Superensemble Forecasts for Weather and Climate

T.N. Krishnamurti, T.S.V. Vijaya Kumar, W.T. Yun, Arun Chakraborty, Lydia Stefanova, Nihat Cubukcu, Eric Williford, Brian Mackey, and J. Sanjay

Department of Meteorology, The Florida State University
Tallahassee, Fl 32306-4520. USA
Email: tnk@io.met.fsu.edu

Abstract

This paper carries a short review of a multimodel/multianalysis superensemble for weather and seasonal climate forecasts. This model was first developed by the authors in 1999 at Florida State University. This entails a large number of forecasts using these multimodels from past data sets, that is called a training phase of the superensemble. During this training phase statistical relation among the model forecasts and the observed fields is obtained using multiple regression methods. This training phase requires roughly 4 months of past daily forecasts for numerical weather prediction (NWP), approximately 6 years of past seasonal forecast and about 60 past hurricane/typhoon/tropical cyclone forecasts from each of the participating member models. The training phase is followed by a forecast phase where the member model forecasts (into the future) use the aforementioned statistics to construct multimodel superensemble forecasts. Our focus on NWP has been to examine the performance of the multimodel superensemble forecast against those of the member models, their ensemble mean and the bias removed ensemble means. We have noted an invariable much superior performance of the multimodel superensemble. We have noted that roughly a minimal number of 7 to 8 models are needed to carry out this exercise. We were also able to improve the database and the statistics of the training phase by rejecting poorer forecast days and optimizing the number of training days. The common metrics for forecast evaluation include root mean square error, anomaly correlation and equitable threat scores. Great impact on real time and experimental forecasts from the superensemble were noted for precipitation, sea level pressure, temperature and 500 hPa geopotential height fields. The improvements in forecasting heavy rains by the multimodel/multianalysis superensemble are found to provide useful guidance in flood events. In hurricane forecasts improvements in track position forecasts of the order of 100 to 250 km were noted in one to three day forecasts. Intensity forecast for hurricanes shows only a marginal improvement. The seasonal climate forecasts show a lower performance from the member models compared to climatology, the multimodel superensemble appears to have skill somewhat above that of climatology.

1. Introduction

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, 2000b and 2001). The novelty of this approach lies in the methodology, which differs from ensemble analysis techniques used elsewhere. This approach yields forecasts with considerable reduction in forecast error compared to the errors of the member models, the 'bias-removed' ensemble averaged forecasts, and the ensemble mean. This technique entails the partition of a time line into two parts. One part is a '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. Specifically, observed anomalies are fit to the member model forecasts according to the classical prescription

$$O' = \sum_{i=1}^{N} \sum_{j=1}^{train} a_i (F_{ij} - \bar{F}_{ij}) + \varepsilon_i \tag{1}$$

where F_{ij} is the i^{th} model forecast (out of N total models), $\overline{F}_{ij}$ is the mean of the i^{th} forecast over the training period (*train*), O′ is the observed anomaly relative to the observed mean over the training period (*train*), the a_i values represent the regression coefficients and ε_i is an error term. The a_i's are determined by requiring the summed squared error integrated over the training period $E = \sum_{i=1}^{N} \varepsilon_i^2$ to be as small as possible. A fit of this sort is performed for all model variables and at all model grid points for which reanalysis observations are available and typically yields close to 7 million regression parameters. This large number arises from the number of transform grid points, number of vertical levels, number of basic variables and the number of models. Over all such locations we have noted diverse performance characteristics of the member models. That arises from differences in: horizontal and vertical descretization, treatment of physics, handling of inhomogenity of land surface, orography, lakes, water bodies, surface physics and boundary conditions. All such peculiarities tend to leave their signature in the error distributions and hence on these weights. These may be thought of as bias correction weights. The second time line part is composed of genuine model predictions, *i.e.* the forecasts of the member models. The superensemble approach combines each of these forecasts according to the weights determined during the training period (*train)* through the formulation

$$S = \overline{O} + \sum_{i=1}^{N} \sum_{j=1}^{train} a_i (F_{ij} - \overline{F}_{ij}) \tag{2}$$

where the notation is defined above except the a_i values, which are the regression coefficients. The determination of a_i follows the well-known Gauss-Jordan elimination method. The prediction '*S*' is referred to as the "superensemble" forecast. This forecast should be contrasted with the more standard anomaly forecasts known as the biased - removed ensemble mean (E) or ensemble mean ($\hat{E}$) forecasts

$$E = \overline{O} + \frac{1}{N} \sum_{i=1}^{N} (F_i - \overline{F}_i) \quad \text{or} \quad \hat{E} = \overline{O} + \frac{1}{N} \sum_{i=1}^{N} (F_i - \overline{O}) \tag{3}$$

The distinction between them comes in the weighting. Assigning all models a weight of 1/N (where N is the number of models) in equation (2) illustrates the connection between the forecasts, and also exemplifies how the training period attempts to identify and highlight good model performance.

The skill of the multi-model superensemble method significantly depends on the error covariance matrix, since the weights of each model are computed from a designed covariance matrix. The current method for the construction of the superensemble utilizes a least square minimization principle within a multiple regression of model output against observed 'analysis' estimates. This entails a matrix inversion that is currently solved by the Gauss Jordan elimination technique. That matrix can be ill conditioned and singular depending on the inter relationships of the member models of the superensemble. We have recently designed a singular value decomposition (SVD) method that overcomes this problem and removes the ill conditioning of the covariance matrix entirely, Yun et al. (2002). Early tests of this method have shown great skills in weather and seasonal climate forecasts compared to the Gauss Jordan elimination method.

In medium range real-time global weather forecasts, the largest skill improvements are seen for precipitation forecasts both regionally and globally. The overall skill of the superensemble is 40 to 120% higher than the precipitation forecast skills of the best global models. For forecasts of variables other than precipitation, the superensemble exhibited major improvements in skill for the divergent part of the wind and the temperature distributions. Tropical latitudes show major improvements in daily weather forecasts. For most variables, we have used the operational ECMWF analysis at 0.5° latitude/ longitude, as the observed benchmark fields, for the training phase. The observed measures of precipitation are derived from the so called 2A12 algorithm of

NASA Goddard that is described in some detail in Krishnamurti et al (2000b, 2001) and within the references stated there in.

The area of seasonal climate simulations has only been addressed recently in the context of atmospheric climate models where the sea surface temperatures and sea ice are prescribed, such as the AMIP data sets. In this context, given a training period of some 8 years and a training data base from the ECMWF, the results exhibit improved skill compared to the member models and the ensemble mean. Preliminary work in this area (Krishnamurti et al. 2002) examines the difficulties involved with prediction of seasonal precipitation anomalies. Most individual member models perform poorly compared to climatology, whereas the superensemble appears to demonstrate precipitation skills slightly above those of climatology. The effectiveness of weather and seasonal climate forecasts from superensemble methodology has also been assessed from measures of standard skill scores such as correlation against observed fields, root mean square (RMS) errors, anomaly correlations and the so-called Brier skill scores (Stefanova and Krishnamurti, 2002) for climate forecasts (assessing skills above those of a climatology).

Training is a major component of this forecast initiative. We have compared training with the best quality 'observed' past data sets versus training deliberately with poorer data sets. This has shown that forecasts are improved when higher quality training data sets are deployed for the evaluation of the multi-model bias statistics. It was felt that the skill during the ‘forecast phase’ could be degraded if the training was executed with either a poorer analysis or poorer forecasts. This was noted in our recent work on precipitation forecasts where we showed that the use of poorer rainfall estimates during the training period affected the superensemble forecasts during the ‘forecast phase’ (Krishnamurti et al. 2000b). In addition, issues on optimizing the number of training days have been addressed from an examination of training with days of high forecast skill versus training with low forecast skill, and training with the best available rain-rate datasets versus those from poor representations of rain. We have learned to improve the forecast skill by selectively improving the distribution of weights for the training phase.

Why does the superensemble generally have higher skill compared to all participating multi-models and the ensemble mean? At each location and for all variables, the ensemble mean assigns a weight of 1/N to all N member models, this includes several poorer-performing models. As a result, assigning the same weight of 1/N to some poorer models was noted to degrade the skill of the ensemble mean. It is possible to remove the bias of models individually (at all locations and for all variables) and to perform an ensemble mean of the bias removed models. Here again the bias removed individual models get equal weight of 1/N and the ensemble mean thus obtained also has somewhat lower skill compared to the superensemble, which carries selective weights distributed in space among all multi-models and for all variables. A poorer model simply does not reach the levels of the best models after its bias removal.

1.1. Experimental Real-time Global Weather Forecasts based on Superensemble

We have developed an experimental real time NWP capability for the forecast of all basic variables such as winds, temperature, surface pressure, geopotential heights and precipitation. These are multianalysis-multimodel superensemble forecasts where eleven models are currently being used on a daily basis. These include the daily operational forecasts from the NCEP, Canadian Weather Service RPN, Australian model for the BMRC, U.S. Navy’s NOGAPS, the Japanese model for JMA and different versions of our in house FSU global spectral model that are physically initialized using different rain rate algorithms (see section 3). In some sense, the construction of the superensemble is a post-processing of multi-model forecasts. The superensemble based forecast is still a viable forecast product that is being prepared experimentally in real time at FSU and is currently available on a real- time basis on the website http://lexxy.met.fsu.edu/rtnwp.

The website http://lexxy.met.fsu.edu/rtnwp has been started with an aim to provide a near real-time multimodel superensemble based weather forecasts over the entire globe up to 6 days in advance. Forecasts for the Mean Sea Level Pressure, Heights at 500 hPa, Surface Temperature, Winds (isotachs) at surface, 850 hPa and 200 hPa are displayed for the whole globe as well as 9 different regions of the globe, these forecasts can be viewed with ease by clicking on a specific region over the World Map provided in the forecast page. Apart from these dynamical variables, 5-day forecasts of the 24-hr total precipitation and a new '5-day accumulated flood potential forecast" is also shown for the entire globe, which again can be viewed over a specific region of interest. Different skill scores computed from these data sets are also shown in the forecast page – they are the RMS errors and systematic errors for forecasts of winds, mean sea level pressure and winds at 850 hPa and 200 hPa; and equitable threat scores, RMS errors and Correlation Coefficients for the precipitation forecasts. The website also features the archives for up to 10 days in the immediate past and provides links to recent publications based on the superensemble technique.

2. The Multimodel Superensemble for Numerical Weather Prediction:

As many as seven global models are being used (Krishnamurti et al, 2000a) for the prediction of weather on a real-time basis. Figure 1 illustrates typical superensemble based weather prediction skills derived from this product. Here the mean RMS forecast errors of 850 hPa winds on day-3 of forecast for various regions of the globe, for the month of August 1998, is shown. The results for member models, an ensemble mean of these member models and that for the superensemble are presented in this figure. Large improvement in reduction of wind forecast errors can be seen over the tropical belt from the superensemble. Basically these results convey what has been stated above on the performance of the superensemble. These results have been confirmed for each month since 1998 to the present time.

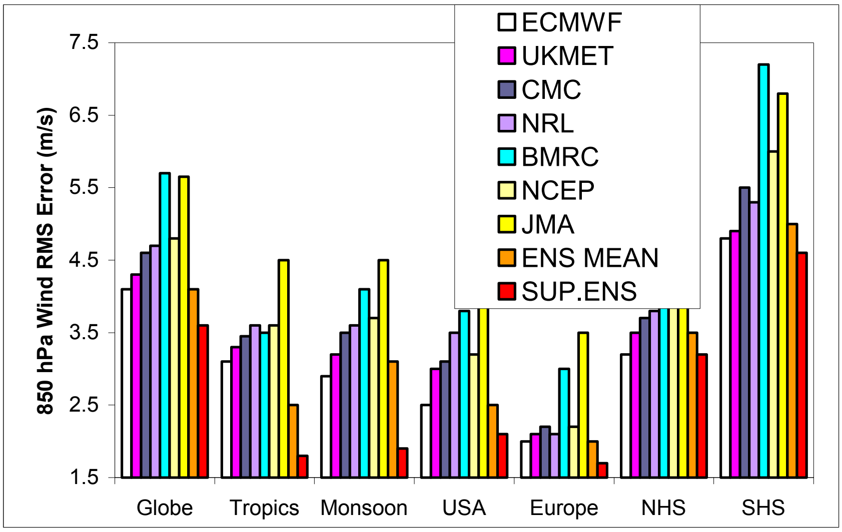


Figure 1. RMS error (ms^{-1}) of 850 hPa winds over different parts of the globe, day-3 forecast, August 1998.

To assess how many models are minimally needed to improve the skill of the multimodel superensemble, we examined sequentially the issue using one to seven models. Results for the mean global wind RMS errors at 850 hPa during August 1998 for day-3 of forecasts are shown in Fig. 2. The models are sequentially added with lower and lower skill while proceeding from one model to seven models. The dashed line shows the error for the ensemble mean and the solid line indicates that of the superensemble. The superensemble skill is higher than that of the ensemble mean for any selection of the number of ensemble members. The skill of the superensemble between 4 and 7 models is small, that is around 3.6 ms^{-1}. The ensemble mean error increases as we add more ensemble members beyond 3 – this is due to the gradual addition of models with lower skill. That rapid increase is not seen for the superensemble since it automatically assigns low weights to the models with low skill. It is also worth noting that half the skill improvement comes from a single model for this procedure. Even if forecasts from a single model were available, it is possible to construct the superensemble of one model. That does differ from the ensemble mean of one model (which is the same as the single model forecast). The training phase of a single model provide linear regression coefficients based on past errors of the model, that is described by an equation of the type y = mx + n, and can even improve the result of a single model forecast via the superensemble approach by as much as 20 percent as shown in Krishnamurti et al (2000a).

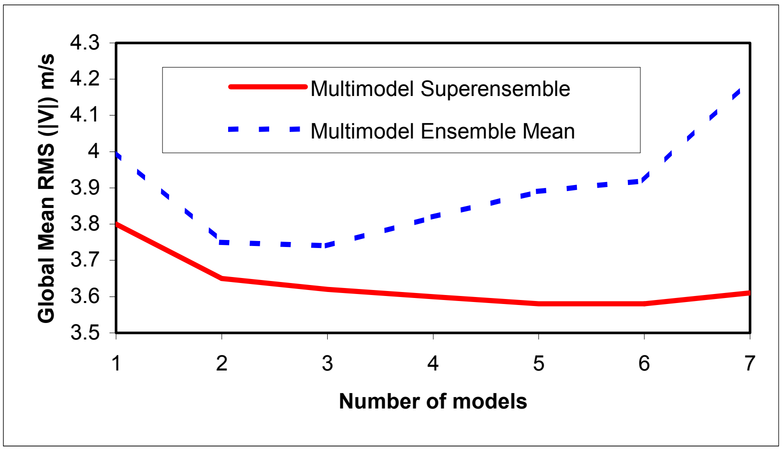


Figure 2. Global Mean RMS error (ms^{-1}) of total wind at 850 hPa during August 1998 for day-3 forecast.

Anomaly correlation of geopotential height at 500 hPa is another stringent measure for assessing the performance of superensemble in the medium range weather forecasts. Table 1 provides some recent results of anomaly correlations at 500 hPa for the global belt obtained from real time superensemble. Here the entries for the anomaly correlation skills covering a forecast period from August 20th to September 17th 2000 are presented. Results for the member models, the ensemble mean and the superensemble are included here. Results for days 1 through 6 of forecasts are provided in this table. A consistent high skill for the superensemble around 0.75 to 0.8 for day 6 is noted in these experimental runs. Also shown in this table are the entrees for the ensemble mean, which lie roughly halfway in between the best model and the superensemble. Thus it appears that a substantial improvement in skill is possible from the use of the proposed superensemble. The overall improvement of the superensemble over the best (available) model is around 10 percent. This improvement of superensemble is a result of the selective weighting of the available models during the training phase. We have also noted that that the skills over the Southern Hemisphere reach those of the Northern Hemisphere from this procedure (Krishnamurti et al. 2003).

	Day-1	Day-2	Day-3	Day-4	Day-5	Day-6
Superensemble	0.992	0.979	0.958	0.928	0.881	0.799
Ensemble Mean	0.983	0.962	0.935	0.891	0.827	0.756
Model-1	09.84	0.967	0.936	0.889	0.824	0.713
Model-2	0.981	0.957	0.932	0.880	0.796	0.623
Model-3	0.963	0.930	0.885	0.815	0.706	0.579
Model-4	0.962	0.925	0.871	0.786	0.697	0.578
Model-5	0.956	0.918	0.858	0.767	0.665	0.549
Model-6	0.941	0.889	0.846	0.739	0.632	

Table 1 500 hPa Global Geopotential Height Anomaly Correlation: 20Aug – 17Sep 2000

3. Precipitation Forecasts from TRMM–SSM/I based Multianalysis Superensemble:

A major improvement in tropical precipitation forecasts has emerged from the use of a TRMM – SSM/I based multi-analysis superensemble (Krishnamurti et al. 2000b). "Multi-analysis" refers to different initial analyses contributing to forecasts from the same model. In this study, the multi-analysis component is based on the FSU GSM initialized with TRMM and SSM/I data sets via a number of rain rate algorithms. Five different initial analysis for each day are deployed that defined the multianalysis component. Those are based on different versions of rain rate estimated derived from TRMM and the DMSP-SSM/I satellites. These rain rate initializations of the different rain rate estimates followed the physical initialization procedures outlined in Krishnamurti et al. (1991). The differences in the analyses arise from the use of these rain rates within a physical initialization. The resulting initialized fields have distinct differences among their initial divergence, heating, moisture, and rain rate descriptions.

The impact of physical initialization on the improvement of precipitation forecast skills was examined in detail by Treadon (1996) where he used the GPI based rain rates for physical initialization. Figure 3 illustrates the correlation of rainfall (observed versus modeled) plotted against the forecast days. Here a very high nowcasting skill of the order of 0.9 is seen, this was a feature of the physical initialization, also noted by Krishamurti et al. (1994). But the forecast skill degrades to 0.6 by day 1 of the forecast and it degrades even more by days 2 and 3 to values such as 0.5 and 0.45 respectively. Using the proposed superensemble approach, it is possible to improve the forecast skills when the TRMM-SSM/I based rain rates are used as a benchmark for the definition of the superensemble statistics and forecast verification.

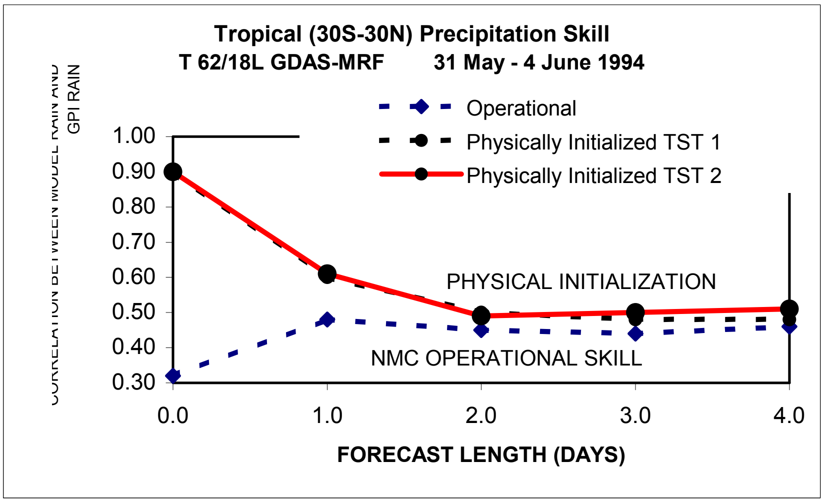


Figure 3. Skill of precipitation forecasts over global tropics based on point correlation (Treadon, 1996).

Figure 4 illustrates the TRMM based forecast skills over the global tropics. Here noticeable improvement of short-range forecasts of precipitation is noted beyond what was obtained in previous studies. The three lines in Fig. 4 show correlations of rainfall (observed versus modeled) as a function of forecast days 0, 1, 2 and 3. The top line in this illustration shows the multianalysis superensemble forecast. The next line, adjacent to the above line, is the forecast from a single global model that utilizes physical initialization of rain rates based on TRMM and SSM/I data sets using the 2A12 and GPROF algorithms respectively. The last (bottom) line with lowest skill represents the results from a control experiment that did not make use of any rain rate initialization. It is clear from these illustrations that the skills from the multianalysis superensemble are higher. These forecast results are based on 5 experiments for each start date during 1-5 August 1998. The day-3 forecast skill reaching as high as 0.7 is indeed a very high skill for rainfall forecasts.

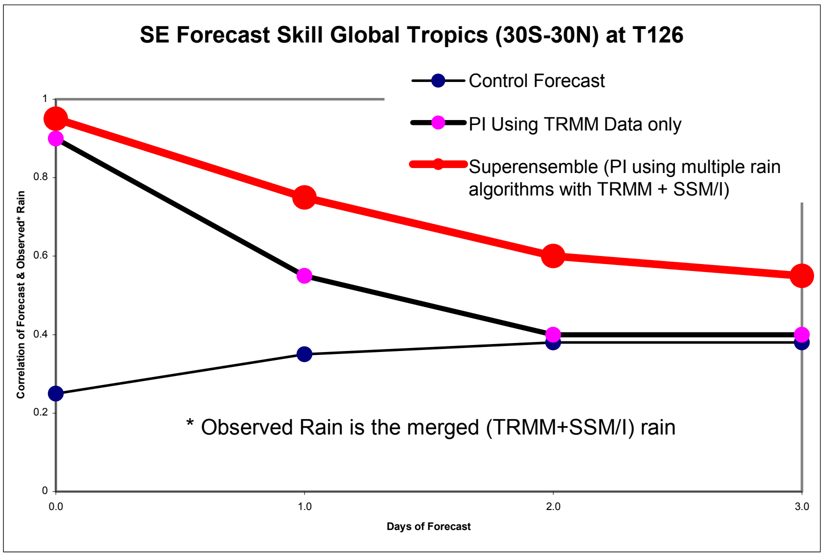


Figure 4: Skill of precipitation forecasts over global tropics (30°S – 30°N) for control forecast, physical initialization using TRMM data only, and superensemble forecasts where the TRMM plus SSM/I rain rates are used as a bench mark.

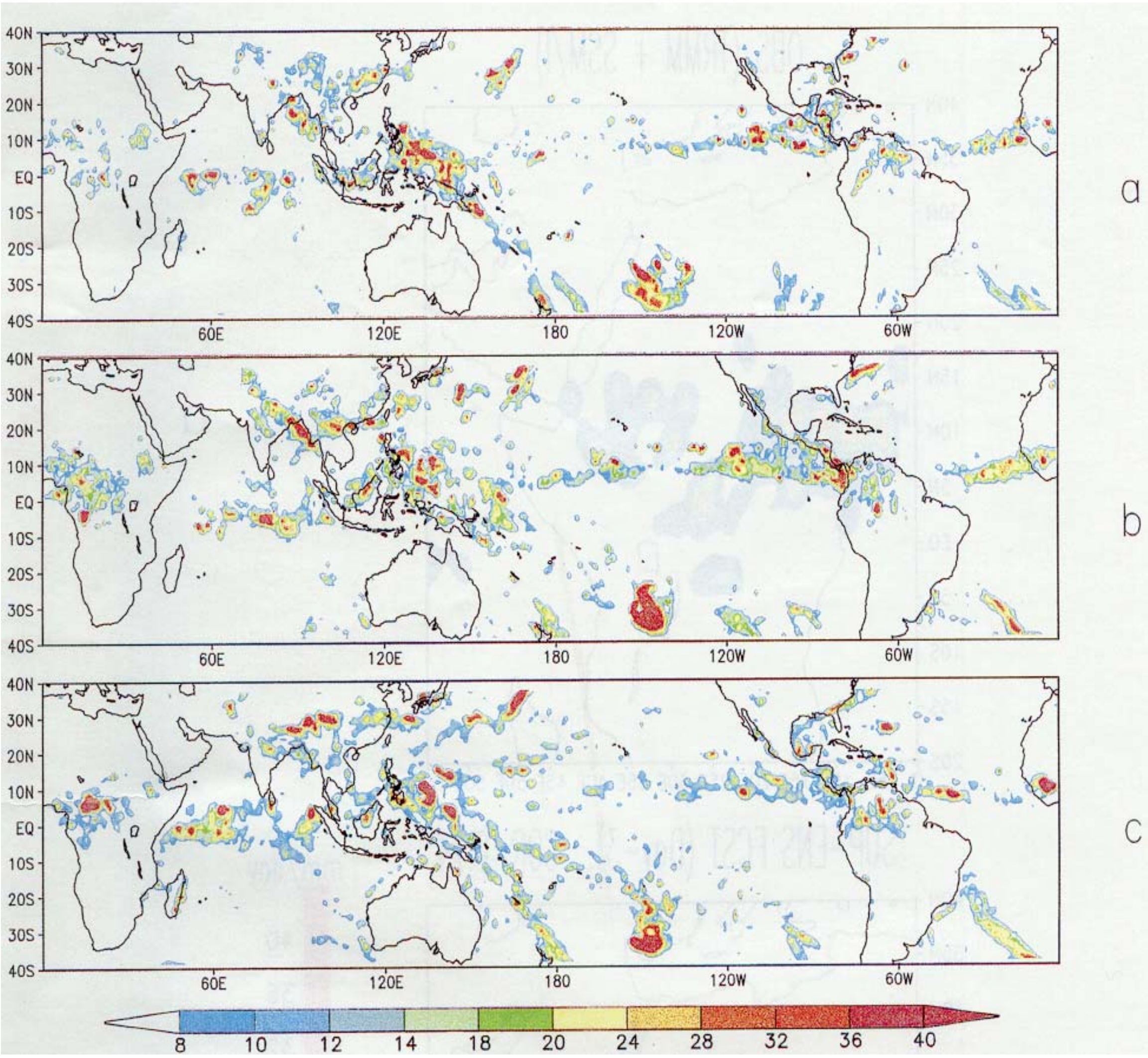


Figure 5. Day-3 rainfall forecast over the global tropical belt, 1200 UTC 15 August 1999: the accumulated rainfall (mm day^{-1}) by (a) observation based on TRMM and SSM/I, (b) multianalysis superensemble and (c) a best single model.

An example of the day-3 forecasts of the precipitation over the global tropical belt is illustrated in Fig. 5 (a, b and c). Figure 5a shows the observed TMI and SSM/I based 24-hr rainfall estimate (mm.day^{-1}) between 1200 UTC 14 August and 1200 UTC 15 August 1999. Figure 5b shows the 3-day forecast form the multianalysis superensemble valid for the same period while Fig. 5c shows the corresponding results from a single best model. The global tropical correlation between the observed and the multianalysis superensemble is 0.55 where the correlation of the best model with respect to the observed estimate is 0.30 for these day-3 forecasts. This reflects a major improvement in rainfall forecasts over the global tropics.

The next area of our research was multianalysis/multimodel superensemble (Krishnamurti et al., 2001). The 12 panels of Fig. 6 illustrate the day 3 rainfall forecasts valid on June 6, 2000. Here the observed rain is shown on the top left panel. The left panels show the multimodel rainfall distributions and the right panels show those from the multianalysis components of the forecasts. The right panel is based on the forecasts from the FSU model at a resolution T126, using different rain rate algorithms in their description of the initial rain. The FSU model's rainfall is, in general, larger than the operational models, and its location and phase errors are generally smaller. Overall, this is the type of multimodel/multianalysis rainfall distributions that we use to construct the superensemble forecasts in our experimental real time forecasts.

Some important results from the 11-model superensemble are presented here. We calculate three measures of skill on a regular basis: i) correlation of model predicted daily rainfall totals and observed estimates; ii) rms errors of model predicted daily rainfall totals; iii) Equitable threat scores for different thresholds of observed and predicted rain. The root mean square errors (RMSE) in precipitation forecasts over the global belt, 50º S to 50º N, covering a forecast period from April 1 to April 15, 2000 are shown in Fig. 7. The training period for these forecasts included the preceding 75 days. The thick black lines denote the RMSE for the multianalysis/multimodel superensemble. The dotted lines show the skills for the selected individual member models, whose skills were high. The thin, solid line shows results for the ensemble mean, with bias removal for individual models. Overall, these results over the global belt show a great promise for the 3-day forecasts of precipitation. It should be pointed out that these results are fairly robust and we see the same skills in the day-to-day real time runs. There is some noticeable improvement in the skill for the superensemble over the ensemble mean. That arises from the fact that the poorer models are assigned weights of 1.0 over the entire globe, whereas the superensemble is more selective regionally (and vertically, for each variable and for each model). Its weights are fractional positive or negative based on the member models' past performance.

We can also look at the correlation of the observed rain (24 hourly totals ending on days 0, 1, 2 and 3 of forecasts) derived from the TRMM-2A12 plus the SSM/I-GPROF based rainfall against the global gridded forecasts of the superensemble-based rains. Those are shown in Fig. 8 for the months September and October 2000. The global forecast correlation skills for days 0, 1, 2 and 3 lie roughly around 0.9, 0.8, 0.62 and 0.55 for these months. These are higher skills compared to what were seen for a single model shown in Fig. 3. Similar results are noted for all the recent months.

As was summarized in Krishnamurti et al. (2001), the one to three days forecast skills of the daily precipitation totals for the three metrics used here are indeed the highest for the superensemble. Table 2 illustrates the results of the threat scores for eight participating members of the real time multimodel/multianalysis system. The threat scores are evaluated covering the precipitation rate intervals greater than 0.2, 10.0, 25.0, 50.0 and 75.0 mm/day. The size of the individual domains is identified within the table. The bias for the member models, ensemble mean and superensemble, were found to be comparable (not shown). The threat scores for the superensemble for all rainfall intervals are the highest compared to the member models and the ensemble mean rainfall. We have also shown the threat scores for the ETA model in the last column over North America. The forecasts for the member models and superensemble are all for August 2000. This covers a 31-day period. The ETA model's equitable threat scores for August over

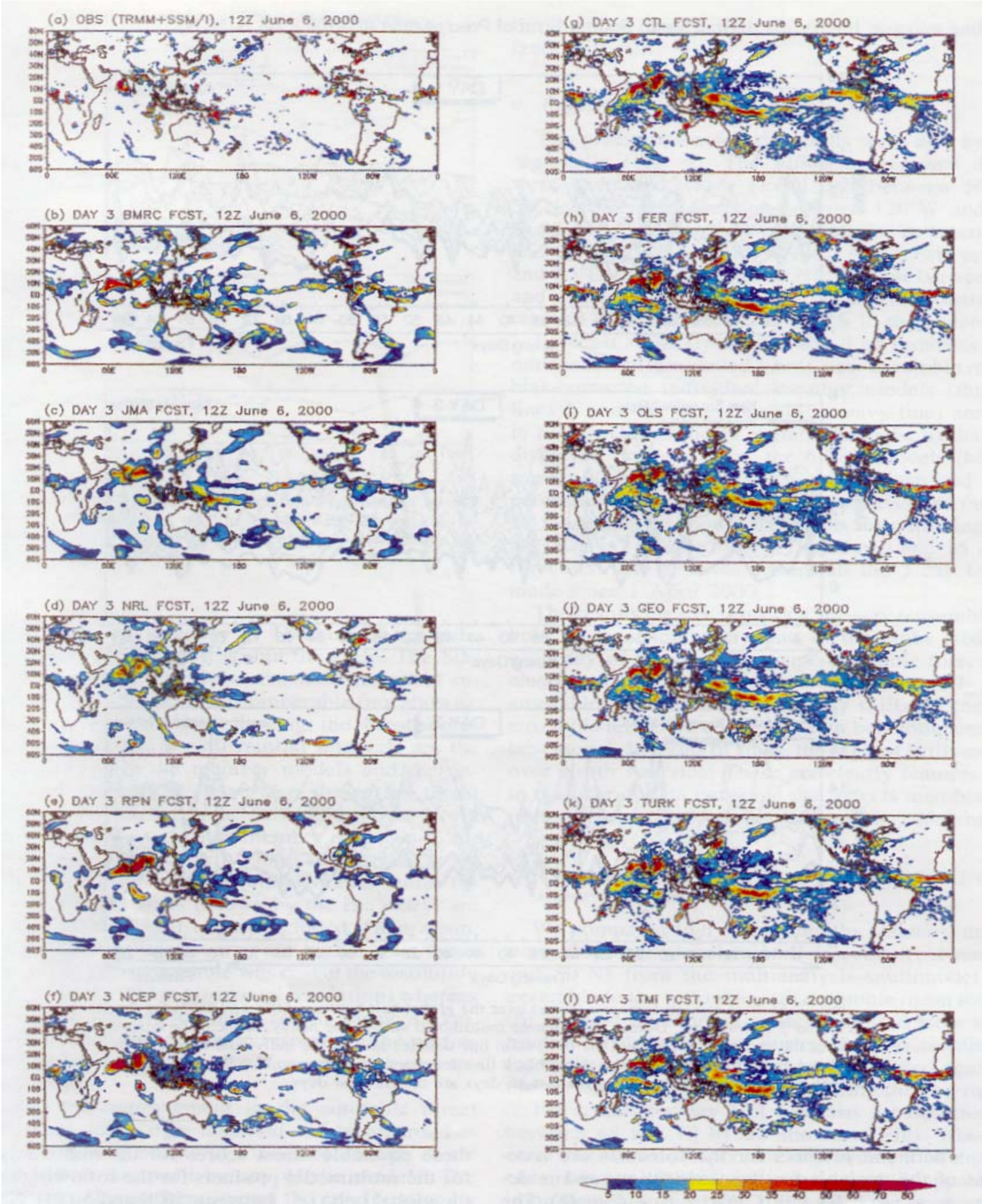


Figure 6. The observed rainfall estimate from the TMI-2A12 and SSM/I-GPROF algorithm for 5 June 2000 is compared with the day-3 forecasts from the 11 member models of the multimodel-multianalysis system.

Precipitation Equitable Threat Scores for August 2000

Pr mm	Member Models 1	2	3	4	5	6	7	8	Ens Mean	Super Ensemble	ETA Model
	Global (50S−50N)										
0.2	0.313	0.295	0.343	0.302	0.296	0.268	0.276	0.273	0.386	0.568	
10	0.237	0.157	0.195	0.132	0.190	0.152	0.174	0.157	0.219	0.312	
25	0.215	0.117	0.153	0.089	0.165	0.114	0.136	0.119	0.148	0.257	
50	0.171	0.088	0.112	0.064	0.145	0.081	0.092	0.080	0.112	0.198	
75	0.073	0.057	0.012	0.000	0.037	0.044	0.055	0.044	0.011	0.272	
	North America (120W−65W, 20N−50N)										
0.2	0.202	0.256	0.200	0.171	0.180	0.222	0.232	0.215	0.305	0.641	0.308/1999
10	0.088	0.062	0.020	0.021	0.014	0.072	0.092	0.076	0.066	0.458	0.288/1995
25	0.054	0.045	0.000	0.012	0.000	0.038	0.066	0.049	0.006	0.425	0.221/1995
50	0.033	0.005	0.000	0.012	0.000	0.021	0.036	0.028	0.008	0.142	0.199/1995
75	0.013	0.000	0.000	0.012	0.000	0.000	0.020	0.014	0.000	0.039	0.131/1991
	South America (110W−10W, 50S−15N)										
0.2	0.340	0.261	0.309	0.325	0.266	0.240	0.248	0.247	0.369	0.594	
10	0.298	0.160	0.222	0.130	0.189	0.161	0.171	0.153	0.243	0.333	
25	0.251	0.118	0.153	0.083	0.148	0.119	0.135	0.114	0.133	0.276	
50	0.166	0.079	0.071	0.040	0.102	0.080	0.087	0.071	0.053	0.216	
75	0.115	0.052	0.026	0.012	0.057	0.048	0.053	0.041	0.018	0.151	
	Asia (50E−120E, 15S−45N)										
0.2	0.390	0.474	0.543	0.426	0.458	0.428	0.459	0.440	0.589	0.636	
10	0.306	0.172	0.270	0.197	0.236	0.165	0.200	0.177	0.246	0.352	
25	0.267	0.131	0.211	0.132	0.170	0.122	0.155	0.133	0.175	0.279	
50	0.198	0.092	0.160	0.045	0.122	0.088	0.112	0.072	0.132	0.198	
75	0.153	0.077	0.117	0.020	0.060	0.075	0.090	0.055	0.041	0.172	
	Africa (20W−55E, 35S−40N)										
0.2	0.411	0.462	0.457	0.416	0.396	0.431	0.447	0.439	0.569	0.692	
10	0.249	0.246	0.189	0.143	0.167	0.261	0.295	0.274	0.248	0.357	
25	0.217	0.167	0.137	0.105	0.131	0.190	0.216	0.204	0.151	0.286	
50	0.141	0.096	0.064	0.052	0.055	0.111	0.101	0.109	0.087	0.185	
75	0.097	0.065	0.052	0.017	0.036	0.085	0.075	0.078	0.025	0.145	
	Australia (110E−160E,40S−0)										
0.2	0.363	0.341	0.368	0.380	0.340	0.292	0.324	0.322	0.364	0.425	
10	0.271	0.191	0.264	0.192	0.242	0.197	0.199	0.185	0.273	0.332	
25	0.218	0.146	0.219	0.114	0.194	0.156	0.165	0.141	0.193	0.285	
50	0.145	0.119	0.130	0.049	0.140	0.111	0.121	0.102	0.130	0.185	
75	0.094	0.116	0.085	0.029	0.111	0.088	0.089	0.076	0.073	0.155	

*** Pr mm denotes precipitation class intervals for rainfall rates greater than the indicated amount in column 1. The threat score for the respective member models over the indicated domain are displayed for the entire month of August 2000. The ETA models threat scores for August of several years (with the highest scores) are shown in the last column for the North American region. The 98 days training period ends with 1 August 2000.**

Table 2 recipitation Equitable Threat Score for the respective member models over the identical domain are displayed for the entire month of August 2000. The Eta Model's threat scores for August of several years (with the highest scores) are shown in the last column for the North American Region.

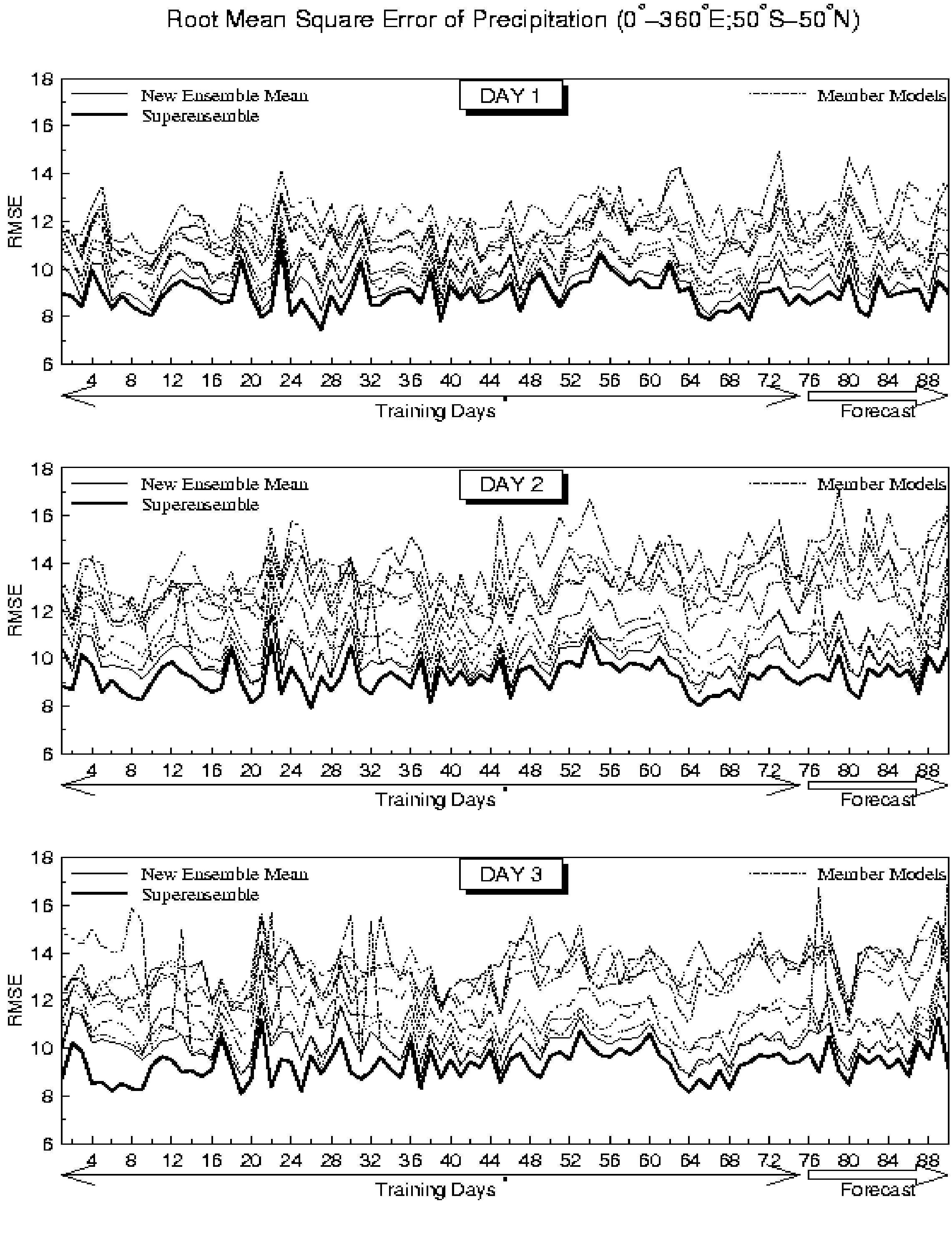


Figure 7: Skill of rainfall forecasts (RMSE) over the global belt between 50oS and 50oN for days 1, 2, and 3 of forecasts. Dotted lines denote multimodel skills. The heavy, dashed line denotes skill of the ensemble mean, and the thin, solid line denotes skill of the individual model's bias removed ensemble mean and the thick, black line denotes the superensemble. The first 75 days denote a training period whereas the last 15 days are the forecast days.

different years (shown by the ETA entry) are shown with their highest scores included. Here again, the superensemble threat scores are higher than those for the ETA. The superensemble was cast at the resolution T126 (i.e. roughly 90km horizontal resolution) whereas the operational ETA model had a resolution of 32km. Considering those differences in resolution, the performance of the superensemble (for these experiments) appears impressive. Although the improvement in the equitable threat scores appear quite large, they should still be regarded as modest. Heavy rain events in excess of 75mm/day are not handled very well by any of the models. The superensemble also underestimates the rain by roughly a factor of 2. We have examined such cases in some detail and it is clear that much further improvement is needed from the member models in order to improve the superensemble based rain. This may require higher resolution modeling for the member models with improved physics and initialization of rain.

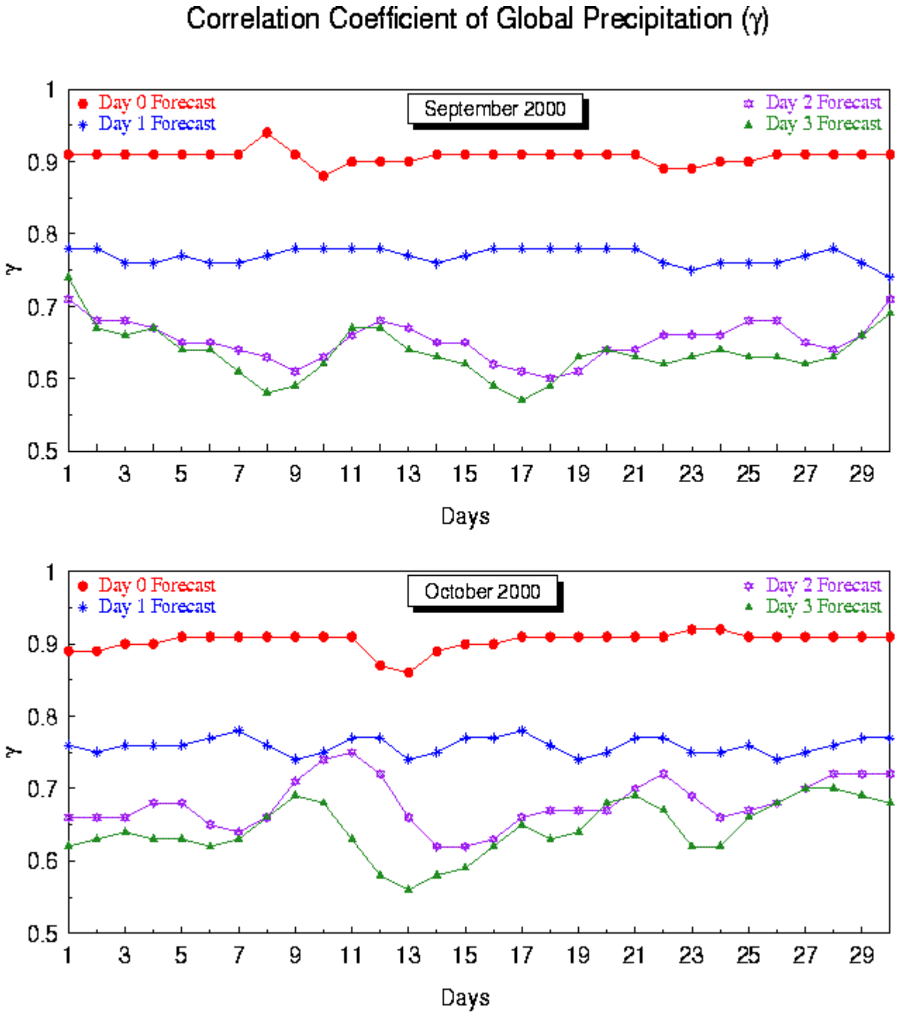


Figure 8. Forecast skill based on correlation of observed rainfall estimates from TRMM-2A12 and the SSM/I-GPROF and the superensemble for day 1, day 2 and day 3 forecasts during September and October 2000.

It is of considerable interest to ask whether the superensemble forecasts of rainfall can provide any useful guidance for floods. Mostly the heavy rains that resulted in the recent Mozambique floods during February and March of the year 2000 resulted from heavy rains over Mozambique and Zimbabwe. The headwaters of the Limpopo River over Zimbabwe experienced the heaviest rainfall that resulted in the cresting of the river over southern Mozambique where the flooding was most severe. Forecasts of rain from this study were projected on Hovmoller diagrams (longitude versus time) and are shown in Fig. 9. Here we show the daily rainfall for the belt 10ºS to 15ºS covering the longitudes 24ºE to 45ºE for the entire month of February (dates are here plotted from bottom to up). The three panels show the 'observed' estimates, those based on the superensemble forecasts (for day 3 of forecasts) valid on the dates of the observed rains of the left panel from the multimodel superensemble and also those predicted for day 3 of forecasts from the best operational model for this region. The best model is determined from the RMSE of rainfall for each model. The units of rainfall are in mm/day. It is clear that the multimodel superensemble carries the 3-day forecasts of heavy rains associated with the Mozambique floods very well. Given the higher rainfall forecast skills from the superensemble, it appears that useful guidance of heavy rain events resulting in floods may be possible from this approach. We have examined the performance of superensemble in flood forecasting issues for about 10 case studies and the results are presented in the real time FSU superensemble NWP website (http://lexxy.met.fsu.edu/rtnwp/pubs/Floods_TRMM.ppt).

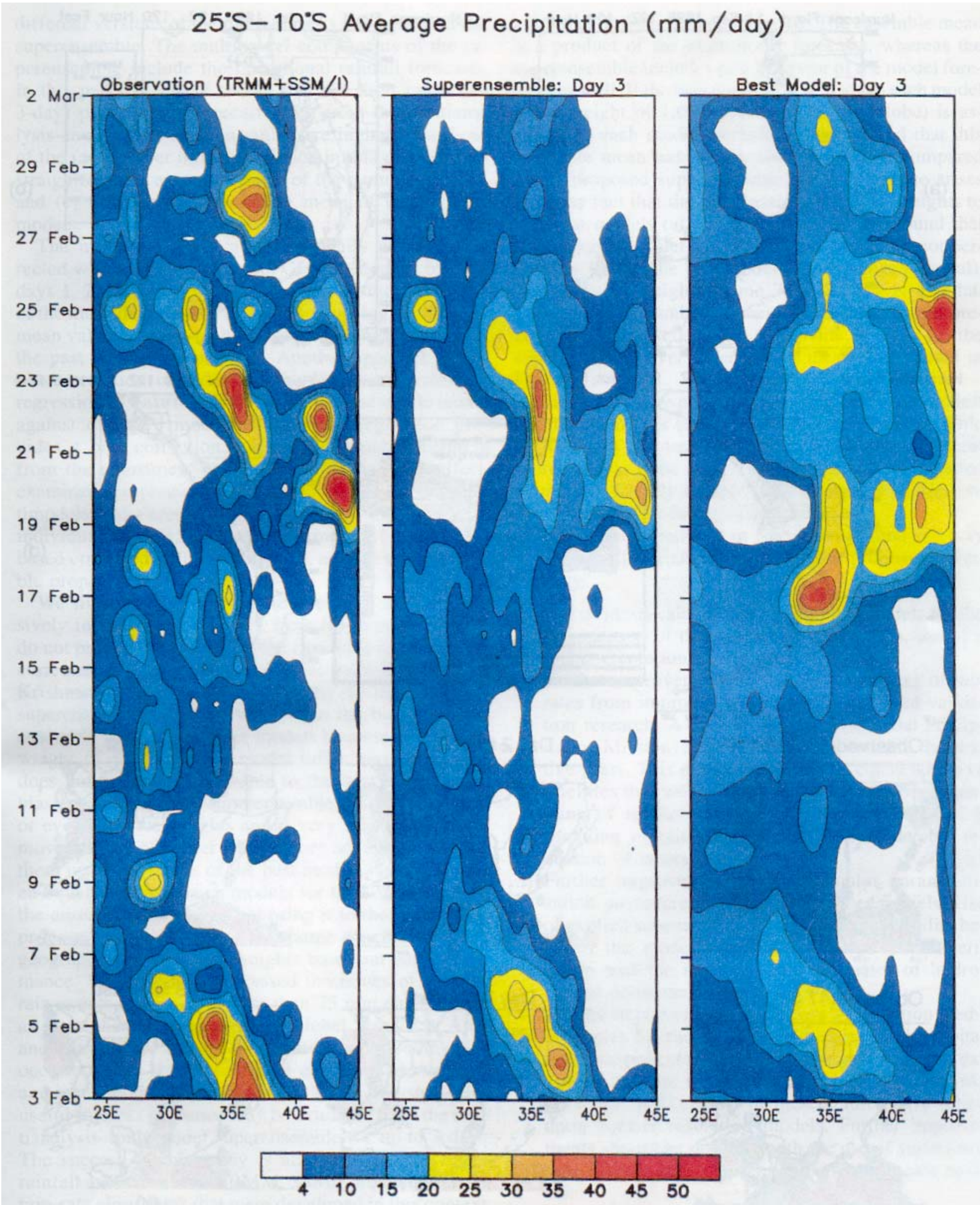


Figure 9. A Hovmoller diagram of daily precipitation (mm day-1) on day 3 of forecasts during the Mozambique floods. Ordinate shows days (bottom to top); abscissa denotes longitude. The three panels denote (left) observed rain (from TRMM-2A12 plus SSM/I GPROF); (middle) superensemble forecasts; (right) best operational model.

A walk-through example of a heavy rainfall and associated flooding over Philippines is illustrated in the following section.

3.1. Walking through a heavy rain computation during a recent flood episode over western Philippines

In Table 3, we present a sequence of computations that highlight the heavy rain forecast over Western Philippines during the passage of Typhoon Halong during July 2002. In this table the first column identifies 9 of the member models. The training coefficient for the superensemble is shown in column 2, these are based on the predicted rain by the member model during the proceeding 120 days (roughly). The predicted rain for day 3 of forecasts (a 24-hour total); the ensemble mean of forecasts, observed rain, the superensemble based rainfall forecast and the bias removed ensemble mean rainfall forecast are also provided in column 2. The third column shows the mean precipitation $\overline{F}_i$ of the member models during the training phase. The fourth column shows the value of superensemble function $a_i\left(F_i - \overline{F}_i\right)$ for each of the member models. The sixth column shows the bias corrected forecast for each model. The last column shows the forecast error for each member model, it also include the errors for the ensemble mean, for the superensemble and for the bias corrected ensemble mean. An examination of this table show that the coefficients range from 0.01 to 0.60. All models underestimate the rainfall on day 3. These coefficients do not reflect this single case, they show the behavior of the member model during the past 120 days.

Multimodel Superensemble Precipitation Forecast (mm.day^{-1}) from h48 to h72
Valid from 20020709/12 UTC thru 20020710/12 UTC
Valid at 15.43 deg N, 120.00 deg E (western shore of Luzon, Philippines)

Model	Coef. a_i	Pcp. F_i	Mean Pcp $\overline{F}_i$	$a_i\left(F_i - \overline{F}_i\right)$	$\overline{O}$ = 26.14 $\overline{O} + \left(F_i - \overline{F}_i\right)$	Error (mm) F_i - OBS
BMRC	0.60191	88.38	31.52	34.23	83.00	-22.06
FSUFER	0.01805	19.38	11.40	0.14	34.12	-91.06
JMA	0.08609	60.46	12.73	4.11	73.87	-49.98
NCEP	0.22313	91.68	15.89	16.91	101.92	-18.76
NRL	0.22343	57.33	10.64	10.43	72.82	-53.11
RPN	0.08697	78.36	20.21	5.06	84.29	-32.08
FSUCTL	0.43546	24.98	18.00	3.04	33.12	-85.46
FSUOLS	0.08800	23.94	11.46	1.10	38.61	-86.50
FSUTRM	0.06130	20.32	11.23	0.56	35.22	-90.12
ENSMEAN		51.65				-58.79
OBS		**110.44**				
SUPENS		101.71				-8.73
BIAS-REM ENSMEAN		61.89				-48.55

Table 3: Walking through a day 3 Superensemble precip. forecast

The mean rain of the training phase $\overline{O}$ was 26.14 mm/day at the location. That added to $\sum_i a_i\left(F_i - \overline{F}_i\right)$, shown in column 5 provides the superensemble forecast. The superensemble forecast of rain for day 3 was 101.71 mm/day. The observed rain 110.44 mm/day. The best model forecast at this location for day 3 of forecast came from the NCEP model. That is not true at all locations and at all ranges of forecasts. The ensemble mean forecast of rain was 51.65 m/day; the bias correlated ensemble mean was slightly better, i.e., 61-89 mm/day. If we were to proceed to an adjacent region, away from this typhoon, on the same day of forecast, we can still see a superior performance of the superensemble although the best model may not be

NCEP. This is shown in Table 4 for comparison purposes. Here the BMRC model exhibited the best forecast among the member models. At this location, the relative spread of statistical weights is somewhat different. This is located at the South China Sea where the observed rainfall was 85.89 mm/day. The overall performance of the superensemble is similar over most regions of the tropics. While the models undergo considerable relative spreads in their forecast from one region to another, the superensemble appears to be more consistent in its overall performance. It its also of interest to note that the forecast errors of the ensemble mean and of the bias removed ensemble mean. It is important to note that no simple relation exists between higher value of the weights (from training) and the forecast of rain in a specific case. For instance, over the South China Sea, the FSU control experiment (FSUCTL) has the highest weight, however, the forecast of rain is poor at this location for this rain event.

Multimodel Superensemble Precipitation Forecast (mm.day^{-1}) from h48 to h72
Valid from 20020709/12 UTC thru 20020710/12 UTC
Valid at 16.36 deg N, 116.25 deg E (South China Sea)

Model	Coef. a_i	Pcp. F_i	Mean Pcp $\overline{F}_i$	$a_i\left(F_i - \overline{F}_i\right)$	$\overline{O}$ = 17.70 $\overline{O} + \left(F_i - \overline{F}_i\right)$	Error (mm) F_i - OBS
BMRC	0.49539	96.73	10.75	42.59	103.68	+10.84
FSUFER	0.16146	17.09	12.71	0.71	22.07	-68.80
JMA	0.21269	34.56	10.59	5.10	41.67	-51.33
NCEP	0.27851	57.93	6.02	14.46	69.61	-27.96
NRL	0.22542	7.01	3.62	0.76	21.09	-78.88
RPN	0.12789	40.31	13.70	3.40	44.30	-45.59
FSUCTL	0.60904	23.30	15.31	4.87	25.69	-62.59
FSUOLS	0.22708	18.64	10.25	1.90	26.08	-67.26
FSUTRM	0.21834	25.91	11.78	3.09	31.83	-59.98
ENSMEAN		35.72				-50.17
OBS		**85.89**				
SUPENS		94.58				+8.68
BIAS-REM ENSMEAN		42.89				-43.00

Table 4: Walking through a day 3 Superensemble precip. forecast

4. Seasonal Climate Forecasts from Multimodel Superensemble:

In the area of seasonal climate forecasting, several papers have been published on the initial development of strategies and application with AMIP (Atmospheric Model Inter-comparison Project) data sets Krishnamurti et al., (1999, 2000a and 2000b) and a first effort with four versions of the FSU Coupled Models, Krishnamurti et al., (2002). Using some arbitrary selection of 8 models from about 31 different global models of AMIP, the superensemble is constructed. All these models had a 10-yr-long integration from 1979 to 1988. The training period consisted of the last 8 years of the data sets while the first two years (1979 and 1980) were subjected to the forecast phase of the superensemble. Monthly mean simulations along with the monthly mean analysis fields provided by ECMWF were used to generate the anomaly multiregression coefficients at each grid point for all vertical levels and all basic variables of the multimodels. Figure 10 shows the time sequence of the RMS error for the meridional wind over the global tropics. One can observe a marked improvement in the skill scores achieved using the superensemble approach.

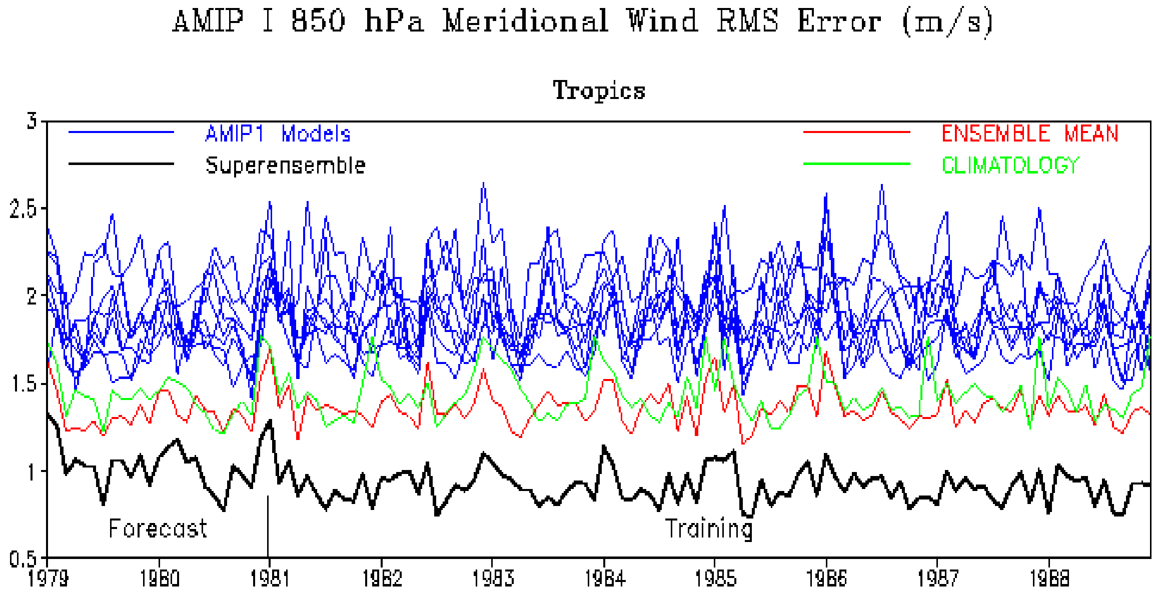


Figure 10. 850 hPa meridional wind RMS error (m.s^{-1}) from AMIP 1 data sets. Training phase is from 1981 to 1988 and forecast phase is from 1979 to 1980. Results from AMIP 1 model forecasts, ensemble mean, superensemble and climatology are shown. (after Krishnamurti et al., 2000a)

Further to this study, several types of model-generated data sets are examined to address the question of seasonal prediction of precipitation over the Asian and the North American monsoon systems (Krishnamurti et al., 2002). The main question we asked is if there is any useful skill in predicting seasonal anomalies (beyond those of climatology). The superensemble methodology is applied here to the anomalies of the predicted multimodel data sets and the observed (analysis) fields. We noted that the superensemble based anomaly forecasts have somewhat higher skill compared to the ensemble mean of member models, individually bias removed ensemble mean of the member models, the climatology and the member models that are being used in this exercise. The illustration for the seasonal correlations (of model rainfall anomalies) with respect to the observed anomalies is presented in Figs. 11a and 11b for Asian Monsoon domain and North American Monsoon domain respectively . The highest anomaly correlations for the seasonal precipitation forecast are generally seen for the multimodel superensemble (heavy line). The other heavy line shows the results for the ensemble mean, the remaining thinner lines show the skill of the member models of AMIP1 and AMIP2. The calculations carried out here used the cross-validation technique, i.e. all years (except the one being forecasted) were used to develop the training data statistics. The skill of forecasts from the superensemble come partly from the forecast performance of multimodels and partly from the training component built into this system that is based on past collective performance of these multimodels. We have separated these components to assess the improvements of the superensemble. Though skill of the forecasts from the superensemble is found to be higher than that of the ensemble mean and has shown some usefulness over the climatology, the issue of forecasting a season in advance in quantitative terms still remains a challenge and demands further advancement in climate modeling studies.

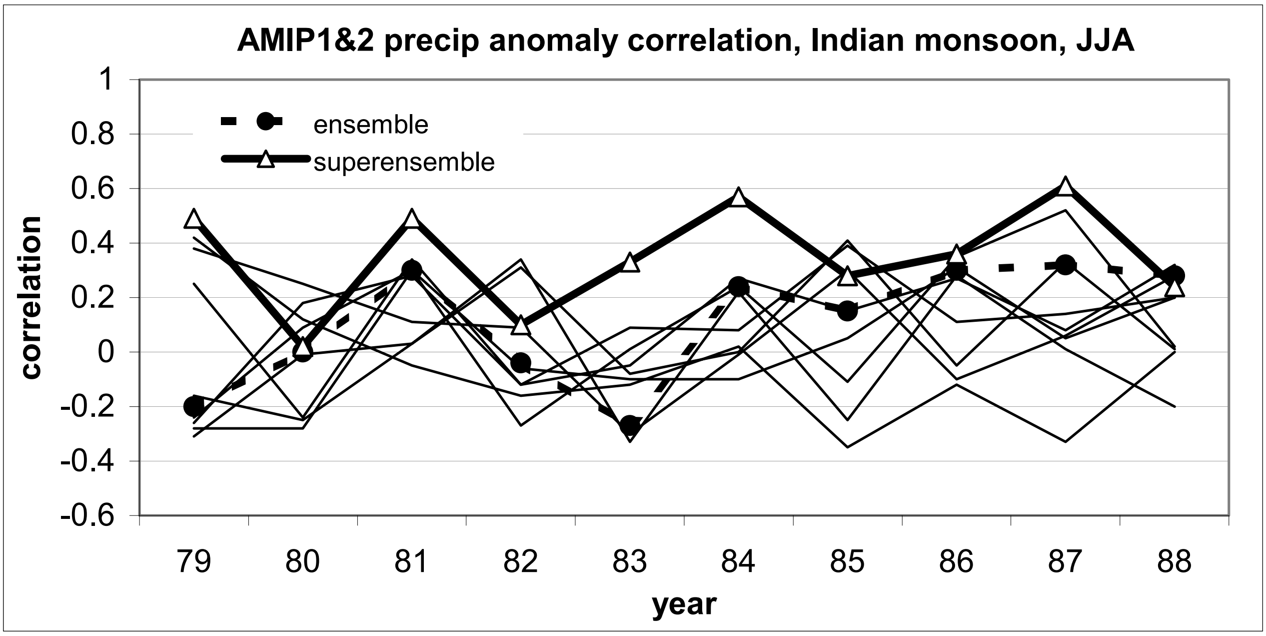


Figure 11. (a) Correlation of seasonal forecasts of precipitation anomalies with respect to observed anomaly estimates based on Xie and Arkin (1996) from the mix of AMIP I and AMIP II data sets. Heavy line at top: Superensemble. Other heavy lone: Bias removed ensemble mean. Thin lines: Member multimodels. Ten years of summer monsoon results are shown here.

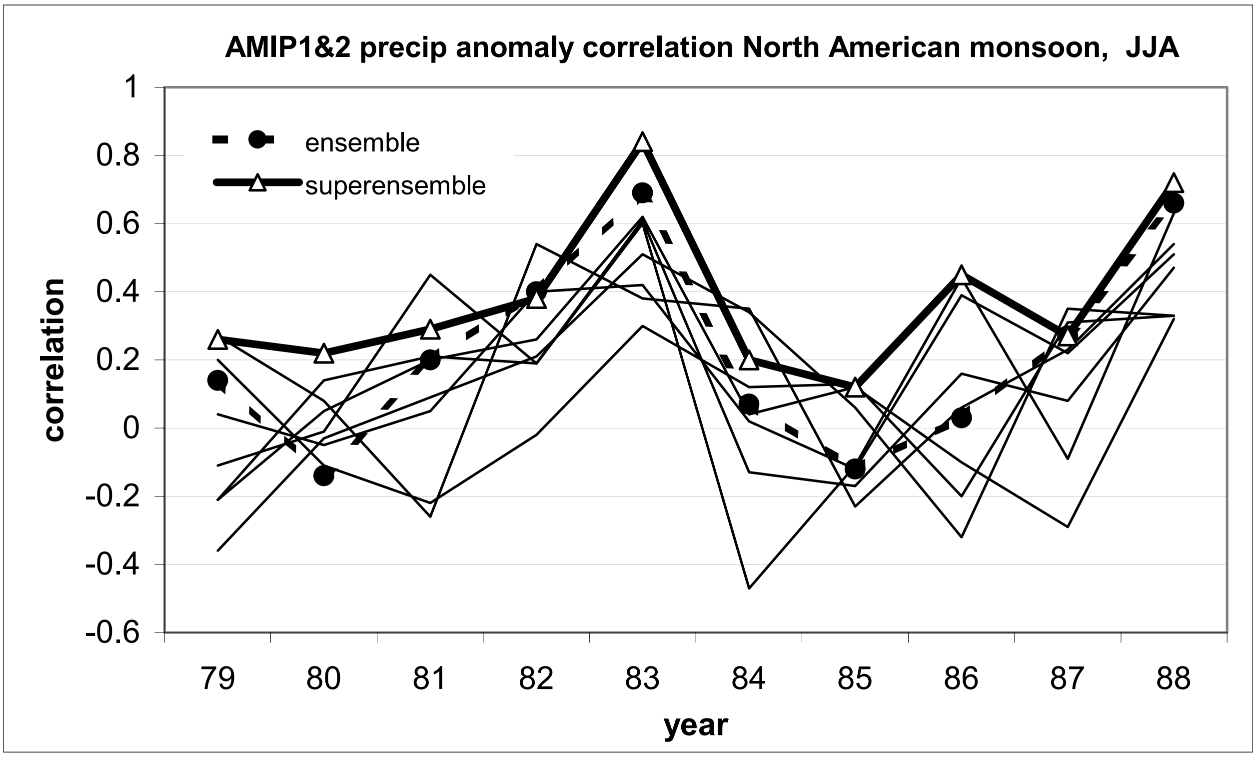


Figure 11 (b) – Same as 11 (a) but for North American Monsoon Domain.

5. Hurricane Forecasts from Multimodel Superensemble:

Real time hurricane forecast is another major component of the superensemble modeling at Florida State University. This approach of training followed by multimodel real time forecasts for tracks and intensity (up to 5 days) provides a superior forecast from the superensemble. Improvements in track forecasts are 25% to 35% better compared to the participating operational forecast models; this has been noted over the Atlantic Ocean basin. The intensity forecasts for hurricanes are only marginally better than those of the best models. Recent real-time tests, during 1999 to 2001, showed marked skills in the forecasts of difficult storms such as Floyd and Lenny where the performance of the superensemble was considerably better than those of the member models (Williford et al., 2002). An example of superensemble track forecasts for Hurricane Lenny is shown in Fig. 12. Here the observed best track for Hurricane Lenny is compared to those of a few member models and the superensemble. In these track forecasts, we note improvements for days 1, 2 and 3 of forecasts for the storm positions, which were of the order of 125, 200 and 350 km. This is an example showing a marked improvement for position forecasts. The illustrations in Fig. 13 show the forecast performance during the year 2001 for the Atlantic hurricane track and intensity. The least error for the superensemble in both categories is a consistent feature compared to all the participating models. This is an area where the use of multimodels (from diverse global modeling units and from FSU) has shown the most promise for forecasts on imminent landfall, tracks and intensity. A superensemble forecast constructed with a suite of some of the finest global models, that are currently available, holds a great promise for the improvements in shore range predictions for the landfall and tracks of hurricanes.

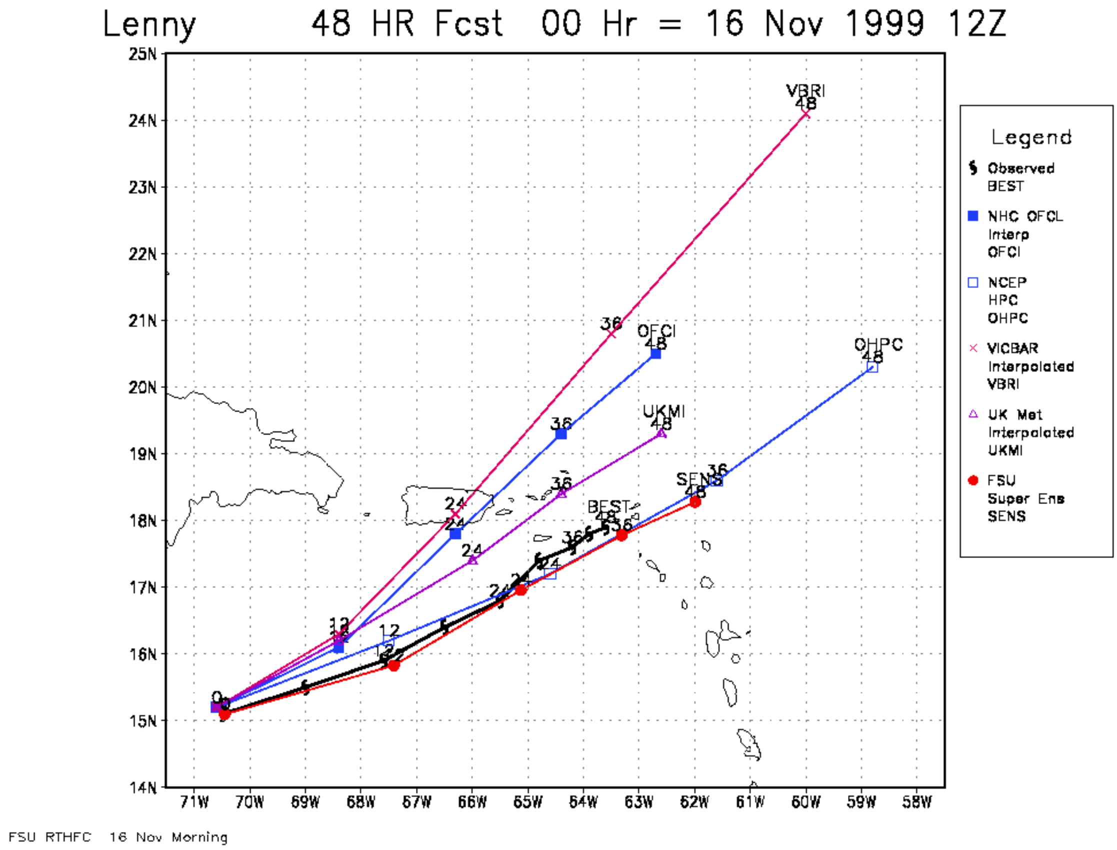
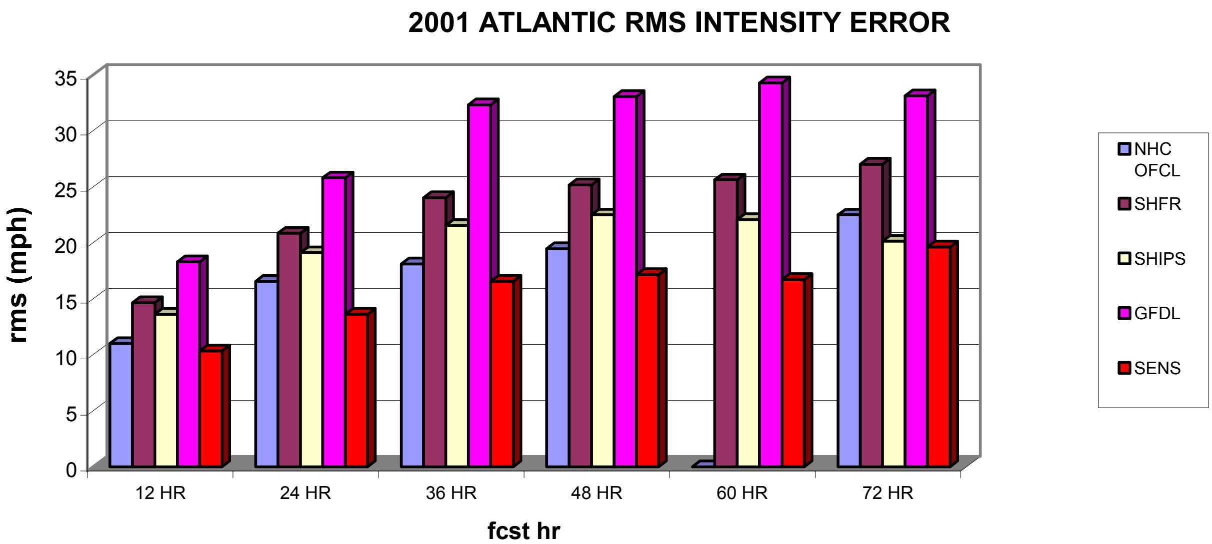


Figure 12. Superensemble track forecast of Lenny. Here the predicted tracks of some member models, ensemble mean and superensemble are shown.

Similar studies on superensemble based track and intensity forecasts for the Pacific region (Vijaya Kumar et al., 2002) have also revealed the usefulness of this methodology displaying considerable improvement of the forecast skills. A summary of the Pacific typhoon track and intensity errors for the years 1998-2000 is provided in Fig. 14 (a and b). Here the position and intensity errors at 12-hour intervals are shown where the skill from the superensemble is consistently high compared to the member models and the ensemble mean. In all of the aforementioned work, preservation of the member model features is an essential requirement during the training and the forecast phases. If drastic changes occur in the member models then the proposed statistical component of the superensemble is invalidated. It is apparent that if no major model changes are invoked during the training and the forecast phases of these forecasts, then we can obtain skill improvements of the order of 61, 138, 159 and 198 km for the typhoon position errors over the best models for forecasts at the end of days 1, 2, 3 and 4 respectively. The corresponding intensity forecast skills (rms errors) at days 1, 2, 3 and 4 of forecasts from the superensemble exceed those of the best models by 5, 10, 13 and 20 knots.



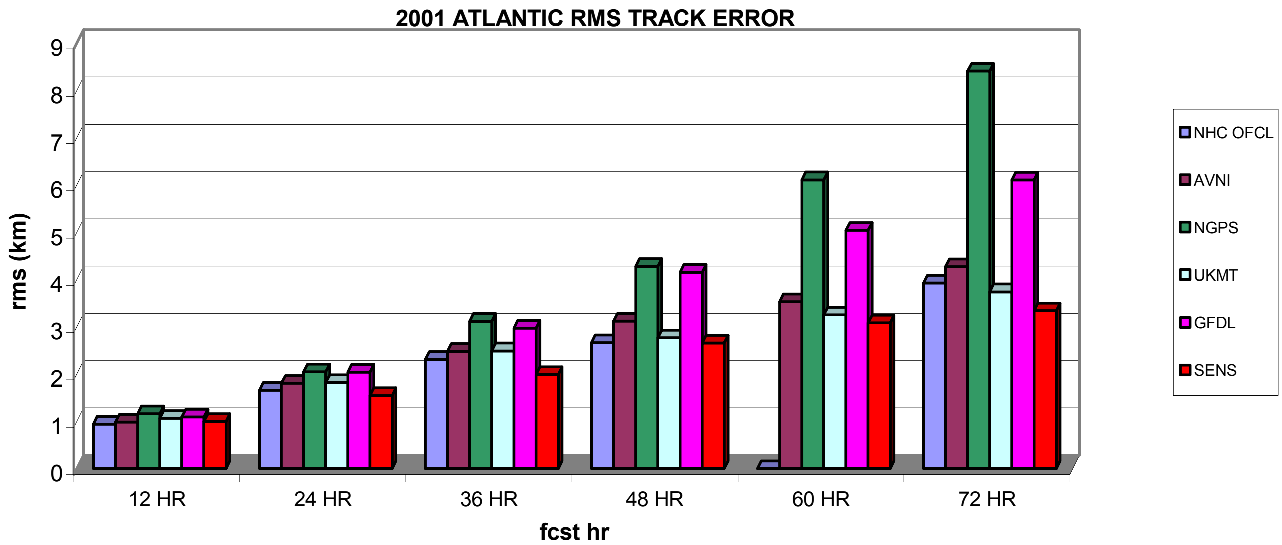
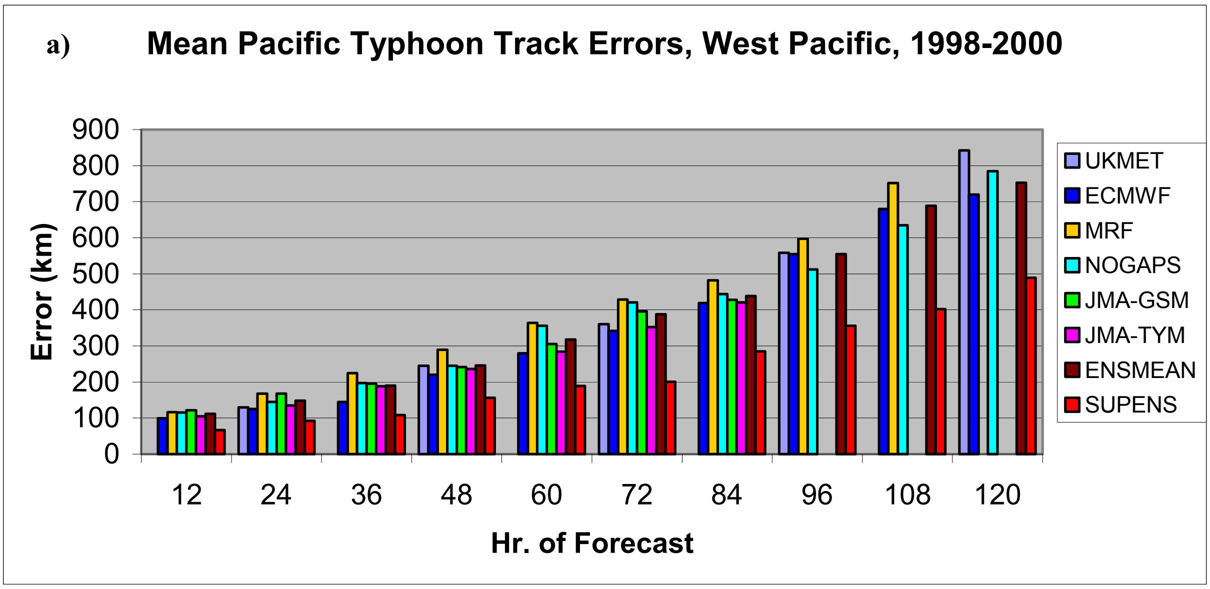


Figure 13. A seasonal summary on the performance of hurricane forecast skill during the year 2001 from various models including the FSU superensemble (SENS). Here the errors for the intensity (mph) and track (km) are displayed for 3-day forecasts.

It is desirable to look at the entire process of making the superensemble forecasts for hurricanes and typhoons. A typical walk-through example of superensemble forecast for Hurricane Lenny and another example of a land-falling typhoon Olga are illustrated in the following section.



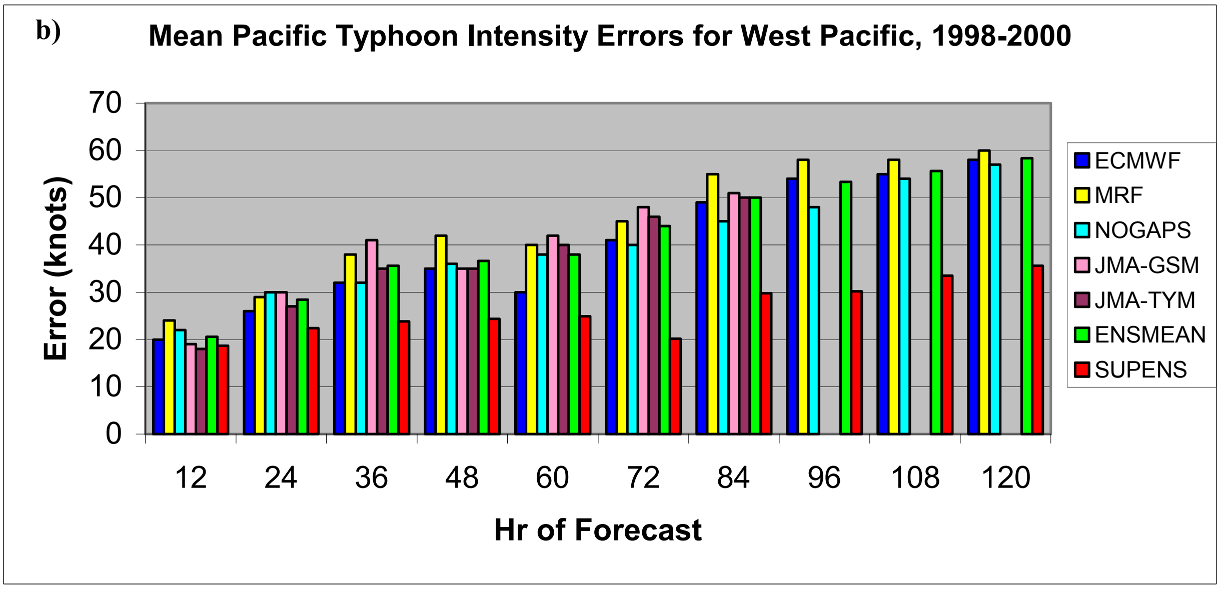


Figure 14: (a) Mean typhoon track errors for the West Pacific (km), 1998-2000. (b) Mean typhoon intensity errors for the West Pacific (km), 1998-2000.

5.1. Walking through a Pacific typhoon superensemble forecast

The proposed methodology described in section 1 is not very complex. It is simple to comprehend but why it works so well has perplexed many. Evidently it is not sufficient to simply state that the training coefficient, derived from a least square minimization principle, removes a bias. Because the classical bias removed defined by forecast mean minus an observational mean is not what this is doing. It is looking at all the member models at the same time and removing a collective bias. Since the least square minimization is a nonlinear computation this is not a simple addictive process. Because this method can easily place a hurricane position nearly a 100 km or more closer to the observed position compared to a best model it does have a huge practical relevance. For this reason we shall show an entire sequence of computation at a day 3 forecast for a typhoon (arbitrarily selected) from our forecast files. Table 5 shows the sequence of computations. This table carries the following entries: Name of participating model (i); superensemble coefficient; predicted latitude (F_i) of Typhoon Olga by the respective model on day 3 of forecasts; latitudinal mean ($\overline{F}$) of all forecasts made by the member model prior to Olga; the value of $a_i\left(F_i - \overline{F}\right)$; the value of $\overline{O} + \sum a_i\left(F_i - \overline{F}\right)$ where Ō is an observed mean that is defined as the observed mean of the latitudes of all storms prior to Olga during 1999. The models included here are the European model (ECMWF), U.S. model (MRF), U.S. Navy's model (NOGAPS), Japanese global model (JMAGSM) and the Japanese regional typhoon model (JMATYM). The other rows on this table include entries for the ensemble mean (EWS MEAN), the observed latitudinal position (OBS) of Typhoon Olga on day 3, the superensemble forecasts (SUPENS) obtained from the equation presented at the top of this table, the bias removed ensemble mean

Day-3 (72-hr) Multimodel Superensemble Forecast of Typhoon Olga (11W)
Valid 1999080212

Position (Latitude)

Model	Coef. a_i	Latitude F_i	Lat. Mean $\overline{F_i}$	$a_i\left(F_i - \overline{F_i}\right)$	$\overline{O} + \left(F_i - \overline{F_i}\right)$	Error (lat.) (in degrees)
ECMWF	0.16706	22.5	31.05	-1.42836	21.485	-3.9
MRF	-0.23966	21.5	29.44	1.9029	22.095	-4.9
NOGAPS	0.03840	22.5	26.68	-0.16052	25.855	-3.9
JMAGSM	-0.06988	22.7	27.71	0.350075	25.025	-3.7
JMATYM	0.08487	24.6	30.28	-0.48204	24.355	-1.8
ENSMEAN		22.76	$\overline{O}$			-3.64
OBS		26.4	25.703			
SUPENS		25.8851				-0.515
BIAS-REM ENSMEAN		23.763				-2.637

Position (Longitude)

Model	Coef. a_i	Longitude F_i	Lon. mean $\overline{F_i}$	$a_i\left(F_i - \overline{F_i}\right)$	$\overline{O} + \left(F_i - \overline{F_i}\right)$	Error (lon.) (in degrees)
ECMWF	0.14756	129.3	135.34	-0.89126	126.63	1.1
MRF	0.79791	129.4	133.71	-3.43899	128.36	1.2
NOGAPS	-0.42747	129.1	130.02	0.393273	131.75	0.9
JMAGSM	-0.88486	131.3	130.55	-0.66364	133.42	3.1
JMATYM	0.46562	129.1	131.24	-0.99643	130.53	0.9
ENSMEAN		129.64	$\overline{O}$			1.44
OBS		128.2	132.67			
SUPENS		127.0729				-1.127
BIAS-REM ENSMEAN		129.46				1.268

Table 5: Walking through one typhoon forecast from the Superensemble

latitudinal forecasts of Typhoon Olga on day 3. The lower half of the Table contains the compounding entries for the longitude forecasting of Typhoon Olga. There is a considerable spread of the regression coefficients for the latitudinal as well as the longitudinal position. These coefficients are not based on any data sets from Typhoon Olga; they are based on the past performance, prior to Olga, of the respective models. These coefficients range in values that are finite positive, finite negative to near zero. That is how the past behavior of the models placed them. What is extremely unique about the superensemble is that past behavior does carry over for the present storm Olga. The superensemble forecasts for both the latitude and longitude places this forecast position closest to the observed position. We clearly see an improvement over the best model by almost 250 km on the day 3 forecast of Olga. All the member model forecasts were moving the storm somewhat too slowly; the "collective bias removal" by the superensemble corrects this deficiency. The ensemble mean is not a good forecast compared to half of these member models. A bias removed ensemble mean is also displayed here that too is not as accurate as the best model. Bias removed ensemble mean assigns equal weights to all member models (after bias removal) that assure that a poor model is equivalent to the best model, which does not seem to hold true. When we look at Fig. 14 where we have displayed the summary of typhoon forecasts for the year 1999 we see the superior performance of the superensemble. That comes about in the same manner as is displayed here in Table 5. We have walked through each of the storm forecasts in this same manner and gotten a better feeling as to why the superensemble is a better product. As stated in Krishnamurti et al (1999) we note temporal resilience of weights (after a reasonable length of the training phase). It is this feature that conveys the high skill of the training phase to the forecast phase.

5.2. Walking through the superensemble forecasts of Hurricane Lenny of 1999

Unlike most Atlantic hurricanes that traverse from east to west, a late season hurricane Lenny of November 1999 encountered steering from westerly flows and moved in an opposite direction. Lenny was located at a rather low latitude around 15^{o}N where this unusual westerly steering was found. A suite of models operational at this time provided their real time forecasts to the National Hurricane Center at Miami. The superensemble tracks, on real time, were constructed using these forecasts based on the predicted increments of latitude and longitude positions at 12 hourly intervals. This was one of the several storms of the year 1999 where the superensemble forecasts were consistently much superior to those of the member models. A purpose of the following display exercise is to provide some insight to the reader on the mechanism of this simple superensemble procedure.

Table 6 (a, b and c) carries sequential entries on the forecasts of Hurricane Lenny's track for hours 36, 72 and 120 of forecasts. In these tables the first column identifies a model number, the second column shows the weights of the superensemble, based on increments, for these respective models from their training phase performance. The next column shows values of the superensemble algorithm function $\left(F_i' - \overline{F}_i'\right)$, i.e., the difference between the model forecast increment and its average from the training phase. The next column prepares the individual entries for the superensemble function $a_i\left(F_i' - \overline{F}_i'\right)$. The fifth column carries the total superensemble forecast for the position increment that now includes the observed mean increment $\overline{O}'$ from the training phase. The forecast latitudes or longitudes (adding all increments to the initial position of the storm) are shown in column 7 and the forecast errors are displayed in column 8. We also show in the rows below the member model forecasts the observed, superensemble forecast and the ensemble mean forecasts in column 7 for the different forecast hours. The various tables display a large similarity in the procedure. Although the member model's training weights do vary from one forecast hour to the next, the overall strength of the superensemble appears to hold steadily. These computations are revealing on the nature of the track forecasts displayed in Fig. 12. It should be noted that this is a one-dimensional superensemble along the tracks of the storms, unlike those of NWP where the grid points along latitude and longitude carry a two-dimensional display of superensemble skills. The number of models for the different forecast hours and even for latitude versus longitude is different. This has to do with the availability of data sets and on the criteria for rejection of training data forecasts where some poorer skills were encountered. It should also be noted that separate training coefficients were derived for each forecast hour. Thus we don't see a uniformity in the number of models in these tables.

6. Prospects for future research and applications:

Currently a number of regional mesoscale models (such as: regional spectral models of FSU, NCEP; ETA; various versions of MM5; ARPS, WRF; RAMS and others) are available that carry out real time forecasts over regional domains. In principle the superensemble methodology can be extended to this class of models and we expect much progress towards more accurate forecasts of severe weather events. The current multimodel superensemble methodology for hurricane tracks and intensity forecasts is also being extended to the tropical cyclone forecasts over the North Indian Ocean and the preliminary results look promising. In the seasonal climate forecasts issue, coupled model based regional models are being run for the North American Monsoon region and Indian Monsoon region using the FSUCGSM with an aim to provide reliable regional climate forecasts and the work is in progress. We are currently running a large number of seasonal forecasts mode with the global and high-resolution regional models at FSU. In order to construct the (8 member) superensemble for regional climate over the North American as well as Indian summer Monsoon regions, numbers of global and regional model runs are going on using FSU coupled and embedded regional spectral Model with different physics at these two monsoon regions. The basic aim of this research is to real time seasonal climate prediction over both the North American and Indian summer monsoon regions.

Acknowledgements: The research work on multimodel/multianlaysis superensemble forecasts for weather, seasonal climate and hurricanes/tropical cyclones is supported by the NOAA Grant numbers NA96GPO400, NA06GPO512, NA16GP1365, FSU Research Foundation Center of Excellence Award, NASA Grant numbers NAG5-9662, NAG8-1537 and NSF Grant numbers ATM-0108741, ATM-9910526. We acknowledge the data support from the European Centre for Medium Range Weather Forecasts especially through the help of Dr. Tony Hollingsworth.

7. References:

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

Krishnamurti, T.N., G.D. Rohaly and H.S. Bedi, 1994: On the improvement of precipitation forecast skill from physical initialization. *Tellus*, **46A**, 598-614.

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 multimodel superensemble. *Science*, 285 (5433), 1548-1550.

Krishnamurti, T.N., C.M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C.E. Williford, S. Gadgil and S. Surendran, 2000a. Multi-model superensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196-4216.

Krishnamurti, T.N., C.M. Kishtawal, D.W. Shin and C.E. Williford, 2000b: Improving Tropical Precipitation Forecasts from a Multianalysis Superensemble. *J. Climate*, **13**, 4217-4227.

Krishnamurti, T.N., S. Surendran, D.W. Shin, R. Correa-Torres, T.S.V. Vijaya Kumar, C.E. Williford, C. Kummerow, R.F. Adler, J. Simpson, R. Kakar, W. Olson and F.J. Turk, 2001. Real Time Multianalysis/Multimodel Superensemble Forecasts of Precipitation using TRMM and SSM/I Products. *Mon. Wea. Rev.,* **129**, 2861-2883.

Krishnamurti, T.N., L. Stefanova, Arun Chakraborty, T.S.V. Vijaya Kumar, Steve Cocke, David Bachiochi and Brian Mackey, 2002: Seasonal Forecasts of precipitation anomalies for North American and Asian Monsoons. Accepted for publication, *Journal of Met. Society of Japan*.

Krishnamurti, T.N., K. Rajendran, T.S.V. Vijaya Kumar, Stephen Lord, Zoltan Toth, Xiaolei Zou, Jon Ahlquist and I. Michael Navon, 2003: Improved Skills for the Anomaly Correlation of Geopotential Heights at 500 hPa. Accepted for publication, *Monthly Weather Review*.

Stefanova, L. and T.N. Krishnamurti, 2002. Interpretation of Seasonal Climate Forecasts Using Brier Skill Score, FSU Superensemble and the AMIP-1 Dataset. *J. Climate*, **15**, 537-544.

Treadon, R.E., 1996: Physical initialization in the NMC global data assimilation system. *Meteorol. Atmos. Phys.,* 60 (1-3), 57-86.

Vijaya Kumar, T.S.V., T.N. Krishnamurti, Michael Fiorino and M. Nagata, 2002: Multimodel Superensemble Forecasting of Tropical Cyclones in the Pacific. In press, *Monthly Weather Review*.

Williford, C.E., T.N. Krishnamurti, Ricardo J. Correa-Torres, Steven Cocke, Zaphiris Christidis and T.S.V. Vijaya Kumar, 2002: Real-Time Multimodel Superensemble Forecasts of Atlantic Tropical Systems of 1999. In press, *Monthly Weather Review*.

Yun, W.T., L. Stefanova and T.N. Krishnamurti, 2002: Improvement of seasonal and long-term forecasts using multimodel superensemble technique. Accepted for publication, *Journal of Climate*.