An Empirical Cumulus Parameterization Scheme for a Global Spectral Model

K. Rajendran [*†]
T. N. Krishnamurti
Department of Meteorology
Florida State University
Tallahassee, Florida

V. Misra
Center for Ocean-Land-Atmosphere Studies
Institute of Global Environment and Society, Inc.
Calverton, Maryland

and

W.-K. Tao
Laboratory for Atmospheres
NASA Goddard Space Flight Center
Greenbelt, Maryland

---

[*]Current affiliation: Meteorological Research Institute, Tsukuba, Ibaraki, Japan

[†]Corresponding author: K. Rajendran, Meteorological Research Institute, 1-1 Nagamine, Tsukuba, Ibaraki-305 0052, Japan. email: *rajend@mri-jma.go.jp*

Abstract

Realistic vertical heating and drying profiles in a cumulus scheme is important for obtaining accurate weather forecasts. A new empirical cumulus parameterization scheme based on a procedure to improve the vertical distribution of heating and moistening over the tropics is developed. The empirical cumulus parameterization scheme (ECPS) utilizes profiles of Tropical Rainfall Measuring Mission (TRMM) based heating and moistening derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) analysis. A dimension reduction technique through rotated principal component analysis (RPCA) is performed on the vertical profiles of heating (Q_1) and drying (Q_2) over the convective regions of the tropics, to obtain the dominant modes of variability. Analysis suggests that most of the variance associated with the observed profiles can be explained by retaining the first three modes. The ECPS then applies a statistical approach in which Q_1 and Q_2 are expressed as a linear combination of the first three dominant principal components which distinctly explain variance in the troposphere as a function of the prevalent large-scale dynamics. The principal component (PC) score which quantifies the contribution of each PC to the corresponding loading profile is estimated through a multiple screening regression method which yields the PC score as a function of the large-scale variables.

The profiles of Q_1 and Q_2 thus obtained are found to match well with the observed profiles. The impact of the ECPS is investigated in a series of short range (1-3 day) prediction experiments using the Florida State University global spectral model (FSUGSM, T126L14). Comparisons between short range ECPS forecasts and those with the modified Kuo scheme show a very marked improvement in the skill in ECPS forecasts. This improvement in the forecast skill with ECPS emphasizes the importance of incorporating realistic vertical distributions of heating and drying in the model cumulus scheme. This also suggests that in the absence of explicit models for convection, the proposed statistical scheme improves the modeling of the vertical distribution of heating and moistening in areas of deep convection.

1. Introduction

One of the most important physical processes that play a crucial role in numerical weather prediction is the cumulus convection. In addition to modifying the heat balance of the tropical atmosphere, the energy release in cumulus convection is strongly linked to the growth of tropical weather systems such as tropical cyclones (Charney and Eliassen 1964). The spatial and temporal distribution of heating in the tropics has an important impact on the large scale circulation and climate variability. It has been shown that vertical profiles of heating and drying of tropical cloud clusters affect the way the large scale circulation responds to deep convection (e.g., Hartmann et al. 1984; Krishnamurti et al. 1995; Zhang and Krishnamurti 1996). The role of diabatic heating on the 30-60 day oscillation in the tropical atmosphere has also been extensively studied and found to be important (e.g., Lau and Peng 1987; Chang and Lim 1988; Sui and Lau 1989). The convective heating serves as forcing for large-scale motions associated with the tropical intraseasonal oscillations and its simulation is found to be sensitive to the heating rates produced by different cumulus schemes (Rajendran et al. 2002). The impact of vertical heating and moistening profiles on the forecast of deep convection is significant especially during extreme events. For example, the space time heating distribution in individual storms is useful for storm diagnostics and forecasting applications. These studies suggest the importance of incorporating realistic vertical heating and drying profiles in the cumulus scheme to obtain accurate weather forecasts.

Krishnamurti et al. (1983) indicated that the modified Kuo scheme implemented in the FSUGSM is limited by the vertical structure function of heating and moistening, which is proportional to the difference between cloud and environmental properties. Johnson (1984) has pointed out that cumulus parameterization schemes that assign vertical distributions to the convective heating and drying should consider the different contributions from separate cumulus and mesoscale components as they have very different heating and moistening profiles. Alexander et al. (1993) analyzed the combined vector (Q_1, Q_2) derived from Australian Monsoon Experiment (AMEX) by applying rotated principal component analysis (RPCA) and showed that peaks in the profiles vary as a function of the life cycle stage of the convection. Here Q_1 and Q_2 are the apparent heat source and apparent moisture sink (Yanai et al. 1973) respectively. RPCA was also applied to the ECMWF analysis by Misra (1997) to extract the basic modes of cumulus heating and drying profiles. Lin and

3

Arakawa (2000) identified the basic modes of cumulus heating and drying profiles through RPCA to be used as a partial closure for Arakawa-Schubert cumulus parameterization scheme. These studies suggest that in the absence of explicit models for convection, statistical schemes might perform a reasonable job in producing profiles of heating and moistening which match closely with observations.

Explicit models of cumulus convection demand very high horizontal and vertical resolution and involve large computational resources. Instead the physical effects of cumulus convection are estimated from known large-scale model parameters which affect deep cumulus convection. In one of the pioneering studies demonstrating the strong dependence of deep cumulus convection on large-scale variables, using photographic mapping of clouds over the tropical Pacific, Malkus and Williams (1963) found that the prevalence of deep cumulus convection coincided with strong low-level synoptic-scale convergence. The composite analysis of 18 tropical disturbances in the equatorial West Pacific by Reed and Recker (1971) illustrated that the heaviest precipitation occurs in the trough region where relative humidity and upward motion were high. The precipitation estimates obtained from large-scale moisture budget, which conformed well with the observed data, suggested that moisture convergence was the single largest contributing factor in the moisture budget. Ogura (1975) found a strong linear correlation between rainfall rate and large-scale vertical velocity in tropical waves and hurricanes. Studies by Krishnamurti et al. (1980, 1983) have emphasized the importance of the control of large-scale wave dynamics on the rainfall associated with tropical mesoscale systems. At higher latitudes, besides upward motion and convergence of water vapor, large-scale thermodynamic structure and vertical wind shear also appear to be important in triggering cumulus convection (Zawadzki and Ro 1978; Zawadzki et al. 1981; Weisman and Klemp 1982). The results of these observational studies provide the large-scale variables which are crucial for cumulus heating and drying distribution in areas of deep convection.

The main objective of the present study is to investigate the forecast improvement in FSUGSM by realistic modeling of the vertical distribution of heating and moistening in areas of deep convection over the global tropics. A statistically based empirical cumulus parameterization scheme (ECPS) which makes use of the observed profiles of apparent heat source (Q_1) and moisture sink (Q_2), is developed. This is achieved by applying RPCA on profiles of Q_1 and Q_2 over active con-

vective regions of the tropics (40°S to 40°N) to obtain dominant modes of variability. The ECPS then uses a statistical approach in which Q_1 and Q_2 are expressed as a linear combination of the first three dominant principal components which distinctly explain variance in the lower, middle and upper troposphere as a function of the prevalent large-scale dynamics. The principal component (PC) score which quantifies the contribution of each PC to the corresponding loading profile is estimated through a multiple screening regression method which yields the PC score as a function of the large-scale variables. To study the impact of modeling realistic vertical distribution of Q_1 and Q_2 associated with deep convection, the ECPS is implemented in the Florida State University global spectral model (FSUGSM, T126L14) and a series of short range forecast experiments are performed. Further, the performance of the ECPS is analyzed by comparing the results of this ensemble of short range forecast experiments started from different initial conditions with the corresponding integrations of the control version of the FSUGSM. The improvement in forecast skill with ECPS is further analyzed against short range experiments in which both Q_1 and Q_2 profiles are based on ECMWF analysis.

The datasets used are described in section 2. Section 3 describes the analysis procedure based on RPCA, the formulation of the ECPS and the analysis results of the statistical approach used for the formulation of the ECPS. The results of short range prediction experiments using FSUGSM are presented in section 4. In section 5, we summarize the results.

2. Datasets: Q_1 and Q_2 profiles

The ECPS utilizes daily observed vertical profiles of apparent heating (Q_1) and moistening (Q_2) over the convective regions of tropics during the period 1-28 February 1998. The Q_1 profiles are derived using TRMM rainfall products (e.g., Tao et al. 2001) and the Q_2 profiles are derived from ECMWF analysis following Yanai et al. (1973).

2.1 TRMM heating (Q_1) profiles

In diagnostic studies (e.g., Yanai et al. 1973), the apparent heat source Q_1 of a large-scale system is defined by averaging horizontally the thermodynamic equation as:

$$Q_1 = C_p \bar{\pi} \left[\frac{\partial \bar{\theta}}{\partial t} + \bar{V} \cdot \nabla \bar{\theta} + \bar{w} \frac{\partial \bar{\theta}}{\partial z} \right],$$ (1)

where

$$\bar{\pi} = (\bar{p}/1000 \ mb)^{\frac{R}{C_p}}.$$

Here θ is the potential temperature, V is the horizontal wind vector, w is the vertical velocity, p is the pressure, R is the dry gas constant, and C_p is the specific heat of air at constant pressure.

The right hand side of Eq. 1 represents the large-scale response to heating from the apparent heat source, Q_1. Q_1 can be estimated either from observations or from grid values in a global or regional scale prediction model. Q_1 can be directly related to the contributions of cloud effects which can be explicitly estimated by a cloud-resolving or cumulus ensemble model (e.g., the Goddard cumulus ensemble model, Simpson and Tao 1993):

$$Q_1 = -C_p \bar{\pi} \frac{1}{\bar{\rho}} \frac{\partial (\overline{\rho w' \theta'})}{\partial z} + L_v (\bar{c} - \bar{e}) + L_f (\bar{f} - \bar{m}) + L_s (\bar{d} - \bar{s}) + \bar{Q}_R,$$ (2)

where ρ is the air density and L_v, L_f and L_s are the latent heats of vaporization, fusion and sublimation respectively. The variables c, e, f, m, d and s are the rates of condensation, evaporation, freezing, melting, deposition and sublimation per unit mass of air respectively. Q_R is the radiative heating rate per unit mass of air and primes indicate deviations from the horizontal averages.

To summarize,
$$Q_1 \ = \ \text{Eddy flux convergence of heat (Term 1)} + \text{Latent heating (Terms 2-4)}$$

$$+ \text{Radiation (Term 5)}.$$

Among the three contributors, the latent heating is found to dominate the eddy flux convergence and radiation contributions to the total Q_1^- (Tao et al. 1993a). Vertically integrated latent heating terms of Eq. 2 can be related to the surface precipitation rate, R, following Yanai et al. (1973) through moisture conservation:

$$\int_{z_{base}}^{z_{top}} \bar{\rho} \left[L_v (\bar{c} - \bar{e}) + L_f (\bar{f} - \bar{m}) + L_s (\bar{d} - \bar{s}) \right] dz \approx \rho_l L_v R,$$ (3)

$$i.e., \int_{z_{base}}^{z_{top}} \overline{\rho} Q_1 dz \approx \rho_l L_v R, \tag{4}$$

where ρ_l is the density of liquid water (1 g cm^{-3}). Thus, it follows that passive microwave estimates of surface rainfall rate can yield reasonable estimates of vertically integrated latent heating and these estimates are close to the vertically integrated Q_1.

a. Convective-Stratiform Heating (CSH) algorithm

The latent heating profiles are then derived using the Convective-Stratiform Heating (CSH) algorithm of Tao et al. (1993b). The flow chart in Fig. 1 shows the detailed procedure for the retrieval of the latent heating profiles. The algorithm uses surface precipitation rates, percentage amount of its stratiform component, the type and location of observed cloud systems and a look-up table containing latent heating profiles associated with various types of cloud systems (e.g., squall vs. non squall or monsoon) at different geographic locations, as inputs. The lookup-table has many sets of stored convective and stratiform latent heating profiles mostly simulated by the Goddard cloud ensemble model. Each set of profiles is normalized by the convective and stratiform rainfall amounts assuming a relation such as (4) applied to both convective and stratiform rainfall regimes. The rain rate, R, is used for limiting the total amount of derived heating (e.g., Eq. 4).

This algorithm takes into account the fact that when cloud latent heating is separated into convective and stratiform regions, the profiles for each region take on a characteristic shape even for systems from different geographic locations. Thus, the success of the method depends on the proper discrimination of convective and stratiform rain regimes which have characteristic vertical motion and heating profile shapes. The strong link between vertical velocity and diabatic heating rate is discussed by Anthes (1982) and Mapes and Houze (1995). Thus, an appropriate set of heating profiles which have been normalized with respect to surface rainfall, R, is selected from this look up table. Then these profiles are multiplied by satellite derived rainfall rates. The percentage of that rate which is stratiform is used in order to properly weigh the convective and stratiform profiles used in the retrieval.

For the retrieval of latent heating data used in this study, daily 0.5°-gridded rainfall and percentage of rainfall classified as stratiform/convective from TRMM Microwave Imager (TMI) produced using Goddard profiling (GPROF) algorithm (Olson et al. 1999) for February 1998 were used as

7

inputs for the CSH algorithm. These CSH derived latent heating profiles are validated with those estimated at well designed field campaigns such as TRMM field programs. The latent heating is found to dominate the eddy flux convergence and radiation contributions to the total heating.

The observed TRMM based February 1998 mean latent heating (Q_1, interpolated to model σ-level) at three vertical levels ($\sigma_{0.3}$, $\sigma_{0.5}$ and $\sigma_{0.8}$) over the global tropics is shown in Fig. 2. The horizontal pattern is very similar to the surface rainfall distribution during this month (not shown) especially at middle and upper levels. Stronger latent heating release ($\gtrsim 10$ K day^{-1}) in the middle and upper troposphere is associated with heavier surface rainfall. The noticeable features are the well defined tropical convergence zone (TCZ) over the East and central Pacific, Atlantic Ocean, a broad belt spanning from the South Indian Ocean through the Indonesian maritime continent, and a distinct south Pacific convergence zone (SPCZ). In addition, broad areas of strong heating are seen over the continental regions of the Southern Hemisphere. The heating over the Pacific Ocean is stronger than the heating over other centers, especially in the upper troposphere. This may be due to the higher tropopause and warmer sea surface temperatures over the Pacific. Over the Pacific and Indian Oceans, strong differential heating between land and ocean can strengthen the gradients in thermodynamic fields and affect the large-scale circulation. Also, at lower troposphere, heat sinks (areas of weak cooling) are seen over the Northern Hemispheric continents and small parts of Indian, Pacific and Atlantic Oceans (mainly at lower troposphere).

2.2   Moistening (Q_2) profiles based on ECMWF analysis

Following Yanai et al. (1973), the apparent moisture sink Q_2 of a large-scale system is defined by horizontally averaging the water vapor equation as:

$$Q_2 = -L \left[\frac{\partial \overline{q}}{\partial t} + \overline{V} \cdot \nabla \overline{q} + \overline{w} \frac{\partial \overline{q}}{\partial z} \right], \tag{5}$$

where $L = L_v$, is the specific latent heat of vaporization.

The datasets used for computing moistening (Q_2) profiles were obtained from ECMWF analysis on a $1° \times 1°$ horizontal grid for one month period extending from 1 through 28 February 1998. The datasets contained zonal and meridional winds, geopotential height, temperature and relative humidity at standard pressure levels. These fields were linearly interpolated to a T126 Gaussian grid in the horizontal and the 14 sigma levels in the vertical corresponding to the FSUGSM

8

(T126L14) resolution. Physical initialization (Krishnamurti et al. 1991, 1993) of FSUGSM using ECMWF analysis and observed SSM/I rain rates, yields FSU analyzed datasets in which vertical distributions of drying (Q_2) are consistent with observed rain rates, surface fluxes and clouds. Q_2 is thus obtained from the FSU analysis fields following Yanai et al. (1973).

The computed February 1998 mean distribution of Q_2 over the global tropics at three vertical levels ($\sigma_{0.3}$, $\sigma_{0.5}$ and $\sigma_{0.85}$) is shown in Fig. 3. The horizontal pattern is close to the surface rainfall distribution especially at lower troposphere. The heat sources (e.g., Fig. 2) in tropics are generally accompanied by moisture sink (Q_2) indicating that the release of latent heating of condensation is the major component of the heat source. The distributions of large positive Q_1 and Q_2 in the tropics are consistent with the distribution of the precipitation indicating that the heating is associated with deep cumulus convection.

2.3   *Validation datasets*

Observed daily precipitation on $1° \times 1°$ grid based on GPCP combined dataset (Huffman et al. 1997) was used for the validation of precipitation forecasts from FSUGSM. Daily ECMWF analysis were used for the validation of circulation fields.

3.   **Empirical cumulus parameterization scheme (ECPS)**

In the ECPS, the initiation and location of convective activity is addressed as in the modified Kuo scheme proposed by Krishnamurti et al. (1983). Thus, the ECPS has similar criteria as that of the modified Kuo scheme in which moisture convergence is considered as the dominant term in the moisture budget of a convective system. This suggests that the deep convection is prevalent if the atmosphere is conditionally unstable and if net moisture convergence is positive, i.e., if,

$$D_c \geq 0.2, \tag{6}$$

$$I_L > 0, \quad \text{and} \quad rainfall > 5 \quad \text{mm/day,}$$

then the ECPS is invoked in the model. Here, I_L is the large-scale moisture convergence and D_c refers to cloud depth which is given by

$$D_c = \frac{p_{LCL} - p_{cloudtop}}{p_{surface}}.$$

9

The condition implied by (6) suggests the following criteria viz.

$$\dot{\sigma} > 0,$$

$$\frac{\partial \theta_e}{\partial \sigma} > 0,$$

where θ_e is the equivalent potential temperature.

The moistening (M) and rain (R) rates are obtained as in modified Kuo scheme (e.g., Krishnamurti and Bedi 1988):

$$\frac{M}{I_L} = a_1 \zeta_{0.7} + b_1 \overline{\omega} + c_1, \tag{7}$$

$$\frac{R}{I_L} = a_2 \zeta_{0.7} + b_2 \overline{\omega} + c_2, \tag{8}$$

where $\zeta_{0.7}$ is the relative vorticity at $\sigma = 0.7$ and $\overline{\omega}$ is the vertically averaged p velocity.

The thermodynamic equation in the x, y, σ framework for the ECPS takes the form:

$$\frac{\partial T}{\partial t} + \nabla \cdot T\mathbf{V} - \dot{\sigma}\gamma - TD + \frac{RT}{C_p}\left[D + \frac{\partial \dot{\sigma}}{\partial \sigma}\right] = \frac{R\alpha_k}{\hat{Q}_1}, \tag{9}$$

and the moisture equation takes the form

$$\frac{\partial q}{\partial t} + \nabla \cdot q\mathbf{V} + \dot{\sigma}\frac{\partial q}{\partial \sigma} - qD = -\frac{(R-M)\beta_k}{\hat{Q}_2}, \tag{10}$$

where

$$\gamma = \frac{RT}{C_p \sigma} - \frac{\partial T}{\partial \sigma}.$$

Here D is the divergence, $\hat{Q}_1$ and $\hat{Q}_2$ are the integrated cumulus heating and drying profiles respectively, and α_k and β_k are the heating and drying components at k^{th} σ-level.

These thermodynamic and moisture equations have two new parameters α_k and β_k compared to the corresponding equations in modified Kuo scheme of Krishnamurti et al. (1983). These parameters are obtained through a statistical analysis procedure which is described in the next section.

3.1 Formulation of the statistical approach

To identify the dominant modes of variability associated with Q_1 and Q_2 profiles, the dimension reduction technique of rotated principal component analysis (RPCA) is applied to the vertical profiles of Q_1 and Q_2 over the active convective regions of the global tropics (40°S-40°N).

10

a. Rotated principal component analysis (RPCA)

One of the most common eigenvectorial techniques used in analyzing meteorological data is principal component analysis (PCA) which is also referred to as empirical orthogonal functions (EOFs, Preisendorfer 1988). With EOF/PCA, the variance of a dataset can be described using a minimal number of extracted factors/components that are spatially and temporally orthogonal to each other. RPCA is a statistical approach to extract preliminary factors through PCA and then obtain the final solution through rotation (a non-singular linear transformation on the PCs/EOFs) of the preliminary factors (Richman 1986). The RPCA solutions thus obtained agree better with the physically realizable structures of the input datasets than EOF/PCA (e.g., Richman 1986). In addition, the rotated principal components provide subdomain stability which is the invariance of modes of variability as sub-portions of the domain are examined. The performance assessment of various rotation methods in recovering the input structure by Richman (1986) suggests that no one specific rotation method will always yield the most accurate results. Our assessment of performance of various rotation methods in recovering the input structures of Q_1 and Q_2 suggested orthogonal varimax rotation as the most appropriate rotation method. Hence, in this study, we have used the orthogonal normal varimax rotation method for applying RPCA.

In PCA, the matrix with EOFs as its column vectors is referred to as the PC loading matrix where the elements of the matrix (PC loadings) represent the correlation between the PC and the variable analyzed. The matrix with PCs as its row vectors is referred to as the PC score matrix where the elements of the matrix (the PC scores) represent the contribution of the PC to the corresponding loading profile. Mathematically RPCA can be written as

$$(\mathbf{A})_{p \times N} = (\mathbf{B})_{p \times p} (\mathbf{T})_{p \times p} (\mathbf{PCC})_{p \times N}, \tag{11}$$

where $\mathbf{A}$ is the input matrix, $\mathbf{B}$ is the unrotated PC loading matrix, $\mathbf{T}$ is the transformation matrix, $\mathbf{PCC}$ is the PC score matrix and N is the total number of points. In this study, the elements of matrix $\mathbf{B}$ are given by the normalization, $\sum_{i=1}^{p} b_{ki}^2 = \lambda_k$, where p is the number of vertical levels and λ_k is the kth eigen value.

To obtain the dominant modes of variability associated with Q_1 and Q_2 profiles, the convectively active regions are delineated by selecting profiles of Q_1 and Q_2 only over points where the

11

integrated value of Q_1 over the depth of the column was a positive definite quantity. It is seen that performing RPCA on selected sets of Q_1 and Q_2 profiles produces EOFs, of which the first three dominant EOFs collectively explain more than 99% of the total variance. For example, the PC loading profiles corresponding to the first three EOFs of Q_1 are shown in Fig. 4a. Here, sigma level 0.99 corresponds to the lowest level above the surface and sigma level 0.05 corresponds to the top of the atmosphere. The first three EOFs explain 66.1%, 30.7% and 2.9% of total variance respectively. The profiles show distinctive peaks in the lower, middle and upper troposphere. These peaks indicate the level of maximum variation in the vertical column. Comparison of the dominant extracted modes with those obtained in previous studies reveals some differences among them which may be due to the differences in the domain, season and methodology of the analysis. While the TRMM based Q_1 rotated EOF1 profile resembles the first mode obtained by the previous studies (e.g., TOGA COARE IOP dataset based modes in Lin and Arakawa 2000, and Tung et. al. 1999), the EOF2 matches reasonably well with the second mode obtained by Alexander et. al. (1993) based on AMEX dataset. The EOF3 closely matches the third mode obtained by Lin and Arakawa (2000) based TOGA COARE IOP dataset. However, the significance of the TRMM based Q_1 rotated EOF3 is low because it contributes only about 3% of the total Q_1 variation.

Figure 4b shows the PC loading profiles corresponding to the first three EOFs with exactly half the number of convectively active data points used for RPCA. It is apparent that the shape of the profiles in the two cases (Figs. 4a and b) match exactly with each other except for a slight difference in the variance explained by individual EOFs. This result shows the subdomain stability of RPCA based on varimax rotation. Similarly, the profiles of the first three EOFs of Q_2, explaining 65.1%, 14.8% and 9.4% of the total variance respectively, are shown in Fig. 4c. Collectively, 89.3% of the total variance is explained by these EOFs. Thus, RPCA is able to isolate the dominant and stable modes of variability associated with the profiles of large scale latent heating and drying over the convectively active regions in the tropics.

Thus, applying RPCA on Q_1 and Q_2 profiles yields the PC loading matrices which can be represented as E_1 and E_2 respectively. These matrices are products of respective unrotated PC loading matrix and the transformation matrix. Similarly, the PC score matrices for Q_1 and Q_2 can

12

be represented as PCC_1 and PCC_2 respectively. Hence,

$$(A)_{p \times 1} = (E_1)_{p \times r}(PCC_1)_{r \times 1} \qquad (12)$$

$$(B)_{p \times 1} = (E_2)_{p \times r}(PCC_2)_{r \times 1} \qquad (13)$$

Here r represents the number of dominant EOFs retained. The elements of column matrix A are α_k and that of column matrix B are β_k, the heating and drying components in Eqs. 9 and 10 respectively. Each of the elements of the column vector PCC_1 given by $PC_{r_{Q_1}}$ and that of PCC_2 given by $PC_{r_{Q_2}}$ are determined through a screening regression analysis which is described in the next subsection.

b. *Multiple screening regression analysis*

In the ECPS formulation, the PC scores are obtained as a function of large-scale variables by applying a multiple screening regression analysis. Three PC scores corresponding to the first three retained EOFs for Q_1 and Q_2 are estimated through the regression analysis. The variables such as cloud depth, moisture convergence and large-scale vertical motion, which are known to vary with convective activity are chosen as the large-scale variables. A screening technique using the best subset of regressions was applied to optimize the number of predictor variables. Several regressions were tested with a varying number of predictors, and for each experiment all the subset of regressions were screened to obtain the optimum number of variables which explain most of the variance. The resulting six regression equations are

$$PC_{1_{Q_1}} = 3.0 - 0.3D_c - 2.1 \times 10^5 \left(\frac{0.5\dot{\sigma}_5 + 0.8\dot{\sigma}_6 + 2\dot{\sigma}_7 + 1.5\dot{\sigma}_8 + 0.7\dot{\sigma}_9 + 0.5\dot{\sigma}_{10}}{6} \right) \qquad (14)$$

$$PC_{2_{Q_1}} = 1.9 - 5.0 \times 10^6 I_L + 0.15D_c - 1.3 \times 10^5 \left(\frac{\dot{\sigma}_{10} + \dot{\sigma}_{11} + \dot{\sigma}_{12} + \dot{\sigma}_{13} + \dot{\sigma}_{14}}{5} \right) \qquad (15)$$

$$PC_{3_{Q_1}} = -2.5 + 9.6 \times 10^5 I_L - 0.38D_c - 5.4 \times 10^5 \dot{\sigma}_{14} \qquad (16)$$

$$PC_{1_{Q_2}} = 1.3 - 3.0 \times 10^6 I_L + 0.17D_c - 1.8 \times 10^5 \left(\frac{\dot{\sigma}_{10} + \dot{\sigma}_{11} + \dot{\sigma}_{12} + \dot{\sigma}_{13} + \dot{\sigma}_{14}}{5} \right) \qquad (17)$$

$$PC_{2_{Q_2}} = 0.55 + 3.8 \times 10^6 I_L + 0.28D_c \qquad (18)$$

$$PC_{3_{Q_2}} = 0.44 - 1.4 \times 10^6 I_L \qquad (19)$$

Here, I_L is the moisture convergence and $\dot{\sigma}_k$ is the vertical velocity at the kth sigma level.

3.2   Reconstructed Q_1 and Q_2 profiles

To analyze the reconstructed profiles, we have chosen five tropical domains (based on the mean February 1998 heating distribution shown in Fig. 2). The three oceanic domains are Central Pacific (CPAC, 15°S-5°N; 180°-140°W), Indian Ocean (IO, 30°S-5°N; 40°E-110°E) and Atlantic Ocean (AO, 15°S-15°N; 45°W-0°) and the two continental areas are Africa (AFR, 35°S-10°N; 0°-40°E) and Australia (AUS, 35°S-15°S; 115°E-155°E). The reconstructed Q_1 profiles for these domains are shown in Fig. 5a. The heating rate is plotted along the x-axis and the model sigma levels are plotted along the y-axis. The retrieved profiles are obtained from a linear combination of the first three dominant rotated EOFs. It is seen that the retrieved profiles match very closely with the corresponding observed profiles. The reconstructed Q_2 profiles for selected two domains (IO and CPAC) are shown in Fig. 5b. For both the domains, the retrieved profiles are close to the observed profile. The results suggest that RPCA is an effective tool for dimension reduction in which the observed vertical profiles of Q_1 and Q_2 are retrieved with a reasonable degree of accuracy by retaining a few dominant EOFs.

4.   Predictive skills of short range forecasts

In order to study the impact of the ECPS, a series of short range (3 day) forecast experiments are carried out with the FSUGSM having in-built physical initialization (Krishnamurti et al. 1991, 1993) using ECMWF analysis and observed SSM/I rain rates. The control version represents the run with the cumulus parameterization based on the modified Kuo scheme of Krishnamurti et al (1983). A brief overview of the FSUGSM control version is given in Appendix A.

4.1   A representative case study

Out of a series of experiments performed with the ECPS, the results of a representative short range forecast started with 5 February 1998 initial condition is compared against corresponding observation and FSUGSM control forecast. The day-3 precipitation forecast over the tropics for 12Z 8 February 1998 is shown in Fig. 6. It is seen that the precipitation distribution is improved globally along the TCZ and regionally over all the major convective centers except over southern Africa. The high forecast rain in the control version becomes comparable with the observed GPCP

14

rain in the ECPS forecast. The longitude-vertical section of day-3 zonal wind averaged over the equatorial belt (5°S-5°N) from the control and the ECPS forecasts are compared with that from the ECMWF analysis in Fig. 7. The circulation, though has some delayed response to the heating profile, shows improvement in the ECPS forecast. The major improvement occurs in the middle and upper troposphere, consistent with the vertical distribution of TRMM based latent heat release (Fig. 2). This suggests that the tropical circulation is sensitive to vertical distribution of large-scale convective heating and drying on this time scale. In addition, this indicates the importance of incorporating ECPS-like cumulus scheme in medium and long range forecasts especially in the context of tropical variabilities associated with Madden-Julian Oscillation (MJO). Figure 8 shows the latitude-vertical section of the day-3 forecast of meridional wind averaged over 40°E-100°E from the control, ECPS and ECMWF analysis. The structure and the magnitude of winds are closer to ECMWF analyzed winds for the ECPS forecast. The ECPS gives rise to a better forecast of Hadley cell associated with Australian monsoon as well as a more realistic Ferrel and Polar cells.

The predictive skill has been evaluated based on the root mean square error (RMSE) of 850 and 200 hPa winds and geopotential height at 500 hPa over six selected tropical domains. Figure 9 shows the estimated RMSE of control and ECPS forecasts for these domains. In general, the error distribution shows a clear improvement in forecast skill from the use of ECPS in FSUGSM. The improvement in forecast skill is quantified by estimating the percentage improvement in RMSE which is estimated as the percentage of the difference between RMSE of ECPS and control divided by the RMSE of control. The percentage improvement in 200 hPa wind and 500 hPa geopotential height fields over different tropical domains (Fig. 10) shows a very marked improvement in the short range forecasts globally and regionally when ECPS is used. The predictive skill seen in this case study is also consistently evident in other short range forecasts.

4.2   Objective verification

To emphasize the improvement in the forecast skill with ECPS, the results from three additional forecast experiments started with initial conditions corresponding to 12 January 2000, 11 March 2000 and 12 March 2000 are analyzed. For all the three experiments, the ECPS scheme was found

15

to yield improved day-1 to day-3 forecasts over the tropics. For example, the day-3 precipitation forecast over the tropics for 12Z 15 January 2000 is shown in Fig. 11. The strength and distribution of precipitation over the global tropics compare better with observation in the ECPS forecast. The precipitation forecasts around the equatorial oceanic convective centers especially over the Indian and western Pacific Oceans are also improved with the ECPS.

The RMSEs of day-1, day-2 and day-3 forecasts of winds at 850 and 200 hPa and geopotential height at 500 hPa from the three additional forecast experiments are compared with the corresponding RMSEs from the control forecasts (Fig. 12). The ECPS forecasts verify better than the control forecasts for all the three variables. In particular, the ECPS provides large improvements for the 500 hPa height forecasts. This consistent improvement in the forecast skill for all the case studies further emphasizes the importance of incorporating realistic vertical distribution of heating and drying in the model cumulus scheme. The results also suggest that in the absence of explicit models for convection, the proposed statistical scheme improves the modeling of the vertical distribution of heating and moistening in areas of deep convection.

4.3   ECPS with Q_1 and Q_2 profiles based on ECMWF analysis

In order to study the relative forecast improvement contributed by the utilization of TRMM based Q_1 dataset in ECPS, the skill in short range forecasts with the present ECPS is compared against those with ECPS utilizing profiles of both Q_1 and Q_2 derived from ECMWF analysis. The day-3 precipitation forecasts from observation and the two forecasts show that the present ECPS yields a better forecast over the global tropics (Fig. 6 and Fig. 13). Although the distribution of ECMWF based ECPS forecast resembles the present ECPS forecast over some parts, the spread of rainfall around the ITCZ region is reduced globally and compares better with the observation in the present ECPS forecast. The pattern correlation coefficient for global tropics is 0.3 for the present ECPS day-3 forecast and 0.27 for the ECPS forecast based on ECMWF analysis. In addition, although large and significant improvement was not evident, the present ECPS is found to yield improved circulation forecasts over most of the tropics compared to ECMWF analysis based ECPS.

16

5. Summary and discussion

A new empirical cumulus parameterization scheme, ECPS, which utilizes observed TRMM based latent heating (Q_1) and moistening (Q_2) derived using ECMWF analysis, has been developed. The profiles of Q_1 and Q_2 for the period of 1-28 February 1998 are used as observations. The dominant modes of variability associated with the vertical profiles of Q_1 and Q_2 over active convective regions of the tropics are obtained through rotated principal component analysis (RPCA) based on normal varimax rotation. Analysis suggests that most of the variance associated with the observed profiles can be explained by retaining the first three modes. The ECPS then applies a statistical approach in which Q_1 and Q_2 are expressed as a linear combination of the first three dominant principal components which distinctly explain variance in the troposphere as a function of the prevalent large-scale dynamics. The principal component (PC) score which quantifies the contribution of each PC to the corresponding loading profile is estimated through multiple screening regression method. This yields the PC score as a function of the large-scale variables. Thus, the dependent variables in the regression equations developed in the scheme modulates the heating about a mean convective and drying profile as a function of large scale variables. The profiles of Q_1 and Q_2 thus obtained are found to match well with observations.

The impact of the ECPS is investigated through a series of short range (1-3 day) forecast experiments using the Florida State University global spectral model (FSUGSM, T126L14). This analysis also serves as a sensitivity study of the tropical circulation to the vertical distribution of large scale convective heating and drying. Forecast comparison with the FSUGSM with modified Kuo scheme shows marked improvement in the short range forecasts using the ECPS. It is seen that the large-scale distribution and amount of precipitation is improved globally along the ITCZ and regionally over most of the convective centers over the tropics. The skill is also evident in the structure and magnitude of winds for the ECPS forecast. The ECPS not only gives rise to a better forecast of Hadley cell but leads to a more realistic Ferrel and Polar cells. This improvement in the forecast skill emphasizes the importance of incorporating realistic vertical distribution of heating and drying in the model cumulus scheme. These results also suggest that in the absence of explicit models for convection, the proposed statistical scheme improves the modeling of the vertical distribution of heating and drying in areas of deep convection.

17

Acknowledgments

APPENDIX A

The Florida State University (FSU) Global Spectral Model

The global model used in this study is identical in all respects to that used in Krishnamurti et al. (1991). The T126 version of the model, however, has been highly vectorized to reduce the model integration time. In addition, moisture variable dew point depression (T-Td) has been replaced by specific humidity and a look-up table is used for moisture calculation to further reduce the computational time. An outline of the model is as follows:

1. Variables:

    a) Independent variables: $\lambda, \theta, \sigma, t$.

    b) Dependent variables: vorticity, divergence, temperature and specific humidity.

2. Resolution:

    a) Horizontal resolution: Triangular spectral truncation; T126 resolution has a 384 x 192 Gaussian transform grid with a horizontal separation of about 100 km at 20 latitude.

    b) Vertical resolution: 14 layers in the vertical between 50 mb and 100 mb. Model variables are staggered in the vertical using Charney-Phillips vertical discretization - vorticity, divergence, wind and geopotential are located at layer interface while temperatures, specific humidity and vertical velocity are assigned at the center of the layer. The vertical grid has higher resolution in stratosphere and in planetary boundary layer.

3. Time integration scheme: The divergence equation, thermodynamic equation and pressure tendency equation are integrated implicitly while for vorticity equation and moisture continuity equation explicit time integration scheme is used. The tendencies of the physical processes are integrated using a forward time integration scheme.

4. Space differing scheme: Spectral in the horizontal; centered differences in the vertical for all variables except moisture which is handled by an upstream differencing scheme.

5. Surface topography is based on envelope orography (Wallace et al. 1983).

6. Parameterization of physical processes:

    a) Deep convection: based on modified Kuo cumulus parameterization scheme (Krishna-

19

murti et al. 1983), where the moistening and mesoscale convergence parameters are obtained from 700 mb vorticity and mean vertical velocity averaged over cloud depth through a regression relation.

b) Shallow convection (Tiedke 1984)

c) Dry convection

d) Large-scale condensation (Kanamitsu 1975). The scheme accounts for evaporation of falling precipitation.

e) Surface fluxes of heat, moisture and momentum are calculated using similarity theory (Businger et al. 1971).

f) Vertical distribution of fluxes in the free atmosphere is based on stability (Richardson number) dependent exchange coefficient (Louis, 1979).

g) 4th order horizontal diffusion (Kanamitsu et al. 1983).

h) Long and shortwave radiative fluxes are based on band model and incorporate the radiative effects of water vapor, carbon dioxide, ozone and clouds (Harshvardan and Corsetti, 1989; Lacis and Hansen, 1974).

i) Parameterization of low, medium and high clouds for radiative transfer calculation is based on threshold relative humidity. Fraction area of various cloud distribution configurations in the vertical is based on random overlap consideration.

j) Surface temperatures: Prescribed over the oceans, while over the land a surface energy balance coupled to the similarity theory determines the surface temperature including its diurnal cycle (Krishnamurti et al. 1991).

7. Initialization: The initialization of the model is achieved in two stages:

a) Nonlinear normal mode initialization: (Kitade, 1983), wherein the tendencies of first 5 modes with phase speed exceeding about 30 m s^{-1} are damped during the initialization. The slow moving higher modes are allowed to adjust freely.

b) Physical initialization wherein moisture field, heat sources and sinks and divergence fields are initialized consistent with observed OLR and rain rates.

References

Alexander, G. D., G. S. Young, and D. V. Ledvina, 1993: Principal component analysis of vertical profiles of Q_1 and Q_2 in the tropics. *Mon. Weather Rev.*, **121**, 535–548.

Anthes, R. A., 1982: Tropical cyclones– Their evolution, structure and effects. *Meteor. Monogr.*, **41**, 1–208.

Businger, J. A., J. C. Wyngard, Y. Izumi, and E. F. Bradley, 1971: Flux profile relationship in the atmospheric surface layer. *J. Atmos. Sci.*, **28**, 181–189.

Chang, C.-P., and H. Lim, 1988: Kelvin wave-CISK: A possible mechanism for the 30-50 day oscillations. *J. Atmos. Sci.*, **45**, 1709–1720.

Charney, J. G., and A. Eliassen, 1964: On the growth of the hurricane depression. *J. Atmos. Sci.*, **21**, 68–75.

Harshvardan, and T.G. Corsetti, 1984: Long-wave parameterization for the UCLA/GLAS GCM. *NASA Tech. Memo. 86072*, Goddard Space Flight Center, Greenbelt, MD 20771, 52 pp.

Hartmann, D. L., H. H. Hendon, and R. A. Houze, Jr., 1984: Some implications of the mesoscale circulations in tropical cloud clusters for large scale dynamics and climate. *J. Atmos. Sci.*, **41**, 113–121.

Huffman, G. J., R. F. Adler, P. A. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5–20.

Johnson, R. H., 1984: Partitioning tropical heat and moisture budgets into cumulus and mesoscale components: Implications for cumulus parameterization. *Mon. Weather Rev.*, **112**, 1590–1601.

Kanamitsu, M., K. Tada, K. Kudo, N. Sato, and S. Isa, 1983: Description of the JMA operational spectral model. *J. Meteor. Soc. Japan*, **61**, 812–828..

Kanamitsu, M., 1975: On numerical prediction over a global tropical belt. *Report No. 75-1*, Dept. of Meteorology, Florida State University, Tallahassee, Florida 32306, pp. 1-282.

Kitade, T. 1983: Nonlinear normal mode initialization with physics. *Mon. Weather Rev.*, **111**, 2194–2213.

Krishnamurti, T. N., S. K. R. Bhowmik, D. Oosterhof, G. Rohaly, and N. Surgi, 1995: Mesoscale signatures within the tropics generated by physical initialization. *Mon. Weather Rev.*, **123**, 2771–2789.

Krishnamurti, T. N., H. S. Bedi, and K. Ingles, 1993: Physical initialization using SSM/I rain rates. *Tellus*, **45A**, 247–269.

Krishnamurti, T. N., J. Xue, H. S. Bedi, K. Ingles, and D. Oosterhof, 1991: Physical initialization for numerical weather prediction over the tropics. *Tellus*, **43**, 53–81.

Krishnamurti, T. N., and H. S. Bedi, 1988: Cumulus parameterization and rainfall rates: Part III. *Mon. Weather Rev.*, **116**, 583–599.

Krishnamurti, T. N., S. Low-Nam, and R. Pasch, 1983: Cumulus parameterization and rainfall rates II. *Mon. Weather Rev.*, **111**, 815–828.

Krishnamurti, T. N., Y. Ramanathan, H. L. Pan, R. J. Pasch, and J. Molinari, 1980: Cumulus parameterization and rainfall rates I. *Mon. Weather Rev.*, **108**, 465–472.

Lacis, A. A., and J. E. Hansen, 1974: A parameterization of the absorption of solar radiation in the earth's atmosphere. *J. Atmos. Sci.*, **31**, 118–133.

Lau, K.-M., and L. Peng, 1987: Origin of the low-frequency (intraseasonal) oscillation in the tropical atmosphere. *J. Atmos. Sci.*, **44**, 950–972.

Lin, C., and A. Arakawa, 2000: Empirical determination of the basic modes of cumulus heating and drying profiles. *J. Atmos. Sci.*, **57**, 3571–3591.

Louis, J.F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Boundary Layer Meteorology*, **17**, 187–202.

Malkus, J. S., and R. T. Williams, 1963: On the interaction between severe storms and large cumulus clouds. *Meteor. Monogr.*, **5**, 59–64.

Mapes, B. E., and R. A. Houze, Jr., 1995: Diabatic divergence profiles in western Pacific warm pool during TOGA COARE. *J. Atmos. Sci.*, **52**, 1807–1828.

22

Misra, V., 1997: A statistically based cumulus parameterization scheme that makes use of heating and moistening rates derived from observations. *Ph. D. Dissertation*, Florida State University, Tallahassee, 184 pp.

Ogura, Y., 1975: On the interaction between cumulus clouds and the larger scale environment. *Pure. Appl. Geophys.*, **13**, 869-889.

Olson, W. S., C. D. Kummerow, Y. Hong, and W.-K. Tao, 1999: Atmospheric latent heating distributions in the Tropics derived from passive microwave radiometer measurements. *J. Appl. Meteor.*, **38**, 633-644.

Preisendorfer, R. W., 1988: *Principal component analysis in meteorology and oceanography*, Elsevier, NewYork.

Rajendran, K., Ravi S. Nanjundiah, and J. Srinivasan, 2002: Comparison of seasonal and instraseasonal variation of tropical climate in NCAR CCM2 GCM with two different cumulus schemes. *Meteor. Atmos. Phys.*, **79**, 57-86..

Reed, R. J., and E. E. Recker, 1971: Structure and properties of synoptic-scale wave disturbances in the equatorial western Pacific. *J. Atmos. Sci.*, **28**, 1117-1133.

Richman, M. B., 1986: Review Article Rotation of principal components. *Int'l. J. of Climatology*, **6**, 293-335.

Simpson, J., and W.-K. Tao, 1993: The Goddard Cumulus Ensemble Model. Part II: Applications for studying cloud precipitating processes and for NASA TRMM. *Terrestrial, Atmospheric and Oceanic Sciences*, **4**, 55-96.

Sui, C.-H., and K.-M. Lau, 1989: Origin of the low-frequency (intraseasonal) oscillations in the tropical atmosphere. Part II: Structure and propagation of mobile wave-CISK modes and their modification by lower boundary forcings. *J. Atmos. Sci.*, **46**, 37-56.

Tao, W.-K., S. Lang, W. S. Olson, R. Meneghini, S. Yang, J. Simpson, C. Kummerow, E. Smith, and J. Halverson, 2001: Retrieved vertical profiles of latent heat release using TRMM rainfall products for February 1998. *J. Appl. Meteor.*, **40**, 957-982.

Tao, W.-K., J. Simpson, C.-H. Sui, B. Ferrier, S. Lang, J. Scala, M.-D. Chou, and K. Picker-

ing, 1993a: Heating, moisture and water budgets of tropical and midlatitude squall lines: Comparisons and sensitivity to longwave radiation. *J. Atmos. Sci.*, **50**, 673–690.

Tao, W.-K., S. Lang, J. Simpson, and R. Adler, 1993b: Retrieval algorithms for estimating the vertical profiles of latent heat release: Their applications for TRMM. *J. Meteor. Soc. Japan*, **71**, 685–700.

Tiedke, M., 1984: The sensitivity of the time-mean large-scale flow to cumulus convection in the ECMWF model. *Workshop on Convection in large-scale numerical models*. ECMWF, 28 Nov. - 1 Dec. 1983, 297–316.

Tung, W.-W., C. Lin, B. Chen, M. Yanai, and A. Arakawa, 1999: Basic modes of cumulus heating and drying observed during TOGA-COARE IOP. *Geophys. Res. Lett.*, **26**, 3117–3120.

Wallace, J. M., S. Tibaldi, and A. J. Simmons, 1983: Reduction of systematic forecast errors in the ECMWF model through the introduction of envelope orography. *Quart. J. Roy. Meteorol. Soc.*, **109**, 683–718.

Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Weather Rev.*, **110**, 504–520.

Yanai, M., S. Esbensen, and J. H. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large scale heat and moisture budgets. *J. Atmos. Sci.*, **30**, 611–627.

Zawadzki, I., E. Torlaschi, and R. Sauvageau, 1981: The relationship between mesoscale thermodynamic variables and convective precipitation. *J. Atmos. Sci.*, **38**, 1535–1540.

Zawadzki, I. I., and C. U. Ro, 1978: Correlations between maximum rate of precipitation and mesoscale parameters. *J. Appl. Meteorol.*, **17**, 1327–1334.

Zhang, Z., and T. N. Krishnamurti, 1996: A generalization of Gill's Heat Induced Tropical Circulation. *J. Atmos. Sci.*, **53**, 1045-1056.

An Empirical Cumulus Parameterization Scheme for a Global Spectral Model

K. Rajendran, T. N. Krishnamurti, V. Misra and W. -K. Tao

Popular Summary

Realistic vertical heating and drying profiles in a cumulus scheme is important for obtaining accurate weather forecasts. A new empirical cumulus parameterization scheme based on a procedure to improve the vertical distribution of heating and moistening over the tropics is developed. The empirical cumulus parameterization scheme (ECPS) utilizes observed profiles of Tropical Rainfall Measuring Mission (TRMM) based apparent heat source (Q_1) and the European Centre for Medium-Range Weather Forecasts (ECMWF) analysis based apparent moisture sink (Q_2). A dimension reduction technique through rotated principal component analysis (RPCA) is performed on the vertical profiles of heating and drying over the convective regions of the tropics, to obtain the dominant modes of variability. Analysis suggests that most of the variance associated with the observed profiles can be explained by retaining the first three modes. The ECPS then applies a statistical approach in which Q_1 and Q_{-2} are expressed as a linear combination of the first three dominant principal components which distinctly explain variance in the troposphere as a function of the prevalent large-scale dynamics. The principal component (PC) score which quantifies the contribution of each PC to the corresponding loading profile is estimated through a multiple screening regression method which yields the PC score as a function of the large-scale variables.

The profiles of Q_1 and Q_2 thus obtained are found to match well with the observed profiles. The impact of the ECPS is investigated in a series of short range (1-3 day) prediction experiments using the Florida State University global spectral model (FSUGSM, T126L14). Comparisons between short range ECPS forecasts and those with the modified Kuo scheme show a very marked improvement in the skill in ECPS forecasts. This improvement in the forecast skill with ECPS emphasizes the importance of incorporating realistic vertical distributions of heating and drying in the model cumulus scheme. This also suggests that in the absence of explicit models for convection, the proposed statistical scheme improves the modeling of the vertical distribution of heating and moistening in areas of deep convection.