Front. Earth Sci. 2024, 18(3): 509–525
https://doi.org/10.1007/s11707-022-1068-0

RESEARCH ARTICLE

A DEM upscaling method with integrating valley lines based on HASM

Mingwei ZHAO[1,5,6,7], Xiaoxiao JU[2], Na ZHAO (✉)[3,4], Chun WANG[1,5,6,7], Yan XU[1,5,6,7], Xiaoran WU[3,4], Weitao LI[1,5,6,7]

1 Geographic Information and Tourism College, Chuzhou University, Chuzhou 239000, China
2 College of Resource, Environment and Tourism, Capital Normal University, Beijing 100048, China
3 State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
4 College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100101, China
5 Anhui Engineering Laboratory of Geo-information Smart Sensing and Services, Chuzhou 239000, China
6 Anhui Center for Collaborative Innovation in Geographical Information Integration and Application, Chuzhou 239000, China
7 School of Geographic Information and Tourism, Chuzhou University, Chuzhou 239000, China

Abstract A new digital elevation model (DEM) upscaling method based on high accuracy surface modeling (HASM) is proposed by combining the elevation information of DEM and the valley lines extracted from DEM with different flow accumulation thresholds. The proposed method has several advantages over traditional DEM upscaling methods. First, the HASM ensures the smoothness of the upscaled DEM. Secondly, several DEMs with different topographic details can be obtained using the same DEM grid size by incorporating the valley lines with different flow accumulation thresholds. The Jiuyuangou watershed in China's Loess Plateau was used as a case study. A DEM with a grid size of 5 m obtained from the local surveying and mapping department was used to verify the proposed DEM upscaling method. We established the surface complexity index to describe the complexity of the topographic surface and quantified the differences in the topographic features obtained from different upscaling results. The results show that topography becomes more generalized as grid size and flow accumulation threshold increase. At a large DEM grid size, an increase in the flow accumulation threshold increases the difference in elevation values in different grids, increasing the surface complexity index. This study provides a new DEM upscaling method suitable for quantifying topography.

Keywords DEM, upscaling, HASM, flow accumulation threshold, surface complexity index

Received January 21, 2023; accepted May 12, 2023

E-mail: zhaon@lreis.ac.cn

1 Introduction

A digital elevation model (DEM) is the digital representation of the earth's surface in raster format (Burrough et al., 1988; Aguilar et al., 2005). It is a data model depicting terrain at a specific scale (Wolock and McCabe, 2000; Horritt and Bates, 2001; Liu, 2008). Since scale is a comprehensive concept, this paper does not discuss this concept in depth, but considers scale as the level of detail of topographic information in DEM. DEM scales with different grid sizes are different; however, this does not mean that DEM scales with the same grid size must be the same. For example, DEM scales derived from different elevation information may be different.

DEMs are widely used in geomorphology, hydrology, ecology, and environmental sciences (Fisher, 1991; Li and Chen, 2005; Murphy et al., 2008; Kawabata and Bandibas, 2009). Different fields or applications may have different requirements for DEM scales (Schoorl et al., 2000; Vaze et al., 2010). Therefore, the scale transformation of DEMs is an important research topic in geography, geographic information systems (GIS), and other fields. Scale transformation includes downscaling and upscaling. Downscaling a DEM means increasing the resolution of the terrain information. Although some scholars have argued that algorithms based on high-resolution interpolation, deep learning, and other methods can handle DEM downscaling without additional data (Zhang et al., 2021; Zhou et al., 2021), these methods simply use a smaller grid size to depict the terrain while maintaining the original data information. In other words, higher resolution information is usually subject to great

uncertainty. In contrast, DEM upscaling discards some of the information in the input DEM and produces a DEM with a lower resolution (Weibel, 1992; Ai et al., 2007). This process does not increase the level of information and has a rigorous mathematical theory foundation, i.e. map generalization. DEM upscaling has important and wide applications. In digital terrain analysis, the optimal analysis scale of different terrain indices may be different, which requires scale transformation based on the original DEM (Xiong et al., 2021, 2022). In addition, DEM upscaling has important applications in the simulation of geomorphological development processes. For example, in areas where water erosion is the main driving force for geomorphological development, such as the Loess Plateau in China, small topographic reliefs can be removed based on DEM upscaling to simulate surface morphology in different historical periods. Therefore, further study of DEM upscaling is still needed (Xiong et al., 2023).

Cartographic generalization is the theoretical basis for DEM upscaling. Thus, this study focuses on simplifying DEM terrain. Numerous studies have been carried out on this subject. For example, Ai and Li (2010) designed a structured DEM synthesis method that used gully system data as auxiliary information and simplified DEM by filling valleys in some small gullies. Zhou and Chen (2011) considered surface points with large slope and aspect values as key points and fused valley line information. A DEM was obtained after creating a triangulated irregular network (TIN) to retain key terrain features. Chen and Li (2013) used the orthogonal least squares method for extracting key surface points for DEM reconstruction to remove surface details and simplify DEM. Chen et al. (2016) combined key surface points of watersheds at different scales to obtain DEMs at different scales. Ma et al. (2017) proposed a method that used surface key points as the center and generated different profiles according to different thresholds in four directions for simplifying a DEM. Wu et al. (2019) reconstructed the terrain by searching for surface key points and used parallel computing based on Compute Unified Device Architecture (CUDA) to improve data processing efficiency. Raposo (2020) used window smoothing and window size determination by calculating terrain information entropy, ensuring that different terrain sections were smoothed differently depending on terrain complexity, i.e., smoother in complex terrain and vice versa. Chen et al. (2020) designed a Laplacian-based DEM upscaling algorithm (GLD). It extracts a skeleton from the DEM and selects lines that depict the main terrain features. Existing studies can be classified into two categories. The first is based on surface key points. Its disadvantage is that it does not consider the constraints of surface feature lines. As a result, key feature line upscaling results may differ substantially from the original data. The second type combines information from different watersheds, valley lines, and other data. However, this method only used valley line or watershed data at different scales.

High-accuracy surface modeling (HASM) is based on the differential geometry surface theory proposed by Yue (2011). This method was initially used for spatial interpolation (Yue et al., 2007; Yue, 2011), and subsequently improved to have the capability for spatial interpolation, spatial data fusion, spatial data assimilation, and scale transformation (Yue et al., 2016; Yue et al., 2020). Scale transformation in HASM uses the original data as the initial value and adds points with high precision for optimization to achieve scale transformation for different levels of precision (Yue et al., 2020). However, the implementation of HASM-based scale transformation is complex and generally depends on other spatial interpolation methods, so it does not effectively integrate the so-called optimization control conditions with the initial value data.

This study optimizes the scaling process in HASM and uses valley lines extracted from the original DEM with different flow accumulation thresholds as input data. The proposed upscaling method can provide DEMs at different scales for a fixed grid size by changing the flow accumulation threshold.

2 Materials and methods

2.1 The improved HASM upscaling method

High accuracy surface modeling (HASM) is a surface modeling method developed in recent years based on differential geometry theory, which provides a new solution to the error problem and multi-scale problem of surface modeling (Yue, 2011). After decades of development, HASM has developed into a set of methods, including spatial interpolation, data assimilation, data fusion, upscaling, downscaling, etc. (Yue et al., 2016; Yue et al., 2020). As is well known, DEM upscaling refers to increasing the grid size of a DEM. Traditional DEM upscaling methods use spatial interpolation to compute elevation values in the new grid based on spatial autocorrelation. Unlike the traditional methods, the DEM upscaling method based on HASM not only considers the detailed features of the local surface, but also considers the macroscopic variation characteristics of the surface. Data and surface morphology characteristics must be considered in HASM-based DEM upscaling. As shown in Fig. 1, the black solid triangle represents the DEM grid after upscaling, and the black solid circles represent the grid centers of the original DEM. A grid in the upscaled DEM contains several grids of the original DEM. These original DEM grids provide the following information. 1) Range of elevation; typically, the raised grid has the average elevation value of the grids in the original DEM.

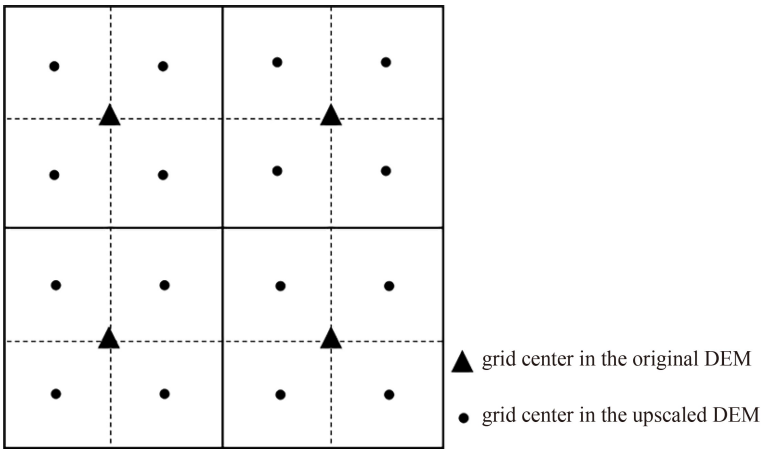

Fig. 1 Locations of the grid centers in the original DEM (black solid triangle) and the grid centers of the upscaled DEM (black solid circle).

2) Maximum elevation; the elevation of the upscaled DEM cannot exceed the elevation in the original DEM. 3) Minimum elevation; the elevation of the upscaled DEM cannot be lower than the elevation in the original DEM.

This information ensures that the upscaled DEM meets the elevation accuracy requirements. The upscaled DEM should accurately describe the terrain, such as the continuity of valleys and ridges. Therefore, a constraint relationship must be established between the grid and its adjacent neighbors. If the average elevation value of the original DEM data in Fig. 1 is M, the maximum elevation value is F, and the minimum elevation value is E, then the elevation value of the raised grid (H) should meet the following conditions:

$$E_{\text{¡}}^{\cdot\cdot 1}H = M + \Delta_{\text{¡}}^{\cdot\cdot 1}F, \tag{1}$$

where Δ is the difference between the grid value of the upscaled DEM and the mean value of the grids in the original DEM. $M+\Delta$ is used in the HASM, because M is known. Thus, the DEM upscaling calculation can be completed as long as Δ is known.

The proposed method integrates the upscaling process and does not execute the two-step strategy of the traditional HASM, i.e., using the HASM to calculate the trend surface, calculating the residual surface, and overlaying the two surfaces. In addition, the accuracy control points can be selected based on the theory of terrain feature analysis to ensure that the upscaled DEM accurately reflects the terrain features.

The optimization of the HASM upscaling model is performed as follows. Three macro-morphology control equations are constructed for each grid:

$$\begin{cases} f_{xx} = \Gamma_{11}^{1} f_x + \Gamma_{11}^{2} f_y + \dfrac{L}{\sqrt{E+G-1}} \\ f_{yy} = \Gamma_{22}^{1} f_x + \Gamma_{22}^{2} f_y + \dfrac{N}{\sqrt{E+G-1}} \\ f_{xy} = \Gamma_{12}^{1} f_x + \Gamma_{12}^{2} f_y + \dfrac{M}{\sqrt{E+G-1}} \end{cases}. \tag{2}$$

The variables in Eq. (2) have been defined in Zhao et al. (2014). The terms on the right side of Eq. (2) are calculated using the original DEM, and those on the left side are expressed as follows:

$$(f_{xx})_{(i,j)} \approx \begin{cases} \dfrac{2\left(fm_{0,j}+\Delta f_{0,j}\right)-5\left(fm_{1,j}+\Delta f_{1,j}\right)+4\left(fm_{2,j}+\Delta f_{2,j}\right)-\left(fm_{3,j}+\Delta f_{3,j}\right)}{h^2}, & i=0 \\ \dfrac{\left(fm_{i+1,j}+\Delta f_{i+1,j}\right)-2\left(fm_{i,j}+\Delta f_{i,j}\right)+\left(fm_{i-1,j}+\Delta f_{i-1,j}\right)}{2h^2}, & i=1,\ldots,I \\ \dfrac{2\left(fm_{I+1,j}+\Delta f_{I+1,j}\right)-5\left(fm_{I,j}+\Delta f_{I,j}\right)+4\left(fm_{I-1,j}+\Delta f_{I-1,j}\right)-\left(fm_{I-2,j}+\Delta f_{I-2,j}\right)}{h^2}, & i=I+1 \end{cases}, \tag{3}$$

$$(f_{yy})_{(i,j)} \approx \begin{cases} \dfrac{2(fm_{i,0}+\Delta f_{i,0})-5(fm_{i,1}+\Delta f_{i,1})+4(fm_{i,2}+\Delta f_{i,2})-(fm_{i,3}+\Delta f_{i,3})}{h^2}, & j=0 \\ \dfrac{\left(fm_{i,j+1}+\Delta f_{i,j+1}\right)-2\left(fm_{i,j}+\Delta f_{i,j}\right)+\left(fm_{i,j-1}+\Delta f_{i,j-1}\right)}{2h^2}, & j=1,\ldots,J \\ \dfrac{2(fm_{i,J+1}+\Delta f_{i,J+1})-5(fm_{i,J}+\Delta f_{i,J})+4(fm_{i,J-1}+\Delta f_{i,J-1})-(fm_{i,J-2}+\Delta f_{i,J-2})}{h^2}, & j=J+1 \end{cases}, \tag{4}$$

$$(f_{xy})_{(i,j)} \approx \begin{cases} \dfrac{(fm_{1,1}+\Delta f_{1,1})-(fm_{1,0}+\Delta f_{1,0})-(fm_{0,1}+\Delta f_{0,1})+(fm_{0,0}+\Delta f_{0,0})}{h^2}, & i=0, j=0 \\ \dfrac{(fm_{1,J+1}+\Delta f_{1,J+1})+(fm_{0,J}+\Delta f_{0,J})-(fm_{1,J}+\Delta f_{1,J})-(fm_{0,J+1}+\Delta f_{0,J+1})}{h^2}, & i=0, j=J+1 \\ \dfrac{(fm_{1,J+1}+\Delta f_{1,J+1})-(fm_{0,J+1}+\Delta f_{0,J+1})+(fm_{0,J-1}+\Delta f_{0,J-1})-(fm_{1,J-1}+\Delta f_{1,J-1})}{2h^2}, & i=0, j=1,\ldots,J \\ \dfrac{(fm_{I+1,1}+\Delta f_{I+1,1})-(fm_{I,0}+\Delta f_{I,0})-(fm_{I,1}+\Delta f_{I,1})+(fm_{I+1,0}+\Delta f_{I+1,0})}{h^2}, & i=I+1, j=0 \\ \dfrac{(fm_{I,J}+\Delta f_{I,J})-(fm_{I+1,J}+\Delta f_{I+1,J})-(fm_{I,J+1}+\Delta f_{I,J+1})+(fm_{I+1,J+1}+\Delta f_{I+1,J+1})}{h^2}, & i=I+1, j=J+1 \\ \dfrac{(fm_{I+1,j+1}+\Delta f_{I+1,j+1})-(fm_{I,j+1}+\Delta f_{I,j+1})+(fm_{I,j-1}+\Delta f_{I,j-1})-(fm_{I+1,j-1}+\Delta f_{I+1,j-1})}{2h^2}, & i=I+1, j=1,\ldots,J \\ \dfrac{(fm_{i+1,1}+\Delta f_{i+1,1})-(fm_{i+1,0}+\Delta f_{i+1,0})+(fm_{i-1,0}+\Delta f_{i-1,0})-(fm_{i-1,1}+\Delta f_{i-1,1})}{2h^2}, & i=1,\ldots,I, j=0 \\ \dfrac{(fm_{i+1,J+1}+\Delta f_{i+1,J+1})-(fm_{i+1,J}+\Delta f_{i+1,J})+(fm_{i-1,J}+\Delta f_{i-1,J})-(fm_{i-1,J+1}+\Delta f_{i-1,J+1})}{2h^2}, & i=1,\ldots,I, j=J+1 \\ \dfrac{(fm_{i,j+1}+\Delta f_{i,j+1})-(fm_{i-1,j+1}+\Delta f_{i-1,j+1})+(fm_{i-1,j-1}+\Delta f_{i-1,j-1})-(fm_{i+1,j-1}+\Delta f_{i+1,j-1})}{4h^2}, & i=1,\ldots,I, j=1,\ldots,J \end{cases}, \tag{5}$$

where $fm_{i,j}$ is the mean value of the grid in the original DEM, $\Delta f_{i,j}$ is the difference between the calculated value of the grid after upscaling and $fm_{i,j}$, which is the unknown variable in the equations, and h is the grid size after upscaling.

After obtaining the three equations in Eq. (2) for each grid, they are expressed as follows:

$$\begin{cases} \boldsymbol{A}z = \boldsymbol{d} \\ \boldsymbol{B}z = \boldsymbol{q} \\ \boldsymbol{C}z = \boldsymbol{p} \end{cases}, \tag{6}$$

where z is a vector composed of the unknowns in each grid; A, B, and C are the coefficient matrices of the unknowns of the three equations, respectively; d, q, and p are the constant vectors on the right side of the three equations, respectively.

Control points are obtained from the original DEM based on the required sampling rate. For the convenience of calculation, the center points of the grids of the upscaled DEM are used. Thus, the following equation is established for each sampling point:

$$fm_{i,j} + \Delta f_{i,j} = f_{i,j}^{s}, \tag{7}$$

where $fm_{i,j}$ is the mean value of the grid in the original DEM, $\Delta f_{i,j}$ is the difference between the calculated value of the grid after upscaling and $fm_{i,j}$, and $f_{i,j}^{s}$ is the elevation value of the sampling point.

Eq. (8) is established after considering all sampling points:

$$\boldsymbol{S}z = \boldsymbol{k}. \tag{8}$$

where S is the coefficient matric of the unknowns of the equation, and k is the constant vector on the right side of the equation.

Eqs. (6) and (8) are combined and then converted into the following equations:

$$\begin{cases} \min \left\| \begin{matrix} \boldsymbol{A} \\ \boldsymbol{B} \\ \boldsymbol{C} \end{matrix} z - \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{q} \\ \boldsymbol{p} \end{bmatrix} \right\|_2 \\ s.t. \quad \boldsymbol{S}z = \boldsymbol{k} \end{cases}. \tag{9}$$

The original DEM provides the maximum and minimum elevations of the DEM after upscaling. Therefore, Eq. (9) can be expressed as

$$\begin{cases} \min \left\| \begin{matrix} \boldsymbol{A} \\ \boldsymbol{B} \\ \boldsymbol{C} \end{matrix} z - \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{q} \\ \boldsymbol{p} \end{bmatrix} \right\|_2 \\ s.t. \quad \begin{matrix} \boldsymbol{S}z = \boldsymbol{k} \\ \boldsymbol{E} \leqslant z \leqslant \boldsymbol{F} \end{matrix} \end{cases}. \tag{10}$$

Eq. (10) can be transformed into Eq. (11) for processing convenience:

$$\mathbf{W}\mathbf{X} = \mathbf{V}. \tag{11}$$

where

$$\mathbf{W} = \mathbf{A}^{\mathrm{T}}\boldsymbol{A} + \mathbf{B}^{\mathrm{T}}\boldsymbol{B} + \mathbf{C}^{\mathrm{T}}\boldsymbol{C} + \mathbf{S}^{\mathrm{T}}\boldsymbol{S},$$
$$\mathbf{V} = \mathbf{A}^{\mathrm{T}}\boldsymbol{d} + \mathbf{B}^{\mathrm{T}}\boldsymbol{q} + \mathbf{C}^{\mathrm{T}}\boldsymbol{p} + \mathbf{S}^{\mathrm{T}}\boldsymbol{k}.$$

The solution of Eq. (11) is then assigned to the grids of the upscaled DEM. Due to the large scale of Eq. (11), iterative methods are needed to solve it. Modified Gauss-Seidel method (MGS), multi-grid method (MG), adaptive method (AM), adjustment method (AD), and preconditioned conjugate gradient method (PCG) were used to solve the linear system (10) (Yue et al., 2010; Yue, 2011; Chen et al., 2012). The results show that although the PCG method requires more computer storage and

computing time, it achieves the highest accuracy (Yue, 2011). With the development of high-performance computing, more attention has been paid to simulation accuracy rather than computing speed and memory usage. Therefore, we use the PCG method to solve the linear system in this study.

2.2 Extraction of valleys from the DEM

The surface overland flow model is typically used to extract valleys from DEMs (O'callaghan and Mark, 1984) by simulating water flowing over the land surface. This model includes several main steps (Tribe, 1992). First, the DEM should be filled to eliminate depressions on the surface, and the flow direction of each grid is determined by calculating the difference in elevation between the center grid and its adjacent grids. Then, the flow path of water over the land surface is calculated. Second, the amount of water flowing through each grid is quantified according to the flow path, which is named as flow accumulation; A grid is considered part of a valley when its flow accumulation value is greater than a threshold. The final step is to find an optimal threshold that can identify all valleys in the study area. This method uses hydrological confluence to identify the flow path and simulate surface runoff. The algorithm is straightforward and generates continuous flow paths. Consequently, this model and its improvements have been widely used to extract valleys from DEMs (Jenson and Domingue, 1988; Tarboton et al., 1991; Turcotte et al., 2001). The key step of the model is to determine a suitable flow accumulation threshold, which determines the starting positions of the valleys. The number and distribution of valleys in a region differ for different flow accumulation thresholds.

Since the objective of this study is to extract valleys in the research area, it is necessary to determine the most appropriate flow accumulation threshold (Tarboton, 1997; Lin et al., 2006). However, because different valleys are used as input data for the DEM upscaling model, it is not necessary to determine the optimal threshold because different valleys with different flow accumulation thresholds are needed.

2.3 Calculating the surface complexity index

Different indices have been used in digital terrain analysis to evaluate surface morphological characteristics at different scales, such as slope, curvature, undulation, etc. (Tang, 2014). The result of DEM upscaling is to remove the small fluctuations of different degrees of the surface, so a parameter is needed to effectively describe the small fluctuation characteristics of the surface. However, there is no effective index to assess this type of surface morphology. Therefore, this paper proposes the surface complexity (SC) index to quantify the difference in surface morphology before and after DEM upscaling.

A grid and its 8 nearest neighbors are selected (Fig. 2). A line connecting the center points of three grids forms an angle in three-dimensional space. A large angle indicates flat terrain in the specified direction, and a small angle indicates complex terrain. Therefore, the SC index can be quantified by the sinusoidal value of the angle. If the spatial coordinates of the three grids are $P_1(x_1, y_1, z_1)$, $P_5(x_5, y_5, z_5)$, and $P_9(x_9, y_9, z_9)$, the sinusoidal value of the angle can be computed as follows:

$$\overrightarrow{P_1P_5} = (a_1, b_1, c_1) = (x_5 - x_1, y_5 - y_1, z_5 - z_1), \tag{12}$$

$$\overrightarrow{P_5P_9} = (a_2, b_2, c_2) = (x_9 - x_5, y_9 - y_5, z_9 - z_5), \tag{13}$$

$$\cos\theta = \frac{a_1.a_2 + b_1.b_2 + c_1.c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2}.\sqrt{a_2^2 + b_2^2 + c_2^2}}, \tag{14}$$

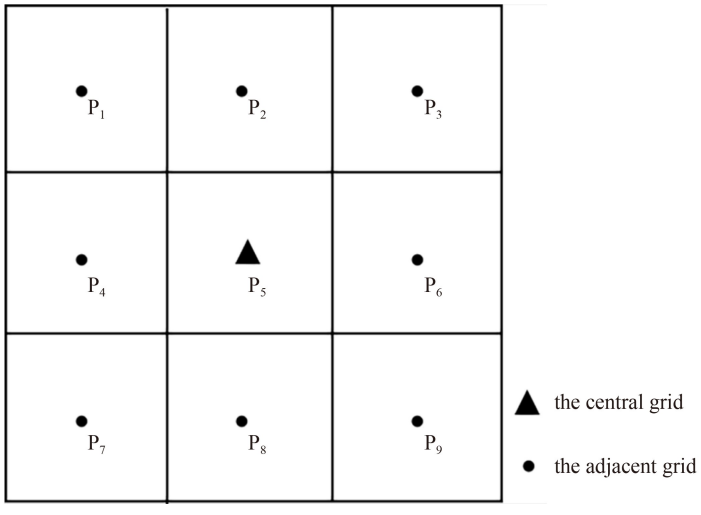

Fig. 2 The series number of central grid and its adjacent grid.

$$\sin\theta = \sqrt{1-(\cos\theta)^2}. \quad (15)$$

One grid is used as the center, and the four directions are determined. The sinusoidal values of the angles in the four directions are computed, and the values are considered the SC index of the grid.

$$T_c = \sin\theta_1 + \sin\theta_2 + \sin\theta_3 + \sin\theta_4. \quad (16)$$

The results is normalized to enable the comparison of the SC index in different regions:

$$\text{SC-Index} = \frac{T_c - \min}{\max - \min}. \quad (17)$$

The maximum value of the SC Index in any direction is 1 (if the angle is 90°), and the minimum value is 0 (if the angle is 180°). Therefore, the maximum and minimum values are, respectively, 4 and 0:

$$\text{SC-Index} = \frac{T_c}{4}. \quad (18)$$

2.4 Study area and data

Jiuyuangou Basin located in Suide County, Shaanxi Province, China, was selected as the test area. This area is located in the hilly region of the Loess Plateau in northern Shaanxi. The altitude range is 985−1491 m, and the relative height difference is 506 m. This area has a flow erosion landform, the gully system is well developed, and belongs to hilly terrain. It is an ideal area to evaluate DEM upscaling. The topography of the test area is shown in Fig. 3. The experimental data used in this study are DEMs produced by the local surveying and mapping department, and the grid size is 5 m.

3 Results

3.1 Upscaling results

The original DEM (5 m grid) was scaled to 30 m and 90 m using the proposed DEM upscaling method. For each grid size, valleys corresponding to flow accumulation thresholds of 500, 5000, and 15000 were selected. Upscaled DEMs with both grid sizes are shown in Figs. 4 and 5. The level of detail is significantly higher for DEM 30 m than for DEM 90 m. As the flow accumulation threshold increases, the number of valleys entering the upscaling model and the output in the upscaled DEM decreases. Since the valleys are formed by runoff erosion in the geomorphological evolution of this region, the proposed upscaling method can be regarded as an anti-erosion process. The larger the flow accumulation threshold, the lower the degree of erosion of the obtained results is due to the simplification of the terrain.

The SRTM-dem of the studied area was adopted in this research as a reference for comparative analysis, the grid size of the first one is 30 m (Fig. 6(a)), and the grid size of the other one is 90 m (Fig. 6(b)). Compared to SRTM-dem, the upscaled DEMs in this study have two advantages. The first is that the upscaled DEMs in this study can maintain the continuity and independence of the valley terrain very well. However, in the case of SRTM-dem, some adjacent valleys were connected to each other or merged into a whole in the area near the

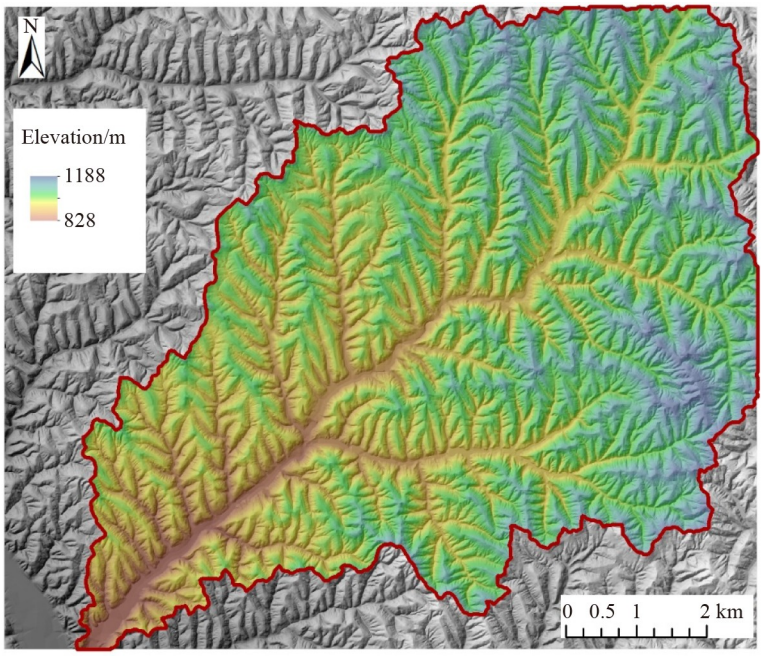

Fig. 3 Topography of the study area.

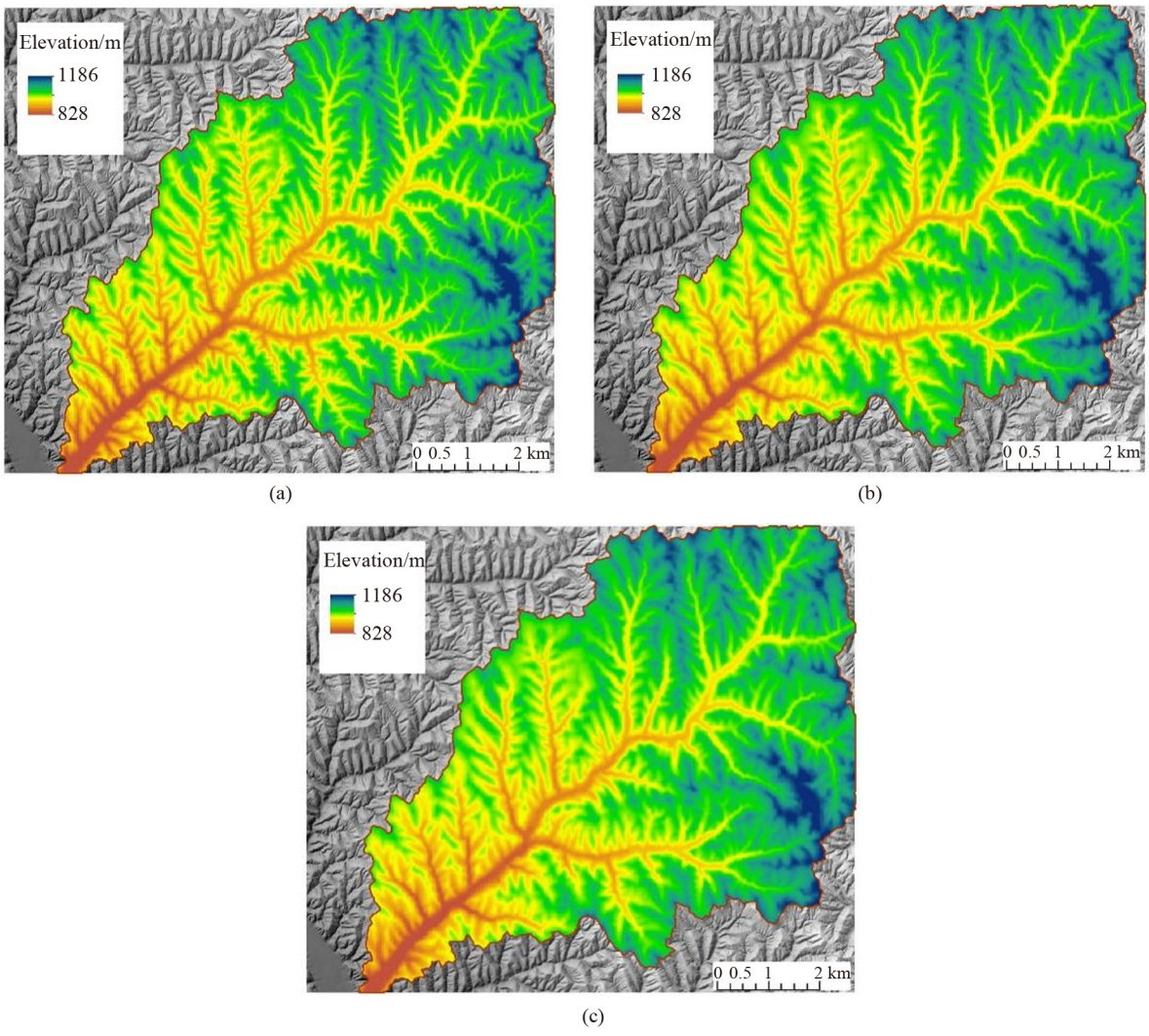

Fig. 4 The upscaling results for a 30 m grid size. (a) Flow accumulation threshold of 500; (b) flow accumulation threshold of 5000; (c) flow accumulation threshold of 15000.

outlet of the watershed, where the valley terrain is well developed, and there are some discontinuous valleys in the upper reaches of the watershed away from the outlet. It is easy to see that a larger grid size leads to more serious problems in the valley terrain. The second advantage of upscaled DEMs in this study is also the main innovation of this study. Upscaled DEMs vary according to the valley used in the upscaling process. The application value of the series of upscaled DEMs is that they can be used to simulate the surface morphology of different periods of landform development, whereas SRTM-dem can only express one surface morphology.

Table 1 provides descriptive statistics and mean slope of original and upscaled DEMs. As the DEM grid size of 5 m is increased to 30 m and 90 m, the minimum value increases, the maximum value decreases, and the mean value increases. The reason is that more valley fillings than peak clippings occur. When the grid size of the upscaled DEM is fixed at 30 m, and the flow accumulation threshold increases, the minimum and maximum values do not change significantly, but the mean value increases. As the grid size of the upscaled DEM and the flow accumulation increase, the standard deviation of the elevation and the mean slope decrease.

To compare the distribution of elevation values of different DEMs well, histograms were generated for the original DEM and all the upscaled DEMs (Fig. 7). It can be seen that the distribution of elevation values is approximately normal for the original DEM and all upscaled DEMs, and the main difference is that the smoothness of the different histograms is different. In terms of the area studied in this paper, the smoothness of the 30 m upscaled DEM is the best, the smoothness of the original DEM is the second, and the smoothness of the 90 m upscaled DEM is the worst. This is because the 30 m upscaled DEM is a small generalization of the original DEM, of which the result is that some small undulating terrains have been removed, so the histogram of the 30 m scaled DEM is smoother. However, when the original DEM was upscaled to 90 m, the elevation disorder of the

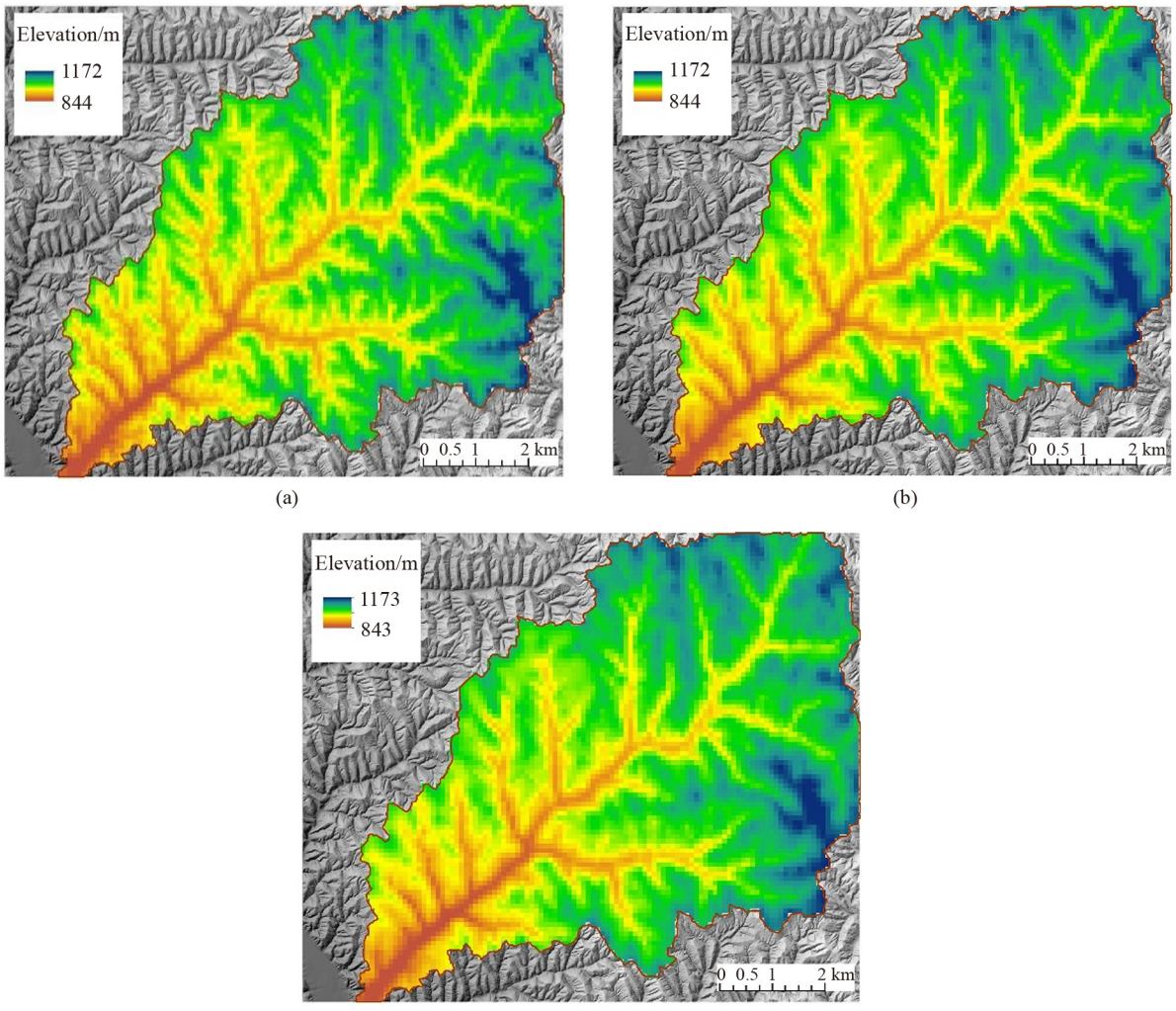

Fig. 5 The upscaling results for a 90 m grid size. (a) Flow accumulation threshold of 500; (b) flow accumulation threshold of 5000; (c) flow accumulation threshold of 15000.

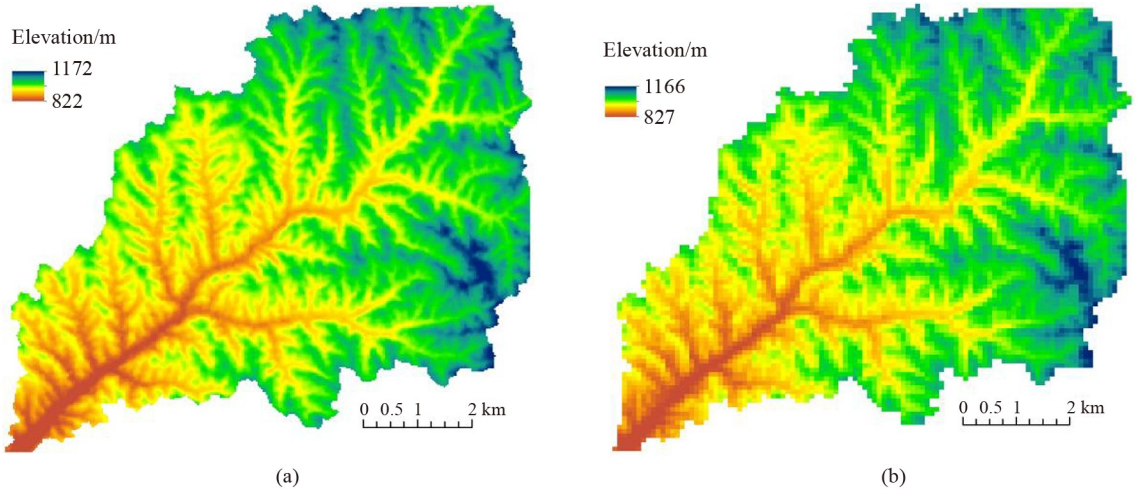

Fig. 6 SRTM-dem of the test area. (a) Grid size is 30 m; (b) grid size is 90 m.

Table 1 Elevation Statistics for Different DEMs

DEM	Min/m	Max/m	Mean/m	Std. Dev.	Slope/(°)
Original DEM	828	1188.3	1003.68	58.24	29.36
Upscaled DEM Grid size:30 m Flow accumulation threshold: 500	827.95	1185.78	1004.37	58.34	22.60
Upscaled DEM Grid size:30 m Flow accumulation threshold: 5000	828.08	1185.74	1007.18	58.21	19.87
Upscaled DEM Grid size:30 m Flow accumulation threshold: 15000	828.08	1185.74	1008.09	57.77	18.79
Upscaled DEM Grid size:90 m Flow accumulation threshold: 500	843.96	1171.78	1004.87	55.72	10.47
Upscaled DEM Grid size:90 m Flow accumulation threshold: 5000	843.41	1173.22	1010.59	56.51	10.16
Upscaled g DEM Grid size:90 m Flow accumulation threshold: 15000	843.40	1173.14	1012.87	56.28	9.44

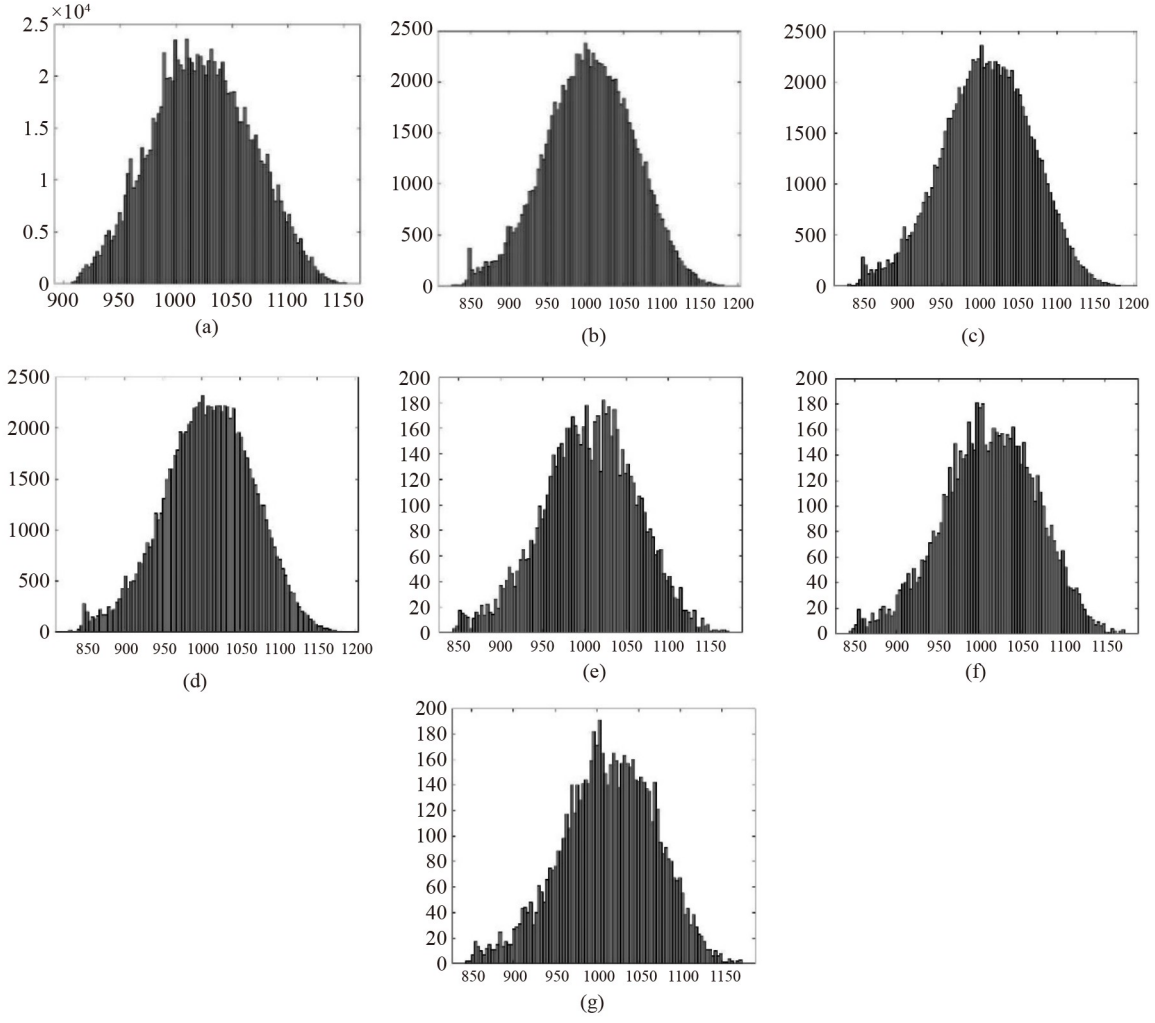

Fig. 7 The histograms of different DEMs. (a) Original DEM; (b) flow accumulation threshold of 500 (grid size: 30 m); (c) flow accumulation threshold of 5000 (grid size: 30 m); (d) flow accumulation threshold of 15000 (grid size: 30 m); (e) flow accumulation threshold of 500 (grid size: 90 m); (f) flow accumulation threshold of 5000 (grid size: 90 m); (g) flow accumulation threshold of 15000 (grid size: 90 m).

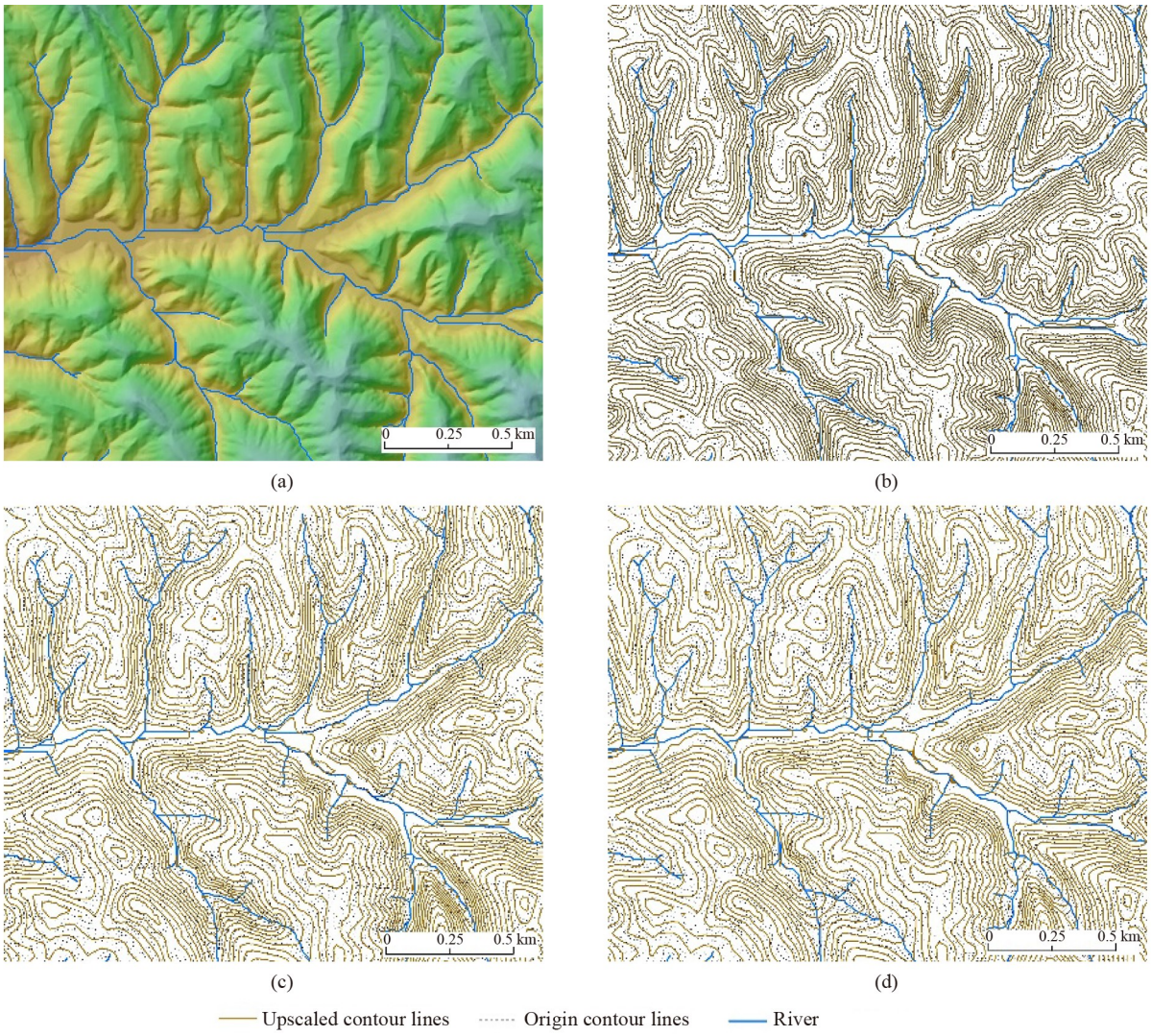

Fig. 8 Contour lines extracted from the upscaled DEM (grid size: 30 m). (a) DEM hillshade; (b) flow accumulation threshold of 500; (c) flow accumulation threshold of 5000; (d) flow accumulation threshold of 15000.

upscaled DEM is significant, so the smoothness of 90 m upscaled DEMs' histograms is also poor.

3.2 Contour maps

Contour maps are critical to describing elevation changes on the earth's surface. The curvature of the contour lines and the relationship between the different contour lines provide information on the terrain. The contour lines extracted from the 30 m and 90 m upscaled DEMs (with flow accumulation thresholds of 500, 5000, and 15000) are shown in Figs. 8 and 9. The blue lines are the valley lines of the test area extracted from the original DEM with a flow accumulation threshold of 300. The largest valley in the small watershed is defined as the first-level valley, and its tributaries are defined as the second-level valleys. Its tributaries are defined as third-level valleys. The main valleys in the test area can be seen on the contour map at a grid size of 30 m. The contour lines in the DEM with a flow accumulation threshold of 500 show the first-level, second-level, and third-level valleys. At a flow accumulation threshold of 5000, the first-level and second-order valleys can be seen, but most of the third-order valleys cannot be detected, or the curvature of the contour lines is difficult to see. Thus, the valley appears as it was before the erosion. These changes are more pronounced at a flow accumulation threshold of 15000, and some of the smaller secondary valleys cannot be detected. Only the first-level valley can be seen at a grid size of 90 m, and the terrain has fewer fluctuations.

3.3 Surface complexity index

The main purpose of DEM upscaling is data generalization, and terrain complexity is an effective measure of DEM generalization. Therefore, a quantitative measure of DEM terrain complexity is essential to evaluate the upscaling results. The SC indices of different

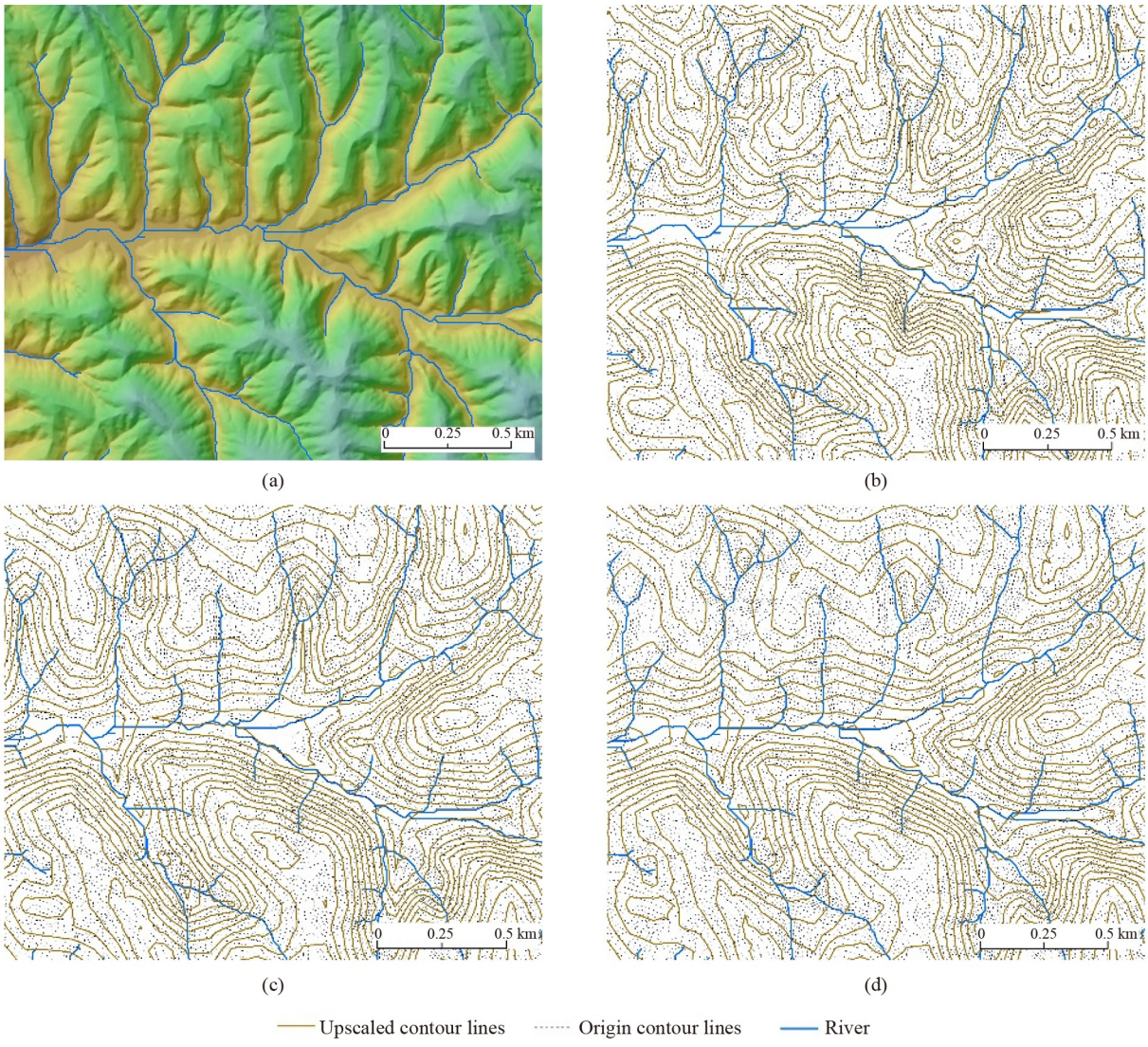

Fig. 9 Contour lines extracted from the upscaled DEM (grid size: 90 m). (a) DEM hillshade; (b) flow accumulation threshold of 500; (c) flow accumulation threshold of 5000; (d) flow accumulation threshold of 15000.

upscaled DEMs were calculated, and the influence of grid size and flow accumulation threshold were analyzed. Figures 10−12 show the SC indix of the original DEM and the 30 m and 90 m upscaled DEMs for flow accumulation thresholds of 500, 5000, and 15000.

It can be seen from Figs. 10−12 that areas with high terrain complexity are located on both sides of the valleys. The SC index is higher on both sides of the valley near the outlet of the watershed than at the top of the watershed. The SC index is higher in regions with irregular valley shape than in regions with regular valley shape. This finding is accurate because the study area has undergone extensive flow erosion. In general, the erosion level is higher in the downstream region, and the SC is higher. A curved valley has complex terrain, resulting in a higher SC index. Upscaled DEM with a grid size of 30 m shows high complexity along the valley. As the flow accumulation threshold increases from 500 to 15000, the valley shrinks, and the number of areas with a high SC index decreases, indicating that the proposed upscaling method is suitable for terrains with high complexity. Upscaled DEM with a grid size of 90 m shows high complexity, but there are fewer terrain fluctuations due to the larger grid size.

Figures 11 and 12 show difference in the spatial distribution of the SC index for the upscaled DEMs with different grid sizes. Table 2 lists the descriptive statistics of the SC index for different DEMs (the minimum is 0 for all results; therefore, it is not listed). The SC index of the upscaled DEM is significantly lower than that of the original DEM (5 m grid size). For example, the maximum SC index of the original DEM is 4.665, where it is about 3 (30 m grid size) and 1.5 (90 m grid size) after upscaling. The mean SC index is also lower after DEM upscaling. In addition, the mean SC index decreases as the flow accumulation threshold increases for both grid

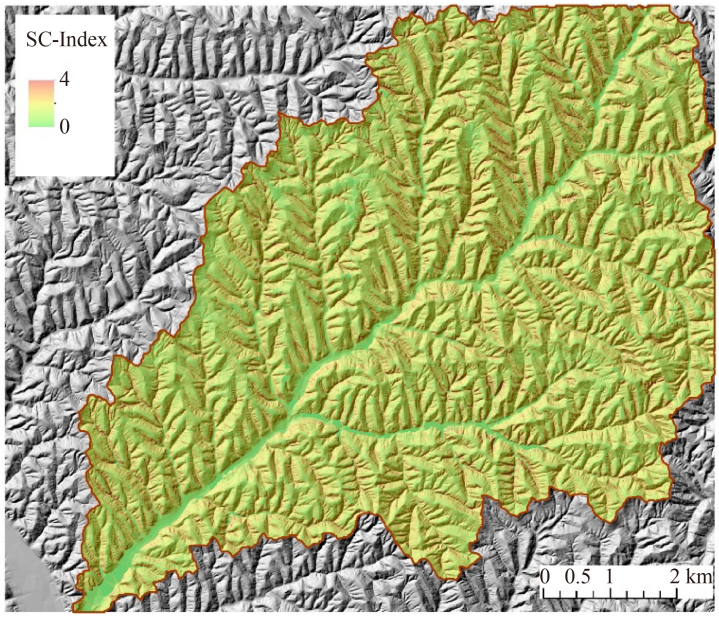

Fig. 10 SC index of the original DEM.

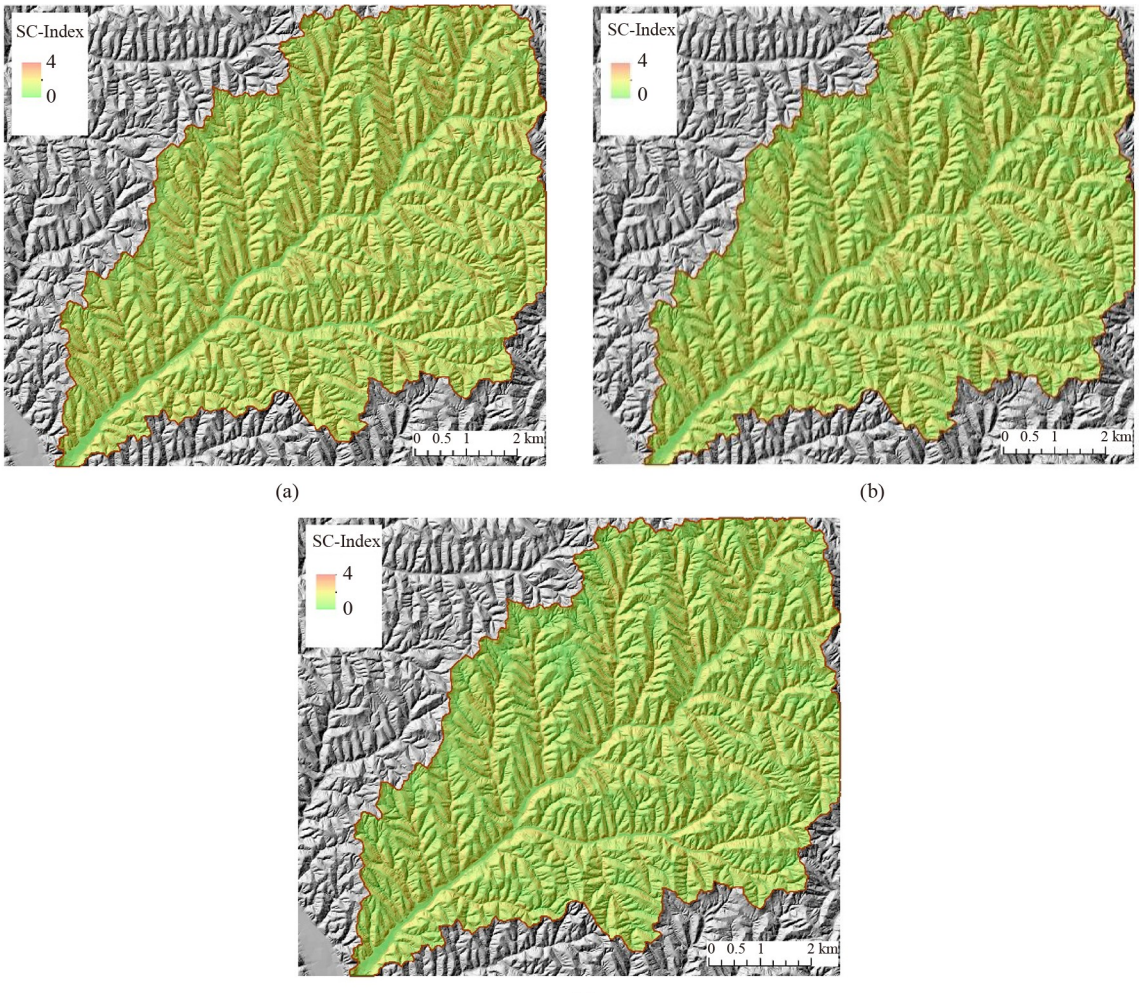

Fig. 11 SC index of the upscaled DEM (grid size: 30 m). (a) Flow accumulation threshold of 500; (b) flow accumulation threshold of 5000; (c) flow accumulation threshold of 15000.

Table 2 Descriptive statistics of the SC Index for Different DEMs (m)

DEM	Max	Mean	Std
Original DEM	4.665	1.533	0.670
Upscaled DEM Grid size:30 m Flow accumulation threshold: 500	3.094	1.297	0.528
Upscaled DEM Grid size:30 m Flow accumulation threshold: 5000	3.079	1.112	0.540
Upscaled DEM Grid size:30 m Flow accumulation threshold: 15000	2.945	1.041	0.530
Upscaled DEM Grid size:90 m Flow accumulation threshold: 500	1.502	0.622	0.321
Upscaled DEM Grid size:90 m Flow accumulation threshold: 5000	1.482	0.586	0.308
Upscaled DEM Grid size:90 m Flow accumulation threshold: 15000	1.419	0.543	0.303

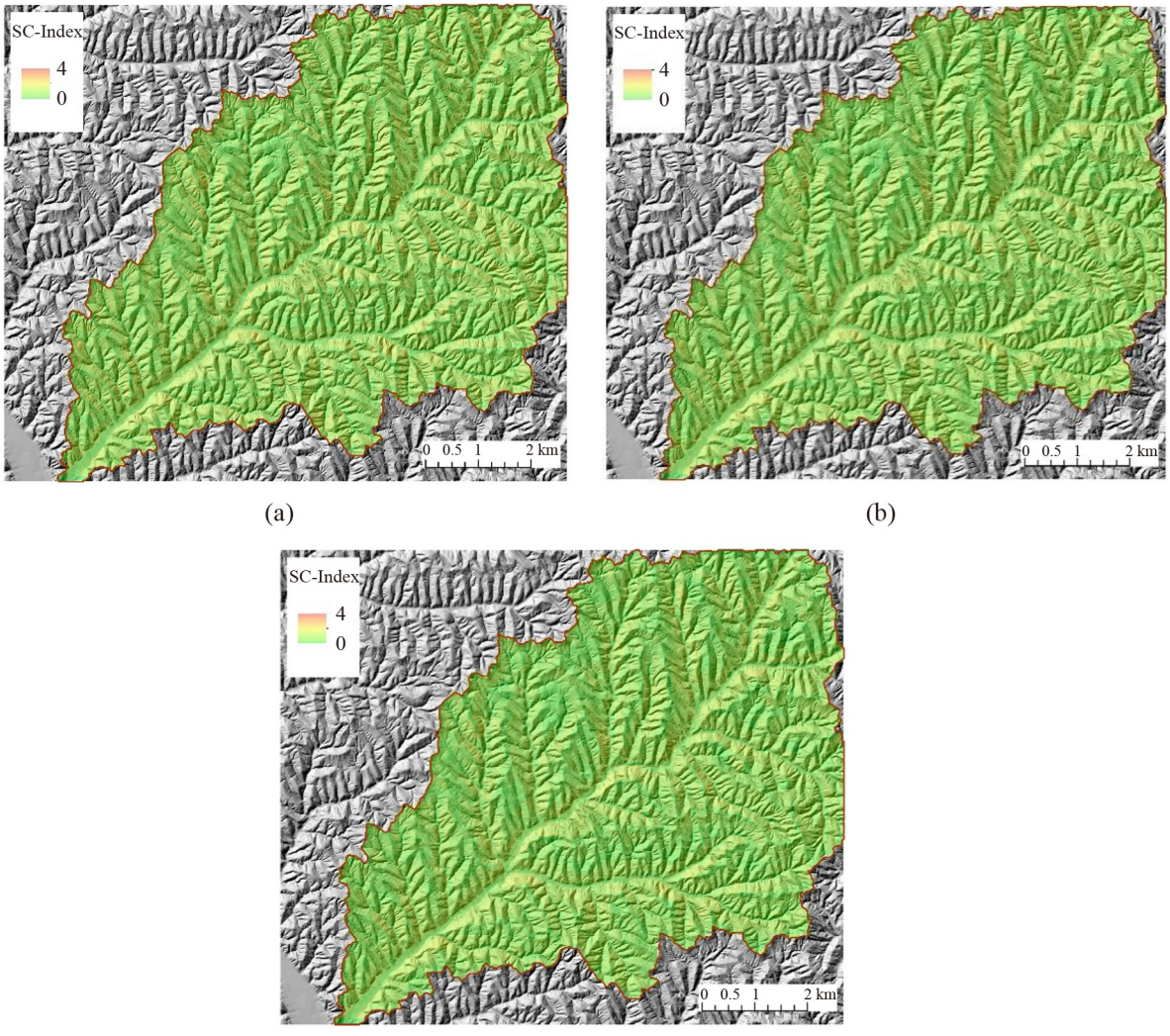

Fig. 12 SC index of the upscaled DEM (grid size: 90 m). (a) Flow accumulation threshold of 500; (b) flow accumulation threshold of 5000; (c) flow accumulation threshold of 15000.

sizes. Compared to the mean value, the difference in the standard deviation between different flow accumulation thresholds is relatively small. The original DEM has the largest standard deviation, followed by 30 m and 90 m grids. However, the standard deviation increases and decreases as the flow accumulation increases for a grid size of 30 m.

Figure 13 shows the histograms of the SC index calculated from the original DEM and each upscaled DEM. It can be seen that the shape of the SC index histogram of the original DEM data is obviously to the right. That is, the distribution on the left is steep and the right is relatively flat. For upscaled DEMs with a grid size of 30 m, the SC index histograms are basically normal, while the histograms show a rightward shape when the grid size increases to 90 m, and the histograms are less smooth for the latter upscaled results, which are different from the original one. The reasons for the differences in the morphology of different histograms can be explained as follows: The original DEM (grid size is 5 m) can better simulate the actual surface, therefore, although the overall slope of the studied area is large, the areas with dramatic terrain changes are still few. That is, the total number of grids with large complexity is small, so the skewness of the corresponding histogram is right. The number of total grids decreases when the grid size becomes 30 m, but the area of valleys and other areas with complex terrain changes decreases less due to the integration of valley information in the upscaling process, so the histograms of these upscaled results are normally distributed. When the grid size continues to increase to 90 m, the total number of grids is further reduced, and the elevation between different grids changes greatly, so the smoothness of the histograms is poor.

A threshold sequence is analyzed to determine the threshold's influence on DEM upscaling. Thresholds of

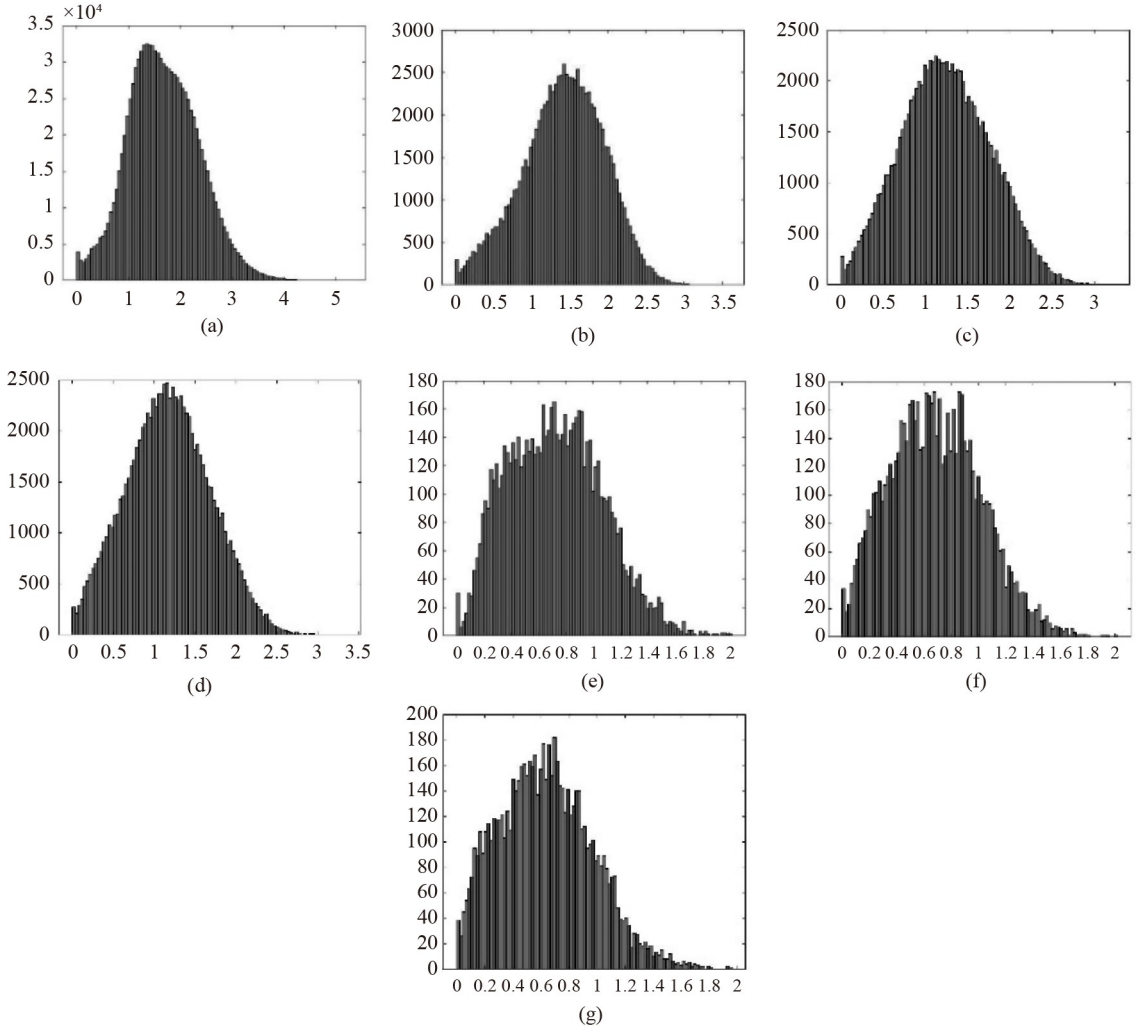

Fig. 13 The histograms of SC index from different DEMs. (a) original DEM; (b) flow accumulation threshold of 500 (grid size: 30 m); (c) flow accumulation threshold of 5000 (grid size: 30 m); (d) flow accumulation threshold of 15000 (grid size: 30 m); (e) flow accumulation threshold of 500 (grid size: 90 m); (f) flow accumulation threshold of 5000 (grid size: 90 m); (g) flow accumulation threshold of 15000 (grid size: 90 m).

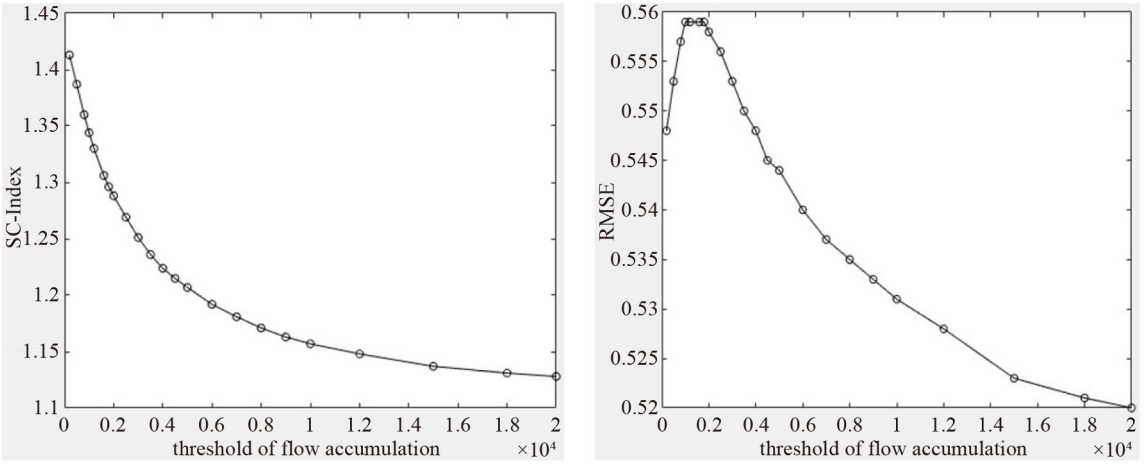

Fig. 14 Mean (left) and RMSE (right) of SC index for different flow accumulation thresholds (grid size: 30 m).

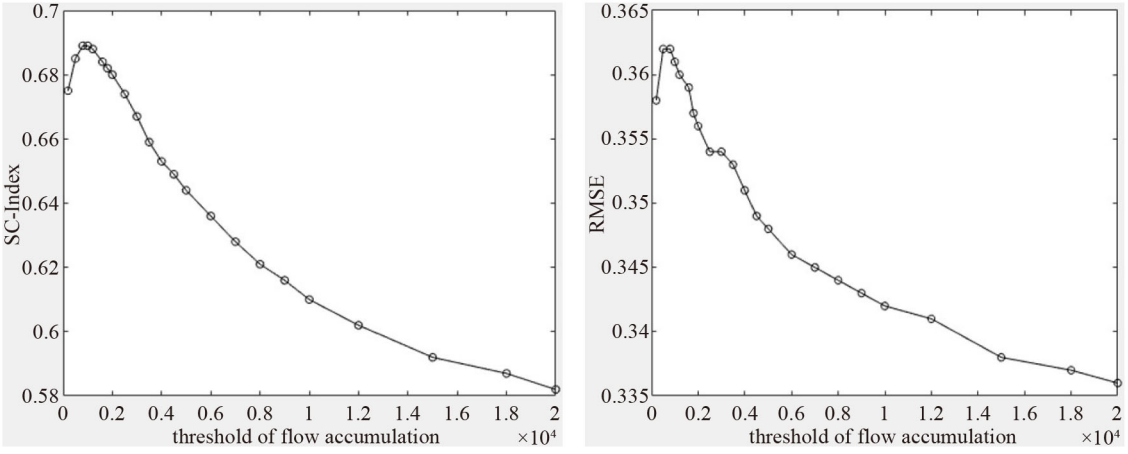

Fig. 15 Mean (left) and RMSE (right) of SC index for different flow accumulation thresholds (grid size: 90 m).

200, 500, 800, 1000, 1200, 1500, 1800, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 6000, 7000, 8000, 9000, 10000, 12000, 15000, 18000, and 20000 are used for DEM grid sizes of 30 m and 90 m, and the SC index is calculated. The mean and root mean square error (RMSE) of the SC index are shown in Figs. 14 and 15.

Trends in the mean SC index for grid sizes of 30 m and 90 m are similar as the flow accumulation threshold increases, but there are some differences in local areas. For a grid size of 30 m, as the flow accumulation threshold increases, the mean SC index decreases rapidly and then stabilizes at a value between 1.1 and 1.15. For a grid size of 90 m, the mean SC index increases first and then decreases with an increase in the flow accumulation threshold. The rate of decline is relatively stable (i.e., approximately linear), and the value eventually stabilizes. The RMSE of the SC index shows similar trends for the 30 m and 90 m grid sizes as the flow accumulation threshold increases, i.e., increase followed by decrease. The only difference is a more gradual change in the RMSE for the 30 m grid size.

The variation characteristics of the SC index as the flow accumulation threshold increases can be explained as follows: A larger flow accumulation threshold means less valley information available for DEM upscaling, so more undulating terrain is abandoned after upscaling, so the complexity of the terrain expressed by the upscaled DEM is generally reduced. However, since available valleys tend to be stable in the series of flow accumulation thresholds, the final SC index is also gradually stable.

4 Discussion and conclusions

The proposed DEM upscaling method was inspired by two facts. First, increasing DEM's grid size in upscaling. However, when the grid size is constant, it does not mean that the DEM scale is fixed. There should still be a scale sequence, and the difference in this scale sequence reflects the key terrain features. Second, valleys are the key terrain features at different scales.

The proposed DEM upscaling method based on HASM is characterized by two innovations. The first is an

improved HASM upscaling method according to the change in elevation before and after DEM upscaling. This approach is reasonable and straightforward. The second is the extraction of valleys at different scales to constrain topographic features. This strategy allows DEMs to be extracted at different scales with the same grid size for different applications. We developed the SC index to quantify the complexity of topographic features of DEMs with different scales. The main conclusions of this paper are as follows.

1) Changing the resolution is not the only DEM upscaling approach. When the grid size of the DEM is fixed, DEMs with different complexity of surface information can be obtained by extracting valleys with different flow accumulation thresholds using the proposed upscaling method. As the flow accumulation threshold increases, the terrain information is simplified when upscaled.

2) As the DEM grid size increases, the minimum DEM value increases, and the maximum value decreases because more valley filling than peak clipping occurs. However, the mean elevation values of the different DEMs show that valley filling is dominant; therefore, the mean value increases with increasing grid size. Standard elevation deviation and mean slope value decrease with increasing grid size. In DEMs with the same grid size, the two parameters also decrease with an increase in the flow accumulation threshold. Similar conclusions can also be drawn from the contours extracted from the different DEMs.

3) A SC index that considers the angle between adjacent grids in the DEM in three-dimensional space was proposed to describe the complexity of surface morphology. As DEM grid size increases, the mean and standard deviation of the SC index decreases significantly, indicating that an increase in the DEM grid size reduces SC index. However, when the grid size is fixed and the flow accumulation threshold increases, SC index trends vary for different grid sizes. At grid size of 30 m, the mean SC index initially decreased sharply, followed by a gentler decline as the flow accumulation increased from 100 to 20000. The RMSE of the SC index increases first and then decreases rapidly, followed by stabilization. At a grid size of 90 m, the mean SC index and its RMSE increase first, then decrease and stabilize.

Acknowledgments We are thankful for all of the helpful comments provided by the reviewers. This research was supported by Anhui Province Universities Outstanding Talented Person Support Project (No. gxyq2022097); Major Project of Natural Science Research of Anhui Provincial Department of Education (Nos. 2022AH040150, KJ2021ZD0130, KJ2021ZD0131); the National Natural Science Foundation of China (Grant No. 42071374); the guiding plan project of Chuzhou Science and Technology Bureau (No. 2021ZD008); "113" Industry Innovation Team of Chuzhou City in Anhui Province.

Competing interests The authors declare that they have no competing interests.

References

Aguilar F J, Aguera F, Aguilar M A, Carvajal F (2005). Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. Photogramm Eng Remote Sensing, 71(7): 805–816

Ai T, Li J (2010). A DEM generalization by minor valley branch detection and grid filling. ISPRS J Photogramm Remote Sens, 65(2): 198–207

Ai T, Liu Y, Huang Y (2007). The hierarchical watershed partitioning and generalization of river network. Acta Geodetica Cartograph Sin, 36(2): 231–236,243

Burrough P A, McDonnell R A, Lloyd C D (1988). Principles of Geographical Information Systems. New York: Oxford University Press

Chen C F, Yue T X, Li Y Y (2012). A high speed method of SMTS. Comput Geosci, 41: 64–71

Chen C, Li Y (2013). An orthogonal least-square-based method for DEM generalization. Int J Geogr Inf Sci, 27(1): 154–167

Chen Y, Ma T, Chen X, Chen Z, Yang C, Lin C, Shan L (2016). A new DEM generalization method based on watershed and tree structure. PLoS One, 11(8): e0159798

Chen Z, Ma X, Yu W, Wu L (2020). An integrated graph Laplacian downsample (IGLD)-based method for DEM generalization. Earth Sci Inform, 13(4): 973–987

Fisher P F (1991). Firest experiments in viewshed uncertainty-The accuracy of the viewshed area. Photogramm Eng Remote Sensing, 57(10): 1321–1327

Horritt M S, Bates P D (2001). Effects of spatial resolution on a raster based model of flood flow. J Hydrol (Amst), 253(1–4): 239–249

Jenson S K, Domingue J O (1988). Extracting topographic structure from digital elevation data for geographic information-system analysis. Photogramm Eng Remote Sensing, 54(11): 1593–1600

Kawabata D, Bandibas J (2009). Landslide susceptibility mapping using geological data, a DEM from ASTER images and an Artificial Neural Network (ANN). Geomorphology, 113(1–2): 97–109

Li J H, Chen W J (2005). A rule-based method for mapping Canada's wetlands using optical, radar and DEM data. Int J Remote Sens, 26(22): 5051–5069

Lin W T, Chou W C, Lin C Y, Huang P H, Tsai J S (2006). Automated suitable drainage network extraction from digital elevation models in Taiwan's upstream watersheds. Hydrol Processes, 20(2): 289–306

Liu X (2008). Airborne LiDAR for DEM generation: some critical issues. Prog Phys Geogr, 32(1): 31–49

Ma T, Chen Y, Hua Y, Chen Z, Chen X, Lin C, Yang C (2017). DEM generalization with profile simplification in four directions. Earth Sci Inform, 10(1): 29–39

Murphy P N C, Ogilvie J, Meng F R, Arp P (2008). Stream network modelling using lidar and photogrammetric digital elevation models: a comparison and field verification. Hydrol Processes,

22(12): 1747–1754

O'Callaghan J F, Mark D M (1984). The extraction of drainage networks from digital elevation data. Comput Vis Graph Image Process, 28(3): 323–344

Raposo P (2020). Variable DEM generalization using local entropy for terrain representation through scale. Intern J Cartograph, 6(1): 99–120

Schoorl J M, Sonneveld M P W, Veldkamp A (2000). Three-dimensional landscape process modelling: the effect of DEM resolution. Earth Surf Process Landf, 25(9): 1025–1034

Tang G (2014). Progress of DEM and digital terrain analysis in China. Acta Geogr Sin, 69(9): 1305–1325

Tarboton D G (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resour Res, 33(2): 309–319

Tarboton D G, Bras R L, Rodriguez-Iturbe I (1991). On the extraction of channel networks from digital elevation data. Hydrol Processes, 5(1): 81–100

Tribe A (1992). Automated recognition of valley lines and drainage networks from grid digital electation models–a review and a new method. J Hydrol (Amst), 139(1–4): 263–293

Turcotte R, Fortin J P, Rousseau A N, Massicotte S, Villeneuve J P (2001). Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network. J Hydrol (Amst), 240(3–4): 225–242

Vaze J, Teng J, Spencer G (2010). Impact of DEM accuracy and resolution on topographic indices. Environ Model Softw, 25(10): 1086–1098

Weibel R (1992). Models and experiments for adaptive computer-assisted terrain generalization. Cartogr Geogr Inf Syst, 19(3): 133–153

Wolock D M, McCabe G J (2000). Differences in topographic characteristics computed from 100- and 1000-m resolution digital elevation model data. Hydrol Processes, 14(6): 987–1002

Wu Q, Chen Y, Wilson J P, Liu X, Li H (2019). An effective parallelization algorithm for DEM generalization based on CUDA. Environ Model Softw, 114: 64–74

Xiong L Y, Li S J, Hu G H, Wang K, Chen M, Zhu A X, Tang G A (2023). Past rainfall-driven erosion on the Chinese loess plateau inferred from archaeological evidence from Wucheng City, Shanxi. Commun Earth Environ, 4(1): 4

Xiong L, Li S, Tang G, Strobl J (2022). Geomorphometry and terrain analysis: data, methods, platforms and applications. Earth Sci Rev, 233: 104191

Xiong L, Tang G, Yang X, Li F (2021). Geomorphology-oriented digital terrain analysis: progress and perspectives. J Geogr Sci, 31(3): 456–476

Yue T X (2011). Surface Modeling: High Accuracy and High Speed Methods. Calabasas: CRC Press

Yue T X, Chen C F, Li B L (2010). An adaptive method of high accuracy surface modeling and its application to simulating elevation surface. Trans GIS, 14(5): 615–630

Yue T X, Du Z P, Song D J, Gong Y (2007). A new method of surface modeling and its application to DEM construction. Geomorphology, 91(1–2): 161–172

Yue T X, Liu Y, Zhao M W, Du Z P, Zhao N (2016). A fundamental theorem of Earth's surface modelling. Environ Earth Sci, 75(9): 751

Yue T X, Zhao N, Liu Y, Wang Y, Zhang B, Du Z P, Fan Z, Shi W, Chen C, Zhao M, Song D, Wang S, Song Y, Yan C, Li Q, Sun X, Zhang L, Tian Y, Wang W, Wang Y, Ma S, Huang H, Lu Y, Wang Q, Wang C, Wang Y, Lu M, Zhou W, Liu Y, Yin X, Wang Z, Bao Z, Zhao M, Zhao Y, Jiao Y, Naseer U, Fan B, Li S, Yang Y, Wilson J P (2020). A fundamental theorem for eco-environmental surface modelling and its applications. Sci China Earth Sci, 63(8): 1092–1112

Zhang R, Bian S, Li H (2021). RSPCN: super-resolution of digital elevation model based on recursive sub-pixel convolutional neural networks. ISPRS Int J Geoinf, 10(8): 501

Zhao N, Yue T X, Zhao M W, Du Z P, Fan Z M, Chen C F (2014). Sensitivity studies of a high accuracy surface modeling method. Sci China Earth Sci, 57(10): 2386–2396

Zhou A, Chen Y, Wilson J P, Su H, Xiong Z, Cheng Q (2021). An enhanced double-filter deep residual neural network for generating super resolution DEMs. Remote Sens (Basel), 13(16): 3089

Zhou Q, Chen Y (2011). Generalization of DEM for terrain analysis using a compound method. ISPRS J Photogramm Remote Sens, 66(1): 38–45