PROGRESS REPORT FOR PROJECT NUMBER NA04NWS4620012
Covers the Period 06/01/2007 - 11/30/2007

Submitted December 17, 2007

TITLE: **Parameterization and Parameter Estimation of Distributed Models For Flash Flood and River Prediction**

PI: **Hoshin V. Gupta[1], Professor**
Co-PI: **Thorsten Wagener[2], Assistant Professor**
[1] Department of Hydrology and Water Resources, The University of Arizona, Tucson, Arizona, 85721, Tel: 520-626-6974
[2] Now at Department of Civil and Environmental Engineering, Penn State Univ., University Park, PA 16802, Tel: 814-865-5673

---

1. PROJECT OBJECTIVES AND ACCOMPLISHMENTS

1.1. Project Objectives

The shift from lumped to distributed models raises many important questions about the proper choice of model parameterization, including the desirable level of model complexity, while significantly increasing the complexity of the parameter estimation problem. The main objective of this project is to collaborate with and support the Hydrologic Modeling team at the NWS Office of Hydrology in the rapid development of an advanced version of the NWS-OH distributed hydrologic model, with particular attention to the issues of parameter estimation, appropriate model structure, supportable model complexity, and model evaluation, diagnosis and improvement. The unifying theme through this proposal is to focus on improving distributed watershed modeling through addressing issues of model parameterization (specification of model components), and estimation of the model parameters in both gauged and ungauged settings. The following tasks were listed under this contract:

1. Parameterization of semi-distributed and distributed hydrologic models within Hydrology Laboratory-Research Modeling System (HL-RMS) framework,
2. Distributed parameter estimation (automated and/or semi-automated) for the above (this work will build on our experience with lumped models, while introducing novel ideas such as regularization that are specifically tailored to distributed models),
3. A priori methods for parameter estimation in ungauged basins using direct inference from watershed properties and statistical regression analysis (existing work by NWS-HL staff will be extended and used to drive this important area of hydrologic modeling research forward).

This work extends our past collaborative work with the NWS by supporting the development of distributed modeling capabilities with particular attention given to ungauged and poorly gauged watersheds, consistent with the aims and future directions of the NWS. This research is being implemented in the context of the HL-RMS thereby maximizing technology transfer and ensuring that the work outcome is of direct value to the NWS.

1.2. Summary of Work Proposed

a. Implement HL-RMS at the University of Arizona as a modeling environment. Incorporate currently available calibration algorithms.
b. Investigate and implement a distributed parameter estimation algorithm based on the concept of regularization
c. Investigate the a priori parameter estimation problem using both bottom-up (incl. the development of relationships between the parameters of the NWS conceptual model components to soil and watershed characteristics) and top-down (regionalization) approaches.
d. Testing of various modifications of the HL-RMS including the comparison to semi-arid specific models.
e. Combining the research on a priori and distributed parameter estimation into a single procedure.
f. Testing the basic equations relating model parameters and watershed properties in a multi-watershed study. Complementing this approach with a statistical regionalisation approach using a minimum of 30 watersheds.
g. Implementing and testing various ensemble-forecasting schemes.
h. Technology transfer through (in person and telephone) meetings.
i. Implementing and testing the new tools for a priori and distributed parameter estimation into the HL-RMS.
j. Extending the Bayesian recursive scheme from lumped to distributed model structures.
k. Continue work on ensemble forecasting schemes.
l. Technology transfer through (in person and telephone) meetings.

1.3. Project Accomplishments During Progress Report Period 06/01/2007 - 11/30/2007

Our activities during last six months focused on: a) improving the strategy for diagnostic evaluation of HL-DHM model, b) extending the regionalization approach introduced by Yadav et al. (2007) to identify behavioral parameter sets.

a) *An automated random sampling-based parameter sensitivity analysis was developed and integrated into the diagnostic model evaluation strategy.*

The HL-DHM model parameterized by Koren a priori parameterization scheme (Koren et al., 2000) was used as a baseline model to benchmark the model performance improvements obtained by diagnostic model evaluation strategy. An automated parameter sampling procedure was devised that incorporates a priori parameter grids as spatial constraints to reduce the dimensionality of the distributed model parameterization problem. It was assumed that the spatial pattern of a priori parameters, in terms of both location and *relative* magnitude of parameters, is well-defined (reasonable and reliable).

Then, the procedure simplifies to that of finding the parameters of a transformation function that modifies the *absolute* magnitude of the parameters in each a priori parameter grid. A novel non-linear transformation function has been devised for this purpose (see Yilmaz, 2007). Proposed non-linear transformation function contains only one parameter, β *(0,2]*. For example, Figure 1 shows that when $\beta = 1$ the transformation function has no effect on the distribution of the UZFWM parameter. However, as $\beta \rightarrow 0$ or $\beta \rightarrow 2$ the parameter values approach their feasible limits and the variance is compressed, while the relative ordering of spatial magnitudes is maintained. The Beta transformation has two important characteristics that make it favorable for the current study. First, the magnitude ordering of the parameter values is maintained (while the grid-to-grid relative parameter ratios may change). Second, the parameter values can now vary fully within the entire feasible region without the need to set subjective thresholds. This procedure was utilized to sample 1600 parameter sets (β) from the HL-DHM model (Figure 2). In Figure 2, each subplot is plotted for a single parameter, each gray dot represents the median value of the transformed parameter values within a parameter grid and x-axis shows the feasible parameter range in the transformed space. The thick, thin and dashed horizontal blue lines correspond to the median, 25th and 10th quantiles of signature measure distribution along the feasible range of the parameters respectively. Therefore, on any scatter plot, the region between two thin blue lines contains 50% of the sampled points. The slope of these quantile lines is therefore an indicator of whether a relationship exists between a signature measure and a parameter. The horizontal lines indicate no relationship, whereas sloping (positive or negative) lines indicate that there exists a relationship between a signature measure and a parameter. It can be seen from Figure 2 that %TBias measure is sensitive to UZTWM and LZTWM, and to some extent to the PFREE parameter because these are the parameters controlling the evapotranspiration from the HL-DHM model. The same parameter sensitivity strategy was also applied to other signature measures and the results are provided in Yilmaz (2007).

Future work will focus on HL-DHM model deficiencies caused by incorrect spatial distribution of parameters.

b) *The new regionalization approach introduced by Yadav et al. (2007) was extended to include a global multi-objective optimization approach to identify behavioral parameter sets.*

Yadav et al. (2007) showed that regionalized indices of watershed response characteristics (e.g. runoff ratio) can be used to considerably reduce the uncertainty in ensemble streamflow predictions at ungauged locations. Indices of hydrologic response were regressed against physical and climatic watershed characteristics including uncertainty. The regression thus provides ranges for response indices at ungauged locations and models (parameter sets) not producing indices within these ranges can be eliminated as unacceptable from ensemble forecasts. One drawback of the proposed strategy was, however, that a very inefficient Monte Carlo approach was used to identify acceptable parameter sets. This limited the approach to very few indices that could be used simultaneously, and to simple watershed model structures. We recently reformulated the problem of finding acceptable parameter sets as a multi-objective optimization problem and solved it using a evolutionary genetic algorithm. Running the algorithm on a parallel cluster, we showed that for 30,000 function evaluations a relatively consistent population of about 10,000 acceptable parameter sets could be found. This improvement to the approach will now allow us to test it on spatially explicit watershed models such as the HL-DHM.

An extension of this approach is in review with WRR. The Monte Carlo approach used initially several lacked efficiency in finding behavioral parameter sets in several of the watersheds. This would render the approach only suitable to very simple models. We thus reformulated the problem so that it can be solved within a multi-objective optimization framework. Zhang et al. (In Review) shows how the sampling with a multi-objective Genetic Algorithm significantly increased the efficiency of the approach. We will test more complex models and a larger number of regionalized constraints in future studies.

2. SUMMARY OF RESEARCH AND EDUCATIONAL EXCHANGES

Scientific exchanges between UA/PSU researchers and NWS-HL personnel have taken place in the form of phone calls and e-mails. The research performed for this project was incorporated into the DMIP-2 experiment organized by the NWS Hydrology Laboratory. Victor Koren, Mike Smith and Zhengtao Cui of HL provided technical assistance for HL-DHM model.

3. PRESENTATIONS AND PUBLICATIONS

Pokhrel, P, Estimation of Spatially Distributed Model Parameters Using a Regularization Approach, MS Thesis, Univ. of Arizona, April 2007.

Reed, P.M., Tang, Y., *Werkhoven, K. van and Wagener, T. 2007. Using Global Sensitivity Analysis to Better Understand How Real-Time Observations Influence Operational Flood Forecasts in the Susquehanna River Basin. ASCE World Water and Environmental Resources Congress, Tampa, Florida, May 2007. (Oral)

Reed, P.M., Tang, Y., Van Werkhoven, K. and Wagener, T. 2007. Using global sensitivity analysis to better understand how real-time observations influence operational flood forecasts in the Susquehanna River Basin. Association of Environmental Engineering and Science Professors, Blacksburg, Virginia July 2007. (Poster)

Tang, Y., Reed, P., Van Werkhoven, K. and Wagener, T. 2007. Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis. *Water Resources Research*, 43, doi:10.1029/2006WR005813.

Tang, Y., Reed, P., Wagener, T. and Van Werkhoven, K. 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrology and Earth System Sciences*, 11, 793-817.

Van Werkhoeven, K., Wagener, T., Reed, P. and Tang, Y. 2007. Sensitivity analysis of a distributed hydrologic model for uncertainty reduction and identification of dominant model controls. 14th IUGG General Assembly, Perugia, Italy, 2-13th July 2007. (Poster)

Van Werkhoven, K., Wagener, T., Tang, Y., and Reed, P. Understanding watershed model behavior across hydro-climatic gradients using global sensitivity analysis. *Water Resources Research*. (In Press)

Van Werkhoven, K., Wagener, T., Tang, Y., and Reed, P. Complexity reduction in multiobjective watershed model calibration. *Water Resources Research*. (In Preparation for Submission in October)

Wagener, T., Gupta, H.V., Yilmaz, K. and Yadav, M. 2007. Evaluation of hydrologic models in ungauged basins using regionalized watershed response characteristics. 14th IUGG General Assembly, Perugia, Italy, 2-13th July 2007. (Oral)

Wagener, T., Reed, P., *Werkhoven, K. van and Tang, Y. 2007. Identification and evaluation of complex environmental systems models using global sensitivity analysis and evolutionary multiobjective optimization. International Workshop on Advances in Hydroinformatics (HIW07), Niagara Falls, Canada, June 4-7th 2007. (Invited Keynote)

Yadav, M., Wagener, T. and Gupta, H.V. 2007. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources*, 30(8), 1756-1774. doi:10.1016/j.advwatres.2007.01.005.

Yilmaz, K., and H.V. Gupta, Towards improved modeling for ungauged and poorly gauged basins: Presented at Middle East Technical University, Ankara, Turkey, June 29, 2007 (Invited Talk)

Yilmaz, K., H.V. Gupta and T. Wagener, Diagnostic Evaluation of a Distributed Watershed Model: A Process-based Approach, Presented at the IUGG General Assembly, Perugia, Italy, July 9–13, 2007 (Oral)

Yilmaz, K., K., 2007. Towards improved modeling for hydrologic predictions in poorly gauged basins, Ph.D. Dissertation, Univ. of Arizona.

Zhang, Z., Wagener, T. and Reed, P. Ensemble streamflow predictions in ungauged basins combining regionalized streamflow indices and multiobjective optimization. *Water Resources Research*. (Pending Revisions)

4. FUTURE WORK

In the light of the experience we have gathered from the above analysis, the following studies will be performed:

a) Measures that are powerful in diagnosing HL-DHM model inadequacies will be further researched. Test runs using spatially lumped Sacramento model is expected give more insight into the diagnostic measure formulation.

b) The utility of the non-linear transformation function discussed in Section 1.3.a will be further explored via incorporating into the regularization framework developed by Pokhrel (2007).

c) Parallel processing techniques developed at the Penn State University will be implemented at the University of Arizona to perform faster model runs required by the optimization algorithms.

d) Implementation of regionalization approach in US watersheds. (ongoing)

e) Connecting sensitivity analysis and model calibration in combined procedure. (Van Werkhoven et al., In Preparation)

5. SUMMARY OF BENEFITS AND PROBLEMS ENCOUNTERED

Benefits that have been experienced during the last year

a. UA/PSU researchers are becoming familiar with the HL-DHM distributed hydrologic model developed by NWS-HL in an effort to contribute and share new ideas. Students are becoming familiar with NWS software, methods and procedures.

b. Project provided research subjects for two master theses (Prafulla Pokhrel, Univ. of Arizona & Maitreya Yadav, PSU) and in part a Ph.D. dissertation (Koray K. Yilmaz, Univ. of Arizona). Sub-projects have been derived from this project for elements of the PhD dissertations by Yong Tang and Katie van Werkhoven (PSU).

c. Fruitful communication between UA/PSU researchers and HL personnel has continued.

Problems encountered

a. No significant problems were encountered during the last 6 months.

6. REFERENCES

Koren, V. I., Smith, M., Wang, D. and Zhang, Z. (2000). Use of soil property data in the derivation of conceptual rainfall-runoff model parameters. 15th Conference on Hydrology, Long Beach, American Meteorological Society, Paper 2.16, USA.

7. FIGURES

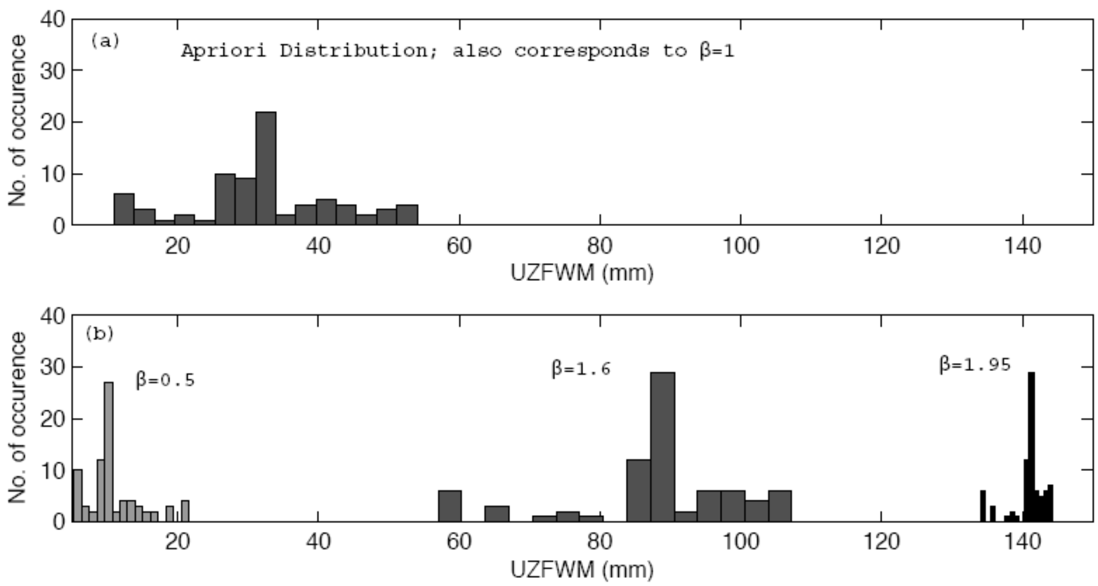

Figure 1. Distribution of parameter values within UZFWM parameter grid, a) A priori specified values b) parameter values after using β-transformation function utilized in this study (for β=0.5, 1.6 and 1.95).

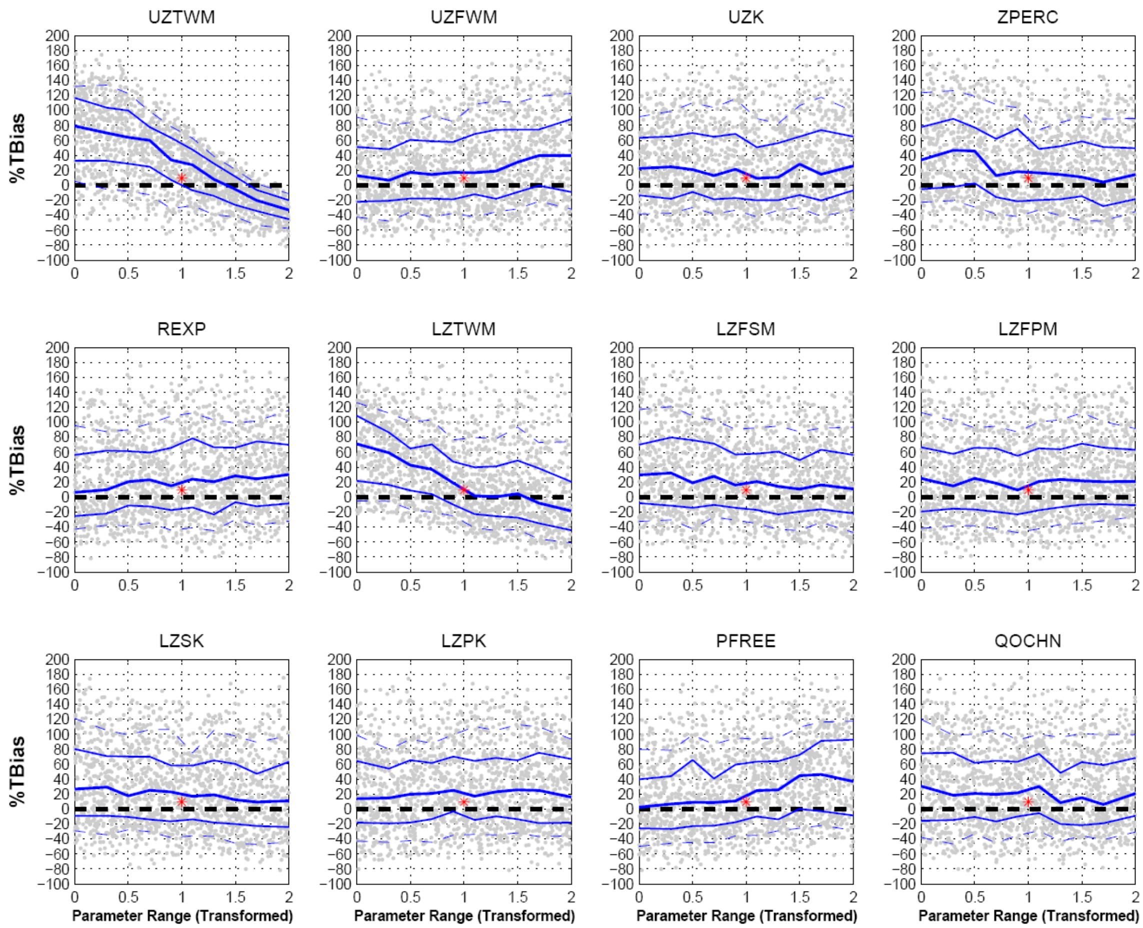

Figure 2. Scatter plots showing the relationship between %TBias (total bias) signature measure (y-axis) and each parameter (x-axis represents β-transformed parameter feasible range) based on 1600 randomly sampled parameter sets (dots) (See Yilmaz, 2007 for more details).