skies used to optimize the team members because clouds decrease SW surface and increase LW fluxes.

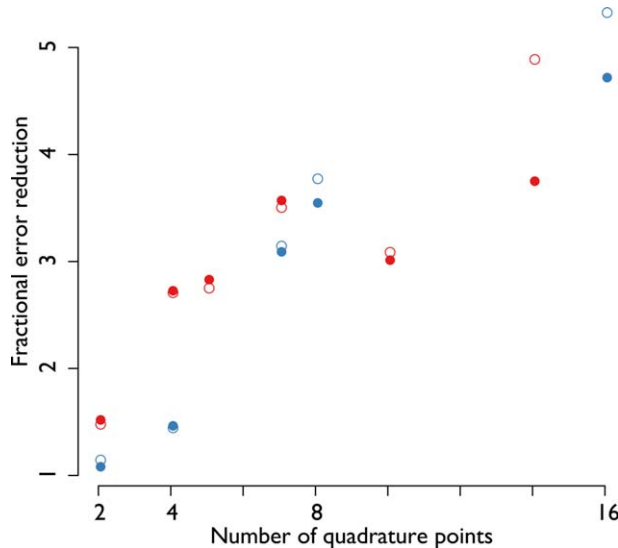


Figure 3. Ratio of 95th percentile error E^{95} in downwelling surface flux using teams of spectral intervals, compared to Monte Carlo sampling of spectral intervals, as a function of the number of samples. Improvements for clear-sky fluxes (open circles) and all-sky fluxes (closed circles) are commensurate. At small sample sizes error in the longwave fluxes (red) is reduced more dramatically by using teams than are shortwave fluxes (blue). The error reduction provided by teams increases with the number of samples but even teams of modest sizes (~ 8 members) provide accuracy comparable to ~ 10 times as many random samples.

package [Mlawer et al., 1997; Iacono et al., 2008] and use the k -distribution from this package to determine gas optical thickness from concentrations, temperature, and humidity. (These k -distributions are well validated and among the most accurate available; see Oreopoulos et al. [2012].) We follow the original RRTMG codes in using the two-stream approximation [after Meador and Weaver, 1980] to compute layer reflectance and transmittance and adding [after Oreopoulos and Barker, 1999] to compute flux profiles in the shortwave; we use the linear-in-tau approximation for the thermal source function [Mlawer et al., 1997] and consider only emission and absorption in the longwave. Cloud and aerosol optical properties are determined from custom-built lookup tables (S. Kinne, et al., A new global aerosol climatology for climate studies, submitted to *Journal of Advances in Modeling Earth Systems*, 2013).

[22] Subgrid-scale variability is treated using “subcolumns” [see, e.g., Räisänen et al., 2004; Pincus et al., 2006]: discrete random samples, each treated as internally homogeneous, that are consistent with the distributions of possible cloud states within each column, including fractional cloudiness in each layer and assumptions about the vertical correlations between layers (“cloud overlap”). This treatment is a generalization of the Monte Carlo Independent Column Approximation [McICA, see Pincus et al., 2003] and may be further generalized to include other kinds of variability, including

the distribution of cloud liquid or ice water content as implied by, for example, the *Tompkins* [2002] cloud scheme.

[23] PSrad supports a range of choices for spectral sampling, including broadband integration (all G quadrature points in order), arbitrary numbers of randomly chosen quadrature points for application of equation (3), and a number of predetermined “leagues” of g -point teams as described in section 2.

[24] Though PSrad is currently intended as a drop-in replacement for RRTMG it was implemented almost entirely from scratch (of the original code, only the subroutine that computes longwave gas optical properties remains). The most important technical difference lies in the organization: each subroutine is vectorized over model columns, which increases computational efficiency on a wide range of platforms even when relatively few spectral intervals are used. Operational centers such as the European Centre for Medium-Range Weather Forecasts have often modified RRTMG in this way [Morcrette et al., 2008].

4. Assessing Approximation Impacts in a Global Model

[25] We have implemented PSrad in ECHAM6 [Stevens et al., 2013], a state-of-the-art atmospheric model used for climate simulations. We perform simulations with a version of the model differing modestly from the version used to produce data for the Fifth Coupled Model Intercomparison Project [Taylor et al., 2012]. The model is run at a horizontal resolution of T63 using 47 levels that extend to 1 hPa. These experiments use a 7.5 min time step.

[26] We consider an ensemble of 29 month-long integrations starting from initial conditions valid at 0Z on 1 April of the years 1976–2004 as simulated by the model in a long integration using specified, time-varying sea-surface temperatures. The benchmark is an integration in which broadband radiation fields are calculated just as frequently as the tendencies from other physical parameterizations (i.e. the radiation and physics time steps are the same). We use this reference forecast to assess the error introduced by increasing the interval between broadband radiation calculations, on the one hand, and by limiting the number of spectral quadrature points used at each time, on the other.

[27] RMS differences with the reference forecast grow with time (Figure 4), but can be divided into roughly three regimes: slow but accelerating error growth in the first 10 days, rapid error growth over the next 10 days, and roughly saturated errors in the last 10 days. This may be very loosely described as the transition from weather (where individual trajectories are followed) to climate (where the statistics of trajectories are of interest), which leads us to examine errors in the first 10 days as one might evaluate forecasts, but to assess errors in the last 10 days statistically, as one might assess climatologies.

[28] The Monte Carlo sampling of fractional cloudiness and overlap introduces noise into the fluxes and

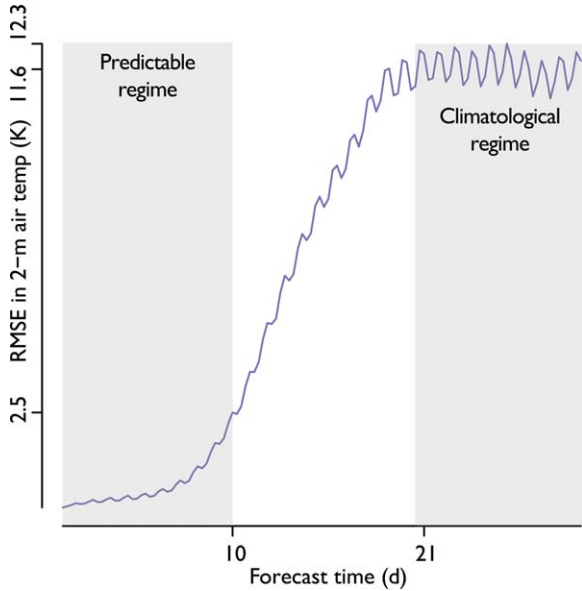


Figure 4. RMS difference in 2 m air temperature, relative to a reference forecast, as a function of forecast lead time, for a simulation making broadband radiation calculations every 2 h while all other physical processes use a 7.5 min time step. The forecasts diverge slowly during the first 10 days (the weather forecasting regime, roughly), so that approximations can be judged by computing point-by-point RMS differences with the reference. After roughly three weeks the forecasts are not predictable in a deterministic sense, so parameterization accuracy is judged by comparing distributions.

causes even a second independent reference forecast to diverge from the benchmark over time: RMS differences between two sets of reference forecasts rise from about 0.05 K during the first day to almost 1.5 K after 10 days (Figure 5, top, purple lines), which we take as the rough limit of deterministic forecasts. Increasing the sparsity of radiation calculations in either time (purple lines, $\Delta t_{\text{rad}} \in [15\text{min}, 1\text{h}, 2\text{h}, 3\text{h}]$) or spectral quadrature points (green lines, $m_{\text{LW}} \in [4, 10]$ and $m_{\text{SW}} \in [7, 16]$) increases the error by modest amounts in an absolute sense. No evidence has been found that such approximation trigger dramatic changes in the simulation [cf. *Pauluis and Emanuel, 2004*], though even hourly radiation calls increases the RMS difference by 50% of the difference introduced by sampling noise (Figure 5, bottom). The mean RMS difference over the first 10 days (Figure 6) is quite tightly related to computational cost (here expressed as the number of shortwave radiation computations per day).

[29] On time scales where the climatology of the model dominates the solution it is more informative to assess the degree to which each approximation produces a statistical distribution of temperatures consistent with the reference forecast. We apply the two-sided Student’s t -test, for each approximation at every grid point and every time step during days 21–31, to compute the likelihood (p -value) that the distribution of 2 m temperatures across the 29 ensemble members is statistically

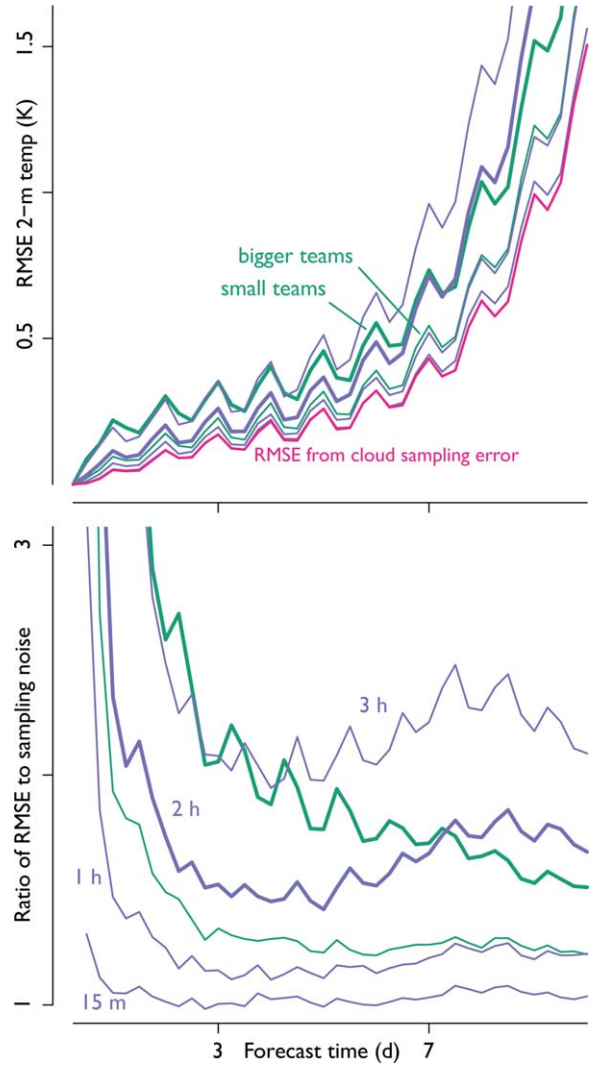


Figure 5. (top) RMS difference in 2 m air temperature, relative to a reference forecast, as a function of forecast lead time for two methods of coupling radiation and dynamics in a global model. Broadband computations applied sparsely in time at intervals ranging from 3 h to 15 min are shown in purple and two examples of “teams” of spectral intervals (as described in the text, using (7, 16) points in the shortwave and (4, 10) in the longwave) are in green. Differences between independent realizations of the reference forecast (in pink) show the differences to be expected from sampling errors in cloud state introduced by the Monte Carlo Independent Pixel Approximation (McICA). (bottom) Ratio of RMS difference of different coupling approximations to the RMS difference due to finite sampling of cloud states. All approximations change model evolution; larger approximations (fewer team members or less frequent sampling in time) degrade the sampling of the diurnal cycle.

indistinguishable at the 95% level between the reference forecast and the forecast using the approximation. Because we perform so many t -tests ($\sim 18,500$ per time step) roughly 5% have p -values corresponding to

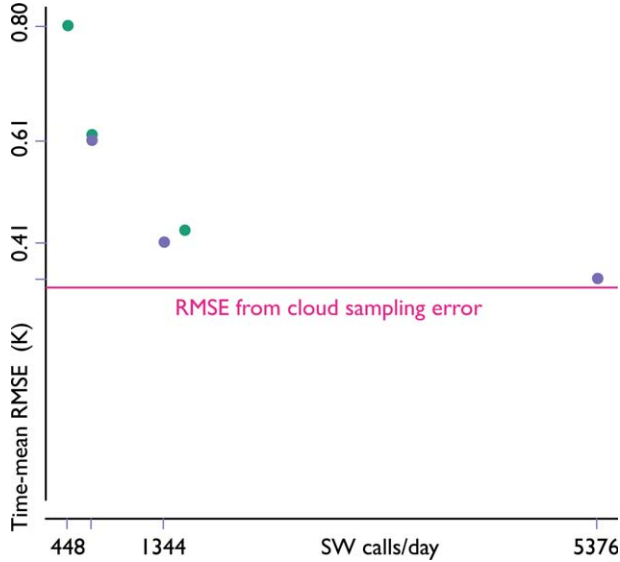


Figure 6. Time-mean error versus a rough measure of computational cost for two methods of coupling radiation and dynamics in a global model. Broadband computations applied sparsely in time are shown in purple and the “teams” of spectral intervals described in the text are in green. Error is measured as the mean over 10 days of the global RMS difference in 2 m air temperature relative to a reference forecast, and computational cost as the daily number of calls to the shortwave solver (which dominates the overall cost). The pink line shows the minimum achievable error (i.e., the error introduced by another realization of cloud states sampled in the Monte Carlo independent pixel approximation). Approximation errors for the two sampling strategies are commensurate for a given computational cost.

“significant” differences (at the 95% level) even if the underlying distribution of 2 m temperatures is the same in both experiments. The large number of false-positives can be reduced using false discovery rate estimation [e.g., *Wilks*, 2006] which exploits the known distribution of p -values expected under the null hypothesis to estimate η_0 , the proportion of uninteresting (or truly insignificant) p -values [*Strimmer*, 2008a], at every time step.

[30] Independent realizations of the reference run (i.e., two runs making broadband radiation calculations at every time step, but using different random number sequences to sample cloud states with McICA) are statistically indistinguishable: the time-mean value of η_0 computed from this pair of experiments is 1. This is almost but not quite true when comparing the reference run to any of the possible approximations (see Table 1). One interpretation of $\bar{\eta}_0$ is as the average fraction of the planet over which a given approximation does not change the simulation significantly. This fraction is greater than 94% for all approximations except MCSI (Table 1, line 1), indicating that changes are detectable but modest. The test statistic is not entirely robust: according to this measure radiation time steps of $T=1h=8\Delta t$ are slightly more

Table 1. Time Mean Fraction of Statistically Insignificant η_0 Differences in 2 m Air Temperature Between a Reference Calculation and Various Approximations for Coupling Radiation to a Global Model^a

Approximation	SW calls/day	$\bar{\eta}_0$
$m_{\text{SW}}=1, m_{\text{LW}}=1$	192	0.804
$m_{\text{SW}}=7, m_{\text{LW}}=4$	1344	0.943
$m_{\text{SW}}=16, m_{\text{LW}}=10$	3072	0.977
$T=15\text{ m}$	10,752	0.971
$T=1\text{ h}$	2688	0.985
$T=2\text{ h}$	1344	0.955
$T=3\text{ h}$	896	0.963

^aThe first three approximations are temporally dense, spectrally sparse calculations using equation (4) (the first is the special case using equation (3)). The latter four make spectrally dense broadband calculations at specified time intervals. All approximations change the simulation of 2 m air temperature by detectable amounts; for a given computational effort, frequent use of spectral teams introduces slightly larger changes than less frequent broadband calculations.

consistent with reference calculations than are $T=15\text{min}=2\Delta t$, which is physically implausible. Given this caveat we note that, in this relatively coarse-resolution model, infrequent broadband radiation calculations introduce slightly smaller changes compared to frequent calls using spectral teams with the same computational cost (cf. lines 2 and 6 of Table 1).

5. Conclusions: Parameterization Error, Simulation Error, and the Coupling of Radiative Transfer to Atmospheric Models

[31] Radiative fluxes respond nearly instantaneously to changes in the optical properties of the atmosphere, so the parameterization of radiation is normally considered a “one-way” problem in which the model provides the state of the atmosphere and the parameterization computes the heating rates and boundary fluxes. In the absence of coupling between radiation and model dynamics one naturally seeks instantaneous radiation calculations that are as accurate as is computationally feasible. Focusing on the accuracy of the overall simulation, including the way radiation calculations are coupled to the model, may allow for other kinds of optimization. As one example, the k -distribution developed for RRTMG was designed, as are most parameterizations, to minimize broadband flux errors. RRTMG aims to balance accuracy with computational cost primarily by minimizing the number of g -points. It would be possible, however, to construct k -distributions using different cost functions, and distributions constructed to minimize errors both across and among teams might be able to achieve greater aggregate accuracy by using more g -points while still limiting noise in surface fluxes. This would open the door to resolving the tension between overall accuracy, limited by of the number of g -points, and computational cost.

[32] Similarly, using spectrally sparse, temporally dense calculations provides a richer set choices in how radiation may be coupled to dynamical models to minimize biases. Both classes of approximations examined

here appear to slightly modify the distribution of temperatures, relative to a reference calculation, at quasi-climatological time scales, in contrast to the introduction of sampling noise in cloud properties [Pincus *et al.*, 2006; Barker *et al.*, 2008]. The modification is small but undesirable, and we are seeking ways to reduce the impact of noise. One promising approach is to sample teams in equation (4) with replacement in time, rather than entirely independently at each location and time.

[33] The results of section 4, especially Figure 6 suggests that computational effort is the primary determinant of accuracy in coupling these two radiative approximations to short-term forecasting models. This comparison is limited, however, since it does not account for true model errors or how such errors might depend on sampling strategy. Some real-life forecast errors, such as the damping of surface heating caused by convective clouds forming in response to initial surface heating, may well be due to under-sampled temporal variability. On the other hand, even modest noise in surface fluxes may lead to forecast errors when the coupling between radiation and atmospheric state is strong, as in nocturnal stable boundary layers. Thus it remains to be seen whether the statistical robustness of equation (4) will translate into improved weather forecasts.

[34] Spectrally sparse, temporally dense radiation calculations, at least as implemented here, disturb simulations with ECHAM6 at least as much as infrequent broadband calculations of the same computational cost. Time steps in ECHAM are relatively short so the comparison may be worse in models with longer time steps. Other considerations may make equation (4) desirable, however, especially the convergence of spectral teams with increasing resolution, more uniform distribution of computation time, and the book-keeping simplifications that arise when some estimate of radiation is computed every time step. In ECHAM, for example, shortwave fluxes are computed for all points (using a very small minimum solar zenith angle for nighttime points) so that temporal interpolation across the sunrise boundary is smooth. Replacing infrequent broadband calculations with frequent spectral samples makes this transition smooth (in aggregate), so that shortwave fluxes are not required at nighttime points, which alone represents a substantial computational savings.

[35] Though radiation ultimately determines earth's climate the coupling between radiative fluxes and the rest of the atmosphere is loose, such that radiation strongly influences the flow only where its effects can accumulate over time, as occurs in descending branches of the general circulation or at the tops of stratocumulus clouds. The approach to radiation calculations proposed here exploits this loose coupling, trading instantaneous accuracy at infrequent intervals for statistical accuracy with more complete sampling of time variability. By traditional measures of error (e.g., the comparison of instantaneous fluxes to benchmark calculations, as in Oreopoulos *et al.*, [2012]), large instantaneous errors make spectral sampling a poor idea. We argue that a more appropriate benchmark is the accumulated effect of approximation errors on the solution as a whole. By

this more holistic measure of accuracy, frequent but sparse sampling becomes much more attractive because the loose coupling of radiation to the flow means that unbiased solutions with large local errors are more desirable than solutions with small biases, even if their local errors are also small. The present work demonstrates a new path toward accuracy that, in some situations, may converge more quickly to the desired solution.

[36] **Acknowledgments.** This work was supported by the Max Planck Society for the Advancement of Science, the National Science Foundation Science and Technology Center for Multi-Scale Modeling of Atmospheric Processes, managed by Colorado State University under cooperative agreement No. ATM-0425247, and by the Office of Naval Research under grant N00014-11-1-0441. Generous computing facilities were provided by the German Climate Computing Center (Deutsches Klimarechenzentrum, DKRZ). We are grateful to the developers of the *fdrtool* software [Strimmer, 2008b] for making our field significance calculations easy. RP appreciates warm hospitality during summer visits to the MPI and practical advice from Thorsten Mauritsen on the care and feeding of ECHAM.

References

- Barker, H. W., J. N. S. Cole, J.-J. Morcrette, R. Pincus, P. Røed, K. von Salzen, and P. A. Vaillancourt (2008), The Monte Carlo Independent Column Approximation: An assessment using several global atmospheric models, *Q. J. R. Meteorol. Soc.*, *134*(635, Part B), 1463–1478, doi:10.1002/qj.303.
- Buizza, R., M. Miller, and T. N. Palmer (1999), Stochastic representation of model uncertainties in the ECMWF ensemble prediction system, *Q. J. R. Meteorol. Soc.*, *125*(560), 2887–2908, doi:10.1002/qj.49712556006.
- Chevallier, F., F. Chéruy, N. A. Scott, and A. Chédin (1998), A neural network approach for a fast and accurate computation of a longwave radiative budget, *J. Appl. Meteorol.*, *37*(11), 1385–1397, doi:10.1175/1520-0450(1998)037<1385:ANNAFA>2.0.CO;2.
- Eckermann, S. D. (2012), Explicitly stochastic parameterization of non-orographic gravity wave drag, *J. Atmos. Sci.*, *68*(8), 1749–1765, doi:10.1175/2011JAS3684.1.
- Evans, K. F., and A. Marshak (2005), Numerical methods, in *3D Radiative Transfer in Cloudy Atmospheres*, edited by A. Marshak and A. B. Davis, pp. 243–282, Springer, New York.
- Fu, Q., and K. N. Liou (1992), On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres, *J. Atmos. Sci.*, *49*, 2139–2156, doi:10.1175/1520-0469(1992)049<2139:OTCDMF>2.0.CO;2.
- Hill, P. G., J. Manners, and J. C. Petch (2011), Reducing noise associated with the Monte Carlo independent column approximation for weather forecasting models, *Q. J. R. Meteorol. Soc.*, *147*(654), 219–228, doi:10.1002/qj.732.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins (2008), Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *J. Geophys. Res.*, *113*, D13103, doi:10.1029/2008JD009944.
- Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi (1983), Optimization by simulated annealing, *Science*, *220*(4598), 671–680, doi:10.1126/science.220.4598.671.
- Krasnopolsky, V. M., M. S. Fox-Rabinovitz, and A. A. Belochitski (2008), Decadal climate simulations using accurate and fast neural network emulation of full, longwave and shortwave, radiation, *Mon. Weather Rev.*, *136*(10), 3683–3695, doi:10.1175/2008MWR2385.1.
- Lacis, A. A., and V. Oinas (1991), A description of the correlated k-distribution method for modeling non-grey gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres, *J. Geophys. Res.*, *96*, 9027–9063.
- Lott, F., L. Guez, and P. Maury (2012), A stochastic parameterization of non-orographic gravity waves: Formalism and impact on the equatorial stratosphere, *Geophys. Res. Lett.*, *39*, L06807, doi:10.1029/2012GL051001.
- 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(639, Part B), 457–468, doi:10.1002/qj.385.
- Mauritsen, T., et al. (2012), Tuning the climate of a global model, *J. Adv. Model. Earth Syst.*, 4, M00A01, doi:10.1029/2012MS000154.
- Meador, W. E., and W. R. Weaver (1980), Two-stream approximations to radiative transfer in planetary atmospheres: A unified description of existing methods and a new improvement, *J. Atmos. Sci.*, 37(3), 630–643, doi:10.1175/1520-0469(1980)037<0630:TSATRT>2.0.CO;2.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough (1997), RRTM, a validated correlated-k model for the longwave, *J. Geophys. Res.*, 103, 16,663–16,682, doi:10.1029/97JD00237.
- Mlawer, E. J., P. D. Brown, S. A. Clough, L. C. Harrison, J. J. Michalsky, P. W. Kiedron, and T. Shippert (2000), Comparison of spectral direct and diffuse solar irradiance measurements and calculations for cloud-free conditions, *Geophys. Res. Lett.*, 27(17), 2653–2656, doi:10.1029/2000GL011498.
- Morcrette, J.-J. (2000), On the effects of the temporal and spatial sampling of radiation fields on the ECMWF forecasts and analyses, *Mon. Weather Rev.*, 128(3), 876–887, doi:10.1175/1520-0493(2000)128<0876:OTEOTT>2.0.CO;2.
- Morcrette, J.-J., H. W. Barker, J. N. S. Cole, M. J. Iacono, and R. Pincus (2008), Impact of a New Radiation Package, McRad, in the ECMWF integrated forecasting system, *Mon. Weather Rev.*, 136(12), 4773–4798, doi:10.1175/2008MWR2363.1.
- Oreopoulos, L., and H. W. Barker (1999), Accounting for subgrid-scale cloud variability in a multi-layer 1D solar radiative transfer algorithm, *Q. J. R. Meteorol. Soc.*, 125(553, Part A), 301–330, doi:10.1002/qj.49712555316.
- Oreopoulos, L., et al. (2012), The continual intercomparison of radiation codes: Results from phase I, *J. Geophys. Res.*, 117, D06118, doi:10.1029/2011JD016821.
- Pauluis, O., and K. Emanuel (2004), Numerical instability resulting from infrequent calculation of radiative heating, *Mon. Weather Rev.*, 132, 673–686, doi:10.1175/1520-0493(2004)132<0673:NIRFIC>2.0.CO;2.
- Pincus, R., and B. Stevens (2009), Monte Carlo spectral integration: A consistent approximation for radiative transfer in large eddy simulations, *J. Adv. Model. Earth Syst.*, 1, 1, doi:10.3894/JAMES.2009.1.1.
- Pincus, R., H. W. Barker, and J.-J. Morcrette (2003), A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields, *J. Geophys. Res.*, 108, 4376, doi:10.1029/2002JD003322.
- Pincus, R., R. S. Hemler, and S. A. Klein (2006), Using stochastically generated subcolumns to represent cloud structure in a large-scale model, *Mon. Weather Rev.*, 134(12), 3644–3656, doi:10.1175/MWR3257.1.
- Räisänen, P., and H. W. Barker (2004), Evaluation and optimization of sampling errors for the Monte Carlo independent column approximation, *Q. J. R. Meteorol. Soc.*, 130(601), 2069–2085, doi:10.1256/qj.03.215.
- Räisänen, P., H. W. Barker, M. F. Khairoutdinov, J. Li, and D. A. Randall (2004), stochastic generation of subgrid-scale cloudy columns for large-scale models, *Q. J. R. Meteorol. Soc.*, 130(601), 2047–2067, doi:10.1256/qj.03.99.
- Stevens, B., et al. (2013), The atmospheric component of the MPI-M earth system model: ECHAM6, *J. Adv. Model. Earth Syst.*, doi:10.1256/qj.06.12, in press.
- Strimmer, K. (2008a), A unified approach to false discovery rate estimation, *BMC Bioinformatics*, 9(1), 303, doi:10.1186/1471–2105-9–303.
- Strimmer, K. (2008b), fdrtool: A versatile R package for estimating local and tail area-based false discovery rates, *Bioinformatics*, 24(12), 1461–1462.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, 93(4), 485–498, doi:10.1175/BAMS-D-11–00094.1.
- Tompkins, A. M. (2002), A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover, *J. Atmos. Sci.*, 59(12), 1917–1942, doi:10.1175/1520-0469(2002)059<1917:APPFTS>2.0.CO;2.
- Turner, D. D., et al. (2004), The QME AERI LBLRTM: A closure experiment for downwelling high spectral resolution infrared radiance, *J. Atmos. Sci.*, 61(22), 2657–2675, doi:10.1175/JAS3300.1.
- Venema, V., A. Schomburg, F. Ament, and C. Simmer (2007), Two adaptive radiative transfer schemes for numerical weather prediction models, *Atmos. Chem. Phys.*, 7(21), 5659–5674, doi:10.5194/acp-7-5659-2007.
- Wilks, D. S. (2006), On “field significance” and the false discovery rate, *J. Appl. Meteorol. Climatol.*, 45(9), 1181–1189, doi:10.1175/JAM2404.1.
- Xu, K. M., and D. A. Randall (1995), Impact of interactive radiative transfer on the macroscopic behavior of cumulus ensembles. Part I: Radiation parameterization and sensitivity tests, *J. Atmos. Sci.*, 52, 785–799, doi:10.1175/1520-0469(1995)052<0785:IOIRTO>2.0.CO;2.

Corresponding author: Robert Pincus, NOAA Earth System Research Lab, Physical Sciences Division, University of Colorado, 325 Broadway R/PSD1, Boulder, CO 80305, USA. (Robert.Pincus@colorado.edu)