Application Research of Parallel Optimization Technology in Hydrological Model

Fang Liu, Shuai Zhou
College Of Science, Xi’an Shiyou University Xi’an 710065, China; 3315245898@qq.com

Abstract. Hydrological model parameters are generally considered to be a simplified representation that characterizes hydrologic processes. As hydrological models continue to deepen the application of hydrological processes in the basin, they face enormous calculations. Meanwhile, in pursuit calibrating the model parameters by optimal algorithms for higher accuracy, the computation burden of optical techniques has become much heavier. Therefore, in order to solve this problem of low efficiency of hydrological model calculation, this paper uses parallel PSO algorithm to calibrate the TOPMODEL model parameters, and then uses parallel computing to process the flow generation in each sub-basin. The results show that the daily runoff simulation value of tangnaihai hydrological station fits well with the measured hydrological process; Whether PSO or sub-basin all can improve computational efficiency by using parallel optimization techniques, the former and the latter increased by 3.22 and 2.57 times, respectively. The results provide a reference for further understanding the application of parallel computing in hydrological models.

1. Introduction

Climate change and human activities have led to significant changes in the watershed hydrological cycle [1]. Hydrological models play a crucial role in simulating the hydrological process and understanding hydrological law of river basins, which has practical significance in the calculation of runoff yield and confluence, flood analysis and optimal allocation of water resources [2-3]. With the development of China’s water conservancy informatization, the demand for high efficiency hydrological model calculation in the industry is becoming more and more urgent. For example, flood forecasting, mountain flood warning and real-time calculation of hydrological models require correct decision-making in a short time, which puts forward higher requirements for the efficiency of hydrological models [4-8]. At present, in addition to the improvement of calculation methods, the performance improvement of computer hardware is also an important method to improve the computational efficiency of hydrological models.

In recent years, domestic and foreign scholars have conducted a lot of research on the improvement of computer hardware [9-12]. Liang et al., [13] use dynamic multi-swarm Particle Swarm Optimizer (DMS-PSO) to solve Multi-objective optimization problems with constraints. Campos et al., [14] use parallel multi-swarm PSO strategies to solve many objective optimization problems, the results show that parallelization has a positive effect on the convergence and diversity of the optimization process for problems with many objectives, but has a positive impact on the single goal. Wang et al., [15] analyze various works illustrating their different architectures, introducing their various parallel patterns for high speed computation. The results show that high parallelism can always be exploited in such applications for the development of high-performance systems. Chen et al., [16] proposes a novel
parallel mutation particle swarm optimization (PMPSO) algorithm based on a master-slave model. The results show that the parallel system based on PMPSO algorithm has a rapid convergence speed, which can effectively correct the wavefront aberration. Liu et al., [17] take Fully Sequential Dependent Hydrological Model (FSDHM) as an example, using a layered approach to parallel computing the runoff generation and confluence process of response unit. The results show that the parallel performance was higher for simulations with large datasets than with small datasets and the maximum speedup ratio reached 12.49. The above research results show that parallel computing can significantly improve computational efficiency. Meanwhile, parallel computing for specific hydrological models can greatly improve computational efficiency, but not all hydrological models can be used in parallel computing.

The TOPMODEL hydrological model takes the sub-basin corresponding to the catchment area as the response unit, and the sub-basin flow generation calculations are repeated many times, and the flow generation processes between the sub-basins are independent of each other. Considering these characteristics of the TOPMODEL hydrological model, parallel calculation just happens to solve this redundant processes for the runoff generation of sub-basins, thereby improve runoff simulation efficiency.

Given the above reasons, this paper builds a TOPMODEL model in the source region of the Yellow River. This paper selected years 2005, 2006–2009 and 2010–2012 as warm-up, calibration, and validation periods, respectively. On this basis, using parallel or non-parallel optimization techniques comparative analysis the computational efficiency of hydrological model, and quantitatively explored the influence of parallel optimization technology on runoff simulation of hydrological model. The results provide a reference for further understanding the application of parallel computing in hydrological models.

2. Study Area and data
The source region of the Yellow River (Fig.1) is located in the northeast Qinghai-Tibet Plateau between longitudes 95°50′E and 103°30′E and latitudes 30°30′E and 35°0′E, covering 12.19 × 10⁴ km² and occupying 16.2% of the entire Yellow River basin (75.24 × 10⁴ km²) [18-19]. The basin is dominated by alpine and semi-humid climates, and the temperature and precipitation are gradually decreasing from southeast to northwest, the annual average precipitation is between 320mm and 750mm, and mostly concentrated in June to October.

The inputs required for the TOPMODEL model include: a DEM (digital elevation model), topographic index data and meteorological data. The DEM is Shuttle radar topographic mission (SRTM) (90) DEM, which comes from the geospatial data cloud (http://www.gscloud.cn). The topographic index data were obtained by hydrological analysis of DEM data using Arcgis software. The meteorological data selected daily average precipitation of 12 meteorological stations in the basin from 2006 to 2012, and the potential evaporation data of each sub-basin surface is calculated by using the Thiessens polygon, and the hydrological data consistent with meteorological data in time.
3. Hydrological model and method

3.1. TOPMODEL model

The TOPMODEL model is a semi-distributed hydrological model proposed by Beven and Kirkby in 1979, which is the principle of variable source area based on topographic index, the model uses the soil moisture content to estimate the size and location of the source area [20]. The whole hydrological process is described by water balance and Darcy's law, and the model is simplified based on three important assumptions.

1. There is stable water supply saturated layer in the basin.
2. Soil hydraulic conductivity and soil moisture deficiency are exponentially decreasing.
3. The hydraulic gradient is approximately the same as the topography slope on the saturated area of the basin. By Darcy’s law, the subsurface runoff at any point can be expressed as:

\[ q_i = T_i \tan \beta_i = R_a \]  

Where \( q_i \) and \( T_i \) represent the unit width of soil interflow and the soil hydraulic conductivity at point \( i \), respectively, \( \tan \beta \) is the slope of the basin, \( R \) is the runoff per unit area, \( a_i \) is the unit width catchment area at point \( i \).

Coupling formula (1) and (2) can be obtained:

\[ z_i = -S_{sw} \ln \frac{R_a}{T_i \tan \beta_i} \]

(2)

\[ \bar{z} = \frac{1}{A} \int_{A} z_i dA = \frac{S_{sw}}{A} \int_{A} \left(-\ln \frac{R_a}{T_i \tan \beta_i}\right) \]

(3)

\[ z_i = \bar{z} - S_{sw} \left(\ln \frac{a_i}{\tan \beta_i} \lambda^*\right) \]

(4)

Where \( A \) is the drainage area, \( \bar{z} \) is the average surface depth of the basin, \( \ln \frac{a_i}{\tan \beta_i} \) is the topographic index at point \( i \), \( \lambda^* \) is the average topographic index.

3.2. Optimization algorithm

The hydrological model has nonlinearity and the input parameters are expressed by empirical functions.
The optimization algorithms have different differences in the optimization mechanism, previous studies have always used optimization algorithms to calibrate hydrological model parameters [21-22]. The basic process of the PSO algorithm: each solution of the optimization problem is called a particle, each particle flies at a certain speed in the n-dimensional search space, and the fitness function is used to measure the advantages and disadvantages of the particle. The well is based on its own flight experience. As well as the flight experience of other particles, the flight speed is dynamically adjusted to fly to the best particle position in the group, so that the optimization problem is optimally solved [23-24].

Assuming that the search space is d-dimensional and there are N_{pop} particles in the population, then the position of the particle i in the group is represented as a d-dimensional vector, the velocity of a particle is defined as the change in position, represented by the vector, and the velocity and position update of particle i can be obtained by the following formula.

\[ v_y(t+1) = \omega \cdot v_y(t) + c_1 r_1 \cdot (p_{best_y}(t) - x_y(t)) + c_2 r_2 \cdot (p_{best_y}(t) - x_y(t)) \]  \hspace{1cm} (5)
\[ x_y(t+1) = x_y(t) + v_y(t+1) \]  \hspace{1cm} (6)

Where \( t \) is the number of particle update iterations, \( \omega \) is the inertia coefficient; \( c_1 \) and \( c_2 \) are the acceleration coefficients; \( r_1 \) and \( r_2 \) are two independent random numbers that are uniformly distributed.

### 3.3. Hydrological model evaluation index

Nash Efficiency Coefficient (NSE), Correlation coefficient (\( R^2 \)) and Relative error (Re) are widely used indicators such as hydrological model, water resources assessment and hydrological frequency analysis [25-26]. Taking into account the advantages of the NSE indicator, this paper selects it as the objective function and uses other indicators to evaluate the model simulation results.

\[ NSE = 1 - \frac{\sum_{i=1}^{n} (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^{n} (Q_{obs,i} - \bar{Q}_{obs})^2} \]  \hspace{1cm} (7)
\[ R^2 = \frac{\sum_{i=1}^{n} [(Q_{sim,i} - \bar{Q}_{sim})(Q_{obs,i} - \bar{Q}_{obs})]^2}{\sum_{i=1}^{n} (Q_{obs,i} - \bar{Q}_{obs})^2 \sum_{i=1}^{n} (Q_{sim,i} - \bar{Q}_{sim})^2} \]  \hspace{1cm} (8)
\[ Re = \frac{\bar{Q}_{sim} - \bar{Q}_{obs}}{\bar{Q}_{obs}} \times 100\% \]  \hspace{1cm} (9)

Where \( Q_{sim,i} \) is the \( i \)th simulated discharge, \( Q_{obs,i} \) is the \( i \)th observed discharge, \( \bar{Q}_{obs} \) is the mean of the observed data, and \( n \) is the simulation period.

### 3.4. Parallel computing of PSO algorithm and sub-basin

This paper takes the source region of the Yellow River as the research object, based on the TOPMODEL hydrological model, the parallel or non-parallel computing of PSO algorithm is used to calibrate the model parameters, and then the results are compared and analyzed. On the basis of this, parallel optimization technology is used to calculate the sub-basin runoff, and thus improving the computational efficiency. The technical route of this paper see Fig.2.
4. Results and analysis

4.1. Calibration period and verification period results

Fig. 3 represents the runoff simulation results of the TOPMODEL model of the Tang Naihai hydrological station calibration and verification periods. The results show that the Tangnaihai hydrological station calibration and verification period, the daily runoff simulation value and the measured value of NSE and $R^2$ are both greater than 0.82, and the relative error $Re$ is less than 30%, this means that the runoff simulation results can meet the requirements. Particularly, the NSE of the calibration period is achieved 0.92, which is higher than 0.82 of the validation period, indicating that the accuracy of the calibration period runoff simulation is better than the validation period. In addition, the trend of daily runoff change is in good agreement with that of runoff change.

---

**Fig. 2 Technical roadmap**

---

**Fig. 3 Runoff simulation results for calibration and validation periods**
4.2. The results of parallel or non-parallel computation of PSO algorithm

4.2.1. Parallel computation of PSO algorithm

The introduction of parallel optimization technology will greatly improve the computational efficiency, this paper takes PSO algorithm as an example to analyze the performance of parallel computation of PSO algorithm for hydrological model parameter calibration. The algorithm is set as follows: Set the population size ($N_{pop}$) is 100, the maximum number of iterations is 50, and the test platform is MATLAB2018, which runs on 4-core PC with 4 GB memory and 2.70 GHz CPU speed.

Because the PSO algorithm is a random search algorithm, the search results are unstable, so the program runs independently five times. The results are shown in Fig.4. The results show that the PSO algorithm approaches convergence when the evolutionary algebra is around 45, and the evolution is not obvious in the later stage. Meanwhile, the optimization results of the objective function value ($NSE$) are basically consistent except for the second calculation, which is as high as 0.92.

![Fig.4 Five iterations parallel computation optimization process](image)

To further evaluate the accuracy of different times of runoff simulation, this paper draws the runoff simulation results on the Taylor diagram, which is shown in Fig.5. The results show that the root mean square error, standard deviation and correlation coefficient of the simulated and measured values of different simulation times are not obvious, and the root mean square error and standard deviation range are between 0.40-0.50, 1.0-1.10, respectively. Meanwhile, the fourth runoff simulation accuracy better than other times, and the second runoff simulation is the worst.

![Fig.5 Different times of runoff simulation accuracy Taylor diagram](image)

4.2.2. Comparative analysis of parallel or non-parallel computing results

Fig.6 is the results of time required for PSO parallel computing or non-parallel computing. It can be seen that the time required is significant difference between PSO parallel computing or not. Without PSO parallel computing, it takes about 14,000 seconds to run once, however, the time required to use parallel computing is only about 4,500 seconds. This means that the PSO parallel algorithm can significantly improve computational efficiency and save resources.
4.3. The results of parallel or non-parallel computation of sub-basin calculation

Fig. 6 is the results of time required for PSO parallel computing or non-parallel computing. The results indicate that sub-basin parallel or non-parallel computing requires less time difference for different times, however, the difference between parallel or non-parallel computing of sub-basin calculation.

In order to verify and analyze the parallel computing performance of the algorithm in the PSO and sub-basin framework, the speedup ratio of the algorithm needs to be calculated. The speedup ratio is the performance gain achieved by reducing the runtime by parallel computing, which is an important indicator for verifying the performance of parallel computing.

It can be seen from Tab.1 that the algorithm is very good speedup ratio after parallelization. Whether PSO or sub-basin all can improve computational efficiency by using parallel optimization techniques, the former and the latter increased by 3.22 and 2.57 times, respectively. Therefore, when calibrating hydrological model parameters or calculating runoff yield and concentration in sub-basins, parallel computing technology must be adopted to improve work efficiency and save resources.

| Calculation category | Speedup ratio |
|----------------------|---------------|
|                      | Minimum value | Maximum value | Average value |
| PSO                  | 2.94          | 3.7           | 3.22          |
| Sub-basin            | 2.27          | 2.76          | 2.57          |

5. Conclusion

This paper takes the Yellow River source region as an example, based on TOPMODEL hydrological model, using PSO parallel or non-parallel algorithm to calibrate hydrological model parameters. On
this basis, parallel optimization technology is used to calculate runoff yield and concentration in sub-basins. The main conclusions are as follows:

(1) The daily runoff simulation value of tangnaihai hydrological station fits well with the measured hydrological process. The $NSE$, correlation coefficient $R^2$ and relative error $Re$ of the model during the calibration and verification period all can meet the satisfactory result, and the correlation coefficients $NSE$ and $R^2$ are above 0.82, which indicates that TOPMODEL model can be well applied in the source period of the Yellow River.

(2) The time required is significant difference using the PSO parallel computing or not and sub-basin parallel computing or not. Whether PSO or sub-basin all can improve computational efficiency by using parallel optimization techniques, the former and the latter increased by 3.22 and 2.57 times, respectively.

(3) When calibrating hydrological model parameters or calculating runoff yield and concentration in sub-basins, parallel computing technology must be adopted to improve work efficiency and save resources.

Acknowledgments
This work was financially supported by the Applied Research Project of Public Welfare Technology of Zhejiang Province (2017C33013).

References
[1] IPCC. Climate Change 2013: The physical science basis[M]. Cambridge: Cambridge University Press, 2013.
[2] Joseph, J. F. and J.H.A. Guillaume. Using a parallelized MCMC algorithm in R to identify appropriate likelihood functions for SWAT [J]. Environmental Modelling & Software, 2013, 46, 292-298.
[3] Li Binquan, Liang Zhongmin, et al. Comparison of parameter uncertainty analysis techniques for a TOPMODEL application [J]. Stochastic Environmental Research & Risk Assessment, 2017, 31(5):1-15.
[4] Wang Hao, Fu Xudong, et al. A common parallel computing framework for modeling hydrological processes of river basins [J]. Parallel Computing, 2011, 37(6):302-315.
[5] Liu Junzhi, Zhu Axing, et al. Estimation of theoretical maximum speedup ratio for parallel computing of grid-based distributed hydrological models [J]. Computers & Geosciences, 2013, 60:58-62.
[6] Vivoni Enrique R, Mascaro Giuseppe, et al. Real-world hydrologic assessment of a fully-distributed hydrological model in a parallel computing environment [J]. Journal of Hydrology, 2011, 409(1-2):483-496.
[7] Rouholahnejad E , Abbaspour K.C., et al. A parallelization framework for calibration of hydrological models [J]. Environmental Modelling & Software, 2012, 31:28-36.
[8] Kurtz Wolfgang, Lapin Andrei, et al. Integrating hydrological modelling, data assimilation and cloud computing for real-time management of water resources [J]. Environmental Modelling & Software, 2017, 93:418-435.
[9] Ouyang Aijia, Tang Zhuo, et al. Parallel hybrid PSO with CUDA for 1D heat conduction equation [J]. Computers & Fluids, 2015, 110:198-210.
[10] Soufi Youcef, Bechouat Mohcene, et al. Fuzzy-PSO controller design for maximum power point tracking in photovoltaic system [J]. International Journal of Hydrogen Energy, 2017, 42:8680-8688.
[11] Gunasundari S , Janakiraman S, et al. Multiswarm Heterogeneous Binary PSO using Win-Win approach for improved Feature Selection in Liver and Kidney disease Diagnosis [J]. Computerized Medical Imaging and Graphics, 2018, 70:135-154.
[12] Kartheeswaran Subramanian, Christopher Durairaj Daniel Dharmaraj. A data-parallelism approach for PSO-ANN based medical image reconstruction on a multi-core system [J].
Informatics in Medicine Unlocked, 2017, 8:21-31.
[13] Liang J J, Pan Quanke, et al. Solving the blocking flow shop scheduling problem by a dynamic multi-swarm particle swarm optimizer [J]. International Journal of Advanced Manufacturing Technology, 2011, 55(5-8):755-762.
[14] Campos Arion de, Pozo Aurora T.R., et al. Parallel multi-swarm PSO strategies for solving many objective optimization problems [J]. Journal of Parallel and Distributed Computing, 2019, 126:13-33.
[15] Wang Tianyi, Qian Kemao. Parallel computing in experimental mechanics and optical measurement: A review (II) [J]. Optics and Lasers in Engineering, 2018, 104:181-191.
[16] Chen Ying, Feng Yong, et al. A parallel system for adaptive optics based on parallel mutation PSO algorithm [J]. Optik, 2014, 125(1):329-332.
[17] Liu Junzhi, Zhu Axing, et al. A layered approach to parallel computing for spatially distributed hydrological modeling [J]. Environmental Modelling & Software, 2014, 51:221-227.
[18] Meng Fanchong, Su Fengge, et al. Impacts of recent climate change on the hydrology in the source region of the Yellow River basin [J]. Journal of Hydrology, 2018, 558, 301–313.
[19] Wang Taihua, Yang Hanbo, et al. Quantifying the streamflow response to frozen ground degradation in the source region of the Yellow River within the Budyko framework [J]. Journal of Hydrology, 2018, 558, 301–313.
[20] Beven Keith. On the generalized kinematic routing method [J]. Water Resources Research, 1979, 15(5):37-44.
[21] Yen Haw, Hoque Yamen, et al. The impact of considering uncertainty in measured calibration/validation data during auto-calibration of hydrologic and water quality models [J]. Stochastic Environmental Research and Risk Assessment, 2015, 29(7):1891-1901.
[22] Xie Shumping, Du Jinkang, et al. A progressive segmented optimization algorithm for calibrating time-variant parameters of the Snowmelt Runoff Model (SRM) [J]. Journal of Hydrology, 2018, 566:470-483.
[23] R.C. Eberhart, J. Kennedy. A new optimizer using particle swarm theory, in: Proceedings Sixth Symposium on Micro Machine and Human Science, Nagoya, Japan, 1995, pp. 39–43.
[24] Bai Tao, Kan Yanbin, et al. Fusing feasible search space into PSO for multi-objective cascade reservoir optimization [J]. Applied Soft Computing, 2017, 51:328-340.
[25] Chang Jianxia, Wang Yimin, et al. Impact of climate change and human activities on runoff in the Weihe River Basin, China [J]. Quaternary International, 2015, (380-381):169-179.
[26] Huang Shenzhi, Chang Jianxia, et al. Monthly streamflow prediction using modified EMD-based support vector machine [J]. Journal of Hydrology, 2014, 511:764-775.