Mapping the average annual runoff depth in Huangshui watershed using DEM

SHAOFENG JIA[1], SHIJUN LIN[1,2] & YAJUN WANG[3]

1 *Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China*
jiasf@igsnrr.ac.cn

2 *Graduate School of the Chinese Academy of Sciences, Beijing 100039, China*

3 *Bureau of Science and Technology of Qinghai Province, Xining 81001, China*

Abstract We found that precipitation and the surface water flows generally increase with elevation, as does the vegetation cover, in Huangshui watershed. This determined that the geographical variables, such as elevation and vegetation cover, can be used to improve the precision of mapping of runoff depth. We focus on discovering an automatic approach for mapping runoff using geographical variables. Using multiple linear regression techniques we create a regression equation that relates runoff depth to altitude and relative distance from the source of water vapour. Data used for this study include runoff data measured at 13 stations, Landsat TM data and DEM data; a 50 m × 50 m grid of runoff depth was created using those models. We used the runoff depth data of 13 drainage zones for validation. Comparing with the empirical model used by local water resource survey institutions, the regression approach appears to give more precise results.

Key words runoff depth; multivariate linear regression; geographical variables; DEM

INTRODUCTION

Runoff depth is one of the most important hydrological data. Mapping runoff depth across space constitutes one of the fundamental tasks in hydrology. The only direct source of information for completing this task is hydrological observation data. However, hydrological series are sometimes too few and/or too short for a reliable determination of runoff statistic characteristics, or even not available at all in ungauged areas. According to the water balance, some studies deduce runoff from simulated spatial distribution of precipitation and evapotranspiration, using the more abundant precipitation data and a rainfall–evapotranspiration model (Morton, 1983). Some studies model the spatial distribution of runoff depth by runoff curve number based on the relation between land use and infiltration and evaporation using remote sensing data (Melesse & Shih, 2002). In recent years, distributed hydrological models using GIS and RS have undergone great development and are used in simulating the discharge of watersheds. Some of these models can obtain runoff information in spatial computation unit (grids or representative unit). TopModel (Beven, 1995) and SWAT (Neitsch *et al*., 2000) are two typical distributed models. But all these models need vast amounts of basic data as input. In areas lacking basic data, the model can not work. Another method of spatial mapping is interpolation. A number of interpolation methods that have been used to model the spatial distribution of runoff depth include IDW, spline, kriging, etc. Usually, every subwatershed's average runoff depth is associated with the centroid of each subwatershed. Then an interpolation method is processed based on those points to model the spatial distribution of runoff depth. However, interpolation methods only consider spatial relationships among sampling points, and do not take into account other properties of the landscape (Gottschalk, 1993; Sauquet *et al*., 2000).

This study sought to map runoff depth across a mountainous region of Huangshui catchment, in Qinghai province of China, in terms of the relationships between runoff depth and a range of geographic variables, including altitude, distance from water vapour intake, and vegetation index. Using multiple linear regression techniques, we create a model to disaggregate the mean annual streamflow measured at the gauge station to basin area. This model is compared to other mapping methods.

METHODS AND DATA

Study area

Huangshui watershed (only the part in Qinghai province, not including the Datong River watershed) is located in the east of Qinghai province. Huangshui River is a first-class tributary of the Yellow River, flowing from northwest to southeast. The Huangshui watershed covers an area of 16 120 km^2, has a shape of a leaf, located in between 36°02′–37°28′N latitude, 100°42′–103°04′E longitude. There are 13 stream-gauging stations in operation in Huangshui watershed (Fig. 1). The annual average runoff of Huangshui watershed is 2.16×10^9 m^3, the average flow rate is 65.7m^3/s, and the annual average runoff depth is 138 mm.

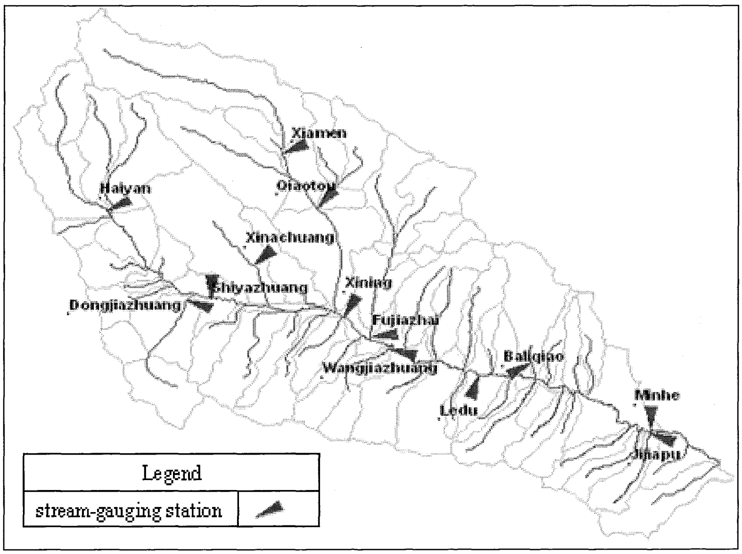

Fig. 1 Location of the study area and positions of the 13 stream-gauging stations.

Huangshui watershed is affected by a southwest water vapour, rising from the Indian Ocean Bay of Bengal. The nearest available sources of moisture are very far from Huangshui watershed. The climate there is classified as semi-humid, with an average annual precipitation of 486.6 mm. The annual precipitation distribution there is very uneven, with a Cv from 0.15 to 0.30. The upstream area receives much more precipitation than the downstream area. The annual precipitation in a valley area between Xining and Minhe varies from 250 mm to 350 mm, while more than 600 mm precipitation is received in the mountain area per year.

Huangshui watershed has a very inclining hypsography, with altitudes ranging from 4900 to 1650 m, rising gradually from southeast to northwest. The minimum elevation is located in the valley near the Gansu-Qinghai border. There are various types of topography in Huangshui watershed characterized by mountains, hills covered by loess and valley plain.

Data

Along the river, we divide the study area between every two gauges. This generates 13 nonoverlapping drainage zones. We used a period of 42 years, between 1958 and 2000, to calculate the average annual runoff depth of every drainage zone (Table 1).

In addition, a digital elevation model (DEM, 50 m × 50 m) and a Landsat TM image (Summer 2000) of Huangshui watershed were used in this study.

Table 1 Average annual runoff depth of drainage zones in Huangshui watershed.

Drainage zone	Drainage area (km^2)	Discharge ($10^9 m^3$)	Runoff depth (mm)
Xinachuan	809	1.633	201.85
Xiamen-Qiaotou	1466	2.696	183.92
Baliqiao	464	0.952	205.26
Xiamen	1308	3.617	276.55
Haiyan	715	0.473	66.21
Haiyan-Shiyazhuang	1732	1.786	103.12
Dongjiazhuang	636	0.84	132.10
Ledu-Minhe	1853	2.133	115.11
Fujiazhai	1112	2.056	184.93
Xining-Ledu	2521	2.506	99.39
Wangjiazhuang	370	0.437	118.10
Jijiabao	192	0.331	172.40
Shiyazhuang-Xining	2356	2.147	91.13

METHOD

Firstly, as in general interpolation methods, we assume that the runoff depth of the centroid point of a drainage zone equals the average runoff depth of the drainage zone and map runoff depth by using the normal interpolation methods (IDW, spline, and kriging). Then we qualitatively revised the contour based on the empirically perceived relation between runoff depth and elevation and vegetation cover. The result shows that the revised mapping is much better than the original interpolated mapping. So we want to construct a mathematical model to use the relations between runoff depth and altitude, vegetation index and other geographical variables to map runoff depth. The multivariate regression method is used in this process. SPSS and ArcInfo are used for regression analysis and spatial mapping separately.

RESULTS AND DISCUSSION

Interpolation of runoff depth

There are three kinds of interpolation methods available in GIS software: IDW, spline, and kriging. Those interpolation methods have been applied to Huangshui watershed by using ArcInfo. The standard deviation of residuals from interpolation methods are shown in Table 2, which shows the spline interpolation, obtains the best result.

Table 2 SD and SDORR of IDW, spline and kriging.

Interpolation method	SD (mm)	SDORR (%)
IDW	27.60	23.05
Spline	17.58	13.42
Kriging	27.36	23.29

SD, the standard deviation of residuals from estimated model. SDORR, the standard deviation of relative residuals from estimated model.

Revision of contour lines based on DEM and Landsat TM

Based on the spline contour, we revised the runoff depth contour by using DEM data. We use a criterion to revise the runoff depth contour that a runoff depth contour line should be nearly paralleled with the elevation contour. The revised runoff depth contour is shown as Fig. 2.

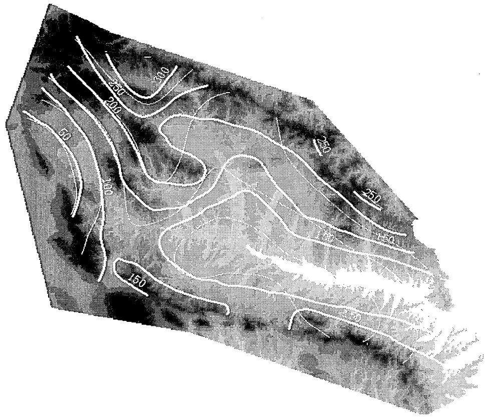

Fig. 2 Revised based on DEM. White line: after revised; white thin line: before revised.

Table 3 SD and SDORR after revision based on DEM.

Runoff mapping method	SD (mm)	SDORR (%)
Spline	17.58	13.42
Revised based on DEM	17.97	12.14

After the contour lines were revised based on the DEM, the residuals became much smaller than those in the spline interpolation model (Table 3). The standard deviation of relative residuals dropped from 13.42% to 12.14%, though the standard deviation of residuals was increased a little (from 17.58 to 17.97). The precision of the estimated values of the relative small runoff depth is improved, apparently at the cost of a little less precision of the relative big runoff depths.

Vegetation is a very important part of the environment. The vegetation can usually reflect the local landform, soil, geology, climate and elevation. Generally, the vegetation cover is thick in the places where precipitation is abundant, and is sparse in the areas lacking precipitation. So, we can assume that more precipitation is expected in the area where the vegetation cover is thick. And we can qualitatively determine that in an area where the vegetation cover is thick, more precipitation is received and runoff is also abundant; and *vice versa*. Based on the vegetation information on Landsat TM images, we revised the runoff depth contour further. Let the distribution of the contour be consistent with the vegetation coverage and get the revised runoff depth contour as shown in Fig. 3. The validation results shows that the revised contour improved further (Table 4).

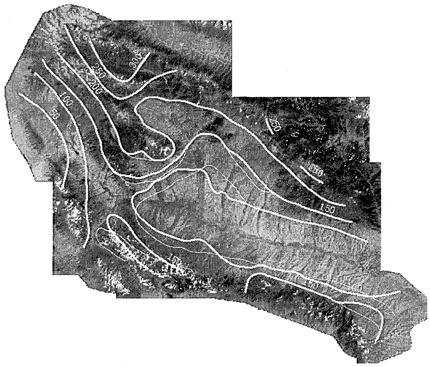

Fig. 3 Revised based on Landsat TM. White line: after revised; white thin line: before revised.

Table 4 SD and SDORR after revision based on Landsat.

Runoff mapping method	SD (mm)	SDORR (%)
Spline	17.58	13.42
Revised based on DEM	17.97	12.14
Revised based on Landsat TM	16.52	11.74

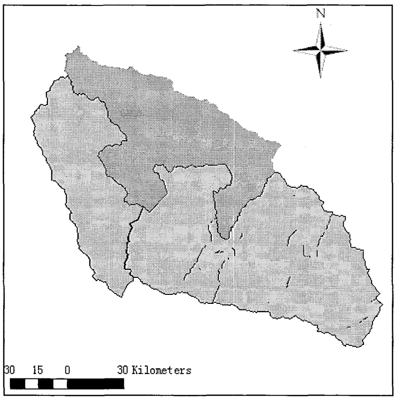

Fig. 4 divided the study area into two parts.

Regression model

According to the statistical characteristics of the runoff depth, we divided Huangshui watershed into two parts: (1) the southern part; (2) the northern part (Fig. 4). The northern part includes Xinachuan, Xiamen-Qiaotou, Xiamen and Fujiazhai drainage zones. The others are included in the southern part. In the regression analysis we model the spatial distribution of runoff depth for the northern part and the southern part, respectively.

The relationships between runoff depth and the independent topographic variables generally fit squared or cubic or higher-power models better than linear ones. So we introduced the squared, cubic and fourth order values of the independent variables to achieve a better fit of the theoretical curve to the data.

We used DEM to provide estimates of elevation, and the Landsat TM data to derive vegetation index. We computed the distance from a drainage zone's centroid to the south side of the watershed (dsouth (m)), and to the east side (deast (m)). The average runoff of the centroids of the 13 subwatersheds described previously was used for multivariate regression analysis using SPSS software.

We took the average runoff depth of a drainage zone as the independent variable, and took the average elevation of a drainage zone (H), the higher power of H (H^2, H^3, and H^4), the average vegetation index of a drainage zone (V), deast, $deast^2$, $deast^3$, $deast^4$, dsouth, $dsouth^2$, $dsouth^3$ and $dsouth^4$ as dependent variables.

For the regression, we followed the criterion of the stepwise approach, selecting the variables backwards. This method has the advantage that a variable selected in one step can be eliminated in another later step. In this way, initially all the variables are introduced in a single step and then they are discarded one by one, based on the outset criteria.

Pearson correlation coefficients for the relationships between runoff depth and the independent variables in the southern area are shown in Table 5. The best runoff depth was estimated with model 2 in Table 5, which is expressed as equation (1).

$$R = 1.71E - 12{*}H^4 - 0.001{*}deast.73.891 \quad (1)$$

For the north area, the final regression model is expressed as equation (2).

$$R = -5.2E - 19{*}dsouth^4.132.42 \quad (2)$$

Statistic test values for this model are shown in Table 6.

Table 5 Statistic values of regression models of southern area.

Model	R	R square	Adjusted R square	Std error of the estimate
1	0.830(a)	0.690	0.651	34.50673
2	0.910(b)	0.829	0.780	27.39156

Table 6 Statistic values of regression model of northern area.

Model	R	R square	Adjusted R square	Std error of the estimate
1	0.999(a)	0.998	0.996	2.12078

Table 7 Estimated average annual runoff depth by regression model.

Drainage zone	Observed (mm)	Estimated (mm)	Difference (mm)	Difference (%)
Xining–Ledu	99.39	118	−18.15	−18.26
Dongjiazhuang	132.1	132	0.52	0.39
Wangjiazhuang	118.1	105	13.39	11.34
Baliqiao	205.26	234	−29.07	−14.16
Xiamen	276.55	250	26.56	9.60
Haiyan	66.21	62	4.21	6.36
Xinachuan	201.85	151	50.95	25.24
Haiyan–Shiyazhuang	103.12	97	6.42	6.23
Shiyazhuang–Xining	91.13	78	13.37	14.67
Jijiabao	172.43	132	40.32	23.39
Xiamen–Qiaotou	183.92	148	35.90	19.52
Ledu–Minhe	115.11	135	−19.83	−17.22
Fujiazhai	184.93	144	41.31	22.34
SD			25.48	15.26

We have noticed that there is no apparent relationship between R and H in the north part. The reason may be that the altitude of the north changes more slowly than the south, the distance from the source of water vapour becomes the major factor of runoff generation in the north part.

Initially, some vegetation variables were included in the regression analysis. But all the vegetation variants were finally eliminated. There are two reasons for the absence of vegetation variables. First, vegetation is strongly correlative with elevation. So when elevation is used in the regression model, the vegetation variables are no longer necessary. Secondly, the main vegetation in the valley plain area is crop, and the crop is hardly correlative with runoff depth.

Residual analysis

The resolution of the DEM used in this study is 50 m. So we can create a 50 m × 50 m grid of runoff depth using the regression model. The value of each grid cell represents the estimated runoff depth of a special 50 m × 50 m area. The observed runoff depth of each 50 m × 50 m area is not available. We can not use the grid to assess the validity of the models directly. But the observed runoff depth data of the 13 drainage zones are available. And the estimated runoff depth of each drainage zone can be easily calculated by using the grid values. So we can use the runoff depth of the 13 drainage zones for validation.

The comparison of the estimated runoff depth of the 13 drainage zones using the regression model and the observed ones are shown in Table 7. The standard deviation of residuals is 25.48 mm, and the standard deviation of relative residuals is 15.26%.

Table 8 SD and SDORR of the six runoff mapping methods.

Runoff mapping method	SD (mm)	SDORR (%)
IDW	27.60	23.05
Spline	17.58	13.42
Kriging	27.36	23.29
Regression	25.48	15.26
Revised manually	16.52	11.74
Empirical model used by local institutions	29.01	18.52

When the runoff depth contour is available, we can create a grid of runoff depth by using GIS software. In order to compare the interpolation models with the regression model, we created a 50 m × 50 m grid of runoff from the runoff depth contour. And each grid created was used to validate the performance of each model. The standard deviation of residuals and relative residuals in the six methods have been calculated and are presented in Table 8.

CONCLUSION

In this study, we succeed in improving the precision of mapping of runoff depth using DEM and related geographic data. By comparing various models we can reach the following conclusions:

(a) The mapping results of runoff depth, computed by using the normal interpolation methods included in GIS software, appear to be better than the empirical mapping which is still used in the local hydrological bureau, according to SD (mm). This indicated that using the GIS software to map runoff depth can not only improve the efficiency, but can also get a more reliable result.

(b) The geographical variables, such as elevation and vegetation cover, can be used to improve the precision of mapping of runoff depth.

(c) Using multiple linear regression techniques we create a regression model that relates runoff depth to altitude and distance from the source of water vapour. And the regression model can achieve a similar precision as that from the normal interpolation methods. But it fails to greatly improve the precision. The reason is that the relation between runoff depth and geographical characteristics is very complicated and it is not always possible to use one regression function to show these complex relations.

(d) One potential merit of the regression model is that it may be extended and used to predict runoff depth in ungauged basins with similar geographical conditions, while interpolation methods do not have this function. Because of the geographical difference between the north and the south in the study area, we used two different equations even in the same basin. This shows that it is very difficult to assess the geographical similarity between various areas, and it needs more effort and study to find a feasible method to decide whether an area's runoff depth can be estimated by the runoff depth equation of another area.

Acknowledgements The research is supported by the National Natural Science Foundation of China (Grant 90302009), and the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences (CXIOG-A05-13).

REFERENCES

Beven, K. J. (1995) TOPMODEL. In: *Computer Models of Watershed Hydrology*, Chapter 18 (ed. by V. P. Singh). Water Resources Publications, Littleton, Colorado, USA.

Gottschalk, L. (1993a) Correlation and 5 covariance of runoff. *Stochastic Hydrol. Hydraul.* **7**, 85–101.

Gottschalk, L. (1993b) Interpolation of runoff applying objective methods. *Stochastic Hydrol. Hydraul.* **7**, 269–281.

Melesse, Assefa, M. S. & Shih, F. (2002) Spatially distributed storm runoff depth estimation using Landsat images and GIS. *Computers and Electronics in Agriculture* **37**, 173–183.

Morton, F. I. (1983) Operational estimate of areal evapotranspiration and their significance to the science and practise of hydrology. *J. Hydrol.* **66**(1–4), 1–76.

Neitsch, S. L., Arnold, J. G., Kiniryj, R. *et al.* (2000) Soil and Water Assessment Tool Theoretical Documentation / Version 2000 [EB.OL] http://www.brc.tamus.edu/swat/.

Sauquet, E., Gottschalk, L. & Leblois, E. (2000) Mapping average annual runoff: a hierarchical approach applying a stochastic interpolation scheme. *Hydrol. Sci. J.* **45**(6), 799–815.