Journal of Quantitative Spectroscopy & Radiative Transfer 173 (2016) 65–71

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy & Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

A fast and accurate PCA based radiative transfer model: Extension to the broadband shortwave region

Pushkar Kopparla [a,*], Vijay Natraj [b], Robert Spurr [c], Run-Lie Shia [a], David Crisp [b], Yuk L. Yung [a]

[a] Division of Geological and Planetary Sciences, MC 150-21, California Institute of Technology, Pasadena, CA 91125, USA
[b] Jet Propulsion Laboratory (NASA-JPL), 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[c] RT Solutions, Inc., 9 Channing St, Cambridge, MA, USA

ARTICLE INFO

Article history:
Received 4 August 2015
Received in revised form
4 December 2015
Accepted 10 January 2016
Available online 16 January 2016

Keywords:
Radiative transfer
Principal component analysis
Performance enhancement
Shortwave broadband
Visible & infrared

ABSTRACT

Accurate radiative transfer (RT) calculations are necessary for many earth–atmosphere applications, from remote sensing retrieval to climate modeling. A Principal Component Analysis (PCA)-based spectral binning method has been shown to provide an order of magnitude increase in computational speed while maintaining an overall accuracy of 0.01% (compared to line-by-line calculations) over narrow spectral bands. In this paper, we have extended the PCA method for RT calculations over the entire shortwave region of the spectrum from 0.3 to 3 microns. The region is divided into 33 spectral fields covering all major gas absorption regimes. We find that the RT performance runtimes are shorter by factors between 10 and 100, while root mean square errors are of order 0.01%.

1. Introduction

The increasing complexity of climate models is a function of the need to include more physical processes and reduce dependence on parameterizations; this is particularly the case for direct and indirect radiative forcing due to clouds and aerosols. Accurate line-by-line (LBL) 1-D multiple scattering (MS) RT models are able to capture the full physics of RT in the shortwave (0.3–1 μm) with simulated radiances showing excellent agreement with observations under cloud free conditions [1]. However such models are computationally very expensive. Therefore, climate models are forced to use fast, approximate RT codes that provide speed as a tradeoff for accuracy. Oreopoulos et al. [2] made an effort to characterize these inaccuracies under the Continual Intercomparison of Radiation Codes program, and they found that the 13 solar RT codes examined tended to overestimate the transmitted and reflected radiances, with large errors ($>10\%$) in the shortwave fluxes at the surface. Another RT model intercomparison experiment that compared 31 RT codes being used in global models, the AeroCom [3], reported that approximate models had biases ranging from -10 to 20% compared to values from LBL models at the top of the atmosphere.

The other typical approaches to dealing with computationally expensive RT in climate modeling are either to update radiative heating calculations only once every few dynamical timesteps or to perform RT on a coarser spatial grid. These approaches are not accurate; they tend to adversely impact model output quite strongly in the presence of clouds and have been shown to affect model climate sensitivity [4]. Pauluis and Emanuel [5] showed that the infrequent RT calculations introduced numerical instabilities; these authors provide a criterion (analogous to the Courant–Friedrichs–Levy criterion) for the maximum allowable time step between two RT calculations.

* Corresponding author.
E-mail address: pkk@gps.caltech.edu (P. Kopparla).

http://dx.doi.org/10.1016/j.jqsrt.2016.01.014
0022-4073/

They also suggest approximate numerical corrections if this criterion is violated due to computational limitations. However, these issues can be avoided if a sufficiently fast and accurate RT model is available.

There have been many approaches to enhance the performance of RT computations, the most widely used being the correlated-k model [6–8]. Other approaches include spectral mapping [9,10], low-stream interpolations [11,12] and low orders of scattering approximations [13]. The idea of using Principal Component Analysis (PCA) for RT was first proposed by Natraj et al. [14] (and later independently by Liu et al. [15]), who reproduced the TOA reflectance spectrum over a small spectral region centered on the O_2 A band at 0.75 μm to accuracies of 0.3% compared to a LBL RT model, while achieving a 10-fold increase in speed. Further development of the PCA model, its expansion to broader spectral regions and the derivation of analytic Jacobians are documented in detail in other papers [16,17].

This paper is organized as follows. In Section 2 we describe the concept of the PCA RT model and its setup. In Section 3, we present runtime and error statistics for TOA radiance calculations, as well as scalings of these parameters for various model settings. In Section 4, we discuss known issues and directions for future model development.

2. The PCA RT technique

2.1. RT models

The PCA RT technique is based on two contrasting RT models. First, for accurate MS calculations we use the LIDORT discrete ordinate RT model [18] which includes the treatment of solar-beam incoming attenuation in a spherically curved atmosphere (the pseudo-spherical approximation), and the use of the delta-M scaling approximation for aerosol scattering with sharply peaked forward scattering. We run LIDORT in MS mode with 32 streams (computational quadrature angles; 16 each for upwelling and downwelling polar directions); the single scatter (SS) contribution is not included in the calculations. A full set of LIDORT MS calculations at every wavelength point is computationally expensive; the key to the PCA RT approach is to drastically limit the number of such full-MS calculations to a reduced set of PCA-determined optical profiles that capture the vast majority of optical information for a given wavelength range.

Secondly, fast RT computations are done using a numerically efficient two-stream-exact single scattering (2S-ESS) RT model [19], which comprises two parts. The 2S part is a fast MS calculation based on a single discrete ordinate in each of the up-welling and down-welling directions; the RT calculation is done analytically except for the multi-layer boundary-value problem (which is also solved using a simple and fast pentadiagonal solver rather than typical matrix inversion techniques). The 2S calculation also uses the pseudo-spherical approximation. The ESS part is an accurate spherical-geometry calculation of the singly scattered radiation computed with the complete scattering phase function (not a truncated form based on a limited number of Legendre polynomial expansion coefficients). The 2S-ESS combination is the "fast" RT calculation; in this, the use of the ESS calculation mitigates bias due to the severe phase function truncation inherent in the 2S approximation to MS. The LIDORT-ESS combination provides the most accurate computation of the complete (SS + MS) radiation field; in the following, we will refer to this benchmark computation as the "Exact RT" calculation, against which the accuracy of the PCA RT model is to be compared.

2.2. Formalism

We will briefly summarize the formalism for the PCA RT model here, as the technique continues to evolve. Earlier versions of the procedure can be found in Refs. [14,16,17]. The first step is to partition a given spectral range into a number of appropriately chosen bins, each bin containing spectral points for which the optical properties are broadly "similar". Binning criteria are often determined by similarities in total atmospheric optical depth; we discuss the criteria in Section 2.3. Each bin, which contains a subset of the optical property data, is then subjected to an independent PCA procedure as described below.

Consider a bin with N wavenumbers in an atmosphere stratified into M optically homogeneous layers. Let $\tau_{i,j}$ be the total optical depth in layer i at wavenumber j and $\omega_{i,j}$ be the layer single scattering albedo. The index i goes from 1 to M and j from 1 to N. Let us denote the logarithm of this set of optical properties by $F_{k,j}$ where $F_{[1,M],j}$ are the optical depths and $F_{[M+1,2M],j}$ are the single scattering albedos and the index k goes from 1 to $2M$. The mean-removed covariance matrix C over all layers has elements given by (k and l are layer indices)

$$C_{k,l} = \overline{(F_{k,j} - \overline{F_k})(F_{l,j} - \overline{F_l})} \tag{1}$$

where the overbar denotes a mean-value over all wavenumbers in the bin. The index l also goes from 1 to $2M$. The $2M$ eigenvectors of this covariance matrix form an orthogonal set that is mutually uncorrelated – this is the set of empirical orthogonal functions (ϵ_k). Each EOF has $2M$ components and unit EOFs are scaled by the square root of the variance, making them dimensionally consistent with the corresponding optical properties. This orthogonal set represents a coordinate system onto which the original optical property data set can be mapped. The projection of the original optical properties onto the EOFs gives a set of weights called the principal component scores (PCs, P_k). Again, there are $2M$ PCs and each PC has N components, corresponding to the number of wavenumbers. Thus, we can express all the original information in our new coordinate system as

$$\ln \tau_{k,j} = \ln \overline{\tau_k} + \sum_{i=1}^{M} P_{i,j}\epsilon_{i,k} \tag{2}$$

$$\ln \omega_{k,j} = \ln \overline{\omega_k} + \sum_{i=M+1}^{2M} P_{i,j}\epsilon_{i,k} \tag{3}$$

Wavenumber information is stored in the PCs, while the