

Prediction of the Diurnal Change Using a Multimodel Superensemble. Part I: Precipitation

T. N. KRISHNAMURTI

Department of Meteorology, The Florida State University, Tallahassee, Florida

C. GNANASEELAN

Department of Meteorology, The Florida State University, Tallahassee, Florida, and Indian Institute of Tropical Meteorology, Pune, India

A. CHAKRABORTY

Department of Meteorology, The Florida State University, Tallahassee, Florida

(Manuscript received 28 March 2006, in final form 4 December 2006)

ABSTRACT

Modeling the geographical distribution of the phase and amplitude of the diurnal change is a challenging problem. This paper addresses the issues of modeling the diurnal mode of precipitation over the Tropics. Largely an early morning precipitation maximum over the oceans and an afternoon rainfall maximum over land areas describe the first-order diurnal variability. However, large variability in phase and amplitude prevails even within the land and oceanic areas. This paper addresses the importance of a multimodel superensemble for much improved prediction of the diurnal mode as compared to what is possible from individual models. To begin this exercise, the skills of the member models, the ensemble mean of the member models, a unified cloud model, and the superensemble for the prediction of total rain as well as its day versus night distribution were examined. Here it is shown that the distributions of total rain over the earth (tropical belt) and over certain geographical regions are predicted reasonably well (RMSE less than 18%) from the construction of a multimodel superensemble. This dataset is well suited for addressing the diurnal change. The large errors in phase of the diurnal modes in individual models usually stem from numerous physical processes such as the cloud radiation, shallow and deep cumulus convection, and the physics of the planetary boundary layer. The multimodel superensemble is designed to reduce such systematic errors and provide meaningful forecasts. That application for the diurnal mode appears very promising. This paper examines some of the regions such as the Tibetan Plateau, the eastern foothills of the Himalayas, and the Amazon region of South America that are traditionally difficult for modeling the diurnal change. In nearly all of these regions, errors in phase and amplitude of the diurnal mode of precipitation increase with the increased length of forecasts. Model forecast errors on the order of 6–12 h for phase and 50% for the amplitude are often seen from the member models. The multimodel superensemble reduces these errors and provides a close match (RMSE < 6 h) to the observed phase. The percent of daily rain and their phases obtained from the multimodel superensemble at 3-hourly intervals for different regions of the Tropics showed a closer match (pattern correlation about 0.4) with the satellite estimates. This is another area where the individual member models conveyed a much lower skill.

1. Introduction

Modeling the diurnal change of precipitation is a challenging scientific problem because of the complex

interactive nature of the physical processes that contribute to determining the diurnal cycle. The Tropical Rainfall Measuring Mission (TRMM) satellite has provided a useful mapping of the diurnal change from its precipitation estimates (Yang and Smith 2006). In this study we use the TRMM satellite-based estimates for the rain rates. The National Aeronautics and Space Administration's (NASA's) TRMM data archive includes a product (called 3B42) that provides 3-hourly rainfall

Corresponding author address: T. N. Krishnamurti, Department of Meteorology, The Florida State University, Tallahassee, FL 32306.

E-mail: tnk@io.met.fsu.edu

DOI: 10.1175/MWR3446.1

totals at a horizontal resolution of about 25 km. When we look at the phase of diurnal change (or the maximum in the diurnal precipitation) in the latitude belt 40°S–40°N from the study of Yang and Smith (2006), we note that the overall picture of afternoon rain over land areas and early morning rain over the ocean area are only partially true. Strong exception to this rule exists over many land and oceanic areas. Over the eastern Tibetan Plateau, during the northern summer season, afternoon showers are prevalent, however if we proceed 300 km to the southeast over the eastern foothills of the Himalayas, rainfall maximum occurs in the early morning hours.

In a preliminary exercise on modeling of the diurnal change of tropical precipitation we deployed a simple radiative transfer algorithm in a global model. This carried an emissivity or absorptivity lookup table-based scheme, Joseph (1970) and Katayama (1972). A first attempt failed to simulate the diurnal differences in precipitation between the Tibetan Plateau and the eastern foothills of the Himalayas. Both regions carried the same afternoon hours for a rainfall maximum in our early efforts on modeling. A quick fix was made to stabilize the afternoon sounding by reducing the albedo of high clouds, reducing the surface sensible heat flux over the sloping terrains, and a nocturnal stabilization of thin high clouds were artificially introduced by reducing the blackbody emittance for high clouds (partial blackbody). Those artificial fixes did produce the correct phases for rain over the eastern Tibetan Plateau (afternoon hours) and the eastern foothills of Himalayas (early morning hours). Those fixes in fact turned out to be disastrous for the rest of the Tropics especially over Brazil where the model failed to provide the afternoon showers. This demonstrated some of the modeling problems for the diurnal change. In general the atmospheric general circulation models (AGCMs) have problems in generating the correct phase of the diurnal cycle compared to the observed precipitation or clouds. The complex interactive nature of the physical processes that contribute in determining the diurnal cycle make it extremely difficult to separate particular components in models that might contribute to the discrepancies. Understanding of the diurnal change, which is so selective in different parts of the Tropics, requires more detailed observational and field experimental thrusts especially for the understanding of the physics involved. Reed et al. (1977) pointed out that the afternoon maximum of rain over the West Africa (near 10°N) extends well into the eastern Atlantic Ocean. Yang and Smith (2006) examined the diurnal change of precipitation using several different rain-rate algorithms. They noted that there were only some minor

differences in the tropical distribution of the phase of diurnal change of precipitation from those different algorithms, and overall there was a general agreement. The preponderance of values over the central Pacific and central Atlantic lie around 0600–0900 local time and confirm the early morning maxima of rain. The phase angle exceeding 1500 local time over Brazil and Africa confirm the afternoon phase for the overland values.

There were many other interesting features within the overall distributions of the phase; however, many of the regions with large zonal asymmetry do not carry any large diurnal amplitude. The aforementioned asymmetries within land or within ocean domain of the Tropics suggest the possible complexity involved in the modeling of these features. In this study our focus will be on addressing the diurnal change issues using multimodels. The datasets for a suite of multimodels can be extracted from an operational suite of models Krishnamurti et al. (2001). Alternatively they could be constructed by varying some components of physics in each model while retaining the rest of the model features.

This notion of developing multimodels from a single base model specifying different physical parameterization schemes for each model has been utilized for modeling studies by several authors, Krishnamurti and Sanjay (2003), Puri and Miller (1990), and others. This is a robust procedure for developing ensemble forecasts. Krishnamurti and Sanjay (2003) developed 6 different cumulus parameterization schemes and performed 720 prediction experiments (short range) with a global model using these 6 versions of cumulus parameterization. The observed precipitation estimates were extracted for the TRMM data archives. A numerical weather prediction (NWP) superensemble was constructed following Krishnamurti et al. (2000a). Here the observed rain rates were used for training and forecast evaluation. We noted in this study that the skill of the superensemble for precipitation forecast was far higher than that of the member models. The present study extends this same idea of multimodels deploying different physics [we used cloud radiative transfer for defining the different versions of the Florida State University Global Spectral Model (FSUGSM)].

Murakami (1976) examined the diurnal change of OLR over the region of West Africa and the eastern tropical Atlantic Ocean. He mapped the western extension of the afternoon phase of maximum convection (minimum OLR) over the eastern Atlantic from West Africa. Reed et al. (1977) and Thompson et al. (1979) noted important phase differences between the occurrence of cloudiness and the start of rains over the tropical Atlantic Ocean. That problem deserves to be exam-

ined in greater detail for the entire tropical belt. Given the International Satellite Cloud Climatology Project (ISCCP) datasets for cloud cover and the TRMM datasets for the precipitation, these observed phase differences could now be mapped geographically. Wallace (1975) examined the phase and amplitudes of U.S. rainfall for the rain gauge datasets. Murakami (1983) defined highly reflective clouds from a very low threshold of OLR for looking at the phase and amplitude of heavy rain. This relates heavy rain derived from a very low value of OLR; however, the study did not distinguish a phase difference between OLR minimum and the heaviest rains.

The newly launched CloudSat is expected to provide some of the finest datasets that might help us in providing better insights for this diurnal change problem. In our present study we use a suite of four FSU global spectral models, all of which utilize a band model for the radiative transfers based on the studies of Lacis and Hansen (1974) for the shortwave component and Harshvardan et al. (1987) for the infrared radiation. These models differ in their cloud specifications. The main focus of the paper is to emphasize the capabilities of multimodel superensemble methodology to a very difficult and most important scientific problem, which is modeling the diurnal change of precipitation. We understand the inherent difficulties in modeling the diurnal change of precipitation using a single model. So we must use an approach such as the one considered in this study to provide a better forecast of the diurnal change of precipitation.

2. Models and methodology

a. Different models for cloud parameterization

Four different versions of the FSUGSM were used in this study, which deal with a common radiative transfer algorithm, Lacis and Hansen (1974). The model has 27 vertical levels with finer levels near the surface and tropopause, and uses triangular truncation at 126 waves (T126), which corresponds to about 0.94° grid in the physical space. The main features of the model are given in appendix A. Our intention is to utilize four models with somewhat different physical parameterization algorithms to improve forecasts. This type of formulation had been successfully implemented by Krishnamurti and Sanjay (2003), Puri and Miller (1990), and several others. Krishnamurti and Sanjay (2003) used six different cumulus parameterization schemes. The present study developed from a similar model construction utilized four different cloud radiative transfer schemes. The use of different cloud radiation schemes in the radiative transfer may not be most ideal for con-

structing multimodels for addressing the issue of diurnal change of rainfall. The idea to use cloud radiation specification for the design of multimodels came from our interest in improving the forecasts of the diurnal change of clouds, Chakraborty et al. (2007). At first we felt that those multimodels may be suitable for addressing improvements of forecasts of the cloud cover and not for the rainfall problem. The design was put together to develop multimodel superensemble-based forecasts for the diurnal change issues for low, middle, and high clouds (Part II of this paper, Chakraborty et al. 2007). It turned out that with those different cloud radiative transfer schemes, we were able to obtain a robust ensemble for the examination of the diurnal change of precipitation. This choice of models in fact provided a stringent test for the diurnal change since the sensitivity of cloud radiative interaction for day 1 of forecasts from this suite of models was quite small. However, we did find substantial differences in these member models for day 5 of the forecasts. We also report here on the large phase and amplitude errors of the diurnal change of precipitation for the individual member models that are largely corrected from the construction of a multimodel superensemble. The cloud parameterization schemes used in all the above four versions of FSUGSM are described below.

1) FSU NEW CLOUD PARAMETERIZATION [MODEL 1 (M1)]

In this parameterization scheme, clouds are allowed to form at any layer of the model. The presence of clouds in a layer is determined by a threshold relative humidity, which varies for three different cloud types of the model. The cloud amount in a layer is a function of the mass-weighted relative humidity of that layer giving that it crosses the threshold value, and is calculated from this relation:

$$cc = \left(\frac{rh - rh_c}{1 - rh_c} \right)^2, \quad (1)$$

where rh is the mass-weighted relative humidity of that layer and rh_c is the threshold relative humidity, which is a function of vertical coordinate (sigma layer values). The specified values for rh_c for three different cloud types are given in Table 1. Furthermore, the high cloud amount [obtained from Eq. (1)] is corrected when the precipitation rate is nonzero, using the following relation [Eq. (2)]:

$$cc = 4(0.247 + 0.126 \ln p - 0.3)^2, \quad (2)$$

TABLE 1. Threshold relative humidity for three different cloud types of model 1 and model 2.

Cloud type	Model 1		Model 2	
	rh _c	α	rh _c	α
High	0.40	1.67	0.40	1.67
Middle	0.45	2.00	0.45	2.00
Low	0.70	2.94	0.66	2.94

where p is the convective precipitation rate. Simulations with this cloud parameterization scheme are referred to hereafter as model 1 (M1).

2) FSUOLD CLOUD PARAMETERIZATION [MODEL 2 (M2)]

This cloud parameterization is similar to the FSUnew scheme (M1) in which the threshold relative humidity is the criteria for forming clouds in the model at any given layer. But it differs from M1 in the formulation of the high cloud. No correction of high cloud amount is made in this scheme depending on the convective precipitation rate as is done for M1. Cloud amount in a grid at a layer is computed from the following formula:

$$cc = [\alpha(rh - rh_c)]^2, \quad (3)$$

where α is a proportionality factor, rh_c is the threshold relative humidity, and rh is the mass-weighted relative humidity in that layer. The values of α and rh_c used by the parameterization scheme for three different cloud types are given in Table 1. This cloud parameterization scheme will be referred to hereafter as model 2 (M2).

3) NCAR-CCM3 CLOUD PARAMETERIZATION [MODEL 3 (M3)]

This cloud parameterization scheme is based on Slingo et al. (2003). In this scheme, convective cloudiness is calculated from the mean convective motion in an atmospheric column. Layered clouds are considered as functions of stability of the atmospheric column. Low clouds are allowed to be present only if the vertical velocity is less than 50 mb day⁻¹. Clouds associated with the low-level inversions are determined from relative humidity criteria. This cloud parameterization scheme will be referred to hereafter as model 3 (M3).

4) PLEIM-XIU CLOUD PARAMETERIZATION [MODEL 4 (M4)]

This cloud parameterization scheme is based on Pleim and Xiu (1995). A cloud is assumed to be present

in a layer if the mean (layer) relative humidity exceeds a threshold value. This threshold relative humidity is a function of the pressure of the layer:

$$rh_c = 1 - 2 \frac{P}{P_s} \left(1 - \frac{P}{P_s} \right) \left[1 + \sqrt{3} \left(\frac{P}{P_s} - 0.5 \right) \right], \quad (4)$$

where P is pressure of the layer and P_s is the surface pressure. In addition, in the nonconvective planetary boundary layer, the maximum relative humidity threshold is set to 0.90. The cloud fraction in a layer is derived from the mass-weighted relative humidity in that layer and the threshold relative humidity using the following relation:

$$cc = \left(\frac{rh - rh_c}{1 - rh_c} \right)^2, \quad (5)$$

where rh is the mass-weighted relative humidity in the model layer and rh_c is the threshold relative humidity derived using Eq. (4). This cloud scheme will be referred to hereafter as model 4 (M4). We noted that the sensitivity to forecasts grew with time and were substantial by day 5 of the forecasts and that was a motivation for the choice of these models.

b. Experimental details

The above four cloud parameterization schemes are used to construct multimodel ensemble forecasts with the FSUGSM. All these schemes calculate cloud fraction diagnostically from the large-scale parameters such as relative humidity. These cloud schemes vary in the exact definition of different cloud types in the model. For example, FSUold does not correct convective cloud amount from the precipitation rate, which is a feature of the FSUnew and the National Center for Atmospheric Research (NCAR) cloud schemes. To construct ensemble forecasts, all other components of the model except the cloud parameterization scheme were kept unchanged between different members. Five-day-long forecasts were carried out with all the four versions of the model starting at 1200 UTC 1 January 2000–1200 UTC 31 March 2000. Initial conditions were extracted from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40). Sea surface temperature (SST) boundary conditions were obtained from the Reynolds and Smith (1994) weekly datasets and interpolated linearly to the model run time. Model outputs were stored at every 3-h interval for understanding the diurnal cycle. Starting at 1500 UTC of the day of the initial condition to 1200 UTC of the next day, all these eight (3-h) forecast time

points together are termed as day-1 forecasts in this study. Day-2 forecasts follow day-1 forecasts and similarly for the other forecast lead times up to day 5. A time series of day-1 forecasts were created for 1500 UTC 1 January 2000–1200 UTC 1 April 2000 by collecting the string of day-1 forecasts valid for those 91 days. Similarly a time series of day-2 forecasts were created for 1500 UTC 2 January 2000–1200 UTC 2 April 2000, and likewise for the other lead times. This construction of the time series of day-1 to -5 forecasts enabled us to study the skills of different forecast products at all lead times for multiple days.

c. Superensemble methodology

The superensemble approach is a recent contribution to the general area of weather and climate forecasting, developed at FSU; this has been discussed in a series of publications, Krishnamurti et al. (1999, 2000a,b, 2001). This technique entails the partition of a time line into two parts. One part is the “training” phase, where forecasts by a set of member models are compared to the observed or the analysis fields with the objective of developing a statistics on the least squares fit of the forecasts to the observations and the second part is the “forecast” phase where estimates for a_i from the training phase are used to create the superensemble prediction. The performance of the individual models is obtained in the training phase using multiple linear regression against observed (analysis) fields. The outcome of this regression is the weights assigned to the individual models in the ensemble, which are then passed on to the forecast phase to construct the superensemble forecasts. In fact, the temporal anomalies (model) of the variables are regressed against the observed anomalies and so in formulating the superensemble forecasts, the weights are multiplied by the corresponding model anomalies. The constructed forecast is

$$S = \bar{O} + \sum_{i=1}^N a_i (F_i - \bar{F}_i), \quad (6)$$

where \bar{O} is the observed climatology over the training period, a_i is the weight for the i th member in the ensemble, and F_i and \bar{F}_i are the i th model’s forecasts and the forecast mean (over the training period), respectively. The N is the number of member models. The weights a_i are obtained by minimizing the error term G :

$$G = \sum_{i=1}^{N_{\text{train}}} (S'_i - O'_i)^2. \quad (7)$$

Here N_{train} is the number of time samples in the training phase, and S'_i and O'_i are the superensemble and

observed field anomalies, respectively, at training time t . This exercise is performed at every model grid points. The skill of the multimodel superensemble method significantly depends on the error covariance matrix (built with the model field anomalies F'_i and F'_j , where F'_i and F'_j are the i th and j th model anomalies, respectively), since the weights of each model are computed from the designed covariance matrix:

$$\mathbf{C} = [c_{i,j}] = \left[\sum_{t=1}^{N_{\text{train}}} F'_i(t)F'_j(t) \right]. \quad (8)$$

The construction of the superensemble utilizes a least squares minimization principle within a multiple regression of model output against observed (analysis) estimates. This entails a matrix inversion that is being solved by the Gauss–Jordan elimination technique. That matrix can be ill conditioned and singular depending on the interrelationships of the member models of the superensemble. Yun et al. (2003) recently designed a singular value decomposition (SVD) method that overcomes this problem and removes the ill conditioning of the covariance matrix completely. Early tests of this method showed great skills in weather and seasonal climate forecasts compared to the Gauss–Jordan elimination method. Using SVD the covariance matrix \mathbf{C} can be written as $\mathbf{C} = \mathbf{U}\mathbf{W}\mathbf{V}^T$, where \mathbf{U} and \mathbf{V} are $N \times N$ orthogonal matrices and \mathbf{W} is an $N \times N$ diagonal matrix. Since \mathbf{C} is a square symmetric matrix, $\mathbf{C}^T = \mathbf{C}$, equivalently $\mathbf{V}\mathbf{W}\mathbf{U}^T = \mathbf{U}\mathbf{W}\mathbf{V}^T$.

This proves that \mathbf{U} and \mathbf{V} are the same and the SVD is equivalent to the principal component analysis. The regression coefficients (or weights a_i) are then obtained by

$$[a_i] = \left\{ \mathbf{V} \left[\text{diag} \left(\frac{1}{w_i} \right) \right] \mathbf{U}^T \right\} [\delta'_i], \quad (9)$$

where $\delta'_i = \sum_{t=1}^{N_{\text{train}}} O'_i F'_i(t)$. This pointwise regression model using the SVD method removes the singularity problem that cannot be entirely solved by the Gauss–Jordan elimination method.

The weighted sum of the member model forecasts can provide a better product in terms of reduced systematic errors where the weights are obtained by regressing the member model forecasts against the observed (or analysis based) counterparts over the training period. These weights obtained from the training phase are different at different grid points (horizontal and vertical), at different forecast times, for each model and each variable, and are passed on to the forecast phase. The superensemble forecast is the weighted sum of the member model forecasts. This procedure removes the collective systematic bias of the member

models and is thus a robust approach for reducing the forecast errors.

A single set of weights is obtained for each of the 3 h of forecasts during a day and at each forecast lead time. The training period was taken from the initial 1500 UTC 1 January 2000–1200 UTC 27 March 2000 simulations and the rest of the simulation period is chosen for the forecast. So the day-1 forecast covers the period of 1500 UTC 27 March–1200 UTC 1 April and the day-2 forecast covers the period of 1500 UTC 28 March–1200 UTC 2 April and so on. Up to day 5 forecasts are constructed in this manner from member models, the ensemble mean, and the superensemble. The weights for day- n ($n = 1, 2, \dots, 5$) forecasts were calculated from day- n simulations of the training period. This was done because of the fact that both bias and the weights of the models depend on the lead times. The Fourier transform of 3-hourly precipitations was carried out separately for every day to calculate diurnal cycle of several days. The diurnal cycle is defined here as the first harmonic of the transformed series. The hour-by-hour average of the first harmonic of the filtered data over the total number of forecast days is the representative diurnal cycle for that period of time. It was found that diurnal cycle obtained using this method is very close (within 1% of errors) to the n th harmonic of the n -day-long time series of data. Hence, n refers to day 1, 2, 3, 4, or 5.

An ensemble mean is defined as the average of all models involved in an ensemble suite. A bias-removed ensemble mean is defined as $EM_{br} = \bar{O} + (1/N) \sum_{i=1}^N (F_i - \bar{F}_i)$. In both of these forecasts equal weights are given to all the member models. Putting $a_i = 1/N$, for $i = 1, 2, \dots, N$ in Eq. (6) one would obtain the bias-removed ensemble mean. A question is being frequently asked whether a simple bias removal of each model followed by a straightforward ensemble mean might have skills equivalent to those of a multimodel superensemble. The bias-removed ensemble mean does not perform as well as a multimodel superensemble because it provides equal weights to all the models, whereas the superensemble is more selective since it assigns different fractional positive and negative weights to different models and these weights are geographically distributed. The skill of the unified cloud model is found superior to the member models, and the RMSE and correlation have been used as skill measures.

d. Precipitation from a unified cloud model

An improved single model with the unified cloud parameterization scheme has been developed from the postprocessing of the superensemble. The precipitation

products (from the unified cloud model) are compared with the individual model precipitation estimates and the precipitation from the ensemble mean and the multimodel superensemble. The notion of the unified model was originally developed for one of our earlier studies (Krishnamurti and Sanjay 2003). The unified cloud scheme is built within a single model where the weighted sum of the cloud schemes is used. This exercise was started with a calculation of weights of the member models for three different cloud types (low, middle, and high clouds) for January and February 2000. The method for the calculation of weights is identical to what we use for the construction of the superensemble. ISCCP cloud fractions were extracted for validation purposes. In the next step, all the four schemes were run in parallel as an integral part of one model to obtain the predicted cloud fractions. Outputs from these cloud parameterization schemes include the cloud fractions for each layer of the model. These eight weights for different cloud fractions, calculated previously, are applied to a single model to construct a unified forecast. All the layers below the 700-hPa level utilize the weights calculated for the low cloud fraction. Layers between 400 and 700 hPa utilize weights for middle clouds and layers above 400 hPa utilize weights for high clouds. Unified cloud fractions were calculated for each of these layers of the model. These cloud fractions are then passed on to the other parts of the model (e.g., shortwave and longwave radiation calculations) and interact fully as the forecast evolves. The unified cloud scheme increases the computing time of the model runs only minimally. This new cloud parameterization scheme is both statistical and physical based. It combines the physically based parameterization schemes based on their local past performances. This cloud scheme is designed to correct the best parameterization scheme of the suite of models. Since this scheme is flexible in terms of the number of models in the ensemble, any number of input member models can, in principle, be used to construct the unified cloud scheme. This new unified scheme is an integral part of one model and thus has the potential to improve forecasts of other parameters of the model in addition to the variable for which the scheme is built upon.

e. TRMM data

The TRMM rainfall product used in this study is based on the TRMM 3B-42 rain-rate algorithm that uses an optimal combination of 2B-31, 2A-12 (Yang and Smith 2006), the Special Sensor Microwave Imager (SSM/I), the Advanced Microwave Scanning Radiometer (AMSR), and the Advance Microwave Sounding Unit (AMSU) precipitation estimates, Adler et al.

(2000). Its temporal coverage starts from January 1998 with a resolution of 3 h, the spatial coverage includes the global belt extending from 50°S to 50°N with a spatial resolution of $0.25^\circ \times 0.25^\circ$.

The 3B-42 estimates are produced in four stages: (i) the microwave-based estimates of precipitation are calibrated and combined, (ii) infrared-based precipitation estimates are created using the calibrated microwave precipitation, (iii) the microwave and IR estimates are combined, and (iv) rescaling to monthly datasets is applied. Each precipitation field is best interpreted as the precipitation rate effective at the nominal observation time.

3. Results and discussion

a. Prediction of diurnal precipitation

In Fig. 1, we first show the forecasts of total rainfall (for day-2 forecasts), which carries the geographical distribution of observed rain and those from forecasts based on the superensemble, the ensemble mean, a model that carried the highest RMSE skill, and those based on the unified cloud model. The member model, the ensemble mean, and the unified cloud model all carry somewhat larger errors, their patterns show wider and more intense rainfall belts over the ITCZ especially over the Pacific Ocean. We also note that the forecast skill of daytime is slightly higher than that of nighttime hours. This may have to do with the physics for the nighttime hours such as the modeling of nocturnal boundary layers and radiative processes that impact the nighttime rainfall forecasts especially over ocean. The spread of rainfall distributions over land area for the member model forecast is somewhat wider than the observed distributions. The superensemble is able to correct such systematic errors to some extent. There are more geographical areas where the superensemble rains are in closer agreement with the observed estimates. There are clearly a few pockets of heavier rains in the TRMM estimates that are not picked up at the same level by the superensemble. However, it should also be noted that most member models carry a much larger number of such pockets of heavy rains thus bringing their RMSE much higher than that of the superensemble. The superensemble forecast further improved by day 5 of the forecast (in total rain as well as day versus night rain), the skill comparison is shown in Table 2. This improvement might be due to the better systematic bias correction in the superensemble.

The predicted phase and amplitude of the diurnal mode over the Tropics for day 2 are illustrated in Fig. 2. The oceans generally carry a phase of 0600–0900 local time (early morning). The land area of South America

and Africa carry a phase of about 1500–1800 local time (late afternoon). These features are best described by the superensemble forecasts. The RMSE of the superensemble for phase is less than 6 h whereas models as well as ensemble mean carry an RMS error close to 8 h. The RMSEs for the amplitudes also show a slight improvement for the superensemble (3.48 mm day^{-1}) compared to that of the best model and the ensemble means (3.84 and 3.7 mm day^{-1} , respectively). Overall the results of prediction of the diurnal modes did not alter very much in days 1–5 of the forecasts. Most of the salient features in the land–ocean differences were captured by the superensemble. The RMSE in diurnal amplitude and phase for day-2 and -5 forecasts are shown in Table 3. The superensemble showed better day-5 forecast skill for both the diurnal amplitude and phase.

b. Regional variation in the diurnal precipitation

In this section we discuss the regional features of phase and amplitude errors in diurnal precipitation forecasts. The diurnal cycle of rainfall over the eastern foothills of the Himalayas located between $24^\circ\text{--}27^\circ\text{N}$ and $90^\circ\text{--}93^\circ\text{E}$ for days 2 and 5 of the forecasts are shown in Figs. 3a,b, respectively. During late March and early April, the observed diurnal change over this region carries a phase of maximum rain during the early morning hours (i.e., close to 0300 local time). We note that most models have very large phase errors and they place this maximum close to 0900 local time. This carries an almost 6-h phase shift with each forecast. The amplitude of the diurnal mode is slightly underestimated by the superensemble. Between days 2 and 5 of the forecasts we note a growth of phase errors for the member models. The superensemble forecast preserves a close agreement with the observed diurnal change of precipitation during the 5 days of forecasts.

This diurnal cycle of rainfall over the Tibetan Plateau located between $31^\circ\text{--}34^\circ\text{N}$ and $89^\circ\text{--}92^\circ\text{E}$ showed that during the period of late March and early April, it experiences an early afternoon rainfall maximum (Figs. 3c,d). The TRMM SSM/I carries an afternoon maximum for the observed rain around 1800 local time. Almost all of the member models as well as the superensemble captured the phase very well. The phase shift was almost negligible for the superensemble and was less than 3 h for all the models and the ensemble mean.

Over the Amazon Valley region between $15^\circ\text{--}5^\circ\text{S}$ and $60^\circ\text{--}50^\circ\text{W}$, the advantage of reasonable day-1 forecast by some of the models (figure not shown) was lost by day 2 of the forecast (Fig. 3e) and large phase and amplitude errors developed by day 5 of the forecasts (Fig. 3f). The observed rainfall maximum is noted around 1800 local time, whereas the member models,

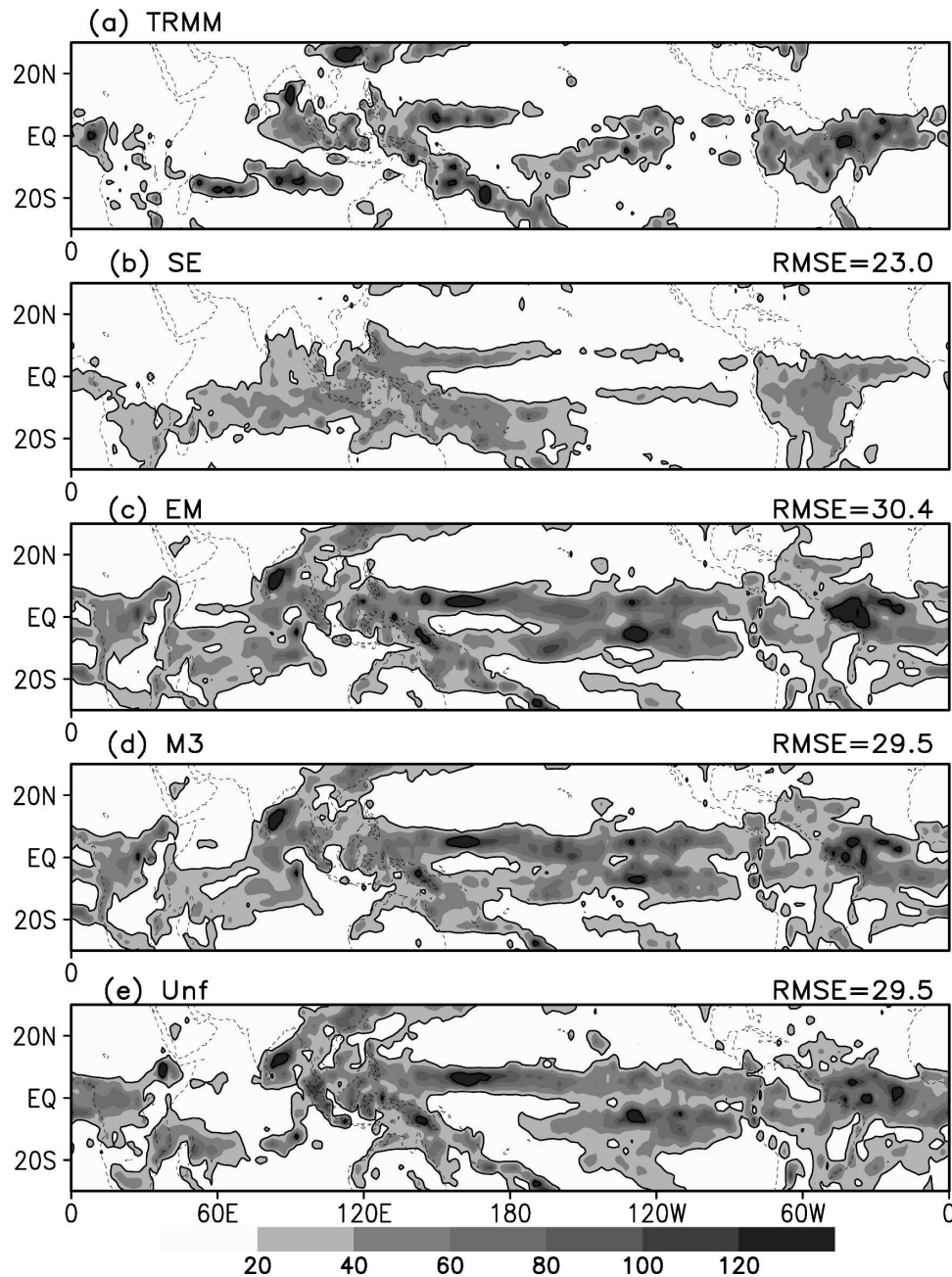


FIG. 1. Distribution of total rainfall accumulation (mm) during 1500 UTC 28 Mar–1200 UTC 2 Apr 2000 (day-2 forecast) for (a) TRMM observed, (b) superensemble (SE), (c) ensemble mean (EM), (d) M3, and (e) the unified cloud model (UNF).

the ensemble mean and the unified model carry the maximum at around the local noon hour. This emphasizes the possibility of serious problems with the physical parameterization for these member models. These errors amplify between days 1 and 5 of the forecasts. At this stage we are not able to find the exact sources of these model errors. But the sources of errors the diurnal change emanates may occur from the improper pa-

rameterization of shallow and deep convection, the planetary boundary layer (PBL) theories, and cloud radiative processes. The superensemble retains an accurate phase and amplitude on day 5 of forecasts (RMSE < 6 h and < 3.4 mm, respectively).

The mechanisms responsible for the observed diurnal variation in convection over land and the ocean appear to be different. Over the warm ocean where deep con-

TABLE 2. RMSE in diurnal precipitation forecast for days 2 and 5.

	Day-2 total rain	Day-5 total rain	Day-2 day rain	Day-5 day rain	Day-2 night rain	Day-5 night rain
SE	23.0	18.8	12.4	10.3	13.0	11.0
EM	30.4	29.6	16.0	15.0	16.9	16.5
M3	29.5	33.2	15.5	17.1	16.5	18.2
UNF	29.4	30.6	15.2	16.0	16.5	17.2

vection occurs, many mechanisms can be involved in controlling the diurnal cycle. The most prominent is probably the direct radiation convection interaction proposed by Randall et al. (1991). In this, the infrared cooling at night from the cloud top is greater than at the cloud base, which results in destabilization of the upper troposphere and hence favoring cloud development in the early morning; in contrast, during day the warming at cloud top due to solar absorption increases stability and therefore restricts convective activity. Another possible mechanism was proposed by Gray and Jacobson (1977) and involves the horizontal cloud versus cloud-free radiation difference. At night the radiative cooling of cirrostratus in the upper troposphere is greater than the radiative cooling of the surrounding less cloudy and cloud-free regions; during the day the situation is reversed. This day–night differential heating cycle gives rise to a diurnal variation in horizontal divergence, which may give rise to a diurnal variation in convective activity. Chen and Houze (1997) have argued that the mechanism is not just a cloud–radiation interaction but rather a much more complicated three-way interaction between the surface, clouds, and radiation. They also argue that cloud life cycle effects are also important with the remnants of clouds from the previous day’s conditioning of the near-surface boundary layer with air that has lower moist static energy and hence producing regions unfavorable for development of convection the following day.

Here we shall consider the major tropical ocean basins (Pacific, Indian, and the Atlantic) between 30°S and 30°N, Fig. 4. The phase errors over the Pacific and the Atlantic Ocean basins were very large for both days 2 and 5 of the forecasts for the member models, the ensemble mean and the unified model. The superensemble forecast holds a distinct superiority over all these model forecasts by removing the systematic errors and providing the correct phase and amplitude. The oceanic maxima of rain occur in the early morning hours between 0600 and 0900 local time for the Pacific and Indian Oceans and around noon local time for the Atlantic Ocean. The model forecasts carried large phase errors of approximately 6–9 h. It should be noted

that these observed phases are based on the entire basin-wide averages and are not entirely representative of each location of a basin. There is considerable variability in phase within a basin as was noted in Fig. 4. The errors for the member models over the Atlantic Ocean, Figs. 4e,f, were some of the largest. The member models carried phase errors as large as 9 h and much larger amplitude for the diurnal change compared to the TRMM SSM/I-based observed estimates. These errors were evident for days 1–5 of the forecasts.

Over land, during the early part of the day the land surface is heated by solar radiation and this increases the air temperature in the lower troposphere and hence instability, which then leads to convection with the resultant maximum in the convective precipitation tending to occur in the evening. At night the strong radiative cooling of the land enhances the stability suppressing convection leading to a minimum in the early morning. Of course local effects, particularly local orography can considerably modify these somewhat idealized concepts.

The observed and predicted tropical diurnal (mode of) precipitations over land areas for days 2 and 5 of the forecasts (covering the period 1500 UTC 28 March–1200 UTC 2 April 2000 and 1500 UTC 31 March–1200 UTC 5 April 2000, respectively) are illustrated in Figs. 4g,h, respectively. These cover the land areas between 30°S and 30°N. The observed maximum of diurnal rainfall (taking the entire landmass average over the Tropics) occurs around 2100 local time. For day 2 of the forecasts, the observed and the superensemble forecasts of the diurnal mode are in excellent agreement. The multimodels, the ensemble mean, and the unified model all carry maxima about 6 h earlier with reasonably comparable amplitudes. For day 5 of the forecasts, the phase and amplitude of the observed rain are clearly replicated in its entirety by the multimodel superensemble. The member models carry very small amplitude for the diurnal mode since they tend to carry opposite phases for different parts of the land areas. The unified model carries larger amplitudes, but its phase for the diurnal maximum of rain is noted at 1500 local time, which is roughly 6 h earlier than the observed.

c. Statistical analysis and skills

The superensemble forecasts are seen to perform better than the bias-removed ensemble mean forecast (Stefanova and Krishnamurti 2002). Although the superensemble provides a deterministic forecast, they showed that the superensemble algorithm carries an equivalent probabilistic forecast as well. A probabilistic forecast is one that estimates the probability of occur-

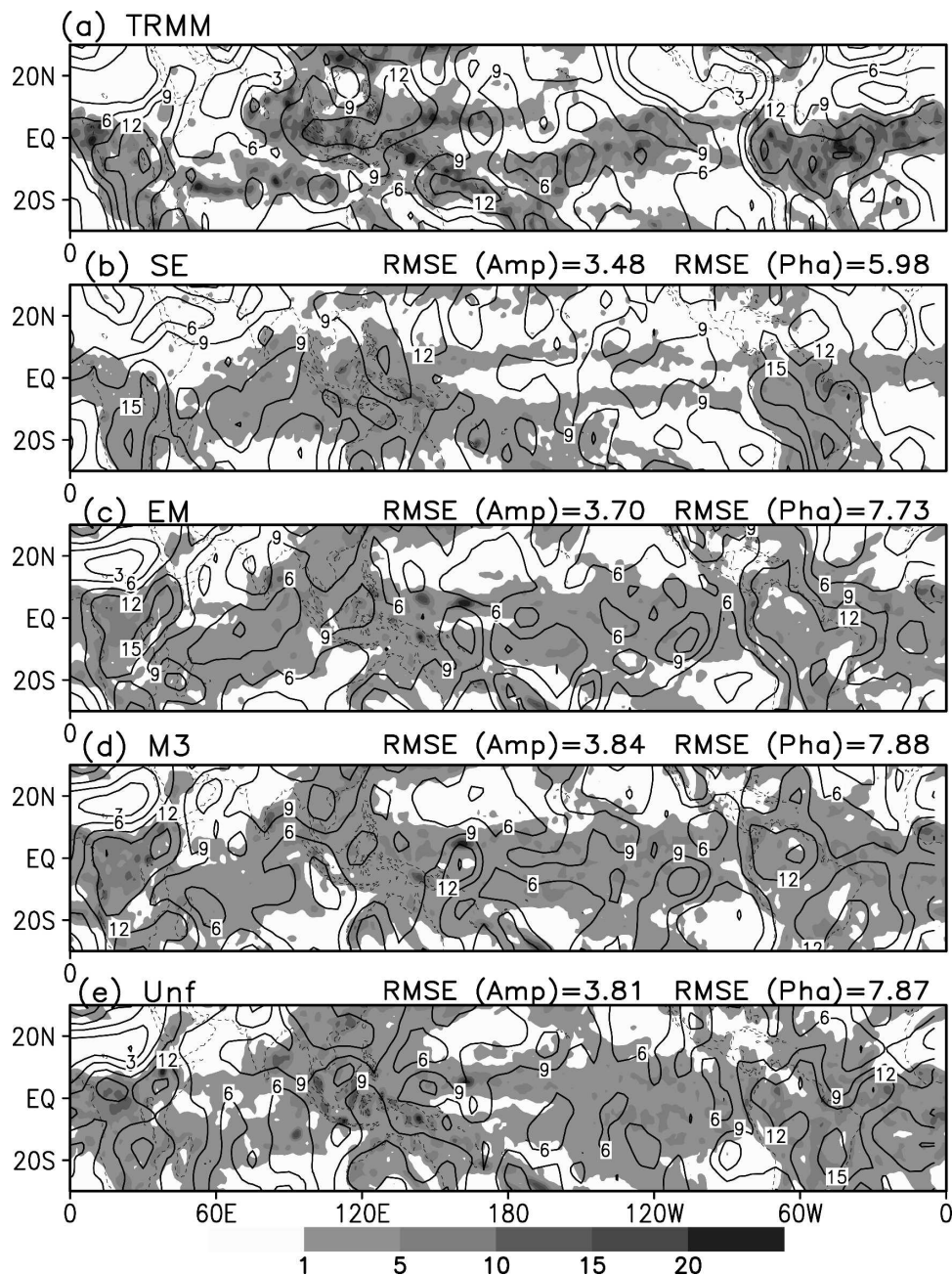


FIG. 2. Distribution of diurnal amplitude (mm day^{-1}), which is shaded, and phase (hours, local time) shown by the contours during 1500 UTC 28 Mar–1200 UTC 2 Apr 2000 (day-2 forecast) for (a) TRMM observed, (b) SE, (c) EM, (d) M3, and (e) UNF.

rence of a chosen event E . The event type selected for this study is the precipitation rate anomaly relative to the mean state exceeding a preselected threshold level. For an ensemble of equally reliable models the probability P of the event E is $(m/N) \times 100\%$, where m is the number of ensemble members forecasting E and N is the total number of ensemble forecasts. One of the most widely used methods for verification of probabil-

TABLE 3. RMSE in diurnal amplitude (mm day^{-1}) and phase (hours, local time) of day-2 and -5 forecasts.

	Day-2 amplitude	Day-5 amplitude	Day-2 phase	Day-5 phase
SE	3.48	3.31	5.98	5.86
EM	3.70	3.84	7.73	8.23
M3	3.84	4.24	7.88	8.20
UNF	3.81	3.94	7.87	8.35

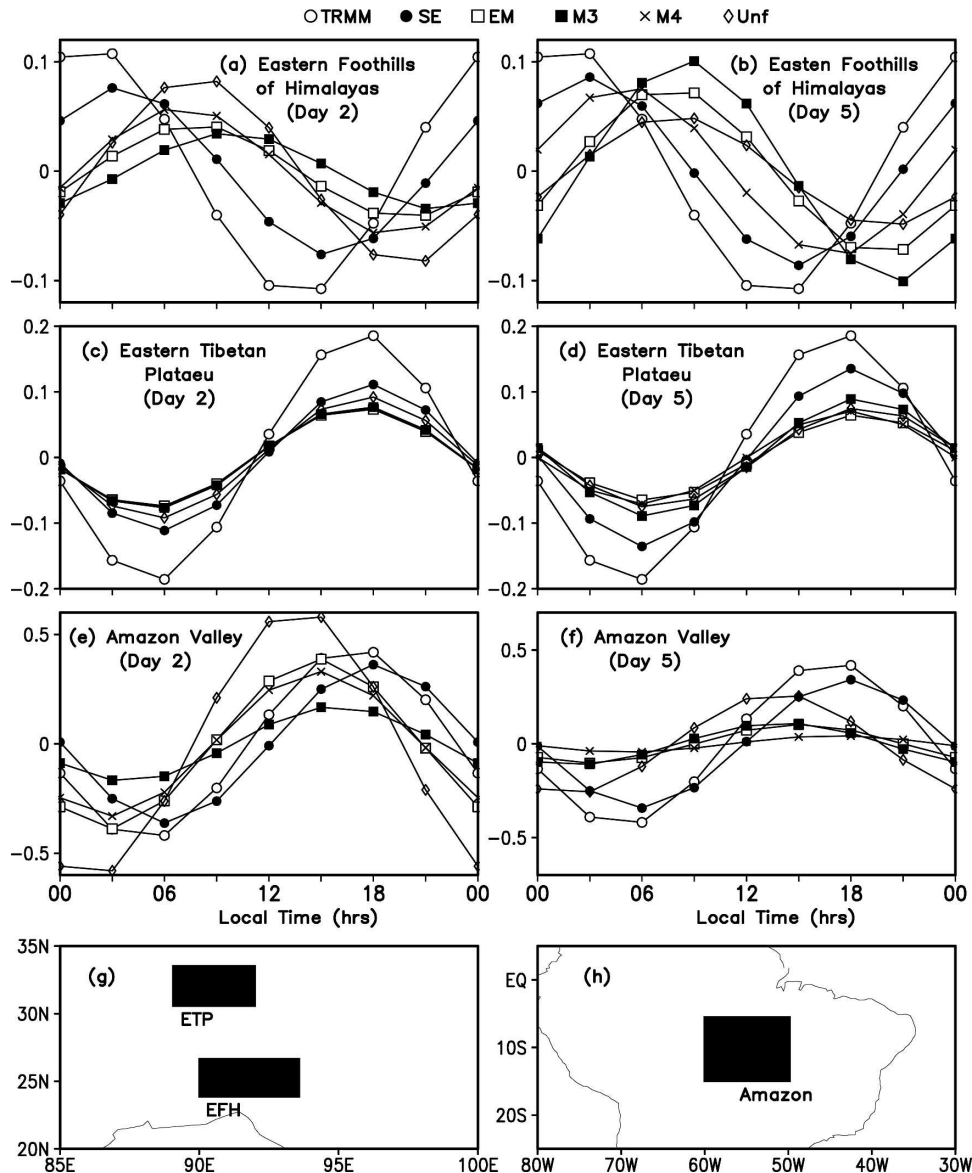


FIG. 3. Diurnal cycle of rainfall [mm (3 h)^{-1}] at local time for TRMM, SE, EM, M3, M4, and UNF over (a) the eastern foothills of the Himalayas (EFH) (24° – 27°N , 90° – 93°E) for day 2, (b) EFH for day 5, (c) the eastern Tibetan Plateau (ETP) (31° – 34°N , 89° – 92°E) for day 2, (d) ETP for day 5, (e) the Amazon Valley (15° – 5°S , 60° – 50°W) for day 2, and (f) the Amazon for day 5; and (g), (h) area maps.

ity forecast is the Brier skill score. The Brier score computation procedure is briefed in appendix B.

The superensemble probability forecasts are compared with the conventionally defined probability forecasts from the multimodel bias-removed ensemble where all N individual unbiased forecasts $F_i - \bar{F} + \bar{O}$ are weighted equally with a uniform weight $1/N$. The corresponding reliability diagrams are shown in Fig. 5. The events considered are precipitation anomalies with respect to the series mean exceeding a threshold (1.5, 2.5, 4.5 mm day^{-1}) for all points of the global Tropics.

The reliability diagrams clearly show that the superensemble gives more reliable forecast (better than the ensemble mean forecast). However, it is not giving the perfect forecast as can be seen from Fig. 5. Both superensemble and ensemble forecasts deviate much from the perfect forecast (i.e., from the diagonal line sloping at 45°). One possible reason for this could be that these scores are computed by grouping the entire global Tropics into one time series and therefore the variability among geographical locations might influence the skill score as the Brier skill score was initially designed

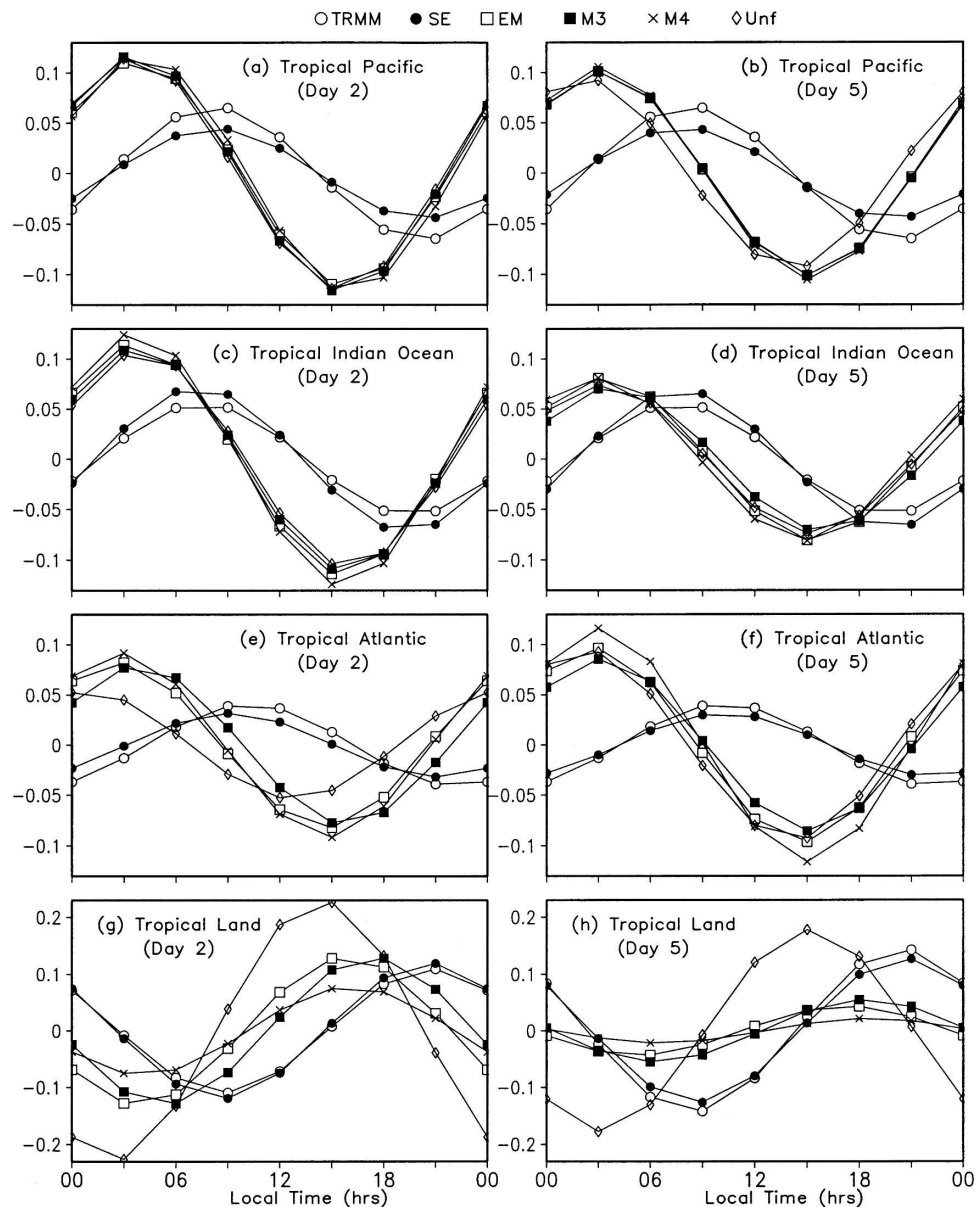


FIG. 4. Diurnal cycle of rainfall [mm (3 h)^{-1}] at local time for TRMM, SE, EM, M3, M4, and UNF for (left) day-2 and (right) day-5 forecasts over the tropical (30°S – 30°N) ocean basins: (a), (b) the tropical Pacific; (c), (d) the tropical Indian Ocean, and (e), (f) the tropical Atlantic; and (g), (h) over the tropical landmass.

for a time series at a single point. However, it is found that superensemble gives much improved probability (or reliability) forecast for all the thresholds. The Brier skill score values for reliability for both days 2 and 5 of the forecasts are shown in Table 4. The reliability scores clearly show that the superensemble is doing much better forecasting, roughly 35% improvement over the ensemble mean for all the thresholds and all the days (up to 5 days) of the forecast. Also from a significant test on difference in RMSE (see appendix

C), it was found that the improvement of superensemble forecast over the ensemble mean forecast is statistically significant at a level generally $>99\%$ for all the days of forecasts.

The equitable threat score (ETS) is the number of correct “yes” forecasts divided by the total number of occasions on which that event was forecasted and/or observed. It can be viewed as a hit rate for the quantity being forecasted, after removing the correct “no” forecasts from consideration. The worst possible threat

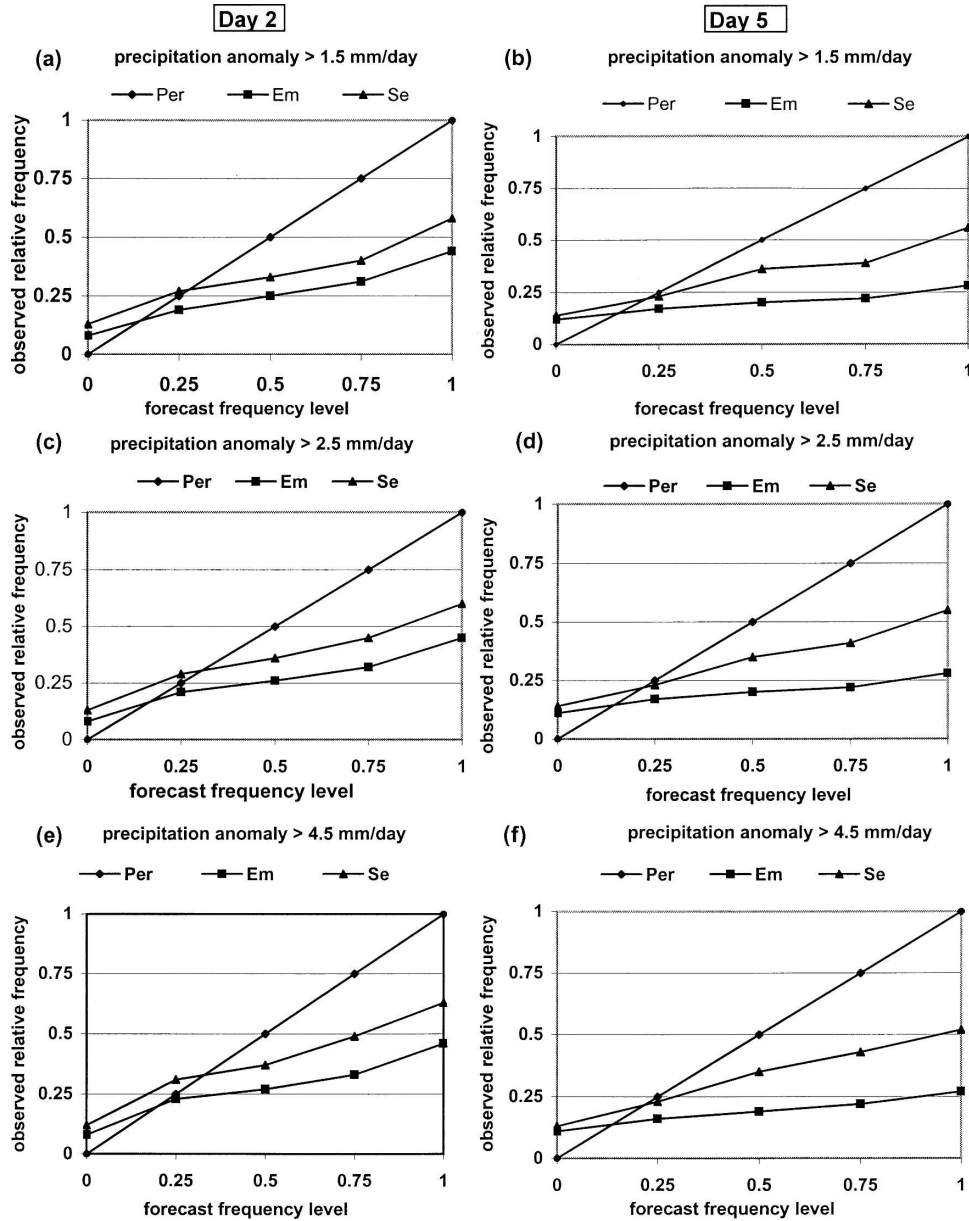


FIG. 5. Reliability diagram for day-2 and -5 forecasts over the tropical belt (30°S–30°N, 0°–60°E) for precipitation anomaly exceeding (a) 1.5 mm day⁻¹ for day 2, $\bar{o} = 0.16$; (b) 1.5 mm day⁻¹ for day 5, $\bar{o} = 0.15$; (c) 2.5 mm day⁻¹ for day 2, $\bar{o} = 0.15$; (d) 2.5 mm day⁻¹ for day 5, $\bar{o} = 0.14$; (e) 4.5 mm day⁻¹ for day 2, $\bar{o} = 0.14$; (f) 4.5 mm day⁻¹ for day 5, $\bar{o} = 0.13$. Here Per, Em, and Se show the perfect, ensemble mean, and superensemble forecasts, respectively, \bar{o} as given in Eq. (B3).

score is zero and the best possible score is one. The bias score (BS) is simply the ratio of the number of yes forecasts to the number of yes observations. Unbiased forecasts exhibit $BS = 1$, indicating that the event was forecast the same number of times that was observed. Bias scores greater than 1 indicate that the event was forecast more often than observed (over forecasting), bias scores less than 1 indicate that the event was fore-

TABLE 4. Brier reliability skill scores for day-2 and -5 forecasts.

Days of forecasts	Day 2			Day 5		
	1.5	2.5	4.5	1.5	2.5	4.5
B_{rel} (EM)	0.58	0.64	0.72	0.49	0.56	0.65
B_{rel} (SE)	0.83	0.85	0.86	0.83	0.84	0.85

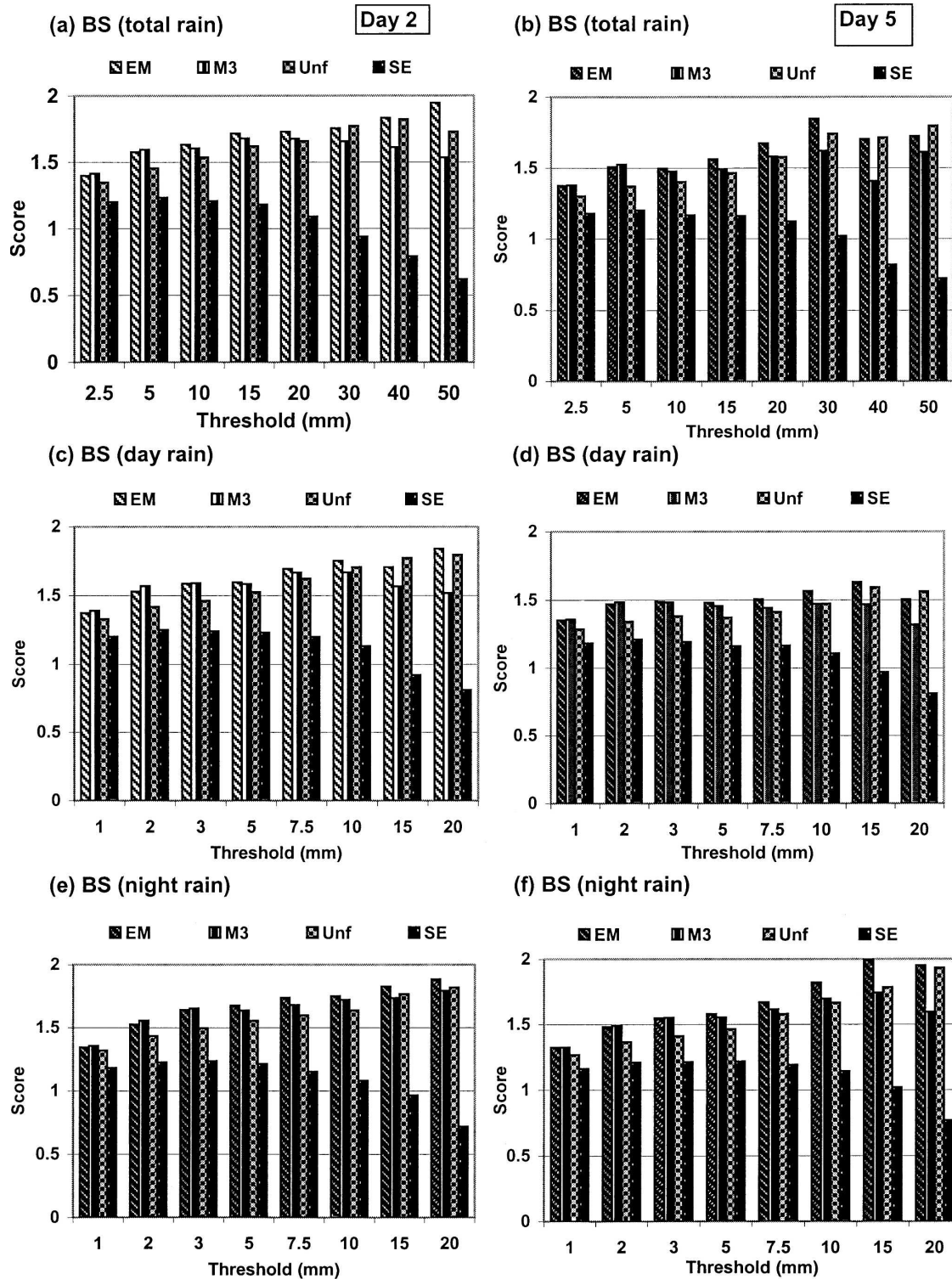


FIG. 6. The BS for day-2 and -5 forecasts over the tropical belt (30°S–30°N, 0°–360°) for EM, M3, UNF, and SE (a) day 2, total rain; (b) day 5, total rain; (c) day 2, day rain; (d) day 5, day rain; (e) day 2, night rain; and (f) day 5, night rain.

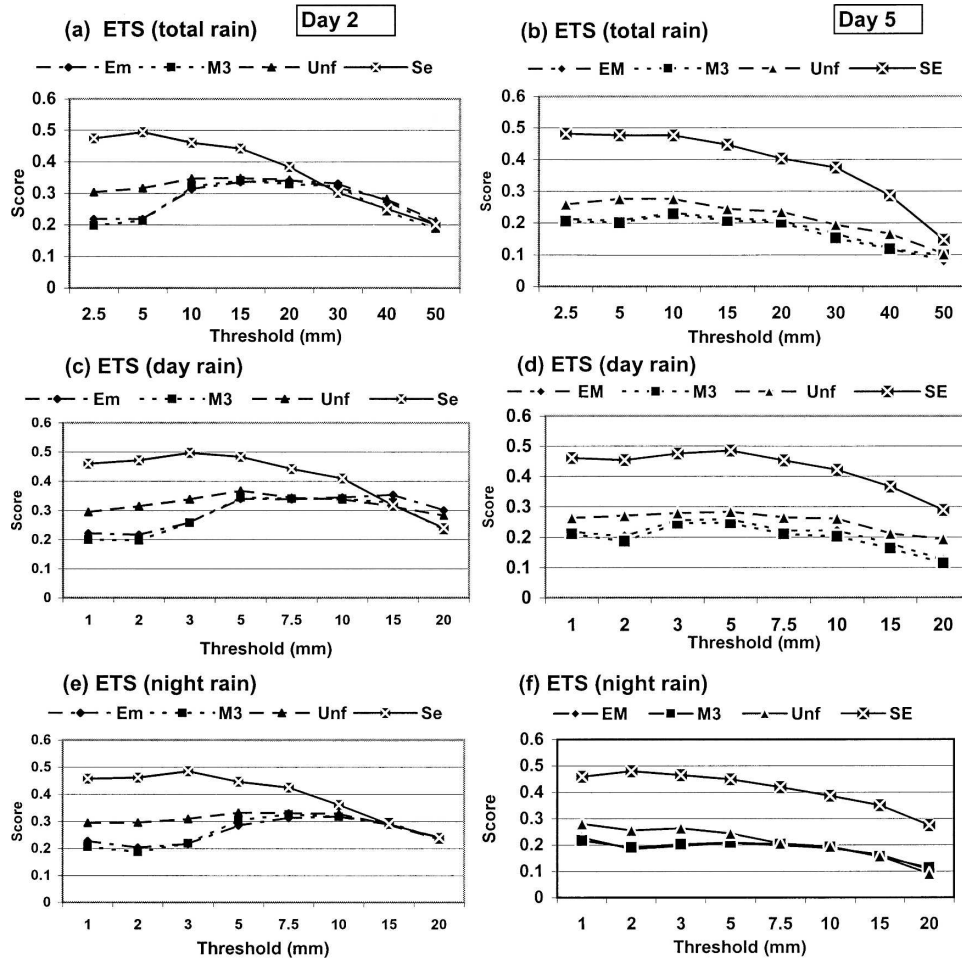


FIG. 7. Same as in Fig. 6, but for the ETS.

cast less often than observed (i.e., underforecasting). In Figs. 6 and 7 we show the bias scores and ETS for total rain, day rain, and night rain for days 2 and 5 of the forecasts. The bar charts show the bias scores (close to 1.0 being a good score) along ordinate and the precipitation thresholds along the abscissa. The superensemble skills for precipitation forecasts for total rain for days 2 and 5 are the highest for all thresholds up to 50 mm day⁻¹. For very heavy rain, the performance of the superensemble is comparable to those of the member models. However for day 5 of the forecasts, the superensemble skill stands out. The separate examination of skills for day versus night also confirms these same findings (i.e., the day-5 skills for the superensemble are indeed very high). The skill of the unified model is better than the member models. The member model bias scores are close to 1.4 for light rain and are as high as 2.0 for heavy rain. This clearly shows that the member models and the ensemble mean overpredict the rainfall events for all the thresholds. Given these over-

all improvements in the skills including those for day and night totals, it was felt that this was a useful database to address the superensemble methodology for the diurnal change. The lower bias in the superensemble emphasizes the increased number of correct forecast events.

Here we follow the work of Janowiak et al. (2005) to illustrate the passage or nonpassage of rainfall along isopleths of local time. Figure 8 shows the time-longitude plots over the tropical belt (10°S–10°N) for the percentage of daily precipitation over each 3-h period covering the forecast period for the day-2 forecast. The diagonal lines are lines of equal local time (isopleths). The aspect of diurnal change that relates to a transient phase follows the sun from east to west during the course of a day. This transient phase undergoes an amplitude modulation during its passage, which follows Fig. 2. This diurnal amplitude modulation of precipitation is related to several complex physical processes such as surface and PBL physics, shallow and deep cu-

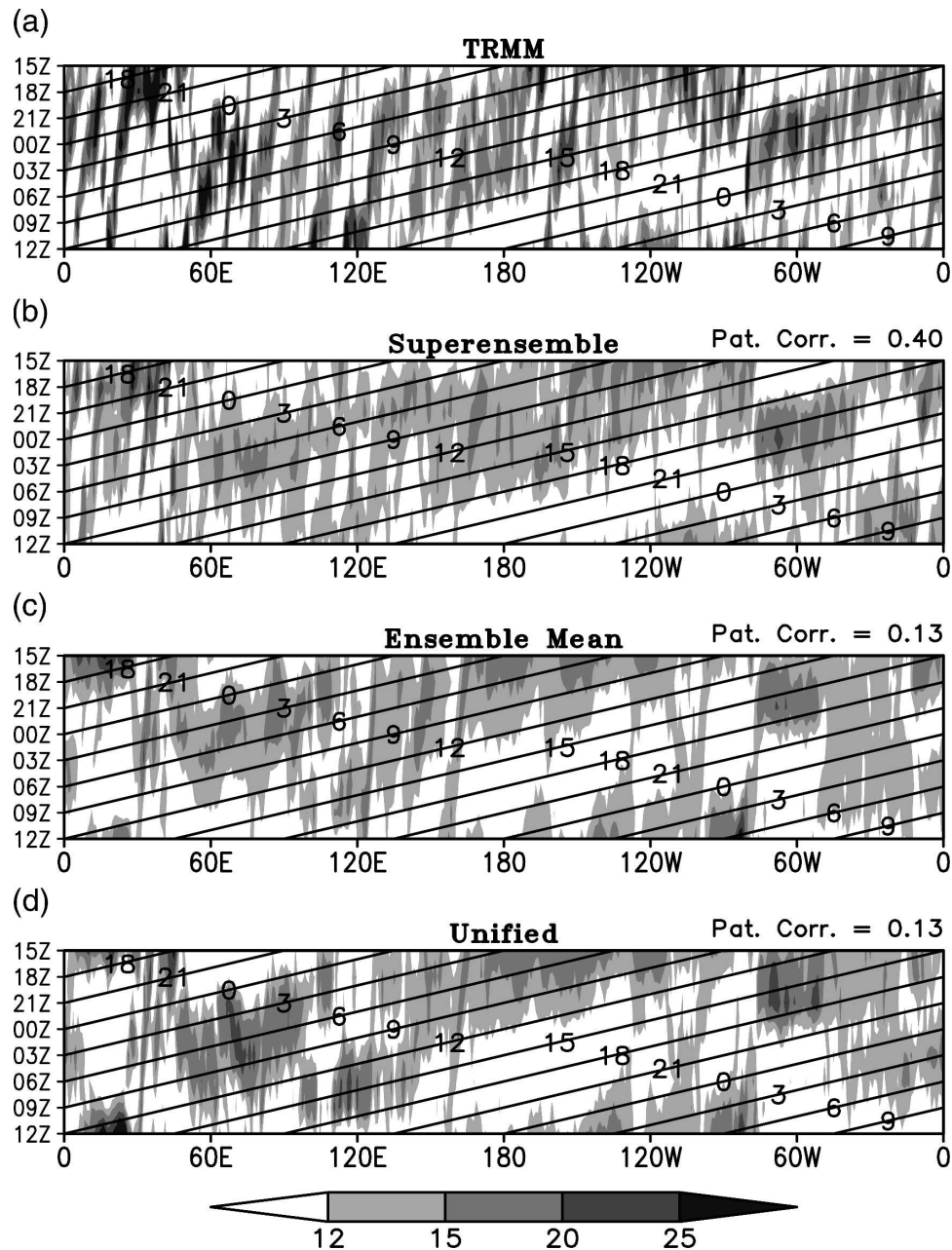


FIG. 8. Percent of daily rainfall for each 3-h period over 10°S–10°N (day-2 forecast) for (a) TRMM, (b) SE, (c) EM, and (d) UNF. The vertical axis indicates time (UTC) and the diagonal lines are lines of equal local time.

mulus convection, and cloud radiative processes. To unravel the cause for the details seen in Fig. 2 it would require a vast numerical experimentation programs, this study is somewhat limited for that level of enquiry. It is important to note here that if rainfall were distributed equally throughout the 24-h period, then 12.5% would be the expected percentage of daily total for each 3-h interval. In the TRMM-observed daily precipi-

tation (Fig. 2a), most of the tropical oceanic region shows early morning and afternoon rainfall maxima, whereas the tropical land areas of South America (70°–40°W) and Africa (10°–40°E) show the late afternoon maxima. Some salient features of the observed precipitation such as the afternoon maxima near 30°E around 1800–2100 UTC and the weaker precipitation (percentage) over the central Pacific Ocean are well captured by

the multimodel superensemble. However, many spurious phase shifted patterns are seen in the ensemble mean and the unified model forecasts. The pattern correlation (which is the spatial correlation designed to detect the similarities in the patterns of the field) of this forecasted variability with the TRMM-observed rainfall variability shows higher skill (pattern correlation = 0.40 on the day-2 forecast). Whereas the respective skills of the member models are found very much lower (figures are not shown here). The pattern correlation was 0.13 for the ensemble mean and 0.13 for the unified model. It is found that over the Atlantic and Pacific Oceans, precipitation follows afternoon hours or early morning hours of local solar passage, respectively.

4. Concluding remarks

In this paper we have discussed a procedure for improving the forecasts of the phase and amplitude of diurnal precipitation for numerical weather prediction. This is a multimodel approach where it is possible to extract additional skills for forecasts beyond those of the member models. By carrying out a training phase, statistical weights were generated that enabled the removal of the collective bias of the member models. Those weights are next used during a forecast phase of the superensemble to obtain these improved forecasts. In spite of the large diurnal errors in phase and amplitude each member model seemed to make the same kind of errors generally. This made it easy for the superensemble to recognize the nature of these errors of each model and to correct them. The nature of similar systematic errors within each single member model runs suggests that the source of these errors may be tractable to the areas of physics and dynamics the model is deficient in.

The multimodels of the current suite were generated from an FSU global spectral forecast model that utilizes different physical parameterization algorithms. That diversity in physics provided us with an ensemble spread of forecasts that was sufficient to derive coefficients for the superensemble. The TRMM data archive of NASA carries a special dataset that provides 3-hourly estimates of rainfall over the global Tropics. Model outputs were also prepared the same 3-hourly intervals. Errors in the modeling of diurnal change arise from several possible physical parameterizations such as the cloud radiation, deep and shallow convection processes, planetary boundary layer physics, and the treatment of the land-ocean surface transfer processes. Large phase and amplitude errors for the diurnal change of precipitation were noted in the member models forecasts. Careful modeling improvements in each of the areas of physical

parameterization are needed for the reduction of these errors. The superensemble is a postprocessing algorithm that has the advantage of providing forecasts for the future 5 days where the errors are considerably reduced by a statistical method. This is being done at all grid locations for all the variables.

In this paper we have shown improved skills for the prediction of total rain over the global tropical and some regional domains. These datasets for the winter season of the year 2000 were next used to extract the diurnal mode (phase and amplitude) for the TRMM and the multimodel forecast database. Some of the regions that carried the largest errors for the diurnal mode of precipitation were the Tibetan Plateau, the eastern foothills of the Himalayas, and Brazil. We found that the large phase transitions between rainfall maxima over short distances (approximately a few hundred kilometers) near the above regions were reasonably handled by the multimodel superensemble. It was possible to carry these same larger skills for the forecasts of the diurnal mode of precipitation through day 5 of the forecasts over these regions.

Using the superensemble methodology it is possible in principle to improve the phase and amplitude of the diurnal change for all variables at all vertical levels. Such a dataset can be very useful for various research activities. The superensemble also carries the best forecast for the total field (not only the diurnal mode) for all variables for short-range forecasts (Krishnamurti et al. 2000a; Mishra and Krishnamurti 2006). There are some well-defined systematic errors in the member models that are being corrected easily (by removing the collective systematic errors of the member models) by the superensemble. What remains to be done is to improve the physics of each model in order to reduce their respective systematic errors.

Acknowledgments. The data used in this study were acquired as part of the Tropical Rainfall Measuring Mission (TRMM). These algorithms were developed by the TRMM Science Team. The data were processed by the TRMM Science Data and Information System (TSDIS) and the TRMM Office; they are archived and distributed by the Goddard Distributed Active Archive Center. TRMM is an international project jointly sponsored by the Japan National Space Development Agency (NASDA) and the U.S. National Aeronautics and Space Administration (NASA) Office of Earth Sciences. We thank Dr. L. Stefanova for her help in the Brier score computation. The study was supported by the following grants: NSF ATM-0419618 and ATM-0311858, NASA NAG5-13563, and NNG05GH81G and NOAA NA16GP1365.

APPENDIX A

A Brief Description of FSUGSM

The model used in this study is the Florida State University Global Spectral Model (FSUGSM) described in Krishnamurti et al. (1998). The horizontal resolution of the model was triangularly truncated at wavenumber T126, which corresponds to a Gaussian grid of 384×192 latitude–longitude, and it has 27 sigma levels. A semi-implicit time integration scheme is used with time step of 450 s to represent the time derivatives. The model uses ECMWF analysis, which includes operational global datasets from the stream of the world weather watch, cloud track winds, commercial aircraft wind reports, surface datasets from marine ships, oceanic buoys surface reports, ocean surface winds from satellite-based scatterometers, and soundings of temperature and humidity from polar-orbiting satellites for the initial analysis. It requires rainfall estimates between hour -24 and 0 . The microwave radiances are provided by TRMM and the Defense Meteorological Satellite Program (DMSP) satellites for deriving rain-rate estimates at roughly 40-km resolution between 50°S and 50°N . An outline of the model is given below:

- 1) Independent variables: x , y , σ , and t ;
- 2) Dependent variables: vorticity, divergence, surface pressure, vertical velocity, temperature, and humidity;
- 3) Vertical resolution: 27 layers between roughly 50 and 1000 hPa;
- 4) Semi-implicit time differencing scheme;
- 5) Envelope orography (Wallace et al. 1983);
- 6) Centered differences in the vertical for all variables except humidity, which is handled by an upstream differencing scheme;
- 7) Fourth-order horizontal diffusion (Kanamitsu et al. 1983);
- 8) Deep convection based on NCEP simplified Arakawa–Schubert cumulus parameterization scheme (Pan and Wu 1995);
- 9) Shallow convection (Tiedke 1984);
- 10) Dry convective adjustment;
- 11) Large-scale condensation (Kanamitsu 1975);
- 12) Surface fluxes via similarity theory (Businger et al. 1971);
- 13) Vertical distribution of fluxes utilizing diffusive formulation where the exchange coefficients are functions of the Richardson number (Louis 1979);
- 14) Longwave and shortwave radiation fluxes based on a band model (Harshvardan and Corsetti 1984; Lacis and Hansen 1974);

- 15) Diurnal cycle with respect to the radiance processes;
- 16) Parameterization of low, middle, and high clouds based on threshold relative humidity for radiative transfer calculations;
- 17) Surface energy balance coupled to the similarity theory (Krishnamurti et al. 1991).

APPENDIX B

A Brief Description of Brier Score Computation

To define the Brier skill score, select an event E that either happens at realization k or does not [$o(k) = 1$ if E occurred, $o(k) = 0$ if it did not] and is forecast to occur with probability $f(k)$. Following Wilks (1995), the Brier score is then defined as

$$b = \frac{1}{n} \sum_{k=1}^n [f(k) - o(k)]^2, \quad (\text{B1})$$

where the index k refers to the forecast–observation pairs and n is the total number of such pairs within the dataset. The lowest possible value of the Brier score is zero, and it can only be achieved with a perfect deterministic forecast. If we let the probabilistic forecast for E be done within i discrete categories y_i , then the frequency with which forecasts of y_i are issued is $p(y_i)$. The frequency within a category y_i forecast with which the event E actually occurs is the conditional frequency $\bar{o}_i = p[o(k) = 1/y_i]$. A reliability diagram is a plot of \bar{o}_i versus y_i . For a perfect forecast, the reliability diagram would be a line at 45° . As suggested by Murphy (1973), it is useful to decompose the Brier score into three terms: reliability, resolution, and uncertainty, as follows:

$$b = \underbrace{\sum_{i=1}^I p(y_i)(y_i - \bar{o}_i)^2}_{\text{reliability}} - \underbrace{\sum_{i=1}^I p(y_i)(\bar{o}_i - \bar{o})^2}_{\text{resolution}} + \underbrace{\bar{o}(1 - \bar{o})}_{\text{uncertainty}}$$

$$= b_{\text{rel}} - b_{\text{res}} + b_{\text{unc}}, \quad (\text{B2})$$

where

$$\bar{o} = \frac{1}{n} \sum_{k=1}^n o(k) \quad (\text{B3})$$

is the unconditional mean frequency of occurrence of the event E , the first term on the RHS is the reliability term, the second one is resolution term, and the third one is uncertainty. The reliability term evaluates the statistical accuracy of the forecast, a perfectly reliable forecast is one for which the observed conditional frequency \bar{o}_i is equal to the forecast probability. The resolution term addresses the distance between the forecast frequency and the unconditional climatological fre-

quency. Forecasts that are always close to the climatological frequency exhibit good reliability because the forecast frequency matches the observed frequency, but has poor resolution because they are not able to distinguish between different regimes. The uncertainty term is a measure of the variability of the system and is not influenced by the forecast. The Brier skill score of reliability can be defined as $B_{\text{rel}} = 1 - (b_{\text{rel}}/b_{\text{unc}})$.

The superensemble defined by Eq. (6) can be rewritten in the following form:

$$S = \left[\sum_{i=1}^N a_i(F_i - \bar{F}_i) \right] + \bar{O} = \frac{1}{N} \sum_{i=1}^N [Na_i(F_i - \bar{F}_i) + \bar{O}]$$

$$= \frac{1}{N} \sum_{i=1}^N F_{S_i}, \quad \text{where } F_{S_i} = Na_i(F_i - \bar{F}_i) + \bar{O} \quad (\text{B4})$$

are the modified forecasts. One possibility for defining the probability of the superensemble then is to treat it as an ensemble of the forecasts F_{S_i} .

APPENDIX C

Significant Test on the Difference in RMSE

Here we assume that the RMSEs of the member models are normally distributed. The compound hypothesis is (Green and Margerison 1978, p. 130)

$$H_0: \text{RMSE}_{\text{em}} - \text{RMSE}_{\text{se}} = 0,$$

$$H_1: \text{RMSE}_{\text{em}} - \text{RMSE}_{\text{se}} \neq 0,$$

where H_0 is the null hypothesis and H_1 is the alternate hypothesis. RMSE_{em} and RMSE_{se} are the population RMSEs of the ensemble mean and superensemble. From the sample estimates of the mean and variances, the t -test parameter can be written as $t_s = (|\text{RMSE}_{\text{em}} - \text{RMSE}_{\text{se}}|/s_{\bar{D}})$, where $s_{\bar{D}} = s_D/\sqrt{n}$ and s_D is the standard deviation of RMSE within the ensemble of n members. The hypothesis H_0 is rejected with $(1 - \alpha)$ 100% confidence if $|t_s| > t_{(1-\alpha/2), (n-1)}$. That is RMSE_{em} and RMSE_{se} differ with a significance level more than $(1 - \alpha)$ 100% when the above condition is satisfied. The significance level was calculated by solving the above equation for α with known values of RMSE_{em} , RMSE_{se} , s_D , and n .

REFERENCES

- Adler, R. F., G. J. Huffman, D. T. Bolvin, S. Curtis, and E. J. Nelkin, 2000: Tropical rainfall distributions determined using TRMM combined with other satellite and rain gauge information. *J. Appl. Meteor.*, **39**, 2007–2023.
- Businger, J. A., J. C. Wungard, Y. Isumi, and E. F. Bradley, 1971: Flux profile relationship in the atmospheric surface layer. *J. Atmos. Sci.*, **28**, 181–189.
- Chakraborty, A., T. N. Krishnamurti, and C. Gnanaseelan, 2007: Prediction of the diurnal change using a multimodel superensemble. Part II: Clouds. *Mon. Wea. Rev.*, in press.
- Chen, S. S., and R. A. Houze Jr., 1997: Diurnal variation and life cycle of deep convective systems over the tropical Pacific warm pool. *Quart. J. Roy. Meteor. Soc.*, **123**, 357–388.
- Gray, W., and R. W. Jacobson, 1977: Diurnal variation of deep cumulus convection. *Mon. Wea. Rev.*, **105**, 1171–1188.
- Green, J. R., and D. Margerison, 1978: *Statistical Treatment of Experimental Data*. 3d ed. Elsevier Scientific, 382 pp.
- Harshvardan, and T. G. Corsetti, 1984: Longwave parameterization for the UCLA/GLAS GCM. Tech. Memo. 86072, 51 pp.
- Harshvardan, R. Davis, D. A. Randall, and T. G. Corsetti, 1987: A fast radiation parameterization for atmospheric circulation models. *J. Geophys. Res.*, **92**, 1009–1016.
- Janowiak, J. E., V. E. Kousky, and R. J. Joyce, 2005: Diurnal cycle of precipitation determined from the CMORPH high spatial and temporal resolution global precipitation analyses. *J. Geophys. Res.*, **110**, D23105, doi:10.1029/2005JD006156.
- Joseph, J. H., 1970: On the solar radiation fluxes in the troposphere. *Sol. Energy*, **13**, 251–261.
- Kanamitsu, M., 1975: On numerical weather prediction over a global tropical belt. Rep. 75-1, 282 pp. [Available from the Department of Meteorology, The Florida State University, Tallahassee, FL 32306.]
- , K. Tada, K. Kudo, N. Sata, and S. Isa, 1983: Description of the JMA operational spectral model. *J. Meteor. Soc. Japan*, **61**, 812–828.
- Katayama, A., 1972: A simplified scheme for computing radiative transfer in the troposphere. Tech. Rep. 6, Department of Meteorology, University of California, Los Angeles, Los Angeles, CA, 77 pp.
- Krishnamurti, T. N., and J. Sanjay, 2003: A new approach to the cumulus parameterization issue. *Tellus*, **55A**, 275–300.
- , K. S. Yap, and D. K. Oosterhof, 1991: Sensitivity of tropical storm forecast to radiative destabilization. *Mon. Wea. Rev.*, **119**, 2176–2205.
- , H. S. Bedi, and V. M. Hardiker, 1998: *An Introduction to Global Spectral Modelling*. Oxford University Press, 272 pp.
- , C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate forecasts from a multimodel superensemble. *Science*, **285**, 1548–1550.
- , —, D. W. Shin, and C. E. Williford, 2000a: Improving tropical precipitation forecasts from a multianalysis superensemble. *J. Climate*, **13**, 4217–4227.
- , —, Z. Zhang, T. LaRow, D. Bachiochi, C. E. Williford, S. Gadgil, and S. Surendran, 2000b: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196–4216.
- , and Coauthors, 2001: Real-time multianalysis–multimodel superensemble forecasts of precipitation using TRMM and SSM/I products. *Mon. Wea. Rev.*, **129**, 2861–2883.
- Lacis, A. A., and J. E. Hansen, 1974: A parameterization for the absorption of solar radiation in the Earth's atmosphere. *J. Atmos. Sci.*, **31**, 118–133.
- Louis, J. F., 1979: A parametric model of the vertical eddy fluxes in the atmosphere. *Bound.-Layer Meteor.*, **17**, 187–202.
- Mishra, A. K., and T. N. Krishnamurti, 2006: Current status of multimodel superensemble and operational NWP forecast of the Indian Summer Monsoon. *J. Earth Syst. Sci.*, in press.
- Murakami, M., 1983: Analysis of the deep convective activity over

- the western Pacific and Southeast Asia. Part I: Diurnal variation. *J. Meteor. Soc. Japan*, **61**, 60–77.
- Murakami, T., 1976: Cloudiness fluctuations during the summer monsoon. *J. Meteor. Soc. Japan*, **54**, 175–181.
- Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteor.*, **12**, 595–600.
- Pan, H.-L., and W.-S. Wu, 1995: Implementing a mass flux convection parameterization package for the NMC medium-range forecast model. NMC Office Note 409, 40 pp.
- Pleim, J. E., and A. Xiu, 1995: Development and testing of a surface flux and planetary boundary layer model for application in mesoscale models. *J. Appl. Meteor.*, **34**, 16–32.
- Puri, K., and M. J. Miller, 1990: Sensitivity of ECMWF analyses-forecasts of tropical cyclones to cumulus parameterization. *Mon. Wea. Rev.*, **118**, 1709–1742.
- Randall, D. A., Harshvardhan, and D. A. Dazlich, 1991: Diurnal variability of the hydrologic cycle in a general circulation model. *J. Atmos. Sci.*, **48**, 40–62.
- Reed, R. J., D. C. Norquist, and E. E. Recker, 1977: The structure and properties of African wave disturbances as observed during Phase III of GATE. *Mon. Wea. Rev.*, **105**, 317–333.
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.
- Slingo, J. M., P. Inness, R. Neale, S. Woolnough, and G.-Y. Yang, 2003: Scale interactions on diurnal to seasonal timescales and their relevance to model systematic errors. *Ann. Geophys.*, **46**, 139–155.
- Stefanova, L., and T. N. Krishnamurti, 2002: Interpretation of seasonal climate forecast using Brier Skill score, the Florida State University superensemble, and the AMIP-I datasets. *J. Climate*, **15**, 537–544.
- Thompson, R. M., Jr., S. W. Payne, E. E. Recker, and R. J. Reed, 1979: Structure and properties of synoptic-scale wave disturbances in the intertropical convergence zone of the eastern Atlantic. *J. Atmos. Sci.*, **36**, 53–72.
- Tiedke, M., 1984: The sensitivity of the time-mean large-scale flow to cumulus convection in the ECMWF model. *Proc. Workshop on Convection in Large-Scale Numerical Models*, Reading, United Kingdom, ECMWF, 297–316.
- Wallace, J. M., 1975: Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. *Mon. Wea. Rev.*, **103**, 406–419.
- , S. Tibaldi, and A. J. Simmons, 1983: Reduction of systematic forecast errors in the ECMWF model through the introduction of an envelope orography. *Quart. J. Roy. Meteor. Soc.*, **109**, 683–717.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. Academic Press, 467 pp.
- Yang, S., and E. A. Smith, 2006: Mechanisms for diurnal variability of global tropical rainfall observed from TRMM. *J. Climate*, **19**, 5190–5226.
- Yun, W. T., L. Stefanova, and T. N. Krishnamurti, 2003: Improvement of multimodel superensemble technique for seasonal forecasts. *J. Climate*, **16**, 3834–3840.