Design and Implementation of Drainage Network Extraction Algorithm Based on Binary Linear Regression

Dandan Su and Guangsheng Liu
Institute of Mineral Prognosis Based on Integrated Information, Jilin University, Changchun, China
Email: dandan198092@163.com, liugs@jlu.edu.cn

Guangda Gao
College of Information Engineering, China University of Geosciences, Beijing, China
Email: gaoguangda@sohu.com

Dandan Su, Jiawei Liu and Yingchen Meng
College of Urban and Environmental Sciences, Changchun Normal University, Changchun, China
Email: {dandan198092, liujiawei54, myc1989}@163.com

***Abstract*—The flow directions of the extracted drainage network are more random based on random flow model. Binary linear regression method is applied to calculate the residual value at every point, and then the normalized residuals are used to prune reasonably as the variables. The terrains are distinguished automatically according to the normalized residuals, and the local catchment area threshold is determined based on each terrain feature. When there are two terrains in the experimental area, source density is large on mountainous terrain, and it is small on flat terrain. Eventually, the extracted result is consistent with the actual drainage network.**

***Index Terms*—random flow model, binary linear regression, standardized residual, normalized residual**

I. Introduction

Along with the integration of computer technology and hydrologic science and geographic information science, the research and development of digital hydrology becomes a hotspot. The researchers have carried on a number of studies on this hotspot, and one of the research directions about digital hydrology is to extract drainage network with ARCGIS [1, 2, 3]. According to the relations between the catchment area threshold and the characteristic parameters of the drainage network, the reasonable catchment area threshold is determined [4,5,6,7], and the extracted drainage network is consistent with the actual drainage network in mountainous terrain. Based on the D8 algorithm, the drainage network is extracted with ARCGIS. Because the influence of the random factors on the drainage network is not considered fully, the extracted drainage is straight, and it includes the parallel branches. In order to make up for this deficiency, the researchers adopt Digital River and Lake Network Method or Trunk River Constraint Method to modify the grid flow directions on flat terrain [8-18]. Thus it can be seen that only one catchment area threshold is used to extract the drainage network, the catchment area threshold is used on mountainous terrain or flat terrain. If the drainage network is extracted on complex terrain, the different catchment area thresholds are needed [19]. In recent years, the researchers have applied binary linear regression method into hydrology. Yunshan Tao has applied iteration method into hydrology calculation in 1991 [20], and this method improves the accuracy of not only binary linear regression but also the result of storm intensity formula. Shuying Wang has improved the traditional least square method, and the improved method is applied to the hydrological analysis and calculation in 2003 [21]. Guoli Wang summarized linear regression analysis which is applied to hydrology in 2004 [22], and he summed up the superiority and the scope of all regression analysis methods. Xiangchun Kong designed the program of multiplication curve matching in 2005 [23], and the program is applied to hydrological analysis. Dongmei Cheng applied EXCEL into linear regression analysis of hydrology in 2012 [24]. In the study, binary linear regression method is applied to calculate the residuals, and the residuals are normalized. And then, according to the normalized residuals, the different terrains are distinguished, and the local catchment area thresholds are determined. Besides, based on random flow model, the randomness of the flow directions is reinforced.

II. Methodologies

A. Random Flow Model

Real flow directions may be arbitrary directions, and they are not limited by the eight directions of D8. In order

Manuscript received April 22, 2013; revised May 20, 2013; accepted June 5, 2013.
Project number: 1212011120445
Corresponding author: Dandan Su

doi:10.4304/jcp.8.9.2277-2283

to reflect the authenticity of the flow directions better, random flow model has been proposed[25]. There are three conditions that need to be met in random flow model. The first one is that random flow direction cannot go against nature. The second one is that “greater the flow speed down, more likely the flow direction is the direction of random flow model”. The third one is that the average of the flow directions in random flow model is the value of the real flow direction. According to three conditions above, the flow directions in random flow model are closer to the real flow directions, and random flow model is more suitable for extracting drainage network in flat terrain to avoid parallel rivers. Rho8 is set as a random variable, and its distribution function is shown in (1). The random flow model sketch is shown in Fig.1. A function is set at the eight neighboring points of point 0 , and it is shown as (2). z (i) is the elevation of the ith neighboring point of point 0 in (2). If the result computed by (2) is the maximum and it is greater than 0 at the ith point, the pointing direction to the ith point is the flow direction at point 0. If the results computed by (2) are the maximum and they are greater than 0 at more than one point, one of those points is randomly chosen as the flow direction at point 0. If the results computed by (2) are less than 0 at all of the eight neighboring points of point 0, they mean that water doesn’t move, or they mean that water moves from point 0 to point 0. As such, the flow directions computed by (2) meet the three conditions of random flow model. The arrow in Fig.1 means that the pointing direction to Point 6 is the flow direction at point 0, and it means that v (6) is equal or more than v(i)(i=1,2,3,4,5,7,8) and v(6) is more than 0.When point 0 is located on the borders or in the corners of the experimental area, some of the neighboring points of point 0 are not shown, and the flow direction at point 0 is computed by the shown neighboring points of point 0. In order to reduce the influence on the flow direction, when point 0 is on the border, some of the computed flow directions which are not parallel to the borders are valid, and others of the computed flow directions which are parallel to the borders are invalid.

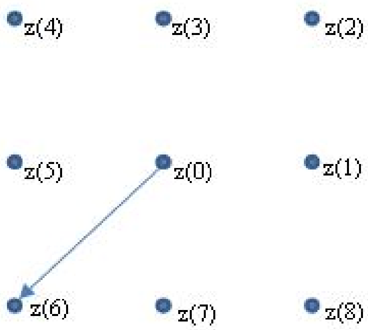

Figure 1. Random flow model sketch

$$P(Rho8 \le x) = \begin{cases} 0 & \text{for } x<0.5 \\ 2-\dfrac{1}{x} & \text{for} 0.5 \le x \le 1 \\ 1 & \text{for } x>1 \end{cases} \tag{1}$$

$$v(i) = \begin{cases} z(0)-z(i) & i=1,3,5,7 \\ (z(0)-z(i)) \times Rho8 & i=2,4,6,8 \end{cases} \tag{2}$$

There are four steps to extract drainage network basing on random flow model. The first step is to compute the flow direction at every point in the experimental area by (1) and (2), and these points are connected as some dendrograms. The second step is to connect some branches to other branches by the flood algorithm, and the branches become bigger. The flood algorithm is to fill the small depressions with water, and water continues to flow down. The third step is to cut off some branches by a pruning algorithm, and the remained branches are the extracted drainage network. The forth step is to smooth the remained branches in order that the extracted drainage network is consistent with the final drainage network. Drainage network extraction algorithm based on binary linear regression is applied in the third step to optimize the pruning process.

B. Binary Linear Regression Residual Calculation and Its Normalization

This is the binary linear regression residual calculation. Equation 3 is Binary linear regression equation [26]. The coefficients which are $\hat{a}$ 、 $\hat{b}_1$ and $\hat{b}_2$ in (3) are computed by (4), (5) and (6). The coefficients which are l_{11} 、 l_{12} 、 l_{21} 、 l_{22} 、 l_{1y} and l_{2y} are computed by (7), (8), (9), (10) and (11). x_{i1} is the value of the ith sample in variable one, and x_{i2} is the value of the ith sample in variable two. y_i is the value of the ith sample in dependent variable. $\bar{x}_1$, $\bar{x}_2$ and $\bar{y}$ which represent the corresponding variable averages are computed by (12), (13) and (14). Binary linear regression residual is computed by (15).

$$\hat{y} = \hat{a} + \hat{b}_1 x_1 + \hat{b}_2 x_2 \tag{3}$$

$$l_{11} b_1 + l_{12} b_2 = l_{1y} \tag{4}$$

$$l_{21} b_1 + l_{22} b_2 = l_{2y} \tag{5}$$

$$a = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2 \tag{6}$$

$$l_{11} = \sum_i x_{i1}^2 - \frac{1}{n}\left(\sum_i x_{i1}\right)^2 \tag{7}$$

$$l_{12} = l_{21} = \sum_i x_{i1} x_{i2} - \frac{1}{n}\left(\sum_i x_{i1}\right)\left(\sum_i x_{i2}\right) \tag{8}$$

$$l_{22} = \sum_i x_{i2}^2 - \frac{1}{n}\left(\sum_i x_{i2}\right)^2 \tag{9}$$

$$l_{1y} = \sum_i x_{i1} y_i - \frac{1}{n}\left(\sum_i x_{i1}\right)\left(\sum_i y_i\right) \tag{10}$$

$$l_{2y} = \sum_i x_{i2} y_i - \frac{1}{n}\left(\sum_i x_{i2}\right)\left(\sum_i y_i\right) \tag{11}$$

$$\bar{x}_1 = \frac{1}{n}\sum_{i=1}^{n} x_{i1} \tag{12}$$

$$\bar{x}_2 = \frac{1}{n}\sum_{i=1}^{n} x_{i2} \tag{13}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{14}$$

$$l_{yy0} = \frac{l_{yy} - b_1 \times l_{1y} - b_2 \times l_{2y}}{n-2-1} \tag{15}$$

Based on binary linear regression, there are the computing processes of binary linear regression residual of every point of the branches and the total residual. Before the third step of extracting drainage network is carried on, binary linear regression residual l_{yy0} of every point of the branches is computed. The points which are used to calculate binary linear regression of the current point are all upstream points of the current point and the current point, and the number of the points is n. The value of variable x_1 is the abscissa of the point, and the value of variable x_{i1} is the abscissa of the ith point. The value of variable x_2 is the ordinate of the point, and the value of variable x_{i2} is the ordinate of the ith point. The value of variable y is the elevation of the point, and the value of variable y_i is the elevation of the ith point. If the number of the points of the branch is less than 4, binary linear regression residual is set to 0. Total residual l is the binary linear regression residual l_{yy0} when the points are all points of the branches.

This is the binary linear regression residual normalization procedure. According to the experiments in extreme terrains, binary linear regression residual is about 5000 in mountainous terrain, and it is below 10 in flat terrain. Fig.2 is the graph about the relationship between the residual of all points and the standardized percentage r. Equation (16) is the equation which represents the relationship between the residual of all points and the standardized percentage r. Here is the standardization procedure of binary linear regression residual. Value R is computed by (17) and (18) according to total residual l of all points in the experimental area, and it is named standardized residual. Here is the normalization procedure of binary linear regression residual. According to standardized residual R, the residual of every point is normalized by (18), and it is named normalized residual.

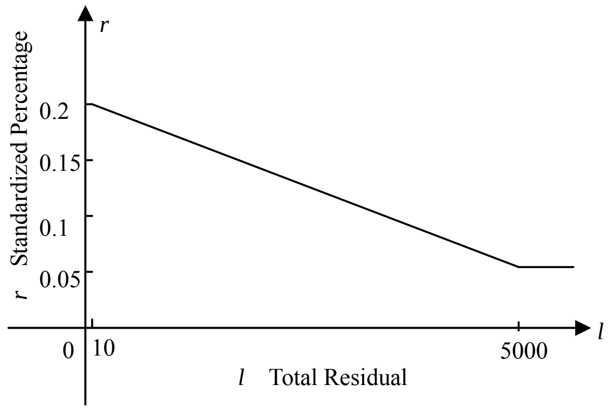

Figure 2. The graph about the relationship between the residual of all points and the standardized percentage

$$r(l) = \begin{cases} 0.2 & l < 10.0 \\ [-0.00003 \times (l - 10.0)] & 10.0 \le l \le 5000.0 \\ 0.05 & l > 5000.0 \end{cases} \tag{16}$$

$$\mathrm{R} = l \times \mathrm{r} \tag{17}$$

$$y_r = \frac{l_{yy0}}{R} \tag{18}$$

C. Design of Drainage Network Extraction Algorithm Based on Binary Linear Regression.

Fig.3 is the dendritic flow figure when the branches are not cut off, and every dendritic structure has a root. The root is the most downstream point. There may be multiple branches from different levels in a dendritic structure. A branch includes the upstream point before the intersection and all other upstream points, and the upstream point before the intersection of the branch is called as the root of the branch. The number of the points of a branch is the size of the branch. The red parts in Fig.3 are the examples of two branches. In the first example, root1 is the root of the first red branch, and the size of the first red branch is 17. In the second example, root2 is the root of the second red branch, and the size of the second red branch is 7.

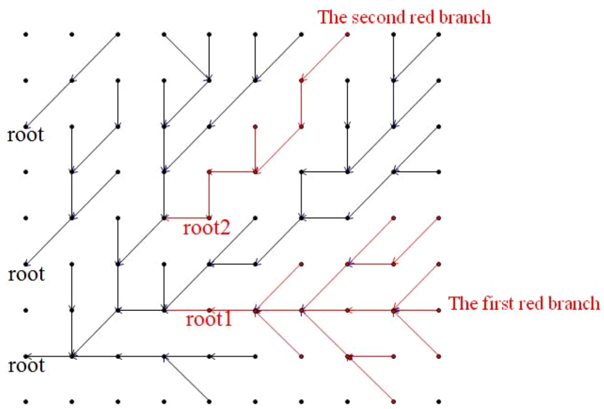

Figure 3. The dendritic flow figure

The processes of pruning are on the base of two factors. The first factor is the size of the branch, and the second factor is the normalized residual of the root of the branch. Two integers are identified which are minbranch and brabch, and minbranch is less than brabch. If the number of the points of the branch is less than minbranch, the branch will be cut. If the number of the points of the branch is greater than brabch, the branch will be remained. If the size of the branch is between minbranch and brabch, the normalized residual of the root point of the branch will be considered. If the normalized residual of the root point of the branch is less than 1.0, the branch will be cut. Otherwise, the branch will be remained. The root of every dendritic structure is the beginning of pruning, and the branches may be cut off in upstream sequence by drainage network extraction algorithm based on binary linear regression. When the searched branch is meet the above-mentioned condition of remaining branch, the branch is remained, and then the root of the branch is regarded as a starting point, and the upstream pruning is recursively carried out from the starting point. When the searched branch is not meet the above-mentioned condition of remaining branch, the branch is cut off. All remained branches have been searched, and pruning has been fulfilled. The remained branches are the extracted drainage network.

III. The Introduction Of The Experimental Area

The experimental area is from 111.0003732°E, 41.95892299° N to 112.4997607°E, 41.54128912 °N, and it roughly includes some parts of SiZiWangQi country and WuChuan country in Inner Mongolia. The area is about 5770 km2. When the experimental area is divided into four equal areas with vertical and horizontal lines, the northwest part (part A) in the experimental area is a flat terrain, and the southeast part (part B) is a mountainous terrain. The actual drainage network in the experimental area is drawn by hand from Google Earth which is shown in Fig.4.

The data sources of the experimental area include geographic coordinates and geographic elevations of all points, which are extracted from Google Earth [27, 28]. The scale of the data source is 1:10000, and the distance between every two points is 60 meters. Then, the geographic coordinates is transformed to BJ-54 coordinates. Finally, the grid spacing in the X direction or the Y direction is 20 meters, and ordinary Kriging method is applied to create contour map. The contour data is taken as the data source of experimental area, and the contour map of the experimental area is shown in Fig.5.

IV. Results

There are the experimental discussions of extracting drainage network with ARCGIS. The main work of extracting drainage network with ARCGIS is to determine the catchment area threshold. When the catchment area threshold is 140, the extracted drainage network is shown in Fig. 6. When the catchment area threshold is 20, the extracted drainage network is shown in Fig. 7. In Fig. 6 and Fig. 7, the red lines are the actual drainage network of the experimental area, and the black ones are the extracted drainage network.

When the catchment area thresholds are 140 and 20, there are some parameters which are shown in TABLE I. When the catchment area threshold is 140, the extracted result in flat terrain (part A) is only in accordance with the actual drainage network which is shown in TABLE I. But the density of the extracted drainage network with ARCGIS in mountainous terrain (part B) is far less than the density of the actual drainage network. When the catchment area threshold is 20, the extraction result in mountainous terrain (part B) is in accordance with the actual drainage network which is shown in TABLE I. But the density of the extracted drainage network with ARCGIS in flat terrain (part A) is far more than the density of the actual drainage network.

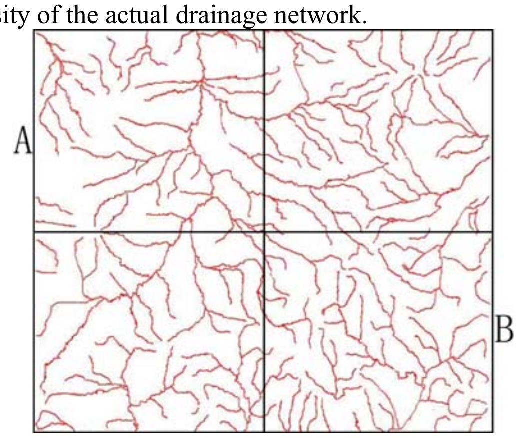

Figure 4. The sketch of the actual drainage network in the experimental area

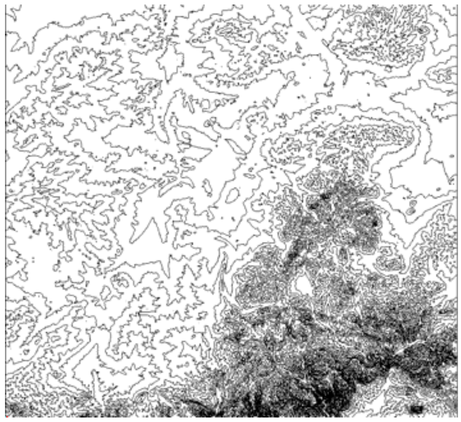

Figure 5. The contour map of the experimental area

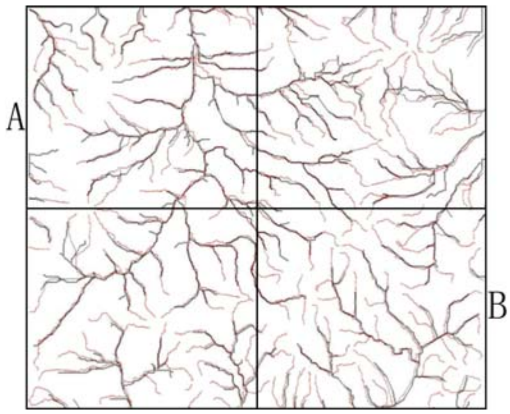

Figure 6. The overlay chart of the actual drainage network and the extracted drainage network with ARCGIS when the catchment area threshold is 140

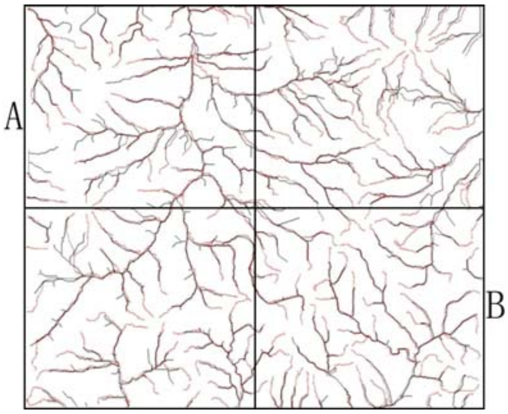

Figure 7. The overlay chart of the actual drainage network and the extracted drainage network with ARCGIS when the catchment area threshold is 20

TABLE I

SOME PARAMETERS OF THE ACTUAL DRAINAGE NETWORK AND THE EXTRACTED DRAINAGE NETWORK WHEN THE CATCHMENT AREA THRESHOLDS ARE 140 AND 20

Catchment area threshold / Drainage parameters	the catchment area threshold is 140	the catchment area threshold is 20
The number of the actual drainage sources in part A	46	46
The density of the actual drainage in part A	0.032	0.032
The number of the extracted drainage with ARCGIS in part A	51	71
The density of the extracted drainage with ARCGIS in part A	0.035	0.049
The number of the actual drainage sources in part B	64	64
The density of the actual drainage in part B	0.044	0.044
The number of the extracted drainage with ARCGIS in part B	37	64
The density of the extracted drainage with ARCGIS in part B	0.026	0.044

V. DISCUSSION

In short, if there are flat terrain and mountainous terrain in an experimental area, two kinds of terrain are not differentiated when the drainage network is extracted with ARCGIS. This is because only one catchment area threshold is used for extracting the drainage network. There are the experimental discussions of the drainage network extraction algorithm based on binary linear regression[29, 30, 31]. In the experimental area, the number of the actual drainage sources is 181, the drainage area is 5770km2, and the actual drainage density is 0.031. When the normalized residual is equal or greater than 1.5, the number of the extracted drainage sources is 183, and the extracted drainage density is 0.032. So, the conclusion is that the extracted drainage network is consistent with the actual drainage network. It is the overlay chart of the actual drainage network and the extracted drainage network which is shown in Fig. 8. When the normalized residual is equal or greater than 1.5, there are the drainage parameters of the extracted drainage and the actual drainage in flat terrain (part A) or mountainous terrain (part B) which is shown in TABLE II. It is shown that the actual drainage density roughly equals to the drainage density in part A in TABLE II. And it is shown that the actual drainage density roughly equals to the extracted drainage density in part B in TABLE II. So, when there are two types of terrain in the experimental area, the terrains are separated automatically from one to another by the algorithm, and then the local catchment area threshold is determined according to the different terrain characteristics.

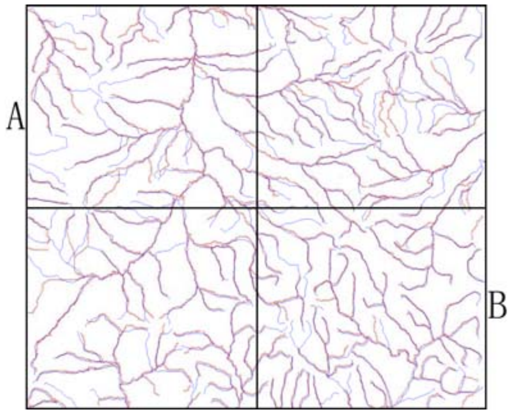

Figure 8. The overlap chart of the extracted drainage network (the blue lines) and the actual drainage network (the red lines)

TABLE II

THE DRAINAGE PARAMETERS OF THE EXTRACTED DRAINAGE AND THE ACTUAL DRAINAGE IN FLAT TERRAIN (PART A) OR MOUNTAINOUS TERRAIN (PART B)

Drainage Sources / Drainage parameters	The actual drainage in part A	The extracted drainage in part A	The actual drainage in part B	The extracted drainage in part B
The number of the sources	46	44	64	66
The drainage area (km^2)	1442.5	1442.5	1442.5	1442.5
The drainage density	0.032	0.031	0.044	0.046

VI. CONCLUSIONS

Only one catchment area threshold is determined to extract the drainage network with ARCGIS. When there are two terrains in the experimental area, the extracted drainage network is only consistent with the actual drainage network on either of two terrains. So the local catchment area thresholds are needed to extract drainage network. Drainage network extraction Algorithm based on binary linear regression is the algorithm that is used to distinguish automatically terrains and determine the local catchment area thresholds, and the proposed method is mainly divided into two steps.

This is the first step. Binary linear regression method is applied to calculate the residual value at every point, and then the normalized residuals are used to prune reasonably as the variables. The terrains are distinguished automatically according to the normalized residuals, and the local drainage area threshold is determined based on each terrain feature.

This is the second step. The pruning conditions are determined. If the size of the branch is less than the size of the smallest area, the branch is cut. If the size of the branch is more than the number of the points in the minimum catchment area, the branch is remained. If the size of the branch is the number between minbranch and brabch, the normalized residual of the root point of the branch will be computed. According to the normalized residual of the root point of the branch, the terrains are distinguished automatically, and the local catchment area thresholds are determined.

According to two steps above, the extracted drainage network is consistent with the actual drainage network when there are two terrains in the experimental area.

ACKNOWLEDGEMENT

The work for this paper is supported by Geology survey information applicable technology of Application research under grant 1212011120445.

REFERENCES

[1] Zhong Zheng, Jun Xiao, Jun Wang, "The application research on engineering hydrology information based on DEM," China Rural Water and Hydropower, vol. 4, pp. 20-22, 2010.

[2] Shiguo Wang, Ying Xiong, "Research on drainage network extraction and classification in depression of nanshan of xinjiang manas river basin," Journal of Anhui Agricultural Science, vol. 38, pp. 17955-17956, 2010.

[3] Xujiao Li, Wuniang Yang, Chao Zeng, "Drainage network extraction in basin of the upper stream of minjiang river," Geospatial Information, vol. 9, pp.58-63, 2010.

[4] Junming Chen, Guangfa Lin, Zhihai Yang, Hanyue Chen, "Parameters Optimization in Extracting Digital Drainage Networks with DEM," OURNAL OF GEOINFORMATIO -N SCIENCE, vol. 13, pp. 32-37, 2011.

[5] Lanyan Ma, Chunping Zhou, Zhuowei Hu., "Research on Drainage Network Extraction in Liao He Basin Based on SRTM DEM and ASTER GDEM", Journal of Anhui Agricultural Science, vol. 39, pp. 2692-2695, 2011.

[6] Yinghui Guan, Fenli Zheng, Bin Wang, Xiaobin Ding, Yibo Liu, "DEM Based Extraction of River Network in Binzhou River Basin of Heilongjiang Province", Bulletin of Soil and Water Conservation, vol. 32, pp. 127-131, 2012.

[7] Xiangyang Song, Faqi Wu, Longshan Zhao, Guangyan Wu, Qingfeng Zhang, "Extraction and analysis of hydrological characteristics of Yanhe watershed based on DEM", Agricultural Research in the Arid Area, vol. 30, pp. 200-206, 2012.

[8] R.Turcotte, J.-P. Fortin, A.N. Rousseau, "Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network", Journal of Hydrology, vol. 240, pp. 225-242, 2000.

[9] Zhenchun Hao, Li Li. "The Method of the Net of River Based on DEM", Journal of Hydrology, vol. 22, pp. 8-10, 2002.

[10] Changfeng Li, Rui Zhao, "The Application of Automatical -ly Extracting Stream Networks in the Xitiaoxi Watershed", TROPICAL GEOGRAPHY, vol. 23, pp. 319-323, 2003.

[11] Changfeng Li, Xuezhi Feng, Rui Zhao, "The Methods and Application of Automatically Extracting Stream Network of Watershed", Journal of Lake Sciences, vol. 15, pp. 205-212, 2003.

[12] Lei Yu, HongBin Liu, Wei Wu, "Extraction of River Networks in the Three Gorges Area in Chongqing Based on DEM", SCIENTIA GEOGRAPHICA SINICA, vol. 26, pp. 616-621, 2006.

[13] Hanwei Xu, Jiang He, Yuanjian She, "Method for extraction of digital drainage network in the Qinhuai River basin based on DEM and remote sensing", Journal of Hohai University (Natural Sciences), vol. 36, pp. 443-447, 2008.

[14] Shijie Zhu, Guoan Tang, Yini Jia, "Research on a method for DEM-derived stream adjustment based on DLG", Modern Surveying and Mapping, vol. 32, pp. 3-4, 2009.

[15] Wenjun Zuo, Jincun Zhang, Chao Jia, "Extraction of Digital Urban River Network Based on DEM", Water Resources and Power, vol. 28, pp. 26-29, 2010.

[16] Jiahong Liu, Dayong Qin, Haihong Li, Zhiyong Yang, "An intensifying method for extracting drainage networks in a plain with violent human activities", Journal of China Institute of Water Resources and Hydropower Research, vol. 8(2), pp. 128-137, 2010.

[17] Wang Chen, Yadong Deng, Hong Liang, Xiaoyun Xie, Jinyi Li, "Application Study on Extracting Karst Area's River System by Digital River Basin Assisted DEM", Journal of Yangtze River Scientific Research Institute, vol. 27, pp. 74-78, 2010.

[18] Wei Zhang, Xin Yang, Guoan Tang, Shijie Zhu, Caili Li, "DEM-based flow direction algorithms study of stream extraction and watershed delineation in the low relief areas", Science of Surveying and Mapping, vol. 37, pp. 94-96, 2012.

[19] Guozhen Zhang, Enping Yan, Yifeng Hong, Li Chen, Dengkui Mo, "Study on the Hydrological Analysis of Dongjiang Lake Scenic Area Based on DEM", Chinese Agricultural Science Bulletin, vol. 29,pp. 172-177, 2013.

[20] Yunshan Tao, "Iteration method solving the nonlinear regression problem and its application in hydrology", Journal of East China Jiao tong University, vol. 1, pp. 42-48, 1991.

[21] Shuying Wang, Yongsheng Gao, "Application of Improved Least Square Method in Hydrological Analysis Calculation", HYDROLOGY, vol. 23, pp. 5-9, 2003.

[22] GuoLi Wang, XiaoFei Chen, Kan Liu, Guoyong Jiang, "Summary on Application of Regression Analysis in Water Science", China Rural Water and Hydropower, vol. 11, pp. 40-44, 2004.

[23] Xiangchun Kong, "Application of multinomial curve fit program in hydrological analysis", Water Resources & Hydropower of Northeast, vol. 23, pp. 34-35, 2005.

[24] Dongmei Cheng, Zhifeng Liu, Linlin Wang, "EXCEL Application in Hydrology Linear Regression Analysis", Heilongjiang Science and Technology of Water Conservancy, vol. 40, pp. 96-98, 2012.

[25] Fairfield.J., Leymarie.P, "Drainage networks from grid digital elevation models", Water Resource Research. vol. 27, pp. 709-717, 1991.

[26] Yiheng Yang, "The processing techniques and approaches of the multidimensional geosciences' data", Beijing: Science Press, pp. 20-28, 2002.

[27] Wenrong Liu, "Development of Integrated Google Earth Elevation Extraction Program and Application in Environ -ment Evaluation", Coal Engineering, vol. 5, pp. 122-124, 2011.

[28] Shengjin Tang, Lidong Zhang, Shaojie Wang, "Elevation data extraction from Google earth", Henan water and South-to-north Water Diversion, vol. 8, pp. 105-106, 2012.
[29] Jianguo Chen, Hao Chen, "A Structured Information Extraction Algorithm for Scientific Papers based on Feature Rules Learning", Journal of Software, vol.8, pp.55-62, 2013.
[30] Bing Gao, Jianpei Zhang, "Density Based Distribute Data Stream Clustering Algorithm", Journal of Software, vol.8, pp.435-442, 2013.
[31] Lu Huang, Haishen Chen, Tingting Hu, "Survey on Resource Allocation Policy and Job Scheduling Algorithms of Cloud Computing1", Journal of Software, vol.8, pp.480-487, 2013.

Dandan Su was born in Changchun, China. She received a B.S. degree in 2003 in computer technology from Jilin Normal University. She received a M.S. degree in 2006 in Earth detecting and information technology from Jilin University. Now she is continually study on Earth detecting studying for a and information technology, and she is Dr. degree in Jilin University. Her research interests in computer science are in the areas of the application of GIS software and algorithm design. She is the first-author of over 7 papers published on journals and is an active consultant for industry in these areas.

Guangsheng Liu was born in Inner Mongolia, China. He received a B.S. degree in 1988 in mathematics and M.S. degree in 1991 in probability and statistics from Northeast Normal University, Changchun, China. He received a D.E. degree in 1999 in earth exploration and information technology from Changchun College of Geology, Changchun, China. In 1991, he joined the Institute of Synthetic Information Mineral Resources Prediction, Changchun college of geology (now be incorporated into Jilin University), Changchun, China, where he is presently an Associate Professor of Mathematical Geology. His research interests are in mathematical geology and he has published over 10 papers on journals.