21.4 A NON-LINEAR FUZZY SET TECHNIQUE FOR COMBINING PRECIPITATION FORECASTS

Brian P. Mackey*
Department of Meteorology
Florida State University, Tallahassee, Florida

1 INTRODUCTION

Floods in the U.S. are among the top two most dangerous natural hazard in terms of frequency and severity. Annual damages from 1990-1999 have averaged $5.56 billion when adjusted to 1997 dollars (NOAA 2000). The ten-year average (1992–2001) number of flood fatalities in the U.S. is 88 per year, second only to heat-related deaths (NOAA 2002). Globally, approximately 1.5 billion people have been affected by floods from 1991–2000 (WMO 2001). Many countries are not prepared to handle the social and economic impacts arising from large flood disasters. However, given more accurate and reliable flood forecasts with enough lead time, governments may be able to act sooner to ameliorate the potential impacts of a flood event instead of just reacting to the situation.

In order to further improve heavy rainfall and flood forecasts, an interdisciplinary and multi-scale approach with state-of-the-art forecasting techniques is needed. In meteorology, as numerical weather prediction (NWP) continues to mature, modeling experiments are now being performed at many different space and time scales. Moreover, ensemble-based predictions have become the standard at many weather services around the world. Multi-model superensemble techniques (Krishnamurti *et al.* 2000a, 2000b, 2001) are also being shown to have great skill. Finally, mesoscale, or limited-area, models have become increasingly valued for predicting flood events. Still, successful quantitative precipitation forecasting (QPF) on the order of two to five days lead time is still very much a challenge, especially for flood events over small areas or over highly varying terrain.

This study, based on fuzzy set theory, aims to improve the accuracy of real-time, global-scale precipitation and flood forecasts from one to five days in advance. In this research, the author simulates a real-time prediction mode so as to demonstrate operational feasibility. This above method utilizes a non-linear classifier system that optimally weights member model precipitation forecasts to produce a more skillful end product. Using this technique, precipitation output can be collected along with other surface forcings from an NWP model and passed on to a spatially distributed hydrology model. Streamflow hydrographs for selected major flood cases can then be generated.

*Corresponding author address: Brian P. Mackey, Dept. of Meteorology, Florida State Univ., Tallahassee, FL 32306-4520; e-mail: mackey@met.fsu.edu.

2 SUPERENSEMBLE METHODOLOGY

In the classical superensemble technique, a benchmark analysis for any variable under study along with several model forecasts from weather services around the world are collected and arranged into a data set that varies in both space and time. Then, temporally, the data are split into two subsets: one for training purposes and the other for the actual forecast verification. Next, multiple linear regression is used on the training subset to determine a set of statistical weights for each model that minimizes the error between the combination of those models (the superensemble) and the observed state. Finally, these weights are applied to the independent forecast subset in order to verify the superensemble. The above process is repeated for each variable, grid point, and forecast time. For example, given discrete daily forecast data valid on days one through six, for seven variables and 65 160 grid points (a global T–126 Gaussian grid), there would be 2 736 720 different sets of coefficient weights and the same number of superensemble forecasts. Figure 1 shows this schematically for only one grid point, one variable, and one forecast time while Equations 1 and 2 describe this method mathematically:

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

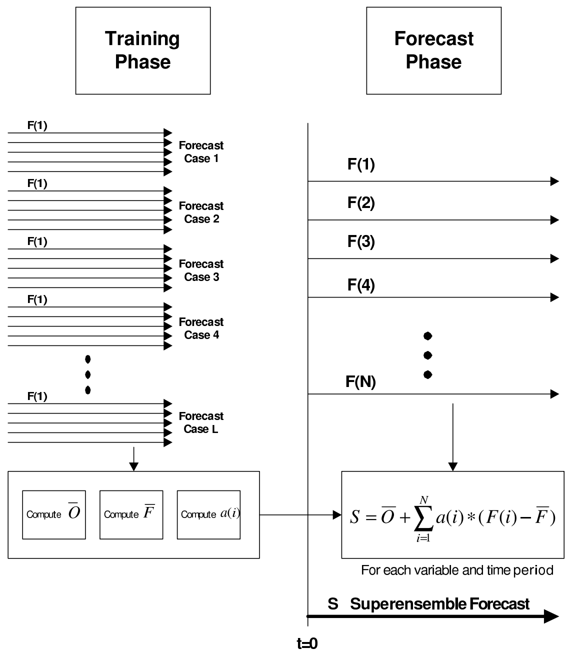

Figure 1: Superensemble forecast methodology

where S is the superensemble forecast, $\bar{O}$ the observed mean over the training period, N the number of members, a_i the ith model regression weight, $F_i(t)$ the ith model forecast, and $\bar{F}_i$ the ith model time mean over the training period. The weights a_i are determined via a least squares minimization of the cost function J:

$$J = \sum_{t=1}^{T} \left(S(t) - O(t)\right)^2 \tag{2}$$

where $O(t)$ is the observed state, and T the length of the training period, which is usually taken to be about 120 days for NWP applications, except precipitation, which uses about 65 training days. In the minimization process, the matrix equation $\mathbf{A} \cdot \vec{x} = \vec{b}$ is solved with the constraint that $\|\vec{b} - \mathbf{A} \cdot \vec{x}\|^2 \longrightarrow 0$. In this formulation, the $N \times N$ matrix $\mathbf{A}$ corresponds to the member model training data for one grid point, one variable, and one forecast time. The $N \times 1$ vector $\vec{b}$ holds the observations corresponding to the model forecast data. The unknown coefficients for each model are contained within the $N \times 1$ vector $\vec{x}$. The solution is $\vec{x} = \mathbf{A}^{-1} \cdot \vec{b}$, that is if $\mathbf{A}^{-1}$ exists (is non-singular).

In developing the superensemble, it is worth noting here that the most effective forecasts are those which are compatible with the training information. In other words, the model data utilized should be consistent for both the training and forecast data sets. Furthermore, if the characteristics of any member model at the code level changes during the training or forecast phases, this would have a negative impact as the regression weights would become rather ineffective.

3 FUZZY SET THEORY METHODOLOGY

The following is a non-linear extension to the superensemble methodology, and it is employed in this study to improve the prediction of global precipitation with special focus on storms that cause floods. This technique takes advantage of a distinct type of fuzzy system — the method of Takagi-Sugeno (1985) — to produce a combination forecast after Fiordaliso 1998 and Xiong *et al.* 2001). While the author does not yet know of any application of this method to NWP of rainfall, in the latter article, Xiong *et al.* utilize this method to forecast river streamflows over a two-year period. Results show that this method outperforms the individual members and is a viable choice for combination forecasting. As in the superensemble, this system is based on a post-processing algorithm that weights each member model forecast in order to produce a superior end product. Unlike the traditional superensemble, however, this non-linear approach uses fuzzy set theory to first classify precipitation regimes, then uses a sort of "smooth-switching" regression technique. Also, the model weights are now calculated at each time step.

Fuzzy set theory is useful for classifying properties of objects via descriptive terms such as "rather cold", "quite heavy", "fast", etc. Empirically, a membership function can be formulated with the result being any real number between zero and one as opposed to only either of those two extremes. This allows for a degree of uncertainty in portraying an object's properties and for a greater ability to classify such properties. In meteorology, for example, to classify rainfall intensity at a given station, one may define a rainfall rate threshold of 20 mm per day such that anything greater is described as "high" and anything less as "low". Now, if one station records a rainfall rate of 15 mm per day while another registers five mm per day, one can conclude that the first station exhibits a high rainfall intensity with some degree of uncertainty. For the second station, one is much less certain of a high rainfall intensity. Empirically, the first station has a membership value much closer to 1.0 than does the second station, where 1.0 represents absolute certainty that the rainfall intensity is high (i.e., greater than or equal to 20 mm per day).

In this research, p multi-model precipitation forecasts are collected at each grid point and for each discrete forecast time i, as shown in the time series $\hat{P}_{1,i}\ , \hat{P}_{2,i}\ , ..., \hat{P}_{p,i}$. A vector can be constructed to hold the input variable data, as in $\hat{P}_i = [\hat{P}_{1,i}\ , \hat{P}_{2,i}\ , ..., \hat{P}_{p,i}]^T$. Given these inputs and an objective univariate clustering algorithm, one can divide each time series at each grid point into any number of partitions (fuzzy sets) that describe the rainfall intensity. In order to keep the number of parameters at a minimum, only two fuzzy sets are created here, one for low and one for high precipitation intensity. Moreover, an arithmetic mean value, $\boldsymbol{\mu}_r$, is computed for both domains. This mean value also serves as the group representative for each domain. This step is termed fuzzification. It is worth noting here that some of the rainfall values can belong to more than one domain, and this is the case in this research.

The next step is called logic decision, which involves specification of if-then inference rules and the computation of a parameter called the "degree of applicability" for each if-then rule. The partitioned input domains (see step one) form an underlying premise for the construction of the if-then rules. Specifying these rules is important because they describe how the process under study is being controlled (Xiong *et al.* 2001). In essence, the question being posed is, "In what ways do the input variables control the output?" The if-then rules are generally descriptive in nature, and they take the following form:

$$\begin{aligned} R_r : IF\ (x_1\ is\ A_r^{(1)},\ x_2\ is\ A_r^{(2)}, ..., x_p\ is\ A_r^{(p)}) \\ THEN\ y_r = f_r(x_1, x_2, ..., x_p), \end{aligned} \tag{3}$$

where R_r is the rth if-then control rule ($r = 1, 2, ..., k$ for k rules), $x_1, ..., x_p$ are the input variables (the multi-model precipitation forecasts), p is the number of input variables (input forecasts), $A_r^{(1)}, ..., A_r^{(p)}$ are the descriptors linked to the fuzzy sets $m_r^{(1)}, ..., m_r^{(p)}$, $f_r(\cdot)$ is the output function corresponding to the rth if-then rule whose result is y_r. This function can be expressed as:

$$\begin{aligned} y_r = f_r(x_1, x_2, ..., x_p) = b_r(0) + b_r(1)x_1 + ... + \\ b_r(p)x_p = b_r(0) + \sum_{j=1}^{p} b_r(j)x_j, \end{aligned} \tag{4}$$

where the coefficients $b_r(j)$ are the unknown parameters to be estimated for each rule.

In this study of multi-model precipitation inputs split into two domains (low and high precipitation), it is possible to have 2^p unique if-then inference rules. To keep parameters at a minimum, however, it is best to limit the number of if-then rules. This, in turn, will lead to better combination forecasts during the forecast phase (Fiordaliso 1998). Hence, only two rules are selected here with the major assumption being that the multi-model inputs are similar in nature to the observations according to the descriptors high and low. In other words, the order of magnitude of model forecasts versus observations is assumed to be the same. Hence, the if-then rules are given as:

$$\begin{aligned} R_1 : IF\ (\hat{P}_{j,i}\ is\ \text{“}low\ precipitation\text{”},\ j = 1, ..., p) \\ THEN\ \hat{P}_{c,i} = b_1(0) + \sum_{j=1}^{p} b_1(j) \cdot \hat{P}_{j,i}\ , \end{aligned} \tag{5}$$

$$\begin{aligned} R_2 : IF\ (\hat{P}_{j,i}\ is\ \text{“}high\ precipitation\text{”},\ j = 1, ..., p) \\ THEN\ \hat{P}_{c,i} = b_2(0) + \sum_{j=1}^{p} b_2(j) \cdot \hat{P}_{j,i}\ . \end{aligned} \tag{6}$$

Rule 1 states that if all model forecasts give low precipitation, thencombine them accordingly. Rule 2 outlines a similar combination algorithm when all models have high precipitation.

Another piece in this formulation is the computation of the "degree of applicability," given by α. It is specified rather arbitrarily depending on the application. For this research, it is taken as the radial basis function (RBF) in Gaussian form after Fiordaliso (1998):

$$\alpha_r(\hat{P}_i) = exp(-\|\hat{P}_i - \mu_r\|_2^2) \tag{7}$$

This parameter serves two purposes. First, it describes how effectively the different inputs $x_1, ..., x_p$ satisfy a given if-then rule. The more an input conforms to a given rule r, the closer α_r is to unity. Second, this parameter also serves as a weighting factor for the final scalar output $\hat{P}_{c,i}$, as seen in the following:

$$\hat{P}_{c,i}\ = \frac{\sum_{r=1}^{k} \alpha_r y_r}{\sum_{r=1}^{k} \alpha_r} = \frac{\sum_{r=1}^{k} \alpha_r f_r(x_1, x_2, ..., x_p)}{\sum_{r=1}^{k} \alpha_r}. \tag{8}$$

Thus, the α_r weights are applied to the "intermediate" output y_r that was already calculated from each of the r if-then rules. As seen in Equation 8, the final desired output $\hat{P}_{c,i}$ is just a weighted average of all the "intermediate" outputs y_r. The unknowns in the system are contained within the y_r term.

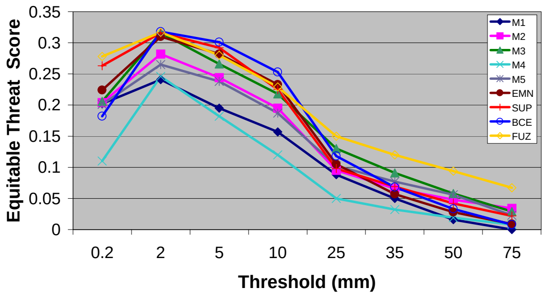

Figure 2: Day 1 equitable threat score (ETS) for 8 thresholds averaged over 15 Aug to 14 Sep, 2003 and from 40°N to 40°S. M1 through M5 denote the 5 multi-models, EMN the regular ensemble mean, SUP the classic superensemble, BCE the bias-corrected ensemble mean, and FUZ the combination forecast based on the fuzzy set method.

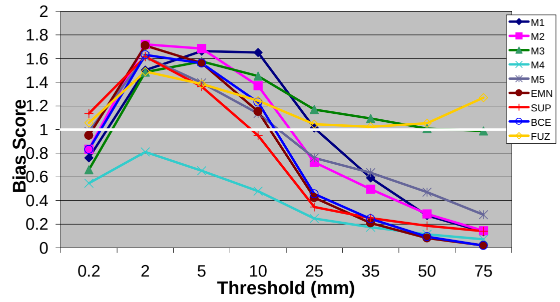

Figure 3: Day 1 bias score for 8 thresholds averaged over 15 Aug to 14 Sep, 2003 and from 40°N to 40°S. M1 through M5 denote the 5 multi-models, EMN the regular ensemble mean, SUP the classic superensemble, BCE the bias-corrected ensemble mean, and FUZ the combination forecast based on the fuzzy set method.

Ultimately, by substituting Equation 4 into 8, it is demonstrated that the input variables are linearly combined to produce the desired end product. The right-hand side can be simplified into $w_{0,i} + w_{1,i} \cdot \hat{P}_{1,i} + w_{2,i} \cdot \hat{P}_{2,i} + \cdots + w_{p,i} \cdot \hat{P}_{p,i}$, where

$$w_{0,i} = \frac{\sum_{r=1}^{k} \alpha_r(\hat{P}_i) \cdot b_r(0)}{\sum_{r=1}^{k} \alpha_r(\hat{P}_i)} , \tag{9}$$

$$w_{j,i} = \frac{\sum_{r=1}^{k} \alpha_r(\hat{P}_i) \cdot b_r(j)}{\sum_{r=1}^{k} \alpha_r(\hat{P}_i)} , \; j = 1, 2, ..., p . \tag{10}$$

The $w_{0,i}, w_{1,i}, ..., w_{p,i}$ weights vary with each time step i, while the unknown parameters $b_r(j)$, $j = 0, 1, ..., p$ for each rule are estimated through a minimization of a quadratic error cost function similar to Equation 2. In this case, however, a multi-dimensional minimization technique, such as the downhill simplex method by Nelder and Mead, must be used if the number of if-then rules is non-trivial (i.e., greater than one). In all, there are $(p+1) \times k$ unknown coefficients, where p is the number of multi-models (five, for this study) and k is the number of if-then rules (two).

4 DATA INPUTS

The benchmark observed analysis used in verifying the precipitation forecasts consists of both TRMM (Tropical Rainfall Measuring Mission) and SSM/I (Special Sensor Microwave Imager) data. The TRMM Microwave Imager (TMI) 2A-12 rainfall algorithm (Kummerow *et al.* 1996 and 2000) provides daily rain rate information within 35° latitude of the equator. Through the DMSP (Defense Meteorology Satellite Program) satellites, the NOAA/NESDIS (National Oceanic and Atmospheric Administration/National Environmental Satellite, Data, and Information Service) SSM/I algorithm (Ferraro and Marks 1995) covers the remainder of the globe so that satisfactory coverage of precipitation areas is attained within 55° latitude of the equator on a daily basis. The data resolution can be as high as 40 km, but for the purposes of this research it is interpolated to a Gaussian grid resolution of T–126, or about 104 km. For regional-scale studies over land, rain gauge data may also be utilized when available to assist in verifying the precipitation forecasts.

The multi-model five-day global forecasts of precipitation from five different cooperating weather services are initialized at 1200 UTC each day. Those centers and models involved in this col-

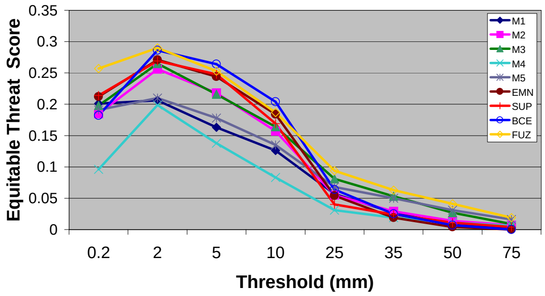

Figure 4: Average day 3 ETS for 15 Aug to 14 Sep, 2003.

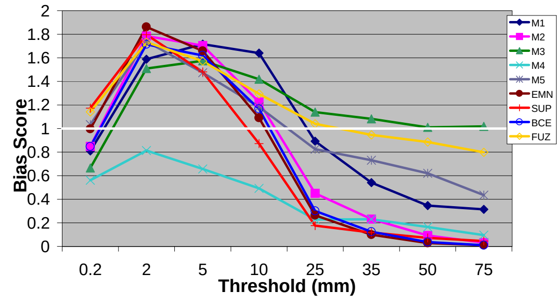

Figure 5: Average day 3 bias score for 15 Aug to 14 Sep, 2003.

laboration are as follows: (1) Bureau of Meteorology Research Center (BMRC) Atmospheric Model (BAM), Australia; (2) Japan Meteorological Agency (JMA) Global Spectral Model (GSM); (3) National Centers for Environmental Prediction (NCEP) Aviation Model (AVN); (4) Naval Research Laboratory (NRL) Navy Operational Global Atmospheric Prediction System (NOGAPS); and (5) Recherche en Prévision Numérique (RPN) Global Environmental Multiscale (GEM) Model, Canada. The reader is directed to the model documentation from those centers for specific model constructs. The output resolutions of the models range from 0.8° to 2.5° latitude/longitude, but for consistency, they are interpolated to T–126. The final superensemble or fuzzy system precipitation forecasts are therefore calculated using the above inputs with the output also at T–126.

5 RESULTS

Using the superensemble and fuzzy set methodologies described above, five-day precipitation forecasts are generated to test the predictability of global rain areas (QPF) at eight different threshold intensities. Several flood case studies are also investigated. All combination forecasts are produced as if in a real-time sense. In other words, observations are used only when available by time zero. For each forecast valid time (days one, two, three, four, and five), 65 training days are used in the conventional superensemble while 35 are used for the fuzzy method, except for the day one forecast (where ten training days are used). These optimal numbers are arrived at upon performing many independent forecast tests not described here.

Two different time periods in 2003 are selected for QPF verification purposes: the full month of February and a 31-day period from 15 August to 14 September. Spatially, all verification statistics are given for a latitude belt of 40° either side of the equator. This includes the global tropics and a good portion of the U.S.

Two skill scores that are used extensively at NCEP to verify QPF are also chosen in this study. They are the equitable threat score (ETS) and the bias score. The ETS is given as

$$ETS = \frac{H - H_{random}}{H + M + FA - H_{random}}, \qquad (11)$$

where

$$H_{random} = \frac{(H + M)(H + FA)}{N}. \qquad (12)$$

In these formulae, H corresponds to a hit, M denotes a miss, FA is the number of false alarms, and N is the total number of forecast points. By using the ETS, a forecast is rewarded for predicting precipitation amounts at least equal to the observed values for a given threshold. However, a prediction is penalized for forecasting precipitation in the wrong place as well as not forecasting it in the right place for that same threshold. In addition, there is an adjustment for hits associated with random chance. The ETS may vary from -0.333 to +1, where zero indicates no skill and one indicates a perfect score.

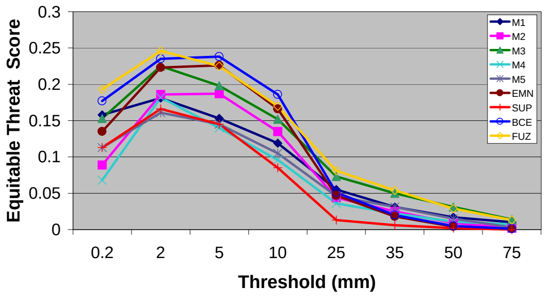

Figure 6: Average day 5 ETS for Feb, 2003.

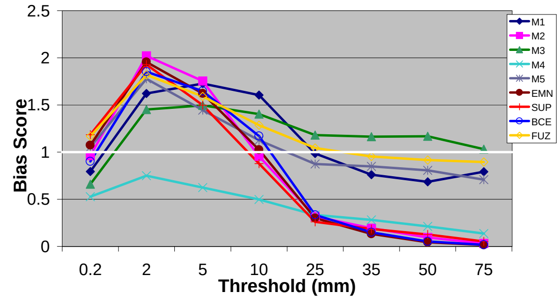

Figure 7: Average day 5 bias score for Feb, 2003.

The bias score, given in Equation 13, only compares areas of predicted and observed rainfall (usually above a given threshold) and does not indicate any degree of accuracy. A score of one indicates no bias. Precipitation amounts are said to be underforecast (overforecast) for a bias score of less (greater) than one.

$$BIAS = \frac{H + FA}{H + M}, \tag{13}$$

Figure 2 shows the average day one ETS over the 80° wide latitude belt and for the time period 15 August to 14 September, 2003. In this figure, the forecasts based on the fuzzy set theory are overall superior to the multi-models (M1 through M5) and the other ensemble techniques (the regular ensemble mean, superensemble, and bias-corrected ensemble mean). The BCE forecast does show slightly greater skill at the five and ten mm per day thresholds, however.

The average day one bias score for the same dates is shown in Figure 3. Here, the fuzzy set theory proves its worth again. Even though there is an overestimation of precipitation for the low to medium amounts, the bias scores for 0.2, 25, 35, and 50 mm per day thresholds are close to the no-bias forecast of one. Hence, this non-linear technique has its greatest impact in removing the very low bias at the higher thresholds. This would seem to indicate that heavy rain amounts would be more realistically forecast over the global tropics compared to the other models, with the exception of model M3. The ensemble mean (EMN) may be suffering from too much smoothing while the classical superensemble, in its attempt to reduce "average error", is sometimes too conservative in its forecast of heavy precipitation.

The day three skill charts, illustrated in Figures 4 and 5 and valid over the same latitude belt, show many of the same features. Yet, for the higher thresholds, the relative ETS skill advantage of the FUZ predictions is lessened compared to day one. Finally, the day five skill charts for the month of February are shown in Figures 6 and 7. These are very similar to the day five August/September skills. The FUZ forecasts are once again, on average, equal to or superior than the remaining predictions for all but the five and ten mm per day thresholds. There is still much room for improvement, however, when considering the perfect forecast ETS is unity.

On the other hand, the FUZ forecasting system does show much higher ETSs for individual flood events over smaller space and time scales. These results are not shown here, but they and other, more in-depth skill score charts will be posted on the following FSU real-time NWP website in the coming months: http://lexxy.met.fsu.edu/rtnwp.

6 SUMMARY AND FUTURE WORK

The preceding results show that a post-processing multi-model combination algorithm based on fuzzy set theory is a viable choice for ensemble forecasting. At the very lowest 0.2 mm per day threshold and also the at the higher rainfall thresholds is where the greatest improvement in skill is noticed.

That combined with the near no-bias forecasts at the 25 mm per day level and up gives an overall superior forecast on average over the global tropics. This is especially true on day one of the forecast. Further results broken down by region and season are being prepared for a more in-depth analysis, as well as flood event case studies.

Furthermore, with assistance and guidance from Dr. Ana Barros, hydrology simulations are taking place with a state-of-the-art distributed numerical model employed to predict basin-scale streamflows (river levels). By using the rainfall/atmospheric data from the both the fuzzy set theory forecast and from the Florida State University Global Spectral Model (FSUGSM), the goal is to more accurately predict basin-scale river flooding on time scales of two to five days in advance. Preliminary results for the Limpopo River Basin over southeastern Africa appear quite promising.

REFERENCES

Ferraro, R. R., and G. F. Marks, 1995: The development of SSM/I rain-rate retrieval algorithms using ground-based radar measurements. *J. Atmos. Oceanic Technol.*, **12**, 755–770.

Fiordaliso, A., 1998: A nonlinear forecasts combination method based on Takagi-Sugeno fuzzy systems. *International Journal of Forecasting*, **14**, 367–379.

Krishnamurti, T. N., C. M. Kishtawal, D. W. Shin and C. E. Williford, 2000a: Multimodel superensemble forecasts for weather and seasonal climate. *J. Climate*, **13**, 4196–4216.

Krishnamurti, T. N., C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil and S. Surendran, 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.

Kummerow, C., W. S. Olson, and L. Giglio, 1996: A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors. *IEEE Trans. on Geosci. and Remote Sensing*, **34**, 1213–1232.

Kummerow, C., and Coauthors, 2000: The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteor.*, **39**, 1965–1982.

NOAA, cited 2000: Compilation of Flood Loss Statistics. [Available online at http://www.nws.noaa.gov/oh/hic/flood_stats/Flood_loss_time_series.htm]

NOAA, cited 2002: Natural Hazard Statistics. [Available online at http://www.nws.noaa.gov/om/hazstats.shtml]

Takagi, T., and M. Sugeno, 1985: Fuzzy identification of systems and its application to modelling and control. *IEEE Trans. Syst. Man. Cybern.*, **15**, 116–132.

WMO, cited 2001: Floods e-brochure. [Available online at http://www.wmo.ch/web/ home/brochure-e/floods.html]

Xiong, L., A. Y. Shamseldin, and K. M. O'Connor, 2001: A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system. *J. Hydrology*, **245**, 196–217.

ACKNOWLEDGEMENTS

The author would like to thank Dr. T.N. Krishnamurti for support and guidance, as well as Dr. Ana Barros. The author would also like to thank all the participating NWP centers around the world for their real-time model forecast data over the past four years and continuing. This research work reported here was supported by NASA TRMM Grant NAG5-9662, NSF Grant ATM-0108741, FSU Research Foundation Cornerstone Award and FSU Research Foundation Award 1338-831-45.