



EPiC Series in Engineering

Volume 3, 2018, Pages 1232–1239

HIC 2018. 13th International  
Conference on Hydroinformatics



# A Parallel Flood Forecasting and Warning Platform Based on HPC Clusters

Ronghua Liu<sup>1,2</sup>, Liang Guo<sup>1,2</sup>, Yali Wang<sup>1,3\*</sup>, Xiaolei Zhang<sup>1,2</sup>, Qi Liu<sup>1,2</sup>,  
Yizi Shang<sup>1</sup>, Xiaoyan Zhai<sup>1,2</sup>, Jiyang Tian<sup>1,2</sup> and Dayong Huang<sup>1,2</sup>

<sup>1</sup> China Institute of Water Resources and Hydropower Research, Beijing, China

<sup>2</sup> Research Center on Flood & Drought Disaster Reduction of the Ministry of Water Resources,  
Beijing, China

<sup>3</sup> College of Hydrology and Water Resources, Hohai University, Nanjing, China  
yaliwang198816@126.com

## Abstract

As floods could be effectively forecasted by distributed hydrological model, their study and application became the key points of flood forecasting and early warning. Based on high performance computing clusters, a parallel flood forecasting and warning platform with the characteristics of partition, classification, and complicated process coupled was established to forecast and warn flood across China, especially for flash flood in China. In addition, the platform was based on China Flash Flood Hydrological Model (CNFF-HM). It used files (not MPI), which based on a shared hierarchical storage system, to pass message to control the start and stop of simulation processes, and the rapid communication among simulation processes was realized; pre-allocation and dynamic allocation methods was together applied to manage the resource of the high performance computing clusters; the automatic switch among different time scale models was realized by simulation driven strategy based on rainfall events; the reboot framework was designed to deal with the process crash and delayed rainfall data. The effectiveness and stability of the platform has been tested by the flood events of 2017. Finally, a case of Weishui catchment in Hunan Province was shown.

## 1 Introduction

Floods have resulted in 19,468 casualties in China since 2000, and 73% of those casualties were caused by flash floods (<http://www.mwr.gov.cn/english/>). The lack of numerous distributed hydrological model computing tool caused great difficulties in flood forecasting and warning for national level, especially for flash flood in China.

---

\* Corresponding author: Yali Wang, PhD, research on flash flood simulation and forecasting.

Flood forecasting and warning platform is to analysis the flood risk comprehensively. Distributed hydrological model has been proved to be an effective tool for flood forecast (Wang, et al., 2017). Compared the forecast or real-time rainfall/runoff with the rainfall/runoff thresholds is an important way to analyze the early warning of flood disaster. Nowadays, the flash floods forecast in China usually based on Rational Method, which is a lumped and conceptual model for flood peak discharge estimate. Moreover, the rainfall threshold is determined by experience. It does not consider the soil moisture, which lead to great limitation together with a large number of false warnings. An effective tool for flood forecasting and early warning was been demanded in China.

Parallel computing technology, cloud computing technology provides a good technical means for the construction of national or basin flash flood warning system to fine warning analysis (Shang, et al., 2016; Liu, et al., 2017). Hydrological simulation cloud computing services using a platform named ParFlow with high performance clusters was provided by Bürger (Bürger, Kollet, Schumacher, & Bösel, 2012). Two model integration methods and a cloud computing based hydraulic platform HydroMP were proposed by Liu et al. to support the large scale and concurrent hydrological and hydraulic modelling (Liu, Wei, Weng, Wang, & Tang, 2014). An integrated river management system was developed by Welsh et al., where basin runoff, river network analysis, submergence analysis, water quality, scheduling, irrigation management, and other models were integrated into a simulation and regulation system for water quantity and quality in river basin (Welsh, et al., 2013). The open modelling interface (OpenMI) provides a method of interoperation between model components, such as the definition of exchanged data and the drive methods (Gregersen, Gijssbers, & Westen, 2007). Besides, it has been reported by Overeem (2013) that the Community Surface Dynamics Modelling System (CSDMS) also aimed to couple different simulation models with the HPC resources using the modularized method, while Rahman et al. (2004) and David et al. (2013) found that the Object Model System (OMS) and Integrated Customer Management System (ICMS) have the ability to integrate the multiple models which have been developed in an object-oriented language and component-based programming (Rahman, Seaton, & Cuddy, 2004; David, et al., 2013).

In this study, a flood forecasting and warning platform in national level was presented, which was based on HPC clusters. And it is organized as follows. Section 2 presents methodology, including the distributed hydrological model, the hydrological model cluster across China, the characteristic of the HPC clusters and the method of parallel computing for the platform computing. Section 3 showed the case of a flood forecasting, the Weishui catchment in Ningxiang, Hunan province, China. Section 4 discusses the result of the flood forecast of Weishui catchment. Finally, Section 5 provides the major conclusion of this study.

## 2 Methods

### 2.1 Distributed hydrological model

The flood forecasting and warning platform was established based on a distributed hydrological model, China Flash Flood Hydrological Model (CNFF-HM), which was design for flash flood forecast by the China Institute of Water Resources and Hydropower Research (IWHR) in 2014. The CNFF-HM was characteristic by modularity, less parameters, and scalability.

The CNFF-HM generalized the catchment into many model elements, including sub-basins, reaches, nodes, reservoirs, etc. on the basis of hydrological processes and geomorphological properties (Figure 1). And the sub-basin was basic calculation unit. The model could perform continuous simulations and forecast flash floods with the time step of 1 day, 1 hour, or 30 minutes. Six calculation modules formed the model, which were used to calculate areal rainfall, evapotranspiration, runoff production, runoff concentration, flow routing, and reservoir regulation. For humid or sub-humid regions in China, the

saturation excess runoff production method (Zhao, 1992) was chosen for runoff production. In addition, for other regions, infiltration-excess runoff (Winchell, Gupta, & Sorooshian, 1998) or vertically-mixed runoff could use for runoff production calculation. The surface flow routing was calculated by distributed unit hydrograph (Maidment, Olivera, Calver, Eatherall, & Fraczek, 1996) (Maidment D. R., 1993), and the river flow routing was used dynamic Muskingum method (Song, Kong, & Zhu, 2011).

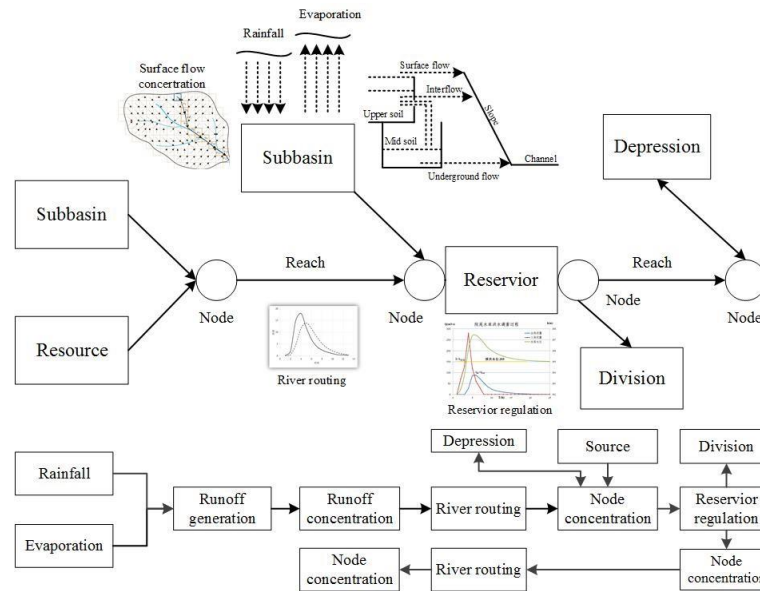


Figure 1: Structure and calculation processes of CNFF-HM

## 2.2 Framework of flood forecasting and warning platform

The flood forecasting and warning platform consisted of area rainfall model, soil moisture model, flood-simulating model, flood warning analysis model, flood routing model of river system (Figure 2). The area rainfall model was designed to fuse multi-source rainfall data, which include gauged station, radar, and numerical weather forecasting. The soil moisture was inverser by a geomorphology-based hydrological model (GBHM) (Yang, Herath, & Musiak, 1998) and remote sensing data in the soil moisture model. Flood simulating model, namely the CNFF-HM, calculated the river discharge continuously and timely. Two series thresholds including rainfall thresholds and runoff thresholds were applied for flash flood warning analysis. The large basin flood was calculated by the flood routing model of river system based on the discharge of sub-basins. For hydrological simulation and flood warning analysis, sixty river systems in China were divided into 132 hydrological model clusters (primary regions) according to climate and geographical features, and those clusters were further divided into 5013 sub-regions (Figure 3).

## 2.3 HPC clusters

The dispatcher of the flood forecast platform was running on a High-Performance Computing (HPC) cluster, which consists of 2 management nodes and 64 computing nodes. And those nodes connected with each other through an InfiniBand network with bandwidth up to 56 GB/s, as illustrated in Figure 4. The computing nodes, each with 12 processors and 64 GB RAM, shared a hierarchical storage system

which consisted of two index controllers and three data controllers. The index controllers were responsible for metadata access request, and the data controllers were used for storage system monitoring. Each data controller included 20 SAS hard drives of 900 GB and 4 SSD hard drives of 240 GB, the aggregation rate of storage nodes reached 3 GB/s. In hierarchical storage system, the most frequently accessed files could be adjusted automatically to SSD high-speed disk for storage so that it provided high speed for communication between each computing node.

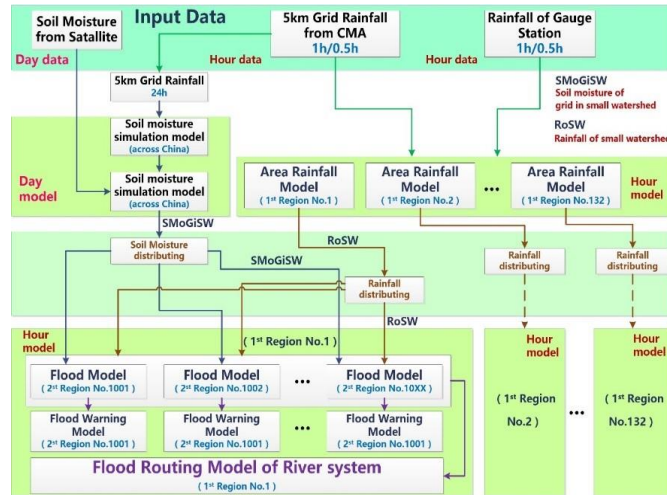


Figure 2: Models and their relationship of flood forecasting and warning platform

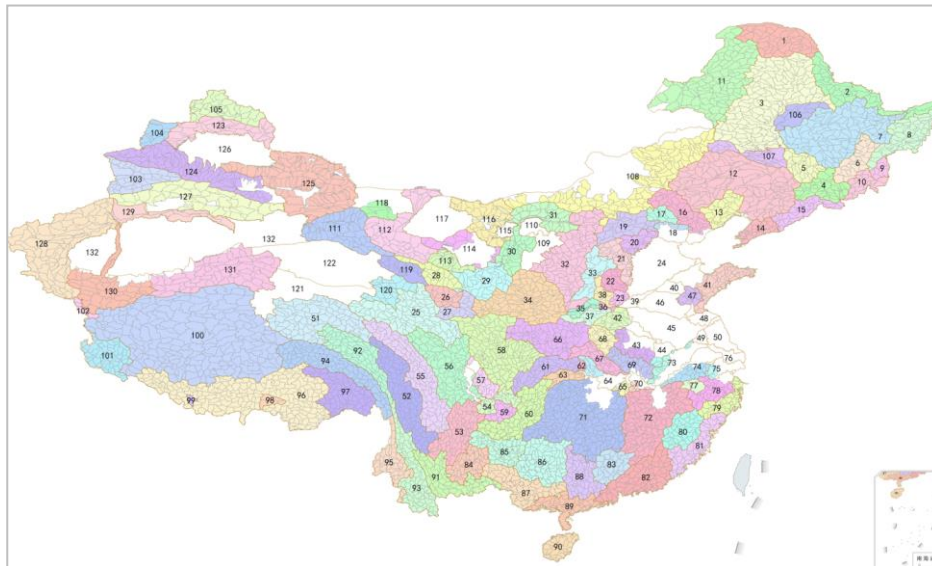


Figure 3: The distribution of hydrological model clusters and sub-regions for flood forecasting and warning

## 2.4 Parallel computing

The hierarchical storage system was used to store various types of model files and communicate rapidly between parallel model processes. As shown in Figure 5, the left side of the red vertical dash line was a brief structure of our scheduling management system, which was a platform based on B/S architecture and WebGIS, and used to manage the models, data, and simulation; the right side was software and simulation processes running on HPC cluster, including Gridview management system, Linux file system, task dispatcher, and various model processes. The

task dispatcher, including five threads for overall scheduling, exception handling, status statistics and Web API for querying task status and flood simulation events, was the most important part of parallel flood simulation engine and was responsible for the scheduling of all kinds of simulation process.

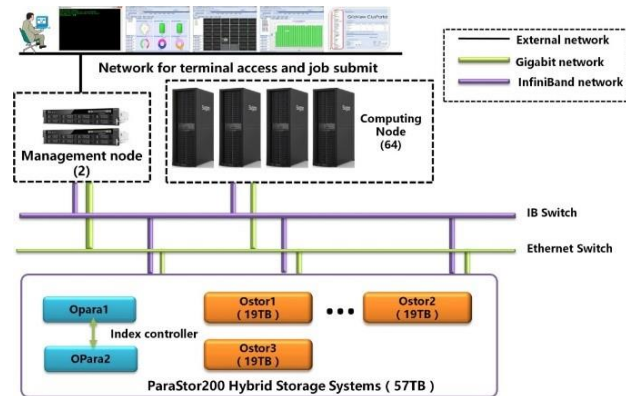
The parallel computing processes of hydrological model communicated with each other by controlling files. The connection of these processes was implemented automatically, the log file of former process was used as the controlling file of later process. The successive process was supposed to read the last record of log file of the preceding process to tell whether to calculation of the current time step every 100 millisecond.

The start and stop of hydrological simulation was completed through a PBS task submitting to the job manager of HPC cluster by the scheduling system, namely dispatcher. When detecting the end time of system or receiving a request from the Web Service API, the scheduling system would signal the HPC job manager to kill all tasks or the specified job, respectively. The time loop of real-time simulation was realized by appending control command to control files, when it was time to compute a new step or a new data file of rainfall was synchronized to the HPC cluster. Thus, the difficult of communication between huge massive processes was conquered.

The national level flash flood forecast and warning system of China had thousands of processes, including resident processes and dynamic processes. The rainfall model, soil moisture model and hydrological model were continuous and resident processes, which were lunched on preassigned computing nodes. And the other processes, i.e. the processes of hydrodynamic model and flash flood warning analysis model was dynamic, which were lunched dynamically according to the use of nodes.

The system could switch automatically from different time scale models according to the detecting or forecast rainfall. Different rainfall events with different rainfall amount were driven by different time step (daily, hourly, or minute time step) hydrological model. The rainfall thresholds for different time step models varied from the different climate regions. In addition, those thresholds were obtained through hydrological experience of local region.

Regarding to rainfall delay and process debug, some features like delay processing function and restart framework were designed in this system. In real-time simulation, the delay processing procedure read the delay data and modified the internal state data and the actual output. By this function, the rainfall delay problem was solved effectively. The restart framework was designed to persist the internal calculation status to the file system. When the simulation program unexpectedly collapses, the scheduling system would try to restart the calculation program after detecting through the HPC cluster



**Figure 4:** Structure of HPC clusters

interface. And the simulation program could recover to the last correct status before crash from the persistence status file, then continues to run.

In addition, the scheduling system would check the status of each computing nodes in the HPC clusters in real time. If a node crash detected, the jobs on that crashed node would be migrated seamlessly to the other backup node. As a result, the robustness of the system was enhanced.

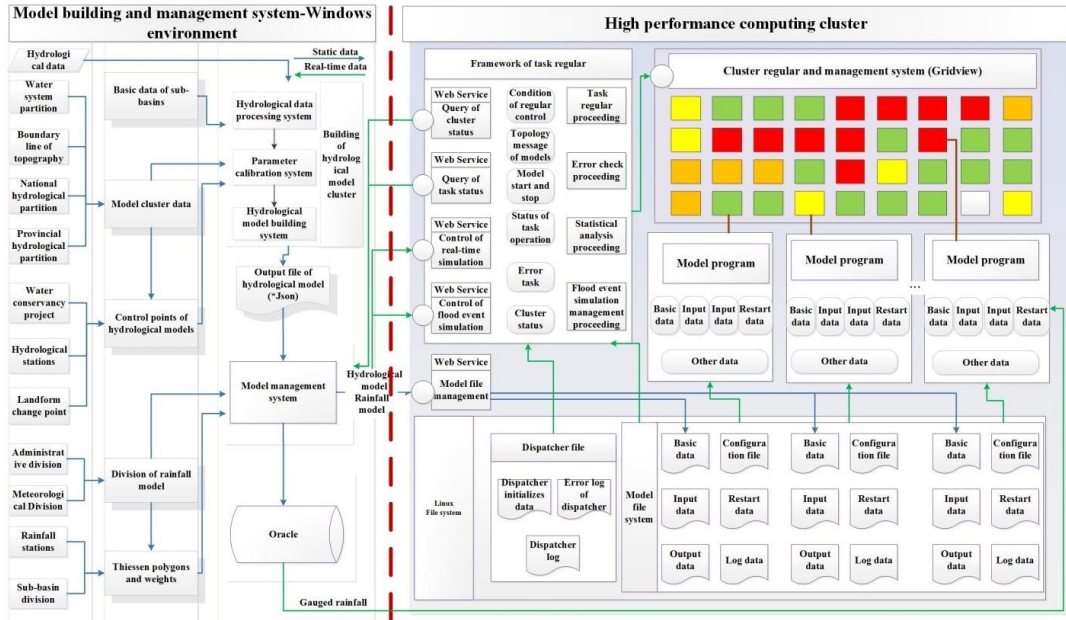


Figure 5: Structure of scheduling system

### 3 Study areas

To test the flood forecasting and warning platform, the flood event of June 30, 2017 in Weishui catchment, which was occurred in Ningxiang, Hunan Province, China was taken as study area (Figure 6). Weishui catchment is 2750 km<sup>2</sup> with river length of 144km. The catchment is located in humid continental monsoon climate region with the characteristic of abundant rainfall, the annual average rainfall was 1358 mm. Most area of this catchment is mountain, flash floods occurred frequently in Ningxiang. In addition, the catchment has 38 rainfall stations and 9 hydrological stations.

### 4 Result and Discussion

Two hydrological models (CNFF-HM) were built in Weishui catchment of Ningxiang for flood forecasting, and the model boundary was shown in Figure 6. The parallel flood forecasting and warning platform continuously forecasted the flood at hourly time step from June 30, 2017, when the monitoring rainfall exceeded 10 mm during 24 h. Actually, the accumulated rainfall was 160.3 mm from June 30, 2017 to July 2, 2017. The flood forecast results of 5:00, 6:00, 8:00, and 9:00 in July 1, 2017 and 8:00 in July 2, 2017 were shown in Figure 7. When the rainfall amount is not large enough, the forecasted flood of the catchment outlet was corresponding small (forecast result of 5:00 and 6:00 in July 1, 2017),

the forecast flood peak discharge was below the alarm discharge. As the rainfall became larger, the forecast flood peak discharge was increasing, and the forecast peak discharge exceeded the alarm discharge, the flood warning message was issued by the platform. Furthermore, all the forecast flood peak time was around 12:00 in July 1, 2017, this time was consistent with the observed flood peak time. All those simulation result indicated the parallel flood forecasting and warning platform could forecast and warn flood timely and correctly, and the HPC clusters for model calculation were effectively.

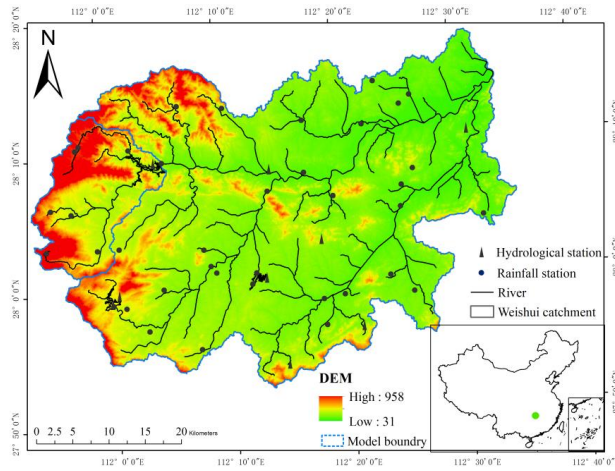


Figure 6: Location and hydrological stations of Weishui

### 5 Conclusion

China is a vast country across different hydro-meteorological region, so the construction of flood early warning system is a challenging work. A parallel flood forecasting and warning platform based on HPC clusters was established for flood forecasting and warning, especially for flash flood across China. This platform was characteristic by rapid communicate, resource effective management, automatic switch among different processes, and reboot framework. In addition, the platform established in China has effectively forecasted and warned the flood in Weishui catchment of Ningxiang in July 1, 2017. The result showed the usability and robustness of the platform for flood forecast and warning.

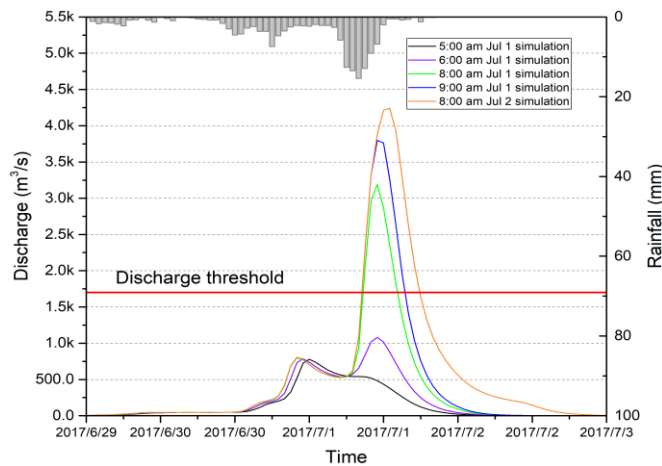


Figure 7: Forecasting results of Weishui catchment

## Acknowledgments

This paper is supported by National key research and development program (Grant No. 2017YFB0203104) and the National Natural Science Foundation of China (Grant No. 51579131).

## Reference

- Bürger, C. M., Kollet, S., Schumacher, J., & Bösel, D. (2012, 12 48). Introduction of a web service for cloud computing with the integrated hydrologic simulation platform ParFlow. *Computers & Geosciences*, pp. 334-336.
- David, O., II, J. A., Lloyd, W. J., Green, T. R., Rojas, K., Leavesley, G. H., & Ahuja, L. (2013, 1). A software engineering perspective on environmental modeling framework design: the object modeling system. *Environmental Modelling & Software*, 39, pp. 201-213.
- Gregersen, J., Gijbbers, P., & Westen, S. (2007, 9). OpenMI: open modelling interface. *Journal of Hydroinformatics*, pp. 175-191.
- Liu, R., Wei, J., Ren, Y., Liu, Q., Wang, G., Shao, S., & Tang, S. (2017, 19 6). HydroMP—a computing platform for hydrodynamic simulation based on cloud computing. *Journal of Hydroinformatics*, 19.
- Liu, R., Wei, J., Weng, Y., Wang, G., & Tang, S. (2014, 5). A cloud computing based platform for hydraulic modeling and simulation service. *Journal of Tsinghua University*, pp. 575-583.
- Maidment, D. R. (1993). Developing a spatially distributed unit hydrograph by using GIS. (p. 181). IWRH Publication.
- Maidment, D. R., Olivera, F., Calver, A., Eatherall, A., & Fraczek, W. (1996, 6). Unit hydrograph derived from a spatially distributed velocity field. *Hydrological Processes*, pp. 831-844.
- Rahman, J. M., Seaton, S. P., & Cuddy, S. M. (2004, 3). Making frameworks more useable: using model introspection and metadata to develop model processing tools. *Environmental Modelling & Software*, pp. 275-284.
- Shang, Y., Guo, Y., Shang, L., Ye, Y., Liu, R., & Wang, G. (2016, 18 5). Processing conversion and parallel control platform: a parallel approach to serial hydrodynamic simulators for complex hydrodynamic simulations. *Journal of Hydroinformatics*, pp. 851-866.
- Song, X., Kong, F., & Zhu, Z. (2011, 4). Application of Muskingum routing method with variable parameters in ungauged basin. *Water Science and Engineering*, pp. 1-12.
- Wang, Y., Liu, R., Guo, L., Tian, J., Zhang, X., Ding, L., . . . Shang, Y. (2017, 9). Forecasting and Providing Warning of Flash floods for Ungauged Mountainous Areas Based on a Distributed Hydrological Model. *Water*, p. 776.
- Welsh, W. D., Vaze, J., Dutta, D., Rassam, D., Rahman, J. M., Jolly, I. D., . . . Lerat, J. (2013, 1). An integrated modelling framework for regulated river systems. *Environmental Modelling & Software*, pp. 81-102.
- Winchell, M., Gupta, H. V., & Sorooshian, S. (1998). On the simulation of infiltration-and saturation-excess runoff using radar-based rainfall estimates: Effects of algorithm uncertainty and pixel aggregation. *Water Resources Research*, 34, pp. 2655-2670.
- Yang, D., Herath, S., & Musiakke, K. (1998, 2). Development of a geomorphology-based hydrological model for large catchments. *Proceedings of Hydraulic Engineering*, (pp. 169-174).
- Zhao, R. (1992). The Xinanjiang model applied in China. *Journal of China Hydrology*, 135, pp. 371-381.