From Model-Based Parameterizations to Lookup Tables: An EOF Approach

GIOVANNI LEONCINI
Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

ROGER A. PIELKE SR.
Cooperative Institute for Research in Environmental Sciences, Program in Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado

PHILIP GABRIEL
Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

(Manuscript received 4 April 2007, in final form 10 February 2008)

ABSTRACT

The goal of this study is to transform the Harrington radiation parameterization into a transfer scheme or lookup table, which provides essentially the same output (heating rate profile and short- and longwave fluxes at the surface) at a fraction of the computational cost. The methodology put forth here does not introduce a new parameterization simply derived from the Harrington scheme but, rather, shows that given a generic parameterization it is possible to build an algorithm, largely not based on the physics, that mimics the outcome of the parent parameterization. The core concept is to compute the empirical orthogonal functions (EOFs) of all of the input variables of the parent scheme, run the scheme on the EOFs, and express the output of a generic input sounding exploiting the input–output pairs associated with the EOFs. The weights are based on the difference between the input and EOFs water vapor mixing ratios. A detailed overview of the algorithm and the development of a few transfer schemes are also presented. Results show very good agreement ($r > 0.91$) between the different transfer schemes and the Harrington radiation parameterization with a very significant reduction in computational cost (at least 95%).

1. Introduction

This study aims to provide an algorithm, or transfer scheme (TS), that can accurately reproduce the output of a generic parameterization—in this specific case the Harrington radiation scheme (Harrington 1997; Harrington et al. 1999)—using a fraction of the computational power required by the parent parameterization. The background of this work has been laid out in Pielke et al. (2006) and is briefly summarized here.

Most numerical weather prediction (NWP) models numerically solve the equations of motion, but also use parameterizations to account for subgrid-scale processes (e.g., turbulence), short- and longwave radiative flux divergence, and other processes that cannot be explicitly simulated within the dynamical core that accounts for the pressure gradient, Coriolis effect, advection, and mass continuity. Land surface interactions are also among the parameterized processes.

Because of computational constrains and limited physical knowledge, parameterizations always involve tunable coefficients (unlike the dynamical core) and are based on approximations. Furthermore most times they are strictly 1D, in the sense that they use values from only one grid column at a time.

Regardless of accuracy, parameterizations represent a considerable share of the total computational burden. Majewski et al. (2002) showed that for high-resolution global NWP models, parameterizations account for 46.8% of the total computational cost while radiation parameterizations make up 57.5% of that number. Similarly Chevallier et al. (1998) reports that the longwave radiation scheme accounts for 10% and 18% of
the total computing time required by, respectively, the
general circulation model at the European Centre for
Medium-Range Weather Forecasts and the climate
model of the Laboratoire de Météorologie Dynamic.
Tests conducted during this study show that the Regional
Atmospheric Modeling System (RAMS) has similar
levels of performance and the Harrington radiation pa-
rameterization occupies 13% of the total CPU time.

Therefore, there would be great gain if a lookup
table (LUT), or a TS, could accurately reproduce the
outputs of a parameterization at a fraction of the com-
putational cost.

The main problem in developing a TS is that a simple
brute force approach would require running the parent
parameterization for all of the possible input cases it
can possibly be given during a simulation. The number
of combinations of all the possible values, aside from
being impractically large, would grossly misrepresent
reality since not every combination is physically pos-
sible. For example, we know that superadiabatic pro-
files of temperature can generally only occur immedi-
ately above a surface such as the ground or a cloud top.
There are also larger-scale constraints such as the gra-
dient wind balance at the synoptic scale in the mid- and
high latitudes. Moreover, parameterizations are very
often a limited representation of reality due to the set
of simplifying assumptions. Therefore a two-stream ra-
diation parameterization, such as the Harrington
scheme, cannot be expected to reproduce the variabil-
ity of the heating rates as a line-by-line code would do,
let alone as nature does. On the contrary, parameter-
izations or “physics packages” must be thought of as
the best possible compromise between the available
computational resources and the accuracy desired. Be-
cause of the limited representation of reality that these
parameterizations offer, their outputs (i.e., physical
variables) are likely to depend on and/or be sensitive to
a more limited number of physical parameters and
quantities as compared to the behavior that the same
physical variables have in nature. The aim of this study
is twofold: first, to show that it is possible to mimic the
behavior of parameterizations at a fraction of their
computational cost; second, this computational speed
increase can be achieved by the use of techniques/
algorithms that are not strictly related to the physics
that parameterizations attempt to represent. The
choice of the Harrington scheme is due mainly to two
factors: its relatively high computational cost and the
limited number of input variables that render it more
easily tractable.

Pielke et al. (2006) suggested ways to reduce the in-
put space of a parameterization, which is usually much
larger than the output space, and in this study the em-
pirical orthogonal functions (EOFs) have been em-
ployed. The two main reasons for this choice are 1) its
ability to identify key “patterns” among all the realistic
inputs to the parent scheme, and 2) the fact that every
input can be expressed as a linear combination of such
patterns, thereby greatly reducing the dimensionality of
the problem. This is possible because the key patterns
(i.e., the EOFs) are obtained through the eigenvalues
and eigenvector of the correlation matrix of the data.
This property of the EOF analysis (or principal com-
ponent analysis) has been exploited widely for decreas-
ing the number of estimators in statistical estimation
(e.g., Davis 1976; Lorenz 1956, 1977). A second benefit
of the EOF analysis is that the physical significance of
the patterns themselves is ranked by the percentage of
the variance of the data they explain, in a linear sense.
This feature has also been exploited widely for several
other applications: for example, to study several aspects
of the large-scale circulation such as the North Ameri-
can monsoon (e.g., Castro et al. 2007b), the Northern
Hemisphere mean flow (Kravtsov et al. 2006), the Arct-
ic Oscillation (Thompson and Wallace 2000), the pre-
dictability of seasonal means (Schubert et al. 2002),
the relation between surface-level humidity and column-
integrated water vapor (Liu et al. 1991), the impact of
microphysics parameterization on a cloud property re-
trieval algorithm (Biggerstaff et al. 2006), and to ap-
proximate the difference in the reflectance of the O$_2$
band obtained from a multiscattering line-by-line code
and a two-stream representation (Natraj et al. 2005).

Furthermore, the EOF analysis has been used widely
not only recently, but since Lorenz (1956) first brought
it to the attention of the atmospheric science commu-
nity, and its sampling errors are well known especially
after the studies of North et al. (1982), von Storch and
Hannoschöck (1985), and more recently Quadrelli et al.
(2005). A detailed description of EOFs and their utility
in the atmospheric sciences can be found in Wilks (2006).

The patterns identified by the EOF analysis are es-
ential for this study since the core concept of the TS is
to compute the EOFs of all the inputs to the Harrington
scheme (HS), obtain the heating rates, as well as the
other outputs associated with the EOFs, and then use
those input–output pairs to approximate the HS out-
puts for every input.

While the words “transfer scheme” better character-
ize the algorithm development herein, traditionally
similar algorithms, or parts of algorithm, have been
called lookup tables, and both nomenclatures are used
interchangeably here.

The details of the algorithm are explained in section
2, while section 3 describes the results of the accuracy
and speed tests, and finally the description and discus-
sion of the results obtained embedding the TS in RAMS are given in section 4.

2. Methodology

a. Overview

The EOFs of the HS input variables (surface albedo, the upwelling longwave radiative flux at the surface, and the vertical profiles of pressure, temperature, and water vapor mixing ratio) are computed for a particular location and a specific time of the day under clear-sky conditions. Then, the EOFs (or, alternatively, the regression of the data onto the principal components) are fed to the offline version of the HS, which, for each individual EOF, provides downwelling surface radiative fluxes and a vertical profile of radiative heating rates. Each EOF is associated with the day exactly at the center of the period over which it has been calculated. When input–output pairs are established, they are used to obtain a synthetic HS output for a generic input at that specific location and time of the day. This process is carried out at each model grid column, and no assumption is made concerning horizontal spatial variations. The same calculation is performed at eight different times of the day every 3 h, starting at 0000 UTC, and the LUT output at intermediate times is obtained by linear interpolation. Because the EOFs are calculated offline prior to the simulation, the only part of this process that must be carried out within a NWP model is the generation of the synthetic output. The TS output will be a weighted average of the output of each individual EOF. The weights are, in turn, determined by an arbitrarily defined “distance” between the EOFs and the input itself, at each grid point. This process is represented schematically in Fig. 1. It is also important to remember that the aim of this study is merely to achieve a reduction in speed, and no physical interpretation of the EOFs or of their significance will be attempted.

Different weighting strategies were tried offline using an independent month of RAMS output (September 2005) and the most successful one was then implemented into RAMS and compared against the parent scheme for a 2-day simulation during September 2005. A few key meteorological fields are then compared from simulations that used the parent scheme, the best TS, and the Chen–Cotton radiation parameterization (Chen and Cotton 1987).

The only difference between the version of the Harrington scheme applied on the EOFs and the version actually used in RAMS is in how the air density is treated. The online (or in RAMS) version uses the background state density, which is constant through each individual run, while the offline version computes
the density from the temperature and pressure profiles, which constitute part of the EOFs, in order to have the thermodynamic quantities as balanced as possible. The differences between the two versions are one order of magnitude smaller than the error introduced by the LUT and thus they have been deemed negligible.

b. The EOFs

Because the EOFs are meant to represent all the possible inputs that a clear-sky atmosphere can provide to the HS, they were first computed using rawinsonde data for a particular station. Unfortunately, the significance of the EOFs at each site varies with the different location histories and missing data periods. To avoid this inconvenience and to take full advantage of the temporal and spatial consistencies that a NWP can offer, the EOFs were computed using the output of series of ad hoc simulations of the RAMS, which has already proved successful in this type of application as shown by Castro et al. (2007a,b).

The choices of the period of the year over which to compute the EOFs and how many years of data to use is a compromise mainly between two competing factors: on one hand the necessity of a small number of EOFs to have a faster algorithm, and on the other the necessity of having a sufficient number of EOFs to explain the variability of the input soundings. Also considering the computational constraints and the fact that each EOF is associated with a fixed day used to compute the zenith angle, along with hour and location, it was decided to split the year into its 12 months. However, results from 1 month only are presented here, since the present study aims to show the feasibility of the LUT concept and the same implementation can be easily carried out for the remaining periods. The effectiveness of the EOFs in building the TS is affected by the day of year associated with the EOFs, because if the corresponding zenith angle is very different from the zenith angle of the profiles input to the TS, the shortwave flux yielded by the TS will be unrealistic. To show that this is not detrimental to the TS, the month chosen for this study is September, which belongs to a transition season, during which the daily range of values of zenith angle at a grid point changes the most during the month.

As has already been stated, this work does not aim to provide a physically meaningful interpretation of the EOFs, but it has been assumed that a more realistic set of EOFs would yield a more realistic set of associated heating rates and surfaces fluxes and therefore lead to a more accurate LUT. Pursuing this line of reasoning, several EOFs are retained, explaining up to 97.5% of the variance. Furthermore, the model climatology data have neither been filtered nor averaged, and the EOFs have been rotated using the varimax method (Kaiser 1958) after scaling the eigenvectors so that each EOF has unity variance, but the principal components are mutually uncorrelated and the EOFs are not orthogonal to each other, thus avoiding unphysical features introduced by this last mathematical constrain (Wilks 2006).

To compute the EOFs for September, model-generated data from 10 Septembers (1995–2004) were used resulting in 300 temporal realizations of the sample because the EOFs are computed at eight times (from 0000 to 2100 UTC every 3 h) using data at the same time of the day (one sample per day per 30 days per 10 months) at each grid point of the RAMS domain. According to Quadrelli et al. (2005), this number is not enough to properly resolve all of the eigenvectors and eigenvalues, especially for the high-order EOFs. This issue, in principle, is even more important in this study, since the TS involves EOFs for up to 97.5% of the total explained variance. Unfortunately, the number of temporal realizations necessary to reduce the problem significantly for so many EOFs (up to 44) is too large for any practical use, and the sample size was arbitrarily chosen. The RAMS simulations all share the same configuration consisting of one grid covering the United States east of the Rocky Mountains (Fig. 2). The radiation scheme was Harrington’s (Harrington 1997; Harrington et al. 1999), a two-stream parameterization, and was called every 20 min. Further details can be found in section 3b(1).

Although the ultimate measures of the TS value are its accuracy and computational cost, not the signifi-
Table 1. Characteristics of the different transfer schemes.

<table>
<thead>
<tr>
<th>TS name</th>
<th>No. of yr</th>
<th>Rotated EOFs</th>
<th>Filter width (days)</th>
<th>Variance retained (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS10RF00</td>
<td>10</td>
<td>Yes</td>
<td>0</td>
<td>97.5</td>
</tr>
<tr>
<td>TS10RF11</td>
<td>10</td>
<td>Yes</td>
<td>11</td>
<td>97.5</td>
</tr>
<tr>
<td>TS10RF21</td>
<td>10</td>
<td>Yes</td>
<td>21</td>
<td>97.5</td>
</tr>
<tr>
<td>TS10RF31</td>
<td>10</td>
<td>Yes</td>
<td>31</td>
<td>97.5</td>
</tr>
<tr>
<td>TS10RF00v75</td>
<td>10</td>
<td>Yes</td>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>TS10RF00v50</td>
<td>10</td>
<td>Yes</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>TS10NRF00</td>
<td>10</td>
<td>No</td>
<td>0</td>
<td>97.5</td>
</tr>
<tr>
<td>TS9RF00</td>
<td>9</td>
<td>Yes</td>
<td>0</td>
<td>97.5</td>
</tr>
</tbody>
</table>

c. The weighting strategy

The output of the TS is a weighted average of the EOFs outputs, and the computation of the weights constitutes the core of the TS.

The orthogonality property of the unrotated EOFs can be exploited to compute the weights for each EOF, as it guarantees that, if the input sounding belongs to the same population that generated the EOFs, there exists one, and only one, linear combination of EOFs that is equal to the input sounding. Unfortunately, despite its simplicity and elegance, this method cannot effectively be applied, because not all the EOFs generated are orthogonal, even using double precision. This happens because, for the cases analyzed, over half of the EOFs (roughly 60) are effectively identical, resulting in a nonorthogonal EOF matrix consistently with the above-mentioned analysis of Quadrelli et al. (2005).

Further weighting strategies have been tested, but all were inversely related to the relative or absolute difference between the input sounding variables and the corresponding EOF variables. At first the weights were determined not only for each EOF, but also at each vertical level, but those strategies did not perform as well as the strategies that produced one weight per EOF. Most likely, this is due to the implicit assumption that heating rates at different levels are independent of the EOFs, they constitute the core of the TS, and to investigate how it is affected by the EOFs, six TSs have been implemented with different sets of EOFs (see Table 1). The control, TS10RF00, is based on rotated, unfiltered EOFs computed using the whole 10-yr model climatology. The number of EOFs retained for each grid point corresponds to a total explained variance of 97.5%. Because the number of retained EOFs greatly affects the TS speed, TS10RF00v75 and TS10RF00v50 used the same set of EOFs, but only up to 75% and 50% of the total explained variance, respectively. Time filtering the data results in fewer and more significant EOFs, so TS10RF11, TS10RF21, and TS10RF31 have been built filtering the data with a Hamming low-pass filter (Oppenheim et al. 1999) with windows of 11, 21, and 31 days. The same EOFs as in TS10RF00, but nonrotated, have been used for TS10NRF00, and the TS9RF00 EOFs are rotated and unfiltered, as in TS10RF00, but based on a 9-yr model climatology, from 1996 to 2004. The name of each TS is related to the characteristics of the EOFs; the first number (10 or 9) refers to the length of the climatology, and R (NR) stands for rotated (nonrotated) EOFs. The letter “F” and the following number indicate the filter window in days: 0 for unfiltered, 11 for an 11-day low-pass filter, etc.

Finally, the EOFs and therefore the TSs are computed only for clear sky to reduce the number of variables involved in the calculations and to have faster turnaround times. A model output column is defined as clear sky if its total cloud water mixing ratio is less than $10^{-8}$ kg kg$^{-1}$. The point is defined as cloudy otherwise. This value has been chosen because reducing it would not decrease the number of clear-sky grid points as depicted in Fig. 3.

RAMS model configuration

RAMS, version 4.3, was used for all the numerical simulations in this study. Most of the choices made were driven by the necessity of using a standard NWP type of configuration, in order to emphasize the usefulness of the LUT for common applications without great computational costs. The basic configuration consists of one grid only of $68 \times 67$ points, with grid spacing of 40 km. The vertical grid has 33 atmospheric levels, and its spacing is stretched with a ratio of 1.12 starting from 50 m up to a maximum of 1500 m. The first level is roughly 24 m AGL while the model top is at 15 800 m. The time step is set to 1 min. The North American Regional Reanalysis (Mesinger et al. 2006) was used to provide initial and lateral boundary conditions, and its total water mixing ratio, horizontal wind, and potential temperature were nudged at the top model level using the Newtonian relaxation method. As mentioned earlier, radiative fluxes were represented by the Harrison scheme and version 2 of the Land–Ecosystem–Atmosphere Feedback model (LEAF2; Walko et al. 2000) parameterized the surface fluxes of heat and moisture. Diffusion was parameterized according to the anisotropic scheme of Smagorinsky (1963). The radiation parameterization (Harrington’s) is called every 20 min. Finally, the convection scheme was the modified version of the Kain–Fritsch convection scheme (Kain and Fritsch 1993; Castro et al. 2002; Castro 2005). The only cloud process allowed was liquid condensation.
from each other. The sum of the difference, as opposed to the absolute value of the difference, has proven to be a poor measure of the similarity between the input sounding and the EOFs, because for a specific input sounding–EOF pair, mixing ratios that are larger for the sounding at a particular vertical level can be compensated by smaller values at different levels. Thus, two EOFs can have very similar weights even though one matches almost perfectly the vertical profile of the water vapor mixing ratio and the other oscillates around the input sounding. The weights for the input sounding $s$, associated with each EOF $e$, which resulted in the best correlation and root-mean-square error (RMSE) between TS and HS outputs, were determined by trial and error and are given by

$$w_s^e = \left( \sum_k |r_k^e - r_k^s| \right)^{-20},$$

where the subscript $k$ indicates the vertical level. The high value of the exponent is necessary to cause the weights of the EOFs, which are very different from the input sounding, to go rapidly to zero. A smaller absolute value exponent would excessively weight the EOFs, which are more distant from the input sounding, producing unrealistic heating rates and surface fluxes. On the other hand, a larger absolute value of the exponent fails to properly weight the EOFs, which are closer to the input sounding, resulting again in unrealistic values of the heating rates and surface fluxes. Because of the very good code optimization at compilation, changing the exponent does not change the overall execution times. Exponentially decaying weights have also been tested, but the sensitivity of the weights to the value of the exponent is much stronger, making the tuning process more difficult and probably location dependent. The weights, which can be only zero or positive, are then normalized at each grid point (i.e., for each input profile) and are the same for all the HS output variables (heating rates, surface short- and longwave downwelling radiation fluxes).

From the tests performed using different weighting strategies, a few conclusions can be drawn:

- A bulk weight per column results in improved correlations when compared against individual weights per grid point, confirming the integral character of the radiative transfer.
- Although all the input variables for the HS scheme (i.e., pressure, water vapor mixing ratio, and temperature profiles, upwelling longwave radiation, and albedo) have been used in the computation of the EOFs, mixing ratio alone works better than any other input variable or combination of variables, even when the combination includes the mixing ratio itself.

![Fig. 3. Percentage of clear-sky grid points against the mixing ratio threshold.](image)
• If the weights computed as described are multiplied by the fractional explained variance and then normalized, the overall performance of the algorithm changes only slightly.

d. Tests

To better describe the performance of the different LUTs implemented here, two kinds of accuracy tests were conducted. First, for each of the TSs, a 2-day simulation was carried out applying the HS at the cloudy points and computing both the TS and the HS at the clear-sky points, but the TS output was driving the simulation. These simulations have been used to derive the error statistics of the LUTs compared with their parent parameterization, at these same grid points, fed with the same input variables.

A second set of three simulations is then carried out, where the only difference is the radiation scheme: S1 uses HS, S2 uses Chen–Cotton, and S3 uses the best LUT and different meteorological fields (250-mb wind, 500-mb geopotential, 500-mb vertical velocity, 2-m temperature, and 10-m wind) from the three simulations are compared. All the above-mentioned simulations have the same setup as was used for the model climatology [see section 2b(1)], where the radiation scheme was called for 20 min and where in the different schemes for S2 and S3 the only relevant difference is the time period: from 0600 UTC 1 September to 0600 UTC 3 September 2005. This period was intentionally chosen to have the largest possible difference between the zenith angle of the input soundings and the EOF one. As in the climatology case, the RAMS simulations are in agreement with the reanalysis (not shown).

### Table 2. Correlation coefficient (r), bias, RMSE, and error standard deviation (std dev) for the eight transfer schemes compared with the original Harrington scheme (see Table 1 for the details on the transfer schemes).

<table>
<thead>
<tr>
<th>Heating rates</th>
<th>TS10RF00</th>
<th>TS10RF11</th>
<th>TS10RF21</th>
<th>TS10RF31</th>
<th>TS10RF00v75</th>
<th>TS10RF00v50</th>
<th>TS10NRF00</th>
<th>TS9RF00</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>0.9110</td>
<td>0.9110</td>
<td>0.9110</td>
<td>0.9110</td>
<td>0.9110</td>
<td>0.9110</td>
<td>0.9110</td>
<td>0.9110</td>
</tr>
<tr>
<td>Bias (K h⁻¹)</td>
<td>0.009 417</td>
<td>0.009 336</td>
<td>0.009 364</td>
<td>0.009 372</td>
<td>0.009 309</td>
<td>0.009 181</td>
<td>0.009 181</td>
<td>0.009 181</td>
</tr>
<tr>
<td>RMSE (W m⁻²)</td>
<td>0.024 309</td>
<td>0.024 337</td>
<td>0.024 378</td>
<td>0.024 420</td>
<td>0.024 092</td>
<td>0.024 268</td>
<td>0.024 211</td>
<td>0.024 118</td>
</tr>
<tr>
<td>Std dev (W m⁻²)</td>
<td>0.022 412</td>
<td>0.022 475</td>
<td>0.022 508</td>
<td>0.022 550</td>
<td>0.022 221</td>
<td>0.022 385</td>
<td>0.022 403</td>
<td>0.022 249</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Longwave flux</th>
<th>TS10RF00</th>
<th>TS10RF11</th>
<th>TS10RF21</th>
<th>TS10RF31</th>
<th>TS10RF00v75</th>
<th>TS10RF00v50</th>
<th>TS10NRF00</th>
<th>TS9RF00</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>0.9624</td>
<td>0.9618</td>
<td>0.9596</td>
<td>0.9596</td>
<td>0.9652</td>
<td>0.9642</td>
<td>0.9633</td>
<td>0.9645</td>
</tr>
<tr>
<td>Std dev (W m⁻²)</td>
<td>8.7835</td>
<td>8.8986</td>
<td>9.117</td>
<td>9.1012</td>
<td>8.4602</td>
<td>8.5585</td>
<td>8.7407</td>
<td>8.518</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Shortwave flux</th>
<th>TS10RF00</th>
<th>TS10RF11</th>
<th>TS10RF21</th>
<th>TS10RF31</th>
<th>TS10RF00v75</th>
<th>TS10RF00v50</th>
<th>TS10NRF00</th>
<th>TS9RF00</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>0.9973</td>
<td>0.9978</td>
<td>0.998</td>
<td>0.9978</td>
<td>0.998</td>
<td>0.9972</td>
<td>0.9977</td>
<td>0.9977</td>
</tr>
<tr>
<td>RMSE (W m⁻²)</td>
<td>45.8934</td>
<td>45.8311</td>
<td>45.7367</td>
<td>45.817</td>
<td>45.9621</td>
<td>45.9066</td>
<td>46.0698</td>
<td>45.7797</td>
</tr>
<tr>
<td>Std dev (W m⁻²)</td>
<td>38.5675</td>
<td>38.5139</td>
<td>38.4633</td>
<td>38.5248</td>
<td>38.596</td>
<td>38.5745</td>
<td>38.7015</td>
<td>38.4936</td>
</tr>
</tbody>
</table>

3. Results

a. Accuracy with respect to the parent scheme

The determination of an acceptable error of the TS, compared with respect to its parent parameterization, depends on the use of the parameterization. For example, if the goal of the parameterization is to obtain diabatic heating rates in a numerical weather prediction model, differences between the different approaches might not matter if the heating rate differences that matter in the prediction of weather in the model were less than about 0.1°C h⁻¹. The ability of the different TSs to reproduce the behavior of the HS is thoroughly evaluated here in this context.

The word “error” in this section, therefore, is used to refer strictly to the differences of the TSs relative to the HS, not against observations. Because of the lack of error analysis against observations for the HS, the uncertainty introduced by the use of the TS is evaluated by comparing meteorological fields obtained with the HS, the most accurate TS, and a second, widely used radiation parameterization (i.e., the Chen–Cotton scheme). A further discussion of the origin and significance of these uncertainties is given in section 4.

The overall correlation coefficients, bias, RMSE, and error standard deviation for the heating rates, and the long- and shortwave surface fluxes are presented in Table 2 for the eight TSs described in section 2b. There is little change among the different across the different TSs, indicating that the HS output is well replicated. For the heating rates, the correlation coefficient (r) is always larger than 0.91, with a bias smaller than 0.009 417 K h⁻¹, and an RMSE of approximately...
0.024 K h$^{-1}$. The longwave flux has a higher correlation coefficient of at least 0.9592, with a negative bias that oscillates between $-10.9737$ and $-12.1454$ W m$^{-2}$, while the RMSE tends to be slightly larger ($-14$ W m$^{-2}$). Although the bias and RMSE for the shortwave fluxes are larger than for the longwave, the correlation coefficient is better for the shortwave fluxes (at least 0.9972). Relative errors (not shown) tend to be smaller and exhibit less variability over time, for the longwave than for the shortwave fluxes because the latter decrease (increase) in value before sunset (sunrise). TS10RF00v75 consistently gives better results across all parameters for the three different outputs.

1) Heating rates

The time evolution of the heating rate correlation coefficients (Fig. 4a) shows a clear daily cycle for all LUTs, with a nighttime maximum and a daytime minimum. TS9RF00 and TS10RF00 have the lowest minima during the day, confirming the value of the choices made for the EOF calculations. Also, the correlation tends to be better close to the surface, although it has a secondary maximum at the third-to-last model level and a secondary minimum at the second level from the ground (Fig. 4d).

The bias (Fig. 4b) shows similar behavior with negative nighttime minima and positive nighttime maxima and virtually no spread among the TSs. Moreover, during the day and at times during the night, the lowest values of the bias (in an absolute sense) correspond to the times when no time interpolation is done and only the EOFs computed at the very same hour are used. This seems to indicate that a higher temporal frequency or a different interpolation scheme could further reduce the bias. The vertical profile of the bias is consistent with the correlation analysis (Fig. 4e). The RMSE behaves very similarly to the bias in time and space (not shown) but is roughly double in magnitude.

The average absolute error is shown in Figs. 4c and 4f, along with the 95th percentile of the distribution. The average absolute error is generally below 0.025 K h$^{-1}$ and it tends to be larger during the night and smaller during the day. The 99th percentile of the absolute error distribution (not shown) has a more accentuated daily cycle between 0.04 and 0.11 K h$^{-1}$ for all the TSs. Throughout the simulation, all the TSs have similar absolute errors distributions such that 0.1% of the absolute errors are four times larger than the 95th percentile as for the absolute error distribution between the online and offline versions of the parent parameterization itself. This last similarity in the error distributions leads to the thought that eliminating the small inconsistency between the version of the HS used in RAMS and the one utilized offline to obtain the heating rates for the EOFs would eliminate the largest errors, although their frequency is too small to significantly improve the other statistics.

2) Longwave fluxes

The minimum value of the correlation coefficient for the longwave fluxes throughout the 2-day experiment is very good (0.925–0.935) for all the TSs. A daily cycle is also present, as for the heating rates, but its amplitude is not nearly as constant (Fig. 5a). The two TSs that employ a smaller fraction of the variance (TS10RF00v50 and TS10RF00v75) tend to have higher correlations with the HS.

The bias also oscillates daily, roughly around $-12$ W m$^{-2}$, but it always stays negative for every TS (Fig. 5b). Given the high correlation coefficient, the RMSE (not shown) is the mirror image of the bias.

The average and maximum absolute error show similar patterns of behavior, (Fig. 5c) as they are relatively large at the beginning of the simulation, then rapidly decrease to a relative minimum at 1500 UTC and increase until the afternoon of the first day. A similar cycle is repeated the second day.

Both the average absolute error and the bias show a pronounced sawtoothlike trend, with peaks corresponding to the hour of the EOF computation, although it is less evident for TS10RF00v75, especially for the averaged absolute error.

3) Shortwave fluxes

The time evolution of the correlation coefficient, shown in Fig. 5a peaks at the hours of EOF computation, and is lower at sunset and dawn, but its worst value is 0.94 and most of the time it is above 0.98 for all TSs. There is virtually no spread among the different TSs, but again TS10RF00v75 proves to be better than the other TS because of its higher values around 1800 UTC 1 September.

The bias is generally negative and has relative minima between the EOF hours (Fig. 5b and 5c). As for the longwave, the RMSE is almost always equal in magnitude to that of the bias. They are only both positive before 1200 UTC and after 0000 UTC. The mean and the 95th percentile of the absolute error behave similarly to the bias (Fig. 5f).

4) Meteorological fields

To show the overall effects of the new algorithm on some common meteorological fields, the outputs of the simulations utilized for the above tests are shown in Figs. 6–10. All the fields are taken at the latest time of
the simulations (2 days) to ensure the maximum divergence between the three different setups of the radiation scheme: the Harrington scheme, the best LUT (TSR10F00v75), and the Chen–Cotton scheme (CCS). The white areas indicate cloudy points, at least for the Harrington or LUT simulations.

The differences are generally small and have different characteristics. The 500-mb vertical velocity (Fig. 6) and the 10-m wind speed (Fig. 7) have average absolute errors that are about half of the corresponding errors of the CCS, while the maxima are more similar (see Table 3), indicating that the LUTs tend to have small absolute errors, with a few points with higher errors (i.e., the distribution has a very long tail). More precisely, for the

Fig. 4. (left) The heating rate error statistics over 2 days of simulation for the (a) correlation coefficient, (b) bias (lower group of lines) and RMSE (upper group of lines), and (c) averaged absolute error (lower group of lines) and 95th percentile of the absolute error distribution (upper group of lines). (right) As in the left panels but per model level: (d) correlation coefficient, (e) bias (left group of lines) and RMSE (right group of lines), and (f) averaged absolute error (left group of lines) and 95th percentile (right group of lines).
two previously mentioned error distributions, the difference between the 99th percentile and the maximum error is at least two times the difference between the 99th and 50th percentiles (not shown). The 2-m temperature (Fig. 8) has lower but more widespread absolute errors: the average is 0.40° C (Table 3), and the difference between the maximum and the 99th percentile is about two-thirds the 99th – 50th percentile difference (not shown). Furthermore, the LUT error ranges from −2.77° to 1.89° C (Fig. 8, top-right panel), while the difference between the HS and the CCS goes from −8.02° to 0.86° C (Fig. 8, bottom right). The 500-mb geopotential height and the 250-mb wind speed have error distributions similar to those of the 10-m wind, but the maximum absolute errors, 1.19 m and 0.29 m s⁻¹, are negligible compared with the actual values of the two fields, and the top and middle-left panels of Figs. 9 and 10 do not show any visible difference.

Fig. 5. (a), (d) The correlation coefficients; (b), (e) the biases; and (c), (f) the averaged absolute error (lower group of lines) and RMSE (upper group) (bottom) for the (a)–(c) longwave and (d)–(f) shortwave fluxes.
FIG. 6. The 500-mb vertical velocity (cm s\(^{-1}\)) at the end of the 2-day test simulation: (left, top to bottom) runs with the HS, the best LUT, and the CCS. (right, top to bottom) The difference between the best LUT and the HS, between the best LUT and the CCS, and between the CCS and HS.
FIG. 7. As in Fig. 6, but for 10-m speed (m s$^{-1}$).
FIG. 8. As in Fig. 6, but for 2-m temperature (°C).
Fig. 9. As in Fig. 6, but for the 500-mb geopotential height (m).
FIG. 10. As in Fig. 6, but for the 250-mb wind speed (m s$^{-1}$).
b. Computational speed

Computational speed has been tested by profiling the four basic model configurations used thus far: the RAMS standard version with HS, the standard version with the CCS, RAMS with TSR10F00, the control TS, and TSR10F00v75, the most accurate TS. The profiling data consist of the number of times each subroutine, within RAMS, has been called, the total time that has been spent on each subroutine (cost), and the time spent on each subroutine excluding the time due to call to or from other functions and subroutines (self-time). This allows a detailed comparison of the execution times accounting not only for the subroutines that actually produce the desired outputs, but also for all those subroutines that preprocess data or carry out computations once per domain, instead of once per grid point.

It was determined that the only nonnegligible operation, besides the schemes, was the loading of the EOF data, which was necessary for the TSs. Every other preprocessing part of the computations, for all schemes, was negligible compared to the core of the corresponding scheme. Table 4 shows the ratio of time spent on the different schemes to the time spent on the HS. For the LUTs a range is given because their execution time is comparable to the accuracy of the time measure. The execution time ratio for the CCS is 0.544, which means it is about twice as fast as the HS, but for the control TS the same ratio is 0.0732, implying a 93% reduction in the execution time. TSR10F00v75 performs even better and the reduction is 96%. The amount of additional data that the TSs require to read at initialization depends on the number of grid points, and with this configuration that amount is sizeable, 388 Mb for the control TS and 124 Mb for TSR10F00v75, but because of the large speed gain, assuming 70% of the domain is cloud free, it takes 40–80 radiation calls to offset the initial overhead for the control TS (i.e., TSR10F00), and only 6–11 for TSR10F00v75. This number varies not only because of the uncertainty of the time measure of the TS, but also because it involves reading from a hard disk whose timing depends on the size of the data, if the disk is mounted over the network, and how many processes are accessing it.

The profiling data discussed above are accurate, but they are also affected by the profiler itself, which modifies the original code in order to measure the timing. In this specific case, calling the LUT alternatively to the HS caused the parent subroutine to increase its self-time even though the LUT was turned on from a configuration file, and the actual executable was not changed. Therefore, the overall duration of the simulations was also measured without any profiler on a dedicated computer to minimize the overhead due to profiling and the input–output (I–O) times. The relative times are shown in the fourth column in Table 4 and are all inclusive. TSR10F00 and TSR10F00v75 result in 10% and 11% reductions in the overall duration of the simulations, respectively. The use of CCS yields a similar reduction of 9% despite the longer time frame. The small difference between the CCS and the TS is due to the additional overhead of the I–O and to the fact that the TS runs only on 70%–80% of the domain, and the HS takes care of the remaining 30%–20%. Most likely an increase in the self-time of the parent subroutine still also occurs, and can be eliminated when a “cloudy sky” version of the LUT will be ready and can completely replace the HS.

4. Discussion

Despite the very high correlation coefficients, the differences between the TSs and the HS for the surface...
heating rates would be affected by the different vertical divergence, so that the comparison between the fluxes that are later converted to heating rates via parameterizations of Fu and Liou (1992), Gabriel et al. the parent parameterization is compared against the errors of the HS, the accuracy of the TSs in reproducing the more EOFs, farther away in time, a polynomial interpolation that takes into account the current time of the day can ensure that the daily maximum of the shortwave fluxes, especially for the daily maximum. Most likely, this behavior can be corrected by requiring that the interpolation of the EOFs at any point in time for the shortwave fluxes happen only when both the preceding and following EOFs have zenith angles that are large enough to trigger the computation of the Harrington scheme, which would be used otherwise. This approach has been attempted. While it slightly decreases the error around the hours of dawn and sunset, it provided little benefit to the overall accuracy performance of the algorithm, and thus, it has not been adopted.

It is noteworthy that all the errors of the shortwave fluxes have their relative maxima always at distances equal in time from the hours at which the EOFs were computed. This happens because the sinelike function that describes the daily behavior of the shortwave fluxes is not very well approximated by the linear interpolation of the relative low-frequency (3 h) values obtained from the EOFs. This results in an underestimation of the shortwave fluxes, especially for the daily maximum. Most likely, this behavior can be corrected by increasing the computational frequency of the EOFs at least during the daytime. In addition, a different weighting strategy can provide some benefit: by use of more EOFs, farther away in time, a polynomial interpolation that takes into account the current time of the day can ensure that the daily maximum of the shortwave fluxes is not underestimated or is underestimated to a lesser extent.

Because of the poor knowledge of the clear-sky errors of the HS, the accuracy of the TSs in reproducing the parent parameterization is compared against the parameterizations of Fu and Liou (1992), Gabriel et al. (2000), and Zhang et al. (2003).

The above radiative transfer schemes predict heat fluxes that are later converted to heating rates via vertical divergence, so that the comparison between the heating rates would be affected by the different vertical grid spacings used for the tests. This is particularly important close to the surface where the vertical grid spacing used in this study (50–1500 m) is certainly smaller than the one regularly used for this kind of test (700 to over 3000 m above the tropopause). Both the Gabriel et al. (2000) and Zhang et al. (2003) radiative transfer schemes have errors of the order of 1 W m$^{-2}$, which at the first RAMS model level translates into a 0.056 K h$^{-1}$ error on the heating rate and compares well against the averaged absolute error of the TS (Figs. 4c and 4f). At higher altitudes, where the resolution is coarser and the density smaller, the same accuracy for the fluxes results in smaller heating rate errors, by about an order of magnitude. Here, the TSs have larger errors (Figs. 4e and 4f). Fu and Liou (1992) used a constant vertical grid spacing of 1 km, and similar back-of-the-envelope calculations using their errors on the heating rate are consistent with the above comparisons.

While the above error analysis provides a very good description of the strengths and weakness of the TSs, the more definitive proof that the uncertainties due to the imperfect reproduction of the HS output by the TS are acceptable lays in the effects on the meteorological fields. Figures 6–10 show that at after 2 days of simulation, the main meteorological fields are not significantly different from the simulations with the parent scheme, even at the higher altitudes where the heating rates errors are relatively larger. The surface temperature differs more than other fields, but it is not unphysically and most of all, its variations from the HS are of the same order of magnitude as another common parameterization (Chen–Cotton), thus strengthening the suitability of the TS for climate simulations.

The computational gain over the HS is 95%, and makes the trade-off error versus accuracy worthwhile. The success of the TS is likely due to the clear-sky condition, which implies weak multiple scattering, and low optical thickness. In this case, the up- and down-welling fluxes decouple and the multiple-scattering problem reduces to Beer’s law. The great computational gains are due to the fact that the use of HS output obtained from the EOFs bypasses the online calculation of the absorption coefficients whose computation is indeed the bottleneck of two-stream parameterizations, such as the HS. Similarly, the selection rule of Gabriel et al. (2001) also exploits this simpler condition to reduce the number of radiative transfer calculations and thus the computational expense. Although this seems to indicate that a TS is likely to work less well when multiple scattering is significant, the study of Natraj et al. (2005) shows the authors were able to accurately simulate the residual of the reflectance in the O$_2$ A band whose absorption coefficients were by a
multiscattering line-by-line code. The number of lines used was selected through an EOF analysis with subsequent radiance calculations initialized by a two-stream model.

5. Conclusions

In this study a methodology to develop a LUT or TS from a parameterization has been presented. It is important to further clarify that this work does not introduce a new parameterization simply derived from a preexisting one, but reduces a parameterization to a TS, whose core concept is to compute the EOFs of the parent scheme input variables, under clear-sky conditions, and run it on the EOFs. Then, the TS output of a generic input is a weighted average of the EOF output, where the weights are based on a form of the distance between the input and each individual EOF. Several TSs have been develop for the Harrington radiation scheme under clear-sky conditions, by using different EOFs, and their errors have been thoroughly analyzed, as has their computational speed. The errors with respect to the parent parameterization, at times, can be larger than what is commonly accepted as error for a radiation parameterization compared against a line-by-line code, but this kind of analysis has not yet been published for the HS or other mesoscale schemes, at least for the clear-sky case. Therefore, it is not possible to know with certainty the error introduced with the TS, and it is suggested that a different weighting strategy is very likely to improve the shortwave flux errors. Furthermore, once the best TS has been implemented into RAMS, the meteorological fields after a 2-day simulation show good agreement with the parent scheme, and a comparison against the meteorological fields obtained by use of the Chen–Cotton scheme indicates that the uncertainties introduced by the TS, as compared with the HS, are less significant than the ones due to the second scheme. Finally, the calculations necessary for the TS are carried out at a fraction of the original cost.

While this study is limited to the Harrington radiation parameterization, it is reasonable to believe that the same methodology can be extended to a cloudy sky and applied to other parameterizations with similar results, as first envisioned in Pielke et al. (2006).

Acknowledgments. During the completion of this work Dr. Pielke and Mr. Leoncini were funded by the U.S. Department of Defense’s Center for Geoscience/Atmospheric Research Grant at Colorado State University (under Cooperative Agreements W911NF-06-2-001 and DAAD19-02-2-0005 with the Army Research Laboratory). Dr Philip Gabriel acknowledges NASA MAP Grant NNG06GC10G and DOE Grant DE-FG02-94ER61748. The authors would also like to thank Dr. Manajit Sengupta for his useful suggestions.

REFERENCES


Lorenz, E. N., 1956: Empirical orthogonal function and statistical weather prediction. Statistical Forecasting Project Science Rep. 1, Dept. of Meteorology, Massachusetts Institute of Technology, 49 pp. [NTIS AD 110268.]


