Research on Watershed Extraction Method Based on GIS

Yanan Li1, 2, 3, *, Na Lei1, 2, 3, a and Yufei Xiong1, 2, 3, b

1Shaanxi Land Engineering Construction Group Co, Ltd, Xi’an, China
2Institute of Land Engineering and Technology, Shaanxi Land Engineering Construction Group Co, Ltd, Xi’an, China
3Key Laboratory of Degraded and Unused Land Consolidation Engineering, the Ministry of Natural Resources, Xi’an, China

*Corresponding author e-mail: 930157635@qq.com, a619648133@qq.com, b295340477@qq.com

Abstract. The Digital Elevation Model (DEM) mainly describes the ground elevation, which contains rich hydrological information and is one of the main data sources for current watershed analysis. This paper takes Fuxian County of Shaanxi Province as the research area, and uses the ASTER GDEM data of 30m resolution, based on the Hydrology module of ArcGIS10.3 platform, after the filling of the earth, the calculation of the flow direction, the calculation of the cumulant accumulation, the determination of the convergence threshold, the extraction of the river network, etc. The method of basin watershed extraction is studied and the characteristics of the basin are comprehensively analysed. The results show that when the convergence threshold is 11000, the extracted river network is closest to the actual river network, and the derivative relationship can be used to determine that the extracted river network reflects the threshold of the real water system

1. Introduction
The Digital Elevation Model (DEM) is a discrete expression of the topography of the Earth's surface. It contains rich hydrological features, topographical information, etc. It has become the main means of digital watershed extraction [1]. The digital elevation model generally has three structures: regular grid DEM, irregular triangulation DEM and contour DEM. The grid structure is easy to combine with remote sensing and grid GIS, and various topographical factors are extracted on the grid DEM. The easiest, so the grid DEM dominates the digital ground model data organization [2]. For example, DEM data with a resolution of 30m and 90m provided by the US Geological Survey (USGS) is widely used in the research of geomorphology, soil, hydrology and ecology in the region. Using DEM data, terrain factors can be extracted, and hydrological analysis and watershed information research can be performed. At this stage, the widely used tools for automatic extraction of watershed information are mainly TOPAZ, River Tools and ArcGIS [3], among which ESRI and the ArcGIS hydrological module developed by the Water Resources Research Center of the University of Austin, Texas, USA, are used for river network extraction and analysis functions. Powerful, can quickly extract watershed information, the use of this module at home and abroad has been a lot of research on the extraction of watershed characteristics of the study of surface waters [4-6].
In this study, the Hydrology module provided by the ArcGIS 10.3 platform was used to utilize the DEM data of 30m resolution in Fu County, Yan'an, to pre-process the DEM data of the study area, determine the flow direction, calculate the cumulative amount of the confluence, determine the confluence threshold, and generate the river network. The processing process and related algorithms of watershed hydrological feature information extraction are studied.

2. Research area overview and data source

2.1. Research area overview
Located in the northern part of Shaanxi Province, Fuxian County in the south of Yan'an City, at the junction of Shaanxi, Gansu and Ningxia provinces. The area is located in the transition zone between the Gaosonggou and the hilly and gully regions in northern Shaanxi. The whole area is 111 km long from east to west, 73.8 km wide from north to south, and the area is about 4182 km². The county has a vast county, rich resources, four distinct seasons, sufficient sunshine and moderate precipitation. The topography of the area includes the river terraces dominated by the Luohe River and the Hulu River, the high-lying gully area in the central part, the hilly gully in the northern part of the area, and the low-lying mountainous areas in the east and west.

2.2. Data source
There are three main ways to obtain DEM data: topographic map digitization, field survey and aerial photogrammetry. Aerial photogrammetry facilitates quick and easy access to large-area data, which is conducive to maintaining the current status of data, as the main way to obtain DEM data at this stage. The DEM data used in this study is derived from the ASTER GDEM V2 data from the Geospatial Data Cloud (http://www.gscloud.cn/), which was produced by NASA based on Earth surface observation data acquired by the Earth observation satellite Terra. Completed, the data covers 99% of the world's range, with UTM/WGS 84 projection and a spatial resolution of 30 m. ASTER GDEM V1 data was first released in 2009, but due to the abnormality of the first version of data in some areas, mainly due to cloud coverage, boundary stacking and other reasons, V2 version is processed on the basis of V1 version, after cloud processing The residual abnormal value is removed, and the abnormal value is replaced by the average value, which effectively eliminates the abnormality of the previous version, and the V1 data is better corrected, and has better precision. Take N35-E108, N36-E108, N35-E109, and N36-E109 four-view images, and mosaic the DEM data in the study area.

3. Research method

3.1. ArcGIS river network extraction process
This paper mainly uses the Hydrology toolbox of the spatial analysis module in ArcGIS 10.3 for watershed extraction. ArcGIS Hydro Tools is a data model developed by ESRI and the Water Resources Research Center of the University of Austin, Texas, for water resources. The hydrological information extraction function can quickly extract a large amount of hydrological information and display it visually. The basic idea of this paper is to analyze and process the DEM data of the study area according to the DEM data and other data in the existing research area, and finally extract the river network information of the study area [7]. The main steps in this paper to extract the watershed include DEM preprocessing, flow direction analysis, confluence unit analysis, river network structure extraction, statistical watershed and river hydraulic eigenvalues, as shown in Figure1.
3.2. **Convergence threshold determination method**

The determination of the convergence threshold is the key to the extraction of river network information, which directly affects the degree of coincidence between the extracted river network and the actual river basin. The larger the catchment area threshold, the sparser the extracted water system; the smaller the threshold, the finer the extracted water system, and some rivers that do not exist may be extracted [8]. In this paper, the threshold will be determined through experiments. 1000-15000 is selected as the catchment area threshold to extract the water system and calculate the river network density. The second derivative of the river network density is used to obtain the threshold of the river network density and the real river network.

4. **Results and analysis**

4.1. **Fill and water flow direction extraction**

The depression refers to the terrain that is approximately closed lower than the surrounding ground. For the DEM data just obtained, the depression refers to the case where the elevation of a certain grid unit is lower than the elevation of the adjacent 8 grid units, often due to accuracy and data anomalies. It will produce false ground, which will reduce the accuracy of river network generation, thus affecting the
quality of small watershed extraction. In the ArcGIS 10.3, select the Fill tool in the Hydrology module to calculate the innocent DEM.

The direction of water flow refers to the direction when the water flows away from the grid. The single flow method and the multi-flow method are currently the main algorithms for calculating the direction of water flow. The single flow method is simple and convenient, and D8 [9] is the most widely used. The principle is to assume that the water flow in a grid flows out from only one direction and then determine the direction of the water flow based on the grid elevation. The water flow direction is extracted by selecting the Hydrogen module's Flow Direction tool in ArcGIS 10.1 using the DEM data after the depression.

4.2. Confluence cumulant calculation and confluence threshold setting

The cumulative amount of the flow is calculated according to the flow direction to obtain the amount of water flowing through the space. The larger the flow rate, the easier it is to form surface runoff. For each grid, the cumulative amount of confluence represents the amount of all unit water flowing into the grid. When the grid cumulant accumulation is greater than a certain threshold, the grid is considered to be located on the waterway and is defined as a river channel. The convergence threshold is the key to the extraction of the river network. For the same traffic accumulation grid, the larger the threshold, the smaller the river network density and the smaller the internal watershed. The setting of this value is subjective. Usually, several values can be set to test to determine the threshold for extracting a suitable river network. This paper will determine the catchment area threshold by experiment, select 1000, 2000, 3000, 1500 as the catchment area threshold to extract the water system and calculate the river network density, and obtain the river network density and Real River through the second derivative of river network density. The closest threshold for the net.

Based on the statistical tools in the ArcGIS 10.3 platform, the length of the river network under different thresholds is obtained, and then the river network density is obtained by calculating the ratio of the river length to the watershed area. The results are shown in Table 1.

| Threshold | River length/km | River network density/(km.km⁻²) |
|-----------|-----------------|---------------------------------|
| 1000      | 831.931         | 0.203                           |
| 2000      | 856.476         | 0.209                           |
| 3000      | 883.53          | 0.216                           |
| 4000      | 913.662         | 0.223                           |
| 5000      | 948.485         | 0.232                           |
| 6000      | 993.709         | 0.243                           |
| 7000      | 1044.791        | 0.255                           |
| 8000      | 1107.378        | 0.270                           |
| 9000      | 1188.949        | 0.290                           |
| 10000     | 1286.186        | 0.314                           |
| 11000     | 1410.011        | 0.344                           |
| 12000     | 1573.886        | 0.384                           |
| 13000     | 1800.833        | 0.440                           |
| 14000     | 2218.932        | 0.542                           |
| 15000     | 3136.772        | 0.766                           |
Figure 2. Curve relationship between river network density and convergence threshold

It can be seen from Fig. 2 that the density of the river network in the study area increases sharply with the increase of the threshold of the convergence, but it gradually becomes slower as the convergence threshold continues to increase. When the convergence threshold is greater than 10000, it tends to be gentle, and the river network is found. The relationship between the density and the sink threshold can be expressed by a power function, $y=22.942x^{-0.493}$, where the correlation coefficient $R^2$ is 0.9997 and the fitness is high.

Figure 3. Second-order derivative of river network density and convergence threshold

It can be seen from Fig. 3 that the convergence threshold has an inflection point in the range of 10000-12000, and the relationship between the river network density and the second-order derivative power function of the convergence threshold is, where the correlation $R^2$ is 1, and the river network and the actual water system are extracted when the convergence threshold is 11000. The most consistent, the results are shown in Figure 4. The method of determining the convergence threshold by obtaining the second derivative of the river network density and the convergence threshold avoids the cumbersome process of extracting the watershed information by DEM and improves the accuracy of the watershed extraction.
5. Conclusion and Outlook
Taking Fuxian County, Yan'an City, Shaanxi Province as the research area, the ArcGIS 10.3 hydrological analysis module was used to study the watershed information in the study area. The results show that when the convergence threshold is 11000 (the catchment threshold is 1.1km²), the river network extracted is the closest to the real water system in the study area, and the ideal results are obtained, which can provide data support for the ecological protection of the basin. Although the watershed extraction method in this paper is scientific and reasonable, it still has some shortcomings and needs to be improved in future research.

Firstly, the automatic extraction of watershed information using GIS and DEM data is current and convenient, and is widely applicable to hydrological analysis of each watershed. However, DEM data with different resolutions will have an impact on the extraction results of the watershed. In the future, multiple resolutions should be emphasized. DEM extracts river network precision comparison research. Secondly, with the increasing maturity of GIS technology related theories and methods, more and more applications in the field of hydrology are applied. However, due to the difference between processing methods and actual terrain, the accuracy of watershed extraction is also affected. A certain impact, improving accuracy is the focus of future research.

References
[1] Zhu Hailing, Yang Xiaohui, Zhang Xuepei, et al. Feature extraction and analysis of upstream watershed of Miyun Reservoir based on DEM [J]. China Soil and Water Conservation Science, 2013, 11 (03): 66 - 72.
[2] Tang G. 2000. A Research on the Accuracy of Digital Elevation Models. Beijing: Science Press.
[3] SONG Xiaomeng, ZHANG Jianyun, ZHAN Chesheng, et al. Research progress of digital watershed feature extraction based on DEM[J]. Progress in Geography, 2013, 32 (1): 31 – 40.
[4] ZHAO Jian, JIA Zhonghua, LUO Wei. Watershed feature extraction based on DEM in ArcGIS environment [J]. Journal of water resources and water engineering, 2006 (01): 74 - 76.
[5] Sun Qingyan, Yu Xinxiang, Hu Shuping, Xiao Yang, Li Jinhai, Wu Jun. Feature extraction and application of DEM watershed based on ArcGIS environment [J]. Journal of Beijing Forestry University, 2008, 30 (S2): 144 - 147.
[6] Jenson S K. Application of hydrologic information automatically extracted from digital elevation...
models [J]. Hydrological process, 1991 (5): 31 - 44.

[7] Tang Guoan, Yang Wei. ArcGIS Geographic Information System Spatial Analysis Experiment Course [M]. Beijing: Science Press, 2006: 478 - 492.

[8] Xu Huixi. Information Extraction of Mountain River Network in Western Sichuan Based on DEM [J]. Geospatial Information, 2015, 13 (04): 65 - 67+13.

[9] He Cancan. Research on digital watershed feature extraction and hydrological network construction based on DEM [D]. Chang'an University, 2017.