10.2 A COMPARISON OF STATISTICAL AND EXPLICIT SHORT-TERM HYDROLOGICAL FORECASTING TECHNIQUES

Glenn Van Knowe*, Kenneth. T. Waight
AWS Truewind, Troy, NY

Manel Ceperuelo, Joan Aymamí, Santi Parés
Meteosim, Barcelona, Spain

Sankar Arumugam and Jesesung Oh
North Carolina State University

INTRODUCTION

Real-time operational hydrological forecasts are needed for several reasons, including hydroelectric power generation, water resource planning and warning of possible water-related disasters. There are several modeling approaches to real-time hydrological forecasts, each having strengths and weaknesses. In an attempt to assess the costs and benefits of one model approach as compared to another, we compared the performance of real-time forecasts produced by a statistically based model with forecasts produced by an explicit distributed hydrological model for both long and short-term forecasting.

The approach employed for the statistical forecasts used a type of constructed analogs approach with nonparametric resampling. The explicit model being used in the evaluation was the Distributed Hydrology Soil Vegetation Model (DHSVM). DHSVM is a spatially distributed hydrological model that explicitly represents the effects of diverse topography and heterogeneous subsurface conditions on the downslope redistribution of subsurface moisture that provides a dynamic representation of the spatial distribution of soil moisture, snow cover, evapotranspiration, and runoff.

The statistical forecasting method has two advantages; it is easier to implement and computationally very efficient. However, initial results indicate that this method may underestimate the rate of runoff and streamflow as compared to the distributed model.

The results of the 2007 water year from historical model simulations run for the Lewis River basin in Washington, with both methods run in forecast mode (that is, only data available in real-time were used in the simulations) are compared in this paper.

2. FORECASTING METHODS

2.1. Statistical Method. The statistical method uses a type of constructed analogs approach (Devineni and Sankarasubramanian, 2010, Sankarasubramanian et al. 2009) as shown in Figure 1. Constructed analogs is a statistical approach to climate downscaling that constructs an analog from a linear combination of past patterns. This methodology has shown good forecasting skill in the past. In this method observational precipitation and streamflow data are used to provide initial conditions. Precipitation from a reforecasted data set is used for future conditions.

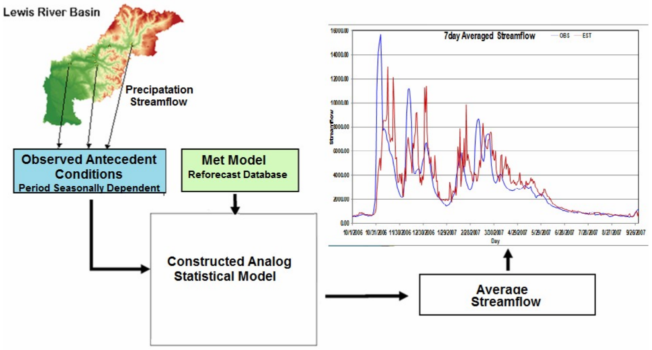

Fig. 1: A schematic illustration of the components of the statistical forecast method.

2.2. Distributive Method. The DHSVM was used as the distributive model. It is a distributed hydrology model that was developed at the University of Washington (Wigmosta, et al., 2002). It has been widely applied both operationally for streamflow prediction for hydropower and in a research capacity to examine the effects of changes in vegetation, forest management, etc. on streamflow.

The DHSVM is a spatially distributed hydrological model that explicitly represents the effects of diverse topography and heterogeneous subsurface conditions. The effects represent the downslope redistribution of subsurface moisture that provides a dynamic representation of the spatial distribution of soil moisture, snow cover, evapotranspiration, and runoff. (Figure 2)

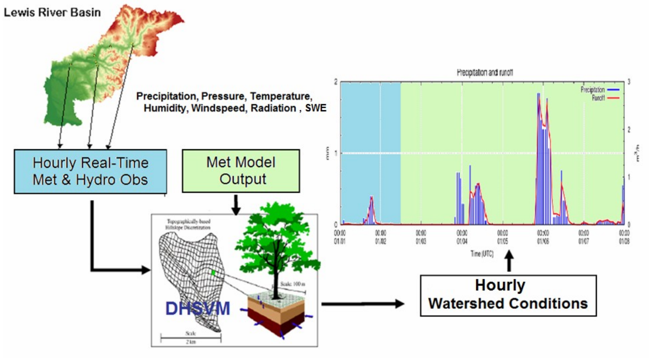

Fig. 2: A schematic illustration of the components of the DHSVM distributed forecast model.

Corresponding author address: Glenn Van Knowe
AWST, 185 Jordan Road., Troy, NY 12180;
email:gevanknowe@meso.com

The observational input data for the DHSVM initial conditions are streamflow, precipitation, snow water equivalent (SWE), temperature, relative humidity, wind speed, temperature lapse rate, and terrain height. The input data for the forecast times were from the numerical model called the Mesoscale Atmospheric Simulation System (MASS) (Kaplan et al. 1982; Manobianco et al. 1996). The variables from the NWP model are precipitation, temperature, relative humidity, wind speed, shortwave radiation, longwave radiation and temperature lapse rate.

3. EVALUATION RESULTS

The comparisons were made for the Lewis River Basin in the southwest corner of Washington (Figure 3). The three sites evaluated were Merwin reservoir, Swift reservoir and Yale reservoir dams. The results for the Merwin reservoir dam are presented as they were representative of the results of all three sites.

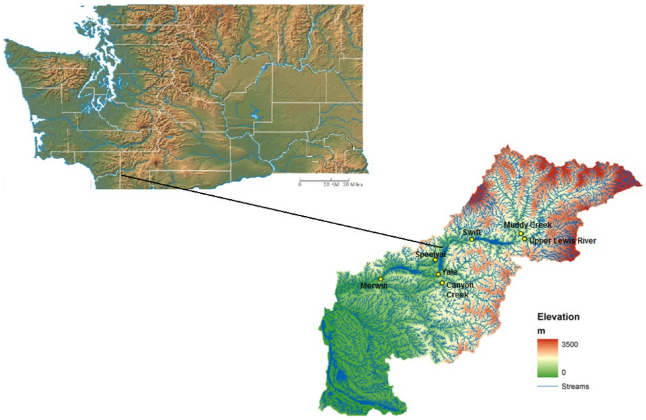

Fig. 3. The Lewis River Basin, location of the study area.

The period of study was for the 2007 hydro year which ran from October 2006 to September 2007. The model forecasts and observational verification were by necessity different for the statistical versus the distributed model. For the statistical model, the seven day average flow rates were used and for the explicit model instantaneous flow rates were used.

The typical results of the performance of the statistical model, with forecast and verification observations (both from the seven day streamflow averages) are provided in Figure 4. The results show that the statistical model does a very good job at representing the seven day averages, but the flow rates are much lower than the peak instantaneous rates typically needed for hydro power prediction, especially during periods of heavy precipitation. The statistical model actually does fairly well for both seven day and as an estimate of instantaneous flow rates during the warm season months when flow rates are low and consistent. But the statistical approach does not do well at representing instantaneous flows during the heavy precipitation events of the cold rainy season, even though the seven day averages are well represented by the statistical model.

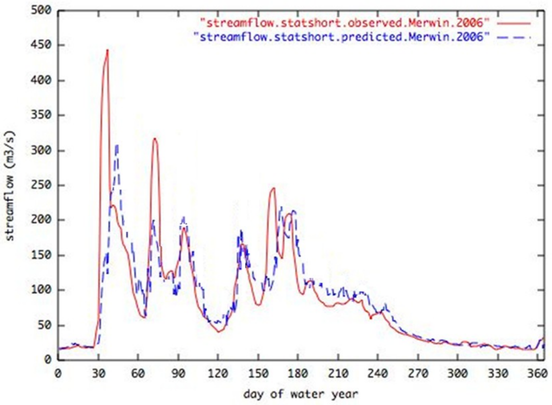

Fig. 4. Statistical model comparison of 24-hr ahead stream flow forecast with observed streamflow. The statistical model forecasts (blue dashed line) are compared with the observed stream flow (red line) for the 2007 water year at Merwin Dam, Washington.

The DHSVM model for instantaneous streamflow rates had very different prediction characteristics from the statistical model as shown in Figure 5. The distributed model provides much more representative instantaneous flow rates. The distributed model also does a better job on capturing the extreme events. However, there is an unexplained high bias in the warm season months.

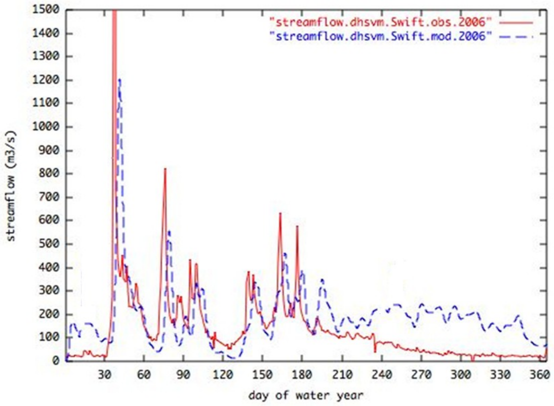

Fig. 5. Distributed model comparison of 24-hr ahead stream flow forecast with observed streamflow. The DHSVM model forecasts (blue dashed line) are compared with the observed stream flow (red line) for the 2007 water year at Merwin Dam, Washington.

Typical results of the performance of the two models are summarized in Figure 6. The figure compares the next day statistical model forecasts and the DHSVM model forecasts with the observed instantaneous stream flow for the 2007 water year at Merwin Dam, Washington. The statistical forecasts are taken from the seven day average forecast. The DHSVM forecasts are for instantaneous streamflow.

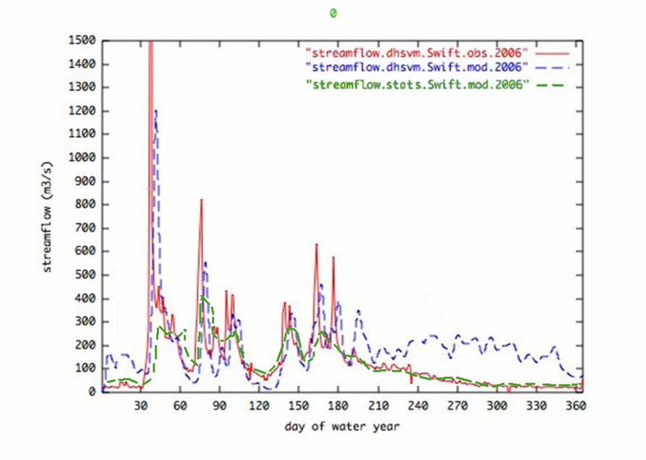

Fig. 6. Statistical model forecast (green dashed line) - DHSVM model forecast (blue dashed line) comparison 24-hr ahead streamflow forecasts with instantaneous streamflow observations (red line) for the 2007 water year at Merwin Dam.

This analysis highlights the fact that the statistical approach is representative of the instantaneous streamflow during times of slow streamflow changes. But the results also show it does not do well in capturing extreme events. It is interesting to note that the statistical model is able to obtain reasonable results with using coarse resolution precipitation (even poor quality) data as input as long as the data was consistent (i.e. consistently low or high).

When comparing the forecast of the two models, the distributed model clearly does a better job on capturing the extreme events, especially during the cold rainy season. However, the distributed model significantly over-predicted the streamflow during the dry warm season, which needs to be corrected in order to make useful streamflow forecasts.

4. SUMMARY AND CONCLUSIONS

The results of this study show that the statistical model does well for averaged flow and times of slow changes. It does not do well in capturing extreme events. Also, it does not provide instantaneous stream flow rates. But it does not require high resolution or quality precipitation data to obtain reasonable results as long as the data is consistent. The statistical model is able to obtain reasonable results with coarse resolution precipitation as input. So this is an advantage in addition to the fact that it is easier to implement and computationally very efficient.

The distributed model does a better job on capturing the extreme events. Also, this option is required for high temporal resolution forecasts. But there are times when the distributed model was consistently worse than the statistical model, such as during the high bias noted during the warm season.

The statistical model could be used to improve distributed model forecasts. For example, the statistical forecast could be used in place of the distributive forecast in the warm dry season months. Another way the distributive model forecast could be improved is by adjustment thorough a Model Output Statistics (MOS) approach technique (Wilks, D.S. 1995). This could be very effective in removing the warm season bias.

6. REFERENCES

Devineni, N. and A. Sankarasubramanian, 2010: Improving the Prediction of Winter Precipitation and Temperature over the Continental United States: Role of ENSO State in Developing Multimodel Combinations, Monthly Weather Review , in press.

Kaplan, M. L., J. W. Zack, V. C. Wong and J. J. Tuccillo, 1982: Initial results from a mesoscale atmospheric simulation system and comparisons with an AVE-SESAME I data set., Monthly Weather Review, Vol 110, 1564-1590.

Manobianco, J., J. W. Zack and G.E. Taylor, 1996: "Workstation-based real-time mesoscale modeling designed for weather support to operations at the Kennedy Space Center and Cape Canaveral Air Station." Bulletin of the American Meteorology Society, 77, 653-672.

Sankarasubramanian, A., U. Lall, N.Devineni and S. Espunevea, 1982: Utility of Operational Streamflow Forecasts in Improving within-season Reservoir Operation, Journal of Applied Climatology & Meteorology, 48(7), 1464–1482.

Wilks, D.S., 1995: Statistical Methods in the Atmospheric Sciences: an Introduction, International Geophysics Series, Vol. 59, Academic Press, 464 pp.

Wigmosta, M.S., B. Nijssen, P. Storck, and D.P. Lettenmaier, 2002: The Distributed Hydrology Soil Vegetation Model, In Mathematical Models of Small Watershed Hydrology and Applications, V.P. Singh, D.K. Frevert, eds., Water Resource Publications, Littleton, CO., 7-42.