# CWRF Optimized Physics Ensemble Improving U.S. 1993 Flood Prediction

Ligang Chen[1*] and Xin-Zhong Liang[1,2]

[1]Division of Illinois State Water Survey, Institute of Natural Resource Sustainability

[2]Department of Atmospheric Sciences

University of Illinois at Urbana-Champaign

## 1. Introduction

No existing model can fully represent the observed climate system. Each model, or each physical configuration of a model, may capture some certain climate signals whereas others don't. Thus, consensus of multiple models or multiple physical configurations of a model have recently been highlighted due to their superior skill over those using a single model or configuration (Rajagopalan et al. 2002). Our previous studies have shown that for precipitation simulation, significant skill improvement can be achieved using an optimal ensemble of multiple cumulus schemes, including regime dependence for activation and relative contribution from the participating schemes (Liang et al. 2007; Liu et al. 2009).

We have finalized the Climate extension of the Weather Research and Forecasting model (CWRF) for the initial release scheduled in this summer (Liang et al. 2010a). This CWRF incorporates over $10^{24}$ combinations of alternative schemes representing each of and interactions among the major physical processes of cloud, aerosol, radiation, surface, planetary boundary layer (PBL), cumulus, and microphysics. It provides an unprecedented tool to explore, analyze, and ultimately solve the ensemble optimization problem for weather and climate prediction.

## 2. Brief introduction to the latest CWRF

The CWRF has been developed on the basis of the Weather Research and Forecasting model (WRF, Skamarock et al. 2008) by incorporating numerous improvements that are crucial to climate scales, including interactions between land–atmosphere–ocean, convection–microphysics and cloud–aerosol–radiation, and system consistency throughout all process modules (Liang et al. 2005a-b; Yuan and Liang 2010). The CWRF improvements have been accomplished through iterative, extensive model refinements, sensitivity experiments, and rigorous evaluations over the past 8 years. As a result, the CWRF has demonstrated greater capability and better performance (with its designated physics configuration) in simulating the U.S. regional climate than our existing CMM5 (Liang et al. 2004, 2007; Zhu and Liang 2005, 2007) and the original WRF. This justifies its initial release for the community use, as scheduled in this year after the first paper on a general model description and basic skill evaluation (Liang et al. 2010a). Detailed description of the model and its evaluation will come along as a series of academic papers. This system provides an ideal platform for developing procedures for optimal ensemble prediction of climate variations.

[*] *Corresponding author address:* Ligang Chen, Illinois State Water Survey, University of Illinois at Urbana-Champaign, 2204 Griffith Drive, Champaign, IL 61820. E-mail: lgchen@illinois.edu

## 3. Optimization Experiment of Multiple Physics Configuration over the U.S.

### 3.1 Selection of the participating physical scheme members for optimization

We first evaluate the CWRF results with selected physics configurations by examining their overall spatial frequency distributions of correlation coefficients (CC) and root mean square error (RMSE) of daily mean precipitation variations as compared with concurrent observations. By comparison, we

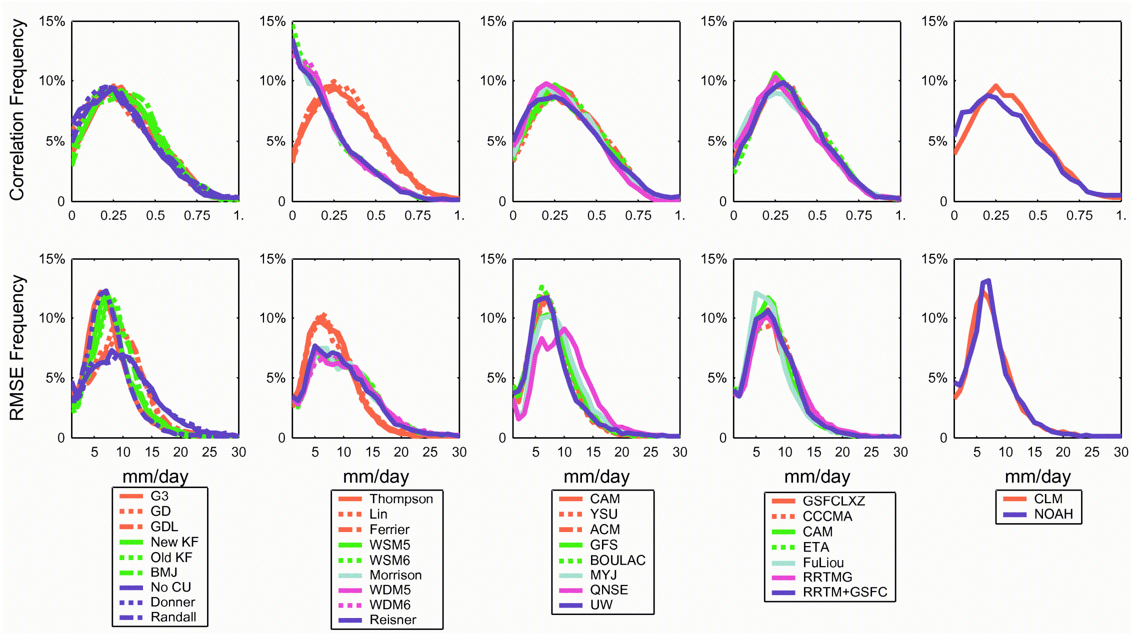


***Figure 1.*** *Point-wise Correlation and RMSE frequencies results of precipitation for 1993 MJJ over U.S. land area from our incorporated physical schemes grouped as five physical processes (left to right column): Cumulus, Microphysics, PBL, Radiation and Land Surface.*

remove those schemes with extremely poor CC and RMSE scores, such as GD and GDL cumulus schemes, Reisner and WDM6 microphysics schemes in microphysics, and QNSE PBL scheme (Figure 1). For the remaining schemes in each group of the five key physical processes (radiation, surface, PBL, cumulus, and microphysics), we compare the overall spatial frequency distributions of CC and RMSE across the schemes (rather than with observations). That comparison allows us to further eliminate those having great similarity (high CC and small RMSE) while choosing one representative scheme. After all the representative members have been selected, we then apply various optimization methods to seek the best prediction of regional climate variations. We have found a rather simple method based on a linear combination of the CC and RMSE scores to define the weights and produce an ensemble prediction of precipitation that has superior skills to individual members as well as the average ensemble using an equal weight. The results are summarized below.

### 3.2 Optimization of Multiple Physics Configurations

The crucial aspect of the ensemble prediction is how to construct the optimal weights for individual members. So far these weights have been optimized to minimize local RMSE or maximize CC against observations. Recently, we have tested three ensemble forecast methods using two member sizes:

1. AVE: the ensemble forecast is a simple average of all members, assuming an equal weight.

2. OPT: the ensemble forecast is produced by applying the optimized weights that are derived by minimizing the RMSE for a previous training period. Here, the optimization solution follows Liang et al. (2007) while the training and subsequent forecast periods are one month each but non-overlapping.

3. LCR: the ensemble forecast is produced by applying the weights as a linear combination of the CC and RMSE scores for individual members from a

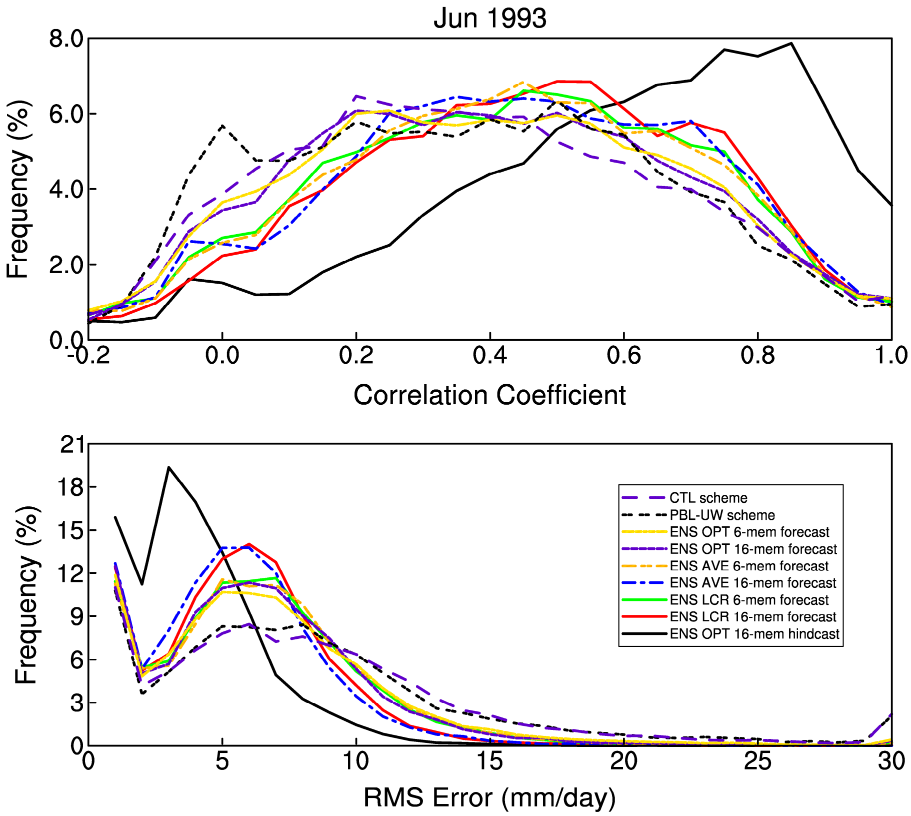


***Figure 2.*** *Spatial frequency distribution of pointwise correlation coefficient and root mean square error (RMSE) of daily mean precipitation as compared with observations during 1993 June over the U.S. land grids simulated by the CWRF using two best schemes (CTL, UW-PBL), and three ensemble forecast methods (OPT, AVE, LCR) using two number sizes (6, 16), as well as the hindcast ensemble optimization.*

previous training period. For member $i$ with $c_i = \mathrm{CC}$ and $r_i = \mathrm{RMSE}^{-1}$ scores, its weight can be defined as: $w_i = \alpha \cdot r_i / \sum_{i=1}^{N} r_i + (1-\alpha) \cdot c_i / \sum_{i=1}^{N} c_i$ . Here $N$ is the total number of members in the ensemble, and $\alpha$ is a relaxation coefficient between 0.0 and 1.0, presently chosen as 0.5. As such, the LCR procedure assigns bigger weights to those members with higher CC and smaller RMSE values.

## 4. Experimental Results

Figures 2-3 illustrate the spatial frequency distributions of pointwise CC and RMSE of daily mean precipitation as compared with observations during 1993 June and July over the U.S. land grids simulated by the CWRF using two best schemes (CTL, UW-PBL), and the three ensemble forecast methods (OPT, AVE, LCR) using two number sizes (6, 16), as well as the hindcast ensemble optimization. The hindcast optimization is the same as OPT except that the training and forecast are made on an identical month, and thus represent the upper limit of the ensemble prediction skill we seek to approach. The 6 members are a subset of the 16-member ensemble, including the CTL plus five representative schemes (best of each physics group) as selected earlier. Our initial conclusions from these results are as follows:

1. The ensemble forecast result, irrespective of the method defining the weights, is always better than those of individual members;

2. The ensemble forecast result is better when more members are included;

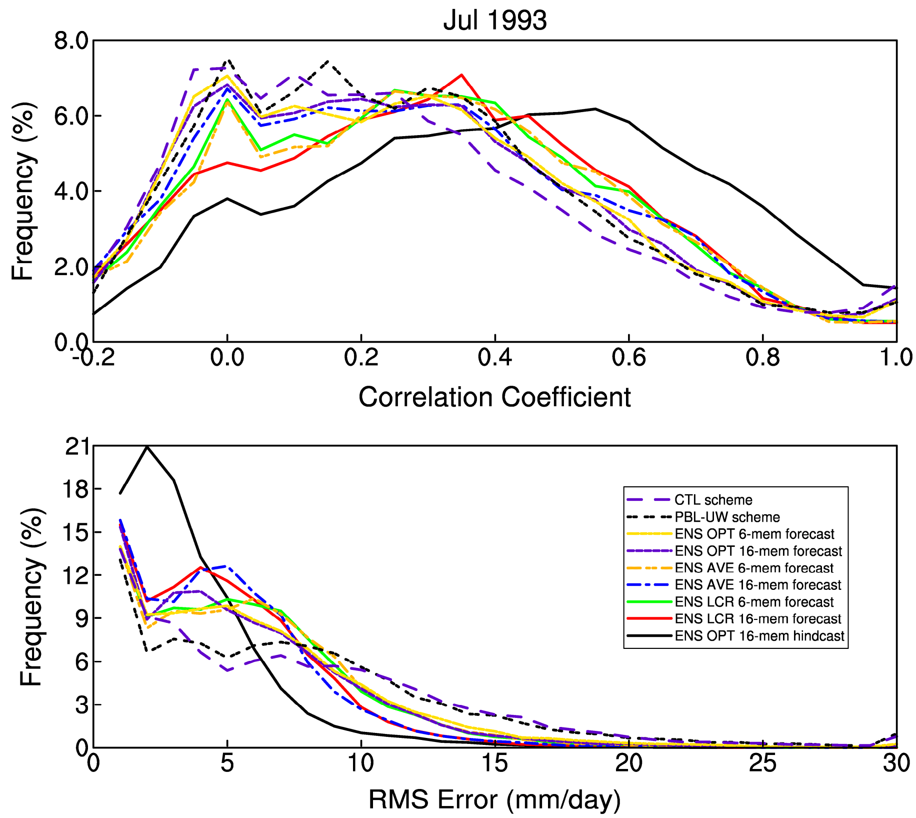


***Figure 3.*** *Spatial frequency distribution of pointwise correlation coefficient and root mean square error (RMSE) of daily mean precipitation as compared with observations during 1993 July over the U.S. land grids simulated by the CWRF using two best schemes (CTL, UW-PBL), and three ensemble forecast methods (OPT, AVE, LCR) using two number sizes (6, 16), as well as the hindcast ensemble optimization.*

3. The ensemble forecast based on the LCR method defining the weights is more realistic than the AVG method, while the OPT method is the worst; and 4. The ensemble forecasts based on three methods (AVG, OPT, LCR) have substantial gaps from the hindcast ensemble optimization, and thus a large room for further improvement.

We strongly believe that the OPT method can be improved significantly by defining a more robust objective function for optimization. This is our next primary focus.

## 5. Conclusions and Future Work

The CWRF model has incorporated and improved a bunch of existed major schemes for each physical process of cloud, aerosol, radiation, surface, planetary boundary layer (PBL), cumulus, and microphysics. They can form over $10^{24}$ combinations, and thus produce same amount of simulation results. This huge data source provides an ideal basis for analyzing and solving the ensemble optimization problem for weather and climate prediction. The ensemble forecast results always demonstrate superior skills over those individual members, and a large room for further improvement exists. Since both observed and simulated precipitation, especially their differences, contain significant small-scale features, their direct use will result in unwanted noises to the optimization solution. We will seek advanced mathematical or statistical tools sort out signals at appropriate scales, and develop a more suitable objective function for minimization to obtain robust weights.

## Acknowledgements

This work was partially supported by the NOAA Climate Prediction Program for the Americas (CPPA) NA08OAR4310575 and NA08OAR4310875, and the DOE Office of Biological and Environmental Research (BER) DE-SC0001683. The model simulations were conducted at the DOE/NERSC, NOAA/ESRL, and UIUC/NCSA facilities.


## References


Liang, X.-Z., L. Li, K.E. Kunkel, M. Ting, and J.X.L. Wang, 2004: Regional climate model simulation of U.S. precipitation during 1982-2002. Part 1: Annual cycle. *J. Climate*, **17**, 3510-3528.

Liang, X.-Z., H. Choi, K.E. Kunkel, Y. Dai, E. Joseph, J.X.L. Wang, and P. Kumar, 2005a: Surface boundary conditions for mesoscale regional climate models. *Earth Interactions*, **9**, 1-28.

Liang, X.-Z., M. Xu, W. Gao, K.E. Kunkel, J. Slusser, Y. Dai, Q. Min, P.R. Houser, M. Rodell, C.B. Schaaf, and F. Gao, 2005b: Development of land surface albedo parameterization bases on Moderate Resolution Imaging Spectroradiometer (MODIS) data. *J. Geophys. Res*., **110**, D11107, doi:10.1029/2004JD005579.

Liang, X.-Z., M. Xu, K. E. Kunkel, G. A. Grell, and J. Kain, 2007: Regional climate model simulation of U.S.-Mexico summer precipitation using the optimal ensemble of two cumulus parameterizations. *J. Climate*, **20**, 5201-5207.

Liang, X.-Z., M. Xu, X. Yuan, T. Ling, H.I. Choi, F. Zhang, L. Chen, S. Liu, S. Su, J.X.L. Wang, K.E. Kunkel, W. Gao, E. Joseph, V. Morris, T.-W. Yu, J. Dudhia, and J. Michalakes, 2010a: Development of CWRF for regional weather and climate prediction: General model description and basic skill evaluation. *J. Climate* (to be submitted).

Liu, S., W. Gao, M. Xu, X. Wang, and X.-Z. Liang, 2009: Regional climate model simulation of China summer precipitation using an optimal ensemble of cumulus parameterization schemes. *Frontiers of Earth Science in China*, **3**(2), 248-257, DOI 10.1007/s11707-009-0022-8.

Rajagopalan, B., U. Lall, and S.E. Zebiak, 2002: Categorical climate forecasts through regularization and optimal combination of multiple GCM ensembles. Mon. Wea. Rev., 130, 1792-1811.

Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, M.G. Duda, X.-Y. Huang, W. Wang, and J.G. Powers, 2008: *A Description of the Advanced Research WRF Version 3*. NCAR Technical Note, NCAR/TN-475+STR, 113 pp.

Yuan, X., and X.-Z. Liang, 2010: Evaluation of a Conjunctive Surface-Subsurface Process model (CCSP) over the United States. *J. Hydrometeor.* (submitted).

Zhu, J., and X.-Z. Liang, 2005: Regional climate model simulation of U.S. soil temperature and moisture during 1982-2002. *J. Geophys. Res*., **110**, D24110, doi:10.1029/2005JD006472.

Zhu, J., and X.-Z. Liang, 2007: Regional climate model simulation of U.S. precipitation and surface air temperature during 1982-2002: Interannual variation. *J. Climate*, **20**, 218-232.