A DISTRIBUTED MODELING SYSTEM FOR SHORT-TERM TO SEASONAL ENSEMBLE STREAMFLOW FORECASTING IN SNOWMELT-DOMINATED BASINS

Mark S. Wigmosta, M. Kashif Gill, André M. Coleman, Rajiv Prasad, and Lance W. Vail [1]

ABSTRACT

This paper describes a distributed modeling system for short-term to seasonal streamflow forecasts with the ability to utilize daily remotely-sensed snow cover products and real-time streamflow and meteorology measurements. Spatial variability in basin characteristics and meteorology is represented using a raster-based computational grid. Canopy interception, snow accumulation and melt, and simplified soil water movement are simulated in each computational unit. The model is run at a daily time-step with surface runoff and subsurface flow aggregated at the basin scale. This approach allows the model to be updated with spatial snow cover and measured streamflow using an Ensemble Kalman-based data assimilation strategy that accounts for uncertainty in weather forecasts, model parameters, and observations used for updating. Model inflow forecasts for the Dworshak Reservoir in north-central Idaho are compared to observations and to April-July volumetric forecasts issued by the Natural Resource Conservation Service (NRCS) for Water Years 2000 – 2006. October 1 and March 1 volumetric forecasts are superior to those issued by the NRCS. The ensemble spread brackets the observed April-July volumetric inflows in all years except for the October 1, 2006 forecast. Short-term (one and three day) forecasts also show excellent agreement with observations.

INTRODUCTION

Up to 80 percent of the water supply in the Western U.S. results from mountain snowmelt where water resource managers require accurate and timely predictions of water supplies to allocate limited resources. Therefore, an accurate characterization of snow accumulation and subsequent melt leads to better management of the system. In the current research, a streamflow river forecasting system is developed for short-term to seasonal forecasts using available meteorological data. For a typical streamflow forecasting approach, observations are used for estimation up to the beginning of the forecast time step and from there on, either past climatic traces or output from a numerical weather prediction model are employed to generate the forecasts. A common problem with such a forecasting scheme is the lack of proper characterization of uncertainty (Slater and Clark, 2006).

This paper provides a method to address uncertainty using an ensemble forecasting approach that also involves the assimilation of observations as they become available. A simplified version of the Distributed Hydrology Soil Vegetation Model (Wigmosta et al., 1994, 2002) was developed at the Pacific Northwest National Laboratory (PNNL) and is used to model snow and moisture states using ensembles of past meteorological traces. Primary inputs to the model include observed precipitation, mean air temperature, elevation, and land cover. The model is run over a spatially-based computational grid that is structured to align with the pixel resolution of various remote-sensing snow products.

The streamflow forecasting methodology is applied to the 6,325 km^2 Dworshak watershed in north-central Idaho State (Figure 1) and compared to Natural Resource Conservation Service (NRCS) volumetric forecasts and to daily inflows estimated by the US Army Corps of Engineers (USACE). Lapse rates are used to distribute meteorological inputs to all model grid cells using an elevation difference approach. An ensemble type approach accounts for uncertainty in weather forecasts, model parameters, and observations.

METHODS

For this preliminary application, model output ensembles are produced using ensembles from past traces of meteorological forcing as a surrogate for the future conditions. The approach also accounts for model parameter uncertainty through a sampling strategy within the calibrated parameter space. There are seven parameters in the hydrologic model that are determined through calibration. An evolutionary computation optimization scheme called

Paper presented Western Snow Conference 2007

[1] Pacific Northwest National Laboratory, Richland, WA, Mark.Wigmosta@pnl.gov

"Particle Swarm Optimization (PSO)" is used to calibrate the model (Eberhart and Kennedy, 1995; Gill et al., 2005). The calibration is done in a way such that multiple initializations are simulated to search within the plausible parameter space, each giving a so-called optimal parameter set; thus, defining a narrowed region (also called feasible region) within the parameter space giving equally feasible solutions. Each ensemble of meteorological input is run with a parameter set sampled from the feasible region, thus giving an ensemble of model output (streamflow in this case). This approach not only accounts for uncertainty within the inputs and the parameters, but also reduces the computational burden required to run individual ensemble members.

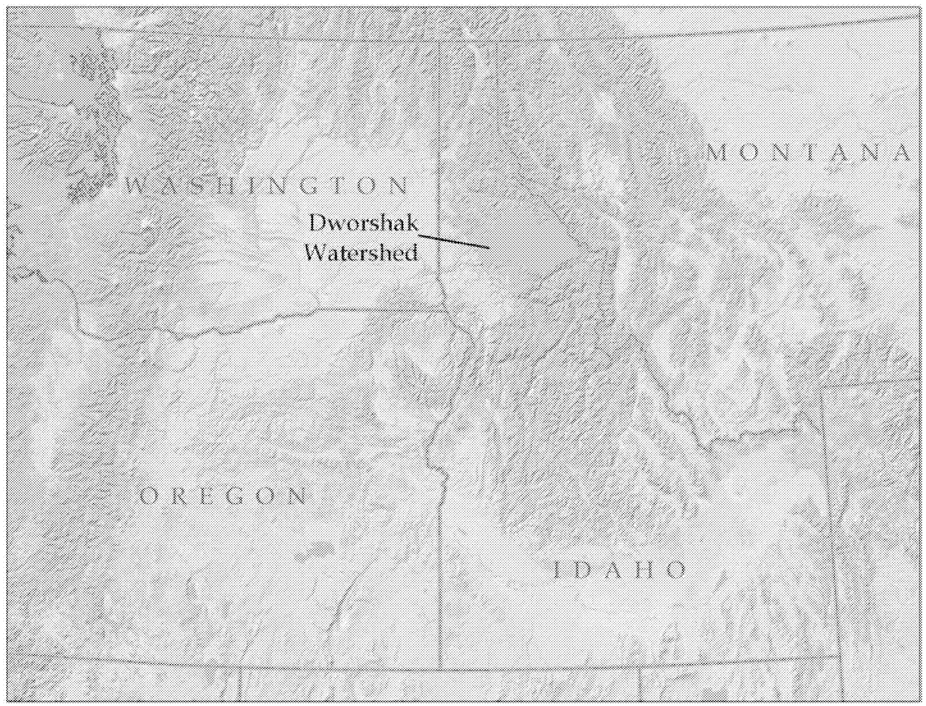

Figure 1. Location of the Dworshak watershed in north-central Idaho

For this study, observed streamflow refers to daily inflow to the Dworshak Reservoir estimated by the USACE using a mass balance approach based primarily on measured changes in pool elevation and the amount of water released from the dam. The potential error associated with back-calculating inflow via mass balance is expressed using an ensemble of estimated inflows. With this approach, the ensemble of model output generated using a meteorological ensemble and sampling from the feasible parameter space, is combined with the ensemble of streamflow observations to update the model state. This is done by using a data assimilation scheme called Ensemble Kalman filter (EnKF). The Ensemble Kalman filter (EnKF) method by Evensen (2002) is a widely popular data assimilation tool in weather, ocean, and hydrologic prediction modeling. The model output ensemble is assimilated with the observation ensemble using a weighting scheme that is governed by the covariance of the model as well as the observation ensemble. Thus, if the observation error covariance approaches zero, the actual measurements (i.e., inflow estimates) are trusted more, while the model output is trusted less. Similarly, as the model estimate error covariance approaches zero the observations are given less and less weight while the predictions are given more weight.

The EnKF assimilation scheme is outlined in Figure 2. During a given time step, a single parameter set is selected from the feasible region and the model is run for the first meteorological ensemble member, then a second parameter set is selected and the model is run for the next meteorological ensemble member. The process is repeated for all ensemble members, generating a model output ensemble. The model output ensemble is then assimilated with the observation ensemble via EnKF to update model states and the process is repeated for the next time step. The EnKF approach developed in this study accounts for: 1) Model Parameter Uncertainty, where the model is calibrated with historical data to obtain best parameter sets that yield equally feasible solutions; 2) Meteorological Forecast Uncertainty, where an ensemble meteorological forecasts is generated by sampling historical record based on similarity (Hausdorff Norm) of current conditions with previous years; the model also allows ensembles of short-term forecasts from numerical weather models to be used; 3) Model Uncertainty, where for each forecast ensemble member a parameter set is drawn from the equally feasible region assuming a uniform distribution and the model is run, generating ensemble output; 4) Observational Uncertainty, where an ensemble of observations is generated using measured values and estimated measurement error (streamflow and snowpack properties), and 5) Data Assimilation with Ensemble Kalman Filter, where the covariance of ensembles is used for weighting to update model soil water storages and model snow properties.

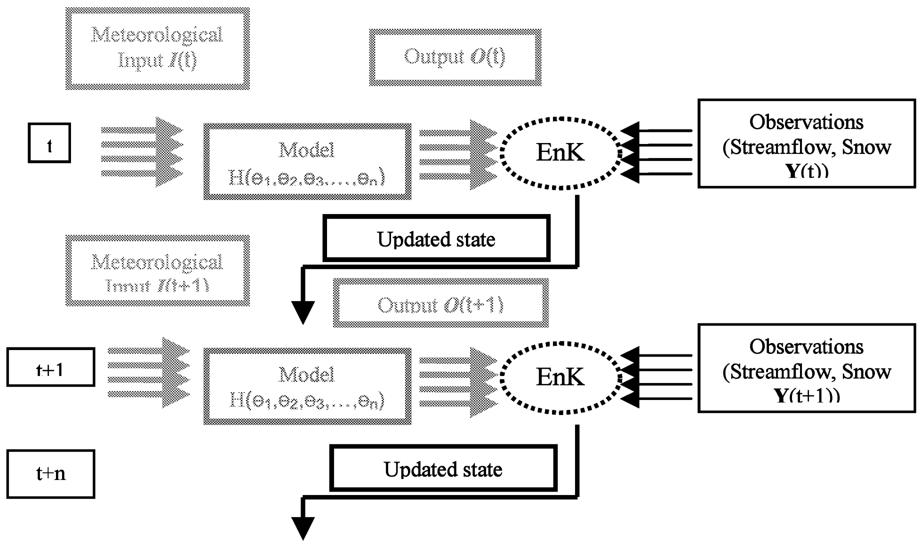

Figure 2. Flow diagram for ensemble updates showing the modeling process

RESULTS AND DISCUSSION

The NRCS issues seasonal streamflow volume forecasts for the Dworshak Reservoir during the melt season (April to July) each month starting in October. These estimates are based on multiple regression techniques using snow, precipitation, and streamflow observations from the previous months. For comparison with the NRCS volumetric forecasts, the daily ensemble forecast volumes are aggregated from April through July to provide total forecast volumes. As will be discussed, the ensemble mean is used for direct comparison to NRCS values, while box and whisker plots are used to compare the ensemble range with observed seasonal volumes obtained by aggregating the daily USACE inflow estimates.

As a first case, the results of the streamflow volume forecasts in this paper are compared to those issued by NRCS on October 1st (without updating). In the second case, the ensembles are updated using streamflow until the end of February, with no update from there on. The volume forecasts from April through July are then compared with results from NRCS issued on March 1st. The cases are shown by a schema in Figure 3. It should be noted that in the preliminary application described, streamflow and meteorology were used to update the model between October 1 and March 1; remotely-sensed snow cover state was not used to update the model.

The October 1st model results along with the observed and NRCS forecasts volumes are shown in Figure 4 for the years 2000 to 2006. The model volume forecasts are generally in good agreement with the observed volumes, and performed better than NRCS forecasts with the exception of years 2004 and 2006. The distribution from the October 1st and March 1st model output ensembles is shown in Figure 5. It is clear that the ensemble spread is bracketing the observations for all the years except the October 1, 2006 forecast. Interestingly, the March 1st NRCS forecasts improve drastically for certain years (i.e. 2001), but overall, the ensemble forecasting method described in this research produced superior results.

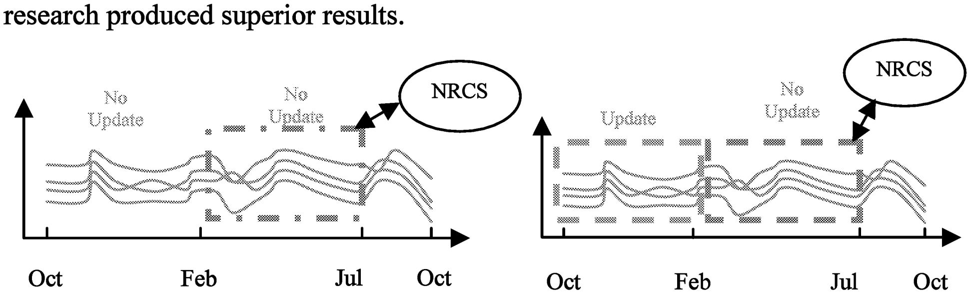

Figure 3. Approach for streamflow updating, (a) no streamflow updating for the October 1 forecasts, (b) observed streamflow was used during the October – February timeframe to update the model prior to March 1 forecasts

Short-term forecasts were also evaluated for Water Year 2006. The results for one-day and three-day forecasts are presented in Figure 6 and Figure 7, respectively. The observations are shown by the dark line where as the model ensemble spread is represented using the shaded polygon. The one-day and three-day forecasts are issued daily after updating the model state with the observations using EnKF. Notice that the model ensemble spread for the three-day forecast is wider than the one-day forecasts, meaning that the confidence in the one-day forecast is higher than the three-day forecast.

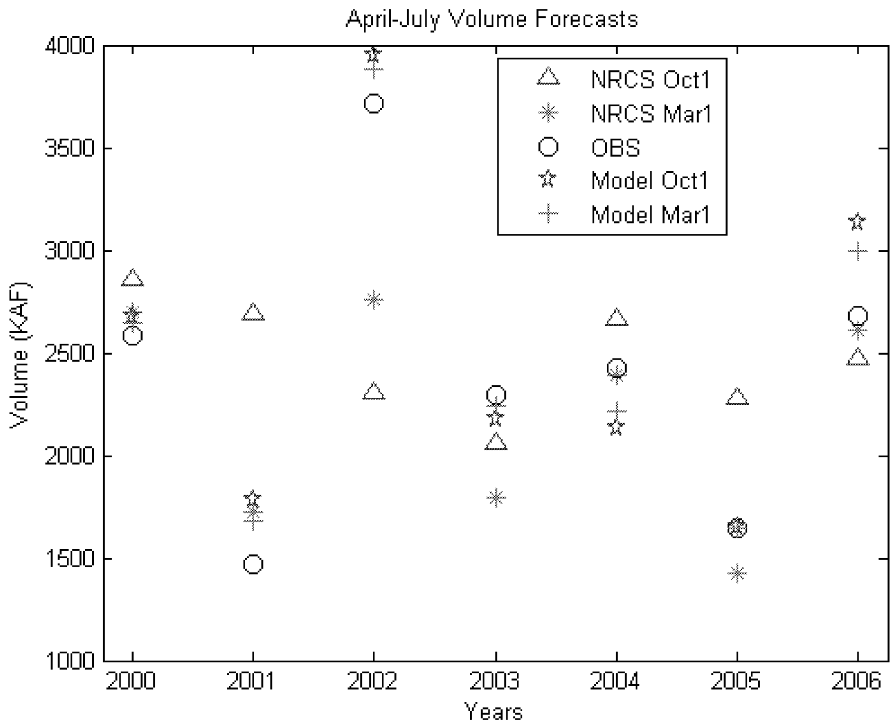

Figure 4. October 1 and March 1 volume forecasts compared with NRCS and observed for April-July

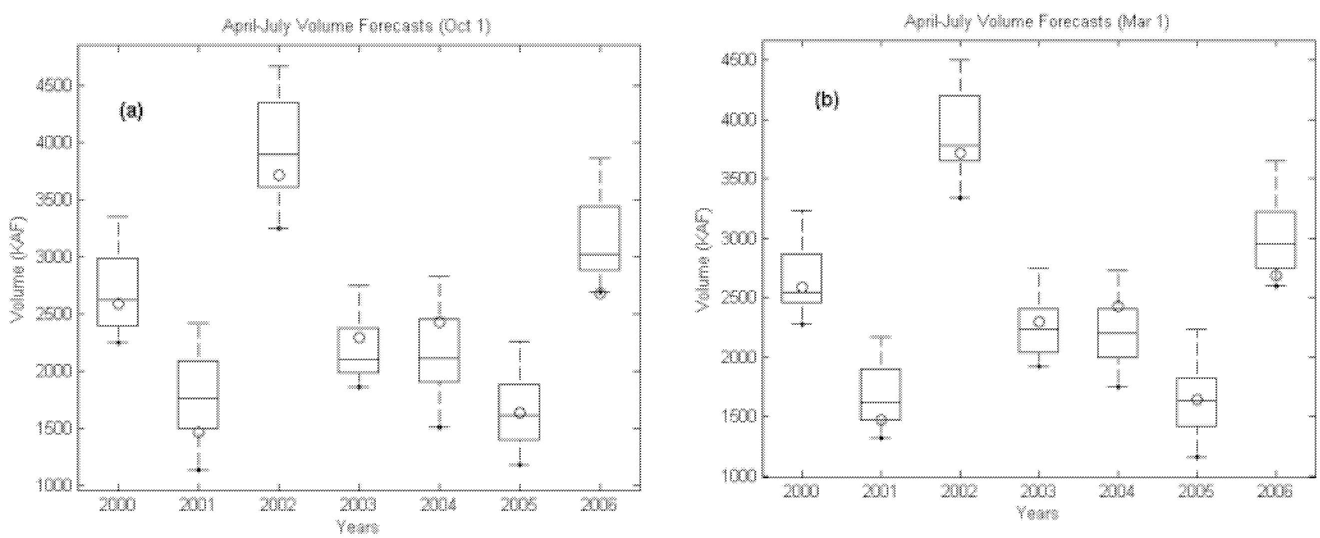

Figure 5. Box-plots of October 1st (a) and March 1st (b) volume forecasts for April-July. Observed volume are shown by the open circle. Ensemble results plotted for the median, 25th and 75th exceedence, and minimum and maximum

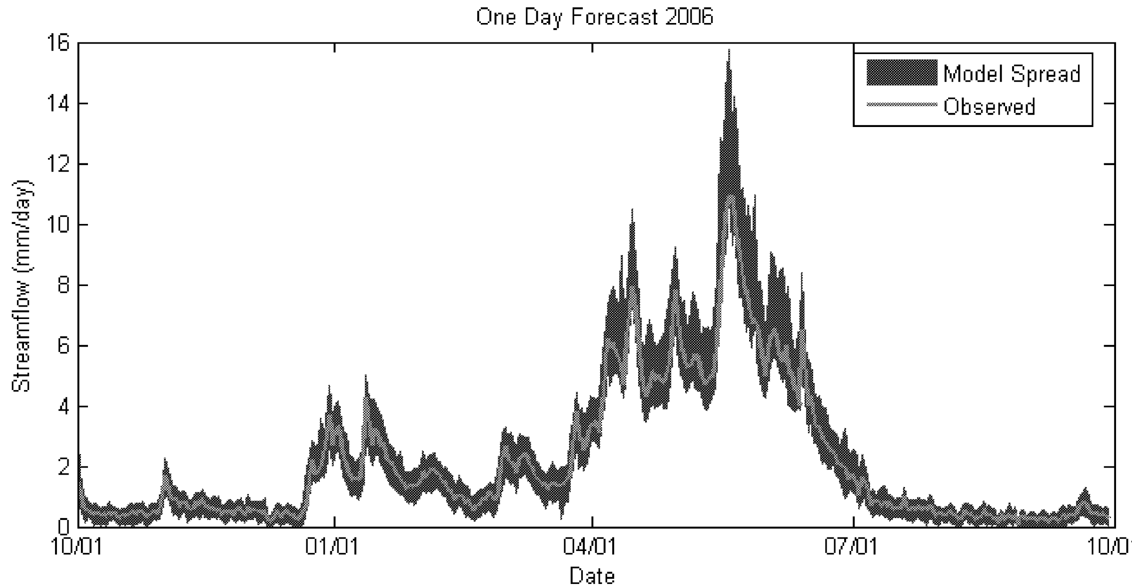

Figure 6. One-day streamflow forecasts for 2006

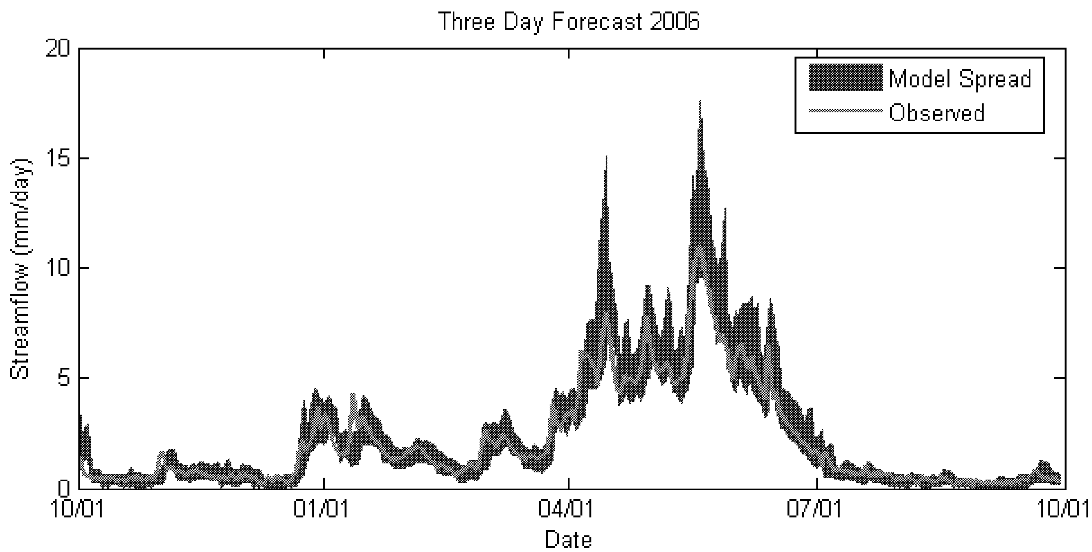

Figure 7. Three-day streamflow forecasts for 2006

CONCLUSIONS

The current research presents an ensemble streamflow forecasting approach with data assimilation that uses a spatially distributed hydrologic model and accounts for uncertainty in weather forecasts, model parameters, and observations. The method is used for short-term to seasonal streamflow forecasts in the snow-dominated Dworshak watershed in north-central Idaho. Forecast volumes for the snowmelt season (April-July) in Water Years 2000-2006, issued at the beginning of the water year (October 1st) are compared to those observed and to those made by the NRCS. The results are found to be in agreement with the observed volumes except in Water Year 2006 and superior to NRCS forecasts in five of the seven years examined.

In the second phase, the model soil water state is updated with the observed streamflow using Ensemble Kalman filter (EnKF) until the beginning of March and no update from there on. A partial improvement in the snowmelt season volume is observed using the updated streamflow. The results were again superior to the NRCS forecasts in five of the seven years. It should be noted that remotely-sensed snow cover was <u>not</u> used to update the model between October 1 and March 1. The current work to incorporate MODIS snow cover, the NOHRSC SNODAS data products, and gridded meteorological forecast data (i.e. National Digital Forecast Database) in the assimilation scheme is in progress. It is anticipated that using an accurate characterization of snow store using remotely-sensed snow products and meteorological forecast data will improve March 1st forecasts. One and three day forecasts with streamflow updating show excellent agreement with observations during Water Year 2006.

ACKNOWLEDGEMENTS

We wish to thank the Pacific Northwest Regional Collaboratory (PNWRC - www.pnwrc.org), Pacific Northwest National Laboratory's Energy Infrastructure Operations Center, and NASA Solutions Network for their support in this work.

LITERATURE CITED

Eberhart, R. C. and J. Kennedy. 1995. A new optimizer using particle swarm theory. In Proceedings, Sixth Symposium on Micro Machine and Human Science, p. 39–43. IEEE Service Center, Piscataway, NJ.

Evensen, G. 2002. Sequential Data Assimilation for Nonlinear Dynamics: The Ensemble Kalman Filter, In Ocean Forecasting: Conceptual basis and applications, edited by N. Pinardi and J. D. Woods, Springer-Verlag Berlin Heidelberg.

Gill, M. K., Y. H. Kaheil, A. Khalil, M. McKee, and L. Bastidas. 2005. Multiobjective Particle Swarm Optimization (MOPSO) for Parameter Estimation in Hydrology. Water Resour. Res., 42, W07417, doi:10.1029/2005WR004528.

Slater, A. G. and M. P. Clark. 2006. Snow Data Assimilation via an Ensemble Kalman Filter. Journal of Hydrometeorology 7(3):478.

Wigmosta, M.S., L. W. Vail, and D. P. Lettenmaier. 1994. A Distributed Hydrology-Vegetation Model for Complex Terrain. Water Resources Research, 30(6):1665-1679.

Wigmosta, M.S., B. Nijssen, and P. Storck. 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., p. 7-42.