Improving Multimodel Weather Forecast of Monsoon Rain

Over China using FSU Superensemble

T.N. Krishnamurti*, Anitha D Sagadevan, Arindam Chakraborty

A. K. Mishra and Anu Simon

Department of Meteorology, Florida State University

Tallahassee, FL 32306

Submitted to Advances in Atmospheric Sciences

Revised and submitted on March 15, 2009

*****Corresponding author address: T.N. Krishnamurti, tnk@io.met.fsu.edu

Abstract

In this paper we present the current capabilities for Numerical Weather prediction of precipitation over China using a suite of ten multimodels and our Superensemble based forecasts. Our suite of models includes the operational suite selected by NCARs TIGGE archives for the THORPEX Program. These are: ECMWF, UKMO, JMA, NCEP, CMA, CMC, BOM, MF, KMA and the CPTEC models.(The acronyms are explained in a table 1 within this paper).The superensemble strategy includes a training and a forecasts phase, for these the periods chosen for this study include the months February through September for the years 2007 and 2008. This paper addresses precipitation forecasts for the medium range i.e. days 1 to 3 and extending out to day 10 of forecasts using this suite of global models. For training and forecasts validations we have made use of an advanced TRMM satellite based rainfall product. We make use of standard metrics for forecast validations that include the RMS errors, spatial correlations and the equitable threat scores. The results of skill forecasts of precipitation clearly demonstrate that it is possible to obtain higher skills for precipitation forecasts for days 1 through 3 of forecasts from the use of the multimodel superensemble as compared to the best model of this suite. Between days 4 to 10 it is possible to have very high skills from the multimodel superensemble for the RMS error of precipitation. Those skills are shown for a global belt and especially over China. Phenomenologically this product was also found very useful for precipitation forecasts for the Onset of the South China Sea monsoon, the life cycle of the Meiyu rains and post typhoon landfall heavy rains and flood events. The higher skills of the multimodel superensemble make it a very useful product for such real time events

Key Words: THORPEX, Ensemble Mean, Multimodels, Superensemble, Monsoon, Meiyu Rain.

1. Introduction

A major improvement of skill for medium range forecast is possible from the construction of a multimodel superensemble. This is a post processing exercise where a large number of forecasts from a suite of models are subjected to training where the bias errors of member models are expressed in terms of statistical weights. These weights are next used in forecasts where the collective biases are reduced to construct a multi model superensemble, Krishnamurti et al (1999, 2000a, 2000b). In a series of papers we have shown that it is possible to take a suite of operational model forecasts and to arrive at forecasts that are somewhat superior to those of the best model (Mishra and Krishnamurti 2007). This paper is a first such application to a China domain where we shall address the skills for day to day forecasts of precipitation and specifically to the forecasts of the Meiyu rains, the onset of the South China Sea monsoon and the prediction of heavy rains arising from the post land fall periods of Typhoons over China. This study became possible because of a new forecasts data archival system called TIGGE at NCAR. (A list of acronyms and their explanations are provided in Table 1). TIGGE is a part of the International THORPEX program. The archives include medium range daily forecasts, through day 10 of forecasts, made by a number of operational global weather modeling groups. This archive includes several years of daily forecasts for all basic variables at all model levels. The TIGGE system is briefly described in Section 2 of this paper.

Forecasting day to day weather is difficult over a China domain compared to many parts of the world. This region is located to the east of major Himalayan mountain chains and the semiarid Gobi desert. The complex zonal motion of the subtropical high of the Pacific Ocean, its incursions into and out of China, impacts the rainfall over this region. Monsoonal moisture enters the Chinese mainland from Bay of Bengal, the South China Sea and even from the Pacific Ocean. Frontal systems of the North West often plunge southwards and contribute to heavy rains. Thus a complex set of surface, orographic, lateral and meridional influences impact the medium range forecasts of rains over China. Among these the influences from the Pacific and those of orography (the eastern slopes of the Himalayas) seem to be some of the major contributors to the forecast errors. This can be seen from the progression of errors (Figure 1) on a day by day basis (from days 1 through 10 of forecasts). Most Errors seem to emanate and expand from the Pacific and the Himalayan regions.

A number of skill score indices are applied for forecast evaluations. The Equitable Threat Scores (ETS) are often used for precipitation forecast skills (see Appendix 1). The typical operational values, for moderate rainfall intensities, of the ETS start out with values like 0.3 initially and fall to values around 0.1 by day 3 of forecasts. Those skills for heavier rains are generally much lower in operations. There is not much skill in numerical weather prediction of rains for medium range forecast i.e. 6 to 10 days in advance. If a suite of the best operational models are used to construct the multimodel superensemble then these score can be raised for forecasts from days 1 to 3. That is demonstrated in this paper for a China domain. Various important features of the superensemble were reviewed by Krishnamurti et al (2006). Having good models in the ensemble suite is certainly very important. For numerical weather prediction improvements the training phase requires as many as 120 past forecasts from each of the member models, generally, for obtaining a stable and useful statistics for the forecast phase of the superensemble (Krishnamurti et. al. 2006). The number ofensemble members suggested for single model based perturbed forecasts are generally of the order of 50, (Palmer et al., 1993 and Toth and Kalnay 1993). The superensemble starts to provide much useful results even from the use of 5 to 10 best operational models. It is possible to design a number of member models by simply changing the physical parameterizations. The following four recent papers, Krishnamurti and Sanjay (2003), Chakraborty et al (2007) , Krishnamurti et al, (2007, 2008) address the use of multiple physical parameterization algorithms such as cumulus parameterization, planetary boundary layer and cloud radiative transfer in suites of multimodels and ways to find systematic errors generated by various such schemes. We show results for the Chinese summer monsoon domain here. Our statistics includes all of China where results of other member models, that are being used here, are included. With the superensemble it is possible to increase the predictability over china by about two days for 5-day forecasts and by more than two days for 10-day forecasts (the RMS errors) compared to the best model. We illustrate here examples of much improved prediction of monsoon rainfall for China from case studies and covering an entire season,

2. TIGGE Datasets

This current paper exploits the TIGGE data sets of THORPEX program that are provided by NCAR see Table 2 below. Table 2 lists the prominent operational weather prediction models from the TIGGE archive for the current period. The table identifies number of ensemble members, model resolution and forecast length. TIGGE datasets are the basis for the member model suite of the Florida State University superensemble. The multimodels included in our study are BOM (Australia), CMA (China), ECMWF (Europe), NCEP (USA) and UKMO (UK). TIGGE was designed at a workshop hosted in 2005 by the European Centre for Medium-Range Weather Forecasts (Richardson et al 2005), with the ten centers officially joining between late 2006 and early 2008. It is an attempt made to link the academic and operational worlds. Such outputs are now available from data archive centers at ECMWF, the China Meteorological Agency, and NCAR. Each day some 240 GB (gigabytes) of data flow from the operational centers into the TIGGE system. The archive now holds more than 100 terabytes.

3. Conventional Superensemble Methodology

The superensemble technique (Krishnamurti et al., 1999) produces a single forecast derived from a multimodel set of forecasts. Forecasts from this methodology do carry the highest skill compared to participating member models of the ensemble, and they carry skills above those of the bias-removed ensemble mean representation. The strategy for the multimodel superensemble partitions the forecast time line into two components. The first of these, called the training phase, utilizes the multimodel forecasts and the observed (analysis) fields to derive model performance statistics. The second phase, called 'the forecast phase', utilizes the multimodel forecasts and the aforementioned statistics to obtain superensemble forecasts into future. During training, with the use of benchmark observed (analysis) fields, past forecasts are used to derive statistics on the past behavior of the models. Given a set of past multimodel forecasts, we have used a multiple regression technique (for the multimodels), in which the model forecasts were regressed against an observed (analysis) field. This utilizes a least squares minimization of the difference between anomalies of the model and the analysis fields in order to determine a distribution of weights. These regression coefficients associated with each individual model conceivably can be interpreted as a measure of that model's relative reliability for the given point over the training period. For each model prognostic variable, the purpose of training is to evaluate model biases geographically and vertically. This being done for ***m*** multimodels at ***n*** grid points (along the horizontal and vertical) for ***p*** variables and ***q*** time intervals constituted as many as $\boldsymbol{m*n*p*q}$ statistical coefficients (which came to around 10^7 weights). These weights are positive, negative and fractional, Krishnamurti et al (2006) which rules out any possibility of over-fitting. In our previous study (Krishnamurti et al 2003) we had cited several examples of heavy rain predictions, when none of the member models were raining heavily. In hurricane track forecasts (Krishnamurti et al 2006) using multimodel superensemble there are numerous examples where all member models show a left bias in track prediction whereas superensemble by correcting this bias shows a track far right of any of the predicted tracks. This degree of detail for the construction of the superensemble was found necessary. The methodology for this conventional procedure consists of a definition of the superensemble forecast:

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

where S is superensemble, $\overline{O}$ is the observed mean field during the training phase; a_i is the weight for the i^{th} member model; F_i and $\overline{F}_i$ are the forecasts and mean forecast fields during the training phase from the i^{th} model. The summation is taken over the N member models of the suite. The weights are computed at each of the grid points by minimizing the objective function ***G*** for the mean square error of the forecasts:

$$G = \sum_{t=0}^{t=train} (S_t - O_t)^2 \quad \ldots \quad \ldots \quad \ldots \quad (2)$$

where ‘t’ denotes the length of a training period.

In this conventional superensemble methodology a collection of a sequence of individual forecasts from several models are subjected to a multiple regression against the observed (or assimilated) counterpart fields. These multi-regression coefficients are collected during the training phase of the superensemble. These statistical weights are separately calculated for each day of forecasts. The length of this training data phase varies for each type of forecast addressed in this paper. These statistics, collected during the training phase are simply passed on to a forecast phase of the superensemble. In this forecast phase, we again have forecasts, from the same member models, that are corrected for their past collective behavior. This type of local bias removal is more effective compared to a conventional bias removed ensemble mean. The later assigns a weight of 1.0 to all models after bias removal. The superensemble includes fractional and even negative weights depending on past behaviors. In a probabilistic sense also, the superensemble probability forecasts are somewhat better than the multimodel bias-removed ensemble at any threshold level (Stefanaova and Krishnamurti, 2002). A consensus of 10-day global forecasts from all models (TIGGE) is prepared following the superensemble strategy. The superensemble includes training and a forecast phase. During the training phase nearly 120 recent past forecasts are used. A simple least square minimization strategy utilizes these multiple forecasts and observed counterparts to obtain statistical weights for each model for each forecast day. Weights for the Superensemble are positive, negative and fractional, which rules out any possibility of over-fitting. In our previous study (Krishnamurti et al. 2006) we had cited several examples of heavy rain predictions, when none of the member models were raining heavily. In hurricane track forecasts using multimodel superensemble there are numerous examples where all member models show a left/right bias in track prediction whereas superensemble by correcting this bias shows a track far right/left of any of the predicted tracks. This degree of detail for the construction of the superensemble was found necessary. Statistics thus obtained are passed on to the forecast phase for the construction of superensemble forecasts

4. Results and Discussions:

TIGGE forecasts and TRMM rain datasets were used to carry out FSU multimodel forecast from 21st April to 21st July 2008 (total 92 days). Model forecasts up to 10 days from the TIGGE archive (1 February to 20 April 2008) are used for training and forecasts are made from 21st April 2008 to 21st July 2008. In Figure 2 we display the time series of RMS errors of precipitation forecasts for days 1, 3, 5, 7, 9 and 10, covering one season of forecast, over China for the domain 90°E-140°E, 15°N-49°N. Fig. 2 shows the RMS errors covering the period February through April 2008. As seen here the RMS errors of the FSU superensemble is around 5 mm d^{-1} for all 10 days of forecasts. RMS errors of the member model forecast increases with the length of forecasts (i.e. forecast day) however irrespective of the forecast day superensemble forecasts RMS errors range from roughly 1.0 to 10.0 mm d^{-1}. RMS errors of the member models forecast reaches as high as 70 mm d^{-1} on day 10, but corresponding superensemble forecast is found less than 10 mm d^{-1}.

This indicates a major improvement for the forecast especially for the forecasts for days 5 to 10. Typical spatial plots of forecasts of precipitation for days 1 and 10 are shown in Fig. 3 (a) and (b) respectively. Here we show that observed rain, as seen from the TRMM files, the member model forecasts, the ensemble mean and the FSU superensemble. Also included on top of each forecast panel are the values for the average rain rate and the spatial correlations.

It became possible to raise these skills well above those of the best model from the construction of the superensemble. If the member models carry consistent and large systematic errors then the superensemble is able to capitalize on these errors and reduce them. Fig. 3 (a) and (b) shows the average rain (in sense rainfall climatology of forecasts) and spatial correlation for the entire period (February–April 2008). Here, again, we see a similar increase of skill from the FSU superensemble. In both these figures (Fig. 3 (a) and (b)), top left panel depicts average observed rain rate from TRMM, rest of the panels are forecast from FSU multimodels where bottom left is ensemble mean and bottom right is superensemble forecast. Numbers mentioned in each forecast panels represents the average rain (AVE=) and spatial correlation (SC=) of the forecast with the observed rain during the period. Figure 3 (a) is prepared using day 1 forecasts of the FSU multimodel and Figure 3 (b) is made of day 10 forecasts. In both the figures 3 (a) and 3 (b) superensemble average rain for the domain is very close to that of TRMM rain. These spatial correlations from the FSU superensemble, for precipitation forecast, are greater than 0.9 for all 10 days of forecast over China. This kind of skill is now possible from the use of the superensemble method for the climatology of daily forecasts. Note that some of the member model forecasts of correlation are as low as 0.3. These show that the multimodel superensemble is able to reduce errors.

4.1. Mei-yu Rains

The Meiyu system, also called Baiu in Japan and Changma in Korea, is a front like precipitation system in East Asia that is found in the summer months. Its onset in late May and June is of considerable interest. It carries a stronger gradient of moisture compared to that of temperature and is not considered a baroclinic frontal system. Heavy rains over eastern China, east of the Yangtze River are known to be associated with the Meiyu rains. Maiyu rains are also of considerable interest in Taiwan, because the south western portion of the Meiyu front often lies over Taiwan and produces heavy rains there. During the months of May through June the Meiyu front often lies over the open Ocean offshore from Southern China and it makes an inland movement during July with heavy rains. In this paper we illustrate days 1 through 3 forecasts of the precipitation forecasts from the FSU multimodel superensemble to show that it is possible to provide some useful guidance for these rainfall forecasts from the data sets of the current suite of operational models. The issue of resolution is still there, the TIGGE suite of models carry a resolution of roughly 80 Km. on the average and only show a somewhat smoother representation of the Meiyu front and its associated rains. However because of the strength of the multi model superensemble it is possible to improve the geographical location and even the amplitude of the predicted rains (as trained and validated with the blended TRMM product resolution). The systematic errors in the geographical locations of the rains are much improved by the superensemble. The amplitude of the predicted rains is corrected towards the TRMM based estimates by this procedure. Here we have selected samples of Day1, Day 5 and Day 7 (fig. 4 a, b, c), to illustrate spatial plots of these forecasts.

We shall illustrate an example of the prediction of Meiyu front rains for 14th June 2008 here. The observed rains from TRMM and the multimodel forecasts of rains are illustrated in Figure 4 (a), (b) and (c). These illustrations carry the member model forecasts and those from the ensemble mean and the multimodel superensemble for days 1, 5 and 7 of forecasts. Our interest here is in the forecasts of the location of the Meiyu front and the amplitude of Meiyu rains. We have included the results for the member models of our FSU multimodels suite along with Ensemble mean (EM) and Superensemble (SE) in this illustration, however all the remaining models of TIGGE could not be used for this experiment because of data availability issues. The features of interest are the elongated line of heavy rains south of Korea and stretching south westwards over Taiwan and South China Sea. The heaviest rains are located immediately south of Korea and the Taiwan region. Most models perform reasonably well in their day 1 forecasts. We will compare the performance of the CMA model and the superensemble here. On day one the RMS errors and spatial correlations for the CMA precipitation forecasts, for the domain shown in Fig. 4 were 14.63 and 0.62, the superensemble was able to improve these numbers to 11.65 for the RMS errors and 0.75 for the spatial correlation. The Superensemble carried consistent higher values for these skills through day 10 of forecasts. The corresponding skills at day 5 for the CMA of 20.41 and 0.27 compare with the better values of 15.63 and 0.49 for the superensemble (Fig. 4 (b)). Similarly for day 7 these skills for CMA were found 24.79 and 0.19 which were improved for superensemble as 16.34 and 0.40, these skills are much higher. The high rainfall rates along the Meiyu front was reasonably predicted during days 1 and 2 by the superensemble, there were some major errors is providing a zonal belt of rains along 30 north extending eastwards in most models by day 5 and day 7 of forecasts that was also reflected in the ensemble mean and the superensemble. All member models along with ensemble mean predicted widespread rains in the south west corner of the domain, which were not reflected in the superensemble forecast and in the TRMM observation.

We have calculated the Equitable Threat Scores (ETS) for the precipitation forecasts for all the models for each day of forecast over a China Domain (Day 1 through 3 are shown in Figure 5). The equitable threat scores and the bias of the precipitation forecasts are shown in Fig. (5 a, b, c, d, e, f), Appendix 1 describes the threat score and the bias scores. The abscissas in these diagrams are the rainfall rate thresholds, i.e. number fifteen (15) denotes all rainfall in excess of 15 mm d^{-1}. The ordinate denotes the equitable threshold scores in panels (a), (c) and (e) and the bias scores in panels (b), (d) and (f). Most operational models start with an equitable threshold score of around 0.30 for day 1 and these scores drop to near 0.15 by day 3 of forecasts. Here we are showing the results for one model, the CMA, the ensemble mean and for the multimodel superensemble. We noted that the best results for rainfall forecasts were obtained from the multimodel superensemble that still carries much higher skills for day 3 of forecasts. Another aspect, we see here, are the skills for heavy rains, i.e. thresholds in excess of 15 mm d^{-1}, those are predicted with higher skills by the multimodel superensemble. This is due to the nature of the consistent systematic errors of the member models; those are easily exploited by the multimodel superensemble in its forecasts. A bias score of 1.0 is considered a perfect score. It is seen from these figures that member models carry large bias errors, those are very much improved by the multimodel superensemble.

4.2. Floods from Post Landfall of a Typhoon Kalmaegi

The prediction of heavy rains and resulting floods from post typhoon landfall is an important research topic. Ideally a suite of high resolution models and the construction of a multimodel superensemble forecast of heavy rains would be expected to provide the most useful product for studying hydrological aspects of floods. It is of considerable interest to ask how far we can go with the TIGGE suite of models for predicting post typhoon landfall heavy rains. During the year 2008, in our data files there were several such typhoons. Of those we have selected one of the important typhoons that caused major flooding from the heavy rains over China. This was Typhoon Kalmeigi, which formed as a tropical depression on 13 July 2008 when it was located to the east of the Philippines. RSMC Tokyo named it Kalmaegi on 15 July 2008; the storm reached its peak winds of 75 knots (139 km/h) on 17 July. Shortly afterwards it made a direct landfall on Taiwan and then moved into China’s Fujan province the next day it emerged into the Taiwan Strait and raced towards North Korea where it became fully extra tropical, resulting in heavy rains and floods over the Fujian province and later over Taiwan, where its effects were much more severe. According to some newspaper reports the storm caused NT$ 300 million worth of damage, and destroyed about 5,100 hectares of orchards and crops. In Xiapu County of Fujian Province and in neighboring Zhejiang Province, 360,000 residents left coastal and low-lying homes to escape the storm. The large scale models of the TIGGE suite and especially the multimodel superensemble provided some very useful forecasts of precipitation for this storm. In Figure 6 (a, b, c) we show precipitation forecasts for days 1, 2 and 3 all valid for July 19 2008, the day of heavy floods over the Xiapu County.

The superensemble forecasts of precipitation carry consistently the highest skills (RMS errors and spatial correlations) compared to all the models and the ensemble mean for all three days of forecasts shown in Fig. 6 (a, b, c). Over this China domain the RMSE and the spatial correlations were 7.79 and 0.44 for the multimodel superensemble for day 3 of forecasts. The corresponding pairs of numbers for the CMA model were 9.82 and 0.34. Most models performed similar to the CMA model. We noted that if regular forecasts were prepared carrying three day forecast totals of rains then the superensemble can consistently provide a very useful product for guidance of heavy rains and floods compared to all the member models. This is a feasible proposition in real time. Overall much useful guidance for heavy rains over the Xiapu County of the Fujian province was possible from the superensemble forecasts.

On June 26 2008 very heavy rains and resulting floods were experienced over the coastal regions of the Guangdong province of China. Those heavy rains are clearly reflected in the TRMM blended rainfall estimates shown in the top left panels of Fig. 7 (a, b, c). These were related to a strong spell of the East Asia monsoon. Moist south westerly flows to the south of this coastal rainfall belt were enclosed by a north westerly drier air to the north of this belt. The heavy rains of the order of 40 to 45 mm d^{-1} (as seen by the TRMM product) were predicted quite well by most models in their day 1 of forecasts, The construction of the superensemble did provide the best product in terms of RMS error and spatial correlations for each of the three days of forecast for this event. By day three of forecasts several of the models, such as the UKMO, carried large errors. The quality and spread of forecast results we see in these Figures 6 and 7 are typical of the large scale model performances that we find in the current day models.

The Equitable threat scores and their bias scores for days 1, 2 and 3 of forecasts for this typhoon landfall event are shown in Fig. 8 (a, b, c). In this example both the ensemble mean and the superensemble produced the best results. Those were much superior to the results on ETS provided by the member models. The ensemble means superior performance was related to an even spread of forecast errors from the member models. The bias scores showed an interesting feature for all 3 days of forecasts, the superensemble carried the best scores closest to 1.0, here we found for moderate rains the bias scores were greater than 1.0 and for heavy rains the superensemble carried a bias lower than 1.0. This feature was seen in all of our forecasts Heavy rains are clearly underestimated somewhat by this consensus forecast. We have also looked at the equitable threat scores for a second typhoon and noted very similar results on rainfall skills during the post landfall 3 day forecast periods (Figure not shown). It is clear that having a superensemble product adds to the value of forecasts because of its consistency of performance towards providing the highest skills.

4.3. Onset of South China Sea Monsoon

The onset of the South China Sea monsoon is a topic of great interest for East Asia. The Myanmar onset of monsoon usually occurs during early May. A number of factors are normally associated with this onset; waves of meridionally propagating intraseasonal waves pass through East Asia and south Asia during the summer months at intervals of roughly 30 to 50 days. One of these passages near 10° N over the Bay of Bengal, generally during early May, is associated with the onset of heavy monsoon rains in Myanmar. Often this onset is triggered by the formation of an onset vortex, which is generally a pre-onset tropical storm over the northern Bay of Bengal. The south westerly broad monsoon current of the Bay of Bengal is strengthened by the southern flanks of the onset vortex, which makes these strong south westerlies to interact with the coastal mountains of Myanmar with very heavy orographic rains. These strengthened westerlies cross towards the South China Sea within a few days after the onset over Myanmar and form the southern flank of the Meiyu front. The onset of the Meiyu rains follows the Myanmar onset very quickly. The onset of the heavy rains along the Meiyu front is of considerable interest in these regions. Ten day forecast guidance, with some improved measurable forecast skills is possible with the current suite of operational models used in the context of the multimodel superensemble. We shall illustrate this aspect of the forecasts in following Figures 9 and 10.

During 2008 the onset of the South China Sea monsoon rains occurred around May 3. In figures 9 (a), (b), (c) and (d). We show our forecasts valid on May 3rd 2008 for days 1, 3, 5 and 10. During the onset the rainfall rates were as large as 30 to 40 mm d^{-1} and the rain belt had entered the inland coastal region as seen in the TRMM based observed estimates in Fig. 9. The member model's forecasts and those from the ensemble mean and the superensemble are also presented in each panel. Also included in the top of each panel are the RMSE and the spatial correlations for each forecast.

The Equitable Threat scores and their bias scores for the onset of South China Sea precipitation forecasts for days 1, 2 and 3 are shown in Figure 10. Noting that the onset of rains comes after a long dry season (i.e. prior to the onset) thus it did not perform as well. Training phase of the superensemble does not provide the most reliable statistical weights. One season of training, based on the same years past 100 days, lowers the skill of the superensemble a little. For further improvements we require many years of pre onset and post onset sets for the construction of superensemble. The ensemble mean ended up performing as well as the superensemble in predicting the monsoon onset rains of the South China Sea coastal areas of China. For day 10 of forecasts the ensemble mean worked better than all models. Over all the forecast skills through day 5 from the superensemble were nearly 50 to 100 percent higher than all of the models (not shown). This suggests that both the location and amplitude of heavy rains during the onset can be predicted very well for nearly a week in advance from the use of the multimodel superensemble. Both the ensemble mean and the superensemble performed much better than the member models. We discussed the skills of one model, the CMA, here in detail, and we see that it is possible to obtain a forecast with superior skill from our post processing. The bias scores also reflect the improved bias from the superior ensemble mean and the superensemble.

5. Conclusions and Future Work

Precipitation forecasts from a suite of mesoscale high resolution models would clearly be the next goal of what has been completed in this study. Here we have deployed a suite of large scale global models of the TIGGE/THORPEX suite to examine through day 10 of forecasts for precipitation over a global tropical belt and especially over China. An ensemble strategy, called the FSU multimodel superensemble, has been used to obtain the best forecasts for precipitation that exceed in skill compared to all member models of this suite. The TIGGE suite provides a unique opportunity for advancing the state of ensemble forecasting for numerical weather prediction. This suite carries forecasts from as many as 10 member models, where many member models are providing a large number of forecasts for each time interval. The superensemble strategy follows our previous work; it carries a large training phase and a forecast phase. During the training phase we generate a statistics on the recent past performances of the member models and their collective bias errors are minimized by this procedure. We prepare different statistical weights for each day of forecast recognizing that some models are more skillful early on in their forecasts, whereas some models carry more skills later in their forecasts. We have illustrated the skills of forecasts over China. For the training and the validation of forecasts better observed estimates of precipitation are needed. For this purpose we have used the TRMM 3B42 database. That is a blended product which utilizes the TRMM microwave radiances as well as IR data sets from geostationary satellites (of the globe) for extracting rainfall estimates. Ideally we should be using a mix of rain gauge over land and TRMM products (Krishnamurti et al 2008); this will be followed up in our future work for rainfall forecasts over China. Our skill Metrics for precipitation forecasts includes the equitable threat scores and the bias scores for each member model and the Superensemble. Using these skill measures we were able to ask about the threat scores for different intensity of rain rates in the forecasts. We noted that it is possible to improve these skills, over China, for all ten days of forecasts from the multimodel forecasts compared to the current best model by about 20 percent. It is possible to diagnose possible areas of the model physics and dynamics that contribute to their systematic errors, Krishnamurti et al (1996). Further work in this area is possible from a suite of mesoscale models. A superensemble based on mesoscale models would be more suitable for addressing forecast issues of local heavy rains and floods. The overestimates of lighter rains and the underestimates of heavy rain bias scores was a common feature in these forecasts over China. This feature can be used to improve future versions of the superensemble which is presently designed towards a minimization of the RMSE only.

Acknowledgements

This study was supported by grants NASA PMM-NNX07AD39G, ATM-0311858 and NSF ATM-0419618. We sincerely thank TIGGE and THORPEX program for providing this very useful data. Thanks are also due to Zaihua Ji from The CISL Research Data Archive (RDA) helping us to download the TIGGE data. Thanks are due to our reviewers for the useful suggestions.

REFERENCES

- Anthes, R. A., 1983: Regional models of the atmosphere in middle latitudes. *Mon. Wea. Rev.*, **111**, 1306-1330.
- Chakraborty, A., T. N. Krishnamurti, and C. Gnanaseelan, 2007: Prediction of the diurnal cycle using a multimodel superensemble. Part II: Clouds. *Mon. Wea. Rev*., 135, 4097– 4116.
- Krishnamurti T.N., T.S.V. Kumar, Won-Tae Yun, A. Chakraborty and L. Stefanova, 2006: Weather and seasonal climate forecasts using the superensemble

approach, *Predictability of Weather and Climate*, Ed. Tim Palmer and Renate Hagedorn.

- Krishnamurti, T. N., and J. Sanjay, 2003: A new approach to the cumulus Parameterization issue. *Tellus*, 55, 275–300.
- Krishnamurti, T. N., C. Gnanaseelan, A. K. Mishra, and A. Chakraborty, 2008: Improved forecasts of diurnal cycle in tropics using multiple global models. Part I: Precipitation. *J. Climate*, 21, 4029–4043.
- Krishnamurti, T. N., C. M. Kishtawal, Shin, D. W., C. E. Williford, 2000b: Improving tropical precipitation forecasts from a multi analysis superensemble. *J. Climate*, 13, 4217–4227.
- Krishnamurti, T. N., H. Bedi, G. Rohaly, and D. Oosterhof, 1996: Partitioning of the seasonal simulation of a monsoon climate. *Mon. Wea. Rev.*, 124, 1499–1520.
- Krishnamurti, T. N., S. Basu, J. Sanjay, and C. Gnanaseelan, 2007: Evaluation of several different planetary boundary layer schemes within a single model, a unified model and a multimodel superensemble. *Tellus A*, 60, 42–61.
- Krishnamurti, T.N., C.M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C.E. Williford, S. Gadgil and S. Surendran, 1999: Improved skills for weather and seasonal climate forecasts from multi-model superensemble. *Science*, **285**(5433), 1548 - 1550.
- Krishnamurti, T.N., C.M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, C.E. Williford, Gadgil, S. and S. Surendran, 2000a: Multi-model superensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196-4216.

- Krishnamurti, T.N., K. Rajendran, T.S.V. Kumar, S. Lord, Z. Toth, X. Zou, J. Ahlquist, S. Cocke, and I.M. Navon, 2003. Improved Skills for the Anomaly Correlation of Geopotential Heights at 500 hPa. *Monthly Weather Review*, **131**, 1082-1102.
- Mishra, A. K., and T. N. Krishnamurti, 2007: Current status of multimodel Superensemble and operational NWP forecast of the Indian summer monsoon. *Journal of Earth System Science*, **116(5)**, 369–384.
- Palmer, T. N., F. Molteni, R. Mureau, R. Buizza, P. Chapelet, and J. Tribbia, 1993: Ensemble prediction. *Proc. Seminar on Validation of Models over Europe,* Vol. 1, Reading, United Kingdom, ECMWF, 21–66
- Richardson, D., R., Buizza and R. Hagedorn, 2005: Final report of the 1st Workshop on the THORPEX Interactive Grand Global Ensemble (TIGGE). WMO TD No. 1273, WWRP-THORPEX No. 5
- Schaefer J.T. (1990): "The critical success index as an indicator of warning skill". Wea. Forecasting, vol 5, pp 570-575.
- Stefanova, L., and T. N. Krishnamurti, 2002: Interpretation of seasonal climate forecast using Brier skill score, FSU superensemble, and the AMIP-I data set. *J. Climate*, 15, 537–544.
- Toth, Z., and Kalnay, E., 1993: *Ensemble forecasting at NMC: The generation of perturbations.* Bull. Amer. Meteor. Soc., ***74,*** *2317–2330*.

List of Figures:

Figure 1: Absolute error in daily precipitation (mm d^{-1}) of ECMWF forecasts on a typical Day (15th June 2008). Dashed line connecting circles show error growth coming from Bay of Bengal, dotted line connects the rectangles showing growth of absolute errors in South China Sea and dark black continuous line connecting elliptical shapes depicts the errors growth in the Pacific Ocean. All forecasts (Day 1, Day 3, Day 7 and Day 8) are valid for June 15th 2008.

Figure 2: Time series of RMS Errors for day 1, 3, 5, 7, 9 and 10 of FSU multimodel forecasts over China domain

Figure 3 (a): Average rain (mm d^{-1}) and spatial correlations of the FSU multimodel forecast of day 1 for the period Feb 2008 to 30th Apr 2008.

Figure 3 (b): Average rain (mm d^{-1}) and spatial correlations of the FSU multimodel forecast of day 10 for the period Feb 2008 to 30th Apr 2008.

Figure 4 (a): FSU multimodel Total Rain (mm) forecast of Day 1 during the passage of Meiyu front on 14th June 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 4 (b): FSU multimodel Total Rain (mm) forecast of Day 5 during the passage of Meiyu front on 14th June 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 4 (c): FSU multimodel Total Rain (mm) forecast of Day 7 during the passage of Meiyu front on 14th June 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 5: Equitable Threat Score (ETS) and Bias scores for Day 1, 2 and 3 forecasts of FSU multimodels during the passage of Meiyu front on June 14th 2008. Left panels (a), (c) and (e) are ETS and right panels are Bias.

Figure 6 (a): FSU multimodel Total Rain (mm) forecast of Day 1 during the passage of typhoon Kalmaegi front on 19th July 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 6 (b): FSU multimodel Total Rain (mm) forecast of Day 2 during the passage of typhoon Kalmaegi front on 19th July 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 6 (c): FSU multimodel Total Rain (mm) forecast of Day 3 during the passage of typhoon Kalmaegi front on 19th July 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 7 (a): FSU multimodel Total Rain (mm) forecast of Day 1 during the heavy rain event induced by the passage of Tropical system Fengshen (Frank) on the eastern side and influence of East Asian monsoon on the western side on 26^{th} June 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 7 (b): FSU multimodel Total Rain (mm) forecast of Day 2 during the heavy rain event induced by the passage of Tropical system Fengshen (Frank) on the eastern side and influence of East Asian monsoon on the western side on 26^{th} June 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 7(c): FSU multimodel Total Rain (mm) forecast of Day 3 during the heavy rain event induced by the passage of Tropical system Fengshen (Frank) on the eastern side and influence of East Asian monsoon on the western side on 26^{th} June 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 8: Equitable Threat Score (ETS) and Bias scores for Day 1, 2 and 3 forecasts of FSU multimodels during the passage of Tropical System Frank on the West Pacific Coast and influence of East Asian Monsoon from the eastern boundary of the domain on June 26^{th} 2008. Left panels (a), (c) and (e) are ETS and right panels are Bias.

Figure 9 (a): FSU multimodel Total Rain (mm) forecast of Day 1 during Onset of the South China Sea Monsoon on May 3rd 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 9 (b): FSU multimodel Total Rain (mm) forecast of Day 3 during Onset of the South China Sea Monsoon on May 3rd 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 9 (c): FSU multimodel Total Rain (mm) forecast of Day 5 during Onset of the South China Sea Monsoon on May 3rd 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 9 (d): FSU multimodel Total Rain (mm) forecast of Day 10 during Onset of the South China Sea Monsoon on May 3rd 2008. Numbers in every forecast panels show the Root Mean Square Errors (RMSE) and Spatial Correlations (SC).

Figure 10: Equitable Threat Score (ETS) and Bias scores for Day 1, 2 and 3 forecasts of FSU multimodels during the onset of the South China Sea Monsoon on May 3rd 2008. Left panels (a), (c) and (e) are ETS and right panels are Bias.

Table 1: List of Acronyms

BOM	Bureau of Meteorology, Australia
CMA	China Meteorological Administration
CMC	Canadian Meteorological Centre
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos (Center for Weather Forecast and Climatic Studies)
ECMWF	European Centre for Medium-Range Weather Forecasts
EM	Ensemble Mean
ETS	Equitable Threat Score
FSU	Florida State University
JMA	Japan Meteorological Agency
KMA	Korea Meteorological Administration
MF	Météo France
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NSF	The National Science Foundation
RMS	Root Mean Square
SC	Spatial Correlation
SE	SuperEnsemble
THORPEX	THe Observing System Research and Predictability EXperiment
TIGGE	THORPEX Interactive Grand Global Ensemble
TRMM	Tropical Rainfall Measuring Mission
UKMO	United Kingdom Meteorological Office

Table 2: TIGGE Models

Center	Ensemble Members	Model Resolution	Forecast Length
ECMWF	51	N200 (Reduced Gaussian)	10 day
ECMWF	51	N128 (Reduced Gaussian)	10-15 day
UKMO	24	1.25 x 0.83 Deg	15 day
JMA	51	1.25 x 1.25 Deg	9 day
NCEP	21	1.00 x 1.00 Deg	16 day
CMA	15	0.56 x 0.56 Deg	10 day
CMC	21	1.00 x 1.00 Deg	16 day
BOM	33	1.50 x 1.50 Deg	10 day
MF	11	1.50 x 1.50 Deg	2.5 day
KMA	17	1.00 x 1.00 Deg	10 day
CPTEC	15	1.00 x 1.00 Deg	15 day

Appendix I

Equitable Threat Scores and Bias Scores

The traditional threat score (TS, e.g., Anthes 1983) measures the accuracy in predicting area of precipitation amounts over any given threshold. The equitable threat score (ETS, Schaefer 1990) measures the skill in predicting the area of precipitation amounts over any given threshold with respect to a random (no skill) control forecast and is defined as

$$ETS = \frac{H - CH}{F + O - H - CH} \qquad \textbf{A.1}$$

and bias score is the ratio of the forecast area (points) to observed area (points) of precipitation amounts over any given threshold (Anthes 1983). It is defined as

$$Bias = \frac{F}{O} \qquad \textbf{A.2}$$

where F is the number of forecast points above a threshold, O is the number of observed points above a threshold, H is the number of hits above threshold, and CH is the expected number of hits in a random forecast of F points for O observed points, which is equal to

$$CH = \frac{F \times O}{NUM} \qquad \textbf{A.3}$$

A value of 1.0 for ETS indicates perfect forecast. The minimum value for ETS can be –1/3. For bias score, The accuracy of a forecast is directly proportional to the ETS value, an ETS > 0.0 denotes a skillful forecast relative to a random forecast, ETS ≤ 0.0, a forecast has no skill. Bias score of 1 is considered to be the perfect match of forecast to observed rain. Interpretation of Bias score alone is not sufficient. While analyzing Bias score one should consider ETS values also.