SWAT Hydrological Model and Big Data Techniques

Xi Guo(✉), Biao Wang, Wei Xiong, and Shisheng Jin

GuiZhou Provincial Meteorological Administration Information Center, GuiYang, GuiZhou, China
guoxiguoxi1989@163.com

Abstract. Meteorological and hydrological data includes a wide range of items with complex format and large scale. In the era of big data, it brings great opportunities and challenges to meteorology and hydrology services. Because isolated information resources and industry data barriers bring incomplete data parameters, which makes the hydrological model simulate inaccurately. This paper mainly discusses the application of big data technology in Hydro-Meteorological industry. Firstly, it introduces the background and principal of hydrological SWAT model. Secondly, this paper proceeds the SWAT simulation and estimates runoff prediction of WangMo river in GuiZhou province, and analyzes simulation results. Finally, it proposes a big data platform architecture design combines with SWAT hydrological model as future research direction. Big data platform will provide libraries of integrated model, method, component, knowledge database for Hydro-Meteorological resources management. It also offers decision-making for flood control, water shortage, water pollution incidents.

Keywords: Hydro-meteorological model · Big data · SWAT model Forecasting model

1 Introduction

Hydrological model is one of the hot spots and important branches in research of hydrology. It is an important tool of studying the hydrology natural law and solving relative practical problems. At present, a new generation of distributed hydrological model has greatly broaden simulation fields, which changes water yield vary simulation in single way to water ecological diversified situation. The rapid development technologies of computer, remote sensing, and geographic information system has greatly promoted hydrological model. However, the distributed hydrological model is not mature enough. It often performs problems in following aspects: parameter estimation and model checking of hydrological

This project has been funded with support by 2017 Youth Technology Foundation of GuiZhou Provincial Meteorological Administration.

K. Long et al. (Eds.): 5GWN 2017, LNICST 211, pp. 79–88, 2018.
https://doi.org/10.1007/978-3-319-72823-0_9

system complexity, different space-time scales of hydrological and spatial parameters variability, different time scales of water cycling dynamic mechanism, and the limitations of data, etc.

At present, facing water issues in the rapid development of society, distributed hydrological model is an effective tool to explore complex hydrological process mechanism for water problems. It becomes an indispensable means for hydrological cycle research, and this will become opportunities and challenges for vast number of Hydro-Meteorological scientists.

2 Hydrological Model SWAT with Big Data Techniques

2.1 Background of SWAT Model

SWAT is a watershed scale model developed by the US Department of Agricultural Research Center USDA-ARS. The purpose of the model is to predict the impact of long-term land management influences on water, sediment and agricultural pollutants in complex watersheds with diversity of soil, land use and management conditions [1]. This section introduces the basic structure of SWAT model, hydrological cycle of land stage, and analyzes the characteristics of model. Finally, in Sect. 3, we simulated WangMo river of climate change on hydrological cycle influenced by SWAT model and put forward the prospects of hydrological model future work in big data field.

2.2 Hydrological Cycle of Land Stage

Water balance is a very important process in SWAT basin simulation. The hydrological simulation of basin can be divided into two main parts: the first part is about land stage hydrological cycle, water quantity control, sediment and nutrients, pesticides; the second part is river hydrological cycle calculation stage, which can be defined as the movement of water and sediment until the export process [2]. Hydrological cycle of SWAT model is mainly composed of following components: climate, hydrology, sediment, crop growth, soil temperature, nutrients, pesticides and agricultural management. The hydrological cycle water balance equation is as follow:

$$SW_t = SW_0 + \sum_{i=1}^{t}(R_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw}) \tag{1}$$

SW_t is final moisture content of soil (mm), SW_0 is soil moisture content at earlier stage (mm), t is step (day), R_{day} is rainfall for the i day (mm), Q_{surf} is surface runoff for the i day (mm), E_a is the evaporation for the i day, W_{seep} is infiltrating and side-flow for the i day of soil bottom section (mm), Q_{gw} is water flow for the i day (mm) [3].

1. Weather and Climate: Watershed climate provides moisture and energy inputs, which controls the water balance. The climatic variables required by SWAT are daily precipitation, maximum/minimum temperature, solar radiation, wind speed and relative humidity. The model reads measured data, and generate climate automatically by weather generator.
2. Hydrological Process: During the precipitation process, it may be intercepted in vegetation canopy or directly drop to soil surface. Soil surface moisture will infiltrate into the soil or generate profile runoff, which moves relatively fast and causes short-term river response when it runs into river. Infiltration of water can be retained in soil, and then be evaporated, or move slowly to surface water system. The physical processes involved include: canopy storage, infiltration, redistribution, evapotranspiration, side runoff, surface runoff, and return flow.
3. Erosion: Calculate the erosion of each HRU and sediment by soil loss equation (MUSLE). MUSLE uses runoff to simulate erosion and sediment yield, which can improve the prediction accuracy of model, and estimate the single storm sediment yield. The hydrological model supports net runoff and peak flow rate, which can be used to calculate runoff erosivity combined with subbasin. Erosion produced by rainfall runoff is calculated by Modified Universal Soil Loss Equation (MUSLE), and the formula is below:

$$Y = 11.8(Q \times pr)^{0.56} K_{USLE} \times C_{USLE} \times P_{USLE} \times LS_{USLE} \tag{2}$$

   Y is quantity of soil erosion (t), Q is surface runoff (mm), pr is peak runoff (m^3/s), K_{USLE} is soil erosion, C_{USLE} is vegetation cover and crop management factor, P_{USLE} is conservation measure factor, LS_{USLE} is terrain factor.
4. Nutrients and Pesticides: SWAT model can track the migration and transformation of nitrogen and phosphorus in several forms. Nutrients run into the river by surface runoff and interflow, and be transported to river downstream. SWAT model simulates the surface runoff carrying pesticides into river, through the leakage into the soil profile and aquifer.
5. Agriculture Management: SWAT model can define in each HRU about the start date of growing season, the time and amount of fertilization, the use of pesticides and irrigation of farming schedule according to the management measures. Beyond these basic management measures, model includes grazing, automatic fertilization and irrigation, as well as every possible water management options. The latest improvement in land management is the integration of sediment and nutrient loads from urban runoff [4].

2.3 Model Structure

Once the SWAT model determines the main river water, sediment, nutrients and pesticides load, river network load is calculated by HYMO command structure. In order to track the material flow in the river, SWAT model simulates the chemical conversion of the river, and computation involves: surface runoff, soil water, groundwater, river, and the confluence of the water storage. The model structure diagram is in Fig. 1. SWAT model has various parameters, mainly include: DEM, spatial distribution of hydro-meteorological station data, soil type and land use, series of evaporation and river data, etc. The list of input parameters could be referred in Table 1.

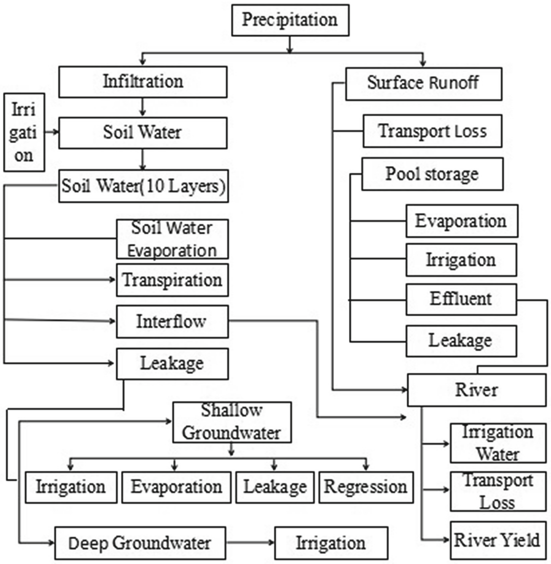

Fig. 1. SWAT model structure

3 WangMo River Simulation on SWAT Model

WangMo river is located in the middle of WangMo county, GuiZhou province, which is slope zone transition of Yunnan-Guizhou plateau to Guangxi hilly, and basin is consisted of karst and normal landform. Annual rainfall is 1249.3 mm, flood season happens concentratively between April and October, and rainy season accounted for 83.14% of the whole year. This river is one of the main branches of BeiPan river, and rises in DaYi village of WangMo county north, run into BeiPan river in WangMo county boundary. The study area is the main stream of river basin, the coordinate is 106°2′–106°12′E, 25°9′–25°23′N. WangMo river is 74 km long, drop height 1050 m, basin area 554 km^2. Main tributaries includes: NaBa river, NaGuo river, NaChao river, SongLin river. River basin belongs to the seasonal mountain river, and flood varies along with the rainfall. Since year of 70s, there has been occurred dozens of severe floods, collapse, landslide, debris flow and other geological disasters. At present, there are few research about WangMo river basin runoff. In this section, we use SWAT model to study various factors on the WangMo basin runoff, which can be based on the analysis of climate change, land use, and land cover.

Table 1. SWAT model input datasets

Type	Shape	Parameter	Source
Digital elevation data	Raster	Elevation, Slope length, River slope direction	STRM3 arc-seconds
Land use data	GRID	Leaf area index, Vegetation root depth, Runoff curve number, Pipe layer height, Manning coefficient	GLC2000 Lucc 1 km
Soil type	GRID	Density, Saturated water conductivity, Water retention rate, Particle content, Root depth	HWSD
Meteorological data	DBF table	Max–Min temperature, Daily precipitation, Relative humidity, Solar radiation and Wind speed	Meteorological CMADS
Hydrological data	DBF table	Daily flow, Monthly flow, etc.	Hydrological site data
Land management info.	x	Farming methods, Vegetation type, Irrigation methods, Fertilization time and quantity	Investigation of statistics

SWAT is a data driven model and it requires several types of data, which are listed in Table 1. Digital Elevation Data (DEM) were obtained from Shuttle Radar Topography Mission (STRM3), NASA and NIMA, U.S. Land cover datasets were adopted from GLC2000 LUCC 1 km CHINA, and China area datasets is cut from the global data [5]. Soil datasets were from Harmonized World Soil Database (HWSD). Weather datasets were collected from China Meteorological Assimilation Driving Datasets for SWAT model (CMADS) [5]. Different processes have been carried out when databases were established. Considering hydrological behavior of the basin with losing productive soil and water as runoff problems, the study took with SWAT2012 integrated with Arc GIS10.3.1 to evaluate the surface runoff watershed of WangMo river [6]. SWAT model is physically based and computationally efficient, uses readily available inputs and enables users to study long-term impacts. To use the model estimates WangMo river basin runoff, firstly, we need to setup a new SWAT project and load the existed WangMo river DEM from disk and define the river network. Figure 2 shows the processes of running SWAT model. The key procedures of modeling is as follow:

- Load or select the SWAT extension checking the relative box in the BASINS Extension Manager (Models category). Add three types of nodes: the rivers export, entry and source point. Generally speaking, one natural basin only has one export.
- Delineate the watershed and define the Hydrological Response Units (HRU). We obtained watershed map when calculated the sub-basin and activated reservoir button.

– Edit SWAT databases and make database index (data including: soil, slope, land use, weather, etc.). Model analyzes land use, soil coverage and HRU and make them overlay together. This is the most complex step, because when establish database, you need calculate lots of parameters by equations.
– Define the weather data with weather generator (data including: relative humidity, solar radiation, and wind speed).
– Set up and run SWAT simulation, after then, apply the default input files writer.
– Analyze, plot and graph SWAT output (SWATOutput.mdb) with Origin.

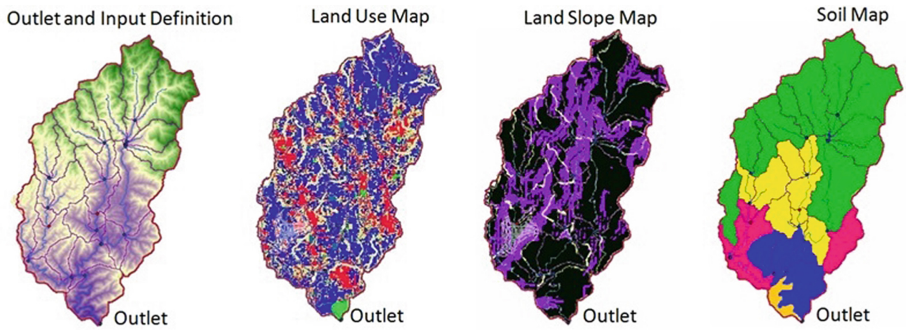

Fig. 2. SWAT model simulation of WangMo river

3.1 Results and Discussion

From the simulation of SWAT model, we obtain relationship between precipitation and runoff in the watershed of WangMo river, which could be referred to Fig. 3, and the simulation period is the average of each month from year 2008 to 2016. Firstly, it reflects the effect of rainfall intensity. When a rainstorm occurs in the watershed, it generates large runoff, which makes the runoff appear large jump in result graph. However, in the period of drought, due to the lack of rainfall, runoff decreases [7,8]. In the case of heavy rainstorm, soil water conservation is difficult to play an effective role. At present, due to human activities, it causes serious damage to the environment, which also makes the relationship between runoff and rainfall become complicated [9]. While upstream desertification phenomenon occurs, it makes groundwater become lower, and reduces runoff storage. The soil moisture content of various effects will make the runoff results different.

When analyzing the relation between precipitation and runoff by SWAT model, we supposed that temperature T is a constant, and then change the value of precipitation. We found that there was a positive correlation between precipitation and runoff and the result is shown in Fig. 4. When precipitation P1 increased by 40%, the runoff increased 115.43 mm, and increased by 71.2% on

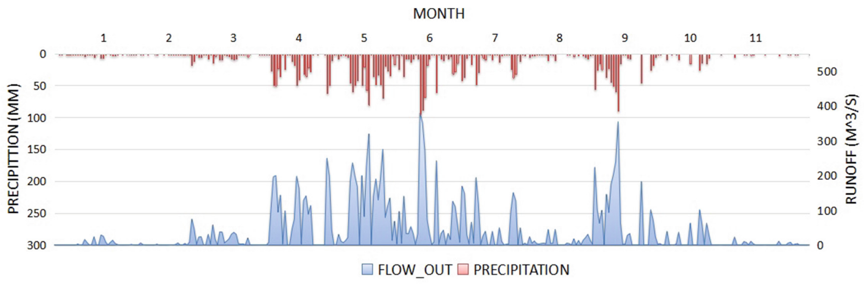

Fig. 3. SWAT model results

the basis of the original runoff; When P2 increased by 20%, the runoff increased 61.42 mm, and increased by 35.23% on the basis of the original runoff; When the P3 was reduced by 15%, the runoff was reduced 41.55 mm, which was reduced by 25.24% on the basis of the original runoff; When the P4 was reduced by 30%, the runoff was reduced 79.48 mm, which was reduced by 54.87% on the basis of the original runoff.

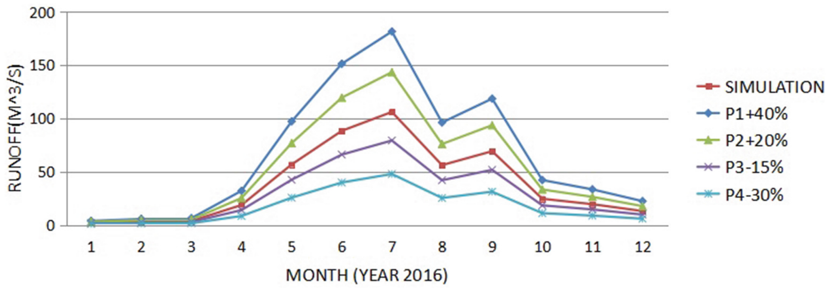

Fig. 4. Runoff variation with precipitation

3.2 Big Data with Hydrological Models

The arrival of big data era has brought hydro-meteorological information opportunities and challenges. This section introduces how to enhance the theoretical basis of hydrological model with big data support, broaden model's applicability, and propose a framework of intelligent real-time hydrological model of big data driven; discuss on the aspects that we can spare effort to study hydrological model forecasting precision of big data techniques. More and more people understand and pay close attention to hydro-meteorological model with big data, which will be widely used in society [10,11]. Big data needs new processing mode to have a strong decision-making capability, process optimization of massive, high growth rate and diversification of datasets. In other words, if we take big data as an industry, the key of achieving profitability such industry is to improve the data processing capacity by making value-added data [12,13].

At present, water resource data center mainly uses relational database to manage structured data and implement geographic spatial data extending; organizing semi-structured or unstructured data by relational database and file storage directory management. For water conservancy data storage and application requirements, the existing data storage structure has a bottleneck on aspects of data processing and analyzing [14]. Therefore, this section puts forward an idea about the structure of hydrological model data center based on big data platform, including data collection, data storage, data processing and analysis, and data application of four layers, which can be referred in Fig. 5.

Hydrological forecasting plays a very important role in flood control and disaster reduction. Due to the climate change and human activities, the basic law of hydrology has changed a lot, and those past hydrological models can not meet the needs of sustainable development of social economy. The hydrological runoff monitoring, rainfall, DEM and soil vegetation, remote sensing images bring increasing rich elements of big data. Research of new hydrology forecast technology driven by big data has become an inevitable way to solve above problems.

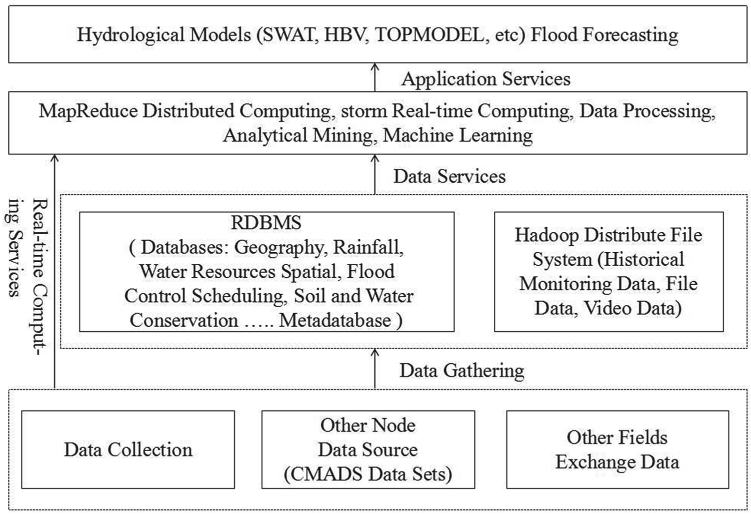

Fig. 5. SWAT model on big data platform

In the background of big data, we can conduct the thorough research from the following several aspects about hydrological law in big data mining, runoff theory method, and the cloud stage of real-time efficient forecasting technology.

– Research on flood characteristic pattern and parameter geography big data mining method. This could study from the characteristics of storm flood and construct a typical watershed model library.
– Research on precise hydrological model and big data driven intelligent real-time forecast model, flood forecast integration and real-time correction method, which is based on water storage capacity curve and terrain index of semi distributed dynamic combination in order to improve the prediction accuracy of semi-humid and semi-arid.
– Study on redistribution of flow along the side and water unit exchange between grid adaptive distributed simulation technology and precise prediction of watershed, which could achieve arbitrary grid unit flow and water level, improve the accuracy of real-time flood forecasting in complex basin.
– Research of water conservancy satellite remote sensing monitoring system and hydrological forecasting cloud platform, in order to achieve an integrate multi system running environment. Multi particle splitting, loose coupling and service collaboration technology could achieve hydro-meteorological flood forecast transferring, which can improve the efficiency, expand forecast period, and improve the accuracy of hydrological forecasting.

Flood forecasting plays a very important role in flood control and disaster reduction. Due to climate change, human activities, social economy development, flood forecasting encounters new problems and challenges. Study on the development of runoff theory, and create model of complex river basin and high accuracy flood forecasting, which has become the key technical points of national flood control. In the future, we can carry out the research and application of flood forecasting methods base on the data mining and big data techniques according to researches aspects above.

4 Conclusion

To develop a suitable model for the hydrological process for a river basin is the most important aspect for water resource management. SWAT hydrological model is applied to WangMo, GuiZhou watershed in order to assess the relationship between runoff and precipitation. Input datasets DEM, STRM3, HWSD, GLC2000, CMADS are quite applicable to run the SWAT model for the basin. The performance and applicability of SWAT model was successfully evaluated through model calibration. Average annual prediction of stream flow is 95.54 mm. As this model has a broaden functions, the model can be utilized not only as a potential tool for water resource management, but also help assess different land management and evaluate the effect of climate change on soil erosion.

Big data technology promotes water conservancy data acquisition, management and application of the rapid development. This paper also proposes a big data technology data center architecture of hydrological model, and to demonstrate how it can improve the management. And MapReduce series big data technology break through traditional data analysis perspective in a completely different way, in order to build the flood forecasting model in a more accurate way.

References

1. Ning, J., Liu, G., Liu, Q., Xie, C.: The hydrologic response unit (HRU) on SWAT model. Adv. Water Sci. **23**(1) (2012)
2. Sun, R., Zhang, X.: The Basin runoff study based on SWAT model. Geogr. Sci. Dev. **30**(3), 28–32 (2010)
3. Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G.: The Soil and Water Assessment Tool Historical Development. Applications and Future Research Directions. Center for Agricultural and Rural Development Lowa State University, Working Paper 07–WP 443
4. Jayakrishnan, R., Srinivasan, R., Santhi, C., Arnold, J.G.: Advances in the application of the SWAT model for water resources management. Hydrol. Process. **19**, 749–762 (2005)
5. This data set is provided by Cold and Arid Regions Science Data Center at Lanzhou. http://westdc.westgis.ac.cn
6. Source for this dataset was Global Land Cover 2000 database. European Commission, Joint Research Centre (2003). http://www-gem.jrc.it/glc2000/
7. Xu, L., Luan, Y., Cheng, X., Cao, X., Chao, K., Gao, J., Jia, Y., Wang, S.: WCDMA data based LTE site selection scheme in LTE deployment. In: International Conference on Signal and Information Processing, Networking and Computers (ICSINC), Beijing, China, pp. 249–260. CRC Press, Taylor and Francis Group (2015)
8. Xu, L., Chen, Y., Chai, K. K., Luan, Y., Liu, D.: Cooperative mobility load balancing in relay cellular networks. In: IEEE International Conference on Communication in China (ICCC), Xi'an, China, pp. 141–146. IEEE Press (2013)
9. Xu, L., Cheng, X., Liu, Y., Chen, W., Luan, Y., Chao, K., Yuan, M., Xu, B.: Mobility load balancing aware radio resource allocation scheme for LTE-advanced cellular networks. In: IEEE International Conference on Communication Technology (ICCT), Hangzhou, China, pp. 806–812. IEEE Press (2015)
10. Cao, Y., Wang, N., Sun, Z., Cruickshank, H.: A reliable and efficient encounter-based routing framework for delay/disruption tolerant networks. IEEE Sens. J. **15**(7), 4004–4018 (2015)
11. Cao, Y., Sun, Z., Wang, N., Riaz, M., Cruickshank, H., Liu, X.: Geographic-based spray-and-relay (GSaR): an efficient routing scheme for DTNs. IEEE Trans. Veh. Technol. **64**(4), 1548–1564 (2015)
12. Rui, X.: Hydrology and big data. In: Advances in Science and Technology of Water Resources, vol. 36. College of Hydrology and Water Resources, Hohai University (2016)
13. Xu, L., Chen, Y., Gao, Y., Cuthbert, L.: A self-optimizing load balancing scheme for fixed relay cellular networks. In: IET International Conference on Communication Technology and Application (ICCTA), Beijing, China, pp. 306–311. IET Press (2011)
14. Xu, L., Luan, Y., Cheng, X., Xing, H., Liu, Y., Jiang, X., Chen, W., Chao, K.: Self-optimised joint traffic offloading in heterogeneous cellular networks. In: IEEE International Symposium on Communications and Information Technologies (ISCIT), Qingdao, China, pp. 263–267. IEEE Press (2016)