Dynamic Model Mixing for Enhancing the Predictability of Hydroclimatic Variables

Chowdhury, Shahadat[1], Ashish Sharma[2]

[1] School of Civil and Environmental Engineering, University of New South Wales, Australia
Email: shahadat@civeng.unsw.edu.au

Keywords: *Dynamic Combination, Sea Surface Temperature, Hierarchical Tree*

EXTENDED ABSTRACT

The option of combining multiple model responses to reduce predictive uncertainty has been investigated in the field of economic forecasting and more lately in climate prediction. The current combination method in hydroclimatology is mainly limited to mixing weights that remain static over time. Recently (Chowdhury and Sharma, 2006) and (Chowdhury et al., 2007) have proposed using time variant mixing weights as an improvement over the static weight combination. The method involves a dynamic hierarchical pair wise combination tree for outputs from multiple predictive models. The model pairs are first matched based on the sample error covariance. Then the pairs are combined by ascertaining a target weight for each time step. This process provides a low dimensional setting for investigating any predictive structure of the relative model strengths.

The weights are predicted using a mixed distribution which is a product of a 'precision ratio' and a 'bias direction'. The precision ratio is the fraction of the squared residual error associated with each of the paired models, and the bias direction represents an indicator of the sign of the two residual errors. The precision ratio is projected forward using a generalised linear autoregressive model and the bias direction is projected by ordered logistic regression.

The method is extended here to combine three climate models, the variables of interest being the monthly global sea surface temperature anomalies at 5° grids for 1956 to 2001. This work is multivariate extension of our earlier univariate (NINO3.4) application (Chowdhury and Sharma, 2006) . The prediction from static weight combination is used as base case for comparison. The predicted sea surface temperature using this dynamic combination algorithm consistently exhibited better accuracy to that of static combination in every season (Figure 1). Improved skill is achieved at 86% of the global grids with the rest showing indifference to static weight skills.

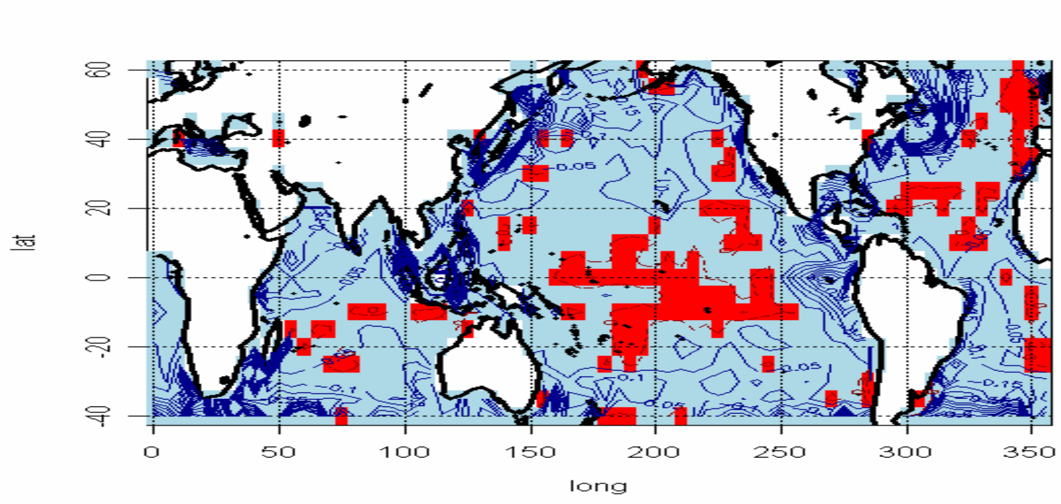

Figure 1. The reduction in prediction error variance due to dynamic weight combination compared to that of static weight combination method.. The lighter shades (blue) and solid contours are the zones with improved prediction. The darker boxes (red) and broken contours where there is no difference. No zones show higher error variance.

1. INTRODUCTION

There are a number of dynamic and stochastic models that predicts the hydroclimatic variables. Each model is subject to certain strengths and limitations. The selection of a single *best* model involves selection uncertainty as well as discards any superior strength of alternative options in certain climatic period. Combining response of various models maximise the strength of individual approach and reduce the variance of predictive uncertainty (Barnston et al., 2003; Colman and Davey, 2003; Greene et al., 2005; Peng et al., 2002; Raftery et al., 2005; Robertson et al., 2004; Sharma and Lall, 2004). Combining models is a well established time series forecasting technique (Clemen, 1989). The method used in this research is referred as the pair wise dynamic model combination (Chowdhury et al., 2007). The term dynamic being used to denote the fact that the mode of combination varies with time, with combination weights being modelled on the basis of the persistence they exhibit; the term pair wise reflects the paired hierarchical tree architecture when the numbers of components models are more than two. The aim is to improve upon the existing method of combining model predictions that overlooks persistence in the individual model skills. The existing developments that set the background of our current research are discussed in next paragraph.

The static mixing weight reflects the accuracy of individual models or maximise the performance of the weighted combination output (Coelho et al., 2004; Doblas-Reyes et al., 2005; Kondrashov et al., 2005; Pavan and Doblas-Reyes, 2000; See and Abrahart, 2001; Xiong et al., 2001). The weighted average combination forms the benchmark against which we compared the performance of our proposed pair wise dynamic combination approach. The proposed dynamic weight method is first applied to univariate predictions, NINO3.4 time series by Chowdhury and Sharma (2006). This 2006 study has been further strengthened by projecting the weights as a mixed distribution of beta binomial conditional on logistic regression and with enhanced hierarchical tree structure (Chowdhury et al., 2007). This paper presents the current state of research to extend the method into multivariate prediction of global sea surface temperature anomalies.

2. THEORY

2.1. Target Weight

Consider a case of combining predictions of a pair comprised of i^{th} and j^{th} component models using dynamic weights. The component predictions at time t are $\hat{u}_{i,t}$ and $\hat{u}_{j,t}$ with residual error of $e_{i,t}$ and $e_{j,t}$, the corresponding true response is y_t where,

$$y_t = \hat{u}_{i,t} + e_{i,t} \quad (1)$$

$$y_t = \hat{u}_{j,t} + e_{j,t} \quad (2)$$

The two models can be combined as follows:

$$y_t = \hat{u}_{i,t}\,\omega_t + \hat{u}_{j,t}\,(1 - \omega_t) + \ddot{e}_t \quad (3)$$

Here $\ddot{e}_t$ is the residual of the combination where the target weight ω_t is available. We assume that the component predictions are unbiased and hence restrict the weights within 0 to 1. One way of estimating the relative skills (and hence ω_t) of the models is by comparing the precision (defined here as the inverse of error variance) of the component predictions (Granger and Newbold, 1977; McLeod et al., 1987). In a two model case, the *precision ratio*, r_t can be estimated as:

$$r_t = e_{j,t}^{\ 2} / (e_{j,t}^{\ 2} + e_{i,t}^{\ 2}) \quad (4)$$

In order to keep $\ddot{e}_t^2 \le Min(e_{i,t}^{\ 2},\ e_{j,t}^{\ 2})$ we propose additional criteria which are based on the direction of the bias of each model. The *bias direction* $\{b_t\ ;\ t=1,2,..t_{\max}\}$ is mapped into three categories $\{mix,\ zero,\ one\}$ as shown in Equation (5):

$$b_t = \begin{cases} mix & 0 > e_{j,t} / e_{i,t} \\ zero & 0 < e_{j,t} / e_{i,t} < 1 \\ one & 1 < e_{j,t} / e_{i,t} \end{cases} \quad (5)$$

The models are combined based on r_t only when $e_{i,t}$ and $e_{j,t}$ have opposing sign i.e. two predictions are bracketing the true value. On the other hand while both predictions exhibit bias in the same direction the better prediction is chosen ignoring r_t. The optimum measure of ω_t is estimated as follows:

$$\omega_t = \begin{cases} r_t & \text{when } b_t = mix \\ 0 & b_t = zero \\ 1 & b_t = one \end{cases} \quad (6)$$

2.2. Forecasting Target Weight

The forecast of weights $\{\omega_t\}$ is done in a two step process. The first step involves predicting $\{r_t\}$, the precision ratio model and the second step is predicting the bias direction $\{b_t \in mix, zero, one\}$.

The precision ratio $\hat{r}_t$ is forecasted using a finite order generalised linear autoregressive (GLAR) model (Shephard, 1995):

$$Logit(\hat{r}_t) = \theta_t + \varphi\, r_{t-.} \quad (7)$$

Where, $r_{t-.}$ is the predictor vector (autoregressive and exogeneous) and θ_t is the periodic intercept. The regression parameter φ is estimated by maximising the likelihood of beta binomial distribution of the response variable.

The second step of this method predicts the direction of bias. An ordered logistic regression (OLR) model (Agresti, 1996) is used as the basis for predicting the categorical bias direction b=$\{b_t\ ; t=1,2,..t_{\max}\}$. The cumulative probability of b, P(b) is estimated as:

$$Logit\,[P(b_t = mix)] = \alpha_1 + x_t\, \beta \quad (8)$$

$$Logit\,[P(b_t = mix \text{ or } zero)] = \alpha_2 + x_t\, \beta.$$

Where x_t are predictor vectors and $\{\alpha_1, \alpha_2, \beta\}$ are intercepts and slope parameters. No third equation is necessary since $P(b = one) = 1 - P(b = mix$ or $zero)$.

2.3. Multivariate Combination

The combination exercise of Equation (3) can be extended to multivariate response, $Y_t = \{y_{t,i}\}$, where i may denote spatial spread or a set of different types of response variable. The example of multivariate response may be global sea surface temperature or a set of indices like {NINO3.4, wind stress, thermo-cline etc}. The multivariate statistic is maintained by designing the predictors $\{r_{t-.}, x_t\}$ that retains the effects of the neighbours.

$$r_{t-.} = \sum_i c_i \bullet r_{t-.,\,i} \quad (9)$$

Where c_i is a measure of influence (eg. correlation coefficient) of the neighbours. Similar relationship of Equation (9) is developed to estimate x_t, which is not shown here.

2.4. Combining Multiple Models

The last three sections presented the basis for combination of two models. In case of higher number of models a paired combination hierarchical tree is used as shown for a four model case in Figure 2. The model pairing is performed by first sorting the models in order of their individual residual variance, and then starting from the lowest variance model and finding its pair as the model with which it has the lowest covariance. This process is repeated for the models that remain until all models are exhausted.

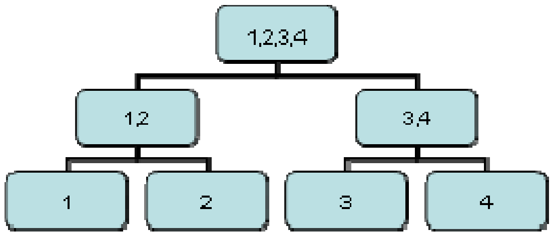

Figure 2. The hierarchical tree of four component models.

3. APPLICATION

The method is applied to improve the prediction of global sea surface temperature anomalies 3 months in advance. Three model responses are combined. The first of the three models was developed at University of California, Los Angles, USA, hereafter referred as UCLA model (Kondrashov et al., 2005). The second of the three models was developed at the Climate Prediction Centre of the National Oceanic and Atmospheric Administration, USA and referred to as the CACPC model (Dool et al., 2003). The third model was prepared by the Demeter project of European Centre for Meteorological Forecast and referred to as the ECMF model (ECMWF, 2004). The hind-casts during the period of January 1956 to December 2001 at 5° by 5° grids of the global sea surface between 60°N to 40°S are used to estimate the dynamic weights. The error variance of the {UCLA, CACPC, ECMF} are {0.70, 1.89, 1.48}. The models UCLA and CACPC are paired first and then ECMF is paired at a higher level of the hierarchical tree.

The first step of this combination method involves calculating the target weights. Figure 3 shows a sample of target weights of UCLA versus CACPC pair, averaged over 1956 to 2001 time series at all grid points. The contour of the target precision weight indicates the spatial variability of the

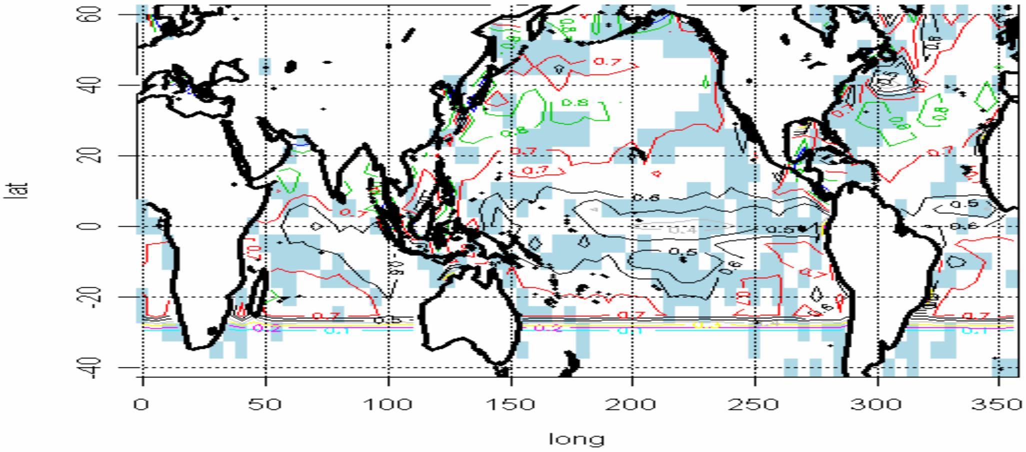

Figure 3. The contour of target precision ratio, the blue shaded area are showing zones where b=*mix*, the white zones are locations where target weights are either 0 or 1. This graph is showing average values across the entire time series

relative model skill. The next step of the combination method forecasts these target weights based on suitable predictors, as detailed in Equations (7) and (8).

3.1. Predictor Selection

The predictors for the precision ratio (r_{t-}) model in Equation (7) are ascertained from lagged values of the response (r_t) over the past 12 time steps (months). Predictors for the categorical bias direction (x_t) of Equation (8) are selected from lagged values of the ratio $e_{i,t}/e_{j,t}$. which is constrained within [-1, 2] avoiding numerical instability when $e_{j,t} \rightarrow 0$. The periodic intercepts are 12 monthly values smoothed over 3 months span. A common set of autoregressive lags are used for the entire sea surface for simplicity. The selected autoregressive lags are mainly in the order of 6, 9 and 12 months.

The spatial characteristics are maintained by adjusting the predictors reflecting the neighbourhood influence. An exhaustive cross correlation analysis of predictors at each grid points against all other grid locations showed that the influence existed only within ±20° distances; Figure 4 displays one such analysis of UCLA and CPC model combination weights at a location. Each predictor vector is smoothed by weighted linear combination of neighbouring predictors within ±20° as shown in Equation (9). The weights $\{c_i\}$ being the correlation coefficient in this study.

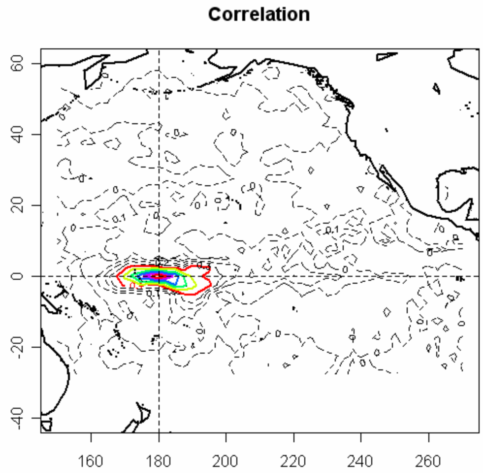

Figure 4. Spatial correlation of weights to 0°N, 180°E grid point. Correlations ≥ 0.4 are drawn in thicker line, and lower values in broken line.

3.2. Static Weight Alternative

The strength of the dynamic weight combination is compared to that of static weight combination. The static weight combination here largely followed the methodology used by Robertson (2004) that can be divided into three steps. Each component model prediction at each grid point is first combined against the climatology prediction. The static weight is estimated by minimising the sum of squared errors of the combined prediction. The weights of each component models are normalised at second step. The above two steps are repeated for all the grid points. At third step, the spatial noise is reduced by smoothing weights at each grid points across the neighbours within ±20° distance.

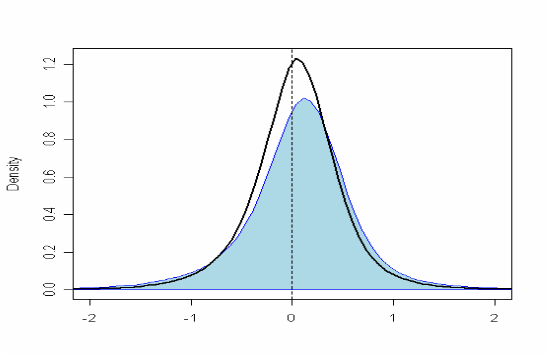

Figure 5.. Density graph of mean of squared error of prediction, the filled graph is the static weight method, the black outer graph is the dynamic weight results.

4. RESULTS AND DISCUSSION

The prediction obtained by this dynamic weight is compared against the best component model, UCLA and the static weight prediction. Mean squared error (MSE) is used as a measure of skill in this study. The density graph of MSE at all the grid point for the entire time series is drawn in Figure 5. The peaky density plot of the dynamic weight demonstrates the smaller error variance (0.298) and with reduced bias (peak is closer to zero vertical) compared to that of static weight method (0.455). Besides, the month by month comparison of MSE demonstrates, in Figure 6, the superior performance of the dynamic weight predictions compared to static weight predictions or best single model predictions. We have also checked the spatial dependence structure of the prediction to that of observed dependence by computing the correlation of each grid point to all other grid points. Very few systematic loss (or gain of) spatial correlation were evident compared to observed values as shown in Figure 7.

The MSE of predictions by static weight method minus that of dynamic weight method are drawn across the global sea surface grid in Figure 1. The cells with positive difference denote improvement and are shown using lighter shades (blue). The figure demonstrates an improvement of the predictions at 86% of the grid boxes across the globe. The decrease of sum of squared error is found statistically significant when analysed by one tailed paired *t test* (p= 2e-16). There are 14% locations where either static or dynamic weights yields similar prediction skill. Note that in absence of any persistence of relative model skills the dynamic weight converges to static weight estimate. In our exercise, the dynamic weight did not worsen the skills of prediction in any location compared to static weights

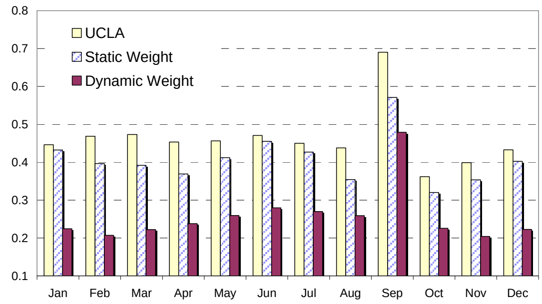

Figure 6. The monthly spread of mean of squared error of prediction. Note the consistence improvement around all seasons.

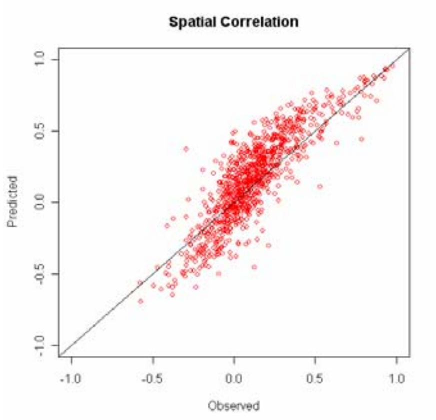

Figure 7. Spatial correlation of SSTA at a grid point to the rest of the sea surface grids.

5. CONCLUSION

This study enhanced a dynamic combination of multiple predictive models to predict multivariate response, such as sea surface temperature anomaly. The pair-wise mixing weight is designed as a product of a 'precision ratio' and a 'bias direction'. The precision ratio is projected forward using a generalised autoregressive model and the bias direction is projected using ordered logistic regression. Three or more models are combined based on a pair wise hierarchical tree. The tree provides a low dimension setting for designing any predictive structure of the relative model strengths. The method is extended to multivariate responses by simply including the neighbours through a linear combination. The method is applied to combine three climate models, the variables of interest being the monthly global sea surface temperature anomalies at 5° grids for a period of 1956 to 2001. The prediction from static weight combination is used as base case for comparison. The predicted sea surface temperature using this dynamic combination algorithm consistently exhibited better accuracy to that of static combination. The spatial statistics of the temperature anomalies are maintained.

6. ACKNOWLEDGEMENT

This research was funded by the Australian Research Council and the Sydney Catchment Authority. We acknowledge the assistance of Upmanu Lall, Andrew Robertson and Lisa Goddard of the Lamont Doherty Earth Observatory Centre of Columbia University, New York. The computation was performed using the freely available R statistical computing platform (R Development Core Team, 2006)

7. REFERENCE:

Agresti, A., 1996. An Introduction to Categorical Data Analysis. Wiley Series in Probability and Statistics, New York.

Barnston, A.G., Mason, S.J., Goddard, L., Dewitt, D.G. and Zebiak, S.E., 2003. Multimodel Ensembling in Seasonal Climate Forecasting at IRI. Bulletin of the American Meteorological Society, 84(12): 1783-1796.

Chowdhury, S. and Sharma, A., 2006. Combining Climate Prediction Models Using Dynamic Weights, 30th Hydrology and Water Resources Symposium, Launceston, TAS.

Chowdhury, S., Sharma, A. and Lall, U., 2007. Sea Surface Temperature Predictions Using a Pair wise Dynamic Model Combination Approach. Journal of Climate, under review.

Clemen, R.T., 1989. Combining forecasts: A review and annotated bibliography. International Journal of Forecasting, 5(4): 559-583.

Coelho, C.A.S., Pezzulli, S., Balmaseda, M., Doblas-Reyes, F.J. and Stephenson, D.B., 2004. Forecast Calibration and Combination: A Simple Bayesian Approach for ENSO. Journal of Climate, 17: 1504-1516.

Colman, A.W. and Davey, M.K., 2003. Statistical prediction of global sea-surface temperature anomalies. International Journal of Climatology, 23(14): 1677-1697.

Doblas-Reyes, F.J., Hagedorn, R. and Palmer, T.N., 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting - II. Calibration and combination. Tellus A, 57(3): 234-252.

Dool, H.v.d., Huang, J. and Fan, Y., 2003. Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981–2001. Journal of Geophysical Research, 108(D16): 8617.

ECMWF, 2004. Development of a European Multi-Model Ensemble System for Seasonal to Inter-Annual Prediction (DEMETER). 434, England.

Granger, C.W.J. and Newbold, P., 1977. Forecasting economic time series. Academic Press, New York.

Greene, A.M., Goddard, L. and Lall, U., 2005. Probabilistic multimodel climate change scenarios for the evolution of regional temperatures, The International Research Institute for Climate and Society, The Earth Institute at Columbia University, New York.

Kondrashov, D., Kravtsov, S., Robertson, A.W. and Ghil, M., 2005. A Hierarchy of Data-Based ENSO Models. Journal of Climate, 18(21): 4425-4444.

McLeod, A.I., Noakes, D.J., Hipel, K.W. and Thompstone, R.M., 1987. Combining hydrologic forecast. Journal of Water Resources Planning Management, 113(1): 29-41.

Pavan, V. and Doblas-Reyes, F.J., 2000. Multi-model seasonal hindcasts over the Euro-Atlantic: skill scores and dynamic features. Climate Dynamics, 16(8): 611-625.

Peng, P., Kumar, A. and Dool, H.v.d., 2002. An analysis of multimodel ensemble predictions for seasonal climate anomalies. Journal of Geophysical Research, 107(D23): 4710.

R Development Core Team, 2006. R: A language and environment for statistical computing. R Foundation for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Raftery, A.E., Geneiting, T., Balabdaoui, F. and Polakowski, M., 2005. Using Bayesian

Model Averaging to Calibrate Forecast Ensembles. Monthly Weather Review, 133: 1155-1174.

Robertson, A.W., Lall, U., Zebiak, S.E. and Goddard, L., 2004. Improved Combination of Multiple Atmospheric GCM Ensembles for Seasonal Prediction. Monthly Weather Review, 132(12): 2732-2744.

See, L. and Abrahart, R.J., 2001. Multi-model data fusion for hydrological forecasting. Computers & Geosciences, 27(2001): 987-994.

Sharma, A. and Lall, U., 2004. Model averaging and its use in probabilistic forecasting of hydrologic variables. In: B. Webb (Editor), Hydrology, Science and practice for the 21st century. British Hydrological Society, Imperial College, London, pp. 372-378.

Shephard, N., 1995. Generalized Linear Autoregressions, Economics Working Paper 8. Nuffield College, Oxford.

Xiong, L., Shamseldin, A.Y. and O'Connor, K.M., 2001. A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system. Journal of Hydrology, 245(1-4): 196-217.